

Molecular characterization of cellular subgroups in the caudal end of the mouse embryo

Inaugural-Dissertation
to obtain the academic degree
Doctor rerum naturalium (Dr. rer. nat.)

submitted to the Department of Biology, Chemistry and
Pharmacy of Freie Universität Berlin

by
Constanze Nandy
from Leonberg

2015

This work was performed at the Max-Planck Institute for Molecular Genetics in Berlin between January 2009 and February 2015, with interruption for maternity leave under the supervision of Prof. Dr. Bernhard G. Herrmann

1st Reviewer: **Prof. Dr. Bernhard G. Herrmann**

Max-Planck Institute for Molecular Genetics, Ihnestr. 73, 14195 Berlin

2nd Reviewer: **Prof. Dr. Stephan Sigrist**

Institute of Biology Freie Universität Berlin Takustr. 6, 14195 Berlin

date of defense: 22.06.2015

Meiner Familie.

Danksagung

Zuerst möchte ich Prof. Bernhard G. Herrmann für die Möglichkeit danken, eine Doktorarbeit in seiner Arbeitsgruppe für Entwicklungsgenetik anzufertigen. Seine Tür stand zu jeder Zeit offen, um Fragen und Probleme zu erörtern. Prof. Stephan Sigrist danke ich für die Übernahme der Position als Gutachter an der Freien Universität Berlin.

Für die Durchführung aller Tetraploid Aggregationen meiner ES-Zellklone bin ich Dr. Lars Wittler, Karol Macura und den übrigen Kollegen der Transgene Abteilung sehr dankbar. Bei Dr. Phillip Grote bedanke ich mich für die Hilfe bei der Zusammenstellung des Konstruktes für die Zellstammbaumanalyse, sowie für die Beantwortung vieler Fragen im Laboralltag. Für die Durchführung der Durchflusszytometrie und seine große Unterstützung bei der RNA-Sequenzierung und Auswertung der Daten, sowie überlassen des Datensatzes der dreifach positiven ES-Zelllinie möchte ich Dr. Frederic Koch ganz besonders danken. Jinhua Liu danke ich sehr für die bioinformatische Auswertung der RNA-Sequenzierung. Bei der Tierpflegerin Christin Franke möchte ich mich für die gute Betreuung der Tiere und die zuverlässige Durchführung der Tamoxifenapplikationen bedanken. Mein Dank geht auch an alle meine Kollegen, die mich während meiner Schwangerschaft unterstützt haben und all jene Arbeitsschritte übernommen haben, die ich während dieser Zeit nicht ausführen durfte. Es ist mir eine Freude, mich bei meinen derzeitigen und ehemaligen Laborkollegen Andrea, Gaby, Lars, Matthias, Shini und Tracie für eine stets angenehme, entspannte und motivierende Arbeitsatmosphäre bedanken. Auch bei allen derzeitigen und ehemaligen Kollegen der ganzen Abteilung bedanke ich mich ganz herzlich für das angenehme Arbeitsklima und die stete Offenheit für auftauchende Fragen oder Probleme. Tracie gebührt meine große Dankbarkeit für die sprachliche Korrektur meiner Arbeit und für viele hilfreiche Kommentare und Diskussionen. Vielen lieben Dank auch für die vielen gemeinsamen Unternehmungen und abendlichen Ablenkungen vom Laboralltag.

Meiner Familie möchte ich für deren beständige Unterstützung während meiner gesamten Promotion bedanken. Mein ganz besonderer Dank gilt meinem Ehemann Tim David, der mir den nötigen Rückhalt, die Stärke und Unterstützung für die Durchführung meiner Arbeit gegeben hat. Meinem lieben Sohn Emil Malo danke ich für die Freude und die Ablenkung vom Alltag, die unser Familienleben auf ganz besonders schöne Weise bereichern.

Contents

1	Introduction	1
1.1	Embryonic development	1
1.1.1	Early embryonic development	1
1.1.2	Gastrulation	2
1.2	Mesoderm patterning	3
1.2.1	Segmentation	4
1.3	Development of endodermal structures	6
1.3.1	Signaling network in embryonic development	7
1.4	Ectoderm development	7
1.4.1	Neural crest development	7
1.4.2	Role of the ventral ectodermal ridge in caudal end development . .	9
1.5	Marker genes-of-interest and their expression profile in the caudal end	10
1.5.1	The caudal end at 9.0 dpc	10
1.5.2	Homeodomain transcriptionfactor MIXL1	12
1.5.3	Homeodomain transcription factor Tlx2	13
1.5.4	Peptide hormone Secretin	13
1.5.5	Paired box transcription factor Pax3	14
1.6	Lineage tracing in the mouse embryo	14
2	Results	17
2.1	Development of mesoderm in the caudal end of the mouse embryo	17
2.2	Generating a Lineage Tracing Construct for <i>in vivo</i> analysis	17
2.2.1	Dre versus Cre recombinase	20
2.3	Lineage tracing driven by the Mixl1 homeodomain transcription factor . .	22
2.4	Lineage tracing driven by the Tlx2 homeodomain transcription factor . .	25
2.5	Lineage tracing driven by the peptide hormone Secretin (<i>Sct</i>)	29
2.6	Lineage tracing driven by the transcription factor <i>Pax3</i>	31
2.7	Transcriptome analysis of caudal end subregions	35
2.7.1	Analysis of canonical pathways in subregions of the caudal end .	36
2.7.2	Clustering of differentially expressed genes from RNA sequencing analysis	40
2.7.3	Analysis of clustered genes for biological processes in subregions of the caudal end	40

2.7.4	Markers for paraxial and lateral mesoderm are expressed in the <i>Mixl1</i> population	46
2.7.5	<i>Noto</i> expression in <i>Mixl1</i> expressing cells in the caudal end	49
2.7.6	<i>Foxa2/T</i> expressing cells representing endoderm	49
2.7.7	Genes expressed in <i>Tlx2</i> expressing cells are involved in endoderm development	50
2.7.8	Genes expressed in <i>Pax3</i> expressing cells are involved in neural crest development	53
3	Discussion	57
3.1	<i>Tlx2</i> descendants contribute to hindgut development	57
3.1.1	Involvement of <i>Tlx2</i> daughter cells in endoderm development . .	58
3.1.2	<i>Tlx2</i> daughter cells do not contribute to the development of the enteric nervous system	61
3.2	<i>Noto</i> expression in <i>Mixl1</i> expressing cells	62
3.3	<i>Mixl1</i> expressing cells contribute to development of paraxial and lateral mesoderm	62
3.4	<i>Mixl1</i> expressing cells and their contribution to development	64
3.5	<i>Pax3</i> descendants contribute to neural crest development	65
3.5.1	<i>Pax3</i> subdomain: Origin of neural crest precursors?	67
3.6	Characteristics of <i>Sct</i> expressing cells	69
3.7	Conclusions	69
3.7.1	Outlook	70
4	Methods	71
4.1	DNA work and cloning	71
4.1.1	Polymerase Chain Reaction (PCR)	71
4.1.2	Cloning vectors and bacterial strains	71
4.1.3	Lineage tracing construct	71
4.1.4	RED/ET recombineering	72
4.1.5	Mini-prep of BAC DNA	72
4.2	Cell culture	73
4.2.1	Culturing embryonic stem cells	73
4.2.2	Electroporation of ES cells	73
4.2.3	Picking of ES cell clones	73
4.2.4	Induction of ES cells with 4-OH tamoxifen	73
4.2.5	X-Gal staining of ES cells	74
4.2.6	Genotyping ES cells by Southern blotting	74
4.2.7	Tetraploid complementation assay using transgenic ES cells . . .	75

Contents

4.3	Mouse work	75
4.3.1	Tamoxifen administration	75
4.4	Histology	75
4.4.1	Imaging of fluorescent embryos	75
4.4.2	WISH (Whole-mount <i>in situ</i> hybridization)	76
4.4.3	X-Gal staining of whole mount embryos	77
4.4.4	Paraffin embedding and sectioning of embryos	77
4.4.5	Counterstaining of sections with Eosin	77
4.5	FACS - Fluorescence activated cell sorting	78
4.5.1	Homogenization of caudal ends for FACS	78
4.5.2	Cell sorting	78
4.6	RNA work	78
4.6.1	Isolation of RNA	78
4.6.2	RNA library preparation	79
4.6.3	RNA sequencing	82
4.6.4	Bioinformatic data analysis	82
4.7	BAC clones used for creating transgenic ES cells and embryos	83
4.8	Buffers, solutions and media	83
5	Summary	87
6	Zusammenfassung	89
References		95
7	Supplementary	111

Chapter 1

Introduction

1.1 Embryonic development

1.1.1 Early embryonic development

About 95% of the mouse and human genomes are identical, and the same also applies to their anatomy and physiology. Since mouse husbandry is relatively affordable, and the generation period is short enough for scientific experiments within a manageable time-frame, *mus musculus* is a commonly used model organism for studying developmental biology.

Fertilization of the oocyte is completed after the sperm enters the cell. About 24 hours after fertilization the first cleavage takes place while the zygote is moving through the oviduct to the uterus. The DNA-methylation pattern of the oocyte is characteristic for totipotent and pluripotent cells at this time-point (Hales et al., 2011). The first cell movements and differentiation starts in the morula at 16 cell stage and leads to the blastula stage (Figure 1.1). This is when *Cdx2* is expressed in the outer cell mass, the trophectoderm, whereas the inner cell mass (ICM) expresses the pluripotency marker *Oct3/4* (*Pou5F1*) (Niwa et al., 2005). Cells located around the outside of the embryo generate the trophoblast, whereas internal cells belong to the inner cell mass (ICM). To implant into the uterine wall, the blastocyst must hatch from the zona pellucida. Now trophoblast integrin binds to uterine collagen and the blastocyst invades into the endometrium of the uterine wall (Wang and Dey, 2006).

After implantation, at around 5 days post-coitum (dpc), the chorion and decidua, the embryonal and maternal parts of the placenta, respectively, begin to form. The placenta nourishes the embryo until birth. The first segregation of cells occurs by the formation of two cell layers, the primitive endoderm and the ICM. The future posterior end of the embryo is marked by the primitive endoderm (Tam and Behringer, 1997b). The visceral endoderm is made up of primitive endodermal cells which contact the epiblast and will give rise to the amnion. As soon as the amnion is developed, it is also filled with amniotic fluid to protect the embryo (Figure 1.1) (Stern and Downs, 2012).

The epiblast is the origin of the embryo proper, and all resulting tissues and structures.

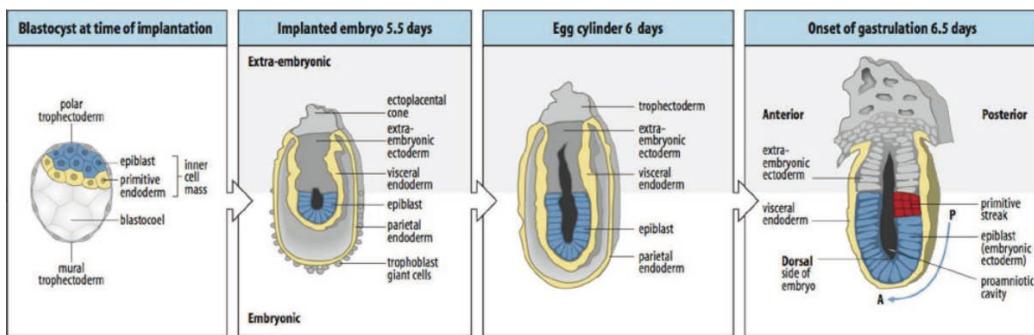


Figure 1.1: Gastrulation and formation of the egg cylinder (Gilbert, 2014)

1.1.2 Gastrulation

Between 5 and 6 dpc the structure of the embryo changes completely, and a cylindric structure emerges. The cup shaped embryo comprises two cell layers, the inner epiblast layer and an outer visceral endoderm, while the inner cavity forms and fills with fluid (Figure 1.1) (Tam and Behringer, 1997b). To set the structure for the body plan all vertebrates undergo the process of gastrulation, which leads to the formation of the three germ layers: endoderm, mesoderm, and ectderm. These three layers will go on to form all body structures. The future posterior end of the embryo is the designated start site of gastrulation and also contains the node, from which the notochord arises. A transient structure, called the primitive streak, appears here and epiblast cells ingress through the primitive streak to build up mesendoderm, the precursor of the mesodermal and endodermal layers (Bellairs, 1986; Tam and Behringer, 1997a). The process of epithelial-mesenchymal transition (EMT) is essential for gastrulation to proceed, as E-cadherin junctions break down to allow the cells to migrate and become mesodermal (Burdsal et al., 1993). For induction and maintenance of the primitive streak, *Nodal*, a member of the TGF β -family but also Wnt and Fgf signaling are necessary (Conlon et al., 1994; Bertocchini et al., 2004; Arkell et al., 2013).

As soon as the structure of the primitive streak is apparent, the anterior and posterior ends of the embryos have to be patterned. Anterior-posterior patterning during gastrulation is performed by two signaling centers. The node is mandatory for proper left-right patterning of the embryo, whereas another signaling center the anterior visceral endoderm (AVE), is important for anterior-posterior axis development. The AVE and the node act in concert to determine the anterior region of the developing embryo (Beddington and Robertson, 1999). In addition, many transcription factors and other genes are expressed in the node, conducting organization and induction of embryonal structures (Davidson and Tam, 2000). Further patterning of the anterior-posterior axis is accomplished by the expression of the Hox genes, which specify the body plan of the embryo along the anterior-posterior axis

1.2 Mesoderm patterning

(Hunt and Krumlauf, 1992).

The ectoderm is the exterior germ layer of the embryo structure. It gives rise to neural tissues, such as the neural plate, which is the origin of the neural crest (NC) and the neural tube-anlagen of the spinal cord. Furthermore, a part of this germ layer will form epidermis. The endoderm becomes implemented in the definitive endoderm and gives rise to the linings of all ducts in the developing embryo, such as the lining of intestine and lungs. The notochord is a midline signaling center which is important for patterning of the neural tube and to assign the body axis. The notochord forms from cells originating from the node (Sulik et al., 1994).

1.2 Mesoderm patterning

Following gastrulation, the mesoderm undergoes a patterning process, in which paraxial mesoderm, lateral plate mesoderm and intermediate mesoderm are generated. Paraxial mesoderm is generated on both sides of the notochord (Tam and Beddington, 1987). As depicted in Figure 1.2, paraxial mesoderm gives rise to somitic mesoderm, where the anterior part of the paraxial mesoderm is source for the head mesoderm. Intermediate mesoderm develops between paraxial and lateral plate mesoderm, and parts of the urogenital system originate from this type of mesoderm. Moreover, the lateral plate mesoderm gives rise to heart and all blood vessels.

BMP signaling and transcription factors from the Forkhead box family are important participants during mesoderm patterning (Wilm et al., 2004). *T* (*Brachyury*) is a transcription factor which is crucial for mesoderm differentiation and axial elongation (Kispert and Herrmann, 1994), and another T-box transcription factor, *Tbx6*, is important for maintaining the presomitic mesoderm (PSM) (Chapman et al., 1996). Between 7.0 and 7.5 dpc, both *T* and *Tbx6* are expressed in the primitive streak. Thereafter *T* is also expressed in the notochord, while *Tbx6* is expressed in the presomitic mesoderm (Wilkinson et al., 1990). Loss of *Tbx6* results in improper specification of paraxial mesoderm, accompanied by the generation of ectopic neural structures (Chapman and Papaioannou, 1998). During segmentation *Tbx6* interacts with genes relevant to the segmentation process, such as *Dll1*, a downstream target of the Notch signaling pathway (White et al., 2003). All developmental processes involving migration of cells, proliferation, and differentiation are regulated by signaling molecules belonging to various signaling pathways. Fibroblast growth factors (FGF) are responsible for cell movement and mesoderm patterning, by regulating other genes important for cell determination, such as *T* (*Brachyury*) or *Tbx6*, reviewed in (Arnold and Robertson, 2009).

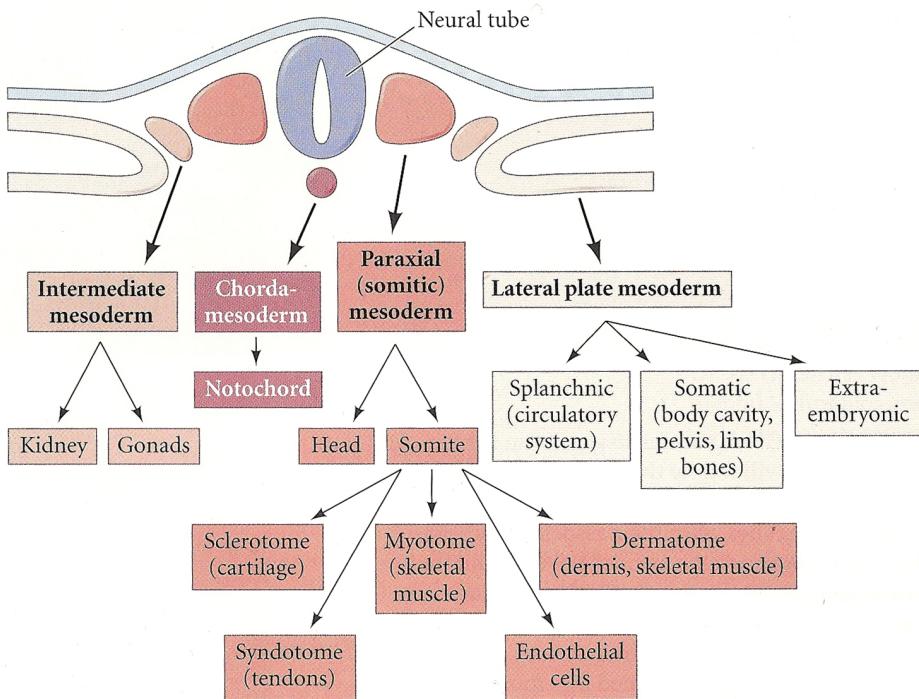


Figure 1.2: Scheme of mesoderm derivatives after mesoderm patterning. Figure modified from (Gilbert, 2014)

1.2.1 Segmentation

At late gastrula stage the embryo is patterned antero-posteriorly by morphogen gradients and resulting region-specific gene expression. *Bmps*, *Wnts*, *Fgfs*, and *Nodal* are expressed in the posterior region of the embryo, while in the anterior half of the embryo WNT and BMP antagonists are expressed. Thus, the highest concentrations of BMP, WNT, and FGF signaling is in the posterior end (Aulehla and Pourquié, 2010). A gradient of *Fgf8* which is only transcribed in the posterior tip of the embryo is established by mRNA decay in newly formed tissue and so converted into a protein gradient (Dubrulle and Pourquié, 2004). Also a retinoic acid (RA) gradient is generated by a balance between decay and synthesis during embryonic development (Sakai, 2001). This RA gradient is important for expression of the *Hox* genes, which are crucial for designating vertebral identity (Allan et al., 2001).

The PSM arises from the primitive streak at the posterior end of the embryo. From the PSM located on both sides of the developing neural tube, somites will form to give rise to important structures: the precursors of vertebrae and rib cage, musculature of the trunk, and skeletal muscle (Christ et al., 2007). After a pair of somites is formed, they bud off synchronously from the PSM in a periodical manner. The periodicity of somite formation and the somite number is specific for each organism (Tam, 1981). Somites begin as bilateral blocks of epithelia which are generated from the PSM (Figure 1.3). Maturation of somites

1.2 Mesoderm patterning

into different characteristic regions starts once somites are formed.

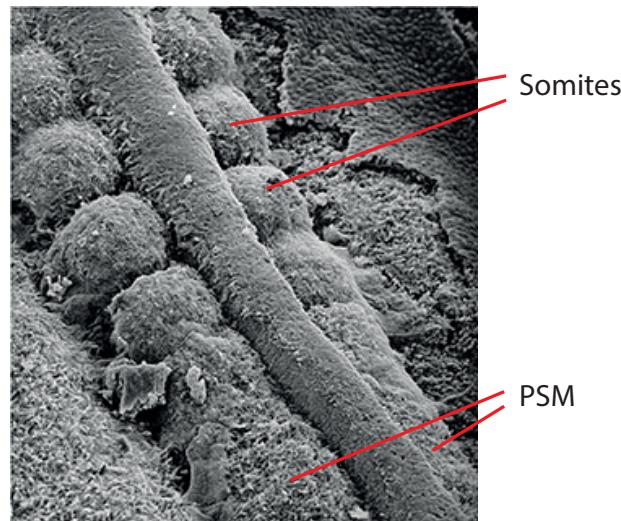


Figure 1.3: Paraxial mesoderm, somites and neural tube after removal of ectoderm. Figure modified from (Gilbert, 2014)

As new somites are being formed, the elongation of the embryo takes place at the same time, while the presomitic mesoderm proliferates.

For proper formation of somites, several signaling pathways are crucial. Through the interaction of signaling gradients and oscillating genes, somites can form periodically from the PSM. Various genes belonging to the WNT, Notch and FGF signaling pathways have been found to be expressed in an oscillating manner (Aulehla and Herrmann, 2004; Dequéant and Pourquié, 2008). Interactions between all of the genes and pathways during segmentation is very complex, and hence not all details are investigated so far. On general, *Fgf8* is transcribed in the caudal end of the embryo. By RNA decay and translation of *Fgf8*, a protein gradient is established. Expression of *Fgf8* is regulated by *Wnt3a*, which is only expressed by cells positioned at the caudal tip of the tail bud. While the caudal end is elongating, cells are gradually positioned more and more anteriorly, and thus cease *Wnt3a* transcription. In this manner a *Wnt3a* gradient is generated. Cells which fall beneath a certain morphogen gradient-threshold, begin the process of differentiating into somites. As the gradient drives the oscillator, cells exposed to a gradient level higher than the threshold activate expression of oscillating genes from the Notch and WNT signaling pathways. For example, *Axin2* which is a negative regulator of the WNT signaling pathway, and *Lfng* from the Notch signaling pathway, are both expressed periodically. Moreover, genes from WNT and Notch signaling pathways oscillate in an alternating fashion. New somites are formed at the segment boundary, which is defined as the border between cells expressing

cycling genes, and those which have these genes switched off (Aulehla and Herrmann, 2004).

1.3 Development of endodermal structures

Definitive endoderm is the section of the endoderm germ layer which is essential for development of intestinal organs like liver, pancreas and gut. During gastrulation, cells ingress at the primitive streak and migrate towards visceral endoderm. Sources of endoderm differ depending on their local position in the embryo. Cells which ingress first during this process will form midgut and foregut, whereas cells migrating later are included into hindgut (Kimura et al., 2006; Zorn and Wells, 2009).

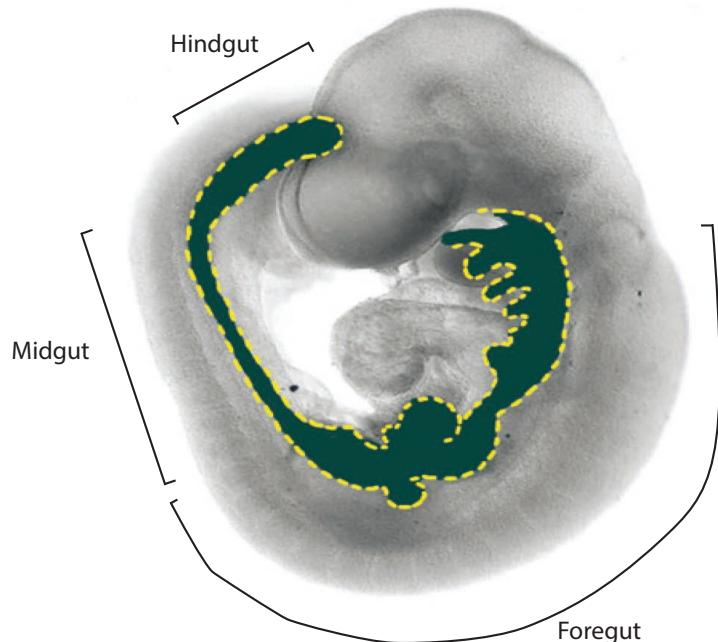


Figure 1.4: 9.5 dpc embryo with gut tube marked in green. Foregut, midgut and hindgut sections are indicated by lines. Figure modified from (Zorn and Wells, 2009)

Determining signaling factors like Nodal and other downstream genes of the TGF β pathway are essential for formation of the endoderm layer from the mesendoderm. Enclosed by mesoderm, the gut tube forms into different sections: foregut, midgut, and hindgut (Zorn and Wells, 2009), which can be distinguished by differential gene expression. The foregut will give rise to esophagus, thyroid, lung, liver, pancreas and stomach, while the midgut forms small intestine, and the hindgut generates the colon and is involved in urogenital development (Figure 1.4) (Kimura et al., 2006). The hindgut is formed by precursor cells from the posterior endoderm, clustered around the anterior primitive streak, which expand

1.4 Ectoderm development

in a rostral to caudal direction (Kimura et al., 2006; Franklin et al., 2008).

Foxa2 and *Sox17* mark the endothelial cell layer along the length of the gut tube (Burtscher and Lickert, 2009; Burtscher et al., 2012). When hindgut and foregut pockets are forming, morphogenesis of the gut starts. Here, *Foxa2* is important for development of the foregut (Chawengsaksophak et al., 2004).

1.3.1 Signaling network in embryonic development

Signaling networks and their connection with each other are crucial for proper formation of embryonal structures and beyond. BMP, TGF β , Wnt and Fgf signaling are major pathways during embryonal development.

BMPs are members of the TGF β -family, and involved in diverse processes such as axis establishing, cell growth and differentiation, apoptosis and formation of the primitive streak (Arnold and Robertson, 2009; Heldin et al., 1997; Conlon et al., 1994). Organization of diverse functions is handled by binding to a variety of receptor complexes (Arnold and Robertson, 2009; Heldin et al., 1997). During gastrulation and mesoderm formation members of the TGF β family are crucial, as it is shown by depletion of *Bmp4*, where most embryos are stalled in egg cylinder stage (Winnier et al., 1995).

Members of the canonical Wnt signaling pathway are involved in diverse functions during embryonal development. Wnt signaling is participating in axis formation, specification of cell fate or somitogenesis. The Wnt signaling gradient, which is established during axis elongation and PSM patterning is crucial for these processes.

FGF signaling is important for mesoderm formation, left-right axis formation, posterior PSM fate and more than that.

All pathways are connected with each other and conduct the interaction for the proper procedure of developmental processes. Especially Wnt and Fgf signaling are important for the segmentation clock (Aulehla and Pourquié, 2010).

1.4 Ectoderm development

1.4.1 Neural crest development

Neural crest (NC) cells were first described by Wilhelm His in 1868 (His, 1868). He discovered cells, migrating from the neural primordial aggregate, to form dorsal root ganglia in chick (Hall, 2008). Depending on the species, neural tube closure occurs before or after the migration of the NC cells. NC cells delaminate from the border between neural plate and epidermis, the neural folds and are also a part of the neural anlagen (Figure 1.5) (Dupin and Le Douarin, 2014). After the NC cells delaminate, they migrate very long distances through the whole embryo to their final destinations. NC cells start to migrate at

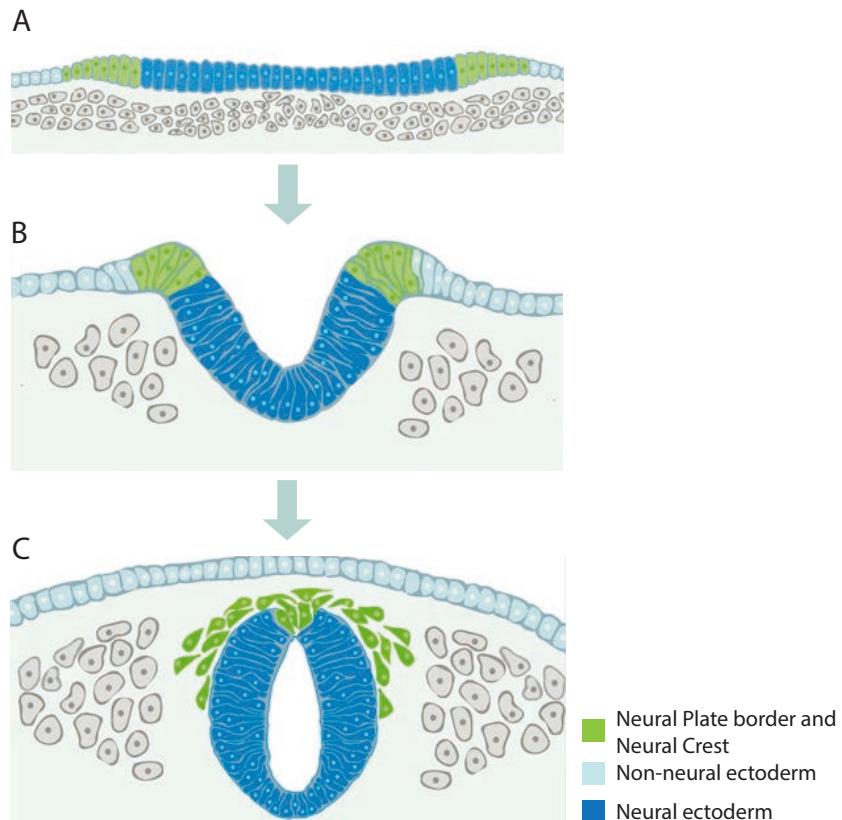


Figure 1.5: Development of Neural Crest cells

From A to C developing embryonic dorsal ectoderm with specified neural plate border in A, pre-migratory neural crest cells in B and migrating neural crest in C. Figure from (Simões Costa and Bronner, 2013)

gastrulation and continue until late organogenesis. NC cells harbour a great differential potential, as more than 100 different cell types arise from the NC lineage, meaning that NC cells are highly pluripotent (Baker and Bronner-Fraser, 1997), and are often referred to as the “4th germ layer”. To understand the role of NC cells better, fundamental experiments on sea urchin were done in 1950 (Horstadius, 1950). In these experiments, prospective NC tissue was transplanted from one amphibian embryo to another. Grafted NC cells were shown to migrate through the embryo to the respective destination in a regional manner. In experiments with vital dye, which was injected into the embryo, the regional connection between origin and destination was verified. Later in the 1960s and 1970s, Nicole Le Douarin used a setup including Japanese quail-chick chimera, since Japanese quail cells can be distinguished from chick cells using feulgen staining. It was stated that NC cells can be traced, and derivatives can be identified at their final destination. With these experiments, a map of populations and their destinations of NC cells in higher organisms could be drawn (Le Lievre and Le Douarin, 1975; Douarin and Kalcheim, 1999).

1.4 Ectoderm development

Tissues which emerge from NC cells are diverse and include neurons, such as sensory nerves of the peripheral nervous system, sympathetic and enteric ganglia, glia cells, like nerve schwann cells and satellite cells, pigment cells (melanocytes) in the epidermis, parts of the mesectoderm for the development of head structures like cartilage and bone, connective tissue, vascular smooth muscle cells and adipocytes (Baker and Bronner-Fraser, 1997; Douarin and Kalcheim, 1999).

The first step in the development of NC cells is specification of the neural plate border, which is the link between non-neural ectoderm and neural ectoderm (Figure 1.5 A). Specification of the neural plate border is assigned by genes like *Msx*, *Pax3*, *Pax7*, *Zic1*, *Dlx3* and *Dlx5* (Simões Costa and Bronner, 2013; Sauka-Spengler and Bronner-Fraser, 2008). *Pax3* has been shown to be crucial for specification of NC in *Xenopus* (Hong and Saint-Jeannet, 2007). *Dlx5* is the one of the earliest genes which designates the neural plate border in mice (Gammill and Bronner-Fraser, 2003). After induction of NC cells, these genes are usually downregulated.

Specification and induction of NC fate is influenced by genes like *Snai1*, *Foxd3*, *Twist*, *Id*, *c-Myc*, *AP-2*, or *Sox8*, *Sox9* and *Sox10* (Figure 1.5 B). *Sox10* is expressed when migration starts and is responsible for maintenance of stemness, and also cooperates with *Pax3* (Gammill and Bronner-Fraser, 2003). In *Xenopus*, *Pax3* and *Zic1* were identified as upstream regulators of neural crest specifiers like *Snai1*, *Foxd3*, *Twist1* and *Tfap2b* (Plouhinec et al., 2014). A major task of the specification process involves regulation of EMT, as a requirement for later delamination and migration. The source of signals to regulate NC cell migration includes paraxial mesoderm, non-neural ectoderm and neuroepithelium (Sauka-Spengler and Bronner-Fraser, 2008).

A large network of genes and pathways is involved in the specification process. WNT1 signaling, along with the BMP and FGF pathways, are involved in the gene regulatory network of NC formation. NC specifier genes control the expression of effector genes, which are responsible for various processes, such as EMT or migration. Specified NC are able to delaminate from the neural tube and migrate through the embryo to their target position (Figure 1.5 C) (Simões Costa and Bronner, 2013).

1.4.2 Role of the ventral ectodermal ridge in caudal end development

The ventral ectodermal ridge (VER) was first described in 1956 by Hans Grüneberg as a thickening of ectoderm on the ventral side of the tail bud, between anterior border of the tail bud mesoderm and the presomitic mesoderm containing region (Grüneberg, 1956). Between 9.0 dpc and 10.0 dpc the VER, which is formed by cells provided from the primitive streak, has its largest dimensions during development (Figure 1.6). In labeling

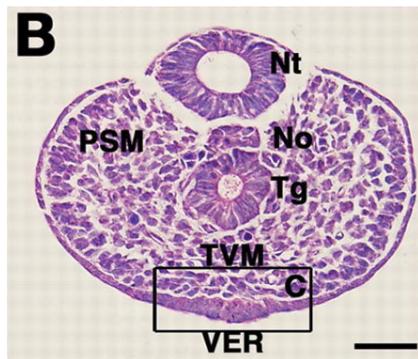


Figure 1.6: Ventral ectodermal ridge (VER)

Transversal section through 10.5 dpc mouse tail.

Figure from (Ohta et al., 2007)

experiments it could be shown that cells from the VER contribute to the dorsal midline. However, thickening of VER is very similar to the apical ectodermal ridge (AER), an important signaling center for growth and differentiation of limbs (Grüneberg, 1956; Goldman et al., 2000; Tam and Beddington, 1987). Due to the analogy of VER to AER, a role as a signaling center was assumed. VER ablation experiments suggest that the VER is involved in regulation of somitogenesis. Besides this, the VER also participates in tail elongation (Goldman et al., 2000). So far there are no details known about particular correlations of this function. Additionally, also a role in shutting down EMT is feasible (Ohta et al., 2007), since the VER is the structural remnant of the primitive streak.

1.5 Marker genes-of-interest and their expression profile in the caudal end

1.5.1 The caudal end at 9.0 dpc

During elongation of the body axis, the caudal end and its progenitor cells play a very important role in maintaining growth. The chordoneural hinge (CNH), a structure in the caudal end where the neural tube and notochord meet, is a source for different progenitor cells, contributing to notochord, neural tube and paraxial mesoderm (Cambray and Wilson, 2002). Experiments using tissue from the caudal end grafted to the kidney capsule of adult kidneys indicated the proliferative and differentiative capacity of the caudal end. Grafted tissue could grow into cells belonging to all three germ layers (Tam, 1984).

In order to prepare for later trunk and organ development, the mesoderm at the caudal end of the embryo undergoes patterning. Instructive signals from organizer cells as well as stem cells are important for the precise patterning and development of trunk and tail. Many different transcription factors and pathways orchestrating these developmental steps.

1.5 Marker genes-of-interest and their expression profile in the caudal end

During patterning of mesoderm several genes are very restricted in their expression patterns, like *Tbx6*, *Mixl1* and *Tlx2*. An important player in mesodermal patterning and determination of cell fate is *Tbx6* (Chapman and Papaioannou, 1998). *Tbx6* is responsible for the commitment of neural or mesodermal fate through interactions with *Sox2* (Li et al., 2002; Takemoto et al., 2011). At 9.5 dpc *Tbx6* is expressed in the presomitic mesoderm of the caudal end, which means it is expressed in the domain where *Mixl1*, *Tlx2* and *Sct* are also expressed.

To investigate processes in the mouse caudal end during axis elongation and trunk development, genes which exhibit very defined expression in the caudal end between 9.0 and 9.5 dpc were chosen (Figure 1.7). *Pax3*, involved in specifying of neural crest is expressed in the neural tube and in the extension of the neural tube around the tip of the caudal-most part of the caudal end. Expression of *Pax3* in the caudal end is seemingly redundant, because there is no neural structure present and the cells expressing *Pax3* in the caudal end are of mesodermal origin. The mesoderm patterning gene *Mixl1* is expressed adjacent to *Pax3* in the caudal end mesoderm and has been mentioned in the context of regulation of stem cell niches (Wolfe and Downs, 2014). *Tlx2* is expressed adjacent to the hindgut pocket and is participating in mesoderm patterning. Also expressed in the caudal end, but in the region of the VER is *Sct*. The function of *Sct* in the caudal end mesoderm is still unknown.

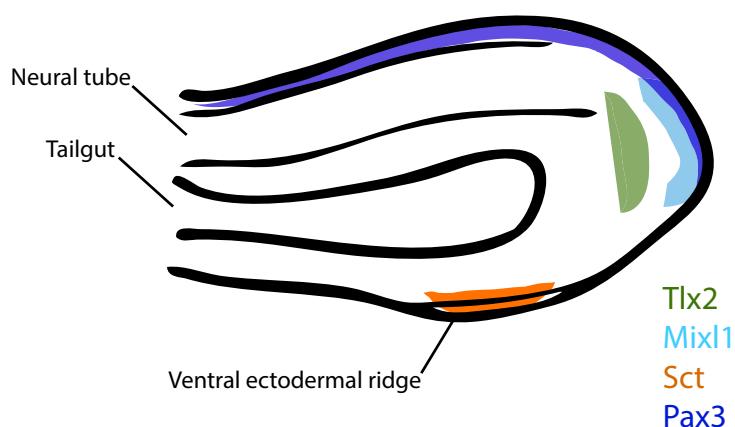


Figure 1.7: Expression domains of marker genes-of-interest. Adapted from (Robb and Tam, 2004)

Many genes identified in the caudal end show very distinct expression domains in different sub-regions of the caudal end. The genes used for our approach were chosen due to their clear and distinct expression patterns in the caudal end (Figure 1.7).

