

**TECHNISCHE UNIVERSITÄT MÜNCHEN**

Lehrstuhl für Experimentelle Genetik

# **Novel mechanisms of endoderm and mesoderm formation**

**Dapeng Yang**

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. M. Hrabě de Angelis

Prüfer der Dissertation:

1.apl. Prof. Dr. J. Beckers

2.Univ.-Prof.Dr. H. Lickert

Die Dissertation wurde am 03.02.2015 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 27.03.2015 angenommen.



---

List of figures .....	iv
List of tables .....	vi
Abstract .....	1
1. Introduction.....	3
1.1 Early embryonic development in the mouse .....	3
1.2 Gastrulation and the formation of three primary germ layers .....	4
1.3 Molecular control of endoderm and mesoderm formation .....	7
1.3.1 Signaling pathways that regulate gastrulation .....	7
1.3.2 miRNAs biogenesis and function in development .....	12
1.3.3 Important transcription factors (TFs) regulating development .....	16
1.3.4 Epithelial-to-mesenchymal transition (EMT) during gastrulation.....	19
1.4 Maintenance and differentiation of mESCs <i>in vitro</i> .....	21
1.4.1 Maintenance of mESCs pluripotency.....	21
1.4.2 Differentiation of mESCs to recapitulate gastrulation .....	23
1.5 Aims of the study .....	25
2. Results.....	27
2.1 Single cell continuous lineage analysis reveals a novel mechanism of endoderm and mesoderm formation .....	27
2.1.1 Generation of a dual knock-in reporter mESC line .....	28
2.1.2 Establishing a mESCs-derived endoderm and mesoderm differentiation system .....	30
2.1.3 Isolation of mESCs-derived endodermal and mesodermal progenitors.....	32
2.1.4 CD-24 is a valuable marker to separate early and late endodermal/ mesodermal progenitors.....	34
2.1.5 Distinct <i>in vitro</i> -derived populations reflect <i>in vivo</i> lineage progenitors.....	37
2.1.6 Delineating the mesendoderm lineage hierachy in culture .....	40
2.1.7 ADE and PDE are derived from distinct progenitor cells .....	45
2.1.8 The Brachyury (T) function is not required for endoderm formation .....	49

2.1.9 Key regulators of classical EMT are not involved in endoderm formation	.51
2.2 miR-335 promotes mesendodermal lineage segregation and shapes a TF gradient in the endoderm	55
2.2.1 Identification of miRNAs regulating mESCs differentiation	56
2.2.2 Spatio-temporal expression of miR-335 during early development	60
2.2.3 miR-335 directly targets <i>Foxa2</i> and <i>Sox17</i>	63
2.2.4 miR-335 overexpression represses endoderm formation	64
2.2.5 Knock-down of miR-335 leads to an increase of endoderm differentiation	67
2.2.6 Mathematical modeling predicts miR-335 function in TF gradient formation	69
2.2.7 miR-335 shapes a TF gradient in the endoderm	71
3. Discussion	75
3.1 Summary of results	75
3.2 <i>In vitro</i> mESCs differentiation system: A valuable tool to investigate mouse gastrulation	79
3.2.1 Establishment of stepwise mESCs differentiation towards endoderm and mesoderm	79
3.2.2 The dual knock-in reporter mESCs line: <i>T<sup>GFP/+</sup>; Foxa2<sup>tagRFP/+</sup></i> , a useful tool to investigate lineage specification during gastrulation	81
3.3 Novel mechanisms of endoderm and mesoderm formation	84
3.3.1 Mesendoderm lineage segregation and cell fate decisions	84
3.3.2 The formation of endoderm does not require EMT	87
3.3.3 miR-335 regulates endoderm TF gradients	88
4. Materials and methods	93
4.1 Materials	93
4.1.1 Instruments	93
4.1.2 Consumables	95
4.1.3 Kits	95
4.1.4 Software	96

---

4.1.5 Chemicals.....	96
4.1.6 Cell culture reagents.....	98
4.1.7 Buffer and solutions.....	98
4.1.8 Antibodies.....	103
4.2 Methods .....	105
4.2.1 Molecular biology.....	105
4.2.2 Cell culture .....	124
4.2.3 Embryology .....	127
4.2.4 Data processing and analysis.....	129
5.References .....	135
6. Abbreviations.....	159
7. Acknowledgements .....	159
8. Curriculum Vitae .....	161

