

GENOMIC ANALYSIS OF EMBRYONIC HEART DEVELOPMENT IN THE MOUSE

by

PAVLE JOSIP VRLJIČAK

B.Sc., McGill University, 2000

M.Sc., McGill University, 2003

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

THE FACULTY OF GRADUATE STUDIES

(Medical Genetics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

September 2010

© Pavle Josip Vrljičak, 2010

ABSTRACT

Malformations of the cardiovascular system are the most common type of birth defect in humans, affecting predominantly the formation of valves and septa. While many studies have addressed the role of specific genes during valve and septa formation, a global understanding is still largely incomplete. To address this deficit we have undertaken a genome-wide transcriptional profiling of the developing heart in the mouse. We generated and analyzed 19 Serial Analysis of Gene Expression (SAGE) libraries representing different regions of the mouse heart at multiple stages of embryonic development.

We speculated that genes important for heart valve development would be differentially expressed in the valve forming regions, and have dynamic temporal expression patterns. We used our dataset to identify a novel list of valve enriched genes. Using k-means cluster analysis we also uncovered 14 distinct temporal gene expression patterns in the developing valves. Unique temporal expression patterns were found to be enriched for specific signalling pathway members and functional categories such as signal transduction, transcription factor activity, proliferation and apoptosis.

The most highly expressed transcription factor within the developing valves was found to be *Twist1*. Analysis of gene expression changes in the *Twist1* null developing valves revealed a novel phenotype consistent with a role of TWIST1 in controlling differentiation of mesenchymal cells following their transformation from endothelium in the mouse. Our data suggests that TWIST1 directly activates valve specific and cell motility gene expression in the atrio-ventricular canal, while suppressing expression of valve maturation markers.

This work provides the first comprehensive temporal and spatial gene expression dataset for heart development during formation of the heart valves. It is a valuable resource for the elucidation of the molecular mechanisms underlying heart development.

TABLE OF CONTENTS

| | |
|--|------|
| ABSTRACT..... | ii |
| TABLE OF CONTENTS..... | iv |
| LIST OF TABLES..... | vi |
| LIST OF FIGURES..... | vii |
| LIST OF ABBREVIATIONS..... | viii |
| ACKNOWLEDGEMENTS..... | ix |
| CO-AUTHORSHIP STATEMENT..... | xi |
| CHAPTER 1 | 1 |
| INTRODUCTION | 1 |
| 1.1. Congenital heart defects..... | 1 |
| 1.1.1. Prevalence of congenital heart defects in the human population..... | 1 |
| 1.1.2. Genetic determinants of congenital heart defects in humans..... | 4 |
| 1.1.3. Animal models of cardiac disease..... | 8 |
| 1.2. Cardiac valve development..... | 10 |
| 1.2.1. Early heart morphogenesis..... | 10 |
| 1.2.2. Valve and septa formation..... | 13 |
| 1.2.3. Valve remodeling..... | 17 |
| 1.3. Signalling pathways in AVC development..... | 19 |
| 1.3.1. Notch signalling pathway..... | 20 |
| 1.3.2. BMP and TGF β signalling pathways..... | 23 |
| 1.4. Role of transcription factors in AVC development..... | 27 |
| 1.4.1. Twist family..... | 29 |
| 1.4.2. Role of Twist family members in development and cancer..... | 32 |
| 1.4.3. Twist family activity in AVC and OFT development..... | 33 |
| 1.5. Transcriptome analysis..... | 34 |
| 1.6. Project goals..... | 39 |
| 1.7. References..... | 41 |
| CHAPTER 2 | 51 |
| GENOMIC ANALYSIS DISTINGUISHES PHASES OF EARLY DEVELOPMENT OF THE MOUSE ATRIO-VENTRICULAR CANAL | 51 |
| 2.1. Introduction..... | 51 |
| 2.2. Materials and methods..... | 53 |
| 2.2.1. Serial analysis of gene expression..... | 53 |
| 2.2.2. Cluster analysis..... | 54 |
| 2.2.3. Mesenchyme-epithelial score..... | 54 |
| 2.3. Results..... | 55 |
| 2.3.1. Clustering of AVC libraries reveals temporal expression patterns..... | 58 |
| 2.3.2. Epithelial/endothelial to mesenchymal transformation in the AVC..... | 64 |
| 2.3.3. Downstream targets of TGF β and Notch signalling in the AVC..... | 67 |
| 2.4. Discussion..... | 72 |
| 2.5. References..... | 76 |

| | |
|--|-----|
| CHAPTER 3 | 79 |
| TWIST1 CONTROLS DIFFERENTIATION OF DEVELOPING ATRIO- VENTRICULAR CANAL CELLS IN THE MOUSE | 79 |
| 3.1. Introduction..... | 79 |
| 3.2. Materials and methods..... | 81 |
| 3.2.1. Tissue collection..... | 81 |
| 3.2.2. Gene expression analysis..... | 82 |
| 3.2.3. Tissue culture..... | 82 |
| 3.2.4. Chromatin immunoprecipitation..... | 83 |
| 3.3. Results..... | 83 |
| 3.3.1. AVC and OFT share significant gene expression..... | 85 |
| 3.3.2. <i>Twist1</i> null AVC shows dramatic gene expression changes..... | 92 |
| 3.3.3. TWIST1 directly regulates AVC gene expression..... | 98 |
| 3.4. Discussion..... | 101 |
| 3.4.1. Identification of AVC- and OFT-enriched genes..... | 101 |
| 3.4.2. Regulation of spatial and temporal expression patterns..... | 102 |
| 3.5. References..... | 105 |
| CHAPTER 4 | 111 |
| SUMMARY, PERSPECTIVES AND FUTURE DIRECTIONS | 111 |
| 4.1. A comprehensive gene expression resource for the study of embryonic heart development..... | 111 |
| 4.2. Study of temporal and spatial restricted gene expression during valve development..... | 117 |
| 4.2.1. Identification of temporal expression patterns during AVC development..... | 118 |
| 4.2.2. Identification of endocardial cushion enriched genes..... | 120 |
| 4.3. Role of TWIST1 in heart development..... | 122 |
| 4.3.1. Direct regulation of transcription by TWIST1..... | 122 |
| 4.3.2. Functional redundancy and compensatory mechanisms for <i>Twist1</i> deletion..... | 124 |
| 4.4. Research impact and future work..... | 135 |
| 4.5. References..... | 137 |
| APPENDICES | 141 |
| Appendix I. PCR primers..... | 141 |
| Appendix II. Epithelium and mesenchymal library pairs in Mouse Atlas database..... | 142 |
| Appendix III. Embryonic SAGE libraries in Mouse Atlas database..... | 143 |
| Appendix IV. Heart development gene expression in AVC libraries..... | 144 |
| Appendix V. Expression of AVC-enriched genes in heart libraries..... | 145 |
| Appendix VI. ChIP-qPCR primers..... | 146 |
| Appendix VII. Expression of known AVC genes in E10.5 heart Tag-seq libraries..... | 147 |
| Appendix VIII. Heart regions sampled for gene profiling..... | 148 |
| Appendix IX. Validation of AVC- and OFT-specific gene expression..... | 149 |
| Appendix X. Differentially expressed genes in <i>Twist1</i> null AVC..... | 150 |
| Appendix XI. <i>Twist1</i> null hearts are smaller than wild-type..... | 156 |
| Appendix XII. Differentially expressed genes in <i>Twist1</i> null AVC..... | 157 |
| Appendix XIII. Chromatin immunoprecipitation..... | 160 |
| Appendix XIV. Animal care certificate..... | 161 |

LIST OF TABLES

| | |
|---|-----|
| Table 1.1. Prevalence of birth defects associated with infant death in human population | 2 |
| Table 1.2. Prevalence of specific heart defects in human population..... | 3 |
| Table 1.3. Representative environmental determinants of congenital heart defects..... | 5 |
| Table 1.4. Representative chromosomal disorders associated with congenital heart defects..... | 6 |
| Table 1.5. Selected genes associated with congenital heart defects in the young | 7 |
| Table 1.6. Selected transcriptional regulators controlling valve morphogenesis | 28 |
| Table 2.1. Tissues and stages sampled..... | 56 |
| Table 2.2. Summary of AVC temporal clusters..... | 61 |
| Table 2.3. Co-expressed genes share specific functions..... | 62 |
| Table 3.1. Overview of Tag-seq libraries | 84 |
| Table 3.2. Selected genes enriched in the E10.5 AVC and OFT..... | 88 |
| Table 3.3. Selected gene expression changes in <i>Twist1</i> null AVC..... | 94 |
| Table 4.1. Mapping efficiency of tag-types in Tag-seq data | 113 |
| Table 4.2. Expression of Twist family members during valve development | 129 |
| Table 4.3. <i>Twist1</i> mutant and <i>Snai2</i> mutant crosses | 134 |

LIST OF FIGURES

| | |
|---|-----|
| Figure 1.1. Early heart morphogenesis | 11 |
| Figure 1.2. Endocardial cushion formation..... | 15 |
| Figure 1.3. Epithelial-to-mesenchymal transformation | 16 |
| Figure 1.4. AVC valve remodeling..... | 18 |
| Figure 1.5. Notch pathway..... | 21 |
| Figure 1.6. BMP and TGF β pathways | 25 |
| Figure 1.7. Twist family homology | 31 |
| Figure 1.8. Serial analysis of gene expression (SAGE)..... | 37 |
| Figure 2.1. Cluster analysis of AVC libraries reveals 14 distinct temporal expression patterns.. | 60 |
| Figure 2.2. Analysis of gene expression by RT-qPCR and <i>in situ</i> hybridization | 63 |
| Figure 2.3. Cluster analysis reveals temporal sequence of EMT in the AVC | 66 |
| Figure 2.4. Multiple distinct phases of TGF β and Notch pathway activity exist in the AVC..... | 70 |
| Figure 3.1. AVC and OFT shared gene expression | 86 |
| Figure 3.2. Altered gene expression in <i>Twist1</i> null AVC..... | 97 |
| Figure 3.3. Chromatin immunoprecipitation | 100 |
| Figure 4.1. <i>Twist1</i> null embryos show normal EMT | 126 |
| Figure 4.2. AVC explant assay shows no defect in cell migration | 127 |
| Figure 4.3. <i>Twist1</i> and <i>Twist2</i> are expressed in mesenchyme cells of AVC and OFT..... | 130 |
| Figure 4.4. TWIST1 and TWIST2 bind similar promoter regions | 131 |
| Figure 4.5. <i>Twist1</i> and <i>Twist2</i> show distinct temporal expression patterns in the AVC | 132 |

LIST OF ABBREVIATIONS

AVC - Atrio-ventricular canal

BHLH - Basic helix loop helix

BMP - Bone morphogenetic protein

CSL - CBF1, suppressor of hairless, Lag-1

ChIP - Chromatin immunoprecipitation

E - Embryonic day

ECM - Extracellular matrix

EMT - Epithelial-to-mesenchymal transformation

FACS - Fluorescence-activated cell sorting

GFP - Green fluorescent protein

GO - Gene ontology

HQ - High quality

MGC - Mammalian gene collection

MIMCD - Mouse inner medullary collecting duct

MORGEN - Mammalian organogenesis – regulation by gene expression networks

OFT - Outflow tract

PCR - Polymerase chain reaction

PHF - Primary heart field

QPCR - Quantitative PCR

RefSeq - Reference sequence

RT-qPCR - Reverse transcription followed by qPCR

SAGE - Serial analysis of gene expression

SHF - Secondary heart field

SVEC - Small vessel endothelial cell

TGF β - Transforming growth factor β

UTR - Untranslated region

ACKNOWLEDGEMENTS

I would like to thank my family, my parents Joza and Elena, and my brother Tomislav for their support over the years. Thank you for being a model of hard work and perseverance and for inculcating in me the love of inquiry. A special thank you goes to Celine Lee for her constant support and encouragement. Thank you each one of you for your faith, hope and love.

I would like to thank my supervisor Pamela Hoodless for taking me into her lab and giving me support and mentoring over the last few years. I have learned a lot about the importance of being focused and of telling a good story. I would also like to thank Aly Karsan and the members of my supervisory committee Liz Conibear, Dixie Mager, and Louis Lefebvre. Your feedback was invaluable in keeping this project grounded.

To all members of the lab, thank you for creating an enjoyable working environment and for feedback. I would like to particularly thank Rebecca Cullum for stimulating discussions and critical review of my work, but most importantly for her friendship over the years. Thank you also to Juan Hou for bringing a mature perspective. To my fellow students, Sam Lee, Ali Hassan, Lynn Mar, Kristen McKnight, thank you for sharing in the joys and pains of grad school. For immense help with experiments I would like to thank the research students working under my supervision: Carol Aiga and Eric Xu. Many thanks also to my collaborators on the heart project Elizabeth Wederell and Victoria Garside, as well as all other members of the lab in the order in which I met them, James Rupert, Rachel Montpetit, Mona Wu, Robin Dickinson, Anita Charters, Lisa Lee, Nicole Kreutz, Wei Wei, and Olivia Alder. Each one collaborated in ways great and small.

I am grateful for collaboration with Aly Karsan's lab, particularly with Alex Chang and Kyle Niessen with whom I exchanged techniques and ideas. Many thanks also to members of the Mouse Atlas and MORGEN teams, particularly Amanda Kotzer. Finally, I would like to thank the St. Mark's and RCAV communities. Your faith and prayers have kept me going.

Financial support during my studies was provided by University of British Columbia graduate fellowships, and the Heart and Stroke foundation.

CO-AUTHORSHIP STATEMENT

Chapter 2

Pavle Vrljicak, Alex C.Y. Chang, Olena Morozova, Elizabeth D. Wederell, Kyle Niessen, Marco A. Marra, Aly Karsan, and Pamela A. Hoodless (2010) Genomic analysis distinguishes phases of early development of the mouse atrio-ventricular canal. *Physiological Genomics* 40(3):150-157.

PV and PH were responsible for the experimental design and manuscript preparation. PV, AC, EW and KN were involved in tissue collection for SAGE library construction. SAGE libraries were constructed at Canada's Michael Smith Genome Sciences Centre. PV and OM performed the cluster analysis shown in Figure 2.1 and Table 2.2. PV was responsible for Gene Ontology analysis shown in Table 2.3, RT-qPCR and *in situ* hybridization shown in Figure 2.2, and data analysis shown in Figures 2.3 and 2.4. AC provided feedback during analysis and interpretation of the data. AK and MM provided critical review of the manuscript.

Chapter 3

Pavle Vrljicak, Eric Xu, Alex C.Y. Chang, Elizabeth D. Wederell, Marco A. Marra, Aly Karsan, and Pamela A. Hoodless (2010) TWIST1 controls differentiation of developing atrio-ventricular canal cells in the mouse. *To be submitted for publication.*

PV and PH were responsible for the experimental design and manuscript preparation. PV was responsible for data analysis and interpretation shown in Tables 3.2 and 3.3, and Figures 3.1 and 3.2, as well as ChIP-qPCR shown in Figure 3.3. PV was also responsible for maintenance of *Twist1* null mouse colony. As a research student working under PV's supervision, EX collaborated in validation of gene expression patterns by collecting tissue and performing RT-qPCR shown in Appendix IX. PV and EW were involved in tissue collection for Tag-seq library construction. Tag-seq libraries were constructed at Canada's Michael Smith Genome Sciences Centre. AC, AK and MM provided critical review of the manuscript.

CHAPTER 1

INTRODUCTION

1.1. Congenital heart defects

1.1.1. Prevalence of congenital heart defects in the human population

Malformations of the cardiovascular system are the most common type of birth defect in humans and the most frequent cause of mortality in infants with birth defects (Table 1.1) (Lee et al., 2001). They have been estimated to occur in 1-2% of live births, with a higher incidence if spontaneous abortions are included (Hoffman, 1995; Hoffman and Kaplan, 2002). Cardiovascular defects can affect most parts of the heart, such as the ventricles, atria and major arteries (Table 1.2). However, an overwhelming majority involve defects of valve and septa formation. The most common defect, present in about 60% of cases, is bicuspid aortic valve, in which the valve dividing the left-ventricle from the aorta has two leaflets instead of three. The second most prevalent heart defect, present in about 15% of cases, is a defect in ventricular septal formation wherein the left and right ventricles are not properly divided leading to left-to-right blood flow.

Table 1.1. Prevalence of birth defects associated with infant death in human population
(adapted from Lee et al., 2001)

| Birth defect | Percent of defects causing infant mortality |
|-------------------------------|--|
| Heart defects | 31% |
| Respiratory defects | 15% |
| Nervous system defects | 13% |
| Multiple abnormalities | 13% |
| Musculoskeletal abnormalities | 7% |

Table 1.2. Prevalence of specific heart defects in human population (adapted from Hoffman and Kaplan, 2002)

| Cardiovascular defect | Prevalence per million |
|--|-------------------------------|
| Bicuspid aortic valve* | 13,556 |
| Ventricular septal defect* | 3,570 |
| Patent ductus arteriosus | 799 |
| Atrial septal defect* | 941 |
| Atrio-ventricular septal defect* | 348 |
| Pulmonic stenosis* | 729 |
| Aortic stenosis* | 401 |
| Coarctation of the aorta | 409 |
| Tetralogy of Fallot* | 421 |
| Complete transposition of the great arteries | 315 |
| Hypoplastic right heart† | 222 |
| Tricuspid atresia* | 79 |
| Ebstein's anomaly* | 114 |
| Pulmonary atresia* | 132 |
| Hypoplastic left heart | 266 |
| Truncus arteriosus * | 107 |
| Double outlet right ventricle* | 157 |
| Single ventricle * | 106 |
| Total anomalous pulmonary venous connection | 94 |
| All defects affecting valves or septa | 20,661 |
| All congenital heart defects | 22,766 |

*Cardiovascular defects affecting valves or septa.

†Hypoplastic right heart includes tricuspid atresia, Ebstein's anomaly, and pulmonary atresia with an intact ventricular septum. The three components of the syndrome do not add up to the total because not all reports included all three.

The heart's activity is required during embryogenesis before it is itself fully formed and it is particularly susceptible to perturbations (Rossant, 1996). Epidemiological data point to many environmental factors contributing to heart defects, including maternal influenza during gestation, and prenatal exposure to non-steroidal anti-inflammatory drugs (Table 1.3) (Jenkins et al., 2007). However, mutations in key regulators of embryonic heart development are thought to be the main causes of congenital heart defects (Pierpont et al., 2007).

1.1.2. Genetic determinants of congenital heart defects in humans

Chromosomal trisomies and microdeletions as well as single-gene mutations have all been linked to congenital heart defects. For example, two of the most prevalent deletion syndromes with heart defects are DiGeorge syndrome and Williams-Beuren syndrome, which are associated with deletions in chromosomes 22q11 and 7q11.23, respectively (Table 1.4) (Ewart et al., 1993; Greenberg, 1993).

A number of single-gene mutations have also been found to be associated with syndromes that include particular congenital heart defects (Table 1.5). Among these are Alagille and Holt-Oram syndromes, which have been linked to mutations in the Notch ligand *JAGGED1* and the transcription factor *TBX5*, respectively (Basson et al., 1994; McCright et al., 2002; Newbury-Ecob et al., 1996; Terrett et al., 1994).

Table 1.3. Representative environmental determinants of congenital heart defects (adapted from Jenkins et al., 2007)

| | Relative risk of heart defects |
|--|---------------------------------------|
| Maternal illness | |
| Phenylketonuria | >6 |
| Pregestational diabetes | 3.1–18 |
| Febrile illness | 1.8–2.9 |
| Influenza | 2.1 |
| Maternal therapeutic drug exposure | |
| Anticonvulsants | 4.2 |
| Nonsteroidal anti-inflammatory drugs | |
| Ibuprofen | 1.86 |
| Sulfasalazine | 3.4 |
| Trimethoprim-sulfonamide | 2.1–4.8 |
| Maternal nontherapeutic drug exposure | |
| Maternal vitamin A | 0.0–9.2 |
| Marijuana | 1.9–2.4 |
| Environmental (maternal) | |
| Organic solvents | 2.3–5.6 |

Table 1.4. Representative chromosomal disorders associated with congenital heart defects
(adapted from Pierpont et al., 2007)

| Chromosomal disorder | Percent with heart defects | Major heart anomalies |
|---|-----------------------------------|---|
| Deletion 4p (Wolf-Hirschhorn syndrome) | 50–65 | ASD, VSD, PDA, TOF |
| Deletion 5p (cri-du-chat) | 30–60 | VSD, ASD, PDA |
| Deletion 7q11.23 (Williams-Beuren syndrome) | 53–85 | AS, PS |
| Trisomy 8 mosaicism | 25 | VSD, PDA, CoA, PS, TA |
| Deletion 8p syndrome | 50–75 | AVSD, PS, VSD, TOF |
| Trisomy 9 | 65–80 | PDA, VSD, TOF/PA, DORV |
| Deletion 10p | 50 | BAV, ASD, VSD, PDA, PS, CoA, TA |
| Deletion 11q (Jacobsen syndrome) | 56 | HLHS, AS, VSD, CoA |
| Trisomy 13 (Patau syndrome) | 80 | ASD, VSD, PDA, HLHS |
| Trisomy 18 (Edwards syndrome) | 90–100 | ASD, VSD, PDA, TOF, DORV, D-TGA, CoA, BAV |
| Deletion 20p12 (Alagille syndrome) | 85–94 | PA, TOF, PS |
| Trisomy 21 (Down syndrome) | 40–50 | AVSD, VSD, ASD, TOF, D-TGA |
| Deletion 22q11 (DiGeorge, velocardiofacial, and conotruncal anomaly face syndrome) | 75 | TA, TOF, VSD |
| Monosomy X (Turner syndrome, 45,X) | 25–35 | CoA, BAV, AS, HLHS |

AS = aortic stenosis; ASD = atrial septal defect; AVSD = atrio-ventricular septal defect; BAV = bicuspid aortic valve; CoA = coarctation of the aorta; DORV = double-outlet right ventricle; HLHS = hypoplastic left heart syndrome; D-TGA = D-transposition of the great arteries; PA = pulmonary atresia; PDA = patent ductus arteriosus; PS = pulmonic stenosis; TA = truncus arteriosus; TOF = tetralogy of Fallot; VSD = ventricular septal defect.

Table 1.5. Selected genes associated with congenital heart defects in the young (adapted from Srivastava, 2006)

| Genetic Mutation | Syndrome Name | Cardiac Disease |
|--------------------------------|----------------------|---|
| Nonsyndromic | | |
| <i>NKX2-5</i> | — | Atrial septal defect, ventricular septal defect, electrical conduction defect |
| <i>GATA4</i> | — | Atrial septal defect, ventricular septal defect |
| <i>MYH6</i> | — | Atrial septal defect |
| <i>NOTCH1</i> | — | Aortic valve disease |
| Syndromic | | |
| <i>TBX5</i> | Holt-Oram | Atrial septal defect, ventricular septal defect, electrical conduction defect |
| <i>TBX1</i> | DiGeorge | Cardiac outflow tract defect |
| <i>TFAP2β</i> | Char | Patent ductus arteriosus |
| <i>JAG1</i> | Alagille | Pulmonary artery stenosis, tetralogy of Fallot |
| <i>PTPN11</i> | Noonan | Pulmonary valve stenosis |
| <i>ELASTIN</i> | William | Supravalvar aortic stenosis |
| <i>FIBRILLIN</i> | Marfan | Aortic aneurysm |

Studies have also shown that single-gene defects can result in non-syndromic congenital heart defects (Table 1.5). For example, mutations in *NKX2-5* were identified in families with atrial septal defects and atrio-ventricular conduction delay (Schott et al., 1998), while mutations of *GATA4* were identified in families with septal defects (Garg et al., 2003) with no apparent syndromic features in either of these cases.

Despite these examples, to date relatively few cardiac abnormalities have been linked to specific gene mutations in humans. The majority of gene factors contributing to these defects remain undetermined.

1.1.3. Animal models of cardiac disease

In recent years, model organisms have been used to investigate the effects of defined genetic mutations during development and maturation of the heart. For example, the *Drosophila* system has proven useful in identifying the regulatory processes that occur during specification and differentiation of the embryonic heart (Bryantsev and Cripps, 2009; Cripps and Olson, 2002). The entire fly genome can be screened for candidate genes thanks to the vast collection of insertional mutants (Cooley et al., 1988). The zebrafish system has also proven useful in large mutagenesis screens, which have revealed specific genetic requirements in multiple aspects of heart development from cardiac fate determination to heart looping and remodeling (Glickman and Yelon, 2002). Finally, the chick embryo has proven useful in the study of heart valve formation and the contribution of neural crest cells to different regions of the heart (Hutson and Kirby, 2007; Martinsen, 2005).

The mouse, however, has become the organism of choice for creating models of human congenital heart defects (Harvey, 2002; Yutzey and Robbins, 2007). In addition to the advantage of dealing with a mammalian heart in terms of the application to human heart defects, our ability to manipulate the mouse genome makes it a good model in which to study the effects of defined genetic changes in the etiology of these defects (Rossant, 1996). Techniques used to manipulate the mouse genome include the use of targeted genetic recombination to create null mutations and drive the expression of marker genes, such as green fluorescent protein (GFP), under the control of tissue specific promoters, as well as the creation of conditional knock-out animals by Cre-lox recombination technology (Adams and van der Weyden, 2008).

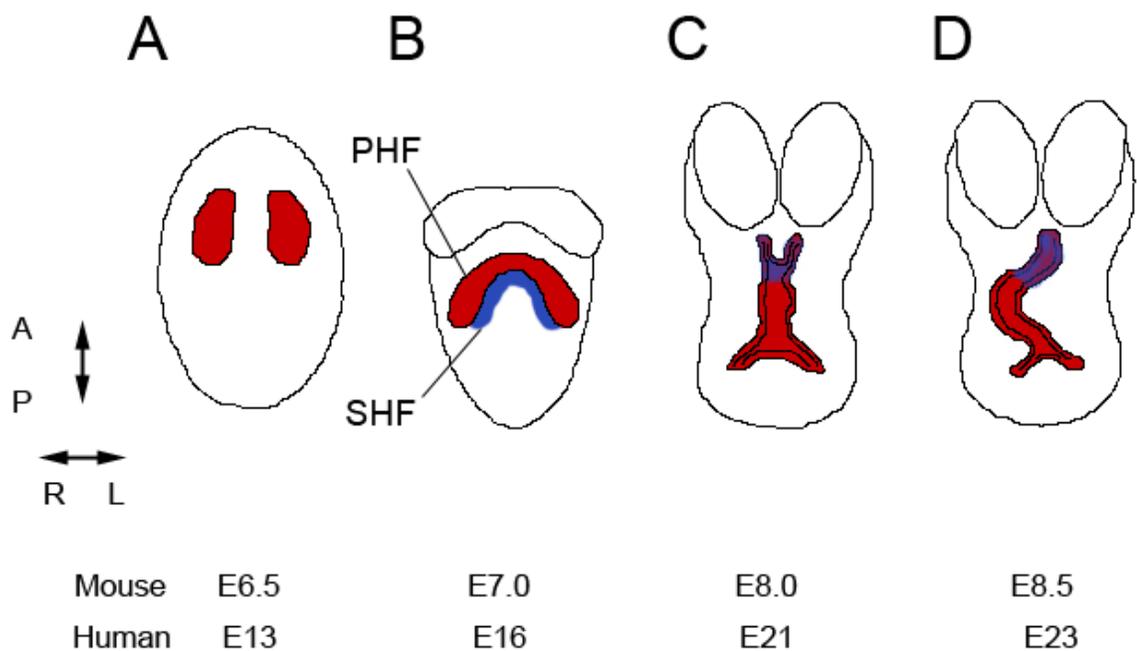
1.2. Cardiac valve development

1.2.1. Early heart morphogenesis

Cardiac development in mammals begins with the formation of the precardiac mesoderm. In this process, a specific subset of mesodermal cells generated during gastrulation migrates to the anterior and lateral portions of the embryo to form two bilateral heart forming regions (Figure 1.1A). Cells within these regions will then differentiate into progenitors for the myocardium (a type of involuntary striated muscle found in the walls of the heart) and the endocardium (a type of epithelial cell that lines the internal cavities of the heart). As the embryo folds, at embryonic day (E) 7.0 in the mouse, the two heart forming regions join to form a cardiac crescent (Figure 1.1B). Between E7.25 and E8.0 in the mouse, fusion of the cardiac crescent at the embryonic midline gives rise to the primary heart tube, consisting of a myocardial outer layer that is separated from an inner endocardial tube by an acellular matrix known as the cardiac jelly (Figure 1.1C) (Eisenberg and Markwald, 1995). At this time, the primitive heart begins to beat and continues to do so throughout the heart remodeling process.

Figure 1.1. Early heart morphogenesis

A) Two bilateral heart forming regions contain myocardial and endocardial progenitors. B) The two heart forming regions join to form a crescent. C) Cardiac crescent fuses at the embryonic midline to give rise to the primary heart tube. An outer myocardial layer is separated from the internal endocardial tube by the cardiac jelly. D) The heart tube loops to the right and specific regions of the myocardium expand to form the future atrial and ventricular chambers. A = anterior; P = posterior; R = right; L= left; PHF = primary heart field; SHF = secondary heart field. Approximate embryonic days (E) are indicated below the drawings.



Recent studies have suggested that a second group of cardiac mesodermal precursor cells, called the secondary heart field (SHF), also plays an important role in heart development in addition to the classic cardiac forming region of the cardiac crescent, the primary heart field (PHF) (Kelly et al., 2001; Mjaatvedt et al., 2001; Waldo et al., 2001). Using quail-chicken chimeras, and mitochondrial fluorescent marking techniques in chicken embryos, it was shown that cells in the pharyngeal mesenchyme are added to the elongating anterior heart tube (Mjaatvedt et al., 2001; Waldo et al., 2001). Moreover, ablation of the PHF in chick embryos did not affect development of the anterior myocardium, further suggesting that it originates from a distinct heart field (Mjaatvedt et al., 2001). A similar field was also identified in mammals. Analysis of a transgenic mouse line in which beta-galactosidase activity was driven by the *Fgf10* gene regulatory sequences revealed shared expression between the anterior most part of the heart tube and the contiguous splanchnic and pharyngeal mesoderm (Kelly et al., 2001). Lineage tracing experiments by DiI labeling in mouse embryos then confirmed movement of these mesoderm cells into the anterior aspect of the heart tube.

Further tracing studies will be needed to determine whether the PHF and SHF are truly two distinct entities or rather various subdivisions of a single primordial heart field (Abu-Issa and Kirby, 2007; Moorman et al., 2007), however it is clear that cells originating from the SHF contribute to heart development at a later stage than those originating from the PHF, and have a different pattern of gene expression (Cai et al., 2003; Kelly et al., 2001; Verzi et al., 2005). It is now generally accepted that the PHF primarily contributes to the atria, atrio-ventricular canal (AVC), and left ventricle, while the outflow tract (OFT), right ventricle, and components of the venous pole are mostly derivatives of the SHF (Srivastava, 2006), although cells derived from the secondary heart field have been found throughout the heart.

As development proceeds, the primary heart tube elongates at both the arterial and venous poles. Starting at about E8.5 in the mouse, the tubular heart loops toward the right (Figure 1.1D) (Manner, 2004) and relatively high levels of proliferation in specific regions of the myocardium expand the future atria and ventricles (Moorman and Christoffels, 2003). As a result of these remodeling events, the future segments of the heart (such as the presumptive right and left ventricles and common atria) become more recognizable, though still arranged in series (De La Cruz et al., 1989).

1.2.2. Valve and septa formation

As looping is occurring in the heart, cardiac jelly begins to accumulate between the endocardium and myocardium in the AVC and OFT (Figure 1.2A), giving rise to local swellings, or cushions, that are rich in extracellular matrix (ECM) components, particularly hyaluronic acid (Camenisch et al., 2000) and versican (Henderson and Copp, 1998; Mjaatvedt et al., 1998). Starting at E9.0-9.5 in the mouse, the initially acellular cushions become populated by mesenchymal cells as a result of an epithelial-to-mesenchymal transformation (EMT) of the endocardial cells in the AVC and OFT regions (Bolender and Markwald, 1979; Markwald et al., 1977). During EMT, polarized and adhesive epithelial cells transform into non-polarized and highly motile mesenchyme cells embedded in the ECM (Markwald et al., 1975) (Figure 1.3). In this multi-step process, groups of cells destined to undergo EMT are first specified and become hypertrophic. These cells then lose polarity markers and intercellular cadherins at adherens junctions. Finally, degradation of the basement membrane by matrix metalloproteinases (such as

MMP2) and cell delamination driven by reorganization of the cytoskeleton leads to the migration and invasion of EMT-generated cells into the cardiac cushion.

Two sets of cushions develop in the AVC and OFT. The first set faces each other on opposing sides of the common AVC and OFT (Figure 1.2B) (Qayyum et al., 2001; Wessels and Sedmera, 2003). In the case of the OFT, fusion of this set of cushions will divide the pulmonary and aortic vessels, while in the case of the AVC, this first set will fuse at the midline and later join with projections from the atria and ventricles to form the inter-atrial and inter-ventricular septa, respectively. After fusion of the first (or major) cushions, another set of endocardial cushions starts to form in the lateral aspects of the AVC and OFT. This second set of cushions (or minor cushions) will give rise to the mitral and tricuspid valves in the case of the AVC cushions and the semilunar valves in the case of the OFT.

Most of our current knowledge on the process of endocardial EMT comes from *in vitro* AVC explant assays, in which EMT is quantified by measuring the invasion of endocardially derived cells into three dimensional collagen gels (Bernanke and Markwald, 1982). Using these assays, it was demonstrated that signals from the adjacent AVC myocardium (such as BMP2) are required for normal EMT to occur (Eisenberg and Markwald, 1995). *In vivo* studies, in which cells were labeled by *Cre* recombinase expression under the control of myocardial, endocardial and neural crest specific promoters (*alphaMHC*, *Tie2* and *Wnt1*, respectively), then conclusively demonstrated that AVC cushion mesenchyme cells originate exclusively from endocardial EMT and that their fate is limited to the fibrous tissues that constitute the valve leaflets and parts of the membranous septa (de Lange et al., 2004).

Figure 1.2. Endocardial cushion formation

A) Cushions are formed in AVC and OFT regions of the heart by secretion of ECM molecules.

B) Two cushions are formed first, facing each other on opposite sides of the common AVC. A second set of cushions appears laterally after the first set has fused. RA = right atrium, LA = left atrium, RV = right ventricle, LV = left ventricle, OFT = outflow tract, AVC = atrio-ventricular canal.

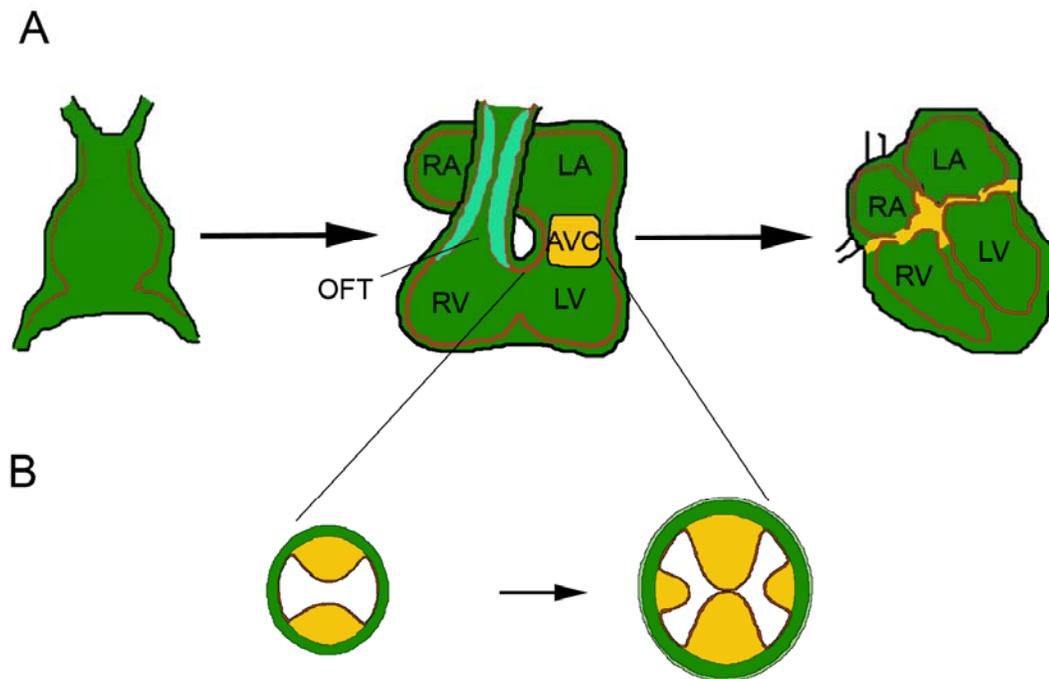
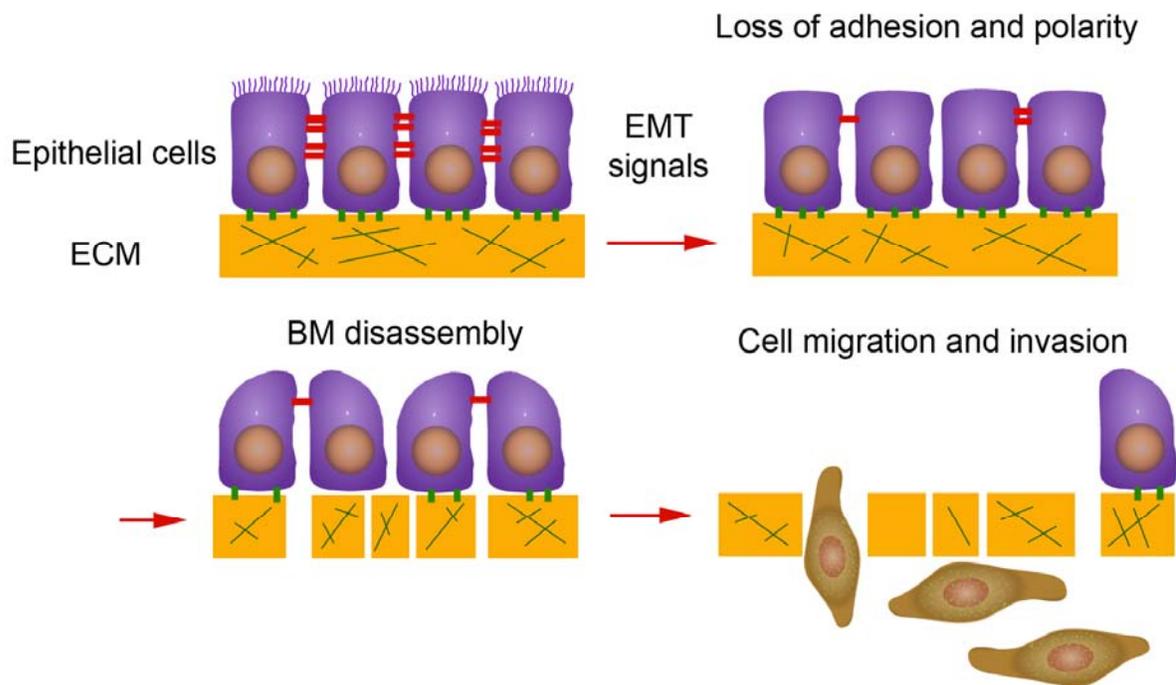


Figure 1.3. Epithelial-to-mesenchymal transformation

EMT is a multistep process. Epithelial cells are first induced to lose adhesion and polarity. Basal membrane (BM) is disassembled. Cells then acquire migratory phenotype and invade the extracellular matrix (ECM) (adapted from Levayer and Lecuit, 2008).



Given their overall histological resemblance, it was thought that the mesenchyme populating the OFT cushions had a similar origin as that of the AVC cushions. However, studies on the contribution of cardiac neural crest cells to OFT development and more recent endocardial-Cre recombination studies, have established that only the most proximal and distal portions of the OFT cushions contain endocardially derived cells, with the rest being derived from cardiac neural crest and SHF cells (Engleka et al., 2005; Hutson and Kirby, 2007; Nakamura et al., 2006; Phillips et al., 1987; Poelmann et al., 1998; Waldo et al., 2001).

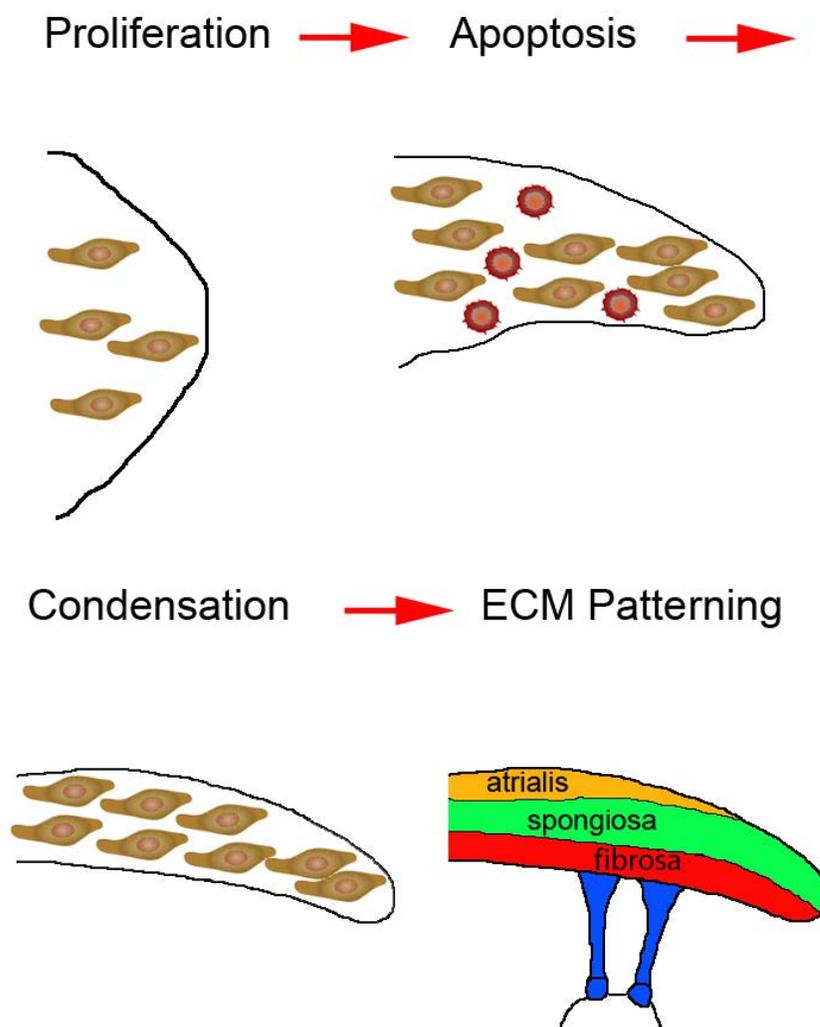
1.2.3. Valve remodeling

While there has been considerable effort in understanding the initial steps of endocardial cushion formation, the mechanisms that induce the remodeling and maturation of the cushions into stress-resistant valve leaflets following EMT are only beginning to be defined. This is due in part to the limited number of mouse mutants with valve defects that are viable past the stage of EMT.

The first step in valve remodeling turns the endocardial cushions into thin mesenchymal leaflets (Figure 1.4). This is accomplished by changes in proliferation and apoptosis in the endocardial cushion (de Lange et al., 2004; Gitler et al., 2003). Higher proliferation rates at distal ends of the future valves lead to elongation of the valve primordia (Hinton et al., 2006; Lincoln et al., 2004). Apoptosis then takes place within the mesenchymal cushion population to remove excess cells (Abdelwahid et al., 2001; Poelmann et al., 2000).

Figure 1.4. AVC valve remodeling

Following EMT, proliferation of cells in the endocardial cushions leads to elongation of the valve primordia. Apoptosis removes excess cells to give rise to thin leaflets. Condensation of mesenchyme cells is followed by differentiation of the mesenchymal cells and patterning of the ECM relative to blood flow.



Following the initial burst of proliferation and apoptosis, an increase in cell density is observed within the mesenchymal leaflets (Kruithof et al., 2007). This condensation of previously dispersed mesenchymal cells is required for the development and patterning of other mesenchymal tissues, such as bone, cartilage, and tendon (Hall and Miyake, 2000). Significantly, recent studies have found that markers of tendon, cartilage and bone development are also expressed in the remodeling AVC suggesting a link between mesenchymal condensation and ECM patterning during valve remodeling (Lincoln et al., 2004; Lincoln et al., 2006). After condensation, the ECM proteins are patterned relative to blood flow, finally leading to adult AVC and OFT valves that are composed of three distinct layers: the fibrosa, spongiosa and atrialis/ventricularis layers. Away from the blood flow, the fibrosa layer is composed of densely packed collagen fibers which provide strength. The spongiosa layer, which is centrally located, has abundant glycosaminoglycans which provide cushioning. Finally the atrialis/ventricularis layer, which is continuous with either the atrial endocardium in the case of the AVC valves or the ventricular endocardium in the OFT valves, is composed of elastic fibers.

1.3. Signalling pathways in AVC development

Over the last two decades there has been intense research interest in studying the factors that contribute to the regulation of valve morphogenesis. This has led to the discovery of many growth factors, intracellular signalling molecules, transcription factors and ECM components that regulate this complex process (reviewed elsewhere: Armstrong and Bischoff, 2004; Person et al., 2005). Chief among these are members of the Notch and BMP/TGF β pathways.

1.3.1. Notch signalling pathway

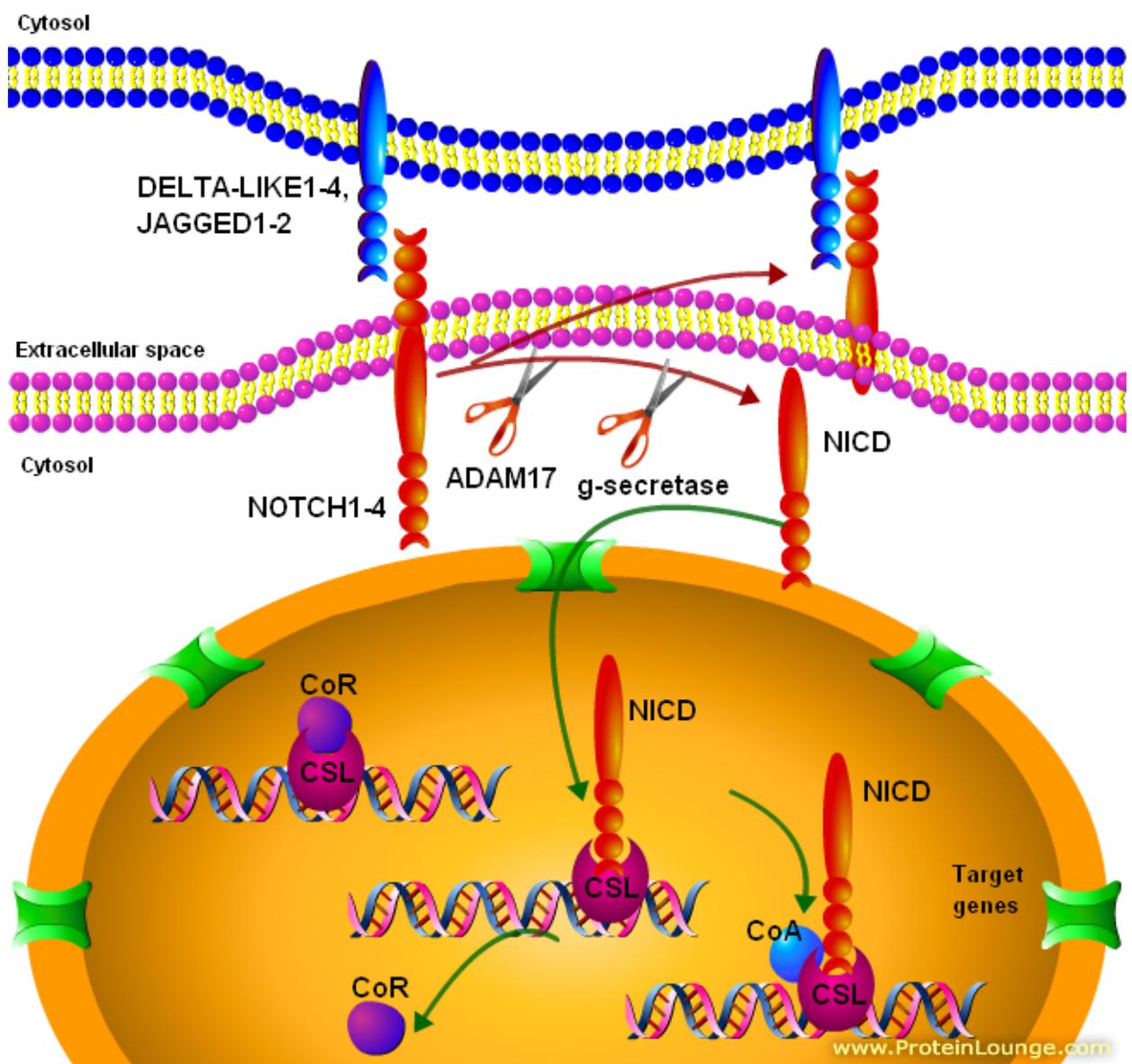
In mammals, the Notch pathway consists of four transmembrane receptors (*Notch1-4*) and five transmembrane ligands, (*Jagged1* and *2* and *Delta-like1, 3, and 4*) (Niessen and Karsan, 2008). Interaction of the Notch ligand with its receptor at the cell surface results in a series of receptor cleavages, mediated first by ADAM17 and then by the γ -secretase complex, which ultimately releases the intracellular domain of the Notch receptor (NICD) and allows it to translocate to the nucleus (Figure 1.5). Once in the nucleus, NICD interacts with a transcriptional repressor called CSL (CBF1, suppressor of hairless, Lag-1; also known as RBP-J κ), which results in activation of target genes, of which the Hes and Hey families of transcription factors are the best defined.