1.5.2 Homeodomain transcriptionfactor MIXL1

In 1989, the homeodomain transcription factor *Mixl1* (Mix1 homeobox-like 1) was first described in *Xenopus* as *Mix.1* (Rosa, 1989). Since the *Mix.1* gene contains a paired-type homeodomain structure in its sequence, a regulating function was obvious (Rosa, 1989). *Mix.1* belongs to a family of transcription factor genes in *Xenopus* with important impact on structuring during gastrulation and endoderm development. *Mix.1* expression in *Xenopus* is regulated by *activin*, *Vg1* and *BMP-4* (Mead et al., 1998).

In mouse, there is only one homolog, *Mixl1*, which is expressed between 5.5 dpc and 11.5 dpc. In the visceral endoderm it is expressed at 5.5 dpc, and one day later at 6.5 dpc the expression is limited to the border of embryonic and extra embryonic tissue, and in the developing mesoderm and primitive streak (Pearce and Evans, 1999). Expression in the posterior primitive streak and crown cells of the node can be observed at 8.5 dpc. From 9.5 dpc until 11.5 dpc, when the signal vanishes, expression can be detected in the tail bud (Pearce and Evans, 1999). *Mixl1* expression in 9.0 dpc embryos, as taken from our in-house MAMEP database (Berlin, Germany; <http://mamep.molgen.mpg.de>), depicts expression in the caudal end mesoderm (Figure 2.3).

To better understand the role of *Mixl1* during embryonic development, experiments using loss-of-function mutations were performed by Hart and colleagues (Hart et al., 2002). Homozygous *Mixl1* mutants died before 10.5 dpc, whereas heterozygous mutants were normal and fertile. Development of germ layers and formation of notochord and node were disturbed in embryos with absent *Mixl1* and, furthermore, *Mixl1* was found to contribute to proper gut morphogenesis. *Mixl1*^{-/-} embryos fail to develop a heart tube and to extend antero-posteriorly. At 9.5 dpc *Mixl1*^{-/-} embryos are arrested in their development (Hart et al., 2002). Lack of the *Mixl1* gene shows that *Mixl1* has a considerable influence on mesoderm patterning. Moreover in *Xenopus*, the *Mixl1* equivalent gene leads to ventralization after injection of Mix.1 RNA into dorsal blastomeres (Mead et al., 1996).

It has been shown that the *Mixl1* promotor is responsive to TGF β , and that *Foxh1* is essential for its activity in HepG2 cells (human liver carcinoma cell line) (Hart et al., 2005). In mouse embryos, *Foxh1* is a negative regulator of *Mixl1* expression by association with Goosecoid, *Gsc*, which has a histone recruitment function (Izzi et al., 2007).

Since the *T* expression domain is expanded in *Mixl1*^{-/-} embryos, it has been suggested that *Mixl1* acts as a repressor of *T*, an important factor for notochord development and mesoderm formation (Kispert et al., 1995). Physical interaction of *Mixl1* and *T* was shown in embryoid bodies (EB), where *Mixl1* in complex with *T* represses *Gsc* expression (Pereira et al., 2011). Also Nodal expression is affected in *Mixl1*^{-/-} embryos, the expression domain appears enlarged (Wolfe and Downs, 2014). Furthermore, induction of *Mixl1* expression by *BMP4* was shown in EB, which are in the process of differentiating into mesodermal cells (Ng et al., 2005).

1.5 Marker genes-of-interest and their expression profile in the caudal end

In the current work we are interested in the *Mixl1* expressing region in the caudal end mesoderm, since this gene is expressed in a very defined region of the caudal end mesenchyme (Figure 2.3), and is associated with mesoderm differentiation.

1.5.3 Homeodomain transcription factor *Tlx2*

In adult mice, the homeodomain transcription factor *Tlx2* is expressed in NC cell descendants, such as enteric nerves and dorsal root ganglia, cranial nerves and sympathetic ganglia (Hatano et al., 1997). In mouse embryos, *Tlx2* is expressed in the primitive streak and a proximal lateral region from 7.0 dpc on. The expression domain becomes restricted during growth of the embryo, so that *Tlx2* is expressed at the very caudal end of the primitive streak at later stages (Tang et al., 1998). At 9.0 dpc *Tlx2* is expressed adjacent to the hindgut pocket and the caudal end (Figure 2.5). To further investigate the function of *Tlx2* during mouse embryonic development, Tang and colleagues generated mice with a targeted mutation, resulting in a complete loss of *Tlx2* expression. Heterozygous embryos appear normal and are vital, whereas homozygous-null embryos die between 7.0 and 7.5 dpc, after they were arrested in development by 7.5 dpc (Tang et al., 1998). Mesoderm formation cannot be observed in the knockout embryos and, furthermore, the ectoderm of these embryos appears thickened. In wild-type embryos, *T*, which is a marker for generated mesoderm, is expressed in the primitive streak at 7.5 dpc and expands through the notochord thereafter. In *Tlx2* deficient embryos, *T* expression does not persist throughout the notochord (Tang et al., 1998).

As for *Mixl1*, *Tlx2* is also involved in the TGF β signaling pathway. In P19 cells *Tlx2* is expressed when BMP4 signaling is activated in those cells (Tang et al., 1998).

Furthermore, *Tlx2* is involved in development of the enteric nervous system of the gut. During this process, *Tlx2* is responsible for controlled apoptosis of enteric ganglia (Borghini et al., 2007).

The *Tlx2* expressing domain in the caudal end of the mouse embryos was chosen for our work, because *Tlx2* is expressed adjacent to the hindgut pocket of the caudal end (Figure 2.5) and is influenced by the BMP signaling pathway.

1.5.4 Peptide hormone Secretin

The peptide hormone Secretin, *Sct* has its main function in the adult, where it stimulates water, electrolytes and bicarbonate release to neutralize stomach contents (Meyer et al., 1970; Watanabe et al., 1986). In adult rats, *Sct* is expressed in brain, lung, heart, kidney and testes (Ohta et al., 1992). In mouse embryos, expression can be detected in different parts of the brain between 10.5 dpc and birth (Lossi et al., 2004). Expression in earlier stage embryos has not been reported. In particular, expression in the ventral ectodermal

ridge (VER) such as is visible in images from the MAMEP database has so far not been described (Figure 2.8). *Sct* was chosen here, due to its specific expression in the VER of 9.0 dpc embryos. Since the VER has an important function as a signaling center during embryonic development, it is suggested that *Sct* may be involved here.

1.5.5 Paired box transcription factor Pax3

Pax3 is a transcription factor that belongs to the group of paired box factors. *Pax3* was first mentioned in the context of neural development (Epstein et al., 1991). During embryonic development, *Pax3* is expressed in the neural tube, anterior presomitic mesoderm, and in developing somites (Figure 2.10) (Goulding et al., 1991; Epstein et al., 1991). In the caudal end, *Pax3* is expressed in the extension of the neural tube. The expression domain is probably paraxial mesoderm. Nothing has been postulated so far about the function of *Pax3* in the non-neural expression domain in the caudal end.

Participation of *Pax3* in muscle development is reported from analysis of the *Splotch* mutant, and from *Pax3* expressing cells of the dermomyotome (Goulding et al., 1991). In the *Splotch* mutant, limb muscles are completely absent, because muscle precursors do not laminate anymore from the hypaxial dermomyotome (Epstein et al., 1991). Besides this, *Splotch* mice show disabilities in migration of NC-derived melanocytes, noticeable by presence of a white belly (Serbedzija and McMahon, 1997).

In development of NC cells, *Pax3* exhibits the function of a neural plate border specifier and is important for migration of NC cells (Plouhinec et al., 2014; Goulding et al., 1991). Interactions of PAX3 with BMP and Wnt signaling has been demonstrated in *Xenopus* embryos (Bang et al., 1999).

In the current study we would like to find out more about the role of the *Pax3* expression domain in the caudal end of the embryo at 9.0 dpc. A role during NC development, especially during induction of prospective NC is postulated.

1.6 Lineage tracing in the mouse embryo

Lineage tracing assays are a common tool used to follow the progeny of cells using β -Galactosidase as a marker. Individual cells marked by the expression of the lacZ reporter, which is passed to each of its descendant cells, can be traced using an enzymatic colourimetric staining reaction. Lineage tracing provides information about how many progeny are generated from a population of cells, and where these cells migrate to during development. The first experiments following lineage of cells were performed by Conklin in 1905, where he benefitted from differences in cell pigmentation in ascidians (Conklin, 1905). Furthermore, in 1929, the first lineage tracing experiments using dye labeled cells were performed by Walter Vogt (Vogt, 1929).

1.6 Lineage tracing in the mouse embryo

Lineage tracing methods, as they are performed today, are non-invasive and assessable using genetic engineering. In order to design a lineage tracing paradigm which was inducible, and thus allowed activation of the lineage tracing construct at any time-point during embryonic development, we utilized a tamoxifen-inducible construct, modified for our purposes (see Figure 2.2) (Kretzschmar and Watt, 2012).

Two different time-points during embryonic development were chosen to compare migration of the cell progeny. The genetically integrated lineage tracing construct was controlled by the promotor of one of the marker genes-of-interest, and lacZ expression was induced by tamoxifen administration to pregnant dams in order to label embryonic cell populations. The approach used here allows immediate monitoring by staining dissected embryos with X-Gal. For a detailed analysis of target tissues, embryos were also sectioned. Thus, progeny from various marker genes could be monitored as long as the expression of the marker gene was sufficient to drive expression of the lacZ reporter.

Chapter 2

Results

2.1 Development of mesoderm in the caudal end of the mouse embryo

In order to get more insight into the function of the individual subregions, the lineage tracing approach shows where lacZ reporter expressing descendant cells of the marker gene expressing tissue are migrating to. In a downstream experiment, cells expressing the Venus reporter under control of the promotor of one of the marker genes, *Mixl1*, *Tlx2*, *Sct* or *Pax3*, were sorted from caudal ends of 9.0 dpc embryos and used for transcriptome analysis to find out, which pathways are active and which developmental programs are switched on.

2.2 Generating a Lineage Tracing Construct for *in vivo* analysis

The lineage tracing approach was performed in a time window during embryonic development, when mesoderm patterning is still an ongoing process. In order to follow descendant cells, 9.0 dpc - 10.5 dpc mouse caudal ends were chosen as the analysis stage to get an idea about the function of the 4 different subregions of the caudal end and their participation in developmental processes of mesoderm patterning. Lineage tracing constructs were inducible with tamoxifen, to allow selection of distinct time points during development. A lacZ reporter was implemented into the ROSA26 locus, to enable tracing of the progeny of marker gene expressing cell populations.

BAC clones (bacterial artificial chromosome), with the lineage tracing cassette introduced following the ATG start site of the particular marker genes were used for random integration into the genome of mice carrying a modified ROSA26R locus (Soriano, 1999). Following induction with tamoxifen, the CRE recombinase translocates to the nucleus of the cell and removes the selection cassette including a stop signal from the ROSA26R locus. BACs used in this approach contained the following genes and promotors: *Tlx2*, *Mixl1*, *Sct* and *Pax3*.

The lineage tracing cassette contained (from 5' to 3') an H2B-Venus reporter, followed by

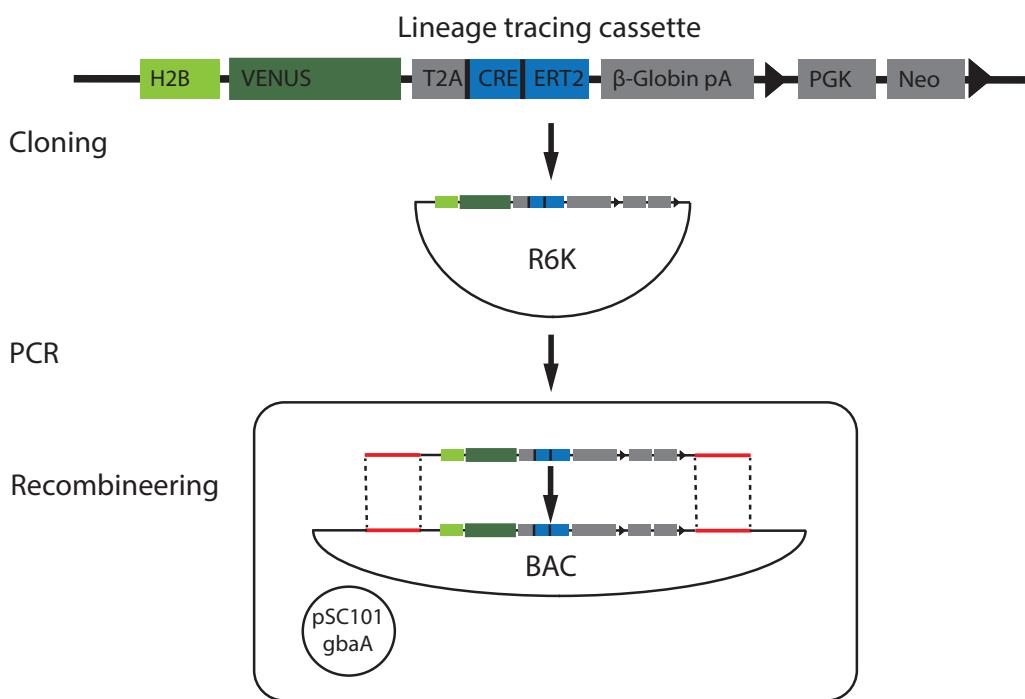


Figure 2.1: Recombineering strategy

Scheme of the recombineering procedure of the lineage tracing cassette into the relevant BACs. Homologous regions used for recombineering are depicted in red.

a T2A and CRE-ERT2, a β -Globin-pA and a frt flanked PGK driven hygromycin resistance gene for selection of the BACs.

To visualize expression of the marker gene, Venus fluorescence under control of the marker gene promotor was used. T2A self cleaving peptide was used to separate reporter protein from CRE-ERT2. CRE was fused to a mutated ligand binding domaine of the estrogen receptor (ERT2) for tamoxifen inducibility, in order to allow removal of the neomycin selection cassette from the ROSA26R locus, which includes a stop signal to prevent translation of the lacZ reporter when there is no tamoxifen.

Cloning of the lineage tracing cassette was performed in the pBluescript II SK+ cloning vector. Before recombineering of the cassette into BAC clones, the construct was subcloned into an R6K vector, which requires *pir* expression for initiation of replication, and thus prevents replication of the vector in BAC-containing bacteria (*pir*-). For recombineering of the lineage tracing cassette into the BAC, homologous sequences (50 bp on 5' and 3') were added via polymerase chain reaction (PCR). The PCR product was then used for recombineering into the BAC. Another recombineering step followed, to exchange the selection-cassette for hygromycin (Figure 2.1). BAC clones containing the lineage tracing cassette were selected by hygromycin resistance and sequenced to check for errors. They

2.2 Generating a Lineage Tracing Construct for *in vivo* analysis

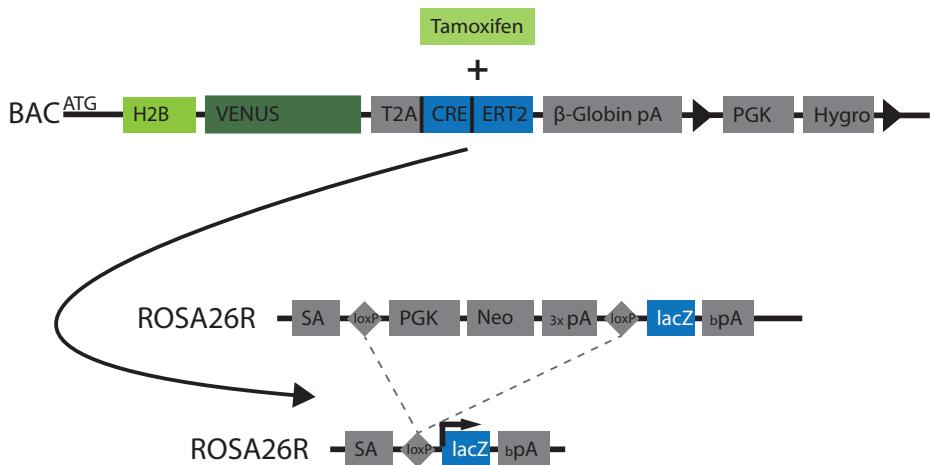


Figure 2.2: Lineage tracing construct

Gene specific BAC with the tracing cassette integrated following ATG of the gene of interest. Constructs include H2B-Venus separated from CRE-ERT2 using a T2A sequence. β -Globin-pA to stop transcription, frt flanked PGK-driven Hygromycin resistance was used for selection. Following tamoxifen treatment, Cre recombinase is activated and allowed removal of PGK-neo-pA from the ROSA26R locus to permit expression of the lacZ reporter.

were then linearized prior to electroporation into ES cells.

129S;b6 Gt(ROSA)26Sor^{tm1Sor} ES cells were used for the integration of the BAC clone including the lineage tracing cassette. The ROSA26R locus of these cells contains a splice acceptor followed by a loxP flanked PGK driven neomycin resistance for selection of the ES cells, followed by a lacZ gene for the lineage tracing procedure. DNA from hygromycin resistant ES cells was either digested with a suitable enzyme and tested by Southern blotting, using a probe hybridizing to a fragment of the hygromycin sequence, or tested using PCR.

ES cells, tested positive were expanded and used for tetraploid aggregation. Mice pregnant with embryos aggregated from modified ES cells outlined above were administrated tamoxifen (2 mg tamoxifen per 40 g body weight by intraperitoneal injection at 7.5 dpc and 8.5 dpc, or 8.5 dpc and 9.5 dpc) to activate expression of the lacZ reporter in marker gene expressing cells. In these cells, tamoxifen was bound to the modified estrogen receptor fused with CRE recombinase and was translocated to the nucleus, where the loxP flanked PGK-neo-pA cassette was excised. Now the lacZ reporter was expressed in all cells, where the lineage tracing cassette was expressed under the control of a marker gene promoter and passed to the progeny of these cells.

Embryos from these experiments were either used for whole mount *in situ* hybridisation (WISH), lineage tracing experiments, or transcriptome analysis. WISH using a probe

detecting *Gfp* (which also recognizes the *Venus* mRNA), or direct H2B-Venus expression by fluorescence microscopy (LSM 710 NLO, Zeiss), was used to verify the expression patterns of the lineage tracing constructs inserted into BAC clones. The expression pattern of the construct should resemble the endogenous expression of the genes for which the transcription start site was used.

In order to get insight into mesoderm patterning and axis development, two individual time-points during development were chosen for the lineage tracing approach to monitor differences of progeny distribution. For the first window, pregnant mice were induced with tamoxifen at both day 7.5 dpc and 8.5 dpc, and embryos prepared at day 9.0 dpc to get more insight into trunk development. In the second case, pregnant mice were induced with tamoxifen at 8.5 dpc and 9.5 dpc and these embryos were dissected at 10.0 dpc, when tail development is taking place. Due to the use of tetraploid aggregation to generate embryos there was often some variability in the age of the embryos. Tamoxifen was administered on each of 2 days prior dissection of embryos, to have enough daughter cells to follow. Only in the case of *Mixl1* were there enough descendants that one injection was sufficient. By alternating time-points of tamoxifen induction during development we can follow marked subregions. The lineage tracing assay was used to see where in the caudal end the descendants of cells were located which had the lacZ reporter expression activated.

2.2.1 Dre versus Cre recombinase

For the first phase of the project, constructs using Dre recombinase instead of Cre were used. Backbone vectors used for BAC assembly exhibit loxP sites in their sequence. If a Cre/loxP system is also used in the cassette introduced into the BAC, it can also interact with the loxP sites from the BAC backbone and remove hereby the whole BAC from the genome. Hence in order to avoid this possibility we thought first to preferentially use the Dre/rox system.

Cre recombinase recombines specifically with *loxP* sites, and Dre with *rox* sites. Toxicity of Cre recombinase is something which is occasionally challenging *in vivo*, and for Dre recombinase this seems to be an issue (Anastassiadis et al., 2009). Apart from that, efficiency of the Dre recombinase is reported to be much weaker than Cre in *Drosophila*, unlike the efficiency of Dre in *E. Coli* (Nern et al., 2011).

As a preliminary *in vitro* experiment, ES cells containing the modified ROSA26 locus with *rox* sites flanking the PGK Hygromycin cassette and a randomly integrated BAC clone with a *Nanog* promotor followed by the Dre lineage tracing cassette were induced using tamoxifen. The ES cells could be positively stained using X-Gal, suggesting that the Dre lineage tracing construct should work (Supplementary Figure 7.1).

As a preliminary *in vivo* experiment, the lineage tracing construct was tested using the *Tlx2* promotor. After induction with tamoxifen, no daughter cells from *Tlx2* expressing cells

2.2 Generating a Lineage Tracing Construct for *in vivo* analysis

could be detected when using Dre recombinase for excision of the selection cassette in the ROSA26 locus. Amount and timing of tamoxifen administrations were modified and tested, yet lacZ expressing cells were never apparent. Thus, we switched our strategy and developed the constructs using the Cre/loxP system, as described above. All data in this thesis results from analysis performed using Cre expressing constructs.

2.3 Lineage tracing driven by the *Mixl1* homeodomain transcription factor

Mixl1 is expressed in the caudal mesoderm of the mouse embryo, where also the mesoderm patterning gene *Tbx6* is expressed. Mesoderm patterning is a crucial step in murine development. We used the lineage tracing assay to get insight into possible regulatory mechanisms of *Mixl1*.

To insure that the lineage tracing construct was being driven by the *Mixl1* promotor, Venus mRNA and *Mixl1* mRNA domains were compared using WISH (Figure 2.3). The *Mixl1* expression data was taken from the MAMEP database (Berlin, Germany; <http://mamep.molgen.mpg.de>). The fluorescence signal from Venus was also directly visualized using a laser scanning microscope (Figure 2.3B). *Mixl1* expression and Venus expression and fluorescence from the lineage tracing construct showed overlapping signals in the same areas in tail mesoderm in 9.0 dpc embryos. For lineage tracing of descendants of the *Mixl1* expressing subdomain, embryos were aggregated from

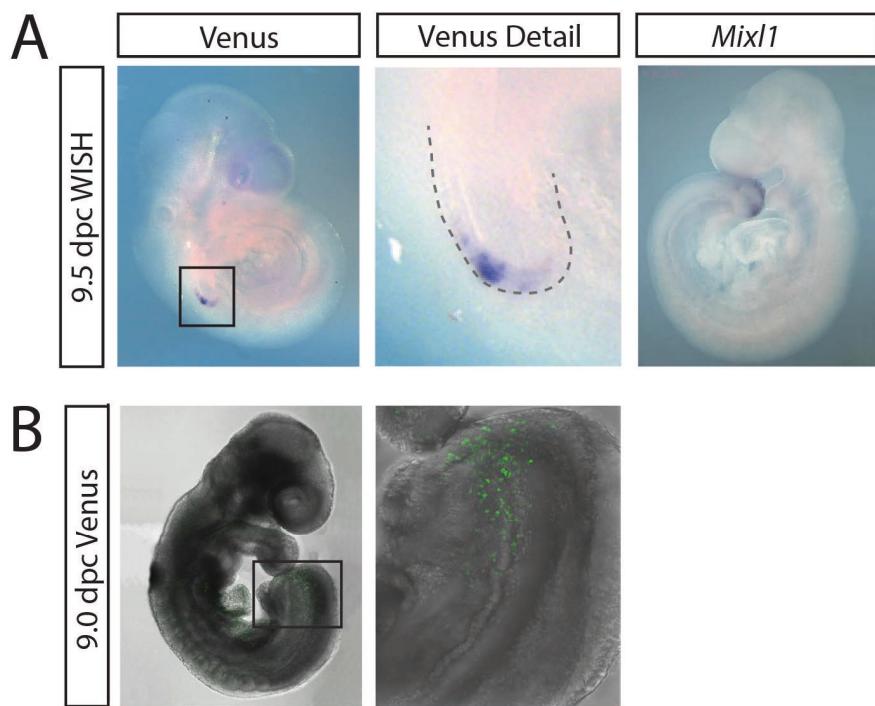


Figure 2.3: *Mixl1* and Venus expression in 9.0dpc embryo

A: left, mouse embryo including the lineage tracing construct with WISH for Venus; middle, detail of the caudal end (with dashed lines); right, WISH *Mixl1* (MAMEP). B: left, fluorescent image of mouse embryo carrying the lineage tracing construct with Venus fluorescence in green; right, detail of caudal end.

2.3 Lineage tracing driven by the *Mixl1* homeodomain transcription factor

129S; b6 Gt(ROSA)26Sor^{tm1Sor} ES cells, carrying a randomly integrated *Mixl1* BAC clone including the lineage tracing cassette. Mice were induced with tamoxifen for one day at 8.0 dpc and dissected at 9.0 dpc. Embryos were then LacZ stained, embedded in paraffin and sectioned.

After tamoxifen induction at 8.0 dpc some cells which express *Mixl1* at that time point also express lacZ, depending on the efficiency of the CRE-ERT2 activity. Two mg of tamoxifen was sufficient here to detect lacZ stained cell progeny. During this time window of 24-30 hours, cells expressing lacZ were found mostly at the posterior end of the notochord (Figure 2.4). Progeny of *Mixl1* expressing cells from tail bud mesenchyme with activated lacZ reporter could also be seen scattered throughout the paraxial and lateral mesoderm (Figure 2.4).

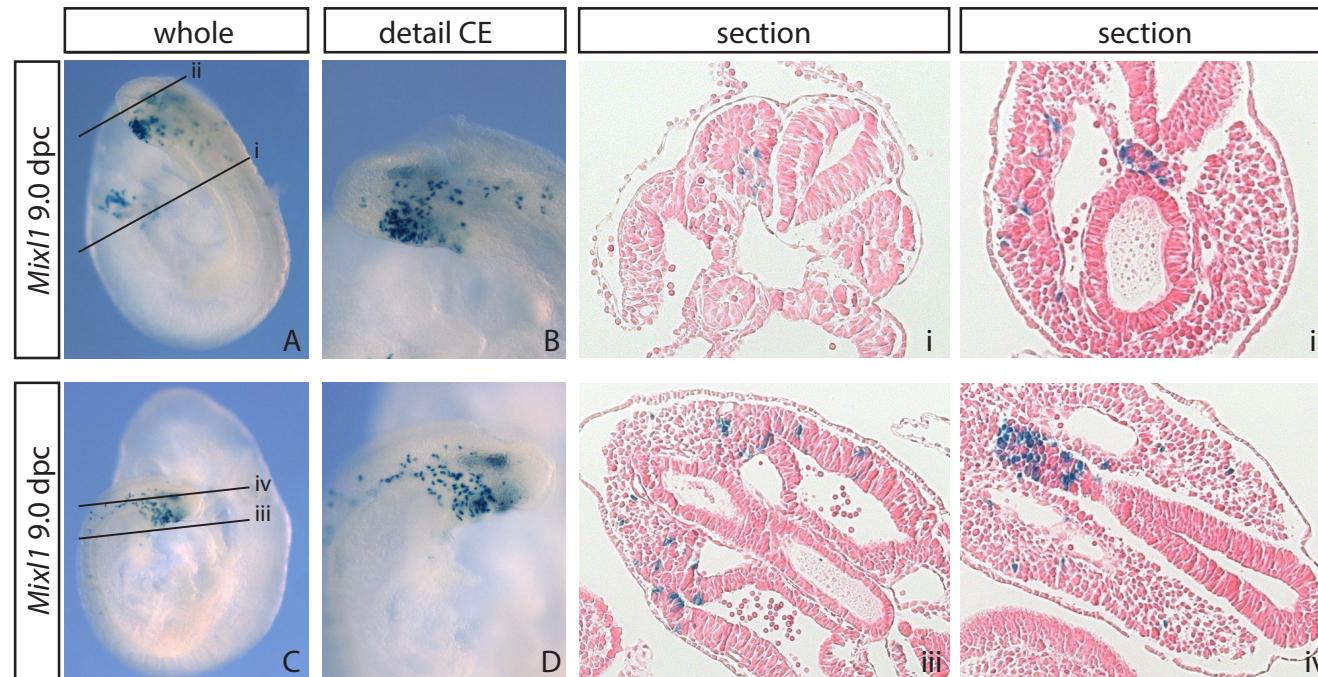


Figure 2.4: *Mixl1* lineage tracing of 9.0 dpc embryos

Lineage tracing of *Mixl1* in embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos stained with X-Gal after 1 day of tamoxifen induction. A and C: Whole embryos with tagged regions of the sections shown in i and ii or iii and iv. B and D: Detail of caudal end, lateral view with traced progeny in notochord and mesoderm. Stained cells are visible in the paraxial mesoderm (i), lateral mesoderm (ii) and notochord (ii). Sections through the ventral part of the caudal end, illustrating lacZ expressing progeny in the lateral mesoderm (iii), notochord and lateral mesoderm (iv). Sections are transversal in (i) and (ii), and longitudinal in (iii) and (iv).

2.4 Lineage tracing driven by the *Tlx2* homeodomain transcription factor

Tlx2 T-cell leukemia homeobox protein 2 transcription factor related to human *HOX11*, was shown to be a downstream target of BMP signaling, a crucial pathway in gastrulation and mesoderm patterning (Kishigami and Mishina, 2005; Tang et al., 1998). At 9.0 dpc *Tlx2* is expressed at the caudal end of the embryo and adjacent to the hindgut pocket (Figure 2.5). From 9.0 dpc on there is also an expression domain in the branchial arches. Later, at 10.5 dpc, *Tlx2* is expressed in the dorsal root ganglia and as well in the branchial arches (Figure 2.5 B). The expression pattern of *Tlx2* by *in situ* hybridization (2.5 B) and fluorescence of H2B Venus within the lineage tracing construct, driven by the *Tlx2*

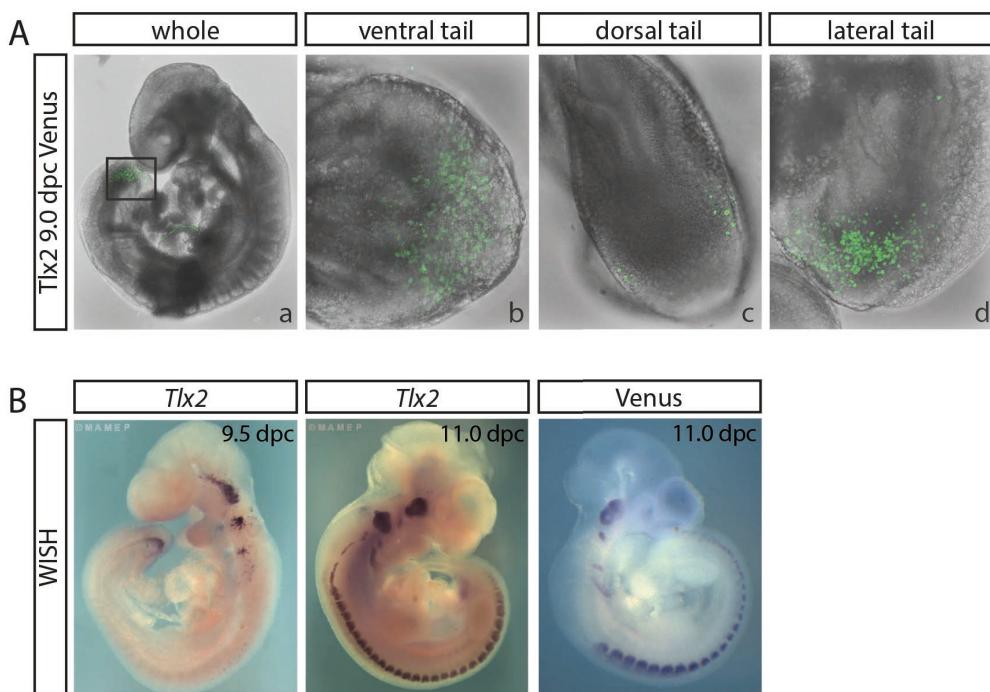


Figure 2.5: *Tlx2* expression in 9.0 dpc embryo

A: Venus expression in embryos containing the construct shown in Figure 2.2 under control of the *Tlx2* promotor, exhibits the expected pattern shown for *Tlx2* by *in situ* hybridization A-a: Whole embryo showing Venus expression in the hindgut pocket of the caudal end and in the caudal end. A-b: caudal end, ventral view. A-c: Caudal end, dorsal view. A-d: Caudal end, lateral view, displaying Venus expressing cells in the hindgut pocket.

B: WISH from the MAMEP database using *Tlx2* probe at 9.5 dpc and 11.0 dpc, and WISH using a Venus probe at 11 dpc show overlapping expression.

Results

promotor (2.5 A), are congruent. Particularly in the dorsal view, the lateral shape of the expression domain is apparent.

As described in section 2.3, 9.0 dpc embryos were dissected and stained with X-Gal. *Tlx2* expression in the caudal end is relatively weak, two days of tamoxifen administration were necessary to induce lacZ expressing progeny from the *Tlx2* expression domain. Pregnant mice were injected twice with tamoxifen when the embryos were 7.0 dpc and 8.0 dpc, and embryos were dissected at 9.0 dpc. When embryos at 10.5 dpc had to be dissected, tamoxifen was administered when embryos were 8.0 dpc and 9.0 dpc.

Embryos dissected at 9.0 dpc show daughter cells from *Tlx2* expressing cells with activated lacZ expression in the paraxial mesoderm (tip of the tail), lateral mesoderm (more anterior in tail) and in the hindgut (Figure 2.6). Contributions of the *Tlx2* expressing domain of the caudal end to the development of the enteric nervous system (nervous system surrounding the gastrointestinal tract, developed from neural crest derived cells) is already known. Progeny of *Tlx2* expressing cells migrate from their expression domain at the dorsal root ganglia (Figure 2.5, B middle and right side) to participate in the process of development of neural crest cells, forming the enteric nervous system (Borghini et al., 2006).

In this lineage tracing assay of 10.5 dpc embryos, even more progeny were found in the hindgut (Figure 2.7). Furthermore, lacZ expressing cells could be found in the dorsal root ganglia of 10.5 dpc embryos. The outcome of the lineage tracing assay shows a participation of progeny of the *Tlx2* expressing subregion of the caudal end in gut development.

2.4 Lineage tracing driven by the *Tlx2* homeodomain transcription factor

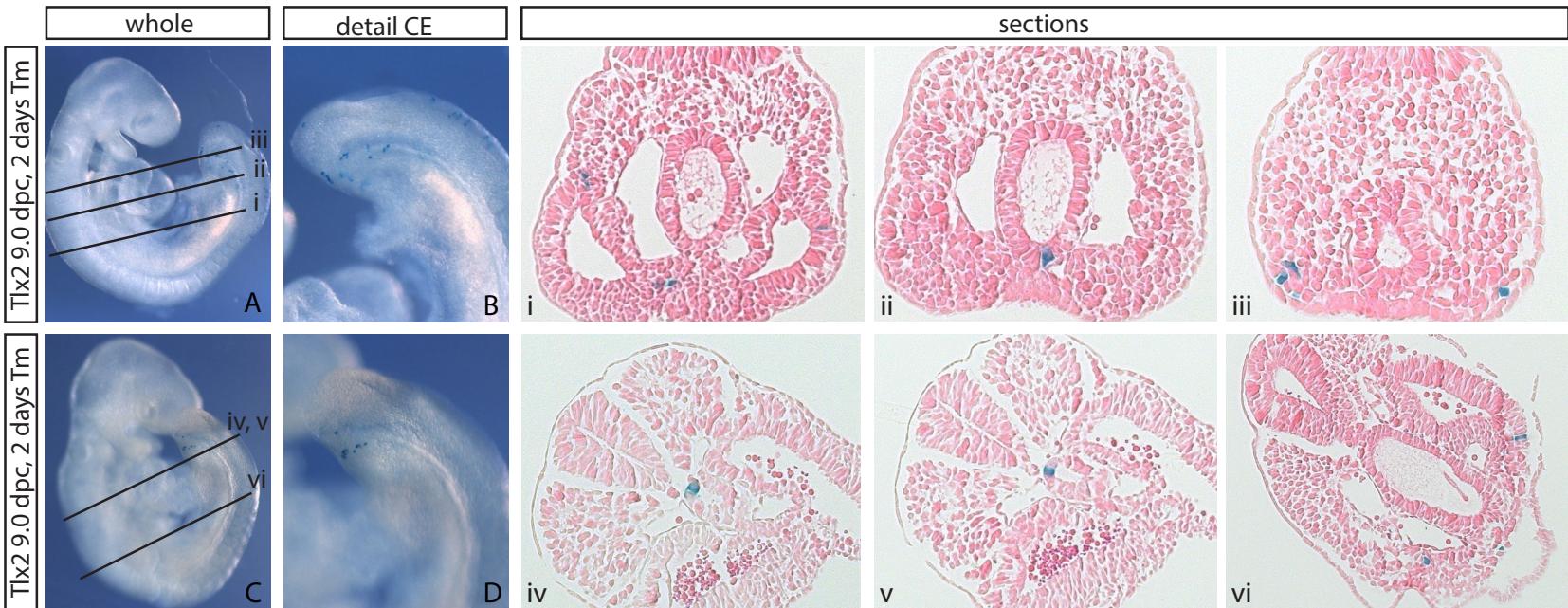


Figure 2.6: Lineage tracing of *Tlx2* in 9.0 dpc embryos

Lineage tracing of *Tlx2* in embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos are shown following 2 days of Tamoxifen induction and Xgal staining. A: Whole embryo, 9.0 dpc, sections A-i-iii marked. B: Detail of lateral view on caudal end stained progeny visible in mesoderm. Sections show occurrence of descendants in lateral mesoderm (A-i, A-iii) and hindgut (A-ii). C: Whole embryo, 9.0 dpc, sections iv-vi marked. D: Detail of lateral view on caudal end stained progeny visible in mesoderm. Sections show occurrence of descendants in hindgut (C-i,C-ii) and lateral mesoderm (C-iii). All sections are transversal.