## List of figures

1.1 Embryonic development of the mouse embryo before gastrulation.....	4
1.2 Foxa2 and T mark the progenitors of endoderm and mesoderm.....	6
1.3 Signals gradients are formed in gastrulation in the mice .....	9
1.4 Biogenesis of canonical miRNAs.....	13
2.1 Generation of dual knock-in reporter T <sup>GFP/+</sup> ; Foxa2 <sup>tagRFP/+</sup> mESCs line .....	29
2.2 Stepwise mESCs differentiatio towards to endoderm and mesoderm.....	31
2.3 Isolation of mESCs-derived endodermal and mesodermal progenitors.....	33
2.4 CD-24 is a valuable marker to show differentiation progress .....	35
2.5 Characterization of mESCs-derived endodermal and mesodermal progenitor .....	39
2.6 <i>In vitro</i> derived populations correlate with the embryonic cell populations formed during gastrulation.....	40
2.7 Single cell time-lapse imaging of endoderm differentiation <i>in vitro</i> .....	42
2.8 Analysis of <i>in vitro</i> mesoderm differentiation dynamics .....	44
2.9 Graphical models of <i>in vitro</i> endoderm and mesoderm formation .....	44
2.10 Cell fate analyses of mESCs–derived early endoderm progenitor cells .....	47
2.11 T function is not required in the endoderm formation .....	50
2.12 Analysis of EMT associated genes in endoderm and mesoderm progenitors .....	53
2.13 Analysis of mRNA and microRNA profiling during endoderm differentiation .....	58
2.14 Identification of miR-335 as a potential regulator of early development .....	59
2.15 Co-expression of Mest/miR-335 in mESCs and mouse embryos.....	61
2.16 Foxa2 and Sox17 are targets of miR-335.....	62
2.17 Comparison of embryonic and overexpressed miR-335 levels.....	64

---

2.18 miR-335 overexpression represses endoderm differentiation .....	66
2.19 Loss of miR-335 leads to the increase in endoderm formation .....	68
2.20 Modeling miRNA mediated protein expression dynamics.....	70
2.21 miR-335 shapes transcription factor gradients in the endoderm .....	73
3.1 Graphical representation of the hypothetical <i>in vitro</i> differentiation model....	76
3.2 Graphical representation of lineage specification during gastrulation.....	78
4.1 Western blot setup.....	122
4.2 Preprocessing of IHC images .....	131
4.3 Foxa2 and Sox17 gradients estimation .....	133
4.4 Gradient analysis.....	134

## List of tables

1.1The schematic depicts the interact of ligand-receptor-Smad in Nodal/Activin and Bmp4 signaling .....	9
2.1List of pluripotency genes used in the multi-class analysis.....	36
2.2List of epiblast genes used in the multi-class analysis.....	36
2.3Number of mRNA and miRNA differentially expressed during mesendoderm and endoderm differentiation .....	60
2.4Target Sequences .....	64
4.1Primary antibody.....	103
4.2Secondary antibody.....	104
4.3Primer pairs for RT-PCR. ....	111
4.4Primer pairs for qPCR assays with SYBR® Green system. ....	112
4.5Genes and assay IDs used in qPCR assays.....	113
4.6Gene lists and assay IDs in the array card .....	114
4.7Primers for 3'UTR cloning .....	120
4.8Primers for cloning and sequencing .....	121



---

## Abstract

The formation of endoderm and mesoderm during gastrulation is well studied; however, the gastrulation lineage tree is still not solved due to the lack of continuous live cell analysis. For this purpose, a dual knock-in  $T^{GFP/+}$ ;  $Foxa2^{tagRFP/+}$  fluorescent reporter mouse embryonic stem cell (mESC) line was generated. Further, an effective strategy for the stepwise differentiation of mESCs into mesendoderm, as well as endoderm and mesoderm lineages by triggering Wnt/ $\beta$ -catenin, Nodal/ActA and Bmp4 signaling was established.

Three distinct mesendoderm lineages are formed from pluripotent mESCs during differentiation. In contrast to the common belief, we demonstrated that anterior definitive endoderm (ADE) is derived from  $Foxa2^+$  progenitors and not from  $T^+$  progenitors, which instead give rise to mesoderm, axial mesoderm and posterior definitive endoderm (PDE). We further uncovered that definitive endoderm (DE) does not require Brachyury(T) transcription factor (TF) function for development.

