NNT : 0000AIXM0000

THÈSE DE DOCTORAT

Soutenue à Aix-Marseille Université
le 07 Octobre 2024 par

Clara GUIJARRO CALVO

Analyse quantitative de l'épithélium des cellules progénitrices cardiaques

Discipline
Biologie santé

Spécialité
Biologie du développement

École doctorale
ED 62 - Sciences de la vie et de la santé

Laboratoire/Partenaires de recherche
IBDM
LIS
MMG

Composition du jury :

Adèle FAUCHERRE — Rapporteuse
CNRS, Institut de Génomique Fonctionnelle, Montpellier, FR

Adrien HALLOU — Rapporteur
Kennedy Institute of Rheumatology, Oxford, UK

Shankar SRINIVAS — Examinateur
Department of Physiology Anatomy and Genetics, Oxford, UK

Sigolène MEILHAC — Examinatrice/Présidente
INSERM, Imagine-Institut Pasteur, Paris, FR

Robert G. KELLY — Directeur de thèse
CNRS, IBDM, Turing Centre for Living Systems, Marseille, FR

Paul VILLOUTREIX — Co-encadrant de thèse
INSERM, LIS, MMG, Turing Centre for Living Systems, Marseille, FR

Affidavit

Je soussigné, Clara Guijarro Calvo, déclare par la présente que le travail présenté dans ce manuscrit est mon propre travail, réalisé sous la direction scientifique de Robert G. Kelly and Paul Villoutreix, dans le respect des principes d'honnêteté, d'intégrité et de responsabilité inhérents à la mission de recherche. Les travaux de recherche et la rédaction de ce manuscrit ont été réalisés dans le respect à la fois de la charte nationale de déontologie des métiers de la recherche et de la charte d'Aix-Marseille Université relative à la lutte contre le plagiat.

Ce travail n'a pas été précédemment soumis en France ou à l'étranger dans une version identique ou similaire à un organisme examinateur.

Fait à Marseille, le 1er juillet 20241er juillet 2024



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Affidavit

I, undersigned, Clara Guijarro Calvo, hereby declare that the work presented in this manuscript is my own work, carried out under the scientific supervision of Robert G.Kelly and Paul Villoutreix, in accordance with the principles of honesty, integrity and responsibility inherent to the research mission. The research work and the writing of this manuscript have been carried out in compliance with both the French national charter for Research Integrity and the Aix-Marseille University charter on the fight against plagiarism.

This work has not been submitted previously either in this country or in another country in the same or in a similar version to any other examination body.

Marseille, 1er juillet 20241er juillet 2024



Cette œuvre est mise à disposition selon les termes de la Licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International.

Liste de publications et/ou brevets et participation aux conférences

1) Liste des publications et/ou brevet [1] réalisées dans le cadre du projet de thèse :

    1. Guijarro, C., Song, S., Aigouy, B., Clement, R., Villoutreix, P., Kelly, R.G. Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart Field. (under revision)
    2. Guijarro, C., Kelly, R.G., 2024. On the involvement of the second heart field in congenital heart defects. C. R. Biol. 347, 9–18. https://doi.org/10.5802/crbiol.151

2) Participation aux conférences[2] et écoles d'été au cours de la période de thèse :
    1. Leducq meeting, Virtual - 2021 - Talk
    2. Leducq meeting, Marseille - 2022 - Talk
    3. Weinstein Cardiovascular Development Conference, Marseille - 2022 - Poster
    4. Centuri Hackathon, Marseille - 2022 - Project presentation & supervisor
    5. Weinstein Cardiovascular Development Conference, San Diego - 2023 - Talk
    6. MecaCell3D, Marseille - 2023 - Talk
    7. Internal seminar IBDM - 2023 - Talk
    8. Internal seminar MMG - 2024 - Talk
    9. Data Science bootcamp - January to April 2024

[1] Cette liste comprend les articles publiés, les articles soumis à publication et les articles en préparation ainsi que les livres, chapitres de livre et/ou toutes formes de valorisation des résultats des travaux propres à la discipline du projet de thèse. La référence aux publications doit suivre les règles standards de bibliographie et doit être conforme à la charte des publications d'AMU.

[2] Le terme « conférence » est générique. Il désigne à la fois « conférence », « congrès », « workshop », « colloques », « rencontres nationales et/ou internationales » … etc.

Indiquer si vous avez fait une présentation orale ou sous forme de poster.

Résumé

Au cours du développement, les cellules adoptent diverses stratégies pour sculpter les tissus épithéliaux en formes 3D caractéristiques avec des régions moléculairement et mécaniquement distinctes. Le tube cardiaque des vertébrés en est un bon exemple, car il s'étend par ajout progressif de cellules progénitrices du deuxième champ cardiaque (SHF) à partir d'une feuille épithéliale dans la paroi péricardique dorsale. Contribuant initialement à l'ensemble du primordium cardiaque à travers le mésocarde dorsal, l'ajout de cellules SHF est limité aux pôles cardiaques après une rupture du mésocarde dorsal. La perturbation du déploiement du SHF entraîne un spectre de malformations cardiaques congénitales (CHD). Le facteur de transcription T-box codant pour les gènes Tbx1 et Tbx5, impliqués dans les syndromes génétiques associés aux maladies coronariennes, régule le développement des pôles artériels et veineux et l'émergence d'une frontière lorsque les cellules progénitrices se séparent en pôles cardiaques alternés. Bien qu'il ait été suggéré que la mécanique épithéliale joue un rôle dans l'élongation du cœur primitif, les mécanismes cellulaires contrôlant le déploiement des cellules progénitrices sont inconnus. Nous présentons ici une analyse morphométrique quantitative des propriétés épithéliales des cellules SHF dans toute la paroi péricardique dorsale murine aux jours embryonnaires 8,5 et 9,5. Les modèles de stress et de tension épithéliales dans le SHF sont intégrés dans notre ensemble de données à l'aide d'une analyse d'images par inférence de force et documentent la relation dynamique entre la forme des cellules et les forces mécaniques dans les embryons de type sauvage avant et après la rupture du mésocarde dorsal. Nous montrons que l'allongement et la tension cellulaires sont les principaux paramètres définissant la morphologie des cellules SHF et identifions les propriétés distinctives des cellules progénitrices adjacentes aux pôles artériels et veineux. De plus, l'analyse morphométrique d'embryons mutants révèle que les modèles d'orientation cellulaire et de stress épithélial dans la paroi péricardique dorsale sont régulés différemment par Tbx1 et Tbx5. En intégrant les résultats de la structuration des caractéristiques à l'échelle cellulaire, du stress mécanique des tissus et de la biologie cellulaire, notre projet fournit de nouvelles informations sur la morphogenèse cardiaque et les origines des CHD.

Mots clés : morphogenèse cardiaque, champs cardiaques, maladies cardiaques congénitales, mécanique des épithéliums, analyse quantitative, gènes T-box.

Abstract

During development, cells adopt diverse strategies to sculpt epithelial tissues into characteristic 3D shapes with molecularly and mechanically distinct regions. The vertebrate heart tube is a good example, as it extends by progressive addition of second heart field (SHF) progenitor cells from an epithelial sheet in the dorsal pericardial wall. Initially contributing to the entire cardiac primordium across the dorsal mesocardium, SHF cell addition is restricted to the cardiac poles after dorsal mesocardial breakdown. Perturbation of SHF deployment results in a spectrum of congenital heart defects (CHD). The T-box transcription factor encoding genes Tbx1 and Tbx5, implicated in genetic syndromes associated with CHD, regulate arterial and venous pole development and the emergence of a boundary as progenitor cells segregate to alternate cardiac poles. While epithelial mechanics has been suggested to play a role in the elongation of the primitive heart, the cellular mechanisms controlling progenitor cells deployment are unknown. Here, we present quantitative morphometric analysis of epithelial properties of SHF cells throughout the murine dorsal pericardial wall at embryonic days 8.5 and 9.5. Patterns of epithelial stress and tension in the SHF are integrated in our dataset using force inference image analysis and document the dynamic relationship between cell shape and mechanical forces in wild-type embryos before and after dorsal mesocardium breakdown. We show that cell elongation and tension are the main parameters defining SHF cell morphology and identify distinguishing properties of progenitor cells adjacent to the arterial and venous poles. Moreover, morphometric analysis of mutant embryos reveals that patterns of cell orientation and epithelial stress in the dorsal pericardial wall are differentially regulated by Tbx1 and Tbx5. By integrating findings from cell-scale feature patterning, mechanical tissue stress and cell biology our project provides new insights into cardiac morphogenesis and the origins of CHD.

Keywords: heart morphogenesis, heart fields, congenital heart disease, epithelial mechanics, quantitative analysis, single-cell morphometrics, T-box genes.

Acknowledgments

I want to thank my previous academic supervisors, Onnik Agbulut, J.C Izpisua Belmonte, and Sigolène Meilhac, for accepting me into their labs, training and recommending me for the PhD program.

I extend my gratitude to Jorge and Adrian for introducing me to 3D animation and creative coding, and to the teachers and the staff from the Data Science bootcamp for simplifying complex concepts and helping me navigate analysis and debugging.

Special thanks to Vincent, Laurent, Fumiyuki and Yuta for their meticulous training and dedication. Your technical expertise and patience have been invaluable in making me more independent in the lab for my PhD.

I am grateful for the interdisciplinary education I received, allowing me to learn from and be inspired by passionate individuals from various fields. I thank the members of both labs, Paul and Robert for their guidance and support, and Solène, Benoit, and Raphael for their help, comments, and feedback during my PhD.

I appreciate the thesis committee for following my PhD journey and the jury members for agreeing to review my thesis work.

I also appreciate the financial support from Centuri, Marmara, G-Research and FRM.

Lastly, I thank my family and friends for their support over the past 4 years. I regret missing important moments but look forward to creating new memories and projects together.

Index

Affidavit in French / English 2-3
List of publications and congresses 4
Résumé / Abstract 5-6
Acknowledgments 7

1. Introduction 10
1.1 Heart development and morphology 11
1.1.1 Morphology and function of the adult heart 11
1.1.2 Congenital heart defects 12
1.1.3 Embryonic development of the heart 15
1.1.4 The second heart field 15
1.1.5 The involvement of the T-box genes in the arterial and venous pole development and their importance for cardiac septation 23
1.1.5.1 Tbx1 and arterial pole development 23
1.1.5.2 Tbx5 and cardiac septation 23
1.2 Epithelia mechanics 27
1.2.1 Mechanics in heart development 28
1.2.2 Epithelia and mechanical properties of cardiac progenitors – Biomechanical model 32
1.2.3 Mechanics regulating tissue morphogenesis and patterning 36
1.2.4 Mechanical forces regulating cell division and proliferation 38
1.2.5 Bayesian force inference 38
1.3 Morphometrics and analysis 40
1.3.1 Morphometrics 40
1.3.2 Machine learning for data analysis 42
1.3.2.1 Supervised learning 44
1.3.2.2 Unsupervised learning 45
1.3.2 Applications of machine learning in biology 49
1.4 Goal and interests of the project 55

2. Materials and methods 57
2.1 Animal model 57
2.2 Embryo dissection 58
2.3 Embryo culture 58
2.4 Immunofluorescence staining on whole mount embryos 58
2.5 RNAscope on wholemount embryos & quantification 59
2.6 In Situ Hybridization riboprobe synthesis 59
2.7 Whole mount in-situ hybridization 60
2.8 Imaging and cell membrane segmentation 60

2.9 Rosette quantification....61
2.10 Image analysis....61
2.11 Data analysis for spatial single-cell morphometrics....61

3. Results....63
3.1 Single-cell morphometrics reveals T-box-dependent patterns of epithelial tension in the SHF....64
Supplementary figures - Single-cell morphometrics reveals T-box-dependent patterns of epithelial tension in the SHF....102
3.2 Boundary formation in the SHF....127
3.3 Cell intercalation and rosette formation in the SHF....133
3.4 Exploring additional data analysis (Annex on Data Science bootcamp)....138

4. Discussion....142
4.1 Discussion on single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the SHF....142
4.2 Discussion on boundary formation in the SHF....151
4.3 Discussion on the intercalation and rosette formation....155
Bibliography....156
Annex - Review- On the involvement of SHF in congenital heart defects....190
List of figures....201
List of abbreviations....204

1. Introduction

"To carry into the study of living beings the laws and lessons of the inorganic should also be considered" (Thomson, 1917).

If we consider that the entire body is an organized society and a cell is an individual, how do the cells know their function in the society, how do they specialize, which are the rules or cues that guide them to execute their job properly and in the right moment?

But what is even more interesting about this society is that the cells, not only they are individuals with specific functions, jobs and tasks but they are also the materials and the building blocks of the society and as such, they respond not only to the rules of genetics but also the rules of physics.

The introduction is divided as follows: a biological part with the principal notions in heart development and morphology (Fig. 1.1A), followed by a biophysics part in epithelia dynamics and Bayesian force inference (Fig. 1.1B), and the third and last part with some concepts in computer science analysis with a special emphasis in clustering methods (Fig. 1.1C-D).

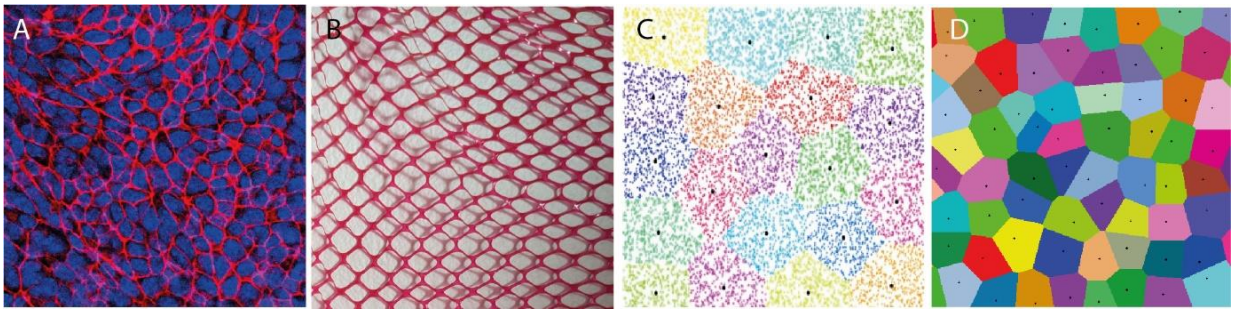


Figure 1.1. Example of a similar shape in different disciplines. A. Cells **B.** Orange bag mesh **C. P**artition of data by k-means **D.** Voronoi diagram

1.1 Heart development and morphology

1.1.1 Morphology and function of the adult heart: the four-chambered pump

The adult heart is a muscular organ responsible for pumping blood through the body. Its morphology and function are intricately connected, contributing to its efficient and continuous operation. It is divided into 4 chambers: two atria (upper chambers) and two ventricles (lower chambers). The heart is enclosed by a double-layered membrane called the pericardium.

The 4 chambers are separated by muscular walls called septa such as the interatrial septum, that separates the right atrium from the left atrium. During fetal development, there is an opening in the interatrial septum called the foramen ovale, which closes shortly after birth. The interventricular septum is the muscular wall that separates the right and left ventricle and plays a crucial role in preventing the mixing of oxygenated and deoxygenated blood. These septa are important structures within the heart as they serve as partitions within the heart, ensuring the separate circulation of blood through the pulmonary and systemic circuits (Fig. 1.2).

The heart wall consists of three layers: the epicardium (outer layer), myocardium (middle layer composed of cardiac muscle), and endocardium (inner layer lining the chambers and valves). The myocardium is especially crucial for the heart's function, as it contracts to pump blood.

From the functional point of view, blood flows through the heart in a unidirectional manner, facilitated by valves that separate the chambers. The tricuspid valve separates the right atrium and right ventricle, while the bicuspid (mitral) valve separates the left atrium and left ventricle. Semilunar valves, namely the pulmonary and aortic valves, regulate blood flow out of the heart into the pulmonary and systemic circulations, respectively.

The cardiac cycle consists of systole (contraction) and diastole (relaxation). During systole, the ventricles contract, forcing blood into the pulmonary artery and aorta. In diastole, the ventricular muscle relaxes, allowing the chambers to fill with blood from the atria. This rhythmic cycle is orchestrated by the sinoatrial (SA) node, the heart's natural pacemaker, the atrioventricular (AV) node, and the ventricular conduction system coordinating and relaying electrical impulses that stimulate muscle contractions.

In summary, the morphology and function of the adult heart are intricately linked to sustain an efficient circulatory system. Problems in the morphology of the heart lead to problems in the establishment of separate pulmonary and systemic blood flows.

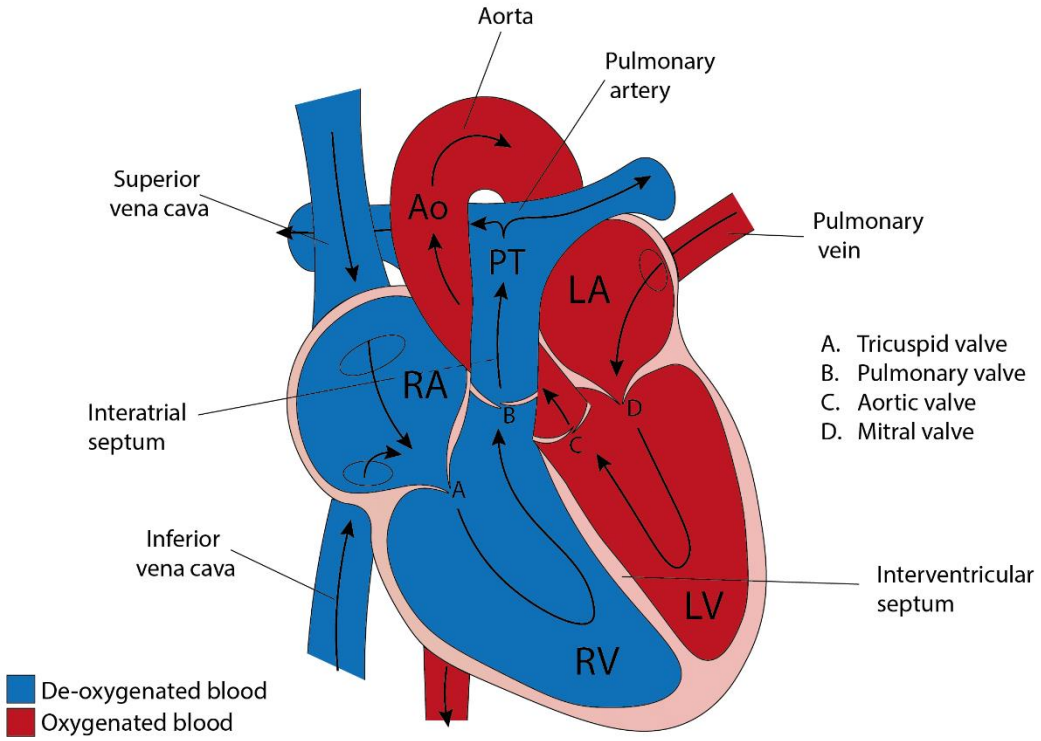


Figure 1.2: Morphology and blood flow of the adult heart. Regions with deoxygenated blood are shown in blue, regions with oxygenated blood are shown in red, the purple arrows indicate the direction of the blood flow. Ao: aorta, PT: pulmonary trunk, RA: right atria, LA: left atria, RV: right ventricle, LV: left ventricle.

1.1.2 Congenital heart defects

Congenital heart defects (CHD) are the most common congenital anomalies in newborns (Bouma & Mulder, 2017). A CHD is defined as a structural abnormality of the heart and great vessels that is present at birth. The prevalence of CHD is between 0.8 and 1.2 % of births (Bouma & Mulder, 2017; van der Linde et al., 2011). The etiology involves both genetic and environmental factors and is incompletely understood due to the fact that these defects arise during embryonic development.

Before 1950, severe CHD were almost fatal, while now advances in medicine and surgery have decreased the mortality and enable the patients to reach adulthood. A meta-analysis from 2020 has revealed that there is a reduction in global child mortality; however, the contribution of congenital anomalies to child mortality is increasing (Glinianaia et al., 2020). Here I present some examples of common forms of CHDs.

Atrial septal defects (ASD) have a prevalence of 1 of 1500 live births and constitute 6-10% of congenital heart defects (Fig. 1.3A). Females show a double predominance compared

to males. In Atrial septal defects, the blood flow comes from the left to the right atrium (D. J. Driscoll, 1999). This increase provokes an excess of volume flow affecting the right structures of the heart: the atrium, ventricle, the pulmonary valve and the pulmonary circulation.

As for Ventricular Septal Defects (VSD), they have a prevalence of 1.5 to 3.5 of 1000 live births, constituting 20% of CHD (Fig. 1.3B). They do not show any gender predominance. In non-severe ventricular septal defects, the blood flows from the left to the right ventricle. This flow will lead to a volume overload of the pulmonary circulation, the left atrium and ventricle. If the patient develops pulmonary vascular obstructive disease, the direction of the blood flow can be reversed and the poorly oxygenated blood will enter the aorta.

Tetralogy of Fallot constitutes around 4-8% of congenital cardiac defects (Fig. 1.3C). It is a compilation of ventricular septal defects, pulmonary stenosis, an aorta that overrides the ventricular septal defect and a right ventricular hypertrophy. Due to the right-to-left switch of the blood flow, the body has a diminished quantity of oxygen (hypoxemia) (Fig. 1.3D and Table 1.1).

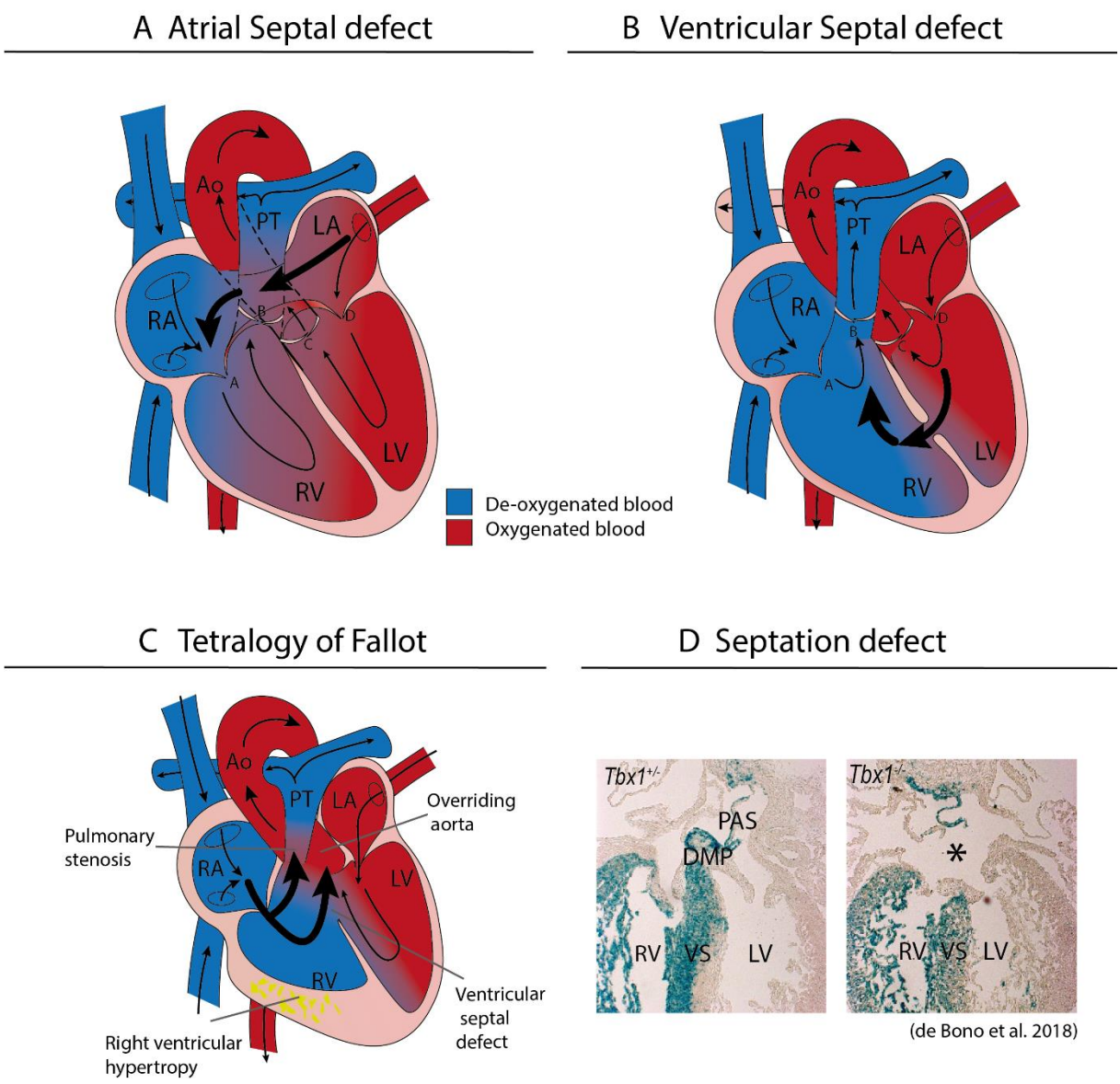


Figure 1.3: Morphology of adult hearts with atrial or ventricular septal defects. Cartoons of **A**. an open frontal adult heart with an Atrial septal defect **B.** an open adult heart with a Ventricular septal defect and **C.** an open frontal adult heart with tetralogy of Fallot**, D.** Example of an atrial/atrioventricular septation defect shown in sections. *Mef2C-AHF-Cre* activated conditional *lacZ* reporter gene in blue (De Bono et al., 2018). Atrial septal structures, the dorsal mesenchymal protrusion and primary atrial septum are defective in *Tbx1* mutant hearts. Black arrows show the direction of the blood flow. FHF: first heart field, SHF: second heart field, A: atria, V: ventricle, PT: pulmonary trunk, Ao: aorta, VS: ventricular septa, RA: right atria, LA: left atria, RV: right ventricle, LV: left ventricle, DMP: dorsal mesenchymal protrusion, PAS: pulmonary atrial septum, A: Tricuspid valve, B: Pulmonary valve, C: Aortic valve, D: Mitral valve.

Defect	Prevalence	% of CHD	Types / Morphological defects	Genes involved
ASD	1/1500 live births	6-10%	Ostium secundum	*Nkx2-5* *Tbx5*+*Gata4* (secundum) *Sema3D* (sinus venosus)
			Ostium primum	
			Sinus venosus	
			Unroofed coronary sinus	
VSD	1.5-3.5/1000 live births	20%	Muscular	*Gata4* / *Tbx1* / *Tbx5*
			Central membranus or perimembranosus: under the aortic valve	
			Outflow tract defect (outlet)	
			Inflow tract defect (inlet)	
Tetralogy of Fallot	3/10000 live births	4-8%	Ventricular septal defect, pulmonary stenosis, overriding aorta, right ventricular hypertrophy	*Gata4* / *Nkx2-5* / *Tbx1* / *Tbx5* / *Jag1*
22q11.2 deletion (DiGeorge) syndrome	~2/10000 live births	~ 2%	Inmmunodeficiency, cardiac defects, craniofacial defects, hypocalcemia, gastrointestinal problems, psychiatric disorders	*Tbx1*
Holt-Oram syndrome	1/100000 live births	< 1%	Upper limbs abnormalities, cardiac defects such as ventricular septal defects, atrial septal defects	*Tbx5*

Table 1.1: Examples of congenital heart defects with their prevalence, the morphological defects associated and the genes involved.

1.1.3 Embryonic development of the heart: from a tube to a four-chambered pump

The heart is formed by two distinct populations of progenitor cells: the first (FHF) and second heart fields (SHF) (Kelly, 2012), corresponding to the first and second waves of mesodermal cells that commit to cardiac fate. Genetic and clonal analysis experiments have shown that the first and second heart fields correspond to distinct lineages defined by their exclusive contributions in the adult heart (Kelly et al., 2014). The Juxtacardiac field (JCF) has also been recognized as a third population that contributes to cardiomyocytes via a FHF transcriptional state as well as to the proepicardium (Tyser et al., 2021) (Fig. 1.4).

The first cells expressing cardiac muscle markers, the FHF cells, are detected prior to the formation of the primitive linear heart tube, they are found under the head folds in a

region known as the cardiac crescent. The primitive linear heart tube forms as a fusion of the cardiac crescent at the midline (Fig. 1.4A).

The primitive heart is a linear heart tube oriented along the antero-posterior axis at embryonic day (E) 8. The arterial and venous poles are also situated in the antero-posterior axis. At this primary time point, the poles are connected to the splanchnic mesoderm by the dorsal mesocardium. SHF progenitors migrate through the dorsal mesocardium to contribute to the growth of the primary heart (Fig. 1.4B1). As the heart tube elongates, the poles are no longer connected by the dorsal mesocardium. The heart tube has detached from the SHF and the cells contribute to the growth of the heart by addition at the arterial and venous poles. At the same time as the heart elongates, it undergoes looping. Cardiac looping will be responsible of the formation of 1) an outer curvature and 2) inner curvature. The outer curvature controls the growth of the ventricles, it undergoes ballooning and trabeculation (Christoffels et al., 2000), while the inner curvature controls the alignment between the inlet and outlet segments (Srivastava, 2006). The right ventricle develops later than the left ventricle by addition of cells from the anterior SHF (Fig. 1.4B2).

After these processes chamber growth and septa formation take place. This step is crucial to maintain correct separation between oxygenated and non-oxygenated blood flow (Fig. 4C). At this stage, the atria and atrioventricular canal are located above the future left ventricle and the outflow tract is located entirely above the future right ventricle.

The interventricular septum forms after cardiac looping, as a myocardial ridge at the posterior part and the apex of the primitive heart (Christoffels et al., 2000; Srivastava, 2006). The ventricular septum originates from both first and second heart fields as it is composed by myocytes of both left and right ventricles(Franco et al., 2006). The inlet septum is controversial. Old studies define the inlet septum as the postero-inferior or caudal part of the ventricular septum (Goor et al., 1970) while more recent studies defend that this structure does not exists in the adult heart (Anderson et al., 2014). The outlet septum forms by the union and the muscularization of the proximal part of the endocardial cushions of the OFT and is taken along leftward by the rotation of the developing aortic valve, to fuse with the upper primitive septum. This rotation depends on the proper growth of the OFT by addition of myocardial cells from the anterior SHF as well as addition of cardiac neural crest cells that are essential to divide the outflow tract into the aorta and pulmonary trunk (Bajolle et al., 2006 ; Bajolle et al., 2008).

When it comes to the atrial septum, from E9.5 to E10, the pSHF give rise to the dorsal mesenchymal protusion (DMP) (Fig. 1.4D). The primary atrial septum develops at E10, growing from the dorsal roof of the common atrium and the DMP (Mommersteeg et al., 2006). The secondary atrial septum (sAS) develops within the roof of the right atrium. This septum grows into the right atrium, without physically fusing with the lower part of

the primary septum to facilitate the embryonic blood circulation. It is only after birth that the two septa will fuse (Deepe et al., 2020).

These events are tightly regulated by various molecular signals, transcription factors, and cellular interactions. Also, heart development depends on the development of adjacent tissues. For example, the endoderm has been thought to play a passive role as the substrate for SHF cells; however, new studies have led to the conclusion that mechanical contraction, shortening and folding of the endoderm can impact myocardial tube formation (Kidokoro et al., 2018; Varner & Taber, 2012; Esteban et al., 2022). Disruptions at any stage can lead CHD, emphasizing the importance of understanding the mechanisms of embryonic heart development for both developmental biology and clinical applications.

This thesis and work will be focused on the study of the heart tube looping and elongation step (Fig.1.4B) and the deployment of the SHF cells to the arterial and venous pole, which is an essential prerequisite for subsequent cardiac septation.

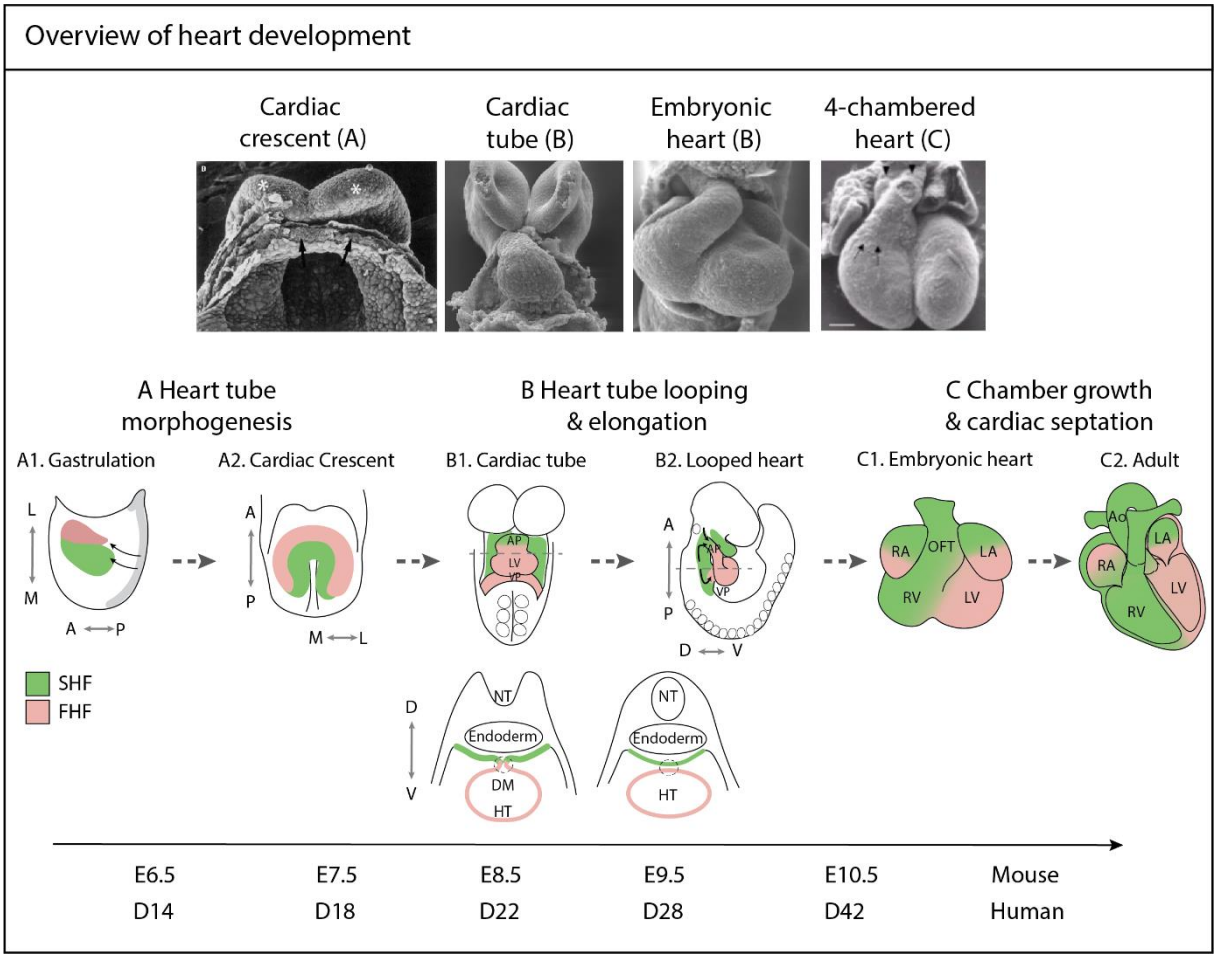


Figure 1.4. Overview of heart development. A. Electron microscopy images of embryonic heart in mouse embryos and scheme showing cardiac first (pink) and second (green) heart field progenitor cells and heart morphology **A.** Heart tube morphogenesis **B.** Heart tube looping and elongation **C.** Chamber growth and cardiac septation**.** FHF: first

heart field, SHF: second heart field, L: Lateral, M: medial, A: anterior, P: posterior, D: dorsal, V: ventral, AP: arterial pole, VP: venous pole, HT: Heart tube, RA: right atria, LA: left atria, RV: right ventricle, LV: left ventricle, OFT: outflow tract, Ao: Aorta, E: mouse embryonic day, D: human embryonic day. Electron microscopy images from (Moorman et al., 2003).

1.1.4 The Second Heart Field

In 1977, it was shown that the avian outflow tract (OFT) is a late addition during heart tube elongation, originating from a source of cells located outside the heart tube (de la Cruz et al., 1977). However, it wasn't until 2001 that the second heart field (SHF) of cardiac progenitor cells was identified through cell labelling experiments in avian embryos (Waldo et al., 2001; Mjaatvedt et al., 2001), and by analysis of an *Fgf10*-enhancer trap transgene expressed in murine pharyngeal mesoderm (Kelly et al., 2001). In fact, the SHF has a common origin with head muscle progenitor cells in cardiopharyngeal mesoderm (CPM) (Diogo et al., 2015).

Initially the SHF contributes to the growth of the heart along the entire length of the cardiac primordium by a transient structure called dorsal mesocardium. Ultimately, SHF addition becomes restricted to the arterial and venous poles after breakdown of the dorsal mesocardium. Described from a temporal and morphological point of view, at E8 the 2 left and right epithelia are connected to the heart tube by the dorsal mesocardium along the antero-posterior axis. After fusion of the bilateral epithelia at E9 the dorsal mesocardium breaks down behind the body of the heart tube and the linear heart tube is connected to the DPW only by the anterior (arterial) and posterior (venous) poles. The distance between the poles is maximal at E9.5 and reduces at E10 (Fig. 1.5A, B).

The SHF provides a pool of proliferating and non-differentiated cardiac progenitor cells that is maintained as the heart grows. It plays an important role in the elongation of the heart tube and SHF-derived cells participate in later cardiac septation. Transcription factors and signaling pathways expressed in these cells play a critical role in regulating the balance between proliferation and differentiation of the cells and their deployment to the cardiac poles.

The SHF is defined by expression of the transcription factor encoding gene *Islet1,* as well as other genes involved in myocardial determination like *Nkx2-5*, *Gata4* or *Mef2c* (Cai et al., 2003). Even if considered as a unique cellular population, the regional heterogeneity in this population is evident. The SHF is patterned on the anterior posterior axis (Fig. 1.5). The regional heterogeneity is especially evident between the anterior and the posterior parts correlating with the contribution to the arterial and the venous pole (Vincent & Buckingham, 2010). Despite the deployment of the cells into anterior and posterior regions, SHF cells added to the arterial and venous pole are clonally related and originate

from common progenitors (Meilhac et al., 2004; Kelly, 2012). The anterior part of the SHF (aSHF) expresses genes including *Fgf8*, *Fgf10* or the *Mef2c* enhancer which mark progenitors of the right ventricle as well as outflow tract myocardium, both the pulmonary trunk and the aorta;. *Tbx1* is also expressed in the aSHF and contributes to myocardium at the base of the pulmonary trunk (Kelly et al., 2001; Verzi et al., 2005; Baldini et al., 2017). *Foxh1* codes for a transcription factor that plays a role in the development of the aSHF, mutations in the gene are linked to arterial pole defects (von Both et al., 2004). The posterior SHF (pSHF) is marked by *Osr1, Hox* genes such as *Hoxb1* or *Tbx5* which is expressed in the pSHF under the control of retinoic acid signaling (Xie et al., 2012; De Bono et al., 2018.The pSHF gives rise to atrial myocardium including the dorsal mesenchymal protrusion. *Tbx18* is expressed in the lateral domain of the SHF where there is an absence of expression of *Nkx2-5* and *Islet1* (Christoffels et al., 2006). The distinction between anterior and posterior SHF domains is made on the basis of gene expression patterns; however, it remains unclear whether cells can move between the domains. In addition to the patterning between the anterior/posterior axis, the SHF also shows left-right patterning. This is driven by Nodal signaling pathway which leads to expression of the transcription factor Pitx2 in cells on the left side of the SHF (Galli et al., 2008).

The SHF cells need to be in a balance between proliferation and differentiation (Vincent & Buckingham, 2010): On the one hand FGF, canonical WNT and hedgehog signaling promote proliferation; on the other hand, antagonizing the proliferative capacities, BMP is the principal signaling regulating differentiation, triggered at the poles of the heart (Hutson et al., 2010). However, at the venous pole, BMP together with WNT and hedgehog signaling promotes migration of the cells to the heart tube at the same time that they proliferate (Xie et al., 2012). These pathways can be activated by the surrounding tissues (Kelly, 2012), such as the neural tube and the notochord, sources of canonical WNT and Hedgehog signaling, promoting proliferation, by the endoderm, a source of hedgehog signaling or by neural crest cells that migrate from the pharyngeal arches to the arterial pole and play an important role in division of the outflow tract into the pulmonary trunk and aorta (Kirby & Waldo, 1990). These pathways can also been activated by autonomous signaling within the SHF (Kelly, 2012; Vincent & Buckingham, 2010). For example, the transcription factor Tbx1 may interfere with BMP signaling pathway to delay differentiation (Kelly, 2012; Chen et al., 2009) (Fig. 1.5-1.6).

In addition to their genetic patterning and signaling mechanisms, SHF cells have been shown to constitute an atypical epithelium in the dorsal pericardial wall, with apical tight and adherens junctions and highly dynamic basal filopodia more typical of mesenchymal cells (Francou et al., 2014) . The epithelial properties of SHF cells are emerging as important characteristics controlling SHF deployment and are regulated by cell adhesion molecules, integrin signaling and components of the planar cell polarity pathway, such as Wnt5a and Vangl2, as well as TBX1, a regulator of Wnt5a, required for focal adhesion and filopodial activity in the SHF (Ramsbottom et al., 2014 ; Cortes et al., 2018 ; Francou et al.,

2014 ; L. Chen et al., 2012). Moreover, analysis of apical junctions, cell shape and actomyosin distribution in the SHF has suggested biomechanical models driving SHF deployment, including planar cell polarity driven push and pull mechanisms and an epithelial tension driven circuit promoting proliferation in the posterior region of the dorsal pericardial wall (D. Li et al., 2016 ; Francou et al., 2017) (Fig. 1.5D-E). Epithelial properties of the SHF cells will be described in more detail in the next section.

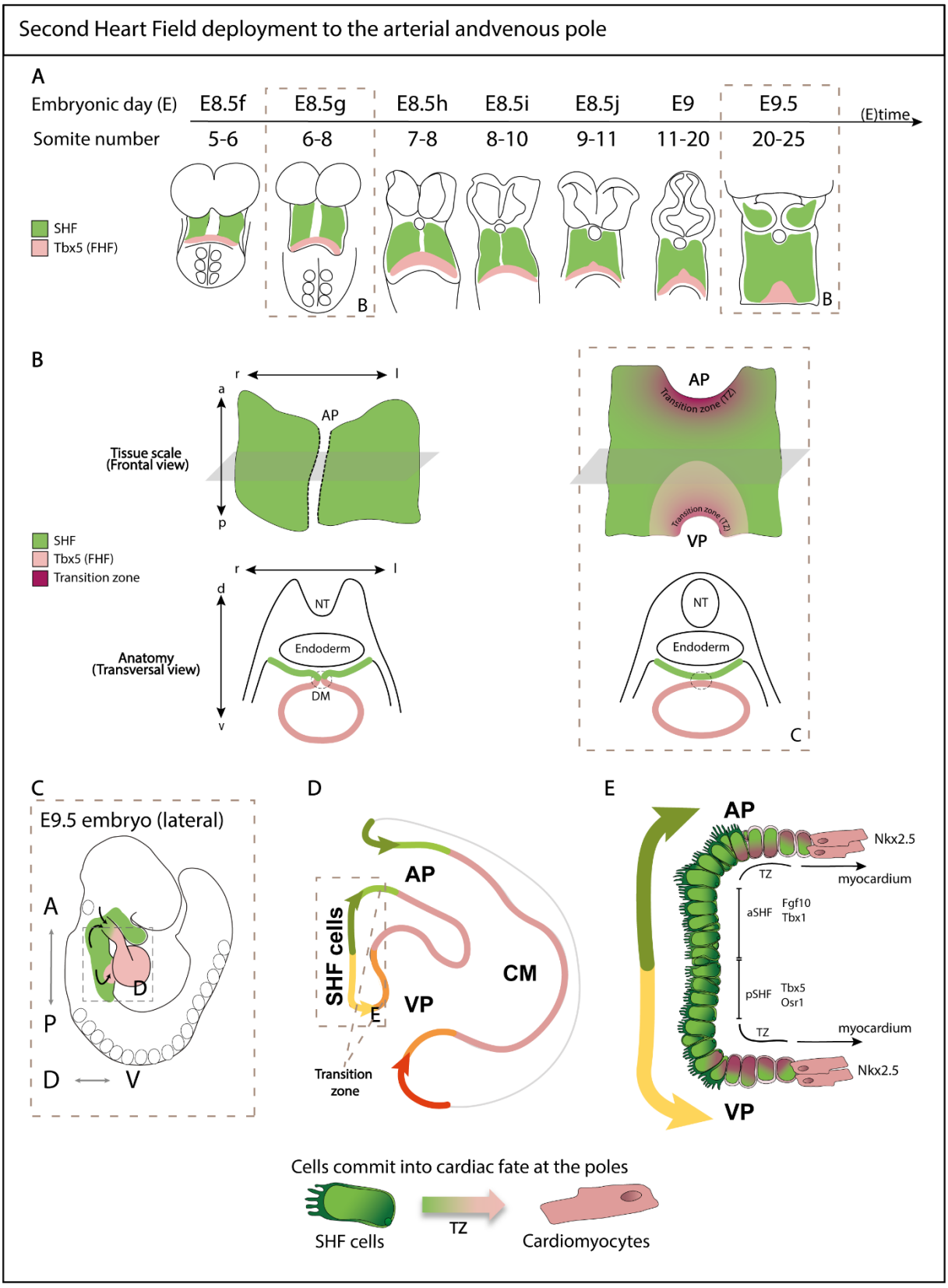


Figure 1.5. SHF deployment and heart tube elongation. A. SHF development from E8.5 to E9.5. **B.** SHF transition between before dorsal mesocaridum breaks down (E8.5) and after dorsal mesocardium breaks down (E9.5). **C.** Contribution of the SHF to the arterial and venous pole of the heart **D.** Sagittal view of the SHF cells with magnification.

Visualization of the atypical nature of the SHF cells, and how as they aproach to the arterial and venous pole they commit into cardiac fate. Pattern of the genes and proteins that determine the different regions in this epithelium. FHF: first heart field, SHF: second heart field, aSHF: anterior SHF, pSHF: posterior SHF, A: anterior, P: posterior, D: dorsal, V: ventral, AP: arterial pole, VP: venous pole, tube, RA: right atria, LA: left atria, RV: right ventricle, LV: left ventricle, NT: Neural tube, TZ: Transition zone. Panels A and B adapted from Santini et al. 2016 and Le Garrec et al. 2017.

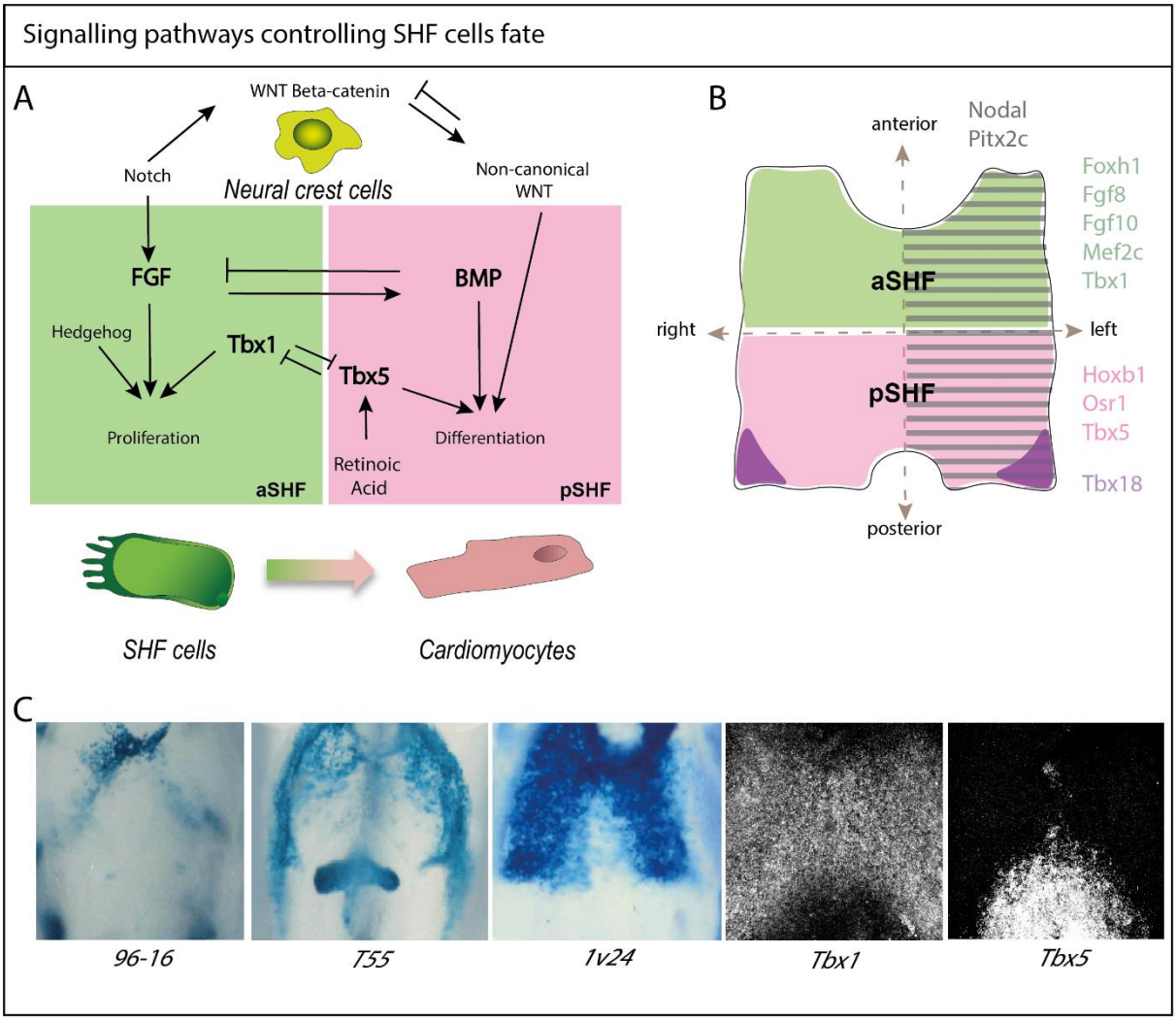


Figure1.6. Signaling pathways controlling proliferation and differentiation of SHF cells and gene expression patterning of the SHF cells. **A.** Schematic of principal pathways involved in the SHF cell differentiation **B.** Schema of the spatial distribution of the most important genes patterning the SHF known to date **C.** Expression patterns of some genes in the SHF, reporter gene expression of the y96-Myf5-nlacZ-16 (*96-16)* transgene making a subpopulation of myocardial cells in the second heart field, *A17-Myf5-nlacZ-T55 (T55),* a transgene at the *Myf5* locus and *Mlc1v-24-nlacZ* transgenic line (*1v24)* transgene, which reflect the expression of Fgf10. Images showing gene expression by X-gal labelling from Kelly et al. 2001, Bajolle et al. 2008, and RNAscope of *Tbx1*, *Tbx5*.

1.1.5 The involvement of T-box genes in arterial and venous pole development and their importance for cardiac septation

1.1.5.1 *Tbx1* and arterial pole development

The T-box transcription factor TBX1 has been identified as a critical regulator of SHF development, promoting progenitor cell proliferation and delaying differentiation in the SHF (Baldini et al., 2017). TBX1 is the major gene implicated in 22q11.2 deletion or DiGeorge syndrome, involving parathyroid aplasia, craniofacial defects such as cleft palate and congenital heart defects such as tetralogy of Fallot or common arterial trunk due to the function of Tbx1 in the SHF but also in the endoderm and ectoderm in the pharyngeal region (Momma, 2010; Unolt et al., 2018). Single-cell RNA-seq experiments have shown that Tbx1 controls myogenic trajectories from a multilineage primed progenitor cell population to cardiac and skeletal muscle fates (Nomaru et al., 2021). Tbx1 null embryos display common arterial trunk associated with defective cardiac neural crest cell development; in addition, myocardium at the base of the pulmonary trunk, the outlet of the right ventricle, fails to contribute to the arterial pole of the heart (Baldini et al., 2017; Théveniau-Ruissy et al., 2008; Rana et al., 2014). Underdevelopment of subpulmonary myocardium has been proposed to be a primary upstream mechanism leading to tetralogy of Fallot (Wilson et al., 2019) (Fig. 1.7A-B). Tetralogy of Fallot is genetically heterogenous but the major known genetic cause is haploinsufficiency for TBX1, 22q11.2 DS accounting for over 15% of patients (Unolt et al., 2018).

Tbx1 is expressed in the anterior SHF and lateral part of the posterior SHF, which retains the early SHF genetic program, and contributes to the myocardium at the base of the pulmonary trunk. Tbx1 may interact with BMP signaling to delay differentiation (Kelly, 2012; L. Chen et al., 2009) and with Pitx2 in the rotation of the outflow tract, which is essential for establishing the correct alignment of the pulmonary trunk and aorta (Bajolle et al., 2006). Even if Tbx1 is more related with arterial pole defects, loss of function of Tbx1 impacts not only the anterior SHF but also the pSHF resulting in venous and arterial pole defects (Rana et al., 2014). Moreover, *Tbx1-Cre* genetic tracing experiments reveal transient expression in progenitor cells giving rise to the atrial septum and dorsal mesenchymal protrusion (Huynh et al., 2007) (Fig. 1.7A,B).

1.1.5.2 *Tbx5* and cardiac septation

Tbx5, a key regulator of the FHF and venous pole differentiation program, is activated in the posterior SHF cells by the influence of retinoic acid (RA) signaling (Bruneau et al., 2001; De Bono et al., 2018). *Tbx5* is upregulated in Tbx1+ second heart field as the dorsal mesocardium breaks down and the heart tube elongates; as a consequence, progenitor cells segregate to alternate poles. Transient co-expression of TBX1 and TBX5 is followed by downregulation of the anterior SHF program, including expression of *Tbx1*, and

establishment of transcriptional boundary within the dorsal pericardial wall separating distinct TBX1 positive arterial pole progenitors in the anterior SHF from TBX5 positive venous pole progenitors in the medial posterior SHF (De Bono et al., 2018) (Fig. 1.7C,D). Retinoic acid signaling is implicated in the formation of the anterior-posterior regions of the mesoderm. Raldh2, an enzyme responsible for the production of RA from retinaldehyde, and encoded by the *Adlh1a2* gene, is expressed in the pSHF (Mic et al., 2002). In *Aldh1a2* null embryos, the expression of aSHF genes is expanded towards the posterior domain, the *Tbx5* expression domain is restricted to the FHF (Sirbu et al., 2008). In these embryos the atria and sinus venosus fail to develop. However, overexpression of Tbx5 and retinoic acid in the pSHF does not expand the expression of the pSHF towards the anterior part of the epithelium (De Bono et al., 2018).

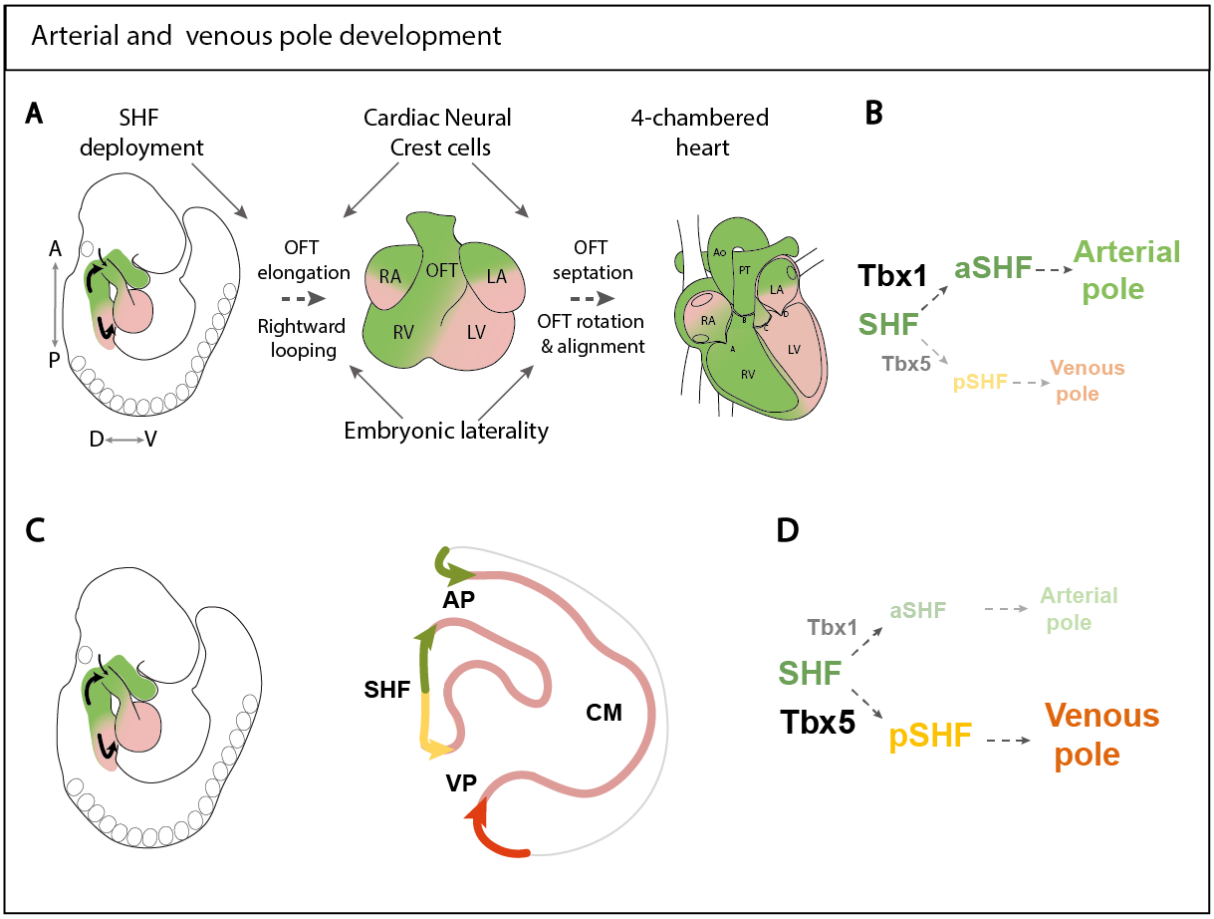


Figure 1.7. Arterial and venous pole development. A. Cartoon showing the principal events for correct arterial pole and outflow tract development. **B.** Tree diagram of the process of SHF deployment with special emphasis in the aSHF branch activated by TBX1. **C.** Scheme of a E9.5 embryo and a high magnification of the second heart field (SHF) segregation into anterior SHF (aSHF) and posterior SHF (pSHF) and its deployment to the arterial and venous poles **D.** Tree diagram of the process of SHF deployment with a special emphasis in the pSHF branch activated by the activation of *Tbx5*. SHF: Second heart field, aSHF: anterior second heart field, pSHF: posterior second heart field, A: anterior, P: posterior, D: dorsal, V: ventral, RA: right atria, LA: left atria, RV: right ventricle, LV: left

ventricle, OFT: outflow tract, Ao: aorta, PT: pulmonary trunk, DPW: Dorsal pericardial wall.

Mutations in TBX5 result in Holt-Oram syndrome, characterized by atrial septal and forelimb defects (Steimle & Moskowitz, 2017). TBX5 has been shown to activate an enhancer of *Aldh1a2*, pointing to a feed forward mechanism in the posterior SHF domain (Rankin et al., 2021). As well, atrial septal defects are observed in conditional mutant mouse embryos lacking Tbx5 in the SHF (Xie et al., 2012). Independent functions of *Tbx5* in the endocardium and the second heart field can explain the involvement of these genes in different pathologies. These functions have been explored in experiments where one or both copies of the gene are deleted in a specific cell type of time period. Conditional knock-out of Tbx5 in endothelial cells increase the number of apoptotic cells in the primary septum (Nadeau et al., 2010). On the other hand, conditional deletion of Tbx5 in the SHF prevents the development of the septum secundum (Xie et al., 2012). Cells expressing both Tbx5 and the SHF program have also been identified in the interventricular septal region using an intersectional lineage reporter approach (Devine et al., 2014). Conditional inactivation using Cre recombinase driven by the *Mef2c* enhancer has shown that *Tbx5* is required in the SHF lineage for formation of the muscular ventricular septum as well as for atrial septation (Xie et al., 2012); moreover, the transcription factors MEF2C and TBX5 co-regulate target genes required for ventricular septal development (Kathiriya et al., 2021). Thus, both atrial and ventricular septal structures form at sites where *Tbx5* expression and the SHF program transiently overlap, implying that this may define a boundary program that is required to initiate septal morphogenesis (Fig. 1.8).

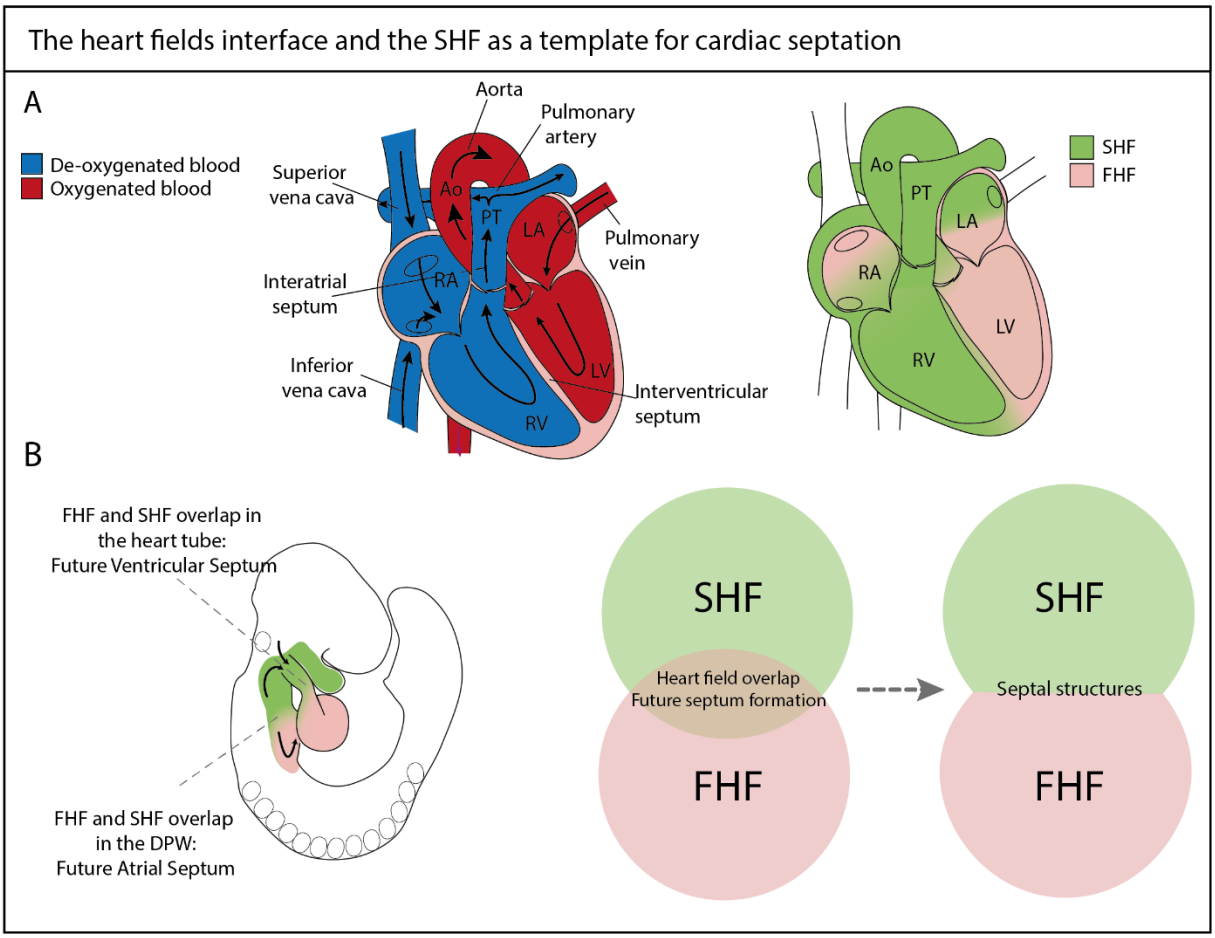


Figure 1.8. The heart field interface and cardiac septation. A. On the left the normal human heart is color coded according to the oxygenation of the blood flow: red if the flow is rich in O_2; blue, if blood rich in CO_2. On the right the normal human heart color is coded according to the contribution of the region from the FHF in pink or from the SHF in green. **B.** The atrial and ventricular septum appear in the boundary between the FHF and the SHF or at the transient co-expression between *Tbx1* and *Tbx5*. The overlap between the two programs at the heart field interface defines the sites of atrial septal structures and the muscular ventricular septum.

1.2 Epithelial mechanics

Because of the epithelial characteristics of cardiac progenitor cells, the formation of the heart can be considered to be a process of cardiac origami (Fig. 1.9) by which a 2-dimensional sheet is folded into a complex 3-dimentional structure. Heart morphogenesis can be driven by epithelial remodeling through a combination of cell and tissue-mechanical forces. This is why the next concept that we will talk about is epithelial mechanics.

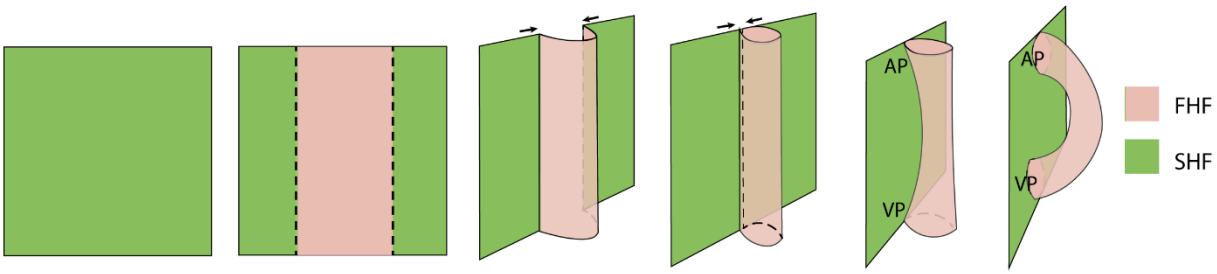


Figure 1.9. Cardiac origami: Scheme of how an epithelial sheet can give rise to the First and Second heart fields and be shaped into a primitive heart tube.

Using an orange mesh bag, Mike Schlaigh could design the TRUMPF pedestrian footbridge in Ditzingen. The bridge is a lightweight shell construction, which due to the high efficiency of the supporting structure is made of only 2cm thin stainless-steel sheets. The holes in the steel sheet indicate the flow of forces to the supports of the bridge. This type of design, beyond its undeniable plasticity, achieves great efficiency in terms of material requirements. It also makes clear how forces travel through the structure from the leads to the foundations. The size and density of the holes depends on the degree of utilization of the supporting structure (Fig. 1.10)

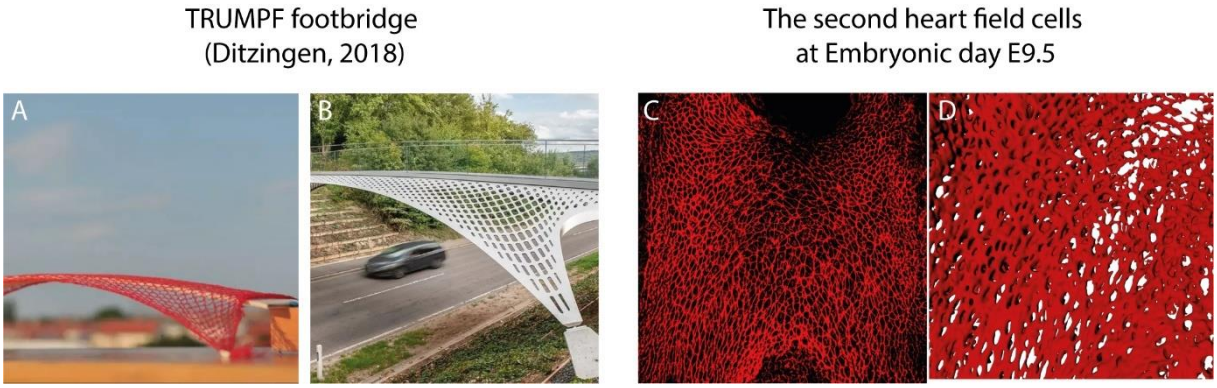


Figure 1.10. The transmission of forces. Example of the TRUMPF bridge and the dorsal pericardial wall. A. The prototype of the bridge used by the engineers **B**. The real bridge in Ditzingen **C.** The phalloidin channel of the SHF cells that contours the apical cell domain **D.** An image magnification of a 3D segmentation of the phalloidin channel.

1.2.1 Mechanics in heart development

Adult and mature cells found in fully differentiated tissues possess an architecture adapted to fulfill their specific functions (Ehler, 2016). In the case of cardiomyocytes, they have a rectangular shape with the sarcomere formed with actin and myosin filaments, rich in mitochondria and endoplasmic reticulum and connected with adjacent cells via intercalated discs. To acquire the shape and architecture best suited for their future role, cardiac cells need to undergo a series of changes during embryonic development and even after birth.

Some studies supporting this hypothesis have shown that after birth, the increase in mechanical performance requires a structural and functional maturation of the myocytes, with the conversion from the fetal to the adult structure of the sarcomeres (Ehler, 2016), a change in shape from polygonal to rod-like, and a sudden reduction of their proliferation with an increase in their volume. Mechanical maturation of the heart, finally, does not only involve maturation of the contractile cell mass. In fact, the cardiac matrix undergoes a significant stiffening from a value <0.2 kPa of the cardiac tube, to a value >20 kPa at birth (Majkut et al., 2013 ; Majkut et al., 2014), when it reaches its physiological stiffness. Interestingly, a coordinated control of cardiac myosin expression in myocytes and collagen I deposition by cardiac fibroblasts orchestrates maturation of the tissue, preparing it for the increase in mechanical load occurring at birth (Majkut et al., 2014). In parallel to the myocyte changes and the ECM stiffness it has been shown that sequestration of YAP due to Hippo pathway-dependent phosphorylation upon matrix stiffness acts as a negative feedback loop preventing further myocyte proliferation (Morikawa et al., 2017). Administration of drugs with a matrix softening activity prolongs the cardiac regeneration period in neonatal mice (Notari et al., 2018). These observations prompt us to consider that the mechanical properties of both the myocardial matrix and myocytes play a crucial role in enhancing heart regeneration during adulthood and regulating cardiac fates in embryonic development.

1.2.2 Epithelia and mechanical properties of cardiac progenitors: a biomechanical model regulating heart elongation

Epithelia play central roles in embryonic development. One example of this is the epithelium formed by the cardiac progenitor population, the Second heart field. The epithelial properties of early cardiac progenitors were first studied in the chick embryo (Linask et al., 2005). During early development, as the cardiac crescent forms, cardiac mesodermal cells acquire an epithelial state and will form part of the splanchnic epithelial wall of the coelomic cavity. As epithelial cells, they acquire apicobasal polarity: N-cadherin accumulates on the apical side and Integrin at the basal cell surface. N-cadherin, Fibronectin and cell-cell and cell-matrix adhesion are important for the initiation of myocardial differentiation.

The epithelial organization and polarity of early myocardial cells have been studied in more detail in Zebrafish (Campinho et al., 2013; Le A. Trinh & Stainier, 2004). Epithelial myocardial precursors move to the midline to form the heart tube. As these cells migrate medially they undergo epithelial maturation and change from a cuboidal to a more elongated shape. As the epithelium matures, the cells form apical tight junctions as shown by accumulation of the tight junction protein ZO-1, the cells form as well an apico-lateral domain where aPKCs accumulate and a basolateral membrane domain defined by β-catenin. During their maturation, these myocardial cells show dynamic expression of Fibronectin. Fibronectin deposition in this region is not crucial for cells to migrate medially but regulates the timing of migration. Moreover, Fibronectin is required for epithelial organization and junction formation and/or maintenance in cardiac precursors, as deletion of Fibronectin leads to abnormal formation or absence of junctions in the tissue. The transcription factor HAND2 is an important player in this process, and regulates epithelial organization by controlling the basal deposition of Fibronectin (Trinh et al., 2005; Garavito-Aguilar et al., 2010). Another study has demonstrated that aPKCι is required cell autonomously in zebrafish myocardial precursors for their epithelial organisation and coherence during heart tube formation (Rohr et al., 2006). In contrast, the molecular and epithelial properties of SHF progenitors in the DPW during heart tube elongation have received less attention.