The importance of the Notch pathway during heart development is borne out by mutations resulting in congenital heart defects in humans. For example, mutations in the Notch ligand *JAGGED1* have been implicated in the rare autosomal dominant disorder Alagille syndrome (Li et al., 1997; Oda et al., 1997), while inactivating mutations of *HEY2* also cause heart defects in humans similar to those found in individuals with Alagille syndrome (Reamon-Buettner and Borlak, 2006).

Recent studies in mice aiming to address the specific role of the Notch pathway during valve morphogenesis have revealed critical functions during specification of the AVC and later during EMT. During AVC specification, the AVC myocardium acquires a distinct gene expression pattern and function from the surrounding chamber myocardium. Unlike the ventricular myocardium, the AVC does not undergo trabeculation, nor does it acquire the high-gap junction density and conductivity of the chamber myocardium (Harrelson et al., 2004).

Figure 1.5. Notch pathway

Notch ligand and receptor interaction at the cell surface results in a series of receptor cleavages. The release of the intracellular domain of the Notch receptor (NICD) allows it to translocate to the nucleus. NICD interacts with a transcriptional repressor called CSL resulting in release of co-repressors (CoR), recruitment of co-activators (CoA) and activation of target genes.



Furthermore, the AVC uniquely expresses both the *Bmp2* ligand and the *Tbx2* transcription factor, which are required for cushion EMT and for the inhibition of chamber specific myocardial gene expression, respectively (Harrelson et al., 2004; Ma et al., 2005; Zhang and Bradley, 1996). Recent studies demonstrated that the Notch target genes *Hey1* and *Hey2* are critically involved in restricting *Bmp2* and *Tbx2* expression to the AVC (Kokubo et al., 2007; Rutenberg et al., 2006). Over-expression of *HEY1* and *HEY2* in chick myocardium led to the repression of *BMP2* transcription, and misexpression of *Hey1* or *Hey2* in the entire cardiac lineage of the mouse resulted in the reduction or absence of the AVC, with poorly defined boundaries and loss of *Bmp2* and *Tbx2* expression. Conversely, *Hey2* disruption in mice and zebrafish resulted in ectopic cardiac *Bmp2* transcription.

Following proper AVC specification, a subset of endocardial cells lining the AVC are induced to undergo EMT. Multiple Notch ligands (*Jagged1* and *Dll4*), receptors (*Notch1*, 2 and 4) and downstream targets (*Hey1*, *Hey2* and *HeyL*) are expressed in the AVC endocardium during this time (Benedito and Duarte, 2005; Fischer et al., 2007; Loomes et al., 2002). Significantly, NOTCH1 activation in cardiac mesodermal cells by Cre recombinase expressed under the control of the *Mesp1* promoter led to increased cushion EMT resulting in hypercellular AVCs and enlarged AV valves (Watanabe et al., 2006). Similarly, transient ectopic expression of activated Notch1 in zebrafish embryos led to hypercellular cardiac valves. In contrast, deficiency of either *Notch1* or *Csl* in the mouse led to significant reduction in AVC EMT *in vivo* and in explant cultures (Timmerman et al., 2004; Watanabe et al., 2006). Finally, *Hey2*-deficient and *Hey1/L* double-deficient mouse embryos also showed severe heart malformations, including membranous ventricular septal defects and dysplastic atrio-ventricular and pulmonary valves caused by defects in endocardial cushion EMT, as confirmed by AVC explants (Fischer et al.,

2007). Thus Notch signalling is critical for both AVC specification and EMT during heart valve formation.

Interestingly, Notch signalling appears to control two different steps of EMT via distinct downstream targets. Notch signalling via the Snail family of transcriptional repressors (*Snail* and *Snai2*) first initiates EMT by repressing expression of *VE-cadherin* and other endothelial markers (Niessen et al., 2008; Timmerman et al., 2004). Hey genes then regulate the acquisition of a mesenchymal phenotype, including the expression of *Mmp2* (Fischer et al., 2007).

1.3.2. BMP and TGF β signalling pathways

Bone morphogenetic protein (BMP) and transforming growth factor β (TGF β) signalling pathways have emerged as vital regulators of multiple aspects of cardiogenesis (Schlange et al., 2000; Shi et al., 2000; Walters et al., 2001). BMP and TGF β ligands are homodimeric proteins involved in many cellular processes such as regulation of cell growth, differentiation and apoptosis. The signalling cascade begins when ligand binding to a type II transmembrane receptor leads to the recruitment and transphosphorylation of a type I receptor (Figure 1.6). The phosphorylated type I receptor then acts as a serine/threonine kinase to phosphorylate and activate Smad proteins, which are the major intracellular mediators of TGF β and BMP signalling (Shi and Massague, 2003). These receptor-regulated Smads (R-Smads - SMAD2 and 3 for TGF β s, and SMAD1, 5 and 8 for BMPs) can now bind to the common-Smad (SMAD4) and accumulate in the nucleus where they control transcription of target genes. Inhibitory Smads regulate the signalling cascade by either competing with R-Smad binding to the type I receptors

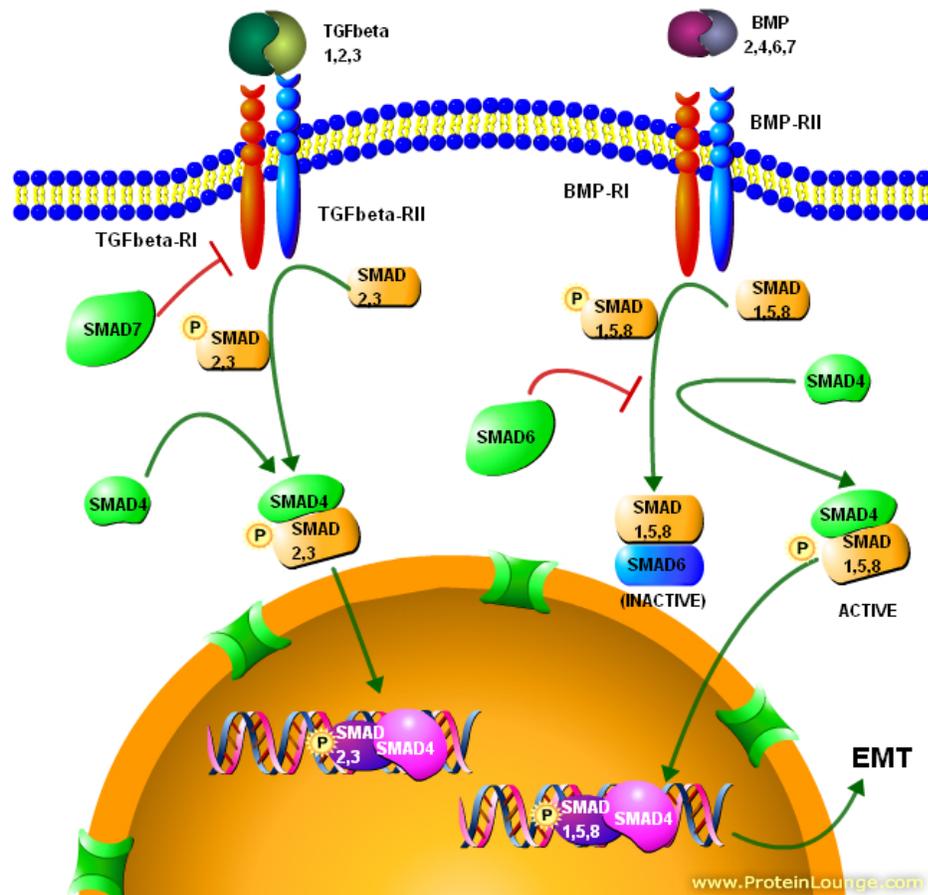
preventing their phosphorylation (as in the case of SMAD7) or by sequestering R-Smads (as in the case of SMAD6 binding to BMP R-Smads 1,5 and 8) (Hata et al., 1998).

During mouse endocardial cushion formation, several BMPs and TGF β s are preferentially distributed within the AVC and OFT regions of the heart. In the AVC and OFT, *Tgfb1* is expressed in endocardial and mesenchymal cells while *Tgfb2* and *Tgfb3* are expressed in the outer myocardium (Akhurst et al., 1990; Dickson et al., 1993). Similarly, *Bmp2* and *Bmp4* are expressed in the myocardium underlying the developing AVC and OFT, while *Bmp6* is expressed in endocardium, myocardium, and mesenchyme of the developing heart, with some enrichment in the AVC (Jones et al., 1991).

Multiple studies have demonstrated that TGF β and BMP signalling are necessary for proper control of EMT. In AVC explant assays, BMP2 can induce EMT in the absence of AVC myocardium (Sugi et al., 2004), and knock-down of *Bmp2* mRNA transcripts results in reduction of mesenchymal cell invasion (Yamagishi et al., 1999). TGF β 2 is also required for mammalian endocardial cushion cell transformation *in vitro* (Camenisch et al., 2002) as demonstrated by use of blocking antibodies. Interestingly *Tgfb2* expression appears to be controlled by BMP2 activity, as AVC endocardium treated with BMP2 expressed elevated levels of *Tgfb2* in the absence of myocardium (Sugi et al., 2004). This BMP2-induced elevation of *Tgfb2* expression in endocardial cells was abolished by treatment with the BMP antagonist, NOGGIN.

Figure 1.6. BMP and TGFβ pathways

Ligand binding to a type II receptor leads to the recruitment and transphosphorylation of a type I receptor. The phosphorylated type I receptor then acts as a serine/threonine kinase to phosphorylate and activate Smad proteins. Receptor-regulated Smads bind to the common-Smad (SMAD4) and accumulate in the nucleus where they control transcription of target genes. SMAD7 inhibits the TGFβ pathway by preventing type I receptor phosphorylation. SMAD6 inhibits BMP pathway by binding to BMP-regulated Smads 1, 5 and 8.



In vivo, disruption of *Bmp2* in the AVC leads to EMT defects (Ma et al., 2005). Inactivation of *Bmp2* in the AVC myocardium by *Nkx2-5* Cre-mediated recombination resulted in lack of EMT indicating that BMP2 signals directly to cushion-forming endocardium to induce EMT. *Bmp6* and *Bmp7* double null mutants also exhibit a delay in OFT formation and hypoplastic cardiac cushions (Kim et al., 2001). In contrast, *Tgfb2* null embryos develop enlarged cushions at E14.5 with elevated levels of EMT markers. Subsequent analysis of *Tgfb2* null AVC explants at E9.5 and E10.5 indicated that EMT, but not cushion cell proliferation, was initially delayed but later remained persistent (Azhar et al., 2009). Therefore, disruption of *Tgfb2* results in a failure to first induce and then limit EMT resulting in hyperplastic valves (Azhar et al., 2009; Bartram et al., 2001; Sanford et al., 1997).

The genetic disruption of BMP receptors causes deregulated cardiac cushion formation as well. Endocardial-specific deletions of *Alk2* or *3* (also known as *Acvr1* and *Bmpr1a*), which encode type I BMP receptors, led to failure of EMT in the AVC (Ma et al., 2005; Song et al., 2007; Wang et al., 2005) while a myocardial-specific disruption of *Alk3* led to hypoplastic cushions that failed to fuse, even though EMT occurred normally (Gaussin et al., 2005). Also, a hypomorphic allele of *Bmpr2* (a type II BMP receptor) that has reduced signalling capability results in failure to form OFT derived valves (Delot et al., 2003). Finally, the increase in BMP signalling caused by deletion of *Smad6* results in hyperplastic valves (Galvin et al., 2000), suggesting that *Smad6* expression, which is induced by BMP2, may be involved in limiting the number of endocardial cells that undergo EMT (Takase et al., 1998).

These studies have demonstrated the importance of the Notch, BMP and TGF β pathways in regulating multiple stages during endocardial cushion formation. In Chapter 2 I will discuss the identification of different sets of downstream targets controlling these different processes during AVC development.

1.4. Role of transcription factors in AVC development

Transcription factors are the intracellular effectors that ultimately control cellular activity by regulating expression of genes. Many transcription factors have been shown to regulate specific aspects of AVC development (Table 1.6). For example, SOX4 activity is critical for endocardial cushion EMT (Ya et al., 1998), while TBX20 plays a role in mesenchymal cell expansion and differentiation (Shelton and Yutzey, 2007). A main focus of my work, as discussed in Chapter 3, has been the transcription factor TWIST1.

Table 1.6. Selected transcriptional regulators controlling valve morphogenesis

| Transcriptional regulator | Role in valve development | References |
|----------------------------------|---|---|
| <i>Fog2</i> | Attenuates endocardial cushion EMT | (Flagg et al., 2007) |
| <i>Hey1</i> | EMT | (Fischer et al., 2007) |
| <i>Hey2</i> | EMT | (Fischer et al., 2007) |
| <i>HeyL</i> | EMT | (Fischer et al., 2007) |
| <i>Msx1</i> | EMT/cushion formation | (Chen et al., 2008) |
| <i>Msx2</i> | EMT/cushion formation | (Chen et al., 2008) |
| <i>Nfatc</i> | Proliferation of valve endothelial cells/valve ECM remodeling | (de la Pompa et al., 1998; Ranger et al., 1998) |
| <i>Scleraxis</i> | Differentiation/ECM organization | (Levay et al., 2008) |
| <i>Smad6</i> | Negative regulation of EMT | (Desgrosellier et al., 2005) |
| <i>Snai2</i> | Promotes migration of transformed endothelial cells | (Niessen et al., 2008; Romano and Runyan, 2000) |
| <i>Sox4</i> | OFT cushion maturation | (Ya et al., 1998) |
| <i>Sox9</i> | Cushion cell expansion and ECM organization | (Lincoln et al., 2007) |
| <i>Tbx2</i> | AVC boundary | (Kokubo et al., 2007) |
| <i>Tbx5</i> | AVC boundary | (Maitra et al., 2009) |
| <i>Tbx20</i> | Cushion cell proliferation and differentiation | (Shelton and Yutzey, 2007) |

1.4.1. Twist family

Twist proteins are highly conserved basic helix-loop-helix (bHLH) transcription factors that have important regulatory functions during embryogenesis. The only Twist family member found in *Drosophila* (dTWIST) is necessary for mesoderm formation during gastrulation, a type of EMT (Simpson, 1983; Thisse et al., 1987). In mammals, several Twist-related genes have been isolated, including *Twist1*, *Twist2*, *Scleraxis*, and *Paraxis* (Burgess et al., 1995; Cserjesi et al., 1995; Li et al., 1995). These proteins share varying degrees of sequence similarity, with TWIST1 and TWIST2 being the closest dTWIST-like proteins (Li et al., 1995; Wolf et al., 1991) followed by SCLERAXIS and PARAXIS. HAND1 and HAND2 are also very highly related bHLH transcription factors that have sometimes been included in the Twist family of transcription factors.

The highest level of sequence conservation among Twist proteins is found in the bHLH domain, which consists of a short stretch of basic amino acids followed by two amphipathic α -helices divided by a loop of varying length (Figure 1.7) (Massari and Murre, 2000). Each of the α -helices allows for protein-protein interactions with other bHLH proteins. Typically, dimerization of the cell-type specific Twist family members with ubiquitously expressed bHLH factors (E-proteins, such as E12) results in the juxtaposition of their basic domains creating a combined DNA binding motif capable of binding to an E-box sequence (3'CANNTG5') (Furumatsu et al., 2010; Wilson-Rawls et al., 2004). Twist proteins then control transcription by forming complexes with non-bHLH proteins such as MEF2, SMAD4, histone deacetylases (HDACs) and acetyltransferases (HATs) (Gong and Li, 2002; Hamamori et al., 1999; Hayashi et

al., 2007; Muir et al., 2008). DNA binding usually leads to the activation of gene expression; however TWIST1 and TWIST2 have also been shown to act as negative regulators of gene expression. They can do this by direct interaction with the DNA binding domains of other transcription factors (Bialek et al., 2004), the basic domain of other bHLHs, or by sequestering E-proteins (Gong and Li, 2002; Hamamori et al., 1997; Lee et al., 2003; Sosic et al., 2003; Spicer et al., 1996).

Twist activity itself can be regulated in a variety of ways. At the transcriptional level, Twist proteins have very specific gene expression domains (see below). Twist activity can also be regulated by heterodimerization with other proteins, such as CREB family members (Hayashi et al., 2007; Massari and Murre, 2000; Muir et al., 2008). Finally, recent work has shown that the phosphorylation of TWIST1 protein at key residues within the bHLH domain can regulate its partner choice leading to changes in DNA binding affinity (Firulli and Conway, 2008).

Figure 1.7. Twist family homology

Mouse amino acid sequences were used for this comparison. “*” denotes same residue, “.” indicates residues with similar properties, “:” Indicates residues are more or less similar.

```

Scleraxis      -MSFAMLRSAPPPGRYLYPEVSPLSEDE----DRGSESSGSDEKPC----- 41
Paraxis       -MAFALLR--PVGAHVLYPDVRLLEDE----ENRSESDASDQS----- 37
Twist1        MMQDVSSSPVSPADDSLSNSEEE PDRQQPASGKRGARKRRSSRRSAGGSAGPGGATGGGI 60
Twist2        -MEEGSSSPVSPVD-SLGTSEEELERQP----KRFRGRKRRYSKKSS----- 40
               *      .      *      .      .:      . . . . .

                                                     <DNA binding

Scleraxis      ----RVHAARCGLQGARR-----RAGGRRRAAGSGPGPGGRPGREPRQRHTANAR 86
Paraxis       ----FGCCEGLEAARR-----GPG--PGSGRRASNGAGPVVVVRQRQAANAR 78
Twist1        GGGDEPGSPAQGKRGKKSAGGGGGGAGGGGGGGSSSSGGGSPQSYEELQTQRVMANVR 120
Twist2        ----EDGSPTPGKRGKK-----GSPSAQSFEELQSQRILANVR 74
               .      *      . . :                               **  **.*

               ><   Helix I   ><   Loop   ><   Helix II   >

Scleraxis      ERDRTNSVNTAFTALRTLIPTEPADRKLSKIETLRLASSYISHLGNVLLVGEACGD-GQP 145
Paraxis       ERDRTQSVNTAFTALRTLIPTEPVDRKLSKIETLRLASSYIAHLANVLLLGELTADDGQP 138
Twist1        ERQRTQSLNEAFAALRKIIPTLPSD-KLSKIQTLKLAARYIDFLYQVLQSDLDKMSC 179
Twist2        ERQRTQSLNEAFAALRKIIPTLPSD-KLSKIQTLKLAARYIDFLYQVLQSDMDNKMTSC 133
               **:***:*:* **:*:*:*:* * * *****:**:*:* ** .* :** .* . .

Scleraxis      CHSGPAFFHSGRAGSPLPPPPPPPLARDGGENTQPKQICTFCLSNQRKLSKDRDRK--- 202
Paraxis       CFR-----AAGGGKSAVP-----AADG---RQPRSICTFCLSNQRKGGSRDLGGSC 182
Twist1        SYV-----AHERLSYAFSVWRMEGAWSMSASH--- 206
Twist2        SYV-----AHERLSYAFSVWRMEGAWSMSASH--- 160
               ..                               : :  :*:. . . .

Scleraxis      -----TAIRS--- 207
Paraxis       LKVRGVAPLRGPRR 196
Twist1        -----
Twist2        -----

```

1.4.2. Role of Twist family members in development and cancer

Despite their sequence and DNA binding similarities, Twist family members have been shown to play unique roles during embryonic development and cancer progression. *Twist1* null mouse embryos die around E11 and display a number of defects, including exencephaly, hypoplastic limb buds, and vascular defects, that are consistent with an important role in mesenchymal cell development (Chen and Behringer, 1995). In addition, *Twist1* heterozygote null mice display a number of partially penetrant phenotypes such as dental malocclusion, craniosynostosis and preaxial polydactyly. These haploinsufficient phenotypes are similar to Saethre-Chotzen syndrome, an autosomal dominant disease in humans which in the majority of documented cases has been shown to involve a loss-of-function mutation in the *TWIST1* gene (Corsi et al., 2002; Firulli and Conway, 2008).

Twist2 knock-out mice die perinatally due to wasting and a failure to thrive (Sosic et al., 2003). They display enhanced proinflammatory cytokine gene expression and increased apoptosis in multiple tissues. Significantly, this phenotype is also observed in *Twist1* and *Twist2* compound heterozygotes suggesting redundancy and dosage dependence in certain contexts (Bialek et al., 2004; Sosic et al., 2003).

While *TWIST1* and *TWIST2* are only expressed in a subset of mesodermally derived tissues, a large fraction of human cancers over-express *TWIST1* and/or *TWIST2* (for example, Kwok et al., 2005; Mironchik et al., 2005; Ohuchida et al., 2007; Yang et al., 2004; Zhang et al., 2007). Significantly, *TWIST1* and *TWIST2* have been associated with the metastatic process which involves a similar loss of polarity and acquisition of migratory phenotype characteristic of

EMT (Ansieau et al., 2008; Yang et al., 2004). Over-expression of *TWIST1* and *TWIST2* is now believed to prevent the premature senescence induced by oncogenes driving the metastatic process, such as RAS (Ansieau et al., 2008).

During mouse embryogenesis, *Scleraxis* null mice show a disruption of tendon differentiation (Murchison et al., 2007), resulting in reduced and disorganized tendon matrix and disrupted cellular organization of force-transmitting and intermuscular tendons. Finally, *Paraxis* null mice show perinatal lethality, with a failure of epithelial cell formation from paraxial mesoderm resulting in disrupted somites and patterning defects of the axial skeleton, peripheral nerves, and skeletal muscles (Burgess et al., 1995).

1.4.3. Twist family activity in AVC and OFT development

Since the start of my project, *TWIST1* and *SCLERAXIS* have been shown to play specific roles during valve formation and remodeling. Both *Twist1* and *Scleraxis* are very highly expressed in the mesenchyme compartment of the AVC and OFT, however they play distinct roles during valve formation. Through the analysis of knock-out mice, it was revealed that *SCLERAXIS* is involved in valve maturation through its role in cell lineage differentiation and matrix distribution in remodeling valve structures (Levay et al., 2008). These alterations in valve precursor cell differentiation and extracellular matrix and collagen fiber organization result in thickened heart valve structures in juvenile *Scleraxis* knock-out mice.

Recent gain and loss of function research performed in primary chicken endocardial cushion cells revealed that *TWIST1* can induce endocardial cushion cell proliferation as well as promote endocardial cushion cell migration through its regulation of *TBX20* expression (Shelton

and Yutzey, 2008). Furthermore, *TWIST1* was found to be subject to BMP2 regulation and to induce expression of cell migration marker genes while repressing markers of valve cell differentiation. In contrast to what was observed in chick endocardial cushion cells, studies in *Twist1* null mice have not found an obvious AVC phenotype (Vincentz et al., 2008). Instead, *Twist1* null mice contain amorphous cellular nodules within their OFT endocardial cushions. This nodular mesenchyme was found to be the result of cardiac neural crest cells that failed to properly differentiate. Thus, the role of TWIST1 in the endocardially derived cells of the mammalian AVC and OFT remains unclear.

1.5. Transcriptome analysis

Techniques developed over the last 15 years have revolutionized biology by allowing the simultaneous expression analysis of thousands of genes (Brenner et al., 2000; DeRisi et al., 1997; Lipshutz et al., 1999; Schena et al., 1995; Velculescu et al., 1995). Perhaps the most widely used to date have been array-based approaches (such as Affymetrix GeneChip), in which labeled transcripts are hybridized to oligonucleotide probes synthesized on a microarray. These have proven useful in the study of heart development and disease leading to the identification of novel genes and regulatory networks (for example, Chakraborty et al., 2008; Moskowitz et al., 2007; Wirrig et al., 2007). However, array-based methods have limited sensitivity and gene discovery potential since they rely on prior knowledge of genome transcripts.

Serial Analysis of Gene Expression (SAGE) is a technique capable of examining gene expression profiles without prior knowledge of the genes involved (Velculescu et al., 1995). SAGE is based on the generation of short nucleotide sequences (tags) from a unique position

within each species of mRNA to evaluate thousands of expressed transcripts in a single assay (Figure 1.8). The 5' ends of tags are generated by digestion of a cDNA pool with a frequent 4 base-pair cutter restriction nuclease (e.g. *NlaIII*), while the 3' ends are generated by cleavage with a type IIS restriction enzyme (e.g. *MmeI*) that cuts at a defined distance from its recognition site. Tags are then sequenced and mapped to the genome, with the frequencies of unique tag sequences (tag-types) corresponding to levels of gene expression.

Traditionally, the cost of sequencing and the labour intensive nature of the SAGE approach made it a less attractive option than microarray analysis for the study of gene expression at a global scale. As the cost of sequencing continues to decrease, SAGE technology has become a more accessible alternative to microarray technology. Unlike microarrays, SAGE data produces absolute values that can be easily compared thus allowing the creation of large public SAGE datasets for numerous tissues and states. Furthermore, SAGE yields very extensive transcriptome information (Smolenski et al., 2004).

The ability of SAGE to provide an accurate measure of gene expression profiles is dependent upon the extent to which the distribution of tag abundances are a true reflection of the distribution of transcripts in the original mRNA population. Several techniques have been developed to address potential biases inherent in the SAGE experiment process.

SAGE tags represent short sequences that need to be appropriately mapped to the right positions in the genome. The traditional SAGE method (shortSAGE) generates 14bp tags which can theoretically identify 4^{14} (268,435,456) unique sequences, allowing the entire transcriptome of any organism to be potentially represented. In practice, repetitive sequences within the genome result in tags that cannot be unambiguously assigned to a particular gene. LongSAGE improves mapping efficiency by extending the length of the resulting tags to 21bp (Li et al.,

2006; Saha et al., 2002). Any remaining ambiguous tags can be removed from the analysis through data analysis software such as DiscoverySpace, a graphical application for bioinformatics data analysis which allows mapping of tags and library comparisons (Robertson et al., 2007).

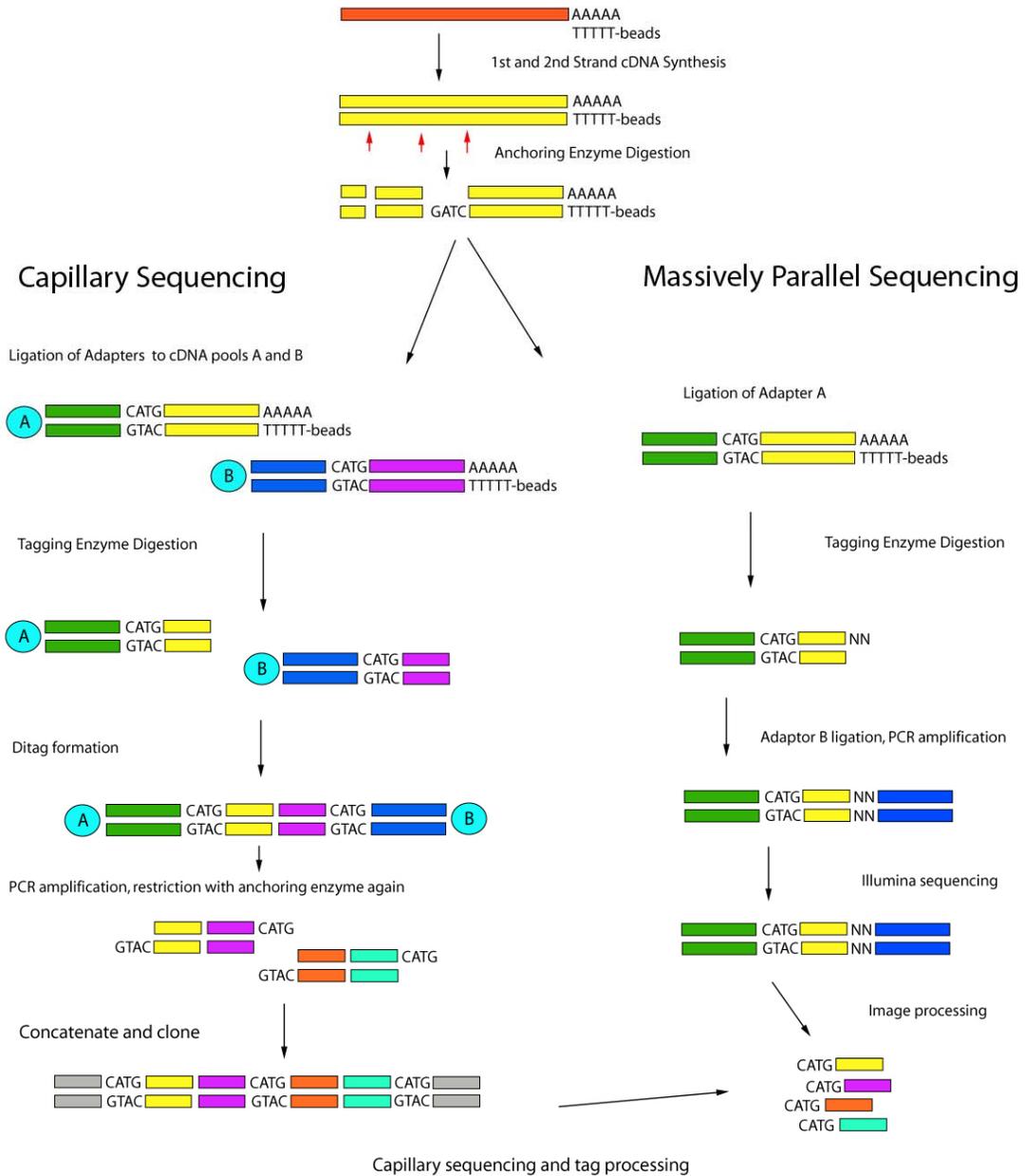
Sampling and sequencing errors disproportionately affect the accuracy and detection of low-abundance transcripts. Recently, a new protocol termed tag sequencing (Tag-seq, also known as digital gene expression, DGE) has been developed to increase the depth of sequencing, thereby increasing sample size (Hanriot et al., 2008; Morrissy et al., 2009). Tag-seq is similar to the LongSAGE protocol, but sequencing employs Illumina's massively parallel sequencing by synthesis protocol in place of conventional Sanger sequencing (Bentley et al., 2008). Typically, a Tag-seq library is sequenced to a depth of 10 million tags, which represents a 100-fold increase in sequencing depth over a typical LongSAGE library. Significantly, the additional depth of sampling provided by Tag-seq leads to a greater number of genes identified in a given tissue, and improves the detectable differences in expression of those genes (Siddiqui et al., 2006). Moreover, comparison of Tag-seq data to LongSAGE data revealed less bias in the sequences observed relative to LongSAGE.

Figure 1.8. Serial analysis of gene expression (SAGE)

Polyadenylated RNA is extracted and transcribed into double-stranded cDNA using a biotinylated oligo(dT) primer. The resulting cDNA is cleaved with a frequent 4bp cutter restriction nuclease (anchoring enzyme – *NlaIII*), and the 3' end fragments are then captured by a streptavidin bead. During the LongSAGE protocol, the cDNA is split into two fractions which are separately ligated to two oligonucleotide adaptors (linkers) A and B. These linkers contain an overhang which is complementary to overhanging ends generated by the anchoring enzyme, a five base recognition sequence for a type IIS restriction nuclease (tagging enzyme cleaving at a defined distance up to 20bp from its recognition site); and some additional sequence to design a specific PCR primer. The linkered cDNA is cleaved with the tagging enzyme resulting in linkers bound to short sequence tags. Pools A and B are then ligated to each other and PCR amplified, resulting in a complex population of sequences each containing two adjacent sequence tags (ditags) derived from the sequences individually coupled to linkers A and B. The ditags can then be released by cleavage with the anchoring enzyme, concatenated and cloned. Capillary sequencing of the resulting clones can then produce long read-outs of sequential ditags of fixed length.

The Tag-seq protocol is similar to the LongSAGE approach, but it avoids ditag production and concatenation, and allows the direct sequencing of tags using massively parallel sequencing on the Illumina Genome Analyzer.

Serial Analysis of Gene Expression



1.6. Project goals

Heart valve development is a complex process. Understanding how gene expression patterns are established and controlled during heart valve formation has important implications for the prevention and treatment of congenital heart defects. Significantly, because many disorders, such as cancer progression, involve the deregulation of processes that occur normally during embryonic development, a detailed characterization of the normal state is essential.

The overall goal of my project was to identify genes that are critical for formation of the heart valves. I hypothesized that these genes would be differentially expressed in the AVC and have dynamic temporal expression patterns.

My approach was to create and analyze SAGE libraries representing different regions and stages of mouse heart development. An original set of 14 heart SAGE libraries was created as part of the Mouse Atlas of gene expression project (Siddiqui et al., 2005). In the Mouse Atlas project over 200 SAGE libraries were created from tissues and stages throughout development, from the single cell zygote to the adult. This comprehensive and quantitative resource was made publicly accessible, and contains data on essentially all genes expressed throughout select stages of mouse development.

Upon the successful completion of the Mouse Atlas project, the MORGEN (Mammalian Organogenesis - Regulation by Gene Expression Networks) project was initiated to investigate the gene networks that regulate the developmental process. As part of the MORGEN project five SAGE libraries were constructed using the massively parallel sequencing technology, focusing on tissues in the heart involved in valve formation and septation of the chambers.

In this project I have used the SAGE technique to analyze the spatial and temporal gene expression patterns controlling the early stages of heart valve formation. In Chapter 2 I will discuss the analysis of temporal expression patterns. In chapter 3 I will discuss the use of deep sequenced SAGE libraries for the identification of cushion specific gene expression. During my analysis, *Twist1* was identified as the most highly expressed transcription factor in the AVC at E10.5. In Chapter 3 I will discuss the creation of a *Twist1* null AVC SAGE library and the analysis of TWIST1's role in controlling differentiation following EMT.

1.7. References

- Abdelwahid, E., Rice, D., Pelliniemi, L. J. and Jokinen, E. (2001) Overlapping and differential localization of Bmp-2, Bmp-4, Msx-2 and apoptosis in the endocardial cushion and adjacent tissues of the developing mouse heart. *Cell Tissue Res.* 305, 67-78.
- Abu-Issa, R. and Kirby, M. L. (2007) Heart field: from mesoderm to heart tube. *Annu.Rev.Cell Dev.Biol.* 23, 45-68.
- Adams, D. J. and van der Weyden, L. (2008) Contemporary approaches for modifying the mouse genome. *Physiol.Genomics* 34, 225-238.
- Akhurst, R. J., Lehnert, S. A., Faissner, A. and Duffie, E. (1990) TGF beta in murine morphogenetic processes: the early embryo and cardiogenesis. *Development* 108, 645-656.
- Ansieau, S., Bastid, J., Doreau, A., Morel, A. P., Bouchet, B. P., Thomas, C., Fauvet, F., Puisieux, I., Doglioni, C., Piccinin, S., *et al.* (2008) Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer.Cell.* 14, 79-89.
- Armstrong, E. J. and Bischoff, J. (2004) Heart valve development: endothelial cell signaling and differentiation. *Circ.Res.* 95, 459-470.
- Azhar, M., Runyan, R. B., Gard, C., Sanford, L. P., Miller, M. L., Andringa, A., Pawlowski, S., Rajan, S. and Doetschman, T. (2009) Ligand-specific function of transforming growth factor beta in epithelial-mesenchymal transition in heart development. *Dev.Dyn.* 238, 431-442.
- Bartram, U., Molin, D. G., Wisse, L. J., Mohamad, A., Sanford, L. P., Doetschman, T., Speer, C. P., Poelmann, R. E. and Gittenberger-de Groot, A. C. (2001) Double-outlet right ventricle and overriding tricuspid valve reflect disturbances of looping, myocardialization, endocardial cushion differentiation, and apoptosis in TGF-beta(2)-knockout mice. *Circulation* 103, 2745-2752.
- Basson, C. T., Cowley, G. S., Solomon, S. D., Weissman, B., Poznanski, A. K., Traill, T. A., Seidman, J. G. and Seidman, C. E. (1994) The clinical and genetic spectrum of the Holt-Oram syndrome (heart-hand syndrome). *N.Engl.J.Med.* 330, 885-891.
- Benedito, R. and Duarte, A. (2005) Expression of Dll4 during mouse embryogenesis suggests multiple developmental roles. *Gene Expr.Patterns* 5, 750-755.
- Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., *et al.* (2008) Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* 456, 53-59.
- Bernanke, D. H. and Markwald, R. R. (1982) Migratory behavior of cardiac cushion tissue cells in a collagen-lattice culture system. *Dev.Biol.* 91, 235-245.
- Bialek, P., Kern, B., Yang, X., Schrock, M., Sosic, D., Hong, N., Wu, H., Yu, K., Ornitz, D. M., Olson, E. N., Justice, M. J. and Karsenty, G. (2004) A twist code determines the onset of osteoblast differentiation. *Dev.Cell.* 6, 423-435.
- Bolender, D. L. and Markwald, R. R. (1979) Epithelial-mesenchymal transformation in chick atrioventricular cushion morphogenesis. *Scan.Electron Microsc.* (3), 313-321.

- Brenner, S., Johnson, M., Bridgham, J., Golda, G., Lloyd, D. H., Johnson, D., Luo, S., McCurdy, S., Foy, M., Ewan, M., *et al.* (2000) Gene expression analysis by massively parallel signature sequencing (MPSS) on microbead arrays. *Nat.Biotechnol.* 18, 630-634.
- Bryantsev, A. L. and Cripps, R. M. (2009) Cardiac gene regulatory networks in *Drosophila*. *Biochim.Biophys.Acta* 1789, 343-353.
- Burgess, R., Cserjesi, P., Ligon, K. L. and Olson, E. N. (1995) Paraxis: a basic helix-loop-helix protein expressed in paraxial mesoderm and developing somites. *Dev.Biol.* 168, 296-306.
- Cai, C. L., Liang, X., Shi, Y., Chu, P. H., Pfaff, S. L., Chen, J. and Evans, S. (2003) Isl1 identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev.Cell.* 5, 877-889.
- Camenisch, T. D., Molin, D. G., Person, A., Runyan, R. B., Gittenberger-de Groot, A. C., McDonald, J. A. and Klewer, S. E. (2002) Temporal and distinct TGFbeta ligand requirements during mouse and avian endocardial cushion morphogenesis. *Dev.Biol.* 248, 170-181.
- Camenisch, T. D., Spicer, A. P., Brehm-Gibson, T., Biesterfeldt, J., Augustine, M. L., Calabro, A., Jr, Kubalak, S., Klewer, S. E. and McDonald, J. A. (2000) Disruption of hyaluronan synthase-2 abrogates normal cardiac morphogenesis and hyaluronan-mediated transformation of epithelium to mesenchyme. *J.Clin.Invest.* 106, 349-360.
- Chakraborty, S., Cheek, J., Sakthivel, B., Aronow, B. J. and Yutzey, K. E. (2008) Shared gene expression profiles in developing heart valves and osteoblast progenitor cells. *Physiol.Genomics*
- Chen, Y. H., Ishii, M., Sucov, H. M. and Maxson, R. E., Jr. (2008) Msx1 and Msx2 are required for endothelial-mesenchymal transformation of the atrioventricular cushions and patterning of the atrioventricular myocardium. *BMC Dev.Biol.* 8, 75.
- Chen, Z. F. and Behringer, R. R. (1995) Twist is Required in Head Mesenchyme for Cranial Neural Tube Morphogenesis. *Genes Dev.* 9, 686-699.
- Cooley, L., Kelley, R. and Spradling, A. (1988) Insertional mutagenesis of the *Drosophila* genome with single P elements. *Science* 239, 1121-1128.
- Corsi, A. K., Brodigan, T. M., Jorgensen, E. M. and Krause, M. (2002) Characterization of a dominant negative *C. elegans* Twist mutant protein with implications for human Saethre-Chotzen syndrome. *Development* 129, 2761-2772.
- Cripps, R. M. and Olson, E. N. (2002) Control of cardiac development by an evolutionarily conserved transcriptional network. *Dev.Biol.* 246, 14-28.
- Cserjesi, P., Brown, D., Ligon, K. L., Lyons, G. E., Copeland, N. G., Gilbert, D. J., Jenkins, N. A. and Olson, E. N. (1995) Scleraxis: a basic helix-loop-helix protein that prefigures skeletal formation during mouse embryogenesis. *Development* 121, 1099-1110.
- De La Cruz, M. V., Sanchez-Gomez, C. and Palomino, M. A. (1989) The primitive cardiac regions in the straight tube heart (Stage 9) and their anatomical expression in the mature heart: An experimental study in the chick embryo. *J.Anat.* 165, 121-131.
- de la Pompa, J. L., Timmerman, L. A., Takimoto, H., Yoshida, H., Elia, A. J., Samper, E., Potter, J., Wakeham, A., Marengere, L., Langille, B. L., Crabtree, G. R. and Mak, T. W. (1998) Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* 392, 182-186.
- de Lange, F. J., Moorman, A. F., Anderson, R. H., Manner, J., Soufan, A. T., de Gier-de Vries, C., Schneider, M. D., Webb, S., van den Hoff, M. J. and Christoffels, V. M. (2004) Lineage and morphogenetic analysis of the cardiac valves. *Circ.Res.* 95, 645-654.

- Delot, E. C., Bahamonde, M. E., Zhao, M. and Lyons, K. M. (2003) BMP signaling is required for septation of the outflow tract of the mammalian heart. *Development* 130, 209-220.
- DeRisi, J. L., Iyer, V. R. and Brown, P. O. (1997) Exploring the metabolic and genetic control of gene expression on a genomic scale. *Science* 278, 680-686.
- Desgrosellier, J. S., Mundell, N. A., McDonnell, M. A., Moses, H. L. and Barnett, J. V. (2005) Activin receptor-like kinase 2 and Smad6 regulate epithelial-mesenchymal transformation during cardiac valve formation. *Dev.Biol.* 280, 201-210.
- Dickson, M. C., Slager, H. G., Duffie, E., Mummery, C. L. and Akhurst, R. J. (1993) RNA and protein localisations of TGF beta 2 in the early mouse embryo suggest an involvement in cardiac development. *Development* 117, 625-639.
- Eisenberg, L. M. and Markwald, R. R. (1995) Molecular regulation of atrioventricular valvuloseptal morphogenesis. *Circ.Res.* 77, 1-6.
- Engleka, K. A., Gitler, A. D., Zhang, M., Zhou, D. D., High, F. A. and Epstein, J. A. (2005) Insertion of Cre into the Pax3 locus creates a new allele of Splotch and identifies unexpected Pax3 derivatives. *Dev.Biol.* 280, 396-406.
- Ewart, A. K., Morris, C. A., Atkinson, D., Jin, W., Sternes, K., Spallone, P., Stock, A. D., Leppert, M. and Keating, M. T. (1993) Hemizyosity at the elastin locus in a developmental disorder, Williams syndrome. *Nat.Genet.* 5, 11-16.
- Firulli, A. B. and Conway, S. J. (2008) Phosphoregulation of Twist1 provides a mechanism of cell fate control. *Curr.Med.Chem.* 15, 2641-2647.
- Fischer, A., Steidl, C., Wagner, T. U., Lang, E., Jakob, P. M., Friedl, P., Knobloch, K. P. and Gessler, M. (2007) Combined loss of Hey1 and HeyL causes congenital heart defects because of impaired epithelial to mesenchymal transition. *Circ.Res.* 100, 856-863.
- Flagg, A. E., Earley, J. U. and Svensson, E. C. (2007) FOG-2 attenuates endothelial-to-mesenchymal transformation in the endocardial cushions of the developing heart. *Dev.Biol.* 304, 308-316.
- Furumatsu, T., Shukunami, C., Amemiya-Kudo, M., Shimano, H. and Ozaki, T. (2010) Scleraxis and E47 cooperatively regulate the Sox9-dependent transcription. *Int.J.Biochem.Cell Biol.* 42, 148-156.
- Galvin, K. M., Donovan, M. J., Lynch, C. A., Meyer, R. I., Paul, R. J., Lorenz, J. N., Fairchild-Huntress, V., Dixon, K. L., Dunmore, J. H., Gimbrone, M. A., Jr, Falb, D. and Huszar, D. (2000) A role for smad6 in development and homeostasis of the cardiovascular system. *Nat.Genet.* 24, 171-174.
- Garg, V., Kathiriya, I. S., Barnes, R., Schluterman, M. K., King, I. N., Butler, C. A., Rothrock, C. R., Eapen, R. S., Hirayama-Yamada, K., Joo, K., *et al.* (2003) GATA4 mutations cause human congenital heart defects and reveal an interaction with TBX5. *Nature* 424, 443-447.
- Gaussin, V., Morley, G. E., Cox, L., Zwijsen, A., Vance, K. M., Emile, L., Tian, Y., Liu, J., Hong, C., Myers, D., *et al.* (2005) Alk3/Bmpr1a receptor is required for development of the atrioventricular canal into valves and annulus fibrosus. *Circ.Res.* 97, 219-226.
- Gitler, A. D., Lu, M. M., Jiang, Y. Q., Epstein, J. A. and Gruber, P. J. (2003) Molecular markers of cardiac endocardial cushion development. *Dev.Dyn.* 228, 643-650.
- Glickman, N. S. and Yelon, D. (2002) Cardiac development in zebrafish: coordination of form and function. *Semin.Cell Dev.Biol.* 13, 507-513.