Using this *in vitro* system and an additional epithelial surface marker CD24, we isolated and characterized the early and late endodermal and mesodermal progenitors by time-resolved molecular profiling. We confirmed that the process of mesoderm formation is a typical epithelial-to-mesenchymal transition (EMT) due to the up-regulation of EMT related key TFs and down-regulation of epithelial markers in the  $T^+$  lineage. In contrast, there was no up-regulation of EMT TFs in the  $Foxa2^+$  lineage. This reveals that the endoderm is formed by a novel process that does not need activation of key regulators of classical EMT, and which we named epithelial-to-epithelial transition (EET).

Finally, we identified that miR-335 fine-tunes TF gradients in the endoderm and promotes mesendodermal lineage segregation. The intronic miR-335 highly accumulates in the mesoderm, but is only transiently expressed in endoderm progenitors. miR-335 overexpression does not affect initial mesendoderm induction, but blocks endoderm differentiation. Conversely, inhibition of miR-335 activity induces  $Foxa2$  and  $Sox17$  protein accumulation and endoderm formation *in vitro*, and shapes TF gradients in the embryonic endoderm.

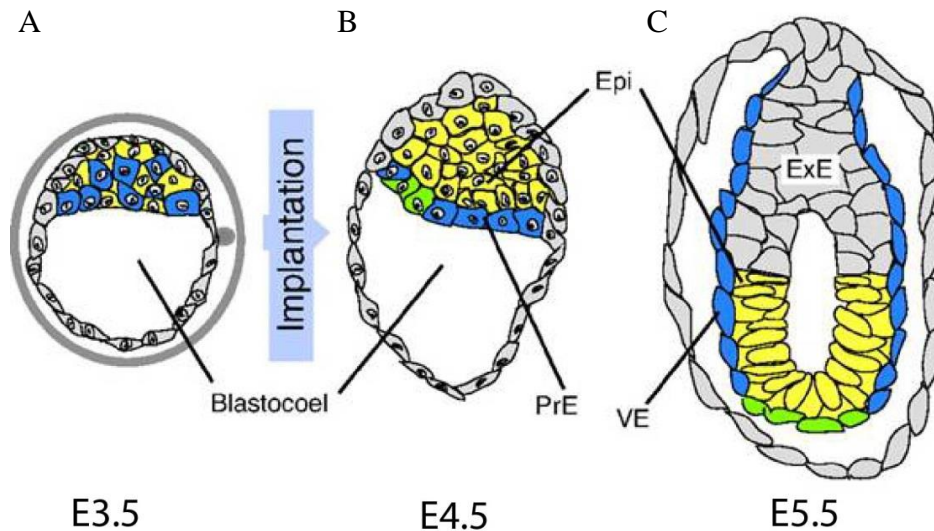
Taken together, these results lead to a better understanding of mesendoderm lineage specification and provide a more efficient way to differentiate functional cells for future tissue replacement.



# 1. Introduction

## 1.1 Early embryonic development in the mouse

Murine development starts from fertilization, the process when a sperm and oocyte fuse together to become a zygote. The fertilized egg subsequently undergoes a cleavage process to form the morula, a multicellular structure with numerous cells. The earliest stage of lineage specification occurs when the morula develops to the blastocyst stage at embryonic day (E) 2.5-3.5 (Rossant and Tam, 2009)(Figure 1.1A). The blastocyst consists of an inner cell mass (ICM) and an outer surrounding epithelial cell layer named trophectoderm (TE), which supports the ICM (Eckert and Fleming, 2008; Tam and Rossant, 2003). In mammalian embryos, the ICM is the source of embryonic stem cells (ESCs) (Evans and Kaufman, 1981; Martin, 1981), which show high expression of pluripotency genes *Oct3/4*, *Sox2*, and *Nanog*, and can give rise to all embryonic tissues (Tang et al., 2010). By E4.5, two cell types including epiblast and primitive endoderm (PrE) appear in the ICM (Figure 1.1B). The pluripotent epiblast cells generate the embryo proper, as well as the yolk sac, allantois, and amnion; whereas the cells of the PrE give rise to two structures: visceral endoderm (VE) and parietal endoderm. The cells from the surrounding TE will form the extra-embryonic ectoderm (ExE) and the ectoplacental cone, progenitor cells of the placenta (Arnold and Robertson, 2009; Gadue et al., 2005). After segregation of the extra-embryonic and embryonic tissues, the embryo undergoes rapid proliferation and differentiates to form the egg cylinder (Figure 1.1C). At this stage, the epiblast joint with the ExE is surrounded by the VE. The VE and epiblast/ExE neighbouring tissue interactions initiate the formation of the proximal-distal (P-D) and anterior-posterior (A-P) axis. From that time on the mouse embryo changes dramatically in size and shape, and the process of gastrulation is initiated.