SHF progenitors were first described as a cuboidal epithelium contiguous to the heart, using light and electronic microscopy (Virágh & Challice, 1973). Since then, further experiments have been addressed to study SHF cells: N-cadherin is important for SHF epithelial organization and prevents premature differentiation of these progenitors in mouse embryo (Soh et al., 2014), *Wnt5a*, an effector of the planar cell polarity (PCP) pathway, is important for the epithelial organization of SHF cells in the DPW (Sinha et al., 2015) while another PCP component, *Vangl2*, controls the epithelial organization of the distal OFT (Ramsbottom et al., 2014). It was not until 2014, that the epithelial properties of the DPW epithelium were characterized in detail (Francou et al., 2017). This study showed that cells in the DPW constitute an atypical apico-basally polarized epithelium with apical epithelial-like surface facing the pericardial cavity and a basal mesenchymal-like surface facing the underlying endoderm: Scribble is localized at the basolateral membrane, the tight junction protein ZO-1 accumulates at the apical cell-cell junctions in the DPW together with Par3 and PATJ, adherens junctions are present at the boundary between apical ($aPKc^{+}$) and basolateral ($Scribble^{+}$) membrane domains. SHF cells are characterized by an irregular basal lamina and a fibronectin-rich extracellular matrix (Fig. 1.11).

At the same time there is a spatial heterogeneity and maturation of the epithelial cells in the DPW cells: at E9.5 ZO-1, aPKC and Scribble are homogeneously distributed throughout the DPW. N-cadherin is expressed in the entire DPW with a gradient increasing towards the venous pole, whereas E-cadherin accumulates only at the anterior part (Fig. 1.11A).

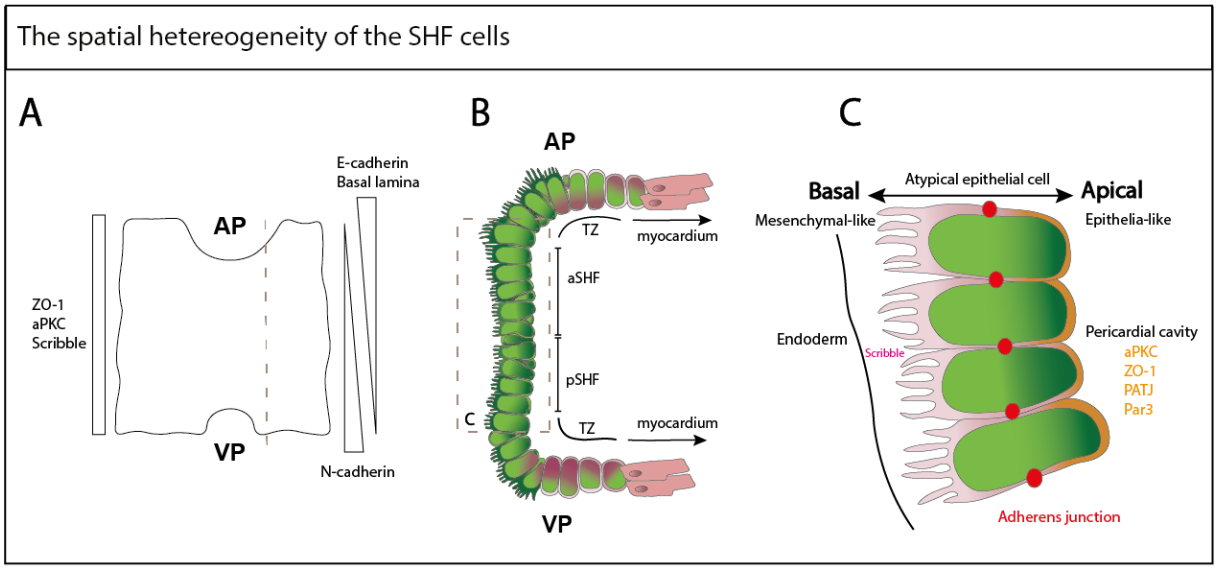


Figure 1.11. The spatial heterogeneity of SHF cells. A. Frontal view of the DPW with the spatial repartition of different proteins **B.** Sagittal view of the SHF in the DPW **C**. Hight magnification of the SHF cells showing their atypical nature and the apical to basal rearrangement of proteins. AP: arterial pole, VP: venous pole, TZ: transition zone, aSHF: anterior second heart field, pSHF: posterior second heart field.

It wasn't until 2017 that the SHF epithelium or Dorsal pericardial wall (DPW) was proposed to be implicated in biomechanical feedback promoting heart tube elongation (Francou et al., 2017). This study reported that the cells in the DPW are oriented towards the OFT, they show polarized actomyosin distribution and nuclear YAP in the posterior DPW (pDPW), indicative of the existence of a mechanical stress in the posterior region. Moreover, cell division and oriented cell growth align with the axis of cell orientation. These results suggested the involvement of epithelial tension in the regulation of heart tube extension (Fig. 1.12). The study proposed a biomechanical model consisting on a feedback of mechanical regulation of heart tube elongation driven by the elevated epithelial tension in the posterior SHF (Fig. 1.13).

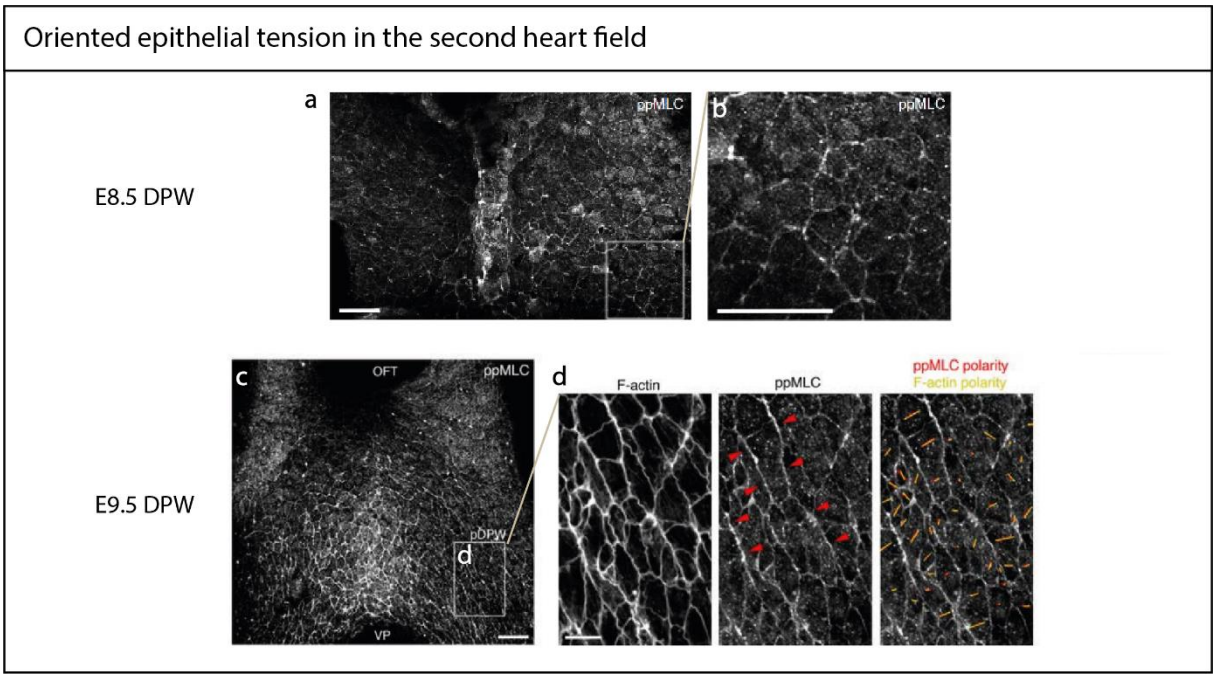


Figure 1.12. Epithelial tension in the DPW accompanies heart tube elongation, oriented epithelial tension appears in the posterior second heart field after heart tube looping. **A.** Whole-mount immunofluorescence of the apical surface view of the DPW at E8.5 showing ppMLC, a read-out of myosin activation. **B.** On the right, a high magnification of the homogeneous ppMLC localization on the membranes of the pSHF cells **C.** Wholemount immunofluorescence of the apical surface view of the DPW at E9.5 showing ppMLC **D.** High magnification of the pDPW at E9.5. of F-actinin and ppMLC where ppMLC accumulates in the elongated membranes of the cells. Red arrows show accumulation of actomyosin cables by ppMLC stainning. Red bars indicate nematic polarity of ppMLC, yello bars indicate nematic polarity of F-actinin.

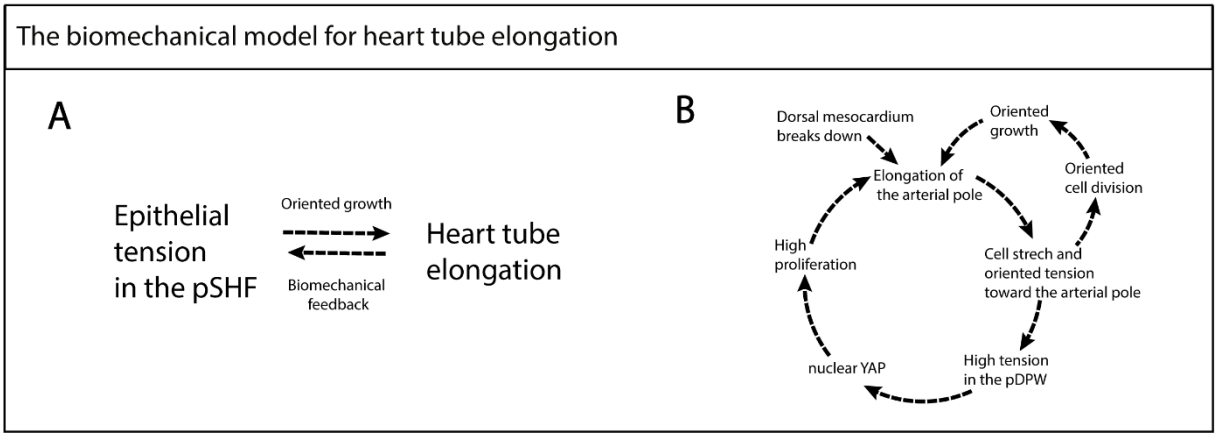


Figure 1.13. A biomechanical model driving heart tube elongation.

1.2.3 Mechanics regulating tissue morphogenesis and patterning

One of the animal models that have been most broadly used to study epithelial mechanics is Drosophila. Studies using this model have shown that the actomyosin network plays an important role in force generation and transmission triggering morphogenesis of epithelial tissues. Local actomyosin dynamics in individual cells are important for the generation and transfer of mechanical forces via cell-cell junctions to neighboring cells. The integration of the local forces into a global force pattern determines the resulting changes in tissue shape (Heisenberg & Bellaïche, 2013) (Fig. 1.14).

During mesoderm invagination, Drosophila cells accumulate apical actin-myosin to stabilize cell shape changes, leading to incremental reductions of the cell apical area. The pulsatile and isotropic actin-myosin flow is not associated with isotropic apical constriction but with constriction along the dorso-ventral axis and global anisotropic stress along the antero-posterior axis (Martin et al., 2010) (Fig. 1.14A).

Similar to this, gastrulation in *C. elegans* and in mice is also triggered by pulsatile, isotropic actin-myosin flows at the apex of the cells that are translated into apical cell constrictions (Roh-Johnson et al., 2012; Francou et al., 2023). Pulsatile anisotropic actin-myosin flows during epithelial elongation driven by cell intercalation, induce junction shortening. The resulting enrichment of acto-myosin at the junction stabilizes the reduction of the junction length and are indicative of epithelial tension (Rauzi et al., 2010 ; Sawyer et al., 2011) (Fig. 1.14B). The accumulation of actin-myosin at the elongated membranes of the cells in the posterior SHF cells appears to represent a similar process preventing cell deformation. Another example is the pulsatile basal myosin flow of the actin network of the base of the *Drosophila* follicular cells that produces anisotropic deformation and have been associated with the elongation of the egg chamber (He et al., 2010) (Fig. 1.14C).

Actin-myosin flows can also be non-pulsatile, as observed in the yolk cell of the gastrulating zebrafish embryo, in which they have been implicated in pulling the enveloping cell layer (EVL). Flow orientation is opposite to the direction of EVL spreading and is associated with the formation of a large actin-myosin cable-like structure located within the yolk cell at the margin of the EVL (Behrndt et al., 2012) (Fig. 1.14D).

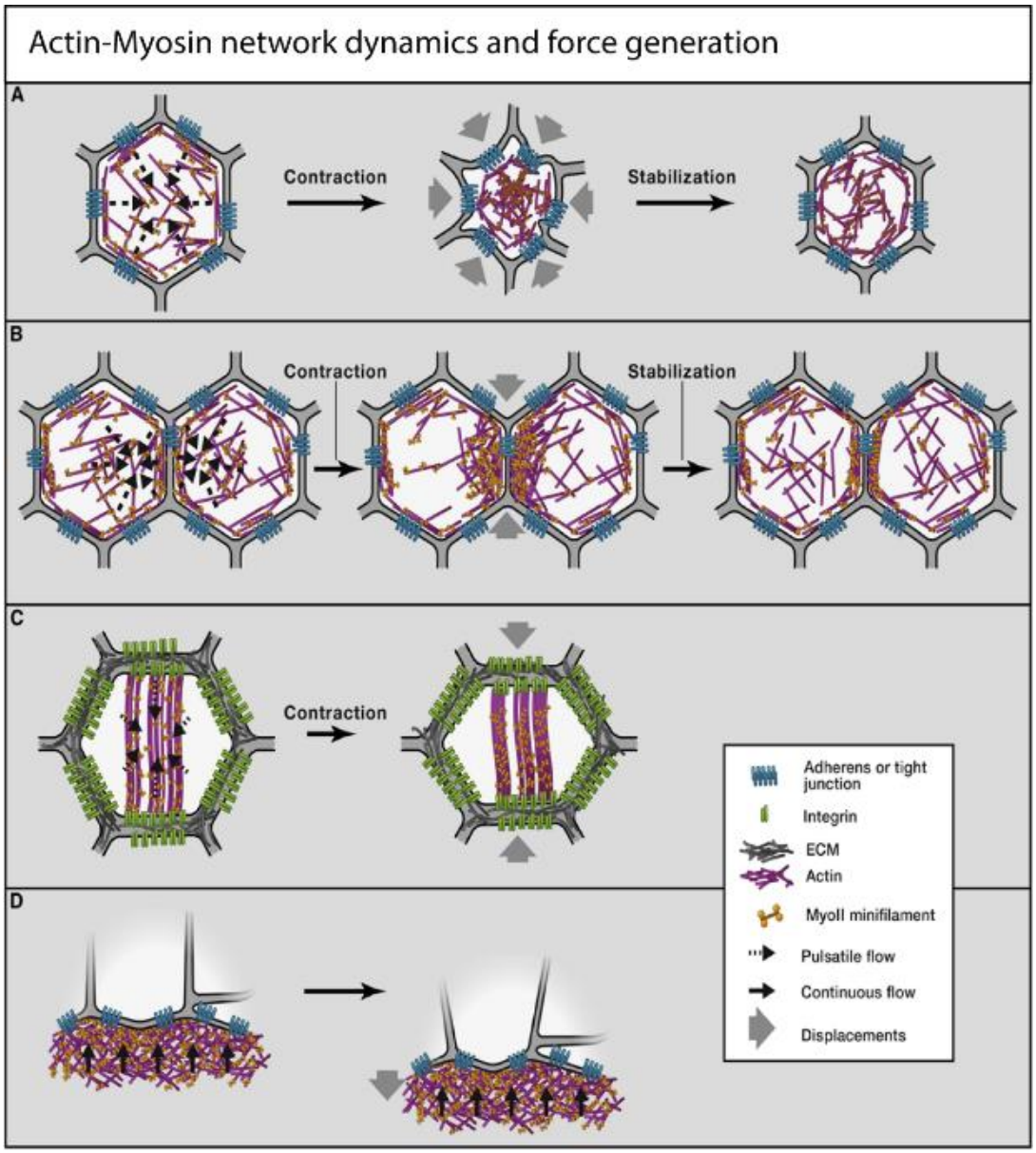


Figure 1.14: Actin-myosin network dynamics and force generation. Image from Forces in tissue morphogenesis and patterning (Heisenberg & Bellaïche, 2013).

Other embryonic processes regulated by mechanics are: (1) intercalation events such as T1 transitions and rosette formation, (2) the planar cell polarity (PCP) pathway which is the signaling pathway that organizes cells in the plane of the tissue, and (3) forces generated by external factors due to interaction with surrounding cells or the extracellular matrix. Here we will introduce these processes.

E-cadherin, armadillo and actin-myosin structures align across multiple cells during intercalation prior to rosette formation in *Drosophila* germband formation. E-cadherin and Armadillo are enriched in dorsal-ventral junctions while actin-myosin complex are enriched in anterior-posterior junctions (Blankenship et al., 2006; Zallen & Wieschaus, 2004). The tension created by the myosin induces further myosin recruitment, creating

supracellular myosin cables that are critical for boundary formation to prevent cell mixing (Landsberg et al., 2009; Monier et al., 2010 ; Aliee et al., 2012). Plus, there are certain studies that show that boundaries between tissue compartments are not only characterized by different gene expression patterns, but also by elevated interfacial tension (Hallou et al., 2023).

Live imaging experiments in the germband revealed that certain cells form vertices at which 5 or more cells meet. These specific cell rearrangements are called multicellular rosettes and indicate a higher-level organization of cell behavior. Rosette formation and resolution proceed in a directional fashion that parallels structural changes at the tissue level. The tracking of this type of intercalation shows that an array of cells that was elongated along the D-V axis can rearrange to form an array elongated along the A-P axis. F-actin enrichment at antero-posterior interfaces is the earliest evidence of these process and the initiation of planar cell polarity (PCP). Subsequently, Myosin II accumulates at antero-posterior interfaces and may coordinate the contraction of linked edges to drive rosette formation. F-actin and Myosin II colocalize at the vertex where 5 or more cells meet. Rosettes resolve in a directional fashion to create contact between cells that were previously separated along the dorsal-ventral axis. Next, E-cadherin association is an early step in new contact formation and coincides with formation of a transient F-actin structure. Finally, Bazooka is recruited to new interfaces after a delay, which may create local differences in adhesion that are important for rosette resolution (Blankenship et al., 2006).

Signaling pathways determining PCP, such as the Wnt/Frizzled and Fat/Dachsous pathways, also play roles in cell intercalation by polarizing subcellular distribution of myosins (Goodrich & Strutt, 2011; Gray et al., 2011) . It has already been shown that this signaling pathway controls neural plate folding and convergent extension movements simultaneously (D. Li et al., 2019). Convergent extension is a conserved morphogenetic event promoting the elongation of the body axis, consisting of the narrowing of the tissue in one dimension while it lengthens in the perpendicular axis due to oriented cell movements (Keller et al., 2000; Wallingford et al., 2002 ; Solnica-Krezel, 2005).

If we take a closer look at the process of epithelial remodeling during neural tube closure (Shparberg et al., 2019), we can see that it has similarities with the process of dorsal mesocardium break down during heart development, with the difference that the heart tube remains attached to the dorsal pericardial wall at the cardiac poles (Fig.). As an example, neural tube closure involves (1) neuroepithelial cell intercalations associated with convergent extension movement of the neural plate and (2) bending of the neural plate along its antero-posterior axis, which is driven by the coordinated apical constriction of neural plate cells close to the neural plate midline, dependent on the activity Shroom3, recruiting Rho-kinase (ROCK), and thereby activating of MyoII at the apex of these cells (Shparberg et al., 2019 ; Hildebrand & Soriano, 1999 ; Haigo et al., 2003). Apical constriction of neural plate cells is anisotropic, and this anisotropic

constriction is, in principle, sufficient to explain both oriented cell intercalation driving convergent extension movements and polarized bending of the neural plate (Nishimura & Takeichi, 2008) (Fig. 1.15).

Another example where cells show planar polarization is during the development of wings in Drosophila. The polarization is mediated by the localization of proteins such as Frizzled to the distal or proximal sides of the wing blade cells. The proximal-distal localization of these proteins is the result of an initial planar polarization of the wing cells toward the wing margin and subsequent rearrangement of the wing cells through anisotropic tension along the proximal-distal axis of the tissue (Aigouy et al., 2010; Sagner et al., 2012).

Cells also perceive extrinsic mechanical forces that can lead to both changes in adhesive and cytoskeletal organization on a short timescale and in fate specification and differentiation on a long timescale (Eyckmans et al., 2011). The force generated by the interaction between cells and the extracellular matrix (ECM) can play important roles in cell proliferation, fate specification and lineage commitment. It is clearly established that stiffness transduced by integrins, transmembrane receptors that help cell adhesion between cells and with the extracellular matrix, can have a major impact on the specification of different stem cells (Discher et al., 2009; Parsons et al., 2010) .

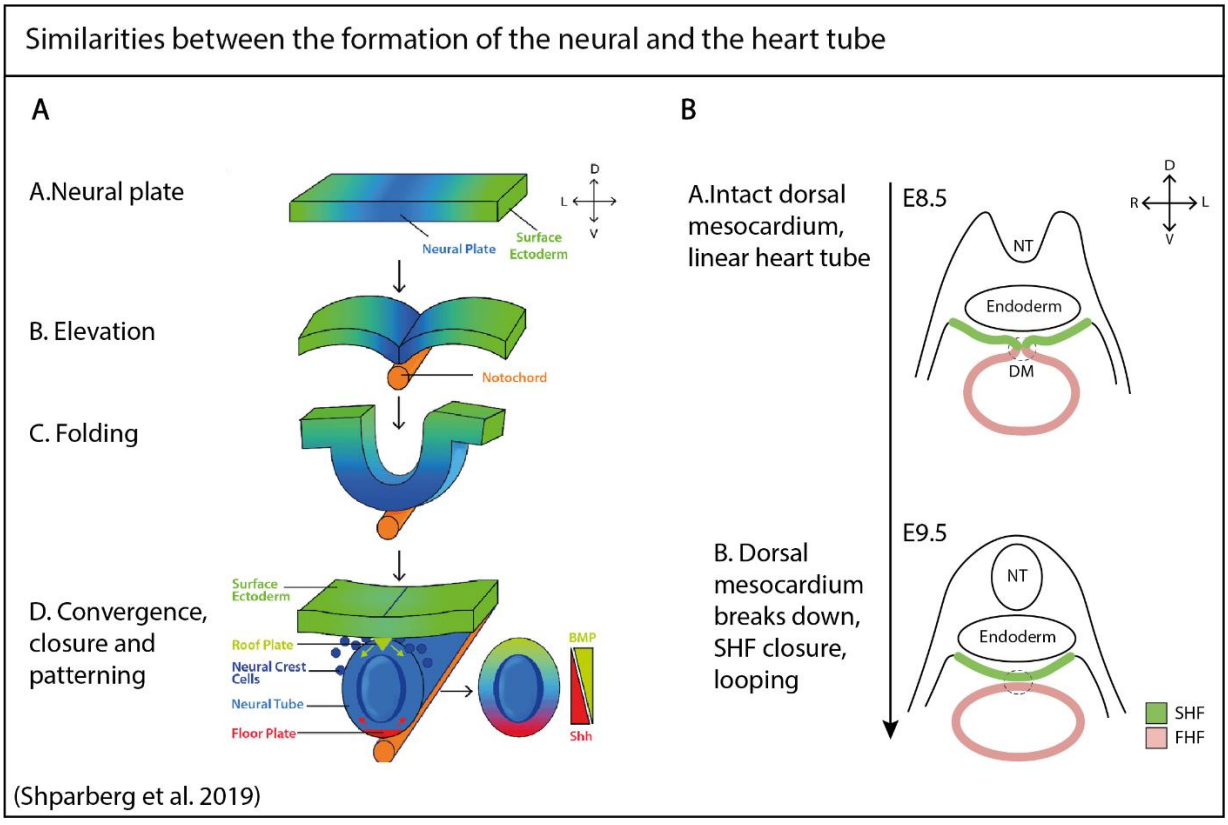


Figure 1.15. **Similarities between the formation of the neural and the heart tube. A. Development of the neural tube**. Figure adapted from (Shparberg et al., 2019). **B.**

Formation of the linear heart tube. V: ventral; D: dorsal; L: left; R: right, NT: neural tube, DM: dorsal mesocardium.

1.2.4 Mechanical forces regulating cell division and proliferation

Mechanical forces have been shown to control the axis of cell division in isolated cells in culture and in developing embryos in vivo (Fink et al., 2011; Campinho et al., 2013). Several studies have provided insights into the mechanisms by which mechanical tension and strain are linked to proliferation and tissue growth (Dupont et al., 2011; Aragona et al., 2013 ; Rauskolb et al., 2014). Two transcriptional co-activators, YAP/TAZ, act as sensors of mechanical cues and mediate this information at the transcriptional level. Applying mechanical stretch to epithelial cells in culture activates YAP/TAZ that on translocation to the nucleus lead to an increase of proliferation. ECM stiffness (Engler et al., 2006), cell geometry (Chen et al., 1997) or cytoskeletal tension (Jaalouk & Lammerding, 2009; Schwartz, 2010) change YAP/TAZ distribution.

The Drosophila wing imaginal disc is an example of how forces can influence cell division and growth through mechanical feedback. The hypothesis is that the differential growth of the tissue is the result of a morphogen gradient. The pouch is patterned by morphogen gradients that establish a Cartesian coordinate system along the anterior-posterior (AP) and dorsal-ventral (DV) axes. As a result of the morphogen gradient the cells in the center have a high growth and become compressed, while the cells in the periphery with a lower growth become stretched. The mechanical stretching of the cell, through the mechanical feedback, lead to an induction of cell growth and proliferation while compression would induce growth arrest (Aragona et al., 2013; Mao & Baum, 2015 ; Mao et al., 2013) (Fig. 1.16A). One of the signaling pathways regulating this hypothesis in the Hippo signaling pathway, where the translocation of the transcriptional coactivator of YAP to the nucleus changes depending on the cell's mechanical environment. According to this hypothesis , stretched cells thus accumulate more nuclear YAP, which leads to the transcription of cell growth and cell differentiation target genes, whereas cell compression would stop cell growth, that could explain the phenomenon of contact inhibition of growth (Aragona et al., 2013 ; Parker & Struhl, 2020). More recent studies in the wing pouch (Dye et al., 2021) have shown that cell morphology patterns can persist and strengthen in the absence of differential growth. The spatial maps of cell morphology reveal that both cell area and cell elongation magnitude are largest at the periphery and decrease toward the center and show that they increase in time in the absence of differential growth. In the wing pouch the radially oriented cell rearrangements balance the tangential cell elongation (Fig. 1.16B).

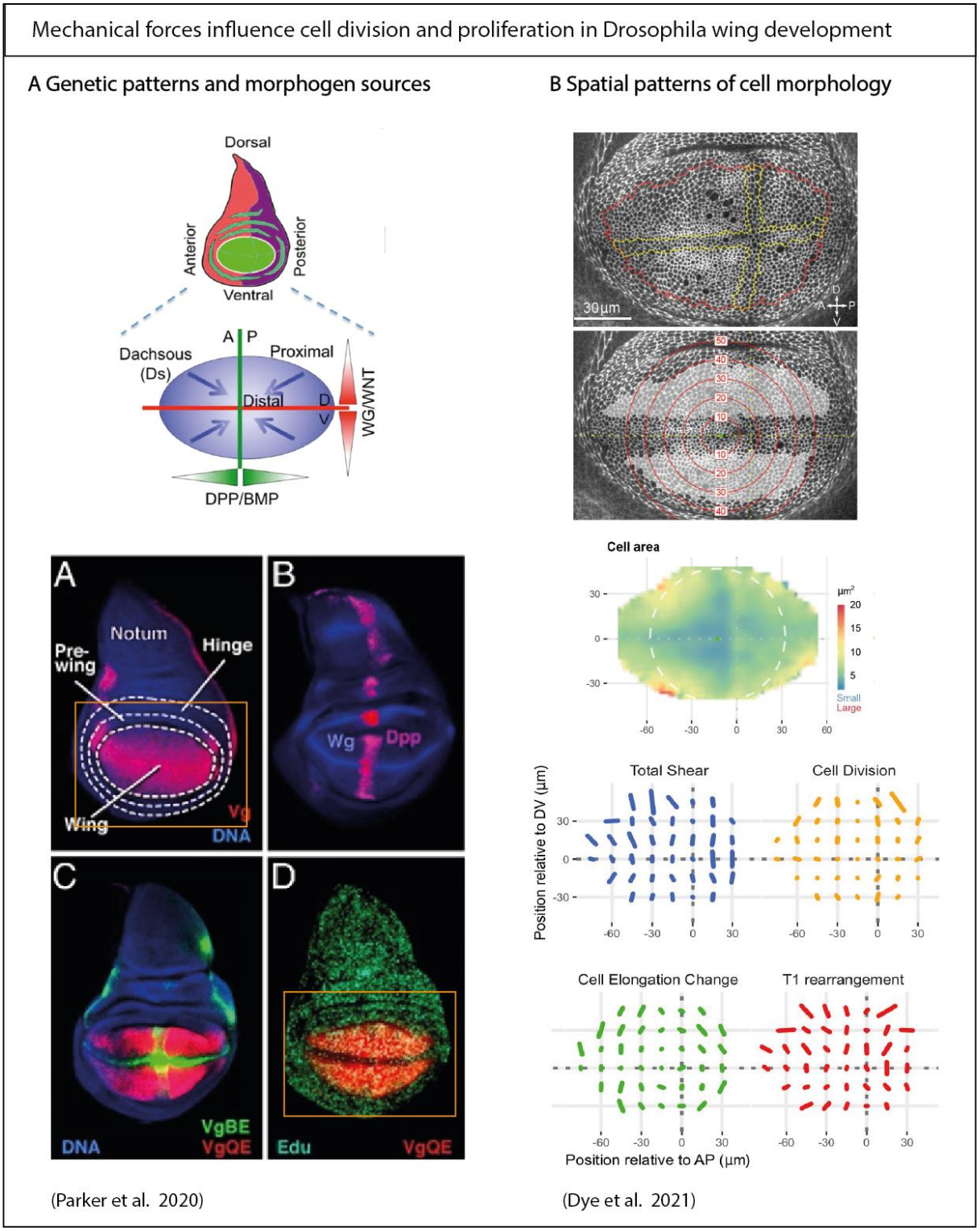


Figure 1.16: Mechanical forces influence cell division and proliferation A. Control of wing growth by Dpp and Wg A. The wing imaginal disc as a model system for studying organ growth control. Morphology, morphogen sources, and proliferative growth of the wing primordium. Figure from (Parker & Struhl, 2020) **B.** Spatial patterns of cell morphology. Figure from (Dye et al., 2021).

1.2.5 Bayesian force inference

To understand the development of the morphology of animals it is essential to analyze how mechanical forces coordinate behaviors of individual cell that underlie tissue deformation. This is why the quantification of mechanical forces and stress is necessary to unravel the physical principles of morphogenesis. Among the already existing methods to measure in vivo mechanics, we find for example laser ablation techniques that evaluate the tension acting on a cell contact surface. However, this approach is invasive and only sheds light on mechanical features at a cellular level. Moreover, it is applicable only to readily imageable epithelia that can be directly accessed by lasers. A complementary approach to the subcellular and invasive approach of laser ablation is force inference, an image-based approach the calculates mechanical features from cell shape and connectivity. This method assumes that the forces acting on epithelial cells are balanced at each vertex (Fig. 1.17A, B). By employing Bayesian statistics, in 2012 a group developed a novel framework to estimate the pressure of each cell and the tension of each contact surface from the observed geometry of the cells (Ishihara & Sugimura, 2012). This approach exemplifies an inference problem, or inverse problem: the pressures and tensions are inferred from the final result or observation.

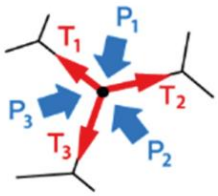
Force inference allows the creation of maps of force and stress in tissue morphogenesis and multi-cellular pattern formation. The method consists of 4 steps: Step1: Acquisition of the image of the epithelia using a membrane marker of the cells. Step2: Image segmentation using the membrane marker. Step3: Acquisition of cell shape and the cell connectivity graph. Step 4: Inference of the edge tensions, cell pressure and tissue stress in MATLAB (Fig. 1.17C). Among the advantages of this method are: it works at single cell resolution, it can analyze more than 10000 cells and it is non-invasive (Ishihara & Sugimura, 2012). Furthermore, it has already been shown to produce consistent data with those obtained using in vivo methods such as laser ablation of cortical actin cables, myosin concentration, photo-elasticity and large-scale tissue ablation (Sugimura et al., 2013).

In 2019, this method has been democratized for a broader public with less physical and coding expertise by the creation of a MATLAB plug-in of open use that has also been used to measure forces within an epithelial layer and produce consistent quantifications with those obtained in in-vivo experiments (Kong et al., 2019). This method has been updated by the developers of the MATLAB plug-in, Wei Kong and Raphael Clement, in 2020, allowing to quantify force inference data at single cell level (Fig. 1.17D). Bayesian force inference works perfectly for epithelial tissues which are relatively flat, however it cannot be used for highly curved tissues and the deflection height of the tissues compared with its size needs to be calculated, ratios below 20-30% are appropriate to be used by Force Inference. It is also important to show that the curvature of the tissue is larger than the radius of a cell, meaning that the curvature of the tissue will have negligible effects on the measurements inferred by the force field. However, many biological systems present important curvature, which is why currently groups are working in developing tools to

measure forces in curved surfaces that could be used for organoids, gastruloids, or the heart tube itself. Due to the sphericity of the early *C.elegans* and *Drosophila* embryos and the first steps of mouse embryonic development, teams have been working on the development of new techniques to infer cellular forces from fluorescence microscopy images of cell membranes. One of the methods is called foambryo which generates 3D meshes of cells' geometry to predict cell surface tensions and pressures (Ichbiah et al., 2023).

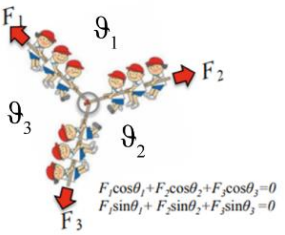
Force inference is based on the assumption that forces are balanced at each vertex

A

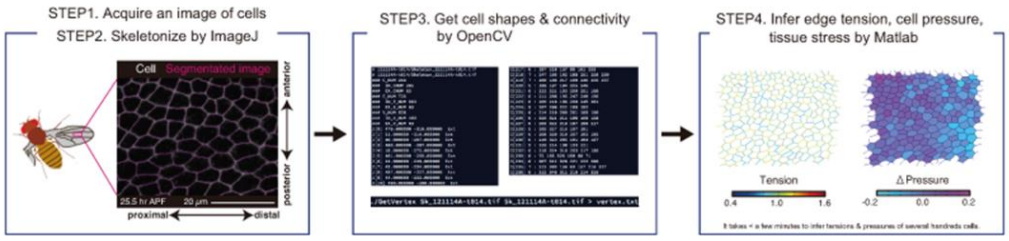


Tension on the cell boundaries (T)
Pressure for each cell (P)

B



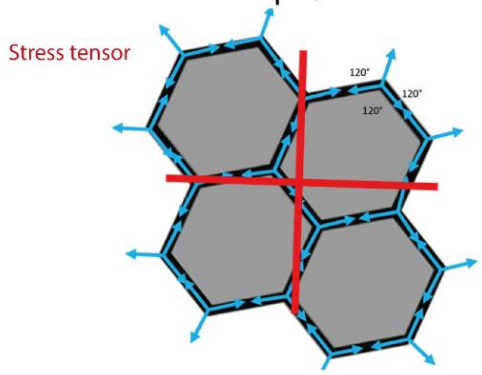
C



(Ishihara et al. 2012)

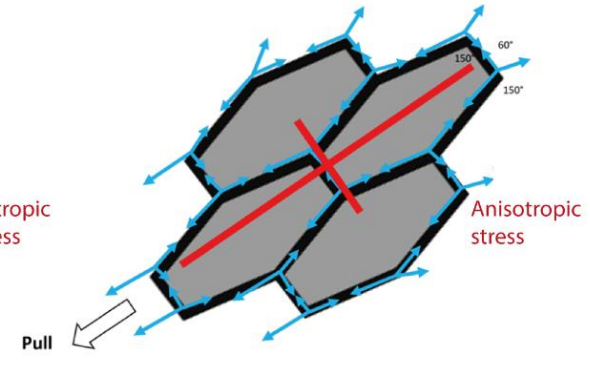
D Outputs of force inference

Relaxed shape / Balanced stress

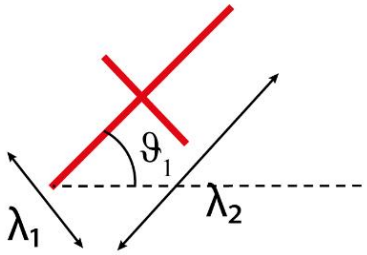


Isotropic stress

Stressed shape



Anisotropic stress



Force inference gives you information about:

1. Principal direction of the stress
2. Amplitudes of the stress
3. Stress anisotropy
4. Pressures

(Kong et al. 2019)

Figure 1.17. A. Force inference assumes that the forces are balanced at each vertex. In red, the tensions of the cell boundaries (T); in blue, the pressure for each cell (P) **B.** Graphical representation of tug-of-war as an example of how forces are balanced at the vertex, if one rope pulls stronger, the force (F) in that rope will be imbalanced and the angles (ϑ) will change. **C.** Scheme of the workflow of force inference. Transformation of a biological image into 2D representation of the epithelium with quantitative data. **D**. Outputs of force inference: On the left, an example of a relaxed epithelia with balanced stress tensor/or isotropic stress; on the right, an example of a stressed shape with an anisotropic stress tensor; on the bottom a representation of a stress tensor and its properties: the principal stress tensor (ϑ_1), the first amplitude of the stress tensor (λ_1), the second amplitude (λ_2), the stress anisotropy (λ_1/ λ_2). In grey, the cell mesh; in blue, the tension at each vertex; in red, the stress tensor.

1.3 Sc-morphometrics and machine learning

The next section introduces the reader to the key points in machine learning that have been used to analyze the data in this thesis. Much as a microscope is a tool to help see things at microscopic scale, machine learning is a tool that can help to reveal and predict patterns in datasets. Already available methods in machine learning have been used as a tool to efficiently analyze images, for example in tasks such as pattern recognition, or in molecular biology to efficiently analyze single-cell RNA seq datasets (Villoutreix, 2021).

1.3.1 Morphometrics

Morphometrics refers to the quantitative analysis of form, which involves size and shape. Size is the physical scale of an object, usually determined by comparing one or more of its spatial dimensions to a reference which serves as a basis for measurement, while shape is the component of form which remains after accounting for differences in size, position and rotation between two or more objects ("Morphometrics," 2017). Bookstein defined morphometrics as the quantitative analysis of form and covariances with form (Bookstein, 1992).

Morphometrics has a long history throughout many areas of biology. The first use of morphometric approaches can be tracked back to the origins of illustrative art, when the accuracy of the description of human and animals was possible thanks to the mathematization of the principles of linear perspective. Perspective is the geometrical procedure in which the angles and lengths of lines on a 2D surface can be adjusted to give the illusion of a 3D space. The perspective procedure was first used by Da Vinci and Albrecht Düer in their geometric studies of the human body: to identify the "ideal" or average proportions of the human body and to express the deviations from this "ideal" via a unified system of geometric transformations that included isotropic or uniform

transformations and anisotropic or non-uniform transformations in all the directions (Fig. 1.18 1A-C). Derived from these artistic foundations, the mathematician D'Arcy Thompson, used morphometric analysis to explain biological shape variations, and discovered that many morphological transformations could be produced with the application of geometric deformations to all regions of the body (Thomson, 1917) (Fig. 1.18 1D). While these procedures were graphical constructs, later mathematical procedures were deployed into computational tools like trend-surface analysis (Sneath1, 1967), bio-orthogonal grids (Bookstein, 1986) and thin plate splines (Bookstein, 1989).

At the same time as the geometrical approaches progressed, statistical approaches in the field of social sciences were developed to characterize the "average man" by measuring different variables that follow a normal distribution, these measurements were later on summarized graphically by plotting their values in a Cartesian space which was then used to interpret the statistical similarities and differences between different individuals (Quetelet, 1835). Other studies followed such as the development of principal component analysis (PCA) by Karl Pearson whose aim was to reduce the dimensionality of multivariate data whilst preserving as much of the relevant information as possible (Bro & K. Smilde, 2014; Jolliffe & Cadima, 2016) or linear discriminant analysis (Jombart et al., 2010) first elaborated by Robert A. Fisher whose purpose is to find a linear combination of features that separates 2 or more classes of objects. He developed statistical analysis using a dataset of Edgar Anderson's including 4 measurements: sepal length, sepal width, petal length and petal width of 3 species of 3 kinds of *Iris* flowers *setosa, versicolor and virginica* (Fig. 1.18 2).

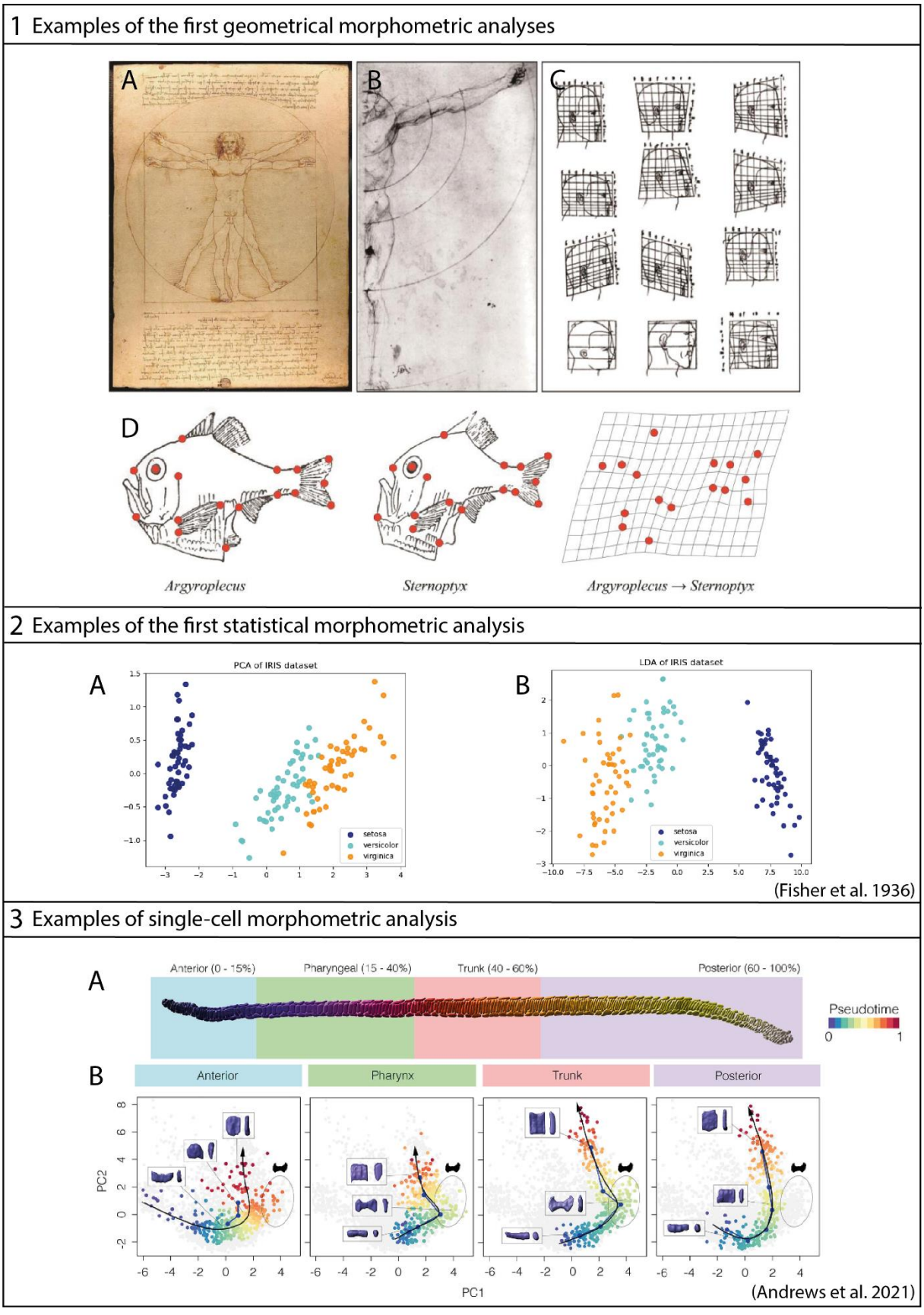


Figure 1.18: 1. Examples of the first geometrical morphometric analysis 1A. The Vitruvian Man, a graphical analysis of human proportions by Leonardo DaVinci **1B.**

Graphical analysis of average arm proportions relative to the body by Albrecht's Düer **1C.** Human facial proportions due to isotropic and anisotropic deformations by Albrecht's Düer **1D.** Set of landmarks collected from 2 fish illustrations from D'Arcy Thompson and used to calculate the deformation of the form by using Bookstein interpolation method of the thin plate spline (TPS) **2. Examples of the first statistical morphometric analysis. 2A.** Principal Component Analysis (PCA) applied to the IRIS dataset **2B.** Linear Discriminant Analysis (LDA) applied to the IRIS dataset **3. Example of a single-cell morphometric analysis 3A.** Segmented 14ss notochord from amphioxus showing exampled regions on the antero-posterior axis for trajectory comparison **3B.** PCA projection of the cells from each region defined in figure 3A, with segmented line connecting the point cloud centers for each stage colored in blue, the inferred pseudo-temporal trajectory colored in black.

Morphometric studies employ objective, quantitative techniques that are preferably computer assisted. With the increase of studies in quantitative methods and the advances in computer graphics, machine learning and artificial intelligence, morphometrics studies are more standardized and accurate and are set to reinvigorate the field again. For example, in the field of cell biology with improvements in microscopy and image analysis there is the possibility of the calculation of features at a single-cell resolution. A good example of this is the single-cell morphometric pipeline used to quantify the geometrical properties of the amphioxus notochord cells, which helped to reveal the interplay between cell shape changes and growth contributing to the full elongation of the notochord length during amphioxus development (Andrews et al., 2021) (Fig. 1.18 3).

1.3.2 Machine learning for data analysis

Artificial intelligence includes machine learning, which in its turn includes deep learning. Artificial intelligence is the general field of computer science that aims to develop intelligent machines. Machine learning concerns the tasks of making accurate predictions from very large datasets with minimal human intervention. Deep learning is a family of models that has been extremely successful at various tasks of machine learning, specially involving images. In its turn, machine learning can be subdivided into supervised, unsupervised and reinforcement learning. In the supervised learning, the points are labelled and a decision criterion is learnt using these labels. In unsupervised learning, the data points are unlabeled and the clusters are learnt from relationships among the points (Fig. 1.19). In reinforcement learning setting the computer interacts with its environment to gain information through reward, and update its state accordingly.

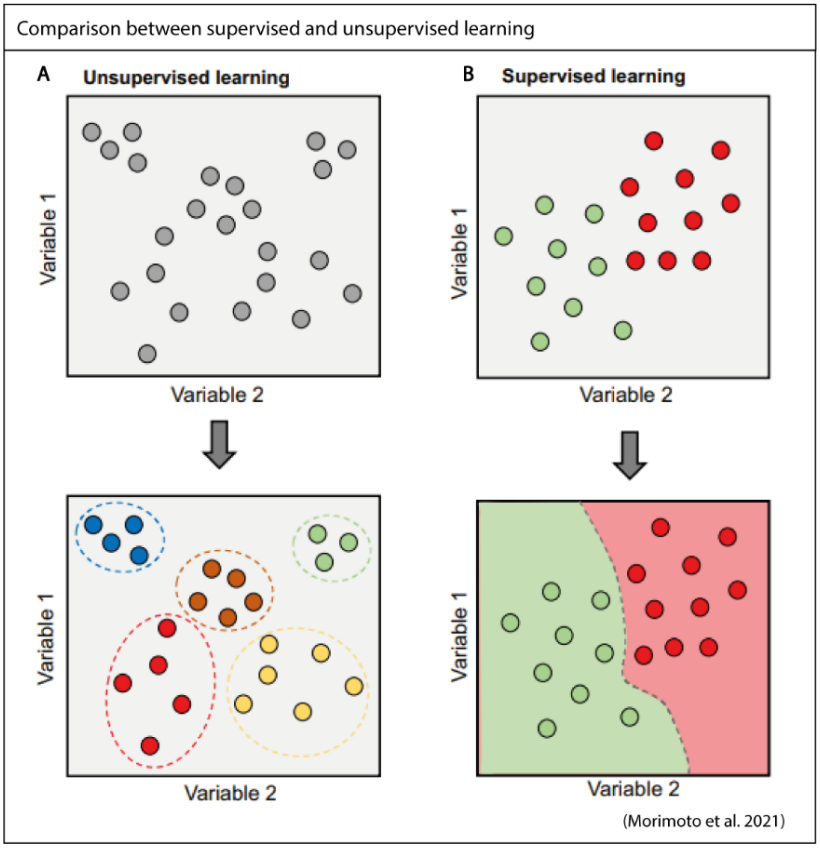


Figure 1.19: Comparison between unsupervised learning (a) and supervised learning tasks (b). Image from (Morimoto & Ponton, 2021).

1.3.2.1 Supervised machine learning

Supervised machine learning involves using labeled datasets to train algorithms for classification or prediction tasks. The two primary tasks in supervised learning are classification and regression. Classification algorithms predict discrete labels, while regression algorithms predict continuous values. Supervised machine learning algorithms are framed as optimization problems, where the goal is to develop a model or function and optimize its parameters to minimize the cost function, or the error between actual and predicted labels. This optimization process enables the model to learn from the data and improve its performance.

To create a supervised model, you first need a dataset with labeled data, referred to as the target. The data is then typically split into training and testing sets in an 80%-20% ratio. Next, you select an appropriate algorithm for the data, train the model, and evaluate its performance using metrics such as accuracy, F1-score, recall, or precision for classification tasks, or the R2 score for regression tasks. Finally, you adjust the model's parameters to enhance its performance.

1.3.2.2 Unsupervised machine learning

Unsupervised learning uses unlabeled data to discover patterns in the datasets. Unsupervised machine learning algorithms includes clustering and dimensionality reduction algorithms. Clustering determines grouping with unlabeled data. It mainly depends on the specific user and the scenario. It has been argued the high number of algorithms could be due to the notion that "cluster" cannot be precisely defined farther than a group of data objects (Estivill-Castro, 2002). A priori, there is no objectively correct clustering algorithm. The most appropriate clustering method for one kind of data will fail on a dataset that contains a radically different kind of data, this is why prior to clustering, manifold learning algorithms are a good way to understand the underlying variables. Interestingly clustering analysis was originated to help anthropology studies to separate ethnic groups according to their cultural traits (Driver & Kroeber, 1932) and later was used in psychology studies to separate people according to their personality traits (Tryon, 1939, Cattell, 1943). Nowadays it is broadly used in many areas including computer graphics, data mining, pattern recognition, image analysis, information retrieval disciplines. In the field of biology, sc-RNAseq analysis, for example, is based on a clustering analysis that aims to group cells according to their gene expression profiles.

There are different types of clustering algorithms (Fig. 1.20):
Connectivity models: Builds models based on distance connectivity, such as hierarchical clustering (Murtagh & Contreras, 2012).
Centroid models: Represents each cluster with a single mean vector, like K-means clustering (Ikotun et al., 2023).
Distribution models: Clusters are modeled using statistical distributions, like gaussian mixture clustering, in this particular case, we assume that the data comprises Gaussian distributions (Weber et al., 2022).
Density models: Define clusters as a connected dense region in data space, like DBSCAN (Deng, 2020) that aims to cluster the data as contiguous regions having high point densities or OPTICS (Ankerst et al., 1999).
Graph-based models: The data of this clustering can be presented as a graph where each element to be clustered is represented as a node and the distance between two elements is modeled by a certain weight on the edge linking the nodes. In graph clustering, elements within a cluster are connected to each other but have no connection to elements outside that cluster (Roy & Chakrabarti, 2017).

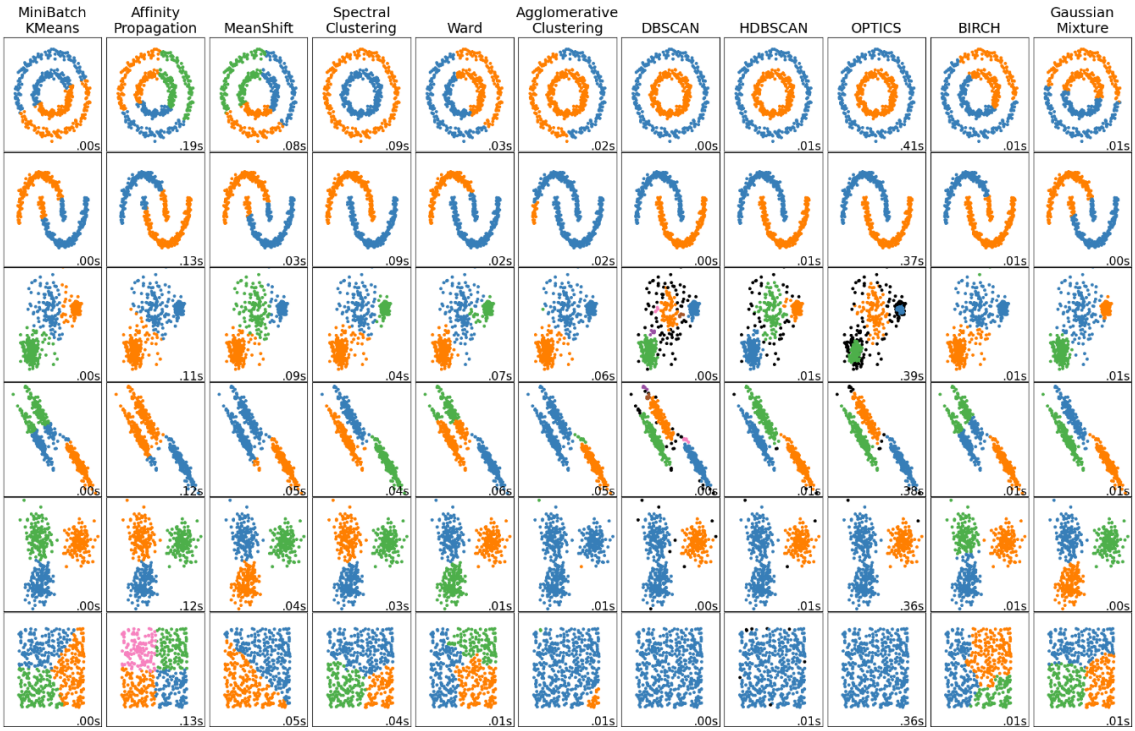


Figure 1.20: Comparison of different clustering algorithms on toy datasets. The rows illustrate the different biological datasets, columns represent the different visualization method applied for each dataset. The data is color coded according to the cluster labels (blue for cluster1, orange for cluster 2, green for cluster 3, pink for cluster 4, black for outliers) obtained depending on the different clustering method used. The figure illustrates that the choice of clustering method depends on the shape of your data, as certain methods are better suited to identifying different types of clusters. Additionally, methods like DBSCAN, HDBSCAN, and OPTICS are particularly suited for detecting outliers. Image from scikit-learn website by Pedregosa et al., 2011.

Among all the existing clustering methods, for our purposes and type of data, we have used the K-means, or centroid-based clustering (Ikotun et al., 2023) and Gaussian mixture (Weber et al., 2022). K-means is solved as an optimization problem: find the K-centroids and assign the objects to nearest centroid such that the distances from the cluster are minimized (Fig. 1.21). K-means clustering starts by specifying a number of clusters or k-centroids (k), randomly defining k-centroids. Then it generates clusters by iteratively running the next 2 tasks: 1) Each data point is assigned to a cluster in such a manner that it is closer to its own cluster center than any other cluster center) For each cluster, the arithmetic mean of all the points belonging to the cluster is calculated, giving a new cluster center. The algorithm stops when the sum of the squared Euclidean distances for all data points is minimized. Clustering and classification algorithms are commonly used in data analysis pipelines to extract information from great amount of data. The difference between K-means and Gaussian mixture models is that the first one is a hard-clustering method, which means that it will associate each point to only one cluster, the limitation is that there is no probability that tells you how much a data point is associated with a

specific cluster. While Gaussian mixture is a soft clustering method, which means that it determines the probability that a given data point belongs to a cluster (Pedregosa et al., 2011). Gaussian mixture is a distribution model composed of several Gaussians or normal distributions. It assumes that all data points are generated from a mixture of a finite number of normal distributions. Mixture models can be seen as an extension of K-means clustering, incorporating information about the covariance structure of the data as well as the center of the latent Gaussians (Fig. 1.22).

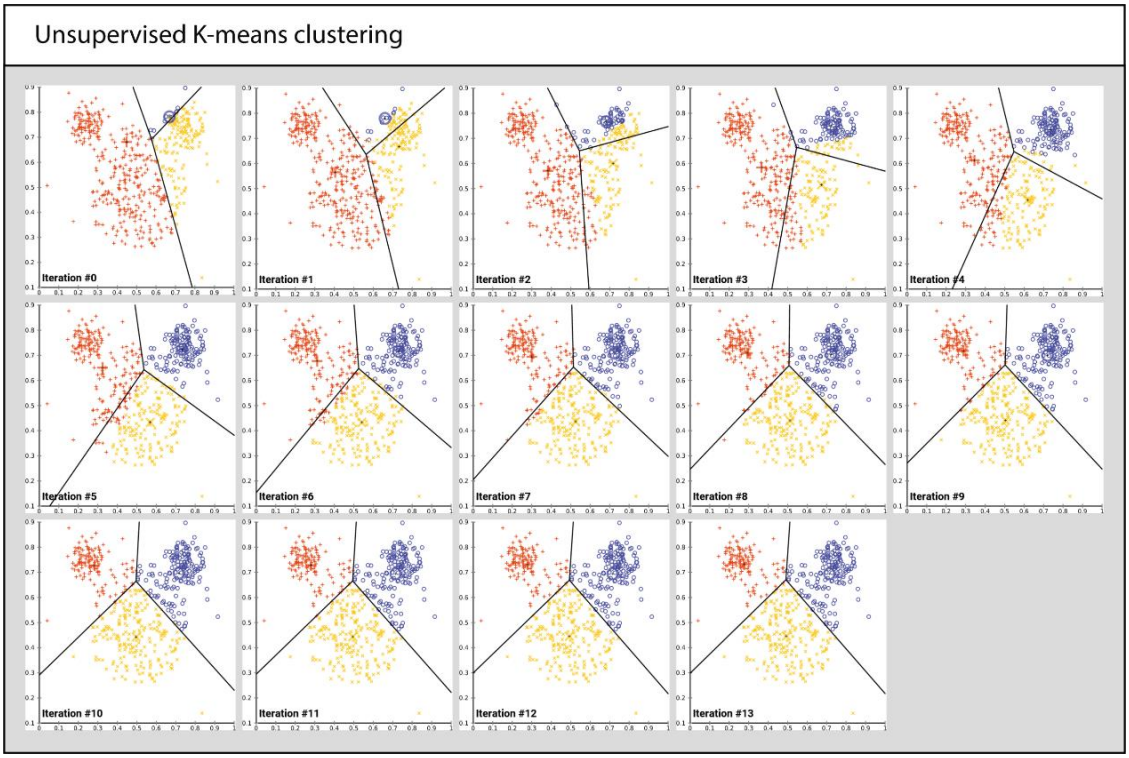


Figure 1.21: Unsupervised K-means clustering. Visualization on the evolution of the clustering partitions of a dataset in 13 iterations. K-means separates data into Voronoi cells, which assumes equal-sized clusters.

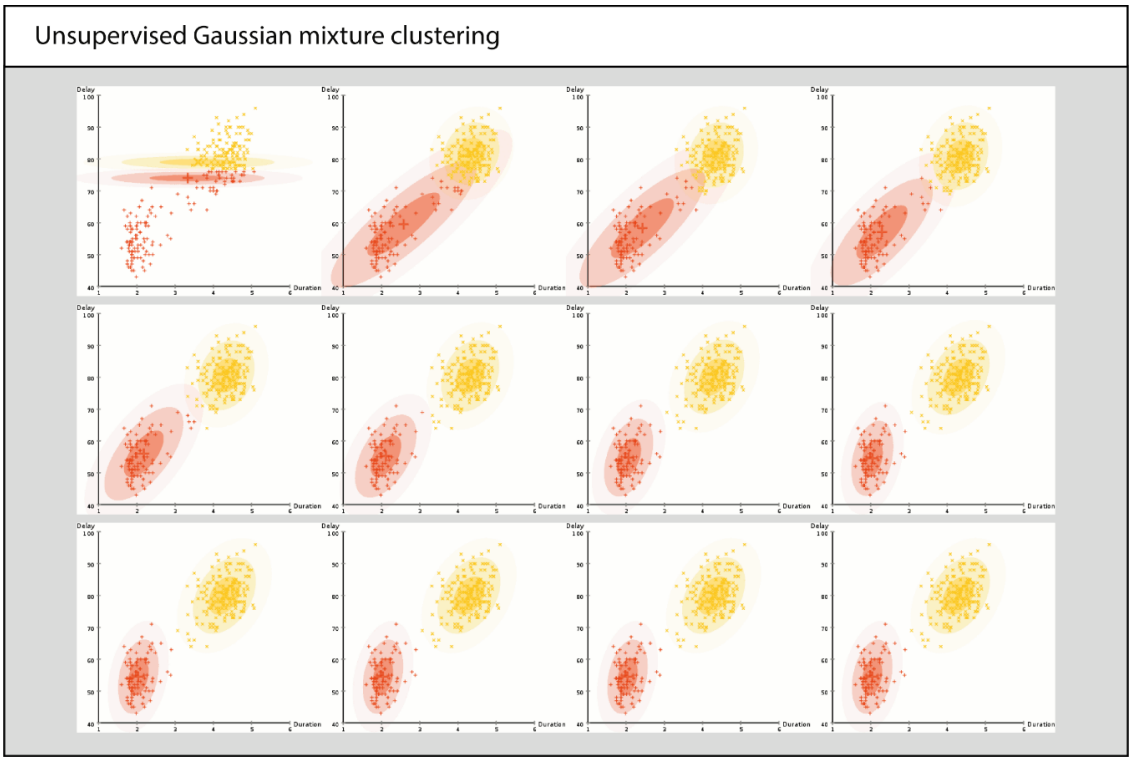


Figure 1.22: Unsupervised Gaussian mixture clustering. Visualization on the evolution of the clustering partitions of a dataset using Gaussian mixture clustering.

Dimensionality-reduction methods are also included among the unsupervised learning methods. One of their goals, among others, is to simplify the data by removing redundant or irrelevant features. This functions as a feature engineering technique, reducing data noise, increasing computational efficiency and helping in visualization (Haghverdi et al., 2016). These methods have gained importance in order to present the information in a visual manner, to understand the shape and structure of high dimensional data (Moon et al., 2019). The most commonly used dimensionality reduction methods are principal component analysis (PCA) (Bro & K. Smilde, 2014; Jolliffe & Cadima, 2016), Uniform Manifold Approximation and Projection for Dimension Reduction or UMAP (Becht et al., 2019) and t-distributes stochastic neighbor embedding or t-SNE (Maaten & Hinton, 2008). PCA is a linear dimensionality reduction method that aims at finding the orthogonal axes that explain the maximum variance in the data, it captures the global patterns and preserves the structure of the data, by reducing its dimensions and its noise, it provides interpretable results and it is a computationally efficient algorithm. On the other hand, UMAP and t-SNA are non-linear techniques preserving both, global and local structures in the data, mainly use for data visualization and exploration purposes, their outcomes are not directly interpretable and they are based on time-consuming algorithms (Table 1.2).

	PCA	t-SNE	UMAP
Linearity	Linear	Non-linear	Non-linear
Preserved structure	Preservation of global structure	Preservation of global structure	Preservation of global structure
Efficiency	Analytical solution	Optimization problem	Optimization problem
Scalability	Less scalable	More scalable	Less scalable
Interpretability	Interpretable results	Not interpretable results	Not interpretable results
Usability	Dimensionality reduction, feature selection, noise reduction	Only visualization, exploratory data analysis	Only visualization, exploratory data analysis
Parameters to be tuned	Small number of parameters to be tuned (9)	More parameters to be tuned (16)	More parameters to be tuned (12)

Table 1.2: Comparison between dimensionality reduction algorithms: PCA, t-SNE and UMAP.

One problem with these methods is that they are not ideal for biomedical data because they cannot handle noisy data effectively. This is why new algorithms are being developed like PHATE, released in 2019, a method that captures both local and global data structure by using an information-geometric distance between the data points (Moon et al., 2019)(Fig. 1.23).

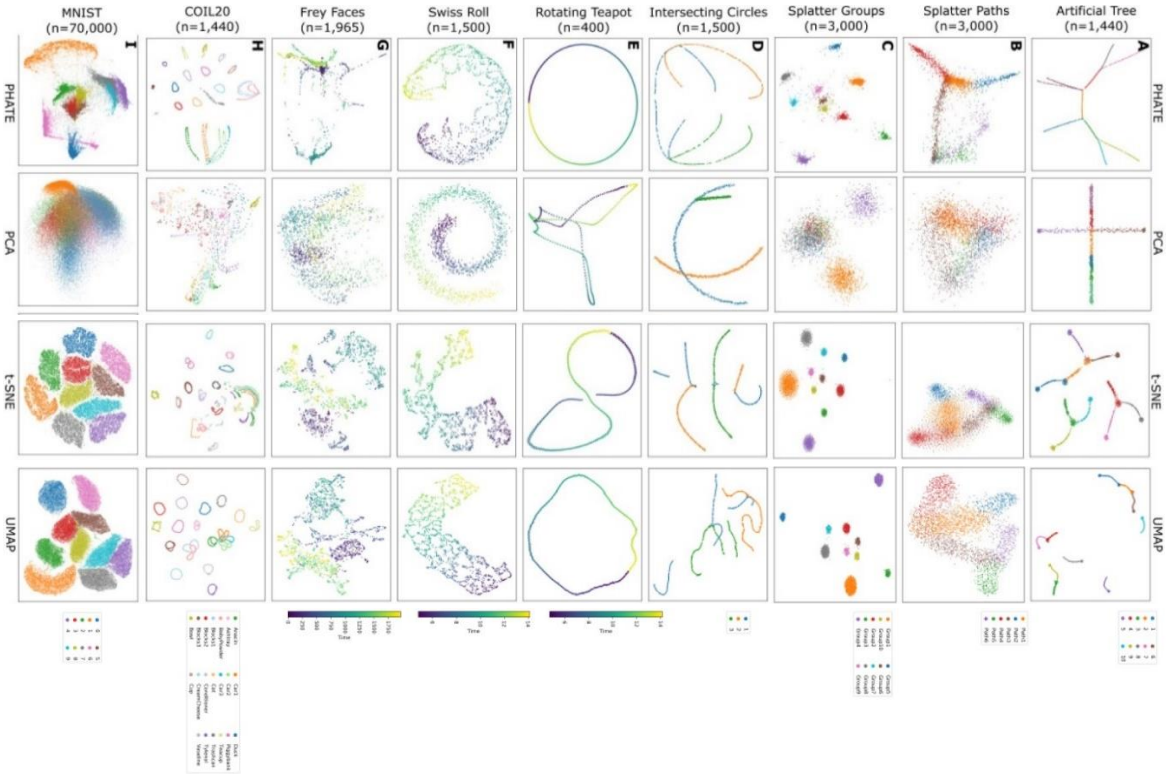


Figure 1.23: Comparison of the dimensionality reduction algorithms PCA, t-SNE, UMAP and PHATE in different datasets and cluster structures. A-I The columns illustrate the different biological datasets, where n represent the number of data points, rows represent the different visualization method applied for each dataset. The color codes correspond either to cluster number (A,B,C,D,H,I) or to trajectory time (E,F,G). The figure shows that the selection of a dimensionality reduction algorithm depends on both the structure of your data and your specific visualization objectives. Image from Moon et al. 2019.

1.3.3 Applications of machine learning in biology

For the development of this thesis we have benefited of already published softwares that help to automatically perform image segmentation, in this case we have used Tissue Analyser (Aigouy et al., 2016). Image segmentation is the process of automatically identify the shape of nuclei or cell membranes. The task of image segmentation has evolved during years, first it was addressed and solved as a task of image thresholding or active contours. Next, the tasks can be addressed and solved as a pixel classification problem with the help of models based on deep learning such as StarDist (Stevens et al., 2022), Cellpose (Stringer et al., 2021; Pachitariu & Stringer, 2022) , Napari (Chiu et al., 2022) or Episeg (Aigouy et al., 2020), based on supervised algorithms and applications that successfully segment 2D and 3D images. The main limitation of this algorithms is that they required a priori large datasets of labelled data to train the models. Once the outlined of the cell or of the nuclei

is obtained, we can then quantify different features of these objects, such as their area, perimeter, cell orientation, elongation axis etc. There are also plenty of applications and plug-in such as Tissue Analyser (Aigouy et al., 2016) a FIJI plug-in used not only for cell quantification but also for cell and bonds tracking in 2D tissues, LocalZprojector, a FIJI plug-in used to measure tissue curvatures, Dproj (Herbert et al., 2021) a MATLAB plug-in for cell quantification in 2D taking into account the tissue slopes and curvatures, or SEGGA (Farrell et al., 2017) a software for analyzing epithelia cell polarity, cell shape and dynamics.

Apart from the uses in image analysis, machine learning algorithms have currently been used in several tasks such as in sc-RNAseq and genetic analysis were cells are being clustered according to the expression of mRNA. These strategies have been used in genetic analysis of embryonic mouse heart at single-cell resolution which revealed a new cardiac progenitor pool genetically distinct from the already known FHF and SHF cardiac progenitors, the JCF (JuxtaCardiac Field) (Fig. 1.25 1)(Tyser et al., 2021). These pipelines have also been used in human genetic studies where European individuals were genotyped and a geographical map of Europe arises naturally as 2-dimensional summary of genetic variation in Europeans (Fig. 1.25 2) (Novembre et al., 2008) or in Korean ethnicities where similarly analysis of ancient and present genome sequences allowed tracking of the genome relationships and origins of actual genotypes (Kim et al., 2020). Lately, there is new upcoming trend in developmental biology to use this type of analysis to classify, cluster and predict the cell identities and delimit territories with similar characteristics according to their morphological features (Andrews et al., 2021), mechanical (Stower et al., 2023; Hallou et al., 2023) behaviors (Stower et al., 2023) or even dynamics (Dominguez et al., 2023). For example, k-means clustering was used to cluster the cells within the myocardium of E9.5 embryos having into account cell volume and cell cycle time, without spatial information, the resulting clusters showed clear spatial continuity (de Boer et al., 2012) (Fig. 1.26). More recently, morphometric data was used to cluster cells according to their morphological and behavioral properties in embryos during gastrulation which lead to patterned phenotypic heterogeneity within the visceral endoderm as well as identifying a new sub-group of cells in distal visceral endoderm (Fig. 26.2) (Stower et al., 2023). Single-cell morphometric analysis has also been used to capture notochord cell shape diversity in amphioxus cells and showed the variation in cell shape trajectories across the antero-posterior axis (Fig. 1.26 3) (Andrews et al., 2021). Due to this increase of interest in the clustering of cells according to their morphological and mechanical properties, there is an increasing number of applications and plug-ins that perform these tasks after cell segmentation such as IMARIS or Napari, whose aim is to democratize the use of machine learning for the non-coding community.

Data science is a discipline that combines computer sciences and statistics and as it has evolved, it has incorporated the concepts of artificial intelligence (M. K. Driscoll & Zaritsky, 2021). A data science pipeline involves the following steps: 1) Problem definition and objectives 2) Data collection 3) Data preprocessing: clean the data by

handling missing values, outliers and inconsistencies. preprocess the data by normalization 4) Data exploration and visualization: perform in-depth exploration of the data to understand patterns, relationships and potential insights using visualization tools and dimensionality reduction. Identify and select the relevant features or create new ones that contribute most effectively to your model and analysis (feature selection/engineering). 5) Choose the appropriate machine learning or statistical model based on the task (regression, classification, clustering) and the characteristics of the data. 6) Asses the performance of the model by evaluating the metrics and finally adjust the parameters to get the best performance. If the model meets the desired performance, deploy it to make predictions on new, unseen data (Fig. 24).