- Gong, X. Q. and Li, L. (2002) Dermo-1, a multifunctional basic helix-loop-helix protein, represses MyoD transactivation via the HLH domain, MEF2 interaction, and chromatin deacetylation. *J.Biol.Chem.* 277, 12310-12317.
- Greenberg, F. (1993) DiGeorge syndrome: an historical review of clinical and cytogenetic features. *J.Med.Genet.* 30, 803-806.
- Hall, B. K. and Miyake, T. (2000) All for one and one for all: condensations and the initiation of skeletal development. *Bioessays* 22, 138-147.
- Hamamori, Y., Sartorelli, V., Ogryzko, V., Puri, P. L., Wu, H. Y., Wang, J. Y., Nakatani, Y. and Kedes, L. (1999) Regulation of histone acetyltransferases p300 and PCAF by the bHLH protein twist and adenoviral oncoprotein E1A. *Cell* 96, 405-413.
- Hamamori, Y., Wu, H. Y., Sartorelli, V. and Kedes, L. (1997) The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol.Cell.Biol.* 17, 6563-6573.
- Hanriot, L., Keime, C., Gay, N., Faure, C., Dossat, C., Wincker, P., Scote-Blachon, C., Peyron, C. and Gandrillon, O. (2008) A combination of LongSAGE with Solexa sequencing is well suited to explore the depth and the complexity of transcriptome. *BMC Genomics* 9, 418.
- Harrelson, Z., Kelly, R. G., Goldin, S. N., Gibson-Brown, J. J., Bollag, R. J., Silver, L. M. and Papaioannou, V. E. (2004) Tbx2 is essential for patterning the atrioventricular canal and for morphogenesis of the outflow tract during heart development. *Development* 131, 5041-5052.
- Harvey, R. P. (2002) Patterning the vertebrate heart. *Nat.Rev.Genet.* 3, 544-556.
- Hata, A., Lagna, G., Massague, J. and Hemmati-Brivanlou, A. (1998) Smad6 inhibits BMP/Smad1 signaling by specifically competing with the Smad4 tumor suppressor. *Genes Dev.* 12, 186-197.
- Hayashi, M., Nimura, K., Kashiwagi, K., Harada, T., Takaoka, K., Kato, H., Tamai, K. and Kaneda, Y. (2007) Comparative roles of Twist-1 and Id1 in transcriptional regulation by BMP signaling. *J.Cell.Sci.* 120, 1350-1357.
- Henderson, D. J. and Copp, A. J. (1998) Versican expression is associated with chamber specification, septation, and valvulogenesis in the developing mouse heart. *Circ.Res.* 83, 523-532.
- Hinton, R. B., Jr, Lincoln, J., Deutsch, G. H., Osinska, H., Manning, P. B., Benson, D. W. and Yutzey, K. E. (2006) Extracellular matrix remodeling and organization in developing and diseased aortic valves. *Circ.Res.* 98, 1431-1438.
- Hoffman, J. I. and Kaplan, S. (2002) The incidence of congenital heart disease. *J.Am.Coll.Cardiol.* 39, 1890-1900.
- Hoffman, J. I. (1995) Incidence of congenital heart disease: II. Prenatal incidence. *Pediatr.Cardiol.* 16, 155-165.
- Hutson, M. R. and Kirby, M. L. (2007) Model systems for the study of heart development and disease. Cardiac neural crest and conotruncal malformations. *Semin.Cell Dev.Biol.* 18, 101-110.
- Jenkins, K. J., Correa, A., Feinstein, J. A., Botto, L., Britt, A. E., Daniels, S. R., Elixson, M., Warnes, C. A., Webb, C. L. and American Heart Association Council on Cardiovascular Disease in the Young. (2007) Noninherited risk factors and congenital cardiovascular defects: current knowledge: a scientific statement from the American Heart Association

- Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 2995-3014.
- Jones, C. M., Lyons, K. M. and Hogan, B. L. (1991) Involvement of Bone Morphogenetic Protein-4 (BMP-4) and Vgr-1 in morphogenesis and neurogenesis in the mouse. *Development* 111, 531-542.
- Kelly, R. G., Brown, N. A. and Buckingham, M. E. (2001) The arterial pole of the mouse heart forms from Fgf10-expressing cells in pharyngeal mesoderm. *Dev.Cell.* 1, 435-440.
- Kim, R. Y., Robertson, E. J. and Solloway, M. J. (2001) Bmp6 and Bmp7 are required for cushion formation and septation in the developing mouse heart. *Dev.Biol.* 235, 449-466.
- Kokubo, H., Tomita-Miyagawa, S., Hamada, Y. and Saga, Y. (2007) Hesr1 and Hesr2 regulate atrioventricular boundary formation in the developing heart through the repression of Tbx2. *Development* 134, 747-755.
- Kruithof, B. P., Krawitz, S. A. and Gaussin, V. (2007) Atrioventricular valve development during late embryonic and postnatal stages involves condensation and extracellular matrix remodeling. *Dev.Biol.* 302, 208-217.
- Kwok, W. K., Ling, M. T., Lee, T. W., Lau, T. C., Zhou, C., Zhang, X., Chua, C. W., Chan, K. W., Chan, F. L., Glackin, C., Wong, Y. C. and Wang, X. (2005) Up-regulation of TWIST in prostate cancer and its implication as a therapeutic target. *Cancer Res.* 65, 5153-5162.
- Lee, K., Khoshnood, B., Chen, L., Wall, S. N., Cromie, W. J. and Mittendorf, R. L. (2001) Infant mortality from congenital malformations in the United States, 1970-1997. *Obstet.Gynecol.* 98, 620-627.
- Lee, Y. S., Lee, H. H., Park, J., Yoo, E. J., Glackin, C. A., Choi, Y. I., Jeon, S. H., Seong, R. H., Park, S. D. and Kim, J. B. (2003) Twist2, a novel ADD1/SREBP1c interacting protein, represses the transcriptional activity of ADD1/SREBP1c. *Nucleic Acids Res.* 31, 7165-7174.
- Levay, A. K., Peacock, J. D., Lu, Y., Koch, M., Hinton, R. B., Jr, Kadler, K. E. and Lincoln, J. (2008) Scleraxis is required for cell lineage differentiation and extracellular matrix remodeling during murine heart valve formation in vivo. *Circ.Res.* 103, 948-956.
- Levayer, R. and Lecuit, T. (2008) Breaking down EMT. *Nat.Cell Biol.* 10, 757-759.
- Li, L., Krantz, I. D., Deng, Y., Genin, A., Banta, A. B., Collins, C. C., Qi, M., Trask, B. J., Kuo, W. L., Cochran, J., *et al.* (1997) Alagille syndrome is caused by mutations in human Jagged1, which encodes a ligand for Notch1. *Nat.Genet.* 16, 243-251.
- Li, L., Cserjesi, P. and Olson, E. N. (1995) Dermo-1: a novel twist-related bHLH protein expressed in the developing dermis. *Dev.Biol.* 172, 280-292.
- Li, Y. J., Xu, P., Qin, X., Schmechel, D. E., Hulette, C. M., Haines, J. L., Pericak-Vance, M. A. and Gilbert, J. R. (2006) A comparative analysis of the information content in long and short SAGE libraries. *BMC Bioinformatics* 7, 504.
- Lincoln, J., Kist, R., Scherer, G. and Yutzey, K. E. (2007) Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev.Biol.* 305, 120-132.
- Lincoln, J., Lange, A. W. and Yutzey, K. E. (2006) Hearts and bones: shared regulatory mechanisms in heart valve, cartilage, tendon, and bone development. *Dev.Biol.* 294, 292-302.
- Lincoln, J., Alfieri, C. M. and Yutzey, K. E. (2004) Development of heart valve leaflets and supporting apparatus in chicken and mouse embryos. *Dev.Dyn.* 230, 239-250.

- Lipshutz, R. J., Fodor, S. P., Gingeras, T. R. and Lockhart, D. J. (1999) High density synthetic oligonucleotide arrays. *Nat.Genet.* 21, 20-24.
- Loomes, K. M., Taichman, D. B., Glover, C. L., Williams, P. T., Markowitz, J. E., Piccoli, D. A., Baldwin, H. S. and Oakey, R. J. (2002) Characterization of Notch receptor expression in the developing mammalian heart and liver. *Am.J.Med.Genet.* 112, 181-189.
- Ma, L., Lu, M. F., Schwartz, R. J. and Martin, J. F. (2005) Bmp2 is essential for cardiac cushion epithelial-mesenchymal transition and myocardial patterning. *Development* 132, 5601-5611.
- Maitra, M., Schluterman, M. K., Nichols, H. A., Richardson, J. A., Lo, C. W., Srivastava, D. and Garg, V. (2009) Interaction of Gata4 and Gata6 with Tbx5 is critical for normal cardiac development. *Dev.Biol.* 326, 368-377.
- Manner, J. (2004) On rotation, torsion, lateralization, and handedness of the embryonic heart loop: new insights from a simulation model for the heart loop of chick embryos. *Anat.Rec.A.Discov.Mol.Cell.Evol.Biol.* 278, 481-492.
- Markwald, R. R., Fitzharris, T. P. and Manasek, F. J. (1977) Structural development of endocardial cushions. *Am.J.Anat.* 148, 85-119.
- Markwald, R. R., Fitzharris, T. P. and Smith, W. N. (1975) Structural analysis of endocardial cytodifferentiation. *Dev.Biol.* 42, 160-180.
- Martinsen, B. J. (2005) Reference guide to the stages of chick heart embryology. *Dev.Dyn.* 233, 1217-1237.
- Massari, M. E. and Murre, C. (2000) Helix-loop-helix proteins: regulators of transcription in eucaryotic organisms. *Mol.Cell.Biol.* 20, 429-440.
- McCright, B., Lozier, J. and Gridley, T. (2002) A mouse model of Alagille syndrome: Notch2 as a genetic modifier of Jag1 haploinsufficiency. *Development* 129, 1075-1082.
- Mironchik, Y., Winnard, P. T., Jr, Vesuna, F., Kato, Y., Wildes, F., Pathak, A. P., Kominsky, S., Artemov, D., Bhujwalla, Z., Van Diest, P., *et al.* (2005) Twist overexpression induces in vivo angiogenesis and correlates with chromosomal instability in breast cancer. *Cancer Res.* 65, 10801-10809.
- Mjaatvedt, C. H., Nakaoka, T., Moreno-Rodriguez, R., Norris, R. A., Kern, M. J., Eisenberg, C. A., Turner, D. and Markwald, R. R. (2001) The outflow tract of the heart is recruited from a novel heart-forming field. *Dev.Biol.* 238, 97-109.
- Mjaatvedt, C. H., Yamamura, H., Capehart, A. A., Turner, D. and Markwald, R. R. (1998) The *Cspg2* gene, disrupted in the *hdf* mutant, is required for right cardiac chamber and endocardial cushion formation. *Dev.Biol.* 202, 56-66.
- Moorman, A. F., Christoffels, V. M., Anderson, R. H. and van den Hoff, M. J. (2007) The heart-forming fields: one or multiple? *Philos.Trans.R.Soc.Lond.B.Biol.Sci.* 362, 1257-1265.
- Moorman, A. F. and Christoffels, V. M. (2003) Cardiac chamber formation: development, genes, and evolution. *Physiol.Rev.* 83, 1223-1267.
- Morrissy, A. S., Morin, R. D., Delaney, A., Zeng, T., McDonald, H., Jones, S., Zhao, Y., Hirst, M. and Marra, M. A. (2009) Next-generation tag sequencing for cancer gene expression profiling. *Genome Res.* 19, 1825-1835.
- Moskowitz, I. P., Kim, J. B., Moore, M. L., Wolf, C. M., Peterson, M. A., Shendure, J., Nobrega, M. A., Yokota, Y., Berul, C., Izumo, S., Seidman, J. G. and Seidman, C. E. (2007) A molecular pathway including *Id2*, *Tbx5*, and *Nkx2-5* required for cardiac conduction system development. *Cell* 129, 1365-1376.

- Muir, T., Wilson-Rawls, J., Stevens, J. D., Rawls, A., Schweitzer, R., Kang, C. and Skinner, M. K. (2008) Integration of CREB and bHLH transcriptional signaling pathways through direct heterodimerization of the proteins: role in muscle and testis development. *Mol.Reprod.Dev.* 75, 1637-1652.
- Murchison, N. D., Price, B. A., Conner, D. A., Keene, D. R., Olson, E. N., Tabin, C. J. and Schweitzer, R. (2007) Regulation of tendon differentiation by scleraxis distinguishes force-transmitting tendons from muscle-anchoring tendons. *Development* 134, 2697-2708.
- Nakamura, T., Colbert, M. C. and Robbins, J. (2006) Neural crest cells retain multipotential characteristics in the developing valves and label the cardiac conduction system. *Circ.Res.* 98, 1547-1554.
- Newbury-Ecob, R. A., Leanage, R., Raeburn, J. A. and Young, I. D. (1996) Holt-Oram syndrome: a clinical genetic study. *J.Med.Genet.* 33, 300-307.
- Niessen, K., Fu, Y., Chang, L., Hoodless, P. A., McFadden, D. and Karsan, A. (2008) Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J.Cell Biol.* 182, 315-325.
- Niessen, K. and Karsan, A. (2008) Notch signaling in cardiac development. *Circ.Res.* 102, 1169-1181.
- Oda, T., Elkahoun, A. G., Pike, B. L., Okajima, K., Krantz, I. D., Genin, A., Piccoli, D. A., Meltzer, P. S., Spinner, N. B., Collins, F. S. and Chandrasekharappa, S. C. (1997) Mutations in the human *Jagged1* gene are responsible for Alagille syndrome. *Nat.Genet.* 16, 235-242.
- Ohuchida, K., Mizumoto, K., Ohhashi, S., Yamaguchi, H., Konomi, H., Nagai, E., Yamaguchi, K., Tsuneyoshi, M. and Tanaka, M. (2007) Twist, a novel oncogene, is upregulated in pancreatic cancer: clinical implication of Twist expression in pancreatic juice. *Int.J.Cancer* 120, 1634-1640.
- Person, A. D., Klewer, S. E. and Runyan, R. B. (2005) Cell biology of cardiac cushion development. *Int.Rev.Cytol.* 243, 287-335.
- Phillips, M. T., Kirby, M. L. and Forbes, G. (1987) Analysis of cranial neural crest distribution in the developing heart using quail-chick chimeras. *Circ.Res.* 60, 27-30.
- Pierpont, M. E., Basson, C. T., Benson, D. W., Jr, Gelb, B. D., Giglia, T. M., Goldmuntz, E., McGee, G., Sable, C. A., Srivastava, D., Webb, C. L. and American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young. (2007) Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 3015-3038.
- Poelmann, R. E., Molin, D., Wisse, L. J. and Gittenberger-de Groot, A. C. (2000) Apoptosis in cardiac development. *Cell Tissue Res.* 301, 43-52.
- Poelmann, R. E., Mikawa, T. and Gittenberger-de Groot, A. C. (1998) Neural crest cells in outflow tract septation of the embryonic chicken heart: differentiation and apoptosis. *Dev.Dyn.* 212, 373-384.
- Qayyum, S. R., Webb, S., Anderson, R. H., Verbeek, F. J., Brown, N. A. and Richardson, M. K. (2001) Septation and valvar formation in the outflow tract of the embryonic chick heart. *Anat.Rec.* 264, 273-283.

- Ranger, A. M., Grusby, M. J., Hodge, M. R., Gravalles, E. M., de la Brousse, F. C., Hoey, T., Mickanin, C., Baldwin, H. S. and Glimcher, L. H. (1998) The transcription factor NF-ATc is essential for cardiac valve formation. *Nature* 392, 186-190.
- Reamon-Buettner, S. M. and Borlak, J. (2006) HEY2 mutations in malformed hearts. *Hum.Mutat.* 27, 118.
- Robertson, N., Oveisi-Fordorei, M., Zuyderduyn, S. D., Varhol, R. J., Fjell, C., Marra, M., Jones, S. and Siddiqui, A. (2007) DiscoverySpace: an interactive data analysis application. *Genome Biol.* 8, R6.
- Romano, L. A. and Runyan, R. B. (2000) Slug is an essential target of TGFbeta2 signaling in the developing chicken heart. *Dev.Biol.* 223, 91-102.
- Rossant, J. (1996) Mouse mutants and cardiac development: new molecular insights into cardiogenesis. *Circ.Res.* 78, 349-353.
- Rutenberg, J. B., Fischer, A., Jia, H., Gessler, M., Zhong, T. P. and Mercola, M. (2006) Developmental patterning of the cardiac atrioventricular canal by Notch and Hairy-related transcription factors. *Development* 133, 4381-4390.
- Saha, S., Sparks, A. B., Rago, C., Akmaev, V., Wang, C. J., Vogelstein, B., Kinzler, K. W. and Velculescu, V. E. (2002) Using the transcriptome to annotate the genome. *Nat.Biotechnol.* 20, 508-512.
- Sanford, L. P., Ormsby, I., Gittenberger-de Groot, A. C., Sariola, H., Friedman, R., Boivin, G. P., Cardell, E. L. and Doetschman, T. (1997) TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. *Development* 124, 2659-2670.
- Schena, M., Shalon, D., Davis, R. W. and Brown, P. O. (1995) Quantitative monitoring of gene expression patterns with a complementary DNA microarray. *Science* 270, 467-470.
- Schlange, T., Andree, B., Arnold, H. H. and Brand, T. (2000) BMP2 is required for early heart development during a distinct time period. *Mech.Dev.* 91, 259-270.
- Schott, J. J., Benson, D. W., Basson, C. T., Pease, W., Silberbach, G. M., Moak, J. P., Maron, B. J., Seidman, C. E. and Seidman, J. G. (1998) Congenital heart disease caused by mutations in the transcription factor NKX2-5. *Science* 281, 108-111.
- Shelton, E. L. and Yutzey, K. E. (2008) Twist1 function in endocardial cushion cell proliferation, migration, and differentiation during heart valve development. *Dev.Biol.* 317, 282-295.
- Shelton, E. L. and Yutzey, K. E. (2007) Tbx20 regulation of endocardial cushion cell proliferation and extracellular matrix gene expression. *Dev.Biol.* 302, 376-388.
- Shi, Y. and Massague, J. (2003) Mechanisms of TGF-beta signaling from cell membrane to the nucleus. *Cell* 113, 685-700.
- Shi, Y., Katsev, S., Cai, C. and Evans, S. (2000) BMP signaling is required for heart formation in vertebrates. *Dev.Biol.* 224, 226-237.
- Siddiqui, A. S., Delaney, A. D., Schnerch, A., Griffith, O. L., Jones, S. J. and Marra, M. A. (2006) Sequence biases in large scale gene expression profiling data. *Nucleic Acids Res.* 34, e83.
- Siddiqui, A. S., Khattra, J., Delaney, A. D., Zhao, Y., Astell, C., Asano, J., Babakaiff, R., Barber, S., Beland, J., Bohacec, S., *et al.* (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc.Natl.Acad.Sci.U.S.A.* 102, 18485-18490.
- Simpson, P. (1983) Maternal-Zygotic Gene Interactions during Formation of the Dorsal-Ventral Pattern in Drosophila Embryos. *Genetics* 105, 615-632.

- Smolenski, A., Schultess, J., Danielewski, O., Garcia Arguinzonis, M. I., Thalheimer, P., Kneitz, S., Walter, U. and Lohmann, S. M. (2004) Quantitative analysis of the cardiac fibroblast transcriptome-implications for NO/cGMP signaling. *Genomics* 83, 577-587.
- Song, L., Fassler, R., Mishina, Y., Jiao, K. and Baldwin, H. S. (2007) Essential functions of Alk3 during AV cushion morphogenesis in mouse embryonic hearts. *Dev.Biol.* 301, 276-286.
- Sosic, D., Richardson, J. A., Yu, K., Ornitz, D. M. and Olson, E. N. (2003) Twist regulates cytokine gene expression through a negative feedback loop that represses NF-kappaB activity. *Cell* 112, 169-180.
- Spicer, D. B., Rhee, J., Cheung, W. L. and Lassar, A. B. (1996) Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* 272, 1476-1480.
- Srivastava, D. (2006) Making or breaking the heart: from lineage determination to morphogenesis. *Cell* 126, 1037-1048.
- Sugi, Y., Yamamura, H., Okagawa, H. and Markwald, R. R. (2004) Bone morphogenetic protein-2 can mediate myocardial regulation of atrioventricular cushion mesenchymal cell formation in mice. *Dev.Biol.* 269, 505-518.
- Takase, M., Imamura, T., Sampath, T. K., Takeda, K., Ichijo, H., Miyazono, K. and Kawabata, M. (1998) Induction of Smad6 mRNA by bone morphogenetic proteins. *Biochem.Biophys.Res.Commun.* 244, 26-29.
- Terrett, J. A., Newbury-Ecob, R., Cross, G. S., Fenton, I., Raeburn, J. A., Young, I. D. and Brook, J. D. (1994) Holt-Oram syndrome is a genetically heterogeneous disease with one locus mapping to human chromosome 12q. *Nat.Genet.* 6, 401-404.
- Thisse, B., el Messal, M. and Perrin-Schmitt, F. (1987) The twist gene: isolation of a Drosophila zygotic gene necessary for the establishment of dorsoventral pattern. *Nucleic Acids Res.* 15, 3439-3453.
- Timmerman, L. A., Grego-Bessa, J., Raya, A., Bertran, E., Perez-Pomares, J. M., Diez, J., Aranda, S., Palomo, S., McCormick, F., Izpisua-Belmonte, J. C. and de la Pompa, J. L. (2004) Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* 18, 99-115.
- Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. (1995) Serial analysis of gene expression. *Science* 270, 484-487.
- Verzi, M. P., McCulley, D. J., De Val, S., Dodou, E. and Black, B. L. (2005) The right ventricle, outflow tract, and ventricular septum comprise a restricted expression domain within the secondary/anterior heart field. *Dev.Biol.* 287, 134-145.
- Vincentz, J. W., Barnes, R. M., Rodgers, R., Firulli, B. A., Conway, S. J. and Firulli, A. B. (2008) An absence of Twist1 results in aberrant cardiac neural crest morphogenesis. *Dev.Biol.* 320, 131-139.
- Waldo, K. L., Kumiski, D. H., Wallis, K. T., Stadt, H. A., Hutson, M. R., Platt, D. H. and Kirby, M. L. (2001) Conotruncal myocardium arises from a secondary heart field. *Development* 128, 3179-3188.
- Walters, M. J., Wayman, G. A. and Christian, J. L. (2001) Bone morphogenetic protein function is required for terminal differentiation of the heart but not for early expression of cardiac marker genes. *Mech.Dev.* 100, 263-273.
- Wang, J., Sridurongrit, S., Dudas, M., Thomas, P., Nagy, A., Schneider, M. D., Epstein, J. A. and Kaartinen, V. (2005) Atrioventricular cushion transformation is mediated by ALK2 in the developing mouse heart. *Dev.Biol.* 286, 299-310.

- Watanabe, Y., Kokubo, H., Miyagawa-Tomita, S., Endo, M., Igarashi, K., Aisaki, K., Kanno, J. and Saga, Y. (2006) Activation of Notch1 signaling in cardiogenic mesoderm induces abnormal heart morphogenesis in mouse. *Development* 133, 1625-1634.
- Wessels, A. and Sedmera, D. (2003) Developmental anatomy of the heart: a tale of mice and man. *Physiol.Genomics* 15, 165-176.
- Wilson-Rawls, J., Rhee, J. M. and Rawls, A. (2004) Paraxis is a basic helix-loop-helix protein that positively regulates transcription through binding to specific E-box elements. *J.Biol.Chem.* 279, 37685-37692.
- Wirrig, E. E., Snarr, B. S., Chintalapudi, M. R., O'neal, J. L., Phelps, A. L., Barth, J. L., Fresco, V. M., Kern, C. B., Mjaatvedt, C. H., Toole, B. P., *et al.* (2007) Cartilage link protein 1 (Crtl1), an extracellular matrix component playing an important role in heart development. *Dev.Biol.* 310, 291-303.
- Wolf, C., Thisse, C., Stoetzel, C., Thisse, B., Gerlinger, P. and Perrin-Schmitt, F. (1991) The M-twist gene of *Mus* is expressed in subsets of mesodermal cells and is closely related to the *Xenopus X-twi* and the *Drosophila twist* genes. *Dev.Biol.* 143, 363-373.
- Ya, J., Schilham, M. W., de Boer, P. A., Moorman, A. F., Clevers, H. and Lamers, W. H. (1998) Sox4-deficiency syndrome in mice is an animal model for common trunk. *Circ.Res.* 83, 986-994.
- Yamagishi, T., Nakajima, Y., Miyazono, K. and Nakamura, H. (1999) Bone morphogenetic protein-2 acts synergistically with transforming growth factor-beta3 during endothelial-mesenchymal transformation in the developing chick heart. *J.Cell.Physiol.* 180, 35-45.
- Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R. A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117, 927-939.
- Yutzey, K. E. and Robbins, J. (2007) Principles of genetic murine models for cardiac disease. *Circulation* 115, 792-799.
- Zhang, H. and Bradley, A. (1996) Mice deficient for BMP2 are nonviable and have defects in amnion/chorion and cardiac development. *Development* 122, 2977-2986.
- Zhang, Z., Xie, D., Li, X., Wong, Y. C., Xin, D., Guan, X. Y., Chua, C. W., Leung, S. C., Na, Y. and Wang, X. (2007) Significance of TWIST expression and its association with E-cadherin in bladder cancer. *Hum.Pathol.* 38, 598-606.

CHAPTER 2

GENOMIC ANALYSIS DISTINGUISHES PHASES OF EARLY DEVELOPMENT OF THE MOUSE ATRIO-VENTRICULAR CANAL

2.1. Introduction

Congenital malformations of the cardiovascular system are observed in at least 1% of newborn babies, with abnormal development of the valves and septal structures accounting for a majority of these defects (Hoffman, 1995). Although many congenital valve defects occur as part of well-defined clinical syndromes, the genetic causes for a large proportion remain undetermined. Given the high morbidity and mortality associated with these defects, an increased molecular understanding of the processes involved in valve formation is crucial to the development of new therapies.

In the embryo, the mitral and tricuspid valves and part of the atrio-ventricular septum develop from cardiac cushions that form in the atrio-ventricular canal (AVC). Beginning at embryonic day (E) 9.0-9.5, signals from the myocardium induce an epithelial to mesenchymal transformation (EMT) of endocardial cells in the AVC, which delaminate and invade the cardiac jelly to form endocardial cushions. The Notch and TGF β pathways have emerged as critical regulators of this process. TGF β 2 from the myocardium induces EMT (Bartram et al., 2001; Camenisch et al., 2002; Sanford et al., 1997), while Notch signalling establishes a boundary for EMT responsiveness (Kokubo et al., 2007; Rutenberg et al., 2006), making endocardial cells

A version of this chapter has been published. Vrljicak et al. (2010) Genomic analysis distinguishes phases of early development of the mouse atrio-ventricular canal. *Physiological Genomics* 40(3):150-157.

lining the AVC competent to respond (Fischer et al., 2007; Timmerman et al., 2004; Watanabe et al., 2006).

Although knowledge about the signalling pathways driving the onset of EMT is accumulating, little is known about the downstream targets of these pathways and the mechanisms governing the events following EMT. After initial invasion, the endocardial cushions undergo remodeling through proliferation, differentiation and apoptosis of the newly formed mesenchyme cells (Person et al., 2005). These highly coordinated events are accompanied by changes in gene expression.

Serial Analysis of Gene Expression (SAGE) examines gene expression profiles without prior knowledge of the genes involved. SAGE permits the simultaneous evaluation of thousands of expressed transcripts (Velculescu et al., 1995), generating absolute values that are easily compared and achieving extensive transcriptome coverage (Smolenski et al., 2004). Previous studies have used SAGE to characterize the transcriptome of heart-derived tissue and cell lines (Anisimov et al., 2002a; Anisimov et al., 2002b; Moskowitz et al., 2007; Smolenski et al., 2004).

We describe the first comprehensive SAGE analysis of mouse embryonic AVC development from E9.5 to E12.5. We identified temporal expression patterns in the developing AVC to determine subsets of transcripts likely involved in key developmental stages. We show that these temporal expression patterns are associated with specific functional processes occurring in the AVC. We also define the temporal distribution of mesenchyme genes during EMT and of specific Notch and TGF β targets. This resource will be useful in the identification of novel AVC genes and the study of their function and regulation during early valve development.

2.2. Materials and methods

2.2.1. Serial analysis of gene expression

C57BL/6J mouse embryos were dissected following procedures approved by the Animal Care Committee at the University of British Columbia. Mice were assigned to the appropriate Theiler stage at the time of tissue collection to ensure uniformity in the classification of developmental stages. To isolate the heart, embryos of the appropriate stage were collected from timed-pregnant females and placed in ice-cold PBS. AVCs containing endothelial, mesenchyme and myocardial cell populations were manually dissected using 30½G needles. Samples from multiple litters were pooled to obtain sufficient RNA for SAGE library construction and avoid possible bias created by spontaneous mutations in the colony. Blood was removed by puncturing the heart chambers and washing the tissue with PBS. Further details are available from the Mouse Atlas of Gene Expression website (<http://www.mouseatlas.org/>). Dissected mouse tissue samples were collected in either RNAlater (Ambion) or TRIzol reagent (Invitrogen). After RNA isolation RNA quality was assessed using an Agilent Bioanalyzer and the RNA was stored at -80 °C until SAGE library construction.

SAGE libraries were constructed using standard protocols (Siddiqui et al., 2005), and the data are available at <http://www.mouseatlas.org/>, CGAP (<http://cgap.nci.nih.gov/>) and GEO (<http://www.ncbi.nlm.nih.gov/geo/>). SAGE data was analyzed and mapped to genes using DiscoverySpace v4.0 (Robertson et al., 2007). Unambiguous sense mappings to the RefSeq

database (<http://www.ncbi.nlm.nih.gov/RefSeq/>) were used. Genes represented by multiple tag-types were manually curated to determine the most likely representative.

2.2.2. Cluster analysis

Poisson-based k-means clustering (PoissonC) for SAGE data (Cai et al., 2004) was conducted over 100 iterations and the data combined to obtain consensus clusters. To estimate the optimal number of clusters (k), the within-cluster dispersion was computed as described (Blackshaw et al., 2004) for values of k from one to 50 over 50 iterations. The largest drop in within-cluster dispersion occurred for k values of five to 15. K=14 was chosen after visual inspection of the resulting patterns.

Gene Ontology (GO) term enrichment analysis was performed using Expression Analysis Systematic Explorer (EASE) as described (Blackshaw et al., 2004; Hosack et al., 2003). Raw EASE scores of less than 0.05 were considered significant.

RT-qPCR and *in situ* hybridization was described previously (McKnight et al., 2007). Primers for qPCR and generation of *in situ* hybridization probes are found in Appendix I. *Twist1*, *Periostin* and *Tbx20* probes were described previously (Chen and Behringer, 1995; Kraus et al., 2001; Kruzynska-Frejtag et al., 2001).

2.2.3. Mesenchyme-epithelial score

To determine the enrichment of a tag-type in epithelium or mesenchyme, six epithelial-mesenchymal library pairs from the Mouse Atlas dataset (Siddiqui et al., 2005) were used to calculate a mesenchyme-epithelial score (Appendix II). These libraries were created by

dissociating the epithelium and mesenchyme from a given organ by trypsin digestion and manual dissection. Epithelium and mesenchyme were separately processed to generate matched pairs. Using Audic-Claverie statistics (Audic and Claverie, 1997), a p-value was generated to quantify the probability that a tag is differentially expressed in a given library pair. One minus the p-value was calculated and a sign was assigned according to the expression pattern of the tag-type with a positive sign for mesenchyme enrichment and a negative sign for epithelium enrichment. The individual pair-wise scores were added up into an overall score which ranged from +6 for a perfect mesenchyme marker to -6 for a perfect epithelial marker.

2.3. Results

To examine global gene expression changes during the initial phase of heart valve formation, we generated and analyzed four SAGE libraries from the mouse AVC at E9.5 to E12.5. These libraries represent endothelial, mesenchyme and myocardial cell expression in the AVC and cover the stages from initiation of EMT through to the beginning of cushion remodeling. For SAGE libraries a sequencing depth of 120,000 tags is comparable to fluorescent-based microarray approaches (Lu et al., 2004), and new transcript discovery reaches a plateau at 300,000 tags (Akmaev and Wang, 2004). We therefore sequenced a minimum of 300,000 tags per library for a total of 1,274,310 tags representing 211,971 different tag-types (Table 2.1).

Table 2.1. Tissues and stages sampled

Atrio-ventricular canal (AVC) tissue was isolated from four time-points for SAGE library construction.

| Library ID | Description | Hearts used ^a | Total tags ^b | Tag-types ^c | RefSeq ^d | Tag count ^e | | | |
|-----------------|---------------------------------|--------------------------|-------------------------|------------------------|---------------------|------------------------|--------|--------|-------|
| | | | | | | 1 | 2-5 | 6-49 | ≥50 |
| SM206 and SM246 | E9.5 AVC (Theiler stage 15) | 79 | 314,093 | 65,103 | 24,840 | 47,395 | 12,003 | 5,065 | 640 |
| SM234 | E10.5 AVC (Theiler stage 17) | 25 | 301,949 | 75,375 | 30,620 | 54,504 | 14,141 | 6,104 | 626 |
| SM008 and SM241 | E11.5 AVC (Theiler stage 19) | 48 | 345,463 | 82,704 | 28,261 | 61,311 | 14,022 | 6,689 | 682 |
| SM238 | E12.5 AVC (Theiler stage 20) | 18 | 312,805 | 71,385 | 28,519 | 51,861 | 13,447 | 5,470 | 607 |
| Total | | | 1,274,310 | 211,971 | 57,629 | 158,211 | 53,613 | 23,328 | 2,550 |

^a Number of hearts used for RNA sample collection.

^b Total number of tags sequenced.

^c Unique tag sequences.

^d Mappings to the RefSeq database.

^e Proportion of tags at different expression levels.

In these libraries, 71% of the tags mapped to known transcripts using the RefSeq database. A further 8% of the tags mapped to the genome, likely representing unannotated transcripts. Of the remaining tags, many were only present once and may have been generated by sequencing, PCR, or other errors (Stollberg et al., 2000). However, 4,869 unmapped tag-types were found at a significant level (> 5 tags) suggesting these tags may represent valid, novel, transcripts. Importantly, these AVC SAGE libraries exhibit a large dynamic range, covering over 3 orders of magnitude in expression levels from genes with greater than 1,000 tags for myosins (e.g. *Myl2*, *Myl4* and *Myl7*), to genes with only a few tags (e.g. transcription factors).

To test whether our SAGE libraries offer an accurate portrayal of transcripts present during the EMT process, we searched for tags representing genes previously shown to be involved in AVC development. We observed tags representing all of these genes, including the highly expressed extracellular matrix (ECM) molecule *Periostin* (Kruzynska-Frejtag et al., 2001), and the lowly expressed transcription factor *Tbx20* (Kraus et al., 2001) (Appendix IV). Furthermore, tags for genes known to be more highly expressed in the AVC were overrepresented in the AVC SAGE libraries compared to other heart libraries in the Mouse Atlas (Siddiqui et al., 2005) (Appendices III and V). These results suggest that our SAGE library dataset is sufficiently comprehensive at 300,000 tags per library to recapitulate patterns of gene expression observed *in vivo* over a large dynamic range. Significantly, the vast majority of tags in our libraries represent genes not previously characterized in the context of AVC development suggesting that our database is a rich source of novel AVC genes.

2.3.1. Clustering of AVC libraries reveals temporal expression patterns

Genes that share temporal expression patterns may participate in similar biological processes. To determine if coordinated patterns of gene expression could be identified in the developing AVC, we performed cluster analysis. To exclude tags with constant expression levels, we conducted pair-wise comparisons of tag-type expression patterns using Audic-Claverie statistics (Audic and Claverie, 1997) which account for different library sizes and was designed for the quantitative, absolute comparison of SAGE gene expression profiles. We identified 8,839 tag-types representing 3,424 genes with at least one significant difference ($p < 0.05$) and grouped them into clusters using a Poisson model-based k-means algorithm designed specifically for SAGE data (Cai et al., 2004). We used this algorithm over 100 iterations to resolve the tag-types into 14 distinct expression patterns (Figure 2.1). These patterns were ordered by visual inspection and hierarchical clustering on the median expression at each time-point.

To examine the genes represented in each cluster, tag-types were mapped to the RefSeq database (Table 2.2). We used EASE (Hosack et al., 2003) to determine if specific GO terms were overrepresented (Table 2.3). Clusters that showed peak expression at E9.5-10.5 and decrease over time (particularly clusters B-D) contained many markers for endothelial cells (e.g. *Vcam1*, *Edg1*, *Edf1* and *Ednra*) while clusters that increase over time (clusters K-N) included factors involved in ECM structure (e.g. *Mmp2*, *Periostin*, *Biglycan* and collagens). Clusters with peak expression at E10.5 (cluster E), E11.5 (cluster J) and both time-points (clusters G) included many transcription factors and signalling components, particularly of the Notch, Wnt, and TGF β pathways such as *Jag1*, *Bmp2*, and *Gsk3 β* . Interestingly, clusters that peak at E11.5 (clusters I

and J) were significantly enriched for genes involved in cell proliferation ($p < 0.05$ - e.g. *Cdc2a*, *Cdc25c*, and *Dusp1*). The cluster with the most pronounced peak at E12.5 (cluster N) was significantly enriched for cell adhesion and apoptosis genes ($p < 0.05$). Both pro-apoptotic and anti-apoptotic genes (e.g. TNF receptor 12a and *Bcl2l1*, respectively) were represented suggesting a balance between pro-apoptotic and anti-apoptotic signals. Thus, our data show dynamic patterns of gene expression in the developing AVC.

To further confirm our SAGE results, we performed quantitative RT-PCR validation on genes with different temporal expression patterns (Figure 2.2A). Our SAGE and RT-qPCR results were highly correlated even for lowly expressed transcription factors (i.e. *Lef1*, *Tbx20*, *Sox9*, *Twist1*). Interestingly, *in situ* hybridization analysis revealed differences in the spatial expression within the AVC region (Figure 2.2B). *Sox9*, *Twist1* and *Periostin* were expressed exclusively in mesenchyme cells throughout the E9.5-E12.5 timeframe, while *Lef1* and *Tbx20* were also expressed in the myocardium.

Figure 2.1. Cluster analysis of AVC libraries reveals 14 distinct temporal expression patterns

Differentially-expressed tag-types were grouped into clusters using a Poisson model-based k-means algorithm (Cai et al., 2004) over 100 iterations with k=14. SAGE libraries are plotted on the x-axis and percent of tag abundance plotted on the y-axis.

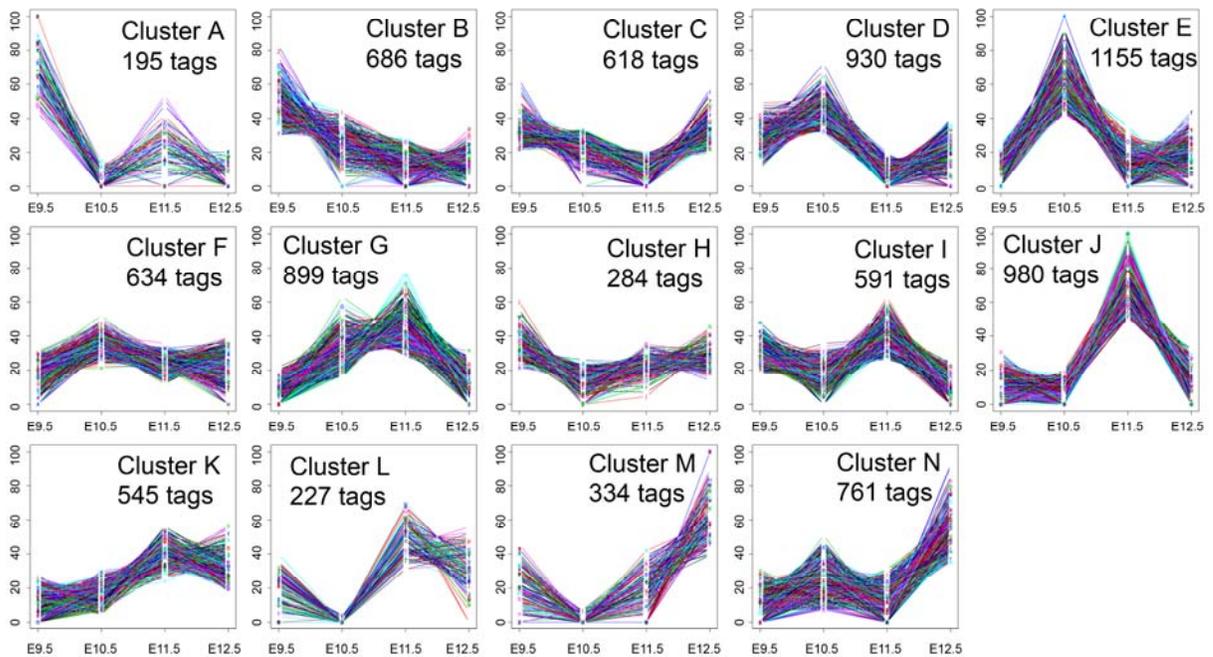


Table 2.2. Summary of AVC temporal clusters

| Cluster | Tags in cluster* | Sense RefSeq mappings† | Selected genes in cluster |
|---------|------------------|------------------------|---|
| A | 195 | 34 | |
| B | 686 | 256 | <i>Nfatc1, Vcam1, Wnt2</i> |
| C | 618 | 257 | <i>Edf1, Itgb1, Msx2, Vegfb</i> |
| D | 930 | 360 | <i>Atf3, Dtx2, Edg1, Ednra, Itga6, Jup, Lef1, Msx1, Nabl, Nkx2-5, Notch1, Tbx2, Tbx20, Tgfb2</i> |
| E | 1,155 | 376 | <i>Bmp2, Gata3, Gsk3β, Hand2, Igf2r, Igfbp5, Jag1, Rras, Smad7, Sox9</i> |
| F | 634 | 312 | <i>Gata6, Igf2, Itgb5, Kitl, Nras, Nrpl, Scx, Smad6, Sox4, Srf, Srl, Vim, Tgfr1, Wnt11</i> |
| G | 899 | 398 | <i>Bmpr1a, Cdh11, Cflar, Dicer, Fbn1, Fgfr3, Foxc2, Igf1, Itga5, Mmp14, Sulfl, Twist1, Vcan, Vegfc, Wt1</i> |
| H | 284 | 127 | <i>Hey2</i> |
| I | 591 | 305 | <i>Pbrm1, Wnt5a</i> |
| J | 980 | 365 | <i>Bmp7, Ctnnb1, Dedd, Eid1, Fdz7, Itga8, Nedd4, Reck, Sod2, Tbx5</i> |
| K | 545 | 283 | <i>Col5a2, Colla1, Colla2, Coll18a1, Col3a1, Col5a1, Col6a2, Id3, Gjal, Krt7, Meox1, Mmp2, Nppa, Smad1, Smad3</i> |
| L | 227 | 81 | <i>Adrbk1, Bmp4, Dlk1, Pkd2, Tgfr2</i> |
| M | 334 | 56 | <i>Cldn5, Efnal</i> |
| N | 761 | 218 | <i>Bcl2l1, Bgn, Cdh13, Col6a1, Dap, Dll4, Dusp1, Elastin, Grina, Itga2b, Itga3, Periostin, Tgfb3, Vwf</i> |
| Total | 8,839 | 3,424 | |

* Number of unique SAGE tags per cluster.

† Fraction of tag-types that mapped unambiguously to the RefSeq database.

Table 2.3. Co-expressed genes share specific functions

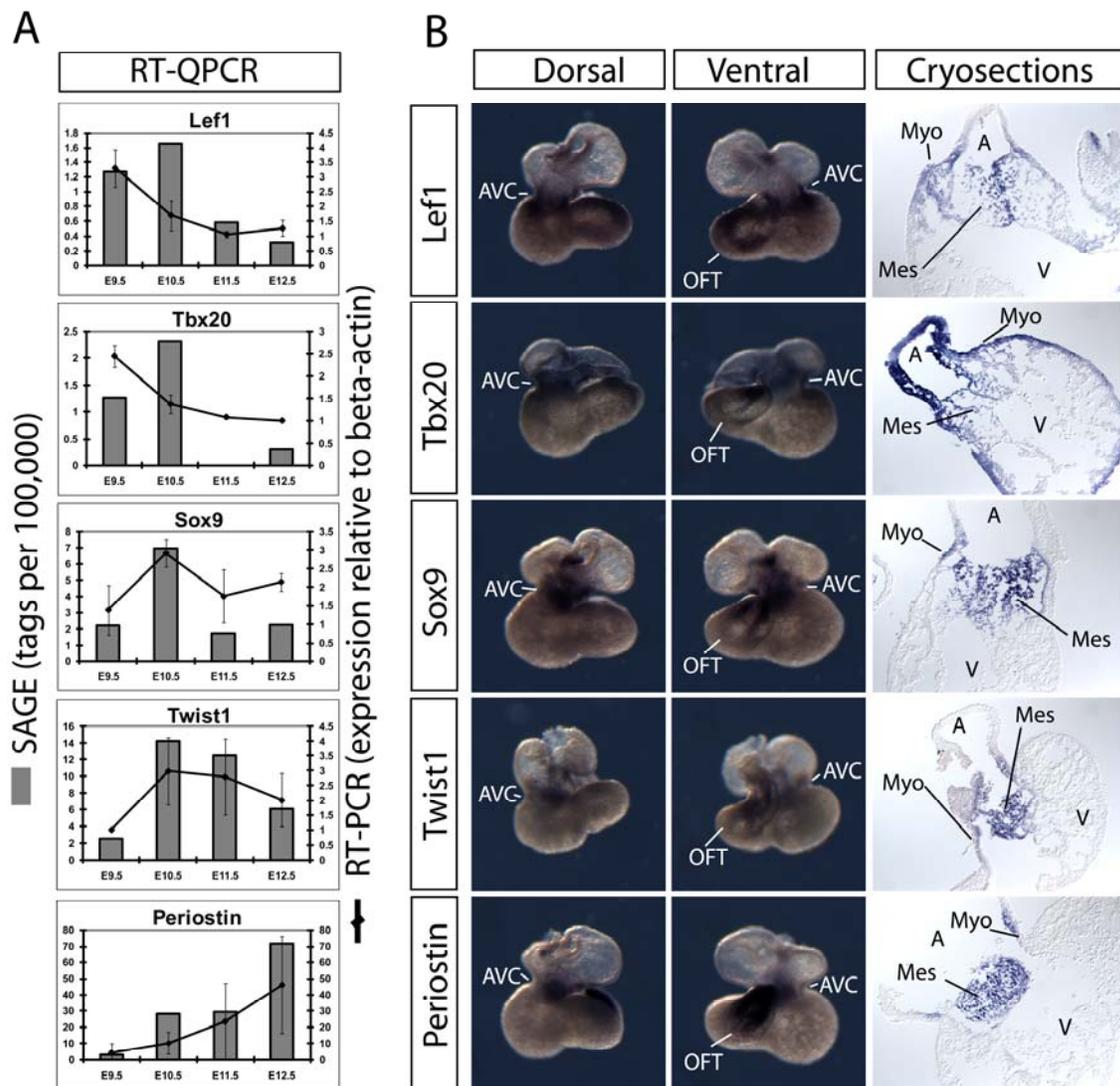
Enriched GO biological process categories were determined for each cluster using EASE (Hosack et al., 2003). P-values represent raw EASE scores for the respective categories. Representative GO categories are indicated.

| Cluster | Biological process |
|---------|---|
| A | |
| B | Biosynthesis p=0.002, amino acid biosynthesis p=0.018 |
| C | Protein biosynthesis p=0.034 |
| D | Protein amino-acid phosphorylation p=0.034, transcription\DNA-dependent p=0.042, energy reserve metabolism p=0.043 |
| E | Morphogenesis p<0.005, skeletal dev. p=0.003, cell communication p=0.009, signal transduction p=0.011 |
| F | Proton transport p=0.043, nucleobase, nucleoside, nucleotide and nucleic acid metabolism p=0.048 |
| G | DNA metabolism p=0.038 |
| H | Protein biosynthesis p=0.010, macromolecule biosynthesis p=0.016 |
| I | Mitotic cell cycle p=0.003, main pathways of carbohydrate metabolism p=0.004, electron transport p=0.021, pyridine nucleotide metabolism p=0.023, protein transport p=0.039 |
| J | Mitosis p=0.020, intracellular signalling cascade p=0.030, intracellular transport p=0.049 |
| K | Fatty acid metabolism p=0.004, protein complex assembly p=0.008, cell adhesion p=0.05 |
| L | |
| M | |
| N | Cell adhesion p=0.002, cell-matrix adhesion p=0.015, communication p=0.024, regulation of apoptosis p=0.033 |

Figure 2.2. Analysis of gene expression by RT-qPCR and *in situ* hybridization

A) Temporal expression patterns of *Lef1*, *Tbx20*, *Sox9*, *Twist1* and *Periostin* as determined by SAGE and RT-qPCR. RT-qPCR results represent average values \pm standard deviations (n=3).