(Figure was modified from Takaoka and Hamada, 2012)

### Figure 1.1 Embryonic development of the mouse embryo before gastrulation

The morphological changes and cell fate specification events are taking place from blastocyst stage to pre-primitive streak stage. The epiblast is derived from ICM (A and B) during blastocyst stage. At E5.5, it joints with the ExE, and is surrounded by the VE(C).

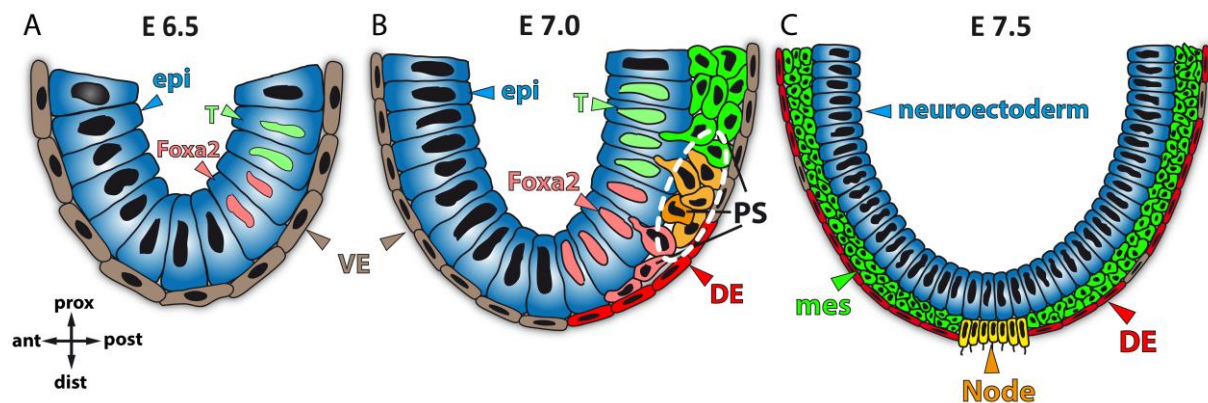
### 1.2 Gastrulation and the formation of three primary germ layers

During gastrulation, the mouse embryo establishes three primary germ layers: ectoderm, mesoderm, and endoderm. In mice, gastrulation is initiated in the epiblast, which is in close contact with the ExE, and surrounded by a single-layered VE (Figure 1.2 A). The epiblast cells at this stage still keep pluripotency, and are the source of epiblast stem cells (EpiSCs). These EpiSCs express high levels of Oct4 and Nanog and are capable to differentiate into the three germ layers both *in vivo* and *in vitro* (Brons et al., 2007). The formation of the primitive streak (PS) is one of the hallmarks of gastrulation. The PS has been shown to contain precursor cells of mesodermal and endodermal lineages that contribute to different germ layers of the embryo. Fate mapping at the early-streak stage of different PS regions shows that, the first mobilized epiblast cells migrate through the posterior PS and contribute to the extra-embryonic mesoderm (Kinder et al., 1999). As gastrulation proceeds, the

cardiac and cranial mesodermal cells are derived from middle and more anterior parts of the PS regions, followed by the formation of paraxial and axial mesoderm within the same region (Kinder et al., 1999, 2001). Whereas, the definitive endoderm (DE) cells are derived from the anterior end of the PS (Murry and Keller, 2008). In contrast, the ectoderm cells are formed directly from the remaining anterior epiblast cells.

DE cells migrate through the PS, intercalate into the overlying VE, and migrate from the posterior part of the embryo to the anterior part to form the anterior definitive endoderm (ADE). In contrast, the endoderm cells which exit the PS later contribute to the posterior definitive endoderm (PDE). Thereby, the ADE is by definition older and earlier generated than the PDE lineage (Yang et al., 2014a). During the process of migration, the endoderm is patterned along the A-P axis with high levels of the transcription factors (TFs) *Foxa2* and *Sox17* accumulating in the ADE (Burtscher et al., 2012). At the end of gastrulation (E7.5), all DE cells are recruited and form an epithelial sheet of approximately 500 cells on the outside of the mouse embryo (Wells and Melton, 1999).