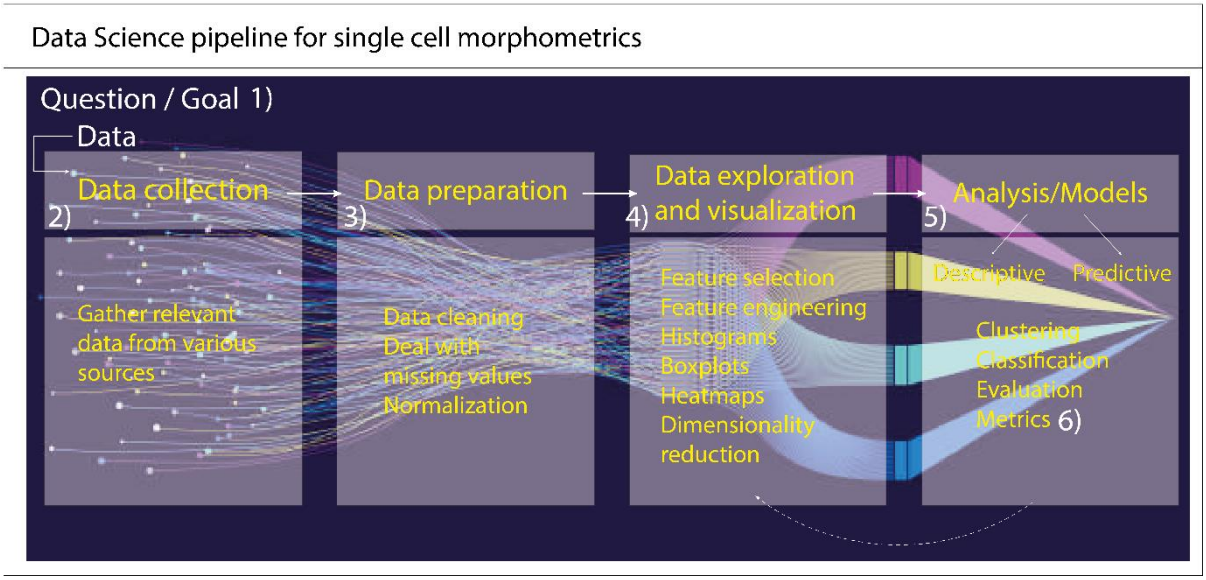


Figure 1.24. Pipeline for a data science analysis.

The next step in biomedical data analysis is the use of Deep learning, which is the state-of-the-art machine learning approach. Efforts are being made to develop robust deep learning tools for medical and clinical diagnosis, for example, for image interpretation to assist radiologists, in the detection of breast cancer (Q. Li & Nishikawa, 2015). Deep learning technics has also been challenged in the diagnosis of congenital heart defects form micro-computed tomography scans to automatically screen normal from malformed hearts (Nguyen et al., 2024).

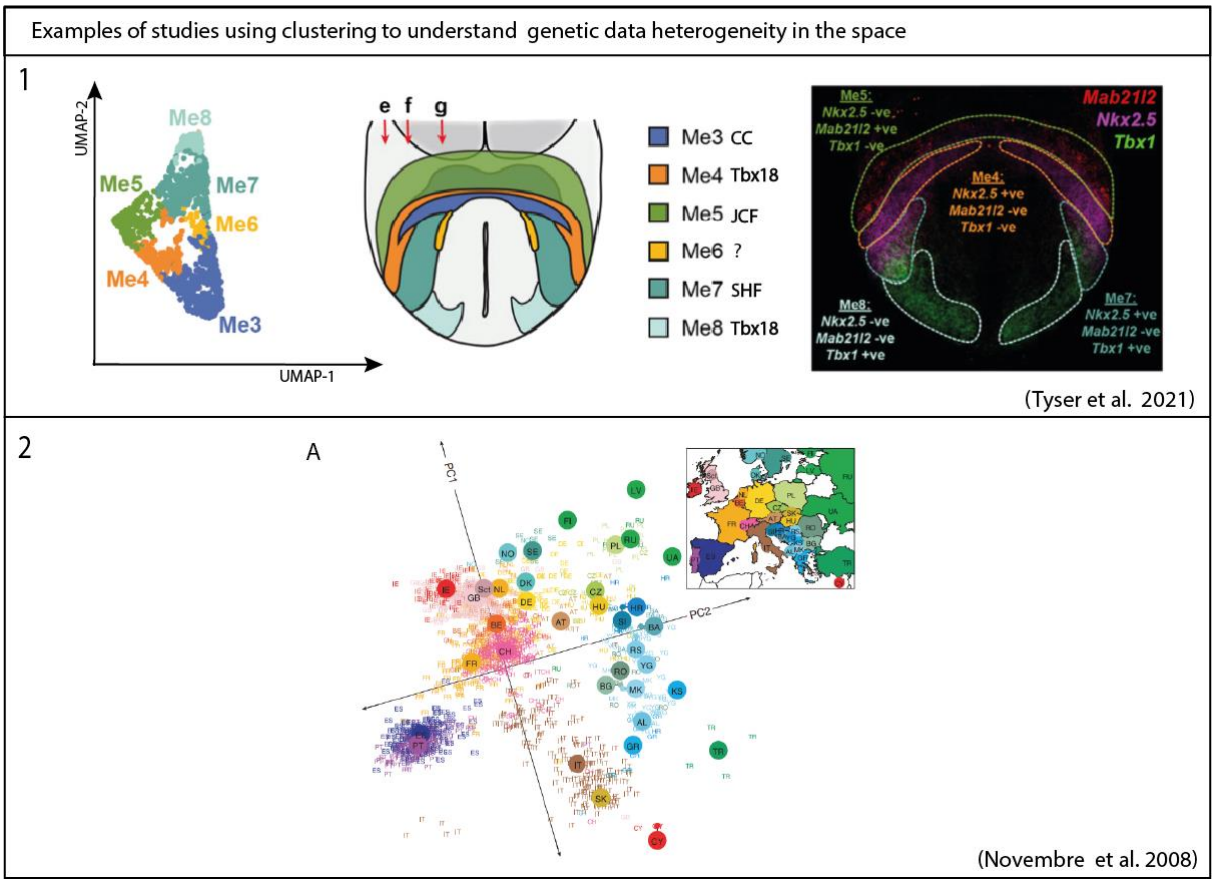


Figure 1.25: Examples of studies using clustering algorithms to understand genetic data heterogeneity in space. **1**. Panel A shows the UMAP space with the 6 clusters, panel B shows the cartoon with the different clusters in the real space in the embryo together with the population of cells the clusters represent, panel C shows RNAscope of a late E7 embryo marked with *Mab212* (red), *Nkx2.5* (pink) and *Tbx1* (green) as well as the rest of the clusters outlined. Me: Mesodermal cell types, CC: Cardiac Crescent, JCF: Juxtacardiac field, SHF: Second Heart field (Tyser et al., 2021) **2.** Statistical summary plotted in the principal component analysis where small colored labels represent and large colored point represent the media PC1 and PC2 for each country (Novembre et al., 2008).

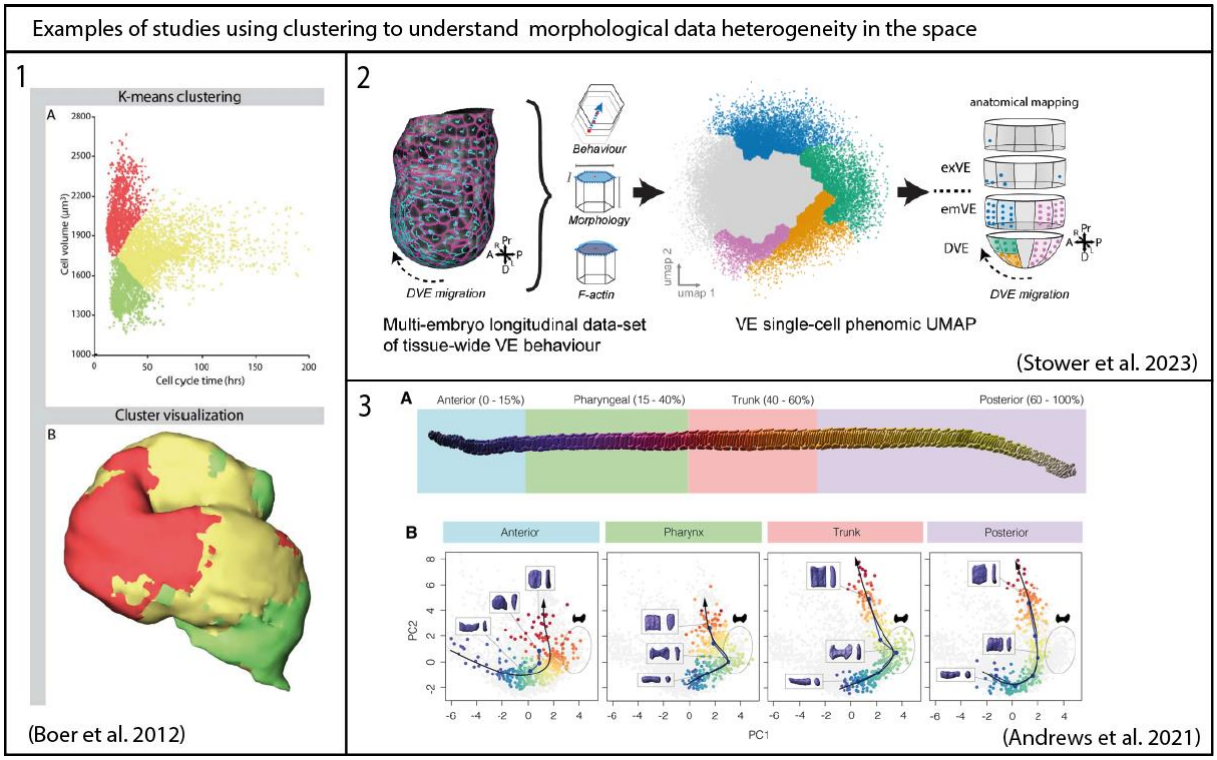


Figure 1.26: Examples of studies using clustering algorithms to understand morphological data heterogeneity in space. **1**. Panel A shows the cluster analysis of boxels based on their volume and the cell cycle length, panel B shows the labelled voxels plotted into the myocardium reconstruction (de Boer et al., 2012) **2.** Single-cell phenomics reveals patterned phenotypic heterogeneity within the distal visceral endoderm (DVE) (Stower et al., 2023) **3.** Single-cell morphometrics reveals cell shape trajectories across the antero-posterior axis (Andrews et al., 2021).

1.4 The goal & interest of the project

The knowledge gap

We do not understand the mechanisms by which SHF cells lying in the progenitor field move into the heart tube. More specifically, we do not understand the mechanisms regulating the segregation of progenitor cells to alternate poles of the heart at Embryonic day (E)9 (Fig. 1.27).

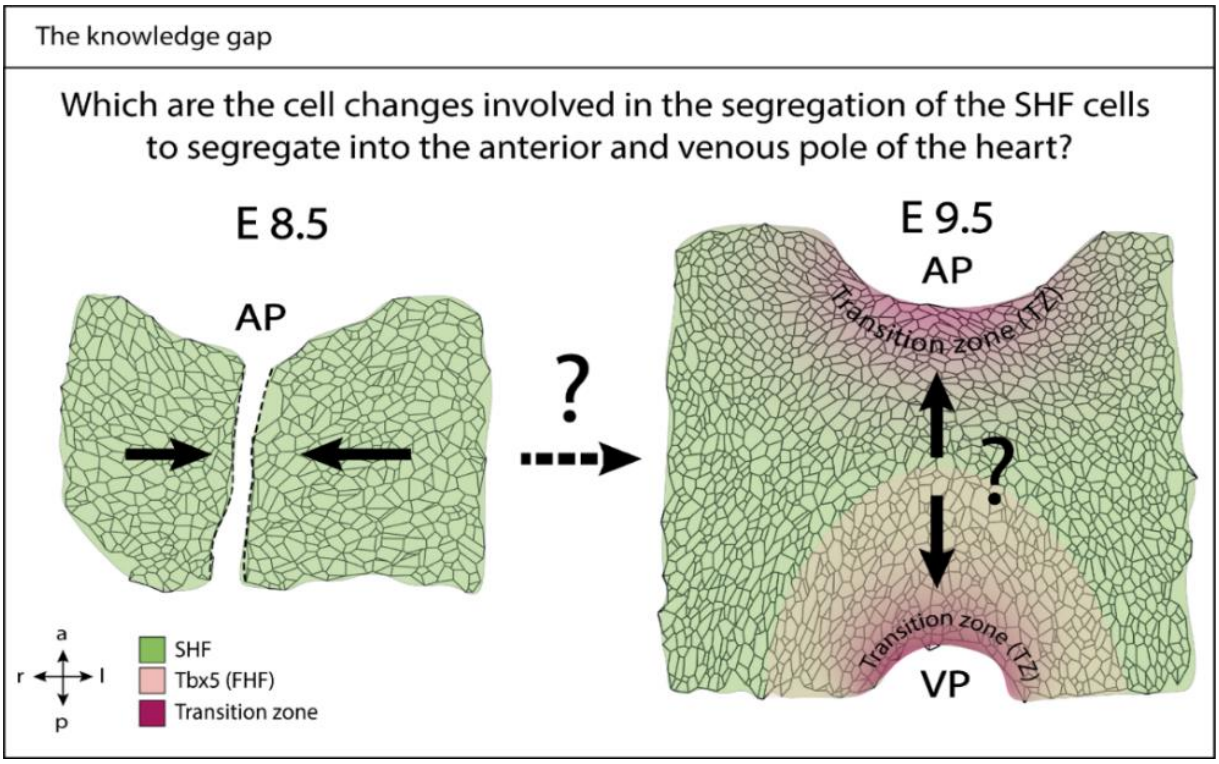


Figure 1.27: The knowledge gap addressed in this study.

The research question

What are the molecular, morphological and mechanical cell changes taking place in the SHF as cells segregate to the anterior and venous poles of the heart? (Fig. 1.27).

Limitations of previous studies in heart development

Existing approaches in studies of morphogenesis (D. Li et al., 2019; Francou et al., 2017), (Esteban et al., 2022) suffer from 3 main limitations:

1) they measure forces in an invasive way, by cutting the OFT or incision of the epithelium,
2) do not analyze the features at the single-cell level,
3) biased choice of certain regions of the epithelium

How have we attempted to overcome previous limitations?

To overcome these problems, we have:

1) used non-invasive image analysis,
2) quantified data at single-cell level over the full epithelial layer,

3) performed the analysis using an unbiased analysis by a clustering approach-based on morphometry that can evaluate the whole epithelium.

So far, we know that there is spatial heterogeneity and maturation of cells within the SHF, despite them being a unique population with a homogeneous distribution of certain genes (such as Isl1, Fgf10) and proteins (such as ZO-1, aPKC, Scribble). On top of the genetic heterogeneity, because of previous studies showing differences in mechanical properties, we have decided to measure morphological and mechanical measurements to perform a single-cell morphometric study of cells in the epithelium and verify if they also segregate into arterial and posterior regions. This study will identify regions with similar properties and establish boundaries. The results from this part of the study can be found in Section 3.1 where I detail the single-cell morphometric analysis performed in the SHF epithelial cells. The analysis was conducted in 2 temporal conditions E8.5 and E9.5, covering the process of dorsal mesocardium break down and heart tube elongation. Additionally, it includes two mutant conditions where *Tbx1* and *Tbx5* were deleted in the SHF cells.

In Section 3.2, I examine the formation of the *Tbx1-Tbx5* boundary. This includes investigating potential boundary genes by single-cell RNA sequencing analysis of SHF cells, as well as the existence of a high-tension boundary between the two domains.

Other mechanisms that may play a role in SHF deployment to the cardiac poles are presented in Section 3.3 including convergent extension, analyzed through rosette formation in the SHF. In Section 3.4, I present the results from classification tasks that combine the morphological properties of the cells with the presence or absence of Tbx5 protein.

2. Materials and methods

2.1 Animal models

The following mouse lines were used: *Tbx1*$^{+/-}$ (Jerome & Papaioannou, 2001), *Mef2c-AHF-Cre* (Verzi et al., 2005), *Tbx5*$^{fl/fl}$ (Bruneau et al., 2001) , *Nkx2-5*$^{+/-}$ (Tanaka et al., 1999), and CD1 mice. Mice were genotyped using DNA extracted from tail-tip, hair or yolk sacs by standard PCR protocol. Embryos were collected at embryonic days (E) 8.5 or 9.5, dated from noon of the day of the vaginal plug (E0.5).

Animal experiments were carried out in agreement with national and European law and approved by the Ethics Committee for Animal Experimentation of Marseille and the French Ministry for National Education, Higher Education and Research (Apafis No. 10266-2017061618121519).

Mouse line	Primer sense	Primer sequence	Product size
MefC-AHF-Cre	**FW**	5' AGG TTC GTT CAC TCA TGG A 3'	Cre = 200 bp
	RV	5' TCG ACC AGT TTA GTT ACC C 3'	
***Tbx5*$^{fl/fl}$**	**FW**	5' AGT CCA ATG ACT TAT CCT GGT GCG C 3'	WT = 203 bp MUT 232bp
	RV	5' AAA TTC CAA CCC CTT CCA CAG ATC GG 3'	
***Tbx1*$^{+/-}$**	**a**	5' CCC AAG GTG GTC TAG CTC TC 3'	WT = 241 bp MUT =300 bp
	b	5' CCT TCC TCT GTA GCC ATA CTG C 3'	
	c	5' ATC TCT CGG GGG ATC ATT C3'	

Table 2.1: PCR primers for genotyping. WT: wild-type, MUT: mutant.

Cre

Temperature	Time	Cycle
98°C	5 min	x1
98°C	20 sec	x32
60°C	20 sec	
72°C	30sec	
72°C	2min	x1

***Tbx5*$^{fl/fl}$**

Temperature	Time	Cycle
98°C	5 min	x1
98°C	20 sec	x35
64°C	20 sec	
72°C	30sec	
72°C	2min	x1

***Tbx1*$^{+/-}$**

Temperature	Time	Cycle
98°C	5 min	x1
98°C	20 sec	x35
60°C	20 sec	
72°C	30sec	
72°C	2min	x1

Tables 2.2-4: PCR programs for the different primers.

2.2 Embryo dissection

Embryos were dissected in PBS and fixed for 1 hour in 4% paraformaldehyde (PFA) for whole-mount staining or overnight for RNAscope or classic In Situ Hybridization. After fixation, the heart was removed to expose the dorsal pericardial wall and embryos processed for whole mount or RNAscope analysis.

For genotyping, adult tail tips, hairs and yolk sacs of embryos were used for genotyping. They were lysed using proteinase to extract DNA and PCR was performed using specific primers. PCR products were separated using agarose gel electrophoresis to detect the specific amplified fragments.

2.3 Embryo culture

E8.5 embryos were dissected in CO_2 independent medium (Invitrogen, 18045-054), 0.04% BSA (Albumin, from bovine serum, SIGMA A88056), 1% penicillin/streptomycin (Invitrogen, 15140-122) without damaging the yolk sac. Embryos were cultured at 37°C in rolling-bottles (30/rotations/minute) for 24 h in 75% rat serum (Janvier Labs), 25% DMEM (Invitrogen) and 1% penicillin/streptomycin in a gas-fitted incubator (BTC Engineering Cambridge, UK). For the first 12 h of culture embryos were exposed to a mixture of 5% CO_2,5%O_2, 90% N_2 and subsequently switched to 5% CO_2, 20% O_2, 75% N_2.

For pharmacological treatment, control embryos were cultured with DMSO at 1 µl/ml and treated embryos with the panRAR antagonist BMS493 at 10^{-5} M (Bristol-Myers-Squibb, Princeton, NJ, USA; B6688 SIGMA). After culture, embryos were removed from the yolk sac and fixed for 1 h in 4% paraformaldehyde prior to whole-mount immunostaining.

2.4 Immunofluorescence staining on whole mount embryos

Fixed embryos with the heart removed were blocked overnight in blocking buffer (PBS/0,1% Triton/3% BSA), incubated overnight at 4 °C with primary antibodies, washed in blocking buffer and incubated overnight at 4 °C with species-specific secondary antibodies. Finally, embryos were stained for F-actin with Phalloidin-Atto 647 and counterstained with Hoechst. After staining, the dorsal pericardial wall was mounted and imaged ventrally using a confocal LSM 780 or 880 microscope. Confocal images were acquired as Z-stacks. The following antibodies were used: goat anti-Tbx5 (1/250, Santa Cruz sc-7866), rabbit anti-Cre (1/500, Novagen 69050), rabbit anti-ppMLC (1/200, Cell Signaling Thr18/Ser19 3674), mouse anti-Golgi (GM-130) (1/100, BD Transduction

Laboratories 610823), mouse anti-YAP (1/100, Santa Cruz sc-101199). Fluorescent secondary antibodies Alexa 488, 568 and 647 from Jackson and Invitrogen were used at 1/200 dilutions. Two fluorescent forms of Phalloidin were used to label F-actin: Phalloidin 647 (1/50, Sigma 65906) or Phalloidin 488 (1/50, Sigma 49409).

Antibody	Species	Dilution	Reference
TBX5	**goat**	1/250	Santa Cruz sc-7866
Cre	**rabbit**	1/500	Novagen 69050
Golgi	**mouse**	1/100	BD Transduction Laboratories 610823
YAP	**mouse**	1/100	Santa Cruz sc-101199
ppMLC2	**rabbit**	1/200	Cell Signaling Thr18/Ser19 3674
Phalloidin	**/**	1/50	Sigma 65906/Sigma 49409

Table 2.5: Primary, secondary antibodies and Phalloidin used for immunostaining.

2.5 RNAscope on whole mount embryos & quantification

After dissection the heart was removed to expose the dorsal pericardial wall. Subsequently, embryos were fixed overnight in 4% paraformaldehyde (PFA) and then dehydrated in methanol baths (75%, 95% and 100%). Whole-mount RNA-FISH was performed according to the protocol of the RNAscope Multiplex Fluorescent v2 Assay (Acbio; cat. No323110)(de Soysa et al., 2019) . Embryos were imaged using a confocal LSM 780 or 880 microscope. The following probes were used: mm-Tbx1-C1 (cat no.481911), mm-Tbx5-C2 (cat no. 519581). For pixel intensity quantification, a ROI line in the middle of the image at E9.5 embryos and in the left part at E8.5 embryos, creating a sagittal section from arterial pole to the venous pole was quantified. For each mRNA channel a plot profile of the ROI was calculated using ImageJ and the algorithm uniform_filtered1d from the scipy library was used to smooth the profile plot.

2.6 In Situ hybridization riboprobe synthesis

Template for the synthesis of *Chn2* and *Cldn11* antisense riboprobe was generated by PCR from E9.5 mouse cDNA using the following primers for *Chn2*: 5' GCC TCT CCA TAT CTC CCG TG 3' and 5' GTG TGT GGG TCA TCT CAG AG 3', for *Cldn11*: 5' GCC AGG CAC TAA AGC CAA AG 3' and 5' ACC CAA GTC AGC AAT GTT CC 3'. PCR products were subcloned in pGEM-T Easy plasmid (Promega) and sequences were verified. Plasmids were linearized with specific restriction endonucleases and antisense riboprobes were transcribed from linearized plasmids with T7, T3 or Sp6 RNA polymerase (Promega) in the presence of Digoxigenin-11dUTP or Fluorescein-12-dUTP (Roche).

2.7 Whole mount in-situ hybridization

E9.5 embryos were dissected in 1x PBS at 4°C and fixed overnight in 4% paraformaldehyde. Embryos were washed in 1x PBS and dehydrated by graded ethanol baths (75%, 95% and 100%). Embryos were bleached with ethanol/H2O2 (5:1, v/v) for 1 hour. Following bleaching, they were rehydrated progressively by ethanol baths and washed in PBS/0.1%Tween. This was followed by proteinase K treatment (20µg/ml, Gibco fungal 25530-031) for 5 minutes. Embryos were then washed twice in glycine at 2mg/ml in PBS/0.1% Tween, fixed in 4% paraformaldehyde/0.2% glutaraldehyde and washed again in PBS/0.1%Tween. This was followed by preincubation of the embryos in hybridization buffer (50% formamide, 5X SSC, 50µg/ml yeast total RNA, 1%SDS, 50µg/mL Heparin) for at least 1 hour. Embryos were then incubated overnight at 70°C in hybridization buffer containing Digoxygenin-labelled probes at 1µg/ml. Embryos were then washed twice for 30 minutes in 50% formamide/5XSSC/1% SDS and 3 times 5 minutes in 0.5M NaCl/10mM Tris-HCl pH7.5/0.1% Tween at 70°C. This was followed by two RNAse A treatments (100µg/ml) 30 minutes each at 37°C followed by 3 washes in 0.5M NaCl/10mM TRIS-HCl pH7.5/0.1% Tween and 3 washes for 30 minutes in 50% Formamide/2X SSC and in TBST (0.14M NaCl, 2.7mM KCl, 25mM Tris-HCl pH7.5, 0.1% Tween). After washing, embryos were blocked with 10% sheep serum and incubated with anti-Dig-AP (Alkaline Phosphatase) (1/2000), overnight at 4°C. Antibodies were washed with TBST over 2 to 3 days at 4°C. Embryos were then washed in high pH buffer (NTMT, 100mM NaCl, 100mM Tris-HCl pH 9.54, 50mM MgCl2, 1% Tween), three times for 20 minutes each. Alkaline phosphatase revelation was performed using BM purple (Roche) at room temperature. Embryos were fixed in 4% paraformaldehyde and imaged in an LSM 770 /880 confocal microscope.

2.8 Imaging and cell membrane segmentation

The DPW was imaged ventrally as a Z-stack using a Zeiss LSM 780 or 880 confocal microscope with a 40X objective. The 20X objective was used to have an overview of the full embryo stage (headfolds). The apical surface of the DPW epithelium was imaged in a ventral view. The apical cortical F-actin belt was used as an apical membrane marker for segmentation. Maximum intensity projection and background noise cleaning per channel was performed prior to segmentation using LocalZprojector (LZP) imageJ plug-in (Herbert et al., 2021). Segmentation of the epithelium was performed using Tissue Analyser software (Aigouy et al., 2016). The software is an ImageJ plugin using a watershed algorithm to segment the cell cortex. Segmentation was performed on the entire DPW. Briefly, base segmentation was automatically performed using Tissue Analyser software and corrected manually.

2.9 Rosette quantification

For rosette quantification, vertexes where more than 4 cells meet were quantified by a custom algorithm developed by Benoit Aigouy. A vertex is considered multicellular when it shares more than 4 cells, a cell is considered to be part with a rosette when it shares at least 1 pixel-length in the vertex. The algorithm counts the number of rosettes in the tissue according to different cutoff going from 1 to 9. The cutoff is the bond length of the cells in the meeting vertex. The larger the cutoff, the less stringent the bond measurement, resulting in a higher count of rosettes. For our analysis, we used the smallest bond length of 3 pixels to identify a rosette. Cells from the same rosette have the same color. Cells shared by two adjacent rosettes are color-coded in a different color. Measurements of the total number of rosettes were performed in E8.5 wild-type embryos, N=4, E9.5 wild-type embryos, N=8, *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos, N = 4, *Tbx1*$^{-/-}$ embryos, N = 4.

2.10 Image analysis – quantification of morphological and mechanical features

Following cell membrane segmentation, apical surface area, perimeter, eccentricity or apical circularity index (0 means that the cell fits a rounded shape, 1 means that the cell fits an ellipse shape), cell orientation and number of neighbors were generated using Deproj MATLAB plug-in (Herbert et al., 2021). Binned cell elongation axis plots were generated using Tissue Analyser. Segmentation data was stored in the form of a .xlsx file. Stress measurements including cell pressures, stress main direction (ϑ_1) and stress anisotropy, which is the ratio between the amplitudes of the stress tensor (λ_1/λ_2) were calculated using force inference MATLAB plug-in (Kong et al., 2019). Stress maps and pressures map were generated using the same plug-in. Stress anisotropy map were generated using a custom-made Python code: if the ratio is close to 1 means that the cells are under isotropic stress and they are color coded in red, if the ratio is close to 0, the cells are under anisotropic stress and they are color coded in yellow. For clarity the plots have been simplified, only the cells with stress anisotropy bellow 0.3 are color coded in yellow. Morphological and spatial data were stored together with mechanical data in the form of a .xlsx and .csv file for all embryos.

2.11 Data analysis for spatial sc-morphometrics

In the article version all features were maintained in the DataFrame, Principal Component Analysis (PCA) and Moran's Index (MI) were assessed to check for the data shape and the spatial coherence of the features. To overcome the fact that PCA is not spatially informative, spatial PCA (sPCA)was used, by combining PCA and maximizing the Moran's

index (MI). Due to the specificity of the spatial PCA for each sample, Solène Song developed a joint spatial PCA (jsPCA) diagonalizing the sPCA matrices across the 6 E9.5 wild-type samples. The joint spatial PC1(jsPCA) vector was used to fit clustering models; in this case, the best algorithm model fitting our data is agglomerative clustering. The results of the labels were assessed by plotting back the labels in the real coordinates of the embryo space. Qualitative and categorical color palettes were chosen following ColorBrewer2.0 (Harrower & Brewer, 2003), and continuous values were plotted following Crameri color palettes (Crameri et al., 2020), two tools for selecting color maps schemes based on the nature of the data.

3. Results

Cell morphology is the form and structure of the cells. It involves characteristics of the cells such as their size, shape and organization of the cellular components. As cells develop and specialize, their morphology undergoes changes that are closely linked to their function and role within the adult tissues and organs.

During embryogenesis stem cells differentiate into specialized cell types such as cardiomyocytes, neurons or epithelial cells. The morphology of progenitor cells changes as they adopt specific shapes and structures suited for their respective functions. For example, cardiomyocytes have an elongated shape with branching structures, allowing for coordinated contraction and electrical communication while neural cells develop long extensions that enable them to transmit electrical signals over long distances or epithelial cells of the digestive tract facilitates absorption (Figure 24). In the subsequent results section, we will examine cardiac progenitor cells before they acquire their final shape, function, and spatial localization. Initially, we will characterize the morphology of the cardiac progenitors, followed by the genetic disruption of these cells. I will also present findings on additional epithelial features of this population and conclude with a genetic characterization of the progenitors with focus on the emergence of a transcriptional boundary within the progenitor cell epithelium.

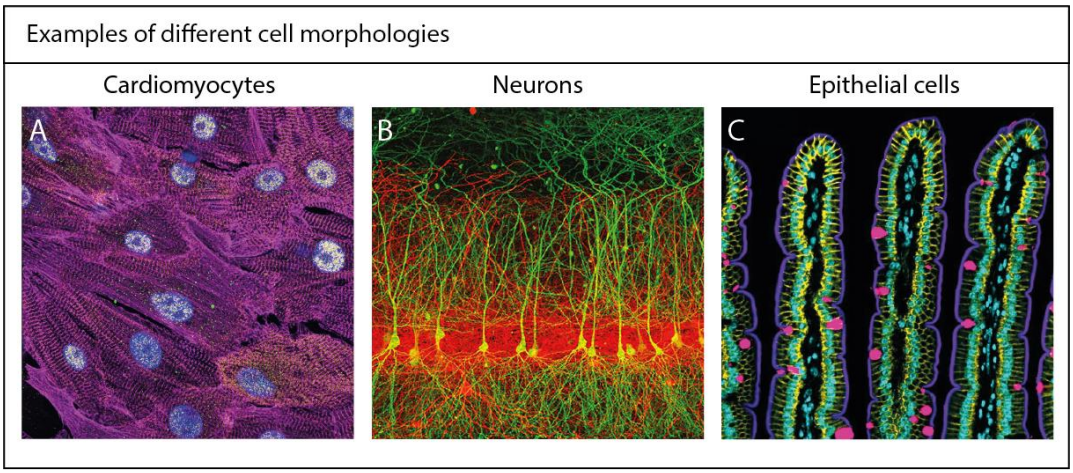


Figure 3.1: Examples of different cell morphologies A. Cardiomyocytes in cell culture. Yap1 in yellow, Amotl1 in green, nuclei in blue and α-actinin in magenta. Image taken when working in the lab of Sigolène Meilhac at Institut Imagine **B.** Pyramidal neurons in green and interneurons in red. Credit: Erik Bloss from Janelia research campus **C.** Intestinal cross section showing its characteristic villus structure. Nuclei in green, gamma-actinin in purple, goblet cells in pink, cell membrane in yellow. Credit: Amy Engevik from Medical University of South Carolina.

3.1 Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart Field

Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart field

Clara Guijarro[1,2,3 +], Solène Song[2,3 +], Benoit Aigouy[1], Raphael Clement[1], Paul Villoutreix[2,3 *], Robert G. Kelly[1 *]

[1] Aix-Marseille Université, CNRS UMR 7288, IBDM, Turing Centre for Living systems, Marseille, France

[2] Aix-Marseille Université, LIS, UMR 7020, Turing Centre for Living systems, Marseille, France

[3] Aix-Marseille Université, MMG, Inserm U1251, Turing Centre for Living systems, Marseille, France

[+] These authors contributed equally

[*] These authors jointly supervised this work

Addresses for correspondence:
Robert G. Kelly,
Aix-Marseille University, CNRS UMR 7288,
IBDM, Luminy Campus Case 907,
73288 Marseille Cedex 9, France
E-mail: Robert.Kelly@univ-amu.fr

Paul Villoutreix,
Marseille Medical Genetics Inserm U 1251,
Turing Centre for Living Systems,
Aix-Marseille University, La Timone Campus,
13005 Marseille, France
E-mail: Paul.Villoutreix@univ-amu.fr

Key Words: Heart morphogenesis, heart fields, congenital heart disease, epithelial mechanics, quantitative analysis, single-cell morphometrics, T-box genes.

Abstract

The vertebrate heart tube extends by progressive addition of epithelial second heart field (SHF) progenitor cells from the dorsal pericardial wall. The interplay between epithelial mechanics and genetic mechanisms during SHF deployment is unknown. Here, we present a quantitative single-cell morphometric analysis of SHF cells during heart tube extension, including force inference analysis of epithelial stress. Joint spatial PCA reveals that cell orientation and stress direction are the main parameters defining apical cell morphology and distinguishes cells adjacent to the arterial and venous poles. Cell shape and mechanical forces display a dynamic relationship during heart tube formation. Moreover, while the T-box transcription factor Tbx1 is necessary for cell orientation towards the arterial pole, activation of Tbx5 in the posterior SHF correlates with the establishment of epithelial stress and SHF deletion of Tbx5 relaxes the progenitor epithelium. Integrating findings from cell-scale feature patterning and mechanical stress provides new insights into cardiac morphogenesis.

Introduction

During embryonic development, cells adopt diverse strategies to sculpt epithelial tissues into organs with molecularly and mechanically distinct regions and characteristic 3D shapes. Organogenesis of the vertebrate heart, originating in epithelial cells in anterior splanchnic mesoderm of the early embryo, is a good example. The cardiac primordium is formed by First Heart Field (FHF) progenitor cells generated by an early wave of mesodermal cell commitment to cardiac fate [1,2,3]. The heart tube then grows by the progressive addition of Second Heart Field (SHF) progenitor cells from contiguous splanchnic mesoderm in the dorsal pericardial wall (DPW). While SHF cells initially contribute along the entire length of the cardiac primordium, progenitor cell addition becomes restricted to the arterial and venous poles of the heart after breakdown of the dorsal mesocardium and formation of the ventral heart tube [4]. SHF cells ultimately give rise to myocardium of the right ventricle and outflow tract at the arterial pole, and atrial myocardium, including atrial septal structures, at the venous pole. As SHF progenitors incorporate, the heart tube elongates, buckles and loops to the right [5,6].

Heart tube extension by progenitor cell deployment creates the template for subsequent cardiac septation events that generate the definitive four-chambered heart. Thus perturbation of SHF addition leads to a spectrum of congenital heart defects including outflow tract and atrial septal defects, observed in human genetic disorders such as 22q11.2 deletion (or DiGeorge) syndrome [7], and Holt-Oram syndrome [8,9]. The T-box transcription factor encoding genes *TBX1* and *TBX5* are implicated in these syndromes, respectively, and have been shown in mouse models to play major roles in regulating progenitor cell addition to the poles of the heart and early heart morphogenesis [10,11,12,13]. Patterning of the SHF epithelium leads to the establishment of distinct TBX1 expressing arterial pole progenitor cells in the anterior SHF and TBX5 expressing venous pole progenitors in the medial posterior SHF. TBX1 coordinates cell addition to the arterial and venous poles [14,15]. It also regulates proliferation, differentiation and epithelial polarity in the SHF [16,17] while TBX5 is required for left ventricular and venous pole development, atrial septation and cell cycle progression in the posterior SHF [11,12]. Despite this clinical significance the cellular mechanisms by which contiguous epithelial progenitor cells are incorporated into the cardiac primordium are poorly understood.

Apicobasal polarity studies revealed that SHF is an atypical polarized epithelium. Cell contacts and epithelial properties are restricted to the apical region of SHF cells while the basolateral

membrane forms dynamic filopodia-like protrusions more commonly observed in mesenchymal cells [17]. Evidence for epithelial tension in the DPW has suggested that SHF cells are subject to mechanical forces that contribute to the process of heart tube extension [18] . In support of such a model, oriented cell growth and cell division together with ppMLC and YAP/TAZ accumulation have been observed in the posterior DPW and wounding experiments have been performed on the epithelium showing oriented deformation consistent with epithelial tension [19]. Further evidence for the importance of epithelial properties of the SHF has emerged from studies of the Planar Cell Polarity (PCP) pathway. Perturbation of *Wnt5a*, encoding a non-canonical Wnt ligand regulating PCP and a transcriptional target of TBX1, impacts SHF addition to the arterial and venous poles, leading to both outflow tract and atrial septation defects [20,21]. In the absence of the PCP gene *Vangl2*, SHF-derived cells are abnormally polarized and disorganized, which leads to a shortening of the OFT and loss of SHF cell progenitor status [22]. In addition, N-cadherin and Fibronectin expression in the SHF are required for epithelial cell architecture and normal heart tube elongation [23,24]. While these studies highlight the importance of epithelial properties of the SHF for progenitor cell deployment, the mechanisms and cues involved in the establishment and regulation of cardiac progenitor cell addition to alternate poles of the heart remain unknown.

Studies more focused on the physical aspects of epithelial morphogenesis have shown that animal cells can actively change their shape through their contractile activity as well as being passively shaped by forces transmitted across the tissue within which they reside [25]. Indeed, contractile forces can propagate across long distances in the plane of the epithelium. Such forces can be counter-balanced by outward pressure generated by tissue growth. Therefore, morphometric analysis and evaluation of mechanical forces are in general essential to make sense of morphogenetic movements and organogenesis [26,27,28,29,30]. Assessing the mechanical state of cells in the SHF is therefore of major interest for understanding how the embryonic heart acquires its shape.

The development of image analysis approaches to segment and quantify apical epithelial parameters [31,32] as well as mechanical features [33] at the level of individual cells, has led to the generation of large datasets. The high degree of complexity of these datasets calls for the use of unsupervised data analysis approaches [34] to explore individual cell features. In addition, the spatial aspect of these types of datasets calls for the development of specific methods that distinguish spatial regions of cells with related features. In a similar way, approaches have been

applied to the analysis of large single cell RNA sequencing datasets to generate tissue gene expression cartography [35]. Previous studies of epithelial features of the SHF epithelium have focused on manually selected regions of the anterior and posterior DPW [19,20]. However, the choice of selected regions might not be representative of the entire tissue level dynamics. A quantitative analysis of morphological and mechanical features at the single-cell level across the entire epithelium coupled with unsupervised approaches such as dimensionality reduction and clustering should reveal dynamic features associated with cell deployment in the progenitor cell epithelium.

The purpose of our study is to create a quantitative map of the morphological and mechanical properties of cardiac progenitors in the DPW and investigate how T-box genes impact on these features to orchestrate the addition of progenitor cells to the cardiac poles. To achieve this aim, we quantified cell morphometrics and inferred forces at single-cell level in the entire SHF epithelium. Once these single-cell features were obtained, we adapted a spatial PCA approach [36] to characterize the major axes of variation and perform unsupervised clustering analysis of cell features. Our results reveal that cell orientation and the principal direction of cell stress are the main features explaining both the variation between cells and the spatial coherence of apical cell morphology in the DPW at embryonic day (E) 9.5. Single cell morphometric analysis prior to dorsal mesocardium breakdown confirms that patterned epithelial tension across the DPW arises as SHF addition become restricted to the cardiac poles. We show that the T-box transcription factors TBX1 and TBX5 differentially impact on tissue stress patterns. Arterial pole progenitor cell behavior correlates with *Tbx1* expression which is required to orient cells around the outflow tract and create a region with low isotropic stress. In contrast, activation of *Tbx5* in the posterior medial region of the DPW between E8.5 and E9.5 correlates with the acquisition of epithelial stress in the DPW. Mutant analysis reveals that Tbx5 activation is required to orient cell shape and cell stress parallel to the antero-posterior axis in venous pole progenitor cells. Together these results reveal that single cell morphometrics can provide insights into how patterned epithelial stress is established in the SHF during early cardiac organogenesis, revealing the interplay between genetics and mechanics during organogenesis.

Results

Single-cell morphometric analysis of the second heart field

The DPW is a heterogeneous epithelial tissue with an organized cellular architecture. Morphogenesis results from dynamic interactions between patterning signals, cell regulatory networks and tissue geometry [37]. Previous studies have suggested the importance of epithelial tension in the SHF for the growth of the heart [18,19,20]. To further investigate the importance of cell morphological features and cell mechanics in SHF deployment, we have developed a new spatial single-cell morphometric analysis pipeline (Fig. 1a). To perform single cell morphometrics, we built a dataset of morphological and mechanical features calculated from Tissue Analyzer [32], an ImageJ package to segment track and quantify 2D epithelial tissues, Deproj, a MATLAB package to yield accurate morphological measurements on 2D projections[31] and Bayesian Force Inference, a MATLAB package to quantify and infer the distribution of forces in a developing tissue from their shape and connectivity[38,39] (Fig. 1a, see Methods). Frontal whole-mount views of microdissected DPWs of 6 E9.5 wild-type embryos staged between 19-24 somites were analyzed with a total of 8618 cells. Apical cell contours were marked by Phalloidin staining. 9 features related to cell morphology and cell mechanics were calculated for each cell of the epithelia. Z-stacks were imaged covering the entire depth of the epithelia (47 μm depth) from the apical to the basal side of the cells, allowing us to generate a 3D acquisition of the cells in the epithelium, used to obtain the 2D projection.

The 9 features comprised 6 morphological features and 3 mechanical features (Fig. 1b,c). The morphological features evaluated were: 1) orientation of the main cell elongation axis, 2) apical cell area, 3) apical cell perimeter, 4) cell eccentricity (apical circularity index), a measure of the elongation of cell shape. In addition, we scored the number of cell neighbors (number of cell apical sides) (5) as a characteristic of cell shape within epithelia [40]. We also evaluated nucleus-to-Golgi polarity (6), which is the angle formed between the vector from the center of the nucleus to the center of the Golgi, relative to the embryonic left-right axis. To measure epithelial stress patterns, we used a force inference MATLAB plug-in [33]. This image analysis-based approach inferes forces, including stress patterns, in epithelial tissues based on the topology and geometry of cell contacts [39,33]. The 3 mechanical features we analyzed were cell pressure, the principal direction of the cell stress tensor (ϑ_1) and stress anisotropy, which is the ratio of the amplitudes of the stress tensor in the principal and secondary directions (λ_1/λ_2).

Nucleus-to-Golgi polarity was included as a potential read-out of planar polarity in DPW cells. In front-rear cell polarity models, such as directed cell migration, centrosomal reorientation towards the leading edge is known to align the Golgi apparatus towards the leading edge [41,42]. Potentially, in the DPW, the Golgi may be oriented towards the arterial or the venous pole. The nucleus-to-Golgi morphological feature was also addressed in 3D reconstructions. The results show that 90% of cells in the DPW have the Golgi positioned apically to the nucleus. However, comparison of the position of the Golgi in the anterior-posterior or left-right axes revealed no preferred localization for the Golgi relative to the nucleus in cells from different regions of the DPW. The Golgi is located in the anterior part of the cell in 50% of the cells and in 50% of the cells Golgi is located in the right part of the cell (Fig. S1). This suggests that apicobasal polarity rather than planar polarity determines Golgi position in DPW cells.

We performed principal component analysis (PCA) on these measurements for cells in the E9.5 DPW. PC1 (Principal component 1) reflects cell size (cell area and perimeter) and cell pressure; PC2 is a combination of cell orientation and cell stress direction; PC3 is composed mainly of stress anisotropy and cell eccentricity while PC4 is composed mainly of number of neighbors (Fig. S2a). PC1 to PC4 are the major contributors to variation: PC1 explains 29% of the heterogeneity of the cells in the DPW according to their morphological and mechanical characteristics, PC2 explains 19% of the variability of the cells, PC3 13% and PC4 11% (Fig. S2b). While PCA is an adapted tool to characterize the intrinsic variation of single cell features, we investigated whether it also accounted for the spatial distribution of cells within the tissue sample. To assess this, we computed the spatial-autocorrelation of each principal component, also known as Moran's Index (MI). We found that PC1 had a lower MI than PC2, and PC2 had the highest MI across all PCs (Fig. S2c). Analysis of correlations between these parameters show that cell area and perimeter have a correlation coefficient of 0.9 and stress direction and cell orientation have a correlation of 0.7 while the remaining features have lower correlation coefficients (Fig. S3). While PCA focuses mainly on feature covariation, our analysis suggests that it is not sufficient to capture the spatial variation within the data.

Development of joint spatial PCA of sc-morphometric data

We investigated if we could identify meaningful spatial patterns for the single cell

morphological and mechanical features within the tissue. For this, we turned to spatial PCA[36], a dimensionality reduction method that identifies the main directions of variation within the data, while considering the spatial distribution. To obtain this double information, we built on a method considering the product matrix between the covariance of the data and Moran's Index. The eigenvectors of this product matrix and their associated eigenvalues reveal the directions that maximize both the covariance and the spatial autocorrelation of linear combinations of variables. This approach generalizes the classical PCA which aims to identify the major direction within the covariance matrix[36]. As shown by Jombart et al.[36], the product matrix between the covariance matrix and Moran's Index can be rewritten in a simple form as a function of the feature matrix and the adjacency matrix, a matrix encoding how close cells are within the tissue. The gain in information by considering jointly the variance in the data and the spatial information comes, however, at the expense of the ability of combining separate tissue samples. Specifically, the product matrix is dependent on the adjacency matrix which is sample-specific and cannot be easily pooled between samples. In order to consider jointly samples in the same conditions, we performed a joint diagonalization [44] of the product matrix (see Methods for details). This approach, which we name joint spatial PCA (jsPCA), yields the joint spatial Principal Components of features within a collection of tissues.

Patterned cell orientation and stress direction in the SHF

We applied the joint spatial PCA pipeline to the single-cell morphometric dataset of cells in the DPW of E9.5 WT embryos. The first joint spatial Principal Component (jsPC1) for each of the sample was composed of cell orientation and stress direction, while the second joint spatial Principal Component (jsPC2) was composed of cell area and cell perimeter (Fig. 1d). These first two jsPC capture most of the variation for all embryos of the cohort (Fig. 1e). Since the aim of jsPCA is to obtain orthogonal directions of variations, we expect the patterns associated with jsPC1 and jsPC2 to be quite different, as represented in Fig. 1f,g. The jsPCA approach was then used to perform dimensionality reduction by projecting the multivariate distribution of cells of each sample on jsPC1. We used the projected data to perform an unsupervised clustering using a Gaussian mixture approach. We found that using 3 clusters provided a relevant way of representing the data which is consistent with the Akaike Information Criterion (Fig. S4). In addition, we were able to use the cluster characteristics learned on E9.5 WT embryos to define comparable clusters in other conditions.

Mapping back cells of the three clusters onto the E9.5 WT embryo revealed that cells in clusters 1 and 2 are distributed in a bilaterally symmetrical manner on the right and left sides of the DPW, alternating around the inferior outflow tract in the anterior DPW, while cells labelled as cluster 3 are predominantly located in the posterior part of the DPW (Figs. 2a,b, S4 and S5). As cell stress direction, orientation, area and perimeter are the main features according to the joint spatial principal component analysis (jsPCA), we calculated the mean values for the 4 parameters for each of the 3 clusters. These mean values serve as feature signatures defining each cluster. Thus, the morphological signatures of the cells labelled as cluster 1 are: cell stress direction: 45° ± 20 and cell orientation: 45.78 °± 20, cell area: 54.4μm^2 ± 32 and cell perimeter: 33.06μm ± 9. The morphological signatures for cells labelled as cluster 2 are: cell stress direction: -47.1° ± 20 and cell main orientation: -47.7° ± 20, cell area: 47.7μm^2 ± 27 and cell perimeter: 30.7μm ± 9. The morphological signatures for cells labelled as cluster 3 are: cell stress direction: 1.46° ± 28 and cell stress main orientation: -0.57° ± 29, cell area: 51.66 μm^2 ± 30 and cell perimeter: 32.34 μm ± 10 (Fig. 2c-f). Mean values for the remaining features are common to clusters 1-3 (Fig. 2g-j). Our single cell morphometrics analysis thus permits visualization of patterned epithelial features indicative of patterned tension across the DPW at the stage of heart tube elongation. Given the bilateral symmetry of clusters 1 and 2 we also classified cells into two clusters computed with cell orientation and stress direction symmetrized with respect to the anterior-posterior axis (Figs. 2k and S4). Although introducing a priori knowledge about the symmetry of the system, and hence some level of supervision, this two cluster analysis highlights the distinct stress patterns in the anterior and posterior DPW and the radial symmetry of anterior DPW cells as they converge on the arterial pole.

Patterned tension in the SHF emerges as the heart tube forms

The connection between the DPW and the heart tube varies over time: at E8.5, the DPW epithelium is connected to the early heart by a mesentery, the dorsal mesocardium, along the antero-posterior axis (Fig. 3a). After the dorsal mesocardium breaks down behind the heart tube, the bilateral right and left sides of the DPW epithelium fuse. As a result, the heart tube remains connected to the SHF only at the arterial and venous poles (Fig. 3b). Single cell morphometrics was performed on whole-mount DPWs microdissected from E8.5 wild-type embryos with 6-8 somites, before breakdown of the dorsal mesocardium (Fig. 3c). The clusters

297 learnt in the E9.5 wild-type embryos were used to classify E8.5 DPW cells.
298
299 Quantitative evaluation of apical cell properties at E8.5 revealed a striking difference in the
300 orientation of cells in the posterior DPW with respect to that at E9.5, before and after dorsal
301 mesocardium breakdown. At E8.5 DPW cells are oriented parallel to the embryonic left-right
302 axis and under isotropic stress (Figs. 3c-c''', S6). In contrast, at E9.5, cells in the posterior DPW,
303 both medially and laterally, are oriented at 90°±10, and display stress anisotropy with the main
304 stress tensor oriented along the anterior-posterior embryonic axis (Fig. 3d-d'''). This reveals the
305 emergence of an E9.5-specific cluster enriched in the posterior DPW between these timepoints,
306 consistent with analysis of cell elongation in selected subregions of the DPW. While in late
307 E9.5 embryos with >20 somites cluster 3 is restricted to the posterior DPW, in embryos with
308 <20 somites an enrichment of cells belonging to cluster 3 was observed in the midline of the
309 DPW, the location of dorsal mesocardial breakdown, expanding towards the anterior dorsal
310 pericardial wall (Fig. S5e,f). Classification of E8.5 cells into two clusters, computed with
311 symmetrized cell orientation and stress direction, also revealed that distinct anterior and
312 posterior clusters emerge between E8.5 and E9.5 (Fig. S6). These results reveal the dynamic
313 nature of cell shape changes and patterns of epithelial tension in the DPW during these stages
314 of heart tube morphogenesis. In particular they highlight the emergence of epithelial features
315 associated with elevated tension in the posterior DPW after breakdown of the dorsal
316 mesocardium.
317
318
319 **Comparison of cluster and T-box gene distribution in the SHF**
320
321 Previous studies have shown, in serial sagittal sections, the emergence of a transcriptional
322 boundary delimiting TBX1 and TBX5 positive domains in the DPW after the dorsal
323 mesocardium breaks down. These domains arise by de novo activation of *Tbx5* expression in
324 the medial posterior region of the DPW at E8.5 followed by TBX1-dependent downregulation
325 of the anterior SHF program in TBX5 positive cells by E9.5 [14]. We validated *Tbx1* and *Tbx5*
326 expression domains using RNAscope fluorescent in situ hybridization on frontal whole-mount
327 views of the DPW and quantification of mRNA intensity (Figs. 4a-e, S7). At early stages (6-8
328 somites), the DPW is entirely *Tbx1*-positive (Fig. 4a,b). As the dorsal mesocardium breaks
329 down behind the central part of the heart tube, *Tbx5* expression is activated in medial posterior
330 SHF cells close to the venous pole (Fig. 4c). *Tbx1* and *Tbx5* are transiently co-expressed (Fig.

4d) prior to the establishment of a boundary as *Tbx1* is downregulated in *Tbx5* expressing cells. At E9.5 distinct *Tbx1*+ and *Tbx5*+ domains of the DPW are well defined (Fig. 4e). *Tbx1*+ *Tbx5*- cells are observed in the anterior DPW and lateral regions of the posterior DPW. Comparison of *Tbx1* and *Tbx5* expression at E9.5 with the sc-morphometric cluster distribution reveals that *Tbx1* expression correlates largely with the distribution of clusters 1 and 2. In contrast, *Tbx5* expressing cells in the posterior DPW at E9.5 fall entirely within cluster 3, that also encompasses *Tbx1*+*Tbx5*- cells in the lateral regions of the posterior DPW (Fig. 4f).

TBX1 and TBX5 pattern apical cell architecture in the SHF

Tbx1 and *Tbx5* are essential for normal development of the arterial and venous poles of the developing heart, respectively. Moreover, haploinsufficiency for *TBX1* or *TBX5* contributes to congenital defects affecting the arterial and venous poles of the heart in 22q11.2 deletion and Holt-Oram syndromes [47,48]. As the emergence of distinct *Tbx1* and *Tbx5* positive territories in the DPW coincides with the breakdown of the dorsal mesocardium, and the expression profile of these genes correlates with the distribution of cells from clusters 1 and 2 (*Tbx1*) and cluster 3 (*Tbx5*), we used single cell morphometrics to investigate the impact of TBX1 and TBX5 loss of function on patterns of apical cell morphology in the DPW. To this end we performed single cell morphometric analysis of the DPW of E9.5 *Tbx1* homozygous null (*Tbx1*$^{-/-}$) embryos [49,50,51,52] in which SHF cell deployment to the arterial pole is defective, and conditional *Tbx5* mutant (*Mef2cAHFCre*;*Tbx5*$^{fl/fl}$) embryos [11,51,53] in which *Tbx5* expression is deleted in the SHF without affecting *Tbx5* in the first heart field and early heart tube. SHF deletion of *Tbx5* results in defective posterior SHF contributions to the venous pole, and later failure of atrial and ventricular septal morphogenesis at the heart field interface [11,51]. Single-cell morphometrics was performed on whole-mount DPWs microdissected from *Tbx1* mutant and *Tbx5* conditional mutant embryos at E9.5. As for the E8.5 dataset, cells from mutant embryos were classified following the clusters learnt from the E9.5 wild-type dataset (Figs. 5, S8-10).

In *Tbx1*$^{-/-}$ embryos, the DPW is hypoplastic and less organized. Single-cell morphometric analysis reveals a loss of the two main bilateral populations, corresponding to clusters 1 and 2, in the aDPW of *Tbx1* mutant embryos (Figs. 5a',b', S9). Cells of cluster 3 are enriched in the posterior DPW, as in E9.5 wildtype embryos, characterized by cells oriented at 90°, stressed parallel to the embryonic anterior-posterior axis and with mean cell area and perimeter of 50μm^2

and 35µm respectively (Fig. 5a',b'). In contrast, in *Mef2cAHFCre;Tbx5$^{fl/fl}$* embryos, cluster 3, which is regionalized at the venous pole in wild type, is no longer enriched in the posterior DPW and distributed throughout the epithelium (Figs. 5c', S10). The 2 remaining clusters (clusters 1 and 2) correspond to the left and right sides of the DPW of *Mef2cAHFCre;Tbx5$^{fl/fl}$* embryos. Quantification of the distribution of cells along the anterior-posterior axis of the DPW revealed significant posterior enrichment of cluster 3 cells in wildtype E9.5 and *Tbx1$^{-/-}$* embryos, but not in wildtype E8.5 or *Mef2cAHFCre;Tbx5$^{fl/fl}$* embryos (Fig. 5d). Single cell morphometric analysis of cells in the DPW of heterozygous *Tbx5* mutant embryos (*Mef2cAHFCre;Tbx5$^{fl/+}$*) revealed similar results to the wild-type situation (Fig. S11). Classification of *Tbx1$^{-/-}$* and *Mef2cAHFCre;Tbx5$^{fl/fl}$* cells into two clusters, computed with symmetrized cell orientation and stress direction, revealed that while two regionalised clusters are maintained in *Tbx1$^{-/-}$* embryos, emergence of a distinct posterior DPW cluster fails in *Mef2cAHFCre;Tbx5$^{fl/fl}$* embryos (Fig. S9, 10). In addition, loss of a distinct anterior cluster after symmetrization in *Mef2cAHFCre;Tbx5$^{fl/fl}$* but not *Tbx1$^{-/-}$* embryos suggests that radial symmetry resulting from convergence of cells on the arterial pole does not depend on TBX5. Together these results demonstrate that *Tbx1* and *Tbx5* differentially impact on patterns of apical cell morphology in the DPW and reveal that activation of *Tbx5* in the SHF is necessary for the regionalization of cluster 3 cells in the posterior DPW at E9.5.

Tbx5 regulates tension and orientation in the posterior SHF

In order to further understand how the absence of TBX1 and TBX5 differentially impact apical cell morphology in the DPW we examined the individual parameters constituting the single cell morphometric datasets, compared to E8.5 and E9.5 wildtype embryos (Fig. 6a,b). In *Tbx1* mutant embryos a larger TBX5+ domain is observed in the DPW (Fig. 6c), characterized by cells oriented in 90° as in wildtype embryos (Fig. 6c), as a result of the loss of the two bilateral domains of clusters 1 and 2 in the aDPW. In conditional *Tbx5* mutant embryos, the TBX5+ domain in the DPW is absent and cells in both the anterior and posterior DPW are oriented in a bilaterally symmetric manner towards the arterial pole (Fig. 6d). Cells oriented parallel to the antero-posterior axis are not regionalized in the posterior DPW of SHF *Tbx5* mutant embryos but are instead observed to be scattered throughout the DPW, both anteriorly and posteriorly (Fig.6d). Quantification of cells oriented horizontally (0°±10, along the embryonic left-right axis) and cells oriented vertically (90°, along the embryonic anterior-posterior axis), reveals an

enrichment of horizontal cells in the anterior DPW and an enrichment of vertical cells in the posterior DPW in the presence of TBX5, but not in E8.5 embryos prior to activation of *Tbx5* in the DPW or in *Tbx5* SHF mutant embryos at E9.5 (Fig. 6e).

A biomechanical feedback loop has been suggested to contribute to elongation of the heart tube through epithelial tension in cardiac progenitor cells in the posterior DPW [18]. As the joint spatial PCA algorithm highlights stress direction as one of the main parameters that account for most of data variance and that is spatially structured, we investigated the impact of loss of T-box gene function on stress patterns in the epithelia. In the presence of TBX5, the *Tbx1*$^{+}$ region correlates with cells under isotropic stress and the *Tbx5*$^{+}$ region correlates with cells under anisotropic stress and a stress tensor aligned parallel to the antero-posterior axis (Fig. 6b). In *Tbx5* conditional mutant embryos, the epithelium is relaxed, as observed from the stress tensor map, and there is an associated loss of cells with anisotropic stress in the posterior lateral DPW (Fig. 6d). As *Tbx5* expression is restricted to the medial region of the posterior DPW, this suggests that *Tbx5* indirectly impacts epithelial tension in a broader region of the epithelium, including the lateral posterior DPW.

The *de novo* activation of *Tbx5* in the posterior DPW has been shown to be blocked by pharmaceutical approaches using a pan-retinoic acid receptor inverse agonist (BMS493) at E7.5 and E8.5, leading to failure of SHF cell contributions to atrial septal structures at the venous pole [14]. We performed single cell morphometric analysis of E9.5 mouse embryos exposed to BMS493 after 24 hours of embryo culture. We observed altered stress patterns and cell orientations similar to those seen after genetic ablation of *Tbx5* in the SHF in 2/3 embryos, consistent with loss of cells corresponding to cluster 3 (oriented at 90° in the posterior DPW) and relaxation of the epithelium. One embryo showed a similar phenotype to wild-type embryos (Fig. S12). While consistent with our findings after conditional *Tbx5* deletion in the SHF, the reduced penetrance possibly reflects heterogeneity in precise embryo stage at the time of BMS493 exposure. Together, these findings point to a previously unidentified role of TBX5 in the acquisition of epithelial tension in the DPW during heart tube elongation.

Tbx5 regulates actomyosin localization in the SHF

In response to epithelial tension, non-muscle myosin can change its subcellular distribution to form active actomyosin complexes that counteract the forces acting on the cell [54,55].

433 Diphosphorylated myosin light chain (ppMLC) is the active form of myosin, and its subcellular
434 localization thus provides a read-out of actomyosin activity. At E9.5 the distribution of ppMLC
435 is concentrated on the long apical membrane of elongated cells in the DPW [19]. Here, we show
436 that the accumulation of ppMLC in the SHF coincides with the region where cells from cluster
437 3 localize, including TBX5 positive cells in the medial posterior DPW (Fig. 7a-a"). At the
438 cellular level ppMLC localizes to the elongated membranes of the cells (Fig. 7b). These results
439 are consistent with previous measurements of the nematic polarity of F-actin and ppMLC [19].
440 Moreover, this result is consistent with the anisotropy maps obtained using force inference
441 revealing an enrichment of anisotropic cells in the posterior DPW. In *Mef2cAHFCre;Tbx5*$^{fl/fl}$
442 mutant embryos, ppMLC accumulation is reduced in the posterior DPW (Fig. 7c-c"). At the
443 cellular level, ppMLC is homogeneously distributed along the cell membranes (Fig. 7d). This
444 altered distribution of ppMLC is consistent with the reduced anisotropy in the posterior DPW
445 of *Tbx5* conditional mutant embryos revealed by force inference. ppMLC localization was also
446 analyzed in embryos exposed to BMS493 in embryo culture. In such embryos (N=4) the
447 polarized pattern of ppMLC distribution that characterizes cells in the posterior dorsal
448 pericardial wall of wild-type embryos was not observed (Fig. S13). Together, these results
449 reinforce the conclusion that in the absence of *Tbx5*, patterned epithelial tension fails to be
450 established in the posterior DPW, leading to relaxion of the progenitor cell epithelium.
451
452 Previous work had indicated that anisotropic ppMLC localisation and cell orientation are
453 reduced in the DPW of *Nkx2-5*$^{-/-}$ embryos, consistent with lack of epithelial tension [19]. The
454 homeodomain containing transcription factor NKX2-5 is a pleiotropic regulator of heart
455 development required for normal heart tube elongation and patterning and *Nkx2-5*$^{-/-}$ embryos
456 die by E9.5 due to cardiovascular defects. Single cell morphometric analysis revealed loss of
457 cells with anterior-posteriorly oriented stress direction and stress anisotopy in *Nkx2-5*$^{-/-}$ embryos
458 (Figs. S14, S15). This is consistent with a relaxed DPW in *Nkx2-5*$^{-/-}$ embryos and with previous
459 observations of loss of cell orientation and homogeneous distribution of ppMLC in *Nkx2-5*$^{-/-}$
460 embryos at E9.5. Given the similarities with our findings in *Tbx5* conditional mutant embryos,
461 this suggests that SHF addition at the venous pole of the heart, a *Tbx5*-driven process, is a
462 critical determinant in the establishment of epithelial tension in the DPW.
463
464
465

Discussion

Morphogenesis is the result of dynamic interactions between the chemical and genetic components (or Gene Regulatory Networks) [62] and the physical and mechanical components (or Cell regulatory networks) of developing embryos [37,63]. The embryonic heart is a 3D organ constructed from the folding of epithelial splanchnic mesoderm. Here we have created a quantitative map of the apical morphology and mechanical state of SHF cardiac progenitors in the DPW during heart tube extension and revealed that two T-box genes, *Tbx1* and *Tbx5*, have distinct effects on patterns of epithelial tension in the progenitor field. This provides new insights into understanding the roles of these factors in orchestrating addition of SHF progenitor cells to the arterial and venous poles of the heart.

In the first part of the study we characterized the shape and mechanics of SHF cells throughout the DPW. We established a pipeline to perform single-cell morphometrics in SHF progenitor cells unifying geometrical measurements of the cells and mechanical measurements using recent image analysis tools, in particular force inference [31,32,33]. The pipeline has helped to pinpoint the main processes clustering the cells in the epithelium: orientation and stress direction. This result highlights the importance of cell mechanics during heart morphogenesis. These forces may arise through progenitor cell addition to the growing heart, initially across the dorsal mesocardium and later at the poles of the heart. Previous work has suggested that a mechanical feedback loop between progenitor cell deployment and epithelial tension in the SHF acts as a motor driving heart tube elongation [18]. The patterns of epithelial tension observed in our single cell morphometric analysis provide further support for such a model. Moreover, our findings suggest that establishment of opposing directionality in trajectories from the SHF to the heart through activation of a *Tbx5*-dependent venous pole progenitor cell program is a critical step in driving epithelial tension in the DPW. Our data also show that nucleus-to-Golgi polarity is among the least important features in the classification of the cells; indeed, the apical position of the Golgi highlights the apicobasal polarity of SHF cells [17]. This strengthens the idea that changes in cell shape and stress in the DPW epithelium are a passive response to external forces rather than being driven by active intrinsic forces, congruent with SHF deployment to the cardiac poles within an apicobasally polarised epithelial sheet.

The single cell morphometric pipeline has allowed us to map SHF progenitor cells in the DPW according to their morphological and mechanical properties. Once the two cardiac poles have

been established at E9.5, our map reveals two bilaterally symmetric clusters alternating around the arterial pole. This result is consistent with previous studies showing the rotation of the cells as they enter the outflow tract, ultimately ensuring correct ventricular arterial alignment [64]. Moreover, genetic studies have revealed differences in gene expression in left and right sides of the DPW, including the asymmetric expression of *Nodal*, *Pitx2c* and *Lefty* in left sided SHF progenitor cells [65,66].

We identified a third cluster labelling cells in the posterior DPW. This cluster identifies a subpopulation of progenitor cells that are oriented and anisotropically stressed at 90°, along the antero-posterior axis. This spatial clustering correlates with the establishment of distinct arterial and venous pole programs within the SHF and provides a read-out of cell properties as cells are added to the opposite poles of the heart, cells in the anterior part (clusters 1 and 2) being added to the arterial pole and TBX5 positive cells in the posterior domain (within cluster 3) being added to the venous pole. Our results thus highlight a symmetry breaking step in the DPW after dorsal mesocardial breakdown as *Tbx5* activation in the posterior DPW establishes a venous pole progenitor cell population with distinct epithelial properties. Indeed, comparison of single-cell morphometric maps of the DPW of E8.5 and E9.5 wild-type embryos as well as *Tbx1*$^{-/-}$ and *Mef2cAHFCre;Tbx5*$^{fl/fl}$ mutant embryos, revealed the *Tbx5* dependence of the regionalization of cluster 3 at the venous pole. In the absence of *Tbx5*, either at E8.5, before *Tbx5* expression in the DPW, or by its genetic disruption in *Tbx5* conditional mutant embryos, cluster 3 is not regionalized. TBX5 is thus necessary to establish the 3rd posterior morphological cluster and to drive epithelial tension and cell orientation parallel to the antero-posterior axis in the posterior DPW. Consistent with this conclusion, symmetrizing cell orientation and stress direction allowed cells be classified into two clusters distributed anteriorally and posteriorally at E9.5 but not E8.5. Moreover, loss of regionalized distribution of these clusters in *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos points to at least two independent forces driving epithelial stress in the DPW: TBX5-dependent patterned tension close to the venous pole and TBX5-independent stress as cells converge bilaterally on the arterial pole.

Genetic perturbation experiments show that TBX1 is important to maintain the maximal cell number within the epithelium, in agreement with previous studies[15,17,16,10]. We observe a loss of the two main bilateral populations in the aDPW; however, the loss of TBX1 does not affect the establishment of epithelial stress in the posterior SHF. On the other hand, we demonstrate that TBX5 is required in the SHF lineage to orient cells in the posterior DPW parallel to the

antero-posterior axis of the embryo at E9.5 and to establish anisotropic stress. This result suggests that the activation of *Tbx5* expression in the SHF cells may directly modulate the mechanical properties of cells in the posterior medial DPW. For example, this may be through regulating ppMLC localization to the elongated membranes of the cells that accompanies cell orientation in 90° in the posterior SHF. Interestingly, YAP and TAZ have been shown to directly interact with TBX5 to regulate transcription; future experiments will investigate whether common targets in the posterior DPW regulate SHF deployment. However, as the orientation and the tension pattern that define cluster 3 is observed in a broader posterior DPW territory than the TBX5+ domain, TBX5 may also or alternatively play an indirect role by exerting forces on the epithelium through TBX5-dependent SHF deployment to the venous pole of the heart. Similarly, loss of NKX2-5, associated with failure of heart tube extension at both poles, leads to loss of posterior enrichment of cells from cluster 3 and relaxation of the progenitor cell epithelium. Together these observations suggest that the onset of SHF cell deployment to the venous pole may drive the establishment of patterned tension in the DPW during heart tube extension. We suggest that the establishment of a distinct TBX5-dependent pathway for cell addition from the DPW to the elongating heart tube through the venous pole would create an opposing force to an earlier established arterial pole pathway, placing the epithelium under mechanical tension.

In early E9.5 embryos (<20 somites) cells from cluster 3 can be observed in the anterior medial region while only in later E9.5 embryos (>20 somites) is the 3rd cluster is restricted to the posterior region. This result provides insights into changing apical cell morphology during the dynamic process of dorsal mesocardium breakdown. Moreover, the dorsal mesocardium is in continuity with the *Tbx5* expressing FHF-derived heart tube at early somite stages [14]. TBX5 may be involved in cell shape changes taking place during dorsal mesocardium breakdown as well as in the later posterior DPW.

With this study, we have created a morphometric map of the SHF cells before and after dorsal mesocardium breakdown and have shown that T-box genes regulating addition of the progenitor cells to the heart tube affect the mechanical state of the DPW epithelium. This provides an example of how links between genetics and mechanical forces may coordinate organogenesis. Future perspectives include coupling single cell morphometric analysis with single cell RNA-seq analysis to generate an integrated map of transcriptional and morphological cell features during SHF deployment. These perspectives have begun to be explored in the mouse brain,

where mechanical patterns have been examined in brain compartments with different transcriptomic profiles [71].

Limitations of our study include the lack of tools to perform live imaging. The data analyzed in this work represent temporal snapshots of DPW development during SHF deployment and intermediate stages have not been addressed. There are significant challenges in dynamically imaging the DPW, which at these stages is located in the central part of the developing embryo and is currently not accessible to imaging over time even using light sheet microscopy. In addition, the single cell morphometric analysis performed in this study has focused on the apical side of the epithelium. The DPW is an atypical epithelium as the cells are epithelial-like apically yet have dynamic mesenchymal-like features baso-laterally [17]. This particular cellular conformation, together with the fact that active actomyosin accumulates on the apical side of cells in the DPW suggest that apical morphometrics provide an accurate readout of epithelial tension. However, exchanges with the extracellular matrix (ECM) baso-laterally, not evaluated here, and or adjacent cell types such as foregut endoderm, may also impact on overall epithelial morphology in the DPW.

Finally, we propose an unsupervised pipeline to delineate regions in space with similar morphometric and mechanical features. This pipeline starts with a dimensional reduction step that enhances the features which, on top of explaining most of the variance, as in classical PCA, also satisfy the criterium of displaying spatial regionalization. This can be applied to different embryos of the same condition, using joint diagonalization, a step termed Joint spatial PCA. The pipeline developed here for analysis of progenitor cells based on image analysis tools and unsupervised learning approaches can be used to analyze single cell measurements with any currently available segmentation and quantification tool.