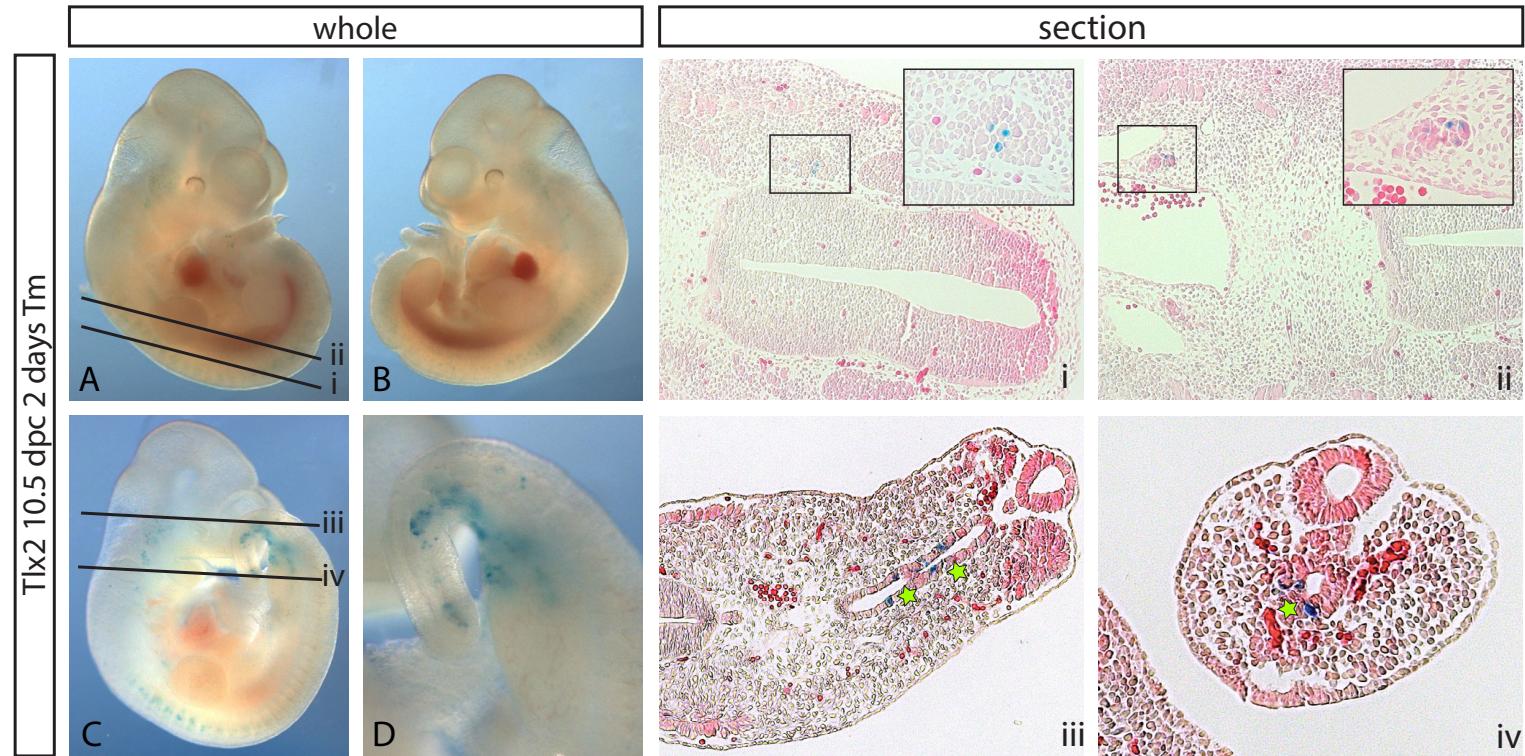


Figure 2.7: Lineage tracing of *Tlx2* in 10.5 dpc embryos

Lineage tracing of *Tlx2* in embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos are shown following 2 days of Tamoxifen induction and Xgal staining. A: Whole embryo, 10.5 dpc from right, sections i,ii marked. B: Whole embryo, 10.5 dpc from left. Sections show details from neural tube region with stained cells in dorsal root ganglia (A-i, A-ii). C: Whole embryo, 10.5 dpc, sections iii, iv marked. D: Detail of the tail. Sections show stained cells in the hindgut (asterisk in C-iii and C-iv) at different positions within the tail. All sections are transversal.

2.5 Lineage tracing driven by the peptide hormone Secretin (*Sct*)

During embryonic development *Sct* is expressed in the ventral ectodermal ridge (VER), an important signaling center for embryonic development. *Sct* is also expressed in heart, brain, lung, testis and kidney. During embryonic development, *Sct* is known to be involved in brain, gut and kidney development (Siu et al., 2005).

In 9.5 dpc embryos, *Sct* is expressed in the forelimb bud, and the ventral side of the embryonic tail around the region of the VER (WISH, Figure 2.8). Expression of H2B-Venus using the lineage tracing construct indicates the same expression pattern as the *in situ* hybridization (Figure 2.8).

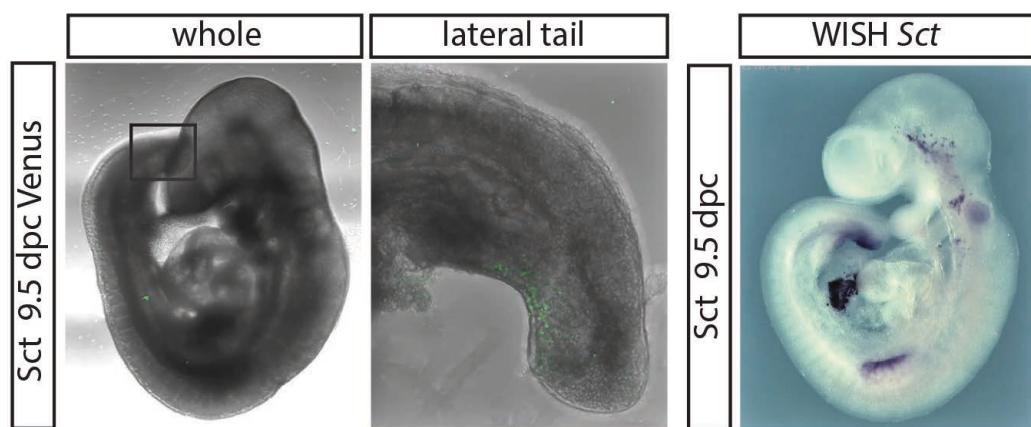


Figure 2.8: Expression of *Sct* in 9.5 dpc embryos Left: Venus expression in embryos containing the construct shown in Figure 2.2, under control of the *Sct* promotor, exhibits the expected pattern shown by *Sct* in *in situ* hybridization in right panel. Middle: Detail of lateral view of caudal end. Right: Whole mount *in situ* hybridisation of 9.5 dpc embryo from MAMEP database using *Sct* probe.

As described in section 2.3, pregnant mice were administrated tamoxifen at 7 and 8 dpc, embryos were taken from the uterus at 9.0 dpc. LacZ expressing daughter cells of *Sct* expressing cells could be uncovered in coelom lining (Figure 2.9 A) and spotted throughout the mesoderm (Figure 2.9B, C). For lineage tracing of 10.0 dpc embryos, mice were injected with tamoxifen at 8.0 dpc and 9.0 dpc. Traces of *Sct* expressing progeny could be found further throughout the embryo, so that progeny could be found now in the limb bud, gut epithelium, ganglia and somites (Figure 2.9). This wide distribution of progeny can be an indicator for diverse roles of *Sct* expressing cells in the early tailbud as embryonic development proceeds.

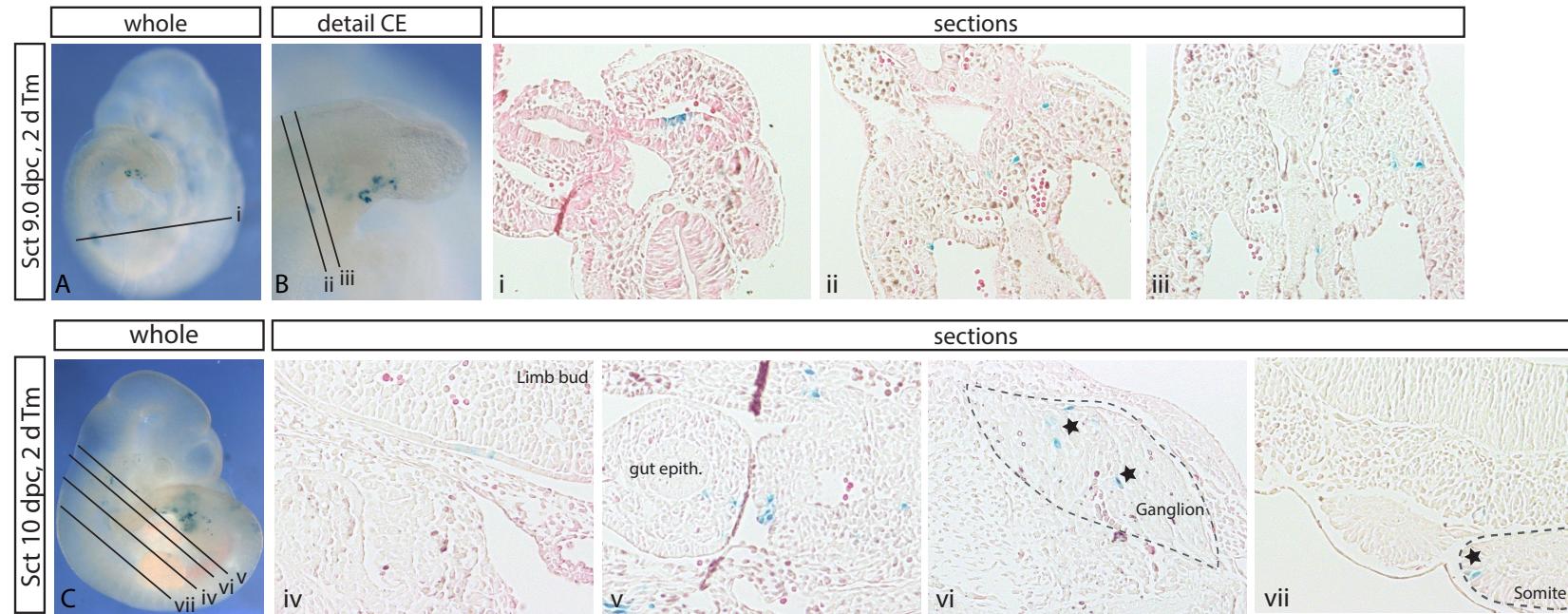


Figure 2.9: Lineage tracing of *Sct* in 9.0 and 10.0 dpc embryos

Lineage tracing of *Sct* in embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos are shown following 2 days of tamoxifen induction and XGal staining. A: Whole embryo, 9.0 dpc, sections i-iii marked. B: Detail of lateral view on caudal end with stained progeny visible in mesoderm. Sections show occurrence of descendants in coelom lining (i) and mesoderm (ii,iii). C: Whole embryo, 10.0 dpc, sections iv-vii marked. Sections show stained descendants in limb bud (iv), mesoderm (v), ganglia (asterisk, vi) and somites (asterisk, vii). All sections are transversal.

2.6 Lineage tracing driven by the transcription factor *Pax3*

The transcription factor *Pax3* is expressed in neural tissue and mainly known as a regulator of myogenesis (Magli et al., 2013) and as a factor for NC formation (Milewski et al., 2004). At 9.5 dpc *Pax3* is expressed in the paraxial mesenchyme and somites, several parts of the future brain, in the future spinal cord, cranial and dorsal root ganglia and in the tail as a narrow, very defined stripe in extension of the neural tube (Figure 2.10, lower panels). Expression of H2B-Venus from the lineage tracing construct (Figure 2.10, upper panels) resembles the expression pattern of *Pax3*, as illustrated by the *in situ* hybridization (Figure 2.10, lower panels).

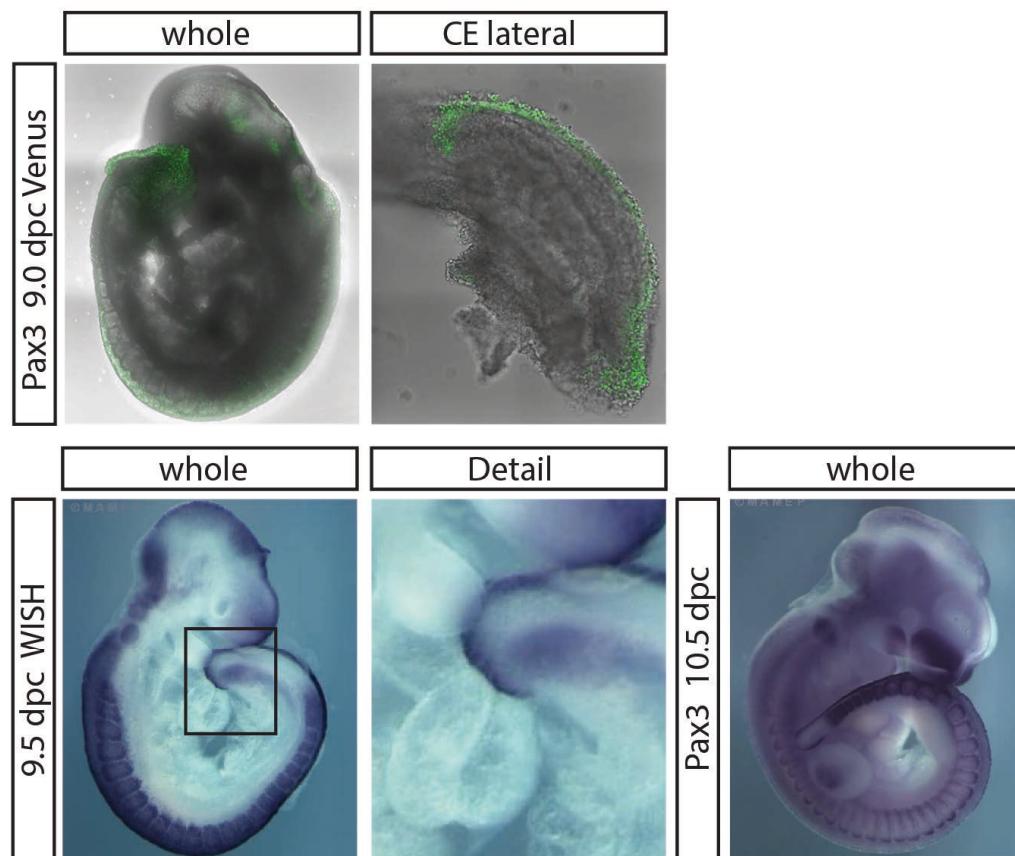


Figure 2.10: Expression of *Pax3* in 9.5 dpc embryos Venus expression in embryos containing the construct shown in Figure 2.2 under control of the *Pax3* promotor, exhibits the expected pattern shown by *Pax3* in *in situ* hybridization in lower panels. Top right: Detail of lateral view on caudal end. Lower: Whole mount *in situ* hybridisation of 9.5 dpc embryo (left) and 10.5 dpc embryo (right) from MAMEP database using *Pax3* probe.

Results

Pregnant mice with embryos carrying the lineage tracing construct driven by the *Pax3* promotor were administrated tamoxifen when the embryos were at 7 and 8 dpc. Embryos were then dissected at 9.0 dpc.

Progeny of *Pax3* expressing cells do not migrate very far at this time point. Thus, migration of neural crest cells has not started yet. Most progeny can be seen in the neural tube (Figure 2.11). Only in the caudal end, lacZ reporter expressing cells are found in the caudal end mesoderm (Figure 2.11 C, D). Very few cells with lacZ reporter expression can be visualized also in the hind gut. Lineage tracing at this time point shows, that *Pax3* progeny remain local at this time-point and are not migrating.

For investigating lineage tracing in 10.0 dpc embryos, tamoxifen administration of the female was done on day 8 and 9 dpc of pregnancy.

Daughter cells of *Pax3* expressing cells were present in the neural tube, but migrating cells were also visible by this stage (Figure 2.12). Progeny of *Pax3* expressing cells began to migrate from the neural tube into ganglion structures. In the caudal end mesoderm almost all cells expressed the lacZ reporter. Only cells surrounding the hindgut and cells within the hindgut itself did not express lacZ, and thus contained no progeny from the *Pax3* expressing subregion. Migrating cells, as they are visible here are cells contributing to muscle development. It is known that *Pax3* is involved in the latter and muscle progenitor cells migrate from the somites to their destination in the embryo.

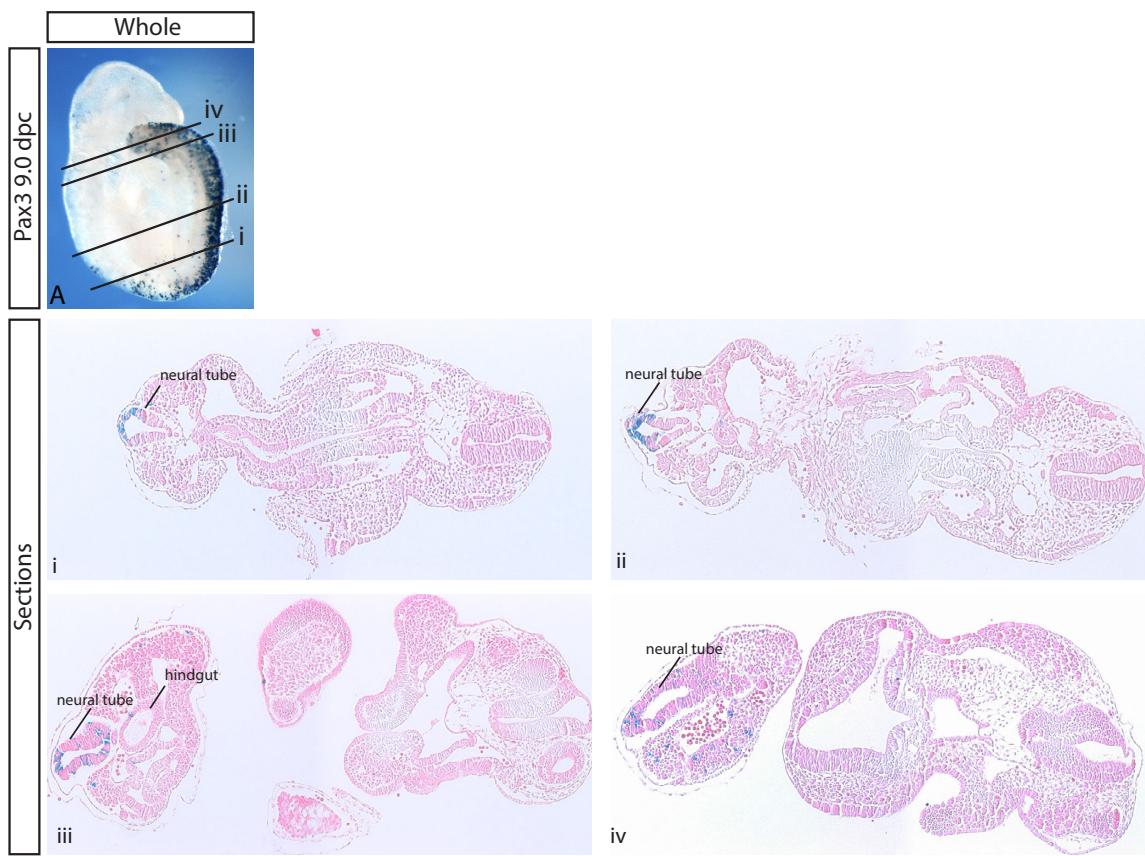


Figure 2.11: Lineage tracing of *Pax3* in 9.0 dpc embryos Lineage tracing of *Pax3* in 9.0 dpc embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos are shown following 2 days of tamoxifen induction and XGal staining. All sections are transversal. A: Whole embryo, 9.0 dpc, sections i-iv marked. Sections show occurrence of descendants in neural tube (i-iv), paraxial (iii-iv) and lateral mesoderm (iii,iv).

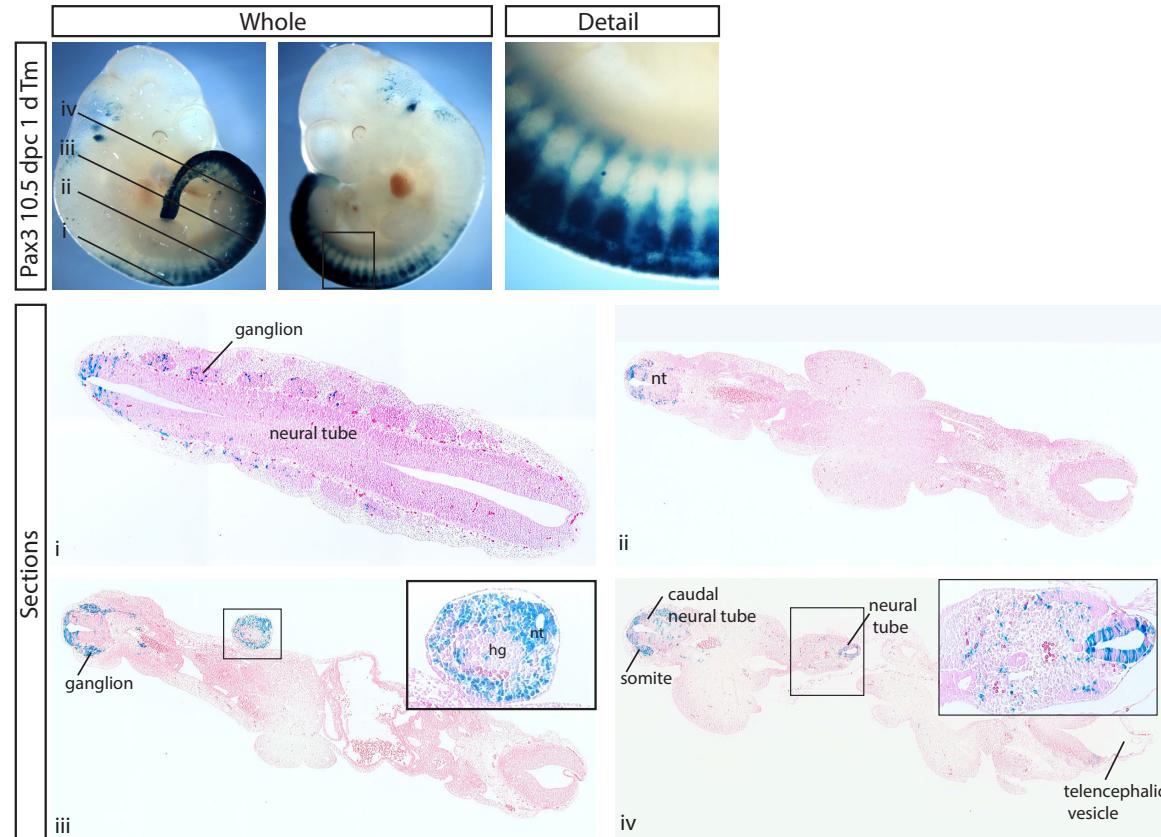


Figure 2.12: Lineage tracing of *Pax3* in 10.5 dpc embryos Lineage tracing of *Pax3* in 10.5 dpc embryos containing the construct shown in Figure 2.2 integrated into R26R ES cells. Embryos are shown following 2 days of tamoxifen induction and X-Gal staining. Top left: Whole embryo from right side, 10.5 dpc, sections i-iv marked. Top middle: Whole embryo from left side, 10.5 dpc. Top right: Magnification of area marked in top middle picture, distribution of descendants dorsally. Sections show occurrence of descendants in neural tube (i-iv), paraxial (iii, iv) and lateral mesoderm (iii, iv). All sections are transversal.

2.7 Transcriptome analysis of caudal end subregions

To learn more about the regions in the caudal end which were investigated using lineage tracing, four subregions expressing one of the marker genes were analyzed on transcriptome level and compared. *Mixl1* is involved in mesoderm differentiation and development, *Tlx2* is also involved in patterning processes of mesoderm but might also contribute to endoderm development. The function and interaction of partners of *Sct* are so far unknown, but the expression pattern in the VER indicates a possible function for *Sct* during embryonic development. *Pax3* is very relevant in NC induction and is an important factor for myogenic development.

Transcriptome analysis was carried out using embryos having between 16-20 somites (~9.0 dpc). The caudal end of the embryo, without somites, was dissected and prepared for fluorescence activated cell sorting (FACS) (Figure 2.13). FACS was performed by Dr. Frederic Koch (MPI-MG) RNA libraries generated from these cells were sequenced with Illumina technology using our in-house sequencing facility. Analysis of RNA sequencing (RNA-Seq) data was performed by Jinhua Liu.

First canonical pathways and their appearance in the different subgroups were analyzed in the context of the *Tbx6* expression domain. Next, all data of *Mixl1*, *Tlx2*, *Sct* and *Pax3* expressing cells were compared and clustered. With information from the lineage tracing assay all clusters were analyzed for genes involved in functions presumed from the migration of progeny during lineage tracing.

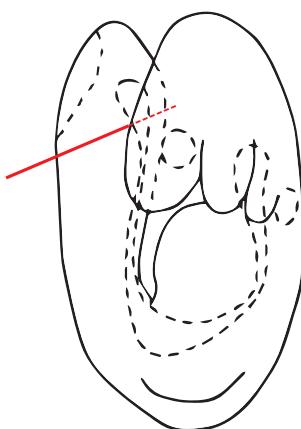


Figure 2.13: Dissection of 9.0 dpc embryo for FACS

Caudal ends of embryos with 16-20 somites were cut from the embryo and used for FACS, followed by RNA-Seq. Picture adapted from EMAP eMouse Atlas Project (<http://www.emouseatlas.org>) (Richardson et al., 2014).

2.7.1 Analysis of canonical pathways in subregions of the caudal end

For verification of the method and to get an idea about developmental processes, canonical signaling pathways for 4 subregions of the caudal end expressing *Mixl1*, *Tlx2*, *Sct* and *Pax3* were examined. Ingenuity pathway analysis (IPA, www.ingenuity.com) was used to analyze RNA sequencing data for canonical pathways. Here, only differential expressed genes of each subgroup were compared to *Tbx6* expressing cells, which represent all presomitic mesodermal cells in the caudal end.

Canonical pathways upregulated in the transcriptome of *Mixl1* expressing cells

The caudal end subregion where *Mixl1* expressing cells reside is crucial in patterning of mesoderm and in heart development. Canonical pathways upregulated here are representative of these functions.

Wnt signaling shows high activity in the posterior end of the paraxial presomitic mesoderm. In the *Mixl1* expressing subgroup Wnt signaling is active. Wnt/β-Catenin signaling plays a crucial role in the segmentation process, where also *Notch* signalling is involved in the organisation of segmentation. EMT, is regulated by many different pathways and factors

Table 2.1: Canonical pathways of differentially expressed genes in the *Mixl1* expressing caudal cell population

Signaling pathway	Function	p-value
Wnt/β-Catenin pathway	Somitogenesis	4.78E-05
Mouse embryonic stem cell pluripotency	Pluripotency	1.26E-02
Notch-signaling	Organization of somitogenesis	3.48E-0.2
Shh signaling	Limb development, segmentation, neural tube patterning	7.37E-02
Role of NANOG in mammalian embryonic stem cell pluripotency	Pluripotency	7.48E-02
Regulation of the epithelial to mesenchymal transition pathway	Regulation of EMT	2.58E-0.3
TGF-β signaling	Regulation of EMT	2.15E-03
NGF-signaling	Neural development	2.42E-02
Gadd45 signaling	Cell survival and apoptosis	2.63E-02
BMP signaling	Different tasks in development	3.81E-02

2.7 Transcriptome analysis of caudal end subregions

which are also expressed in this subgroup. TGF β signaling, SHH, and Wnt signaling, amongst others, are important pathways in this process (Lamouille et al., 2014). Neural development controlled by NGF signaling is also ongoing in this particular subgroup. Innervation and cutaneous sensory development are organized by NGF signaling (Wickramasinghe et al., 2008). GADD45 (Growth Arrest and DNA Damage-inducible 45) is a pathway involved in stress response and was also represented in this analysis.

The presence of all these pathways in the *Mixl1* expressing subregion of the caudal end brings together the findings of *Mixl1* deficient mice, where mesoderm patterning is defective, and the outcome of the lineage tracing, where the progeny of this subgroup is apparent in lateral and particularly paraxial mesoderm. As *Mixl1* deficient mice fail to extend antero-posteriorly, *Mixl1* seems to play an important role in this elongating process.

Canonical pathways in differentially expressed genes of the *Tlx2* expressing subgroup

Besides *Mixl1*, also *Tlx2* plays an important role in mesoderm patterning. Hence, canonical pathways which are active in this subgroup reflect this function.

Both the Wnt/ β -Catenin pathway and Shh signaling are crucial for somitogenesis. Since *Tlx2* expressing cells are found in the primitive streak, these pathways are also active in this subgroup of cells. Many organizing signals are relevant for development of cells and

Table 2.2: Canonical pathways of differentially expressed genes in the *Tlx2* expressing caudal end population as found by RNA-Seq analysis

Signaling pathway	Function	p-value
Wnt/ β -Catenin pathway	somitogenesis	3.18E-04
Transcriptional Regulatory Network in Embryonic Stem Cells	Pluripotency and differentiation	5.64E-03
Regulation of the epithelial to mesenchymal transition pathway	EMT regulation	6.43E-03.
Mouse embryonic stem cell pluripotency	Pluripotency	2.13E-02
TGF- β signaling	Regulation of EMT	7.73E-02
Axonal Guidance Signaling	Axon outgrowth	
Shh signaling	limb development, segmentation, neural tube patterning	7.37E-02
Gadd45 signaling	cell survival and apoptosis	2.63E-02
NGF-signaling	Neural development	3.89E-02

maintenance of pluripotency. Such pathways are combined in "Transcriptional Regulatory Network in Embryonic Stem Cells" and "Mouse Embryonic Stem Cell Pluripotency", as pluripotency is necessary for differentiation into diverse lineages. EMT is crucial for several steps in development, like mesoderm formation during gastrulation. Genes which are involved in EMT regulating TGF- β signaling are expressed in this caudal end subgroup. *Tlx2* is also involved in neural development, particularly in enteric nervous system development, which is shown by expression of pathway members of axon guidance signaling and NGF signaling.

Canonical pathways expressed in the *Tlx2* expressing subgroup of the caudal end reflect many fundamental functions which occur in caudal end mesoderm during developmental processes.

Canonical pathways upregulated in the *Sct* expressing subpopulation

Nothing is known so far about the role of *Sct* in caudal end development. Canonical pathways which are upregulated in *Sct* expressing cells are typical for mesoderm development. Wnt/ β -Catenin pathways like Shh signaling are very important in somitogenesis. Furthermore Notch signaling is part of the organization of segmentation during embryonic development. This analysis shows the involvement of *Sct* in diverse processes in the caudal mesoderm.

Table 2.3: Canonical pathways of differentially expressed genes in the *Sct* expressing caudal end population as found by RNA-Seq analysis

Signaling pathway	Function	p-value
Wnt/ β -Catenin pathway	Somitogenesis	1.25E-06
Regulation of the epithelial to meso- enchymal transition pathway	EMT regulation	1.12E-04
Mouse embryonic stem cell pluripotency	Pluripotency	2.38E-03
Transcriptional Regulatory Net- work in Embryonic Stem Cells	Pluripotency and differentiation	3.03E-03
TGF- β signaling	Regulation of EMT	1.38E-02
Shh signaling	Limb development, segmentation	2.51E-02
Gadd45 signaling	Cell survival and apoptosis	3.45E-02
Notch-signaling	Organization of somitogenesis	4.83E-02

2.7 Transcriptome analysis of caudal end subregions

Canonical pathways upregulated in *Pax3* expressing cells

The *Pax3* expressing subgroup represents cells residing in the caudal mesoderm of the embryo in extension of the neural tube. Canonical pathways represented in the *Pax3* expressing subpopulation (Table 2.4) are mainly involved in neural development, somitogenesis, and regulation of somitogenesis.

Axonal guidance signaling is fundamental for growth of axons to their destination in the embryo. For growth and proliferation of cells, NGF signaling is crucial. Wnt signaling and Notch signaling are essential for regulation of EMT and somitogenesis. Together with TGF- β and BMP signaling they fulfill multiple tasks in driving developmental programs and regulation of cell proliferation and differentiation. These signaling pathways show the mesodermal character of the *Pax3* expressing cell group and also indicates to a later implementation of these cells into the neural lineage.

Table 2.4: Top canonical pathways upregulated in the *Pax3* expressing caudal end population as determined by GO analysis from RNA-Seq data of *Pax3* expressing cells at 9.0 dpc.

Signaling pathway	Function	p-value
Notch-signaling	organization of somitogenesis	1.59E-03
Wnt/ β -Catenin pathway	somitogenesis, EMT regulation	1.87E-03
NGF-signaling	Neural development	1.08E-02

2.7.2 Clustering of differentially expressed genes from RNA sequencing analysis

RNA-Seq datasets from *Mixl1*, *Tlx2*, *Sct*, *Pax3* and *Foxa2/-T* sorted caudal cells were compared. Cells sorted for *Foxa2* but without *T* expression were used as an endoderm control for *Tlx2* expressing cells, to get insight into possible endoderm development in *Tlx2* expressing tissue in the caudal end (see Chapter 2.7.6).

With this heat map comparison of all 5 domains (Figure 2.14) for shared expression and differences is possible. Gene lists with differentially expressed genes from this table were investigated further in terms of developmental contributions. RNA-Seq data was utilized for clustering using Cluster 3.0 (Stanford University 1998-99). The data was normalized, and a k-means clustering algorithm was used, in which the Pearson correlation was applied to measure relative distances in the data set. Best results were obtained by setting the cluster number to 10. Visualization of the data was performed by using Java Treeview (Saldanha, 2004).

Each cluster represents a combination of upregulated genes or even one predominantly upregulated gene. Biological functions of each cluster are presented in section 2.7.3.

2.7.3 Analysis of clustered genes for biological processes in subregions of the caudal end

Genes with similar expression behavior were clustered together in groups (Figure 2.14). For each group a tendency in relation to function was discovered and listed (Table 2.5). Biological processes analyzed for each cluster were selected by relevance to embryonic development, in which only processes with the lowest possible p-values were listed (complete tables of all processes in the supplementary section 7).

In cluster I, genes expressed in subgroups expressing *Mixl1*, *Tlx2*, *Sct* and *Pax3* are upregulated, whereas genes in subgroup expressing *Foxa2/-T* are mostly downregulated (see cluster I in 2.14). Biological processes involved in somitogenesis (represented by these genes: *Nrarp*, *Dkk1*, *Tbx6*, *Msgn1*, *Dll1* and *Lfng*) and further somite development, as well as anterior/posterior pattern specification (represented by: *Ifitm1*, *Hoxd11*, *Nrarp*, *Dkk1*, *Hoxa11*, *Ets2*, *Hoxd10*, *Tbx6*, *Msgn1*, *Dll1* and *Lfng*) are represented here. Furthermore, processes involved in development of epithelium are active in this cluster.

Cluster II represents genes which are upregulated in *Mixl1*, *Tlx2* and *Sct* expressing cells. In this cluster several tendencies are visible. Many biological processes related to circuit development as cardiovascular system development are active and some genes included in this cluster are *Foxc2*, *Tbx4*, *Hand2*, *Snai2*, *Gata6*, *Hand1*, *Sox6*, *Gata4*, *Fgfr2*, *Twist1*, *Pdgfra*, *Smad6*, *Tbx3*, *Gata2* and *Foxf1*. Moreover mesenchyme development and several processes involved in organ and anatomical structure formation such as digestive system,

2.7 Transcriptome analysis of caudal end subregions

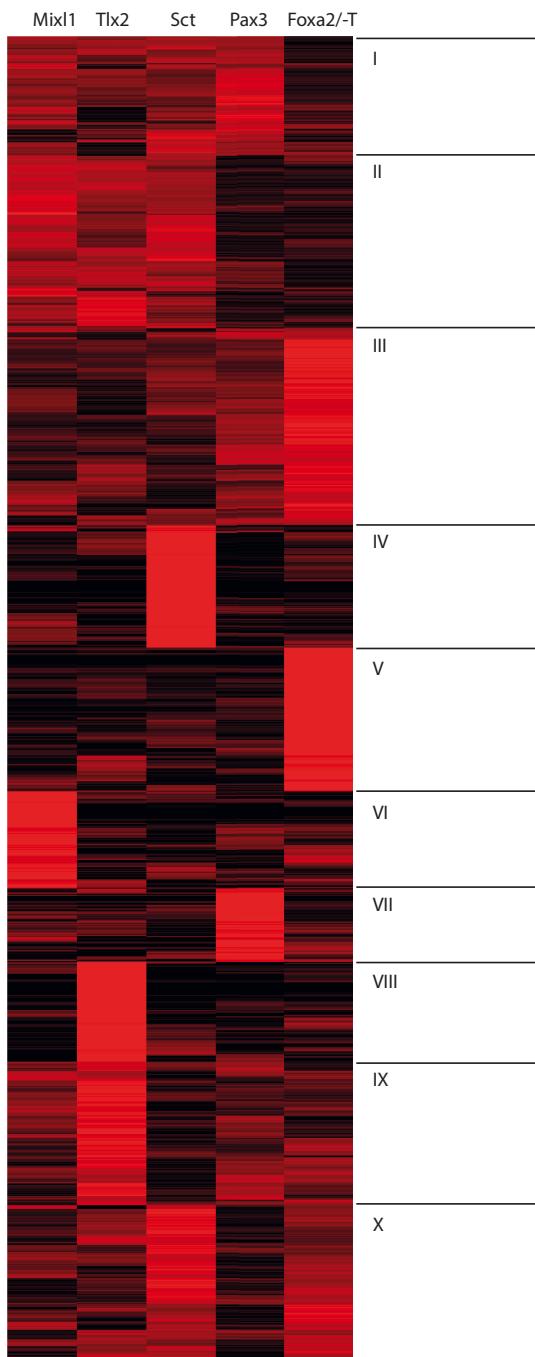


Figure 2.14: Clustered heatmap of differentially expressed genes of subpopulations of cells from the caudal end by RNA-Seq. FPKM values for each gene were normalized and clustered based on euclidian distance. Each row represents one gene and each column represents cells isolated from transgenic embryos using indicated promotor driving Venus expression. Black intensity represents downregulation and red intensity depicts upregulation.