B) Whole-mount *in situ* hybridization was performed on E10.5 hearts. *Sox9*, *Twist1* and *Periostin* are expressed in the mesenchyme cells of the AVC, while *Lef1* and *Tbx20* are also expressed in the myocardium. (A = atria; V = ventricles; AVC = atrio-ventricular canal; Myo = myocardium; Mes = mesenchyme)



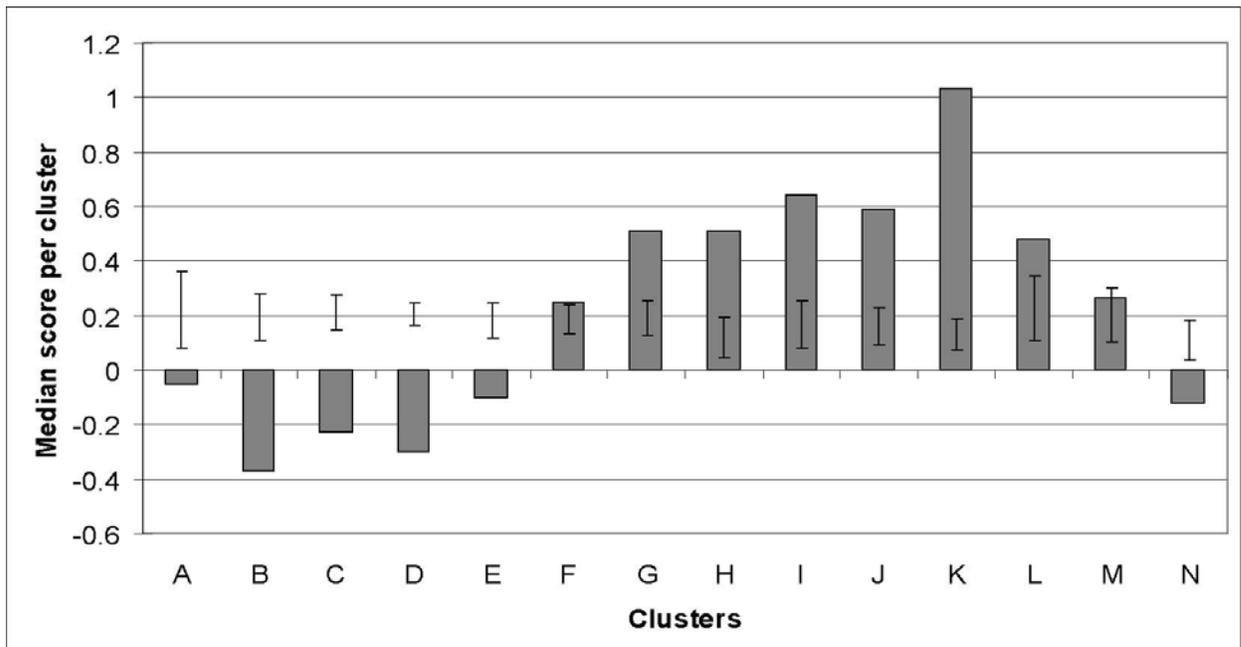
2.3.2. Epithelial/endothelial to mesenchymal transformation in the AVC

To examine the transition of gene expression from epithelial to mesenchymal morphology in the AVC on a global scale, and to address the gene expression differences within the various cell-types of the AVC, we used epithelial-mesenchymal library pairs from our SAGE data collection (Siddiqui et al., 2005) to examine the temporal expression pattern of epithelial- and mesenchymal-enriched genes. Six pairs of libraries, representing the dissociated epithelial and mesenchymal components of kidney, lung, male and female urogenital sinus, as well as large and small intestine, were used to calculate a mesenchyme-epithelial score ranging from +6 (mesenchyme enrichment) to -6 (epithelial enrichment). A mesenchyme-epithelial score was assigned to each tag-type and the median score per cluster was calculated (Figure 2.3). Clusters peaking at E9.5 (i.e. clusters A-C) were more likely to contain tag-types with higher expression in epithelial cells compared to mesenchyme. Clusters with peaks at E10.5 or E11.5 (e.g. clusters G-J) included tag-types more commonly expressed in mesenchymal tissues, supporting the production of mesenchyme in the AVC at these time-points. GO analysis of the mesenchyme-enriched genes with mesenchyme-epithelial score above +1.5 in the clusters demonstrated that genes in these clusters were more likely to be involved in signal transduction and transcriptional regulation (data not shown). The cluster with highest mesenchymal score (cluster K) contained tag-types that consistently increased in expression from E9.5 to E11.5 with expression maintained at E12.5. This cluster is significantly enriched for many of the extracellular matrix components necessary to support mesenchymal cells ($p = 0.009$). Finally, clusters peaking at E12.5 (cluster N) included tag-types with higher expression in epithelium suggesting that at

E12.5 the gene expression program is dramatically changing. The increased proportion of epithelial genes with expression peaking at E12.5 indicates that mesenchyme production is down-regulated by E12.5. Interestingly, GO analysis of the genes with positive mesenchyme scores suggests that clusters which peak at E12.5 are significantly enriched in genes involved in bone remodeling ($p = 0.034$) supporting work demonstrating that heart valve maturation shares regulatory mechanisms active in developing cartilage, tendon and bone (Chakraborty et al., 2008; Lange and Yutzey, 2006). Taken together, this global analysis of epithelial and mesenchymal markers defines the temporal events of EMT in the AVC and supports the functional relationships between the genes grouped in the gene expression clusters. Moreover, since our libraries were from the whole AVC which included both the endothelial and mesenchymal cells, this analysis allows prediction of the cellular expression of genes in the clusters.

Figure 2.3. Cluster analysis reveals temporal sequence of EMT in the AVC

Tag-types were assigned a mesenchyme-epithelial score based on their distribution across six epithelial-mesenchymal library pairs in the Mouse Atlas dataset (Siddiqui et al., 2005). A positive score indicates mesenchyme enrichment while a negative score indicates epithelial enrichment. Median scores across all tag-types are presented for each cluster. Error bars represent standard deviations obtained from a set of randomized clusters of similar size (n=10).



2.3.3. Downstream targets of TGF β and Notch signalling in the AVC

The TGF β pathway has been shown to be critical for EMT in the AVC. Though many components of the pathway (e.g. *Acvr1b*, *Smad2* and *Smad4*) are expressed at similar levels throughout AVC development, a number of TGF β pathway members showed distinct temporal expression patterns (Table 2.2), suggesting that they might control different processes during valve formation. For instance, expression of *Tgfb2* was found to be high at E9.5-10.5 and then decrease over time (cluster D), while *Tgfb3* expression peaked at E12.5 (cluster N) paralleling previous expression studies (Camenisch et al., 2002; Molin et al., 2003). While TGF β 2 can control the initiation and cessation of EMT in AVC explant assays and *in vivo* (Azhar et al., 2009; Bartram et al., 2001; Camenisch et al., 2002), the *in vivo* roles of TGF β 1 and TGF β 3 remain unclear. To determine which genes might be regulated by TGF β in the AVC and how they are related to the temporal patterns, we used published microarray data of Human Umbilical Vein Endothelial Cells (HUVECs) treated with TGF β 1 (Wu et al., 2006). TGF β 1, 2 and 3 can act through the same receptors to activate intracellular effectors. Of the 336 genes showing at least 1.5 fold change in expression in response to TGF β 1 treatment, 129 (35%) were detected in our AVC SAGE libraries, and 80% of these (104 - 30% of total) showed a dynamic temporal expression pattern in the AVC and were present in one of our clusters. Interestingly, TGF β responsive genes are overrepresented in specific clusters (Figure 2.4A) forming a pattern that suggests multiple phases of TGF β activity in the AVC. The first phase corresponds to tag-types whose expression is highest at E10.5 (clusters D, E and F). It includes transcription factors involved in cell differentiation such as *Atf3* and *Nabl*. A second phase of TGF β responsive genes

corresponds to tag-types whose expression peaks at E11.5 (cluster J) and includes regulators of apoptosis (e.g. *Sod2* and *Dedd*). Finally a third phase corresponds specifically to tag-types whose expression dramatically increases at E12.5 (cluster N). It includes ECM components *Biglycan* and Von Willebrand factor homolog (*Vwf*).

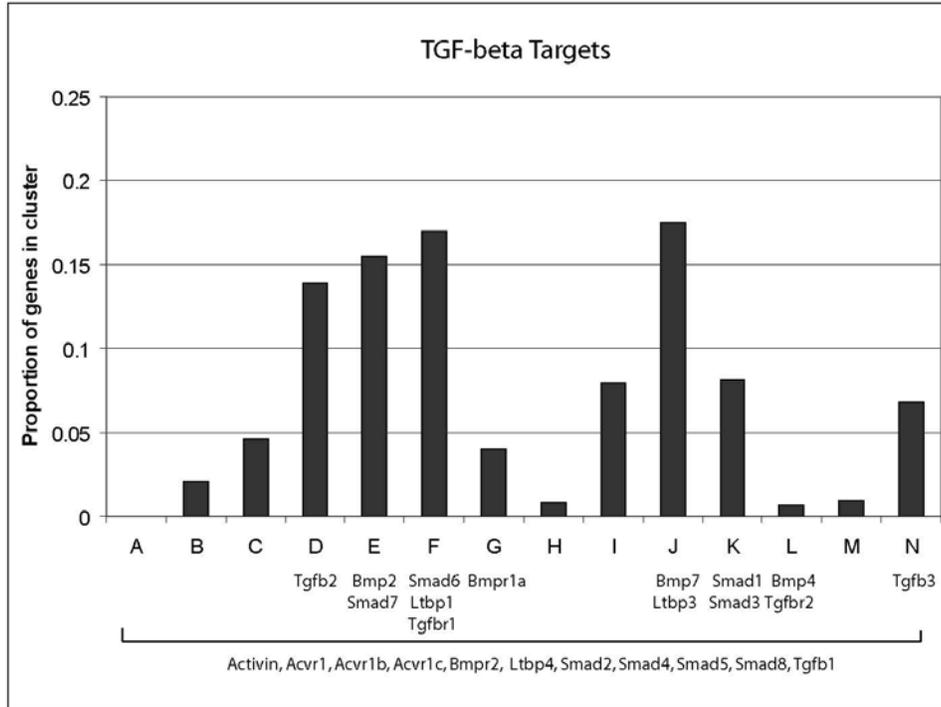
The Notch pathway has also been implicated in EMT during AVC development. To identify genes regulated by Notch and to determine if targets of the Notch pathway show similar patterns to those of TGF β , we analyzed microarray data generated from HUVECs activated by over-expression of the Notch ligand *Dll4* (Harrington et al., 2008). From the 695 genes showing at least a 1.5 fold change in the microarray experiments, 535 (77%) were detected in the AVC SAGE libraries, and 242 (35% of the total) showed differential expression over time and were present in one of our clusters. As with TGF β , clusters with peak expression at E10.5, E11.5 and E12.5 showed a higher proportion of targets determined by microarray analysis indicating high signalling activity at these time-points. Four phases of Notch activity were observed in our clusters (Figure 2.4B). The first phase of Notch responsive genes, which includes signalling pathway members *Tgf β 2*, *Jag1*, *Smad7* and *Igfbp5*, corresponds to genes with highest expression at E10.5 (cluster D and E). The second phase corresponds to tag-types whose expression peaks at E10.5 and E11.5 (cluster G). Genes in this peak include cell-adhesion molecules such as *Itga5*, the regulator of apoptosis, *Cflar* and the microRNA processing enzyme *Dicer*. A third phase corresponds to tag-types whose expression peaks at E11.5 (cluster J) including the endopeptidase inhibitor *Reck*, a modulator of the Notch pathway. Finally a fourth phase corresponds to tag-types whose expression dramatically increases at E12.5 (cluster N). It includes ECM components *Elastin* and *Periostin*. As with TGF β , the distribution of Notch signalling targets suggests that there are multiple phases of signalling activity occurring in the AVC, each associated with

distinct functions. Interestingly, there was considerable overlap between the phases of Notch and TGF β signalling activity since both ligands activated a high number of genes in cluster E (peaking at E10.5), supporting a role for both ligands in the early EMT phase of AVC development. However, distinct differences were observed between Notch and TGF β activity since cluster G (peaking at E10.5 and E11.5) contained a high number of Notch responsive genes but only a few TGF β responsive genes. Cluster J which peaks at E11.5 and is associated with GO categories related to proliferation, contained the highest proportion of TGF β responsive genes suggesting that TGF β may play an important role during this phase of AVC development. Importantly, 28 genes (8%) were found to be regulated by both TGF β 1 and DLL4, including members of the Notch and TGF β pathways (*Jag1* and *Tgfbr2*). These observations support cross-talk between these pathways and suggest that these two ligands also have distinct functions in AVC development.

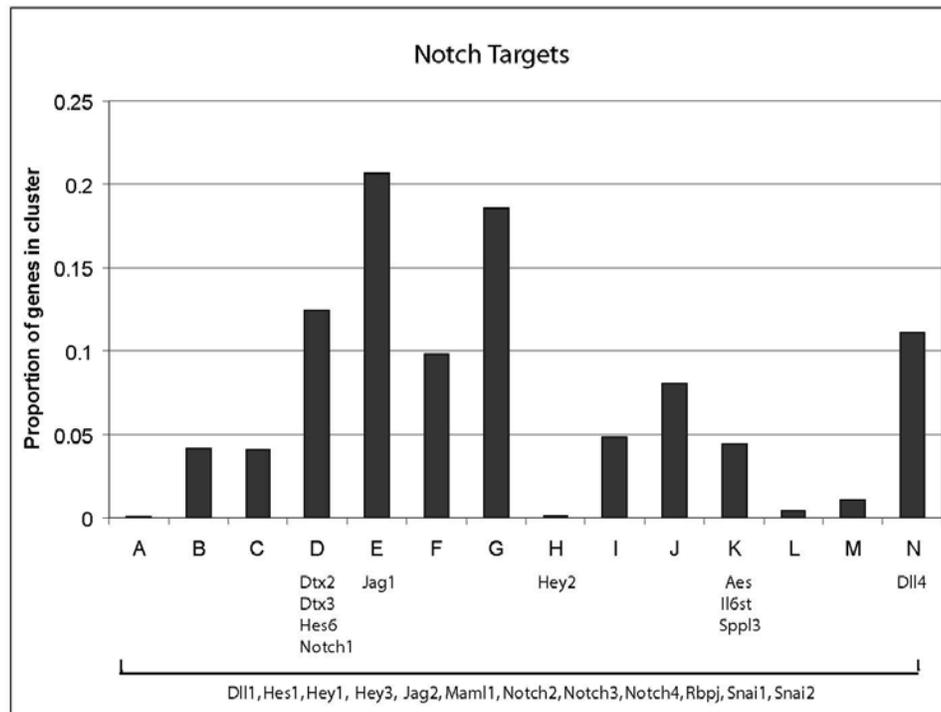
Figure 2.4. Multiple distinct phases of TGF β and Notch pathway activity exist in the AVC

TGF β (A) and Notch (B) target genes were obtained from two microarray studies (Harrington et al., 2008; Wu et al., 2006) in which TGF β and Notch pathways were activated in HUVECs by 4hrs of TGF β 1 treatment or *Dll4* over-expression, respectively. Target genes were defined by a minimum of 1.5 fold up- or down-regulation. The proportion of genes in each cluster is plotted on the y-axis, normalized to cluster size. Cluster distribution of known pathway members is indicated below the graph and members that are expressed in the AVC but do not significantly change in expression across AVC development are listed below the square bracket.

A



B



2.4. Discussion

We present the first comprehensive temporal dataset for the initiation of heart valve development in the mouse. This global and unbiased analysis characterizes AVC development from the initiation of EMT at E9.5 through to the beginning of valve remodeling at E12.5. Tags for genes known to be involved in AVC development were reliably observed including highly expressed structural genes, and comparatively lowly expressed transcription factors, indicating that our SAGE dataset recapitulates gene expression patterns observed *in vivo*. Furthermore, 29% of tags did not map to known transcripts using the RefSeq database indicating that our SAGE libraries are a rich source of potentially novel transcripts.

We resolved the tag-types into 14 distinct temporal expression patterns. These patterns varied significantly in the number of genes they contained suggesting that certain expression patterns are more prevalent during AVC development. Specifically, a large proportion (34%) of genes showed peaks of expression at E10.5, E11.5, or both (clusters E, G and J). These clusters contained many TGF β , Notch and BMP pathway members supporting the importance of these pathways at these stages.

Our AVC SAGE libraries include mixed populations of cells; therefore, a change in the proportion of transcripts in our libraries could reflect both a change of expression within cells as well as a change in the proportion of the different cell populations in the AVC. To investigate the distribution of mesenchyme-enriched genes, we compared six epithelial-mesenchymal library pairs (Siddiqui et al., 2005) to calculate the mesenchyme enrichment of tag-types. This permitted the temporal evaluation of mesenchyme-enriched genes on a global scale. Previous studies have

focused on E9.5-10.5 as the critical time in EMT in the AVC (Camenisch et al., 2002). We show that some mesenchyme-enriched genes begin to be expressed at E10.5 and later (clusters G-J) suggesting that mesenchyme differentiation continues throughout the entire timeframe.

Following EMT, the mesenchyme of the AVC expands through proliferation (Hinton et al., 2006; Kruihof et al., 2007). Though proliferation decreases as valves mature, no studies have shown the temporal changes in proliferation between E9.5 and E12.5. In clusters I and J which peak at E11.5, the expression of genes associated with mitosis and cell division are remarkably overrepresented compared to the other stages, suggesting that a burst of proliferative activity occurs at E11.5. As EMT concludes, the AVC is remodeled through differentiation of mesenchymal cells. Mesenchymal cells near the endocardial layer remain undifferentiated and are highly proliferative (Gitler et al., 2003; Sugi et al., 2003) while cells near the myocardial layer are less proliferative and express markers of differentiation (Mjaatvedt et al., 1999). We observed that genes correlated with an undifferentiated state, such as *Msx1* and *Twist1*, show a peak of expression at E10.5-11.5 (cluster D and G respectively). *Meox1*, a marker of differentiating mesenchyme, increased gradually after E10.5 (cluster K) (Candia et al., 1992). Moreover, genes involved in bone remodeling were enriched in the mesenchymal fraction at E12.5 indicating that the program leading to remodeling of the ECM has initiated by this time-point.

Finally, apoptosis removes excess cells and remodels the valves into leaflets. In the developing mouse heart, apoptotic cells were absent from E9.5-11.5 AVCs, but began to appear in the fusion seam of the AV cushion at E12.5-13.5 (Zhao and Rivkees, 2000). Correspondingly, we found that the cluster with the most dramatic increase in expression at E12.5 (cluster N) is

enriched for genes that regulate apoptosis. Together, our data provide molecular signatures distinguishing key morphogenetic events during early AVC development.

TGF β and Notch pathways control EMT during AVC development. Less is known about their role after EMT. Thus, we characterized the temporal expression of TGF β and Notch pathway members. Interestingly, different ligands of the TGF β and Notch pathways showed distinct expression patterns in our dataset suggesting that there are multiple phases of TGF β and Notch pathway activity over the course of AVC development. For example, expression of *Tgfb2* was found to be high at E9.5-10.5 and then decrease over time (cluster D), while *Tgfb3* expression peaked at E12.5 (cluster N). Previous work has found *Tgfb2* to be essential for EMT in mouse knockout embryos and *in vitro* explant assays (Azhar et al., 2009; Camenisch et al., 2002; Sanford et al., 1997). Significantly, the high expression of *Tgfb3* at E12.5 suggests that it plays a role in the remodeling process. Similarly, expression of the Notch ligand *Jag1* peaked at E10.5 (cluster E) while the Notch ligand *Dll4* was highest at E12.5 (cluster N) suggesting that they control different processes during AVC development.

Using endothelial cell microarray data, we overlaid our temporal expression patterns with targets of TGF β and Notch pathways. Out of 1031 genes differentially expressed in the microarray experiments, 664 were found in our AVC SAGE libraries and 346 had a dynamic temporal expression pattern. Notably, there was a significant difference between the types of TGF β and Notch target genes peaking at different time-points. For example, TGF β responsive genes peaking at E9.5-10.5 (cluster D) included many genes involved in transcription factor activity or chromatin remodeling, while those peaking at E12.5 included many ECM proteins and cell-adhesion molecules. Significantly, several members of the TGF β pathway were found

downstream of Notch signalling and vice-versa indicating that there is considerable cross-talk between the pathways.

In conclusion, we describe the creation of a database for AVC development which we have used to analyze the temporal expression of genes. This resource is valuable for the elucidation of the molecular mechanisms underlying heart development.

2.5. References

- Akmaev, V. R. and Wang, C. J. (2004) Correction of sequence-based artifacts in serial analysis of gene expression. *Bioinformatics* 20, 1254-1263.
- Anisimov, S. V., Tarasov, K. V., Riordon, D., Wobus, A. M. and Boheler, K. R. (2002a) SAGE identification of differentiation responsive genes in P19 embryonic cells induced to form cardiomyocytes in vitro. *Mech.Dev.* 117, 25-74.
- Anisimov, S. V., Tarasov, K. V., Stern, M. D., Lakatta, E. G. and Boheler, K. R. (2002b) A quantitative and validated SAGE transcriptome reference for adult mouse heart. *Genomics* 80, 213-222.
- Audic, S. and Claverie, J. M. (1997) The significance of digital gene expression profiles. *Genome Res.* 7, 986-995.
- Azhar, M., Runyan, R. B., Gard, C., Sanford, L. P., Miller, M. L., Andringa, A., Pawlowski, S., Rajan, S. and Doetschman, T. (2009) Ligand-specific function of transforming growth factor beta in epithelial-mesenchymal transition in heart development. *Dev.Dyn.* 238, 431-442.
- Bartram, U., Molin, D. G., Wisse, L. J., Mohamad, A., Sanford, L. P., Doetschman, T., Speer, C. P., Poelmann, R. E. and Gittenberger-de Groot, A. C. (2001) Double-outlet right ventricle and overriding tricuspid valve reflect disturbances of looping, myocardialization, endocardial cushion differentiation, and apoptosis in TGF-beta(2)-knockout mice. *Circulation* 103, 2745-2752.
- Blackshaw, S., Harpavat, S., Trimarchi, J., Cai, L., Huang, H., Kuo, W. P., Weber, G., Lee, K., Fraioli, R. E., Cho, S. H., *et al.* (2004) Genomic analysis of mouse retinal development. *PLoS Biol.* 2, E247.
- Cai, L., Huang, H., Blackshaw, S., Liu, J. S., Cepko, C. and Wong, W. H. (2004) Clustering analysis of SAGE data using a Poisson approach. *Genome Biol.* 5, R51.
- Camenisch, T. D., Molin, D. G., Person, A., Runyan, R. B., Gittenberger-de Groot, A. C., McDonald, J. A. and Klewer, S. E. (2002) Temporal and distinct TGFbeta ligand requirements during mouse and avian endocardial cushion morphogenesis. *Dev.Biol.* 248, 170-181.
- Candia, A. F., Hu, J., Crosby, J., Lalley, P. A., Noden, D., Nadeau, J. H. and Wright, C. V. (1992) Mox-1 and Mox-2 define a novel homeobox gene subfamily and are differentially expressed during early mesodermal patterning in mouse embryos. *Development* 116, 1123-1136.
- Chakraborty, S., Cheek, J., Sakthivel, B., Aronow, B. J. and Yutzey, K. E. (2008) Shared gene expression profiles in developing heart valves and osteoblast progenitor cells. *Physiol.Genomics*
- Chen, Z. F. and Behringer, R. R. (1995) Twist is Required in Head Mesenchyme for Cranial Neural Tube Morphogenesis. *Genes Dev.* 9, 686-699.
- Fischer, A., Steidl, C., Wagner, T. U., Lang, E., Jakob, P. M., Friedl, P., Knobloch, K. P. and Gessler, M. (2007) Combined loss of Hey1 and HeyL causes congenital heart defects because of impaired epithelial to mesenchymal transition. *Circ.Res.* 100, 856-863.
- Gitler, A. D., Lu, M. M., Jiang, Y. Q., Epstein, J. A. and Gruber, P. J. (2003) Molecular markers of cardiac endocardial cushion development. *Dev.Dyn.* 228, 643-650.

- Harrington, L. S., Sainson, R. C., Williams, C. K., Taylor, J. M., Shi, W., Li, J. L. and Harris, A. L. (2008) Regulation of multiple angiogenic pathways by Dll4 and Notch in human umbilical vein endothelial cells. *Microvasc.Res.* 75, 144-154.
- Hinton, R. B., Jr, Lincoln, J., Deutsch, G. H., Osinska, H., Manning, P. B., Benson, D. W. and Yutzey, K. E. (2006) Extracellular matrix remodeling and organization in developing and diseased aortic valves. *Circ.Res.* 98, 1431-1438.
- Hoffman, J. I. (1995) Incidence of congenital heart disease: II. Prenatal incidence. *Pediatr.Cardiol.* 16, 155-165.
- Hosack, D. A., Dennis, G., Jr, Sherman, B. T., Lane, H. C. and Lempicki, R. A. (2003) Identifying biological themes within lists of genes with EASE. *Genome Biol.* 4, R70.
- Kokubo, H., Tomita-Miyagawa, S., Hamada, Y. and Saga, Y. (2007) *Hes1* and *Hes2* regulate atrioventricular boundary formation in the developing heart through the repression of *Tbx2*. *Development* 134, 747-755.
- Kraus, F., Haenig, B. and Kispert, A. (2001) Cloning and expression analysis of the mouse T-box gene *tbx20*. *Mech.Dev.* 100, 87-91.
- Kruithof, B. P., Krawitz, S. A. and Gaussin, V. (2007) Atrioventricular valve development during late embryonic and postnatal stages involves condensation and extracellular matrix remodeling. *Dev.Biol.* 302, 208-217.
- Kruzynska-Frejtag, A., Machnicki, M., Rogers, R., Markwald, R. R. and Conway, S. J. (2001) Periostin (an osteoblast-specific factor) is expressed within the embryonic mouse heart during valve formation. *Mech.Dev.* 103, 183-188.
- Lange, A. W. and Yutzey, K. E. (2006) NFATc1 expression in the developing heart valves is responsive to the RANKL pathway and is required for endocardial expression of cathepsin K. *Dev.Biol.* 292, 407-417.
- McKnight, K. D., Hou, J. and Hoodless, P. A. (2007) Dynamic expression of thyrotropin-releasing hormone in the mouse definitive endoderm. *Dev.Dyn.* 236, 2909-2917.
- Mjaatvedt, C., Yamamura, H., Wessels, A., Ramsdell, A., Turner, D. and Markwald, R. (1999) Mechanisms of segmentation, septation, and remodeling of the tubular heart: endocardial cushion fate and cardiac looping. 159-177.
- Molin, D. G., Bartram, U., Van der Heiden, K., Van Iperen, L., Speer, C. P., Hierck, B. P., Poelmann, R. E. and Gittenberger-de-Groot, A. C. (2003) Expression patterns of *Tgfbeta1-3* associate with myocardialisation of the outflow tract and the development of the epicardium and the fibrous heart skeleton. *Dev.Dyn.* 227, 431-444.
- Moskowitz, I. P., Kim, J. B., Moore, M. L., Wolf, C. M., Peterson, M. A., Shendure, J., Nobrega, M. A., Yokota, Y., Berul, C., Izumo, S., Seidman, J. G. and Seidman, C. E. (2007) A molecular pathway including *Id2*, *Tbx5*, and *Nkx2-5* required for cardiac conduction system development. *Cell* 129, 1365-1376.
- Person, A. D., Klewer, S. E. and Runyan, R. B. (2005) Cell biology of cardiac cushion development. *Int.Rev.Cytol.* 243, 287-335.
- Robertson, N., Oveisi-Fordorei, M., Zuyderduyn, S. D., Varhol, R. J., Fjell, C., Marra, M., Jones, S. and Siddiqui, A. (2007) DiscoverySpace: an interactive data analysis application. *Genome Biol.* 8, R6.
- Rutenberg, J. B., Fischer, A., Jia, H., Gessler, M., Zhong, T. P. and Mercola, M. (2006) Developmental patterning of the cardiac atrioventricular canal by Notch and Hairy-related transcription factors. *Development* 133, 4381-4390.

- Sanford, L. P., Ormsby, I., Gittenberger-de Groot, A. C., Sariola, H., Friedman, R., Boivin, G. P., Cardell, E. L. and Doetschman, T. (1997) TGFbeta2 knockout mice have multiple developmental defects that are non-overlapping with other TGFbeta knockout phenotypes. *Development* 124, 2659-2670.
- Siddiqui, A. S., Khattri, J., Delaney, A. D., Zhao, Y., Astell, C., Asano, J., Babakaiff, R., Barber, S., Beland, J., Bohacec, S., *et al.* (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc.Natl.Acad.Sci.U.S.A.* 102, 18485-18490.
- Smolenski, A., Schultess, J., Danielewski, O., Garcia Arguinzonis, M. I., Thalheimer, P., Kneitz, S., Walter, U. and Lohmann, S. M. (2004) Quantitative analysis of the cardiac fibroblast transcriptome-implications for NO/cGMP signaling. *Genomics* 83, 577-587.
- Stollberg, J., Urschitz, J., Urban, Z. and Boyd, C. D. (2000) A quantitative evaluation of SAGE. *Genome Res.* 10, 1241-1248.
- Sugi, Y., Ito, N., Szebenyi, G., Myers, K., Fallon, J. F., Mikawa, T. and Markwald, R. R. (2003) Fibroblast growth factor (FGF)-4 can induce proliferation of cardiac cushion mesenchymal cells during early valve leaflet formation. *Dev.Biol.* 258, 252-263.
- Timmerman, L. A., Grego-Bessa, J., Raya, A., Bertran, E., Perez-Pomares, J. M., Diez, J., Aranda, S., Palomo, S., McCormick, F., Izpisua-Belmonte, J. C. and de la Pompa, J. L. (2004) Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* 18, 99-115.
- Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. (1995) Serial analysis of gene expression. *Science* 270, 484-487.
- Watanabe, Y., Kokubo, H., Miyagawa-Tomita, S., Endo, M., Igarashi, K., Aisaki, K., Kanno, J. and Saga, Y. (2006) Activation of Notch1 signaling in cardiogenic mesoderm induces abnormal heart morphogenesis in mouse. *Development* 133, 1625-1634.
- Wu, X., Ma, J., Han, J. D., Wang, N. and Chen, Y. G. (2006) Distinct regulation of gene expression in human endothelial cells by TGF-beta and its receptors. *Microvasc.Res.* 71, 12-19.
- Zhao, Z. and Rivkees, S. A. (2000) Programmed cell death in the developing heart: regulation by BMP4 and FGF2. *Dev.Dyn.* 217, 388-400.

CHAPTER 3

TWIST1 CONTROLS DIFFERENTIATION OF DEVELOPING ATRIO- VENTRICULAR CANAL CELLS IN THE MOUSE

3.1. Introduction

The transformation of the heart tube into a four-chambered organ divided by valves and septa is a critical event during mammalian heart development that is required for its proper function. Initially, the embryonic heart is a simple tube composed of endocardial and myocardial cell layers separated by an acellular extracellular matrix (ECM) termed the cardiac jelly. To form the heart valves and septa, the cardiac jelly begins to accumulate in the atrio-ventricular canal (AVC) and outflow tract (OFT) regions giving rise to endocardial cushions (Person et al., 2005). At around embryonic (E) day 9 in the mouse and day 26 in humans, inductive signals from the myocardium mediated by members of the Notch, Wnt and TGF β pathways activate endocardial cells in these regions to undergo an epithelial-to-mesenchymal transformation (EMT) (Camenisch et al., 2002; Liebner et al., 2004; Timmerman et al., 2004; Watanabe et al., 2006). EMT is a multi-step process in which polarized and adhesive endocardial cells transform into non-polarized and highly motile mesenchyme cells embedded in the cardiac jelly (Markwald et al., 1975). After invasion of the ECM, the newly transformed mesenchyme cells proliferate and undergo further differentiation (Hinton et al., 2006; Kruithof et al., 2007). Finally, removal of excess mesenchyme cells by apoptosis starting at E12.5 in mice, and patterning of ECM molecules relative to direction of blood flow turns the endocardial cushions into thin stress-

resistant AV valve leaflets and semilunar valve cusps (Armstrong and Bischoff, 2004; Zhao and Rivkees, 2000).

The process of endocardial cushion transformation demonstrates tight temporal and spatial gene expression specificity controlled by a network of transcription factors. Single-gene studies have helped to determine the important role of a number of transcription factors during this process, including SOX9 (Lincoln et al., 2007), NFATc (de la Pompa et al., 1998; Ranger et al., 1998), SNAI2 (Niessen et al., 2008) and β -CATENIN (Liebner et al., 2004). A more systematic approach to the study of endocardial cushion development has been attempted with microarray analysis, and has proven useful in the identification of AVC-specific gene expression (Wirrig et al., 2007), and in the comparison of wild-type and mutant phenotypes in the developing AVC (Rivera-Feliciano et al., 2006) and OFT (Zhu et al., 2007). However, microarrays have important limitations in their sensitivity and potential for novel gene or transcript discovery ('t Hoen et al., 2008).

New advances in sequencing technology and the completion of genome sequencing projects allow for a comprehensive analysis of the transcriptome (Morrissy et al., 2009). Serial analysis of gene expression (SAGE) is a technique which resolves transcripts into unique short 21-base sequence tags that can be mapped to the genome (Saha et al., 2002; Velculescu et al., 1995). It is independent of previous transcript knowledge and can be used for novel gene and transcript variant discovery. In previous work from our lab we have used the SAGE technique to describe the temporal expression changes occurring from E9.5 to E12.5 during early AVC development in the mouse and found that E10.5 was a time-point of increased signalling and transcription factor activity (Vrljicak et al., 2009).

In this study we use tag sequencing (Tag-seq), a combination of the SAGE technique with massively parallel sequencing technology (Morrissy et al., 2009), to create gene expression libraries from atria, ventricles, AVC and OFT of the mouse heart at E10.5. Using this dataset we identified a high-confidence AVC- and OFT-enriched gene set at E10.5 containing novel endocardial cushion genes and transcription factors. We focus here on the basic helix-loop-helix (bHLH) transcription factor *Twist1* as the most highly expressed DNA-binding transcription factor in the E10.5 AVC endocardial cushions. Our data on the gene expression differences between *Twist1* mutant and wild-type AVC at E10.5 suggests that TWIST1 plays a critical role in mesenchymal differentiation post-EMT in the mouse, by directly regulating expression of AVC- and OFT-enriched transcription factors such as *Klf4*, *Tb11x*, *Hes6*, *Id2* and *Lef1*, while suppressing expression of valve maturation markers.

3.2. Materials and methods

3.2.1. Tissue collection

Mouse work was carried out following protocols approved by the Animal Care Committee at the University of British Columbia. To isolate specific regions of the heart, E10.5 embryos (Theiler stage 17) were collected from timed-pregnant C57BL/6J females and manually dissected using 30½G needles. Multiple litters were pooled to obtain at least 400ng of total RNA for Tag-seq library construction and to avoid possible bias created by spontaneous mutations in the colony. Blood was removed by puncturing the heart chambers and washing the tissue with PBS. Dissected mouse heart samples were collected in TRIzol reagent (Invitrogen). After

isolation, RNA quality was assessed using an Agilent Bioanalyzer and the RNA was stored at -80°C until Tag-seq library construction. *Twist1* null mice have been described previously (Chen and Behringer, 1995) and were maintained in an ICR (Taconic Farms) background.

3.2.2. Gene expression analysis

Tag-seq libraries were constructed as described (Morrissey et al., 2009). Tag-seq data was analyzed and mapped to genes using DiscoverySpace v4.0 (Robertson et al., 2007b). Unambiguous sense mappings to the RefSeq database (<http://www.ncbi.nlm.nih.gov/RefSeq/>) were used. Expression from multiple tag-types was pooled to obtain overall expression for a particular gene. Gene ontology analysis was performed on DAVID (Dennis et al., 2003; Huang et al., 2009).

3.2.3. Tissue culture

The mouse kidney epithelial cell line, mIMCD3, and the mouse endothelial cell line, SVEC, were obtained from the American Type Culture Collection and kept in DMEM supplemented with 10% fetal bovine serum. The *Twist1* over-expression vector was constructed by subcloning a myc-tagged cDNA (gift from Dr. A. Firulli) (Firulli et al., 2005) into the pRRL.PPT.SF.IRES-VENUSnucmem lentiviral vector (gift from Dr. T. Schroeder). This lentiviral vector contains a self-inactivating long terminal repeat, the 118-bp polypurine tract, and directs expression of the myc-tagged *Twist1* cDNA, an internal ribosomal entry site (IRES),

and Venus nuclear membrane (VENUSnucmem) from an internal spleen focus-forming virus (SFFV) promoter.

3.2.4. Chromatin immunoprecipitation

Chromatin immunoprecipitation followed by quantitative PCR (ChIP-qPCR) was carried out as described (Wederell et al., 2008), using either an anti-c-myc antibody (Covance, MMS-150P) or a mouse non-specific IgG (Sigma, F2883). The fold enrichment of each target site was calculated as 2 to the power of the cycle threshold (cT) difference between the IgG immunoprecipitated sample and the specific antibody immunoprecipitated sample. Primers used for ChIP-qPCR are listed in Appendix VI.

3.3. Results

To study the regional distribution of genes in the developing heart we constructed Tag-seq gene expression libraries from atria, ventricles, AVC and OFT of E10.5 mouse hearts (Table 3.1). Each library was sequenced to a minimum depth of 7 million tags for a total of over 34 million tags. To exclude low abundance transcripts and tags that might have been generated by library construction or sequencing errors we focused on 10,670 RefSeq genes expressed at more than 5 tags in at least one library.

Table 3.1. Overview of Tag-seq libraries

Atria, atrio-ventricular canal, ventricles and out-flow tract were isolated from E10.5 (Theiler stage 17) mouse hearts for Tag-seq library construction. High quality (HQ) tag-types were present at greater than 5 tags per library.

| Library ID | Description | RNA used | All Tags | HQ Tags | HQ Tag-types |
|------------|--|----------|------------|------------|--------------|
| MM0265 | E10.5 Atria | 1µg | 8,303,915 | 4,666,514 | 35,110 |
| MM0263 | E10.5 Atrio-ventricular canal | 500ng | 10,445,112 | 6,075,421 | 58,407 |
| MM0266 | E10.5 Ventricles | 1µg | 8,284,532 | 4,223,810 | 32,822 |
| MM0264 | E10.5 Outflow tract | 500ng | 7,072,418 | 5,267,618 | 56,744 |
| | Total | | 34,105,977 | 20,233,363 | 101,847 |
| MM0513 | E10.5 <i>Twist1</i> ^{-/-} Atrio-ventricular canal | 400ng | 17,262,266 | 12,977,237 | 79,271 |

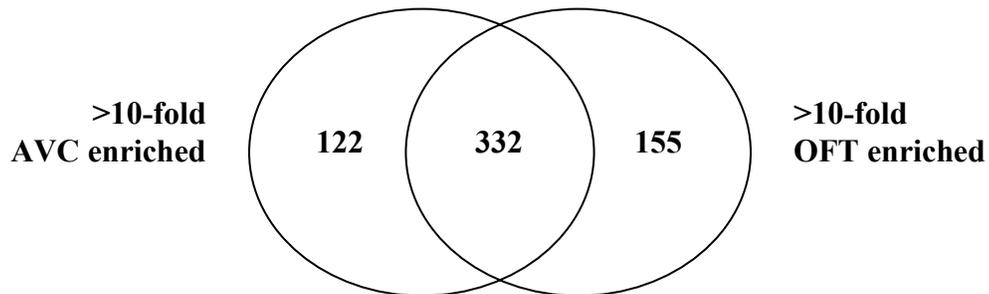
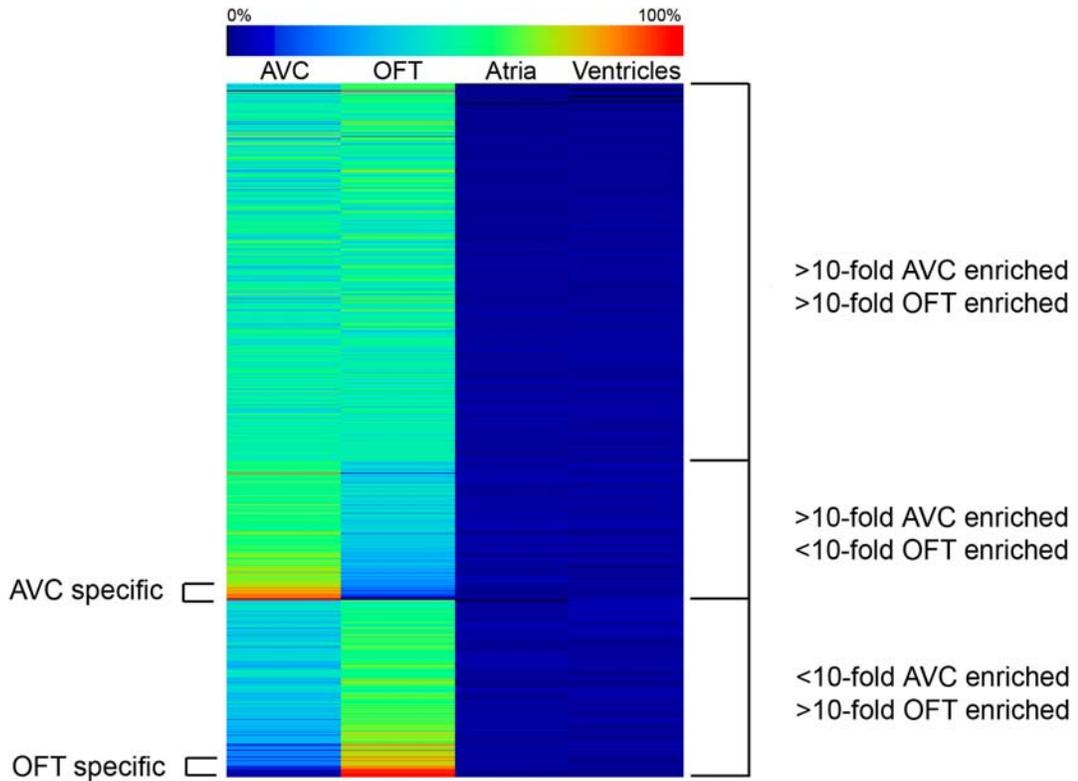
To test whether our Tag-seq libraries faithfully captured known patterns of gene expression, we examined 17 genes previously characterized in the context of AVC development (Appendix VII). Expression of these genes in the AVC Tag-seq library spanned two orders of magnitude ranging from 9 tags per million for the transcription factor *Fog2* to 1,473 tags per million for the structural molecule *Vimentin*. Significantly, each of the genes we examined was found to be differentially expressed in the AVC when compared to the atria and ventricles. Furthermore, their AVC enrichment (calculated as the ratio of AVC expression over atria and ventricles) ranged from two-fold for *Vimentin*, to 124-fold for the transcription factor *Sox9*, indicating that the Tag-seq libraries represent a reliable source of gene expression information over a wide dynamic range.

3.3.1. AVC and OFT share significant gene expression

Since the OFT undergoes a similar process of endocardial cushion formation and EMT as the AVC, we speculated that genes critical for early valve formation would also be enriched in the OFT. We therefore searched for genes enriched in both AVC and OFT to identify additional genes involved in this process. To calculate the AVC enrichment of genes we averaged the ratio of AVC expression over atria and ventricles after normalizing to library size. Similarly, we calculated OFT enrichment by averaging the ratio of OFT expression over atria and ventricles. We first identified 454 genes with a 10-fold or greater AVC enrichment, and compared them against 487 genes with a 10-fold or greater OFT enrichment. Comparison of AVC- and OFT-enriched genes revealed a significant gene expression overlap, with 332 genes enriched at least 10-fold in both AVC and OFT (Figure 3.1).

Figure 3.1. AVC and OFT shared gene expression

Overlap in genes with 10-fold or higher AVC- or OFT-enrichment. Expression of all tag-types mapping in the sense direction to the same RefSeq gene were pooled. AVC- or OFT- enrichment was calculated as the averaged ratio of AVC or OFT gene expression over atria and ventricles.



Furthermore, we observed considerable OFT enrichment of the AVC-enriched genes and vice-versa, with 96% of AVC-enriched genes showing a greater than 5-fold enrichment in the OFT, and 95% of OFT-enriched genes showing a greater than 5-fold enrichment in the AVC. In addition, a number of genes specific to the OFT or AVC were also identified, such as *Galanin* and *Adamts19* in the AVC and *Rgs5* and *Gabra4* in the OFT (Table 3.2 and Appendix IX), indicating that our dataset can distinguish gene expression differences underlying unique AVC and OFT developmental processes.

For our subsequent analysis we focused on shared AVC and OFT gene expression by selecting 565 genes showing either a 10-fold or greater AVC enrichment and a 5-fold or greater OFT enrichment, or a 10-fold OFT enrichment and 5-fold or greater AVC enrichment (Table 3.2). This list included known endocardial cushion development genes such as *Bmp2* (Ma et al., 2005), *Jag1* (Nosedá et al., 2004), *Gata4* (Crispino et al., 2001), and *Sox9* (Lincoln et al., 2007), providing confidence in our selection criteria. Furthermore, our list captured both highly expressed genes like *Col3a1* (Peacock et al., 2008) (expressed at 1442 tags per million in the AVC), and relatively lower expressed transcription factors such as *Tbx20* (Shelton and Yutzey, 2007) (expressed at 51.7 tags per million in the AVC), suggesting that we can identify endocardial cushion genes at various expression levels. Significantly a majority of the genes in our AVC- and OFT- enriched list have not been described in the context of endocardial cushion development, indicating that our dataset represents a novel list of endocardial cushion genes.