Methods

Mouse strains and embryo generation. The following mouse lines were used: *Tbx1*$^{+/-}$ [7], *Mef2c-AHF-Cre* [53], *Tbx5*$^{fl/fl}$ [9], *Nkx2.5*$^{-/-}$ [67] and CD1 mice. Mice were maintained on a mixed C57Bl/6 and CD1 background and genotyped by PCR using standard protocols from DNA extracted from hair, tail tips, or yolk sacs as described in the above references and Supplementary Table 1. Embryos were collected at embryonic days (E) 8.5 or 9.5, dated from noon of the day of the vaginal plug (E0.5). Embryos were not genotyped for sex. Animal experiments were carried out in agreement with national and European law and approved by the Ethics Committee for Animal Experimentation of Marseille and the French Ministry for National Education, Higher Education and Research (Apafis No. 10266-2017061618121519).

Embryo dissection and whole-mount immunostaining. Embryos were collected and fixed for 1 h in 4% paraformaldehyde (PFA) for routine whole-mount immunostaining. After fixation, the heart was removed to expose the dorsal pericardial wall and embryos processed for whole mount staining. For immunostaining, embryos were blocked overnight in blocking buffer (PBS/0,1% Triton/3% BSA), incubated overnight at 4 °C with primary antibodies, washed in blocking buffer and incubated overnight at 4 °C with species-specific secondary antibodies. Finally, embryos were stained for F-actin with phalloidin and counterstained with Hoechst. After staining, the pharyngeal region was microdissected and the dorsal pericardial wall mounted and imaged ventrally using a confocal LSM 780 or 880 microscope. Confocal images were acquired as Z-stacks. The following antibodies were used: goat anti-Tbx5 (1/250, Santa Cruz sc-7866), rabbit anti-Cre (1/500, Novagen 69050), rabbit anti-ppMLC (1/200, Cell Signaling Thr18/Ser19 3674), mouse anti-Golgi (GM-130) (1/100, BD Transduction Laboratories 610823). Fluorescent secondary antibodies Alexa 488, 568 and 647 from Jackson and Invitrogen were used at 1/200. Two fluorescent forms of phalloidin were used to label F-actin: phalloidin 647 (1/50, Sigma 65906) or phalloidin 488 (1/50, Sigma 49409). In this work we included single cell morphometric analysis of images from a previous study [19] for 2 *Tbx1*$^{-/-}$ embryos; 4 *Nkx2.5*$^{-/-}$ embryos and 3 E8.5 wild-type embryos.

Mouse embryo culture. E8.5 embryos were dissected in CO_2 independent medium (Invitrogen, 18045-054), 0.04% BSA (Albumin, from bovine serum, SIGMA A88056), 1% penicillin/streptomycin (Invitrogen, 15140-122) without damaging the yolk sac. Embryos were cultured in rolling-bottles for 24 h in 75% rat serum (Harlan SR-0100 or IGBMC Strasbourg),

25% DMEM (Invitrogen) and 1% penicillin/streptomycin. For the first 12 h of culture embryos were exposed to a mixture of 5% CO_2,5%O_2, 90% N_2 and subsequently switched to 5% CO_2, 20% O_2, 75% N_2. For pharmacological treatment, control embryos were cultured with DMSO at 1 µl/ml and treated embryos with the panRAR antagonist BMS493 at 10^{-5} M (Bristol-Myers-Squibb, Princeton, NJ, USA; B6688 SIGMA). After culture, embryos were removed from the yolk sac and fixed for 1 h in 4% paraformaldehyde prior to whole-mount immunostaining.

RNAscope on whole-mount embryos. Embryos were fixed overnight in 4% paraformaldehyde (PFA) and then dehydrated in methanol. After fixation the heart was removed to expose the dorsal pericardial wall. Whole-mount RNA-FISH was performed according to the protocol of the RNAscope Multiplex Fluorescent v2 Assay (Acbio; cat. No323110)[68]. Embryos were imaged using a confocal LSM 780 or 880 microscope. The following probes were used: mm-Tbx1-C1 (cat no.481911), mm-Tbx5-C2 (cat no. 519581). For pixel intensity quantification, a ROI line in half of the left part of the embryos was drawn, creating a sagittal section from arterial pole to the venous pole. For each mRNA channel the intensity was measured along the ROI. The measurements were normalized using the min-max-normalization, scaling the values in the array range from 0 to 1. The profile plot of the measurements was visualized using ImageJ. The algorithm uniform_filtered1d from the scipy library was used to smooth the profile plot.

Confocal microscopy, image analysis and quantification. The dorsal pericardial wall was imaged ventrally as a Z-stack using a Zeiss LSM 780 or 880 confocal microscope with a 40X objective. The 20X objective was used to have an overview of the full embryo stage (headfolds). The apical surface of the DPW epithelium was imaged in a ventral view. The apical cortical F-actin belt was used as an apical membrane marker for segmentation. Max intensity projection and background noise cleaning per channel was performed prior to segmentation using LocalZprojector (LZP) imageJ plug-in[31]. Segmentation of the epithelium was performed using Tissue Analyser software [32]. The software is an ImageJ plugin using a watershed algorithm to segment the cell cortex. Segmentation was performed on the entire DPW. Briefly, base segmentation was automatically performed by the Tissue Analyser software and corrected manually. Apical surface area, perimeter, eccentricity or apical circularity index (0 means that the cell fits a rounded shape, 1 means that the cell fits an ellipse shape), cell orientation and number of neighbors were generated using Deproj MATLAB plug-in [31]. Binned cell elongation axis plots were generated using Tissue Analyser. Segmentation data was stored in the form of

a .xlsx file. Stress measurements including cell pressures, stress main direction (ϑ_1) and stress anisotropy, which is the ratio between the amplitudes of the stress tensor (λ_1/λ_2) were calculated using force inference MATLAB plug-in [33]. Force inference curvature tests were performed to calculate the radius of tissue and cell curvature (Fig. S16). In specific cases (Figs. 2k, S4b, S5, 6, 9 and 10), we have computed our pipeline on AP-symmetrized orientation and stress-direction. AP-symmetrized means that we consider absolute value of the (orientation – 90°) and the absolute value of the stress direction, making these two features symmetric with respect to the antero-posterio axis. Stress maps and pressure maps were generated using the same plug-in. Stress anisotropy map were generated using a custom-made Python code: if the ratio is close to 1 means that the cells are under isotropic stress and they are color coded in red, if the ratio is close to 0, the cells are under anisotropic stress and they are color coded in yellow. Morphological and spatial data were stored together with mechanical data in the form of a .xlsx and .csv file for each embryo.

Data analysis for single-cell morphometrics

The pipeline is made of two steps. The first step, called Joint spatial PCA, projects the cells along the axis that maximize the variance in the data while at the same time enhancing features that are spatially regionalized. The second step is training a Gaussian model on the projected data points to distinguish clusters, that are then mapped back onto physical space to reveal the location of the different regions. The pipeline is applied on the E9.5 WT embryos. Then, the cells from other conditions embryos are projected on the same axis, and the Gaussian model trained on the E9.5 WT data is used to cluster the additional embryos.

The input for the training on E9.5 WT embryos is the respective adjacency matrices (n x n) that express the neighbourhoods of cells in space and the features matrices (n x 8) where n in the number of cells in the embryo. Each column of the features matrices are centered and scaled for each embryo separately. Then, for each embryo, the matrix M_k (k=1:8) is expressed as $(X^t(L+L^t)X)/2n$ where L is the row normalized adjacency matrix. This matrix M_k is analogous to the covariance matrix in classical PCA. The difference is that the eigenvectors associated with highest eigenvalues not only explain most of the variance but also displays high spatial auto-correlation [36]. These matrices are then jointly diagonalized by a function made available by Cardoso et al. [44] in MATLAB. This consists in finding the optimal common eigenvectors for the set of matrices that approximate their diagonalizations. We rank the eigenvalues and select the highest, and take its associated eigenvector as projection axis jsPC1 (Figure 1d,e). We project all the datapoints on this axis. These main features contributing to this axis are cell

orientation and cell main stress direction. Then, we fit a Gaussian mixture model on all cells from embryos of E9.5 WT condition projected on the same axis jsPC1. The Akaike score gives 3 as the optimal number of clusters (Fig. S4). The cells from other conditions, all centered and scaled individually, are then clustered using the parameters of the previously fitted Gaussian mixture model learnt on E9.5 WT embryos. The different steps of the pipeline are implemented in Python, except for the function for approximate joint diagonalization implemented in MATLAB. Additional cluster maps showing each cluster independently are shown in Figures S17-S20.

Statistics and reproducibility. Segmentation and analysis was performed on three or more embryos of each of the different developmental stages, genotypes or manipulations, as follows (N is the number of embryos, n the number of cells): E8.5 wild-type embryos, N=6 , n=2996; E9.5 wild-type embryos, N=6, n=8618; *Tbx1*$^{-/-}$ embryos, N=3, n=2551 ; *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos, N=3, n=3695; *Nkx2.5*$^{-/-}$ N=4 , n=5568 ; BMS treated embryos, N=3 , n=3432. All data are included in the manuscript and supplementary figures. The results of single cell morphometrics where obtained by pooling the 6 E9.5 wild-type embryos. The results were used to predict the cells labels in the different temporal and genetically modified embryos. For cluster signatures the means were calculated and plotted in the distribution of the values for each feature and cluster number. Statistical analysis of mean values for antero-posterior and left-right distance was performed using two-tailed *T*-test. Mann Whitney U test from scipy.stats.mannwhitneyu was used to calculate the p-values in cluster boxplots (Fig. 5). Wilcoxon signed-rank test from scipy.stats.wilcoxon was used for vertical (> 90°), horizontal (<10°) boxplots and anisotropic cells (<0.3) in Fig. 6. For all boxplots, the y-positions were normalized as follows: $n_{ormalized}$=-2(y-<y>)/ (Y_{max}- Y_{min}); ns, non-significant; **P<0.01; ***P<0.001. Error bars represent standard deviation (SD).

Data availability. Source data are provided with this paper as a Source Data file and available on the Github: sc-morphometrics-SHF.

Code availability. The measurement of Golgi-to-nucleus polarity in 3D and the sc-morphology analysis is available on: sc-morphometrics-SHF, the code of the Bayesian force-inference has been published [38], and the Joint Spatial PCA is available on the Github: JointSpatialPCA.

731
732 **Figure legends**
733
734 **Figure 1. Spatial single-cell morphometrics highlights cell size, orientation and stress**
735 **direction as the main parameters defining apical cell morphology in the dorsal pericardial**
736 **wall. a.** Workflow for the single-cell morphometric analysis of cells in the dorsal pericardial
737 wall. **b-c.** Description of the cell features used for the single-cell morphometrics. **d.** Heatmap
738 showing the coefficients for each feature that compose each joint spatial principal component
739 axis (jsPCA). **e**. Heatmap showing the eigenvalues for each principal component. **f.**
740 Visualization of the main joint spatial PCA1 parameters: cell orientation and stress tensor in
741 real space. **g.** Visualization of the main jsPCA2 parameters: cell area and cell perimeter, in real
742 space. A: anterior, P: posterior, R: right, L: left, λ: stress amplitude, ϑ: stress main orientation,
743 e: embryo, JSPC: Joint spatial principal component. N = 6 embryos, n = 8618 cells. Source data
744 are provided as a Source Data file.
745
746 **Figure 2. Spatial single-cell morphometrics define regions in the DPW based on**
747 **morphological and mechanical cell properties. a.** E9.5 wild-type cells plotted back to their
748 physical space and color coded with the cluster number; cluster 1 cells are in green, cluster 2
749 cells are in orange and cluster 3 cells are in purple. **b.** Scatter plots of the cells by their cluster
750 label in the real space. **c-j.** Kernel density estimate plots of the different features defining each
751 cluster label. The measurements for cells belonging to cluster 1 are in green, for cluster 2 are in
752 orange and for cluster 3 in purple. **k.** E9.5 wild-type cells plotted back to their physical space
753 and color coded after classification into two clusters following symmetrization of cell
754 orientation and stress direction with respect to the anterior-posterior axis. Cluster 1 cells are in
755 green and cluster 2 cells are in pink. N = 6 embryos, n = 8618 cells.
756
757
758 **Figure 3. Spatial single-cell morphometrics reveals that patterned tension in the second**
759 **heart field emerges during heart tube formation. a, b.** Scheme of the architecture of the SHF
760 at embryonic days 8.5 (a) and 9.5 (b) showing breakdown of the dorsal mesocardium (DM)
761 between these timepoints. **c-c'''.** Frontal view of E8.5 wild-type whole mount dorsal pericardial
762 wall (c), main cell orientation map (c'), stress map (c'') and cluster map (c'''). **d-d'''.** Frontal
763 view of E9.5 wild-type whole-mount dorsal pericardial wall (d), main cell orientation map (d'),
764 stress map (d'') and cluster map (d'''). Scale bars: 100µm. NT, neural tube; AP: arterial pole,

765 VP: venous pole; r, right; l, left; d' dorsal; v, ventral; a, anterior; p, posterior. Wild-type E8.5:
766 N = 6 embryos, n = 2994 cells; wild-type E9.5: N = 6 embryos, n = 8618 cells.
767
768
769 **Figure 4. Clusters defined by spatial single-cell morphometrics correlate with the**
770 **distribution of T-box transcription factors in the dorsal pericardial wall. a-e.** Frontal
771 whole-mount views of the dorsal pericardial wall after RNAscope fluorescent in situ
772 hybridization showing *Tbx1* (green) and *Tbx5* (magenta) mRNA distribution between E8 (6
773 somites) and E9.5 (25 somites), before and after dorsal mesocardial breakdown. Nuclei are
774 stained using Hoescht (blue) **f.** Summary of the expression profiles of *Tbx1* and *Tbx5* based on
775 the schemas in Fig. 3a and b, showing the links between *Tbx1* and *Tbx5* expression domains
776 and clusters 1-3, and the input of retinoic acid signaling [14]. AP: arterial pole, VP: venous pole,
777 aSHF: anterior second heart field, pSHF: posterior second heart field. E8-E8.5: N=5 embryos,
778 E9: N = 3 embryos, E9.5 : N= 6 embryos. Scale bars: 100µm.
779
780
781 **Figure 5. T-box transcription factors TBX1 and TBX5 pattern apical cell architecture in**
782 **the dorsal pericardial wall. a-c.** Ventral whole-mount views of E9.5 wild-type, *Tbx1*$^{-/-}$, and
783 *Tbx5* conditional mutant DPWs stained with phalloidin. **a'-c'.** Cell segmentation of E9.5 wild-
784 type, *Tbx1* mutant and *Tbx5* conditional mutant and cells labelled according to their cluster
785 number. **d.** Boxplot quantifications of the spatial coordinates of cells belonging to each of the
786 clusters along the antero-posterior axis (with y coordinates normalized from -1 (VP) to 1 (AP))
787 for E8.5 (p=3.55, p=1) and E9.5 wild-type (***p=5.89e-35, ***p=5.91e-64) , *Tbx1*$^{-/-}$
788 (***p=5.00e-29, ***p=1.08e-52) and *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos (p=3.42, p=2.99); p-
789 values are based on Mann Whitney U test and test whether the spatial distribution of a cluster
790 is regionalised along the antero-posterior axis. Centre lines show the median, box limits show
791 the first and third quartiles, and whiskers show maximum and minimum values. For the
792 statistical test, the following biological replicas were analyzed: wild-type E8.5: N = 6 embryos,
793 n = 2994 cells; wild-type E9.5: N = 6 embryos, n = 8618 cells; *Tbx1*$^{-/-}$: N = 3 embryos, n = 2551
794 cells; *Mef2cAHFCre;Tbx5*$^{fl/fl}$: N = 4 embryos, n = 2809 cells. Scale bars: 50µm.
795
796
797 **Figure 6. TBX5 induces stress anisotropy and orients cells parallel to the antero-posterior**
798 **axis in the posterior dorsal pericardial wall. a-d.** Whole-mount immunofluorescence of the

dorsal pericardial wall. Phalloidin in white, TBX5 in red and nuclei in blue. **a-d'**. Cell orientation map. **a-d".** Spatial plot of the cells whose main orientation angle is 90°±10 compared with left-right axis of the embryo on the segmentation mask of the apical cell membrane. **a-d"'.** Cell stress map. **a-d''''**. Cell anisotropy quantification map. **A-D.** Morphometric map. **e.** Boxplot quantification of the spatial coordinate of cells oriented vertically (90°±10) or horizontally (0/180°) along the anterior-posterior axis (with y coordinates normalized from -1 (VP) to 1 (AP)); p-values are calculated using a Wilcoxon signed-rank test. For horizontal cells: E9.5wild-type, ***p=5.45e-41, Tbx1, ***p=4.34e-5, Tbx5, p=0.02. For vertical cells: E9.5WT: ***p=1.25e-39, Tbx1: ***p=5.11e-27, Tbx5: ***p=3.56. **f.** Boxplot quantification of cells with anisotropic stress (< 0.3) along the antero-posterior axis (with y coordinates normalized from -1 (VP) to 1 (AP)). p-values are calculated using a Wilcoxon signed-rank test and test whether the spatial distribution of cells deviates from a symmetric distribution. Error bars represent standard deviation. In boxplots in panels e and f, centre lines show the median, box limits show the first and third quartiles, and whiskers show maximum and minimum values. E9.5 wild-type, ***p=5.45e-41, Tbx1, ***p=4.34e-5, Tbx5, p=0.39. For the statistical tests, the following biological replicas were analyzed: wild-type E8.5: N = 6 embryos, n = 2994 cells; wild-type E9.5: N = 6 embryos, n = 8618 cells; *Tbx1*$^{-/-}$: N = 3 embryos, n = 2551; *Mef2cAHFCre;Tbx5*$^{fl/fl}$: N = 4 embryos, n = 2809 cells. Scale bars: 50μm.

Figure 7. Deletion of Tbx5 disrupts the cellular localization of diphosphorylated myosin light chain in the dorsal pericardial wall. a-a". Ventral whole-mount views of E9.5 wildtype DPW stained with phalloidin (white), ppMLC (blue) and Tbx5 (red). **b.** High magnification view of the posterior DPW in wild-type DPW embryos of the ppMLC; white arrows indicate the accumulation of ppMLC at the long membranes of the cells. **c-c".** Ventral whole mount views of E9.5 DPW *Mef2cAHFCre;Tbx5*$^{fl/fl}$ mutants stained with phalloidin (white), ppMLC (blue) and TBX5 (red). **d.** High magnification view of the posterior DPW region in *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos of the ppMLC, showing no accumulation of ppMLC at the long membranes of the cells. Wild-type E9.5: N=5 embryos; *Mef2cAHFCre;Tbx5*$^{fl/fl}$: N=4 embryos. Scale bars: 50μm.

References

1. Kelly, R. G., Buckingham, M. E. & Moorman, A. F. Heart Fields and Cardiac Morphogenesis. *Cold Spring Harb. Perspect. Med.* **4**, a015750 (2014).
2. Srivastava, D. Making or Breaking the Heart: From Lineage Determination to Morphogenesis. *Cell* **126**, 1037–1048 (2006).
3. Miyamoto, M., Gangrade, H. & Tampakakis, E. Understanding Heart Field Progenitor Cells for Modeling Congenital Heart Diseases. *Curr. Cardiol. Rep.* **23**, 38 (2021).
4. Kelly, R. G. The second heart field. *Curr. Top. Dev. Biol.* **100**, 33–65 (2012).
5. Le Garrec, J.-F. *et al.* A predictive model of asymmetric morphogenesis from 3D reconstructions of mouse heart looping dynamics. *eLife* **6**, e28951 (2017).
6. Esteban, I. *et al.* Pseudodynamic analysis of heart tube formation in the mouse reveals strong regional variability and early left–right asymmetry. *Nat. Cardiovasc. Res.* **1**, 504–517 (2022).
7. Jerome, L. A. & Papaioannou, V. E. DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nat. Genet.* **27**, 286–291 (2001).
8. Bruneau, B. G. *et al.* Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Dev. Biol.* **211**, 100–108 (1999).
9. Bruneau, B. G. *et al.* A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell* **106**, 709–721 (2001).
10. Parisot, P., Mesbah, K., Théveniau–Ruissy, M. & Kelly, R. G. Tbx1, subpulmonary myocardium and conotruncal congenital heart defects. *Birt. Defects Res. A. Clin. Mol. Teratol.* **91**, 477–484 (2011).
11. Xie, L. *et al.* Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation. *Dev. Cell* **23**, 280–291 (2012).
12. Zhou, L. *et al.* Tbx5 and Osr1 interact to regulate posterior second heart field cell cycle progression for cardiac septation. *J. Mol. Cell. Cardiol.* **85**, 1–12 (2015).
13. Nomaru, H. *et al.* Single cell multi-omic analysis identifies a Tbx1-dependent multilineage primed population in murine cardiopharyngeal mesoderm. *Nat. Commun.* **12**, 6645 (2021).
14. De Bono, C. *et al.* T-box genes and retinoic acid signaling regulate the segregation of arterial and venous pole progenitor cells in the murine second heart field. *Hum. Mol. Genet.* **27**, 3747–3760 (2018).
15. Rana, M. S. *et al.* Tbx1 coordinates addition of posterior second heart field progenitor cells to the arterial and venous poles of the heart. *Circ. Res.* **115**, 790–799 (2014).
16. Chen, L., Fulcoli, F. G., Tang, S. & Baldini, A. Tbx1 Regulates Proliferation and Differentiation of Multipotent Heart Progenitors. *Circ. Res.* **105**, 842–851 (2009).
17. Francou, A., Saint-Michel, E., Mesbah, K. & Kelly, R. G. TBX1 regulates epithelial polarity and dynamic basal filopodia in the second heart field. *Development* **141**, 4320–4331 (2014).

18. Cortes, C., Francou, A., De Bono, C. & Kelly, R. G. Epithelial Properties of the Second Heart Field. *Circ. Res.* **122**, 142–154 (2018).
19. Francou, A., De Bono, C. & Kelly, R. G. Epithelial tension in the second heart field promotes mouse heart tube elongation. *Nat. Commun.* **8**, 14770 (2017).
20. Li, D., Angermeier, A. & Wang, J. Planar cell polarity signaling regulates polarized second heart field morphogenesis to promote both arterial and venous pole septation. *Dev. Camb. Engl.* **146**, dev181719 (2019).
21. Li, D. *et al.* Spatial regulation of cell cohesion by Wnt5a during second heart field progenitor deployment. *Dev. Biol.* **412**, 18–31 (2016).
22. Ramsbottom, S. A. *et al.* Vangl2-Regulated Polarisation of Second Heart Field-Derived Cells Is Required for Outflow Tract Lengthening during Cardiac Development. *PLOS Genet.* **10**, e1004871 (2014).
23. Soh, B.-S. *et al.* N-cadherin prevents the premature differentiation of anterior heart field progenitors in the pharyngeal mesodermal microenvironment. *Cell Res.* **24**, 1420–1432 (2014).
24. Arriagada, C. & Astrof, S. The mesodermal source of fibronectin is required for heart morphogenesis and cardiac outflow tract elongation by regulating cell shape, polarity, and mechanotransduction in the second heart field. 2022.10.28.514299 Preprint at https://doi.org/10.1101/2022.10.28.514299 (2023).
25. Mao, Y. & Baum, B. Tug of war—The influence of opposing physical forces on epithelial cell morphology. *Dev. Biol.* **401**, 92–102 (2015).
26. Rohlf, F. J. Morphometrics. *Annu. Rev. Ecol. Syst.* **21**, 299–316 (1990).
27. Mitteroecker, P. & Schaefer, K. Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. *Am. J. Biol. Anthropol.* **178**, 181–210 (2022).
28. Andrews, T. G. R., Pönisch, W., Paluch, E. K., Steventon, B. J. & Benito-Gutierrez, E. Single-cell morphometrics reveals ancestral principles of notochord development. *Development* **148**, (2021).
29. Aigouy, B. *et al.* Cell Flow Reorients the Axis of Planar Polarity in the Wing Epithelium of Drosophila. *Cell* **142**, 773–786 (2010).
30. Rubin, S. *et al.* Application of 3D MAPs pipeline identifies the morphological sequence chondrocytes undergo and the regulatory role of GDF5 in this process. *Nat. Commun.* **12**, 5363 (2021).
31. Herbert, S. *et al.* LocalZProjector and DeProj: a toolbox for local 2D projection and accurate morphometrics of large 3D microscopy images. *BMC Biol.* **19**, 136 (2021).
32. Aigouy, B., Umetsu, D. & Eaton, S. Segmentation and Quantitative Analysis of Epithelial Tissues. *Methods Mol. Biol. Clifton NJ* **1478**, 227–239 (2016).
33. Kong, W. *et al.* Experimental validation of force inference in epithelia from cell to tissue scale. *Sci. Rep.* **9**, 14647 (2019).

902 34. Pedregosa, F. *et al.* Scikit-learn: Machine Learning in Python. *Mach. Learn. PYTHON.*
903 35. Nitzan, M., Karaiskos, N., Friedman, N. & Rajewsky, N. Gene expression cartography. *Nature*
904 **576**, 132–137 (2019).
905 36. Jombart, T., Devillard, S., Dufour, A.-B. & Pontier, D. Revealing cryptic spatial patterns in
906 genetic variability by a new multivariate method. *Heredity* **101**, 92–103 (2008).
907 37. Gorfinkiel, N. & Martinez Arias, A. The cell in the age of the genomic revolution: Cell
908 Regulatory Networks. *Cells Dev.* **168**, 203720 (2021).
909 38. Kong, W. *et al.* Experimental validation of force inference in epithelia from cell to tissue scale.
910 *Sci. Rep.* **9**, 14647 (2019).
911 39. Ishihara, S. & Sugimura, K. Bayesian inference of force dynamics during morphogenesis. *J.*
912 *Theor. Biol.* **313**, 201–211 (2012).
913 40. Gibson, M. C., Patel, A. B., Nagpal, R. & Perrimon, N. The emergence of geometric order in
914 proliferating metazoan epithelia. *Nature* **442**, 1038–1041 (2006).
915 41. Ravichandran, Y., Goud, B. & Manneville, J.-B. The Golgi apparatus and cell polarity: Roles of
916 the cytoskeleton, the Golgi matrix, and Golgi membranes. *Curr. Opin. Cell Biol.* **62**, 104–113
917 (2020).
918 42. Burute, M. *et al.* Polarity Reversal by Centrosome Repositioning Primes Cell Scattering during
919 Epithelial-to-Mesenchymal Transition. *Dev. Cell* **40**, 168–184 (2017).
920 43. Becht, E. *et al.* Dimensionality reduction for visualizing single-cell data using UMAP. *Nat.*
921 *Biotechnol.* **37**, 38–44 (2019).
922 44. Cardoso, J.-F., Cardoso, J.-F. & Souloumiac, A. Jacobi Angles For Simultaneous Diagonalization.
923 *SIAM J. Matrix Anal. Appl.* **17**, (1996).
924 45. Thorndike, R. L. Who belongs in the family? *Psychometrika* **18**, 267–276 (1953).
925 46. Palla, G., Fischer, D. S., Regev, A. & Theis, F. J. Spatial components of molecular tissue biology.
926 *Nat. Biotechnol.* **40**, 308–318 (2022).
927 47. Steimle, J. D. & Moskowitz, I. P. TBX5: A Key Regulator of Heart Development. *Curr. Top.*
928 *Dev. Biol.* **122**, 195–221 (2017).
929 48. Baldini, A., Fulcoli, F. G. & Illingworth, E. Tbx1: Transcriptional and Developmental Functions.
930 *Curr. Top. Dev. Biol.* **122**, 223–243 (2017).
931 49. Liao, J. *et al.* Identification of downstream genetic pathways of Tbx1 in the second heart field.
932 *Dev. Biol.* **316**, 524–537 (2008).
933 50. Théveniau-Ruissy, M. *et al.* The del22q11.2 Candidate Gene Tbx1 Controls Regional Outflow
934 Tract Identity and Coronary Artery Patterning. *Circ. Res.* **103**, 142–148 (2008).
935 51. Koshiba-Takeuchi, K. *et al.* Reptilian heart development and the molecular basis of cardiac
936 chamber evolution. *Nature* **461**, 95–98 (2009).
937 52. Jerome, L. A. & Papaioannou, V. E. DiGeorge syndrome phenotype in mice mutant for the T-box
938 gene, Tbx1. *Nat. Genet.* **27**, 286–291 (2001).

939 53. Verzi, M. P., McCulley, D. J., De Val, S., Dodou, E. & Black, B. L. The right ventricle, outflow
940 tract, and ventricular septum comprise a restricted expression domain within the
941 secondary/anterior heart field. *Dev. Biol.* **287**, 134–145 (2005).
942 54. LeGoff, L., Rouault, H. & Lecuit, T. A global pattern of mechanical stress polarizes cell divisions
943 and cell shape in the growing Drosophila wing disc. *Development* **140**, 4051–4059 (2013).
944 55. Mao, Y. *et al.* Differential proliferation rates generate patterns of mechanical tension that orient
945 tissue growth. *EMBO J.* **32**, 2790–2803 (2013).
946 56. Pan, D. The hippo signaling pathway in development and cancer. *Dev. Cell* **19**, 491–505 (2010).
947 57. Halder, G. & Johnson, R. L. Hippo signaling: growth control and beyond. *Dev. Camb. Engl.* **138**,
948 9–22 (2011).
949 58. Varelas, X. The Hippo pathway effectors TAZ and YAP in development, homeostasis and
950 disease. *Development* **141**, 1614–1626 (2014).
951 59. Dupont, S. *et al.* Role of YAP/TAZ in mechanotransduction. *Nature* **474**, 179–183 (2011).
952 60. Zhao, B. *et al.* Inactivation of YAP oncoprotein by the Hippo pathway is involved in cell contact
953 inhibition and tissue growth control. *Genes Dev.* **21**, 2747–2761 (2007).
954 61. Aragona, M. *et al.* A mechanical checkpoint controls multicellular growth through YAP/TAZ
955 regulation by actin-processing factors. *Cell* **154**, 1047–1059 (2013).
956 62. Turing, A. M. The chemical basis of morphogenesis. *Bull. Math. Biol.* **52**, 153–197 (1990).
957 63. Thomson, J. A. On Growth and Form. *Nature* **100**, 21–22 (1917).
958 64. Bajolle, F. *et al.* Rotation of the myocardial wall of the outflow tract is implicated in the normal
959 positioning of the great arteries. *Circ. Res.* **98**, 421–428 (2006).
960 65. Dykes, I. M. Left Right Patterning, Evolution and Cardiac Development. *J. Cardiovasc. Dev. Dis.*
961 **1**, 52–72 (2014).
962 66. Desgrange, A., Le Garrec, J.-F. & Meilhac, S. M. Left-right asymmetry in heart development and
963 disease: forming the right loop. *Development* **145**, dev162776 (2018).
964 67. Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N. & Izumo, S. The cardiac homeobox gene
965 Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Dev.*
966 *Camb. Engl.* **126**, 1269–1280 (1999).
967 68. de Soysa, T. Y. *et al.* Single-cell analysis of cardiogenesis reveals basis for organ level
968 developmental defects. *Nature* **572**, 120–124 (2019).
969 69. Harrower, M. & Brewer, C. A. ColorBrewer.org: An Online Tool for Selecting Colour Schemes
970 for Maps. *Cartogr. J.* **40**, 27–37 (2003).
971 70. Crameri, F., Shephard, G. E. & Heron, P. J. The misuse of colour in science communication. *Nat.*
972 *Commun.* **11**, 5444 (2020).
973 71. Adrien Hallou, Ruiyang He, Benjamin D. Simons, & Bianca Dumitrascu. A computational
974 pipeline for spatial mechano-transcriptomics. *bioRxiv* 2023.08.03.551894 (2023)
975 doi:10.1101/2023.08.03.551894.

Acknowledgements

The project leading to this publication has received funding from France 2030, the French Government program managed by the French National Research Agency (ANR-16-CONV-0001) and from the Excellence Initiative of Aix-Marseille University A*MIDEX (Turing Centre for Living Systems (P.V., C.G.). We are grateful to the Agence National pour la Recherche (Heartbox ANR-18-CE13-011 and Heartbound ANR-22-CE13 projects (R.G.K.), Fondation pour la Recherche Médicale DEQ20150331717 (R.G.K.) and FDT202304016591 (C.G.), AFM-Telethon and Fondation Leducq Transatlantic Network of Excellence 15CVD01 (R.G.K.) for financial support and the France-BioImaging/PICsL infrastructure (ANR-10-INSB-04-01).

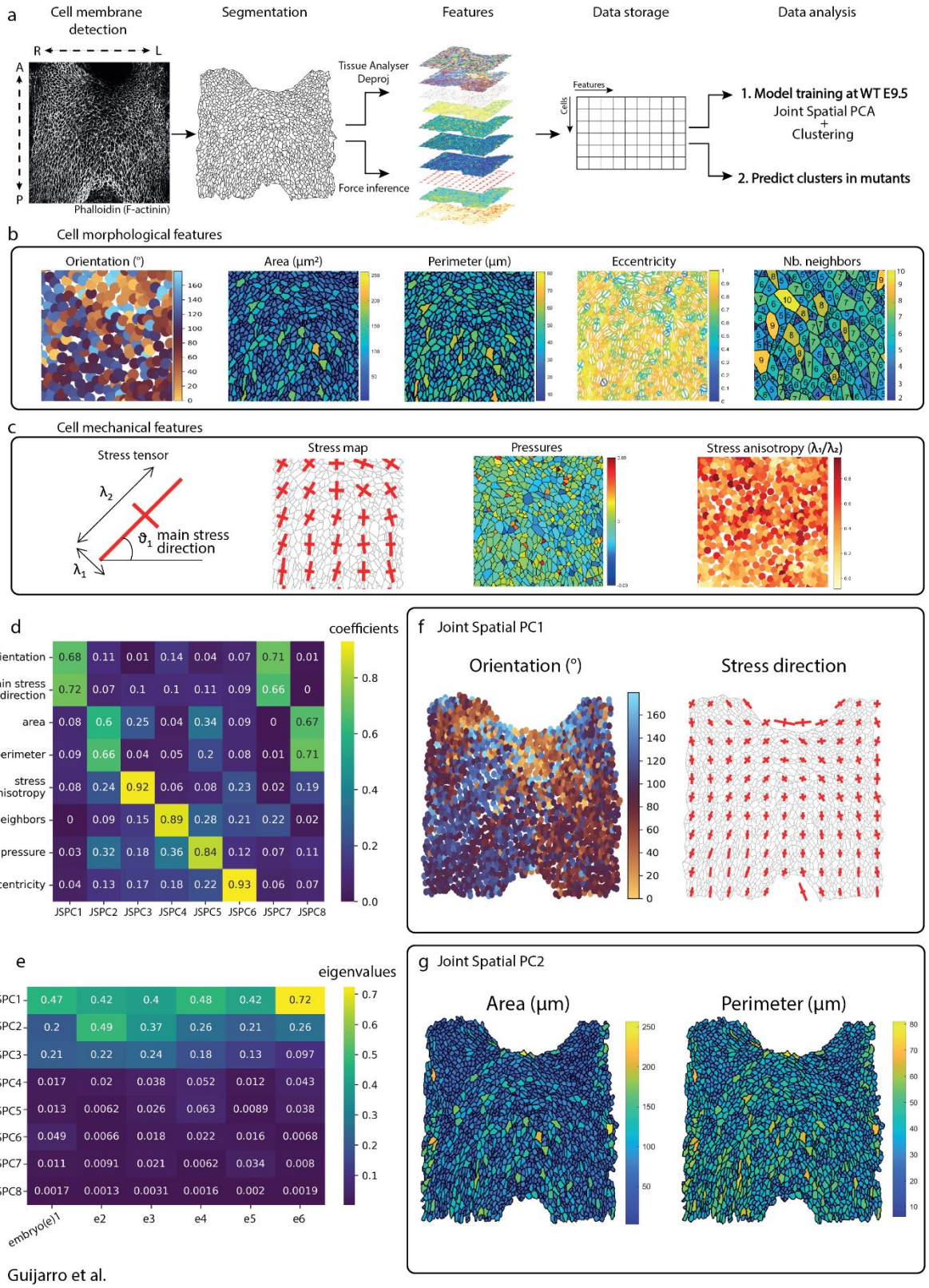
Author contributions

C.G., P.V. and R.G.K. designed experiments. Experiments were performed by C.G. C.G, S.S., B.A and R.C. developed codes for data analysis and visualization. C.G. and S.S. performed the analysis and quantifications. C.G., S.S., R.G.K. and P.V. wrote the manuscript. All authors provided intellectual input, read and approved the manuscript. R.G.K. and P.V. managed the project.

Competing interests

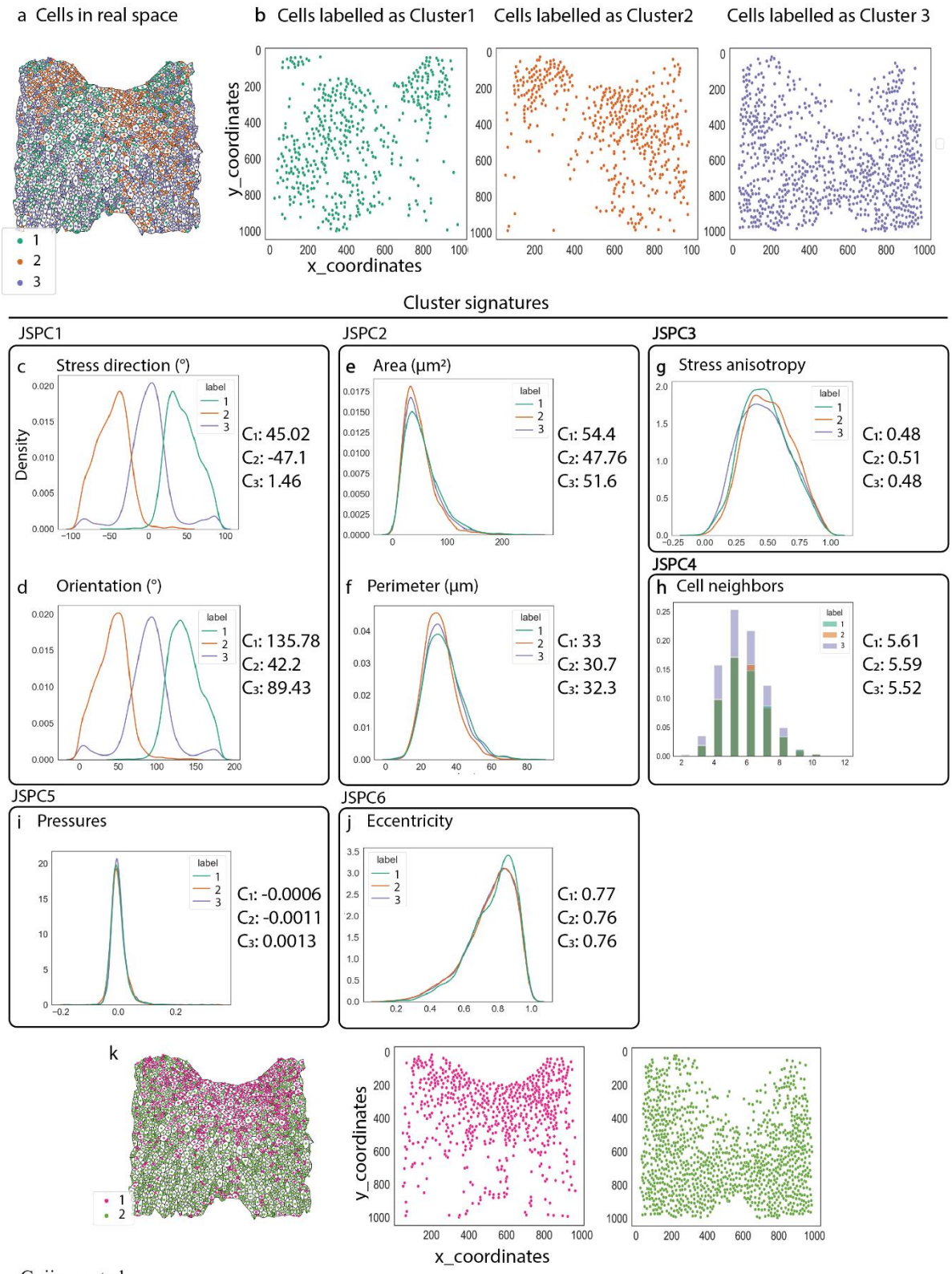
The authors declare no competing interests.

Figure 1



Guijarro et al.

Figure 2



Guijarro et al.

Figure 3

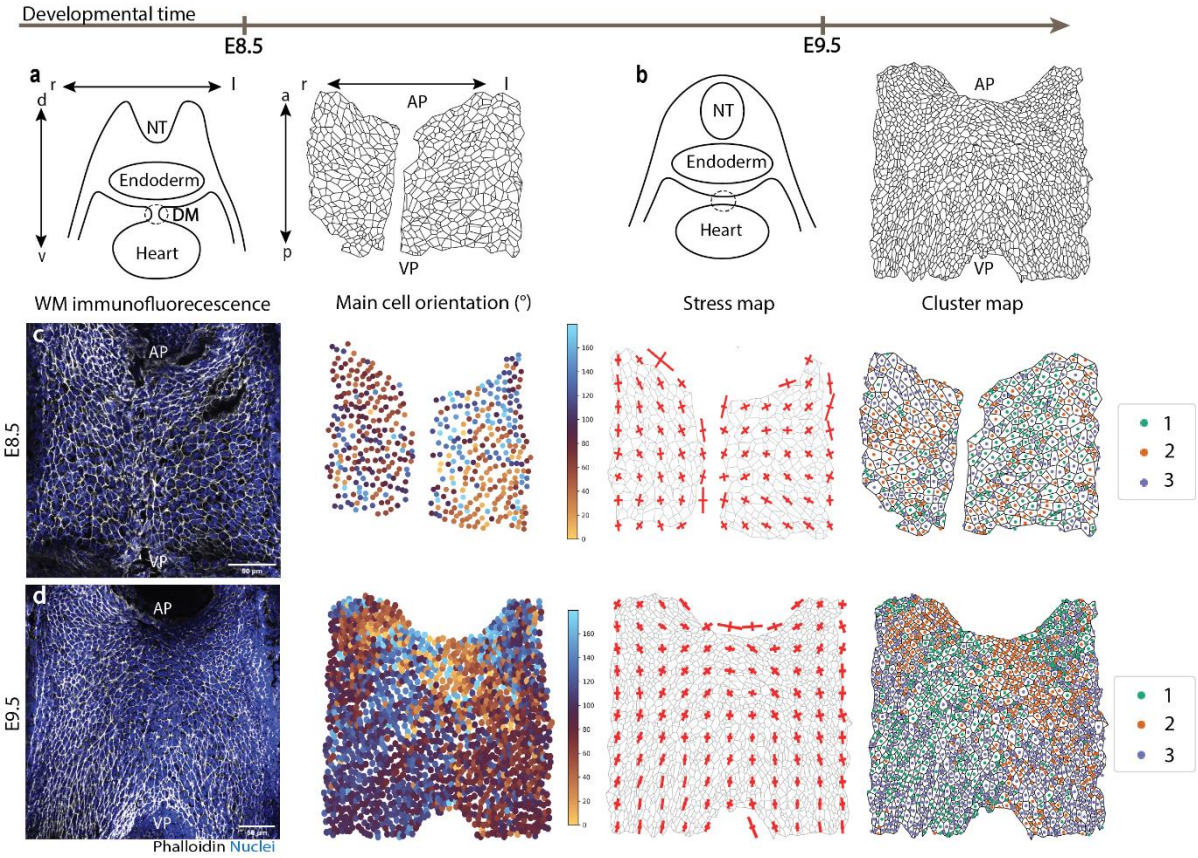
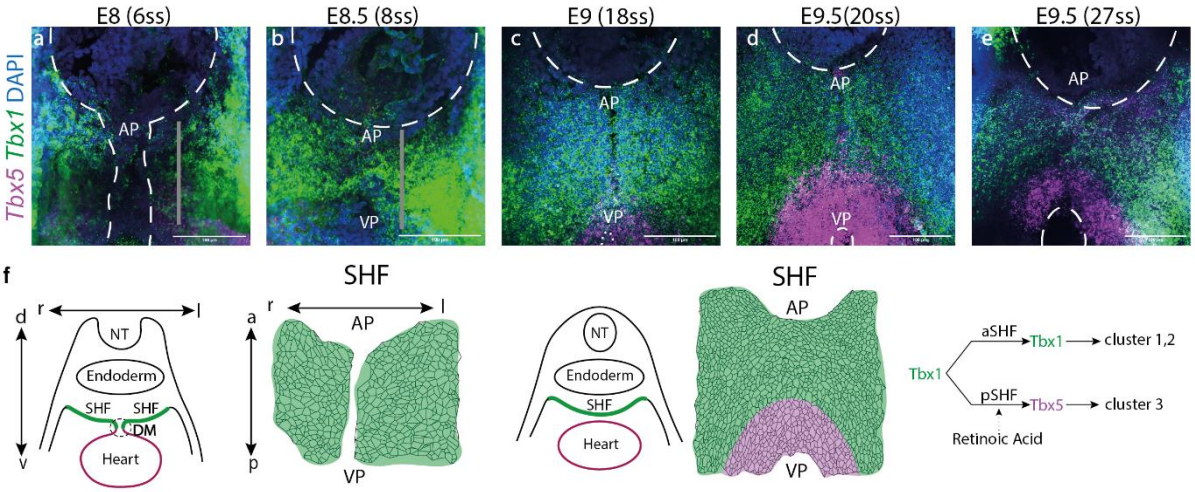

Developmental time
E8.5
E9.5
a
r
l
d
v
NT
Endoderm
DM
Heart
a
p
AP
VP
b
NT
Endoderm
Heart
AP
VP
WM immunofluorecescence
Main cell orientation (°)
Stress map
Cluster map
c
E8.5
AP
VP
d
E9.5
AP
VP
Phalloidin Nuclei
1
2
3

Figure 4



Guijarro et al.

Figure 5

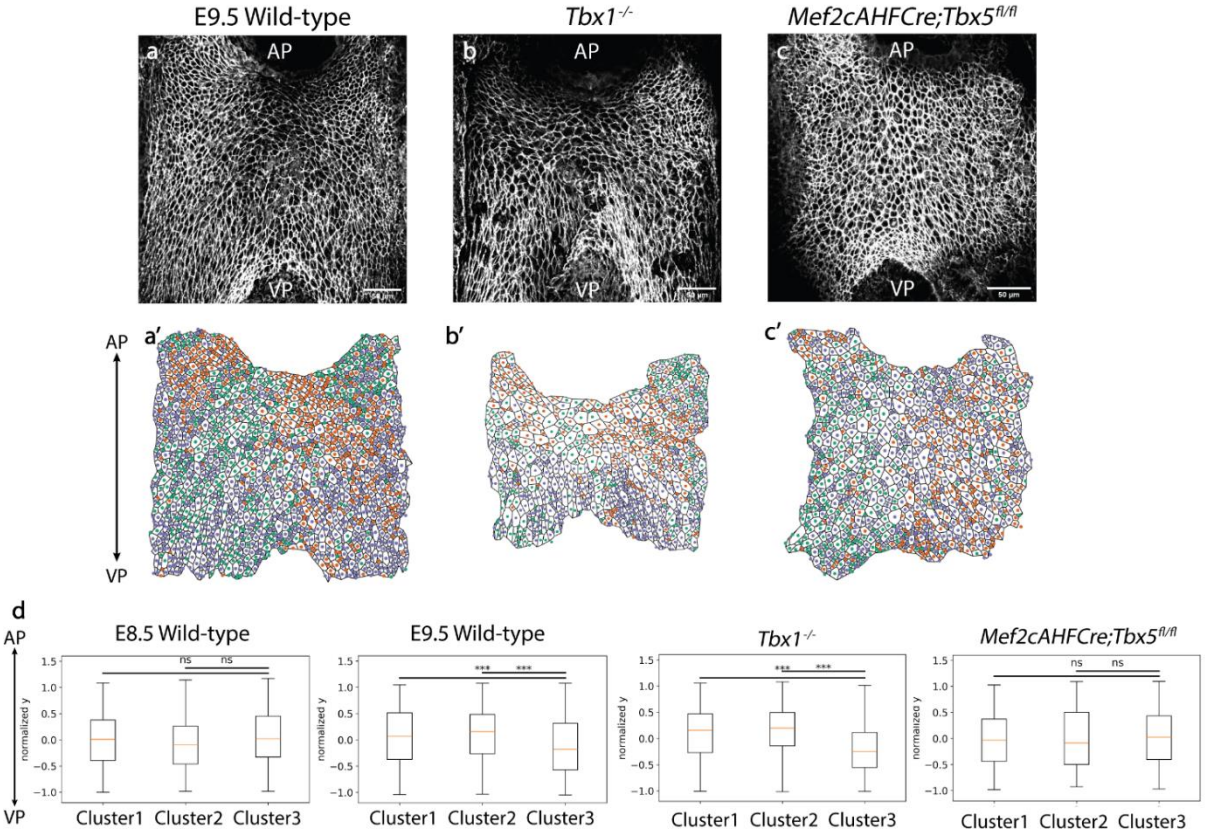

E9.5 Wild-type
Tbx1 -/-
Mef2cAHFCre;Tbx5 fl/fl
a
b
c
AP
VP
a′
b′
c′
d
E8.5 Wild-type
E9.5 Wild-type
Tbx1 -/-
Mef2cAHFCre;Tbx5 fl/fl
normalized y
ns
***
Cluster1
Cluster2
Cluster3

Figure 6

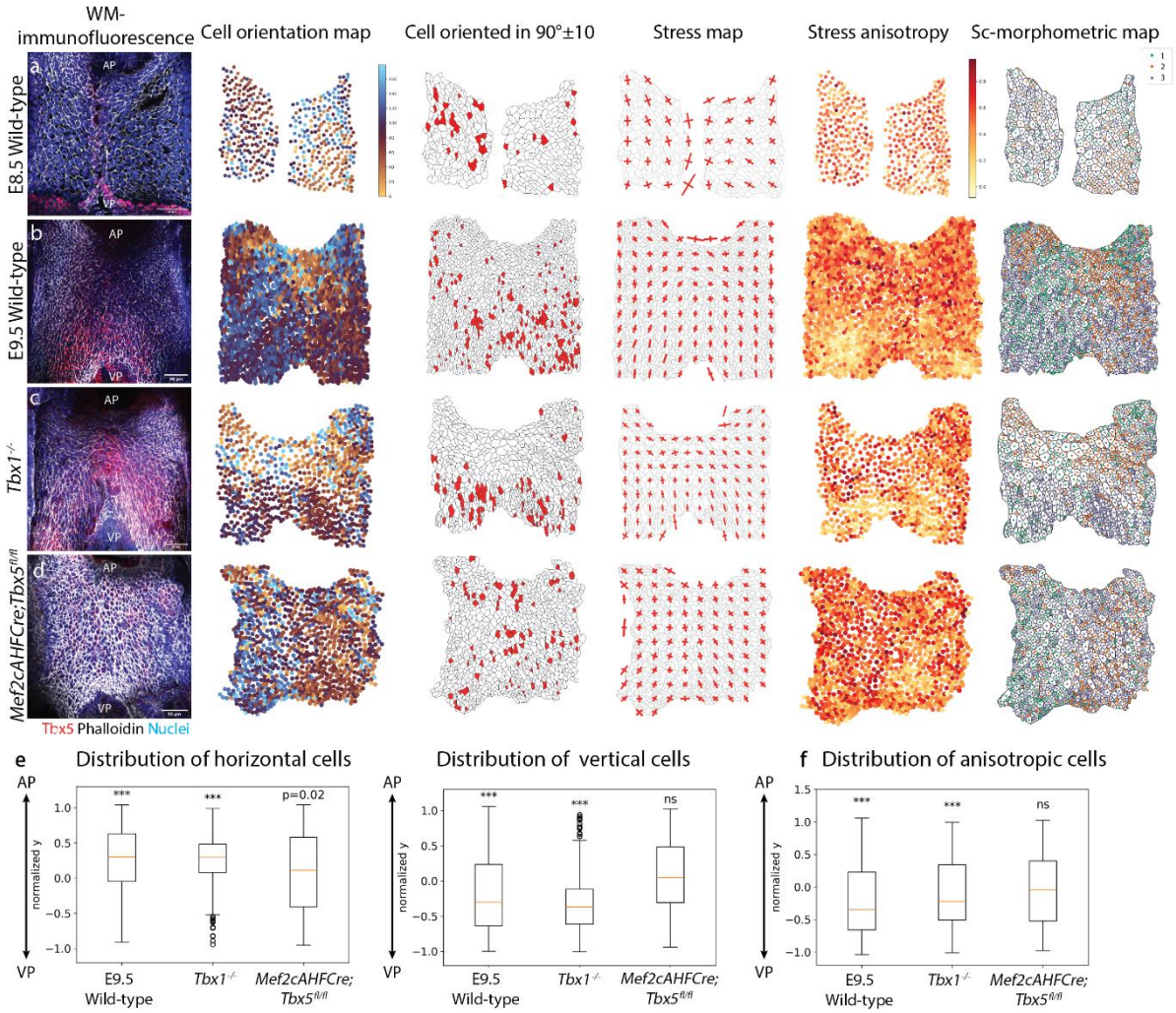
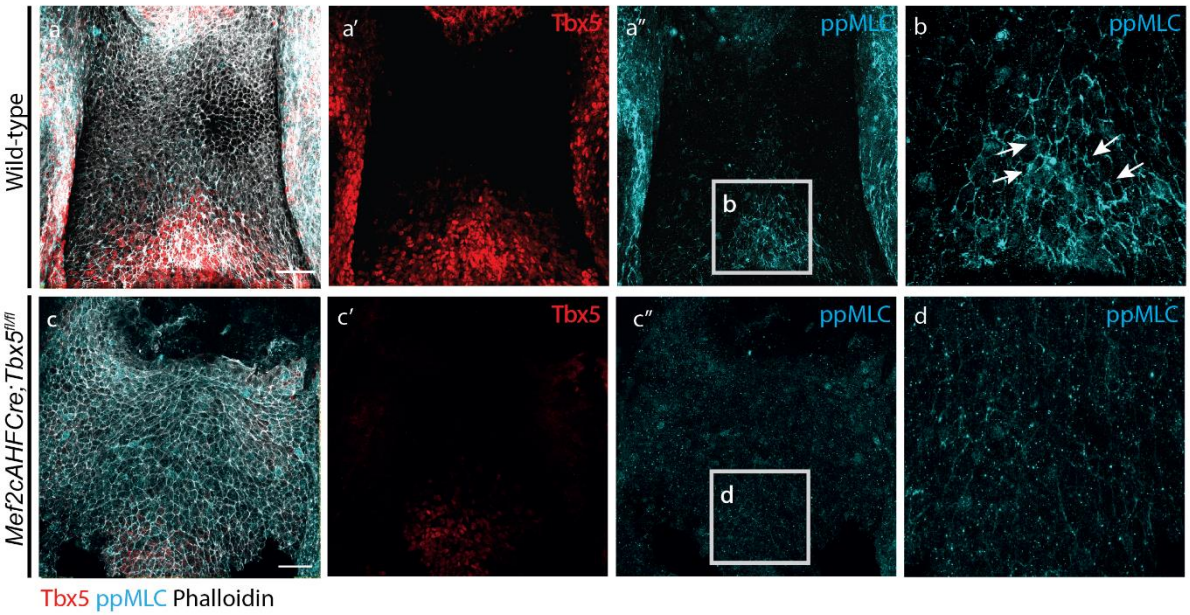

WM-immunofluorescence
Cell orientation map
Cell oriented in 90°±10
Stress map
Stress anisotropy
Sc-morphometric map
E8.5 Wild-type
E9.5 Wild-type
Tbx1-/-
Mef2cAHFCre;Tbx5fl/fl
AP
VP
Tbx5 Phalloidin Nuclei
e Distribution of horizontal cells
Distribution of vertical cells
f Distribution of anisotropic cells
normalized y
***
p=0.02
ns
E9.5 Wild-type
Tbx1-/-
Mef2cAHFCre; Tbx5fl/fl

Figure 7



Guijarro et al.

Supplementary figures - Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart Field

Supplementary Information

Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart field

Clara Guijarro[1,2,3 +], Solène Song[2,3 +], Benoit Aigouy[1], Raphael Clement[1], Paul Villoutreix[2,3 *], Robert G. Kelly[1 *]

[1] Aix-Marseille Université, CNRS UMR 7288, IBDM, Turing Centre for Living systems, Marseille, France

[2] Aix-Marseille Université, LIS, UMR 7020, Turing Centre for Living systems, Marseille, France

[3] Aix-Marseille Université, MMG, Inserm U1251, Turing Centre for Living systems, Marseille, France

[+] These authors contributed equally

[*] These authors jointly supervised this work

Addresses for correspondence:
Robert G. Kelly,
Aix-Marseille University, CNRS UMR 7288,
IBDM, Luminy Campus Case 907,
73288 Marseille Cedex 9, France
E-mail: Robert.Kelly@univ-amu.fr

Paul Villoutreix,
Marseille Medical Genetics Inserm U 1251,
Turing Centre for Living Systems,
Aix-Marseille University, La Timone Campus,
13005 Marseille, France
E-mail: Paul.Villoutreix@univ-amu.fr

Fig. S1: Characterization of the nucleus-to-Golgi polarity in cells of the dorsal pericardial wall. **a.** The E9.5 dorsal pericardial wall visualized after 3D (left) and apical (right) segmentation, showing the axes and coordinates referred to in this study. Phalloidin (red), Golgi (green), nuclei (blue). **b.** Maximum intensity projection of the apical part of the DPW epithelium, high magnification of the epithelium corresponding to the boxed region and scheme of the calculation of the nucleus-to Golgi polarity used for single-cell morphometrics (the angle formed between the vector from the center of the nucleus to the center of the Golgi and the left-right position (X) within the tissue. **c.** Stack count-plots of the localization of nucleus-to Golgi polarity in 3D. Golgi (green), nuclei (blue), nucleus-to-golgi vector (red) **d.** Spatial distribution of the vectors of each cell of the epithelium color coded according to the angle as already described in Rubin et al., 2020[30]. A: anterior; P: posterior; R: right; L: left. Source data are provided as a Source Data file. E8.5 N=5 embryos, E8.5 n=2141cells ; E9.5 N=4 embryos, E9.5 n= 6140 cells.

Fig. S2: Principal component analysis of cells in the dorsal pericardial wall of E9.5 wild-type embryos. a. Heatmap showing the coefficients of each feature for each principal component (PCA). **b.** Bar plot showing the PCA explained variance. **c**. Moran index for each classical PC on E9.5 wild-type embryos. **d.** PCA visualization using PC1 as x-axis and PC2 as y-axis and color coded depending on the dispersion of the values for each feature. Source data are provided as a Source Data file. Error bars represent SD. N = 6 embryos, n = 8618 cells.

Fig. S3: Feature correlations. a. Correlation matrix. **b.** Correlation plots per features (only the most correlated are shown). **c.** Correlation graphs between all features. Source data are provided as a Source Data file.

Fig. S4: Cluster signatures in E9.5 wild-type. a. Cluster maps and kernel density distribution plots for stress direction and orientation using the all the raw values (control). **b.** Cluster maps and kernel density distribution plots using the AP-symmetrized values of orientation and stress direction for three and two clusters. **c.** Optimal number of clusters according to Akaike Information Criterion (AIC). N = 6 embryos, n = 8618 cells. Source data are provided as a Source Data file.

Fig. S5: Morphometrics reproducibility in E9.5 wild-type embryos. a-f. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (Tbx5 in red, phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map, stress anisotropy map and cell cluster distribution shown for three clusters and two clusters after symmetrizing for cell orientation and strezzss direction. Embryos a-d have >20 somites and embryos e and f < 20 somites. Embryo f has not been stained with TBX5 antibody. N = 6 embryos, n = 8618 cells.

Fig. S6: Morphometrics in E8.5 wild-type embryos. a-d. Showing for each embryo, in the following order, the apical projection of the dorsal pericardial wall after whole-mount immunofluorescence (TBX5 in red, phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map, stress anisotropy map and cell cluster distribution shown for three clusters and two clusters after symmetrizing for cell orientation and stress direction. N = 6 embryos, n = 2994 cells.

Fig. S7: *Tbx1* and *Tbx5* mRNA distribution and quantification from E8 to E9.5. a-e. Frontal whole-mount views of the DPW of wildtype embryos after RNAscope fluorescent in situ hybridization showing nuclei (blue), *Tbx1* (green) and *Tbx5* (magenta) mRNA distribution between E8 (6 somites) and E9.5 (25 somites) wild-type embryos, before and after dorsal mesocardial breakdown. The white line delimits the arterial pole of the heart (AP). Grey areas represent the area where the pixel intensity has been measured. **f.** Quantification of mRNA pixel intensity of *Tbx1* and *Tbx5*. **g.** Frontal whole-mount views of the DPW of *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos after RNAscope fluorescent in situ hybridization showing nuclei (blue), *Tbx1* (green) and *Tbx5* (magenta) mRNA distribution. Quantification areas have a width of 300 pixels. Scale bars:100µm. Source data are provided as a Source Data file. E8-E8.5 : N = 5 embryos, E9 : N = 3 embryos, E9.5 : N = 6 embryos.

Fig. S8: Morphometrics of E9.5, 8.5 wildtype and E9.5 *Tbx1* and *Tbx5* conditional mutant embryos. a-a'''. Cluster distribution in real space in wild-type E8.5 and E9.5 embryos, E9.5 *Tbx1*$^{-/-}$ and *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos. **b.** Cluster signatures and kernel density distributions of *Tbx1*$^{-/-}$, *Mef2cCreAHFCre*; *Tbx5*$^{fl/fl}$ and E8.5 wild-type embryos. Source data are provided as a Source Data file. E8.5 wild-type: N = 6 embryos, n = 2994 cells; E9.5 wild-type: N = 6 embryos, n = 8618 cells; E9.5 *Tbx1*$^{-/-}$: N = 3 embryos, n = 2551; E9.5 *Mef2cAHFCre;Tbx5*$^{fl/fl}$: N = 4 embryos, n = 2809 cells.

Fig. S9: Morphometrics reproducibility in E9.5 *Tbx1*$^{-/-}$ embryos. a-c. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (TBX5 in red, phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map, the stress anisotropy map and cell cluster distribution shown for three clusters and two clusters after symmetrizing for cell orientation and stress direction. N = 3 embryos, n = 2551 cells.

Fig. S10: Morphometrics reproducibility in E9.5 *Mef2cAHFCre;Tbx5*$^{fl/fl}$ embryos. a-c. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (TBX5 in red, phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map, the stress anisotropy map and cell cluster distribution shown for three clusters and two clusters after symmetrizing for cell orientation and stress direction. N = 4 embryos, n = 2809 cells.

Fig. S11: Morphometrics reproducibility in E9.5 *Mef2cAHFCre;Tbx5*$^{fl/+}$ embryos. a-b. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (phalloidin in red, nuclei in blue, TBX5 in white), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map and the stress anisotropy map. N = 3.

Fig. S12: Morphometrics of the dorsal pericardial wall of E9.5 embryos treated with the retinoic acid antagonist BMS493 in embryo culture. a-d. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (TBX5 in red, Golgi in green, phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map and the stress anisotropy map. N = 4 BMS493 treated embryos.

Fig. S13: Diphosphorylated Myosin light chain 2 (ppMLC) distribution in E9.5 embryos treated with retinoic acid inhibitor BMS493. a. Maximum intensity projection of the apical part of the epithelia **b'-b'''.** Magnification of the aDPW region (box b in panel a). **c'-c'''.** Magnification of the pDPW region (box c in panel a). Phalloidin (white), ppMLC (green) and TBX5 (red). N = 5 wild-type embryos; N = 4 BMS493 treated embryos.

Fig. S14: Morphometrics reproducibility in E9.5 *Nkx2.5*$^{-/-}$ embryos. a-d. Showing, for each embryo, in the following order: the apical projection of the DPW after whole-mount immunofluorescence (phalloidin in white, nuclei in blue), the cell orientation map, cells oriented horizontally (in blue) along the left-right axis of the embryo on the segmentation channel, cells oriented vertically (in red) along the anterior-posterior axis of the embryo, the stress tensor map, the cell pressure map and the stress anisotropy map. N = 4 embryos, n = 5568 cells.

Fig. S15: Cluster signatures of *Nkx2.5*$^{-/-}$ embryos. Kernel density and histogram distributions of each morphometrical feature for *Nkx2.5*$^{-/-}$ embryos. Source data are provided as a Source Data file. Cluster1 in green, cluster2 in orange, cluster3 in purple. N = 4 embryos, n = 5568 cells.

Fig. S16: Force inference curvature tests a. Map of the mean, first and second principal curvature of an E9.5 wild-type embryo. **b.** Histogram plot distribution of the absolute values for the first and second radius of curvature from all E9.5 wild-type embryos, with a threshold set at 4000 μm used for visualization purposes. **c.** Histogram plot showing the distribution of the cell radius from all E9.5 wild-type embryos. N = 6 embryos, n = 8618 cells.

Fig. S17: Cluster maps for E9.5 wild-type embryos showing each cluster independently. Scatter plots with axes representing the x-y coordinates of the cells, showing each cluster independently. Cluster1 in green, cluster2 in orange, cluster3 in purple. N = 6 embryos, n = 8618 cells.

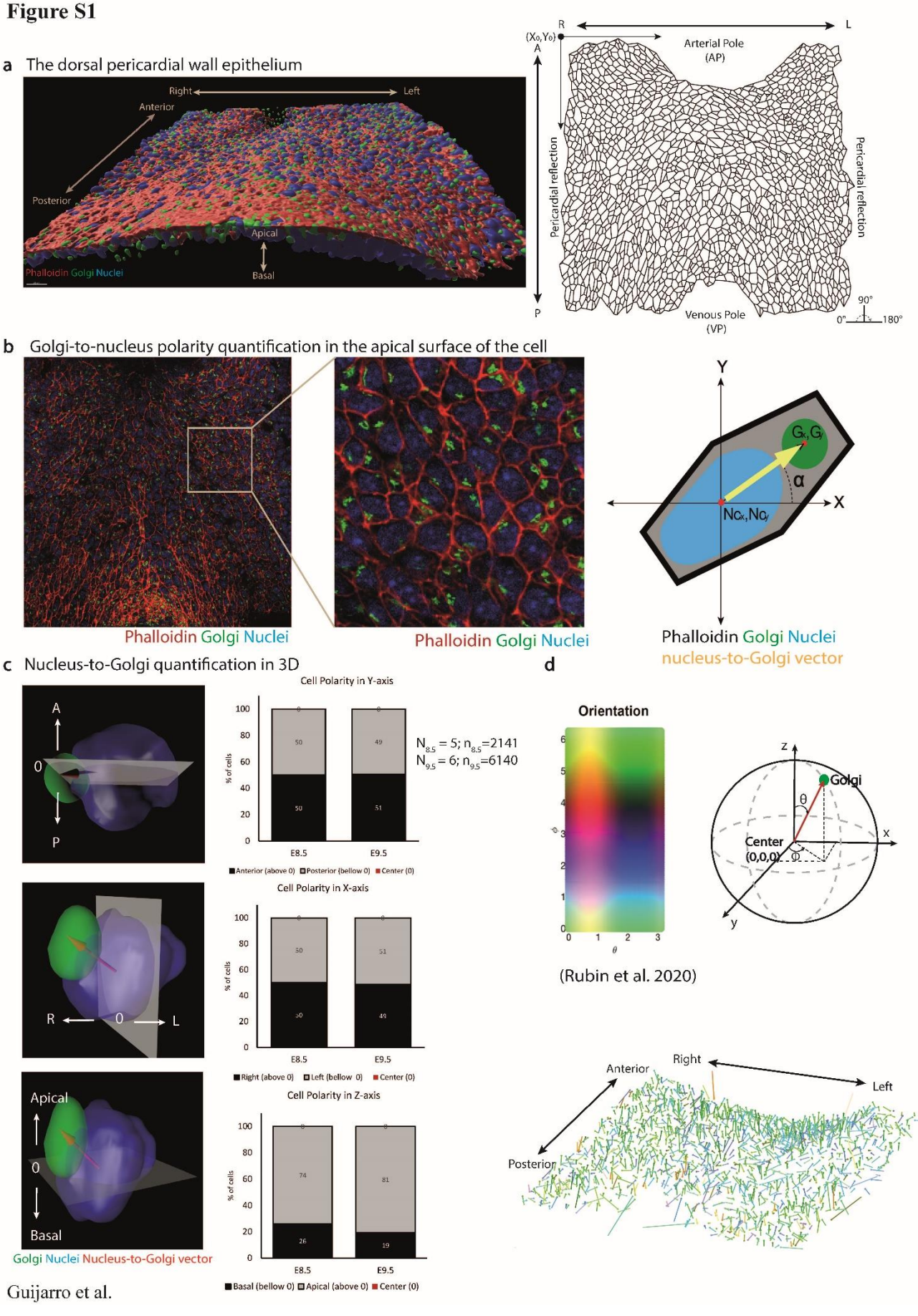
Fig. S18: Cluster maps for E8.5 wild-type embryos showing each cluster independently. Scatter plots with axes representing the x-y coordinates of the cells, showing each cluster independently. Cluster1 in green, cluster2 in orange, cluster3 in purple. N = 6 embryos, n = 2994 cells.

Fig. S19: Cluster maps for E9.5 *Mef2cAHFCre;Tbx5*$^{fl/fl}$ and *Tbx1*$^{-/-}$ embryos showing each cluster independently. Scatter plots with axes representing the x-y coordinates of the cells, showing each cluster independently. Cluster1 in green, cluster2 in orange, cluster3 in purple. *Tbx1*$^{-/-}$: N = 3 embryos, n = 2551; *Mef2cAHFCre;Tbx5*$^{fl/fl}$: N = 4 embryos, n = 2809 cells.

Fig. S20: Cluster maps for E9.5 *Nkx2.5*$^{-/-}$ embryos showing each cluster independently. Scatter plots with axes representing the x-y coordinates of the cells, showing each cluster independently. Cluster1 in green, cluster2 in orange, cluster3 in purple. N = 4 embryos, n = 5568 cells.

Supplementary Table1: Primers list.

Figure S1



Guijarro et al.