Results

urogenital system and limb development are apparent. Neural crest development is also represented in this cluster. Genes involved here are: *Nrp1*, *Isl1*, *Ednra*, *Twist1*, *Foxc2*, *Hand2*, *Snai2*, *Aldh1a2*. Auxiliary for different developmental steps, endothelial cell migration is also positively regulated in cluster II.

Genes highly expressed in *Foxa2/T* expressing cells and slightly upregulated in *Pax3* expressing subgroups are represented in cluster III. Processes relevant for neural tube development and rostrocaudal tube patterning (*Hes3*, *Pax6* and *Hes1*), as well as locomotion processes are apparent in this cluster. Genes involved into neural tube patterning are: *Zic2*, *Sfrp2*, *Hes3*, *Hes5*, *Pax6*, *Grl3*, *Foxb1*, *Zic5* and *Hes1*. All processes listed for cluster III have very low p-values.

Cluster IV includes genes upregulated in *Sct* expressing subregions. Biological processes analyzed here are all associated with development and differentiation of blood cells and blood related processes like gas transport. Moreover rostral/caudal axis specification of the somites (*Mesp2*) is also related to this cluster.

Development and differentiation of neurons dominates biological processes dedicated to cluster V, for example *Neurog2*, *Sox3*, *Sox2*, *Sox1* and *Shh*. Genes classified in cluster V are upregulated in the *Foxa2/T* expressing subgroup.

Genes upregulated in *Mixl1* expressing cells are ranked into cluster VI. Main processes represented here are biological processes involved in development of the heart, including muscle tissue development. Genes important in these processes are: *Pitx2*, *Wnt2* and *Dand5*.

Cluster VII represents genes upregulated in *Pax3* expressing cells. Genes ranked here are involved in fibroblast growth factor signaling, a process involved in many different developmental processes of organ structures. Furthermore, patterning and elongation of the caudal end (post- anal tail morphogenesis) and regulation of heart induction. Among those genes involved in positive regulation of catenin import into nucleus are *Pou5f1* and *Wnt3a*. Moreover, processes involved in mesoderm formation and BMP signaling are listed in this cluster, represented by *Pou5f1*, *Wnt3a* and *T*.

Genes highly expressed in *Tlx2* expressing cells are integrated in cluster VIII. This cluster displays biological processes, contributing to development of immune cells. *Lgals3*, *Spp1*, *Ccl2* and *Fcer1g* contribute to chemotaxis and migration of neutrophils and granulocytes. Phagocytosis and endocytosis processes are also represented here. These can be processes involved in the development of hematopoietic stem cells in the embryo (Li et al., 2014).

Cluster IX contains genes upregulated in cells expressing *Tlx2* and in lower amounts, *Mixl1*. Biological processes which are included here are all involved in organ development. Genes which are involved in these processes are: *Ngfr*, *Bmp4*, *Hoxd13*, *Hoxc11*, *Cdh1* and *Pax2*. Regulation of mophogenesis of a branching structure is a process participating in development of the uretric bud, a process which is regulated from *Bmp4* and *Pax2*, which

2.7 Transcriptome analysis of caudal end subregions

are expressed in this cluster. For processes listed in this cluster, the p-values were not as significant as for the other groups.

Cluster X includes genes highly expressed in *Sct* and, in a lower amount, in *Foxa2/T* expressing cells. Lowest levels have genes expressed in *Pax3* expressing cells. Processes represented in this cluster are involved in development of cardiovascular system (among these genes: *Foxh1*, *Foxp4*, *Cdh5*, *Wt1* and *Pdgfb*) and epithelium (e.g. of the heart tube). Furthermore, tube development and cell adhesion processes are listed here as well as regulation of locomotion.

Table 2.5: Biological processes for Clusters I to X from RNA-Seq analysis

Biological process	GO #	# of Genes	p-value
Cluster I			
Somitogenesis	GO:0001756	7/62	2.75E-06
Epithelium development	GO:0060429	19/863	2.89E-06
Somite development	GO:0061053	7/70	6.24E-06
Anterior/Posterior pattern specification	GO:0009952	10/204	6.60E-06
Cluster II			
Blood vessel development	GO:0001568	38/420	6.35E-28
Circulatory system development	GO:0072359	46/715	8.62E-28
Cardiovascular system development	GO:0072358	46/715	8.62E-28
Regulation of cell proliferation	GO:0042127	42/1225	7.74E-15
Mesenchyme development	GO:0060485	17/152	2.05E-12
Digestive system development	GO:0055123	14/116	2.06E-10
Urogenital system development	GO:0001655	18/249	4.57E-10
Limb development	GO:0060173	14/165	2.07E-08
Neural crest cell development	GO:0014032	8/53	4.04E-06
Positive regulation of endothelial cell migration	GO:0010595	8/54	4.66E-06
Renal system vasculature development	GO:0061437	5/11	1.18E-05
Cluster III			
Neural tube development	GO:0021915	8/53	1.13E-02
Locomotion	GO:0040011	19/835	1.83E-02
Rostrocaudal neural tube patterning	GO:0021903	3/13	4.26E-02

Results

Biological process	GO #	# of Genes	p-value
Cluster IV			
Erythrocyte development	GO:0048821	4/32	1.42E-03
Somite rostral/caudal axis specification	GO:0032525	3/13	2.71E-03
Erythrocyte differentiation	GO:0030218	5/77	2.87E-03
Erythrocyte homeostasis	GO:0034101	5/86	4.81E-03
Gas transport	GO:0015669	3/20	9.64E-03
Cluster V			
Neurogenesis	GO:0022008	26/1219	2.20E-06
Nervous system development	GO:0007399	30/1626	3.61E-06
Neuron differentiation	GO:0030182	20 /769	5.71E-06
Generation of neurons	GO:0048699	24/1146	1.23E-05
Anatomical structure morphogenesis	GO:0009653	26/1835	4.64E-03
Organ development	GO:0048513	31/2423	4.91E-03
Cluster VI			
cardiac atrium development	GO:0003230	3/29	5.32E-03
atrial cardiac muscle tissue development	GO:0003228	2/5	5.66E-03
extracellular regulation of signal transduc- tion	GO:1900115	2/12	3.22E-02
Cluster VII			
fibroblast growth factor receptor signaling pathway	GO:0008543	4/40	2.05E-04
post-anal tail morphogenesis	GO:0036342	3/23	1.55E-03
regulation of heart induction	GO:0090381	2/7	7.32E-03
positive regulation of catenin import into nucleus	GO:0035413	2/11	1.80E-02
cardiac cell fate commitment	GO:0060911	2/11	1.80E-02
mesoderm formation	GO:0001707	3/60	2.57E-02
BMP signaling pathway	GO:0030509	3/61	2.69E-02

2.7 Transcriptome analysis of caudal end subregions

Biological process	GO #	# of Genes	p-value
Cluster VIII			
neutrophil chemotaxis	GO:0030593	4/39	4.84E-04
neutrophil migration	GO:1990266	4/40	5.34E-04
positive regulation of phagocytosis	GO:0050766	4/42	6.46E-04
granulocyte chemotaxis	GO:0071621	4/46	9.20E-04
granulocyte migration	GO:0097530	4/48	1.09E-03
regulation of endocytosis	GO:0030100	5/155	8.52E-03
immune system process	GO:0002376	13/1389	2.59E-02
Cluster IX			
regulation of morphogenesis of a branching structure	GO:0060688	5/58	1.62E-03
regulation of organ morphogenesis	GO:2000027	6/151	1.67E-02
metanephric collecting duct development	GO:0072205	2/4	2.07E-02
embryonic digit morphogenesis	GO:0042733	4/60	3.07E-02
Cluster X			
blood vessel development	GO:0001568	21/420	6.05E-10
blood vessel morphogenesis	GO:0048514	18/352	1.69E-08
cardiovascular system development	GO:0072358	23/715	2.92E-07
tube development	GO:0035295	16/546	3.70E-04
cell adhesion	GO:0007155	18/685	3.74E-04
regulation of locomotion	GO:0040012	17/615	3.78E-04
epithelium development	GO:0060429	19/863	2.37E-03

2.7.4 Markers for paraxial and lateral mesoderm are expressed in the *Mixl1* population

Paraxial mesoderm flanks the neural tube on both sides, where somites develop from unsegmented presomitic mesoderm by epithelialization and differentiate into the vertebral column and skeletal muscles (Rossant and Tam, 2002). In between paraxial mesoderm and lateral plate mesoderm is the intermediate mesoderm, which gives rise to kidney and some urogenital organs.

In the lineage tracing assay for *Mixl1* descendant cells, progeny was indicated in lateral and paraxial mesoderm (section 2.3). The lateral mesoderm gives rise to heart, blood cells and blood vessels, whereas somites originate from paraxial mesoderm. To understand the role that *Mixl1* expressing cells are playing in the process of formation and organization of structures derived from mesoderm, all genes that were upregulated in the *Mixl1* expressing region and involved in development of paraxial or lateral mesoderm were analyzed. These genes are all represented in cluster II and cluster VI (Figure 2.5). Cluster II represents genes highly expressed in *Mixl1*, *Tlx2* and *Sct* expressing cells, whereas cluster VI represents genes highly expressed in *Mixl1* only. As depicted in Figure 2.15 B, all listed genes are

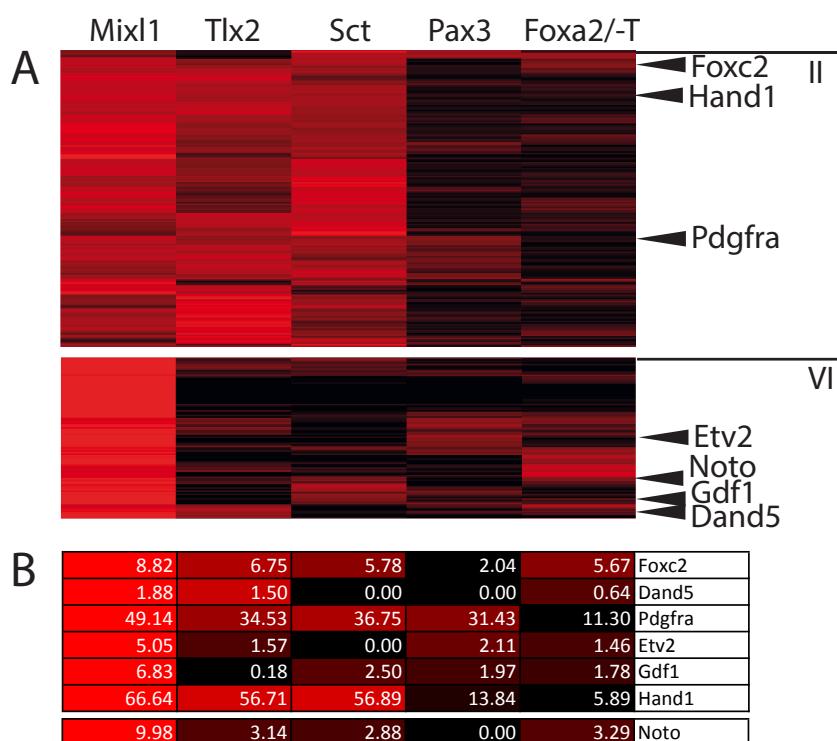


Figure 2.15: Cluster II and cluster VI and FPKM values with depicted genes involved in paraxial and lateral mesoderm development, as well as notochord development

2.7 Transcriptome analysis of caudal end subregions

clearly upregulated in *Mixl1* expressing cells. In *Pax3* and *Foxa2/-T* expressing subregions, these genes are either equivalently expressed or downregulated. Figure 2.15 A shows in which clusters these genes, are included. The expression patterns of genes involved here are shown in Figure 2.16.

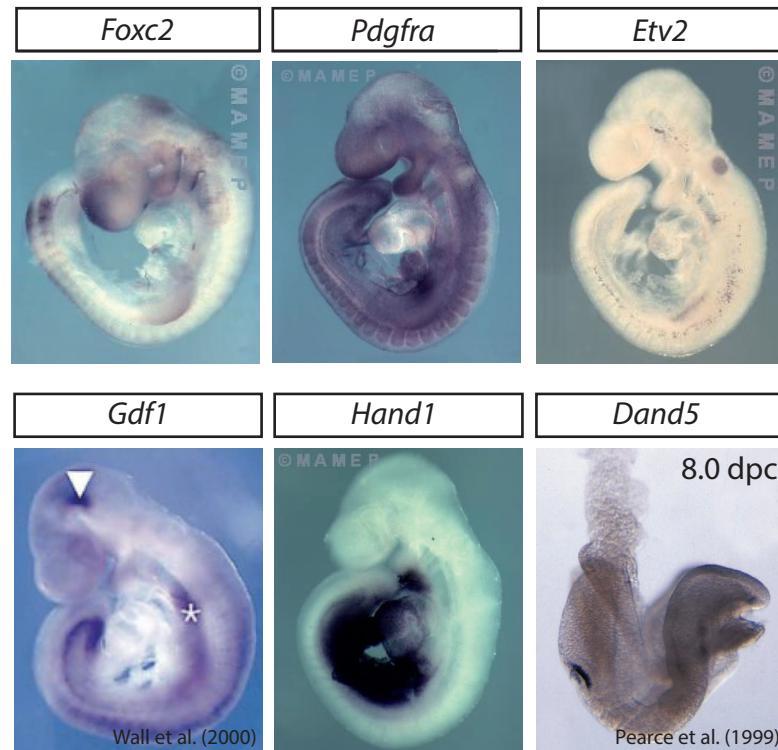


Figure 2.16: Expression patterns of marker genes at midgestation for paraxial and lateral mesoderm formation which are upregulated in the *Mixl1* expressing subpopulation

Foxc2 is expressed in paraxial and intermediate mesoderm. *Pdgfra* is expressed in paraxial mesoderm and somites among other regions. *Etv2* is expressed in paraxial mesoderm and somites. *Gdf1* expression is located in notochord, neural tube, ventrolateral mesoderm and tailbud (Wall et al., 2000). *Hand1* is expressed in trunk mesenchyme, gut, arterial system and heart. *Dand5* expression is not available for at 9.5 dpc embryos, but in 8.0 dpc embryos *Dand5* is asymmetrically expressed in the node. All WISH excluding *Gdf1* and *Dand5* from MAMEP, *Gdf1* from (Wall et al., 2000), *Dand5* from (Pearce et al., 1999).

For specification to paraxial or intermediate mesoderm *Foxc2* is an important factor which acts as a negative regulator of intermediate mesoderm formation (Wilm et al., 2004). *Foxc2* expression is elevated in *Mixl1* and *Tlx2* expressing cells, in which expression is higher in the *Mixl1* expressing subgroup. In *Pax3* expressing cells, *Foxc2* expression is the lowest of all subgroups analyzed.

Dand5 is expressed in the node at 8.0 dpc, is able to bind *Nodal*, and is important for left/right patterning and direct binding of BMP, acting as a BMP antagonist (Marques et al., 2004). *Dand5* shows very low expression levels in *Mixl1* and *Tlx2* expressing cells. In *Sct*, *Pax3* and *Foxa2/-T* cells, *Dand5* is either not or very lowly expressed.

Platelet-derived growth factor *Pdgfra* is involved in many different processes during embryonic development for gastrulation and in the postnatal mouse. It is expressed in many tissues, including paraxial mesoderm (Takakura et al., 1997). *Pdgfra* is expressed at the highest level in *Mixl1* expressing cells, but is downregulated only in the *Foxa2/-T* expressing domain, in the subgroups expressing *Tlx2*, *Sct*, or *Pax3*, *Pdgfra* shows roughly equivalent expression levels.

Etv2 is important in the morphogenesis of blood vessels and induces vascular mesoderm formation (Kataoka et al., 2011). In *Mixl1* expressing tissue, *Etv2* shows its highest expression level. In *Sct* expressing cells, *Etv2* is not expressed, whereas in regions expressing *Tlx2*, *Foxa2/-T* or *Pax3*, it is expressed at a low level.

Gdf1 regulates formation of the body axis during development through induction of mesoderm and is crucial for left/right patterning. *Gdf1* is expressed in ventrolateral mesoderm (Wall et al., 2000). *Gdf1* showed increased expression in only the *Mixl1* cell population, whereas in all other subregions shown here it was downregulated. This effect was strongest in the *Tlx2* expressing domain, where *Gdf1* was not expressed.

Heart and neural crest derivatives expressed transcript 1, *Hand1*, is expressed in lateral mesoderm, heart and neural crest derivatives. *Hand1* is required during vasculogenesis, placental development, and plays a role in both enteric smooth muscle development and ventral body wall closure (Maska et al., 2010). *Hand1* was upregulated in domains expressing *Mixl1*, *Tlx2* or *Sct*, but strongest in the *Mixl1* expressing domain. The *Pax3* and *Foxa2/-T* cell populations show downregulation of *Hand1*.

In the *Mixl1* expressing subregion, a variety of genes are expressed which are necessary for paraxial and lateral mesoderm development (Figure 2.16). This analysis shows the influence of the *Mixl1* expressing region in developmental processes of the caudal mesoderm. However, this analysis reveals a participation of all subgroups tested here in caudal mesoderm development, to some degree. This is not surprising, since signalling and regulation between the different regions in the caudal end are essential for successful development.

2.7.5 *Noto* expression in *Mixl1* expressing cells in the caudal end

In the lineage tracing approach, progeny of *Mixl1* expressing cells are present in the caudal end of the notochord. Since there are many cells visible, it is feasible that genes expressed in the *Mixl1* expressing region are involved in notochord development.

In *Mixl1* expressing cells, *Noto* was found to be upregulated compared to other analyzed subdomains of the caudal end. *Noto* is important for morphogenesis of the tailbud, and is expressed in the notochord at 9.5 dpc (Figure 2.17). *Noto* is arranged in cluster VI, where mainly genes are found which are expressed more highly in the *Mixl1* expressing cells, than in other analyzed subgroups of the caudal end (Figure 2.15 A). Highest expression of *Noto* can be found in *Mixl1* expressing cells (Figure 2.15 B), in all other subgroups *Noto* expression is very low or absent.

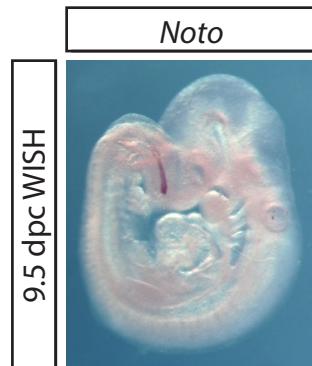


Figure 2.17: Whole mount *in situ* hybridization for *Noto*

Expression of *Noto* by WISH of 9.5 dpc embryo. Image courtesy of Dr. Tracie Pennimpede (MPI-MG).

2.7.6 *Foxa2/-T* expressing cells representing endoderm

Progeny of *Tlx2* expressing cells were found, amongst other places, in the hindgut of the embryos at 9.0 dpc and 10.5 dpc (Figures 2.6 and 2.7). To compare data from the *Tlx2* domain with genes expressed in cells involved in endoderm development, another subgroup of cells was used which express *Foxa2*, but not *Brachyury* (*T*). *Foxa2* is expressed throughout the developing endoderm and in the notochord. Since *T* is also expressed in the notochord, cells expressing *Foxa2* but not *T*, are cells from the endoderm layer important for gut and organ development, excluding *Foxa2* expressing cells of the notochord. A cell line with 3 different fluorescent markers was used to sort these cells (Figure 2.18). This cell line includes the *Brachyury* promotor upstream of the mCherry reporter (Figure 2.18 B), CFP (Cyan fluorescent protein) driven by *Foxa2* (Figure 2.18 C) and Venus controlled by *Sox2* (Figure 2.18 D). This cell line was generated and kindly provided by Dr. Frederic

Koch (MPI-MG). Embryos aggregated from this cell line were dissected at 9.0 dpc and caudal ends without somites were collected, sorted for cells expressing *Foxa2* (CFP) but not *T* (mCherry) using FACS, and used for RNA sequencing.

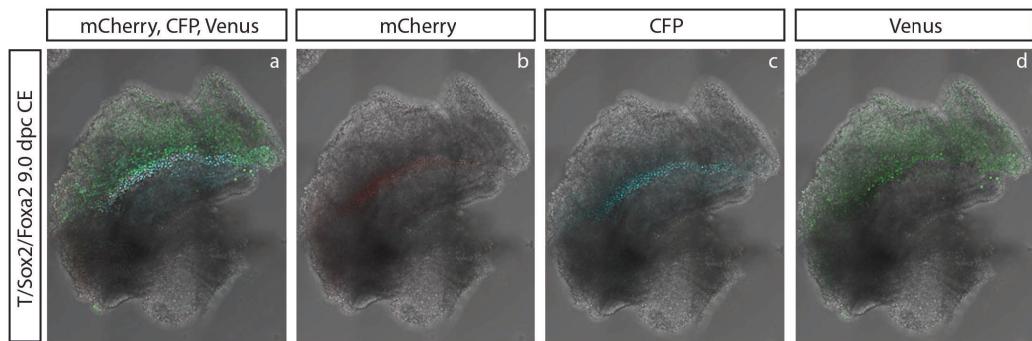


Figure 2.18: Expression of *Foxa2*, *T* and *Sox2* in lateral view of 9.0 dpc caudal ends
 a: overlay of all fluorescent proteins. mCherry under control of *T*, CFP under control of *Foxa2* and Venus under control of *Sox2*. b: mCherry illustrating *T* expression in the notochord. c: CFP representing *Foxa2* expression in notochord and future gut region. d: Venus indicating the *Sox2* expression in neural tube and foregut.

2.7.7 Genes expressed in *Tlx2* expressing cells are involved in endoderm development

From all genes highly expressed in *Tlx2*, or in *Tlx2* and *Foxa2/T* expressing subregions, a few genes were found which also have a function in endoderm development (Table 2.6). FPKM values of these genes are depicted in Figure 2.19 B.

Genes involved in endoderm development are arranged in clusters III and V, as well as in clusters VIII and XI. In clusters III and V genes are clustered which show their highest expression levels in cells expressing *Foxa2/T*. Genes with their highest expression levels in cells expressing *Tlx2*, were grouped to clusters VIII and XI (Figure 2.19 A).

Expression pattern of these genes are depicted in Figure 2.20. All genes discussed here are expressed in the caudal end, dissected for this assay, except *Cdh1*, which is expressed in the endoderm at 8.5 dpc (Stemmler et al., 2003). *Pax9* is a transcription factor which is used as a marker for definitive endoderm (Sherwood et al., 2007). Furthermore it is involved in the cellular response of cells to growth factor stimuli. *Pax9* is only expressed in *Tlx2* expressing cells, in other subregions it is only represented in only very low amounts. Another transcription factor expressed mainly in *Foxa2-T* expressing cells and in lower amounts in all other subgroups of the caudal end shown here is *Cdx2*. *Cdx2* belongs to a group of homeobox genes, responsible for development of antero-posterior orientation in the caudal region of the embryo (Silberg et al., 2000).

2.7 Transcriptome analysis of caudal end subregions

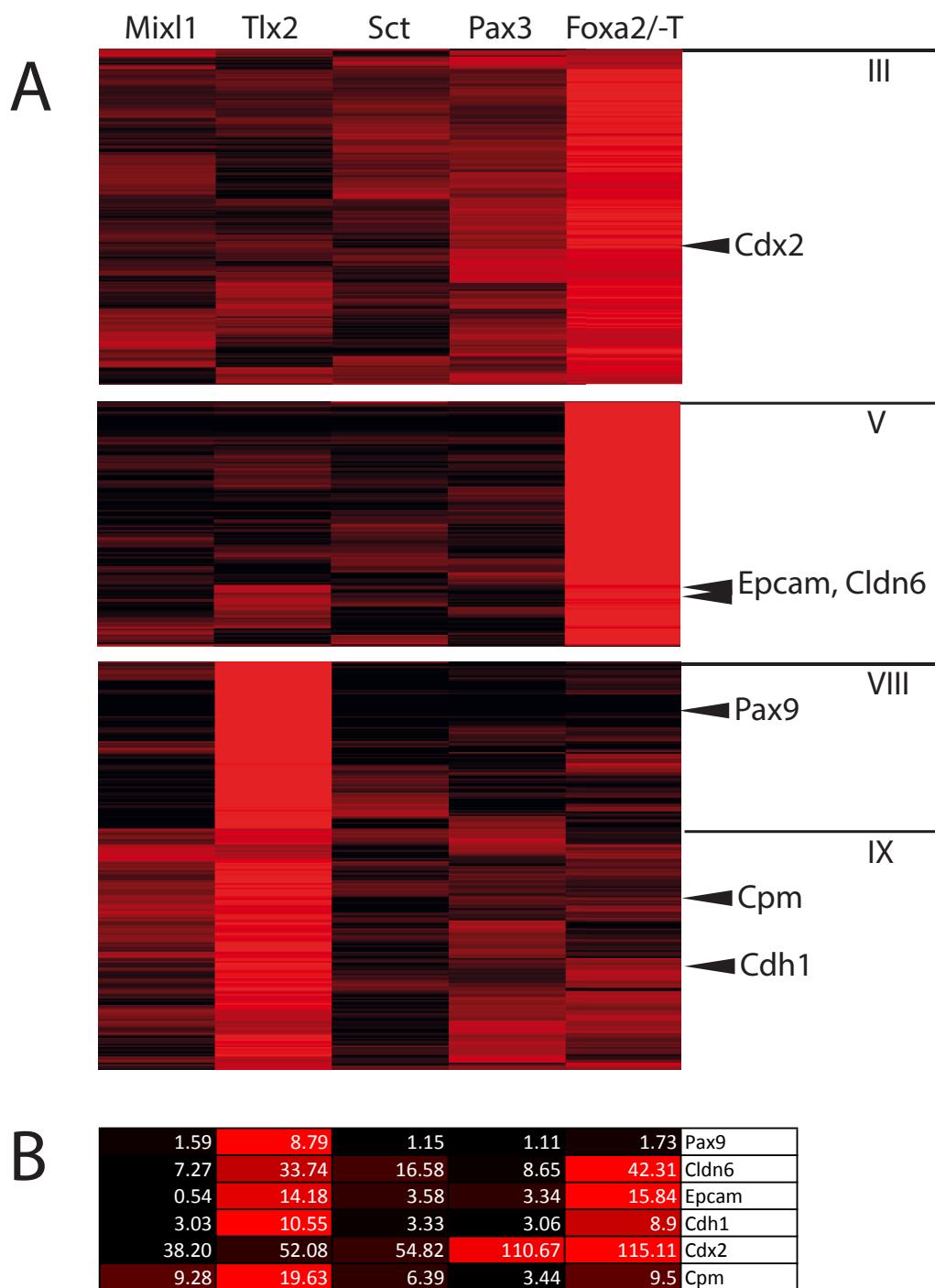


Figure 2.19: Cluster with tagged genes and FPKM values for endoderm development in *Tlx2* expressing cell group.

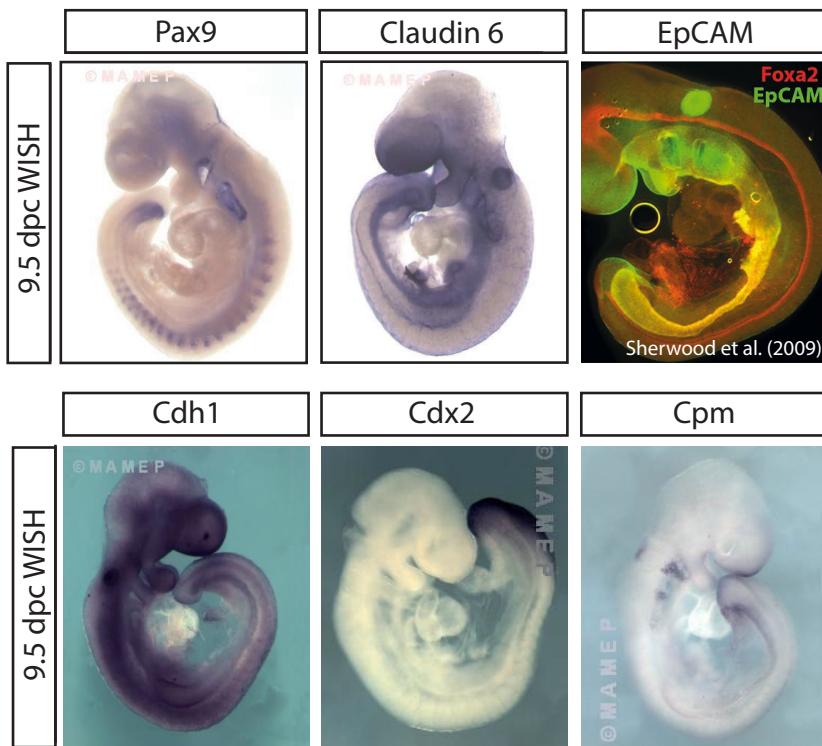


Figure 2.20: Expression of genes involved in endoderm development, which are upregulated in the *Tlx2* expressing cell group

Expression of genes involved in endoderm development and upregulated in the *Tlx2* expressing cell group are shown here. *Pax9* is expressed in branchial pouch, hindgut, paraxial mesoderm, and oral cleft. *Cldn6* is expressed strongest in the gut, in extraembryonic components, otocyst and ectoderm. EPCAM whole mount immunofluorescence, showing EPCAM in green and FOXA2 in red. EPCAM is expressed in endoderm where it overlaps with FOXA2 (Sherwood et al., 2009). *Cdh1* (E-Cadherin) is expressed in brain, branchial arches. *Cdx2* is expressed in the neural tube, caudal end, hindgut and gut. *Cpm* is expressed in branchial pouches, neural tube, paraxial mesenchyme and gut. All images excluding EPCAM are from MAMEP.

Cldn6 is known to be involved in cell junctions, in particular tight junctions. This gene is involved in organization of junctions and is an integral component of the membrane and described as an epithelialization factor in development of embryonic bodies (Turksen and Troy, 2001). Highest expression levels are present in *Tlx2* and *Foxa2-T* subregions, in all other subregions *Cldn6* is also expressed at lower amounts.

Also an integral component of the membrane is the epithelial adhesion molecule *Epcam*. Beside its function in cell adhesion, it is also involved in stem cell differentiation. Like *Cldn6*, *Epcam* is also expressed more highly in *Tlx2* and *Foxa2-T* expressing cells. *Cdh1*, also known as E-Cadherin, is part of the cadherin family, a group of transmembrane

2.7 Transcriptome analysis of caudal end subregions

adhesion proteins. *Cdh1* plays a role in epithelial cell morphogenesis, organization and assembly of tight junctions. Moreover it is part of intestinal epithelial cell development and is a negative regulator of the Wnt signaling pathway. For *Cdh1* the maximum expression levels are present in subregions expressing *Tlx2* and *Foxa2/-T*.

Cpm, Carboxypeptidase M is a membrane coupled component in the membrane, associated with macrophage development. In (Tamplin et al., 2008) it is stated as a novel definitive endoderm marker. Its highest expression is detectable in the *Tlx2* expressing subdomain, in other subgroups *Cpm* shows a weaker expression level (The Jackson Laboratory, Bar Harbor, Maine, <http://www.informatics.jax.org>).

Multiple genes listed here are involved in tight junctions or at least junctions and their assembly or persistence. The impact of the *Tlx2* expressing region to gut development is shown here on the basis of genes involved in tight junctions and gut development (Table 2.6).

Table 2.6: Function of genes involved in endoderm development

Gene symbol	Gene name	Function	Reference
<i>Pax9</i>	paired box 9	endoderm development	(Sherwood et al., 2007)
<i>Cldn6</i>	Claudin 6	epithelial differentiation	(Turksen and Troy, 2001)
<i>Epcam</i>	epithelial cell adhesion molecule	liver development	(González et al., 2009)
<i>Cdh1</i>	E-Cadherin	intestinal epithelial development	(Cano et al., 2000)
<i>Cdx2</i>	Homeobox protein CDX-2	directing intestinal development	(Gao et al., 2009)
<i>Cpm</i>	Carboxypeptidase M	function unknown	(Tamplin et al., 2008)

2.7.8 Genes expressed in *Pax3* expressing cells are involved in neural crest development

Neural crest cells express *Pax3*, and migrate from this region at a later time point. At 9.0 dpc migration markers are not highly expressed but other genes typical for NC development and specification are. Hence, it seems that NC are still developing at this time-point, but migration of NC cells is not taking place so far. In the lineage tracing assay migrating progeny from *Pax3* expressing cells were apparent at 10.0 dpc (Figure 2.12). At 9.0 dpc, the lineage tracing did not show migrating descendants of the *Pax3* expressing subgroup.

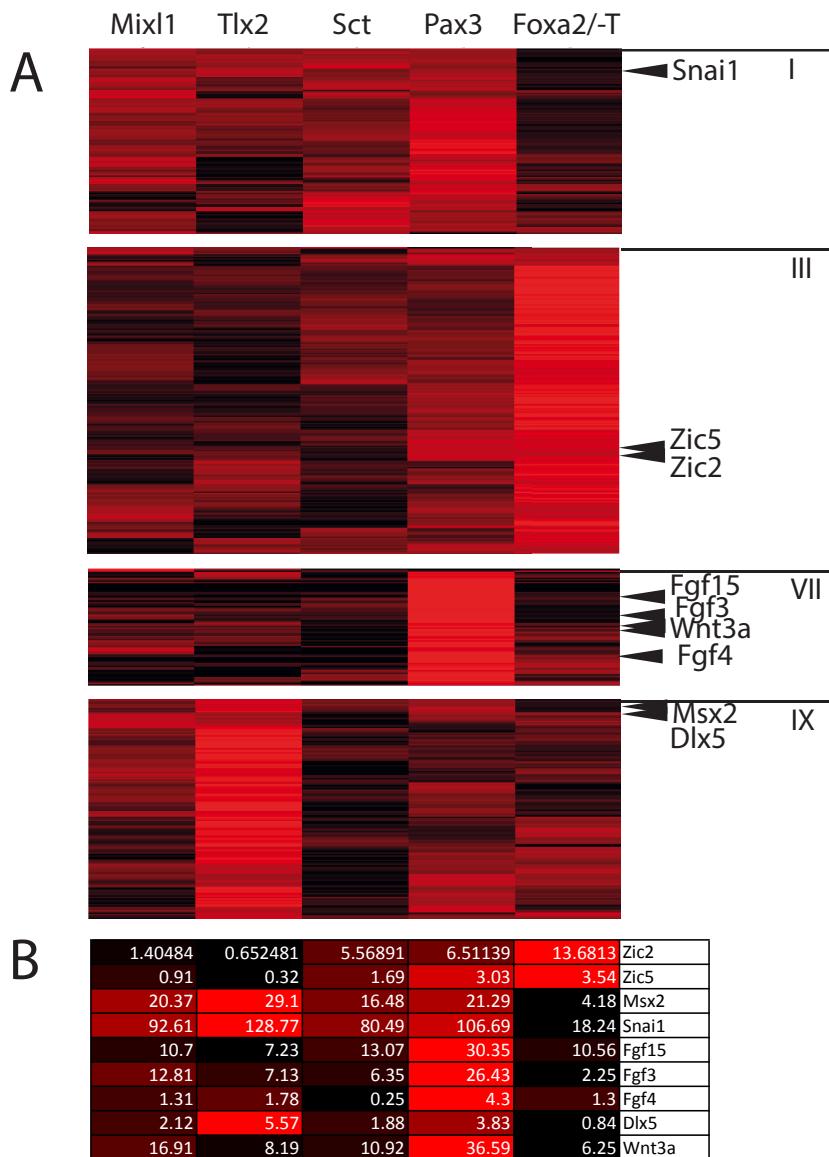


Figure 2.21: Clusters I, III, VII and IX and FPKM values with depicted genes involved in neural crest development.

Expression pattern of genes involved in NC development and expressed in *Pax3* expressing cells, are depicted in Figure 2.22. All genes discussed here are expressed in the caudal end. FPKM values of these genes are depicted in Figure 2.21 B.