Table 3.2. Selected genes enriched in the E10.5 AVC and OFT

Expression levels represent all sense tags for a gene and are shown as tags per million. AVC- or OFT- enrichment was calculated as the averaged ratio of AVC or OFT expression over atria and ventricles.

| Gene symbol | Gene name | RefSeq accession | AVC | OFT | AVC enr. | OFT enr. |
|---|---|------------------|--------|--------|----------|----------|
| A. Genes enriched in both AVC and OFT | | | | | | |
| Transcription regulators | | | | | | |
| <i>Twist1</i> | Twist homolog 1 | NM_011658 | 918 | 299.1 | 15.5 | 5 |
| <i>Sox9</i> | SRY-box containing gene 9 | NM_011448 | 514.4 | 376.7 | 123.6 | 90.5 |
| <i>Aes</i> | Amino-terminal enhancer of split | NM_010347 | 232.4 | 304.8 | 125.2 | 164.2 |
| <i>Gata4</i> | GATA binding protein 4 | NM_008092 | 227.4 | 122.8 | 32.9 | 17.8 |
| <i>Zeb1</i> | Zinc finger E-box binding homeobox 1 | NM_011546 | 185.4 | 196.1 | 25.4 | 26.9 |
| <i>Tead2</i> | TEA domain family member 2 | NM_011565 | 120 | 96.9 | 13.7 | 11.1 |
| <i>Klf4</i> | Kruppel-like factor 4 | NM_010637 | 32.1 | 47.8 | 23.7 | 35.4 |
| <i>Tbllx</i> | Transducin (beta)-like 1 X-linked | NM_020601 | 24.2 | 19.9 | 17.9 | 14.8 |
| <i>Hes6</i> | Hairy and enhancer of split 6 | NM_019479 | 18.6 | 41.2 | 13.8 | 30.5 |
| <i>Lef1</i> | Lymphoid enhancer binding factor 1 | NM_010703 | 18.1 | 11 | 13.4 | 8.1 |
| ECM structural molecules and modifiers | | | | | | |
| <i>Col3a1</i> | Collagen, type III, alpha 1 | NM_009930 | 1441.9 | 2334.3 | 14 | 23 |
| <i>Fbn2</i> | Fibrillin 2 | NM_010181 | 327 | 232 | 10.8 | 7.7 |
| <i>Hspg2</i> | Perlecan (heparan sulfate proteoglycan 2) | NM_008305 | 122.9 | 134.9 | 17.2 | 18.9 |
| <i>Col9a3</i> | Collagen, type IX, alpha 3 | NM_009936 | 120.5 | 38.2 | 62.5 | 19.8 |
| <i>Lama4</i> | Laminin, alpha 4 | NM_010681 | 110.1 | 66.6 | 62.1 | 37.6 |
| <i>Mmp14</i> | Matrix metalloproteinase 14 | NM_008608 | 89.3 | 127.2 | 66.1 | 94.1 |
| <i>Col5a1</i> | Collagen, type V, alpha 1 | NM_015734 | 65.1 | 92.2 | 10.1 | 14.4 |
| <i>Bgn</i> | Biglycan | NM_007542 | 49.4 | 36.9 | 36.5 | 27.3 |
| <i>Tnc</i> | Tenascin C | NM_011607 | 47.7 | 523.1 | 23.4 | 256 |
| Signalling molecules | | | | | | |
| <i>Igfbp5</i> | Insulin-like growth factor binding protein 5 | NM_010518 | 1563.9 | 1149.7 | 60.4 | 44.4 |
| <i>Spnb2</i> | Spectrin beta 2 | NM_175836 | 243 | 200.6 | 16.7 | 13.8 |
| <i>Bmp2</i> | Bone morphogenetic protein 2 | NM_007553 | 219.6 | 27.1 | 101.7 | 12.6 |
| <i>Bmper</i> | BMP-binding endothelial regulator | NM_028472 | 173.5 | 303.9 | 60.2 | 105.5 |
| <i>Adam9</i> | A disintegrin and metalloproteinase domain 9 | NM_007404 | 90.2 | 75.9 | 19.2 | 16.1 |
| <i>Htra1</i> | HtrA serine peptidase 1 | NM_019564 | 70.3 | 31.9 | 52 | 23.6 |
| <i>Fzd1</i> | Frizzled homolog 1 | NM_021457 | 64 | 79.7 | 27.9 | 34.7 |
| <i>Csnk1e</i> | Casein kinase 1, epsilon | NM_013767 | 43.6 | 36.3 | 32.3 | 26.8 |
| <i>Btc</i> | Betacellulin | NM_007568 | 34.9 | 16.3 | 25.8 | 12.1 |
| <i>TGFβ1</i> | Transforming growth factor beta 1 | NM_011577 | 30.5 | 19.6 | 11.3 | 7.3 |
| <i>Kdr</i> | Kinase insert domain protein receptor | NM_010612 | 24.6 | 17.2 | 18.2 | 12.8 |
| <i>Jag1</i> | Jagged 1 | NM_013822 | 12 | 27.91 | 8.9 | 20.7 |
| Proliferation | | | | | | |
| <i>Ccnd2</i> | Cyclin D2 | NM_009829 | 594.3 | 1117.7 | 8.3 | 15.6 |
| <i>Irs1</i> | Insulin receptor substrate 1 | NM_010570 | 122.4 | 73.1 | 23.4 | 13.9 |
| <i>Cdca7</i> | Cell division cycle associated 7 | NM_025866 | 114.1 | 157.7 | 27.6 | 38.2 |
| <i>Serpinfl</i> | Serine (or cysteine) peptidase inhibitor, clade F, member 1 | NM_011340 | 24.9 | 38.2 | 11.1 | 17.1 |

| Gene symbol | Gene name | RefSeq accession | AVC | OFT | AVC enr. | OFT enr. |
|--|---|------------------|--------|--------|----------|----------|
| Apoptosis | | | | | | |
| <i>Bat3</i> | HLA-B-associated transcript 3 | NM_057171 | 319.5 | 361.4 | 23.3 | 26.4 |
| <i>Ccar1</i> | Cell division cycle and apoptosis regulator 1 | NM_026201 | 195.4 | 171.4 | 25.6 | 22.5 |
| <i>Traf7</i> | TNF receptor-associated factor 7 | NM_153792 | 82.3 | 114.1 | 12.3 | 17.1 |
| <i>Traf4</i> | TNF receptor associated factor 4 | NM_009423 | 70.2 | 48.8 | 17.1 | 11.9 |
| <i>Daxx</i> | Fas death domain-associated protein | NM_007829 | 63.1 | 126.9 | 9.7 | 19.6 |
| Ribosome | | | | | | |
| <i>Rrp1</i> | Ribosomal RNA processing 1 homolog | NM_010925 | 115.4 | 91.5 | 14.2 | 11.3 |
| <i>Exosc6</i> | Exosome component 6 | NM_028274 | 86.4 | 182 | 12 | 25.4 |
| <i>Rplp0</i> | Ribosomal protein, large, P0 | NM_007475 | 82.4 | 97 | 61 | 71.8 |
| <i>Gnl3l</i> | Guanine nucleotide binding protein-like 3 (nucleolar)-like | NM_198110 | 43.2 | 77.5 | 31.9 | 57.3 |
| <i>Ddx51</i> | DEAD (Asp-Glu-Ala-Asp) box polypeptide 51 | NM_027156 | 29.2 | 24.1 | 21.6 | 17.8 |
| Bone, cartilage, tendon development | | | | | | |
| <i>Papss2</i> | 3'-phosphoadenosine 5'-phosphosulfate synthase 2 | NM_011864 | 218.9 | 115.5 | 161.9 | 85.4 |
| <i>Pdlim7</i> | PDZ and LIM domain 7 | NM_001114088 | 101.1 | 122.8 | 40.8 | 49.6 |
| <i>Ankrd11</i> | Ankyrin repeat domain 11 | NM_001081379 | 89.1 | 82.2 | 36.3 | 33.5 |
| Chromatin modifiers | | | | | | |
| <i>Hmga2</i> | High mobility group AT-hook 2 | NM_010441 | 1798.8 | 1498.3 | 16.6 | 13.8 |
| <i>Chd4</i> | Chromodomain helicase DNA binding protein 4 | NM_145979 | 238.3 | 386.9 | 15.2 | 24.7 |
| <i>Chd8</i> | Chromodomain helicase DNA binding protein 8 | NM_201637 | 46.9 | 29.8 | 29.2 | 18.6 |
| B. AVC-specific genes | | | | | | |
| <i>Gal</i> | Galanin | NM_010253 | 38.2 | <1.1 | 28.3 | 0.8 |
| <i>Tceal6</i> | Transcription elongation factor A (SII)-like 6 | NM_025355 | 20.4 | <1.1 | 15.1 | 0.8 |
| <i>Nppa</i> | Natriuretic peptide precursor type A | NM_008725 | 316.24 | 31.31 | 14 | 1.4 |
| <i>6330442E1</i> | RIKEN cDNA 6330442E10 | NM_178745 | 17.9 | 2.47 | 13.2 | 1.8 |
| <i>ORik</i> | Calcium channel, voltage-dependent, alpha 2/delta subunit 2 | NM_020263 | 16.31 | 2.09 | 12.1 | 1.6 |
| C. OFT-specific genes | | | | | | |
| <i>Isl1</i> | ISL1 transcription factor, LIM/homeodomain | NM_021459 | 2.308 | 72.88 | 1.58 | 49.96 |
| <i>Rgs5</i> | Regulator of G-protein signalling 5 | NM_009063 | 1.48 | 40.99 | 1.09 | 30.32 |
| <i>Mybpc1</i> | Myosin binding protein C, slow-type | NM_175418 | <1 | 26.06 | 0.73 | 19.28 |
| <i>Psd</i> | Pleckstrin and Sec7 domain containing | NM_028627 | 1.15 | 25.1 | 0.85 | 18.57 |
| <i>Gabra4</i> | Gamma-aminobutyric acid (GABA) A receptor, subunit alpha 4 | NM_010251 | 0.988 | 24.29 | 0.73 | 17.97 |
| <i>Stmn4</i> | Stathmin-like 4 | NM_019675 | 1.32 | 17.32 | 0.98 | 12.81 |
| <i>Lamc2</i> | Laminin, gamma 2 | NM_008485 | <1 | 15.6 | 0.73 | 11.54 |

The Notch, TGF β and Wnt pathways have been shown to be critically involved in controlling endocardial cushion EMT, and their activity must be tightly regulated (Armstrong and Bischoff, 2004; Person et al., 2005). Significantly, modulators of these pathways were particularly represented in our AVC and OFT enriched gene list (Table 3.2). Among the regulators of Wnt signalling we found *Csnkle* (Kim et al., 2010), a casein kinase responsible for DISHEVELLED phosphorylation, and *R-spondin3* (Kim et al., 2008), an inhibitor of Wnt receptor internalization, both of which control β -CATENIN dependent transcriptional activation. Among the genes controlling TGF β signalling we found *Bmper*, a secreted factor that directly interacts with BMP ligands (Kelley et al., 2009; Moser et al., 2003), and *Htra1*, a secreted serine protease that inhibits TGF β family members by its proteolytic activity (Heinke et al., 2008; Oka et al., 2004). Finally, among the regulators of the Notch pathway we found the bHLH transcription factor *Hes6* (Fior and Henrique, 2005; Pissarra et al., 2000), and the Groucho related transcriptional co-repressors *Aes* and *Tle2* (Chen and Courey, 2000).

The most highly expressed signalling pathway modulator in the AVC and OFT was found to be *Igfbp5*, which has been shown to promote cartilage anabolism and osteoblast proliferation (Andress and Birnbaum, 1992). This is significant in light of research showing shared gene expression in developing heart valves, cartilage, bone and tendons (Chakraborty et al., 2008). Other genes with shared expression between these tissues include *Papss2* (Stelzer et al., 2007), a critical enzyme for sulfation of proteoglycans during chondrocyte growth and development, and *Ankrd11*, a transcriptional co-regulator thought to recruit histone deacetylases and other chromatin modifiers to transcriptional complexes (Neilsen et al., 2008; Zhang et al., 2004).

Intriguingly, ribosomal processing factors, such as *Rrp1*, and ribosomal protein genes, such as *Rplp0*, were found to be enriched in the AVC and OFT. This is interesting given recent research showing the effects of knock-down or mutation of ribosomal proteins on development of several model systems (Amsterdam et al., 2004; MacInnes et al., 2008; McGowan et al., 2008; Uechi et al., 2006).

Following EMT, newly formed mesenchymal cells undergo proliferation resulting in the expansion of the endocardial cushions. In our list of AVC- and OFT-enriched genes we identified several mediators of cell proliferation such as *Ccnd2* and *Cdca7* (Rojas et al., 2008). Mesenchyme cells then undergo further differentiation characterized by expression of complex ECM molecules and matrix metalloproteinases (Armstrong and Bischoff, 2004; Chakraborty et al., 2008; Lincoln et al., 2004). Previously characterized endocardial cushion ECM and structural proteins such as *Tenascin C* (Lincoln et al., 2004), *Perlecan* (Sasse et al., 2008), and several collagens (e.g. *Col3a1*, *Col9a3*, *Col5a1*) (Peacock et al., 2008) were highly enriched in the AVC and OFT gene expression libraries at E10.5. Significantly, our list of AVC and OFT enriched genes also contained novel endocardial cushion ECM proteins and modifiers such as the matrix metalloproteinase 14 (*Mmp14*).

Interestingly, while remodeling of the valve tissue by apoptosis has been suggested to begin at E12.5 in the mouse, several regulators of apoptosis were highly enriched in the E10.5 developing cushions. These include positive regulators of apoptosis, such as *Traf4* (Sax and El-Deiry, 2003) and *Traf7* (Xu et al., 2004), and inhibitors such as *Bat3* (Tsukahara et al., 2009) and *Syvn1* (Yamasaki et al., 2007), suggesting a balance of pro- and anti-apoptotic signals.

Finally, gene ontology categories related to transcription regulatory activity were significantly ($p < 0.05$) overrepresented in our AVC- and OFT-enriched gene list suggesting that

our list is able to capture key regulators of AVC- and OFT-specific gene expression. Highly represented among the regulators of transcription were members of the bHLH transcription factor family including *Twist1*, *Twist2*, *Hes6*, and *Mlxip*, supporting an important role for this transcription factor family in AVC development. In particular, the bHLH transcription factor *Twist1* was found to be the most highly expressed DNA-binding transcription factor in the AVC at E10.5. Members of the Gata family (*Gata2*, *3*, *4* and *5*) were also highly represented in our list of endocardial cushion enriched transcription factors. Both GATA4 and GATA3 have been shown to play important roles during endocardial cushion development leading to AVC and OFT defects when mutated in mice (Crispino et al., 2001; Raid et al., 2009). In contrast, *Gata2* and *5* have not previously been described in the context of endocardial cushion development. Other transcription factors of interest not previously identified during valve development include *Klf4* and *Zeb1* (an inhibitor and an inducer of EMT, respectively) (Vandewalle et al., 2009; Yori et al., 2010), and the mediator of Wnt signalling, *Tblix* (Li and Wang, 2008).

3.3.2. *Twist1* null AVC shows dramatic gene expression changes

The bHLH transcription factor *Twist1* was found to be the most highly expressed DNA-binding transcription factor in the AVC at E10.5. Previous research from our lab and others established that *Twist1* is specifically expressed in mesenchyme cells of the AVC and OFT, and has a temporal expression pattern that peaks at E10.5 before decreasing at E12.5 (Ma et al., 2005; Vrljicak et al., 2009). Research in chick AVC development has suggested a role for TWIST1 in the promotion of proliferation and migration of endocardial cushion cells, coupled with an inhibition of their differentiation (Shelton and Yutzey, 2008). However, research in

Twist1 null mice has only revealed defects in neural crest cell contribution to the OFT (Vincentz et al., 2008b). The role of TWIST1 in mammalian AVC endocardial cushion development remains unclear due in part to the embryonic lethality of the *Twist1* null mutation at about E11 (Chen and Behringer, 1995), prior to valve maturation.

To determine which genes could be regulated by TWIST1 during AVC development, we created a Tag-seq library from the AVC of E10.5 *Twist1* null mice. Comparison of the *Twist1* null and wild-type AVC libraries demonstrated dramatic gene expression changes, with 60 genes up-regulated at least 10-fold in the *Twist1* null AVC, and 78 genes down-regulated at least 10-fold (Table 3.3). Interestingly, genes down-regulated in the *Twist1* null AVC were highly enriched in the wild-type AVC and OFT, while genes up-regulated in the *Twist1* null library were more likely to be enriched in atria and ventricles libraries (Figure 3.2A and Appendix X). This suggests that the *Twist1* null AVC had lost endocardial cushion specific gene expression and become more atria and ventricle-like.

Furthermore, gene ontology analysis revealed significant differences in the types of genes up- and down-regulated in the *Twist1* null AVC. Genes down-regulated in the *Twist1* null AVC were significantly enriched for ECM proteins ($p=2.7e-3$), such as *Col9a2*, *Coll6a1*, *Decorin* and *Biglycan*. Down-regulated genes were also significantly enriched for signalling molecules ($p=2.6e-2$), such as *Wnt9b*, *Nrg1*, and *Crebbp*, and included a high proportion of regulators of cell motility ($p=5.4e-2$), such as the Cd9 antigen and *Pf4*. In contrast, genes up-regulated in the *Twist1* null AVC were significantly enriched for nucleotide biosynthetic process ($p=1.3e-2$), and mitochondrial membrane ($p=6.5e-3$) categories. Genes up-regulated in the *Twist1* null library were also highly enriched for regulators of apoptosis ($p=6.7e-2$) such as *Vnn1* and *Caspase 9*.

Table 3.3. Selected gene expression changes in *Twist1* null AVC

Expression levels represent all sense tags for a gene and are shown as tags per million. *Twist1* enrichment was calculated as ratio of gene expression in *Twist1* null library over wild-type AVC.

| Gene symbol | Gene name | RefSeq accession | AVC | OFT | Atria | Ventricles | <i>Twist1</i> ^{-/-} | Twist1 enr. |
|--|---|------------------|-------|-------|-------|------------|------------------------------|-------------|
| A. Genes down-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| Transcription factors | | | | | | | | |
| <i>Twist1</i> | Twist homolog 1 | NM_011658 | 918 | 299.1 | 63.4 | 55.9 | 4.6 | 0.005 |
| <i>Tb11x</i> | Transducin (beta)-like 1 X-linked | NM_020601 | 24.2 | 20 | <1.3 | <1.4 | 0.6 | 0.026 |
| <i>Taf15</i> | TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor | NM_027427 | 104.6 | 92 | 7.3 | 10.9 | 7.8 | 0.076 |
| Signalling molecules | | | | | | | | |
| <i>Atxn1</i> | Ataxin 1 | NM_009124 | 9.6 | 11.8 | 1.7 | <1.4 | 0.8 | 0.081 |
| <i>Nrg1</i> | Neuregulin 1 | NM_178591 | 7.9 | 16.3 | <1.3 | <1.4 | <0.5 | 0.059 |
| <i>Smpd3</i> | Sphingomyelin phosphodiesterase 3 | NM_021491 | 6.8 | 2.9 | 1.9 | 1.4 | <0.5 | 0.068 |
| <i>Unc13b</i> | Unc-13 homolog B | NM_001081413 | 16.6 | 17.1 | <1.3 | <1.4 | 1.6 | 0.098 |
| <i>Wnt9b</i> | Wingless-type MMTV integration site 9B | NM_011719 | 8.1 | <1.1 | <1.3 | <1.4 | 0.5 | 0.067 |
| <i>Crebbp</i> | CREB binding protein | NM_001025432 | 6.9 | 8.4 | <1.3 | <1.4 | 0.6 | 0.089 |
| <i>Tnfrsf12a</i> | Tumor necrosis factor receptor superfamily, member 12a | NM_013749 | 12.5 | 17.3 | <1.3 | <1.4 | 1.2 | 0.098 |
| ECM components | | | | | | | | |
| <i>Bgn</i> | Biglycan | NM_007542 | 49.4 | 36.9 | <1.3 | <1.4 | 4.6 | 0.092 |
| <i>Col9a2</i> | Collagen, type IX, alpha 2 | NM_007741 | 14.8 | 17.5 | <1.3 | <1.4 | 0.7 | 0.047 |
| <i>Coll16a1</i> | Collagen, type XVI, alpha 1 | NM_028266 | 7.7 | 4 | 3 | 1.4 | 0.5 | 0.07 |
| <i>Dcn</i> | Decorin | NM_007833 | 60.4 | 41.8 | 33.6 | 33.9 | 4.7 | 0.078 |
| <i>Fmod</i> | Fibromodulin | NM_021355 | 5.6 | <1.1 | 3 | <1.4 | 0.5 | 0.096 |
| Bone, cartilage and tendon development | | | | | | | | |
| <i>Papss2</i> | 3'-phosphoadenosine 5'-phosphosulfate synthase 2 | NM_011864 | 218.9 | 115.5 | <1.3 | <1.4 | 12 | 0.055 |
| <i>Mgp</i> | Matrix Gla protein | NM_008597 | 12.3 | 1.9 | 2.6 | 1.7 | <0.5 | 0.038 |
| Other | | | | | | | | |
| <i>S100a10</i> | S100 calcium binding protein A10 (calpactin) | NM_009112 | 158.1 | 209.7 | 61.3 | 54 | 14.6 | 0.093 |
| <i>Syng2</i> | Synaptogyrin 2 | NM_009304 | 41.4 | 27.4 | 15.2 | 15.2 | 2.2 | 0.054 |

| Gene symbol | Gene name | RefSeq accession | AVC | OFT | Atria | Ventricles | <i>Twist1</i> ^{-/-} | <i>Twist1</i> enr. |
|--|--|------------------|-------|-------|-------|------------|------------------------------|--------------------|
| B. Genes up-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| Transcription factors | | | | | | | | |
| <i>Gata1</i> | GATA binding protein 1 | NM_008089 | 1 | <1.1 | 12.2 | 4.7 | 10.2 | 10.2 |
| <i>Zfp652</i> | Zinc finger protein 652 | NM_201609 | 1 | <1.1 | <1.3 | 5.2 | 11.6 | 11.7 |
| <i>Taf2</i> | TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor | NM_001081288 | 3.5 | 4.9 | <1.3 | <1.4 | 47.4 | 13.7 |
| <i>Irf2</i> | Interferon regulatory factor 2 | NM_008391 | 1 | 10.5 | 7.3 | 15.4 | 23.6 | 23.9 |
| Signalling molecules | | | | | | | | |
| <i>Fst</i> | Follistatin | NM_008046 | 1 | 1.1 | 15.9 | <1.4 | 15.3 | 15.5 |
| Apoptosis | | | | | | | | |
| <i>Casp9</i> | Caspase 9 | NM_015733 | <1 | <1.1 | <1.3 | <1.4 | 13.9 | 14 |
| <i>Prkdc</i> | Protein kinase, DNA activated, catalytic polypeptide | NM_011159 | 3 | 6.3 | 20.8 | 32.7 | 33.2 | 11.2 |
| <i>Vnn1</i> | Vanin 1 | NM_011704 | 2.1 | 16.3 | <1.3 | 5.9 | 25.5 | 11.9 |
| <i>Zc3hc1</i> | Zinc finger, C3HC type 1 | NM_172735 | 8.7 | 9.7 | 51.4 | 96.8 | 92.3 | 10.6 |
| Other | | | | | | | | |
| <i>Sf3b4</i> | Splicing factor 3b, subunit 4 | NM_153053 | 22.1 | 51.2 | 185.4 | 225.1 | 384.8 | 17.4 |
| <i>Wdr75</i> | WD repeat domain 75 | NM_028599 | 5.8 | 20.3 | 12 | 15.4 | 74.1 | 12.9 |
| <i>1110059 E24Rik</i> | RIKEN cDNA 1110059E24 gene | NM_025423 | 4.1 | 2.5 | 1.9 | 8.8 | 57.6 | 14 |
| <i>Pi4ka</i> | Phosphatidylinositol 4-kinase, catalytic, alpha polypeptide | NM_001001983 | 7.9 | 13.6 | 81.4 | 76.2 | 110.1 | 13.9 |
| <i>Adss1l</i> | Adenylosuccinate synthetase like 1 | NM_007421 | 21.3 | 25 | 58.1 | 146 | 216 | 10.2 |
| <i>Wdr74</i> | WD repeat domain 74 | NM_134139 | 7.2 | 14.2 | 34.5 | 32.2 | 124 | 17.1 |
| C. Putative TWIST1 targets from literature | | | | | | | | |
| <i>Acan</i> | Aggrecan | NM_007424 | <1 | 1 | <1.3 | <1.4 | <0.5 | 0.468 |
| <i>Cdh11</i> | Cadherin 11 | NM_009866 | 55.1 | 33.4 | 31.1 | 35.7 | 33.7 | 0.611 |
| <i>Mmp2</i> | Matrix metalloproteinase 2 | NM_008610 | 194.2 | 237.2 | 26.8 | 29.6 | 105.8 | 0.545 |
| <i>Postn</i> | Periostin, osteoblast specific factor | NM_015784 | 144.1 | 138.6 | 52.5 | 32.9 | 29.0 | 0.201 |
| <i>Tbx20</i> | T-box 20 transcript variant 1 | NM_194263 | 51.7 | 40.1 | 1.5 | <1.4 | 43.3 | 0.837 |

Interestingly, the putative TWIST1 targets *Periostin*, *Cdh11* and *Mmp2* (Shelton and Yutzey, 2008) showed an 80%, 60% and 50% decrease in expression levels, respectively, while other putative TWIST1 targets identified in chick endocardial cushion cells (*Tbx20* and *Aggrecan*) showed less than 15% expression changes within the *Twist1* null library (Table 3.3). This suggests that TWIST1 might not be sufficient to regulate these genes in the mouse, but that it may act as a modulator in combination with other AVC-specific transcription factors.

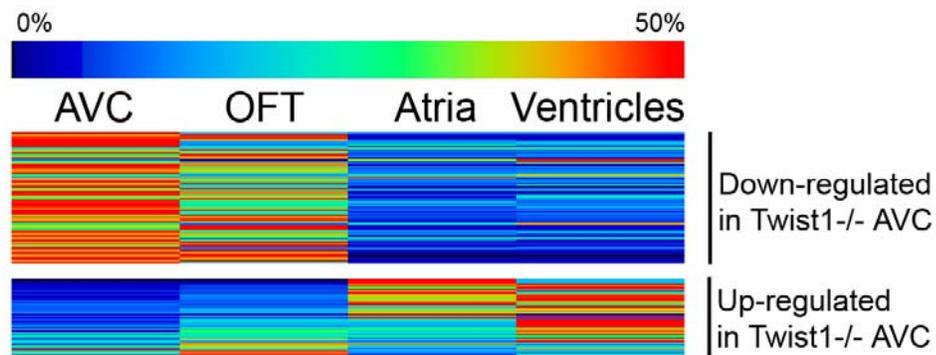
The Tag-seq libraries generated by our lab represent a heterogeneous population of cells; therefore overall differences in gene expression could reflect either changes in expression within a cell-type or changes in the proportion of cells within the cell population. Notably, *Twist1* null embryos did not appear to show a significant developmental delay, containing similar number of somites at the time of dissection as their wild-type littermates, but the hearts of *Twist1* null embryos were 15-20% smaller overall (Appendix XI). This overall heart size difference was reflected in 30% smaller AVC endocardial cushions in the *Twist1* null hearts, raising the possibility that the changes in gene expression observed were an indirect consequence of the embryonic phenotype caused by the *Twist1* null mutation.

To address this possibility, we examined the temporal expression pattern of the genes differentially expressed in the *Twist1* null AVC. We first identified genes showing at least a 10-fold differential gene expression in the *Twist1* null AVC and a dynamic temporal expression pattern from E9.5 to E12.5 (Vrljicak et al., 2009). We then grouped these genes according to their peak expression over the E9.5 to E12.5 timeframe (Figure 3.2B and Appendix XII).

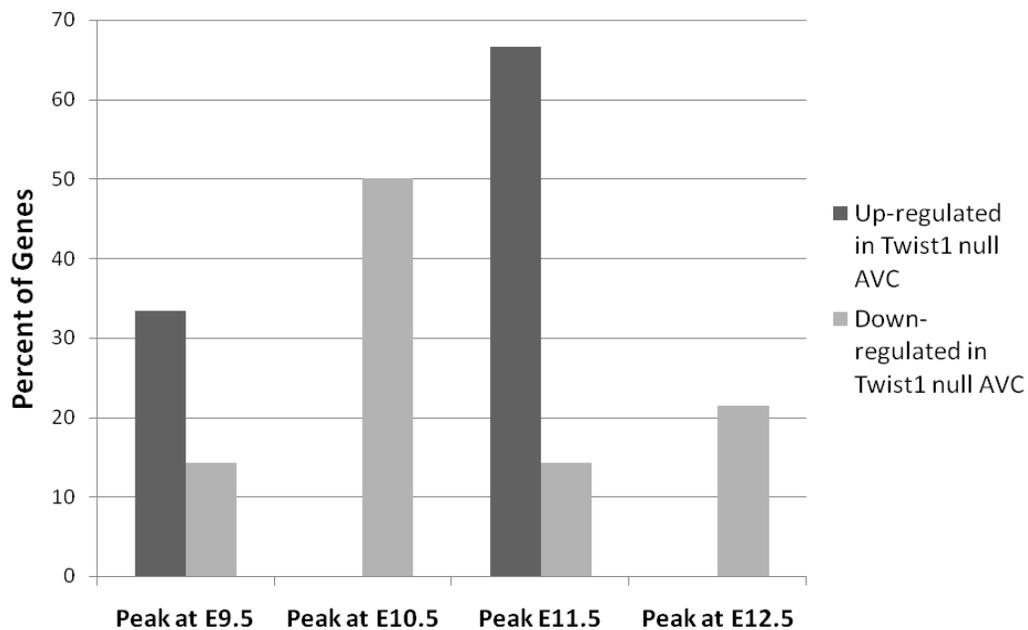
Figure 3.2. Altered gene expression in *Twist1* null AVC

A. Genes down-regulated in the *Twist1* null AVC are enriched in wild-type AVC and OFT, while genes up-regulated in *Twist1* null AVC tend to be enriched in atria and ventricles. B. Genes down-regulated in *Twist1* null AVC show peak expression at E10.5, while up-regulated genes show peak expression at E11.5. Expression of all tag-types mapping in the sense direction to the same RefSeq gene were pooled.

A



B



Genes that were down-regulated in the *Twist1* null AVC were more likely to have peak expression at E10.5, matching the peak expression of *Twist1*. In contrast, genes that were up-regulated in the mutant AVC were more likely to show peak expression at E11.5. Because a developmental delay would have been expected to result in a *Twist1* null AVC expression more similar to the wild-type E9.5 AVC, these results indicate that the gene expression changes observed in the *Twist1* null AVC cannot be solely explained by an overall developmental delay. Furthermore, they support a role of TWIST1 in the inhibition of AVC endocardial cushion cell maturation following EMT.

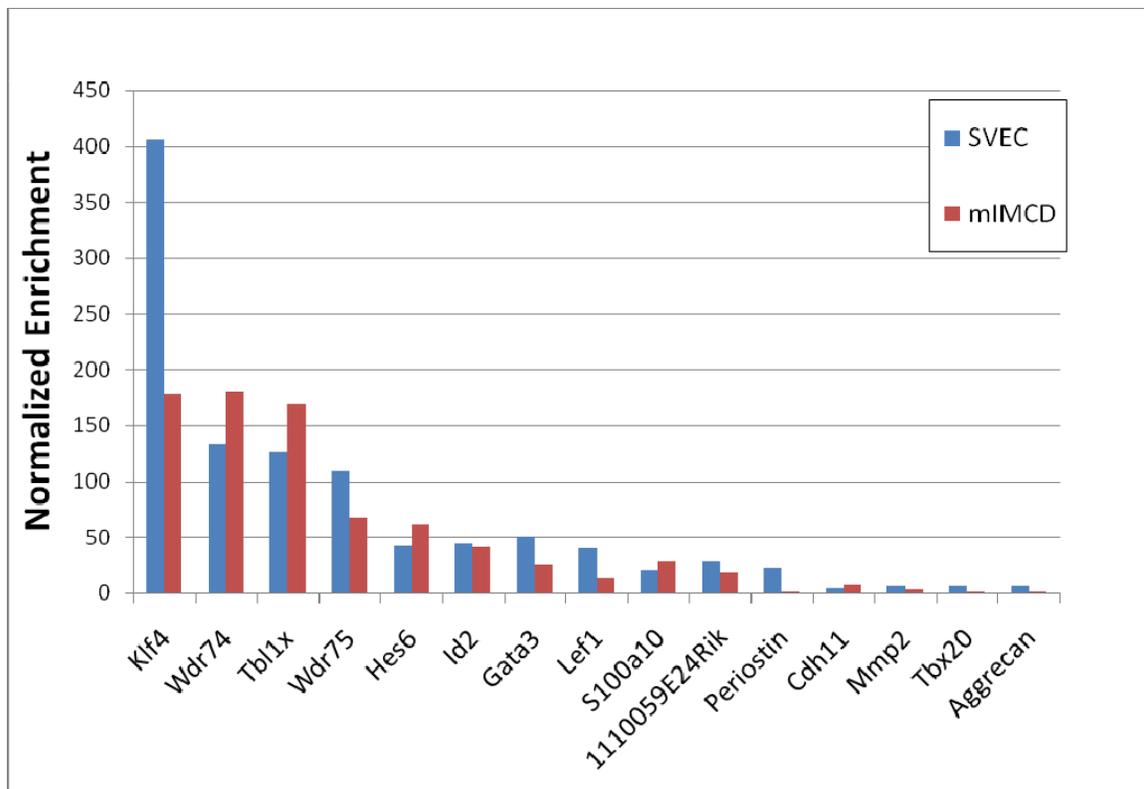
3.3.3. TWIST1 directly regulates AVC gene expression

Typically, dimerization of TWIST1 with ubiquitously expressed bHLH factors (E-proteins, such as E12) results in the juxtaposition of their basic domains creating a combined DNA binding motif capable of binding to an E-box sequence (3'CANNTG5') (Laursen et al., 2007). DNA binding usually leads to the activation of gene expression; however TWIST1 can also act as a negative regulator of gene expression by direct interaction with the basic domain of other bHLHs or by sequestering E-proteins (Hamamori et al., 1997; Spicer et al., 1996). To test whether the gene expression changes observed in the *Twist1* null AVC were a result of direct binding of TWIST1 to DNA we sought to identify direct activity of TWIST1 at the promoter level. We concentrated our analysis on selected genes with 10-fold differential gene expression in the *Twist1* null AVC and dynamic temporal expression patterns, as well as selected AVC- and OFT-enriched transcription factors, and five putative TWIST1 targets identified in chick AVC endocardial cushion cells (Shelton and Yutzey, 2008).

First the promoters of possible TWIST1 target genes were analyzed for E-box sequences. All the promoters from our selected list were found to contain multiple E-box sequences within -2kb and +200bp of the transcriptional start site (data not shown). To identify which of these promoters were directly bound by TWIST1 we performed ChIP-qPCR on endothelial and epithelial cell lines (SVEC and mIMCD, respectively) over-expressing a myc-tagged version of *Twist1* (Figure 3.3 and Appendix XIII). Ten of the 40 gene promoters we investigated were enriched over 20-fold after ChIP-qPCR, indicating that TWIST1 was bound at these DNA sequences and suggesting a direct role of TWIST1 in their regulation. Interestingly, six of the eight most enriched promoters bound by TWIST1 were transcription factors (*Klf4*, *Tb11x*, *Hes6*, *Id2*, *Lef1* and *Gata3*) suggesting that TWIST1 is acting as a master regulator of transcription in developing endocardial cushions. Unexpectedly, we observed no evidence for direct binding of TWIST1 to the promoters of the five putative TWIST1 targets (*Periostin*, *Cdh11*, *Mmp2*, *Tbx20* and *Aggrecan*). However, it is possible that TWIST1 binding could be occurring at regulatory sites outside of the sequences tested, such as in distant enhancers or introns, or that the necessary co-factors are not expressed in the cell lines used. Taken together these results suggest that TWIST1 plays a critical role in endocardial cushion development, by directly regulating spatially and temporally restricted gene expression.

Figure 3.3. Chromatin immunoprecipitation

Myc-tagged *Twist1* was over-expressed in mIMCD and SVEC cells, and immunoprecipitated with anti-myc antibody. Enrichment was calculated as 2 to the power of the cycle threshold (cT) difference between the IgG immunoprecipitated sample and the anti-myc antibody immunoprecipitated sample. Enrichment was normalized to lowest value over 40 promoters analyzed. See Appendix XIII for complete results.



3.4. Discussion

The morphogenetic process of endocardial cushion transformation and maturation that occurs in the AVC and OFT of the rudimentary heart is critical for normal valve development and cardiac septation leading to a mature, four-chambered heart. In this study we describe the generation and analysis of five Tag-seq libraries from atria, ventricles, OFT, as well as wild-type and *Twist1* null AVCs of the developing E10.5 mouse heart. These Tag-seq libraries were sequenced at a depth of over 7 million tags per library. SAGE libraries sequenced at a depth of 120,000 tags have comparable sensitivity to fluorescent-based microarray approaches (Lu et al., 2004). Furthermore, studies on mouse and human cell lines have found a significant decrease in new transcript discovery once 300,000 tags have been attained (Akmaev and Wang, 2004) with unique tag discovery approaching zero at ~650,000 tags collected (Velculescu et al., 1999). Therefore, we have likely reached saturation of the heart transcriptome at this time-point, improving our capacity for gene and transcript discovery.

3.4.1. Identification of AVC- and OFT-enriched genes

Using these Tag-seq libraries we identified a list of 565 high-confidence AVC- and OFT-enriched genes. A majority of AVC-enriched genes overlapped substantially with an OFT-enriched gene list, highlighting the shared mechanisms underlying the development of these two regions. However, there were some notable exceptions with several genes enriched in the AVC but not the OFT, and vice versa. These gene expression differences between the AVC and OFT

could reflect differences in timing of cushion formation, the cells populating these regions or the cardiogenic lineage giving rise to them. For example, the most differentially expressed gene in the OFT was found to be *Isl1*, a marker of the secondary heart field that contributes cells to the OFT and right ventricle (Cai et al., 2003), while the most highly AVC specific gene was found to be *Galanin*, a neuropeptide known to be expressed in the AVC before becoming restricted to AV-node and AV-rings (Schweickert et al., 2008).

A majority of the genes identified in our list of AVC- and OFT-enriched genes have not been described in the context of endocardial cushion development. They included ECM structural proteins and modifiers (e.g. *Mmp14*) as well as regulators of apoptosis and proliferation (*Bat3*, *Traf4* and *Traf7*, as well as *Ccnd2* and *Cdca7*, respectively). They also included novel endocardial cushion signalling molecules (such as *Igfbp5* and *Csnkle*) and transcription factors (*Tbl1x*, *Klf4* and *Tead2*). Interestingly, while a number of these AVC- and OFT-enriched genes have not been described in the context of heart development (such as the transcription factors *Zeb1*, *Calcoco1*, and *Taf6*), others (such as *Foxm1* and *Mef2c*) have known roles during cardiac morphogenesis (Ramakrishna et al., 2007; Vincentz et al., 2008a), suggesting that they might play multiple functions during heart development.

3.4.2. Regulation of spatial and temporal expression patterns

Twist1 was the highest expressed DNA-binding transcription factor in the E10.5 AVC where its expression is restricted to the mesenchyme cell population (Ma et al., 2005; Vrljicak et al., 2009). Research in chick AVC development has suggested a role for TWIST1 in the promotion of proliferation and migration of endocardial cushion cells, coupled with an inhibition

of their differentiation (Shelton and Yutzey, 2008). However, no obvious morphological differences were observed in the *Twist1* null mouse AVC (Vincentz et al., 2008b). In addition TWIST1 has also been shown to regulate apoptosis in other systems (Maestro et al., 1999). Comparison of *Twist1* null and wild-type AVC expression revealed significant expression changes consistent with these proposed roles for TWIST1. Genes down-regulated in the *Twist1* null AVC were enriched in AVC and OFT suggesting that TWIST1 activity is required for proper AVC and OFT gene expression. Down-regulated genes were enriched for cell motility and ECM molecules expressed during differentiation. Up-regulated genes were enriched in apoptosis categories. Furthermore, up-regulation of genes normally expressed at E11.5 suggests that TWIST1 might act to block maturation of mesenchyme cells.

Surprisingly, from the genes previously identified as TWIST1 targets in the chick AVC (Shelton and Yutzey, 2008), only *Periostin*, *Cdh11* and *Mmp2* showed a significant expression change in the *Twist1* null AVC. This suggests that in the mouse TWIST1 might not be sufficient to regulate the transcription of *Tbx20* and *Aggrecan*, but that it might act as a modulator. Alternatively, other members of the bHLH transcription factor family, such as TWIST2 (Sosic et al., 2003), could compensate for lack of TWIST1 activity in the mouse.

Using ChIP-qPCR we analyzed the ability of TWIST1 to bind to the promoters of our list of differentially expressed genes and identified a number of novel direct targets of TWIST1. Genes whose promoters are bound by TWIST1 could be either up-regulated in the *Twist1* null AVC, such as *Wdr74* and *Wdr75*, or down-regulated, such as *Tblix*. This suggests that TWIST1 could act both as an activator and inhibitor of transcription in the context of AVC development.

Several of our novel direct and indirect targets of TWIST1 suggest a role for TWIST1 as a critical regulator of multiple pathways important for AVC development. For example, TWIST1

was found to control the expression of *Lef1* and *Tbllx*, members of the Wnt pathway important for AVC endocardial cushion formation (Liebner et al., 2004). Similarly, TWIST1 was found to regulate expression of the bHLH transcription factor *Hes6*, a regulator of HES1 downstream of Notch signalling (Bae et al., 2000). Finally, TWIST1 is known to interfere with BMP pathway induced gene expression (Hayashi et al., 2007; Reinhold et al., 2006). The regulation by TWIST1 of *Decorin* and *Sf3b4*, two genes known to regulate the BMP pathway (Nishanian and Waldman, 2004; Ruiz-Lozano et al., 1997; Takeuchi et al., 1994; Watanabe et al., 2007), suggests that TWIST1 might modulate BMP signalling through their activity.

Finally, while our TWIST1 binding analysis was restricted to promoter regions, gene regulatory regions have been found within exons, introns and intergenic regions (Wederell et al., 2008). Therefore an absence of ChIP-qPCR enrichment could reflect regulation by sequences outside of the regions tested. The use of the ChIP technique in combination with massively parallel sequencing would be necessary to establish TWIST1-DNA associations at the genome level (Robertson et al., 2007a).

In summary, we have used the SAGE technique coupled with deep sequencing technology to obtain transcriptome information for atria, ventricles, AVC and OFT of the E10.5 developing mouse heart. We used this resource to identify region specific gene expression and the transcriptional regulation of these patterns. Our findings are consistent with a role for TWIST1 in the differentiation of AVC mesenchyme post-EMT in the mouse, and suggested that its activity is required for inhibiting maturation of the newly formed mesenchymal cells in the AVC.

3.5. References

- Akmaev, V. R. and Wang, C. J. (2004) Correction of sequence-based artifacts in serial analysis of gene expression. *Bioinformatics* 20, 1254-1263.
- Amsterdam, A., Sadler, K. C., Lai, K., Farrington, S., Bronson, R. T., Lees, J. A. and Hopkins, N. (2004) Many ribosomal protein genes are cancer genes in zebrafish. *PLoS Biol.* 2, E139.
- Andress, D. L. and Birnbaum, R. S. (1992) Human osteoblast-derived insulin-like growth factor (IGF) binding protein-5 stimulates osteoblast mitogenesis and potentiates IGF action. *J.Biol.Chem.* 267, 22467-22472.
- Armstrong, E. J. and Bischoff, J. (2004) Heart valve development: endothelial cell signaling and differentiation. *Circ.Res.* 95, 459-470.
- Bae, S., Bessho, Y., Hojo, M. and Kageyama, R. (2000) The bHLH gene *Hes6*, an inhibitor of *Hes1*, promotes neuronal differentiation. *Development* 127, 2933-2943.
- Cai, C. L., Liang, X., Shi, Y., Chu, P. H., Pfaff, S. L., Chen, J. and Evans, S. (2003) *Isl1* identifies a cardiac progenitor population that proliferates prior to differentiation and contributes a majority of cells to the heart. *Dev.Cell.* 5, 877-889.
- Camenisch, T. D., Molin, D. G., Person, A., Runyan, R. B., Gittenberger-de Groot, A. C., McDonald, J. A. and Klewer, S. E. (2002) Temporal and distinct TGFbeta ligand requirements during mouse and avian endocardial cushion morphogenesis. *Dev.Biol.* 248, 170-181.
- Chakraborty, S., Cheek, J., Sakthivel, B., Aronow, B. J. and Yutzey, K. E. (2008) Shared gene expression profiles in developing heart valves and osteoblast progenitor cells. *Physiol.Genomics*
- Chen, G. and Courey, A. J. (2000) Groucho/TLE family proteins and transcriptional repression. *Gene* 249, 1-16.
- Chen, Z. F. and Behringer, R. R. (1995) Twist is Required in Head Mesenchyme for Cranial Neural Tube Morphogenesis. *Genes Dev.* 9, 686-699.
- Crispino, J. D., Lodish, M. B., Thurberg, B. L., Litovsky, S. H., Collins, T., Molkentin, J. D. and Orkin, S. H. (2001) Proper coronary vascular development and heart morphogenesis depend on interaction of GATA-4 with FOG cofactors. *Genes Dev.* 15, 839-844.
- de la Pompa, J. L., Timmerman, L. A., Takimoto, H., Yoshida, H., Elia, A. J., Samper, E., Potter, J., Wakeham, A., Marengere, L., Langille, B. L., Crabtree, G. R. and Mak, T. W. (1998) Role of the NF-ATc transcription factor in morphogenesis of cardiac valves and septum. *Nature* 392, 182-186.
- Dennis, G., Jr, Sherman, B. T., Hosack, D. A., Yang, J., Gao, W., Lane, H. C. and Lempicki, R. A. (2003) DAVID: Database for Annotation, Visualization, and Integrated Discovery. *Genome Biol.* 4, P3.
- Fior, R. and Henrique, D. (2005) A novel *hes5/hes6* circuitry of negative regulation controls Notch activity during neurogenesis. *Dev.Biol.* 281, 318-333.
- Firulli, B. A., Krawchuk, D., Centonze, V. E., Vargesson, N., Virshup, D. M., Conway, S. J., Cserjesi, P., Laufer, E. and Firulli, A. B. (2005) Altered Twist1 and Hand2 dimerization

- is associated with Saethre-Chotzen syndrome and limb abnormalities. *Nat.Genet.* 37, 373-381.
- Hamamori, Y., Wu, H. Y., Sartorelli, V. and Kedes, L. (1997) The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol.Cell.Biol.* 17, 6563-6573.
- Hayashi, M., Nimura, K., Kashiwagi, K., Harada, T., Takaoka, K., Kato, H., Tamai, K. and Kaneda, Y. (2007) Comparative roles of Twist-1 and Id1 in transcriptional regulation by BMP signaling. *J.Cell.Sci.* 120, 1350-1357.
- Heinke, J., Wehofsits, L., Zhou, Q., Zoeller, C., Baar, K. M., Helbing, T., Laib, A., Augustin, H., Bode, C., Patterson, C. and Moser, M. (2008) BMPER is an endothelial cell regulator and controls bone morphogenetic protein-4-dependent angiogenesis. *Circ.Res.* 103, 804-812.
- Hinton, R. B., Jr, Lincoln, J., Deutsch, G. H., Osinska, H., Manning, P. B., Benson, D. W. and Yutzey, K. E. (2006) Extracellular matrix remodeling and organization in developing and diseased aortic valves. *Circ.Res.* 98, 1431-1438.
- Huang da, W., Sherman, B. T. and Lempicki, R. A. (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat.Protoc.* 4, 44-57.
- Kelley, R., Ren, R., Pi, X., Wu, Y., Moreno, I., Willis, M., Moser, M., Ross, M., Podkova, M., Attisano, L. and Patterson, C. (2009) A concentration-dependent endocytic trap and sink mechanism converts Bmper from an activator to an inhibitor of Bmp signaling. *J.Cell Biol.* 184, 597-609.
- Kim, K. A., Wagle, M., Tran, K., Zhan, X., Dixon, M. A., Liu, S., Gros, D., Korver, W., Yonkovich, S., Tomasevic, N., Binnerts, M. and Abo, A. (2008) R-Spondin family members regulate the Wnt pathway by a common mechanism. *Mol.Biol.Cell* 19, 2588-2596.
- Kim, S. Y., Dunn, I. F., Firestein, R., Gupta, P., Wardwell, L., Repich, K., Schinzel, A. C., Wittner, B., Silver, S. J., Root, D. E., *et al.* (2010) CK1epsilon is required for breast cancers dependent on beta-catenin activity. *PLoS One* 5, e8979.
- Kruithof, B. P., Krawitz, S. A. and Gaussin, V. (2007) Atrioventricular valve development during late embryonic and postnatal stages involves condensation and extracellular matrix remodeling. *Dev.Biol.* 302, 208-217.
- Laursen, K. B., Mielke, E., Iannaccone, P. and Fuchtbauer, E. M. (2007) Mechanism of transcriptional activation by the proto-oncogene Twist1. *J.Biol.Chem.* 282, 34623-34633.
- Li, J. and Wang, C. Y. (2008) TBL1-TBLR1 and beta-catenin recruit each other to Wnt target-gene promoter for transcription activation and oncogenesis. *Nat.Cell Biol.* 10, 160-169.
- Liebner, S., Cattelino, A., Gallini, R., Rudini, N., Iurlaro, M., Piccolo, S. and Dejana, E. (2004) Beta-catenin is required for endothelial-mesenchymal transformation during heart cushion development in the mouse. *J.Cell Biol.* 166, 359-367.
- Lincoln, J., Kist, R., Scherer, G. and Yutzey, K. E. (2007) Sox9 is required for precursor cell expansion and extracellular matrix organization during mouse heart valve development. *Dev.Biol.* 305, 120-132.
- Lincoln, J., Alfieri, C. M. and Yutzey, K. E. (2004) Development of heart valve leaflets and supporting apparatus in chicken and mouse embryos. *Dev.Dyn.* 230, 239-250.
- Ma, L., Lu, M. F., Schwartz, R. J. and Martin, J. F. (2005) Bmp2 is essential for cardiac cushion epithelial-mesenchymal transition and myocardial patterning. *Development* 132, 5601-5611.