Figure S2

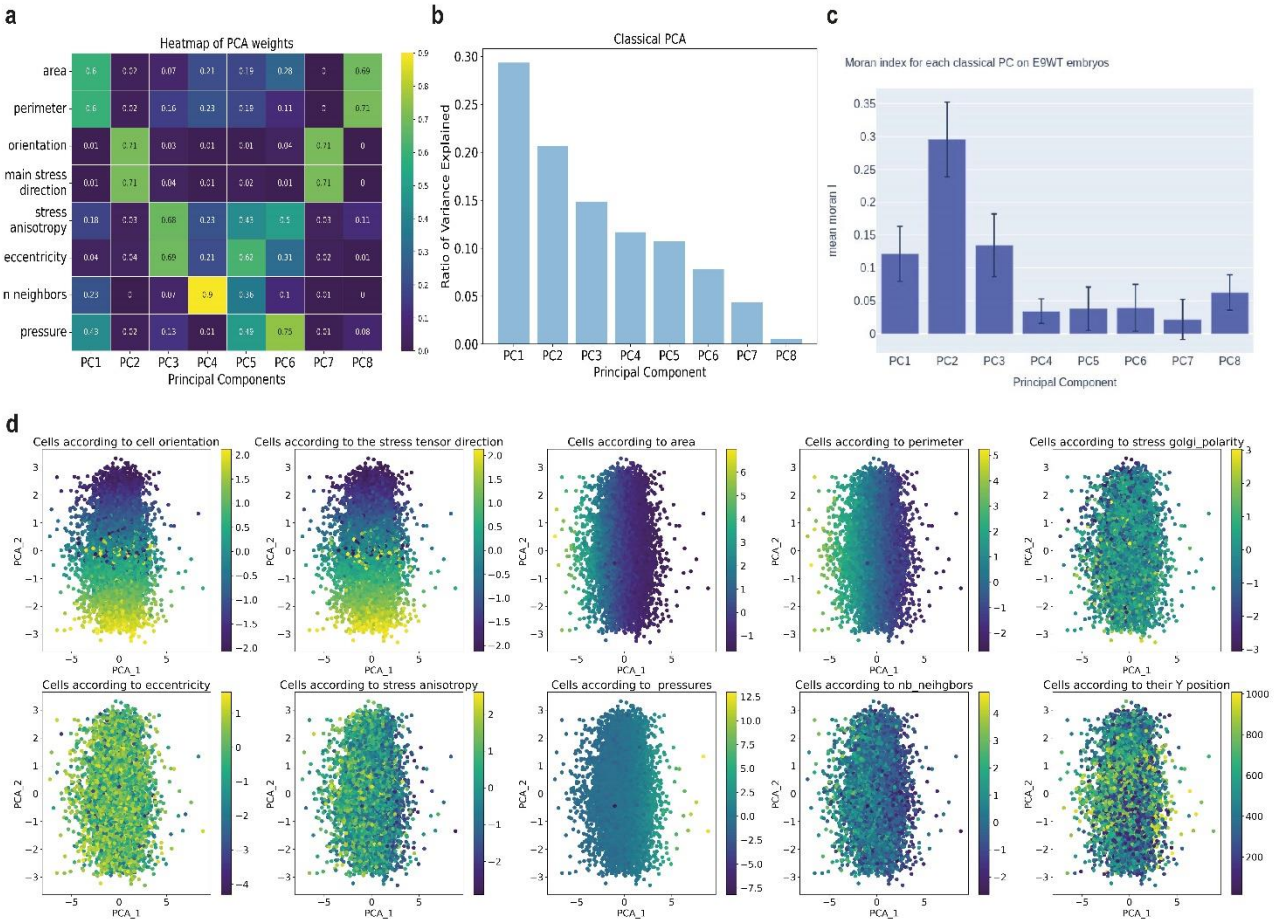
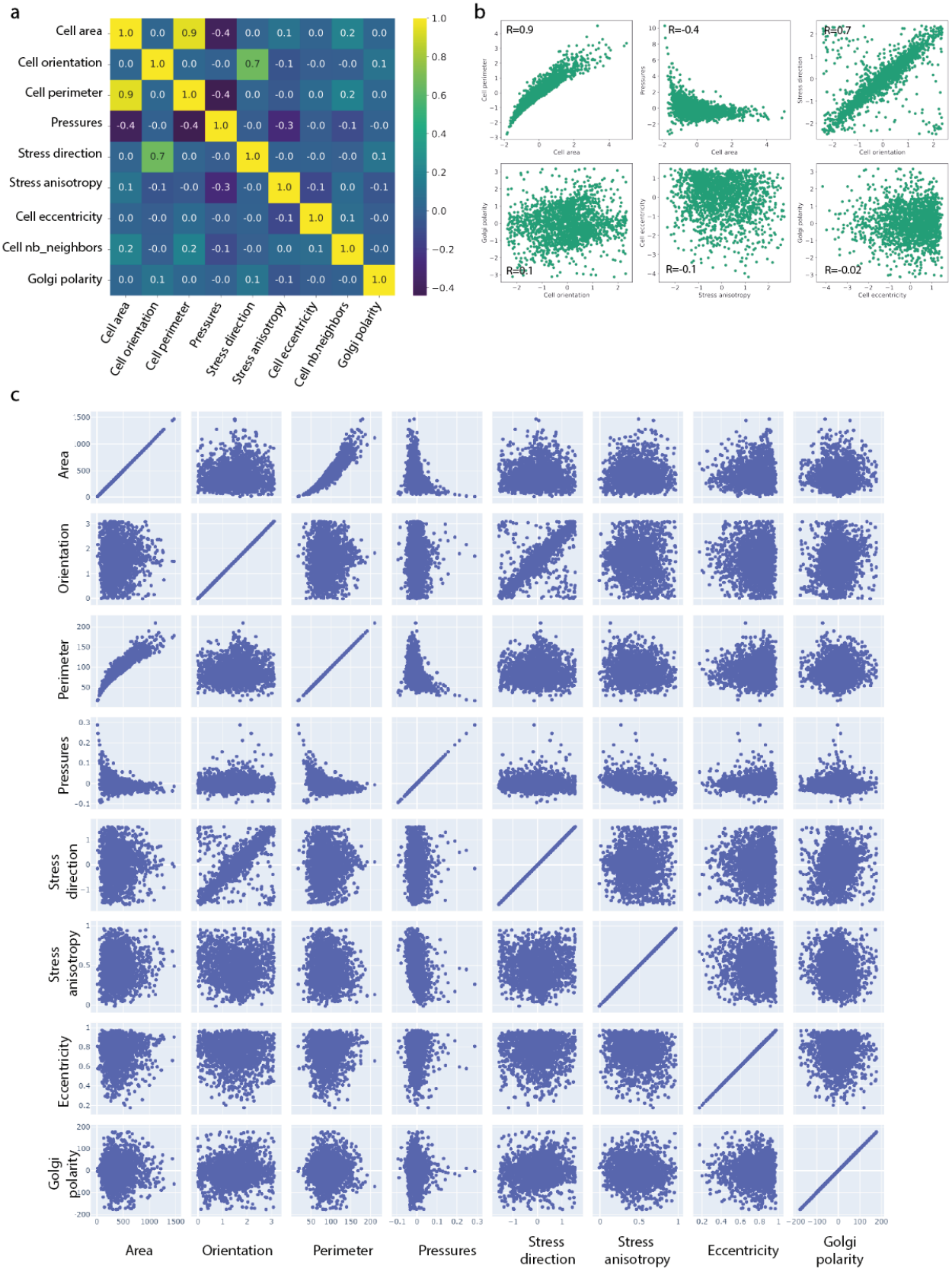
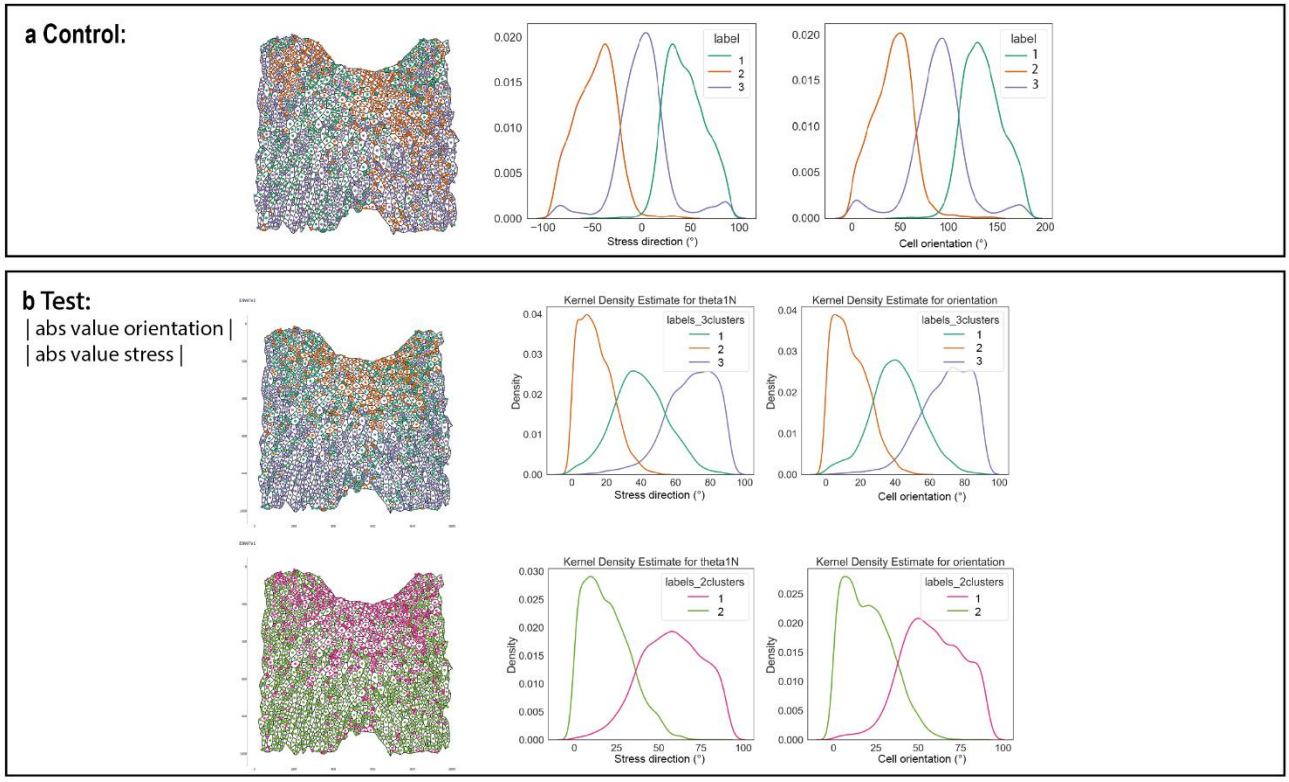

a
Heatmap of PCA weights
area
perimeter
orientation
main stress direction
stress anisotropy
eccentricity
n neighbors
pressure
PC1 PC2 PC3 PC4 PC5 PC6 PC7 PC8
Principal Components
b
Classical PCA
Ratio of Variance Explained
Principal Component
c
Moran index for each classical PC on E9WT embryos
mean moran I
Principal Component
d
Cells according to cell orientation
Cells according to the stress tensor direction
Cells according to area
Cells according to perimeter
Cells according to stress golgi_polarity
Cells according to eccentricity
Cells according to stress anisotropy
Cells according to pressures
Cells according to nb_neihgbors
Cells according to their Y position
PCA_1
PCA_2

Figure S3



Guijarro et al.

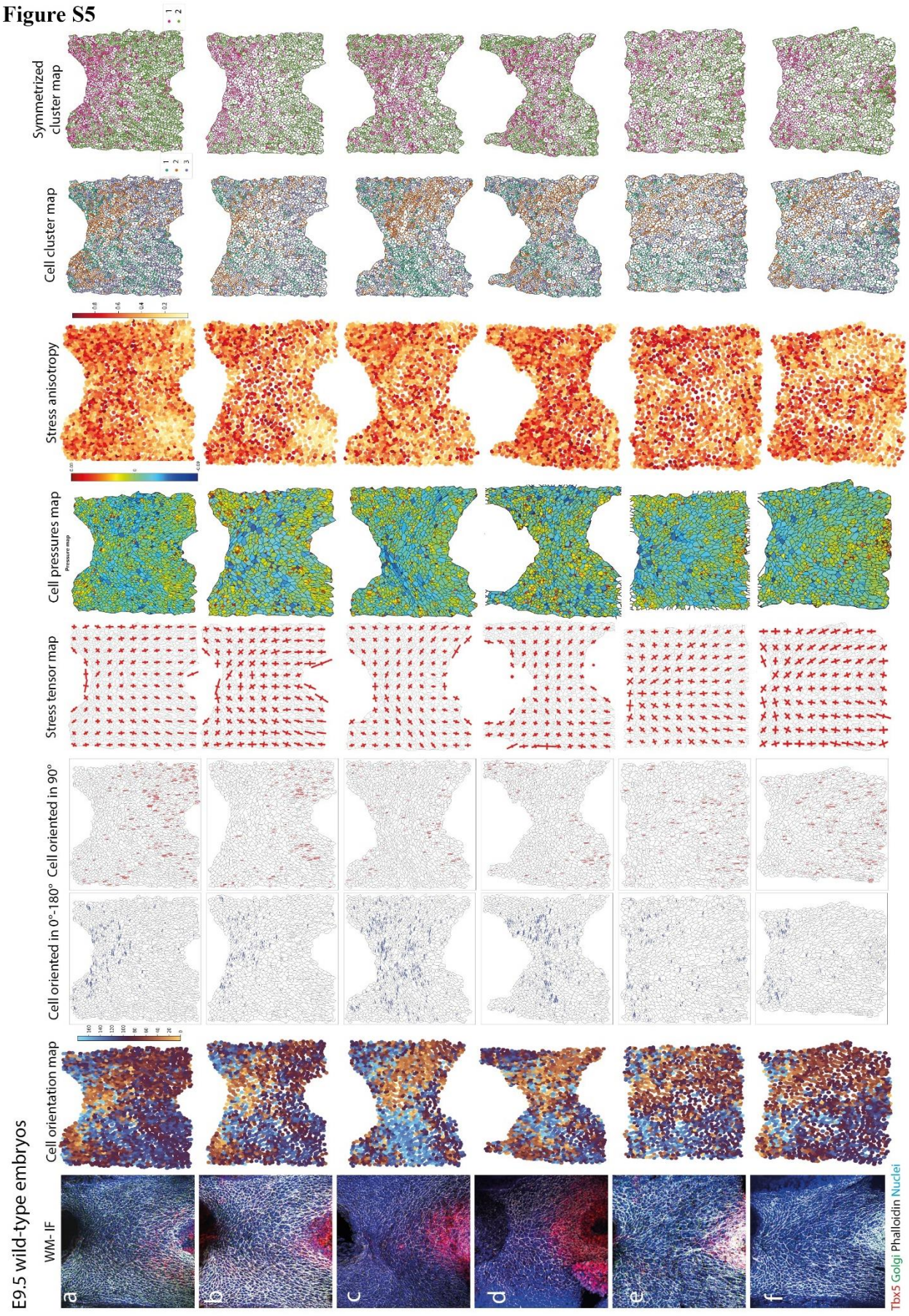
Figure S4



c

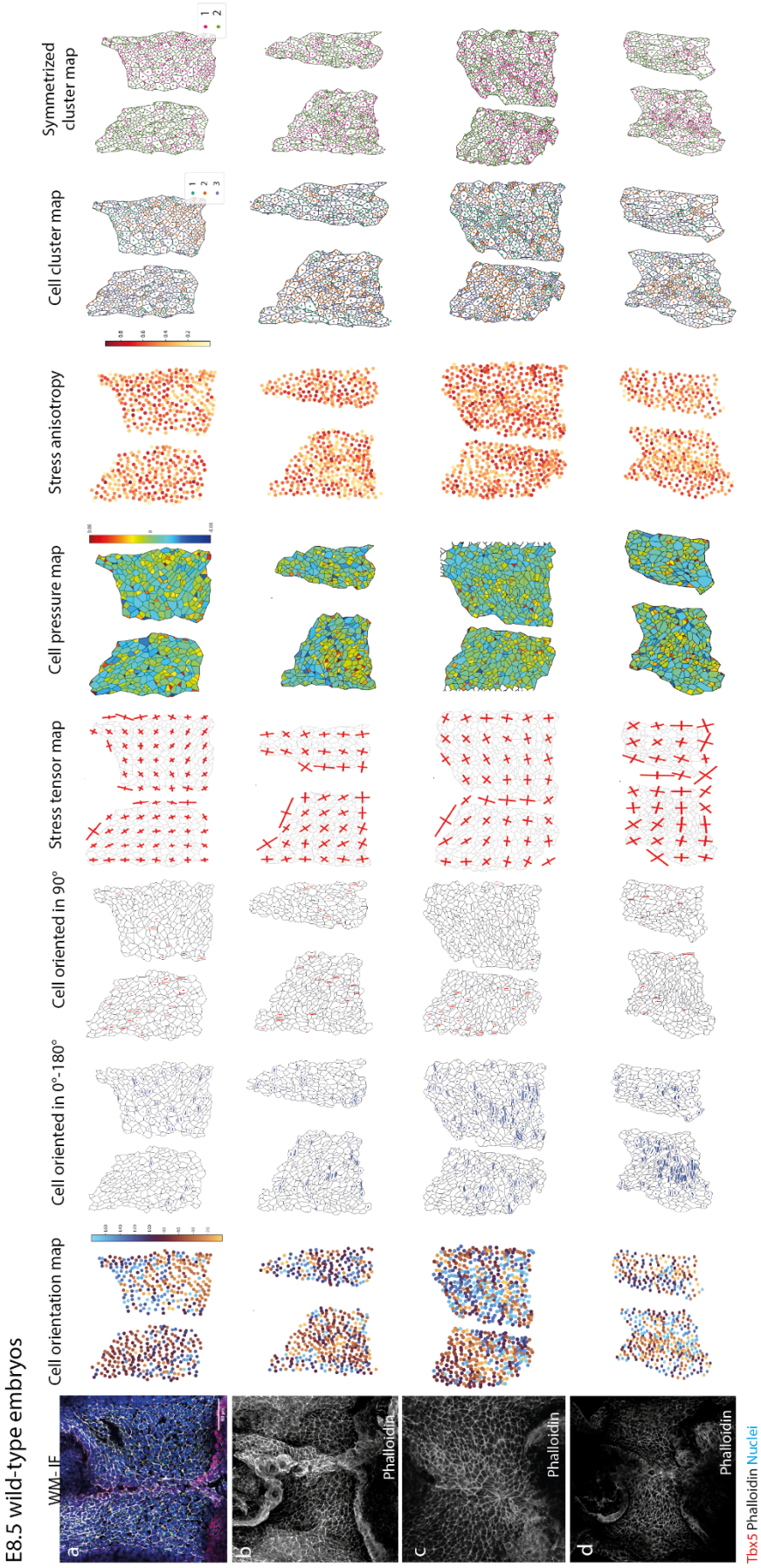
n	AIC
1	28532
2	28380
3	28366
4	28389
5	28377
6	28374
7	28382
8	28376
9	28377

Figure S5



Guijarro et al.

Figure S6



Guijarro et al.

Figure S7

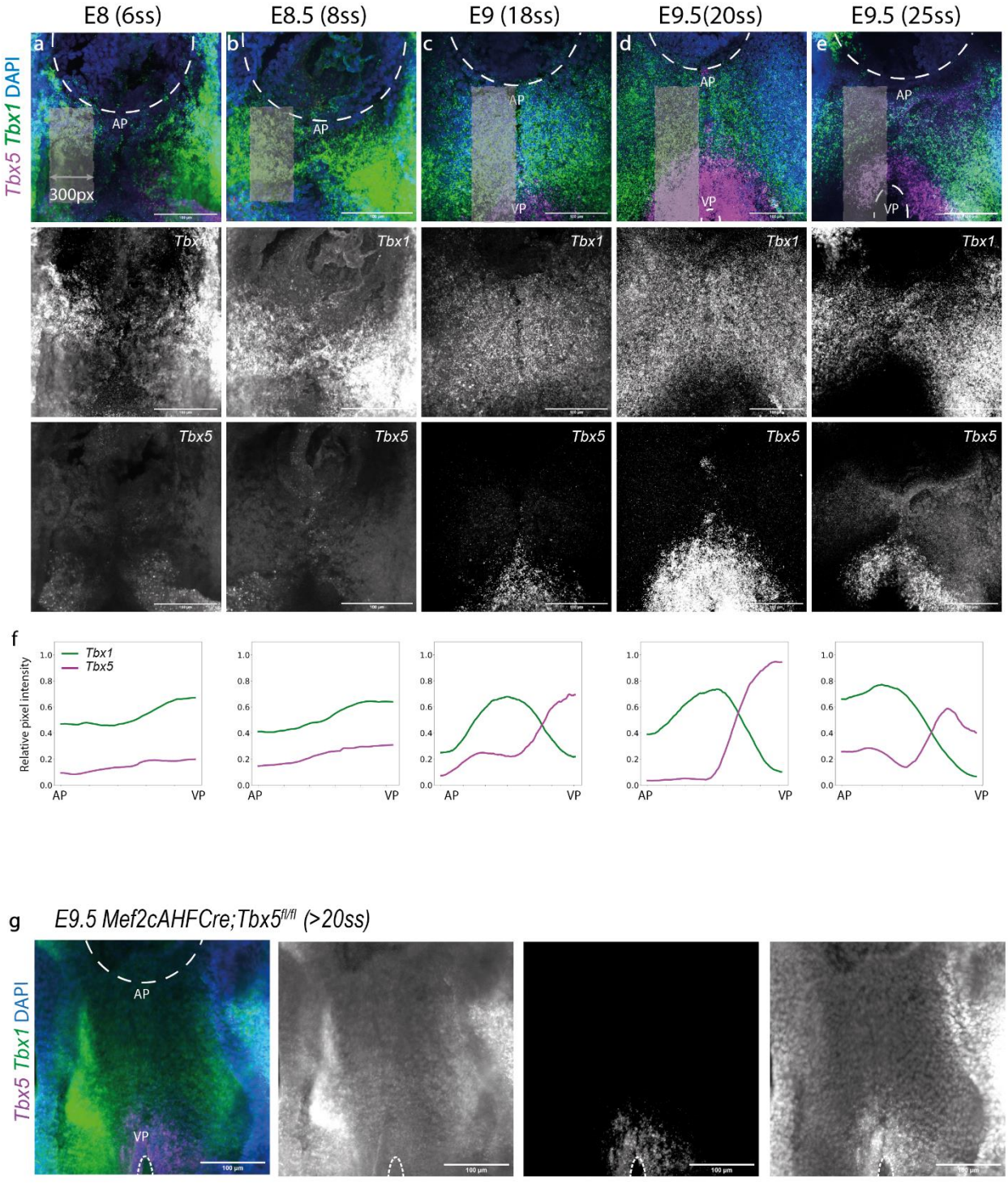

E8 (6ss)
E8.5 (8ss)
E9 (18ss)
E9.5(20ss)
E9.5 (25ss)
Tbx5 Tbx1 DAPI
AP
VP
300px
Tbx1
Tbx5
f
Relative pixel intensity
Tbx1
Tbx5
g
E9.5 Mef2cAHFCre;Tbx5fl/fl (>20ss)

Figure S8

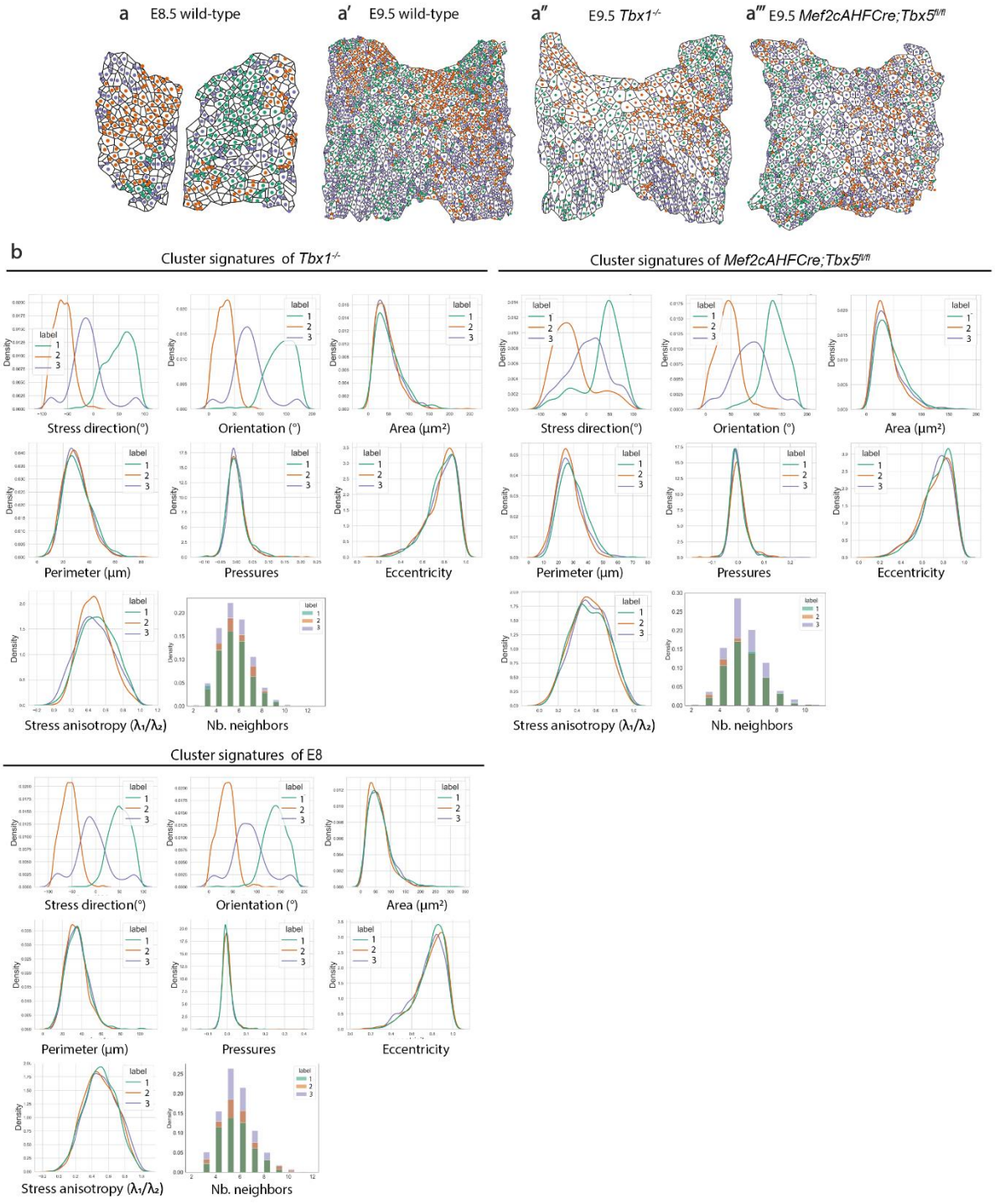
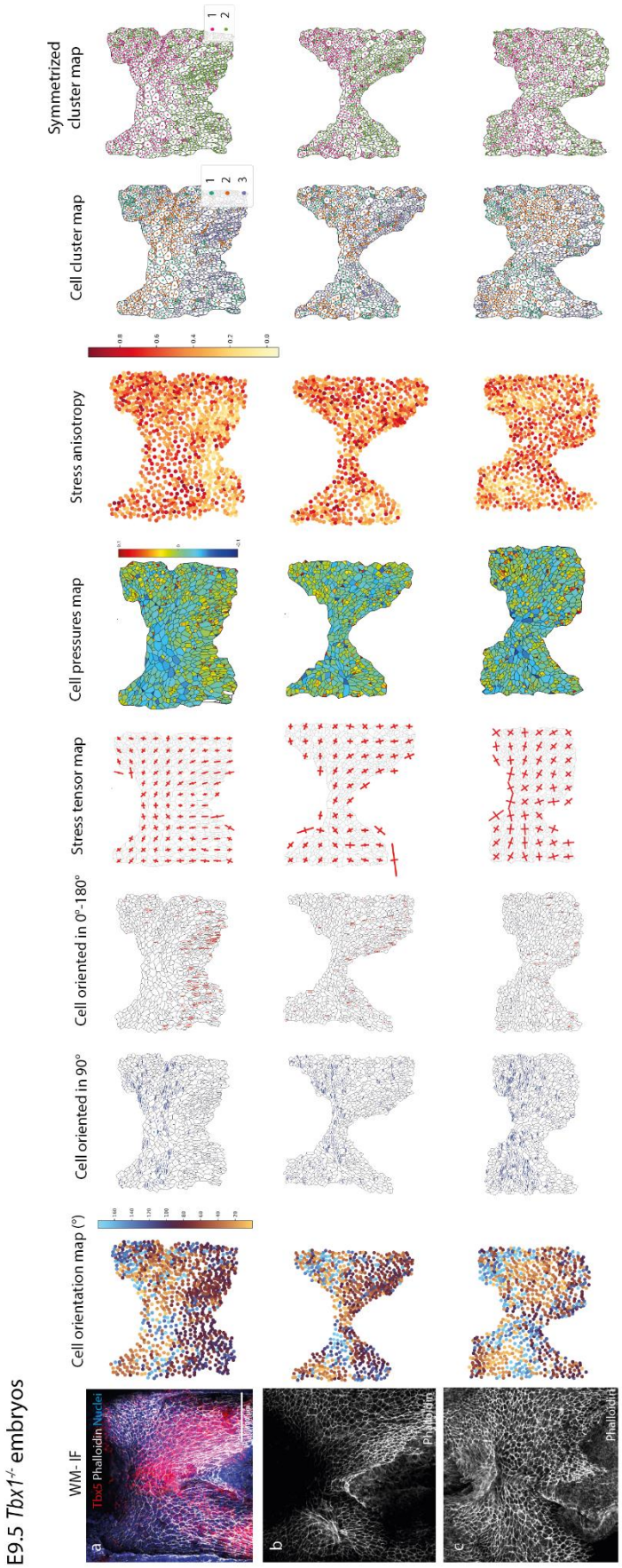

a E8.5 wild-type
a′ E9.5 wild-type
a″ E9.5 Tbx1-/-
a‴ E9.5 Mef2cAHFCre;Tbx5fl/fl
b
Cluster signatures of Tbx1-/-
Cluster signatures of Mef2cAHFCre;Tbx5fl/fl
Cluster signatures of E8
Stress direction(°)
Orientation (°)
Area (µm²)
Perimeter (µm)
Pressures
Eccentricity
Stress anisotropy (λ1/λ2)
Nb. neighbors
label
Density

Figure S9



Guijarro et al.

Figure S10

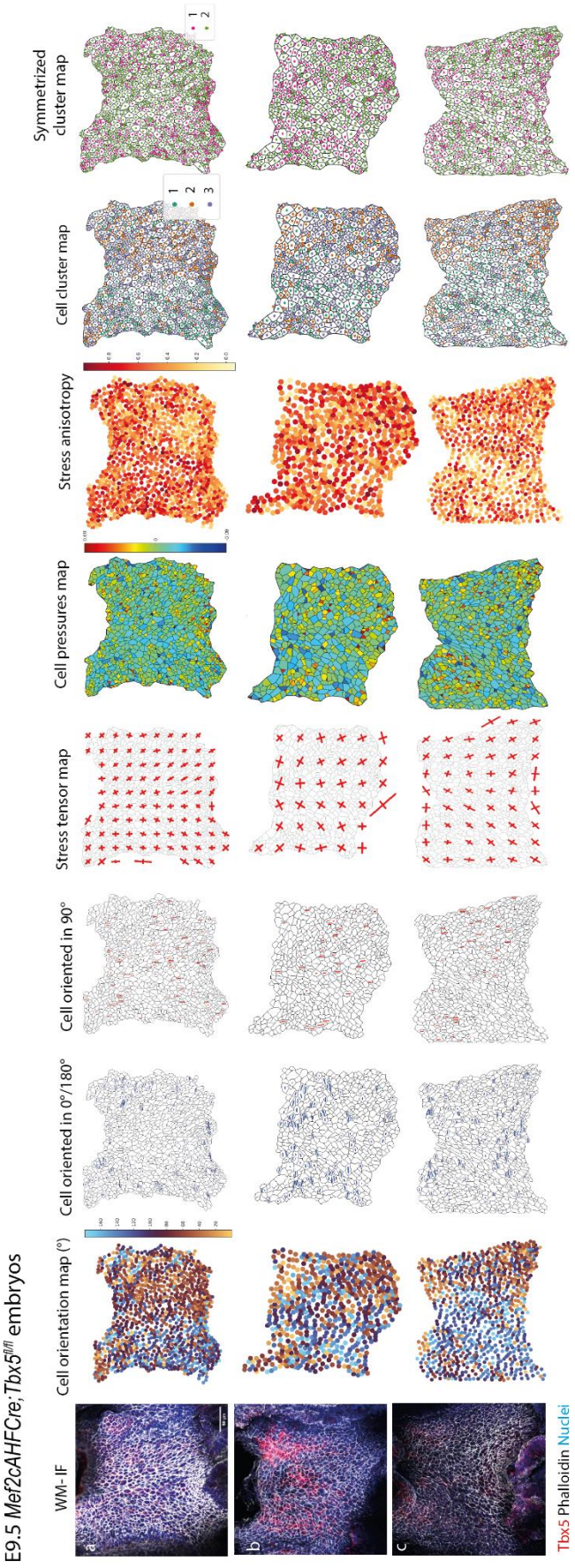

E9.5 Mef2cAHFCre;Tbx5fl/fl embryos
Symmetrized cluster map
Cell cluster map
Stress anisotropy
Cell pressures map
Stress tensor map
Cell oriented in 90°
Cell oriented in 0°/180°
Cell orientation map (°)
WM-IF
a
b
c
Tbx5 Phalloidin Nuclei

Figure S11

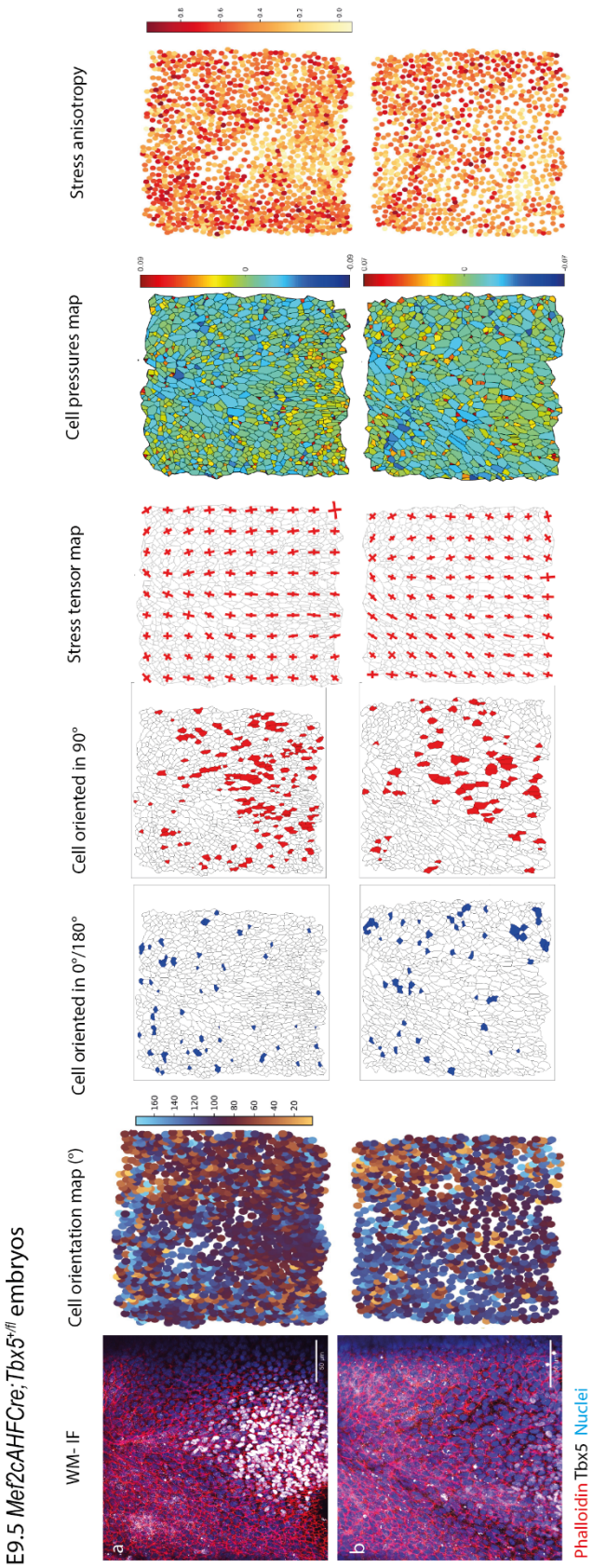
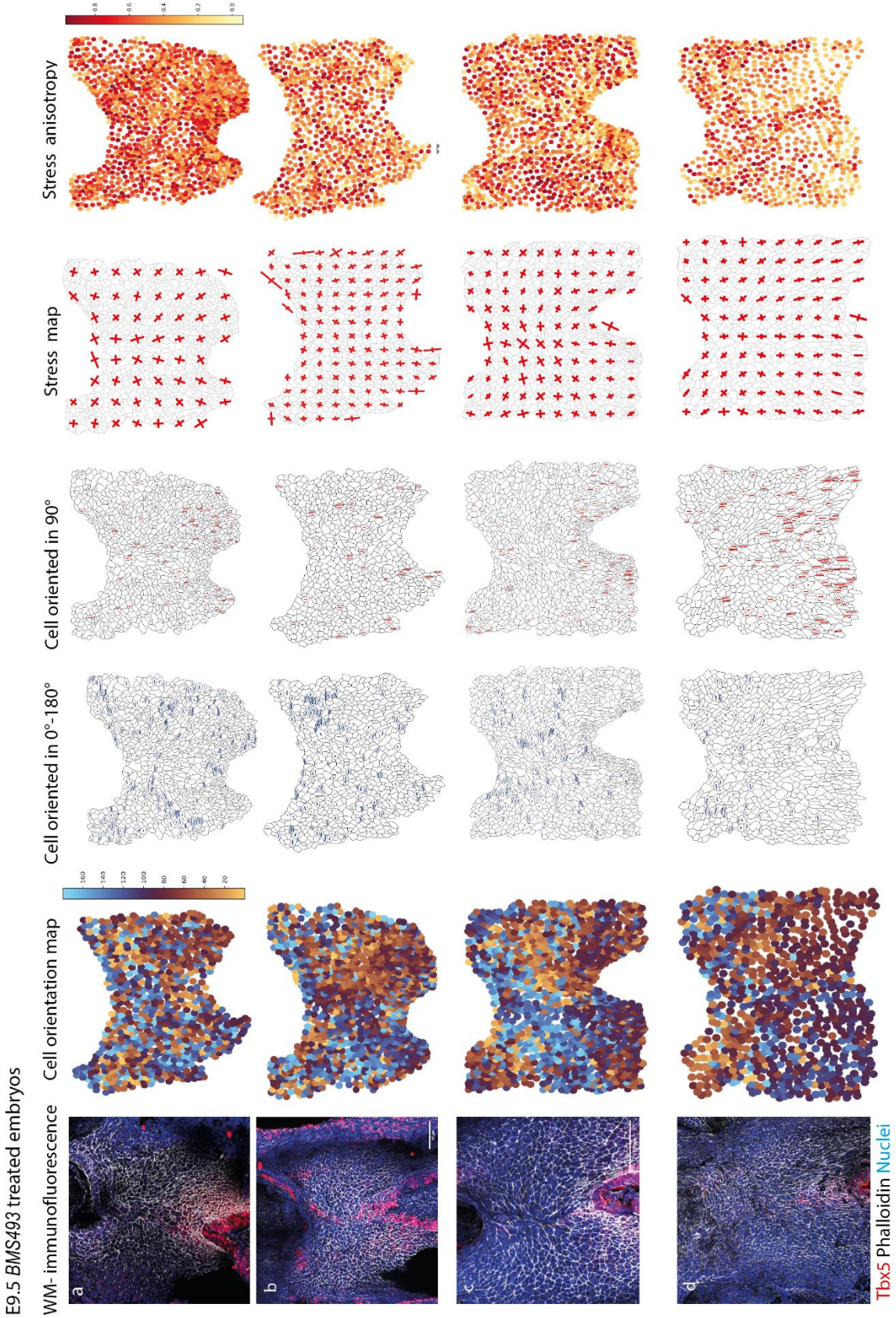

E9.5 Mef2cAHFCre;Tbx5+/fl embryos
WM- IF
Cell orientation map (°)
Cell oriented in 0°/180°
Cell oriented in 90°
Stress tensor map
Cell pressures map
Stress anisotropy
Phalloidin Tbx5 Nuclei
a
b

Figure S12



Guijarro et al.

Figure S13

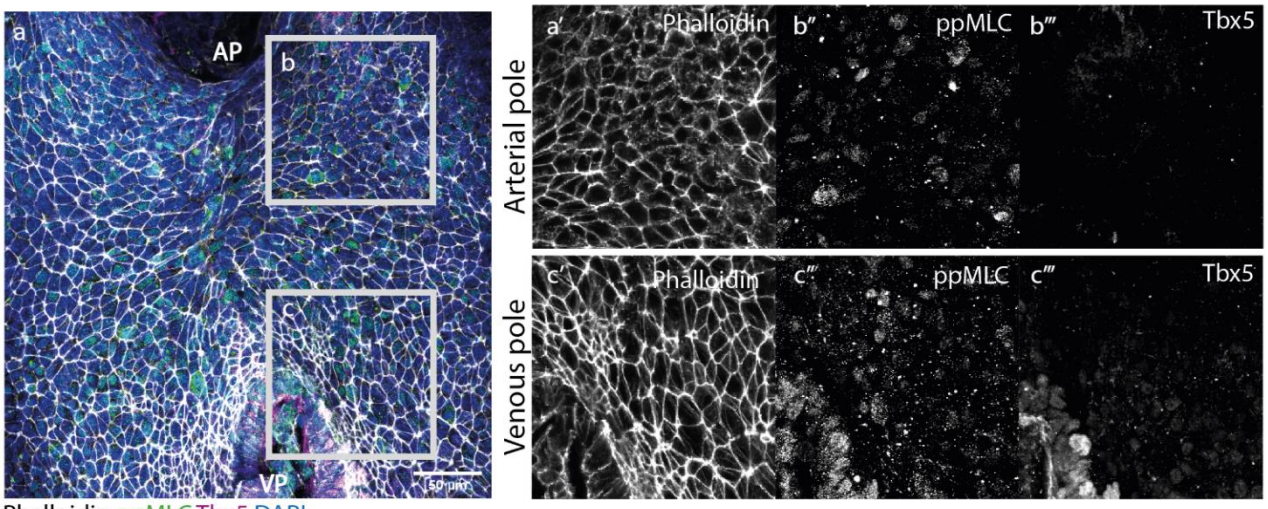
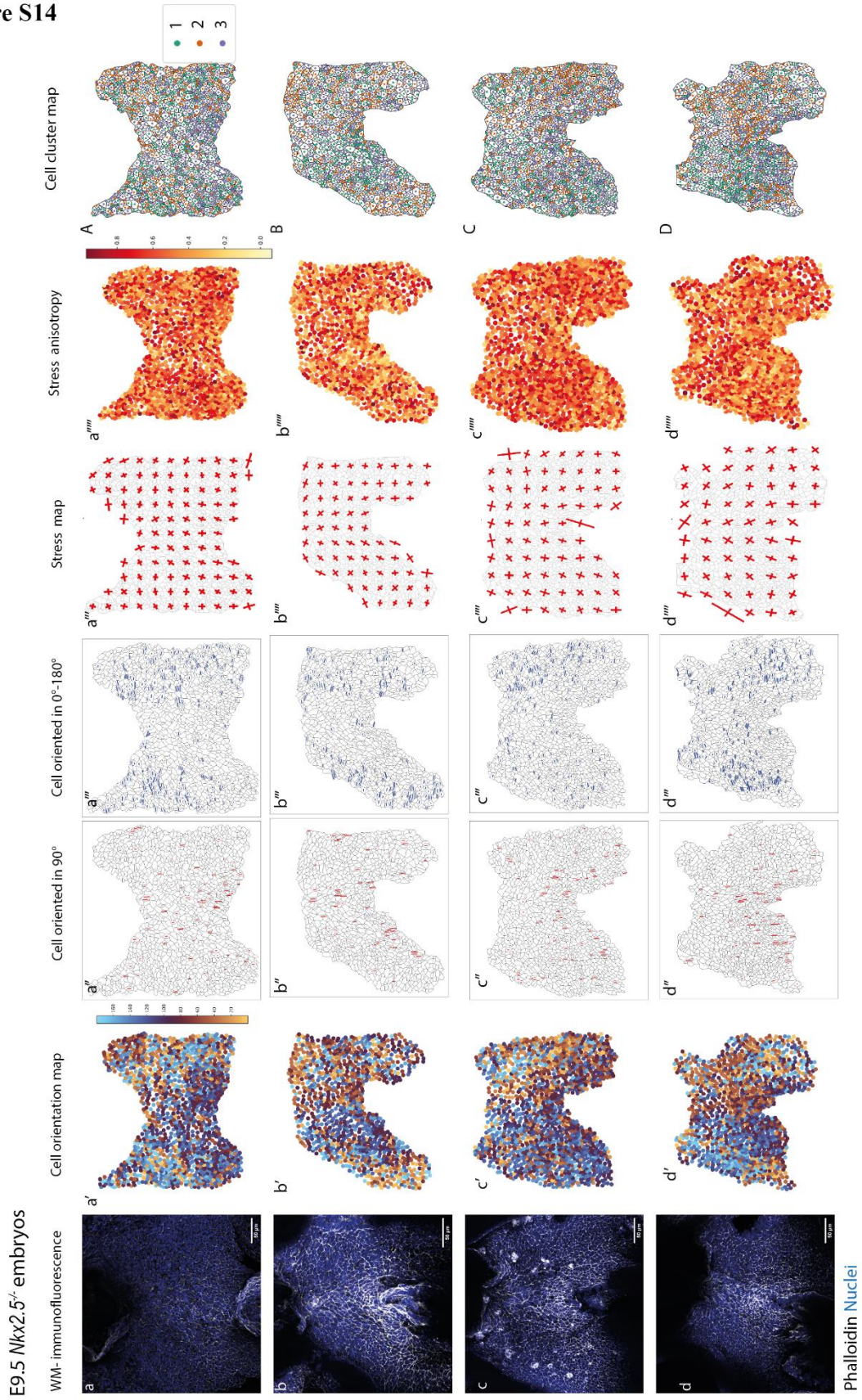

a
AP
b
c
VP
50 μm
Phalloidin ppMLC Tbx5 DAPI
Arterial pole
Venous pole
a′
Phalloidin
b″
ppMLC
b‴
Tbx5
c′
Phalloidin
c″
ppMLC
c‴
Tbx5

Figure S14



Guijarro et al.

Figure S15

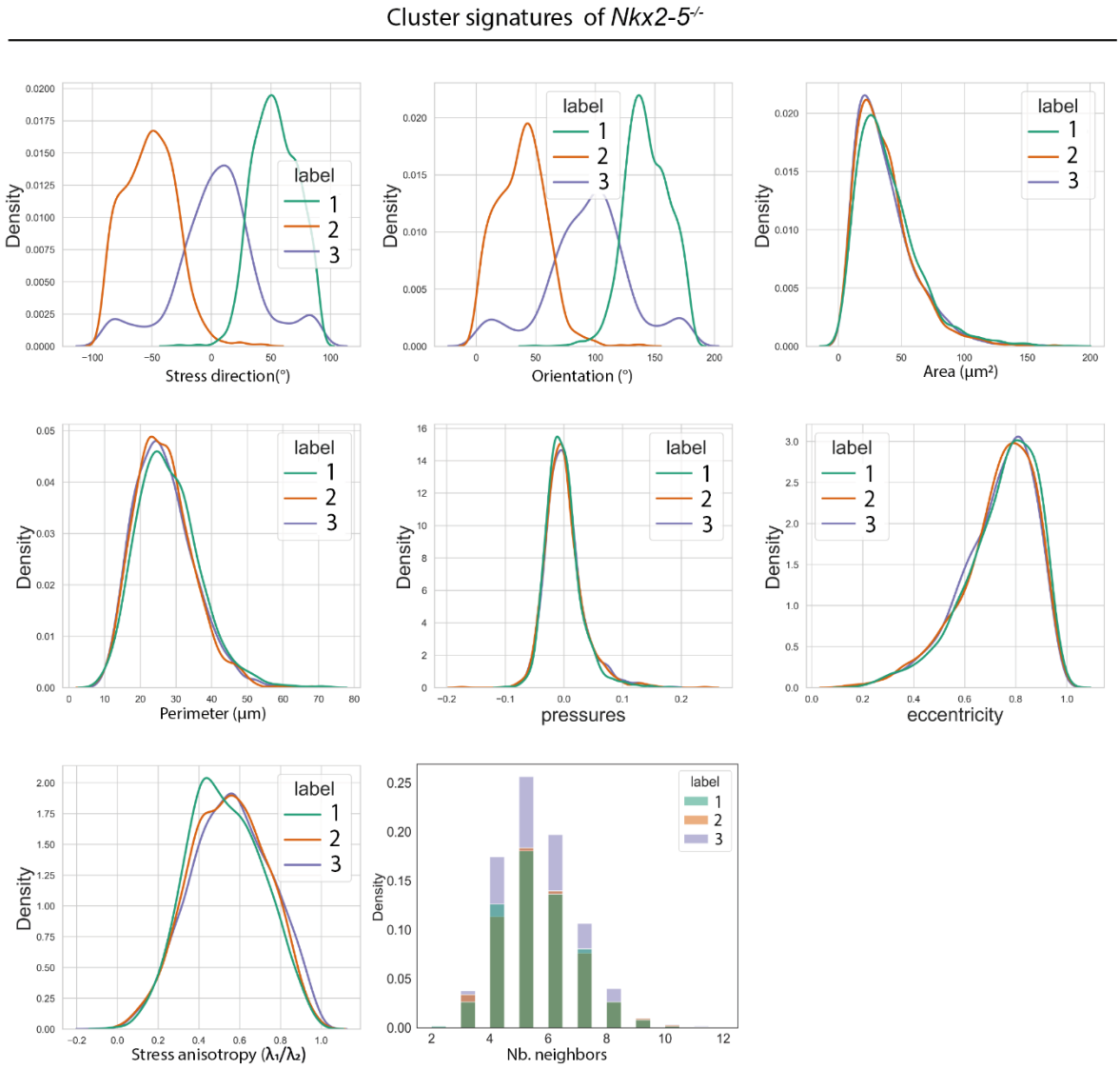

Cluster signatures of Nkx2-5-/-
Density
label
1
2
3
Stress direction(°)
Orientation (°)
Area (µm²)
Perimeter (µm)
pressures
eccentricity
Stress anisotropy (λ1/λ2)
Nb. neighbors

Figure S16

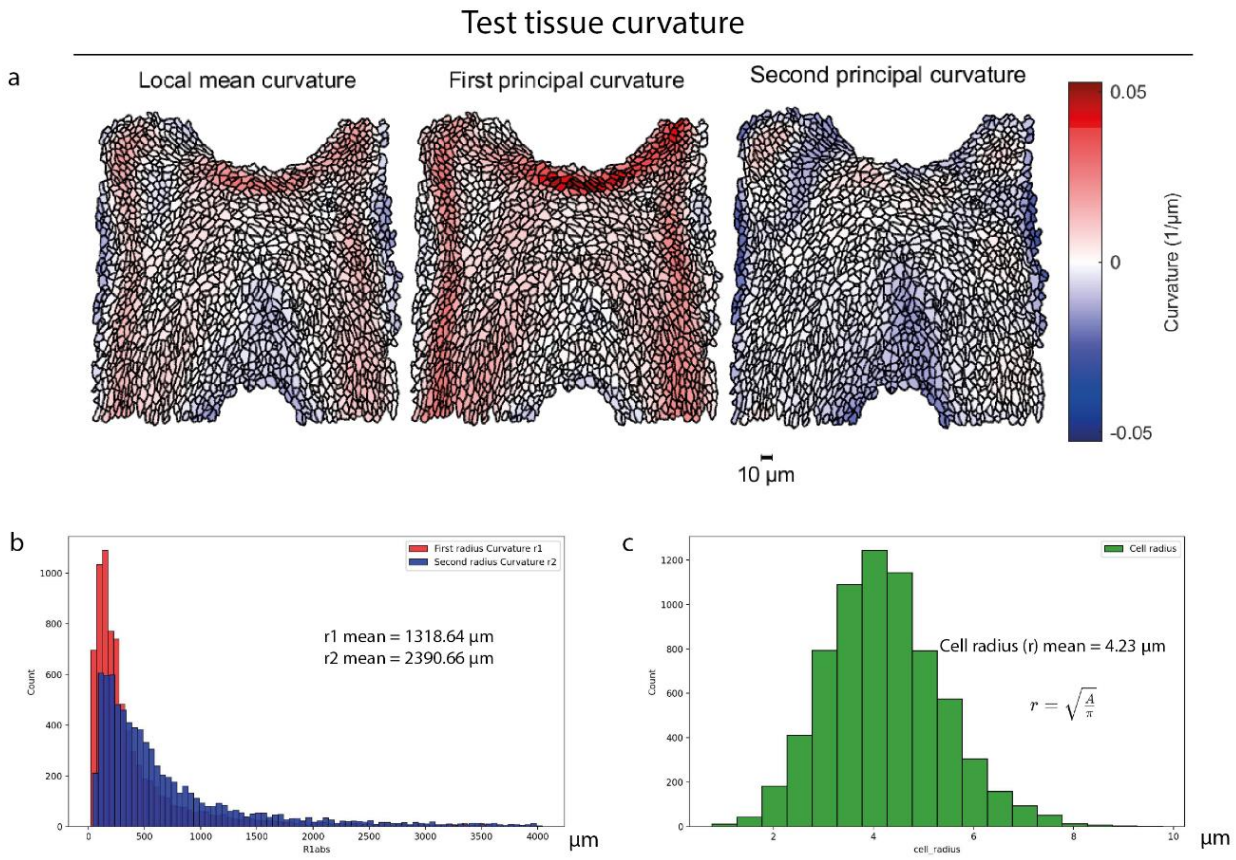

Test tissue curvature
a
Local mean curvature
First principal curvature
Second principal curvature
0.05
0
-0.05
Curvature (1/µm)
10 µm
b
First radius Curvature r1
Second radius Curvature r2
r1 mean = 1318.64 µm
r2 mean = 2390.66 µm
Count
R1abs
µm
c
Cell radius
Cell radius (r) mean = 4.23 µm
$r = \sqrt{\frac{A}{\pi}}$
Count
cell_radius
µm

Figure S17

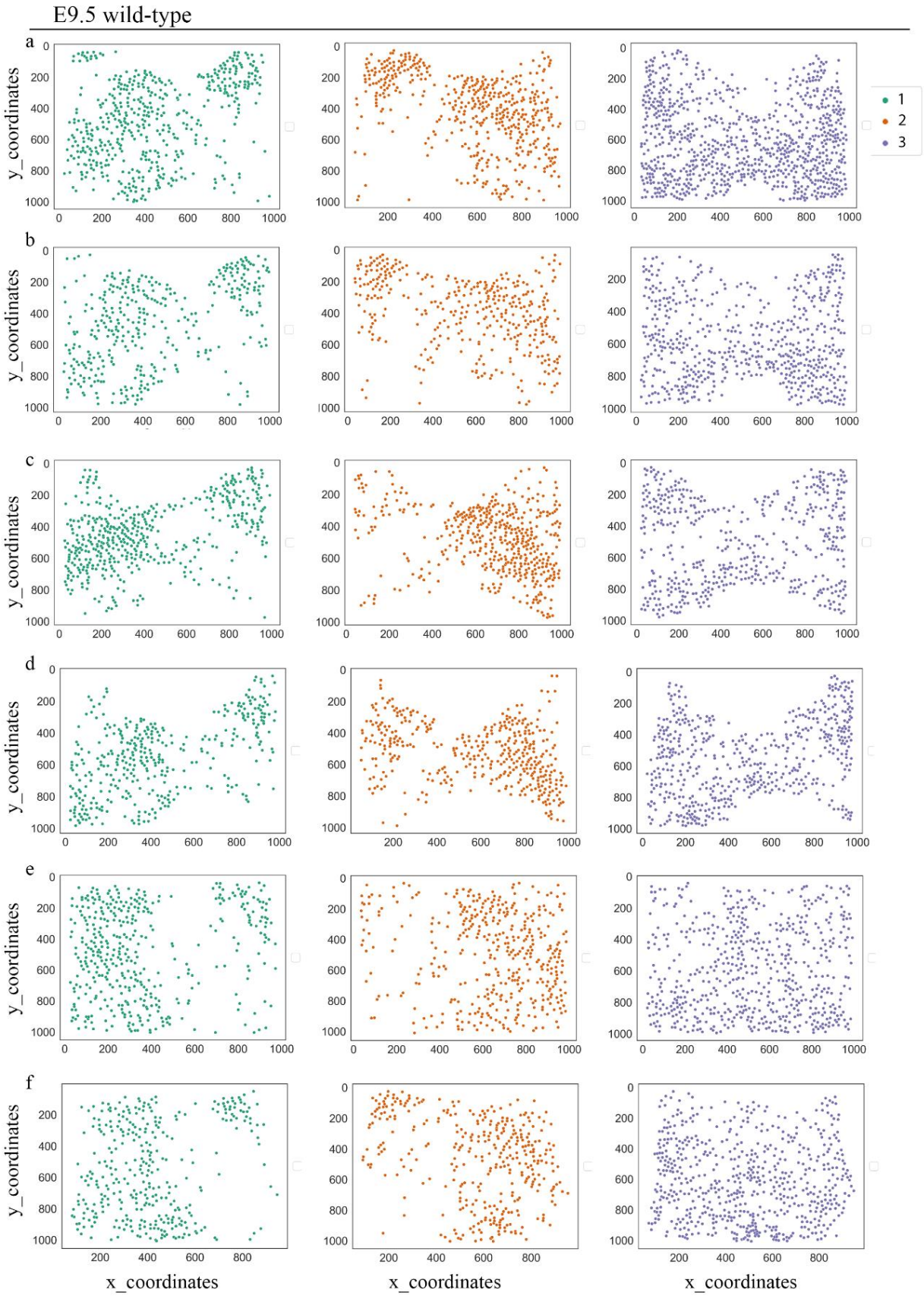


Figure S18

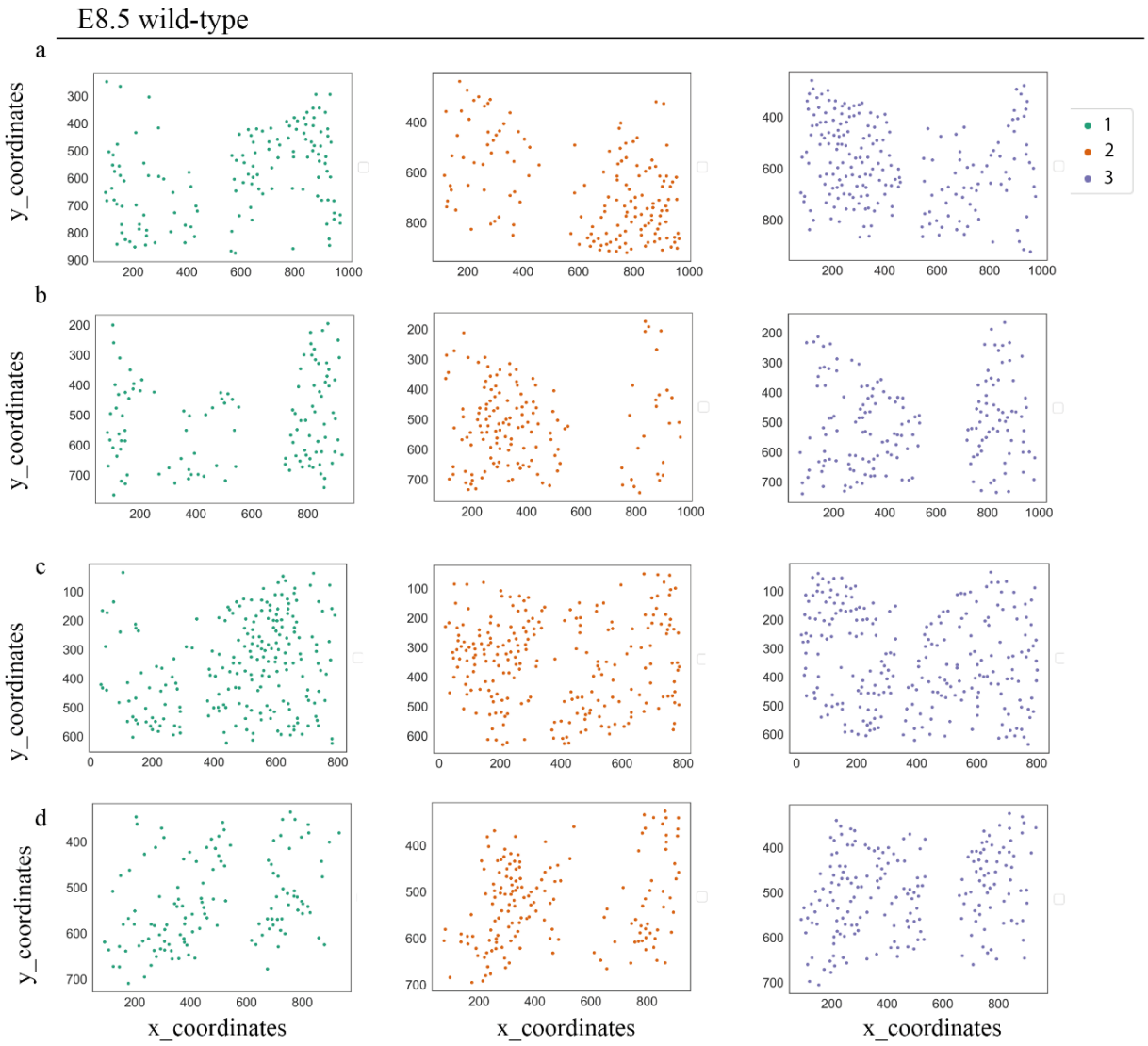
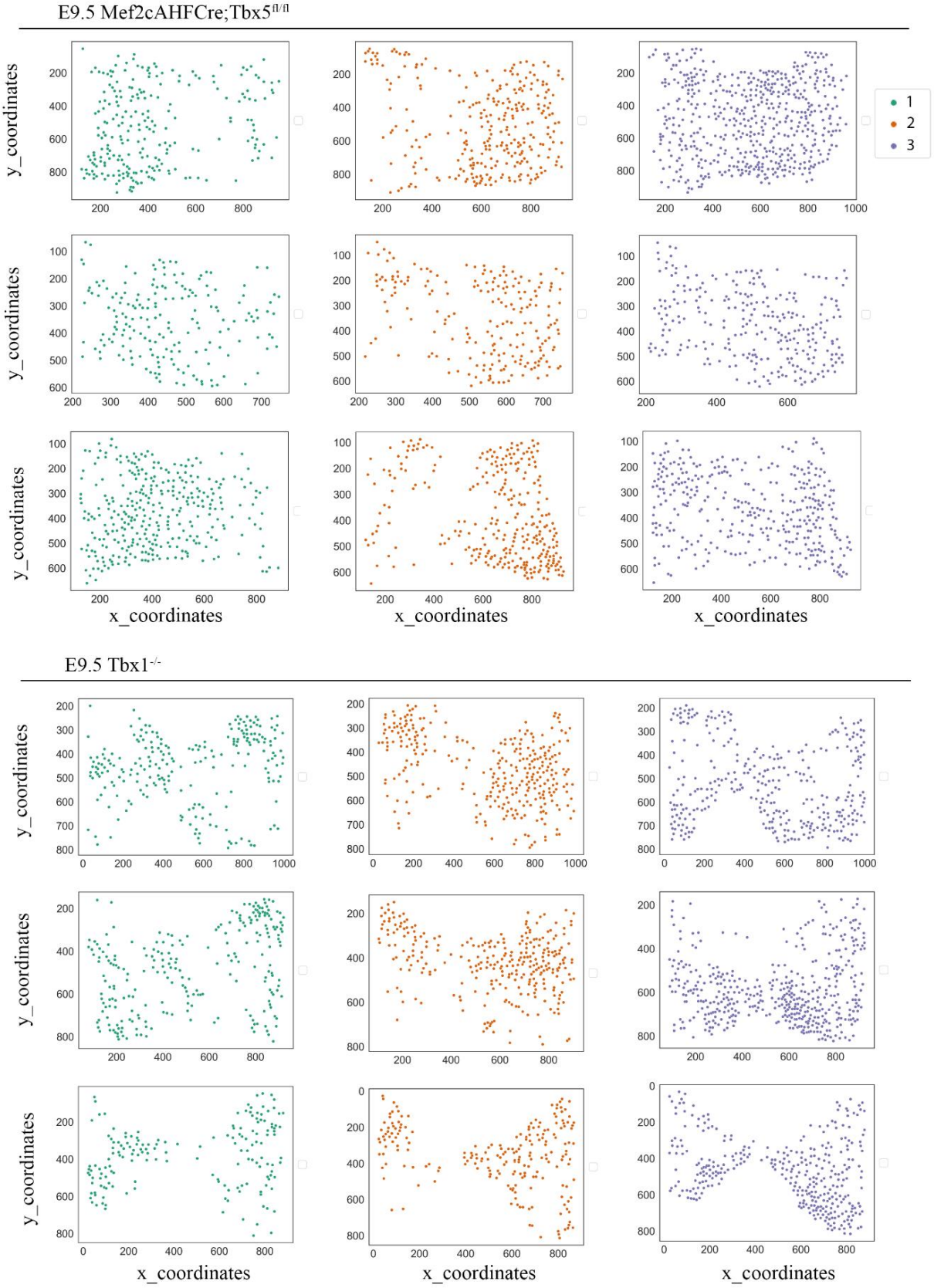
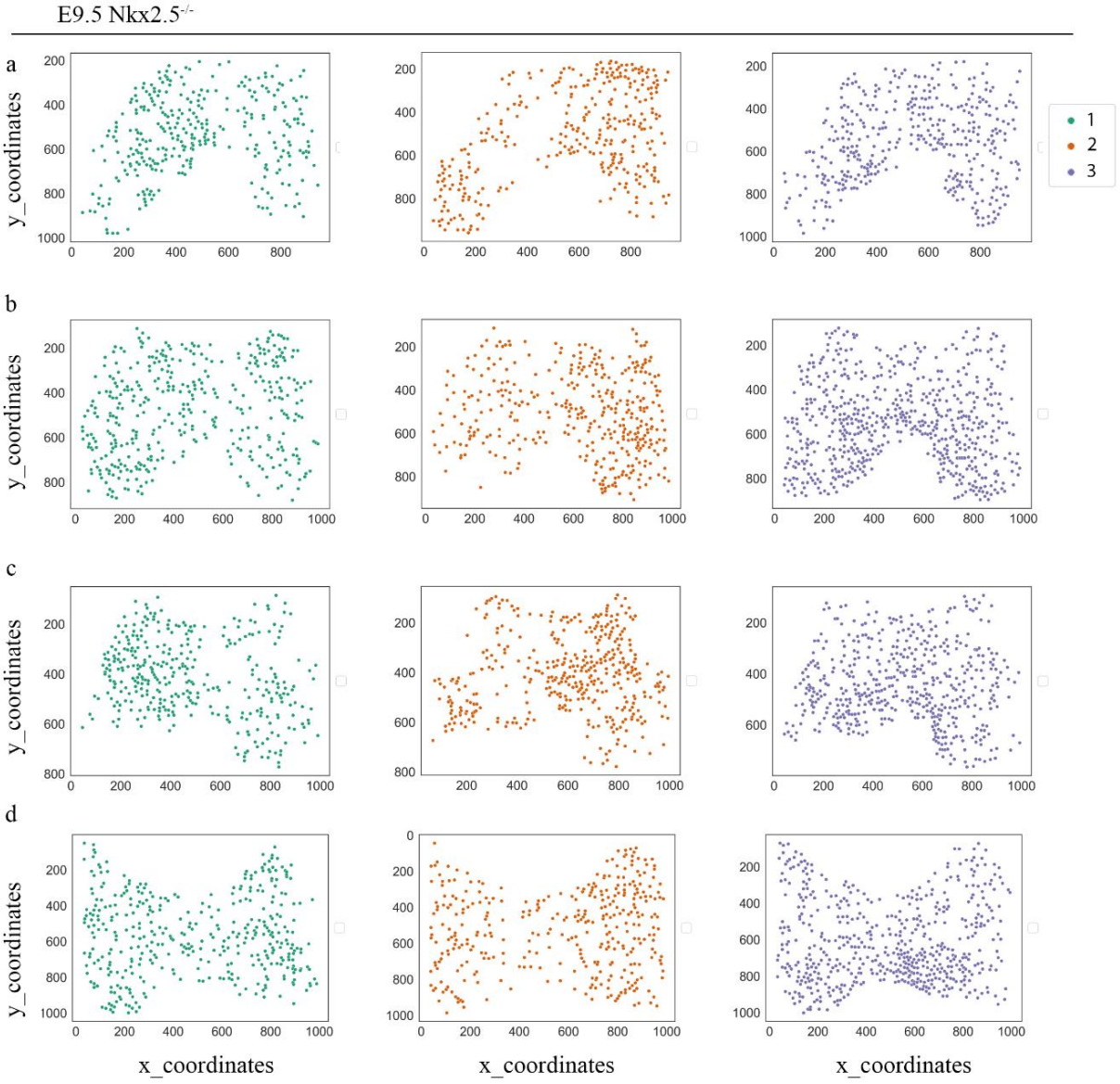


Figure S19



Guijarro et al.

Figure S20

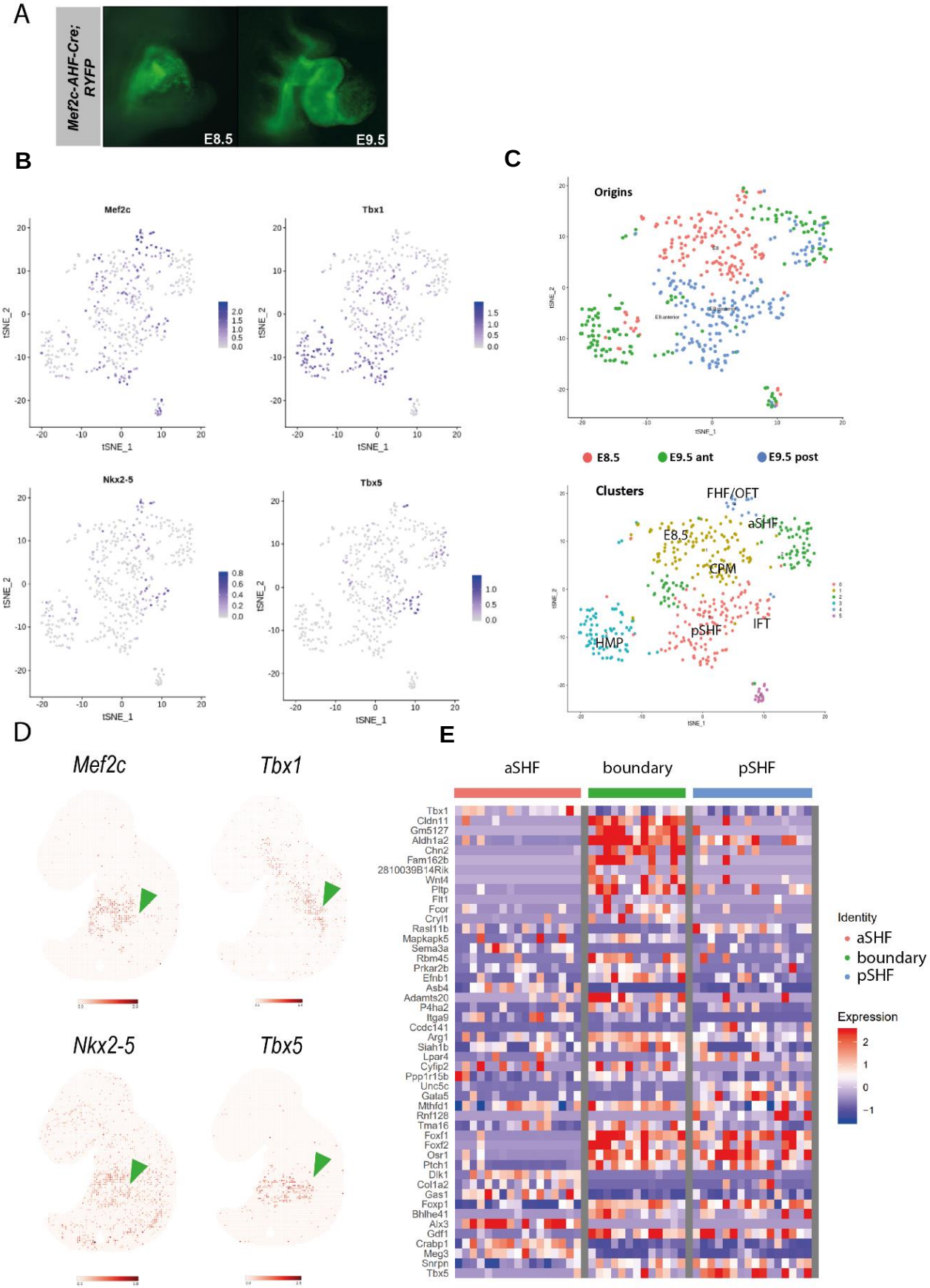


3.2 Boundary formation in the Second Heart Field

According to previous studies (De Bono et al., 2018) and my RNAscope results, there is a progressive emergence of the transcriptional boundary between *Tbx1* and *Tbx5* in the SHF between E8.5 and E9.5. During heart development *Tbx1* and *Tbx5* expression demarcate exclusive regions except for a short time window at E9 where both genes are expressed simultaneously in cells at the interface between *Tbx1* and *Tbx5* expression domains. Tbx1 expression is then downregulated in co-expressing cells and thus the anterior limit of this region of co-expression will become, at E9.5, a sharp boundary between *Tbx1* positive cells anteriorly and *Tbx5* positive cells posteriorly and medially.

To investigate the mechanisms involved in the formation of both distinct *Tbx1* and *Tbx5* populations and to test the hypothesis that there might be genes transiently expressed in the *Tbx1-Tbx5* co-expression domain that may be involved in the establishment of the transcriptional boundary, single-cell RNA sequencing (scRNAseq) of SHF cells at E8.5 and E9.5 was performed in collaboration with the Sperling laboratory (Berlin).