Specification of the neural plate border requires the presence of several factors, like *Zic5*, *Msx2* and *Pax3* (Meulemans and Bronner-Fraser, 2004). Expression of *Zic5* and *Pax3* is shown in Figure 2.22. It is stated, that *Fgf4* is able to induce expression of some neural plate border markers, like *Dlx5* (Yardley and García-Castro, 2012). *Dlx5* is an early NC marker, which specifies the neural plate border (Yardley and García-Castro, 2012). *Fgf4*

2.7 Transcriptome analysis of caudal end subregions

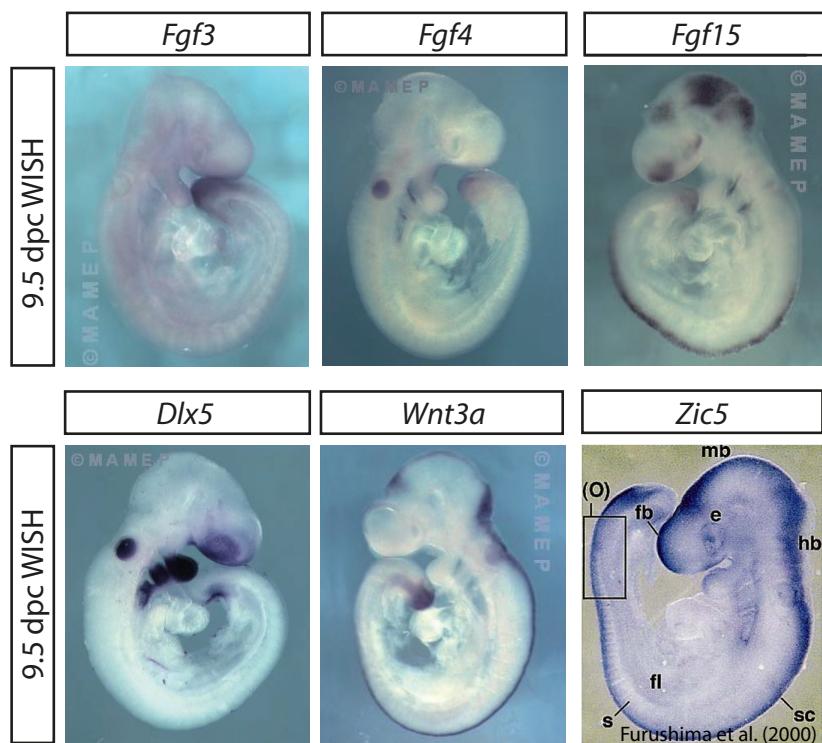


Figure 2.22: Expression pattern of genes involved in neural crest development, which are also upregulated in the *Pax3* expressing subregion of 9.0 dpc embryos WISH of 9.5 dpc embryos from MAMEP, except *Zic5* (Furushima et al., 2000) . *Fgf3* is expressed in the embryo caudal end, paraxial mesoderm, neural plate and caudal neuropore. *Fgf4* is expressed in paraxial mesenchyme, branchial arches, ectoderm and tail mesoderm. *Fgf15* is expressed in paraxial mesoderm, neural tube, future brain, branchial arches and tail. *Dlx5* is expressed in the hindgut, neural tube, ectoderm, frontonasal process, otocyst, forelimb bud and otic pit. *Wnt3a* is expressed in neuroectoderm and tail bud. *Zic5* is expressed in presomitic mesoderm, the dorsal side of somites, future brain and the developing eye.

shows very low expression in all subgroups tested here and has its highest expression in *Pax3* expressing subgroup of cells. *Dlx5* has its highest expression in *Tlx2* and *Pax3* expressing cells, in remaining subgroups the expression is very low.

Zic5 is a member of a family of zinc finger proteins and is equally expressed in the *Pax3* and *Foxa2/T* expressing subgroups. In all other sorted cell groups of the caudal end it is almost not expressed. *Msx2* is a muscle segment homeobox gene, which is a regulator of cell survival of NC derived cells. Here, it is expressed in *Pax3*, *Mixl1*, *Tlx2* and *Sct* expressing subgroups, in which its highest expression is in the *Tlx2* expressing subgroup. *Snai1* is a transcription factor which plays a role in NC specification and mesoderm patterning (Simões Costa and Bronner, 2013; Cano et al., 2000). In *Tlx2* and

Pax3 expressing cells, *Snai1* has its highest expression levels, but it is also expressed both in *Mixl1* and *Sct* expressing cells and *Foxa2/-T* expressing cell groups at a very low level. *Fgf15* is a fibroblast growth factor which is expressed in the neural ectoderm and involved in migration of NC during development of the cardiac outflow tract (Wright et al., 2004; Vincentz et al., 2005).

Fgf3 has its maximum expression in *Pax3* expressing cells and its lowest expression in the *Foxa2/-T* expressing subgroup. In all other subgroups, *Fgf3* is equally expressed. *Fgf3* is involved in neural crest participation in heart development (Urness et al., 2011). Wnt signals play a role in induction of NC, *Wnt3a* is thought to play a role in regulation of proliferation of neural crest, but this has to be tested further (Dongkyun et al., 2010). In this assay, *Wnt3a* is expressed strongest in *Pax3* expressing cells, but is also present in all other sorted subregions of the caudal end.

Table 2.7: Genes involved in neural crest development also upregulated in the *Pax3* expressing subpopulation.

Gene symbol	Gene name	Function	Reference
<i>Fgf3</i>	fibroblast growth factor 3	induction of NC	(Yardley and García-Castro, 2012)
<i>Fgf4</i>	fibroblast growth factor 4	induction of NC	(Yardley and García-Castro, 2012)
<i>Fgf15</i>	fibroblast growth factor 15	induction of NC migration	(Vincentz et al., 2005)
<i>Dlx5</i>	distal-less homeobox 5	NC marker	(Yardley and García-Castro, 2012)
<i>Wnt3a</i>	wingless-type family, member 3A	induction of NC	(Saint-Jeannet et al., 1997)
<i>Zic5</i>	zinc finger protein of the cerebellum 5	NC precursor marker in <i>Xenopus</i>	(Nakata, 2000)

Chapter 3

Discussion

Four different marker genes expressed in the caudal end were used to control expression of a genetically integrated lineage tracing construct. Two different time-points during mouse development were chosen to compare the location of the cellular progeny containing these constructs. Cells were sorted using Venus fluorescence regulated by the promotor of the marker gene and used for RNA-Seq analysis.

The lineage tracing assay revealed two new findings.

Cells expressing *Tlx2* are involved in hindgut development.

The lineage tracing assay revealed descendant cells from *Tlx2* expressing cells in the hindgut. This new finding implicates cells from the caudal end in hindgut development. Genes involved in endoderm development were also expressed in sorted *Tlx2* expressing cells. Together, findings from the lineage tracing and RNA-Seq strongly suggest that cells from the caudal end are contributing to hindgut development.

The *Mixl1* expressing domain of the caudal end participates in notochord formation.

Progeny of cells expressing *Mixl1* were found not only in mesodermal tissues like lateral and paraxial mesoderm, but also in the posterior end of the notochord. Transcription analysis of sorted *Mixl1* expressing cells supported these findings. *Noto*, participating in the development of the posterior notochord was found to be expressed mainly in *Mixl1* expressing cells in the caudal end.

Analysis of RNA-Seq data uncovered the *Pax3* expressing domain in the caudal end contributes to NC development. The *Pax3* expressing domain in the caudal end has not been described so far. This expression domain is thought to be different from the expression domain in the neural tissue, since these cells might be of mesenchymal origin. The role of this subgroup is not well understood as yet.

3.1 *Tlx2* descendants contribute to hindgut development

Current knowledge about development of the hindgut states that there is outgrowth of the hindgut structure in a rostro-caudal direction. However, a contribution of cells migrating later from the caudal end into the hindgut has so far not been described.

In the lineage tracing assay used here, progeny from *Tlx2* expressing cells with activated

lacZ reporter were found in the hindgut at 9.0 dpc (Figure 2.6 C-i, C-ii) and also one day later at 10.5 dpc (Figure 2.7 C-i, C-ii). Even more cells could be visualized in the hindgut at 10.5 dpc, meaning the migration of *Tlx2* expressing progeny into the hindgut started before 9.5 dpc and is probably not complete at 10.5 dpc. To find out more about endoderm development and the contribution of *Tlx2* expressing cells into this process, RNA-Seq data of *Tlx2* expressing cells were analyzed.

Cluster VIII and IX represent groupings with genes mainly upregulated in *Tlx2* expressing cells. Biological processes represented in cluster VIII and IX are mainly processes related to development of the immune system (in Cluster VIII), or processes involved in regulation of organ development. P-values of the processes listed in cluster IX are very high. Genes clustered in these processes could probably also be related to other processes, which were not shown in the performed analysis. Processes of development of the immune system with multipotential stem cells as precursors for mature immune cells are thought to be in the mesenchyme surrounding the heart and in the yolk sac and not in the caudal end (Landreth, 2002).

3.1.1 Involvement of *Tlx2* daughter cells in endoderm development

Since the lineage tracing using *Tlx2* showed descendants of cells expressing the lacZ reporter in the developing gut tube at 9.0 dpc and one day later at 10.5 dpc, participation of these cells in endoderm development is likely. In order to compare the gene expression pattern of sorted cells which express the Venus reporter under control of the *Tlx2* promoter to the gene expression pattern of endodermal cells, we sorted cells from the caudal end at 9.0 dpc which express CFP under control of *Foxa2*, but which are negative for mCherry expression under the control of the *T* promoter.

Foxa2/T sorted cells were expected to represent general endoderm development. At 9.0 dpc *Foxa2* is expressed in the node and notochord, as well as the foregut and hindgut (Sasaki and Hogan, 1993). By using cells which express *Foxa2*, but not the mesodermal marker *T*, the only cells which remain represent the endoderm region. Genes characteristic for endoderm development were expected to be expressed in both cells, cells expressing Venus under control of *Tlx2* and *Foxa2/T* sorted cells.

RNA-Seq analysis revealed genes expressed in both sorted cell pools, which are involved in endoderm development. These genes included *Cldn6*, *EpCAM*, *Cdh1*, *Cdx2* and *Cpm*. All of the genes applied here were assigned to 4 different clusters (Table 2.19). Cluster III, represents genes with highest expression in *Foxa2/T* cells, but which are also expressed in other subgroups of sorted cells. This group includes the gene *Cdx2*. Cluster V represents genes which are mainly expressed in *Foxa2/T* cells, and includes *Cldn6* and *EpCAM*. Cluster VII and cluster IX contain genes with highest expression in the *Tlx2* expressing subgroup of cells, whereas cluster IX includes genes highly expressed in *Tlx2* expressing

3.1 *Tlx2* descendants contribute to hindgut development

cells, and with lower amounts in remaining subgroups. *Pax9* is included in cluster VIII, *Cpm* and *Cdh1* are included in cluster IX.

Only *Pax9* is expressed predominantly in cells expressing *Tlx2*. All other genes are expressed in both cell groups (Table 2.19). Genes representative of indicators for a participation of the *Tlx2* expressing cell group and its progeny into endoderm formation can be distributed into two main groups: 1) Integral membrane components as *Cldn6*, *EpCAM*, *E-Cadherin* and *Cpm*, and 2) transcription factors like *Pax9* or *Cdx2*.

Abundance of integral membrane proteins in *Tlx2* expressing cells

Four genes, *Cldn6*, *EpCAM*, *E-Cadherin*, and *Cpm* are integral components of the membrane.

Cldn6 is a transmembrane protein, located at tight junctions. During embryonic development, *Cldn6* is known as an early marker and activator of epithelialization in embryonic bodies (Turksen and Troy, 2001; Sugimoto et al., 2013). In particular, *Cldn6* is also a marker for stem cells during embryonal development, where it was found to be expressed together with other stem cell markers like *Fut4* (SSEA-1), *Oct4*, and *Nanog*. As soon as the cells start to differentiate, expression levels of *Cldn6* become reduced (Wang et al., 2012). In adult tissue *Cldn6* is not expressed, except for cells in the kidney (Zhao et al., 2008). Here, in connection with cells migrating from the hindgut pocket into the hindgut itself, *Cldn6* expression in cells contributing to hindgut development indicate that there can be stem cells for later epithelial development present in *Tlx2* expressing cells of the caudal end.

At 9.5 dpc *EpCAM* is expressed exclusively in endoderm. During the blastula stage, *EpCAM* is expressed in embryonic stem cells, which are similar to epiblast cells (Sherwood et al., 2007). *EpCAM* is a transmembrane glycoprotein, involved in epithelial cell adhesion, and is expressed in several epithelial cells. Similarly to *Cldn6*, also *EpCAM* is also expressed in progenitor- or stem- cells, in which the role *EpCAM* plays during embryonic development is not fully elucidated. However, it has been shown that *EpCAM* is involved in maintenance of the stem cell type (González et al., 2009). Furthermore *EpCAM* expression in the endoderm is very specific, so that it was also used for cell sorting experiments, to separate between endodermal and mesodermal cells (Sherwood et al., 2009).

E-Cadherin is a Ca^{2+} dependent cell adhesion molecule. During EMT *E-Cadherin* is down regulated in mesenchymal cells to allow migration of the cells (Cano et al., 2000). Though it is expressed in cells of the definitive endoderm and also later in the gut endoderm epithelium, when morphogenesis proceeds (Viotti et al., 2014). *E-Cadherin* is an appropriate indicator for participation of the *Tlx2* expressing cell group in endoderm and gut development.

Tamplin et al. (2008) described another marker for definitive endoderm, *Cpm* (carboxypeptidase M), expressed in the gut at 8.5 dpc. Due to the ventral expression pattern of *Cpm*, this gene is thought to be involved in dorso-ventral patterning of the gut. Besides this, *Cpm* is a membrane anchored enzyme, which catalyzes hydrolysis of C-terminal amino acids (Tamplin et al., 2008).

Endoderm specific transcription factors in *Tlx2* expressing cells

Two transcription factors, *Pax9* and *Cdx2*, expressed in *Tlx2* expressing cells were identified as genes involved in endoderm development. Whereas *Pax9* is expressed mainly in *Tlx2* expressing cells, *Cdx2* is expressed mostly in *Pax3* and *Foxa2/T* expressing cells along with all remaining cells, including the cells sorted for *Tlx2* expression.

Pax9, a transcription factor of the Pax family is known to be expressed in the paraxial mesoderm, somites, branchial pouch and also in the epithelium of the hindgut (Neubüser et al., 1995). Moreover, *Pax9* was described as a marker of definitive endoderm, since its expression is higher in definitive endoderm, than in visceral endoderm (Sherwood et al., 2007). *Pax9*, a regulator of the foregut, is involved in pharyngeal (Lewis and Tam, 2006), craniofacial and limb development. During tooth development *Pax9* is involved in BMP signaling (Peters et al., 1998). The impact of *Pax9* seems to be wide and diverse. In endoderm development, *Pax9* is involved in development of thymus and parathyroid glands.

In embryos with loss of *Cdx2*, the *Pax9* expression domain is expanded into the posterior gut tube, at 14.5 dpc *Cdx2* promotes intestinal fate, and acts as an antagonizing factor to the foregut differentiation program (Gao et al., 2009). At 9.5 dpc *Cdx2* is expressed in the neural tube, and the posterior gut endoderm, including hindgut (Chawengsaksophak et al., 2004). FPKM values from RNA-Seq data indicate expression of *Cdx2* in the caudal end, with differences in the particular domains of cells sorted for this experiment. Notably, in later embryonal development, *Cdx2* is involved to intestinal epithelial development (Silberg et al., 2000).

Two genes of transmembrane proteins presented here, *EpCAM* and *Cldn6*, are expressed in stem cells during development. Furthermore, *Cpm* and *E-Cadherin* are two other transmembrane proteins. *Cpm* is specifically expressed in the endoderm during development, and *E-Cadherin* is expressed in the definitive endoderm and in the gut epithelium later. Besides these integral membrane proteins, two transcription factors, also involved in later endoderm formation were indicated from the RNA-Seq analysis. Taken together these findings support the idea that cells from the caudal end which express these markers for endoderm formation, contribute to gut development by migrating into the hindgut during embryonic development.

3.1 *Tlx2* descendants contribute to hindgut development

3.1.2 ***Tlx2* daughter cells do not contribute to the development of the enteric nervous system**

At embryonic day 9.5 dpc sacral neural crest cells start to migrate from the neural tube into the hindgut to form the enteric nervous system (Wang et al., 2011). *Tlx2* is an important factor, crucial for normal development of the enteric nervous system (Borghini et al., 2007). *Tlx2*^{-/-} mice first develop normally, but obtain megacolon a few weeks later. Megacolon is a hyperinnervation of the enteric nerves, caused by missing ganglia in the enteric nervous system (Hatano et al., 1997; Newgreen and Young, 2002). In embryos, *Tlx2* is expressed in the dorsal root ganglia and enteric nerve ganglia (Hatano et al., 1997).

Progeny of lacZ reporter expressing cells regulated by the *Tlx2* promoter were found in the hindgut. Since it is known that cells migrating into gut to form the enteric nervous system enter the foregut region (Cheeseman et al., 2014), the latter cells are most probably no cells contributing to the enteric nervous system. It seems that cells migrating into the hindgut originated from the caudal end region and not from dorsal root ganglia. It is possible that migration of cells for formation of enteric nervous system occurs at a later time-point, which was not examined here.

3.2 Noto expression in *Mixl1* expressing cells

The lineage tracing assay for *Mixl1* descendant cells showed many descendants to be located at the caudal end of the notochord (Figure 2.4). Highest expression of *Noto* can be found in *Mixl1* expressing cells (Table 2.15 B), in all other subgroups Noto expression is very low or absent like in *Pax3* expressing cells.

The notochord structure is important for patterning of the neural tube and arises from the node. Between 6.5 and 7.0 dpc *Mixl1* expression can be detected in the ventral side of the node (Wolfe and Downs, 2014). LacZ expressing descendants of *Mixl1* expressing cells could be either cells from the node participating in the notochord structure, or cells from the caudal end which also contribute to the notochord, as it is assumed by Yamanaka et al. (Yamanaka et al., 2007). Embryos with deleted *Noto*, form a notochord, but not in the tail region (Ben Abdelkalek et al., 2004). *Noto* expressing cells are found to contribute to the posterior part of notochord and are important for tail bud morphogenesis (Yamanaka et al., 2007; Zizic Mitrecic et al., 2010). *Mixl1*^{-/-} embryos also do not develop a normal notochord structure, in place of notochord they exhibit a structure of condensed mesoderm, which ends caudally with a bulge like structure (Hart et al., 2002).

Expression of *T* is, amongst other processes, necessary for proper development of the notochord (Kispert and Herrmann, 1994). Moreover, the expression domain of *T* is expanded in embryos lacking *Mixl1*, and Pereira and colleagues (2011) could verify a direct interaction between *T* and MIXL1 protein in EB (Pereira et al., 2011). From current and previous findings we can suggest that *Mixl1* expressing cells are contributing to notochord development and are possibly participating in the *T* expression network, involved here.

3.3 *Mixl1* expressing cells contribute to development of paraxial and lateral mesoderm

Mixl1 is a downstream gene of the Activin/Nodal pathway which is important for primitive streak development and induction of endoderm and mesoderm (Pereira et al., 2012; Conlon et al., 1994).

Expression of *Mixl1* can first be found at 5.5 dpc in the visceral endoderm. It is also expressed in epiblast, nascent mesoderm and primitive streak at 6.5 dpc (Robb et al., 2000). A faint domain of expression can be detected in the tail bud mesoderm (Figure 2.3), and is visible until 11.5 dpc (Pearce and Evans, 1999).

In the lineage tracing assay shown here, *Mixl1* expressing descendants were located in the paraxial and lateral mesoderm of the embryo (Figure 2.4). Cells were scattered throughout

3.3 Mixl1 expressing cells contribute to development of paraxial and lateral mesoderm

the mesoderm, and clusters of cells in the mesoderm were not existent. The paraxial and presomitic mesoderm neighbor the neural tube. Unsegmented presomitic mesoderm is the precursor tissue for the somites. Organs derived from somites are skeletal muscle, ribs and vertebrae. As *Mixl1* null mutants only develop until 9.0 dpc, restrictions in development from this tissue are not visible, since somites are formed regularly (Hart et al., 2002).

More descendants of *Mixl1* expressing cells can be traced in the lateral mesoderm (Figure 2.4). Structures originating from the lateral mesoderm are heart, blood vessels and blood cells. In *Mixl1* null mutants no heart tube develops, which verifies the involvement of *Mixl1* in heart development (Hart et al., 2002).

Thus, already known findings from literature, along with the analysis of migrating cells using the lineage tracing assay, support a mesoderm patterning function of *Mixl1*.

RNA sequencing analysis illustrates tendencies of *Mixl1* expressing cells

The RNA sequencing analysis from sorted *Mixl1* expressing cells at 9.0 dpc gives insight into pathways which are activated in this cellular subtype at this stage of embryo development. RNA sequencing gives insight into genes which are expressed in *Mixl1* expressing cells of the embryo. Pereira and colleagues (2011) demonstrated the ability of *Mixl1* to bind several other proteins, which are important during embryonic development (Pereira et al., 2011). For example, in *Mixl1*^{-/-} embryos it was shown already, that *T* is upregulated when *Mixl1* is absent. Direct inhibition of *T* by *Mixl1* would require physical contact between the proteins, which was shown in an in vitro ES cells system, where *T* and *Mixl1* can bind each other (Pereira et al., 2011). Another in vitro assay gives a similar outcome, in an ES culture system, inhibition of *T* is shown following RNAi -mediated knockdown of *Mixl1* (Izumi et al., 2007). But, expression detected by RNA-Seq here shows normal levels of *T* expression in all subregions investigated. Canonical pathways from cells expressing *Mixl1* indicated the expression of genes implemented in signaling pathways involved in somitogenesis (Wnt and Notch), BMP signaling, and EMT.

Genes which are both annotated for development of lateral and paraxial mesoderm and expressed in *Mixl1* expressing cells, are involved in different mesoderm patterning tasks during development. Blood and blood vessels are generated from paraxial and lateral mesoderm in the caudal end (Dumont et al., 1995). This domain covers *Mixl1*, *Tlx2* and *Sct* expressing subregions of the caudal end. BMP, Notch and Wnt signaling is required for regular function of blood and blood vessel formation (Lee et al., 2008). Furthermore, some genes including *Etv2*, *Hand1* expressed here are participate in heart, blood and smooth muscle development (Wareing et al., 2012; Maska et al., 2010). Another factor involved in hematopoiesis and expressed in paraxial mesoderm is *Pdgfra*, the platelet-derived growth factor receptor α . *Pdgfra* is a receptor, which binds all kinds of platelet derived growth factors and is important for multiple processes during embryonic organ development

(Takakura et al., 1997). In human embryonic stem cells, *Pdgfra* and *Mixl1* are highly enriched, and co-expressed (Davis et al., 2008). *Etv2* is expressed in *Pdgfra* expressing potential hematopoietic tissue and gives rise to endothelial and hematopoietic cells (Ding et al., 2013). Moreover also *Foxc2*, a forkhead transcription factor was found to be essential for transcription of genes needed for development of anterior presomitic mesoderm and is involved in heart and blood vessel development as well as in somitogenesis (Kume et al., 2001).

Besides the activation of genes involved in hematopoietic development, cells expressing *Mixl1* also show activation of genes involved in regulation of left and right patterning. *Gdf1*, a Notch regulated gene which is part of the network responsible for left-right signaling is expressed here together with *Dand5* (*Cerl2*) (Kitajima et al., 2013; Marques et al., 2004). Characteristic for *Dand5* is the differential distribution of mRNA and protein in the left-right embryo, which is dynamic in the node structure (Inácio et al., 2013). *Gdf1* is also expressed in the crown cells of the node (Tanaka et al., 2007). Together these genes are involved into left-right patterning, which occurs earlier in the embryo at 7.5 dpc.

A clustering analysis revealed groupings of biological processes within *Mixl1* expressing cells, which absolutely suit previous findings. Somitogenesis or axial patterning processes, as well as processes involved in blood circulatory system and heart development were specified. All these findings support the participation of *Mixl1* in these processes, although direct modulation of processes or genes can not be shown with this assay.

3.4 *Mixl1* expressing cells and their contribution to development

Cells from the *Mixl1* expressing subgroup of cells are contributing to both development of paraxial and lateral mesoderm and the posterior end of the notochord. The processes these cells are contributing to are diverging. Therefore, the question arises whether cells contributing to such divergent processes are from the same origin. There is the possibility, that the domain of *Mixl1* expressing cells exists as two different subgroups of cells, contributing to either notochord, or mesoderm development. The other option is, *Mixl1* expressing cells are mesodermal precursor cells, which can contribute to paraxial, lateral and axial mesoderm. Overall these characteristics are not well understood as yet.

3.5 ***Pax3* descendants contribute to neural crest development**

The lineage tracing essay shown here exhibits lacZ expressing progeny of *Pax3* expressing cells at 9.0 dpc in the caudal end mesoderm, neural tube and later, at 10.0 dpc also in ganglia. The role of *Pax3* during myogenesis and during neural plate border specification is widely known (Williams and Ordahl, 1994; Serbedzija and McMahon, 1997).

At 9.0 dpc *Pax3* was expressed in the extension of the neural tube in the caudal end of the embryo. The expression domain of *Pax3* in the tail bud is probably paraxial mesoderm. It is reasonable to assume that the function of cells expressing *Pax3* in the tail bud mesoderm is separate from those in the neural *Pax3* expression domain. The function of these cells in the caudal end has not been studied yet. If these cells in the caudal end are from the same origin as *Pax3* expressing cells from the neural tube, is not known. Since *Pax3* is known to be involved into the development of NC, RNA-Seq data was analyzed for genes relevant to this issue.

Nine genes were found to contribute to NC development. Most of the genes (*Fgf15*, *Fgf3*, *Wnt3a* and *Fgf4*) are included in the same cluster, cluster VII, which contains genes which are highly expressed in *Pax3* expressing cells, but with much lower level of expression in all other analyzed cell populations. *Snai1* is classified to cluster I where all cell subtypes except *Foxa2/T* show expression, *Zic5* is expressed in cluster III, where genes highly expressed in *Foxa2/T* cells are clustered, but with high expression in some other subgroups. Cluster IX contains genes highly expressed mainly in *Tlx2* expressing cells, but genes from this cluster are also upregulated in *Mixl1* or *Pax3* expressing cells, like *Msx2* and *Dlx5*.

Biological processes analyzed in clusters with genes expressed highly in *Pax3* expressing cells are often not related to processes *Pax3* is directly involved in as processes of neural development. These are present in cluster I, VII and partly in cluster III and IX. Cluster III includes processes related to neural development, since the p-values of these processes are relatively high, the correlative power is lower.

Fibroblast growth factors and their impact to NC cell activation

Fibroblast growth factors found here (*Fgf15*, *Fgf3* and *Fgf4*) are all involved in the induction of NC cells in chick and zebrafish. Since *Fgf* genes are involved in induction of other signaling pathways which are crucial for NC induction, a participation in NC formation in mouse is feasible (Yardley and García-Castro, 2012). It has been shown in chick that *Fgf4* is able to induce *Pax7*, *Msx1* and *Dlx5* at the non neural ectoderm, by interaction with BMP and Wnt pathways (Yardley and García-Castro, 2012).

Fgf15 is from 9.0 dpc on involved in cardiac modeling of the outflow tract. Here cardiac

neural crest are crucial for proper formation of the outflow tract. *Fgf15* plays a role in the migration pathway of the cardiac neural crest, in *Fgf15^{-/-}* mice, a defect in neural crest migration could be observed (Vincentz et al., 2005). Beside its function in pharyngeal arch and heart development (Wright et al., 2004), *Fgf15* could be also involved into NC development of the caudal end, as it is differentially expressed here. There is a clear up regulation of *Fgf15* apparent in *Pax3* expressing cells (Table 2.21 B).

Fgf3 is expressed in different tissues in the embryo during development, as well as in the caudal end at 9.0 dpc. In chick it was tested, whether *Fgf2*, *Fgf3*, *Fgf4* or *Fgf8* were able to induce NC alone, but it was determined that they could not (Yardley and García-Castro, 2012). Fgfs are a small piece in the NC cell activation network, and since there are many family-members, it is likely that functions of the *Fgf* genes can be compensated for by other family members (Yardley and García-Castro, 2012).

Common members of the NC cell developmental network

Different genes participating in NC development could be detected in RNA-Seq data. In general, all genes shown here were expressed at a very low level in the mouse caudal end. It is possible, that this is due to the status of NC development in the caudal end at the time-point chosen for the analysis.

Before specification of the neural border takes place, markers for the neural plate border are expressed, such as *Dlx5*. It has been shown to be expressed in the neural plate (McLarren et al., 2003) and appeared mainly in *Tlx2* and *Pax3* expressing cells in the RNA-Seq data presented here. However, *Dlx5* was described as an early neural crest marker, expressed at the neural plate, but does not contribute to NC induction (Gammill and Bronner-Fraser, 2003).

The transcription factor *Snai1* is involved in EMT (Carver et al., 2001). During this process *Snai1* is a repressor of *E-Cadherin*, a transmembrane protein responsible for cell adhesion (Cano et al., 2000). On the one hand, *Snai1* is used as a marker for the early mesodermal lineage. On the other hand, *Snai1* functions as a neural crest marker, when NC cells are developing. In *Xenopus* the *Snai1* equivalent is required for NC specification and induction and as well for migration of the NC (Aybar, 2003; LaBonne and Bronner-Fraser, 2000). Here, *Snai1* was expressed in all sorted cells, but the highest expression was indicated in *Tlx2* and *Pax3* expressing cells. This expanded expression pattern is possibly due to the diverse functions of *Snai1* during EMT and NC cell development.

Wnt3a is expressed in the primitive streak and caudal neural plate and plays a role in somitogenesis, notochord and tail bud formation (Takada et al., 1994; Takemoto et al., 2011). Furthermore, *Wnt3a* is involved in neural tube patterning and in the differentiation process of NC cells (Saint-Jeannet et al., 1997). In mouse ES cells it was shown, that *Wnt3a* plays an important role during induction of the NC cells. Moreover *Wnt3a*

3.5 *Pax3* descendants contribute to neural crest development

maintains the undifferentiated state of NC cells during induction (Fujita et al., 2014). Here, *Wnt3a* was expressed in all cells sorted, but is highest in *Pax3* expressing cells. In the *Pax3* expressing region in the caudal end *Wnt3a* is possibly involved in arresting prospective NC in an undifferentiated state.

Participation of *Zic* genes in development of neural crest was first examined in *Xenopus*. *Zic1*, *Zic2* and *Zic3* were described by their function during induction of NC (Nakata et al., 1998). *Zic5* was identified as a gene, involved in development of the cerebellum. In *Xenopus* *Zic5* was discovered as an important factor inducing NC development, where *Zic5* appears as a marker for prospective NC cells (Nakata, 2000). *Zic3* is expressed in the primitive streak (Nagai et al., 1997) and expressed in all subgroups shown here, therefore it was not included in the cluster analysis. Overall, the presence of these transcripts in the *Pax3* expressing subgroup suggests a role for *Pax3* in specifying NC precursors.

3.5.1 *Pax3* subdomain: Origin of neural crest precursors?

As genes expressed in the cells isolated from the expression domain containing *Pax3* in the caudal end indicate, this domain in the caudal end is probably involved in development of NC. As NC are part of the neural anlagen and delaminate from the neural folds, the tail bud *Pax3* expression domain is particularly interesting. Cluster VII, where genes upregulated in *Pax3* cells are grouped indicated an mesodermal origin of cells, since typical mesodermal genes are expressed here, for example *Pou5f1*, *Wnt3a* and *T*. This expression domain in the tail bud is not yet part of the *bona fide* neural tube. As there are no markers available in mouse which are unique for prospective NC, it can be stated finally, that NC specification can occur in the caudal end. In chicken, *Slug* is expressed in prospective NC cells (Gammill, 2002), but in mouse the equivalent gene is not crucial for NC development and not expressed in NC precursors (Jiang et al., 1998). *Zic5* could be a marker for prospective NC, but its role in mice still has to be investigated.

Expression levels of some of the genes represented here were very low. Since we were not investigating several time points using RNA-Seq, these genes might be upregulated later, or were already down regulated. Some *in situ* hybridizations of genes found here also depicted weak signals, which would indicate that the expression of these genes is very low in the *Pax3* expressing subdomain of the caudal end. All genes found here in connection with NC development indicated that induction and specification of NC development took place, but migration of the NC had not started yet.

It is probable that the tail bud expression domain of *Pax3* represents a pre-patterning process for the elongation of the neural tube. *Pax3* is known to be involved in a signaling network with β -catenin and *Cdx2*, which is crucial for elongation of the neural tube (Zhao

Discussion

et al., 2014). Transcriptome data support this idea, as patterning and elongation of the tail are biological processes where genes upregulated in Cluster VII are involved in.

3.6 Characteristics of *Sct* expressing cells

The expression pattern of *Sct* in the caudal end at 9.0 dpc is very local, overlapping with the position of the ventral ectodermal ridge (VER). LacZ expressing progeny of *Sct* expressing cells were discovered distributed through the mesoderm, in gut epithelium, ganglia and somites. Biological processes which appeared in the analysis of cluster IV with genes highly expressed in *Sct* expressing cells exhibit a contribution of genes expressed in these cells in the development of erythrocytes.

Genes characteristic for VER were also expressed in all other investigated subgroups shown here, like *Bmp2* (Goldman et al., 2000). *Bmp2* is expressed broadly in the caudal end with no significant up regulation in *Sct* expressing cells. Since there are no characteristic genes for this tissue apparent in the list of genes with highest expression in *Sct* expressing cells, it is difficult to say whether there is a cluster of genes which truly defines the VER. Given that many genes expressed in *Sct* expressing cells and all other subgroups are involved in somitogenesis, *Sct* expressing cells appear as a component of the mesoderm, without any notable functional distinctions.

3.7 Conclusions

Analysis of different cellular subgroups in the caudal end of the embryo reveal diverse characteristics. *Mixl1* is important for contribution to mesoderm patterning. The lacZ expressing progeny of *Mixl1* expressing cells could be detected in lateral and paraxial mesoderm, a structure, where heart, blood vessels and blood cells originate from. In addition, progeny was indicated in the posterior notochord. The analysis from sorted *Mixl1* expressing cells exhibited genes upregulated in these cells, required for blood and blood vessel formation. Biological processes, mainly represented in sorted *Mixl1* expressing cells also display processes involved in blood circulatory system formation or heart development. Upregulation of *Noto* in *Mixl1* expressing cells, points to a participation of these cells into notochord development.

Tlx2 is, like *Mixl1* also involved in mesoderm patterning. The lineage tracing assay indicated lacZ expressing progeny of *Tlx2* expressing cells in the hindgut. Cells contributing from the caudal end to hindgut development are a novel finding which has not been reported so far. The RNA seq analysis revealed integral membrane proteins and transcription factors involved in endoderm development. These were upregulated in sorted *Tlx2* expressing cells. These essential facts support the hypothesis of cells from the *Tlx2* expressing subregion contributing to hindgut development.

Nothing is reported yet about the subgroups of *Pax3* expressing cells in the caudal end. Due to their participation into NC development it was considered, that prospective NC

cells could occur here. Several genes, upregulated in sorted *Pax3* expressing cells, support this idea. To date there are no marker genes for prospective NC to date. Hence, one can assume from the genes involved into NC development and which are upregulated here, that there are cells present in this subgroup, contributing to NC cells.

3.7.1 Outlook

In order to investigate further details of the mentioned subgroups, different experimental setups are possible.

Inducable overexpression of genes could help to better understand characteristic correlations of developmental processes. Ablation of descendant cells are supposable to investigate, whether these cells have an impact in embryonic development. Besides these, deeper analysis of RNA seq data with subsequent testing using real time PCR would be workable.

Chapter 4

Methods

4.1 DNA work and cloning

4.1.1 Polymerase Chain Reaction (PCR)

Polymerase chain reactions were performed using PrimeSTAR (TAKARA), Phusion (Fisher Scientific) or Expand long template system (Roche), according to manufacturer's instructions.

4.1.2 Cloning vectors and bacterial strains

DNA fragments were purified and blunt-end ligated into EcoRV-linearised pBluescript II SK+ (Stratagene), using T4 Ligase (NEB). Ligated vectors were transformed into chemical competent DH5 α *E. coli*. Fragments for recombineering were ligated into the R6K vector and transformed into DH10 β F'DOT chemically competent *E. coli*.

4.1.3 Lineage tracing construct

All parts of the construct were subcloned into pBluescript II SK+, and the cassette was then cloned into the R6K vector. The original R6K vector contains a kanamycin resistance, which had to be exchanged for a ampicillin resistance, because a neomycin resistance cassette, which also provides resistance to kanamycin, is used in the lineage tracing cassette. The PL451 vector was used as a template for the Neo cassette flanked by two FRT sites driven by a prokaryotic em7 and a eukaryotic PGK promotor. For recombineering of the cassette into the BACs, primers were generated which contained 20-25 bp homologous to the ends of the targeting construct, along with 50-75 bp of sequence immediately 3-prime and 5-prime to the integration point, following the ATG of each gene from the BAC sequence. This generated a PCR product which contained the cassette flanked by regions homologous to the gene of interest. The cassette flanked by the gene homologous regions was then used for RED/ET recombineering. Exchange of Dre recombinase for Cre recombinase was performed by Dr. Heiner Schrewe (MPI-MG), cassette cloning was done in cooperation with Dr. Mikhail Sukchev (MPI-MG).

4.1.4 RED/ET recombineering

RED/ET recombineering was performed using pSC101 gbaA plasmid employing all Bacteriophage λ genes which are essential for recombineering. pSC101 gbaA only replicates at 30 °C (Zhang et al., 1998). For recombination of the lineage tracing construct from the R6K vector into the BAC vector (bacterial artificial chromosome), bacteria containing the BAC were grown in overnight cultures, and fresh LB media was inoculated with 30 μ l of overnight culture. Bacteria were grown until OD₆₀₀ ~0.3, centrifuged at 10 000 rpm at 4 °C for 30 seconds. The pellet was resuspended with 1 ml cell culture grade water, and the wash step repeated. Supernatant was discarded by tilting, ~100 ng pSC101 gbaA plasmid was added to competent cells and electroporated at 200 Ω , 1800 V, and 25 μ F. Bacteria were plated on agar plates containing 34 μ g/ml chloramphenicol and 10 μ g/ml tetracycline, and were incubated at 30 °C.