- MacInnes, A. W., Amsterdam, A., Whittaker, C. A., Hopkins, N. and Lees, J. A. (2008) Loss of p53 synthesis in zebrafish tumors with ribosomal protein gene mutations. *Proc.Natl.Acad.Sci.U.S.A.* 105, 10408-10413.
- Maestro, R., Dei Tos, A. P., Hamamori, Y., Krasnokutsky, S., Sartorelli, V., Kedes, L., Doglioni, C., Beach, D. H. and Hannon, G. J. (1999) Twist is a potential oncogene that inhibits apoptosis. *Genes Dev.* 13, 2207-2217.
- Markwald, R. R., Fitzharris, T. P. and Smith, W. N. (1975) Structural analysis of endocardial cytodifferentiation. *Dev.Biol.* 42, 160-180.
- McGowan, K. A., Li, J. Z., Park, C. Y., Beaudry, V., Tabor, H. K., Sabnis, A. J., Zhang, W., Fuchs, H., de Angelis, M. H., Myers, R. M., Attardi, L. D. and Barsh, G. S. (2008) Ribosomal mutations cause p53-mediated dark skin and pleiotropic effects. *Nat.Genet.* 40, 963-970.
- Morrissy, A. S., Morin, R. D., Delaney, A., Zeng, T., McDonald, H., Jones, S., Zhao, Y., Hirst, M. and Marra, M. A. (2009) Next-generation tag sequencing for cancer gene expression profiling. *Genome Res.* 19, 1825-1835.
- Moser, M., Binder, O., Wu, Y., Aitsebaomo, J., Ren, R., Bode, C., Bautch, V. L., Conlon, F. L. and Patterson, C. (2003) BMPER, a novel endothelial cell precursor-derived protein, antagonizes bone morphogenetic protein signaling and endothelial cell differentiation. *Mol.Cell.Biol.* 23, 5664-5679.
- Neilsen, P. M., Cheney, K. M., Li, C. W., Chen, J. D., Cawrse, J. E., Schulz, R. B., Powell, J. A., Kumar, R. and Callen, D. F. (2008) Identification of ANKRD11 as a p53 coactivator. *J.Cell.Sci.* 121, 3541-3552.
- Niessen, K., Fu, Y., Chang, L., Hoodless, P. A., McFadden, D. and Karsan, A. (2008) Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J.Cell Biol.* 182, 315-325.
- Nishanian, T. G. and Waldman, T. (2004) Interaction of the BMPR-IA tumor suppressor with a developmentally relevant splicing factor. *Biochem.Biophys.Res.Comm.* 323, 91-97.
- Nosedá, M., Chang, L., McLean, G., Grim, J. E., Clurman, B. E., Smith, L. L. and Karsan, A. (2004) Notch activation induces endothelial cell cycle arrest and participates in contact inhibition: role of p21Cip1 repression. *Mol.Cell.Biol.* 24, 8813-8822.
- Oka, C., Tsujimoto, R., Kajikawa, M., Koshiba-Takeuchi, K., Ina, J., Yano, M., Tsuchiya, A., Ueta, Y., Soma, A., Kanda, H., Matsumoto, M. and Kawaichi, M. (2004) HtrA1 serine protease inhibits signaling mediated by Tgfbeta family proteins. *Development* 131, 1041-1053.
- Peacock, J. D., Lu, Y., Koch, M., Kadler, K. E. and Lincoln, J. (2008) Temporal and spatial expression of collagens during murine atrioventricular heart valve development and maintenance. *Dev.Dyn.* 237, 3051-3058.
- Person, A. D., Klewer, S. E. and Runyan, R. B. (2005) Cell biology of cardiac cushion development. *Int.Rev.Cytol.* 243, 287-335.
- Pissarra, L., Henrique, D. and Duarte, A. (2000) Expression of *hes6*, a new member of the Hairy/Enhancer-of-split family, in mouse development. *Mech.Dev.* 95, 275-278.
- Raid, R., Krinka, D., Bakhoff, L., Abdelwahid, E., Jokinen, E., Karner, M., Malva, M., Meier, R., Pelliniemi, L. J., Ploom, M., *et al.* (2009) Lack of *Gata3* results in conotruncal heart anomalies in mouse. *Mech.Dev.* 126, 80-89.
- Ramakrishna, S., Kim, I. M., Petrovic, V., Malin, D., Wang, I. C., Kalin, T. V., Meliton, L., Zhao, Y. Y., Ackerson, T., Qin, Y., *et al.* (2007) Myocardium defects and ventricular

- hypoplasia in mice homozygous null for the Forkhead Box M1 transcription factor. *Dev.Dyn.* 236, 1000-1013.
- Ranger, A. M., Grusby, M. J., Hodge, M. R., Gravallesse, E. M., de la Brousse, F. C., Hoey, T., Mickanin, C., Baldwin, H. S. and Glimcher, L. H. (1998) The transcription factor NF-ATc is essential for cardiac valve formation. *Nature* 392, 186-190.
- Reinhold, M. I., Kapadia, R. M., Liao, Z. and Naski, M. C. (2006) The Wnt-inducible transcription factor Twist1 inhibits chondrogenesis. *J.Biol.Chem.* 281, 1381-1388.
- Rivera-Feliciano, J., Lee, K. H., Kong, S. W., Rajagopal, S., Ma, Q., Springer, Z., Izumo, S., Tabin, C. J. and Pu, W. T. (2006) Development of heart valves requires Gata4 expression in endothelial-derived cells. *Development* 133, 3607-3618.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., *et al.* (2007a) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat.Methods* 4, 651-657.
- Robertson, N., Oveisi-Fordorei, M., Zuyderduyn, S. D., Varhol, R. J., Fjell, C., Marra, M., Jones, S. and Siddiqui, A. (2007b) DiscoverySpace: an interactive data analysis application. *Genome Biol.* 8, R6.
- Rojas, A., Kong, S. W., Agarwal, P., Gilliss, B., Pu, W. T. and Black, B. L. (2008) GATA4 is a direct transcriptional activator of cyclin D2 and Cdk4 and is required for cardiomyocyte proliferation in anterior heart field-derived myocardium. *Mol.Cell.Biol.* 28, 5420-5431.
- Ruiz-Lozano, P., Doevendans, P., Brown, A., Gruber, P. J. and Chien, K. R. (1997) Developmental expression of the murine spliceosome-associated protein mSAP49. *Dev.Dyn.* 208, 482-490.
- Sasse, P., Malan, D., Fleischmann, M., Roell, W., Gustafsson, E., Bostani, T., Fan, Y., Kolbe, T., Breitbach, M., Addicks, K., *et al.* (2008) Perlecan is critical for heart stability. *Cardiovasc.Res.* 80, 435-444.
- Sax, J. K. and El-Deiry, W. S. (2003) Identification and characterization of the cytoplasmic protein TRAF4 as a p53-regulated proapoptotic gene. *J.Biol.Chem.* 278, 36435-36444.
- Schweickert, A., Deissler, K., Britsch, S., Albrecht, M., Ehmann, H., Mauch, V., Gaio, U. and Blum, M. (2008) Left-asymmetric expression of Galanin in the linear heart tube of the mouse embryo is independent of the nodal co-receptor gene cryptic. *Dev.Dyn.* 237, 3557-3564.
- Shelton, E. L. and Yutzey, K. E. (2008) Twist1 function in endocardial cushion cell proliferation, migration, and differentiation during heart valve development. *Dev.Biol.* 317, 282-295.
- Shelton, E. L. and Yutzey, K. E. (2007) Tbx20 regulation of endocardial cushion cell proliferation and extracellular matrix gene expression. *Dev.Biol.* 302, 376-388.
- Siddiqui, A. S., Khattra, J., Delaney, A. D., Zhao, Y., Astell, C., Asano, J., Babakaiff, R., Barber, S., Beland, J., Bohacec, S., *et al.* (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc.Natl.Acad.Sci.U.S.A.* 102, 18485-18490.
- Sosic, D., Richardson, J. A., Yu, K., Ornitz, D. M. and Olson, E. N. (2003) Twist regulates cytokine gene expression through a negative feedback loop that represses NF-kappaB activity. *Cell* 112, 169-180.
- Spicer, D. B., Rhee, J., Cheung, W. L. and Lassar, A. B. (1996) Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* 272, 1476-1480.

- Stelzer, C., Brimmer, A., Hermanns, P., Zabel, B. and Dietz, U. H. (2007) Expression profile of Papss2 (3'-phosphoadenosine 5'-phosphosulfate synthase 2) during cartilage formation and skeletal development in the mouse embryo. *Dev.Dyn.* 236, 1313-1318.
- 't Hoen, P. A., Ariyurek, Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H., de Menezes, R. X., Boer, J. M., van Ommen, G. J. and den Dunnen, J. T. (2008) Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res.* 36, e141.
- Takeuchi, Y., Kodama, Y. and Matsumoto, T. (1994) Bone matrix decorin binds transforming growth factor-beta and enhances its bioactivity. *J.Biol.Chem.* 269, 32634-32638.
- Timmerman, L. A., Grego-Bessa, J., Raya, A., Bertran, E., Perez-Pomares, J. M., Diez, J., Aranda, S., Palomo, S., McCormick, F., Izpisua-Belmonte, J. C. and de la Pompa, J. L. (2004) Notch promotes epithelial-mesenchymal transition during cardiac development and oncogenic transformation. *Genes Dev.* 18, 99-115.
- Tsukahara, T., Kimura, S., Ichimiya, S., Torigoe, T., Kawaguchi, S., Wada, T., Yamashita, T. and Sato, N. (2009) Scythe/BAT3 regulates apoptotic cell death induced by papillomavirus binding factor in human osteosarcoma. *Cancer.Sci.* 100, 47-53.
- Uechi, T., Nakajima, Y., Nakao, A., Torihara, H., Chakraborty, A., Inoue, K. and Kenmochi, N. (2006) Ribosomal protein gene knockdown causes developmental defects in zebrafish. *PLoS One* 1, e37.
- Vandewalle, C., Van Roy, F. and Berx, G. (2009) The role of the ZEB family of transcription factors in development and disease. *Cell Mol.Life Sci.* 66, 773-787.
- Velculescu, V. E., Madden, S. L., Zhang, L., Lash, A. E., Yu, J., Rago, C., Lal, A., Wang, C. J., Beaudry, G. A., Ciriello, K. M., *et al.* (1999) Analysis of human transcriptomes. *Nat.Genet.* 23, 387-388.
- Velculescu, V. E., Zhang, L., Vogelstein, B. and Kinzler, K. W. (1995) Serial analysis of gene expression. *Science* 270, 484-487.
- Vincentz, J. W., Barnes, R. M., Firulli, B. A., Conway, S. J. and Firulli, A. B. (2008a) Cooperative interaction of Nkx2.5 and Mef2c transcription factors during heart development. *Dev.Dyn.* 237, 3809-3819.
- Vincentz, J. W., Barnes, R. M., Rodgers, R., Firulli, B. A., Conway, S. J. and Firulli, A. B. (2008b) An absence of Twist1 results in aberrant cardiac neural crest morphogenesis. *Dev.Biol.* 320, 131-139.
- Vrljicak, P. J., Chang, A. C., Morozova, O., Wederell, E. D., Niessen, K., Marra, M. A., Karsan, A. and Hoodless, P. A. (2009) Genomic Analysis Distinguishes Phases of Early Development of the Mouse Atrio-Ventricular Canal. *Physiol.Genomics*
- Watanabe, H., Shionyu, M., Kimura, T., Kimata, K. and Watanabe, H. (2007) Splicing factor 3b subunit 4 binds BMPR-IA and inhibits osteochondral cell differentiation. *J.Biol.Chem.* 282, 20728-20738.
- Watanabe, Y., Kokubo, H., Miyagawa-Tomita, S., Endo, M., Igarashi, K., Aisaki, K., Kanno, J. and Saga, Y. (2006) Activation of Notch1 signaling in cardiogenic mesoderm induces abnormal heart morphogenesis in mouse. *Development* 133, 1625-1634.
- Wederell, E. D., Bilenky, M., Cullum, R., Thiessen, N., Dagpinar, M., Delaney, A., Varhol, R., Zhao, Y., Zeng, T., Bernier, B., *et al.* (2008) Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Res.* 36, 4549-4564.

- Wirrig, E. E., Snarr, B. S., Chintalapudi, M. R., O'neal, J. L., Phelps, A. L., Barth, J. L., Fresco, V. M., Kern, C. B., Mjaatvedt, C. H., Toole, B. P., *et al.* (2007) Cartilage link protein 1 (Crtl1), an extracellular matrix component playing an important role in heart development. *Dev.Biol.* 310, 291-303.
- Xu, L. G., Li, L. Y. and Shu, H. B. (2004) TRAF7 potentiates MEKK3-induced AP1 and CHOP activation and induces apoptosis. *J.Biol.Chem.* 279, 17278-17282.
- Yamasaki, S., Yagishita, N., Sasaki, T., Nakazawa, M., Kato, Y., Yamadera, T., Bae, E., Toriyama, S., Ikeda, R., Zhang, L., *et al.* (2007) Cytoplasmic destruction of p53 by the endoplasmic reticulum-resident ubiquitin ligase 'Synoviolin'. *EMBO J.* 26, 113-122.
- Yori, J. L., Johnson, E., Zhou, G., Jain, M. K. and Keri, R. A. (2010) Kruppel-like factor 4 (KLF4) inhibits epithelial-to-mesenchymal transition through regulation of E-cadherin gene expression. *J.Biol.Chem.*
- Zhang, A., Yeung, P. L., Li, C. W., Tsai, S. C., Dinh, G. K., Wu, X., Li, H. and Chen, J. D. (2004) Identification of a novel family of ankyrin repeats containing cofactors for p160 nuclear receptor coactivators. *J.Biol.Chem.* 279, 33799-33805.
- Zhao, Z. and Rivkees, S. A. (2000) Programmed cell death in the developing heart: regulation by BMP4 and FGF2. *Dev.Dyn.* 217, 388-400.
- Zhu, H., Cabrera, R. M., Wlodarczyk, B. J., Bozinov, D., Wang, D., Schwartz, R. J. and Finnell, R. H. (2007) Differentially expressed genes in embryonic cardiac tissues of mice lacking *Folr1* gene activity. *BMC Dev.Biol.* 7, 128.

CHAPTER 4

SUMMARY, PERSPECTIVES AND FUTURE DIRECTIONS

Tremendous advances in medical and surgical care of children with congenital heart defects over the past decade have made survival into adulthood a reality (Pierpont et al., 2007). One of the consequences of these clinical successes is the estimated growth in the proportion of children being born with congenital heart defects (<http://www.phac-aspc.gc.ca/publicat/cac-acc02/index-eng.php>). Therefore, understanding the process of heart formation during embryogenesis will have a great impact on preventative measures, genetic counseling and the prospects of *in vitro* engineering.

4.1. A comprehensive gene expression resource for the study of embryonic heart development

Single-gene approaches have been useful in the identification of specific genes involved in embryonic heart development; however a global understanding of the genes involved had been lacking. To address this deficit, we created a comprehensive resource for the study of gene expression changes during embryonic heart development in the mouse. Using the SAGE technique, 19 gene expression libraries were created from developing mouse heart tissue (Appendix VIII). These libraries cover different regions of the heart (e.g. atria, ventricles, AVC and OFT) at multiple time-points, starting from the linear heart tube at E8.5 up to the beginning of the valve remodeling process at E12.5.

The gene expression data presented in this thesis faithfully captures known and novel patterns of gene expression. As described in chapters 2 and 3, all known heart development genes analyzed were identified in our database, and, overall, SAGE was able to closely reflect known temporal and spatial expression patterns. Furthermore, novel gene expression patterns identified by SAGE were also validated by RT-qPCR and *in situ* hybridization, suggesting that our database was a reliable source of novel gene expression information.

An important strength of our SAGE data is that it does not rely on previous gene knowledge therefore it has the ability to identify novel genes and transcript variants. In this project we concentrated on tags mapping to annotated transcripts. However, a significant number of tag-types did not match any known entries in the gene databases. For example, out of 285,875 tag-types in the Tag-seq libraries that mapped to the genome, 202,686 could not be mapped to RefSeq, MGC, or Ensembl, a set of widely used gene databases (Table 4.1). Although some of these unannotated tags might be the result of background transcription or technical errors, 52,023 of these tag-types are expressed at a significant level (≥ 5 tags per library) suggesting that they might represent biologically relevant transcriptional events. In an experiment performed with data from the Mouse Atlas project, a collection of over 200 SAGE libraries, 77% of SAGE library singletons were able to be detected by RT-PCR suggesting that even these low abundance tags could correspond to real transcriptional events (Siddiqui et al., 2005).

Table 4.1. Mapping efficiency of tag-types in Tag-seq data

| Library | SA | MA | MG | MT | MR |
|-------------------------|-----------|-----------|-----------|-----------|-----------|
| E10.5 AVC | 3,451,899 | 208,467 | 205,242 | 70,966 | 57,905 |
| E10.5 OFT | 1,428,184 | 114,136 | 112,090 | 48,235 | 39,689 |
| E10.5 Atria | 2,744,689 | 61,319 | 60,437 | 25,388 | 20,294 |
| E10.5 Ventricles | 3,134,073 | 67,944 | 66,972 | 26,814 | 21,441 |
| All libraries | 9,767,087 | 289,869 | 285,875 | 83,189 | 67,480 |
| singletons | 8,667,158 | 102,595 | 101,703 | 13,718 | 10,927 |
| 2-4 | 784,400 | 84,795 | 83,606 | 20,928 | 16,924 |
| 5-9 | 166,711 | 40,650 | 39,962 | 13,573 | 11,000 |
| 10-19 | 72,651 | 22,624 | 22,240 | 9,482 | 7,598 |
| 20-49 | 41,246 | 16,608 | 16,277 | 8,499 | 6,760 |
| ≥ 50 | 34,921 | 22,597 | 22,087 | 16,989 | 14,271 |

SA= sans-adaptors (number of tag-types remaining after adaptor sequences were removed).

MA= mapped all (number of tag-types mapping to the genome or the gene databases).

MG=mapped to the genome (number of tag-types mapping to the mouse genome).

MT= mapped to transcript databases (number of tag-types mapping to either RefSeq, MGC, or Ensembl).

MR= mapped to RefSeq (number of tags mapping to the curated RefSeq database).

* For pooled library data (all libraries) tag-types were grouped according to raw tag counts.

Further research will be required to address the significance of the novel tag-types identified in our dataset. One possibility is that tag-types mapping to unannotated regions of the genome could represent alternative ends of transcripts, or genes currently undetected by gene prediction algorithms (Dike et al., 2004). One approach to study whether these unannotated tag-types represent alternative transcripts or novel genes is by use of the new technique of RNA sequencing (RNA-seq) (Morin et al., 2008a; Mortazavi et al., 2008; Nagalakshmi et al., 2008). In this technique a population of RNA molecules is converted to a library of cDNA fragments and deeply sequenced. The resulting sequencing reads are then assembled to produce a genome-scale transcription map representing the structure of transcripts and their level of expression. RNA-seq does not depend on the presence of particular restriction sites within the cDNA, therefore it can detect any transcripts lacking the recognition sequence for *NlaIII* (3'CATG5') used as anchoring enzyme in the creation of our SAGE libraries. RNA-seq can also reveal the precise exon structure of genes including splice junctions, transcriptional start sites, and 3' ends (Morin et al., 2008a; Mortazavi et al., 2008; Nagalakshmi et al., 2008). Currently, the RNA-seq technique is gaining ground as an alternative gene expression profiling method. However, RNA-seq has limitations as each of the fragmentation methods (such as RNA hydrolysis or cDNA sonication) required to make large RNA molecules compatible with most deep-sequencing technologies creates a different type of bias in the resulting data. For example, RNA fragmentation results in data depleted for transcript ends, while cDNA fragmentation is usually biased towards sequences from the 3' end of transcripts. Furthermore, in RNA-seq longer transcripts generate more reads than shorter transcripts of similar expression levels, resulting in more statistical power to detect differential expression for long transcripts compared to short ones (Oshlack and Wakefield,

2009). At current levels of sequencing, most low expressed genes will not be detected. Therefore, in my view, a full representation of the transcriptome will require a combination of the quantification power of SAGE with the gene structure analysis of RNA-seq.

Unannotated transcripts could also represent primary transcript precursors for short non-coding RNAs (ncRNAs) (Kapranov et al., 2007). One family of ncRNAs, the microRNAs, is emerging as a group of critical regulators of gene expression during heart development and homeostasis (Liu and Olson, 2010). These small (20-30 nucleotide) RNAs are transcribed by RNA polymerase II and, after a series of processing steps, are incorporated into a multiprotein RNA-induced silencing complex called RISC (van Bakel and Hughes, 2009). RISC typically recognizes the 3' untranslated regions of specific target mRNAs and induces post-transcriptional gene silencing. Because our SAGE library construction protocol used oligo-dT primers for cDNA synthesis, processed microRNAs lacking a poly-A tail were not captured in our gene expression libraries. Creation of microRNA libraries, where small RNAs are directly sequenced after adaptor ligation, should help describe the full complement of RNA species in the tissue or time-point of interest (Morin et al., 2008b).

Of the tag-types that mapped to gene databases, as many as 20% mapped in the antisense direction (i.e. they mapped to the strand opposite the coding sequence). Although some of these antisense tags could be due to the presence of unannotated genes on the opposite strand as well as technical artifacts of the SAGE library construction process, an increasing number of reports has illustrated the existence of naturally occurring antisense RNAs (Chen et al., 2004; He et al., 2008; Katayama et al., 2005; Quere et al., 2004). Experimental evidence suggests that antisense transcripts may govern the expression of their sense counterparts by affecting transcription, maturation, transport, stability and translation of mRNA (Vanhee-Brossollet and Vaquero, 1998).

Recently, a mechanism of gene regulation has been proposed involving the cleavage and processing of co-expressed sense and antisense transcripts into single-stranded short RNAs that could then act in a similar fashion to microRNAs (Tam et al., 2008; Watanabe et al., 2008). Further analysis of the antisense and sense pairs present in our dataset could uncover novel regulatory mechanisms in heart development.

Since the SAGE libraries created as part of this project were derived from a heterogeneous population of cells, it is not possible to determine from our heart libraries alone which genes are expressed in the endocardial, mesenchymal or myocardial compartments. However, one critical advantage of SAGE data is that it is digital and absolute allowing the easy comparison of expression across different tissues and stages. In chapter 2, we used six epithelial-mesenchymal SAGE library pairs from the Mouse Atlas resource to identify genes whose expression is likely enriched in the endocardial or mesenchymal populations of the AVC (see Figure 2.3). However, this approach could not identify the myocardial cell population, or genes with epithelial expression in one context and mesenchymal expression in another.

One strategy that can be used for the isolation of specific cell populations for SAGE library construction is fluorescence-activated cell sorting (FACS) (Hoffman et al., 2008; Khattra et al., 2007). In this technique, different cell populations are separated according to their surface markers or based on the expression of a marker protein (such as GFP) driven by a cell specific promoter. One disadvantage of this method is the potential change in gene expression caused by the disruption of cell contacts during the cell sorting protocol or the time elapsed between dissection and sample collection. Another technique that can be used to isolate specific subpopulations of cells is laser-capture microdissection (LCM). In this technique, specific cells of interest are isolated from microscopic regions of tissue that has been sectioned. This strategy

has been successfully applied to the analysis of specific brain regions (Siddiqui et al., 2005), and could be used even in the absence of well defined molecular markers, such as in the comparison of the major and minor cushions of the AVC. At the current level of technology, however, RNA amplification needs to be performed on the limited amount of material obtained by FACS and LCM, introducing limits to the comparison of libraries made from amplified and non-amplified material.

In summary, the resource presented in this thesis constitutes a significant advance in the characterization of the transcriptome and of its regulation during embryonic heart development. Proteome analysis would be an important complement to our data by determining properties of biological systems that are not apparent by DNA or mRNA sequence analysis alone, such as the quantity of protein expression, the subcellular location of proteins, or their state of modification (Macri and Rapundalo, 2001).

4.2. Study of temporal and spatial restricted gene expression during valve development

Defects in valve and septa formation constitute a large majority of congenital heart defects; therefore we sought to use our gene expression resource to identify novel valve development genes. Our underlying hypothesis was that genes important for heart valve development would be differentially expressed in the AVC and OFT, and have dynamic temporal expression patterns.

4.2.1. Identification of temporal expression patterns during AVC development

As described in chapter 2, we first investigated the temporal expression patterns critical for early valve development by analyzing four SAGE libraries created from E9.5, E10.5, E11.5 and E12.5 mouse AVCs. Using cluster analysis, we uncovered 14 distinct temporal expression patterns falling into 5 main categories: an increase over time, a decrease over time, a peak at E10.5, a peak at E11.5 and a peak at both E10.5 and E11.5.

The temporal expression patterns identified in the developing AVC were highly correlated with function (see Table 2.2). For example, genes that decreased from E9.5 to E12.5 were more likely endothelial expressed genes, while genes that increased over this time period were more likely mesenchyme markers, indicating that we could identify gene expression changes driving EMT. Similarly, clusters with peak expression at E11.5 were enriched for genes involved in cell proliferation, whereas clusters with peak expression at E12.5 were enriched for genes involved in apoptosis. Therefore the co-regulation uncovered in our study can be used to help predict the functions of novel genes, although gain and loss of function experiments would be required to establish their specific roles.

Interestingly, a high proportion of mesenchyme enriched genes were found to peak at E10.5 and E11.5, suggesting that mesenchyme differentiation continues past the E10.5 time-point. Previous studies using collagen explant assays have focused on E9.5-10.5 as the critical time in EMT in the AVC (Camenisch et al., 2002). Developed in the early 1980s (Bernanke and Markwald, 1982; Runyan and Markwald, 1983), the collagen AVC explant assay has been instrumental in the analysis of the EMT process. However, other important processes of valve development, such as mesenchyme cell maturation and condensation, as well as leaflet

elongation and remodeling by apoptosis, have not been successfully studied in these assays. Recently, a new three dimensional collagen tube culturing system has been described that recapitulates morphological and molecular changes occurring during later stages of valve development (Goodwin et al., 2005; Norris et al., 2009). Over-expression or knock-down of candidate genes in these three dimensional collagen assays may help determine the factors controlling AVC remodeling following EMT.

Two pathways with potential roles in regulating AVC maturation are the Notch and TGF β pathways. These pathways have emerged as critical regulators of AVC specification and EMT, however their precise role at later stages in AVC development is not currently known. Interestingly, different ligands of the TGF β and Notch pathways showed distinct patterns during our temporal expression analysis, suggesting that there are multiple phases of TGF β and Notch pathway activity over the course of AVC development. For example, expression of *Tgfb2* was found to be high at E9.5-10.5 and then decrease over time, while *Tgfb3* expression peaked at E12.5. Similarly, expression of the Notch ligand *Jag1* peaked at E10.5 while the Notch ligand *Dll4* was highest at E12.5. To study the possible roles of these pathways at different time-points we overlaid our temporal expression patterns with putative targets of TGF β and Notch pathways identified from the literature (see Figure 2.4). We found a significant difference between the types of TGF β and Notch target genes peaking at different time-points. For example, TGF β responsive genes peaking at E9.5-10.5 included many genes involved in transcription factor activity or chromatin remodeling, while those peaking at E12.5 included many ECM proteins and cell-adhesion molecules. Conditional mutants that inactivate signalling pathway activity at different time-points of valve development could help uncover the role of these pathways following EMT.

4.2.2. Identification of endocardial cushion enriched genes

As described in chapter 2, our temporal expression analysis revealed a significant number of transcription factors and signalling pathway members with peak expression at E10.5, suggesting this was an important time-point for AVC development. In chapter 3 we focused on endocardial cushion enriched gene expression at E10.5 by creating four gene expression libraries from atria, ventricles, AVC and OFT. For this study we used the recently developed Tag-seq technique, which allowed for a significant increase in library depth over traditional LongSAGE libraries. To determine AVC fold enrichment we calculated the ratio of AVC gene expression over the atria and ventricles. Similarly, we determined OFT enrichment by calculating the ratio of OFT gene expression over the expression in atria and ventricles.

Both the AVC and OFT undergo a process of endocardial cushion formation and EMT; however differences between these tissues have been reported. In collagen explant assays, the OFT initiates EMT up to one day later than the AVC (Camenisch et al., 2002). In addition, cardiac neural crest cells begin to invade the OFT starting at E10.5 and contribute substantially to the future OFT valves (de Lange et al., 2004). Interestingly, the question of gene expression similarities and differences between the early AVC and OFT cushions had not been systematically studied. During our analysis, comparison of AVC enriched genes against OFT enriched genes showed a significant gene expression overlap indicating that these tissues share many regulatory mechanisms. Our list of shared AVC and OFT enriched genes included highly expressed ECM molecules, signalling molecules, modulators of proliferation and apoptosis, and transcription factors, most notably the bHLH transcription factors *Twist1* and *Twist2*.

Significantly, they also contained a majority of genes not previously described in the context of heart development including many Rikens cDNAs and other genes with little functional information. In addition, some of the genes identified have known roles during earlier stages of cardiac development, suggesting that they play multiple roles during heart development. In the case of some of these genes, such as *Mef2c*, lethality before endocardial cushion formation has precluded the analysis of their role in valve development (Lin et al., 1998; Vincentz et al., 2008a). The use of conditional mutations could reveal their specific roles during valve formation.

Our analysis also detected genes that were differentially expressed in either the AVC or OFT. For example, *Rgs5* and *Gabra4* were differentially expressed in the OFT but not the AVC. In contrast, *Galanin* and *Adamts19* expression were found enriched in the AVC but not the OFT. These gene expression differences could reflect differences in timing of OFT and AVC cushion development, invasion of neural crest into the OFT, or different cell populations arising in the AVC and OFT. For example, the AVC will give rise to the AV node and rings involved in the transmission of electrical impulses (Christoffels et al., 2010). *In situ* hybridization or immunohistochemistry analyses should help identify the source of these overall gene expression differences.

Finally, a significant contribution of our work to the field of heart development has been the identification of spatially and temporally co-expressed genes. Further analysis of the regulatory sequences shared by these co-expressed genes can help define cis-elements necessary for driving specific gene expression during heart development (Habets et al., 2003). These regulatory modules could then be exploited in transgenic experimental models to drive expression of genes within specific spatio-temporal domains (Yang et al., 2009).

4.3. Role of TWIST1 in heart development

As described in chapter 3, the most highly expressed DNA-binding transcription factor in the AVC at E10.5 was found to be *Twist1*. Expression of *Twist1* in the heart is specific to the mesenchyme cells of the AVC and OFT. Temporally, *Twist1* expression peaks at E10.5 before decreasing at E12.5. This spatial and temporal expression pattern suggested that TWIST1 could play a role in controlling endocardial cushion cell phenotype following EMT. However, analysis of the hearts of *Twist1* null mice by us (unpublished results, see below) and others (Vincentz et al., 2008b) had not found an obvious AVC phenotype, even though research in chick AVC endocardial cushion cells had suggested a role of TWIST1 in proliferation, migration and maturation (Shelton and Yutzey, 2008).

In our study, comparison of the Tag-seq *Twist1* mutant library against the wild-type AVC library showed dramatic gene expression changes, representing the first time an AVC phenotype is identified in the *Twist1* knock-out mouse. Analysis of up- and down-regulated gene expression suggested that TWIST1 plays a role in determining endocardial cushion specific gene expression, specially the acquisition of a migratory phenotype, while blocking apoptosis and maturation of mesenchyme cells.

4.3.1. Direct regulation of transcription by TWIST1

TWIST1 can regulate transcription in a variety of ways, either through direct binding to DNA or through interaction with other proteins (Hamamori et al., 1997; Laursen et al., 2007; Spicer et al., 1996). In our study, we were able to identify TWIST1 direct binding to the

promoters of genes differentially expressed in the *Twist1* null AVC by performing chromatin-immunoprecipitation followed by quantitative PCR (ChIP-qPCR). At the time of this research the lack of a good ChIP-quality commercial antibody for TWIST1 made the study of endogenous protein binding in heart tissue impractical. We therefore performed ChIP-qPCR on a myc-tagged version of *Twist1* over-expressed in two mouse cell lines (one epithelial and one endothelial). Out of 40 gene promoters analyzed, 10 were enriched over 20-fold after TWIST1 ChIP-qPCR suggesting direct TWIST1 binding. Significantly, 6 of the top 8 genes bound by TWIST1 were found to be transcription factors (*Klf4*, *Tb11x*, *Hes6*, *Id2*, *Lef1* and *Gata3*) suggesting that TWIST1 could act as a master regulator of endocardial cushion gene expression. Interestingly, TWIST1 binding was observed at the promoters of both up- and down-regulated genes in the *Twist1* null AVC, suggesting that it can act to either activate or repress the transcription of specific promoters. Further analysis will be required to understand the mode of action of TWIST1 at these promoters, including the role of other transcription factors and chromatin modifiers.

Lack of ChIP-qPCR amplification at some of the promoters tested could reflect the absence of required co-factors in the cell lines used, therefore analysis of TWIST1 binding in mouse endocardial cushion cells *in vivo* remains important. Alternatively, TWIST1 could act indirectly through other transcription factors, or form part of a complex without binding DNA directly. Finally, absence of binding of TWIST1 at the promoter regions could be the result of regulatory sequences outside the regions analyzed, since these sequences can be located at significant distances from transcriptional start sites as well as in introns.

A new technique called ChIP-seq has been developed to explore binding across the genome (Robertson et al., 2007; Wederell et al., 2008). In ChIP-seq, chromatin-

immunoprecipitation is followed by massively parallel sequencing of the resulting fragments. These fragments are then overlaid over the genome to create a map of transcription factor binding. ChIP-seq performed with anti-TWIST1 antibodies could uncover binding elements outside of the promoter regions studied in this thesis. Moreover, analysis of TWIST1 binding at a genome level might uncover sequence differences between activated and repressed genes, such as flanking binding sequences for other transcription factors.

In the future, creation of ChIP-seq libraries for other AVC and OFT enriched transcription factors identified in this project will provide information on the gene regulatory networks important for valve development. Moreover, the use of the ChIP-seq technique need not be restricted to transcription factors but can be expanded to histone modifications or other epigenetic marks (Robertson et al., 2008). These could help distinguish between active and repressed promoters to identify the role of TWIST1 in a particular gene context.

4.3.2. Functional redundancy and compensatory mechanisms for *Twist1* deletion

Our study of AVC development in the *Twist1* null embryo revealed dramatic gene expression changes. Interestingly, these changes were not accompanied by any obvious morphological difference in the AVC of *Twist1* null mice at E10.5 (Figure 4.1). To explore the possibility of subtle effects of the *Twist1* null mutation, we performed collagen explant assays (Figure 4.2), but we observed no obvious EMT or cell migration phenotype in *Twist1* null AVCs in these assays. Unfortunately, the embryonic lethality of the *Twist1* null mice at E11 has so far precluded the analysis of defects at later stages of development. Therefore, the *Twist1* mutation could still result in defects of valve maturation not detected at E10.5. The use of a recently

described conditional null *Twist1* allele would be important to dissect these later phenotypes (Chen et al., 2007). Alternatively, there could be compensatory mechanisms for the absence of TWIST1 activity in the mouse.

Twist2 is a possible gene that could compensate for the *Twist1* null mutation. TWIST1 and TWIST2 share extensive sequence similarity, and nearly identical bHLH domains (see Figure 1.7). Furthermore, TWIST1 and TWIST2 have been shown to play similar roles during cancer progression and development. During cancer progression, TWIST1 and TWIST2 both can override oncogene-induced premature senescence and cooperate with RAS to transform mouse embryonic fibroblasts (Ansieau et al., 2008; Maestro et al., 1999). There is also evidence for functional redundancy of TWIST1 and TWIST2 during development. For example, over-expression of either *Twist1* or *Twist2* in osteoblastic cells causes similar changes in their morphology, gene expression, and biochemical response to cytokines (Lee et al., 2000). Significantly, the *Twist2* null phenotype of enhanced pro-inflammatory cytokine gene expression and increased apoptosis in multiple tissues is also observed in *Twist1* and *Twist2* compound heterozygotes suggesting redundancy and dosage dependence in certain contexts (Bialek et al., 2004; Susic et al., 2003).

Figure 4.1. *Twist1* null embryos show normal EMT

Wild-type and *Twist1* null embryos were dissected at E10.5. Hematoxylin and Eosin staining was performed after paraffin sectioning. *Twist1* null embryos show defects in neural tube closure and limb development. AVC endocardial cushions of *Twist1* null hearts are populated by mesenchymal cells.

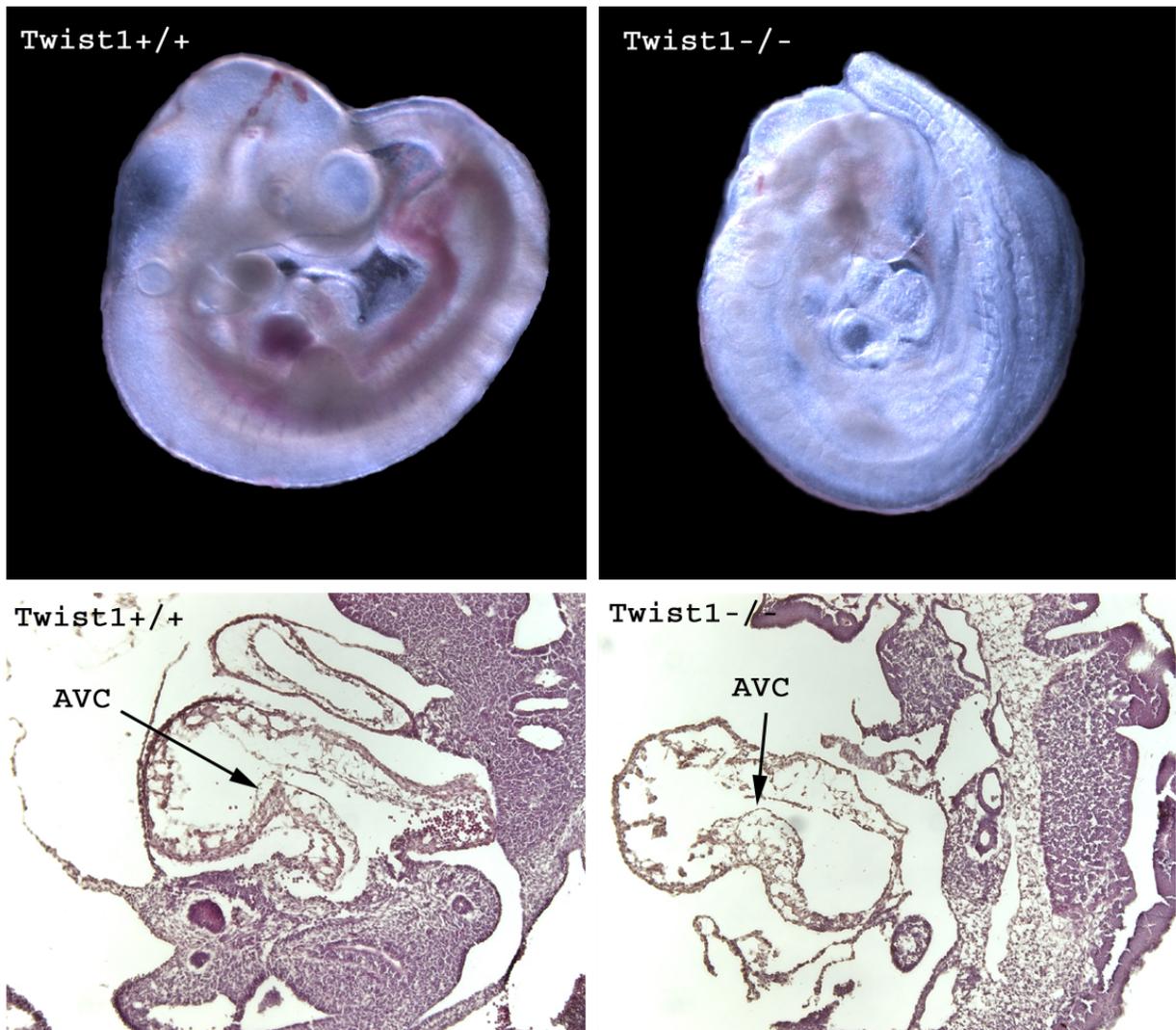
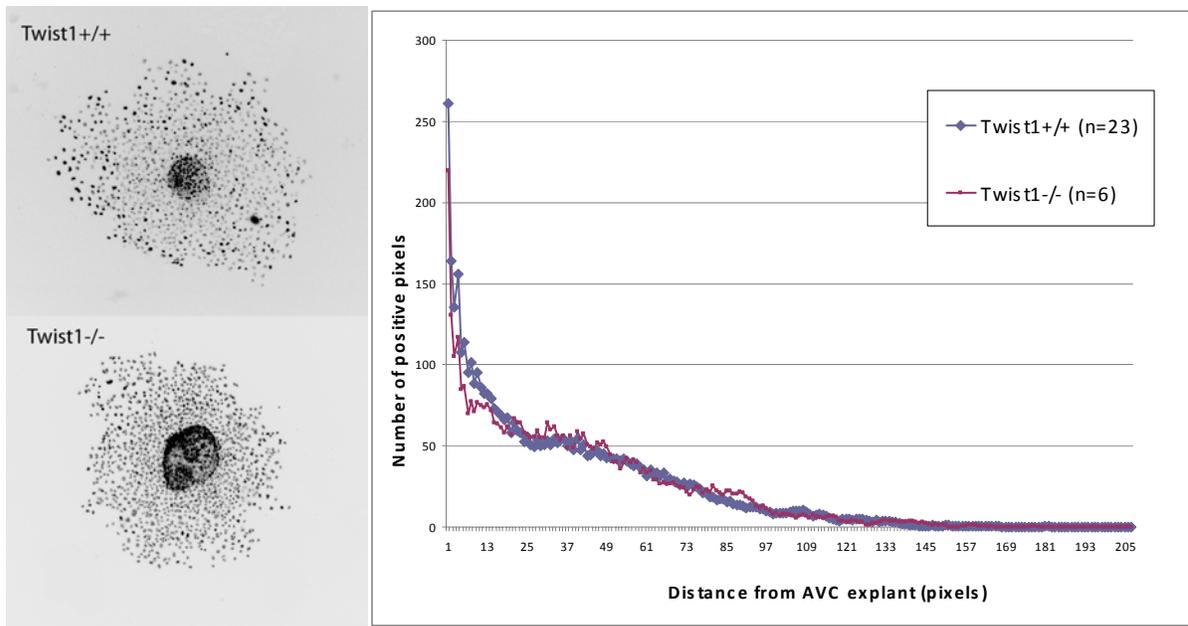


Figure 4.2. AVC explant assay shows no defect in cell migration

Collagen gel assays were performed as described (Niessen et al., 2008). Briefly, AVCs were dissected at E9.5 and placed in a type I collagen matrix for 72hrs. After fixation, explants were stained with DAPI and imaged under fluorescent microscopy. Number of positive fluorescent pixels was used as a measure of cell number at a particular distance from the original AVC explant.



In the mouse embryonic heart, *Twist2* and *Twist1* have overlapping spatial gene expression domains, as determined from our Tag-seq data (Table 4.2), and *in situ* hybridization analysis (Figure 4.3), although *Twist2* expression was much lower than *Twist1*. To test whether TWIST2 could potentially regulate the same genes as TWIST1, we performed ChIP-qPCR on cells over-expressing a flag-tagged version of *Twist2*. Preliminary experiments show that TWIST2 can bind to the same promoter regions as TWIST1 (Figure 4.4), supporting the possibility that TWIST2 activity could compensate for TWIST1 in the developing heart. These findings would need to be expanded to other promoters to see if other TWIST1 targets can be bound by TWIST2.

Interestingly, RT-qPCR analysis showed differences in the temporal expression patterns for *Twist1* and *Twist2* in the developing AVC. While *Twist1* expression peaks at E10.5 and E11.5, *Twist2* expression peaks at E9.5 and decreases at E10.5 (Figure 4.5). Furthermore, *Twist2* expression was down-regulated in the *Twist1* knock-out AVC by about 30% (Table 4.2). Therefore it is not clear if TWIST2 activity would be sufficient to compensate for TWIST1 activity in the mouse AVC. Unfortunately, the creation of double knock-outs for *Twist1* and *Twist2* is complicated by the early lethality of the compound heterozygote, therefore knock-down of *Twist2* in *Twist1* null AVC explants, or conditional deletion would be required to address this question.

Table 4.2. Expression of Twist family members during valve development

| Tag-seq data | | | | |
|--|---------------|---------------|------------------|----------------|
| | <i>Twist1</i> | <i>Twist2</i> | <i>Scleraxis</i> | <i>Paraxis</i> |
| E10.5 AVC | 918 | 14.8 | 3.8 | 0 |
| E10.5 OFT | 299.1 | 35.9 | 2.9 | 1.3 |
| E10.5 Atria | 63.4 | 1.3 | 1.5 | 0 |
| E10.5 Ventricles | 55.9 | 1.7 | 2.4 | 0 |
| E10.5 <i>Twist1</i> ^{-/-} AVC | 4.6 | 9.8 | 2.9 | 0 |

| LongSAGE data | | | | |
|----------------------|---------------|---------------|------------------|----------------|
| | <i>Twist1</i> | <i>Twist2</i> | <i>Scleraxis</i> | <i>Paraxis</i> |
| E9.5 AVC | 36 | 0 | 4 | 0 |
| E10.5 AVC | 213 | 0 | 33 | 0 |
| E11.5 AVC | 179 | 15 | 19 | 0 |
| E12.5 AVC | 87 | 0 | 25 | 8 |

* Expression normalized per million.

Figure 4.3. *Twist1* and *Twist2* are expressed in mesenchyme cells of AVC and OFT

In situ hybridization was performed on E11.5 mouse hearts followed by cryosectioning. A= atria, V= ventricles, AVC= atrio-ventricular canal, OFT= outflow tract.

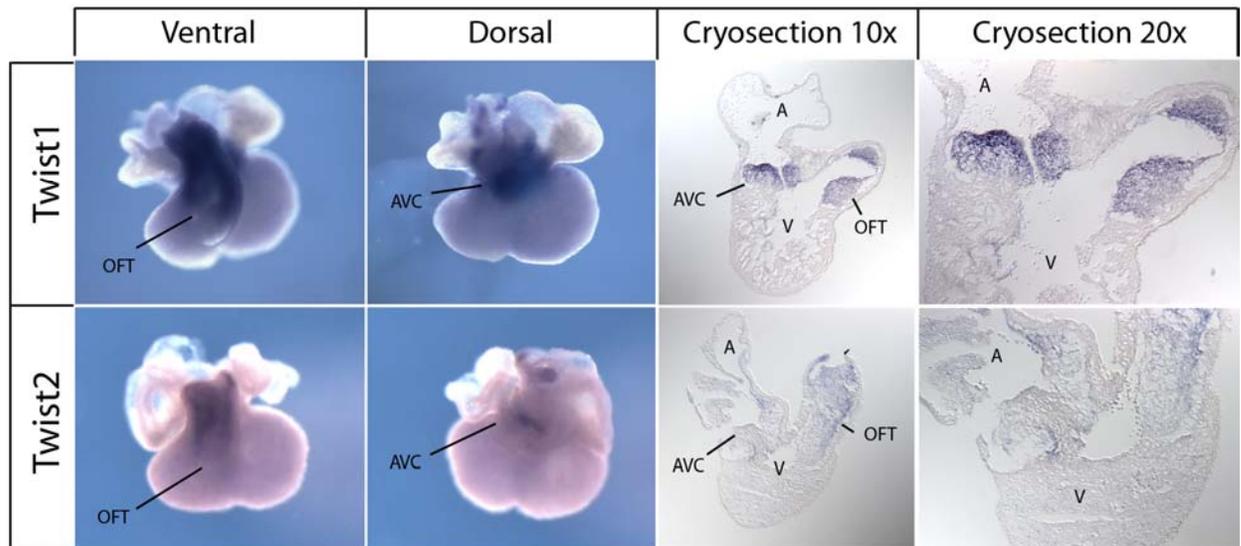


Figure 4.4. TWIST1 and TWIST2 bind similar promoter regions

Chromatin-immunoprecipitation followed by PCR was performed as described in chapter 3. Either a myc-tagged version of *Twist1*, or a flag-tagged version of *Twist2* were over-expressed in mIMCD cells. Anti-myc or anti-flag antibodies were used for ChIP. Results were normalized to Troponin C (Tnnc).