The sequencing was performed on cells from the E8.5, E9.5 anterior and E9.5 posterior SHF of *Mef2cAHFCre; RYFP* embryos. After microdissection and removal of the heart tube and tissue dorsal to the foregut, Magali Théveniau-Ruissy in the lab performed FACS (Fluorescent-Activated Cell Sorting) was used to purify YFP positive cells, separating *Mef2c-AHF* lineage positive SHF cells. ScRNAseq was performed using Smart seq, allowing more depth of sequence and more reads, although for a smaller number of cells, compared to the 10X Genomics approaches. The dataframe created for analysis has a shape of 21,769 features (genes) across 572 samples (cells). 3 dimensionality reduction algorithms were used to plot the cells PCA, UMAP, t-SNE. t-SNE was chosen to visualize the data, as it is the dimensionality reduction algorithm better recapitulating the morphology of the data and generating more distinct clusters. The data was then clustered according to their genetic profiles using K-nearest neighbor (KNN) graph. 6 clusters were obtained and they were labelled according to the expression of important genes involved in heart development. Cluster0: posterior Second Heart Field or Inflow Tract (pSHF/IFT), cluster1: Cardiopharyngeal mesoderm (CPM) the progenitor pool that will give rise to both, heart and head muscle cells, cluster2: anterior Second heart field (aSHF), cluster3: head muscle progenitor cells, cluster4: First Heart Field or Outflow tract (FHF/OFT), cluster5: Endothelial cells (Fig. 3.2.1).


A
Mef2c-AHF-Cre; RYFP
E8.5
E9.5
B
Mef2c
Tbx1
Nkx2-5
Tbx5
tSNE_1
tSNE_2
C
Origins
E8.5
E9.5 ant
E9.5 post
Clusters
FHF/OFT
aSHF
E8.5
CPM
pSHF
IFT
HMP
D
Mef2c
Tbx1
Nkx2-5
Tbx5
E
aSHF
boundary
pSHF
Tbx1
Cldn11
Gm5127
Aldh1a2
Chn2
Fam162b
2810039B14Rik
Wnt4
Pltp
Flt1
Fcor
Cryl1
Rasl11b
Mapkapk5
Sema3a
Rbm45
Prkar2b
Efnb1
Asb4
Adamts20
P4ha2
Itga9
Ccdc141
Arg1
Siah1b
Lpar4
Cyfip2
Ppp1r15b
Unc5c
Gata5
Mthfd1
Rnf128
Tma16
Foxf1
Foxf2
Osr1
Ptch1
Dlk1
Col1a2
Gas1
Foxp1
Bhlhe41
Alx3
Gdf1
Crabp1
Meg3
Snrpn
Tbx5
Identity
aSHF
boundary
pSHF
Expression
2
1
0
−1

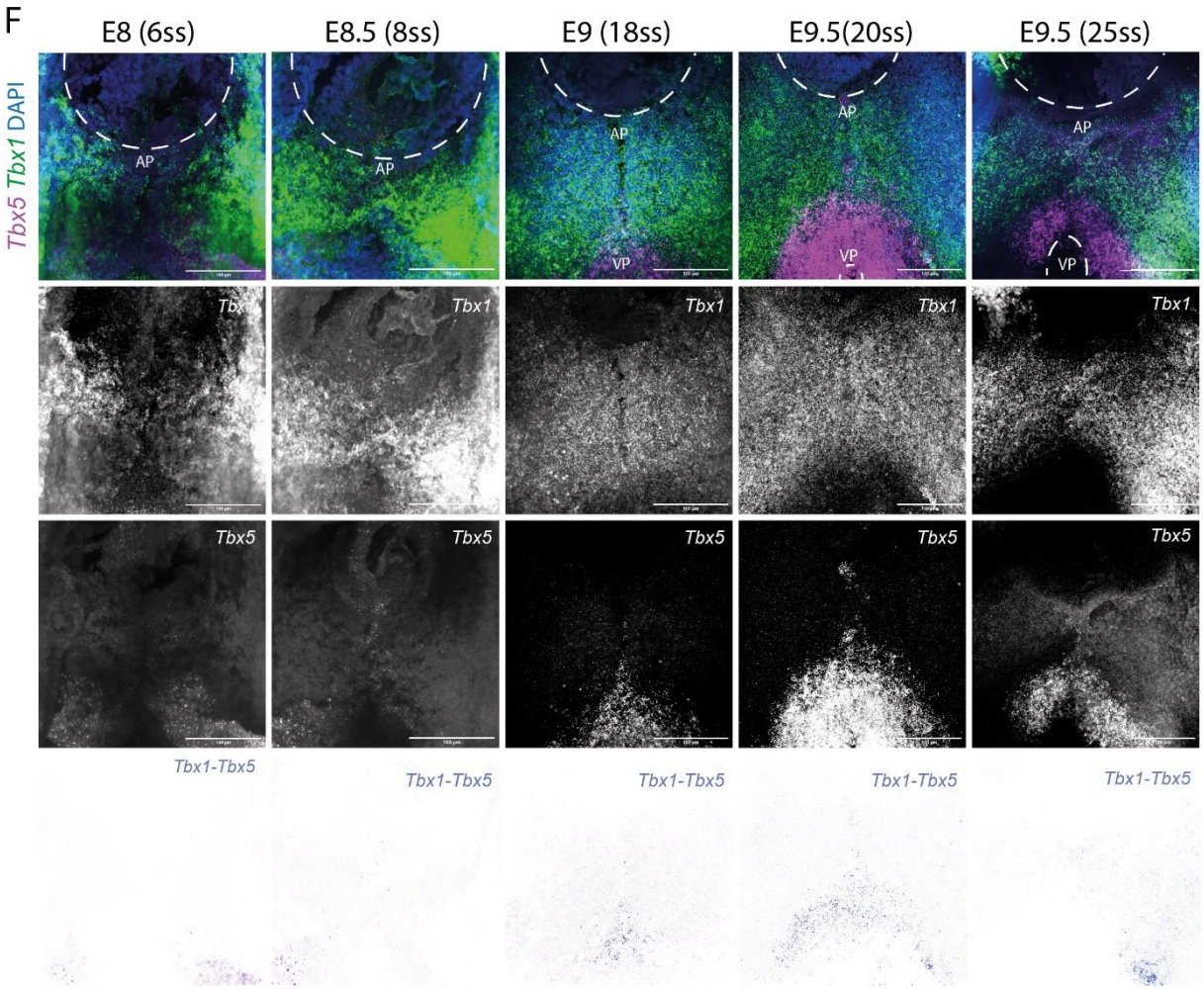


Figure 3.2.1: Single-cell RNA-seq of the second heart field cells. A. Immunofluorescence of E8.5 and E9.5 *Mef2cAHFCre; RYFP* embryos **B.** t-SNE scatterplots of the SHF cells color-coded according to the level of expression of different genes *Mef2c, Tbx1, Tbx5* and *Nkx2.5* **C**. t-SNE projection of SHF cells, with cells color-coded according to their origin or according to the cluster to which they belong **D.** Scatterplot of the cells in the real embryo localizations in a sagittal section at E9.5 using the MOSTA spatial transcriptomics browser, Mouse Organogenesis Spatiotemporal Transcriptomic Atlas-MOSTA (A. Chen et al., 2022), color coded according to the level of expression of *Mef2c, Tbx1, Tbx5* and *Nkx2.5*. Green arrows indicate the overlapping region between Tbx1 and Tbx5 **E.** Heatmap showing the expression levels of different genes, classified based on their localization in the aSHF, pSHF or boundary zone. **F.** Tbx1 and Tbx5 mRNA distribution and quantification from E8 to E9.5 in the Dorsal pericardial wall, showing frontal whole-mount views of the DPW of the wild-type embryos after RNAscope fluorescent in-situ hybridization showing nuclei (blue), *Tbx1* (green) and *Tbx5* (magenta). The panels bellow show the separate channels of *Tbx1* or *Tbx5* and lastly the spatial superposition of the double *Tbx1-Tbx5* mRNA (linked with the quantifications in FigureS7).

Consistent with the expression data obtained by RNAscope analysis *Tbx5* and *Tbx1* expressing cells were identified in proximity in the t-SNE plot. Moreover, the *Tbx5* positive cells were located in the posterior DPW cluster and had downregulated expression of *Mef2c-AHF-Cre* at E9.5, similarly to the situation in the embryo (De Bono et al, 2018). Focusing on the t-SNE plot and the expression profiles the data was mined to check for cells co-expressing both *Tbx1* and *Tbx5* at the interface between these expression domains. According to the results there are 78 cells that co-express *Tbx1* and *Tbx5*. 307 that only express *Tbx1*, 29 that only express *Tbx5*, and 48 that do not express neither *Tbx1* nor *Tbx5*.

The 78 cells were further analyzed. Of these, cells in the boundary region between *Tbx1* and *Tbx5* positive cells were selected and the most highly expressed genes in this population of cells relative to adjacent *Tbx1*+*Tbx5*- and *Tbx1*-*Tbx5*+ cells were identified. The expression profile was compared with the rest of the dataset and those genes which are highly expressed in comparison to the rest of the dataset were chosen and plotted in the t-SNE. Among the candidates obtained we find *Cldn11, Chn2, Gm5127, Aldh1a2, Gdnf, or Pltp*. Interestingly, these genes are membrane proteins or components of tight junctions working as physical barriers regulating proliferation, polarity and signal transduction.

In preliminary experiments I performed a first validation of the results using diogoxygenin labelled probes in wholemount in situ hybridization. Only 2 genes were investigated: *Cldn11* and *Chn2*. Preliminary validation using classic ISH results did not detect expression of either Cldn11 and Chn2 in the SHF, although a positive control for these two probes was missing. Further validation will be performed using other in situ hybridisation techniques including RNAscope and MERFISH that allow validation of 3 or more probes at once, allowing direct comparison of the expression domains of these genes with those of *Tbx1* and *Tbx5* (Fig. 3.2.2).

CHN2: This gene encodes a guanosine triphosphate (GTP)-metabolizing protein. This protein plays role in proliferation and migration of smooth muscle cells. Mutations in these gene are associated with cancer and schizophrenia (Leung et al., 1994; Caloca et al., 1997).

CLDN11: Claudins are integral membrane proteins and components of tight junction strands. Tight junction strands serve as a physical barrier to prevent solutes and water from passing freely through the paracellular space between epithelial or endothelial cell sheets. It also play critical roles in maintaining cell polarity and signal transductions. In the central neural system, this protein plays an important role in regulating proliferation and migration of oligodendrocytes (Denninger et al., 2015; Gow et al., 2004; Devaux & Gow, 2008).

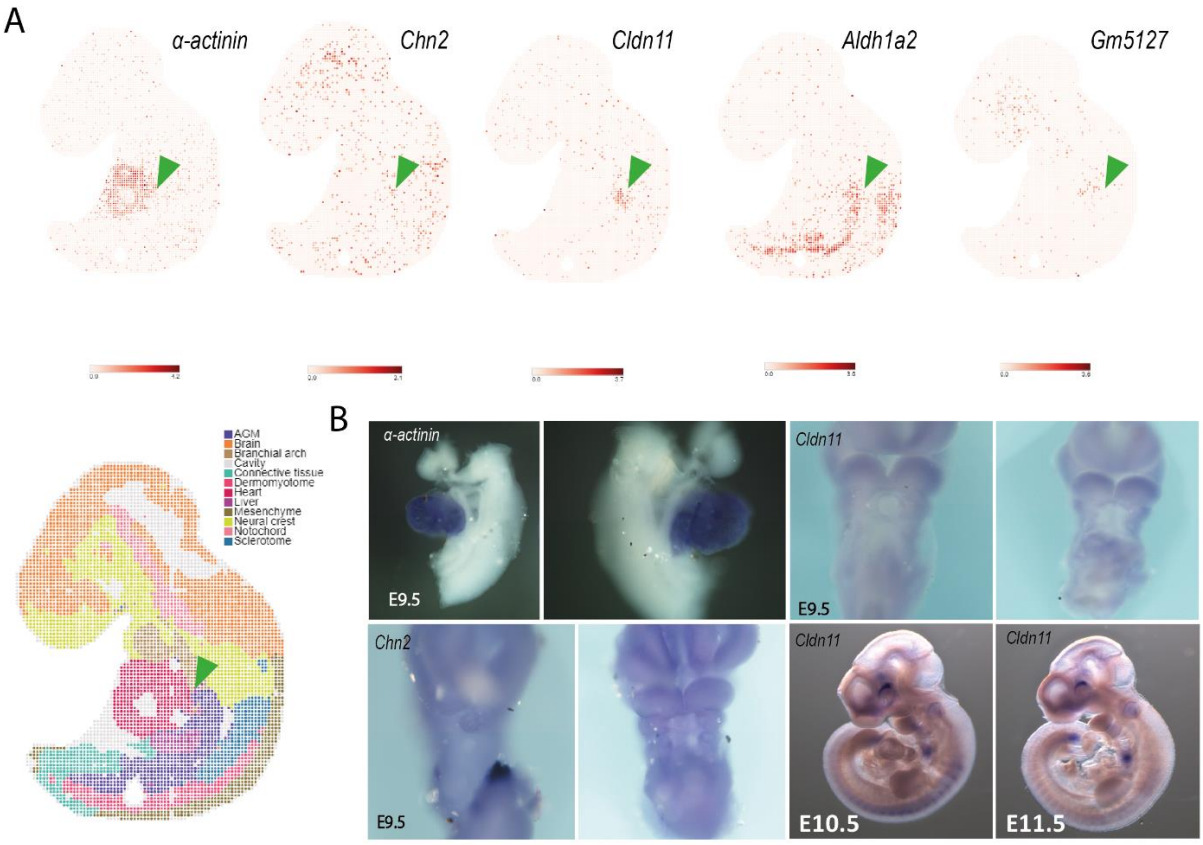


Figure 3.2.2: Validation of candidate boundary genes in E9.5 embryos A. Scatterplot of the cells in the real embryo localizations and color coded according the level of expression of *Actn2, Chn2, Cldn11, Aldhn1a2* and Gm5127. The left bottom panel shows the embryo at E9.5 with the regions color-coded according to cell MOSTA. Green arrows indicated the overlapping region between *Tbx1* and *Tbx5.* **B.** In situ hybridization of E9.5 embryos with α-actinin, *Chn2, Cldn11* probes as well a published in situ hybridization of *Cldn11* in E10.5 and E11.5 embryos (Lewandowski et al., 2015).

Another interesting candidate gene to be addressed is *Aldh1a*, coding for the aldehyde dehydrogenase protein, an enzyme that catalyzes the synthesis of retinoic acid (RA) from retinaldehyde. The expression of this protein in the region between the *Tbx1*+ and *Tbx5*+ region, is consistent with the activation of retinoic acid signaling in the venous pole before the upregulation of *Tbx5*. Interestingly, diseases associated with *ALDH1A2* are diaphragmatic hernia with cardiovascular defects, a congenital defect leading to pulmonary hypoplasia and cardiovascular defects such as conotruncal anomalies or the formation of aortopulmonary window and spina bifida, a birth defect not directly related with a heart defect but involving the incomplete closing of the spine and spinal cord membranes (Schreiner et al., 2021; Pavan et al., 2009).

In line with genetic boundary formation, we assessed using my single-cell morphometric dataset, whether a boundary of tension could be present between the Tbx1+ and Tbx5+

compartments. It has been suggested that boundaries between tissue compartments are characterized by both specific gene expression patterns and elevated interfacial tension (Hallou et al., 2023). Following the reasoning of this and other studies (Fagotto, 2020), cells are able to discriminate their contacts: contacts between same cell types are called homotypic contacts, and contacts between different cell types are called heterotypic contacts. To determine if a genetic boundary forms in the SHF, we conducted preliminary experiments. We found that cells with homotypic contacts, whether in the *Tbx1* or *Tbx5* region (Tbx5 region assessed by its protein expression and Tbx1 assessed by the absence of *Tbx5*), exhibit no significant difference in tension. However, there is a significant increase in tension at the boundary, or heterotypic contacts, which connect Tbx5 and Tbx1 expressing cells (Fig. 3.2.3). These results were observed in 2 out of 4 embryos; further embryos need to be processed for confirmation. It would be interesting to pursue this analysis and investigate if the boundary compartments identified after sc-morphometrics display higher tension.

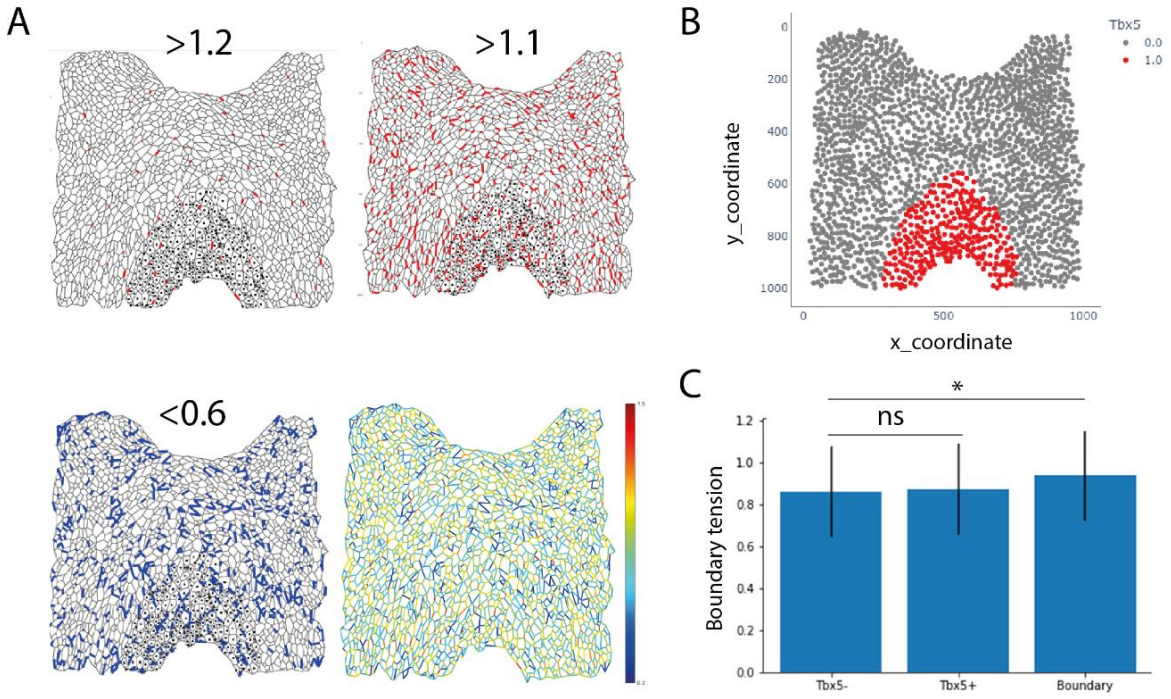


Figure 3.2.3: Quantification of homotypic and heterotypic tension in SHF cells at the *Tbx1 Tbx5* boundary region A. Spatial visualization of the tension map. For visual purposes the tension contacts have been separated: in red the "high" tension contacts, with tensions higher than 1.2,1.2 and in blue the contacts with "low" tensions, lower than 0.6 **B.** The SHF cells color-coded according to if they express *Tbx5* (red) or not (grey) **C.** Boxplot showing the average tensions of the cell homotypic contacts in Tbx1 region, Tbx5 region and the boundary or heterotypic contacts MOSTA (A. Chen et al., 2022) .

3.3 Cell intercalation in the Second Heart Field

In addition to biomechanical forces, we hypothesized that other events might be controlling the addition of SHF cells to the cardiac poles. Convergent extension and cell intercalation occur during the processes of cell movement and axis elongation (Fig. 3.3.1) (Wallingford et al., 2002; Keller et al., 2000 ; Rauzi et al., 2010). For example, throughout gastrulation or neurulation. Convergent extension is a process in which a tissue narrows along one axis and lengthens in the perpendicular axis. This process has been reported to be important as cells converge during early heart tube formation in avians (Kidokoro et al., 2018; Blankenship et al., 2006).

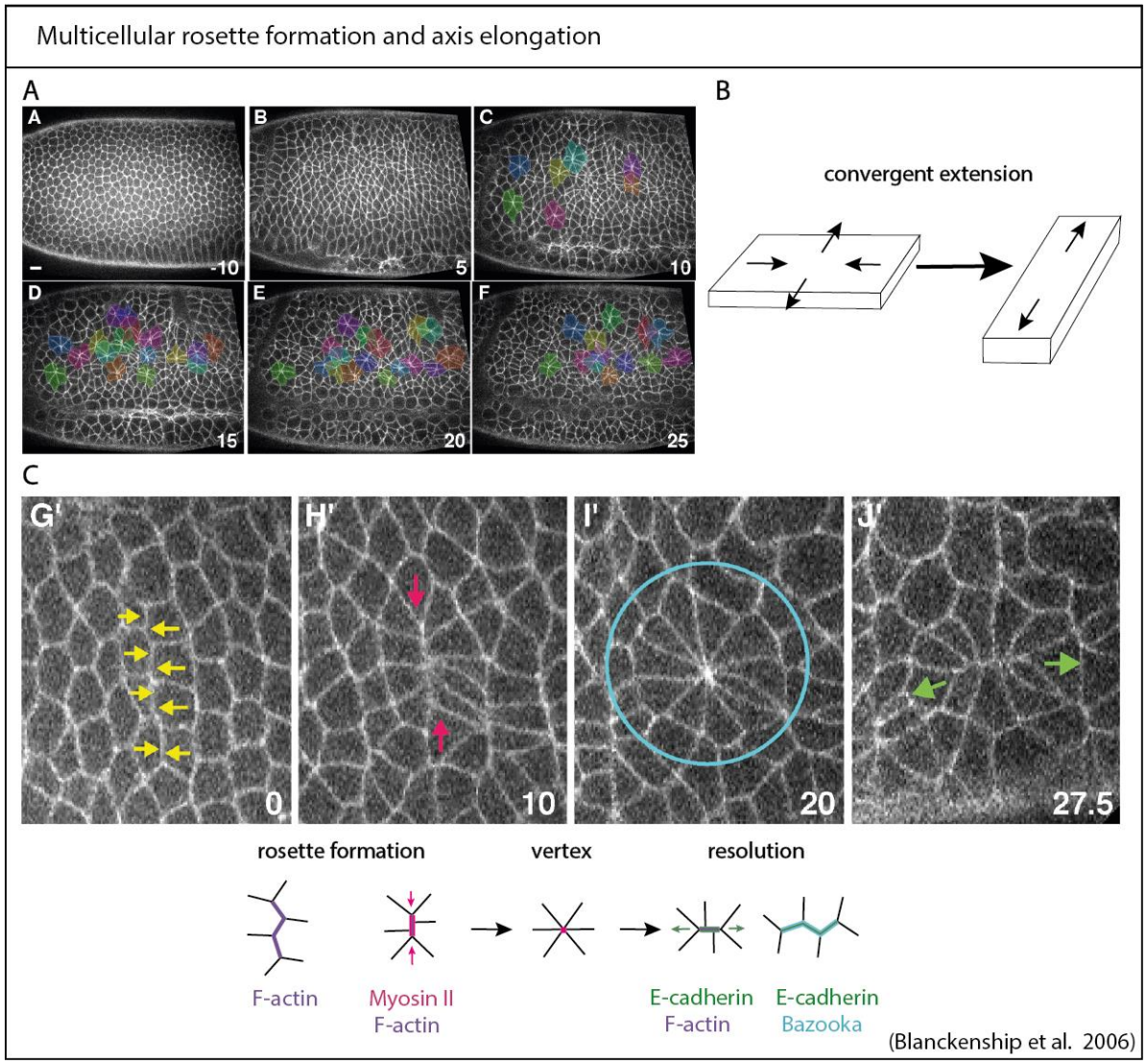


Figure 3.3.1 Multicellular rosette formation in intercalating cells and axis elongation. A. (A–F) Time-lapse images of the germband. Rosettes of six or more cells are highlighted. Colors do not indicate the same cells in sequential images **B.** Scheme recapitulating convergent extension. Arrows show the direction of the forces **C.** Time lapse of images of a single rosette and bellow and schematic sequence of events taking

place during rosette formation**. G'.** During rosette formation, two columns of cells initially align (edges that will contract are indicated by yellow arrows, **H'.** Adjacent pairs of cells constrict their shared interfaces in concert during rosette formation, the direction of convergence is indicated by red arrows, **I'.** Multiple cells come into simultaneous contact at a high-order vertex, the 11-cell rosette is indicated by a blue circle, **J'.** This apparently symmetric structure resolves in a strictly directional fashion during rosette resolution, the direction of extension is indicated by green arrows. Figure adapted from (Blankenship et al., 2006).

To verify the occurrence of convergent extension, we quantified the percentage of multicellular rosettes, structures closely associated with the essential cell behaviors that drive axis elongation (Blankenship et al., 2006) (Fig. 3.3.1). The quantifications show that rosettes are present in the DPW, but that there is no significant difference in the number of rosettes in E9.5 wild-type, compared to E8.5 wild-type, *Tbx1*$^{-/-}$, *Mef2cAHFCre; Tbx5*$^{fl/fl}$ or *Nkx2.5*$^{-/-}$ mutant embryos. This last mutant is interesting as it has a perturbed SHF deployment that prevents heart tube extension (Francou et al., 2017). For quantifications, a threshold of 3 pixels was used for the bond length at cell vertices. A higher threshold means reduced stringency in rosette calling, resulting in more rosettes appearing in the tissue. Including high bond lengths will encompass vertices that may not be truly correspond to rosettes or it would correspond to rosettes resolving (Fig. 3.3.1J', Fig. 3.3.2A). However, rosette proportions between the different genotypes did not change if higher bond lengths were chosen (Fig. 3.3.2B).

To look for regional areas with differential rosette formation, we quantified the rosettes between the anterior and posterior SHF (dividing the image at the medium point) as well as between the left and right SHF (dividing the image in 2 halves). These quantifications show that there is no significant difference between the proportion of rosettes between the anterior and posterior domains nor the left and right domains in wild-type embryos. Only in *Tbx1*$^{-/-}$ mutants was a significant difference between the left-right domains, detected with a decrease in the number of rosettes in the right SHF that remains to be understood (Fig. 3.3.3A, D).

Cell intercalation thus is a process taking place dorsal closure of the dorsal pericardial wall and is maintained during the elongation of the DPW. However, my results suggest that this process is not dependent on *Tbx1*, *Tbx5* or *Nkx2-5*. Cell intercalation and establishment of tension in the pSHF thus appear to be independent processes. As a corollary, T-box genes and Nkx2.5 impact the morphology of the SHF cells by affecting tension patterns, but not rosette formation.

Lastly, the overall dimensions of the dorsal pericardial wall were evaluated to understand the changes in the tissue (Fig. 3.3.3E). From E8.5 to E9.5, there is an increase in both the arterial-to-venous pole distance and the left-right distance. However, in the mutants, while the left-right distance remains unchanged, the arterial-to-venous pole distance

decreases in all cases, being most significant in *Tbx1* mutant embryos and less significant in *Nkx2.5* and conditional *Tbx5* mutant embryos (p=0.08). The ratio between the arterial-to-venous pole distance and the orthogonal axis shows that from E8.5 to E9.5, the anterior-posterior axis increases compared to the left-right axis, indicating that the tissue is converging as the dorsal mesocardium closes and extending into the poles. In the three mutants, extension and growth of the anterior-posterior axis are reduced, the mechanisms by which this reduction happens has to be different from convergent extension mechanism. The reduction of the elongation may be by impairment of proliferation or the perturbation of the tension pattern in the SHF.

Rosette formation is a dynamic process, and a more detailed study of this phenomenon will require live imaging to accurately capture and analyze its progression. It would be interesting to study the direction of the resolution of the rosettes.

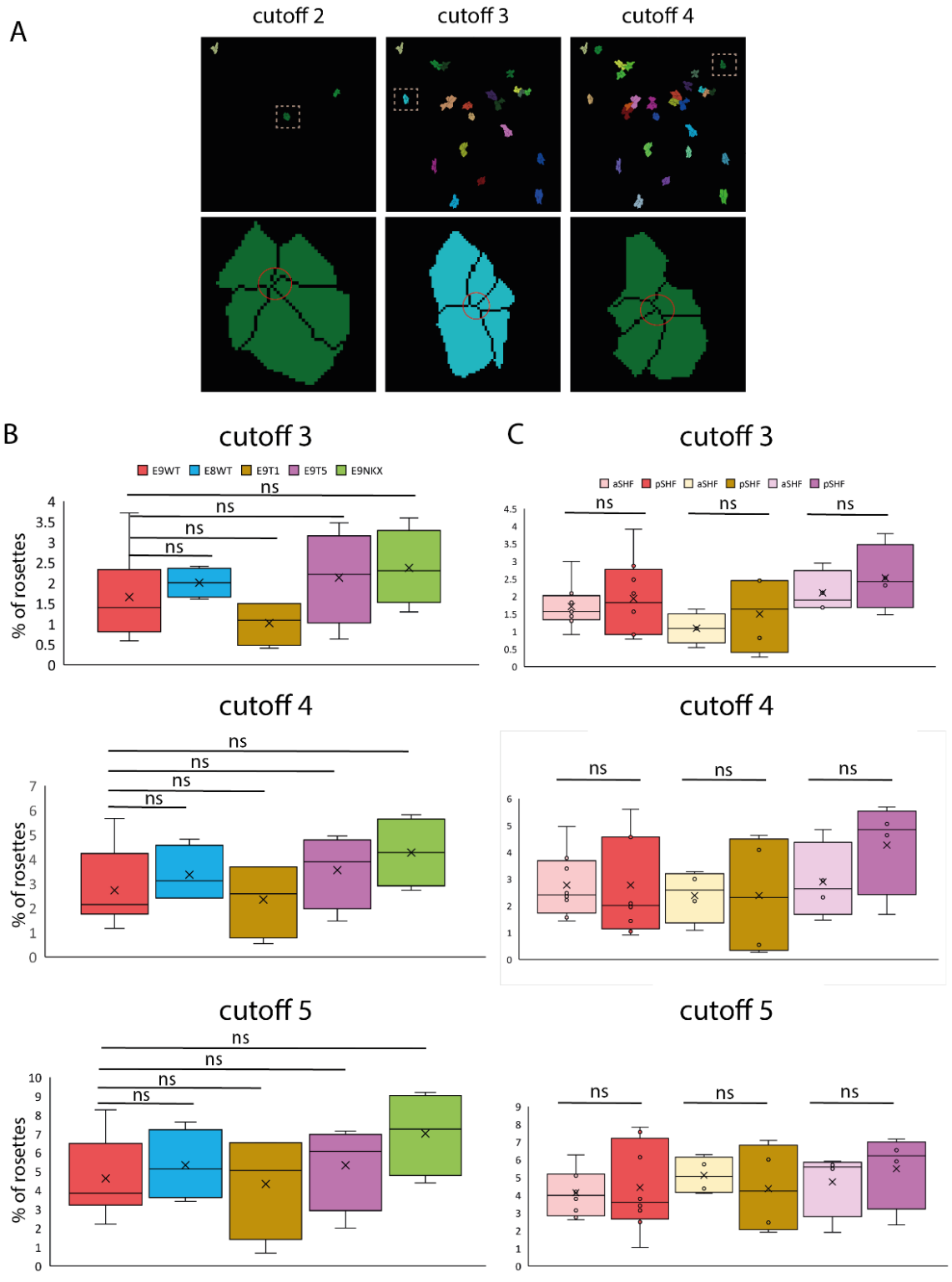


Fig. 3.3.2: Rosette quantification according to different bond length cutoffs A. Bond length with different cutoff **B.** Boxplots showing the quantification of the % of rosettes in the DPW of E8.5, E9.5 wild-type, *Tbx1*$^{-/-}$, *Mef2cAHFCre; Tbx5*$^{fl/fl}$ and *Nkx2.5* $^{-/-}$ **C.** Boxplots of the **%** of rosettes comparing the anterior and posterior DPW E9.5 wild-type, *Tbx1*$^{-/-}$ and *Mef2c-AHF-Cre; Tbx5*$^{fl/fl}$.

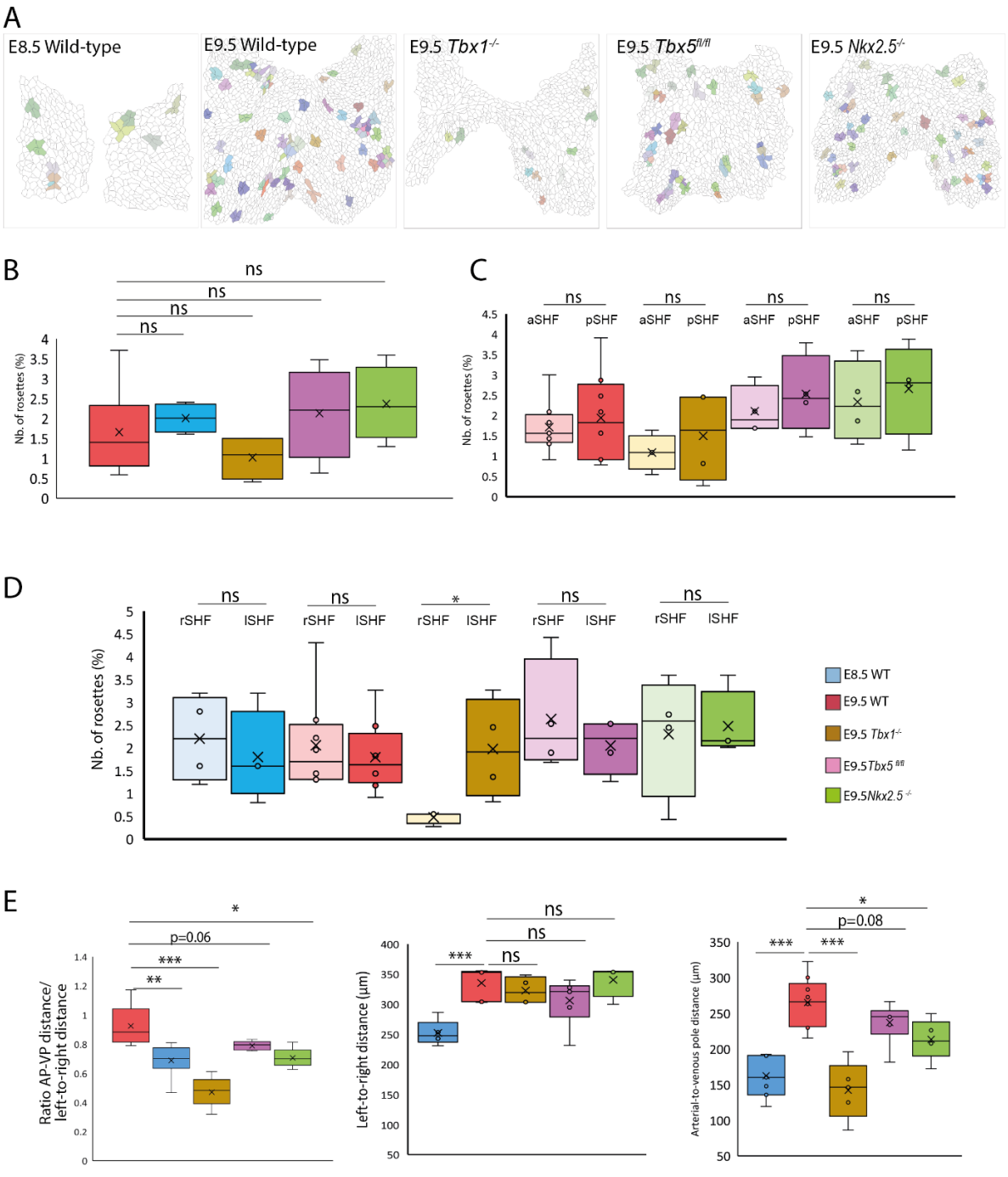
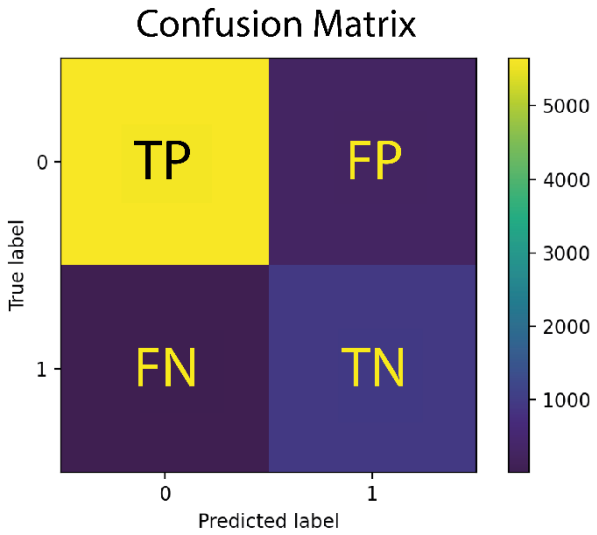


Figure 3.3.3: Rosette quantifications A. Segmentation masks where rosettes has been color-coded. The bond length cut off to determine a rosette vertex is step up to 3. Cells belonging to the same rosette have the same color; on the top right, boxplots of the quantifications of the percentage of rosettes **B.** Quantification of the number of rosettes for each mutation **C.** Quantification of the number of rosettes between the anterior and posterior SHF. Light color represents the anterior SHF and dark color represent the posterior SHF **D.** Quantification of the number of rosettes for each mutation **E.** Quantification of the left-to-right distance, the arterial-to-venous pole distance and the ratio between both distances. AP: arterial pole; VP: venous pole; R: right; L: left, ns = non-significant; *P<0.05, **P<0.01; ***P<0.001.

3.4 Exploring additional data analysis approaches

3.4.1 Classification task: Prediction of Tbx5 expression

In the initial section, we explored the diversity within the SHF and approached the issue through unsupervised methods using clustering algorithms. Subsequently we introduced a new dataset column indicating whether cells are positive for *Tbx5*. This column serves as our target variable. With this target variable, our revised problem statement is: Can we predict *Tbx5* expression in cells based on their apical morphological characteristics? This newly framed problem now involves a supervised learning task. This section of my thesis was performed as a small project during the Data science & Engineering bootcamp training I followed on machine learning.



To investigate this question, we trained multiple classification algorithms and evaluated their performance using accuracy and F1-score metrics, along with the confusion matrix. Accuracy measures the proportion of correctly predicted instances out of the total, while the F1-score is the harmonic mean of precision and recall. Precision represents the ratio of true positive predictions to all positive predictions, and recall measures the model's ability to identify true positives among actual positives. Finally, the confusion matrix (Fig. 3.4.1), that helps identify misclassifications by displaying the number of true positives (TP) and true negatives (TN), false positives (FN) and false negatives (FN).

The data set consists of 23 control images and 14 mutant images, only of the phalloidin channel. The data was divided into training and test sets, with an 80-20% split. During preprocessing, we observed a high correlation (0.94) between the area and perimeter features, prompting us to remove one of these parameters from the dataset. Highly correlated variables can redundantly complicate the model without enhancing predictive capability, potentially leading to overfitting.

Among the algorithms tested, XGBoost, a scalable tree boosting algorithm (T. Chen & Guestrin, 2016), showed the highest performance, achieving an accuracy of 0.97 on the test set and 0.99 on the training set, with F1-scores of 0.974 and 0.907, respectively. In our initial classification, in addition to morphological and mechanical characteristics, we incorporated the spatial coordinates (X and Y) of the cells. Feature selection algorithms highlighted the spatial coordinates as the most significant factors influencing the

classification of cells into Tbx5+ or Tbx5-. Thus, unsurprisingly, the spatial location of cells within the tissue emerged as the primary determinant for predicting Tbx5 expression, followed by the direction of stress tensor and cell orientation.

When we exclude the spatial coordinates (X and Y) and analyze only the morphological and mechanical features, we observe a decrease in both accuracy and F1-score. The primary features influencing cell classification are orientation and the primary stress direction, followed by pressure, stress anisotropy, perimeter, polarity eccentricity, and the number of cell neighbors. The metrics (Accuracy, F1-score and the confusion matrix) indicate that algorithms without spatial position information generally perform worse compared to those that include spatial coordinates, except for XGBoost, which maintains similar performance metrics as the previous dataset. Additionally, all models exhibit overfitting, as test metrics are higher than those of the training set, indicating that the models are memorizing the datasets rather than generalizing to new data (Fig. 3.4.2, Table 1and 2). These findings are consistent with the observation that clustering algorithms perform better when using joint spatial PCA rather than classic PCA. Spatial information plays a crucial role in determining cell fate and morphology. In conclusion, predictions of Tbx5 expression achieve better performance when considering spatial positions, with XGBoost being the best model, despite its tendency to overfit. To achieve better performance metrics and fight overfitting more data should be collected.

Model	Accuracy train	Accuracy Test	F1 score train	F1 score test
Logistic Regression	0.793	0.781	0.538	0.52
Decision Tree	0.793	0.945	0.538	0.805
Random Forest	0.959	0.931	0.869	0.781
SVM with radial basis function	0.97	0.923	0.903	0.75
Adaboost	0.978	0.96	0.927	0.866
Gradient Boosting	0.941	0.938	0.763	0.752
XGBoost	0.993	0.974	0.974	0.907

Table 3.4.1: Table with the performance metrics obtained with different models considering the spatial positions and the morphological features of the cells.

Model	Accuracy train	Accuracy Test	F1 score train	F1 score test
Logistic Regression	0.862	0.863	0.004	0
Decision tree	0.864	0.863	0.041	0.033
Random forest	0.984	0.833	0.944	0.205
SVM with radial basis function	0.999	0.92	0.998	0.619
Adaboost	1	0.842	1	0.056
Gradient boosting	0.932	0.941	0.719	0.717
XGBoost	0.9993	0.982	0.997	0.923

Table 3.4.2: Table with the performance metrics obtained with different models not taking into account the spatial coordinates of the cells.

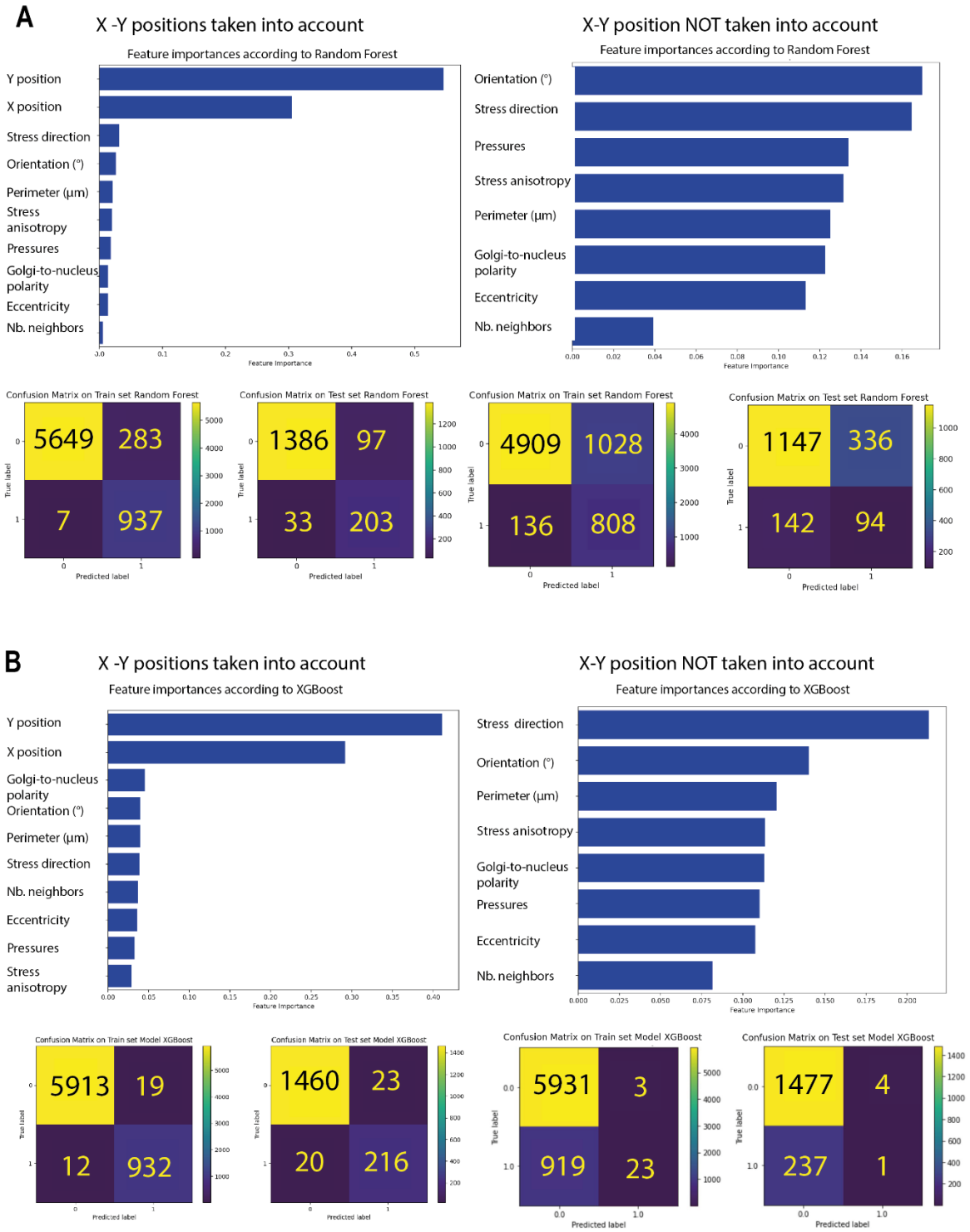


Figure 3.4.2: Feature selection and confusion matrix of random forest (A) and XGBoost (B) models. The left panels show the results considering the coordinates of the cells, the right panels show the results not considering the x-y coordinates of the cells.

3.4.2 Optional clustering method

Other possible ways of clustering methods where explored during the bootcamp. In this instance, an exploratory data analysis of the features was conducted prior to clustering, using a correlation matrix and pair plot to examine all the features in the dataset. Highly correlated features, with a correlation coefficient higher to 0.9 were discarded, maintaining only one of the pairs in the dataset. Missing values were drop as well, due to their small number (<30 rows) and outliers were maintained in the dataset. The data was centered and normalized using Standard Scaler method. Dimensionality reduction methods (PCA and UMAP)(Becht et al., 2019; Pedregosa et al., 2011.), were performed to understand the relationship and the shape of the data, using sk-learn and seaborn Python libraries (Pedregosa et al., 2011).

Once the data followed all the previous preprocessing steps, the elbow and the silhouette methods were used to choose the optimal number of clusters (k). The elbow method works as follows: it runs k-means clustering on the dataset for a range of values for k (say from 1-10) and then for each value of k computes an average score for all clusters. By default, the distortion score is computed, the sum of square distances from each point to its assigned centre (Thorndike, 1953, Pedregosa et al., 2011). On the other hand, the silhouette methods are used to evaluate the quality of clustering by studying the separation distance between the resulting clusters. It displays a measure of how close each point in one cluster is to points in the neighboring clusters. The scores range from -1 to 1. Positive scores mean that most of the points are clustered appropriately; scores close to 0, mean that clusters are overlapping or poorly defined, while negative scores points are assigned when points are assigned to the wrong clusters. Once the number of clusters is stablished clustering algorithms were fitted on the E9.5 wild-type data. For reproducibility, random state was fixed to 0. The best spatial results were obtained when the first Principal Component is used in the clustering. K-means is the best model clustering the data, respecting the spatial positions of the cells. The results were assessed by plotting back the labels in the real coordinates of the embryo space. To understand the cluster signatures distribution, histogram plots for every feature were assess. The steps and code can be found in: ClaraGCH/CardiacMap

Unlike the jsPCA pipeline, this method discards highly correlated features and does not consider the Moran's index or the spatial coherence of the features. As a result, this pipeline is more general and suitable for non-biological problems. In contrast, jsPCA provides additional information about the spatial significance of features and it is better suited for spatially coherent issues, such as those found in biological contexts like spatial transcriptomics or single-cell morphometrics.

4. Discussion

4.1 Discussion on single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the SHF

In this thesis we have developed and applied a single-cell morphometric pipeline to characterize the temporal progression of the dorsal pericardial wall structure from E8.5 to E9.5, a critical period in heart development. Additionally, we have evaluated alterations in the wall's architecture under three different genetic conditions.

The morphological and mechanical cartography serves dual purposes: it visually depicts and characterizes the epithelium while also highlighting the correlation between *Tbx5* activation in the wall and the acquisition of specific morphological and mechanical characteristics. *Tbx5* plays a crucial role in patterning and establishing the venous pole of the heart. Thus, the allocation of epithelial cardiac progenitor cells to the arterial and venous poles is influenced not only by two distinct genetic programs but also by different tension and orientation patterns.

Although some past limitations have been successfully addressed, there are still some constraints that need to be addressed: 1) One of the biggest limitations on the study of the second heart field cells is the absence of live imaging techniques. For this project, we have used snapshots at different time points during heart development from E8 to E9.5. Currently it is not possible to record deep enough (>400µm) and at a cellular resolution, for long time periods and keep the sample intact. To meet this requirement, a light-sheet microscope with the resolution of a confocal microscope must be engineered, integrated with an embryo culture system that does not necessitate rolling the embryos. The most advance microscopes so far are the light-sheet microscope developed in Janelia capable of adapting the optical properties to the changes in size and shape of the post-implantation mouse embryo from gastrulation to early organogenesis at cell level (McDole et al., 2018). More recently at the Gladstone Institute, live light-sheet fluorescence microscopy has allowed whole-embryo time-lapse imaging of mouse post-implantation development for between 9 and 18 hours (Dominguez et al., 2023). However, these experiments present limitations such as phototoxicity, divergence with the normal development environment, tracking inaccuracies and restrictions on long duration imaging, and do not yet extend into the time of extensive SHF deployment after E8.5. Similarly, embryo culture technics have advanced, allowing appropriate development of embryos from E5.5 until E11 (Aguilera-Castrejon et al., 2021) and covering the heart tube elongation process. The development of approaches for live imaging in the second heart field have led to new discoveries in the heart development field such as the slice culture experiments that help to characterized the epithelial properties of the dorsal pericardial wall, specially the dynamics of the basal filopodia in

the second heart field (Francou et al., 2014). Live imaging experiments were used to study the migration patterns of the distal visceral endoderm (DVE) as well as to identify a new subgroup of cells within the DVE (Stower et al., 2023). Similarly, live imaging of the mammalian epiblast reveals that, despite being temporal and spatially close, endothelial precursors diverge early from other mesodermal derivatives; in particular, cardiomyocytes and endocardial cells show a limited lineage relationship (Sendra et al., 2024) 2) The lack of confocal microscopes allow imaging of up to 4 channels, so the imaging and measurement of cellular morphometrics are restricted to the 4 proteins measured. Novel spectral microscopes allow to increase the number of channels but still, they are limited to only 10 proteins maximum (Lelouard et al., 2018). Quantifying the localization of proteins involved in cell polarity and proteins characteristic of epithelial cells, such as pericentrin, gamma-tubulin, aPKC, par3, E-cadherin, or N-cadherin, would be valuable for assessing further morphological parameters of the cells. Additionally, imaging transcription factors involved in cardiac development would enable the categorization of cells and facilitate correlation and classification tasks between transcription factor expression and specific morphological features 3) The lack of consideration consideration for adjacent structures and tissues. Heart development depends on the adjacent tissues and structures such as the extracellullar matrix (Arriagada & Astrof, 2023) and the endoderm (Kidokoro et al., 2018; Varner & Taber, 2012; Esteban et al., 2022). These structures may play passive roles as a substrate for the SHF, that we have not consider when studying the elongation of the heart tube and the addition to the poles.

This study highlights a different role of *Tbx5* beyond the one involved in cardiac specification and differentiation, *Tbx5* is required to establish epithelial tension and polarized ppMLC2 accumulation in the pSHF. *Tbx5* has always been associated with the FHF and its role in cell maturation into cardiomyocytes. Works supporting a role of *Tbx5* in tension include recent studies by the Yelon group in Zebrafish. They show that *tbx5a* regulates cardiac chamber morphology, specifically the ventricular morphology of the inner and outer curvature cells, by regulating the actin cytoskeleton. *Zebrafish tbx5a* mutants exhibit disruption of the actomyosin cytoskeleton. The outer curvature of *tbx5a* mutant hearts undergo less lateral extension and more apico-basal extension than wild-type outer curvature cells; while cells in the inner curvature of *tbx5a* mutant hearts undergo less apico-basal extension than wild-type inner curvature cells. Experiments perturbing ppMLC polarization, such as laser ablations, the use of myosin inhibition drugs such as Blebbistain or the perturbation of myosin light chain (MLC) phosphorylation by Y27632 drug to inhibit Rho kinase, would be interesting to investigate whether the tension pattern can control the activation of *Tbx5* gene or to verify that ppMLC polarization is required for the addition of the pSHF cells to the venous pole. Other mechanisms for testing actomyosin contraction would be the use of optogenetics to recruit Rho kinases, as in a study where selective optogenetic activation on the apical side of dorsal embryonic cells recruited RhoGEF to the membrane of the cells in *Drosophila* leading to tissue invagination demonstrating that apical constriction is sufficient to induce

deformation in the context of Drosophila invagination (Izquierdo et al., 2018 ; Martínez-Ara et al., 2022).

Along the same lines it would be interesting to overexpress *Tbx5*, for example using AAV vectors, in different areas of the SHF and analyze the effects on cell morphology (cell orientation, stress direction and cell stress anisotropy) or by the engineering of optogenetic versions of the Tbx5 transcription factor (Singh et al., 2022). Perturbation of Tbx5 expression at a precise location and time would be interesting for studying the Tbx5 downstream targets genes involved in the arousal of tension. Apart from the direct role that Tbx5 has on the cells expressing this gene, we have observed that the effects extend the non-Tbx5 region. The posterior lateral part of the SHF also experiences the same morphological changes as the Tbx5 region, showing an autonomous and a non-autonomous effect of Tbx5 in the wall. The tension created in the pSHF may be the responsible of the cell addition at the venous pole and draw the epithelium in opposing directions. In Nkx2.5 mutants, a mutant where tension is abolished, addition to the poles is blocked.

As there might be a new role of Tbx5 related to cell mechanics. It would be interesting to investigate whether the transcription factors and partners of TBX5 in FHF and posterior SHF are the same, as well as Tbx5 target genes. Here, we have addressed the changes that the absence of Tbx5 induce in the SHF epithelium. However, due to its importance in CHD, Tbx5 has already been extensively studied (Bruneau et al., 2001 ; Steimle & Moskowitz, 2017; Bruneau et al., 1999 ; Mori et al., 2006), it is known to be involved in cardiac conduction system, in early specification of cardiac progenitor cells, as well as in atria and left ventricle development. Mutations in *TBX5* are associated with Holt-Oram syndrome characterized by upper limb and heart defects. *Tbx5* null mouse embryos die in utero around E10.5, similar to the *Mef2cCre* conditional *Tbx5* mutant embryos. Previous studies have shown that homozygous null embryos exhibit a grossly abnormal linear heart tube and complete absence of forelimbs (Bruneau et al., 2001) as if the growth of the heart and the deployment to the venous fails in absence of *Tbx5*. On the contrary, partial loss of Tbx5 make the hearts undergo heart looping with the formation of hypoplastic left and right atria (Mori et al., 2006). In this study, we did not examine the impact of *Tbx5* loss on the cardiac structures, already studied in other laboratories (Bruneau et al., 2001; Steimle & Moskowitz, 2017). However, our results regarding the absence of tension parallel to the antero-posterior axis in the posterior SHF support the hypoplasia of atrial structures observed in earlier studies. This suggests that the cellular changes in the SHF, including the absence of the tension and the incorrect alignment of the cells, may contribute to preventing posterior SHF cells from integrating into the developing heart.

Tbx5 is a transcription factor that acts as part of a multifactor transcriptional complex. The first identified partner of Tbx5 is Nkx2-5, it also interacts with other major cardiac transcription factors such as GATA4, GATA6, TBX20, MEF2C and myocarin (Steimle & Moskowitz, 2017). Interestingly, for cardiac reprogramming studies, a cocktail containing

the key transcription factors regulating cardiac development has already been established, including the SHF venous pole marker TBX5 but also GATA4, MEF2c and HAND2 transcription factors (Qian et al., 2012; Chen et al., 2017). Meaning that these transcription factors are sufficient to give a cardiomyocyte identity although with a low reprogramming efficiency, not even reaching 50% of efficiency (Chen et al., 2017). To increase the efficiency rate, new combinations with transcription factors, small molecules and miRNAs have been created. However, none of these strategies address the role of mechanical tension in their reprogramming protocols. Furthermore, if mechanical tension and forces are important in maintaining distinct anterior and posterior fates, it would be interesting that in vitro reprogramming techniques include them in the protocol. For example, changes in substrate and the absence of surrounding tissues may be a cause of reduced reprogramming efficiency.

Other strategies consist in stimulating the cardiac regeneration potential by stimulating proliferation and dedifferentiation through the forced expression of YAP an effector of the Hippo pathway signaling, a coactivator that it has been shown to localize in the nucleus of the pSHF cells in presence of *Tbx5* (Francou et al., 2017). In *Tbx5* conditional mutant, YAP's nuclear localization is observed in a salt-and-pepper pattern (data not shown). YAP and TAZ transcriptional coactivators directly associate with *Tbx5* and stimulate expression of *Tbx5*-dependent genes (Murakami et al., 2005). Studies supporting the identification of YAP/TAZ partner of Tbx5 help to explain the correlation between the high colocalization between nuclear YAP and *Tbx5* expression. However, all reprogramming strategies have low efficiency rates and ignore the mechanical and physical inputs that cells require to specialize in cardiomyocytes and correctly shape the adult heart (Fig. 4.1A). Interestingly, studies on zebrafish, have highlighted the role of the mechanosensory Piezo proteins in outflow tract and aortic valve development (Duchemin et al., 2019; Faucherre et al., 2020). Piezo can detect mechanical changes in the environment and influence pathways like Hippo pathway, further supporting the importance of mechanical cues in cardiac regeneration and development.

Descriptive and mechanistic studies in heart development are essential to develop effective and coherent regenerative strategies. Gaining insight into the formation of the heart, the genetic and the physical cues, the timing and sequence of the rules affecting progenitor cells in order to create the final heart architecture are key information for the field of heart bioengineering. This emerging field aims to advance the development of gastruloids (Rossi et al., 2021) and cardioids (Schmidt et al., 2023) as new models for studying cardiogenesis, heart defects and drug testing (Fig. 4.1B,C).

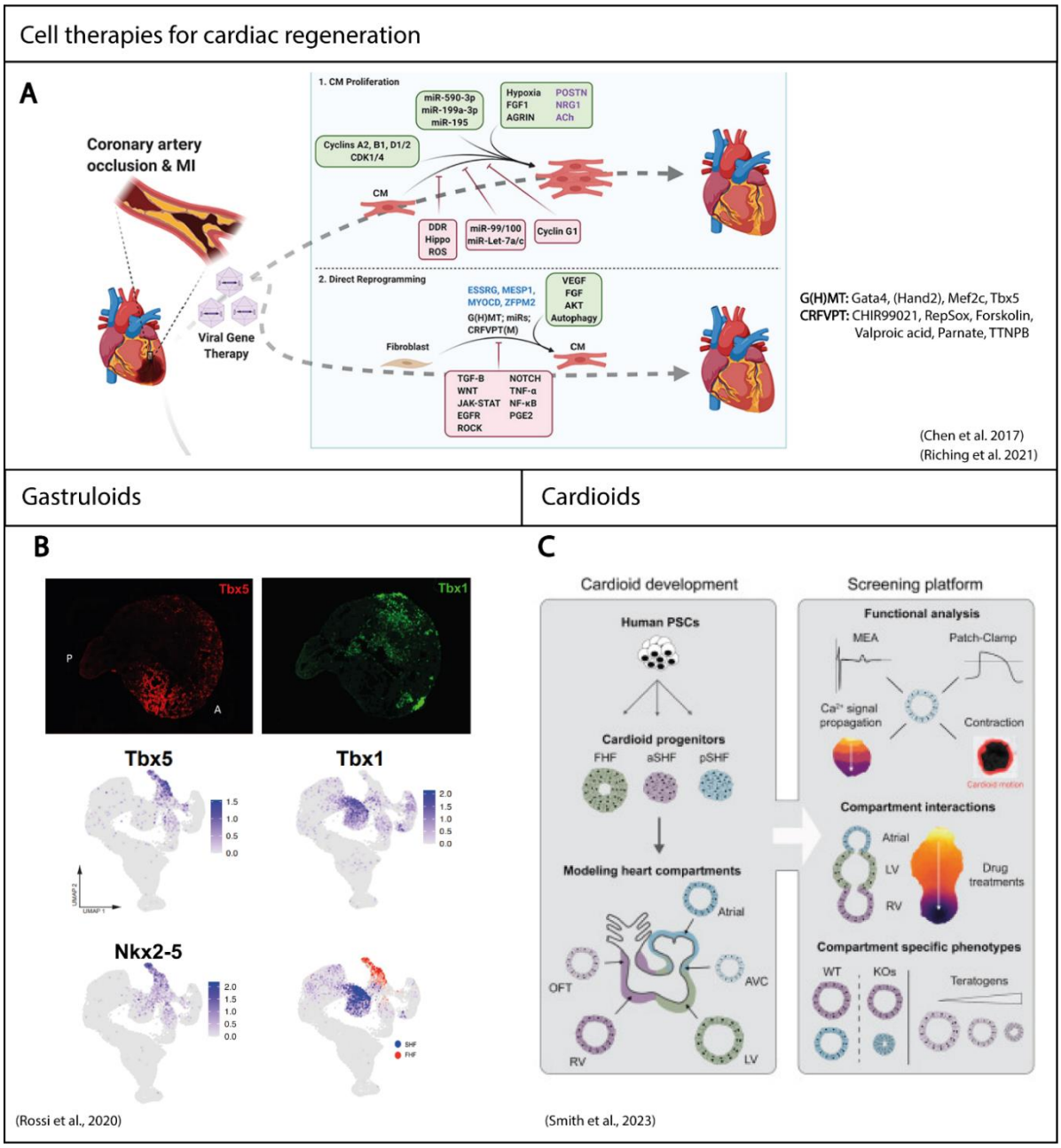


Figure 4.1. Perspectives in the cardiac field fed by the discoveries in heart development. A. Cell therapy strategies for heart regeneration. Schematic indicating factors that regulate CM proliferation and direct reprogramming of fibroblasts to CMs to regenerate the myocardium following MI. Green boxes indicate environmental, signaling, and genetic factors that promote CM proliferation/fibroblast reprogramming. Red boxes indicate factors that impair these processes. Factors that have conflicting reports regarding their roles in inducing CM proliferation are listed in purple text. Blue text indicates factors that enhance direct reprogramming in human fibroblasts (Riching & Song, 2021). DDR: DNA damage response, GHMT: GATA4, HAND2n MEF2C, TBX5, CRFVPT: CHIR99021, RepSox, Forskolin, Valproic acid, Parnate, TTNPB. DDR: DNA damage response, GHMT: GATA4, HAND2n MEF2C, TBX5, CRFVPT: CHIR99021, RepSox, Forskolin, Valproic acid, Parnate, TTNPB. **B.** Gastruloids mimic early cardiac morphology,

form distinct first and second heart fields and recapitulate beating. **C.** Cardioids are formed of progenitor cells belonging to the FHF, aSHF and pSHF and model heart compartments. PSCs: pluripotent stem cells, FHF: first heart field, SHF: second heart field, aSHF: anterior second heart field, pSHF: posterior second heart field, WT: wild-type, KOs: Knock out, OFT: outflow tract, AVC atrioventricular canal, RV: right ventricle, LV: left ventricle, MEA: multiple electrode arrays.

To address the complexity of the DPW, we have followed an unsupervised approach by clustering the cells according to some of their morphological features (Fig. 4.2), this analysis receives the name of single-cell morphometrics. Clustering analysis is part of the single-cell morphometric analysis, it is an unsupervised learning technique where the algorithm identifies patterns and groups in the data based on similarity or distance metrics. I used clustering because I thought it would be a good solution to discover hidden structures in the data without prior information about the groups. In the nature, data is often continuous and complex, and clustering offers is a way to fractionate the continuous data to make it more interpretable. Sc-morphometric analysis has already been used in the field of biology to study the development of amphioxus notochord, where 9 features obtained from the 3 dimensions of the cells were used for the analysis (angle of major ellipsoid axis, number of triangles, volume, relative AP-DV-ML dimensions, AP anisotropy, nuclear displacement from the centre of mass, cuboidness, sphericity, flatness) , PCA and clustering , clusters were defined as 6 bins of equal length across the AP axis of the 14ss notochord, and for the trajectory inference K-means clustering was used instead of the somite stage (Andrews et al., 2021). A more recent study using not morphological but mechanical and behavioral properties of the cells, use a similar analysis by collecting 14 features or phenomic signatures and used a k-means clustering on the 2D UMAP coordinates, that was used as an input for hierarchical clustering with Ward linkage, helping to discover a previously unknown sub-group within the distal visceral endoderm (Stower et al., 2023).

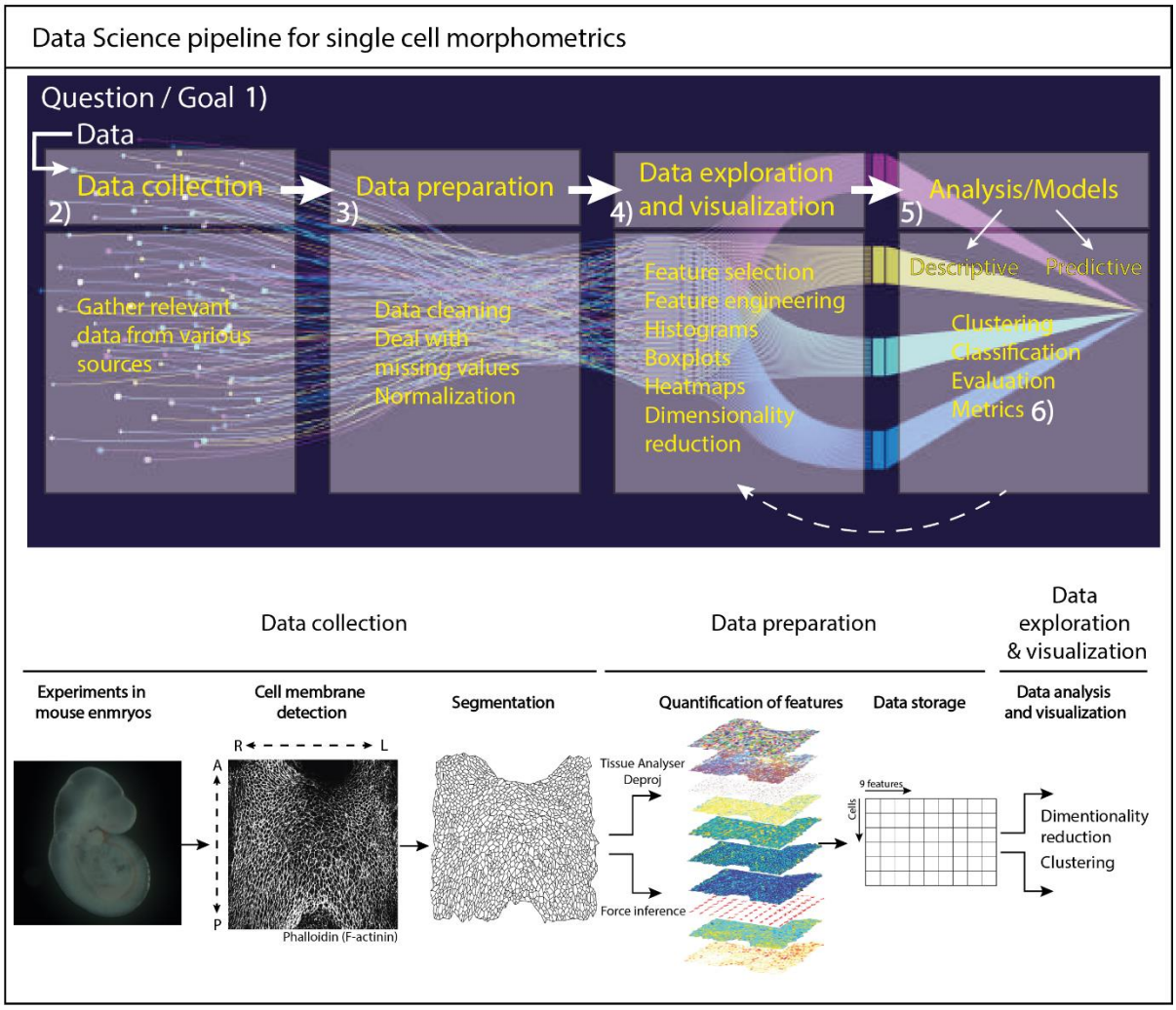


Figure 4.2. Data science pipelines incorporated in biological research.

If we compare the 2 cluster methods used: k-means (Ikotun et al., 2023) and agglomerative clustering (Murtagh & Contreras, 2012), k-means clustering is more sensitive to the outliers than agglomerative clustering. K-means creates spherical clusters and it will always use all the datapoints and include them into a cluster, with the consequence of changing the clusters center and the distribution of the clusters. On the other hand, agglomerative clustering is less sensible to the outliers. According to my dataset a k-means clustering works well and the clustering happens in one dimension (stress and orientation). If a tissue would not be under stress, then these parameters would not be informative, maybe other features would be more important as in the case of the quail embryo (Kong et al., 2019) where the main features are area and perimeter.

Based on our findings, Cluster 1 and Cluster 2 cells are associated with the arterial pole, while Cluster 3 cells are primarily found in the venous pole. Although it could be argued that some anterior cells are labelled as Cluster 3, the main group of Cluster 3 cells is localized in the venous pole. Clustering is a global algorithm, which might have problems in some missassignments. Additionally, outliers and the intrinsic noise of the data can introduce biases. It would be overly simplistic to expect algorithms to perfectly capture

the complexity of natural biological systems. Therefore, the clusters we have identified not as absolute truths but rather as indicators or trends in cell patterning. In our system, there is 2 main clusters created by two colliding stresses; one tangential to the outflow tract with lower stress at the arterial pole and another perpendicular oriented parallel to the anterior-posterior axis at the venous pole. The anterior cluster can be subdivided in 2 clusters that alternate, due to rotation of the outflow tract and the maintenance of the left-right symmetry. Depending on the properties and the phenomena shaping the tissue, the results would be different. Analyzing this type of data requires either an expertise in both data science and biology. Dual expertise is essential to ensure accurate analysis and meaningful interpretation of the data.

The concept of using clustering to create a map to define territories and boundaries is rooted in 2 different ideas: 1) Understanding clustering and the natural tendency of similar entities to cluster together, particularly in light of the recent widespread adoption of dimensionality reduction and clustering analysis when treating sc-RNAseq data 2) Considering the body as an organized society where each cell is viewed as an individual.

What it comes to the dimensionality reduction algorithms, these algorithms are performed in the preprocessing step to have an idea of the distribution of the data. UMAP and PCA are the dimensionality reduction algorithms mainly used, whose purpose is to find a simplified space that explains the highest variance of the data, they can be considered as a visualization tool that can give you a hint about the distribution of the data and can bring to the surface clusters that are separated in space (Jolliffe & Cadima, 2016). The goal of the algorithms is to find the 2 axis that explains the better the variance of the data. If we take as an example PC1 in our data, this axis is explained by area, perimeter and pressures while PC2 is formed by stress direction and orientation. Together, these 2 axes explain nearly the 50% of the variance of the data. In our paper a new dimensionality reduction algorithm has been used, we called joint spatial PCA (jsPCA). In the case of the morphological and mechanical data of the DPW, Moran's Index, which is a readout for spatial coherence, shows that the classical PCA is not spatially informative. This is why we used a spatial PCA (sPCA) that combines the classic PCA and the Moran's Index; in other words, in spatial PCA, the principal components are the product of the covariance matrix and the connectivity matrix. However, the spatial PCA is specific for each tissue sample. To overcome this problem, we developed the joint sPCA (jsPCA) by jointly diagonalizing the sPCA matrix from all the E9.5 wild-type embryos. This new dimensionality reduction algorithm gives as a result that the jsPC1 is formed by orientation and stress direction while the jsPC2 is formed by area and perimeter, results aligned with the real spatial distribution of the cells.

The fact that cells are able to respond to the rules of genes and mechanical constraints raises the point of which is stronger: the genes or the mechanical constraints. In the SHF and according to our results *Tbx5* is modulating the mechanics of the epithelium. The triggering event is the activation of Tbx5 in this population. *Tbx5* is triggering, stress,

polarizing ppMLC, inducing proliferation and changing the orientation of the cells parallel to the antero-posterior axis at the pSHF (Fig. 4.3A, B).

Combining the gene expression territories and the morphological features of the second heart field over time would yield valuable information. Integrating data from spatial transcriptomics, MERFISH, MERSCOPE and other technics, that enable to perform single-cell transcriptomics in the context of the tissue microenvironment, together with, spectral microscopes that can image up to 9 fluorophores in a single image, could provide a complete characterization map of the SHF (Fig. 4.3A).