Bacterial colonies containing BAC and pSC101 gbaA plasmid were picked and grown overnight at 30 °C . One ml of fresh media was inoculated with these cultures and grown until OD₆₀₀ ~0.2.-3.0 in LB without antibiotics. Once the cells reached the linear growth phase, 30 μ l of 10% arabinose was added to induce recombineering proteins from the pSC101 gbaA plasmid, which contains an arabinose inducible BAD promotor (Zhang et al., 1998). Cultures were now grown at 37 °C for 45-60 min, to constrain the R6K vector from further replication. Cultures were then transferred to 4 °C and centrifuged at 10 000 rpm for 30 sec. The pellet was resuspended in ice-cold cell culture grade water, and this washing step was repeated twice. In the final step the supernatant was discarded by inverting the tube, leaving about 50 μ l water in the tube. About 300 to 600 ng purified PCR product were added and the mixture electroporated in a 1 mm cuvettes (200 Ω , 1800 V and 25 μ). Bacteria were plated on LB agar plates with a minimal amount of antibiotic For hygromycin low salt LB medium (10 g Bacto-tryptone, 5 g Bacto-yeast extract, 5 g NaCl, in 1 L) was used.

4.1.5 Mini-prep of BAC DNA

Colonies from recombineering were picked from agar plates and were grown overnight in 4 ml LB medium with appropriate antibiotic. Two ml of culture were centrifuged at maximum speed for 1 min, supernatant was discarded, and the pellet resuspended in 200 μ l P1 buffer. For cell lysis 220 μ l P2 was added and after 5 minutes the lysis process was stopped by adding 220 μ l P3 (all buffers from Qiagen). Tubes were left on ice for 5 minutes and after centrifugation at highest speed for 10 minutes, the clear supernatant was transferred into a new tube. Samples were precipitated with 550 μ l isopropanol and centrifuged for 10 min at highest speed. To wash the pellet, isopropanol was discarded, 550 μ l 70% ethanol added and the mixture centrifuged for 10 min at highest speed. The

4.2 Cell culture

ethanol was discarded carefully and the pellet was dried for approximately 10 min at room temperature. Purified BAC DNA was dissolved in 15 µl water and used for restriction digest or PCR.

4.2 Cell culture

4.2.1 Culturing embryonic stem cells

Embryonic stem cells (ES), (3×10^6 cells per 6 cm dish) were cultured in ES cell medium with 15 % fetal calf serum (FCS) on gelatinized dishes coated with murine feeder cells ($3-4 \times 10^4$ cells/cm²) at 37 °C in a humidified incubator 7,5 % CO₂. Medium was changed daily and the cells were split every two or three days. For splitting, cells were washed with PBS twice and incubated in Trypsin-EDTA for 5 minutes.

Cells were frozen in ES cell medium with 10 % FCS and 10 % DMSO and stored in liquid nitrogen.

4.2.2 Electroporation of ES cells

ES cells containing a modified gene trap ROSA locus with a splice acceptor followed by a loxP flanked PGK Neo 3xpA and lacZ pA were used for BAC integration (129S;b6 Gt(ROSA)26Sor^{tm1Sor}). For random integration of the modified BAC into these ES cells, 2×10^6 cells and 3 µg PISceI linearized BAC DNA were used. After washing trypsinized cells with medium, cells were resuspended in 0.8 ml PBS, mixed with 3 µl linearized BAC DNA and electroporated in 4 mm electroporation cuvettes with a GenePulser (240 V, 500 µF). Electroporated cells were plated in 3 different concentrations on 6 cm dishes with feeder cells. Two days after electroporation selection medium was added to the cells containing 150 µg/ml hygromycin. Cells were grown 4-7 days until individual colonies became visible.

4.2.3 Picking of ES cell clones

Individual clones were picked into 20 µl trypsin-EDTA solution, and incubated at 37 °C for 5 minutes. The trypsin reaction was stopped by adding 100 µl ES cell medium with appropriate antibiotic for selection (150 µg/ml hygromycin, 250 µg/ml G418). Cells were grown in 12-well plates on feeder cells until they were frozen into tubes.

4.2.4 Induction of ES cells with 4-OH tamoxifen

Cells were treated overnight with 5 µM 4-OH-tamoxifen (10 mM in EtOH), the cells were fixed and stained following X-Gal staining of ES cells protocol.

4.2.5 X-Gal staining of ES cells

Cells were washed twice in PBS, before they were fixed in 2 % PFA/ 0.2 % glutaraldehyde at 4 °C for 10 minutes. Following fixation, cells were washed twice in PBS and kept overnight in the dark, in filtered staining solution (5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM MgCl₂, 0.01 % SDS, 0.02 %, NP-40, 0.1 % X-Gal in PBS) at 30 °C. Staining was stopped with two PBS washes and was documented using Axio Observer (Zeiss).

4.2.6 Genotyping ES cells by Southern blotting

Picked clones were split onto gelatinized 96-well plates and grown to confluence. Cells were washed twice with PBS and were lysed at 60 °C overnight in 50 µl TENS buffer (10 mM Tris-HCl (pH 7.5), 10 mM EDTA (pH 8.0), 10 mM NaCl, 0.5 % Sarcosyl, with 1 mg/ml Proteinase K added prior to use). On the following day, 100 µl of ice-cold 100% ethanol/75 mM NaCl was added per well, and left for 30 min at room temperature to allow the DNA to precipitate. The plate was then centrifuged for 5 min at 1000 rpm and the solution discarded by inverting the plate. Wells were washed 3 times by adding 200 µl 70 % ethanol, each time the plate was centrifuged and the ethanol discarded by inverting the plate. After drying of the DNA for several minutes, it was dissolved in 25 µl 0.1 TE per well and used for restriction digest or PCR.

Restriction digests were done overnight at 37 °C in the appropriate buffer with BSA, using 50 U enzyme per well and adding 100 mM spermidine and 100 µg/ml RNase A.

Fragmented DNA was separated on a 0.75 % agarose gel for 4 h at 80 V (gel size: 14 x 23 cm). The gel was washed in denaturation buffer (1.5 mM NaCl, 0.5 N NaOH) for 30 min, then in neutralization buffer (1.5 M NaCl, 0.5 M Tris-HCl (pH 8)) for 30 min, and finally in 20x SSC for 10 min to help enable large fragments to transfer to the membrane. DNA was then transferred overnight to a nylon membrane (Hybond-N, Amersham) using Whatman Turboblot equipment according to the manufacturer's protocol.

On the following day, the DNA was crosslinked to the membrane using UV radiation (500 µJ/cm² radiation: UV Stratalinker 2400, Stratagene) and dried at room temperature. For radioactive labelling of the probe, 25 ng template fragment (hygromycin probe template for Cre lineage tracing cassette and neomycin probe template for Dre lineage tracing cassette) was diluted in 45 µl TE buffer and denatured at 95 °C for 1 min. Denatured template was then added to a random labelling mix (Rediprime, Random Prime Labelling System, GE Healthcare) with 5 µl α³²P dCTP (GE Healthcare Bio-Sciences) and incubated at 37 °C for 45 min to 1 h to allow the labelling reaction, catalysed by the Klenow fragment of the DNA polymerase I, to take place. To remove unincorporated radioactive nucleotides, G-25 MicroSpin Column (illustrate ProbeQuant, GE Healthcare) were used. Follow-

4.3 Mouse work

ing another denaturation step for 1 min at 95 °C the probe was added to the membrane, which was already in ExpressHyb Hybridisation Solution (Clonetech) and hybridized overnight with rotating at 63 °C. The next day, the membrane was washed 2 times for 10 min in 2xSSC/0.1 % SDS and 2x 10 min in 0.2% SSC/0.1 % SDS at 63 °C. After rinsing the membrane with water it was packed in cling film and exposed to a phosphoimager screen for 24 h. The Storm® gel and blot imaging system was used to detect signal from phosphoimager screens.

4.2.7 Tetraploid complementation assay using transgenic ES cells

ES cells with integrated BAC clones were used for tetraploid aggregation, a well-established method in our lab and performed in the transgenic unit supervised by Dr. Lars Wittler (MPI-MG) (Tanaka et al., 2009; Behringer et al., 2014). All aggregations were carried out by Karol Macura and colleagues.

4.3 Mouse work

All mouse experiments were done with strict adherence to German animal protection law (LAGeSo licence number G0043/12). Mice were sacrificed by cervical dislocation. After dissection of the uterus, preparation of the embryos was performed in chilled PBS for *in situ* hybridization or X-Gal staining, and further fixation was done according to whole-mount *in situ* or X-Gal staining protocols. For FACS experiments, uteruses were dissected into cold PBS, and embryos were then immediately stored in cold M2 media (Sigma-Aldrich).

4.3.1 Tamoxifen administration

Tamoxifen was dissolved in corn oil, (preheated to 42°C) at 37°C. The final concentration for injection was 10 mg/ml. Tamoxifen was placed at 4°C for up to one month. Two-hundered µl (2 mg) tamoxifen were used per intraperitoneal injection, carried out by the animal caretaker. Tamoxifen was administrated on each of two days before dissection between 8 and 10 am.

4.4 Histology

4.4.1 Imaging of fluorescent embryos

Embryos were dissected in cold PBS (137 mM NaCl, 2.7 mM KCl, 10 mM Na₂HPO₄, 1.8 mM KH₂PO₄) and were immediately transferred to cold M2 media in 12-well slides

(ibidi GmbH). Imaging was done using a multi photon laser-scanning microscope (LSM 710 NLO, Zeiss) by Matthias Marks (MPI-MG).

4.4.2 WISH (Whole-mount *in situ* hybridization)

Embryos were prepared in cold PBS and transferred to cold 4 % PFA (40 g PFA, 120 µl 1 M NaOH, 100 ml 10x PBS, filtered in 1 L) for overnight fixation at 4°C. Embryos were washed in cold PBT (PBS/0.1 % Tween-20) twice for 5 min and were then taken through a dehydration series: 25 % methanol/PBT, 50 % methanol/PBT, 75 % methanol/PBT and 100 % methanol for 5-15 minutes each. Embryos were either stored in methanol at -20°C or used right away. Rehydration was performed using 75 % methanol/PBT, 50 % methanol/PBT, and 25 % methanol/PBT for 5-15 min each, followed by 2 PBT washes. For bleaching and inactivation of endogenous phosphatase activity, embryos were incubated in 6 % H₂O₂/PBT for 30-40 min and washed in PBT 3 times for 5-10 min. To allow better penetration of the labelled probe, the embryo was digested with proteinase K/PBT (10 mg/ml proteinase K) for 8 min (9.5 dpc), or 10 min (10.5 dpc). The reaction was stopped using glycine/PBT (2 mg/ml) solution for 5 minutes. After 2 washes with PBT, embryos were post-fixed using 0.2 % glutaraldehyde/4 % PFA and washed 2 times in PBT for 5 min each.

Embryos were incubated in hybridization solution for 1 h, and then the denatured labelled probe was added (~1 µg/ml) after denaturation at 95 °C for 5 min. Hybridization was performed at 68 °C overnight.

To wash off the un-bound probe, embryos were incubated twice in Solution 1 for 30 min at 68 °C once in Solution 3T for 30 min at 68 °C and once in Solution 3T/MABT (1:1, v:v) for 20 min at room temperature to allow the specimen to cool down. Three washes with MABT at room temperature for 15 min each followed.

Blocking was done in MABT with 2 % Boehringer Blocking Solution (Roche) and with 20 % lamb serum for 1 hour at room temperature. Anti-digoxigenin-AP conjugated antibody (Roche) was added to the Blocking Solution in a dilution of 1:2000 and incubated overnight at 4°C. The next day, embryos were extensively washed with MABT at room temperature: 3x 5 min each, 3x 15 min, 2x 30 min, and then every hour. After an overnight step with MABT at 4 °C, embryos were washed 4 more times every hour. For staining, embryos were first incubated twice in NTMT for 10 min each and then once for 40 min. BM Purple (Roche) was added and embryos were incubated in the dark until staining became visible. Staining was stopped with 3 washes of PBT and then 4 % PFA was added. Embryos were photographed using a SteREO Discovery.V12 microscope (Zeiss), or MZ 16a (Leica) and an AxioCam Color camera (Zeiss), and processed using the AxioVision 4.6 software (Zeiss). Embryos were stored in 4 % PFA at 4 °C.

4.4 Histology

4.4.3 X-Gal staining of whole mount embryos

Embryos were dissected in ice-cold PBS and washed twice before they were fixed in 2 % PFA/ 0.2 % glutaraldehyde at 4 °C (9.0 dpc embryos were fixed 30-60 minutes and 10.0 dpc embryos for 60 minutes). Following fixation, specimens were washed twice in PBS and kept overnight in the dark, in filtered staining solution (5 mM potassium ferricyanide, 5 mM potassium ferrocyanide, 2 mM MgCl₂, 0.01 % SDS, 0.02 %, NP-40, 0.1 % X-Gal in PBS) at 30 °C. Staining was stopped with two PBS washes, and embryos were post-fixed in 4 % PFA overnight at 4 °C.

4.4.4 Paraffin embedding and sectioning of embryos

X-Gal stained embryos were washed in PBST twice to remove fixative and dehydrated in 3 steps using 20 % ethanol 10 min, 40 % ethanol 10 min and 70 % ethanol 10 min (storage at -20 °C possible here). Embryos were now processed using program 9 (Table 4.1) in an STP 120 Spin Tissue Processor (Thermo Scientific) and embedded in paraffin wax (Histowax; Leica), using an EC 350-1 embedding station (MICROM). A rotary microtome (HM 355 S; MICROM) was used to generate 4 µm sections which were transferred onto adhesion microscope slides (SuperFrost Ultra Plus; Menzel). After several hours of drying on 37 °C plate, sections were stored at room temperature until counterstained with eosin.

Solution	Time	Agitation
70 % ethanol	2 x 15 min	2
96 % ethanol	10 min	2
100 % ethanol	3 x 3 min	2
Xylene	2 x 5 min	2
Paraffin	120 min	1
Paraffin	overnight	1

Table 4.1: STP 120 Spin Tissue Processor program 9

4.4.5 Counterstaining of sections with Eosin

Sections were baked at 63 °C for 2 h before de-paraffinizing in UltraClear (J.T. Baker) for 2 x 3 min and once at 30 sec. Sections were rehydrated for one minute in 2 x 100 % and once at 95 %, 90 %, 80 %, 70 %, and 50 % ethanol, stained for ~5 sec in alcoholic Eosin Y Solution (Sigma-Aldrich), washed in ddH₂O for 30 sec and dehydrated in 80 % and 2 x 100 % ethanol for 1 min each. After 2 x 2 min in UltraClear and a short dip in UltraClear,

sections were mounted with Entellan (Merck) and covered with glass coverslips. Sections were documented using Axio Observer (Zeiss).

4.5 FACS - Fluorescence activated cell sorting

4.5.1 Homogenization of caudal ends for FACS

Embryos were dissected in ice-cold PBS and immediately transferred to ice-cold M2 media (Sigma-Aldrich). Caudal ends were dissected from embryos with 20-22 somites using a tungsten needle. Only the unsegmented caudal ends were taken and collected in the smallest possible volume of M2. Ten μ l of 10% BSA/PBS (sterile-filtered) and 100 μ l Trypsin/EDTA were added and carefully pipetted up and down, to homogenize cells. Samples were kept on ice for 2-5 min and were mixed from time-to-time until solutions appeared clear. The trypsin reaction was stopped by adding 200 μ l 10% BSA/PBS. Samples were kept on ice until they were analyzed by FACS.

4.5.2 Cell sorting

For cell sorting, a FACS Aria II (BD biosciences) was used. Cell sorting was operated by Dr. Frederic Koch (MPI-MG). Prior to loading, samples were filtered twice with a 35 μ m nylon mesh cell strainer (Corning). Sorting was performed, using an 85 μ m nozzle, and fluorescent cells were directly sorted into low-binding tubes, pre-filled with 500 μ l RLT buffer (Qiagen) plus 10 μ l/ml β -Mercaptoethanol (Sigma). Samples were vortexed and stored at -80 °C until processed for RNA.

4.6 RNA work

4.6.1 Isolation of RNA

In order to isolate RNA from cells sorted by FACS, samples were thawed and processed according to the RNeasy Micro protocol (Qiagen). All steps were performed according to the protocol with on-column DNase I digest. Briefly 50 μ l DNase I stock solution (Qiagen) and 1 μ l DNase (Roche) were added to 350 μ l Buffer RDD and mixed by inverting the tube. The DNaseI incubation mix was added directly to the spin column and incubated for 15 minutes. Elution was performed in 13 μ l RNase-free water. All samples were stored at -80°C until further use.

4.6.2 RNA library preparation

The RNA library was prepared according to the TotalScript RNA-Seq Kit (epicentre), with a few modifications. Before starting, RNA quality was analyzed using a Bioanalyzer (Agilent). Quality of RNA samples were RIN>7. Five ng of total RNA in 11.3 µl RNase-free water was used here.

11.3 µl Total RNA

1.95 µl TotalScript Optimized Buffer

0.78 µl Random Hexamer Primer

Samples were heated for 2 minutes to 65 °C and then kept at 4 °C. Samples were placed on ice.

First strand synthesis

For first strand synthesis Actinomycin D was added to the reaction, to improve strandedness.

1.95 µl DTT

0.39 µl dNTPs

0.78 µl RiboGuard RNase Inhibitor

0.78 µl Actinomycin D (250 ng/µl)

0.78 µl EpiScript Reverse Transcriptase

Samples were incubated using this program:

5 minutes at 25°C

25 minutes at 42°C

15 minutes at 70°C

hold 4°C

Second strand synthesis

For second strand synthesis,

0.78 µl

19.5 µl 2 nd Strand Master Mix

were added and incubated for 1 hour at 16 °C. Enzymes were heat inactivated at 80 °C for 15 minutes and the samples set to 4 °C.

Tagmentation approach

A Tn5 transposase variant was used to add index primer sequences and to control the fragmentation of cDNA. By this step, directional sequencing of the library is feasible.

10 µl TotalScript Tagment Buffer

1 µl TotalScript Enzyme

Samples were incubated using this program:

5 minutes at 55 °C

Hold at 4 °C

5 µl TotalScript Stop Solution was added and samples were incubated at RT for 5 minutes.

Purification of samples

Samples were purified using AMPure XP magnetic beads. Sixty-five μ l AMPure beads were added to each reaction, samples were vortexed and left at RT for 5 minutes. Tubes were placed on a magnetic stand and left there for 5 minutes before unbound material was removed. Without moving the tubes, beads were washed twice with 250 μ l 80 % ethanol. Beads were allowed to dry for 5 minutes, and the tubes were removed from the magnetic stand and resuspended in 15 μ l elution buffer (10 mM Tris-HCl; pH 8). All samples were placed on the mangetic stand and beads were captured for 5 minutes. Tagmented cDNA from the elution phase was transferred to a fresh tube at RT.

Oligo replacement

In order to fill gaps and add index primers, reagents and primers were added after a quick centrifugation step.

14 μ l purified tagmented cDNA

4 μ l Gap-Fill Buffer

1 μ l Index according to Table 4.2

Samples were incubated for 1 minute at 45 °C and afterwards for 30 minutes at 37 °C.

1 μ l Gap-Fill Enzyme

was added to the reaction and all samples incubated at 37 °C for 30 minutes.

Index #	Sorted cells
Index 1	Tlx2
Index 1	Mixl1
Index 2	Sct
Index 2	Pax3
Index 3	Foxa2/-T

Table 4.2: Indeces used in library preparation for RNA-Seq

Purification of samples

Samples were purified using AMPure XP magnetic beads. 25 μ l AMPure beads were added to each reaction, samples were vortexed and left at RT for 5 minutes. Tubes were placed on magnetic stand and left there for 5 minutes, before unbound material was removed. Without moving the tubes, beads were washed twice with 250 μ l 80 % ethanol. Beads were allowed to dry for 5 minutes, and the tubes were removed from the magnetic stand and resuspended in 25 μ l elution buffer (10 mM Tris-HCl; pH 8). All samples were placed on the mangetic stand and beads were captured for 5 minutes. Gap filled cDNA from the elution phase was transferred to a fresh tube and put on ice.

PCR amplification

Samples were centrifuged quickly and set on ice.

24 µl Gap-filled cDNA

1 µl TotalScript PCR Primer cocktail

25 µl 2x Phusion High-Fidelity PCR master mix

PCR was performed, using cycle conditions listed in Table 4.3

Temperature	Time
95°C	2 minutes
Number of cycles: 14	
94°C	10 seconds
60°C	30 seconds
72°C	1 minute
72°C	5 minutes

Table 4.3: PCR program used for amplification of cDNA

Purification of amplified samples

Samples were purified using AMPure XP magnetic beads. Forty-five µl AMPure beads were added to each reaction, samples were vortexed and left at RT for 5 minutes. Tubes were placed on a magnetic stand and left there for 5 minutes, before unbound material was removed. Without moving the tubes, beads were washed twice with 250 µl 80 % ethanol. Beads were allowed to dry for 5 minutes, and the tubes were removed from the magnetic stand and resuspended in 20 µl water. All samples were placed on the magnetic stand and beads were captured for 5 minutes. Nine µl amplified cDNA from the elution phase was transferred to a fresh tube and put on ice.

Analysis and quantification of the library

The finished TotalScript RNA-Seq library was now analyzed with respect to fragment size and quality using the Bioanalyzer (Agilent).

4.6.3 RNA sequencing

RNA sequencing was performed by the Sequencing Core Facility team of Dr. Bernd Timmermann (Service Department, Max-Planck Institute for Molecular Genetics, Berlin) using the ultra-high-throughput sequencing system HiSeq 2500 (Illumina).

4.6.4 Bioinformatic data analysis

All sequencing reads from RNA-Seq were mapped to the mouse genome (mm10) using TopHat v2.0.8b. Cuffdiff was used to retrieve the normalized FPKM values. Genes differentially expressed among the 5 samples were selected based on $\text{abs}(\log_2 \text{FC}) \geq 1.5$ for at least one pair comparison. Bioinformatic data analysis was done by Jinhua Liu. Data was utilized for clustering using Cluster 3.0 (Stanford University 1998–99). Data was normalized and k-means clustering algorithm was performed by setting cluster number to 10. Visualization was done by Java Treeview (Saldanha, 2004).

4.7 BAC clones used for creating transgenic ES cells and embryos

4.7 BAC clones used for creating transgenic ES cells and embryos

Gene symbol	BAC clone name
Tlx2	RP23-453P10
Mixl1	RP23-394F3
Sct	RP23-134L4
Pax3	RP23-168B8

Table 4.4: Applied BACs

4.8 Buffers, solutions and media

ES cell medium	composition for 600 ml
Knockout Dulbecco's Modified Eagles Medium (DMEM) 4,500 mg/ml glucose, with sodium pyruvate (Gibco)	500 ml
ES cell tested fetal calf serum (FCS)	90 ml
100x glutamine (Lonza)	6 ml
100x penicillin (500U/ml)/streptomycin (500µg/ml) (Lonza)	6 ml
100x non-essential amino acids (Gibco)	6 ml
100x nucleosides (Chemicon)	6 ml
β-mercaptoethanol (Sigma)	4,3 µl
LIF (Murine Leukemia Inhibitory Factor ESGRO (10^7) (Chemicon)	60 µl
LB-media	
Bacto-trypton	10 g
Bacto-yeast extract	5 g
NaCl	10 g
fill up to 1L	

Southern blot

Denaturation buffer

NaCl 1.5M	300 ml
NaOH 0.5M	100 ml
fill up to 1L	

Neutralization buffer

NaCl 1.5M	300 ml
Tris 0.5M pH8	500 ml
fill up to 1L	

WISH

Prehybridisation mix

Formamid	250 ml
20x DEPC SSC pH 5	125 ml
SDS 20%	25 ml
yeast RNA	500 µl
Heparin	500 µl
fill up to 500 ml with DEPC water, store at -20°C	

Solution 1

Formamide	250 ml
20x SSC pH 4.5	125 ml
SDS 10%	50 ml
water	75 ml
store at -20°C	

Solution 3T

Formamide	250 ml
20x SSC pH 4.5	50 ml
water	200 ml
Tween-20	500 µl
store at -20°C	

4.8 Buffers, solutions and media

MAB(T)

maleic acid 23.22 g

NaCl 17.5 g

Ajust pH to 7.5 using 10 M NaOH

Fill up to 2 L, filter and store at 4°C

(Tween-20 0.1%)

NT(M)T

Tris 1M pH 9.5 50 ml

(MgCl₂ 1 M 25 ml)

NaCl 5 M 10 ml

Tween-20 10% 5 ml

water (410 ml) 435 ml

Chapter 5

Summary

The caudal end of the midgestational mouse embryo comprises several important subdomains which are essential for morphogenesis and tissue maintenance during trunk and tail elongation. This study investigates different cellular subdomains from the caudal ends of 9.0 and 10.5 days post coitum (dpc) murine embryos. Each of these subgroups is characterized by the specific expression of a different marker gene. Four different marker genes were chosen for this study: *Tlx2*, *Mixl1*, *Sct* and *Pax3*.

Lineage tracing was performed to reveal the progeny of each subgroup. In order to compare the localization of subgroup descendants at these two distinct time-points during embryonic development, an inducible construct was employed. Two different reporter genes were genetically integrated into ES cells for the generation of reporter embryos. For lineage tracing analysis, a lacZ gene was integrated in the ROSA26R locus, able to be triggered by tamoxifen induction. For RNA-Seq analysis, a Venus fluorescent protein under control of the marker gene promotor was used for cell sorting.

The lineage tracing assay revealed *Mixl1* expressing cells in lateral and paraxial mesoderm, and in the posterior notochord. *Tlx2* expressing progeny were found in the hindgut, and progeny of *Sct* expressing cells were spotted throughout the mesoderm. Descendants of *Pax3* expressing cells were located in the neural tube and later in ganglion structures.

The transcriptomes of each cellular subgroup were examined using RNA-Seq. Data from the lineage tracing approach was integrated in order to find upregulated genes for directed processes within each subgroup of sorted cells.

Transcriptome analysis of the *Mixl1* subgroup indicated upregulation of *Noto*, which is involved in notochord development. Additionally, genes involved in processes characteristic for mesoderm were found. Thus *Mixl1* expressing cells could contain two different subgroups or could be precursors for different domains of mesoderm.

Since progeny of *Tlx2* expressing cells were found in the hindgut, genes indicating participation in endoderm development were examined and genes coding for integral membrane proteins were identified. My observation that progeny from *Tlx2* expressing cells contribute to hindgut development has not been described before. Previous research detailed hindgut formation as a process expanding in a rostral-to-caudal direction. Here I depict the novel finding that a second origin for the hindgut exists, derived from cells from the caudal end. For the subdomain of *Pax3* expressing cells, neural crest cell genes were found to be

upregulated. This novel finding shows that neural crest cells arise from the mesodermal *Pax3* expressing region in the caudal end to contribute to NC development.

Transcriptome analysis of cells sorted for *Sct* expression identifies this subdomain as a component of the mesoderm, without any additional functional distinctions.

Overall, the work contained in this thesis uses an innovative genetic approach to provide functional information about various cellular subpopulations within the developing mouse caudal end. Presented results confirm not only common correlations but give novel insights into so far unknown processes of embryonic mouse development.

Chapter 6

Zusammenfassung

Das kaudale Ende des 9.0 Tage alten Mausembryos besteht aus unterschiedlichen Teilbereichen, welche wichtige Funktionen während der Rumpf- und Schwanzentwicklung inne haben. In dieser Arbeit wurden mehrere Untergruppen des kaudalen Endes von 9 und 10 Tage alten Mausembryonen untersucht. Jede dieser Untergruppen ist durch die spezifische Expression eines Gens charakterisiert. Für diese Arbeit wurden 4 unterschiedliche Markergene, *Mixl1*, *Tlx2*, *Sct* und *Pax3* augewählt.

Eine Zellstammbaumanalyse, bei welcher eine per Tamoxifen induzierbare Kassette verwendet wurde, ermöglichte Untersuchungen zu ausgewählten Zeitpunkten während der Embryonalentwicklung. Zu diesem Zweck wurden zwei unterschiedliche Reportergene verwendet. Das Venusreporterogen, welches vom Promotor des Markergens kontrolliert wurde, sowie ein lacZ-Gen. Das lacZ-Gen, war in den ROSA26R Genlokus integriert und wurde somit nur nach erfolgter Applikation von Tamoxifen durch Rekombinase vermitteltes Ausschneiden einer Stoppkassette exprimiert. Venus exprimierende Zellen wurden zusätzlich mittels fluoreszenzbasierter Durchflusszytometrie sortiert und anschließend für RNA-Sequenzierung verwendet.

Die Zellstammbaumanalyse zeigte die Lokalisation der Tochterzellen der unterschiedlichen durch die Markergene charakterisierten Subdomänen. Im lateralen und paraxialen Mesoderm, sowie im Notochord fanden sich Abkömmlinge der *Mixl1* exprimierenden Zellen. Tochterzellen von *Tlx2* exprimierenden Zellen wurden im Hinterdarm nachgewiesen. Nachkommen von *Sct* exprimierenden Zellen wurden verstreut im Mesoderm identifiziert. Nachfahren von *Pax3* exprimierenden Zellen wurden im Neuralrohr, sowie zu einem späteren Zeitpunkt der Embryonalentwicklung in Nervenknoten gefunden.

Die Transkriptomanalyse von sortierten Zellen mit *Mixl1* Expression zeigte das Gen *Noto* in diesen Zellen erhöht exprimiert. Außerdem waren verstärkt exprimierte Gene enthalten, welche in Entwicklungsprozesse involviert sind, die charakteristisch für das Mesoderm sind. Folglich könnten die *Mixl1* exprimierenden Zellen aus zwei unterschiedlichen Untergruppen zusammengesetzt sein, oder Vorläufer unterschiedlicher Mesodermdomänen enthalten.

Nachkommen *Tlx2* exprimierender Zellen wurden im Hinterdarm gefunden. In der Transkriptomanalyse konnten Gene die zu Entwicklung des Endeoderms beitragen und insbesondere für transmembrane Proteine kodierende Gene identifiziert werden. Das Zellen,

insbesondere *Tlx2* exprimierende Zellen aus dem kaudalen Ende des Mausembryos zur Entwicklung des Hinterdarmes beitragen, ist bisher noch nicht beschrieben worden. Nach bisheriger Auffassung wächst der Hinterdarm vom anterioren in Richtung des Kaudalen Endes. Die Ergebnisse dieser Arbeit legen nahe, dass es zwei Richtungen gibt aus denen Zellen zur Hinterdarmentwicklung beitragen.

Bei Zellen aus der Subregion des kaudalen Endes mit *Pax3* Expression konnte gezeigt werden, dass diese Zellen Gene exprimieren, die in die Entwicklung von Neuralleistenzellen involviert sind. Diese neuen Ergebnisse zeigen, dass sich auch aus der mesodermalen Expressionsdomäne von *Pax3* Neuralleistenzellen heranbilden.

Die Analyse des Transcriptomes von *Sct* exprimierenden Zellen wies diese Domäne als Teil des Mesoderm, aber ohne eine für diese Domäne besondere Funktion aus.

In der vorliegenden Forschungsarbeit wurde ein innovativer Ansatz gewählt, um neue Erkenntnisse über funktionelle Zusammenhänge in zellulären Subpopulationen des kaudalen Endes der Maus zu gewinnen. Die hier vorgelegten Ergebnisse untermauern einerseits bestehende Erkenntnisse, zeigen aber vor allem bisher unbekannte, neue Zusammenhänge der Embryonalentwicklung der Maus auf.

List of Figures

1.1	Gastrulation and formation of the egg cylinder (Gilbert, 2014)	2
1.2	Scheme of mesoderm derivatives after mesoderm patterning. Figure modified from (Gilbert, 2014)	4
1.3	Paraxial mesoderm, somites and neural tube after removal of ectoderm. Figure modified from (Gilbert, 2014)	5
1.4	9.5 dpc embryo with gut tube marked in green. Foregut, midgut and hindgut sections are indicated by lines. Figure modified from (Zorn and Wells, 2009)	6
1.5	Development of Neural Crest cells	8
1.6	Ventral ectodermal ridge (VER)	10
1.7	Expression domains of marker genes-of-interest. Adapted from (Robb and Tam, 2004)	11
2.1	Recombineering strategy	18
2.2	Lineage tracing construct	19
2.3	<i>Mixl1</i> and Venus expression in 9.0 dpc embryo	22
2.4	textit{Mixl1} lineage tracing of 9.0 dpc embryos	24
2.5	<i>Tlx2</i> expression during embryogenesis	25
2.6	Lineage tracing of <i>Tlx2</i> in 9.0 dpc embryos	27
2.7	Lineage tracing of <i>Tlx2</i> in 10.5 dpc embryos	28
2.8	Expression of <i>Sct</i> in 9.5 dpc embryos	29
2.9	Lineage tracing of <i>Sct</i> in 9.0 and 10.0 dpc embryos	30
2.10	Expression of <i>Pax3</i> in 9.5 dpc embryos	31
2.11	Lineage tracing of <i>Pax3</i> in 9.0 dpc embryos	33
2.12	Lineage tracing of <i>Pax3</i> in 10.5 dpc embryos	34
2.13	Dissection of 9.0 dpc embryo for FACS	35
2.14	Clustered heatmap of differentially expressed genes of subpopulations of cells from the caudal end by RNA-Seq.	41
2.15	Cluster II and cluster VI with depicted genes involved in paraxial and lateral mesoderm development, as well as notochord development	46
2.16	Expression patterns of marker genes at midgestation for paraxial and lateral mesoderm formation which are upregulated in the <i>Mixl1</i> expressing subpopulation	47

List of Figures

2.17 Whole mount <i>in situ</i> hybridization for <i>Noto</i>	49
2.18 Expression of <i>Foxa2</i> , <i>T</i> and <i>Sox2</i> in lateral view of 9.0 dpc caudal ends	50
2.19 Cluster with tagged genes and FPKM values for endoderm development in the <i>Tlx2</i> expressing cell group.	51
2.20 Expression of genes involved in endoderm development, which are upregulated in the <i>Tlx2</i> expressing cell group	52
2.21 Clusters I, III, VII and IX and FPKM values with depicted genes involved in neural crest development.	54
2.22 Expression pattern of genes involved in neural crest development, which are also upregulated in the <i>Pax3</i> expressing subregion of 9.0 dpc embryos	55
7.1 In <i>vitro</i> induction of the lineage tracing construct using Dre recombinase	111

List of Tables

2.1	Canonical pathways of differentially expressed genes in the <i>Mixl1</i> expressing caudal cell population	36
2.2	Canonical pathways of differentially expressed genes in the <i>Tlx2</i> expressing caudal end population as found by RNA-Seq analysis	37
2.3	Canonical pathways of differentially expressed genes in the <i>Sct</i> expressing caudal end population as found by RNA-Seq analysis	38
2.4	Top canonical pathways upregulated in the <i>Pax3</i> expressing caudal end population as determined by GO analysis from RNA-Seq data of <i>Pax3</i> expressing cells at 9.0 dpc.	39
2.5	Biological processes for Clusters I to X from RNA-Seq analysis	43
2.6	Function of genes involved in endoderm development	53
2.7	Genes involved in neural crest development also upregulated in the <i>Pax3</i> expressing subpopulation.	56
4.1	STP 120 Spin Tissue Processor program 9	77
4.2	Indeces used in library preparation for RNA-Seq	80
4.3	PCR program used for amplification of cDNA	81
4.4	Applied BACs	83
7.1	Biological processes of genes listed in Cluster I	111
7.2	Biological processes of genes listed in Cluster II	112
7.3	Biological processes of genes listed in Cluster III	119
7.4	Biological processes of genes listed in Cluster IV	121
7.5	Biological processes of genes listed in Cluster V	122
7.6	Biological processes of genes listed in Cluster VI	123
7.7	Biological processes of genes listed in Cluster VII	124
7.8	Biological processes of genes listed in Cluster VIII	125
7.9	Biological processes of genes listed in Cluster IX	126
7.10	Biological processes of genes listed in Cluster X	127

References

- Allan, D., Houle, M., Bouchard, N., Meyer, B. I., Gruss, P., and Lohnes, D. RARgamma and Cdx1 interactions in vertebral patterning. *Developmental biology*, **240**(1):46–60, 2001.
- Anastassiadis, K., Fu, J., Patsch, C., Hu, S., Weidlich, S., Duerschke, K., Buchholz, F., Edenhofer, F., and Stewart, a. F. Dre recombinase, like Cre, is a highly efficient site-specific recombinase in *E. coli*, mammalian cells and mice. *Disease models & mechanisms*, **2**(9-10):508–15, 2009.
- Arkell, R. M., Fossat, N., and Tam, P. P. Wnt signalling in mouse gastrulation and anterior development: new players in the pathway and signal output. *Current opinion in genetics & development*, pages 1–7, 2013.
- Arnold, S. J. and Robertson, E. J. Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nature reviews Molecular cell biology*, **10**(2):91–103, 2009.
- Aulehla, A. and Herrmann, B. G. Segmentation in vertebrates: clock and gradient finally joined. *Genes Dev*, **18**(17):2060–2067, 2004.
- Aulehla, A. and Pourquié, O. Signaling gradients during paraxial mesoderm development. *Cold Spring Harbor perspectives in biology*, **2**(2):a000869, 2010.
- Aybar, M. J. Snail precedes Slug in the genetic cascade required for the specification and migration of the *Xenopus* neural crest. *Development*, **130**(3):483–494, 2003.
- Baker, C. V. and Bronner-Fraser, M. The origins of the neural crest. Part I: embryonic induction. *Mechanisms of Development*, **69**(1-2):3–11, 1997.
- Bang, A. G., Papalopulu, N., Goulding, M. D., and Kintner, C. Expression of Pax-3 in the lateral neural plate is dependent on a Wnt-mediated signal from posterior nonaxial mesoderm. *Developmental biology*, **212**(2):366–80, 1999.
- Beddington, R. and Robertson, E. Axis development and early asymmetry in mammals. *Cell*, **96**:195–209, 1999.
- Behringer, R., Gertsenstein, M., Nagy, K. V., and Nagy, A. *Manipulating the Mouse Embryo*. Cold Spring Harbor Laboratory Press, 2014. ISBN 978-1-936113-01-9.