The DE and mesoderm cells are derived from epiblast cells ingression through different regions of PS, indicating that the cell fate is specified before the PS forms. Fate maps and whole-mount immunohistochemistry (IHC) analysis of the early streak epiblast show that the cell fate is committed in the posterior-distal region of the epiblast, where it contains progenitors of the mesodermal and endodermal lineages (Burtscher and Lickert, 2009; Lawson et al., 1991). The T-box transcription factor *Brachyury* (T) and the winged helix/forkhead box TF 2 (*Foxa2*) have been shown to mark these progenitors within distinct regions of the epiblast (Figure 1.2) (Burtscher and Lickert, 2009). T and *Foxa2* proteins are synthesized in two mutually exclusive cell populations in the posterior epiblast in the pre-streak stage embryo (Figure 1.2A). At mid-streak stage, these *Foxa2*<sup>+</sup> and T<sup>+</sup> populations migrate into PS and form three distinct populations (*Foxa2*<sup>+</sup>, T<sup>+</sup> and *Foxa2*<sup>+</sup>T<sup>+</sup>; Figure 1.2 B). As gastrulation proceeds, these cells migrate out of the PS and form three distinct populations including T<sup>+</sup> mesodermal cells, *Foxa2*<sup>+</sup>T<sup>+</sup> axial mesoderm, and *Foxa2*<sup>+</sup> endodermal cells (Figure 1.2 C).



(Figure modified from Biospectrum 2011)

**Figure 1.2 Foxa2 and T mark the progenitors of endoderm and mesoderm**

- (A) T and Foxa2 are synthesized in two distinct cell populations in the posterior epiblast in the early-streak stage mouse embryo at E6.5.
- (B) After the formation of the PS, three cell populations can be distinguished: T<sup>+</sup> mesoderm progenitor cells (green) in the posterior PS, Foxa2<sup>+</sup>T<sup>+</sup> axial mesoderm progenitor cells (yellow) in the APS, and Foxa2<sup>+</sup> DE progenitor cells (red) in anterior end of the PS region.
- (C) After gastrulation is completed, endoderm (Foxa2<sup>+</sup>), mesoderm (T<sup>+</sup>), and node/notochord (Foxa2<sup>+</sup>T<sup>+</sup>) are formed.

### **1.3 Molecular control of endoderm and mesoderm formation**

Formation of endoderm and mesoderm is a very important event in vertebrate development, giving rise to many essential features of the body (Wilkinson, Bhatt, 1990; Zorn and Wells, 2009). In mammals, DE cells contribute to the lining of the digestive and respiratory tubes and their associated organs, such as the pancreas, lung, thyroid, liver, thymus, and biliary system. These organs provide many essential functions including glucose homeostasis, gas exchange, detoxification, digestion, nutrient absorption, and blood clotting (Zorn and Wells, 2009). In contrast, mesoderm generates organs between endoderm tissues and ectoderm wall, such as notochord, somites, heart, blood vessels, and urogenital system organs including kidneys and the gonads (Gilbert, *Developmental biology* 9<sup>th</sup>), which are important for development. A better understanding of mechanisms which are involved in endoderm and mesoderm formation could provide fundamental insights in controlling organ development.

The temporal and spatial segregations of progenitors of endoderm and mesoderm that occur in gastrulation are not random, but are mainly regulated by the signal gradients produced by the embryo itself. The cell populations in different epiblast and PS regions acquire different signaling environments and therefore contribute to different fates. Many studies have shown that members of the TGF- $\beta$  superfamily including Nodal and Bmp4 and the members of Wnt family, such as Wnt3 and Wnt3a, as well as their inhibitors, TFs, miRNAs, are responsible for the induction and specification of lineages (Liu et al., 1999; Mallanna and Rizzino, 2010; Schier, 2003; Vincent et al., 2003; Winnier et al., 1995; Yanagisawa, 1990).