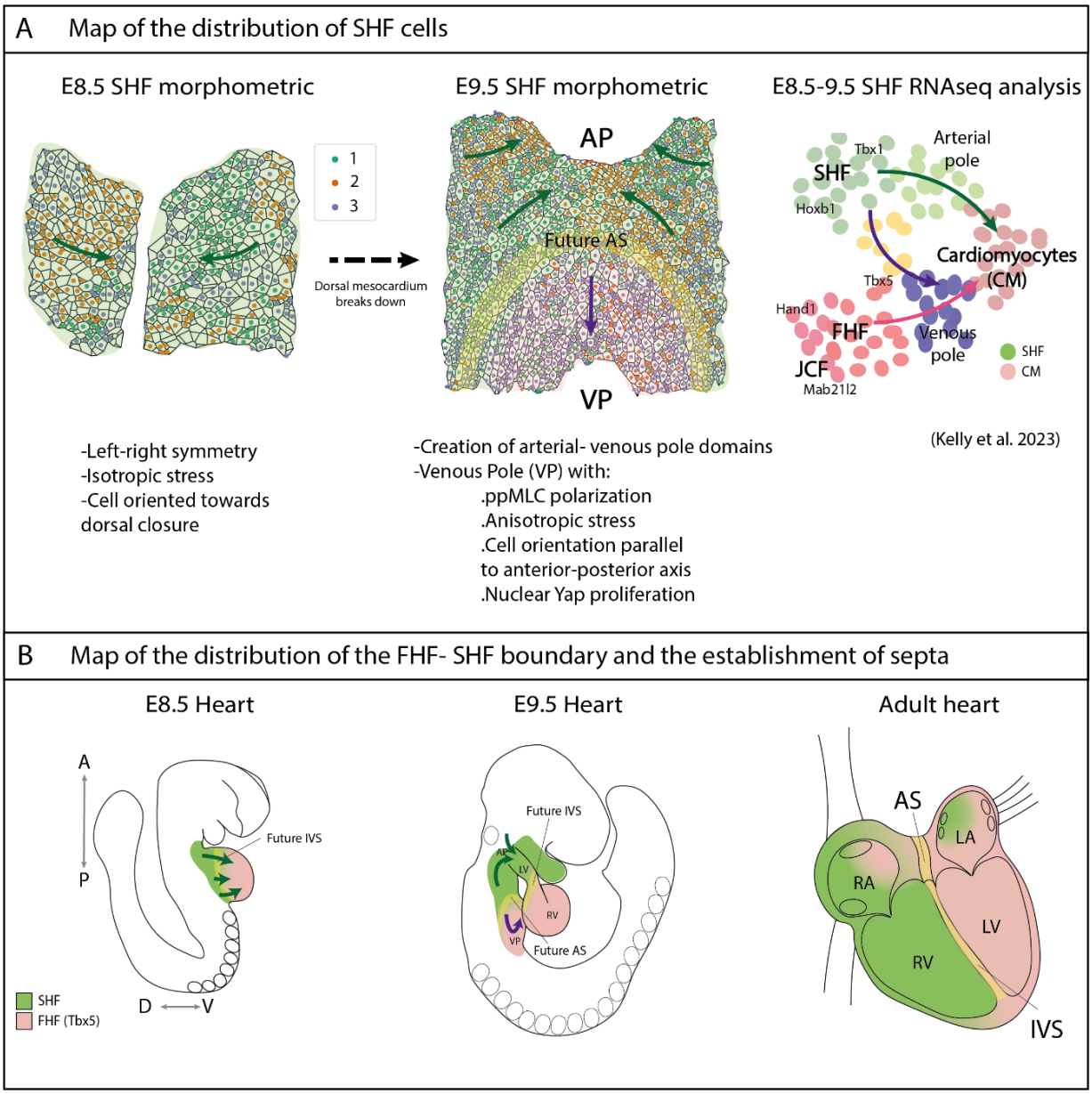


Figure 4.3. Map and trajectories of the Second Heart Field cells. A. Map showing the potential regions and distributions according to the sc-morphometrics and the *Tbx1-Tbx5* domains of the Second Heart field cells at E8.5 and E9.5. On the right, a two-dimensional representation of scRNA-seq data based on (Tyser et al., 2021) and (Zhang et al., 2021)

studies showing inferred Juxta-cardiac JCF/FHF and SHF trajectories to myocardium through intermediate venous and arterial pole differentiation states (Kelly, 2023) **B.** Cartoons showing the lateral view of the embryos, highlighting the contribution of the multilineage progenitor cells to the growth of the heart, at E8.5 through the dorsal mesocardium and at E9.5 by alternating the poles of the heart, in yellow the overlap between the 2 heart fields is highlighted. On the right, a carton of the adult heart showing the atrial and septal septa formed coinciding with the boundary between the first and second heart fields.

4.2 Discussion on boundary formation in the SHF

The analysis of the morphological and mechanical properties of the SHF, along with sc-RNAseq analyses, suggest the existence of a boundary and the formation of distinct populations within the SHF in the Dorsal pericardial wall. This boundary can be visually detected and forms where the *Tbx1* and *Tbx5* intersect and the 2 stress and cell orientation patterns collide.

Concerning analysis of the sc-RNAseq analysis, the initial dataset contains 21,769 features (genes) across 572 samples (cells), due to the nature of the data we need to have some considerations: 1) The number of features (21769 genes in our case) can affect the performance of the clustering. High-dimensional data can lead to the curse of dimensionality (Altman & Krzywinski, 2018), where the distance between points becomes less meaningful and clustering algorithms may struggle to find meaningful clusters. Plus, these effects are amplified by poor data quality 2) To overcome the first consideration, the use of preprocessing steps or feature selection improve the clustering algorithm. The main to-take-home idea from point 1-2 is that having more characteristics or quantifying more features in your dataset does not give you better and more meaningful clusters. For example, in the study of Iris plants only 4 features were necessary to cluster and differentiate the 3 types of flowers (Jombart et al., 2010) (Fig.1.18.2AB). Irrelevant features increase noise in the results. Preprocessing steps like reducing the number of dimensions and removing outliers are necessary. Performing the analysis using the principal components help on finding the trends and patterns of the data. However, reducing too much the features may lead to the loss of information. In our case, in the sc-RNAseq experiment, during the preprocessing step cells with over 20000 transcripts, less than 500 transcripts and cells whose 15% of their transcripts are mitochondrial were removed. The single-cell RNA seq dataset for analysis then comprises 7636 features (genes) and 460 samples (cells) after preprocessing. Next, Principal component analysis is performed showing that the 10 principal components are capturing the highest variability of the dataset. Then KNN clustering is performed in the 10 principal components and K is set to 6, revealed by dimensional reduction visualization. 3) As in the single-cell morphometric clustering, sc-RNAseq clustering methods are largely exploratory, the interpretation of the clusters depends on the domain

of knowledge, the results should not be considered as a truth but more as a hint for next experiments. The results of the clustering have to be always verified by RNA scope experiments, in situ hybridization and the analysis of other databases.

Two types of boundaries have been defined depending on the behavior of the cells in the boundary in response to a perturbation: the non-lineage boundaries and lineage or compartment boundaries. At non-lineage boundaries, cells can move across gene expression boundaries and adapt their identity to that of their neighbors, in this case fate determination requires an upstreaming signaling input that continuously instruct the cells. According to the observations of this and other studies (Francou et al., 2017; De Bono et al., 2018) the *Tbx1-Tbx5* boundary of the second SHF can be considered as a non-lineage boundary driven by retinoic acid signaling. At the onset of the formation of the continuous SHF epithelium, the tissue is undifferentiated, the cells undergo events such as convergent extension or proliferation and move across gene expression boundaries. At this stage the SHF is a dynamic tissue with constant cell movements and rearrangements. Venous pole fate determination is instructed by retinoic acid signaling at the posterior medial part of the SHF. This differentiation creates a physical boundary that will subsequently restrict cell-mixing and transform the boundary into a lineage boundary.

At lineage or compartment boundaries, the identities of cells are inherited and displaced cells are sorted back into the territory of cells with the same identity; in this case, there might already be signatures in the tissue and the cells to restrict the mixing and maintain straight boundary between the populations. To recognize this type of compartments lineage-tracing experiments were used where single cells grow into large patches of cells, called clones. Similarly, in the SHF the analysis of cell clones with *Mesp1-Cre/Rosa-Confetti* embryos show an anisotropic and oriented shape directed towards the arterial pole of the heart (Francou et al., 2017). Such clusters appear to respect the borders observed by the sc-morphometrics. A border once established may thus be similar to a compartment boundary, the 2 cell populations representing separate compartments. Sometimes a compartment or lineage boundary corresponds to a gene expression domain boundary, as in Drosophila where *engrailed* expression controlling the development of the posterior compartment cells (Kornberg et al., 1985) . These compartments have a defined identity thanks to what are called selector genes, the loss of which eliminates the identity of the territory and their ectopic expression induce their identity in the new territory. The concept of selector genes is principally related to compartment boundaries, but it can also be found in the non-lineage boundaries. An example of a selector gene is *engrailed*, which is specifically expressed in all cells of the posterior compartments of *D. Melanogaster*, preventing their mixing with the anterior cells (Kornberg et al., 1985, Garcia-Bellido & Santamaria, 1972). Other example of a selector gene in a non-lineage boundary is *Mesp2* which is a central player in the pre-patterning of the paraxial mesoderm (PSM). Loss of Mesp2 in the mouse embryo causes a loss of somites and caudalization of the PSM (Takahashi et al., 2000), while its overexpression drives rostralization of the PSM (Saga et al., 1997). Similarly, in the case of SHF development, *Tbx5* and retinoic acid signaling

could be considered a selector gene/signaling pathway regulating the pSHF. Loss of Tbx5 in the SHF causes failure of the pSHF to identify as a venous pole domain, and a similar phenotype is observed when retinoic acid (RA) signaling is blocked. However, contrary to what happen with Mesp2, overexpression of RA show no posteriorization of the anterior domain (De Bono et al., 2018). This lack of effect could be due the expression of Cyp genes in the anterior SHF, that degrade RA, rather than a physical barrier (Rydeen & Waxman, 2016).

After the boundary is formed, it needs to be maintained. In addition to the activity of selector genes; signaling pathways and physical interactions are other sources of boundary maintenance. For example, Eph-ephrin signaling has been shown to be important in the process of somite formation in zebrafish, lack of EphA4 drives an absence of somites (Barrios et al., 2003) or ephrinB2, the ligand is involved in the coordination of boundary formation and cell epithelization in chick embryos (Watanabe et al., 2009). Whether ephrins are playing a role in the boundary maintenance between *Tbx1* and *Tbx5* remains unknown.

Physical mechanisms can also play a role in the boundary maintenance by: reducing cell proliferation, deposing extracellular matrix, differential cell adhesion and mechanical tension. In the case of the SHF, the arterial and venous pole boundary is a sharp and straight boundary. The higher the tension at the interface, the 'straighter' the interface became (Landsberg et al., 2009). A computational pipeline for spatial mechano-transcriptomics in which boundaries between compartments are characterized by elevated interfacial tension was used in the boundary analysis of neural crest cells-midbrain region and midbrain-cranial mesoderm (Hallou et al., 2023). The existence of a high-tension boundary between the *Tbx1* and *Tbx5* domains was tested in preliminary experiments, revealing that 2 out of 4 E9.5 wild-type embryos had a higher tension level on the boundary cells compared with the rest of the domain. A different way to measure the mechanical tension on cell borders is to check the velocity of cell vertex displacement after laser ablation (Landsberg et al., 2009). A possible experiment that could be proposed to this end, would be the ablation of cell borders of $Tbx1^{+}$ cells and $Tbx5^{+}$ cells and measurement of the velocity of cell vertex displacement, and the ablation of cells borders between the anterior and posterior domain. Collaboration with theoreticians to simulate mechanisms of tissue growth, such as proliferation or convergent extension, differential tension and pressures by the use of vertex models or other computational models, can provide insights into the mechanism of dorsal closure and the later elongation and deployment of the SHF cells. Some investigations have even incorporated computational model describing morphogen transport coupled to a vertex model (Smith et al., 2012). This study showed that the growth rate of each cell increases linearly with the local concentration of a morphogen. Another example is the use of geoadditive structural equation modeling in order to predict the mechanical behavior of the cells by identifying gene modules (Hallou et al., 2023).

Here, we have shown that physical mechanisms like tension and cell elongation divide the SHF into territories depending whether they are under the arterial or venous pole influence. To study the existence of further selector genes or signaling pathways or a mechanism for maintenance of the *Tbx1-Tbx5* boundary, RNA-seq experiment and single-cell analysis, help to pinpoint genes expressed in the double positive cells, such as *Chn2*, *Cldn11*, *Aldh1a2* or *Gm5127*. *Chn2,* known to be involved in proliferation and migration; *Cldn11*, membrane proteins and components of tight junctions important as physical barriers and maintenance of cell polarity; *Aldh1a2*, an enzyme that synthetize retinoic acid and *Gm5127* predicted to be involved in the regulation of Rho protein. 2 out of the top 4 candidates genes expressed in the double positive *Tbx1* and *Tbx5* population were tested in situ hybridization technic: *Chn2* and *Cldn11*. Initial in-situ hybridization results did not validate the single-cell analysis predictions. Probe redesign will need to be tested again as already published single-cell analysis and in situ hybridization experiments such as those in The Mouse Organogenesis Spatiotemporal Transcriptomic Atlas (MOSTA) (A. Chen et al., 2022), a database containing 53 sagittal sections of wild-type mouse embryos from E9.5 to E16.5 spatially mapping gene expression, show that these genes are located in the correct spatial embryo coordinates and thus merit further investigation.

Moreover, it would be interesting to track and follow the contributions of the cells that belong to cluster 3 but are Tbx1+. Do they contribute to the arterial or venous pole? Previous studies show that posterior SHF genes (*Hoxb1, Hoxa1* and *Hoxa3*) contribute to both the arterial an venous poles of the heart, descendants of cells expressing *Hoxb1* are found in the OFT and the atria , while *Hoxa1* and *Hoxa3* descendants are found in the distal OFT and some regions of the atria (Bertrand et al., 2011).The genetic lineage tracing experiments of *Hox* genes align with the single-cell morphometric analysis where cells spatially located in the posterior SHF are not only labelled with cluster 3, predicted to contribute to the venous pole. Additionally, cells located laterally are labelled as cluster 1 and 2, which are predicted to contribute to the arterial pole. Exploring this lateral-posterior region, negative for *Tbx5*, yet acquiring Tbx5-dependent patterns of epithelial tension implying indirect roles for *Tbx5*, will provide insights into how mechanical and genetic cues interact to define the fate and direction of deployment of SHF cells.

The distinct gene expression patterns, different mechanical properties, distinct fates of cells on each side of the boundary, as well as the possibility of a boundary maintenance mechanism such as the expression of boundary-specific genes or the presence of a physical barrier boundary and cell lineage and the signaling molecule/morphogen restriction suggest the idea that the *Tbx1-Tbx5* border may function as a compartment boundary during cardiac development (Fig. 4.4).

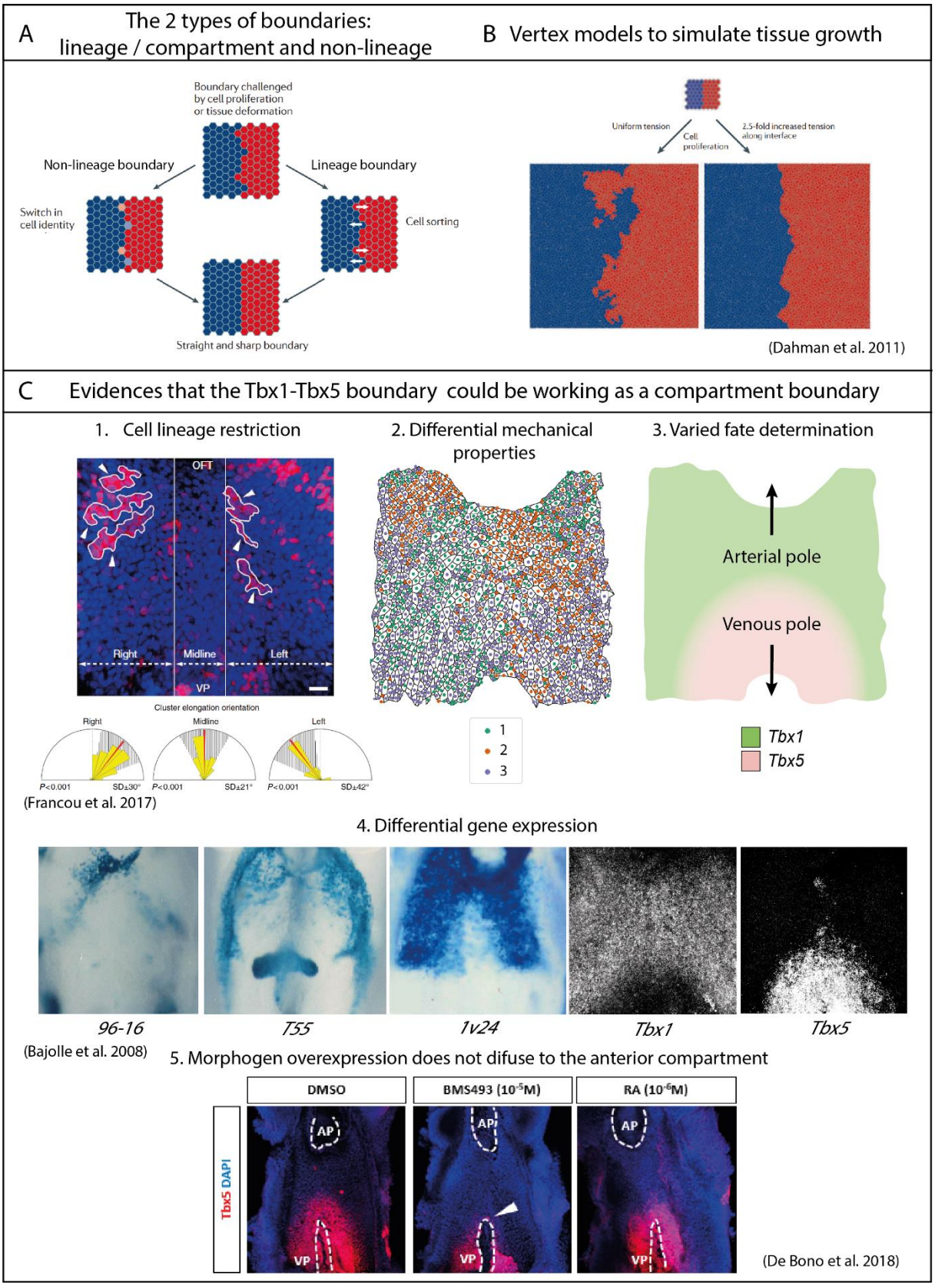


Figure 4.4: Evidence that *Tbx1-Tbx5* boundary is a compartment boundary. A. The lineage and non-lineage boundary **B.** Example of a vertex model to simulate the changes in boundary after proliferation **C.** Evidences that the *Tbx1-Tbx5* boundary may serve as a

compartment boundary: 1) Cell linage restriction where *Mesp1-Cre;Rosa-Confetti* cell clusters in the DPW at E9.5, showing isolated RFP-positive cell clusters highlighted with white arrows, are oriented towards the OFT from Francou et al. 2017. Bellow in yellow Distribution plots of the circularity index and maximal length of the cell clusters. 2) Single-cell morphometric map 3) Varied fate determination $Tbx1^{+}$ cells mapping the arterial pole and $Tbx5^{+}$ cells mapping the venous pole. 4) Differential gene expression of some SHF genes: reporter gene expression of the y96-Myf5-nlacZ-16 (*96-16)* transgene making a subpopulation of myocardial cells in the second heart field, *A17-Myf5-nlacZ-T55 (T55),* a transgene at the *Myf5* locus and *Mlc1v-24-nlacZ* transgenic line (*1v24)* transgene, which reflect the expression of Fgf10. Images showing gene expression by X-gal labelling from Kelly et al. 2001, Bajolle et al. 2008, and RNAscope of *Tbx1*, *Tbx5.* 5) Morphogen over expression does not difuse to anterior compartment, under overexpression of RA Tbx5 expression (in red) does not expand to the anterior compartment, cell nuclei are stained in blue.

4.3 Discussion on the intercalation and rosette formation

Interestingly, parallel to the activation of *Tbx5* and the establishment of tension in the SHF, we found that multicellular rosettes are present in the SHF. The appearance of rosettes depends on the chosen cutoff for bond length, a higher cutoff result in a greater number of rosettes. Rosettes are indicative of a cell mechanism related to cell intercalation during elongation of epithelial and mesenchymal tissues such us convergent extension. However, we have not shown which pathways or genes are involved in this process. No significant difference in the number of rosettes have been shown among any of the mutants analyzed here. Knowing that cell intercalation involves a huge variety of behaviors including differential adhesion, cell-shape changes or protein polarization. It would be interesting to investigate which of these mechanisms and proteins are upstream of the intercalation process (Blankenship et al., 2006; Bertet et al., 2004).

Planar cell polarity (PCP) is known to play a role in the correct formation of the heart, impairment of the PCP pathway causes a spectrum of cardiac outflow tract defects (Ramsbottom et al., 2014). It has been shown that in the undifferentiated cells from the SHF move into the outflow tract where SHF cells become more polarized, turn off progenitor markers and differentiate into cardiomyocytes. This zone receive the name of transition zone and it is marked by Vangl2. Absence of Vangl2 produces an abnormal polarized and disorganized SHF transition zone that traduce in a shortened and thickened outflow tract, similarly with the hypoplastic phenotype seen in *Tbx1* mutants (Ramsbottom et al., 2014). A more recent study showed that PCP signaling polarize the cells of the second heart field to promote deployment of the arterial and also the venous pole (Li et al., 2019). In *Wnt5a* mutant embryos, the *Wnta* lineage fails to extend into the arterial and venous pole leading to both outflow tract and atrial septal defects. Oriented actomyosin cables in the medial SHF is highlighted as the potential mechanism controlling

the length-width proportions of the SHF, promote lengthening and restricting the widening, which in other words is the process of convergent extension. In another study, *Wnt5a* has been shown to regulate cell intercalation in the SHF and generating the pushing force involved in the deployment of the SHF into the OFT. Mutant mice of *Wnt5a* show similar phenotypes to that of *Tbx1* mutant including outflow tract malformations (Sinha et al., 2015). In addition, *Wnt5a* is a transcriptional target of Tbx1. However, rosette formation and intercalation process does not appear to be affected directly by the T-box transcription factors. It would be interesting to quantify the process of rosette formation in Wnta mutant and other effectors of the pathway to see if both processes are related and dependent. However, to investigate this process it would be highly necessary to count on the development of time-lapse imaging experiments that enable to follow-up and track the rosette formation and their resolution.

In summary, this thesis demonstrates that applying single-cell morphometrics is a powerful method for uncovering cell patterns. Although the current methodology and pipeline are arduous and time-consuming, the development of more powerful algorithms will likely reduce the process, potentially enabling direct insights from the images and reducing the steps involved in data collection and preparation.

Returning to the initial question about the cellular changes involved in the segregation of SHF cells into the anterior and venous poles of the heart. We have shown that *Tbx5* have both an autonomous and a non-autonomous role in the pSHF cells. It alters the morphology of the SHF, creating two colliding stresses; one tangential to the outflow tract with lower stress at the arterial pole and another oriented parallel to the anterior-posterior axis at the venous pole. The change of orientation and tension in the pSHF, after the dorsal mesocardium breaks down, is vital for cell addition at the venous pole and drawing the epithelium in opposing directions. Its absence disturbs heart tube elongation and lead to the development of congenital heart defects as seen in *Nkx2.5* and conditional *Tbx5* mutants. Furthermore, the collision of both stresses correlates with the frontier with *Tbx1* and *Tbx5* expression where we hypothesize is important for future atrial structures. However, the principal limitations of this project are the absence of live imaging and the lack of consideration for the underlying endoderm and extracellular matrix, that could also play a role in heart tube elongation.

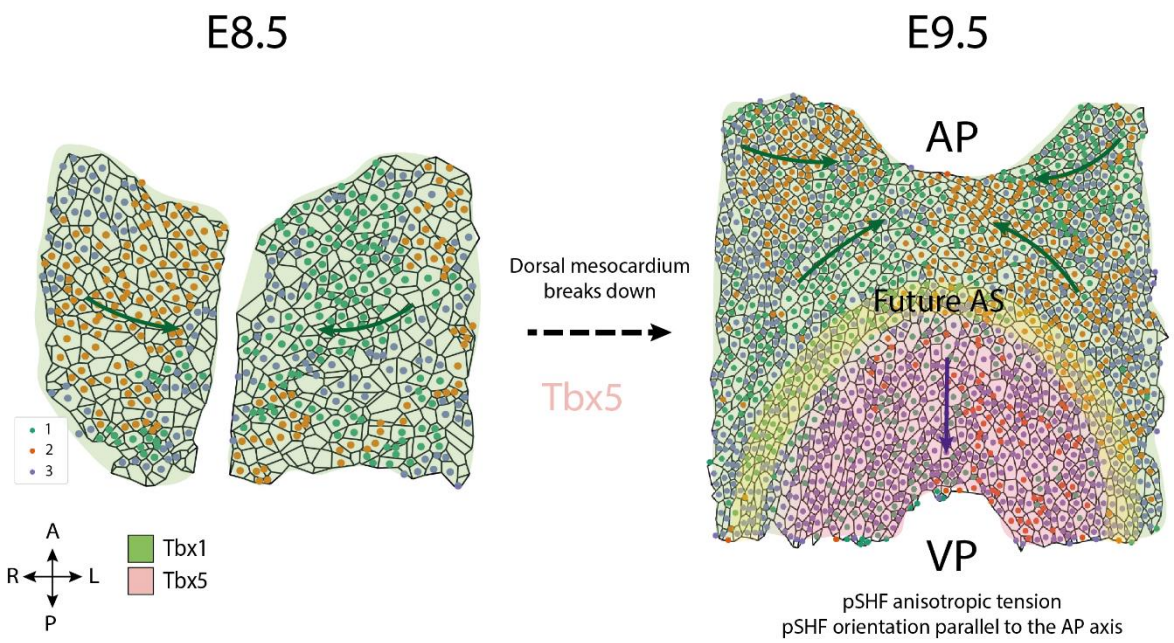


Figure 4.5: Changes in the apical morphology of SHF cells that contribute to the deployment of the cells to both poles and consequently the proper elongation of the heart tube.

Bibliography

Aguilera-Castrejon, A., Oldak, B., Shani, T., Ghanem, N., Itzkovich, C., Slomovich, S., Tarazi, S., Bayerl, J., Chugaeva, V., Ayyash, M., Ashouokhi, S., Sheban, D., Livnat, N., Lasman, L., Viukov, S., Zerbib, M., Addadi, Y., Rais, Y., Cheng, S., … Hanna, J. H. (2021). Ex utero mouse embryogenesis from pre-gastrulation to late organogenesis. *Nature*, *593*(7857), 119–124. https://doi.org/10.1038/s41586-021-03416-3

Aigouy, B., Cortes, C., Liu, S., & Prud'Homme, B. (2020). EPySeg: A coding-free solution for automated segmentation of epithelia using deep learning. *Development (Cambridge, England)*, *147*(24), dev194589. https://doi.org/10.1242/dev.194589

Aigouy, B., Farhadifar, R., Staple, D. B., Sagner, A., Röper, J.-C., Jülicher, F., & Eaton, S. (2010). Cell Flow Reorients the Axis of Planar Polarity in the Wing Epithelium of Drosophila. *Cell*, *142*(5), 773–786. https://doi.org/10.1016/j.cell.2010.07.042

Aigouy, B., Umetsu, D., & Eaton, S. (2016). Segmentation and Quantitative Analysis of Epithelial Tissues. *Methods in Molecular Biology (Clifton, N.J.)*, *1478*, 227–239. https://doi.org/10.1007/978-1-4939-6371-3_13

Aliee, M., Röper, J.-C., Landsberg, K. P., Pentzold, C., Widmann, T. J., Jülicher, F., & Dahmann, C. (2012). Physical mechanisms shaping the Drosophila dorsoventral compartment boundary. *Current Biology: CB*, *22*(11), 967–976. https://doi.org/10.1016/j.cub.2012.03.070

Altman, N., & Krzywinski, M. (2018). The curse(s) of dimensionality. *Nature Methods*, *15*(6), 399–400. https://doi.org/10.1038/s41592-018-0019-x

Anderson, R. H., Spicer, D. E., Brown, N. A., & Mohun, T. J. (2014). The development of septation in the four-chambered heart. *Anatomical Record (Hoboken, N.J.: 2007)*, *297*(8), 1414–1429. https://doi.org/10.1002/ar.22949

Andrews, T. G. R., Pönisch, W., Paluch, E. K., Steventon, B. J., & Benito-Gutierrez, E. (2021). Single-cell morphometrics reveals ancestral principles of notochord development. *Development*, *148*(16). https://doi.org/10.1242/dev.199430

Ankerst, M., Breunig, M., Kröger, P., & Sander, J. (1999). *OPTICS: Ordering Points to Identify the Clustering Structure*. *28*, 49–60. https://doi.org/10.1145/304182.304187

Aragona, M., Panciera, T., Manfrin, A., Giulitti, S., Michielin, F., Elvassore, N., Dupont, S., & Piccolo, S. (2013). A mechanical checkpoint controls multicellular growth through YAP/TAZ regulation by actin-processing factors. *Cell*, *154*(5), 1047–1059. https://doi.org/10.1016/j.cell.2013.07.042

Arriagada, C., & Astrof, S. (2023). *The mesodermal source of fibronectin is required for heart morphogenesis and cardiac outflow tract elongation by regulating cell shape, polarity, and mechanotransduction in the second heart field* (p. 2022.10.28.514299). bioRxiv. https://doi.org/10.1101/2022.10.28.514299

Bajolle, F., Zaffran, S., Kelly, R. G., Hadchouel, J., Bonnet, D., Brown, N. A., & Buckingham, M. E. (2006). Rotation of the myocardial wall of the outflow tract is implicated in the normal positioning of the great arteries. *Circulation Research*, *98*(3), 421–428. https://doi.org/10.1161/01.RES.0000202800.85341.6e

Bajolle, F., Zaffran, S., Meilhac, S. M., Dandonneau, M., Chang, T., Kelly, R. G., & Buckingham, M. E. (2008). Myocardium at the base of the aorta and pulmonary trunk is prefigured in the outflow tract of the heart and in subdomains of the second heart field. *Developmental Biology*, *313*(1), 25–34. https://doi.org/10.1016/j.ydbio.2007.09.023

Baldini, A., Fulcoli, F. G., & Illingworth, E. (2017). Tbx1: Transcriptional and Developmental Functions. *Current Topics in Developmental Biology*, *122*, 223–243. https://doi.org/10.1016/bs.ctdb.2016.08.002

Barrios, A., Poole, R. J., Durbin, L., Brennan, C., Holder, N., & Wilson, S. W. (2003). Eph/Ephrin Signaling Regulates the Mesenchymal-to-Epithelial Transition of the Paraxial Mesoderm during Somite Morphogenesis. *Current Biology*, *13*(18), 1571–1582. https://doi.org/10.1016/j.cub.2003.08.030

Becht, E., McInnes, L., Healy, J., Dutertre, C.-A., Kwok, I. W. H., Ng, L. G., Ginhoux, F., & Newell, E. W. (2019). Dimensionality reduction for visualizing single-cell data using UMAP. *Nature Biotechnology*, *37*(1), Article 1. https://doi.org/10.1038/nbt.4314

Behrndt, M., Salbreux, G., Campinho, P., Hauschild, R., Oswald, F., Roensch, J., Grill, S. W., & Heisenberg, C.-P. (2012). Forces driving epithelial spreading in zebrafish gastrulation. *Science (New York, N.Y.)*, *338*(6104), 257–260. https://doi.org/10.1126/science.1224143

Bertet, C., Sulak, L., & Lecuit, T. (2004). Myosin-dependent junction remodelling controls planar cell intercalation and axis elongation. *Nature*, *429*(6992), Article 6992. https://doi.org/10.1038/nature02590

Bertrand, N., Roux, M., Ryckebüsch, L., Niederreither, K., Dollé, P., Moon, A., Capecchi, M., & Zaffran, S. (2011). Hox genes define distinct progenitor sub-domains within the second heart field. *Developmental Biology*, *353*(2), 266–274. https://doi.org/10.1016/j.ydbio.2011.02.029

Blankenship, J. T., Backovic, S. T., Sanny, J. S. P., Weitz, O., & Zallen, J. A. (2006). Multicellular rosette formation links planar cell polarity to tissue morphogenesis.

Developmental Cell, *11*(4), 459–470. https://doi.org/10.1016/j.devcel.2006.09.007

Bookstein, F. L. (1986). Size and Shape Spaces for Landmark Data in Two Dimensions. *Statistical Science*, *1*(2), 181–222. https://doi.org/10.1214/ss/1177013696

Bookstein, F. L. (1989). Principal warps: Thin-plate splines and the decomposition of deformations. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*(6), 567–585. https://doi.org/10.1109/34.24792

Bookstein, F. L. (1992). *Morphometric Tools for Landmark Data: Geometry and Biology*. Cambridge University Press. https://doi.org/10.1017/CBO9780511573064

Bouma, B. J., & Mulder, B. J. M. (2017). Changing Landscape of Congenital Heart Disease. *Circulation Research*, *120*(6), 908–922. https://doi.org/10.1161/CIRCRESAHA.116.309302

Bro, R., & K. Smilde, A. (2014). Principal component analysis. *Analytical Methods*, *6*(9), 2812–2831. https://doi.org/10.1039/C3AY41907J

Bruneau, B. G., Logan, M., Davis, N., Levi, T., Tabin, C. J., Seidman, J. G., & Seidman, C. E. (1999). Chamber-specific cardiac expression of Tbx5 and heart defects in Holt-Oram syndrome. *Developmental Biology*, *211*(1), 100–108. https://doi.org/10.1006/dbio.1999.9298

Bruneau, B. G., Nemer, G., Schmitt, J. P., Charron, F., Robitaille, L., Caron, S., Conner, D. A., Gessler, M., Nemer, M., Seidman, C. E., & Seidman, J. G. (2001). A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease. *Cell*, *106*(6), 709–721. https://doi.org/10.1016/s0092-8674(01)00493-7

Cai, C.-L., Liang, X., Shi, Y., Chu, P.-H., Pfaff, S. L., Chen, J., & Evans, S. (2003). Isl1 Identifies a Cardiac Progenitor Population that Proliferates Prior to Differentiation and

Contributes a Majority of Cells to the Heart. *Developmental Cell*, *5*(6), 877–889. https://doi.org/10.1016/S1534-5807(03)00363-0

Caloca, M. J., Fernandez, N., Lewin, N. E., Ching, D., Modali, R., Blumberg, P. M., & Kazanietz, M. G. (1997). B2-Chimaerin Is a High Affinity Receptor for the Phorbol Ester Tumor Promoters *. *Journal of Biological Chemistry*, *272*(42), 26488–26496. https://doi.org/10.1074/jbc.272.42.26488

Campinho, P., Behrndt, M., Ranft, J., Risler, T., Minc, N., & Heisenberg, C.-P. (2013). Tension-oriented cell divisions limit anisotropic tissue tension in epithelial spreading during zebrafish epiboly. *Nature Cell Biology*, *15*(12), 1405–1414. https://doi.org/10.1038/ncb2869

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology*, *38*(4), 476–506. https://doi.org/10.1037/h0054116

Chen, A., Liao, S., Cheng, M., Ma, K., Wu, L., Lai, Y., Qiu, X., Yang, J., Xu, J., Hao, S., Wang, X., Lu, H., Chen, X., Liu, X., Huang, X., Li, Z., Hong, Y., Jiang, Y., Peng, J., … Wang, J. (2022). Spatiotemporal transcriptomic atlas of mouse organogenesis using DNA nanoball-patterned arrays. *Cell*, *185*(10), 1777-1792.e21. https://doi.org/10.1016/j.cell.2022.04.003

Chen, C. S., Mrksich, M., Huang, S., Whitesides, G. M., & Ingber, D. E. (1997). Geometric control of cell life and death. *Science (New York, N.Y.)*, *276*(5317), 1425–1428. https://doi.org/10.1126/science.276.5317.1425

Chen, L., Fulcoli, F. G., Ferrentino, R., Martucciello, S., Illingworth, E. A., & Baldini, A. (2012). Transcriptional control in cardiac progenitors: Tbx1 interacts with the BAF chromatin remodeling complex and regulates Wnt5a. *PLoS Genetics*, *8*(3), e1002571. https://doi.org/10.1371/journal.pgen.1002571

Chen, L., Fulcoli, F. G., Tang, S., & Baldini, A. (2009). Tbx1 Regulates Proliferation and Differentiation of Multipotent Heart Progenitors. *Circulation Research*, *105*(9), 842–851. https://doi.org/10.1161/CIRCRESAHA.109.200295

Chen, T., & Guestrin, C. (2016). XGBoost: A Scalable Tree Boosting System. *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 785–794. https://doi.org/10.1145/2939672.2939785

Chen, Y., Yang, Z., Zhao, Z.-A., & Shen, Z. (2017). Direct reprogramming of fibroblasts into cardiomyocytes. *Stem Cell Research & Therapy*, *8*, 118. https://doi.org/10.1186/s13287-017-0569-3

Chiu, C.-L., Clack, N., & the napari community. (2022). napari: A Python Multi-Dimensional Image Viewer Platform for the Research Community. *Microscopy and Microanalysis*, *28*(S1), 1576–1577. https://doi.org/10.1017/S1431927622006328

Christoffels, V. M., Habets, P. E., Franco, D., Campione, M., de Jong, F., Lamers, W. H., Bao, Z. Z., Palmer, S., Biben, C., Harvey, R. P., & Moorman, A. F. (2000). Chamber formation and morphogenesis in the developing mammalian heart. *Developmental Biology*, *223*(2), 266–278. https://doi.org/10.1006/dbio.2000.9753

Christoffels, V. M., Mommersteeg, M. T. M., Trowe, M.-O., Prall, O. W. J., de Gier-de Vries, C., Soufan, A. T., Bussen, M., Schuster-Gossler, K., Harvey, R. P., Moorman, A. F. M., & Kispert, A. (2006). Formation of the venous pole of the heart from an Nkx2-5-negative precursor population requires Tbx18. *Circulation Research*, *98*(12), 1555–1563. https://doi.org/10.1161/01.RES.0000227571.84189.65

Cortes, C., Francou, A., De Bono, C., & Kelly, R. G. (2018). Epithelial Properties of the Second Heart Field. *Circulation Research*, *122*(1), 142–154. https://doi.org/10.1161/CIRCRESAHA.117.310838

Crameri, F., Shephard, G. E., & Heron, P. J. (2020). The misuse of colour in science communication. *Nature Communications*, *11*(1), Article 1. https://doi.org/10.1038/s41467-020-19160-7

de Boer, B. A., van den Berg, G., Soufan, A. T., de Boer, P. A. J., Hagoort, J., van den Hoff, M. J. B., Moorman, A. F. M., & Ruijter, J. M. (2012). Measurement and 3D-Visualization of Cell-Cycle Length Using Double Labelling with Two Thymidine Analogues Applied in Early Heart Development. *PLoS ONE*, *7*(10), e47719. https://doi.org/10.1371/journal.pone.0047719

De Bono, C., Thellier, C., Bertrand, N., Sturny, R., Jullian, E., Cortes, C., Stefanovic, S., Zaffran, S., Théveniau-Ruissy, M., & Kelly, R. G. (2018). T-box genes and retinoic acid signaling regulate the segregation of arterial and venous pole progenitor cells in the murine second heart field. *Human Molecular Genetics*, *27*(21), 3747–3760. https://doi.org/10.1093/hmg/ddy266

de la Cruz, M. V., Sánchez Gómez, C., Arteaga, M. M., & Argüello, C. (1977). Experimental study of the development of the truncus and the conus in the chick embryo. *Journal of Anatomy*, *123*(Pt 3), 661–686.

de Soysa, T. Y., Ranade, S. S., Okawa, S., Ravichandran, S., Huang, Y., Salunga, H. T., Schricker, A., Del Sol, A., Gifford, C. A., & Srivastava, D. (2019). Single-cell analysis of cardiogenesis reveals basis for organ level developmental defects. *Nature*, *572*(7767), 120–124. https://doi.org/10.1038/s41586-019-1414-x

Deepe, R., Fitzgerald, E., Wolters, R., Drummond, J., Guzman, K. D., Hoff, M. J. B. van den, & Wessels, A. (2020). The Mesenchymal Cap of the Atrial Septum and Atrial and Atrioventricular Septation. *Journal of Cardiovascular Development and Disease*, *7*(4), Article 4. https://doi.org/10.3390/jcdd7040050

Deng, D. (2020). DBSCAN Clustering Algorithm Based on Density. *2020 7th International Forum on Electrical Engineering and Automation (IFEEA)*, 949–953. https://doi.org/10.1109/IFEEA51475.2020.00199

Denninger, A. R., Breglio, A., Maheras, K. J., LeDuc, G., Cristiglio, V., Demé, B., Gow, A., & Kirschner, D. A. (2015). Claudin-11 Tight Junctions in Myelin Are a Barrier to Diffusion and Lack Strong Adhesive Properties. *Biophysical Journal*, *109*(7), 1387–1397. https://doi.org/10.1016/j.bpj.2015.08.012

Devaux, J., & Gow, A. (2008). Tight junctions potentiate the insulative properties of small CNS myelinated axons. *The Journal of Cell Biology*, *183*(5), 909–921. https://doi.org/10.1083/jcb.200808034

Devine, W. P., Wythe, J. D., George, M., Koshiba-Takeuchi, K., & Bruneau, B. G. (2014). Early patterning and specification of cardiac progenitors in gastrulating mesoderm. *eLife*, *3*, e03848. https://doi.org/10.7554/eLife.03848

Diogo, R., Kelly, R. G., Christiaen, L., Levine, M., Ziermann, J. M., Molnar, J. L., Noden, D. M., & Tzahor, E. (2015). A new heart for a new head in vertebrate cardiopharyngeal evolution. *Nature*, *520*(7548), Article 7548. https://doi.org/10.1038/nature14435

Discher, D. E., Mooney, D. J., & Zandstra, P. W. (2009). Growth factors, matrices, and forces combine and control stem cells. *Science (New York, N.Y.)*, *324*(5935), 1673–1677. https://doi.org/10.1126/science.1171643

Dominguez, M. H., Krup, A. L., Muncie, J. M., & Bruneau, B. G. (2023). Graded mesoderm assembly governs cell fate and morphogenesis of the early mammalian heart. *Cell*, *186*(3), 479-496.e23. https://doi.org/10.1016/j.cell.2023.01.001

Driscoll, D. J. (1999). LEFT-TO-RIGHT SHUNT LESIONS. *Pediatric Clinics of North America*, *46*(2), 355–368. https://doi.org/10.1016/S0031-3955(05)70123-1

Driscoll, M. K., & Zaritsky, A. (2021). Data science in cell imaging. *Journal of Cell Science*, *134*(7), jcs254292. https://doi.org/10.1242/jcs.254292

Driver, H. E., & Kroeber, A. L. (1932). *Quantitative expression of cultural relationships, by H.E. Driver and A.L. Kroeber*. University of California Press. http://catalogue.bnf.fr/ark:/12148/cb320450716

Duchemin, A.-L., Vignes, H., & Vermot, J. (n.d.). Mechanically activated piezo channels modulate outflow tract valve development through the Yap1 and Klf2-Notch signaling axis. *eLife*, *8*, e44706. https://doi.org/10.7554/eLife.44706

Dupont, S., Morsut, L., Aragona, M., Enzo, E., Giulitti, S., Cordenonsi, M., Zanconato, F., Le Digabel, J., Forcato, M., Bicciato, S., Elvassore, N., & Piccolo, S. (2011). Role of YAP/TAZ in mechanotransduction. *Nature*, *474*(7350), Article 7350. https://doi.org/10.1038/nature10137

Dye, N. A., Popović, M., Iyer, K. V., Fuhrmann, J. F., Piscitello-Gómez, R., Eaton, S., & Jülicher, F. (2021). Self-organized patterning of cell morphology via mechanosensitive feedback. *eLife*, *10*, e57964. https://doi.org/10.7554/eLife.57964

Ehler, E. (2016). Cardiac cytoarchitecture—Why the "hardware" is important for heart function! *Biochimica et Biophysica Acta*, *1863*(7Part B), 1857–1863. https://doi.org/10.1016/j.bbamcr.2015.11.006

Engler, A. J., Sen, S., Sweeney, H. L., & Discher, D. E. (2006). Matrix Elasticity Directs Stem Cell Lineage Specification. *Cell*, *126*(4), 677–689. https://doi.org/10.1016/j.cell.2006.06.044

Esteban, I., Schmidt, P., Desgrange, A., Raiola, M., Temiño, S., Meilhac, S. M., Kobbelt, L., & Torres, M. (2022). Pseudodynamic analysis of heart tube formation in the mouse reveals strong regional variability and early left–right asymmetry. *Nature*

Cardiovascular Research, *1*(5), Article 5. https://doi.org/10.1038/s44161-022-00065-1

Estivill-Castro, V. (2002). Why so many clustering algorithms: A position paper. *ACM SIGKDD Explorations Newsletter*, *4*(1), 65–75. https://doi.org/10.1145/568574.568575

Eyckmans, J., Boudou, T., Yu, X., & Chen, C. S. (2011). A hitchhiker's guide to mechanobiology. *Developmental Cell*, *21*(1), 35–47. https://doi.org/10.1016/j.devcel.2011.06.015

Fagotto, F. (2020). Tissue segregation in the early vertebrate embryo. *Seminars in Cell & Developmental Biology*, *107*, 130–146. https://doi.org/10.1016/j.semcdb.2020.05.020

Farrell, D., Weitz, O., Magnasco, M., & Zallen, J. (2017). SEGGA: A toolset for rapid automated analysis of epithelial cell polarity and dynamics. *Development*, *144*, 1725–1734. https://doi.org/10.1242/dev.146837

Faucherre, A., Moha Ou Maati, H., Nasr, N., Pinard, A., Theron, A., Odelin, G., Desvignes, J.-P., Salgado, D., Collod-Béroud, G., Avierinos, J.-F., Lebon, G., Zaffran, S., & Jopling, C. (2020). Piezo1 is required for outflow tract and aortic valve development. *Journal of Molecular and Cellular Cardiology*, *143*, 51–62. https://doi.org/10.1016/j.yjmcc.2020.03.013

Fink, J., Carpi, N., Betz, T., Bétard, A., Chebah, M., Azioune, A., Bornens, M., Sykes, C., Fetler, L., Cuvelier, D., & Piel, M. (2011). External forces control mitotic spindle positioning. *Nature Cell Biology*, *13*(7), Article 7. https://doi.org/10.1038/ncb2269

Franco, D., Meilhac, S. M., Christoffels, V. M., Kispert, A., Buckingham, M., & Kelly, R. G. (2006). Left and right ventricular contributions to the formation of the

interventricular septum in the mouse heart. *Developmental Biology*, *294*(2), 366–375. https://doi.org/10.1016/j.ydbio.2006.02.045

Francou, A., Anderson, K. V., & Hadjantonakis, A.-K. (2023). A ratchet-like apical constriction drives cell ingression during the mouse gastrulation EMT. *eLife*, *12*, e84019. https://doi.org/10.7554/eLife.84019

Francou, A., De Bono, C., & Kelly, R. G. (2017). Epithelial tension in the second heart field promotes mouse heart tube elongation. *Nature Communications*, *8*, 14770. https://doi.org/10.1038/ncomms14770

Francou, A., Saint-Michel, E., Mesbah, K., & Kelly, R. G. (2014). TBX1 regulates epithelial polarity and dynamic basal filopodia in the second heart field. *Development*, *141*(22), 4320–4331. https://doi.org/10.1242/dev.115022

Galli, D., Domínguez, J. N., Zaffran, S., Munk, A., Brown, N. A., & Buckingham, M. E. (2008). Atrial myocardium derives from the posterior region of the second heart field, which acquires left-right identity as Pitx2c is expressed. *Development (Cambridge, England)*, *135*(6), 1157–1167. https://doi.org/10.1242/dev.014563

Garavito-Aguilar, Z. V., Riley, H. E., & Yelon, D. (2010). Hand2 ensures an appropriate environment for cardiac fusion by limiting Fibronectin function. *Development (Cambridge, England)*, *137*(19), 3215–3220. https://doi.org/10.1242/dev.052225

Garcia-Bellido, A., & Santamaria, P. (1972). Developmental Analysis of the Wing Disc in the Mutant Engrailed of DROSOPHILA MELANOGASTER. *Genetics*, *72*(1), 87–104.

Glinianaia, S. V., Morris, J. K., Best, K. E., Santoro, M., Coi, A., Armaroli, A., & Rankin, J. (2020). Long-term survival of children born with congenital anomalies: A systematic review and meta-analysis of population-based studies. *PLOS Medicine*, *17*(9), e1003356. https://doi.org/10.1371/journal.pmed.1003356

Goodrich, L. V., & Strutt, D. (2011). Principles of planar polarity in animal development. *Development (Cambridge, England)*, *138*(10), 1877–1892. https://doi.org/10.1242/dev.054080

Goor, D. A., Edwards, J. E., & Lillehei, C. W. (1970). The Development of the Interventricular Septum of the Human Heart; Correlative Morphogenetic Study. *Chest*, *58*(5), 453–467. https://doi.org/10.1378/chest.58.5.453

Gow, A., Davies, C., Southwood, C. M., Frolenkov, G., Chrustowski, M., Ng, L., Yamauchi, D., Marcus, D. C., & Kachar, B. (2004). Deafness in Claudin 11-Null Mice Reveals the Critical Contribution of Basal Cell Tight Junctions to Stria Vascularis Function. *The Journal of Neuroscience*, *24*(32), 7051–7062. https://doi.org/10.1523/JNEUROSCI.1640-04.2004

Gray, R. S., Roszko, I., & Solnica-Krezel, L. (2011). Planar cell polarity: Coordinating morphogenetic cell behaviors with embryonic polarity. *Developmental Cell*, *21*(1), 120–133. https://doi.org/10.1016/j.devcel.2011.06.011

Haghverdi, L., Büttner, M., Wolf, F. A., Buettner, F., & Theis, F. J. (2016). Diffusion pseudotime robustly reconstructs lineage branching. *Nature Methods*, *13*(10), 845–848. https://doi.org/10.1038/nmeth.3971

Haigo, S. L., Hildebrand, J. D., Harland, R. M., & Wallingford, J. B. (2003). Shroom induces apical constriction and is required for hingepoint formation during neural tube closure. *Current Biology: CB*, *13*(24), 2125–2137. https://doi.org/10.1016/j.cub.2003.11.054

Hallou, A., He, R., Simons, B. D., & Dumitrascu, B. (2023). *A computational pipeline for spatial mechano-transcriptomics* [Preprint]. Genomics. https://doi.org/10.1101/2023.08.03.551894

Harrower, M., & Brewer, C. A. (2003). ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps. *The Cartographic Journal*, *40*(1), 27–37. https://doi.org/10.1179/000870403235002042

He, L., Wang, X., Tang, H. L., & Montell, D. J. (2010). Tissue elongation requires oscillating contractions of a basal actomyosin network. *Nature Cell Biology*, *12*(12), 1133–1142. https://doi.org/10.1038/ncb2124

Heisenberg, C.-P., & Bellaïche, Y. (2013). Forces in Tissue Morphogenesis and Patterning. *Cell*, *153*(5), 948–962. https://doi.org/10.1016/j.cell.2013.05.008

Herbert, S., Valon, L., Mancini, L., Dray, N., Caldarelli, P., Gros, J., Esposito, E., Shorte, S. L., Bally-Cuif, L., Aulner, N., Levayer, R., & Tinevez, J.-Y. (2021). LocalZProjector and DeProj: A toolbox for local 2D projection and accurate morphometrics of large 3D microscopy images. *BMC Biology*, *19*(1), 136. https://doi.org/10.1186/s12915-021-01037-w

Hildebrand, J. D., & Soriano, P. (1999). Shroom, a PDZ domain-containing actin-binding protein, is required for neural tube morphogenesis in mice. *Cell*, *99*(5), 485–497. https://doi.org/10.1016/s0092-8674(00)81537-8

Hutson, M. R., Zeng, X. L., Kim, A. J., Antoon, E., Harward, S., & Kirby, M. L. (2010). Arterial pole progenitors interpret opposing FGF/BMP signals to proliferate or differentiate. *Development (Cambridge, England)*, *137*(18), 3001–3011. https://doi.org/10.1242/dev.051565

Huynh, T., Chen, L., Terrell, P., & Baldini, A. (2007). A fate map of Tbx1 expressing cells reveals heterogeneity in the second cardiac field. *Genesis (New York, N.Y.: 2000)*, *45*(7), 470–475. https://doi.org/10.1002/dvg.20317

Ichbiah, S., Delbary, F., McDougall, A., Dumollard, R., & Turlier, H. (2023). Embryo mechanics cartography: Inference of 3D force atlases from fluorescence

microscopy. *Nature Methods*, *20*(12), 1989–1999. https://doi.org/10.1038/s41592-023-02084-7

Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhaija, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences*, *622*, 178–210. https://doi.org/10.1016/j.ins.2022.11.139

Ishihara, S., & Sugimura, K. (2012). Bayesian inference of force dynamics during morphogenesis. *Journal of Theoretical Biology*, *313*, 201–211. https://doi.org/10.1016/j.jtbi.2012.08.017

Izquierdo, E., Quinkler, T., & De Renzis, S. (2018). Guided morphogenesis through optogenetic activation of Rho signalling during early Drosophila embryogenesis. *Nature Communications*, *9*(1), 2366. https://doi.org/10.1038/s41467-018-04754-z

Jaalouk, D. E., & Lammerding, J. (2009). Mechanotransduction gone awry. *Nature Reviews Molecular Cell Biology*, *10*(1), Article 1. https://doi.org/10.1038/nrm2597

Jerome, L. A., & Papaioannou, V. E. (2001). DiGeorge syndrome phenotype in mice mutant for the T-box gene, Tbx1. *Nature Genetics*, *27*(3), 286–291. https://doi.org/10.1038/85845

Jolliffe, I. T., & Cadima, J. (2016). Principal component analysis: A review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *374*(2065), 20150202. https://doi.org/10.1098/rsta.2015.0202

Jombart, T., Devillard, S., & Balloux, F. (2010). Discriminant analysis of principal components: A new method for the analysis of genetically structured populations. *BMC Genetics*, *11*(1), 94. https://doi.org/10.1186/1471-2156-11-94

Kathiriya, I. S., Rao, K. S., Iacono, G., Devine, W. P., Blair, A. P., Hota, S. K., Lai, M. H., Garay, B. I., Thomas, R., Gong, H. Z., Wasson, L. K., Goyal, P., Sukonnik, T., Hu, K. M., Akgun, G. A., Bernard, L. D., Akerberg, B. N., Gu, F., Li, K., … Bruneau, B. G. (2021). Modeling Human TBX5 Haploinsufficiency Predicts Regulatory Networks for Congenital Heart Disease. *Developmental Cell*, *56*(3), 292-309.e9. https://doi.org/10.1016/j.devcel.2020.11.020

Keller, R., Davidson, L., Edlund, A., Elul, T., Ezin, M., Shook, D., & Skoglund, P. (2000). Mechanisms of convergence and extension by cell intercalation. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *355*(1399), 897–922. https://doi.org/10.1098/rstb.2000.0626

Kelly, R. G. (2012). The second heart field. *Current Topics in Developmental Biology*, *100*, 33–65. https://doi.org/10.1016/B978-0-12-387786-4.00002-6

Kelly, R. G. (2023). The heart field transcriptional landscape at single-cell resolution. *Developmental Cell*, *58*(4), 257–266. https://doi.org/10.1016/j.devcel.2023.01.010

Kelly, R. G., Brown, N. A., & Buckingham, M. E. (2001). The Arterial Pole of the Mouse Heart Forms from Fgf10-Expressing Cells in Pharyngeal Mesoderm. *Developmental Cell*, *1*(3), 435–440. https://doi.org/10.1016/S1534-5807(01)00040-5

Kelly, R. G., Buckingham, M. E., & Moorman, A. F. (2014). Heart Fields and Cardiac Morphogenesis. *Cold Spring Harbor Perspectives in Medicine*, *4*(10), a015750. https://doi.org/10.1101/cshperspect.a015750

Kidokoro, H., Yonei-Tamura, S., Tamura, K., Schoenwolf, G. C., & Saijoh, Y. (2018). The heart tube forms and elongates through dynamic cell rearrangement coordinated with foregut extension. *Development*, *145*(7). https://doi.org/10.1242/dev.152488

Kim, J., Jeon, S., Choi, J.-P., Blazyte, A., Jeon, Y., Kim, J.-I., Ohashi, J., Tokunaga, K., Sugano, S., Fucharoen, S., Al-Mulla, F., & Bhak, J. (2020). The Origin and Composition of Korean Ethnicity Analyzed by Ancient and Present-Day Genome Sequences. *Genome Biology and Evolution*, *12*(5), 553–565. https://doi.org/10.1093/gbe/evaa062

Kirby, M. L., & Waldo, K. L. (1990). Role of neural crest in congenital heart disease. *Circulation*, *82*(2), 332–340. https://doi.org/10.1161/01.cir.82.2.332

Kong, W., Loison, O., Chavadimane Shivakumar, P., Chan, E. H., Saadaoui, M., Collinet, C., Lenne, P.-F., & Clément, R. (2019). Experimental validation of force inference in epithelia from cell to tissue scale. *Scientific Reports*, *9*(1), 14647. https://doi.org/10.1038/s41598-019-50690-3

Kornberg, T., Sidén, I., O'Farrell, P., & Simon, M. (1985). The engrailed locus of drosophila: In situ localization of transcripts reveals compartment-specific expression. *Cell*, *40*(1), 45–53. https://doi.org/10.1016/0092-8674(85)90307-1

Landsberg, K. P., Farhadifar, R., Ranft, J., Umetsu, D., Widmann, T. J., Bittig, T., Said, A., Jülicher, F., & Dahmann, C. (2009). Increased cell bond tension governs cell sorting at the Drosophila anteroposterior compartment boundary. *Current Biology: CB*, *19*(22), 1950–1955. https://doi.org/10.1016/j.cub.2009.10.021

Le A. Trinh, & Stainier, D. Y. R. (2004). Fibronectin Regulates Epithelial Organization during Myocardial Migration in Zebrafish. *Developmental Cell*, *6*(3), 371–382. https://doi.org/10.1016/S1534-5807(04)00063-2

Lelouard, H., Mailfert, S., & Fallet, M. (2018). *A ten-color spectral imaging strategy to reveal localization of gut immune cell subsets*. https://doi.org/10.13140/RG.2.2.18840.88324

Leung, T., How, B. E., Manser, E., & Lim, L. (1994). Cerebellar beta 2-chimaerin, a GTPase-activating protein for p21 ras-related rac is specifically expressed in granule cells

and has a unique N-terminal SH2 domain. *Journal of Biological Chemistry*, *269*(17), 12888–12892. https://doi.org/10.1016/S0021-9258(18)99959-X

Lewandowski, J. P., Du, F., Zhang, S., Powell, M. B., Falkenstein, K. N., Ji, H., & Vokes, S. A. (2015). Spatiotemporal regulation of GLI target genes in the mammalian limb bud. *Developmental Biology*, *406*(1), 92–103. https://doi.org/10.1016/j.ydbio.2015.07.022

Li, D., Angermeier, A., & Wang, J. (2019). Planar cell polarity signaling regulates polarized second heart field morphogenesis to promote both arterial and venous pole septation. *Development (Cambridge, England)*, *146*(20), dev181719. https://doi.org/10.1242/dev.181719

Li, D., Sinha, T., Ajima, R., Seo, H.-S., Yamaguchi, T. P., & Wang, J. (2016). Spatial regulation of cell cohesion by Wnt5a during second heart field progenitor deployment. *Developmental Biology*, *412*(1), 18–31. https://doi.org/10.1016/j.ydbio.2016.02.017

Li, Q., & Nishikawa, R. M. (2015). *Computer-Aided Detection and Diagnosis in Medical Imaging*. Taylor & Francis.

Linask, K. K., Manisastry, S., & Han, M. (2005). Cross Talk between Cell–Cell and Cell–Matrix Adhesion Signaling Pathways during Heart Organogenesis: Implications for Cardiac Birth Defects. *Microscopy and Microanalysis*, *11*(3), 200–208. https://doi.org/10.1017/S1431927605050440

Maaten, L. van der, & Hinton, G. (2008). Visualizing Data using t-SNE. *Journal of Machine Learning Research*, *9*(86), 2579–2605.

Majkut, S., Dingal, P. C. D. P., & Discher, D. E. (2014). Stress Sensitivity and Mechanotransduction during Heart Development. *Current Biology : CB*, *24*(10), R495–R501. https://doi.org/10.1016/j.cub.2014.04.027

Majkut, S., Idema, T., Swift, J., Krieger, C., Liu, A., & Discher, D. E. (2013). Heart-specific stiffening in early embryos parallels matrix and myosin expression to optimize beating. *Current Biology: CB*, *23*(23), 2434–2439. https://doi.org/10.1016/j.cub.2013.10.057

Mao, Y., & Baum, B. (2015). Tug of war—The influence of opposing physical forces on epithelial cell morphology. *Developmental Biology*, *401*(1), 92–102. https://doi.org/10.1016/j.ydbio.2014.12.030

Mao, Y., Tournier, A. L., Hoppe, A., Kester, L., Thompson, B. J., & Tapon, N. (2013). Differential proliferation rates generate patterns of mechanical tension that orient tissue growth. *The EMBO Journal*, *32*(21), 2790–2803. https://doi.org/10.1038/emboj.2013.197

Martin, A. C., Gelbart, M., Fernandez-Gonzalez, R., Kaschube, M., & Wieschaus, E. F. (2010). Integration of contractile forces during tissue invagination. *The Journal of Cell Biology*, *188*(5), 735–749. https://doi.org/10.1083/jcb.200910099

Martínez-Ara, G., Taberner, N., Takayama, M., Sandaltzopoulou, E., Villava, C. E., Bosch-Padrós, M., Takata, N., Trepat, X., Eiraku, M., & Ebisuya, M. (2022). Optogenetic control of apical constriction induces synthetic morphogenesis in mammalian tissues. *Nature Communications*, *13*(1), 5400. https://doi.org/10.1038/s41467-022-33115-0

McDole, K., Guignard, L., Amat, F., Berger, A., Malandain, G., Royer, L. A., Turaga, S. C., Branson, K., & Keller, P. J. (2018). In Toto Imaging and Reconstruction of Post-Implantation Mouse Development at the Single-Cell Level. *Cell*, *175*(3), 859-876.e33. https://doi.org/10.1016/j.cell.2018.09.031

Meilhac, S. M., Esner, M., Kelly, R. G., Nicolas, J.-F., & Buckingham, M. E. (2004). The clonal origin of myocardial cells in different regions of the embryonic mouse heart.

Developmental Cell, *6*(5), 685–698. https://doi.org/10.1016/s1534-5807(04)00133-9

Mic, F. A., Haselbeck, R. J., Cuenca, A. E., & Duester, G. (2002). Novel retinoic acid generating activities in the neural tube and heart identified by conditional rescue of Raldh2 null mutant mice. *Development (Cambridge, England)*, *129*(9), 2271–2282. https://doi.org/10.1242/dev.129.9.2271

Mjaatvedt, C. H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R. A., Kern, M. J., Eisenberg, C. A., Turner, D., & Markwald, R. R. (2001). The Outflow Tract of the Heart Is Recruited from a Novel Heart-Forming Field. *Developmental Biology*, *238*(1), 97–109. https://doi.org/10.1006/dbio.2001.0409

Momma, K. (2010). Cardiovascular anomalies associated with chromosome 22q11.2 deletion syndrome. *The American Journal of Cardiology*, *105*(11), 1617–1624. https://doi.org/10.1016/j.amjcard.2010.01.333

Mommersteeg, M. T. M., Soufan, A. T., de Lange, F. J., van den Hoff, M. J. B., Anderson, R. H., Christoffels, V. M., & Moorman, A. F. M. (2006). Two distinct pools of mesenchyme contribute to the development of the atrial septum. *Circulation Research*, *99*(4), 351–353. https://doi.org/10.1161/01.RES.0000238360.33284.a0

Monier, B., Pélissier-Monier, A., Brand, A. H., & Sanson, B. (2010). An actomyosin-based barrier inhibits cell mixing at compartmental boundaries in Drosophila embryos. *Nature Cell Biology*, *12*(1), 60–69. https://doi.org/10.1038/ncb2005

Moon, K. R., Van Dijk, D., Wang, Z., Gigante, S., Burkhardt, D. B., Chen, W. S., Yim, K., Elzen, A. V. D., Hirn, M. J., Coifman, R. R., Ivanova, N. B., Wolf, G., & Krishnaswamy, S. (2019). Visualizing structure and transitions in high-dimensional biological data. *Nature Biotechnology*, *37*(12), 1482–1492. https://doi.org/10.1038/s41587-019-0336-3

Moorman, A., Webb, S., Brown, N., Lamers, W., & Anderson, R. (2003). Development of the heart: (1) Formation of the cardiac chambers and arterial trunks. *Heart (British Cardiac Society)*, *89*, 806–814.

Mori, A. D., Zhu, Y., Vahora, I., Nieman, B., Koshiba-Takeuchi, K., Davidson, L., Pizard, A., Seidman, J. G., Seidman, C. E., Chen, X. J., Henkelman, R. M., & Bruneau, B. G. (2006). Tbx5-dependent rheostatic control of cardiac gene expression and morphogenesis. *Developmental Biology*, *297*(2), 566–586. https://doi.org/10.1016/j.ydbio.2006.05.023

Morikawa, Y., Heallen, T., Leach, J., Xiao, Y., & Martin, J. F. (2017). Dystrophin-glycoprotein complex sequesters Yap to inhibit cardiomyocyte proliferation. *Nature*, *547*(7662), 227–231. https://doi.org/10.1038/nature22979

Morimoto, J., & Ponton, F. (2021). Virtual reality in biology: Could we become virtual naturalists? *Evolution: Education and Outreach*, *14*. https://doi.org/10.1186/s12052-021-00147-x

Morphometrics: History, development methods and prospects. (2017). *Zoological Systematics*, *42*(1), Article 1. https://doi.org/10.11865/zs.201702

Murakami, M., Nakagawa, M., Olson, E. N., & Nakagawa, O. (2005). A WW domain protein TAZ is a critical coactivator for TBX5, a transcription factor implicated in Holt-Oram syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(50), 18034–18039. https://doi.org/10.1073/pnas.0509109102

Murtagh, F., & Contreras, P. (2012). Algorithms for hierarchical clustering: An overview. *WIREs Data Mining and Knowledge Discovery*, *2*(1), 86–97. https://doi.org/10.1002/widm.53

Nadeau, M., Georges, R. O., Laforest, B., Yamak, A., Lefebvre, C., Beauregard, J., Paradis, P., Bruneau, B. G., Andelfinger, G., & Nemer, M. (2010). An endocardial pathway involving Tbx5, Gata4, and Nos3 required for atrial septum formation. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(45), 19356–19361. https://doi.org/10.1073/pnas.0914888107

Nguyen, H., Desgrange, A., Ochandorena-Saa, A., Benhamo, V., Meilhac, S. M., & Zimmer, C. (2024). *Deep learning-based detection of murine congenital heart defects from µCT scans* (p. 2024.04.06.588383). bioRxiv. https://doi.org/10.1101/2024.04.06.588383

Nishimura, T., & Takeichi, M. (2008). Shroom3-mediated recruitment of Rho kinases to the apical cell junctions regulates epithelial and neuroepithelial planar remodeling. *Development (Cambridge, England)*, *135*(8), 1493–1502. https://doi.org/10.1242/dev.019646

Nomaru, H., Liu, Y., De Bono, C., Righelli, D., Cirino, A., Wang, W., Song, H., Racedo, S. E., Dantas, A. G., Zhang, L., Cai, C.-L., Angelini, C., Christiaen, L., Kelly, R. G., Baldini, A., Zheng, D., & Morrow, B. E. (2021). Single cell multi-omic analysis identifies a Tbx1-dependent multilineage primed population in murine cardiopharyngeal mesoderm. *Nature Communications*, *12*(1), Article 1. https://doi.org/10.1038/s41467-021-26966-6

Notari, M., Ventura-Rubio, A., Bedford-Guaus, S. J., Jorba, I., Mulero, L., Navajas, D., Martí, M., & Raya, Á. (2018). The local microenvironment limits the regenerative potential of the mouse neonatal heart. *Science Advances*, *4*(5), eaao5553. https://doi.org/10.1126/sciadv.aao5553

Novembre, J., Johnson, T., Bryc, K., Kutalik, Z., Boyko, A. R., Auton, A., Indap, A., King, K. S., Bergmann, S., Nelson, M. R., Stephens, M., & Bustamante, C. D. (2008). Genes mirror

geography within Europe. *Nature*, *456*(7218), 98–101. https://doi.org/10.1038/nature07331

Pachitariu, M., & Stringer, C. (2022). Cellpose 2.0: How to train your own model. *Nature Methods*, *19*(12), Article 12. https://doi.org/10.1038/s41592-022-01663-4

Parker, J., & Struhl, G. (2020). Control of Drosophila wing size by morphogen range and hormonal gating. *Proceedings of the National Academy of Sciences*, *117*(50), 31935–31944. https://doi.org/10.1073/pnas.2018196117

Parsons, J. T., Horwitz, A. R., & Schwartz, M. A. (2010). Cell adhesion: Integrating cytoskeletal dynamics and cellular tension. *Nature Reviews. Molecular Cell Biology*, *11*(9), 633–643. https://doi.org/10.1038/nrm2957

Pavan, M., Ruiz, V. F., Silva, F. A., Sobreira, T. J., Cravo, R. M., Vasconcelos, M., Marques, L. P., Mesquita, S. M., Krieger, J. E., Lopes, A. A., Oliveira, P. S., Pereira, A. C., & Xavier-Neto, J. (2009). ALDH1A2 (RALDH2) genetic variation in human congenital heart disease. *BMC Medical Genetics*, *10*(1), 113. https://doi.org/10.1186/1471-2350-10-113

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., & Cournapeau, D. (n.d.). Scikit-learn: Machine Learning in Python. *MACHINE LEARNING IN PYTHON*.