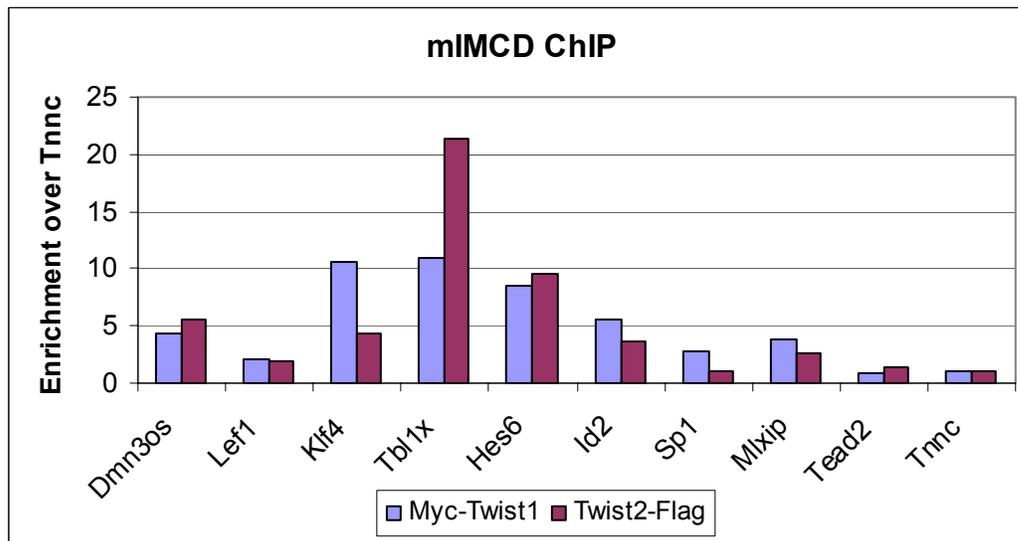
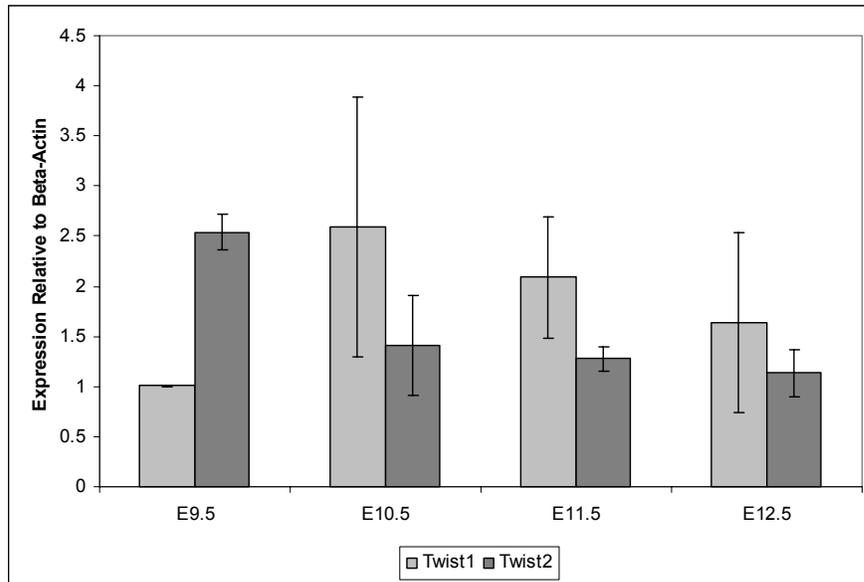


Figure 4.5. *Twist1* and *Twist2* show distinct temporal expression patterns in the AVC

RT-qPCR was performed on AVC samples from E9.5, E10.5, E11.5 and E12.5 mouse embryonic hearts. Expression levels were normalized to lowest value. Error bars represent \pm standard deviation (n=3).



Snai2 (also known as *Slug*) is another gene that could compensate for TWIST1 activity. *Snai2* is a member of the Snail family of zinc finger transcription factors which drive EMT during development and tumour progression (Nieto, 2002). *Snai2* is enriched in the endocardial and mesenchymal cells of the developing AVC and OFT. Furthermore, analysis of *Snai2* deficient mice revealed a role in endocardial cushion EMT (Niessen et al., 2008). *Snai2* has also been shown to genetically interact with *Twist1*, as *Snai2/Twist1* compound heterozygous mice show increased penetrance of the craniosynostosis and polydactyl phenotype caused by the *Twist1* mutation (Oram and Gridley, 2005).

To study the possible role of SNAI2 in compensating for TWIST1 activity in the AVC we crossed *Twist1* mutant and *Snai2* null mice (Table 4.3). *Twist1/Snai2* compound heterozygotes were viable and fertile, and were obtained at the expected Mendelian ratios (Table 4.3). In preliminary compound heterozygote crosses, we obtained one *Twist1*^{+/-};*Snai2*^{-/-} embryo at E9.5 which did not appear grossly different from wild-type littermates. Similarly, *Twist1*^{-/-};*Snai2*^{+/-} embryos dissected at E10.5 were grossly indistinguishable from *Twist1* homozygote mutants. These embryos contained normal looking AVC cushions populated by presumptive mesenchyme cells, although further characterization might reveal subtle phenotypes. Interestingly, no *Twist1/Snai2* double mutants have been observed, suggesting that *Snai2* deficiency interacts with the *Twist1* mutation to produce embryonic lethality prior to E9.5. Because this lethality occurs before cushion formation, we could not examine the effect of the double mutation on valve formation. Therefore conditional deletion will be required to further investigate this interaction.

Table 4.3. *Twist1* mutant and *Snai2* mutant crosses

| A. Progeny of <i>Twist1</i>^{+/-} x <i>Snai2</i>^{+/-} cross at weaning (n=5) | | | | | |
|--|-------------------|----------------------|-------------------|------------------------|-------------------|
| Genotype | % expected | Observed | % observed | | |
| Wild type | 25 | 8 | 25.8 | | |
| <i>Twist1</i> ^{+/-} | 25 | 7 | 22.6 | | |
| <i>Snai2</i> ^{+/-} | 25 | 8 | 25.8 | | |
| <i>Twist1</i> ^{+/-} <i>Snai2</i> ^{+/-} | 25 | 8 | 25.8 | | |
| B. Progeny of <i>Twist1</i>^{+/-}/<i>Snai2</i>^{+/-} x <i>Twist1</i>^{+/-}/<i>Snai2</i>^{+/-} cross | | | | | |
| Genotype | % expected | Weaning (n=4) | | E9.5-10.5 (n=3) | |
| | | Observed | % observed | Observed | % observed |
| Wild type | 6.3 | 3 | 18.8 | 0 | 0 |
| <i>Twist1</i> ^{+/-} | 12.5 | 1 | 6.3 | 1 | 3.5 |
| <i>Snai2</i> ^{+/-} | 12.5 | 5 | 31.3 | 4 | 14.3 |
| <i>Twist1</i> ^{+/-} <i>Snai2</i> ^{+/-} | 25 | 6 | 37.5 | 13 | 46.4 |
| <i>Twist1</i> ^{-/-} <i>Snai2</i> ^{+/-} | 12.5 | 0 | 0 | 8 | 28.6 |
| <i>Twist1</i> ^{+/-} <i>Snai2</i> ^{-/-} | 12.5 | 0 | 0 | 1 | 3.6 |
| <i>Twist1</i> ^{-/-} | 6.3 | 0 | 0 | 1 | 3.6 |
| <i>Snai2</i> ^{-/-} | 6.3 | 1 | 6.3 | 0 | 0 |
| <i>Twist1</i> ^{-/-} <i>Snai2</i> ^{-/-} | 6.3 | 0 | 0 | 0 | 0 |

In summary, our research has shown a role of TWIST1 in AVC development. While the gene expression changes observed have not been correlated with morphological changes, it is likely that more research will uncover maturation defects not observed at early stages of valve development. Moreover, the identification of *Twist2* expression in the heart suggests a possible compensatory mechanism. Further research will clarify the role of TWIST1 and TWIST2 in valve development.

4.4. Research impact and future work

This thesis describes the creation of a comprehensive resource for the study of gene expression changes during embryonic heart development, and its use for the analysis of early valve formation. This resource could also be used to address other important questions during heart development, such as the determination of heart specific gene expression, the specification of the atria and ventricles, and the trabeculation of the chamber myocardium.

Due to the digital nature of the SAGE data, the resource presented in this thesis can be easily built upon by the addition of more gene expression libraries. Following EMT the valves begin to be remodeled by proliferation, apoptosis, condensation and ECM stratification (Armstrong and Bischoff, 2004). Extending the gene expression resource to later embryonic stages would allow these critical processes to be further investigated. Similarly, while in this thesis we concentrated on *Twist1* by creating a gene expression library of mutant tissue; additional libraries should be made from different knock-out embryos for comparison.

Another opportunity lies in the comparison of diseased and normal adult valve tissue. One of the most common pathological changes occurring in adult aortic valves is their

calcification, which leads to stenosis or regurgitation and can ultimately require surgical valve replacement (Schoen, 2008). Interestingly, a congenital malformation of the aortic valve (such as bicuspid aortic valve) is a major risk factor for the development of calcification (Otto, 2002). It is significant that in chapters 2 and 3, markers of bone development were found to be expressed during valve formation, suggesting links between embryonic development and pathological changes occurring in the adult.

Few of the genetic causes of human congenital heart defects have been identified so far. In humans, individuals sharing the same mutations can show remarkable variations in penetrance and expressivity of congenital heart phenotypes. Similarly, gene knock-out mice often show phenotypic variation within a litter. This is believed to be due to the role of modifier loci to either buffer perturbations on cardiac development or direct manifestation of a defect. Recent research in *Nkx2-5* knock-out mice placed in different mouse strain backgrounds has shown that different modifier loci impact the effect of mutations in cardiac development (Winston et al., 2010). Our data provides a novel source of candidate genes that can be analyzed in genetic association studies, as single determinants of heart phenotype or as modifier loci.

Finally, our data could impact the area of tissue engineering. An estimated 275,000 patients around the world undergo heart valve replacement each year. This is most often done by introducing bioprostheses, usually of porcine origin, which have limited durability (Schoen, 2008). Therefore, creation of bioengineered valves with live cells capable of repair could provide lifelong solutions. The data provided in this thesis can help make this a reality by increasing understanding of the pathways that can be exploited to drive cells into specific lineages *in vitro*.

4.5. References

- Ansieau, S., Bastid, J., Doreau, A., Morel, A. P., Bouchet, B. P., Thomas, C., Fauvet, F., Puisieux, I., Doglioni, C., Piccinin, S., *et al.* (2008) Induction of EMT by twist proteins as a collateral effect of tumor-promoting inactivation of premature senescence. *Cancer.Cell.* 14, 79-89.
- Armstrong, E. J. and Bischoff, J. (2004) Heart valve development: endothelial cell signaling and differentiation. *Circ.Res.* 95, 459-470.
- Bernanke, D. H. and Markwald, R. R. (1982) Migratory behavior of cardiac cushion tissue cells in a collagen-lattice culture system. *Dev.Biol.* 91, 235-245.
- Bialek, P., Kern, B., Yang, X., Schrock, M., Sobic, D., Hong, N., Wu, H., Yu, K., Ornitz, D. M., Olson, E. N., Justice, M. J. and Karsenty, G. (2004) A twist code determines the onset of osteoblast differentiation. *Dev.Cell.* 6, 423-435.
- Camenisch, T. D., Molin, D. G., Person, A., Runyan, R. B., Gittenberger-de Groot, A. C., McDonald, J. A. and Klewer, S. E. (2002) Temporal and distinct TGFbeta ligand requirements during mouse and avian endocardial cushion morphogenesis. *Dev.Biol.* 248, 170-181.
- Chen, J., Sun, M., Kent, W. J., Huang, X., Xie, H., Wang, W., Zhou, G., Shi, R. Z. and Rowley, J. D. (2004) Over 20% of human transcripts might form sense-antisense pairs. *Nucleic Acids Res.* 32, 4812-4820.
- Chen, Y. T., Akinwunmi, P. O., Deng, J. M., Tam, O. H. and Behringer, R. R. (2007) Generation of a Twist1 conditional null allele in the mouse. *Genesis* 45, 588-592.
- Christoffels, V. M., Smits, G. J., Kispert, A. and Moorman, A. F. (2010) Development of the pacemaker tissues of the heart. *Circ.Res.* 106, 240-254.
- de Lange, F. J., Moorman, A. F., Anderson, R. H., Manner, J., Soufan, A. T., de Gier-de Vries, C., Schneider, M. D., Webb, S., van den Hoff, M. J. and Christoffels, V. M. (2004) Lineage and morphogenetic analysis of the cardiac valves. *Circ.Res.* 95, 645-654.
- Dike, S., Balija, V. S., Nascimento, L. U., Xuan, Z., Ou, J., Zutavern, T., Palmer, L. E., Hannon, G., Zhang, M. Q. and McCombie, W. R. (2004) The mouse genome: experimental examination of gene predictions and transcriptional start sites. *Genome Res.* 14, 2424-2429.
- Goodwin, R. L., Nesbitt, T., Price, R. L., Wells, J. C., Yost, M. J. and Potts, J. D. (2005) Three-dimensional model system of valvulogenesis. *Dev.Dyn.* 233, 122-129.
- Habets, P. E., Moorman, A. F. and Christoffels, V. M. (2003) Regulatory modules in the developing heart. *Cardiovasc.Res.* 58, 246-263.
- Hamamori, Y., Wu, H. Y., Sartorelli, V. and Kedes, L. (1997) The basic domain of myogenic basic helix-loop-helix (bHLH) proteins is the novel target for direct inhibition by another bHLH protein, Twist. *Mol.Cell.Biol.* 17, 6563-6573.
- He, Y., Vogelstein, B., Velculescu, V. E., Papadopoulos, N. and Kinzler, K. W. (2008) The antisense transcriptomes of human cells. *Science* 322, 1855-1857.
- Hoffman, B. G., Zavaglia, B., Witzsche, J., Ruiz de Algora, T., Beach, M., Hoodless, P. A., Jones, S. J., Marra, M. A. and Helgason, C. D. (2008) Identification of transcripts with enriched expression in the developing and adult pancreas. *Genome Biol.* 9, R99.

- Kapranov, P., Cheng, J., Dike, S., Nix, D. A., Duttagupta, R., Willingham, A. T., Stadler, P. F., Hertel, J., Hackermuller, J., Hofacker, I. L., *et al.* (2007) RNA maps reveal new RNA classes and a possible function for pervasive transcription. *Science* 316, 1484-1488.
- Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., Nishida, H., Yap, C. C., Suzuki, M., Kawai, J., *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science* 309, 1564-1566.
- Khattra, J., Delaney, A. D., Zhao, Y., Siddiqui, A., Asano, J., McDonald, H., Pandoh, P., Dhalla, N., Prabhu, A. L., Ma, K., *et al.* (2007) Large-scale production of SAGE libraries from microdissected tissues, flow-sorted cells, and cell lines. *Genome Res.* 17, 108-116.
- Laursen, K. B., Mielke, E., Iannaccone, P. and Fuchtbauer, E. M. (2007) Mechanism of transcriptional activation by the proto-oncogene Twist1. *J.Biol.Chem.* 282, 34623-34633.
- Lee, M. S., Lowe, G., Flanagan, S., Kuchler, K. and Glackin, C. A. (2000) Human Dermo-1 has attributes similar to twist in early bone development. *Bone* 27, 591-602.
- Lin, Q., Lu, J., Yanagisawa, H., Webb, R., Lyons, G. E., Richardson, J. A. and Olson, E. N. (1998) Requirement of the MADS-box transcription factor MEF2C for vascular development. *Development* 125, 4565-4574.
- Liu, N. and Olson, E. N. (2010) MicroRNA regulatory networks in cardiovascular development. *Dev.Cell.* 18, 510-525.
- Macri, J. and Rapundalo, S. T. (2001) Application of proteomics to the study of cardiovascular biology. *Trends Cardiovasc.Med.* 11, 66-75.
- Maestro, R., Dei Tos, A. P., Hamamori, Y., Krasnokutsky, S., Sartorelli, V., Kedes, L., Doglioni, C., Beach, D. H. and Hannon, G. J. (1999) Twist is a potential oncogene that inhibits apoptosis. *Genes Dev.* 13, 2207-2217.
- Morin, R., Bainbridge, M., Fejes, A., Hirst, M., Krzywinski, M., Pugh, T., McDonald, H., Varhol, R., Jones, S. and Marra, M. (2008a) Profiling the HeLa S3 transcriptome using randomly primed cDNA and massively parallel short-read sequencing. *BioTechniques* 45, 81-94.
- Morin, R. D., O'Connor, M. D., Griffith, M., Kuchenbauer, F., Delaney, A., Prabhu, A. L., Zhao, Y., McDonald, H., Zeng, T., Hirst, M., Eaves, C. J. and Marra, M. A. (2008b) Application of massively parallel sequencing to microRNA profiling and discovery in human embryonic stem cells. *Genome Res.* 18, 610-621.
- Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L. and Wold, B. (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat.Methods* 5, 621-628.
- Nagalakshmi, U., Wang, Z., Waern, K., Shou, C., Raha, D., Gerstein, M. and Snyder, M. (2008) The transcriptional landscape of the yeast genome defined by RNA sequencing. *Science* 320, 1344-1349.
- Niessen, K., Fu, Y., Chang, L., Hoodless, P. A., McFadden, D. and Karsan, A. (2008) Slug is a direct Notch target required for initiation of cardiac cushion cellularization. *J.Cell Biol.* 182, 315-325.
- Nieto, M. A. (2002) The snail superfamily of zinc-finger transcription factors. *Nat.Rev.Mol.Cell Biol.* 3, 155-166.
- Norris, R. A., Potts, J. D., Yost, M. J., Junor, L., Brooks, T., Tan, H., Hoffman, S., Hart, M. M., Kern, M. J., Damon, B., Markwald, R. R. and Goodwin, R. L. (2009) Periostin promotes a fibroblastic lineage pathway in atrioventricular valve progenitor cells. *Dev.Dyn.* 238, 1052-1063.

- Oram, K. F. and Gridley, T. (2005) Mutations in snail family genes enhance craniosynostosis of Twist1 haplo-insufficient mice: implications for Saethre-Chotzen Syndrome. *Genetics* 170, 971-974.
- Oshlack, A. and Wakefield, M. J. (2009) Transcript length bias in RNA-seq data confounds systems biology. *Biol.Direct* 4, 14.
- Otto, C. M. (2002) Calcification of bicuspid aortic valves. *Heart* 88, 321-322.
- Pierpont, M. E., Basson, C. T., Benson, D. W., Jr, Gelb, B. D., Giglia, T. M., Goldmuntz, E., McGee, G., Sable, C. A., Srivastava, D., Webb, C. L. and American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young. (2007) Genetic basis for congenital heart defects: current knowledge: a scientific statement from the American Heart Association Congenital Cardiac Defects Committee, Council on Cardiovascular Disease in the Young: endorsed by the American Academy of Pediatrics. *Circulation* 115, 3015-3038.
- Quere, R., Manchon, L., Lejeune, M., Clement, O., Pierrat, F., Bonafoux, B., Commes, T., Piquemal, D. and Marti, J. (2004) Mining SAGE data allows large-scale, sensitive screening of antisense transcript expression. *Nucleic Acids Res.* 32, e163.
- Robertson, A. G., Bilenky, M., Tam, A., Zhao, Y., Zeng, T., Thiessen, N., Cezard, T., Fejes, A. P., Wederell, E. D., Cullum, R., *et al.* (2008) Genome-wide relationship between histone H3 lysine 4 mono- and tri-methylation and transcription factor binding. *Genome Res.* 18, 1906-1917.
- Robertson, G., Hirst, M., Bainbridge, M., Bilenky, M., Zhao, Y., Zeng, T., Euskirchen, G., Bernier, B., Varhol, R., Delaney, A., *et al.* (2007) Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing. *Nat.Methods* 4, 651-657.
- Runyan, R. B. and Markwald, R. R. (1983) Invasion of mesenchyme into three-dimensional collagen gels: a regional and temporal analysis of interaction in embryonic heart tissue. *Dev.Biol.* 95, 108-114.
- Schoen, F. J. (2008) Evolving concepts of cardiac valve dynamics: the continuum of development, functional structure, pathobiology, and tissue engineering. *Circulation* 118, 1864-1880.
- Shelton, E. L. and Yutzey, K. E. (2008) Twist1 function in endocardial cushion cell proliferation, migration, and differentiation during heart valve development. *Dev.Biol.* 317, 282-295.
- Siddiqui, A. S., Khattri, J., Delaney, A. D., Zhao, Y., Astell, C., Asano, J., Babakaiff, R., Barber, S., Beland, J., Bohacec, S., *et al.* (2005) A mouse atlas of gene expression: large-scale digital gene-expression profiles from precisely defined developing C57BL/6J mouse tissues and cells. *Proc.Natl.Acad.Sci.U.S.A.* 102, 18485-18490.
- Sosic, D., Richardson, J. A., Yu, K., Ornitz, D. M. and Olson, E. N. (2003) Twist regulates cytokine gene expression through a negative feedback loop that represses NF-kappaB activity. *Cell* 112, 169-180.
- Spicer, D. B., Rhee, J., Cheung, W. L. and Lassar, A. B. (1996) Inhibition of myogenic bHLH and MEF2 transcription factors by the bHLH protein Twist. *Science* 272, 1476-1480.
- Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., Hodges, E., Anger, M., Sachidanandam, R., Schultz, R. M. and Hannon, G. J. (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature* 453, 534-538.

- van Bakel, H. and Hughes, T. R. (2009) Establishing legitimacy and function in the new transcriptome. *Brief Funct. Genomic Proteomic* 8, 424-436.
- Vanhee-Brossollet, C. and Vaquero, C. (1998) Do natural antisense transcripts make sense in eukaryotes? *Gene* 211, 1-9.
- Vincentz, J. W., Barnes, R. M., Firulli, B. A., Conway, S. J. and Firulli, A. B. (2008a) Cooperative interaction of Nkx2.5 and Mef2c transcription factors during heart development. *Dev. Dyn.* 237, 3809-3819.
- Vincentz, J. W., Barnes, R. M., Rodgers, R., Firulli, B. A., Conway, S. J. and Firulli, A. B. (2008b) An absence of Twist1 results in aberrant cardiac neural crest morphogenesis. *Dev. Biol.* 320, 131-139.
- Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., Chiba, H., Kohara, Y., Kono, T., Nakano, T., *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature* 453, 539-543.
- Wederell, E. D., Bilenky, M., Cullum, R., Thiessen, N., Dagpinar, M., Delaney, A., Varhol, R., Zhao, Y., Zeng, T., Bernier, B., *et al.* (2008) Global analysis of in vivo Foxa2-binding sites in mouse adult liver using massively parallel sequencing. *Nucleic Acids Res.* 36, 4549-4564.
- Winston, J. B., Erlich, J. M., Green, C. A., Aluko, A., Kaiser, K. A., Takematsu, M., Barlow, R. S., Sureka, A. O., Lapage, M. J., Janss, L. L. and Jay, P. Y. (2010) Heterogeneity of Genetic Modifiers Ensures Normal Cardiac Development. *Circulation*
- Yang, G. S., Banks, K. G., Bonaguro, R. J., Wilson, G., Dreolini, L., de Leeuw, C. N., Liu, L., Swanson, D. J., Goldowitz, D., Holt, R. A. and Simpson, E. M. (2009) Next generation tools for high-throughput promoter and expression analysis employing single-copy knock-ins at the Hprt1 locus. *Genomics* 93, 196-204.
- Yang, J., Mani, S. A., Donaher, J. L., Ramaswamy, S., Itzykson, R. A., Come, C., Savagner, P., Gitelman, I., Richardson, A. and Weinberg, R. A. (2004) Twist, a master regulator of morphogenesis, plays an essential role in tumor metastasis. *Cell* 117, 927-939.

APPENDICES

Appendix I. PCR primers

| QPCR primers | | |
|---------------------------------------|---------------------------|---------------------------|
| | Forward primer (5' to 3') | Reverse primer (5' to 3') |
| <i>Lef1</i> | TCACTGTCAGGCGACACTTC | TGAGGCTTCACGTGCATTAG |
| <i>Tbx20</i> | ACCAGTGCCTGGCCTAGTTTCAAT | CTACTGTCAGGAGATTTCCCGTATC |
| <i>Sox9</i> | AGCCCTAAGTGCCCAAGCACATT | AACGCTGGTATTCAGGGAGGTACA |
| <i>Twist1</i> | AGTCTGAACACTCGTTTGTGTCCC | ATGCCTTTCCTGTCAGTGGCTGAT |
| <i>Periostin</i> | GAGGTCTCCAAGGTCACAAAGTTC | TCTTCTTTGCAGGTGTGTCTCCCT |
| <i>β-actin</i> | GCTCTTTTCCAGCCTTC | CGGATGTCAACGTCACA |
| Primers for cloning into pCRII | | |
| | Forward primer (5' to 3') | Reverse primer (5' to 3') |
| <i>Lef1</i> | ATGTCGTAGCTGAGTGCACGCTAA | AAGAGGTGGCAGTGACTGTGTCTT |
| <i>Sox9</i> | GGTAACGATTGCTGGGATTC | CTGGTTGCAAGGAAGGCTAA |

Appendix II. Epithelium and mesenchymal library pairs in Mouse Atlas database

| Library ID | Description | Total RNA | Number of organs | Method | Total tags | Tag-types | Singletons |
|-----------------|--|-----------------|------------------|----------|------------|-----------|------------|
| SM169 | E12.0 Kidney Epithelium (TS20) | 1.5µg | 154 | LongSAGE | 47,568 | 15,502 | 10,608 |
| SM124 | E12.0 Kidney Mesenchyme (TS 20) | 2.5µg | 154 | LongSAGE | 90,457 | 28,345 | 19,240 |
| SM155 and SM156 | E12.0 Large Intestine Epithelium (TS 20) | 2.5µg and 0.5µg | 69 | LongSAGE | 106,733 | 21,581 | 12,404 |
| SM078 | E12.0 Large Intestine Mesenchyme (TS 20) | 10µg | 66 | LongSAGE | 86,078 | 23,095 | 15,126 |
| SM127 | E11.5 Lung Epithelium (TS 19) | 2.5µg | 64 | LongSAGE | 88,520 | 27,369 | 18,738 |
| SM135 | E11.5 Lung Mesenchyme (TS 19) | 5.5µg | 64 | LongSAGE | 99,197 | 25,036 | 16,058 |
| SM151 and SM159 | E12.0 Small Intestine Epithelium (TS 20) | 5µg and 0.5µg | 53 | LongSAGE | 88,407 | 20,905 | 13,401 |
| SM095 | E12.0 Small intestine Mesenchyme (TS 20) | 5µg | 53 | LongSAGE | 95,041 | 21,341 | 13,931 |
| SM065 | E16.5 Male Urogenital Sinus Epithelium (TS 24) | 10µg | 44 | LongSAGE | 122,991 | 28,066 | 17,153 |
| SM036 | E16.5 Male Urogenital Sinus Mesenchyme (TS 24) | 5µg | 44 | LongSAGE | 104,141 | 32,421 | 21,975 |
| SM163 | E16.5 Female Urogenital Sinus Epithelium (TS 24) | 2.5µg | 47 | LongSAGE | 119,087 | 30,788 | 20,565 |
| SM170 | E16.5 Female Urogenital Sinus Mesenchyme (TS 24) | 2.5µg | 47 | LongSAGE | 91,498 | 25,315 | 16,573 |

TS = Theiler stage

Appendix III. Embryonic SAGE libraries in Mouse Atlas database

| Library ID | Description | Total RNA | Hearts used | Method | Total tags | Tag-types | Singletons |
|-----------------|---------------------------------------|-----------|-------------|----------|------------|-----------|------------|
| SM145 | E8.0 Whole (TS 13) | 2.5µg | 8 | LongSAGE | 103,647 | 30,241 | 22,335 |
| SM006 | E9.0 Atria (TS 15) | 1.6µg | 72 | LongSAGE | 106,470 | 31,831 | 23,178 |
| SM206 and SM246 | E9.5 Atrio-ventricular canal (TS 15) | 550ng | 79 | LongSAGE | 314,093 | 65,103 | 47,395 |
| SM007 | E9.0 Left ventricle (TS 15) | 4.7µg | 78 | LongSAGE | 126,254 | 37,002 | 27,129 |
| SM005 | E9.0 Bulbus cordis (TS 15) | 5µg | 79 | LongSAGE | 107,191 | 30,010 | 20,928 |
| SM236 | E9.0 Outflow tract (TS 17) | 2.5µg | 84 | LongSAGE | 300,489 | 62,263 | 41,913 |
| SM234 | E10.5 Atrio-ventricular canal (TS 17) | 2.5µg | 25 | LongSAGE | 301,949 | 75,375 | 54,504 |
| SM235 | E10.5 Outflow tract (TS 17) | 2.5µg | 25 | LongSAGE | 313,516 | 68,453 | 49,163 |
| SM004 | E11.5 Atria (TS 19) | 5µg | 48 | LongSAGE | 108,532 | 28,469 | 20,265 |
| SM008 and SM241 | E11.5 Atrio-ventricular canal (TS 19) | 6.3µg | 48 | LongSAGE | 345,463 | 82,704 | 61,311 |
| SM051 | E11.5 Ventricles (TS 19) | 11.1µg | 48 | LongSAGE | 133,691 | 34,328 | 24,012 |
| SM002 | E12.5 Atria (TS 20) | 8.95µg | 34 | LongSAGE | 99,197 | 23,860 | 17,090 |
| SM238 | E12.5 Atrio-ventricular canal (TS 20) | 2.5µg | 18 | LongSAGE | 312,805 | 71,385 | 51,861 |
| SM003 | E12.5 Ventricles (TS 20) | 11.9µg | 10 | SAGE | 113,010 | 25,248 | 16,609 |

TS= Theiler stage

Appendix IV. Heart development gene expression in AVC libraries

| Name | Accession | Tag-type | E9.5 AVC | E10.5 AVC | E11.5 AVC | E12.5 AVC |
|---------------------------------|--------------|--------------------|-------------|--------------|--------------|--------------|
| <i>Bmp2</i> | NM_007553 | GAAGGTTGCTGAGCAAA | 5.7 | 17.1 | 1.5 | 2.9 |
| <i>Cdh11</i> | NM_009866 | TACGAAAATGTTACAGC | 0.0 | 6.3 | 13.0 | 6.4 |
| <i>Coup-TF2</i> | NM_009697 | ACAAAATAAACACTGTTG | 1.6 | 1.7 | 2.9 | 3.5 |
| <i>Cx45</i> | NM_008122 | TTCTCAGCAATAATGCA | 1.9 | 4.3 | 4.3 | 2.2 |
| <i>Edil3</i> | NM_001037987 | TAGGATGTTGTAAACTC | 1.6 | 4.3 | 8.4 | 1.6 |
| <i>ErbB2</i> | NM_001003817 | ACATCCAGGGCAGCCGG | 2.9 | 12.3 | 1.2 | 3.5 |
| <i>Fog2</i> | NM_011766 | GATGGAATAAAATTCCA | 0.6 | 0.3 | 0.6 | 0.0 |
| <i>Foxc2</i> | NM_013519 | GCTTTGTACAGTAGATG | 0.6 | 3.0 | 2.0 | 1.3 |
| <i>Hey2</i> | NM_013904 | TACTCTTATGCACTTCA | 4.1 | 1.3 | 1.4 | 2.6 |
| <i>Id1</i> | NM_010495 | TGTTCCAGCCGACGATC | 7.0 | 13.6 | 12.7 | 8.6 |
| <i>Id3</i> | NM_008321 | TGATGTATATTAAACTT | 3.8 | 9.6 | 36.5 | 18.9 |
| <i>Igfl</i> | NM_010512 | CCCAAGACTCAGAAGGA | 1.6 | 2.7 | 2.0 | 0.6 |
| <i>Jag1</i> | NM_013822 | GATAAACACCAGCAGAA | 0.0 | 2.6 | 0.6 | 0.6 |
| <i>Meox1</i> | NM_010791 | TATCAGTTTTCCCCTAC | 0.0 | 1.3 | 2.0 | 2.2 |
| <i>Mmp2</i> | NM_008610 | GGAAATGGCAAACAAGT | 3.8 | 10.6 | 16.2 | 12.1 |
| <i>Msx1</i> | NM_010835 | CTGGTGCTTCACCAAGG | 9.2 | 11.3 | 2.9 | 5.8 |
| <i>Msx2</i> | NM_013601 | TGCTTGAGTTGCTGGAG | 2.5 | 3.0 | 0.6 | 3.2 |
| <i>Nfl</i> | NM_010897 | TTGACGTCTGTCGCAGA | 1.9 | 2.3 | 3.2 | 1.0 |
| <i>NFATc</i> | NM_198429 | AACCATTCTTAGTAGAC | 1.0 | 1.7 | 1.5 | 1.6 |
| <i>PDGFRα</i> | NM_011058 | TTTTGTTTTAAAAAGTG | 1.0 | 1.0 | 6.4 | 1.6 |
| <i>Periostin</i> | NM_015784 | CAATGTGGGTTTCCTGC | 4.1 | 32.1 | 36.5 | 86.0 |
| <i>Prrx2</i> | NM_009116 | GCCAACAGCATCGCCAG | 12.7 | 10.3 | 6.9 | 5.8 |
| <i>Smad6</i> | NM_008542 | TCTCCGGATGCCACCAA | 7.0 | 11.9 | 4.9 | 3.5 |
| <i>Snai1</i> | NM_011427 | AATAATGGCCATCACTT | 0.6 | 2.0 | 2.6 | 1.3 |
| <i>Snai2</i> | NM_011415 | TGATGGATGCAGTAATA | 0.0 | 0.0 | 0.6 | 1.0 |
| <i>Sox9</i> | NM_011448 | GAGGACGATTGGAGAAT | 2.2 | 7.0 | 1.7 | 2.2 |
| <i>Tbx20</i> | NM_194263 | GCAGACACTGCAGGTGA | 1.3 | 2.3 | 0.0 | 0.3 |
| <i>Tbx5</i> | NM_011537 | TGAGATGTCTACGAACG | 0.3 | 0.3 | 1.4 | 0.0 |
| <i>Transgelin</i> | NM_011526 | GGCAGCTCCCACCTATC | 83.1 | 59.6 | 26.9 | 55.9 |
| <i>Twist1</i> | NM_011658 | GTAAAATGCAAATAGAT | 2.9 | 16.9 | 14.8 | 8.0 |
| <i>Vegfr1</i> | NM_010228 | GGTACCTGCTCCCCTGT | 0.0 | 0.0 | 0.3 | 0.3 |
| <i>Vimentin</i> | NM_011701 | AAGGAAGAGATGGCTCG | 53.5 | 90.1 | 77.3 | 113.8 |

*Expression levels expressed as tags per 100,000

Appendix V. Expression of AVC-enriched genes in heart libraries

| Name | Accession | Tag-type | E9.5 AVC | E9.0 LV | E11.5 Atria | E11.5 AVC | E11.5 V | E12.5 Atria | E12.5 AVC | E12.5 V |
|------------------|--------------|-------------------|-------------|------------|----------------|--------------|------------|----------------|--------------|------------|
| <i>Edil3</i> | NM_001037987 | TAGGATGTTGTAAACTC | 1.6 | 4 | 1.8 | 8.4 | 0 | 0 | 1.6 | 0.9 |
| <i>Fog2</i> | NM_011766 | GATGGAATAAAATCCA | 0.6 | 0 | 0 | 0.6 | 0 | 0 | 0 | 0.9 |
| <i>Id1</i> | NM_010495 | TGTTCCAGCCGACGATC | 7.0 | 4.8 | 8.3 | 12.7 | 3 | 4 | 8.6 | 0 |
| <i>Jag1</i> | NM_013822 | GATAAACACCAGCAGAA | 0 | 0 | 0 | 0.6 | 0 | 0 | 0.6 | 0.9 |
| <i>Mmp2</i> | NM_008610 | GGAAATGGCAAACAAGT | 3.8 | 4 | 11.1 | 16.2 | 7.5 | 8.1 | 12.1 | 9.7 |
| <i>Msx1</i> | NM_010835 | CTGGTGCTTCACCAAGG | 9.2 | 3.2 | 0 | 2.9 | 0 | 0 | 5.8 | 0.9 |
| <i>Msx2</i> | NM_013601 | TGCTTGAGTTGCTGGAG | 2.6 | 0 | 0 | 0.6 | 0 | 0 | 3.2 | 0 |
| <i>NFATc</i> | NM_198429 | AACCATTCTTAGTAGAC | 1 | 0.8 | 0.9 | 1.5 | 0 | 2 | 1.6 | 0.9 |
| <i>Periostin</i> | NM_015784 | CAATGTGGGTTTCCTGC | 4.1 | 0.8 | 2.8 | 36.5 | 3.7 | 0 | 86 | 7.6 |
| <i>Smad6</i> | NM_008542 | TCTCCGGATGCCACCAA | 7.0 | 0.8 | 0 | 4.9 | 0 | 2 | 3.5 | 1.8 |
| <i>Snai1</i> | NM_011427 | AATAATGGCCATCACTT | 0.6 | 0 | 1.8 | 2.6 | 0 | 0 | 1.3 | 0 |
| <i>Sox9</i> | NM_011448 | GAGGACGATTGGAGAAT | 2.2 | 0 | 1.8 | 1.7 | 0.7 | 0 | 2.2 | 0.9 |
| <i>Tbx20</i> | NM_194263 | GCAGACACTGCAGGTGA | 1.3 | 0.8 | 0 | 0 | 0 | 0 | 0.3 | 0 |
| <i>Vimentin</i> | NM_011701 | AAGGAAGAGATGGCTCG | 53.5 | 31.7 | 33.2 | 77.3 | 22.4 | 66.5 | 113.8 | 50.4 |

*Expression levels expressed as tags per 100,000. AVC = Atrio-ventricular canal; LV = left ventricle; V = ventricles.

Appendix VI. ChIP-qPCR primers

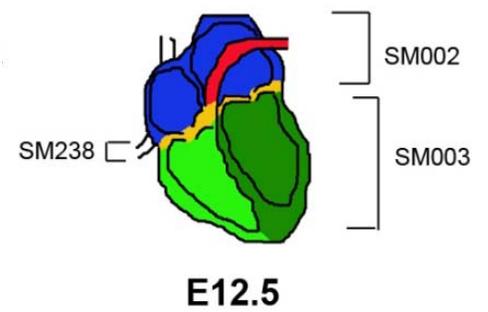
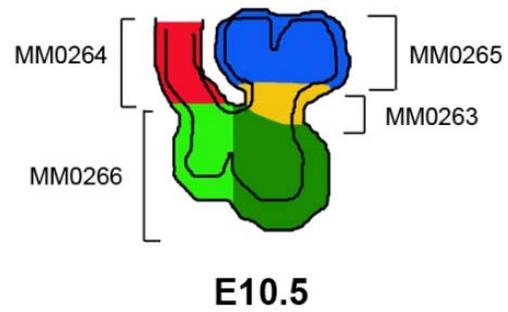
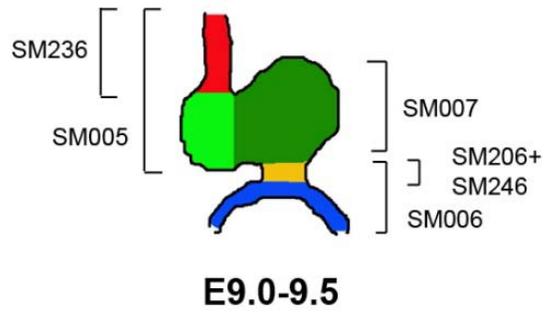
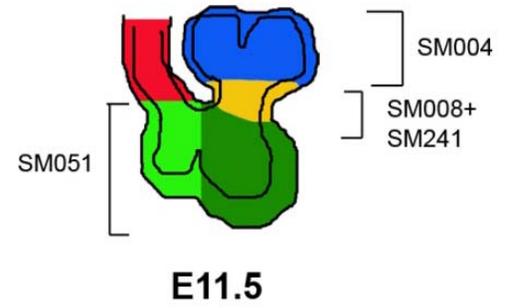
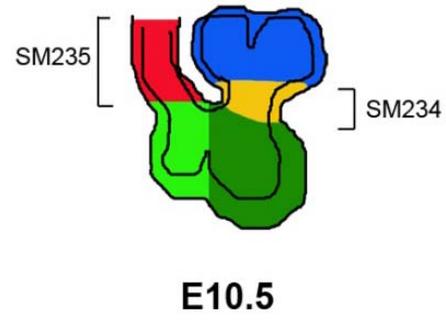
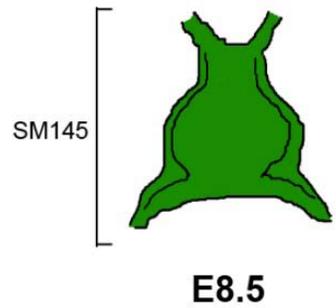
| Gene | Forward primer (5' to 3') | Reverse primer (5' to 3') |
|------------------|----------------------------|---------------------------|
| <i>Adssl1</i> | TTGAATCCTCTCTGCTCCTGTGT | ATATGCAGCTCTTGGGTGGCTGTA |
| <i>Aes</i> | TGACATGATGTTTCCGCAAAGCCG | TGCAGTGAGCTGAGCTGGG |
| <i>Aggrecan</i> | TATGTATGTGTCAACGCGCACCAT | AGAGAGAGTGAGGGACCTTGAGC |
| <i>Calcocol</i> | CCTCCTATAGTCGGAGTGCTGTTT | CACCCTCAGTTAGGCCAAGTTTCT |
| <i>Cdh11</i> | AACTTCAGTCGCTCCAGACTGTGT | AGATGCCAAACCCAGGAATAGGTG |
| <i>Decorin</i> | TCCTGAACTCTGCTACAGGCTA | AGCACAGCTTCCTTCTCGTTCCTT |
| <i>Erf</i> | GCGGGCACAGTGTCTCCAT | TCCCGGGTTAATATCGCACTCCG |
| <i>Foxm1</i> | AGACAAGCCGGTGCCGATT | CCGCTTTCAGTTGTTCCGCTGTTT |
| <i>Gata3</i> | GGGTTTGGGTTGCAGTTTCCTTGT | GCGACGCAACTTAAGGAGGTTCTA |
| <i>Gata4</i> | TTTCTGGGAAACTGGAGCTGGC | GCGGACTTGTGAGTTTCTCCTGC |
| <i>Gata5</i> | AGGGACCGACTAGAAGAGAGAAGG | TTGTAGACTGGGTCTCCACAGAGC |
| <i>Hand1</i> | GCTTAGAATTGTGTGGCGCTTGCT | CTTTGATGTCAACCTCTAGCCGGA |
| <i>Hcfc1</i> | CCAGCCTCTGACACGCCTTTATTA | TTGTGGAAGCCGCCATCTTGAAAC |
| <i>Hes6</i> | GCGAAAGGAAGTACAAGCACACCA | ACAGCCAGACCTCCTGCCACTTA |
| <i>Hmg20b</i> | AATGCGTCACCTGGCCAACATTAC | AGGGTTAGGGTGTTCATCCCAT |
| <i>Id2</i> | AGCTCTGGGAATTGGAATAAGGCG | GGCAAATTGAGTACAGTGTGCGCT |
| <i>Jund</i> | AGACGCACAGATGAGGTCCAGTTT | TCTGTCCCTACCCTGCTGTTTCTT |
| <i>Klf4</i> | AGACGACAGGACAAGCGCGTA | TCGAAAGTCCTGCCACGGGAA |
| <i>Lef1</i> | ATTCGGACATTCCCGGAGCCTT | CCCTCGGTCAAAGCAAAGAGCTG |
| <i>Mgp</i> | ACATAGGGTGGCCACAATTTCTGC | TGGCTTGGCACTAACTTGCCTTG |
| <i>Mlxip</i> | TTGTCCGCCCCGAGAAGAAA | TCCTCGTAACGTCCTGTCGTCTT |
| <i>Msx1</i> | AGCACAGCCCAATGGTTCTCT | TGTTAATAAGGCAAGGCCAGCGTG |
| <i>Nfatc1</i> | TTCCCTCTTGTACACCTTTGCCCA | GTAATGAGGACACAAGGGACTGGA |
| <i>Nkx2-5</i> | CCCTGCGTTTAGACTCAGCATAAC | TGCAAGGAGATTGCTCCTCGTTAG |
| NM_025423 | AACTTCTTCCGGGTCCTGTCTT | TGCCTTCTGGGAACTCATACCA |
| NM_025556 | AATCAAAGCAGCATGTCTCCAGGC | ACCCGCACTACAGCATCCCTTAAT |
| <i>Periostin</i> | TCTGTAAGGCCATCGCAAGCTTCA | TATCACACAGGAACAGCAGCAGCA |
| <i>Pi4ka</i> | AGAGAGGAACAGCTTGAGGCTTCT | CAAAGTGCGGCCTGTATTTGGGAA |
| <i>S100a10</i> | TCCCGAGACGGCTGGATTCTTATT | TCCCTGCGGAAGAATGCTCTTAGT |
| <i>Serbf2</i> | AGGCCTTAGAATGATCGAGGTTTCCC | ATCCCATCCGAGACTTAACAGGGT |
| <i>Sf3b4</i> | TGCACTTAGGACACAACCTCTGCGT | TGCTTCCAGACTGTTAGAGGCTCA |
| <i>Sirt1</i> | TTTAAATCTCCCGCAGCCGAGC | GCCATCTTCCAACCTGCCTCTCT |
| <i>Smad6</i> | AAGCGCTTTGTGCTCGTGTACC | ACCAGCGAAACGATGCTAGAGACA |
| <i>Sox4</i> | GCTTCTCATTGCACGCGGAGATTA | AGCCAATCAGCCGCTGTAACCTAAC |
| <i>Sox9</i> | TTGGCCCGAGGTATCTAACGTGAA | TGGTAAAGTTGTCTGCTCCACAGA |
| <i>Spl</i> | TACGCAACTTGCTCTTACACGCCT | TGAAAGGAGGGCCTTGACAGAGAAA |
| <i>Syng2</i> | TCCTTTGGTCTGGAGCAGAATGGT | AGACACCCAAATCCCACCAATCCA |
| <i>Taf6</i> | GCGGGCCCTTTAAATTTAGGCGAA | GCATGCGCTGACCGTTCTATGATT |
| <i>Tbl1x</i> | TGTAGAGAGCCTCTTCCAGGTGTCT | ACGCGAGCAAGCCTACCCA |
| <i>Tbx2</i> | TCAGATCGGTCTCTGCGCTTT | ATATTAACCAATGACGGGCCAGCG |
| <i>Tbx20</i> | AGTCTGGAAGCAGTGACGTGAGA | TGCCTCGCGCTTAATTTGCT |
| <i>Tead2</i> | TGGTTGGACCAGATACAGCTCTGA | AGAAGTGTGGGTGTGGGTGTGTAA |
| <i>Tnnc</i> | AGATACAGGTGCCAAGGTCTTCAG | TTACCGATGGCAACCATGAGTGGA |
| <i>Tpi1</i> | ATTTGGCCAAGGTCTGACAACCTGC | AAAGCAAAGTCACGGTTGAGGAGC |
| <i>Wdr74</i> | ACCGGCATCCTGAAAGGTGAGTAT | GACGGTGTGAAATTTGCCGCATGT |
| <i>Wdr75</i> | AAGGCAACGGCTTCACTAGGAGAA | TCATACCAGATGCCGCTACAGCTA |
| <i>Zeb1</i> | ATTCAAACCTCTGCAGCGTCCAAGG | ACTCGAGGCTTTACGACATCACCT |
| <i>Zfx3</i> | CTGCCTTTGCATAACAGAACGCCA | TCTACCACCAGGATCAAACCCAAC |

Appendix VII. Expression of known AVC genes in E10.5 heart Tag-seq libraries

| Name | Accession | Atria | AVC | Ventricles | OFT | AVC Enrichment |
|------------------|-----------|-------|--------|------------|-------|----------------|
| <i>Fog2</i> | NM_011766 | 1.9 | 9.4 | 8.3 | 3.4 | 3 |
| <i>Gata4</i> | NM_008092 | 9.4 | 227.4 | 5.5 | 122.8 | 33 |
| <i>Id1</i> | NM_010495 | 12.4 | 53 | 14.4 | 62.3 | 4 |
| <i>Id2</i> | NM_010496 | 63 | 1086.3 | 102.1 | 959.5 | 13.1 |
| <i>Jag1</i> | NM_013822 | 0 | 12 | 0 | 27.9 | 12 |
| <i>Mmp2</i> | NM_008610 | 26.8 | 194.2 | 29.6 | 237.2 | 6.9 |
| <i>Msx1</i> | NM_010835 | 4.5 | 92.4 | 1.42 | 24.7 | 42.8 |
| <i>NFATc</i> | NM_198429 | 17.1 | 46.4 | 13.5 | 41 | 3.1 |
| <i>Periostin</i> | NM_015784 | 52.5 | 144.1 | 32.9 | 138.6 | 3.6 |
| <i>Smad6</i> | NM_008542 | 0 | 53.5 | 0 | 61.5 | 53.5 |
| <i>Snai1</i> | NM_011427 | 2.1 | 43.5 | 1.42 | 18.4 | 25.5 |
| <i>Snai2</i> | NM_011415 | 0 | 13.3 | 1.4 | 10.8 | 11.3 |
| <i>Sox4</i> | NM_009238 | 0 | 98.8 | 0 | 86.5 | 68.5 |
| <i>Sox9</i> | NM_011448 | 4.1 | 514.4 | 4.3 | 376.7 | 123.6 |
| <i>Tbx2</i> | NM_009324 | 0 | 20.1 | 0 | 6.1 | 9.7 |
| <i>Tbx20</i> | NM_194263 | 1.5 | 51.7 | 0 | 40.1 | 43.1 |
| <i>Vimentin</i> | NM_011701 | 561 | 1473.1 | 790 | 1610 | 2.3 |

*Expression levels expressed as tags per million.