#### **1.3.1 Signaling pathways that regulate gastrulation**

In the early stage of mouse embryonic development, the anterior VE (AVE), the ExE, the posterior epiblast, and the node are important sources of signals including Nodal, Wnt3, and Bmp4 (Figure 1.3) (Conlon et al., 1994; Liu et al., 1999; Varlet et al., 1997). Nodal and Wnt3 are essential for the formation of PS, endoderm, and mesoderm. Mouse embryos lacking Nodal or Wnt3 fail to form the PS and its derivatives (Conlon

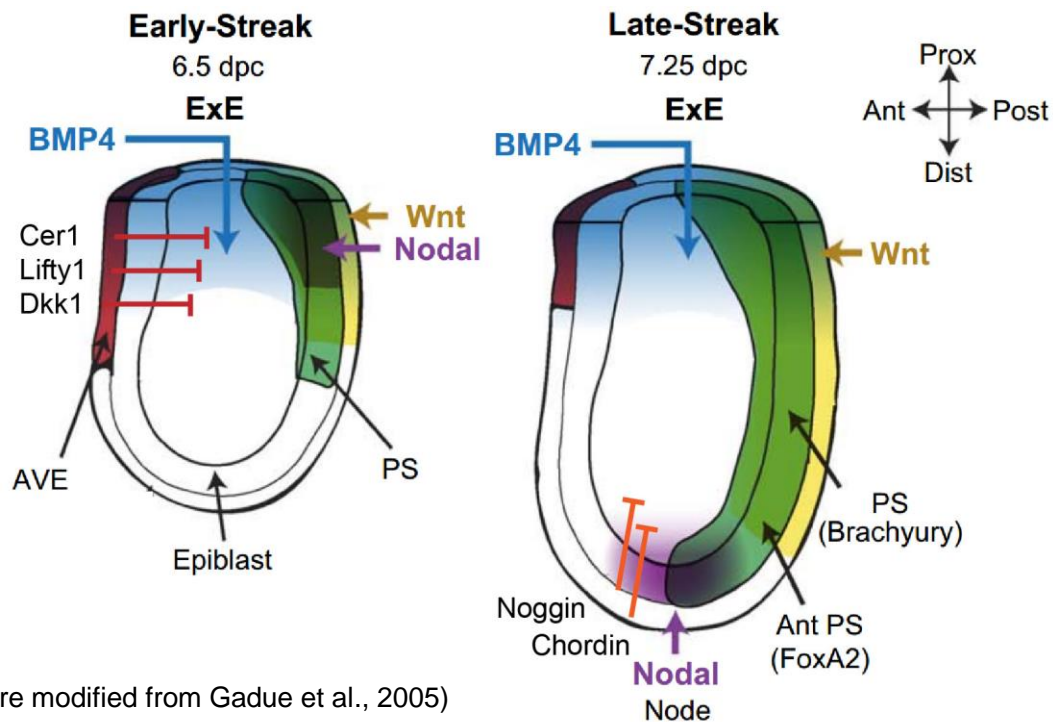
et al., 1994; Liu et al., 1999). Bmp4, another signal produced by the ExE, has been shown to induce expression of PS and mesoderm markers (Beppu et al., 2000; Ohinata et al., 2009; Winnier et al., 1995).

In parallel, the Nodal antagonists including Cerberus-like protein 1 (Cer-1) and left-right determination factor 1 (Lefty1) as well as the Wnt antagonist Dickkopf homologue 1 (Dkk1) are produced by the AVE (Belo et al., 2000; Glinka et al., 1998; Perea-Gomez et al., 2002; Thomas and Beddington, 1996). These inhibitors attenuate the influence of Nodal and Wnt signaling in the anterior part of the embryo, and restrict signaling activity to the posterior region. Embryos lacking Nodal antagonists show expansion of the APS and its derivatives or leads to the formation of ectopic PS on the anterior side (Perea-Gomez et al., 2002). This indicates that the balance between the activity of Nodal/Wnts and their antagonists is necessary for the AP patterning of the epiblast (Figure 1.3A). In contrast, the Bmp antagonists Chordin (chrd) and Noggin are expressed in the node (Bachiller et al., 2000; Klingensmith et al., 1999). The gradients formed by Bmp4 and its antagonists along the P-D region of the embryos are very important for mesoderm and endoderm germ layer specification (Figure 1.3B). Thereby, analyses of the signaling components that shape different domains are necessary to study the mechanisms involved in patterning and lineage specification.