Qian, L., Huang, Y., Spencer, C. I., Foley, A., Vedantham, V., Liu, L., Conway, S. J., Fu, J., & Srivastava, D. (2012). In vivo reprogramming of murine cardiac fibroblasts into induced cardiomyocytes. *Nature*, *485*(7400), 593–598. https://doi.org/10.1038/nature11044

Quetelet, A. (1835). *Sur l'homme et le développement de ses facultés: Ou, Essai de physique sociale*. Paris : Bachelier, imprimeur-libraire, quai des Augustins, no 55. http://archive.org/details/surlhommeetled00quet

Ramsbottom, S. A., Sharma, V., Rhee, H. J., Eley, L., Phillips, H. M., Rigby, H. F., Dean, C., Chaudhry, B., & Henderson, D. J. (2014). Vangl2-Regulated Polarisation of Second Heart Field-Derived Cells Is Required for Outflow Tract Lengthening during Cardiac Development. *PLOS Genetics*, *10*(12), e1004871. https://doi.org/10.1371/journal.pgen.1004871

Rana, M. S., Théveniau-Ruissy, M., De Bono, C., Mesbah, K., Francou, A., Rammah, M., Domínguez, J. N., Roux, M., Laforest, B., Anderson, R. H., Mohun, T., Zaffran, S., Christoffels, V. M., & Kelly, R. G. (2014). Tbx1 coordinates addition of posterior second heart field progenitor cells to the arterial and venous poles of the heart. *Circulation Research*, *115*(9), 790–799. https://doi.org/10.1161/CIRCRESAHA.115.305020

Rankin, S. A., Steimle, J. D., Yang, X. H., Rydeen, A. B., Agarwal, K., Chaturvedi, P., Ikegami, K., Herriges, M. J., Moskowitz, I. P., & Zorn, A. M. (2021). Tbx5 drives Aldh1a2 expression to regulate a RA-Hedgehog-Wnt gene regulatory network coordinating cardiopulmonary development. *eLife*, *10*, e69288. https://doi.org/10.7554/eLife.69288

Rauskolb, C., Sun, S., Sun, G., Pan, Y., & Irvine, K. D. (2014). Cytoskeletal tension inhibits Hippo signaling through an Ajuba-Warts complex. *Cell*, *158*(1), 143–156. https://doi.org/10.1016/j.cell.2014.05.035

Rauzi, M., Lenne, P.-F., & Lecuit, T. (2010). Planar polarized actomyosin contractile flows control epithelial junction remodelling. *Nature*, *468*(7327), 1110–1114. https://doi.org/10.1038/nature09566

Riching, A. S., & Song, K. (2021). Cardiac Regeneration: New Insights Into the Frontier of Ischemic Heart Failure Therapy. *Frontiers in Bioengineering and Biotechnology*, *8*. https://www.frontiersin.org/articles/10.3389/fbioe.2020.637538

Roh-Johnson, M., Shemer, G., Higgins, C. D., McClellan, J. H., Werts, A. D., Tulu, U. S., Gao, L., Betzig, E., Kiehart, D. P., & Goldstein, B. (2012). Triggering a cell shape change by exploiting preexisting actomyosin contractions. *Science (New York, N.Y.)*, *335*(6073), 1232–1235. https://doi.org/10.1126/science.1217869

Rohr, S., Bit-Avragim, N., & Abdelilah-Seyfried, S. (2006). Heart and soul/PRKCi and nagie oko/Mpp5 regulate myocardial coherence and remodeling during cardiac morphogenesis. *Development (Cambridge, England)*, *133*(1), 107–115. https://doi.org/10.1242/dev.02182

Rossi, G., Broguiere, N., Miyamoto, M., Boni, A., Guiet, R., Girgin, M., Kelly, R. G., Kwon, C., & Lutolf, M. P. (2021). Capturing Cardiogenesis in Gastruloids. *Cell Stem Cell*, *28*(2), 230-240.e6. https://doi.org/10.1016/j.stem.2020.10.013

Roy, S. G., & Chakrabarti, A. (2017). Chapter 11—A novel graph clustering algorithm based on discrete-time quantum random walk. In S. Bhattacharyya, U. Maulik, & P. Dutta (Eds.), *Quantum Inspired Computational Intelligence* (pp. 361–389). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-804409-4.00011-5

Rydeen, A. B., & Waxman, J. S. (2016). Cyp26 Enzymes Facilitate Second Heart Field Progenitor Addition and Maintenance of Ventricular Integrity. *PLoS Biology*, *14*(11), e2000504. https://doi.org/10.1371/journal.pbio.2000504

Saga, Y., Hata, N., Koseki, H., & Taketo, M. M. (1997). Mesp2: A novel mouse gene expressed in the presegmented mesoderm and essential for segmentation initiation. *Genes & Development*, *11*(14), 1827–1839. https://doi.org/10.1101/gad.11.14.1827

Sagner, A., Merkel, M., Aigouy, B., Gaebel, J., Brankatschk, M., Jülicher, F., & Eaton, S. (2012). Establishment of global patterns of planar polarity during growth of the Drosophila wing epithelium. *Current Biology: CB*, *22*(14), 1296–1301. https://doi.org/10.1016/j.cub.2012.04.066

Sawyer, J. K., Choi, W., Jung, K.-C., He, L., Harris, N. J., & Peifer, M. (2011). A contractile actomyosin network linked to adherens junctions by Canoe/afadin helps drive convergent extension. *Molecular Biology of the Cell*, *22*(14), 2491–2508. https://doi.org/10.1091/mbc.E11-05-0411

Schmidt, C., Deyett, A., Ilmer, T., Haendeler, S., Torres Caballero, A., Novatchkova, M., Netzer, M. A., Ceci Ginistrelli, L., Mancheno Juncosa, E., Bhattacharya, T., Mujadzic, A., Pimpale, L., Jahnel, S. M., Cirigliano, M., Reumann, D., Tavernini, K., Papai, N., Hering, S., Hofbauer, P., & Mendjan, S. (2023). Multi-chamber cardioids unravel human heart development and cardiac defects. *Cell*, *186*(25), 5587-5605.e27. https://doi.org/10.1016/j.cell.2023.10.030

Schreiner, Y., Schaible, T., & Rafat, N. (2021). Genetics of diaphragmatic hernia. *European Journal of Human Genetics*, *29*(12), 1729–1733. https://doi.org/10.1038/s41431-021-00972-0

Schwartz, M. A. (2010). Integrins and Extracellular Matrix in Mechanotransduction. *Cold Spring Harbor Perspectives in Biology*, *2*(12), a005066. https://doi.org/10.1101/cshperspect.a005066

Sendra, M., McDole, K., Jimenez-Carretero, D., Hourcade, J. de D., Temiño, S., Raiola, M., Guignard, L., Keller, P. J., Sánchez-Cabo, F., Domínguez, J. N., & Torres, M. (2024). *Epigenetic priming of embryonic lineages in the mammalian epiblast* (p. 2024.01.11.575188). bioRxiv. https://doi.org/10.1101/2024.01.11.575188

Shparberg, R., Glover, H., & Morris, M. (2019). Modeling Mammalian Commitment to the Neural Lineage Using Embryos and Embryonic Stem Cells. *Frontiers in Physiology*, *10*. https://doi.org/10.3389/fphys.2019.00705

Singh, A. P., Wu, P., Ryabichko, S., Raimundo, J., Swan, M., Wieschaus, E., Gregor, T., & Toettcher, J. E. (2022). Optogenetic control of the Bicoid morphogen reveals fast

and slow modes of gap gene regulation. *Cell Reports*, *38*(12), 110543. https://doi.org/10.1016/j.celrep.2022.110543

Sinha, T., Li, D., Théveniau-Ruissy, M., Hutson, M. R., Kelly, R. G., & Wang, J. (2015). Loss of Wnt5a disrupts second heart field cell deployment and may contribute to OFT malformations in DiGeorge syndrome. *Human Molecular Genetics*, *24*(6), 1704–1716. https://doi.org/10.1093/hmg/ddu584

Sirbu, I. O., Zhao, X., & Duester, G. (2008). Retinoic Acid Controls Heart Anteroposterior Patterning by Downregulating Isl1 through the Fgf8 Pathway. *Developmental Dynamics : An Official Publication of the American Association of Anatomists*, *237*(6), 1627–1635. https://doi.org/10.1002/dvdy.21570

Smith, A. M., Baker, R. E., Kay, D., & Maini, P. K. (2012). Incorporating chemical signalling factors into cell-based models of growing epithelial tissues. *Journal of Mathematical Biology*, *65*(3), 441–463. https://doi.org/10.1007/s00285-011-0464-y

Sneath1, P. H. A. (1967). Trend-surface analysis of transformation grids. *Journal of Zoology*, *151*(1), 65–122. https://doi.org/10.1111/j.1469-7998.1967.tb02866.x

Soh, B.-S., Buac, K., Xu, H., Li, E., Ng, S.-Y., Wu, H., Chmielowiec, J., Jiang, X., Bu, L., Li, R. A., Cowan, C., & Chien, K. R. (2014). N-cadherin prevents the premature differentiation of anterior heart field progenitors in the pharyngeal mesodermal microenvironment. *Cell Research*, *24*(12), 1420–1432. https://doi.org/10.1038/cr.2014.142

Solnica-Krezel, L. (2005). Conserved Patterns of Cell Movements during Vertebrate Gastrulation. *Current Biology*, *15*(6), R213–R228. https://doi.org/10.1016/j.cub.2005.03.016

Srivastava, D. (2006). Making or Breaking the Heart: From Lineage Determination to Morphogenesis. *Cell*, *126*(6), 1037–1048. https://doi.org/10.1016/j.cell.2006.09.003

Steimle, J. D., & Moskowitz, I. P. (2017). TBX5: A Key Regulator of Heart Development. *Current Topics in Developmental Biology*, *122*, 195–221. https://doi.org/10.1016/bs.ctdb.2016.08.008

Stevens, M., Nanou, A., Terstappen, L. W. M. M., Driemel, C., Stoecklein, N. H., & Coumans, F. A. W. (2022). StarDist Image Segmentation Improves Circulating Tumor Cell Detection. *Cancers*, *14*(12), 2916. https://doi.org/10.3390/cancers14122916

Stower, M., Zhou, F., Hathrell, H., Yeung, J., Thowfeequ, S., Godwin, J., Schneider, F., Lagerholm, C., Fritzsche, M., Thiyagalingam, J., Lu, X., Rittscher, J., & Srinivas, S. (2023). *Single-cell phenomics reveals behavioural and mechanical heterogeneities underpinning collective migration during mouse anterior patterning* (p. 2023.03.31.534937). bioRxiv. https://doi.org/10.1101/2023.03.31.534937

Stringer, C., Wang, T., Michaelos, M., & Pachitariu, M. (2021). Cellpose: A generalist algorithm for cellular segmentation. *Nature Methods*, *18*(1), Article 1. https://doi.org/10.1038/s41592-020-01018-x

Sugimura, K., Bellaïche, Y., Graner, F., Marcq, P., & Ishihara, S. (2013). Robustness of force and stress inference in an epithelial tissue. *Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE Engineering in Medicine and Biology Society. Annual International Conference*, *2013*, 2712–2715. https://doi.org/10.1109/EMBC.2013.6610100

Takahashi, Y., Koizumi, K., Takagi, A., Kitajima, S., Inoue, T., Koseki, H., & Saga, Y. (2000). Mesp2 initiates somite segmentation through the Notch signalling pathway. *Nature Genetics*, *25*(4), 390–396. https://doi.org/10.1038/78062

Tanaka, M., Chen, Z., Bartunkova, S., Yamasaki, N., & Izumo, S. (1999). The cardiac homeobox gene Csx/Nkx2.5 lies genetically upstream of multiple genes essential for heart development. *Development (Cambridge, England)*, *126*(6), 1269–1280. https://doi.org/10.1242/dev.126.6.1269

Théveniau-Ruissy, M., Dandonneau, M., Mesbah, K., Ghez, O., Mattei, M.-G., Miquerol, L., & Kelly, R. G. (2008). The del22q11.2 Candidate Gene Tbx1 Controls Regional Outflow Tract Identity and Coronary Artery Patterning. *Circulation Research*, *103*(2), 142–148. https://doi.org/10.1161/CIRCRESAHA.108.172189

Thomson, J. A. (1917). On Growth and Form. *Nature*, *100*(2498), Article 2498. https://doi.org/10.1038/100021a0

Thorndike, R. L. (1953). Who belongs in the family? *Psychometrika*, *18*(4), 267–276. https://doi.org/10.1007/BF02289263

Trinh, L. A., Yelon, D., & Stainier, D. Y. R. (2005). Hand2 regulates epithelial formation during myocardial diferentiation. *Current Biology: CB*, *15*(5), 441–446. https://doi.org/10.1016/j.cub.2004.12.083

Tryon, R. C. (1939). *Cluster analysis: Correlation profile and orthometric (factor) analysis for the isolation of unities in mind and personality*. Edwards brother, inc., lithoprinters and publishers.

Tyser, R. C. V., Ibarra-Soria, X., McDole, K., Jayaram, S. A., Godwin, J., Brand, T. A. H. van den, Miranda, A. M. A., Scialdone, A., Keller, P. J., Marioni, J. C., & Srinivas, S. (2021). Characterization of a common progenitor pool of the epicardium and myocardium. *Science*. https://doi.org/10.1126/science.abb2986

Unolt, M., Versacci, P., Anaclerio, S., Lambiase, C., Calcagni, G., Trezzi, M., Carotti, A., Crowley, T. B., Zackai, E. H., Goldmuntz, E., Gaynor, J. W., Digilio, M. C., McDonald-McGinn, D. M., & Marino, B. (2018). Congenital heart diseases and cardiovascular

abnormalities in 22q11.2 deletion syndrome: From well-established knowledge to new frontiers. *American Journal of Medical Genetics. Part A*, *176*(10), 2087–2098. https://doi.org/10.1002/ajmg.a.38662

van der Linde, D., Konings, E. E. M., Slager, M. A., Witsenburg, M., Helbing, W. A., Takkenberg, J. J. M., & Roos-Hesselink, J. W. (2011). Birth prevalence of congenital heart disease worldwide: A systematic review and meta-analysis. *Journal of the American College of Cardiology*, *58*(21), 2241–2247. https://doi.org/10.1016/j.jacc.2011.08.025

Varner, V. D., & Taber, L. A. (2012). Not just inductive: A crucial mechanical role for the endoderm during heart tube assembly. *Development (Cambridge, England)*, *139*(9), 1680–1690. https://doi.org/10.1242/dev.073486

Verzi, M. P., McCulley, D. J., De Val, S., Dodou, E., & Black, B. L. (2005). The right ventricle, outflow tract, and ventricular septum comprise a restricted expression domain within the secondary/anterior heart field. *Developmental Biology*, *287*(1), 134–145. https://doi.org/10.1016/j.ydbio.2005.08.041

Villoutreix, P. (2021). What machine learning can do for developmental biology. *Development (Cambridge, England)*, *148*(1), dev188474. https://doi.org/10.1242/dev.188474

Vincent, S. D., & Buckingham, M. E. (2010). How to make a heart: The origin and regulation of cardiac progenitor cells. *Current Topics in Developmental Biology*, *90*, 1–41. https://doi.org/10.1016/S0070-2153(10)90001-X

Virágh, S., & Challice, C. E. (1973). Origin and differentiation of cardiac muscle cells in the mouse. *Journal of Ultrastructure Research*, *42*(1), 1–24. https://doi.org/10.1016/S0022-5320(73)80002-4

von Both, I., Silvestri, C., Erdemir, T., Lickert, H., Walls, J. R., Henkelman, R. M., Rossant, J., Harvey, R. P., Attisano, L., & Wrana, J. L. (2004). Foxh1 is essential for development of the anterior heart field. *Developmental Cell*, *7*(3), 331–345. https://doi.org/10.1016/j.devcel.2004.07.023

Waldo, K. L., Kumiski, D. H., Wallis, K. T., Stadt, H. A., Hutson, Mary. R., Platt, D. H., & Kirby, M. L. (2001). Conotruncal myocardium arises from a secondary heart field. *Development*, *128*(16), 3179–3188. https://doi.org/10.1242/dev.128.16.3179

Wallingford, J. B., Fraser, S. E., & Harland, R. M. (2002). Convergent Extension: The Molecular Control of Polarized Cell Movement during Embryonic Development. *Developmental Cell*, *2*(6), 695–706. https://doi.org/10.1016/S1534-5807(02)00197-1

Watanabe, T., Sato, Y., Saito, D., Tadokoro, R., & Takahashi, Y. (2009). EphrinB2 coordinates the formation of a morphological boundary and cell epithelialization during somite segmentation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(18), 7467–7472. https://doi.org/10.1073/pnas.0902859106

Weber, C. M., Ray, D., Valverde, A. A., Clark, J. A., & Sharma, K. S. (2022). Gaussian mixture model clustering algorithms for the analysis of high-precision mass measurements. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, *1027*, 166299. https://doi.org/10.1016/j.nima.2021.166299

Wilson, R., Ross, O., & Griksaitis, M. J. (2019). Tetralogy of Fallot. *BJA Education*, *19*(11), 362–369. https://doi.org/10.1016/j.bjae.2019.07.003

Xie, L., Hoffmann, A. D., Burnicka-Turek, O., Friedland-Little, J. M., Zhang, K., & Moskowitz, I. P. (2012). Tbx5-hedgehog molecular networks are essential in the second heart

field for atrial septation. *Developmental Cell*, *23*(2), 280–291. https://doi.org/10.1016/j.devcel.2012.06.006

Zallen, J. A., & Wieschaus, E. (2004). Patterned Gene Expression Directs Bipolar Planar Polarity in Drosophila. *Developmental Cell*, *6*(3), 343–355. https://doi.org/10.1016/S1534-5807(04)00060-7

Zhang, Q., Carlin, D., Zhu, F., Cattaneo, P., Ideker, T., Evans, S. M., Bloomekatz, J., & Chi, N. C. (2021). Unveiling Complexity and Multipotentiality of Early Heart Fields. *Circulation Research*, *129*(4), 474–487. https://doi.org/10.1161/CIRCRESAHA.121.318943

Annex - Review

ACADÉMIE DES SCIENCES
INSTITUT DE FRANCE

Comptes Rendus. Biologies
2024, Vol. 347, p. 9-18
https://doi.org/10.5802/crbiol.151

Review article

On the involvement of the second heart field in congenital heart defects

Clara Guijarro [a] **and Robert G. Kelly** [*, a]

[a] Aix-Marseille Université, CNRS UMR 7288, IBDM, Marseille, France
E-mail: Robert.Kelly@univ-amu.fr (R. G. Kelly)

Abstract. Congenital heart defects (CHD) affect 1 in 100 live births and result from defects in cardiac development. Growth of the early heart tube occurs by the progressive addition of second heart field (SHF) progenitor cells to the cardiac poles. The SHF gives rise to ventricular septal, right ventricular and outflow tract myocardium at the arterial pole, and atrial, including atrial septal myocardium, at the venous pole. SHF deployment creates the template for subsequent cardiac septation and has been implicated in cardiac looping and in orchestrating outflow tract development with neural crest cells. Genetic or environmental perturbation of SHF deployment thus underlies a spectrum of common forms of CHD affecting conotruncal and septal morphogenesis. Here we review the major properties of SHF cells as well as recent insights into the developmental programs that drive normal cardiac progenitor cell addition and the origins of CHD.

Keywords. Heart development, Embryology, Cardiac progenitor cells, Congenital heart defects.

Note. Robert G. Kelly is the recipient of the prix Lamonica de Cardiologie - Fondation pour la Recherche Biomédicale PCL 2022 from the Académie des sciences.

Funding. Agence National pour la Recherche (PP-heart ANR-20-CE13-0029-01 and Heartbound ANR-22-CE13 projects), AFM-Telethon, French National Research Agency (ANR-16-CONV-0001), Excellence Initiative of Aix-Marseille University A*MIDEX (Turing Centre for Living Systems) and the Fondation pour la Recherche Médicale.

Published online: 15 March 2024

1. Introduction

Defects in the complex process of cardiac morphogenesis underlie congenital heart defects (CHD) that affect 1 in 100 live births [1]. The heart is derived from mesoderm, the central of the three embryonic germ layers, and originates in the anterior region of the early embryo as a tubular structure with venous (inflow) and arterial (outflow) poles. Cardiac progenitor cells located in splanchnic mesoderm, overlying future foregut endoderm, contribute progressively to growth of the heart. These cells, known as the second heart field (SHF), ultimately give rise to ventricular septal, right ventricular and outflow tract myocardium at the arterial pole, and atrial, including atrial septal myocardium, at the venous pole. Deployment of SHF cells is coincident with looping morphogenesis, an essential conserved process during which the heart tube rotates to the right to correctly position the future regions of the definitive heart. Subsequent steps of heart development involve cardiomyocyte proliferation driving chamber morphogenesis, development of the coronary arteries and cardiac conduction system coordinating the heartbeat, and cardiac septation. During septation the formation of ventricular and atrial septa generates the four chambered heart, and the embryonic outflow tract is divided into the outlets of the left and right ventricles, the ascending aorta and pulmonary trunk. The initially tubular embryonic heart

* Corresponding author.

thus acquires independent right atrioventricular and left ventriculoarterial junctions. Cardiac septation is complete at birth when the pulmonary duct and the interatrial foramen close, coincident with pulmonary perfusion, isolating the pulmonary and systemic circulatory systems.

Heart tube elongation by addition of SHF cells creates the template for cardiac septation and therefore perturbation of SHF deployment in human patients and animal models results in a spectrum of common forms of CHD [2, 3]. These include ventricular and atrial septal defects as well as outflow or conotruncal defects, and lead to failure to correctly separate the systemic and pulmonary circulation. Dissecting the developmental programs driving normal cardiac progenitor cell deployment is thus essential to understand how defects in these processes result in CHD, as well as for directing stem cell derived cardiomyocytes to different regional fates for drug testing and tissue repair following cardiac damage. Here we review the contribution of the SHF during heart development and its involvement in CHD, with focus on the roles of two T-box transcription factors implicated in human CHD, TBX1 and TBX5.

2. Early heart development and the second heart field

Nascent mesodermal cardiac progenitor cells emerge from the primitive streak at gastrulation and migrate to the anterior lateral region of the embryo (Figure 1A). Two distinct waves of progenitor cells have been identified that independently activate the gene encoding the early mesodermal transcription factor MESP1 [4–6]. The first wave generates first heart field (FHF) progenitor cells that differentiate to form bilateral cardiac primordia that converge in the ventral midline as the foregut closes, giving rise to the cardiac crescent and ultimately to left ventricular and atrial myocardium.

The cardiac crescent is transformed into a transiently linear heart tube that subsequently elongates and loops to the right, generating the embryonic heart at midgestation in the mouse, equivalent to the 5th week of human development (Figure 1B). Heart tube elongation is driven by the progressive addition of SHF cells, derived from the second wave of *Mesp1* expression, located in pharyngeal mesoderm dorsomedial to the early heart tube, including the dorsal wall of the pericardial cavity [5]. The SHF was first identified in 2001 through analysis of a *Fgf10* enhancer trap transgene expressed in murine pharyngeal mesoderm [7] and by cell labelling experiments in avian embryos [8, 9]. These findings built on pioneering embryological evidence for the addition of extracardiac cells to the growing cardiac poles [10, 11]. Initially contributing along the entire length of the cardiac primordium, SHF addition becomes restricted to the arterial and venous poles after breakdown of the dorsal mesocardium separates the heart tube from the dorsal pericardial wall (Figure 1B). Genetic tracing and retrospective lineage experiments have shown that the first and second heart fields correspond to distinct genetic lineages defined by their exclusive contributions to the outflow tract and left ventricle respectively, and that SHF progenitor cells adding to the arterial and venous poles are clonally related [12–14]. Subsequent labeling experiments and genetic lineage analyses have revealed the progressive addition of SHF progenitor cells to the elongating heart tube, giving rise successively to the right ventricle, proximal and distal outflow tract, and atrial myocardium, as well as smooth muscle cells of the pulmonary trunk (Figure 1C) [15–18]. Clonal relationships have been established between these SHF subpopulations and branchiomeric skeletal muscles of the head and neck, consistent with the existence of common cardiac and skeletal muscle progenitor cells in pharyngeal mesoderm [19–22]. This is supported by the recent identification of multilineage primed progenitor cells that contribute to skeletal muscle and myocardium, using single-cell RNA-seq analysis [23]. Single-cell RNA sequencing is a revolutionary approach that has yielded unprecedented insights into heterogeneity and developmental trajectories of cardiac progenitor cell populations [24]. These include the identification of a population of FHF cells termed the juxtacardiac field, located at the lateral boundary between the embryo and extraembryonic mesoderm, that contributes to growth of the early cardiac primordium along its ventrolateral margin, in contrast to SHF contributions along the dorsomedial margin [25–27]. Single-cell transcriptomics has also revealed differences in gene expression between left and right SHF cells, consistent with the recent finding that the SHF is a driver of looping morphogenesis [28, 29]. Defects in cardiac looping, or in

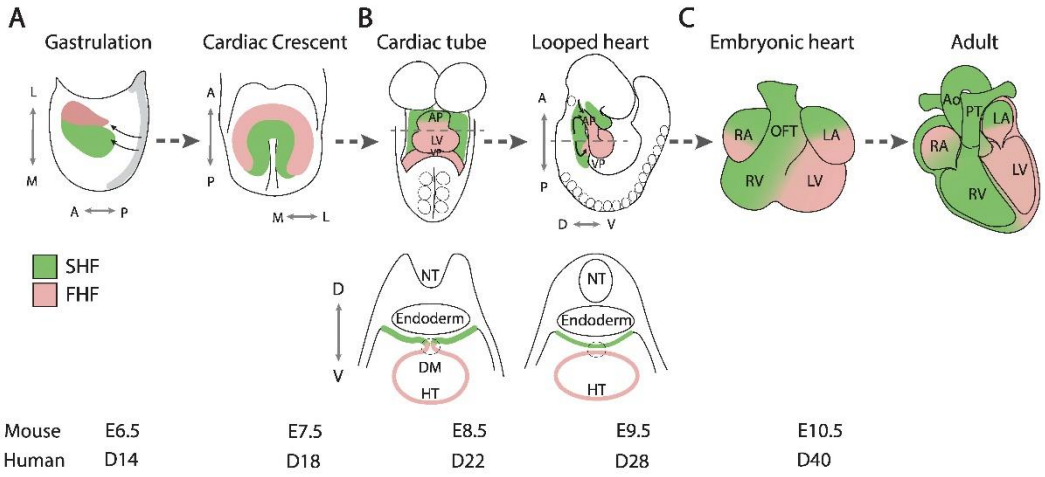


Figure 1. Cardiac development. Cartoon showing the contribution of first (pink) and second (green) heart field progenitor cells to the developing mouse heart. (A) Heart tube morphogenesis, (B) heart tube looping and elongation and (C) chamber growth and cardiac septation. FHF, first heart field; SHF, second heart field; L, lateral; M, medial; A, anterior; P, posterior; D, dorsal; V, ventral; AP, arterial pole; VP, venous pole; HT, heart tube; RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle; NT, neural tube; DM, dorsal mesocardium; OFT, outflow tract; Ao, aorta; PT, pulmonary trunk; E: mouse embryonic day, D: human embryonic day.

the prior establishment of embryonic laterality, have been implicated in a spectrum of CHD [29, 30]. Dynamic imaging using two-photon and laser scanning fluorescent microscopy have begun to confirm the trajectories taken by cardiac progenitor cells during growth of the early heart. These experiments have revealed a temporal delay between first and second heart field cell deployment and distinct patterns of cell behavior in the juxtacardiac and second heart fields during heart tube assembly [6, 31]. In addition, quantitative morphometric approaches provide high-resolution understanding of the complex events associated with early heart development [32].

3. TBX1 and arterial pole development

Progenitor cells adding to the arterial pole of the heart, known as the anterior SHF, give rise to right ventricular and outflow tract myocardium. As these cells converge on the growing arterial pole, they move through a transition zone in which progenitor cell gene expression is downregulated and myocardial differentiation activated (Figure 2A) [33, 34]. While lack of SHF deployment results in severe early lethal phenotypes, milder defects in this process generate an outflow tract that is too short to allow the normal alignment between the left ventricle and ascending aorta during cardiac septation. Consequently, anterior SHF defects contribute to conotruncal CHD, that account for 30% of CHD, ranging from partial ventriculoarterial misalignment as in the case of overriding aorta and tetralogy of Fallot, to double outlet right ventricle (Figure 2B). Defective growth of the ascending aorta has also been implicated in transposition of the great arteries [35].

Work in many groups worldwide has identified critical transcriptional regulators, intercellular signaling pathways and epithelial features that maintain SHF progenitor cell status and regulate progressive differentiation at the cardiac poles [36–39]. In addition to *Fgf10* expression, SHF cells are demarcated by expression of the transcription factors ISL1, TBX1 and an enhancer from the *Mef2c* locus [7, 15, 16, 40]. These transcription factors integrate the activity of signaling events including autocrine signaling from pharyngeal mesoderm as well as signal

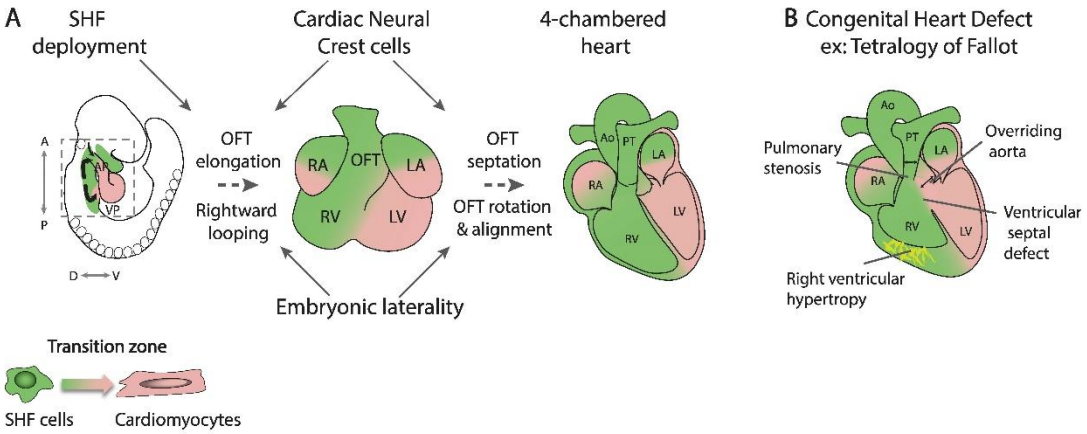


Figure 2. Arterial pole development and outflow tract septation. (A) While SHF cells contribute to the myocardial wall of the outflow tract at the transition zone, SHF addition is regulated by cardiac neural crest cells that also mediate outflow tract septation. Embryonic laterality impacts on rightward looping and ventriculoarterial alignment. (B) An example of a conotruncal congenital heart defect, tetralogy of Fallot. SHF, Second heart field; AP, arterial pole; VP, venous pole; A, Anterior; P, Posterior; D, Dorsal; V, ventral; RA, right atrium; LA, left atrium; RV, right ventricle; LV, left ventricle; OFT, outflow tract; Ao, aorta; PT, pulmonary trunk.

exchange with surrounding cell types, in particular adjacent ventral pharyngeal endoderm and neural crest-derived mesenchyme. Heart tube elongation is regulated by a balance between pro-proliferative FGF signals and pro-differentiation BMP signaling; FGF10 expression in the SHF, for example, acts together with FGF8 to control arterial pole development [41]. As SHF cells differentiate in the transition zone, *Fgf10* expression is downregulated by NKX2-5, that blocks ISL1-mediated activation of an intronic *Fgf10* enhancer [42]. SHF cells in the dorsal pericardial wall have been shown to constitute an atypical epithelium, with apical tight and adherens junctions and highly dynamic basal filopodia more typical of mesenchymal cells [43]. The epithelial properties of SHF cells are emerging as an important regulatory node controlling SHF deployment and are regulated by cell adhesion molecules and components of the planar cell polarity pathway, such as WNT5A and VANGL2, as well as TBX1, required for focal adhesion and filopodial activity in the SHF [33, 38, 43, 44]. Moreover, analysis of apical junctions, cell shape and actomyosin distribution in the SHF has suggested biomechanical models driving SHF deployment, including planar cell polarity driven push and pull mechanisms and an epithelial tension driven circuit promoting proliferation in the posterior region of the dorsal pericardial wall [45, 46].

Cardiac neural crest cells, like the SHF, move through the pharyngeal region into the developing arterial pole of the heart, playing an important role in dividing the embryonic outflow tract into the pulmonary trunk and ascending aorta [47]. Ablation of neural crest cells thus leads to persistent common arterial trunk due to a failure of outflow tract septation; cardiac neural crest cell ablation also leads to a shortened outflow tract through reduced SHF deployment (Figure 2A) [48, 49]. This observation in avian embryos was among the first to show the significance of SHF addition for cardiac looping and outflow tract alignment. These phenotypes reflect signal exchange between SHF and neural crest cells as they converge on the arterial pole of the heart, orchestrating normal outflow tract elongation. The T-box transcription factor TBX1 has been identified as a critical regulator of SHF development, promoting progenitor cell proliferation and delaying differentiation in the SHF [40]. *TBX1* is the major gene implicated in 22q11.2 deletion or DiGeorge syndrome, characterized by craniofacial and cardiovascular defects, including a range

of conotruncal CHD [50, 51]. Single-cell RNA-seq experiments have shown that Tbx1 controls myogenic trajectories from the multilineage primed progenitor cell population to cardiac and skeletal muscle fates [23]. *Tbx1* null embryos display common arterial trunk associated with defective cardiac neural crest cell development; in addition, cells giving rise to myocardium at the base of the pulmonary trunk, the outlet of the right ventricle, fail to contribute to the arterial pole of *Tbx1* mutant hearts [40, 52, 53]. Underdevelopment of subpulmonary myocardium has been proposed to be a primary upstream mechanism leading to tetralogy of Fallot [54]. Indeed, tetralogy of Fallot is genetically heterogenous but the major known genetic cause is haploinsufficiency for *TBX1*, 22q11.2 DS accounting for over 15% of patients [50]. Moreover, SHF ablation in avian embryos leads to a tetralogy of Fallot-like phenotype [55]. Distinct genetic programs have been defined in subpulmonary and subaortic myocardium, reflecting earlier differences in the distribution of progenitor cells in the SHF; recently *Pparg* and *Dlk1* have been shown to demarcate subpulmonary and subaortic myocardium respectively [53, 56, 57]. Future subpulmonary myocardium, initially located in the inferior wall of the midgestation outflow tract, becomes positioned ventrally during outflow tract septation through counterclockwise rotation of the myocardial wall driven by ongoing SHF addition (Figure 2A) [58, 59].

Genetic studies have identified mutations associated with sporadic and familial CHD, including mutations in genes encoding T-box transcription factors and chromatin regulators [60]. Examples include *TBX1* and *CHD7*, the latter implicated in cardiovascular defects associated with CHARGE syndrome. CHD7 is an ATP-dependent chromatin remodeler required for normal contributions of both neural crest and SHF lineages during heart development, including regulation of *Fgf10* expression in the SHF through the ISL1-bound intronic enhancer [61]. Integration of genetic studies with single cell RNA-seq datasets can provide further mechanistic insights and has, for example, highlighted the importance of outflow tract progenitor cells in tetralogy of Fallot [62]. However, genetic defects only account for one third of CHD cases to date, suggesting an important, yet poorly understood, effect of *in utero* exposure to environmental factors. Indeed, recent studies in the mouse have identified the SHF as the target of short-term gestational hypoxia, maternal iron deficiency and prenatal alcohol exposure, in each case leading to reduced deployment of SHF progenitor cells and a shortened OFT, resulting in a range of conotruncal CHD [63–65].

4. TBX5 and cardiac septation

Dye labelling and genetic lineage tracing experiments have shown that subpulmonary myocardium originates in *Hoxb1* expressing cells in the posterior lateral region of the dorsal pericardial wall, corresponding to the location of proliferative multilineage primed progenitor cells [23, 53, 66]. Downregulation of *Hoxb1* expression is required for normal arterial pole and right ventricular development [67]. Cardiac progenitor cells in the *Isl1*, *Hoxb1*, *Tbx1*, *Wnt5a* and *Mef2c* enhancer genetic lineages contribute not only to the arterial pole but also to the venous pole of the heart, making critical contributions to atrial septal structures including the primary atrial septum and dorsal mesenchymal protrusion that bridges the atrial septum with the atrioventricular cushions (Figure 3A) [53, 67–70]. Defects in dorsal mesenchymal protrusion development result in primum type atrial septal defects, a class of atrioventricular septal defect [71].

In contrast to the anterior SHF, which retains the early SHF genetic program, posterior SHF cells activate *Tbx5*, a key regulator of the FHF and venous pole differentiation program, under the influence of retinoic acid signaling [72, 73]. *Tbx5* is upregulated in TBX1 positive pharyngeal mesoderm as the dorsal mesocardium breaks down and progenitor cells segregate to alternate poles (Figure 3B). Transient co-expression of TBX1 and TBX5 is followed by down-regulation of the anterior SHF program and the establishment of a transcriptional boundary within the dorsal pericardial wall, separating TBX1 positive arterial pole progenitors in the anterior SHF from TBX5 positive venous pole progenitors in the medial posterior SHF [73]. Mutations in *TBX5* result in Holt-Oram syndrome, characterized by atrial septal CHD and forelimb defects [74]. TBX5 has recently been shown to activate an enhancer of *Aldh1a2*, encoding a retinoic acid biosynthetic enzyme, pointing to a feed-forward mechanism between TBX5 and retinoic acid signaling as the posterior SHF domain is established in the dorsal pericardial wall [75]. The

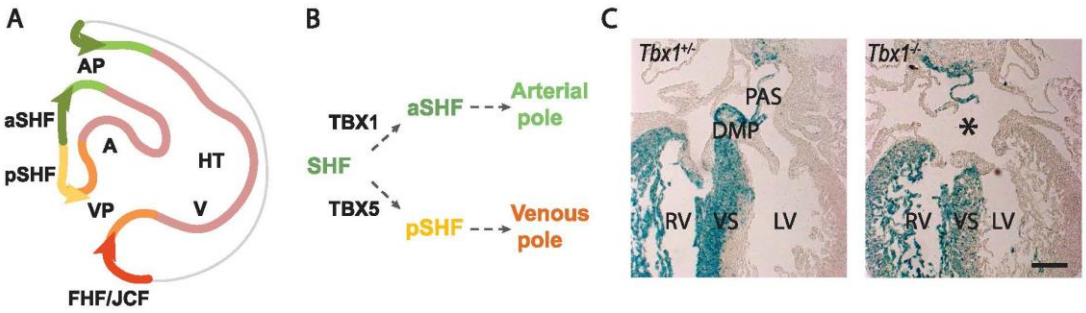


Figure 3. Venous pole development and atrial septation. (A) Scheme of the region boxed in Figure 2A, showing progenitor cell segregation into anterior SHF (aSHF) and posterior SHF (pSHF) domains contributing to the arterial and venous poles. Note the contribution of first heart field (FHF) cells from the juxtacardiac field (JCF) to the venous pole. (B) Diagram showing that Tbx1 expression is maintained in the aSHF, while Tbx5 is activated in the pSHF. (C) Left: the SHF genetic lineage (blue) labelled by a conditional β-galactosidase reporter gene activated by Cre recombinase in the SHF, showing labeling in the right ventricle (RV), interventricular septum (VS) and atrial septal structures: the primary atrial septum (PAS) and dorsal mesenchymal protrusion (DMP). Right: the DMP fails to form in *Tbx1* mutant hearts resulting in a primum type atrial septal defect (asterisk). AP, arterial pole; VP, venous pole; A, atrium; V, ventricle; RV, right ventricle; LV, left ventricle; HT, heart tube. Scale bar in C: 100 μm

boundary between anterior and posterior domains is perturbed in the absence of *Tbx1*, resulting in failure of dorsal mesenchymal protrusion development and atrial septal defects in 30% of *Tbx1* null embryos, potentially accounting for a low frequency of atrial septal defects in 22q11.2 DS patients (Figure 3C) [53]. Similar atrial septal defects are observed in conditional mutant mouse embryos lacking *Tbx5* in the SHF [76]. The posterior SHF shares lineage contributions with pulmonary mesenchyme and signal exchange between the SHF and adjacent pulmonary endoderm plays an important role in venous pole development [75, 77]. TBX5 activates Wnt ligand expression in the posterior SHF that signals to pulmonary endoderm, in turn controlling the timing of differentiation in the posterior SHF through the Hedgehog signaling pathway [78, 79].

Cells expressing both *Tbx5* and the SHF program have also been identified in the interventricular septal region using an intersectional lineage reporter approach [4]. Conditional inactivation using Cre recombinase driven by the *Mef2c* enhancer has shown that *Tbx5* is required in the SHF lineage for formation of the muscular ventricular septum as well as for atrial septation [76, 80]; moreover, the transcription factors MEF2C and TBX5 coregulate target genes required for ventricular septal development [81]. Thus both atrial and ventricular septal structures form at sites where *Tbx5* expression and the SHF program transiently overlap, suggesting that this defines a boundary program in the developing heart required to initiate septal morphogenesis. Atrial and ventricular CHD together account for over 50% of CHD [82]. Further dissection of this regulatory program, operating at the interface between the heart fields, will be essential to increase our understanding of the origins of septal CHD.

5. Conclusions and perspectives

Since the discovery of the SHF extensive evidence has revealed the importance of this spatiotemporally heterogenous progenitor cell population in the etiology of CHD affecting development of the cardiac poles and septa. While severe heart tube elongation defects result in early embryonic lethality, milder perturbation of cell addition at the cardiac poles results in alignment defects and failure to separate the systemic and pulmonary circulation at birth, contributing to a spectrum of common forms of CHD encountered in human patients. Technological advances, such as single-cell transcriptomics and quantitative

morphometric approaches, as well as dynamic imaging and establishment of a variety of stem cell derived cardiac organoid models, are providing new insights into the regulation of SHF development [31,83]. Current objectives include elucidation of the regulatory networks controlling distinct SHF subpopulations implicated in CHD, the identification of downstream pathways effecting regional heart morphogenesis and cardiac septation, and dissection of the disease mechanisms by which genetic and environmental factors interact in the SHF to cause CHD.

Declaration of interests

The authors do not work for, advise, own shares in, or receive funds from any organization that could benefit from this article, and have declared no affiliations other than their research organizations.

Funding

Work in RK's group is supported by the Agence National pour la Recherche (PP-heart ANR-20-CE13-0029-01 and Heartbound ANR-22-CE13 projects) and the AFM-Telethon. CG acknowledges the support of the French Government program managed by the French National Research Agency (ANR-16-CONV-0001) and the Excellence Initiative of Aix-Marseille University A*MIDEX (Turing Centre for Living Systems) and the Fondation pour la Recherche Médicale. RK is an Inserm research fellow.

References

[1] B. G. Bruneau, "The developmental genetics of congenital heart disease", *Nature* **451** (2008), p. 943-948.

[2] J. A. Epstein, H. Franklin, "Epstein lecture. Cardiac development and implications for heart disease", *New Engl. J. Med.* **363** (2010), p. 1638-1647.

[3] R. G. Kelly, M. E. Buckingham, A. F. Moorman, "Heart fields and cardiac morphogenesis", *Cold Spring Harb. Perspect. Med.* **4** (2014), article no. a015750.

[4] W. P. Devine, J. D. Wythe, M. George, K. Koshiba-Takeuchi, B. G. Bruneau, "Early patterning and specification of cardiac progenitors in gastrulating mesoderm", *Elife* **3** (2014), article no. e03848.

[5] F. Lescroart, S. Chabab, X. Lin, S. Rulands, C. Paulissen, A. Rodolosse, H. Auer, Y. Achouri, C. Dubois, A. Bondue, B. D. Simons, C. Blanpain, "Early lineage restriction in temporally distinct populations of Mesp1 progenitors during mammalian heart development", *Nat. Cell Biol.* **16** (2014), p. 829-840.

[6] K. Ivanovitch, P. Soro-Barrio, P. Chakravarty, R. A. Jones, D. M. Bell, S. N. Mousavy Gharavy, D. Stamataki, J. Delile, J. C. Smith, J. Briscoe, "Ventricular, atrial, and outflow tract heart progenitors arise from spatially and molecularly distinct regions of the primitive streak", *PLoS Biol.* **19** (2021), article no. e3001200.

[7] R. G. Kelly, N. A. Brown, M. E. Buckingham, "The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm", *Dev. Cell* **1** (2001), p. 435-440.

[8] K. L. Waldo, D. H. Kumiski, K. T. Wallis, H. A. Stadt, M. R. Hutson, D. H. Platt, M. L. Kirby, "Conotruncal myocardium arises from a secondary heart field", *Development* **128** (2001), p. 3179-3188.

[9] C. H. Mjaatvedt, T. Nakaoka, R. Moreno-Rodriguez, R. A. Norris, M. J. Kern, C. A. Eisenberg, D. Turner, R. R. Markwald, "The outflow tract of the heart is recruited from a novel heart-forming field", *Dev. Biol.* **238** (2001), p. 97-109.

[10] S. Viragh, C. E. Challice, "Origin and differentiation of cardiac muscle cells in the mouse", *J. Ultrastruct. Res.* **42** (1973), p. 1-24.

[11] M. V. de la Cruz, C. Sanchez Gomez, M. M. Arteaga, C. Arguello, "Experimental study of the development of the truncus and the conus in the chick embryo", *J. Anat.* **123** (1977), p. 661-686.

[12] S. M. Meilhac, M. Esner, R. G. Kelly, J. F. Nicolas, M. E. Buckingham, "The clonal origin of myocardial cells in different regions of the embryonic mouse heart", *Dev. Cell* **6** (2004), p. 685-698.

[13] F. Lescroart, T. Mohun, S. M. Meilhac, M. Bennett, M. Buckingham, "Lineage tree for the venous pole of the heart: clonal analysis clarifies controversial genealogy based on genetic tracing", *Circ. Res.* **111** (2012), p. 1313-1322.

[14] R. G. Kelly, "The second heart field", *Curr. Top. Dev. Biol.* **100** (2012), p. 33-65.

[15] C. L. Cai, X. Liang, Y. Shi, P. H. Chu, S. Pfaff, J. Chen, S. Evans, "Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart", *Dev. Cell* **5** (2003), p. 877-889.

[16] M. P. Verzi, D. J. McCulley, S. De Val, E. Dodou, B. L. Black, "The right ventricle, outflow tract, and ventricular septum comprise a restricted expression domain within the secondary/anterior heart field", *Dev. Biol.* **287** (2005), p. 134-145.

[17] L. Tirosh-Finkel, H. Elhanany, A. Rinon, E. Tzahor, "Mesoderm progenitor cells of common origin contribute to the head musculature and the cardiac outflow tract", *Development* **133** (2006), p. 1943-1953.

[18] Z. Zhou, J. Wang, C. Guo, W. Chang, J. Zhuang, P. Zhu, X. Li, "Temporally distinct Six2-positive second heart field progenitors regulate mammalian heart development and disease", *Cell Rep.* **18** (2017), p. 1019-1032.

[19] F. Lescroart, R. G. Kelly, J. F. Le Garrec, J. F. Nicolas, S. M. Meilhac, M. Buckingham, "Clonal analysis reveals common lineage relationships between head muscles and second heart field derivatives in the mouse embryo", *Development* **137** (2010), p. 3269-3279.

[20] F. Lescroart, W. Hamou, A. Francou, M. Theveniau-Ruissy, R. G. Kelly, M. Buckingham, "Clonal analysis reveals a common origin between nonsomite-derived neck muscles and

heart myocardium", *Proc. Natl. Acad. Sci. USA* **112** (2015), p. 1446-1451.

[21] R. Diogo, R. G. Kelly, L. Christiaen, M. Levine, J. M. Ziermann, J. L. Molnar, D. M. Noden, E. Tzahor, "A new heart for a new head in vertebrate cardiopharyngeal evolution", *Nature* **520** (2015), p. 466-473.

[22] F. Lescroart, C. E. Dumas, N. Adachi, R. G. Kelly, "Emergence of heart and branchiomeric muscles in cardiopharyngeal mesoderm", *Exp. Cell Res.* **410** (2022), article no. 112931.

[23] H. Nomaru, Y. Liu, C. De Bono, D. Righelli, A. Cirino, W. Wang, H. Song, S. E. Racedo, A. G. Dantas, L. Zhang, C. L. Cai, C. Angelini, L. Christiaen, R. G. Kelly, A. Baldini, D. Zheng, B. E. Morrow, "Single cell multi-omic analysis identifies a Tbx1-dependent multilineage primed population in murine cardiopharyngeal mesoderm", *Nat. Commun.* **12** (2021), article no. 6645.

[24] C. Trapnell, "Defining cell types and states with single-cell genomics", *Genome Res.* **25** (2015), p. 1491-1498.

[25] R. C. V. Tyser, X. Ibarra-Soria, K. McDole, S. Arcot Jayaram, J. Godwin, T. A. H. van den Brand, A. M. A. Miranda, A. Scialdone, P. J. Keller, J. C. Marioni, S. Srinivas, "Characterization of a common progenitor pool of the epicardium and myocardium", *Science* **371** (2021), article no. eabb2986.

[26] Q. Zhang, D. Carlin, F. Zhu, P. Cattaneo, T. Ideker, S. M. Evans, J. Bloomekatz, N. C. Chi, "Unveiling complexity and multipotentiality of early heart fields", *Circ. Res.* **129** (2021), p. 474-487.

[27] R. G. Kelly, "The heart field transcriptional landscape at single-cell resolution", *Dev. Cell* **58** (2023), p. 257-266.

[28] R. C. V. Tyser, X. Ibarra-Soria, M. Pedroza, A. M. A. Miranda, T. A. H. van den Brand, A. Scialdone, J. C. Marioni, S. Srinivas, "Moelcular characterization of left-right symmetry breaking in the mouse embryo", 2022, bioRxiv, https://doi.org/10.1101/2022.12.26.521947.

[29] A. Desgrange, J. F. Le Garrec, S. Bernheim, T. H. Bonnelykke, S. M. Meilhac, "Transient nodal signaling in left precursors coordinates opposed asymmetries shaping the heart loop", *Dev. Cell* **55** (2020), p. 413-431.

[30] G. C. Gabriel, C. B. Young, C. W. Lo, "Role of cilia in the pathogenesis of congenital heart disease", *Semin. Cell Dev. Biol.* **110** (2021), p. 2-10.

[31] M. H. Dominguez, A. Leigh Krup, J. M. Munice, B. G. Bruneau, "Graded mesoderm assembly governs cell fate and morphogenesis of the early mammalian heart", *Cell* **186** (2023), p. 479-496.

[32] I. Esteban, P. Schmidt, A. Desgrange, M. Raoila, S. Temino, S. M. Meilhac, L. Kobbelt, M. Torres, "Pseudodynamic analysis of heart tube formation in the mouse reveals strong regional variability and early left–right asymmetry", *Nat. Cardiovasc. Res.* **1** (2022), p. 504-517.

[33] S. A. Ramsbottom, V. Sharma, H. J. Rhee, L. Eley, H. M. Phillips, H. F. Rigby, C. Dean, B. Chaudhry, D. J. Henderson, "Vangl2-regulated polarisation of second heart field-derived cells is required for outflow tract lengthening during cardiac development", *PLoS Genet.* **10** (2014), article no. e1004871.

[34] R. Jain, D. Li, M. Gupta, L. J. Manderfield, J. L. Ifkovits, Q. Wang, F. Liu, Y. Liu, A. Poleshko, A. Padmanabhan, J. C. Raum, L. Li, E. E. Morrisey, M. M. Lu, K. J. Won, J. A. Epstein, "HEART development. Integration of Bmp and Wnt signaling by Hopx specifies commitment of cardiomyoblasts", *Science* **348** (2015), article no. aaa6071.

[35] S. O. Omer, F. M. Alhabshan, A. M. Z. Jijeh, N. C. Caimbon, C. C. Enriquez, J. Manner, T. M. Yelbuz, "Is transposition of the great arteries associated with shortening of the intrapericardial portions of the great arterial trunks? An echocardiographic analysis on newborn infants with simple transposition of the great arteries to explore an animal model-based hypothesis on human beings", *J. Am. Heart Assoc.* **10** (2021), article no. e019334.

[36] F. Rochais, K. Mesbah, R. G. Kelly, "Signaling pathways controlling second heart field development", *Circ. Res.* **104** (2009), p. 933-942.

[37] L. A. Dyer, M. L. Kirby, "The role of secondary heart field in cardiac development", *Dev. Biol.* **336** (2009), p. 137-144.

[38] C. Cortes, A. Francou, C. De Bono, R. G. Kelly, "Epithelial properties of the second heart field", *Circ. Res.* **122** (2018), p. 142-154.

[39] M. Rowton, A. Guzzetta, A. B. Rydeen, I. P. Moskowitz, "Control of cardiomyocyte differentiation timing by intercellular signaling pathways", *Semin. Cell Dev. Biol.* **118** (2021), p. 94-106.

[40] A. Baldini, F. G. Fulcoli, E. Illingworth, "Tbx1: transcriptional and developmental functions", *Curr. Top. Dev. Biol.* **122** (2017), p. 223-243.

[41] Y. Watanabe, S. Miyagawa-Tomita, S. D. Vincent, R. G. Kelly, A. M. Moon, M. E. Buckingham, "Role of mesodermal FGF8 and FGF10 overlaps in the development of the arterial pole of the heart and pharyngeal arch arteries", *Circ. Res.* **106** (2010), p. 495-503.

[42] Y. Watanabe, S. Zaffran, A. Kuroiwa, H. Higuchi, T. Ogura, R. P. Harvey, R. G. Kelly, M. Buckingham, "Fibroblast growth factor 10 gene regulation in the second heart field by Tbx1, Nkx2-5, and Islet1 reveals a genetic switch for down-regulation in the myocardium", *Proc. Natl. Acad. Sci. USA* **109** (2012), p. 18273-18280.

[43] A. Francou, E. Saint-Michel, K. Mesbah, R. G. Kelly, "TBX1 regulates epithelial polarity and dynamic basal filopodia in the second heart field", *Development* **141** (2014), p. 4320-4331.

[44] L. Chen, F. G. Fulcoli, R. Ferrentino, S. Martucciello, E. A. Illingworth, A. Baldini, "Transcriptional control in cardiac progenitors: Tbx1 interacts with the BAF chromatin remodeling complex and regulates Wnt5a", *PLoS Genet.* **8** (2012), article no. e1002571.

[45] D. Li, T. Sinha, R. Ajima, H. S. Seo, T. P. Yamaguchi, J. Wang, "Spatial regulation of cell cohesion by Wnt5a during second heart field progenitor deployment", *Dev. Biol.* **412** (2016), p. 18-31.

[46] A. Francou, C. De Bono, R. G. Kelly, "Epithelial tension in the second heart field promotes mouse heart tube elongation", *Nat. Commun.* **8** (2017), article no. 14770.

[47] M. L. Kirby, K. L. Waldo, "Role of neural crest in congenital heart disease", *Circulation* **82** (1990), p. 332-340.

[48] T. M. Yelbuz, K. L. Waldo, D. H. Kumiski, H. A. Stadt, R. R. Wolfe, L. Leatherbury, M. L. Kirby, "Shortened outflow tract leads to altered cardiac looping after neural crest ablation", *Circulation* **106** (2002), p. 504-510.

[49] T. M. Yelbuz, K. L. Waldo, X. Zhang, M. Zdanowicz, J. Parker,

T. L. Creazzo, G. A. Johnson, M. L. Kirby, "Myocardial volume and organization are changed by failure of addition of secondary heart field myocardium to the cardiac outflow tract", *Dev. Dyn.* **228** (2003), p. 152-160.

[50] K. Momma, "Cardiovascular anomalies associated with chromosome 22q11.2 deletion syndrome", *Am. J. Cardiol.* **105** (2010), p. 1617-1624.

[51] M. Unolt, P. Versacci, S. Anaclerio, C. Lambiase, G. Calcagni, M. Trezzi, A. Carotti, T. B. Crowley, E. H. Zackai, E. Goldmuntz, J. W. Gaynor, M. C. Digilio, D. M. McDonald-McGinn, B. Marino, "Congenital heart diseases and cardiovascular abnormalities in 22q11.2 deletion syndrome: from well-established knowledge to new frontiers", *Am. J. Med. Genet. A* **176** (2018), p. 2087-2098.

[52] M. Theveniau-Ruissy, M. Dandonneau, K. Mesbah, O. Ghez, M. G. Mattei, L. Miquerol, R. G. Kelly, "The del22q11.2 candidate gene Tbx1 controls regional outflow tract identity and coronary artery patterning", *Circ. Res.* **103** (2008), p. 142-148.

[53] M. S. Rana, M. Theveniau-Ruissy, C. De Bono, K. Mesbah, A. Francou, M. Rammah, J. N. Dominguez, M. Roux, B. Laforest, R. H. Anderson, T. Mohun, S. Zaffran, V. M. Christoffels, R. G. Kelly, "Tbx1 coordinates addition of posterior second heart field progenitor cells to the arterial and venous poles of the heart", *Circ. Res.* **115** (2014), p. 790-799.

[54] R. Van Praagh, "The first Stella van Praagh memorial lecture: the history and anatomy of tetralogy of Fallot", *Semin. Thorac. Cardiovasc. Surg. Pediatr. Card. Surg. Annu.* (2009), p. 19-38.

[55] C. Ward, H. Stadt, M. Hutson, M. L. Kirby, "Ablation of the secondary heart field leads to tetralogy of Fallot and pulmonary atresia", *Dev. Biol.* **284** (2005), p. 72-83.

[56] F. Bajolle, S. Zaffran, S. M. Meilhac, M. Dandonneau, T. Chang, R. G. Kelly, M. E. Buckingham, "Myocardium at the base of the aorta and pulmonary trunk is prefigured in the outflow tract of the heart and in subdomains of the second heart field", *Dev. Biol.* **313** (2008), p. 25-34.

[57] M. Rammah, M. Theveniau-Ruissy, R. Sturny, F. Rochais, R. G. Kelly, "PPARgamma and NOTCH regulate regional identity in the murine cardiac outflow tract", *Circ. Res.* **131** (2022), p. 842-858.

[58] F. Bajolle, S. Zaffran, R. G. Kelly, J. Hadchouel, D. Bonnet, N. A. Brown, M. E. Buckingham, "Rotation of the myocardial wall of the outflow tract is implicated in the normal positioning of the great arteries", *Circ. Res.* **98** (2006), p. 421-428.

[59] R. W. Scherptong, M. R. Jongbloed, L. J. Wisse, R. Vicente-Steijn, M. M. Bartelings, R. E. Poelmann, M. J. Schalij, A. C. Gittenberger-De Groot, "Morphogenesis of outflow tract rotation during cardiac development: the pulmonary push concept", *Dev. Dyn.* **241** (2012), p. 1413-1422.

[60] S. U. Morton, D. Quiat, J. G. Seidman, C. E. Seidman, "Genomic frontiers in congenital heart disease", *Nat. Rev. Cardiol.* **19** (2022), p. 26-42.

[61] A. Stathopoulou, P. Wang, C. Thellier, R. G. Kelly, D. Zheng, P. J. Scambler, "Charge syndrome-associated CHD7 acts at ISL1-regulated enhancers to modulate second heart field gene expression", *Cardiovasc. Res.* **119** (2023), p. 2089-2105.

[62] C. S. M. Tang, M. Mononen, W. Y. Lam, S. C. Jin, X. Zhuang, M. M. Garcia-Barcelo, Q. Lin, Y. Yang, M. Sahara, E. Eroglu, K. R. Chien, H. Hong, P. K. H. Tam, P. J. Gruber, "Sequencing of a Chinese tetralogy of Fallot cohort reveals clustering mutations in myogenic heart progenitors", *JCI Insight* **7** (2022), article no. e152198.

[63] H. Shi, V. C. O'Reilly, J. L. Moreau, T. R. Bewes, M. X. Yam, B. E. Chapman, S. M. Grieve, R. Stocker, R. M. Graham, G. Chapman, D. B. Sparrow, S. L. Dunwoodie, "Gestational stress induces the unfolded protein response, resulting in heart defects", *Development* **143** (2016), p. 2561-2572.

[64] J. I. Kalisch-Smith, N. Ved, D. Szumska, J. Munro, M. Troup, S. E. Harris, H. Rodriguez-Caro, A. Jacquemot, J. J. Miller, E. M. Stuart, M. Wolna, E. Hardman, F. Prin, E. Lana-Elola, R. Aoidi, E. M. C. Fisher, V. L. J. Tybulewicz, T. J. Mohun, S. Lakhal-Littleton, S. De Val, E. Giannoulatou, D. B. Sparrow, "Maternal iron deficiency perturbs embryonic cardiovascular development in mice", *Nat. Commun.* **12** (2021), article no. 3447.

[65] D. C. Harvey, P. De Zoysa, O. Toubat, J. Choi, J. Kishore, H. Tsukamoto, S. R. Kumar, "Concomitant genetic defects potentiate the adverse impact of prenatal alcohol exposure on cardiac outflow tract maturation", *Birth Defects Res.* **114** (2022), p. 105-115.

[66] N. Bertrand, M. Roux, L. Ryckebusch, K. Niederreither, P. Dolle, A. Moon, M. Capecchi, S. Zaffran, "Hox genes define distinct progenitor sub-domains within the second heart field", *Dev. Biol.* **353** (2011), p. 266-274.

[67] S. Stefanovic, B. Laforest, J. P. Desvignes, F. Lescroart, L. Argiro, C. Maurel-Zaffran, D. Salgado, E. Plaindoux, C. De Bono, K. Pazur, M. Theveniau-Ruissy, C. Beroud, M. Puceat, A. Gavalas, R. G. Kelly, S. Zaffran, "Hox-dependent coordination of mouse cardiac progenitor cell patterning and differentiation", *Elife* **9** (2020), article no. e55124.

[68] B. S. Snarr, J. L. O'Neal, M. R. Chintalapudi, E. E. Wirrig, A. L. Phelps, S. W. Kubalak, A. Wessels, "Isl1 expression at the venous pole identifies a novel role for the second heart field in cardiac development", *Circ. Res.* **101** (2007), p. 971-974.

[69] M. M. Goddeeris, S. Rho, A. Petiet, C. L. Davenport, G. A. Johnson, E. N. Meyers, J. Klingensmith, "Intracardiac septation requires hedgehog-dependent cellular contributions from outside the heart", *Development* **135** (2008), p. 1887-1895.

[70] D. Li, A. Angermeier, J. Wang, "Planar cell polarity signaling regulates polarized second heart field morphogenesis to promote both arterial and venous pole septation", *Development* **146** (2019), article no. dev181719.

[71] L. E. Briggs, J. Kakarla, A. Wessels, "The pathogenesis of atrial and atrioventricular septal defects with special emphasis on the role of the dorsal mesenchymal protrusion", *Differentiation: Res. Biol. Diversity* **84** (2012), p. 117-130.

[72] B. G. Bruneau, G. Nemer, J. P. Schmitt, F. Charron, L. Robitaille, S. Caron, D. A. Conner, M. Gessler, M. Nemer, C. E. Seidman, J. G. Seidman, "A murine model of Holt-Oram syndrome defines roles of the T-box transcription factor Tbx5 in cardiogenesis and disease", *Cell* **106** (2001), p. 709-721.

[73] C. De Bono, C. Thellier, N. Bertrand, R. Sturny, E. Jullian, C. Cortes, S. Stefanovic, S. Zaffran, M. Theveniau-Ruissy, R. G. Kelly, "T-box genes and retinoic acid signaling regulate the segregation of arterial and venous pole progenitor cells in the murine second heart field", *Hum. Mol. Genet.* **27** (2018), p. 3747-3760.

[74] J. D. Steimle, I. P. Moskowitz, "TBX5: a key regulator of heart development", *Curr. Top. Dev. Biol.* **122** (2017), p. 195-221.

[75] S. A. Rankin, J. D. Steimle, X. H. Yang, A. B. Rydeen, K. Agar-

wal, P. Chaturvedi, K. Ikegami, M. J. Herriges, I. P. Moskowitz, A. M. Zorn, "Tbx5 drives Aldh1a2 expression to regulate a RA-Hedgehog-Wnt gene regulatory network coordinating cardiopulmonary development", *Elife* **10** (2021), article no. e69288.

[76] L. Xie, A. D. Hoffmann, O. Burnicka-Turek, J. M. Friedland-Little, K. Zhang, I. P. Moskowitz, "Tbx5-hedgehog molecular networks are essential in the second heart field for atrial septation", *Dev. Cell* **23** (2012), p. 280-291.

[77] T. Peng, Y. Tian, C. J. Boogerd, M. M. Lu, R. S. Kadzik, K. M. Stewart, S. M. Evans, E. E. Morrisey, "Coordination of heart and lung co-development by a multipotent cardiopulmonary progenitor", *Nature* **500** (2013), p. 589-592.

[78] J. D. Steimle, S. A. Rankin, C. E. Slagle, J. Bekeny, A. B. Rydeen, S. S. Chan, J. Kweon, X. H. Yang, K. Ikegami, R. D. Nadadur, M. Rowton, A. D. Hoffmann, S. Lazarevic, W. Thomas, E. A. T. Boyle Anderson, M. E. Horb, L. Luna-Zurita, R. K. Ho, M. Kyba, B. Jensen, A. M. Zorn, F. L. Conlon, I. P. Moskowitz, "Evolutionarily conserved Tbx5-Wnt2/2b pathway orchestrates cardiopulmonary development", *Proc. Natl. Acad. Sci. USA* **115** (2018), p. E10615-E10624.

[79] M. Rowton, C. Perez-Cervantes, S. Hur, J. Jacobs-Li, E. Lu, N. Deng, A. Guzzetta, A. D. Hoffmann, M. Stocker, J. D. Steimle, S. Lazarevic, S. Oubaha, X. H. Yang, C. Kim, S. Yu, H. Eckart, M. Koska, E. Hanson, S. S. K. Chan, D. J. Garry, M. Kyba, A. Basu, K. Ikegami, S. Pott, I. P. Moskowitz, "Hedgehog signaling activates a mammalian heterochronic gene regulatory network controlling differentiation timing across lineages", *Dev. Cell* **57** (2022), p. 2181-2203.

[80] K. Koshiba-Takeuchi, A. D. Mori, B. L. Kaynak, J. Cebra-Thomas, T. Sukonnik, R. O. Georges, S. Latham, L. Beck, R. M. Henkelman, B. L. Black, E. N. Olson, J. Wade, J. K. Takeuchi, M. Nemer, S. F. Gilbert, B. G. Bruneau, "Reptilian heart development and the molecular basis of cardiac chamber evolution", *Nature* **461** (2009), p. 95-98.

[81] I. S. Kathiriya, K. S. Rao, G. Iacono, W. P. Devine, A. P. Blair, S. K. Hota, M. H. Lai, B. I. Garay, R. Thomas, H. Z. Gong, L. K. Wasson, P. Goyal, T. Sukonnik, K. M. Hu, G. A. Akgun, L. D. Bernard, B. N. Akerberg, F. Gu, K. Li, M. L. Speir, M. Haeussler, W. T. Pu, J. M. Stuart, C. E. Seidman, J. G. Seidman, H. Heyn, B. G. Bruneau, "Modeling human TBX5 haploinsufficiency predicts regulatory networks for congenital heart disease", *Dev. Cell* **56** (2021), p. 292-309.

[82] Y. Liu, S. Chen, L. Zuhlke, G. C. Black, M. K. Choy, N. Li, B. D. Keavney, "Global birth prevalence of congenital heart defects 1970–2017: updated systematic review and meta-analysis of 260 studies", *Int. J. Epidemiol.* **48** (2019), p. 455-463.

[83] G. Rossi, N. Broguiere, M. Miyamoto, A. Boni, R. Guiet, M. Girgin, R. G. Kelly, C. Kwon, M. P. Lutolf, "Capturing cardiogenesis in gastruloids", *Cell. Stem Cell* **28** (2021), p. 230-240, e236.

List of figures

1.Introduction

Figure 1.1 - Example of the repetition of a shape in different disciplines……..10
Figure 1.2 - Morphology and architecture of the adult heart……12
Figure 1.3 - Cardiac progenitors patterning during mouse heart development ……..14
Figure 1.4 - Overview of heart development ……..17
Figure 1.5 - SHF development and heart tube development……21
Figure 1.6 - pathways controlling proliferation and differentation of SHF cells and gene expression patterning of the SHF …… 22
Figure 1.7 - Arterial & venous pole development……24
Figure 1.8 - The heart fields interface and cardiac septation …… 26
Figure 1.9 - The cardiac origami ……27
Figure 1.10 - The transmission of forces. Example of the TRUMPF bridge and the dorsal pericardial wall ……27
Figure 1.11 - The spatial heterogeneity of the SHF cells ……30
Figure 1.12 - The biomechanical model driving heart tube elongation ……31
Figure 1.13 - Epithelial tension in the DPW accompanies heart tube elongation …… 31
Figure 1.14 - Actin-Myosin Network Dynamics and Force Generation …… 33
Figure 1.15 - Similarities between the formation of the neural and the heart tube……35
Figure 1.16 - Mechanical forces influence cell division and proliferation……37
Figure 1.17 - Force inference scheme based on the assumption that the forces are balanced at each vertex……39
Figure 1.18 - Examples of the first's graphical and statistical morphometric analysis……42
Figure 1.19 - Comparison between unsupervised and unsupervised learning ……44
Figure 1.20 - Pipeline for a data scientist analysis……46
Figure 1.21 - Comparison of different clusterin algorithms on toy datasets……47
Figure 1.22 - Unsupervised k-means clustering ……48
Figure 1.23 - Unsupervised Gaussian mixture clustering ……50
Figure 1.24 - Comparison of the dimensionality reduction algorithms PCA, t-SNE, UMAP and PHATE in different datasets ……52
Figure 25 - Examples of studies using clustering algorithms to understand data heterogeneity in space ……53
Figure 26 - Examples of studies using clustering algorithms to understand data heterogeneity in space …… 54
Figure 27 - The knowledge gap …… 55

Table 1.1- Summary of congenital heart defects with their prevalence, the morphological defects associated and the genes involved…… 15
Table 1.2 - Comparison between dimensionality reduction algorithms …… 48

2.Materials & methods

Table 2.1 - PCR primers for genotyping……..57
Table 2.2 - PCR programs for different primers …… 57
Table 2.3 - Primary, secondary antibodies and phalloidin used for immunostaining... 59

3.Results

Figure 3.1 **-** Examples of different cell morphologies ……..63

3.1 Single-cell morphometrics reveals T-box dependent patterns of epithelial tension in the Second Heart Field

Main figures

Figure 1 - Spatial single-cell morphometrics highlights cell size, orientation and stress direction as the main parameters defining apical cell morphology in the Dorsal Pericardial Wall……95
Figure 2 - Figure 2. Spatial single-cell morphometrics define regions in the DPW based on morphological and mechanical cell properties…… 96
Figure 3 - Spatial single-cell morphometrics reveals that patterned tension in the second heart field emerges during heart tube formation……97
Figure 4 - Clusters defined by spatial single-cell morphometrics correlate with the distribution of T-box transcription factors in the dorsal pericardial wall ……98
Figure 5 - T-box transcription factors *Tbx1* and *Tbx5* pattern apical cell architecture in the Dorsal Pericardial Wall ……99
Figure 6 - *Tbx5* induces stress anisotropy and orients the cells parallel to the antero-posterior Dorsal Pericardial Wall …… 100
Figure 7- Deletion of *Tbx5* disrupts the cellular localization of diphosphorylated myosin light chain in the Dorsal Pericardial Wall……101

Supplementary figures …… 107-126

3.2 Genetic boundary formation in the Second Heart Field

Figure 3.2.1 - Single-cell RNAseq analysis of the Second Heart Field cells ……128
Figure 3.2.2 - Validation of boundary genes ……131
Figure 3.2.3 - Quantification of the homotypic and heterotypic tensions in the SHF....132

3.3 Cell intercalation and rosette formation in the SHF

Figure 3.3.1 - Multicellular rosette formation in intercalating cells and axis elongation……133
Figure 3.3.2 - Rosette quantification according to different bond length cutoff ……136
Figure 3.3.3 - Rosette quantifications …… 137

3.4 Exploring additional data analysis

Figure 3.4.1 - Confusion matrix 138
Figure 3.4.2 - Feature selection and confusion matrix of random forest and XGBoost140

Table 3.4.1 - Performance metrics obtained with different models considering the spatial positions and the morphological features of the cells 139
Table 3.4.2 - Performance metrics obtained with different models, not considering the spatial coordinates of the cells 139

Discussion
Figure 4.1 - Perspectives in the cardiac field fed by the discoveries in heart development.....146
Figure 4.2 - Data science pipeline incorporated in biological research 148
Figure 4.3 - Map and trajectories of the Second Heart Field cells..... 150
Figure 4.4 - Evidence that *Tbx1-Tbx5* boundary is a compartment boundary 155
Figure 4.5 - Final model 158

Annex - Review - On the involvement of the Second Heart Field in congenital heart defects.....190-200

List of abbreviations

A: Atria
Ao: Aorta
A-P: Antero- Posterior axes
AP: Arterial Pole
ASD: Atrial Septal Defect
aSHF: aSHF
AV: AtrioVentricular
BMP: Bone Morphogenetic Protein
CHD: Congenital Heart defect
CPM: Cardiopharyngeal Mesoderm
DM: Dorsal Mesocardium
DMP: Dorsal Mesenchymal Protrusion
DPW: Dorsal Pericardial Wall
D-V: Dorso-Ventral axes
DVE: Distal Visceral Endoderm
ECM: Extracellular matrix
FGF: Fibroblast Growth Factor
FHF: First Heart Field
IFT: Inflow Tract
JCF: Juxta Cardiac Field
LA: Left Atria
L-R: Left-Right axes
LV: Left Ventricle
MET: Mesenchymal-to-epithelial transition
MUT: Mutant
NT: Neural Tube
OFT: Outflow Tract
PAS: Pulmonary Atrial Septum
PCA: Principal Component Analysis
PCP: Planar Cell Polarity
ppMLC: di-phosphorylated Myosin Light Chain
pSHF: pSHF
PT: Pulmonary Trunk
RA: Retinoic Acid
RA: Right Atria
RV: Right Ventricle
SA: SinoAtrial
sAS: secondary Atrial Septum
SHF: Second Heart Field
TAZ: Transcription coactivator with PDZ-binding motif
Tbx: T-box transcription Factor

TOF: Tetrallogy Of Fallot
t-SNE: t-distributed Stochastic Neighbor Embedding
TZ: Transition Zone
UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction
V : Ventricle
VP : Venous Pole
VS : Ventricular Septa
VSD: Ventricular Septal Defect
WNT: Wingless-related Integration Site
WT: Wild-type
YAP: Yes-Associated Protein