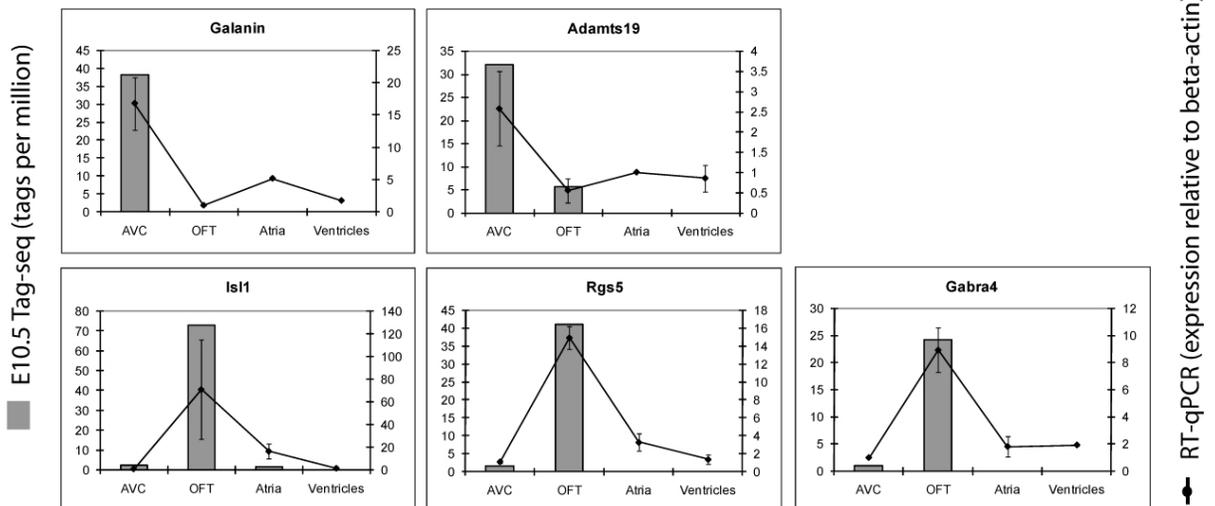
Appendix VIII. Heart regions sampled for gene profiling



Appendix IX. Validation of AVC- and OFT-specific gene expression

Selected AVC and OFT specific genes were validated by RT-qPCR with the following primers (*Galanin* forward primer 5' ATGCCTGCAAAGGAGAAGAGAGGT3', reverse primer 5' GTGAGGCCATGCTTGTGCTGCTAAAT3'; *Adams19* forward primer 5' TGAACACAGCACTGTGAGCCTGAT3', reverse primer 5' AGGTGTTATAGGGAGCTGCCACTT3'; *Isl1* forward primer 5' GATGCGCTCATGAAGGAGCAACTA3', reverse primer 5' GTTTCTTGCCTTGCACCGCTTGT3'; *Rgs5* forward primer 5' AAGAATTCATCCAGACAGAGGCC3', reverse primer 5' AGACGGTTCCACCAGGTTCTTCAT3'; *Gabra4* forward primer 5' AGGTGGGAAATCACTCCAGCAAGA3', reverse primer 5' ACATAGTCTCAGCTGCATTTGCC3'; β -actin forward primer 5' GCTCTTTTCCAGCCTTC3', reverse primer 5' CGGATGTCAACGTCACA3')

RT-qPCR results are represented as average values from three independent E10.5 hearts samples \pm standard deviations.



Appendix X. Differentially expressed genes in *Twist1* null AVC

Expression normalized per million.

| Gene symbol | Gene name | RefSeq | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|--|--|--------------|-------|------|------|-------|--------|-------------|
| A. Genes up-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| | | | | | 4411 | | | |
| <i>Mt2</i> | Metallothionein 2 | NM_008630 | 6.1 | 8.2 | .5 | 886 | 425.8 | 69.9 |
| <i>2900062</i> | RIKEN cDNA | | | | | | | |
| <i>L11Rik</i> | 2900062L11 gene | NR_003642 | 1.3 | <1.1 | <1.3 | <1.4 | 65.8 | 49.8 |
| <i>LOC674321</i> | PREDICTED: similar to Gcsh protein | XM_973659 | <1 | <1.1 | <1.3 | <1.4 | 45.3 | 45.9 |
| <i>LOC100047762</i> | PREDICTED: similar to Aspartate aminotransferase, cytoplasmic (Transaminase A) (Glutamate oxaloacetate transaminase 1) | XM_001478835 | <1 | <1.1 | <1.3 | <1.4 | 45 | 45.5 |
| <i>LOC100046877</i> | PREDICTED: similar to developmentally regulated RNA-binding protein 1 | XM_001477072 | <1 | <1.1 | <1.3 | <1.4 | 36.3 | 36.7 |
| <i>Clk1</i> | CDC-like kinase 1 | NR_027854 | <1 | <1.1 | <1.3 | <1.4 | 33 | 33.4 |
| <i>LOC100048483</i> | PREDICTED: similar to cytochrome c oxidase subunit VIII | XM_001480453 | 6.3 | 13.7 | 6 | 39.8 | 191 | 30.6 |
| <i>Hbb-y</i> | Hemoglobin Y, beta-like embryonic chain | NM_008221 | 597.8 | 1095 | 1553 | 35381 | 15444 | 25.8 |
| <i>Slc4a1</i> | solute carrier family 4 (anion exchanger), member 1 | NM_011403 | <1 | 4 | 46.9 | 10.4 | 24.5 | 24.8 |
| <i>Irf2</i> | Interferon regulatory factor 2 | NM_008391 | 1 | 10.5 | 7.3 | 15.4 | 23.6 | 23.9 |
| <i>Cat</i> | Catalase | NM_009804 | 4 | 6.7 | 32.1 | 24.9 | 90.5 | 22.9 |
| <i>LOC100044980</i> | PREDICTED: similar to Thoc7 protein | XM_001473442 | <1 | <1.1 | <1.3 | <1.4 | 20.9 | 21.2 |
| <i>Tpi1</i> | Triosephosphate isomerase 1 | NM_009415 | 189.3 | 258. | 1318 | 1669. | 3913.6 | 20.7 |
| <i>Adrm1</i> | Adhesion regulating molecule 1 | NM_019822 | 1.8 | 2.3 | <1.3 | <1.4 | 36.1 | 19.9 |
| <i>Uqcr</i> | Ubiquinol-cytochrome c reductase (6.4kD) subunit | NM_025650 | 67.3 | 112 | 1420 | 1680 | 1330 | 19.8 |
| <i>LOC100045884</i> | PREDICTED: hypothetical protein LOC100045884 | XM_001475215 | <1 | <1.1 | <1.3 | <1.4 | 18.9 | 19.1 |
| <i>Gm1943</i> | Predicted gene 1943 | NR_002928 | <1 | <1.1 | <1.3 | <1.4 | 18.5 | 18.8 |
| <i>Ankrd37</i> | Ankyrin repeat domain 37 | NM_001039562 | 3 | 2.1 | 17.6 | 8.3 | 53.5 | 18.1 |
| <i>Laptm4b</i> | Lysosomal-associated protein transmembrane 4B | NM_033521 | <1 | <1.1 | <1.3 | <1.4 | 17.5 | 17.7 |
| <i>Sf3b4</i> | Splicing factor 3b, subunit 4 | NM_153053 | 22.1 | 51.2 | 185. | 4 | 225.1 | 17.4 |
| <i>Mogat2</i> | Monoacylglycerol O-acyltransferase 2 | NM_177448 | <1 | 1.7 | 7.5 | 7.8 | 17.2 | 17.4 |
| <i>Cfc1</i> | Cripto, FRL-1, cryptic family 1 | NM_007685 | 1 | 2.5 | <1.3 | 4.5 | 17 | 17.2 |

| Gene symbol | Gene name | RefSeq | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|------------------|--|--------------|------|------|------|------|--------|-------------|
| <i>Rwdd2b</i> | RWD domain containing 2B | NM_016924 | 1.3 | 1.1 | 6.2 | 7.3 | 22.7 | 17.2 |
| <i>Wdr74</i> | WD repeat domain 74 | NM_134139 | 7.2 | 14.2 | 34.5 | 32.2 | 124 | 17.1 |
| <i>Gspt2</i> | G1 to S phase transition 2 | NM_008179 | 1.2 | 5.7 | 5.4 | 14.7 | 19.1 | 16.6 |
| <i>Tusc4</i> | Tumor suppressor candidate 4 | NM_018879 | <1 | 1.9 | 3.4 | 4.7 | 16.3 | 16.5 |
| <i>Fst</i> | Follistatin | NM_008046 | 1 | 1.1 | 15.9 | <1.4 | 15.3 | 15.5 |
| <i>LOC675857</i> | PREDICTED: similar to valosin | XM_985571 | 4 | <1.1 | <1.3 | <1.4 | 60.6 | 15.3 |
| <i>Thyn1</i> | Thymocyte nuclear protein 1 | NM_144543 | 2 | 1.9 | 3.2 | 3.6 | 29.6 | 14.9 |
| <i>Ptpn7</i> | Protein tyrosine phosphatase, non-receptor type 7 | NM_177081 | <1 | 1.1 | <1.3 | <1.4 | 14.3 | 14.5 |
| <i>Casp9</i> | Caspase 9 | NM_015733 | <1 | <1.1 | <1.3 | <1.4 | 13.9 | 14 |
| <i>I110059</i> | RIKEN cDNA 1110059E24 gene | NM_025423 | 4.1 | 2.5 | 1.9 | 8.8 | 57.6 | 14 |
| <i>Pi4ka</i> | Phosphatidylinositol 4-kinase, catalytic, alpha polypeptide | NM_001001983 | 7.9 | 13.6 | 81.4 | 76.2 | 110.1 | 13.9 |
| <i>Taf2</i> | TAF2 RNA polymerase II, TATA box binding protein (TBP)-associated factor | NM_001081288 | 3.5 | 4.9 | <1.3 | <1.4 | 47.4 | 13.7 |
| <i>LOC675032</i> | PREDICTED: similar to polymyositis scleroderma overlap syndrome (PM-SCL) antigen 1 a | XM_983011 | <1 | <1.1 | <1.3 | <1.4 | 13 | 13.1 |
| <i>Phlpp2</i> | PH domain and leucine rich repeat protein phosphatase 2 | NM_001122594 | 1.2 | <1.1 | <1.3 | 1.9 | 15 | 13 |
| <i>Wdr75</i> | WD repeat domain 75 | NM_028599 | 5.8 | 20.3 | 12 | 15.4 | 74.1 | 12.9 |
| <i>Sms</i> | Spermine synthase | NM_009214 | 1.5 | 5.1 | <1.3 | 5.5 | 18.5 | 12.5 |
| <i>Atp5e</i> | ATP synthase, H ⁺ transporting, mitochondrial F1 complex, epsilon subunit | NM_025983 | <1 | 2.9 | 1.3 | 1.4 | 12.2 | 12.4 |
| <i>LOC636885</i> | PREDICTED: similar to ATP binding domain 1 family, member B | XR_033998 | <1 | <1.1 | <1.3 | <1.4 | 11.9 | 12 |
| <i>Sirt6</i> | Sirtuin 6 | NM_181586 | <1 | <1.1 | <1.3 | <1.4 | 11.8 | 11.9 |
| <i>Vnn1</i> | Vanin 1 | NM_011704 | 2.1 | 16.3 | <1.3 | 5.9 | 25.5 | 11.9 |
| <i>Zfp652</i> | Zinc finger protein 652 | NM_201609 | 1 | <1.1 | <1.3 | 5.2 | 11.6 | 11.7 |
| <i>Nsun2</i> | NOL1/NOP2/Sun domain family member 2 | NM_145354 | <1 | 5.9 | 6.4 | <1.4 | 11.6 | 11.7 |
| <i>Stard10</i> | START domain containing 10 | NM_019990 | 12.3 | 9.5 | 272 | 14.7 | 141.6 | 11.5 |
| <i>Arsg</i> | Arylsulfatase G | NM_028710 | <1 | 2.3 | 3 | 6.9 | 11.1 | 11.2 |
| <i>Prkdc</i> | Protein kinase, DNA activated, catalytic polypeptide | NM_011159 | 3 | 6.3 | 20.8 | 32.7 | 33.2 | 11.2 |
| <i>Slc25a22</i> | Solute carrier family 25 (mitochondrial carrier, glutamate), member 22 | NM_026646 | 3 | 12.5 | 37.9 | 50.9 | 32.4 | 11 |

| Gene symbol | Gene name | RefSeq | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|--|--|--------------|-------|-------|------|------|--------|-------------|
| <i>1700052</i> | RIKEN cDNA | | | | | | | |
| <i>K11Rik</i> | 1700052K11 gene | NR_027956 | 1.7 | 3.8 | <1.3 | <1.4 | 18.1 | 10.9 |
| <i>Ipp</i> | IAP promoted placental gene | NM_008389 | 1 | 5.7 | 4.7 | 8.8 | 10.7 | 10.9 |
| <i>Trim72</i> | Tripartite motif-containing 72 | NM_001079932 | 2.6 | 2.9 | 5.8 | 23 | 28.2 | 10.7 |
| <i>Cmtm4</i> | CKLF-like MARVEL transmembrane domain containing 4 | NM_153582 | <1 | <1.1 | 3 | 2.8 | 10.6 | 10.7 |
| <i>Cldn12</i> | Claudin 12 | NM_022890 | 4.5 | 10.8 | 12.9 | 29.4 | 47.7 | 10.7 |
| <i>Ankrd54</i> | Ankyrin repeat domain 54 | NM_144849 | 3 | 7 | 7.1 | 28.2 | 31.5 | 10.6 |
| <i>Zc3hcl</i> | Zinc finger, C3HC type 1 | NM_172735 | 8.7 | 9.7 | 51.4 | 96.8 | 92.2 | 10.6 |
| <i>Flad1</i> | RFad1, flavin adenine dinucleotide synthetase | NM_177041 | 1 | 2.5 | 2.4 | 8.3 | 10.4 | 10.5 |
| <i>Hbb-bh1</i> | Hemoglobin Z, beta-like embryonic chain | NM_008219 | 603 | 701 | 0 | 8910 | 6320 | 10.5 |
| <i>Gata1</i> | GATA binding protein 1 | NM_008089 | 1 | <1.1 | 12.2 | 4.7 | 10.2 | 10.3 |
| <i>Ncrna00086</i> | Non-protein coding RNA 86 | NR_028086 | <1 | <1.1 | <1.3 | <1.4 | 10.1 | 10.2 |
| <i>Adssl1</i> | Adenylosuccinate synthetase like 1 | NM_007421 | 21.3 | 25 | 58.1 | 146 | 216 | 10.2 |
| B. Genes down-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| <i>Twist1</i> | Twist homolog 1 | NM_011658 | 918.0 | 299.1 | 63.4 | 55.9 | 4.6 | 0.005 |
| <i>Rn18s</i> | 18S RNA | NR_003278 | 659.6 | 908.6 | 7.3 | 4.5 | 14.1 | 0.021 |
| <i>Ntm</i> | Neurotrimin | NM_172290 | 18.9 | 13.7 | <1.3 | <1.4 | <0.5 | 0.024 |
| <i>Tbllx</i> | Transducin (beta)-like 1 X-linked | NM_020601 | 24.2 | 19.9 | <1.3 | <1.4 | 0.6 | 0.026 |
| <i>LOC236598</i> | 28S ribosomal RNA | NR_003279 | 694.3 | 981.9 | 2.8 | <1.4 | 18.5 | 0.027 |
| <i>Peg3</i> | Paternally expressed 3 | NM_008817 | 16.8 | 17.3 | <1.3 | <1.4 | <0.5 | 0.028 |
| <i>Chd8</i> | Chromodomain helicase DNA binding protein 8 | NM_201637 | 46.9 | 29.8 | <1.3 | 2.1 | 1.4 | 0.030 |
| <i>Tnni3k</i> | TNNI3 interacting kinase | NM_177066 | 12.5 | 15.3 | <1.3 | <1.4 | <0.5 | 0.037 |
| <i>Mgp</i> | Matrix Gla protein | NM_008597 | 12.3 | 1.9 | 2.6 | 1.7 | <0.5 | 0.038 |
| <i>Pf4</i> | Platelet factor 4 | NM_019932 | 23.7 | 3.0 | 1.9 | 3.8 | 0.9 | 0.039 |
| <i>Ptprb</i> | Protein tyrosine phosphatase, receptor type, B | NM_029928 | 10.7 | 20.1 | <1.3 | <1.4 | <0.5 | 0.043 |
| <i>LOC10047427</i> | PREDICTED: similar to thyroid hormone receptor | XM_001478949 | 10.4 | 10.3 | 1.5 | 1.4 | <0.5 | 0.045 |
| <i>Col9a2</i> | Collagen, type IX, alpha 2 | NM_007741 | 14.8 | 17.5 | <1.3 | <1.4 | 0.7 | 0.047 |
| <i>Miip</i> | Migration and invasion inhibitory protein | NM_001025365 | 15.5 | 8.2 | 2.6 | <1.4 | 0.8 | 0.050 |
| <i>LOC674324</i> | PREDICTED: similar to glyceraldehyde-3-phosphate dehydrogenase | XR_032032 | 19.8 | 2.9 | 14.4 | 13.5 | 1.0 | 0.051 |
| <i>Syngn2</i> | Synaptogyrin 2 | NM_009304 | 41.4 | 27.4 | 15.2 | 15.2 | 2.2 | 0.054 |

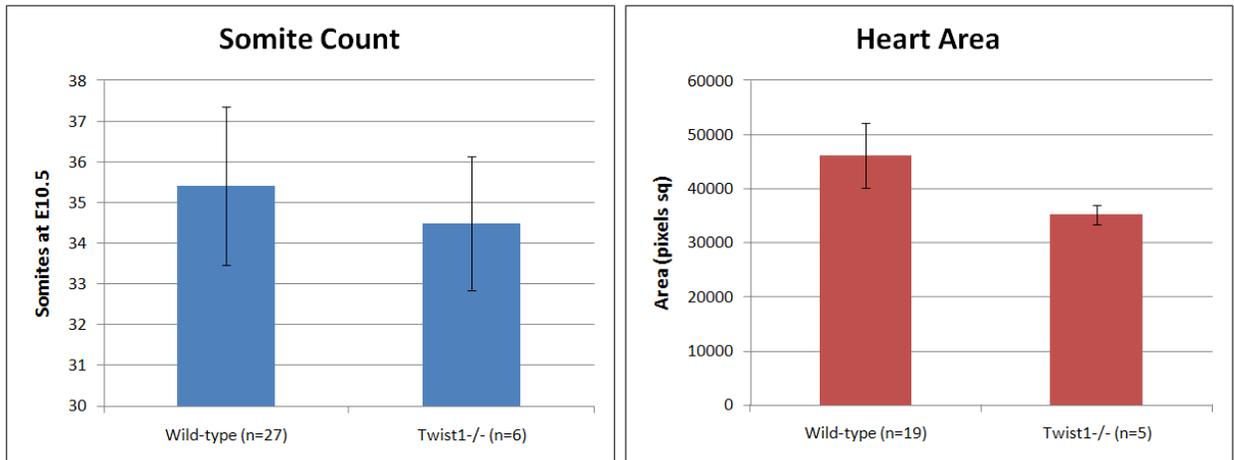
| Gene symbol | Gene name | RefSeq | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|-----------------|--|-----------|-------|-------|------|------|--------|-------------|
| <i>Papss2</i> | 3'-phosphoadenosine 5'-phosphosulfate synthase 2 | NM_011864 | 218.9 | 115.5 | <1.3 | <1.4 | 12.0 | 0.055 |
| <i>Zc3h13</i> | Zinc finger CCCH type containing 13 | NM_026083 | 32.9 | 31.9 | 8.8 | 5.5 | 1.8 | 0.056 |
| <i>Dhrs4</i> | Dehydrogenase/reductase (SDR family) member 4 | NM_030686 | 8.1 | 5.7 | 2.6 | <1.4 | 0.5 | 0.057 |
| <i>Mks1</i> | Meckel syndrome, type 1 | NM_001039 | 8.1 | 3.4 | 1.5 | 3.8 | 0.5 | 0.057 |
| <i>Nrg1</i> | Neuregulin 1 | NM_178591 | 7.9 | 16.3 | <1.3 | <1.4 | <0.5 | 0.058 |
| <i>5031439</i> | RIKEN cDNA | NM_001033 | | | | | | |
| <i>G07Rik</i> | 5031439G07 gene | 273 | 26.0 | 39.3 | 1.5 | 1.7 | 1.5 | 0.059 |
| <i>Creb5</i> | cAMP responsive element binding protein 5 | NM_172728 | 10.4 | 16.0 | <1.3 | <1.4 | 0.6 | 0.059 |
| <i>Car13</i> | Carbonic anhydrase 13 | NM_024495 | 7.7 | 2.7 | 6.9 | 12.1 | 0.5 | 0.060 |
| <i>Tcofl</i> | Treacher Collins Franceschetti syndrome 1, homolog | NM_011552 | 11.4 | 7.0 | <1.3 | <1.4 | 0.7 | 0.061 |
| <i>Maged2</i> | Melanoma antigen, family D, 2 | NM_030700 | 10.0 | 13.5 | <1.3 | <1.4 | 0.6 | 0.062 |
| <i>Cox10</i> | COX10 homolog, cytochrome c oxidase assembly protein, heme A: farnesyltransferase | NM_178379 | 14.8 | 16.7 | 1.9 | 2.6 | 0.9 | 0.063 |
| <i>Nrk</i> | Nik related kinase | NM_013724 | 7.3 | 6.1 | <1.3 | <1.4 | <0.5 | 0.064 |
| <i>Zcchc12</i> | Zinc finger, CCHC domain containing 12 | NM_028325 | 7.2 | 11.4 | <1.3 | <1.4 | 0.5 | 0.064 |
| <i>Kcne11</i> | Potassium voltage-gated channel, Isk-related family, member 1-like | NM_021487 | 10.5 | 4.0 | <1.3 | 1.9 | 0.7 | 0.066 |
| <i>Pigs</i> | Phosphatidylinositol glycan anchor biosynthesis, class S | NM_201406 | 23.1 | 15.7 | 1.9 | 2.1 | 1.5 | 0.067 |
| <i>Wnt9b</i> | Wingless-type MMTV integration site 9B | NM_011719 | 8.1 | <1.1 | <1.3 | <1.4 | 0.5 | 0.067 |
| <i>Slc25a19</i> | Solute carrier family 25 (mitochondrial thiamine pyrophosphate carrier), member 19 | NM_026071 | 6.9 | 6.7 | <1.3 | <1.4 | <0.5 | 0.067 |
| <i>Cd9</i> | CD9 antigen | NM_007657 | 29.7 | 19.6 | 3.6 | 10.2 | 2.0 | 0.067 |
| <i>Acsf2</i> | Acyl-CoA synthetase family member 2 | NM_153807 | 6.8 | 3.0 | <1.3 | <1.4 | <0.5 | 0.068 |
| <i>Smpd3</i> | Sphingomyelin phosphodiesterase 3 | NM_021491 | 6.8 | 2.9 | 1.9 | 1.4 | <0.5 | 0.068 |
| <i>Casp3</i> | Caspase 3 | NM_009810 | 7.7 | 4.9 | <1.3 | <1.4 | 0.5 | 0.070 |
| <i>Coll16a1</i> | Collagen, type XVI, alpha 1 | NM_028266 | 7.7 | 4.0 | 3.0 | 1.4 | 0.5 | 0.070 |
| <i>Trim16</i> | Tripartite motif-containing 16 | NM_053169 | 6.6 | 11.0 | 1.5 | 3.3 | <0.5 | 0.070 |
| <i>Ramp1</i> | Receptor (calcitonin) activity modifying protein 1 | NM_016894 | 12.0 | 13.3 | 1.3 | 8.1 | 0.8 | 0.071 |

| Gene symbol | Gene name | RefSeq | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|-----------------|---|--------------|-------|-------|------|------|--------|-------------|
| <i>Gp5</i> | Glycoprotein 5 (platelet) | NM_008148 | 6.4 | 3.4 | <1.3 | 2.6 | <0.5 | 0.072 |
| <i>Taf15</i> | TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor | NM_027427 | 104.6 | 92.0 | 7.3 | 10.9 | 7.8 | 0.074 |
| <i>Map3k3</i> | Mitogen-activated protein kinase kinase kinase 3 | NM_011947 | 21.4 | 18.2 | <1.3 | <1.4 | 1.6 | 0.076 |
| <i>Dcn</i> | Decorin | NM_007833 | 60.4 | 41.8 | 33.6 | 33.9 | 4.7 | 0.078 |
| <i>Mapk11</i> | Mitogen-activated protein kinase 11 | NM_011161 | 5.9 | 10.3 | <1.3 | <1.4 | 0.5 | 0.078 |
| <i>Col9a3</i> | Collagen, type IX, alpha 3 | NM_009936 | 120.5 | 38.2 | 3.0 | <1.4 | 9.4 | 0.078 |
| <i>Hnrnpul2</i> | Heterogeneous nuclear ribonucleoprotein U-like 2 | NM_001081196 | 5.8 | 2.5 | <1.3 | 3.6 | <0.5 | 0.080 |
| <i>Atxn1</i> | Ataxin 1 | NM_009124 | 9.6 | 11.8 | 1.7 | <1.4 | 0.8 | 0.081 |
| <i>6330408</i> | RIKEN cDNA | | | | | | | |
| <i>A02Rik</i> | 6330408A02 gene | NM_177312 | 5.6 | 4.6 | 1.3 | <1.4 | 0.5 | 0.083 |
| <i>Rnf187</i> | Ring finger protein 187 | NM_022423 | 5.6 | 3.0 | <1.3 | <1.4 | 0.5 | 0.083 |
| <i>Galnt10</i> | UDP-N-acetyl-alpha-D-galactosamine:polypeptide N-acetylgalactosaminyltransferase 10 | NM_134189 | 15.6 | 16.7 | 5.1 | 1.9 | 1.3 | 0.084 |
| <i>Tbc1d20</i> | TBC1 domain family, member 20 | NM_024196 | 5.4 | 4.9 | <1.3 | <1.4 | <0.5 | 0.085 |
| <i>Xrn1</i> | 5'-3' exoribonuclease 1 | NM_011916 | 5.4 | 4.6 | <1.3 | <1.4 | 0.5 | 0.085 |
| <i>Prpf19</i> | PRP19/PSO4 pre-mRNA processing factor 19 homolog | NM_134129 | 5.4 | 2.1 | <1.3 | <1.4 | <0.5 | 0.085 |
| <i>Nxn</i> | Nucleoredoxin | NM_008750 | 5.4 | 2.9 | 4.5 | 7.6 | <0.5 | 0.085 |
| <i>Plek2</i> | Pleckstrin 2 | NM_013738 | 5.4 | 8.9 | 19.3 | 8.3 | <0.5 | 0.085 |
| <i>Tatdn2</i> | TatD DNase domain containing 2 | NM_001033463 | 8.9 | 7.6 | <1.3 | <1.4 | 0.8 | 0.087 |
| <i>Slc6a17</i> | Solute carrier family 6 (neurotransmitter transporter), member 17 | NM_172271 | 5.3 | 2.7 | <1.3 | <1.4 | <0.5 | 0.088 |
| <i>Aldh2</i> | Aldehyde dehydrogenase 2, mitochondrial | NM_009656 | 34.0 | 28.3 | 2.1 | <1.4 | 3.0 | 0.089 |
| <i>Crebbp</i> | CREB binding protein | NM_001025432 | 6.9 | 8.4 | <1.3 | <1.4 | 0.6 | 0.089 |
| <i>Mfsd8</i> | Major facilitator superfamily domain containing 8 | NM_028140 | 6.9 | 5.1 | 2.4 | 8.5 | 0.6 | 0.089 |
| <i>EG668479</i> | PREDICTED: predicted gene, EG668479 | XM_001001706 | 16.1 | 1.3 | 55.3 | 83.1 | 1.5 | 0.091 |
| <i>Bgn</i> | Biglycan | NM_007542 | 49.4 | 36.9 | <1.3 | <1.4 | 4.6 | 0.092 |
| <i>S100a10</i> | S100 calcium binding protein A10 (calpactin) | NM_009112 | 158.1 | 209.7 | 61.3 | 54.0 | 14.6 | 0.092 |
| <i>2410022</i> | RIKEN cDNA | | | | | | | |
| <i>L05Rik</i> | 2410022L05 gene | NM_025556 | 43.1 | 73.5 | 11.8 | 18.0 | 4.0 | 0.093 |
| <i>Dagla</i> | Diacylglycerol lipase, alpha | NM_198114 | 13.2 | 4.8 | 1.3 | <1.4 | 1.2 | 0.093 |

| Gene symbol | Gene name | RefSeq` | AVC | OFT | A | V | Twist1 | Twist1 enr. |
|------------------|---|--------------|------|------|------|------|--------|-------------|
| <i>BC030307</i> | cDNA sequence BC030307 | NM_001003910 | 4.9 | 7.6 | 1.3 | <1.4 | <0.5 | 0.094 |
| <i>Trim2</i> | Tripartite motif-containing 2 | NM_030706 | 4.9 | 4.8 | <1.3 | <1.4 | 0.5 | 0.094 |
| <i>Scpep1</i> | Serine carboxypeptidase 1 Brain and acute leukemia, | NM_029023 | 47.3 | 39.7 | 65.6 | 64.2 | 4.5 | 0.095 |
| <i>Baalc</i> | cytoplasmic | NM_080640 | 14.6 | 5.7 | <1.3 | <1.4 | 1.4 | 0.095 |
| <i>Fmod</i> | Fibromodulin | NM_021355 | 5.6 | <1.1 | 3.0 | <1.4 | 0.5 | 0.096 |
| <i>Chtf8</i> | CTF8, chromosome transmission fidelity factor 8 homolog | NM_145412 | 4.8 | <1.1 | <1.3 | <1.4 | <0.5 | 0.097 |
| <i>Chtf18</i> | CTF18, chromosome transmission fidelity factor 18 homolog | NM_145409 | 4.8 | <1.1 | <1.3 | 1.7 | <0.5 | 0.097 |
| <i>Apitd1</i> | Apoptosis-inducing, TAF9-like domain 1 | NM_027263 | 13.5 | 10.4 | 1.9 | <1.4 | 1.3 | 0.097 |
| <i>Dync1h1</i> | Dynein cytoplasmic 1 heavy chain 1 | NM_030238 | 28.6 | 28.7 | 1.5 | <1.4 | 2.8 | 0.097 |
| <i>Unc13b</i> | Unc-13 homolog B | NM_001081413 | 16.6 | 17.1 | <1.3 | <1.4 | 1.6 | 0.098 |
| <i>Tnfrsf12a</i> | Tumor necrosis factor receptor superfamily, member 12a | NM_013749 | 12.5 | 17.3 | <1.3 | <1.4 | 1.2 | 0.098 |
| <i>Pqhc2</i> | PQ loop repeat containing 2 | NM_145384 | 6.3 | 1.9 | <1.3 | <1.4 | 0.6 | 0.099 |

Appendix XI. *Twist1* null hearts are smaller than wild-type

Wild-type and *Twist1* null AVCs were dissected at E10.5 and imaged. Wild-type and *Twist1* null embryos had similar number of somites at time of dissection. Heart size was calculated by measuring the area in microscopic images. Error bars represent \pm standard deviations.



Appendix XII. Differentially expressed genes in *Twist1* null AVC

Expression normalized to 100,000.

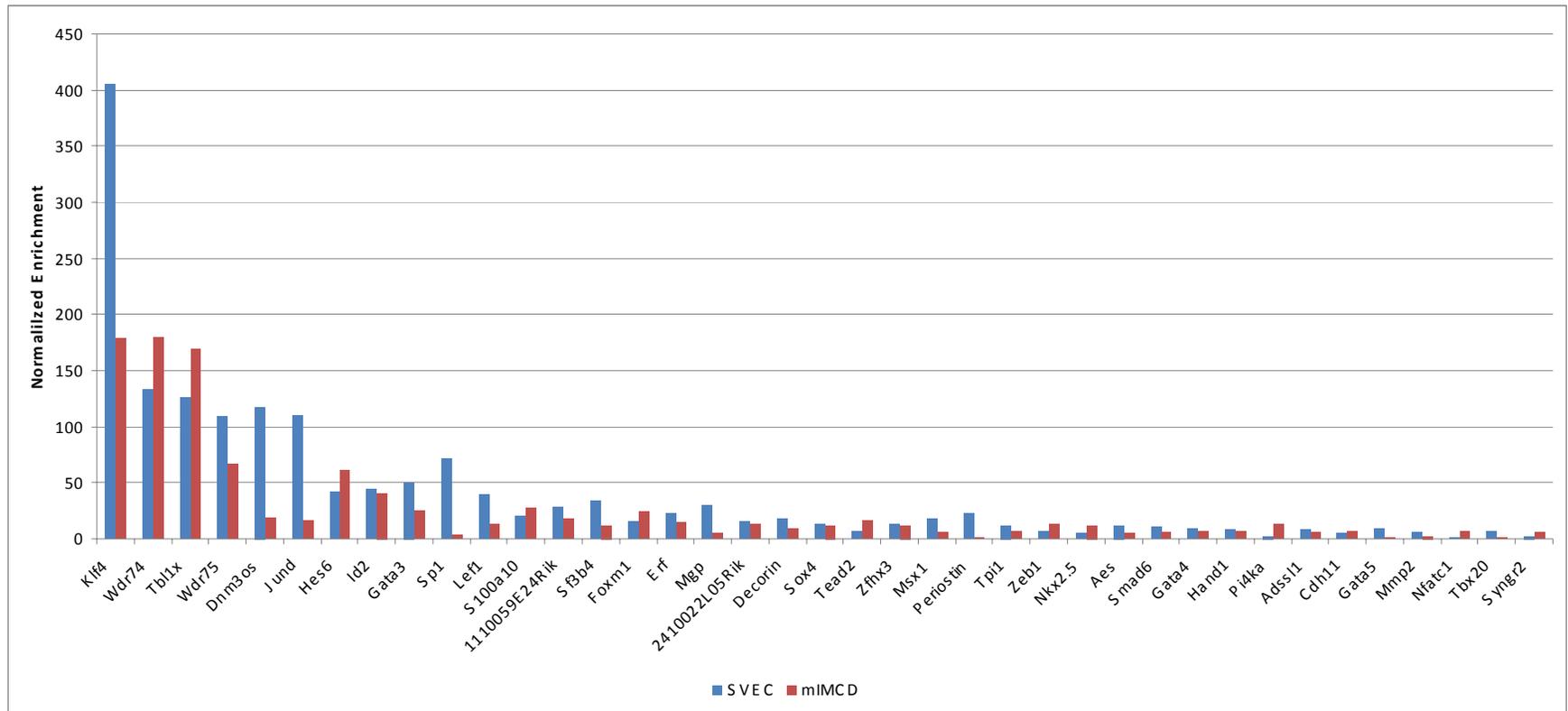
| Gene symbol | Gene name | Accession | Twist1 enr. | Temporal patterns | E9 | E10 | E11 | E12 |
|--|---|--------------|-------------|-------------------|-------|------|-------|-------|
| A. Genes up-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| <i>Mt2</i> | Metallothionein 2 | NM_008630 | 69.9 | Peak at E9 | 32.2 | 4.3 | 27.8 | 2.9 |
| <i>Adssl1</i> | Adenylosuccinate synthetase like 1 | NM_007421 | 10.2 | Peak at E9 | 6.4 | 1.7 | 2.3 | 2.6 |
| <i>Uqcr</i> | Ubiquinol-cytochrome c reductase (6.4kD) subunit | NM_025650 | 19.8 | Peak at E9 | 194.8 | 75.2 | 79.3 | 169.8 |
| <i>Wdr74</i> | WD repeat domain 74 | NM_134139 | 17.1 | Peak at E9 | 6.4 | 2 | 5.2 | 3.5 |
| <i>2900062L11Rik</i> | RIKEN cDNA 2900062L11 gene | NR_003642 | 49.8 | Peak at E11 | 1.6 | 2.3 | 4.6 | 1.3 |
| <i>Pi4ka</i> | Phosphatidylinositol 4-kinase, catalytic, alpha polypeptide | NM_001001983 | 13.9 | Peak at E11 | 1.3 | 1.3 | 2.9 | 0.6 |
| <i>Tpi1</i> | Triosephosphate isomerase 1 | NM_009415 | 20.7 | Peak at E11 | 71.3 | 41.7 | 121.3 | 60.7 |
| <i>Sf3b4</i> | Splicing factor 3b, subunit 4 | NM_153053 | 17.4 | Peak at E11 | 2.9 | 3.3 | 7.5 | 1.3 |
| <i>Wdr75</i> | WD repeat domain 75 | NM_028599 | 12.9 | Peak at E11 | 1.9 | 1.3 | 3.8 | 1 |
| <i>Zc3hcl1</i> | Zinc finger, C3HC type 1 | NM_172735 | 10.6 | Peak at E11 | 2.9 | 1 | 6.7 | 1.9 |
| <i>Slc4a1</i> | Solute carrier family 4 (anion exchanger), member 1 | NM_011403 | 24.8 | Peak at E11 | 1.3 | 0 | 7.8 | 2.6 |
| <i>1110059E24Rik</i> | RIKEN cDNA 1110059E24 gene | NM_025423 | 14 | Peak at E11 | 0 | 0 | 1.4 | 0 |
| B. Genes down-regulated in <i>Twist1</i> null AVC | | | | | | | | |
| <i>Rn18s</i> | 18S RNA | NR_003278 | 0.021 | Peak at E9 | 182.7 | 71.9 | 82.2 | 63 |
| <i>2410022L05Rik</i> | RIKEN cDNA 2410022L05 gene | NM_025556 | 0.093 | Peak at E9 | 15.9 | 7.3 | 3.5 | 3.8 |
| <i>LOC236598</i> | 28S ribosomal RNA | NR_003279 | 0.027 | Peak at E9 | 102.3 | 74.9 | 101 | 74.9 |
| <i>Ramp1</i> | Receptor (calcitonin) activity modifying protein 1 | NM_016894 | 0.071 | Peak at E9 | 5.6 | 1.7 | 2.2 | 5.4 |
| <i>Ntm</i> | Neurotrimin | NM_172290 | 0.024 | Peak at E10 | 3.2 | 5.6 | 2 | 2.6 |
| <i>Trim16</i> | Tripartite motif-containing 16 | NM_053169 | 0.07 | Peak at E10 | 0.4 | 1.7 | 0 | 1.7 |

| Gene symbol | Gene name | Accession | Twist1 enr. | Temporal patterns | E9 | E10 | E11 | E12 |
|----------------------|---|--------------|-------------|-------------------|------|------|-----|------|
| <i>Col9a2</i> | Collagen, type IX, alpha 2 | NM_007741 | 0.047 | Peak at E10 | 0 | 2.5 | 0.7 | 1.7 |
| <i>Papss2</i> | 3'-phosphoadenosine 5'-phosphosulfate synthase 2 | NM_011864 | 0.055 | Peak at E10 | 0.4 | 8.7 | 4.1 | 3.7 |
| <i>5031439G07Rik</i> | RIKEN cDNA 5031439G07 gene | NM_001033273 | 0.059 | Peak at E10 | 0 | 2.5 | 0.7 | 1.2 |
| <i>Col9a3</i> | Collagen, type IX, alpha 3 | NM_009936 | 0.078 | Peak at E10 | 0.4 | 7.5 | 3.4 | 2.5 |
| <i>Hnrnpul2</i> | Heterogeneous nuclear ribonucleoprotein U-like 2 | NM_001081196 | 0.08 | Peak at E10 | 0.4 | 1.7 | 0 | 0 |
| <i>Aldh2</i> | Aldehyde dehydrogenase 2, mitochondrial | NM_009656 | 0.089 | Peak at E10 | 0 | 5.8 | 2.2 | 2.5 |
| <i>Crebbp</i> | CREB binding protein | NM_001025432 | 0.089 | Peak at E10 | 0.4 | 2.1 | 0 | 0.4 |
| <i>Pqlc2</i> | PQ loop repeat containing 2 | NM_145384 | 0.099 | Peak at E10 | 0 | 1.3 | 0 | 0 |
| <i>Syng2</i> | Synaptogyrin 2 | NM_009304 | 0.054 | Peak at E10 | 3.2 | 7.9 | 3 | 7 |
| <i>Pigs</i> | Phosphatidylinositol glycan anchor biosynthesis, class S | NM_201406 | 0.067 | Peak at E10 | 2.0 | 4.6 | 1.1 | 3.7 |
| <i>Taf15</i> | TAF15 RNA polymerase II, TATA box binding protein (TBP)-associated factor | NM_027427 | 0.074 | Peak at E10 | 2.8 | 5.4 | 3.7 | 0.8 |
| <i>Unc13b</i> | Unc-13 homolog B | NM_001081413 | 0.098 | Peak at E10 | 0.4 | 2.5 | 1.5 | 1.7 |
| <i>Chd8</i> | Chromodomain helicase DNA binding protein 8 | NM_201637 | 0.03 | Peak at E11 | 0.4 | 2.1 | 4.5 | 1.7 |
| <i>Fmod</i> | Fibromodulin | NM_021355 | 0.096 | Peak at E11 | 0.4 | 2.1 | 4.5 | 2.5 |
| <i>Dcn</i> | Decorin | NM_007833 | 0.078 | Peak at E11 | 0.8 | 4.6 | 6.3 | 5.4 |
| <i>Kcne1l</i> | Potassium voltage-gated channel, Isk-related family, member 1-like | NM_021487 | 0.066 | Peak at E11 | 0.4 | 0 | 1.9 | 0.8 |
| <i>S100a10</i> | S100 calcium binding protein A10 (calpactin) | NM_009112 | 0.092 | Peak at E12 | 16.2 | 15.6 | 9.3 | 16.9 |
| <i>Mgp</i> | Matrix Gla protein | NM_008597 | 0.038 | Peak at E12 | 0.4 | 1.3 | 4.5 | 31.9 |

| Gene symbol | Gene name | Accession | Twist1 enr. | Temporal patterns | E9 | E10 | E11 | E12 |
|--------------------|--|------------------|--------------------|--------------------------|-----------|------------|------------|------------|
| <i>Pf4</i> | Platelet factor 4 | NM_019932 | 0.039 | Peak at E12 | 1.2 | 2.1 | 6 | 10.8 |
| <i>Gp5</i> | Glycoprotein 5 (platelet) | NM_008148 | 0.072 | Peak at E12 | 0 | 0.8 | 1.1 | 2.1 |
| <i>Bgn</i> | Biglycan | NM_007542 | 0.092 | Peak at E12 | 0.8 | 6.7 | 4.8 | 15.3 |
| <i>Tnfrsf12a</i> | Tumor necrosis factor receptor superfamily, member 12a | NM_013749 | 0.098 | Peak at E12 | 0.4 | 1.7 | 1.9 | 3.3 |

Appendix XIII. Chromatin immunoprecipitation

Myc-tagged *Twist1* was overexpressed in mIMCD and SVEC cells, and immunoprecipitated with anti-myc antibody. Enrichment was calculated as 2 to the power of the cycle threshold (cT) difference between the IgG immunoprecipitated sample and the anti-myc antibody immunoprecipitated sample. Enrichment was normalized to lowest value.



Appendix XIV. Animal care certificate



THE UNIVERSITY OF BRITISH COLUMBIA

ANIMAL CARE CERTIFICATE BREEDING PROGRAMS

Application Number: A08-0703

Investigator or Course Director: [Pamela Hoodless](#)

Department: Medical Genetics

Animals:

Mice C57BL6/J 250
Mice Foxh1tmlJlw 400
Mice B6.129-Twist1tmlBlm/J 50
Mice R26R-YFP 150
Mice Foxa2tmlDnl 100
Mice Pdx1-GFP 170
Mice Hipp1 100
Mice R26R-LacZ 150
Mice Ngn-Cre-ER 150
Mice TGIF 50
Mice Ngn-cre 150
Mice ICR 250

Approval Date: December 9, 2009

Funding Sources:

Funding Agency: Canadian Institutes of Health Research (CIHR)
Funding Title: Formation and patterning of the definitive endoderm

Funding Agency: Canadian Institutes of Health Research (CIHR)
Funding Title: Characterization of regulatory regions, modules and elements in mammalian genomes

Funding Agency: Heart and Stroke Foundation of British Columbia and Yukon
Funding Title: Valve formation in the mouse atrioventricular canal

Funding Agency: Genome Canada

| | |
|------------------------|--|
| Funding Title: | Dissecting Gene Expression Networks in Mammalian Organogenesis |
| Funding Agency: | Genome British Columbia |
| Funding Title: | Dissecting Gene Expression Networks in Mammalian Organogenesis |
| Unfunded title: | N/A |

The Animal Care Committee has examined and approved the use of animals for the above breeding program.

This certificate is valid for one year from the above approval date provided there is no change in the experimental procedures. Annual review is required by the CCAC and some granting agencies.

A copy of this certificate must be displayed in your animal facility.

Office of Research Services and Administration
102, 6190 Agromony Road, Vancouver, BC V6T 1Z3
Phone: 604-827-5111 Fax: 604-822-5093