- Bellairs, R. The primitive streak. *Anatomy and embryology*, **174**(1986):1–14, 1986.
- Ben Abdelkalek, H., Beckers, A., Schuster-Gossler, K., Pavlova, M. N., Burkhardt, H., Lickert, H., Rossant, J., Reinhardt, R., Schalkwyk, L. C., Müller, I., Herrmann, B. G., Ceolin, M., Rivera-Pomar, R., and Gossler, A. The mouse homeobox gene Not is required for caudal notochord development and affected by the truncate mutation. *Genes and Development*, **18**:1725–1736, 2004.
- Bertocchini, F., Skromne, I., Wolpert, L., and Stern, C. D. Determination of embryonic polarity in a regulative system: evidence for endogenous inhibitors acting sequentially during primitive streak formation in the chick embryo. *Development (Cambridge, England)*, **131**(14):3381–90, 2004.
- Borghini, S., Bachetti, T., Fava, M., Di Duca, M., Cargnini, F., Fornasari, D., Ravazzolo, R., and Ceccherini, I. The TLX2 homeobox gene is a transcriptional target of PHOX2B in neural-crest-derived cells. *The Biochemical journal*, **395**(2):355–361, 2006.
- Borghini, S., Di Duca, M., Santamaria, G., Vargiu, M., Bachetti, T., Cargnini, F., Pini Prato, A., De Giorgio, R., Lerone, M., Stanghellini, V., Jasonni, V., Fornasari, D., Ravazzolo, R., and Ceccherini, I. Transcriptional regulation of TLX2 and impaired intestinal innervation: possible role of the PHOX2A and PHOX2B genes. *European journal of human genetics : EJHG*, **15**(8):848–855, 2007.
- Burdsal, C. a., Damsky, C. H., and Pedersen, R. a. The role of E-cadherin and integrins in mesoderm differentiation and migration at the mammalian primitive streak. *Development (Cambridge, England)*, **118**(1993):829–844, 1993.
- Burtscher, I. and Lickert, H. Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development (Cambridge, England)*, **136**(6):1029–38, 2009.
- Burtscher, I., Barkey, W., Schwarzfischer, M., Theis, F. J., and Lickert, H. The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development. *Genesis (New York, N.Y. : 2000)*, **50**(6):496–505, 2012.
- Cambray, N. and Wilson, V. Axial progenitors with extensive potency are localised to the mouse chordoneural hinge. *Development*, **129**:4855–4866, 2002.
- Cano, A., Pérez-Moreno, M. A., Rodrigo, I., Locascio, A., Blanco, M. J., del Barrio, M. G., Portillo, F., and Nieto, M. A. The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nature cell biology*, **2**(2):76–83, 2000.

References

- Carver, E. A., Jiang, R., Lan, Y., Oram, K. F., and Gridley, T. The mouse snail gene encodes a key regulator of the epithelial-mesenchymal transition. *Molecular and cellular biology*, **21**(23):8184–8, 2001.
- Chapman, D. L. and Papaioannou, V. E. Three neural tubes in mouse embryos with mutations in the T-box gene Tbx6. *Nature*, **391**(6668):695–7, 1998.
- Chapman, D. L., Agulnik, I., Hancock, S., Silver, L. M., and Papaioannou, V. E. Tbx6, a mouse T-Box gene implicated in paraxial mesoderm formation at gastrulation. *Developmental biology*, **180**(2):534–42, 1996.
- Chawengsaksophak, K., de Graaff, W., Rossant, J., Deschamps, J., and Beck, F. Cdx2 is essential for axial elongation in mouse development. *Proceedings of the National Academy of Sciences of the United States of America*, **101**(20):7641–5, 2004.
- Cheeseman, B. L., Zhang, D., Binder, B. J., Newgreen, D. F., and Landman, K. A. Cell lineage tracing in the developing enteric nervous system: superstars revealed by experiment and simulation. *Journal of the Royal Society, Interface / the Royal Society*, **11**(93):20130815, 2014.
- Christ, B., Huang, R., and Scaal, M. Amniote somite derivatives. *Developmental dynamics : an official publication of the American Association of Anatomists*, **236**(9):2382–96, 2007.
- Conklin, E. G. The organization and cell-lineage of the ascidian egg. *Philadelphia : [Academy of Natural Sciences]*, 1905.
- Conlon, F., Lyons, K., Takaesu, N., Barth, K., Kispert, A., Herrmann, B., and Robertson, E. A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development*, **120**(7):1919–1928, 1994.
- Davidson, B. P. and Tam, P. P. L. The node of the mouse embryo. *Current Biology*, **10**(17):617–619, 2000.
- Davis, R. P., Ng, E. S., Costa, M., Mossman, A. K., Sourris, K., Elefanty, A. G., and Stanley, E. G. Targeting a GFP reporter gene to the MIXL1 locus of human embryonic stem cells identifies human primitive streak-like cells and enables isolation of primitive hematopoietic precursors. *Blood*, **111**(4):1876–1884, 2008.
- Dequéant, M. L. and Pourquié, O. Segmental patterning of the vertebrate embryonic axis. *Nature reviews Genetics*, **9**(5):370–382, 2008.
- Ding, G., Tanaka, Y., Hayashi, M., Nishikawa, S.-I., and Kataoka, H. PDGF receptor alpha+ mesoderm contributes to endothelial and hematopoietic cells in mice.

Developmental dynamics : an official publication of the American Association of Anatomists, **242**(3):254–68, 2013.

Dongkyun, K., Jinsoo, S., and Jin, E.-J. Wnt-3 and Wnt-3a play different region-specific roles in neural crest development in avians. *Cell biology international*, **34**(7):763–8, 2010.

Douarin, N. L. and Kalcheim, C. *The Neural Crest*, volume 28. Cambridge University Press, 1999. ISBN 0521620104.

Dubrulle, J. and Pourquié, O. fgf8 mRNA decay establishes a gradient that couples axial elongation to patterning in the vertebrate embryo. *Nature*, **427**(6973):419–422, 2004.

Dumont, D. J., Fong, G. H., Puri, M. C., Gradwohl, G., Alitalo, K., and Breitman, M. L. Vascularization of the mouse embryo: a study of flk-1, tek, tie, and vascular endothelial growth factor expression during development. *Developmental dynamics : an official publication of the American Association of Anatomists*, **203**(1):80–92, 1995.

Dupin, E. and Le Douarin, N. M. The neural crest, A multifaceted structure of the vertebrates. *Birth defects research. Part C, Embryo today : reviews*, 2014.

Epstein, D. J., Vekemans, M., and Gros, P. splotch (Sp2H), a mutation affecting development of the mouse neural tube, shows a deletion within the paired homeodomain of Pax-3. *Cell*, **67**(4):767–774, 1991.

Franklin, V., Khoo, P. L., Bildsoe, H., Wong, N., Lewis, S., and Tam, P. P. Regionalisation of the endoderm progenitors and morphogenesis of the gut portals of the mouse embryo. *Mechanisms of development*, **125**(7):587–600, 2008.

Fujita, K., Ogawa, R., Kawawaki, S., and Ito, K. Roles of chromatin remodelers in maintenance mechanisms of multipotency of mouse trunk neural crest cells in the formation of neural crest-derived stem cells. *Mechanisms of development*, 2014.

Furushima, K., Murata, T., Matsuo, I., and Aizawa, S. A new murine zinc finger gene, Opr. *Mechanisms of Development*, **98**(1-2):161–164, 2000.

Gammill, L. S. Genomic analysis of neural crest induction. *Development*, **129**(24):5731–5741, 2002.

Gammill, L. S. and Bronner-Fraser, M. Neural crest specification: migrating into genomics. *Nature reviews. Neuroscience*, **4**(10):795–805, 2003.

Gao, N., White, P., and Kaestner, K. H. Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Developmental cell*, **16**(4):588–99, 2009.

References

- Gilbert, S. F. *Developmental Biology*. Sinauer Associates, Inc., 2014. ISBN 978-0-087893-978-7.
- Goldman, D., Martin, G., and Tam, P. Fate and function of the ventral ectodermal ridge during mouse tail development. *Development*, **127**:2113–2123, 2000.
- González, B., Denzel, S., Mack, B., Conrad, M., and Gires, O. EpCAM is involved in maintenance of the murine embryonic stem cell phenotype. *Stem cells (Dayton, Ohio)*, **27**(8):1782–91, 2009.
- Goulding, M. D., Chalepakis, G., Deutsch, U., Erselius, J. R., and Gruss, P. Pax-3, a novel murine DNA binding protein expressed during early neurogenesis. *The EMBO journal*, **10**(5):1135–47, 1991.
- Grüneberg, H. A Ventral Ectodermal Ridge of the Tail in Mouse Embryos. *Nature*, **177** (4513):787–788, 1956.
- Hales, B. F., Grenier, L., Lalancette, C., and Robaire, B. Epigenetic programming: from gametes to blastocyst. *Birth defects research. Part A, Clinical and molecular teratology*, **91**(8):652–65, 2011.
- Hall, B. K. The neural crest and neural crest cells : discovery and significance for theories of embryonic organization. *J. Biosci.*, **33**(December):781–793, 2008.
- Hart, A. H., Hartley, L., Sourris, K., Stadler, E. S., Li, R. L., Stanley, E. G., Tam, P. P. L., Elefanty, A. G., and Robb, L. Mixl1 is required for axial mesendoderm morphogenesis and patterning in the murine embryo. *Development (Cambridge, England)*, **129**(15): 3597–3608, 2002.
- Hart, A. H., Willson, T. A., Wong, M., Parker, K., and Robb, L. Transcriptional regulation of the homeobox gene Mixl1 by TGF-beta and FoxH1. *Biochemical and biophysical research communications*, **333**(4):1361–9, 2005.
- Hatano, M., Aoki, T., Dezawa, M., Yusa, S., Itsuka, Y., Koseki, H., Taniguchi, M., and Tokuhisa, T. A novel pathogenesis of megacolon in Ncx/Hox11L.1 deficient mice. *The Journal of clinical investigation*, **100**(4):795–801, 1997.
- Heldin, C. H., Miyazono, K., and ten Dijke, P. TGF-beta signalling from cell membrane to nucleus through SMAD proteins. *Nature*, **390**(December):465–471, 1997.
- His, W. Untersuchungen über die erste Anlage des Wirbelthierleibes: Die erste Entwicklung des Hühnchens im Ei. *Verlag von F. C. W. Vogel, Leipzig*, 1868.

References

- Hong, C.-S. and Saint-Jeannet, J.-P. The activity of Pax3 and Zic1 regulates three distinct cell fates at the neural plate border. *Molecular biology of the cell*, **18**(6):2192–202, 2007.
- Hunt, P. and Krumlauf, R. Hox codes and positional specification in vertebrate embryonic axes. *Annual review of cell biology*, **8**:227–256, 1992.
- Inácio, J. M., Marques, S., Nakamura, T., Shinohara, K., Meno, C., Hamada, H., and Belo, J. A. The Dynamic Right-to-Left Translocation of Cerl2 Is Involved in the Regulation and Termination of Nodal Activity in the Mouse Node. *PloS one*, **8**(3):e60406, 2013.
- Izumi, N., Era, T., Akimaru, H., Yasunaga, M., and Nishikawa, S.-I. Dissecting the molecular hierarchy for mesendoderm differentiation through a combination of embryonic stem cell culture and RNA interference. *Stem cells (Dayton, Ohio)*, **25**(7):1664–74, 2007.
- Izzi, L., Silvestri, C., von Both, I., Labbé, E., Zakin, L., Wrana, J. L., and Attisano, L. Foxh1 recruits Gsc to negatively regulate Mixl1 expression during early mouse development. *The EMBO journal*, **26**(13):3132–3143, 2007.
- Jiang, R., Lan, Y., Norton, C. R., Sundberg, J. P., and Gridley, T. The slug gene is not essential for mesoderm or neural crest development in mice. *Developmental Biology*, **198**(2):277–285, 1998.
- Kataoka, H., Hayashi, M., Nakagawa, R., Tanaka, Y., Izumi, N., Nishikawa, S., Jakt, M. L., Tarui, H., and Nishikawa, S.-I. Etv2/ER71 induces vascular mesoderm from Flk1+PDGFR α + primitive mesoderm. *Blood*, **118**(26):6975–86, 2011.
- Kimura, W., Yasugi, S., Stern, C. D., and Fukuda, K. Fate and plasticity of the endoderm in the early chick embryo. *Developmental biology*, **289**(2):283–95, 2006.
- Kishigami, S. and Mishina, Y. BMP signaling and early embryonic patterning. *Cytokine & growth factor reviews*, **16**(3):265–78, 2005.
- Kispert, A. and Herrmann, B. G. Immunohistochemical analysis of the Brachyury protein in wild-type and mutant mouse embryos. *Developmental biology*, **161**(1):179–93, 1994.
- Kispert, A., Koschorz, B., and Herrmann, B. G. The T protein encoded by Brachyury is a tissue-specific transcription factor. *The EMBO journal*, **14**(19):4763–72, 1995.
- Kitajima, K., Oki, S., Ohkawa, Y., Sumi, T., and Meno, C. Wnt signaling regulates left-right axis formation in the node of mouse embryos. *Developmental biology*, **380**(2):222–32, 2013.

References

- Kretzschmar, K. and Watt, F. M. Lineage tracing. *Cell*, **148**(1-2):33–45, 2012.
- Kume, T., Jiang, H., Topczewska, J. M., and Hogan, B. L. The murine winged helix transcription factors, Foxc1 and Foxc2, are both required for cardiovascular development and somitogenesis. *Genes & development*, **15**(18):2470–82, 2001.
- LaBonne, C. and Bronner-Fraser, M. Snail-related transcriptional repressors are required in Xenopus for both the induction of the neural crest and its subsequent migration. *Developmental biology*, **221**(1):195–205, 2000.
- Lamouille, S., Xu, J., and Derynck, R. Molecular mechanisms of epithelial-mesenchymal transition. *Nature reviews. Molecular cell biology*, **15**(3):178–96, 2014.
- Landreth, K. S. Critical windows in development of the rodent immune system. *Human & experimental toxicology*, **21**(9-10):493–8, 2002.
- Le Lievre, C. S. and Le Douarin, N. M. Mesenchymal derivatives of the neural crest: analysis of chimaeric quail and chick embryos. *J Embryol Exp Morphol*, **34**(1):125–154, 1975.
- Lee, D., Park, C., Lee, H., Lugus, J. J., Kim, S. H., Arentson, E., Chung, Y. S., Gomez, G., Kyba, M., Lin, S., Janknecht, R., Lim, D.-S., and Choi, K. ER71 acts downstream of BMP, Notch, and Wnt signaling in blood and vessel progenitor specification. *Cell stem cell*, **2**(5):497–507, 2008.
- Lewis, S. L. and Tam, P. P. Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Developmental dynamics : an official publication of the American Association of Anatomists*, **235**(9):2315–2329, 2006.
- Li, J., Ning, Y., Hedley, W., Saunders, B., Chen, Y., Tindill, N., Hannay, T., and Subramaniam, S. The Molecule Pages database. *Nature*, **420**(6916):716–7, 2002.
- Li, Y., Esain, V., Teng, L., Xu, J., Kwan, W., Frost, I. M., Yzaguirre, A. D., Cai, X., Cortes, M., Maijenburg, M. W., Tober, J., Dzierzak, E., Orkin, S. H., Tan, K., North, T. E., and Speck, N. A. Inflammatory signaling regulates embryonic hematopoietic stem and progenitor cell production. *Genes & development*, pages gad.253302.114–, 2014.
- Lossi, L., Bottarelli, L., Candusso, M. E., Leiter, A. B., Rindi, G., and Merighi, A. Transient expression of secretin in serotoninergic neurons of mouse brain during development. *The European journal of neuroscience*, **20**(12):3259–69, 2004.
- Magli, A., Schnettler, E., Rinaldi, F., Bremer, P., and Perlingeiro, R. C. R. Functional dissection of Pax3 in paraxial mesoderm development and myogenesis. *Stem cells (Dayton, Ohio)*, **31**(1):59–70, 2013.

References

- Marques, S., Borges, A. C., Silva, A. C., Freitas, S., Cordenonsi, M., and Belo, J. A. The activity of the Nodal antagonist Cerl-2 in the mouse node is required for correct L/R body axis. *Genes & development*, **18**(19):2342–7, 2004.
- Maska, E. L., Cserjesi, P., Hua, L. L., Garstka, M. E., Brody, H. M., and Morikawa, Y. A Tlx2-Cre mouse line uncovers essential roles for hand1 in extraembryonic and lateral mesoderm. *Genesis (New York, N.Y. : 2000)*, **48**(8):479–84, 2010.
- McLarren, K. W., Litsiou, A., and Streit, A. DLX5 positions the neural crest and preplacode region at the border of the neural plate. *Developmental Biology*, **259**(1):34–47, 2003.
- Mead, P. E., Brivanlou, I. H., Kelley, C. M., and Zon, L. I. BMP-4-responsive regulation of dorsal-ventral patterning by the homeobox protein Mix.1. *Nature*, **382**(6589):357–60, 1996.
- Mead, P. E., Zhou, Y., Lustig, K. D., Huber, T. L., Kirschner, M. W., and Zon, L. I. Cloning of Mix-related homeodomain proteins using fast retrieval of gel shift activities, (FROGS), a technique for the isolation of DNA-binding proteins. *Proceedings of the National Academy of Sciences*, **95**(19):11251–11256, 1998.
- Meulemans, D. and Bronner-Fraser, M. Gene-regulatory interactions in neural crest evolution and development. *Developmental cell*, **7**(3):291–9, 2004.
- Meyer, J., Way, L., and Grossman, M. Pancreatic response to acidification of various lengths of proximal intestine in the dog. *Am J Physiol – Legacy Content*, **219**(4):971–977, 1970.
- Milewski, R. C., Chi, N. C., Li, J., Brown, C., Lu, M. M., and Epstein, J. A. Identification of minimal enhancer elements sufficient for Pax3 expression in neural crest and implication of Tead2 as a regulator of Pax3. *Development (Cambridge, England)*, **131**(4):829–37, 2004.
- Nagai, T., Aruga, J., Takada, S., Günther, T., Spörle, R., Schughart, K., and Mikoshiba, K. The expression of the mouse Zic1, Zic2, and Zic3 gene suggests an essential role for Zic genes in body pattern formation. *Developmental biology*, **182**(2):299–313, 1997.
- Nakata, K. A novel member of the Xenopus Zic family, Zic5, mediates neural crest development. *Mechanisms of Development*, **99**(1-2):83–91, 2000.
- Nakata, K., Nagai, T., Aruga, J., and Mikoshiba, K. Xenopus Zic family and its role in neural and neural crest development During submission of this paper, Mizuseki et al., reported the Xenopus Zic-related-1 gene which was highly homologous to mouse Zic1

References

- gene (Mizuseki et al., 1998). Accession No. Zic1, AB. *Mechanisms of Development*, **75** (1-2):43–51, 1998.
- Nern, A., Pfeiffer, B. D., Svoboda, K., and Rubin, G. M. Multiple new site-specific recombinases for use in manipulating animal genomes. *Proceedings of the National Academy of Sciences of the United States of America*, **108**(34):14198–203, 2011.
- Neubüser, A., Koseki, H., and Balling, R. Characterization and developmental expression of Pax9, a paired-box-containing gene related to Pax1. *Developmental biology*, **170**(2):701–16, 1995.
- Newgreen, D. and Young, H. M. Enteric Nervous System: Development and Developmental Disturbances—Part 1. *Pediatric and Developmental Pathology*, **5**(3):224–247, 2002.
- Ng, E. S., Azzola, L., Sourris, K., Robb, L., Stanley, E. G., and Elefanty, A. G. The primitive streak gene Mixl1 is required for efficient haematopoiesis and BMP4-induced ventral mesoderm patterning in differentiating ES cells. *Development (Cambridge, England)*, **132**(5):873–884, 2005.
- Niwa, H., Toyooka, Y., Shimosato, D., Strumpf, D., Takahashi, K., Yagi, R., and Rossant, J. Interaction between Oct3/4 and Cdx2 determines trophectoderm differentiation. *Cell*, **123**(5):917–29, 2005.
- Ohta, M., Funakoshi, S., Kawasaki, T., and Itoh, N. Tissue-specific expression of the rat secretin precursor gene. *Biochemical and Biophysical Research Communications*, **183** (2):390–395, 1992.
- Ohta, S., Suzuki, K., Tachibana, K., Tanaka, H., and Yamada, G. Cessation of gastrulation is mediated by suppression of epithelial-mesenchymal transition at the ventral ectodermal ridge. *Development (Cambridge, England)*, **134**(24):4315–4324, 2007.
- Pearce, J. J., Penny, G., and Rossant, J. A mouse cerberus/Dan-related gene family. *Developmental biology*, **209**(1):98–110, 1999.
- Pearce, J. J. and Evans, M. J. Mml, a mouse Mix-like gene expressed in the primitive streak. *Mechanisms of Development*, **87**(1-2):189–192, 1999.
- Pereira, L. A., Wong, M. S., Lim, S. M., Sides, A., Stanley, E. G., and Elefanty, A. G. Brachyury and related Tbx proteins interact with the Mixl1 homeodomain protein and negatively regulate Mixl1 transcriptional activity. *PloS one*, **6**(12):e28394, 2011.

- Pereira, L. A., Wong, M. S., Mei Lim, S., Stanley, E. G., and Elefanti, A. G. The Mix family of homeobox genes—key regulators of mesendoderm formation during vertebrate development. *Developmental biology*, **367**(2):163–77, 2012.
- Peters, H., Neubüser, A., Kratochwil, K., and Balling, R. Pax9-deficient mice lack pharyngeal pouch derivatives and teeth and exhibit craniofacial and limb abnormalities. *Genes & development*, **12**(17):2735–47, 1998.
- Plouhinec, J.-L., Roche, D. D., Pegoraro, C., Figueiredo, A. L., Maczkowiak, F., Brunet, L. J., Milet, C., Vert, J.-P., Pollet, N., Harland, R. M., and Monsoro-Burq, A. H. Pax3 and Zic1 trigger the early neural crest gene regulatory network by the direct activation of multiple key neural crest specifiers. *Developmental biology*, **386**(2):461–72, 2014.
- Richardson, L., Venkataraman, S., Stevenson, P., Yang, Y., Moss, J., Graham, L., Burton, N., Hill, B., Rao, J., Baldock, R. A., and Armit, C. EMAGE mouse embryo spatial gene expression database: 2014 update. *Nucleic acids research*, **42**(Database issue):D835–44, 2014.
- Robb, L., Hartley, L., Begley, C. G., Brodnicki, T. C., Copeland, N. G., Gilbert, D. J., Jenkins, N. A., and Elefanti, A. G. Cloning, expression analysis, and chromosomal localization of murine and human homologues of a *Xenopus mix* gene. *Developmental dynamics : an official publication of the American Association of Anatomists*, **219**(4):497–504, 2000.
- Robb, L. and Tam, P. P. L. Gastrula organiser and embryonic patterning in the mouse. *Seminars in cell & developmental biology*, **15**(5):543–54, 2004.
- Rosa, F. M. Mix.1, a homeobox mRNA inducible by mesoderm inducers, is expressed mostly in the presumptive endodermal cells of *Xenopus* embryos. *Cell*, **57**(6):965–974, 1989.
- Rossant, J. and Tam, P. P. *Mouse Development*. Academic Press, 2002. ISBN 978-0-12-597951-1.
- Saint-Jeannet, J. P., He, X., Varmus, H. E., and Dawid, I. B. Regulation of dorsal fate in the neuraxis by Wnt-1 and Wnt-3a. *Proceedings of the National Academy of Sciences of the United States of America*, **94**(25):13713–13718, 1997.
- Sakai, Y. The retinoic acid-inactivating enzyme CYP26 is essential for establishing an uneven distribution of retinoic acid along the antero-posterior axis within the mouse embryo. *Genes & Development*, **15**(2):213–225, 2001.

References

- Saldanha, A. J. Java Treeview—extensible visualization of microarray data. *Bioinformatics (Oxford, England)*, **20**(17):3246–8, 2004.
- Sasaki, H. and Hogan, B. Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development*, **118**(1):47–59, 1993.
- Sauka-Spengler, T. and Bronner-Fraser, M. A gene regulatory network orchestrates neural crest formation. *Nature reviews. Molecular cell biology*, **9**(7):557–68, 2008.
- Serbedzija, G. N. and McMahon, A. P. Analysis of neural crest cell migration in Splotch mice using a neural crest-specific LacZ reporter. *Developmental biology*, **185**(2):139–47, 1997.
- Sherwood, R. I., Jitianu, C., Cleaver, O., Shaywitz, D. A., Lamenzo, J. O., Chen, A. E., Golub, T. R., and Melton, D. A. Prospective isolation and global gene expression analysis of definitive and visceral endoderm. *Developmental biology*, **304**(2):541–55, 2007.
- Sherwood, R. I., Chen, T.-Y. A., and Melton, D. A. Transcriptional dynamics of endodermal organ formation. *Developmental dynamics : an official publication of the American Association of Anatomists*, **238**(1):29–42, 2009.
- Silberg, D. G., Swain, G. P., Suh, E. R., and Traber, P. G. Cdx1 and Cdx2 expression during intestinal development. *Gastroenterology*, **119**(4):961–971, 2000.
- Simões Costa, M. and Bronner, M. E. Insights into neural crest development and evolution from genomic analysis. *Genome research*, **23**(7):1069–80, 2013.
- Siu, F. K., Sham, M. H., and Chow, B. K. Secretin, a known gastrointestinal peptide, is widely expressed during mouse embryonic development. *Gene expression patterns : GEP*, **5**(3):445–451, 2005.
- Soriano, P. Generalized lacZ expression with the ROSA26 Cre reporter strain. *Nature genetics*, **21**(1):70–71, 1999.
- Stemmler, M. P., Hecht, A., Kinzel, B., and Kemler, R. Analysis of regulatory elements of E-cadherin with reporter gene constructs in transgenic mouse embryos. *Developmental dynamics : an official publication of the American Association of Anatomists*, **227**(2):238–45, 2003.
- Stern, C. D. and Downs, K. M. The hypoblast (visceral endoderm): an evo-devo perspective. *Development (Cambridge, England)*, **139**(6):1059–69, 2012.

- Sugimoto, K., Ichikawa-Tomikawa, N., Satohisa, S., Akashi, Y., Kanai, R., Saito, T., Sawada, N., and Chiba, H. The tight-junction protein claudin-6 induces epithelial differentiation from mouse F9 and embryonic stem cells. *PloS one*, **8**(10):e75106, 2013.
- Sulik, K., Dehart, D. B., Inagaki, T., Carson, J. L., Vrablic, T., Gesteland, K., and Schoenwolf, G. C. Morphogenesis of the Murine Node and Notochordal Plate. *Developmental dynamics : an official publication of the American Association of Anatomists*, **201**:260–278, 1994.
- Takada, S., Stark, K. L., Shea, M. J., Vassileva, G., McMahon, J. A., and McMahon, A. P. Wnt-3a regulates somite and tailbud formation in the mouse embryo. *Genes & Development*, **8**(2):174–189, 1994.
- Takakura, N., Yoshida, H., Ogura, Y., Kataoka, H., Nishikawa, S., and Nishikawa, S.-I. PDGFR Expression During Mouse Embryogenesis: Immunolocalization Analyzed by Whole-mount Immunohistostaining Using the Monoclonal Anti-mouse PDGFR Antibody APA5. *Journal of Histochemistry & Cytochemistry*, **45**(6):883–893, 1997.
- Takemoto, T., Uchikawa, M., Yoshida, M., Bell, D. M., Lovell-Badge, R., Papaioannou, V. E., and Kondoh, H. Tbx6-dependent Sox2 regulation determines neural or mesodermal fate in axial stem cells. *Nature*, **470**(7334):394–8, 2011.
- Tam, P. P. The control of somitogenesis in mouse embryos. *Journal of embryology and experimental morphology*, **65 Suppl**:103–128, 1981.
- Tam, P. P. and Behringer, R. R. Mouse gastrulation: the formation of a mammalian body plan. *Mech Dev*, **68**(1-2):3–25, 1997a.
- Tam, P. P. L. The histogenetic capacity of tissues in the caudal end of the embryonic axis of the mouse. *J Embryol Exp Morphol*, **82**(1):253–266, 1984.
- Tam, P. P. L. and Beddington, R. S. P. The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development*, **99**:109–126, 1987.
- Tam, P. P. and Behringer, R. R. Mouse gastrulation: the formation of a mammalian body plan. *Mechanisms of Development*, **68**(1-2):3–25, 1997b.
- Tamplin, O. J., Kinzel, D., Cox, B. J., Bell, C. E., Rossant, J., and Lickert, H. Microarray analysis of Foxa2 mutant mouse embryos reveals novel gene expression and inductive roles for the gastrula organizer and its derivatives. *BMC genomics*, **9**:511, 2008.
- Tanaka, C., Sakuma, R., Nakamura, T., Hamada, H., and Saijoh, Y. Long-range action of Nodal requires interaction with GDF1. *Genes & development*, **21**(24):3272–82, 2007.

References

- Tanaka, M., Hadjantonakis, A.-K., Vintersten, K., and Nagy, A. Aggregation chimeras: combining ES cells, diploid, and tetraploid embryos. *Methods in molecular biology (Clifton, N.J.)*, **530**:287–309, 2009.
- Tang, S. J., Hoodless, P. A., Lu, Z., Breitman, M. L., McInnes, R. R., Wrana, J. L., and Buchwald, M. The Tlx-2 homeobox gene is a downstream target of BMP signalling and is required for mouse mesoderm development. *Development*, **125**, 1998.
- Turksen, K. and Troy, T. C. Claudin-6: a novel tight junction molecule is developmentally regulated in mouse embryonic epithelium. *Developmental dynamics : an official publication of the American Association of Anatomists*, **222**(2):292–300, 2001.
- Urniss, L. D., Bleyl, S. B., Wright, T. J., Moon, A. M., and Mansour, S. L. Redundant and dosage sensitive requirements for Fgf3 and Fgf10 in cardiovascular development. *Developmental Biology*, **356**(2):383–397, 2011.
- Vincentz, J. W., McWhirter, J. R., Murre, C., Baldini, A., and Furuta, Y. Fgf15 is required for proper morphogenesis of the mouse cardiac outflow tract. *Genesis (New York, N.Y. : 2000)*, **41**(4):192–201, 2005.
- Viotti, M., Nowotschin, S., and Hadjantonakis, A.-K. SOX17 links gut endoderm morphogenesis and germ layer segregation. *Nature cell biology*, **16**(12):1146–1156, 2014.
- Vogt, W. Gestaltungsanalyse am Amphibienkeim mit Örtlicher Vitalfärbung. *Wilhelm Roux' Archiv für Entwicklungsmechanik der Organismen*, **120**:384–706, 1929.
- Wall, N. A., Craig, E. J., Labosky, P. A., and Kessler, D. S. Mesendoderm induction and reversal of left-right pattern by mouse Gdf1, a Vg1-related gene. *Developmental biology*, **227**(2):495–509, 2000.
- Wang, H. and Dey, S. K. Roadmap to embryo implantation: clues from mouse models. *Nature reviews. Genetics*, **7**(3):185–99, 2006.
- Wang, L., Xue, Y., Shen, Y., Li, W., Cheng, Y., Yan, X., Shi, W., Wang, J., Gong, Z., Yang, G., Guo, C., Zhou, Y., Wang, X., Zhou, Q., and Zeng, F. Claudin 6: a novel surface marker for characterizing mouse pluripotent stem cells. *Cell research*, **22**(6):1082–5, 2012.
- Wang, X., Chan, A. K. K., Sham, M. H., Burns, A. J., and Chan, W. Y. Analysis of the sacral neural crest cell contribution to the hindgut enteric nervous system in the mouse embryo. *Gastroenterology*, **141**(3):992–1002.e1–6, 2011.

- Wareing, S., Eliades, A., Lacaud, G., and Kouskoff, V. ETV2 expression marks blood and endothelium precursors, including hemogenic endothelium, at the onset of blood development. *Developmental dynamics : an official publication of the American Association of Anatomists*, **241**(9):1454–64, 2012.
- Watanabe, S., Chey, W. Y., Lee, K. Y., and Chang, T. M. Secretin is released by digestive products of fat in dogs. *Gastroenterology*, **90**(4):1008–17, 1986.
- White, P. H., Farkas, D. R., McFadden, E. E., and Chapman, D. L. Defective somite patterning in mouse embryos with reduced levels of Tbx6. *Development (Cambridge, England)*, **130**:1681–1690, 2003.
- Wickramasinghe, S. R., Alvania, R. S., Ramanan, N., Wood, J. N., Mandai, K., and Ginty, D. D. Serum response factor mediates NGF-dependent target innervation by embryonic DRG sensory neurons. *Neuron*, **58**(4):532–45, 2008.
- Wilkinson, D. G., Bhatt, S., and Herrmann, B. G. Expression pattern of the mouse T gene and its role in mesoderm formation. *Nature*, **343**(6259):657–9, 1990.
- Williams, B. and Ordahl, C. Pax-3 expression in segmental mesoderm marks early stages in myogenic cell specification. *Development*, **120**(4):785–796, 1994.
- Wilm, B., James, R. G., Schultheiss, T. M., and Hogan, B. L. M. The forkhead genes, Foxc1 and Foxc2, regulate paraxial versus intermediate mesoderm cell fate. *Developmental biology*, **271**(1):176–89, 2004.
- Winnier, G., Blessing, M., Labosky, P. A., and Hogan, B. L. Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes & Development*, **9**(17):2105–2116, 1995.
- Wolfe, A. D. and Downs, K. M. Mixl1 localizes to putative axial stem cell reservoirs and their posterior descendants in the mouse embryo. *Gene expression patterns : GEP*, **15**(1):8–20, 2014.
- Wright, T. J., Ladher, R., McWhirter, J., Murre, C., Schoenwolf, G. C., and Mansour, S. L. Mouse FGF15 is the ortholog of human and chick FGF19, but is not uniquely required for otic induction. *Developmental biology*, **269**(1):264–75, 2004.
- Yamanaka, Y., Tamplin, O. J., Beckers, A., Gossler, A., and Rossant, J. Live imaging and genetic analysis of mouse notochord formation reveals regional morphogenetic mechanisms. *Developmental cell*, **13**(6):884–96, 2007.
- Yardley, N. and García-Castro, M. I. FGF signaling transforms non-neural ectoderm into neural crest. *Developmental Biology*, **372**(2):166–177, 2012.

References

- Zhang, Y., Buchholz, F., Muyrers, J. P. P., and Stewart, A. F. A new logic for DNA engineering using recombination in *Escherichia coli*. *Nature new technology*, **20**(2):123–128, 1998.
- Zhao, L., Yaoita, E., Nameta, M., Zhang, Y., Cuellar, L. M., Fujinaka, H., Xu, B., Yoshida, Y., Hatakeyama, K., and Yamamoto, T. Claudin-6 localized in tight junctions of rat podocytes. *American journal of physiology. Regulatory, integrative and comparative physiology*, **294**(6):R1856–62, 2008.
- Zhao, T., Gan, Q., Stokes, A., Lassiter, R. N. T., Wang, Y., Chan, J., Han, J. X., Pleasure, D. E., Epstein, J. a., and Zhou, C. J. β -catenin regulates Pax3 and Cdx2 for caudal neural tube closure and elongation. *Development*, **141**(2014):148–157, 2014.
- Zizic Mitrecic, M., Mitrecic, D., Pochet, R., Kostovic-Knezevic, L., and Gajovic, S. The mouse gene Noto is expressed in the tail bud and essential for its morphogenesis. *Cells, tissues, organs*, **192**(2):85–92, 2010.
- Zorn, A. M. and Wells, J. M. Vertebrate endoderm development and organ formation. *Annual review of cell and developmental biology*, **25**:221–51, 2009.

Chapter 7

Supplementary

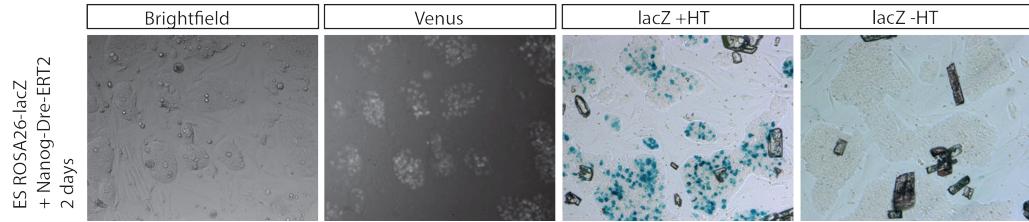


Figure 7.1: In vitro induction of the lineage tracing construct using Dre recombinase

Cells were incubated 2 days with Hydroxy-Tamoxifen. Venus fluorescence of the *Nanog* controlled H2B-Venus was monitored. After staining with X-Gal, cells were fixed and pictures captured. Left: Cells in brightfield; second from left: Venus expression of the cells; second from right: lacZ staining of Hydroxy-Tamoxifen treated cells; right: lacZ staining of untreated cells.

Table 7.1: Biological processes of genes listed in Cluster I

Biological process	GO #	# of Genes	p-value
anatomical structure morphogenesis	GO:0009653	30/1835	1.41E-07
tissue development	GO:0009888	26/1415	2.17E-07
chordate embryonic development	GO:0043009	17/636	1.07E-06
embryo development ending in birth or egg hatching	GO:0009792	17/643	1.26E-06
somitogenesis	GO:0001756	7/62	2.75E-06
epithelium development	GO:0060429	19/863	2.89E-06
embryo development	GO:0009790	20/988	4.66E-06
somite development	GO:0061053	7/70	6.24E-06
anterior/posterior pattern specification	GO:0009952	10/204	6.60E-06
segmentation	GO:0035282	7/80	1.53E-05
pattern specification process	GO:0007389	13/430	1.71E-05
tissue morphogenesis	GO:0048729	14/516	2.03E-05
regulation of multicellular organismal process	GO:0051239	29/2156	2.08E-05

Biological process	GO #	# of Genes	p-value
anatomical structure formation involved in morphogenesis	GO:0048646	17/815	3.64E-05
regionalization	GO:0003002	11/315	4.06E-05
anatomical structure development	GO:0048856	40/3847	4.88E-05
single-organism developmental process	GO:0044767	43/4327	4.97E-05
developmental process	GO:0032502	43/4339	5.38E-05

Table 7.2: Biological processes of genes listed in Cluster II

Biological process	GO #	# of Genes	p-value
anatomical structure morphogenesis	GO:0009653	70/1835	4.45E-29
blood vessel development	GO:0001568	39/420	6.35E-28
circulatory system development	GO:0072359	47/715	8.62E-28
cardiovascular system development	GO:0072358	47/715	8.62E-28
vasculature development	GO:0001944	39/452	9.25E-27
system development	GO:0048731	87/3278	1.71E-26
anatomical structure development	GO:0048856	93/3847	6.04E-26
developmental process	GO:0032502	97/4339	7.18E-25
single-organism developmental process	GO:0044767	96/4327	3.19E-24
organ morphogenesis	GO:0009887	45/785	5.04E-24
multicellular organismal development	GO:0007275	90/3829	7.82E-24
blood vessel morphogenesis	GO:0048514	33/352	2.35E-23
embryo development	GO:0009790	48/988	8.86E-23
organ development	GO:0048513	71/2423	1.25E-22
heart development	GO:0007507	33/414	3.43E-21
chordate embryonic development	GO:0043009	37/636	1.93E-19
embryo development ending in birth or egg hatching	GO:0009792	37/643	2.78E-19
tissue development	GO:0009888	50/1415	6.73E-18
skeletal system development	GO:0001501	29/389	1.38E-17
muscle structure development	GO:0061061	29/396	2.22E-17
tube development	GO:0035295	32/546	1.38E-16
cell differentiation	GO:0030154	66/2822	2.12E-15
regulation of biological process	GO:0050789	129/9748	2.75E-15
cellular developmental process	GO:0048869	67/2964	6.06E-15
regulation of cell proliferation	GO:0042127	43/1225	7.74E-15

Biological process	GO #	# of Genes	p-value
biological regulation	GO:0065007	131/10145	8.40E-15
multicellular organismal process	GO:0032501	98/6004	9.39E-15
embryonic organ development	GO:0048568	27/426	1.65E-14
regulation of multicellular organismal process	GO:0051239	56/2156	2.19E-14
anatomical structure formation involved in morphogenesis	GO:0048646	35/815	3.37E-14
single-multicellular organism process	GO:0044707	95/5830	5.96E-14
positive regulation of cellular process	GO:0048522	75/3833	7.85E-14
embryonic morphogenesis	GO:0048598	29/544	8.88E-14
positive regulation of biological process	GO:0048518	79/4248	1.31E-13
heart morphogenesis	GO:0003007	20/211	1.60E-13
epithelium development	GO:0060429	35/863	1.87E-13
in utero embryonic development	GO:0001701	25/396	2.91E-13
cell fate commitment	GO:0045165	20/232	9.35E-13
tissue morphogenesis	GO:0048729	27/516	1.66E-12
mesenchyme development	GO:0060485	17/152	2.05E-12
regulation of cellular process	GO:0050794	121/9335	2.21E-12
regulation of macromolecule metabolic process	GO:0060255	77/4399	1.22E-11
cell development	GO:0048468	40/1338	2.46E-11
morphogenesis of an epithelium	GO:0002009	23/402	3.46E-11
muscle tissue development	GO:0060537	20/283	3.59E-11
cardiac muscle tissue development	GO:0048738	16/154	4.05E-11
tube morphogenesis	GO:0035239	21/327	5.11E-11
regulation of developmental process	GO:0050793	46/1806	6.44E-11
regulation of transcription from RNA polymerase II promoter	GO:0006357	41/1454	7.46E-11
regulation of metabolic process	GO:0019222	83/5177	8.46E-11
angiogenesis	GO:0001525	19/261	9.25E-11
positive regulation of macromolecule biosynthetic process	GO:0010557	39/1329	9.70E-11
negative regulation of cellular process	GO:0048523	65/3418	1.02E-10
regulation of locomotion	GO:0040012	27/615	1.02E-10
regulation of anatomical structure morphogenesis	GO:0022603	29/729	1.37E-10
striated muscle tissue development	GO:0014706	19/268	1.46E-10

Biological process	GO #	# of Genes	p-value
regulation of multicellular organismal development	GO:2000026	40/1420	1.63E-10
artery morphogenesis	GO:0048844	11/52	1.83E-10
epithelial tube morphogenesis	GO:0060562	20/310	1.87E-10
negative regulation of biological process	GO:0048519	68/3751	2.05E-10
digestive system development	GO:0055123	14/116	2.06E-10
regulation of macromolecule biosynthetic process	GO:0010556	61/3118	2.40E-10
cellular response to growth factor stimulus	GO:0071363	18/240	2.50E-10
positive regulation of macromolecule metabolic process	GO:0010604	49/2114	2.71E-10
sensory organ development	GO:0007423	24/493	2.90E-10
skeletal system morphogenesis	GO:0048705	17/211	3.69E-10
urogenital system development	GO:0001655	18/249	4.57E-10
response to growth factor	GO:0070848	18/249	4.57E-10
regulation of gene expression	GO:0010468	66/3630	4.89E-10
vasculogenesis	GO:0001570	12/77	5.31E-10
embryonic organ morphogenesis	GO:0048562	19/289	5.35E-10
artery development	GO:0060840	11/58	5.85E-10
regulation of biosynthetic process	GO:0009889	62/3295	7.85E-10
cardiac chamber development	GO:0003205	14/129	8.36E-10
regulation of cellular component movement	GO:0051270	26/622	8.79E-10
cardiocyte differentiation	GO:0035051	13/105	1.04E-09
pattern specification process	GO:0007389	22/430	1.09E-09
regulation of nucleobase-containing compound metabolic process	GO:0019219	63/3418	1.18E-09
regulation of apoptotic process	GO:0042981	35/1171	1.25E-09
positive regulation of biosynthetic process	GO:0009891	39/1458	1.65E-09
regulation of programmed cell death	GO:0043067	35/1187	1.82E-09
regulation of cellular macromolecule biosynthetic process	GO:2000112	58/3007	2.12E-09
cardiac chamber morphogenesis	GO:0003206	13/112	2.30E-09
regulation of nitrogen compound metabolic process	GO:0051171	63/3471	2.32E-09
extracellular matrix organization	GO:0030198	15/171	2.61E-09
regulation of RNA metabolic process	GO:0051252	56/2846	2.67E-09

Biological process	GO #	# of Genes	p-value
extracellular structure organization	GO:0043062	15/172	2.83E-09
regulation of RNA biosynthetic process	GO:2001141	55/2765	2.93E-09
regulation of cellular metabolic process	GO:0031323	76/4777	3.11E-09
regulation of cell motility	GO:2000145	24/556	3.48E-09
regulation of cellular biosynthetic process	GO:0031326	60/3229	3.67E-09
regulation of primary metabolic process	GO:0080090	74/4601	4.29E-09
renal system development	GO:0072001	16/211	4.31E-09
coronary vasculature development	GO:0060976	9/36	5.89E-09
regulation of transcription, DNA-templated	GO:0006355	54/2748	7.89E-09
regulation of nucleic acid-templated transcription	GO:1903506	54/2748	7.89E-09
morphogenesis of a branching epithelium	GO:0061138	15/186	8.37E-09
regulation of cell migration	GO:0030334	23/533	9.87E-09
muscle cell differentiation	GO:0042692	16/224	1.03E-08
regulation of cell death	GO:0010941	35/1266	1.07E-08
regulation of localization	GO:0032879	43/1857	1.10E-08
regulation of signal transduction	GO:0009966	46/2096	1.11E-08
cardiac ventricle development	GO:0003231	12/102	1.32E-08
morphogenesis of a branching structure	GO:0001763	15/193	1.39E-08
branching morphogenesis of an epithelial tube	GO:0048754	14/160	1.39E-08
regulation of epithelial cell migration	GO:0010632	13/130	1.42E-08
positive regulation of epithelial cell migration	GO:0010634	11/79	1.53E-08
outflow tract morphogenesis	GO:0003151	10/58	1.56E-08
positive regulation of cellular biosynthetic process	GO:0031328	37/1429	1.67E-08
positive regulation of transcription from RNA polymerase II promoter	GO:0045944	29/894	1.81E-08
digestive tract development	GO:0048565	12/105	1.83E-08
limb development	GO:0060173	14/165	2.07E-08
appendage development	GO:0048736	14/165	2.07E-08
coronary vasculature morphogenesis	GO:0060977	8/27	2.16E-08
positive regulation of nucleic acid-templated transcription	GO:1903508	33/1163	2.25E-08
positive regulation of transcription, DNA-templated	GO:0045893	33/1163	2.25E-08

Biological process	GO #	# of Genes	p-value
muscle organ development	GO:0007517	16/237	2.34E-08
positive regulation of RNA biosynthetic process	GO:1902680	33/1166	2.40E-08
positive regulation of gene expression	GO:0010628	38/1521	2.42E-08
mesenchymal cell differentiation	GO:0048762	12/109	2.79E-08
positive regulation of cellular component movement	GO:0051272	18/324	3.22E-08
positive regulation of locomotion	GO:0040017	18/329	4.11E-08
positive regulation of RNA metabolic process	GO:0051254	33/1200	5.03E-08
positive regulation of multicellular organismal process	GO:0051240	33/1200	5.03E-08
connective tissue development	GO:0061448	14/178	5.51E-08
enzyme linked receptor protein signaling pathway	GO:0007167	20/430	5.99E-08
regulation of signaling	GO:0023051	48/2376	6.22E-08
limb morphogenesis	GO:0035108	13/147	6.29E-08
appendage morphogenesis	GO:0035107	13/147	6.29E-08
regulation of cell communication	GO:0010646	48/2381	6.68E-08
positive regulation of cell proliferation	GO:0008284	25/712	9.22E-08
cardiac ventricle morphogenesis	GO:0003208	10/70	9.43E-08
transcription from RNA polymerase II promoter	GO:0006366	22/547	1.03E-07
positive regulation of developmental process	GO:0051094	29/964	1.04E-07
mesenchymal cell development	GO:0014031	11/95	1.05E-07
kidney development	GO:0001822	14/192	1.45E-07
positive regulation of cell motility	GO:2000147	17/313	1.56E-07
cellular response to chemical stimulus	GO:0070887	34/1327	1.58E-07
muscle organ morphogenesis	GO:0048644	10/75	1.82E-07
single-organism process	GO:0044699	129/11868	2.04E-07
positive regulation of cellular metabolic process	GO:0031325	48/2464	2.12E-07
cell migration	GO:0016477	22/570	2.19E-07
cellular response to endogenous stimulus	GO:0071495	22/571	2.27E-07
regulation of response to stimulus	GO:0048583	50/2653	2.53E-07
positive regulation of metabolic process	GO:0009893	50/2659	2.73E-07

Biological process	GO #	# of Genes	p-value
stem cell development	GO:0048864	14/204	3.14E-07
reproductive structure development	GO:0048608	18/375	3.22E-07
reproductive system development	GO:0061458	18/380	3.95E-07
gland development	GO:0048732	17/335	4.28E-07
locomotion	GO:0040011	26/835	4.82E-07
cellular response to organic substance	GO:0071310	29/1036	5.32E-07
nervous system development	GO:0007399	37/1626	5.95E-07
positive regulation of nucleobase-containing compound metabolic process	GO:0045935	37/1627	6.05E-07
negative regulation of apoptotic process	GO:0043066	24/723	6.55E-07
stem cell differentiation	GO:0048863	15/257	6.74E-07
transcription, DNA-templated	GO:0006351	40/1878	7.36E-07
nucleic acid-templated transcription	GO:0097659	40/1878	7.36E-07
RNA biosynthetic process	GO:0032774	40/1885	8.18E-07
palate development	GO:0060021	10/88	8.27E-07
positive regulation of cell migration	GO:0030335	16/306	8.91E-07
negative regulation of programmed cell death	GO:0043069	24/736	9.25E-07
positive regulation of nitrogen compound metabolic process	GO:0051173	37/1656	9.73E-07
cardiac septum development	GO:0003279	9/66	1.14E-06
cardiac right ventricle morphogenesis	GO:0003215	6/16	1.46E-06
localization of cell	GO:0051674	22/635	1.56E-06
cell motility	GO:0048870	22/635	1.56E-06
ameboidal-type cell migration	GO:0001667	11/124	1.62E-06
embryonic skeletal system development	GO:0048706	11/127	2.07E-06
response to endogenous stimulus	GO:0009719	25/833	2.16E-06
regulation of cell differentiation	GO:0045595	32/1321	2.18E-06
embryonic limb morphogenesis	GO:0030326	11/128	2.24E-06
embryonic appendage morphogenesis	GO:0035113	11/128	2.24E-06
cardiac septum morphogenesis	GO:0060411	8/51	3.01E-06
regulation of transmembrane receptor protein serine/threonine kinase signaling pathway	GO:0090092	12/170	3.89E-06
neural crest cell development	GO:0014032	8/53	4.04E-06
aromatic compound biosynthetic process	GO:0019438	42/2168	4.20E-06
negative regulation of cell death	GO:0060548	24/797	4.23E-06

Biological process	GO #	# of Genes	p-value
cell migration involved in heart development	GO:0060973	5/9	4.37E-06
positive regulation of endothelial cell migration	GO:0010595	8/54	4.66E-06
neurogenesis	GO:0022008	30/1219	4.99E-06
nucleobase-containing compound biosynthetic process	GO:0034654	41/2100	5.33E-06
response to organic substance	GO:0010033	35/1611	5.84E-06
negative regulation of transcription from RNA polymerase II promoter	GO:0000122	21/632	7.36E-06
neural crest cell differentiation	GO:0014033	8/59	9.16E-06
eye development	GO:0001654	15/314	9.38E-06
macromolecule biosynthetic process	GO:0009059	46/2591	9.94E-06
transmembrane receptor protein tyrosine kinase signaling pathway	GO:0007169	14/271	1.08E-05
kidney vasculature development	GO:0061440	5/11	1.18E-05
renal system vasculature development	GO:0061437	5/11	1.18E-05
positive regulation of ERK1 and ERK2 cascade	GO:0070374	10/117	1.18E-05
heterocycle biosynthetic process	GO:0018130	41/2162	1.20E-05
negative regulation of nucleobase-containing compound metabolic process	GO:0045934	27/1052	1.27E-05
regulation of cellular response to growth factor stimulus	GO:0090287	11/152	1.27E-05
organic cyclic compound biosynthetic process	GO:1901362	42/2266	1.48E-05
regulation of endothelial cell migration	GO:0010594	9/90	1.59E-05
cellular component movement	GO:0006928	25/928	1.75E-05
macromolecule metabolic process	GO:0043170	75/5609	1.83E-05
negative regulation of response to stimulus	GO:0048585	27/1072	1.86E-05
negative regulation of nitrogen compound metabolic process	GO:0051172	27/1074	1.93E-05
embryonic cranial skeleton morphogenesis	GO:0048701	7/43	2.07E-05
cellular nitrogen compound biosynthetic process	GO:0044271	41/2206	2.10E-05
response to acid chemical	GO:0001101	11/162	2.40E-05
muscle tissue morphogenesis	GO:0060415	8/67	2.41E-05

Biological process	GO #	# of Genes	p-value
single-organism cellular process	GO:0044763	113/10325	2.41E-05
determination of bilateral symmetry	GO:0009855	9/95	2.50E-05
vascular endothelial growth factor signaling pathway	GO:0038084	5/13	2.68E-05
specification of symmetry	GO:0009799	9/96	2.73E-05
embryonic skeletal system morphogenesis	GO:0048704	9/97	2.98E-05
developmental process involved in reproduction	GO:0003006	19/565	3.06E-05
positive regulation of cell adhesion	GO:0045785	11/167	3.25E-05
patterning of blood vessels	GO:0001569	7/46	3.26E-05
endocardial cushion development	GO:0003197	6/28	3.88E-05
negative regulation of RNA biosynthetic process	GO:1902679	24/900	4.04E-05
regulation of epithelial cell proliferation	GO:0050678	13/258	4.58E-05
negative regulation of cell proliferation	GO:0008285	18/525	5.19E-05
cartilage development	GO:0051216	10/138	5.38E-05
vasculogenesis involved in coronary vascular morphogenesis	GO:0060979	5/15	5.41E-05
embryonic heart tube development	GO:0035050	8/75	5.64E-05
endocrine system development	GO:0035270	9/106	6.24E-05
regionalization	GO:0003002	14/315	6.67E-05
negative regulation of multicellular organismal process	GO:0051241	23/863	7.74E-05
negative regulation of RNA metabolic process	GO:0051253	24/933	7.76E-05
lung development	GO:0030324	11/183	8.01E-05
negative regulation of cellular biosynthetic process	GO:0031327	26/1081	8.27E-05
placenta development	GO:0001890	10/145	8.44E-05
respiratory tube development	GO:0030323	11/186	9.39E-05

Table 7.3: Biological processes of genes listed in Cluster III

Biological process	GO #	# of Genes	p-value
anatomical structure morphogenesis	GO:0009653	33/1835	5.18E-03
single-organism developmental process	GO:0044767	59/4327	1.02E-02

Biological process	GO #	# of Genes	p-value
developmental process	GO:0032502	59/4339	1.10E-02
neural tube development	GO:0021915	8/153	1.13E-02
cell differentiation	GO:0030154	43/2822	1.36E-02
cellular component organization	GO:0016043	52/3712	1.74E-02
regulation of neuron apoptotic process	GO:0043523	9/210	1.75E-02
locomotion	GO:0040011	19/835	1.83E-02
negative regulation of glial cell differentiation	GO:0045686	4/27	1.92E-02
multicellular organismal development	GO:0007275	53/3829	1.99E-02
cellular developmental process	GO:0048869	44/2964	2.01E-02
cellular component organization or biogenesis	GO:0071840	53/3832	2.03E-02
organ development	GO:0048513	38/2423	2.23E-02
cellular component movement	GO:0006928	20/928	2.46E-02
system development	GO:0048731	47/3278	2.50E-02
regulation of timing of cell differentiation	GO:0048505	3/11	2.61E-02
negative regulation of nervous system development	GO:0051961	9/222	2.65E-02
negative regulation of forebrain neuron differentiation	GO:2000978	2/2	3.02E-02
positive regulation of developmental process	GO:0051094	20/964	4.06E-02
regulation of development, heterochronic	GO:0040034	3/13	4.26E-02
rostrocaudal neural tube patterning	GO:0021903	3/13	4.26E-02
apoptotic process	GO:0006915	18/828	4.94E-02
tube development	GO:0035295	14/546	4.99E-02
negative regulation of gliogenesis	GO:0014014	4/35	5.16E-02
blood vessel development	GO:0001568	12/420	5.38E-02
regulation of cell proliferation	GO:0042127	23/1225	5.58E-02
negative regulation of cell development	GO:0010721	9/246	5.63E-02
cell projection organization	GO:0030030	17/767	5.94E-02
regulation of nervous system development	GO:0051960	15/629	6.44E-02
pattern specification process	GO:0007389	12/430	6.65E-02
negative regulation of cellular process	GO:0048523	47/3418	6.70E-02
negative regulation of inner ear receptor cell differentiation	GO:2000981	2/3	6.75E-02

Biological process	GO #	# of Genes	p-value
negative regulation of pro-B cell differentiation	GO:2000974	2/3	6.75E-02
negative regulation of mechanoreceptor differentiation	GO:0045632	2/3	6.75E-02
trochlear nerve development	GO:0021558	2/3	6.75E-02
auditory receptor cell fate determination	GO:0042668	2/3	6.75E-02
regulation of transcription from RNA polymerase II promoter involved in spinal cord motor neuron fate specification	GO:0021912	2/3	6.75E-02
programmed cell death	GO:0012501	18/852	6.95E-02
regulation of neuron death	GO:1901214	9/254	7.10E-02
sensory perception of chemical stimulus	GO:0007606	1/1223	7.13E-02
negative regulation of neurogenesis	GO:0050768	8/201	7.13E-02
tissue development	GO:0009888	25/1415	7.39E-02
negative regulation of neuron differentiation	GO:0045665	7/153	7.43E-02
regionalization	GO:0003002	10/315	7.90E-02
anatomical structure development	GO:0048856	51/3847	8.30E-02
negative regulation of biological process	GO:0048519	50/3751	8.50E-02
regulation of cell division	GO:0051302	8/207	8.65E-02
regulation of cellular component organization	GO:0051128	27/1604	8.97E-02
regulation of biological quality	GO:0065008	35/2326	9.61E-02

Table 7.4: Biological processes of genes listed in Cluster IV

Biological process	GO #	# of Genes	p-value
erythrocyte development	GO:0048821	4/32	1.42E-03
somite rostral/caudal axis specification	GO:0032525	3/13	2.71E-03
erythrocyte differentiation	GO:0030218	5/77	2.87E-03
erythrocyte homeostasis	GO:0034101	5/86	4.81E-03
gas transport	GO:0015669	3/20	9.64E-03
myeloid cell development	GO:0061515	4/53	9.87E-03
myeloid cell homeostasis	GO:0002262	5/104	1.16E-02
mesodermal cell migration	GO:0008078	2/4	1.29E-02
somitogenesis	GO:0001756	4/62	1.79E-02

Biological process	GO #	# of Genes	p-value
iron ion homeostasis	GO:0055072	4/65	2.14E-02
homeostasis of number of cells	GO:0048872	6/193	2.67E-02
somite development	GO:0061053	4/70	2.82E-02
mesoderm migration involved in gastrulation	GO:0007509	2/6	2.88E-02
Notch signaling pathway	GO:0007219	4/79	4.43E-02
embryonic axis specification	GO:0000578	3/34	4.51E-02
segmentation	GO:0035282	4/80	4.64E-02
homeostatic process	GO:0042592	14/1076	5.07E-02
metal ion homeostasis	GO:0055065	8/397	5.17E-02
transition metal ion homeostasis	GO:0055076	4/94	8.39E-02

Table 7.5: Biological processes of genes listed in Cluster V

Biological process	GO #	# of Genes	p-value
cell fate specification	GO:0001708	8/68	1.75E-06
cellular developmental process	GO:0048869	43/2964	1.86E-06
neurogenesis	GO:0022008	26/1219	2.20E-06
nervous system development	GO:0007399	30/1626	3.61E-06
cell differentiation	GO:0030154	41/2822	4.55E-06
neuron differentiation	GO:0030182	20/769	5.71E-06
anatomical structure development	GO:0048856	49/3847	7.74E-06
negative regulation of biological process	GO:0048519	48/3751	9.96E-06
generation of neurons	GO:0048699	24/1146	1.23E-05
negative regulation of cellular process	GO:0048523	44/3418	4.01E-05
single-organism developmental process	GO:0044767	51/4327	4.46E-05
developmental process	GO:0032502	51/4339	4.87E-05
cell development	GO:0048468	25/1338	5.40E-05
system development	GO:0048731	42/3278	9.77E-05
epithelial tube branching involved in lung morphogenesis	GO:0060441	5/28	1.50E-04
central nervous system development	GO:0007417	16/616	1.55E-04
cell fate commitment	GO:0045165	10/232	2.60E-04
neuron fate specification	GO:0048665	5/33	3.32E-04
multicellular organismal development	GO:0007275	45/3829	3.66E-04
neuron fate commitment	GO:0048663	6/75	1.17E-03

Biological process	GO #	# of Genes	p-value
single-multicellular organism process	GO:0044707	58/5830	1.36E-03
lung morphogenesis	GO:0060425	5/47	1.82E-03
morphogenesis of a branching epithelium	GO:0061138	8/186	3.06E-03
multicellular organismal process	GO:0032501	58/6004	3.52E-03
morphogenesis of a branching structure	GO:0001763	8/193	3.97E-03
negative regulation of transcription from RNA polymerase II promoter	GO:0000122	14/632	4.56E-03
anatomical structure morphogenesis	GO:0009653	26/1835	4.64E-03
organ development	GO:0048513	31/2423	4.91E-03
brain development	GO:0007420	12/480	5.63E-03
developmental process involved in reproduction	GO:0003006	13/565	6.12E-03
single-organism process	GO:0044699	93/11868	7.40E-03
negative regulation of mesenchymal cell apoptotic process	GO:2001054	3/11	7.57E-03
negative regulation of cellular carbohydrate metabolic process	GO:0010677	4/32	7.82E-03
endocrine system development	GO:0035270	6/106	8.02E-03
regulation of transcription from RNA polymerase II promoter	GO:0006357	22/1454	8.20E-03
branching morphogenesis of an epithelial tube	GO:0048754	7/160	9.35E-03
negative regulation of gene expression	GO:0010629	21/1363	9.64E-03

Table 7.6: Biological processes of genes listed in Cluster VI

Biological process	GO #	# of Genes	p-value
cardiac atrium development	GO:0003230	3/29	5.32E-03
atrial cardiac muscle tissue morphogenesis	GO:0055009	2/5	5.66E-03
atrial cardiac muscle tissue development	GO:0003228	2/5	5.66E-03
sequestering of extracellular ligand from receptor	GO:0035581	2/7	1.11E-02
determination of bilateral symmetry	GO:0009855	4/95	1.14E-02
specification of symmetry	GO:0009799	4/96	1.19E-02
membrane organization	GO:0061024	7/468	3.20E-02

Biological process	GO #	# of Genes	p-value
extracellular negative regulation of signal transduction	GO:1900116	2/12	3.22E-02
extracellular regulation of signal transduction	GO:1900115	2/12	3.22E-02
in utero embryonic development	GO:0001701	6/396	7.36E-02

Table 7.7: Biological processes of genes listed in Cluster VII

Biological process	GO #	# of Genes	p-value
fibroblast growth factor receptor signaling pathway	GO:0008543	4/40	2.05E-04
response to fibroblast growth factor	GO:0071774	4/51	5.30E-04
cellular response to fibroblast growth factor stimulus	GO:0044344	4/51	5.30E-04
post-anal tail morphogenesis	GO:0036342	3/23	1.55E-03
enzyme linked receptor protein signaling pathway	GO:0007167	7/430	7.17E-03
regulation of heart induction	GO:0090381	2/7	7.32E-03
positive regulation of catenin import into nucleus	GO:0035413	2/11	1.80E-02
cardiac cell fate commitment	GO:0060911	2/11	1.80E-02
regulation of transcription from RNA polymerase II promoter involved in heart development	GO:1901213	2/13	2.50E-02
mesoderm formation	GO:0001707	3/60	2.57E-02
BMP signaling pathway	GO:0030509	3/61	2.69E-02
mesodermal cell fate commitment	GO:0001710	2/14	2.89E-02
stem cell differentiation	GO:0048863	5/257	2.92E-02
mesoderm morphogenesis	GO:0048332	3/63	2.95E-02
cell surface receptor signaling pathway involved in heart development	GO:0061311	2/18	4.75E-02
mesodermal cell differentiation	GO:0048333	2/19	5.28E-02
regulation of cell fate specification	GO:0042659	2/19	5.28E-02
organ induction	GO:0001759	2/20	5.84E-02
formation of primary germ layer	GO:0001704	3/86	7.18E-02
regulation of catenin import into nucleus	GO:0035412	2/23	7.68E-02

Biological process	GO #	# of Genes	p-value
positive regulation of nucleocytoplasmic transport	GO:0046824	3/91	8.43E-02
cell fate commitment involved in formation of primary germ layer	GO:0060795	2/25	9.04E-02
stem cell development	GO:0048864	4/204	9.94E-02

Table 7.8: Biological processes of genes listed in Cluster VIII

Biological process	GO #	# of Genes	p-value
neutrophil chemotaxis	GO:0030593	4/39	4.84E-04
neutrophil migration	GO:1990266	4/40	5.34E-04
positive regulation of phagocytosis	GO:0050766	4/42	6.46E-04
granulocyte chemotaxis	GO:0071621	4/46	9.20E-04
granulocyte migration	GO:0097530	4/48	1.09E-03
regulation of phagocytosis	GO:0050764	4/57	2.11E-03
cell chemotaxis	GO:0060326	5/123	2.90E-03
antigen processing and presentation of peptide antigen	GO:0048002	4/64	3.30E-03
myeloid leukocyte migration	GO:0097529	4/66	3.71E-03
positive regulation of dendritic cell antigen processing and presentation	GO:0002606	2/5	7.01E-03
regulation of endocytosis	GO:0030100	5/155	8.52E-03
leukocyte chemotaxis	GO:0030595	4/83	8.91E-03
regulation of dendritic cell antigen processing and presentation	GO:0002604	2/6	1.01E-02
positive regulation of endocytosis	GO:0045807	4/91	1.26E-02
antigen processing and presentation	GO:0019882	4/94	1.43E-02
eosinophil chemotaxis	GO:0048245	2/8	1.78E-02
multicellular organismal iron ion homeostasis	GO:0060586	2/8	1.78E-02
eosinophil migration	GO:0072677	2/9	2.25E-02
immune system process	GO:0002376	13/1389	2.59E-02
positive regulation of immune system process	GO:0002684	8/563	2.60E-02

Biological process	GO #	# of Genes	p-value
positive regulation of myeloid leukocyte cytokine production involved in immune response	GO:0061081	2/10	2.78E-02
inflammatory response	GO:0006954	6/317	3.20E-02
positive regulation of tumor necrosis factor production	GO:0032760	3/50	3.30E-02
positive regulation of antigen processing and presentation	GO:0002579	2/11	3.35E-02
positive regulation of tumor necrosis factor superfamily cytokine production	GO:1903557	3/51	3.50E-02
macrophage chemotaxis	GO:0048246	2/12	3.98E-02
immune response	GO:0006955	8/610	4.38E-02
regulation of mononuclear cell migration	GO:0071675	2/13	4.66E-02
leukocyte migration	GO:0050900	4/133	5.21E-02
defense response	GO:0006952	9/785	5.27E-02
antigen processing and presentation of exogenous peptide antigen via MHC class II	GO:0019886	2/14	5.39E-02
regulation of antigen processing and presentation	GO:0002577	2/15	6.18E-02
phagocytosis, engulfment	GO:0006911	2/15	6.18E-02
antigen processing and presentation of peptide antigen via MHC class II	GO:0002495	2/16	7.02E-02
antigen processing and presentation of peptide or polysaccharide antigen via MHC class II	GO:0002504	2/18	8.84E-02

Table 7.9: Biological processes of genes listed in Cluster IX

Biological process	GO #	# of Genes	p-value
regulation of morphogenesis of a branching structure	GO:0060688	5/58	1.62E-03
regulation of organ morphogenesis	GO:2000027	6/151	1.67E-02
metanephric collecting duct development	GO:0072205	2/4	2.07E-02
embryonic digit morphogenesis	GO:0042733	4/60	3.07E-02
nephric duct formation	GO:0072179	2/5	3.22E-02

Biological process	GO #	# of Genes	p-value
negative regulation of metanephros development	GO:0072217	2/7	6.26E-02
mesenchymal cell differentiation involved in kidney development	GO:0072161	2/7	6.26E-02
mesenchymal cell differentiation involved in renal system development	GO:2001012	2/7	6.26E-02
regulation of branching involved in prostate gland morphogenesis	GO:0060687	2/7	6.26E-02
embryonic limb morphogenesis	GO:0030326	5/128	6.29E-02
embryonic appendage morphogenesis	GO:0035113	5/128	6.29E-02
nephric duct morphogenesis	GO:0072178	2/8	8.15E-02
nephric duct development	GO:0072176	2/8	8.15E-02
collecting duct development	GO:0072044	2/8	8.15E-02

Table 7.10: Biological processes of genes listed in Cluster X

Biological process	GO #	# of Genes	p-value
blood vessel development	GO:0001568	21/420	6.05E-10
vasculature development	GO:0001944	21/452	2.34E-09
blood vessel morphogenesis	GO:0048514	18/352	1.69E-08
multicellular organismal development	GO:0007275	60/3829	5.04E-08
porphyrin-containing compound metabolic process	GO:0006778	8/37	7.13E-08
tetrapyrrole metabolic process	GO:0033013	8/42	1.91E-07
anatomical structure morphogenesis	GO:0009653	38/1835	2.19E-07
circulatory system development	GO:0072359	23/715	2.92E-07
cardiovascular system development	GO:0072358	23/715	2.92E-07
single-organism developmental process	GO:0044767	63/4327	2.92E-07
system development	GO:0048731	53/3278	3.27E-07
developmental process	GO:0032502	63/4339	3.27E-07
anatomical structure formation involved in morphogenesis	GO:0048646	24/815	6.85E-07
anatomical structure development	GO:0048856	57/3847	1.52E-06
organ development	GO:0048513	42/2423	4.41E-06
regulation of vasculature development	GO:1901342	12/198	4.42E-06

Biological process	GO #	# of Genes	p-value
porphyrin-containing compound biosynthetic process	GO:0006779	6/25	6.63E-06
tetrapyrrole biosynthetic process	GO:0033014	6/25	6.63E-06
angiogenesis	GO:0001525	13/261	1.09E-05
heme metabolic process	GO:0042168	6/29	1.58E-05
cofactor biosynthetic process	GO:0051188	9/106	1.68E-05
single-multicellular organism process	GO:0044707	71/5830	3.27E-05
single-organism process	GO:0044699	114/11868	6.59E-05
positive regulation of cellular process	GO:0048522	53/3833	6.79E-05
multicellular organismal process	GO:0032501	71/6004	1.10E-04
positive regulation of biological process	GO:0048518	56/4248	1.30E-04
regulation of angiogenesis	GO:0045765	10/183	1.74E-04
cofactor metabolic process	GO:0051186	11/238	2.46E-04
pigment metabolic process	GO:0042440	6/50	3.68E-04
tube development	GO:0035295	16/546	3.70E-04
cell adhesion	GO:0007155	18/685	3.74E-04
regulation of locomotion	GO:0040012	17/615	3.78E-04
cellular component organization	GO:0016043	50/3712	3.81E-04
biological adhesion	GO:0022610	18/693	4.39E-04
aromatic compound biosynthetic process	GO:0019438	35/2168	4.72E-04
cellular nitrogen compound biosynthetic process	GO:0044271	35/2206	6.97E-04
regulation of biological quality	GO:0065008	36/2326	8.55E-04
tissue morphogenesis	GO:0048729	15/516	8.71E-04
hematopoietic or lymphoid organ development	GO:0048534	16/587	9.22E-04
immune system process	GO:0002376	26/1389	9.49E-04
cellular component organization or biogenesis	GO:0071840	50/3832	9.64E-04
cellular developmental process	GO:0048869	42/2964	1.02E-03
platelet aggregation	GO:0070527	5/35	1.20E-03
heterocycle biosynthetic process	GO:0018130	34/2162	1.21E-03
nephron development	GO:0072006	6/62	1.25E-03
organic cyclic compound biosynthetic process	GO:1901362	35/2266	1.26E-03
regulation of cell migration	GO:0030334	15/533	1.28E-03
morphogenesis of an epithelium	GO:0002009	13/402	1.28E-03

Biological process	GO #	# of Genes	p-value
tissue development	GO:0009888	26/1415	1.32E-03
hemopoiesis	GO:0030097	15/539	1.46E-03
immune system development	GO:0002520	16/612	1.55E-03
myeloid cell differentiation	GO:0030099	9/186	1.65E-03
morphogenesis of a branching epithelium	GO:0061138	9/186	1.65E-03
cell differentiation	GO:0030154	40/2822	1.84E-03
regulation of cell motility	GO:2000145	15/556	2.10E-03
morphogenesis of a branching structure	GO:0001763	9/193	2.21E-03
epithelium development	GO:0060429	19/863	2.37E-03
positive regulation of cell migration	GO:0030335	11/306	2.60E-03
regulation of response to stimulus	GO:0048583	38/2653	2.61E-03
positive regulation of metabolic process	GO:0009893	38/2659	2.75E-03
sprouting angiogenesis	GO:0002040	5/42	2.86E-03
heme biosynthetic process	GO:0006783	4/20	2.92E-03
regulation of cellular response to growth factor stimulus	GO:0090287	8/152	2.94E-03
single organismal cell-cell adhesion	GO:0016337	9/201	3.03E-03
positive regulation of angiogenesis	GO:0045766	7/110	3.10E-03
positive regulation of cell motility	GO:2000147	11/313	3.20E-03
regulation of cell proliferation	GO:0042127	23/1225	3.45E-03
regulation of leukocyte migration	GO:0002685	7/115	4.10E-03
branching morphogenesis of an epithelial tube	GO:0048754	8/160	4.23E-03
metanephric nephron development	GO:0072210	4/22	4.23E-03
positive regulation of cellular component movement	GO:0051272	11/324	4.38E-03
positive regulation of gene expression	GO:0010628	26/1521	4.62E-03
positive regulation of locomotion	GO:0040017	11/329	5.04E-03
muscle structure development	GO:0061061	12/396	5.56E-03
cell morphogenesis involved in differentiation	GO:0000904	13/463	5.59E-03
regulation of glomerular mesangial cell proliferation	GO:0072124	3/8	5.82E-03
membrane organization	GO:0061024	13/468	6.24E-03
positive regulation of leukocyte migration	GO:0002687	6/84	6.80E-03
regulation of immune system process	GO:0002682	18/854	7.22E-03

Supplementary

Biological process	GO #	# of Genes	p-value
positive regulation of cellular metabolic process	GO:0031325	35/2464	7.53E-03
regulation of cellular component movement	GO:0051270	15/622	7.62E-03
organonitrogen compound biosynthetic process	GO:1901566	12/410	7.76E-03
regulation of body fluid levels	GO:0050878	9/228	8.04E-03
single organism cell adhesion	GO:0098602	9/229	8.31E-03
homotypic cell-cell adhesion	GO:0034109	5/53	8.62E-03
single-organism membrane organization	GO:0044802	12/415	8.71E-03