### **1.3.1.1 Nodal/Activin A**

Nodal is a member of the transforming growth factor-beta (TGF- $\beta$ ) family of growth factor. It is essential for gastrulation and germ layer formation. Nodal binds Activin type I (Alk4 or Alk7) and type II (ActRIIA or ActRIIB) transmembrane serine/threonine kinase receptors and an EGF-CFC (epidermal growth factor-Cripto-FRL1-Cryptic) co-receptor to form an active signaling complex. This complex leads to the phosphorylation of Smad2/3, which subsequently binds to Smad4. The Smad2/3-Smad4 complex undergoes nuclear translocation and regulates target gene expression (Schier, 2003). Activin-A (ActA), another member of TGF- $\beta$  family, initiates signaling through the same receptors and shares a common downstream signaling cascade with Nodal (Table 1.1).





**Figure 1.3 Signaling gradients are formed during mouse gastrulation**

During gastrulation, signals including Nodal, Wnt and Bmp4 are produced from the regions of AVE, ExE, PS, and the node. The expression of Nodal and Wnt inhibitors (Cer1, Lefty1 and Dkk1) in the AVE results in the restriction of signaling activity to the posterior side of the embryo; whereas the expression of Bmp4 antagonists (Noggin and Chordin) in the node restricts its signaling activity to the proximal region of the embryo.

Ligands	Receptors		TFs
	Type II	Type I	
Nodal	ActRIIA and B	ALK4	Smad2/3, Smad4
Activin	ActRIIA and B	ALK4	Smad2/3, Smad4
Bmp4	BmpRII	Alk2, 3, and 6	Smad 1/ 5/ 8, Smad4

**Table 1.1 The schematic depicts the interact of ligand-receptor-Smad in Nodal/Activin and Bmp4 signaling**

Nodal is first detected in the embryonic ectoderm and PrE, and then at low level throughout the epiblast and VE at E5.5 (Varlet et al., 1997). At the early PS stage, *Nodal* mRNA expression is restricted to the posterior epiblast and the PS region (Varlet et al., 1997) (Figure 1.3 A). Nodal is only produced in the node at the late PS stage to regulate left-right (L-R) body axis (Collignon et al., 1996; Conlon et al., 1994) (Figure 1.3B). In mouse, Nodal is a very important mesendoderm inducer during gastrulation. Mouse embryos fail to gastrulate and exhibit defects in forming mesoderm and DE after implantation because of lack of Nodal or its downstream TFs Smad2/3 (Brennan et al., 2001; Dunn et al., 2004; Liu et al., 2004). In many species, such as *zebrafish* and mouse, the DE cells arise from Nodal-expressing cells, while the mesoderm cells are induced in the adjacent cells (Ben-Haim et al., 2006; Hagos and Dougan, 2007; Zorn and Wells, 2009). This indicates that patterning of anterior and posterior PS, as well as the subsequent DE and mesoderm formation relies on Nodal in a concentration-dependent manner.

### 1.3.1.2 Bmp4

Bone morphogenetic proteins (Bmps), members of the TGF- $\beta$  superfamily, are important signaling factors in mouse development. From the large gene family only Bmp2 and Bmp4 have been shown to have roles in embryonic development. Similar to Nodal, Bmp4 is also a member of the TGF- $\beta$  family and acts through the activation of the Smad TF complex. Instead of binding to type II receptors and leading to activation of Smad2/3, Bmp4 signals through the phosphorylation of Smad1/5/8. Then the activated Smad interacts with Smad4 to form an activated complex, entering the nucleus, and functioning as TFs (Feng and Derynck, 2005; Gadue et al., 2005; Kitisin et al., 2007)(Table 1.1).

Previous studies in *Xenopus*, *zebrafish* and mouse have shown that Bmp signals are required for gastrulation, especially for mesoderm formation (Beppu et al., 2000; Munoz-Sanjuan, 2002). In the mouse, Bmp4 is detected in the ExE before gastrulation. At E6.5, Bmp4 is detected in posterior PS with a relative low expression. The highest expression levels are found in the posterior PS, allantois, and amnion by E7.5 (Winnier et al., 1995). Most of homozygous null Bmp4 embryos die around gastrulation stage due to the loss of embryonic mesoderm. Surviving embryos

---

develop to the head fold or early somite stage, but are developmentally retarded and show disorganized posterior body structures (Lawson et al., 1999; Winnier et al., 1995). *Bmpr1a(Alk3)* mutant embryos fail to undergo gastrulation, and show a down-regulation of *Oct4*, *Nanog*, and *FGF5* expression accompanying by a neural differentiation in the epiblast, which indicates that Bmp4 is required to maintain pluripotency in the epiblast (Di-Gregorio et al., 2007).

### 1.3.1.3 Wnts

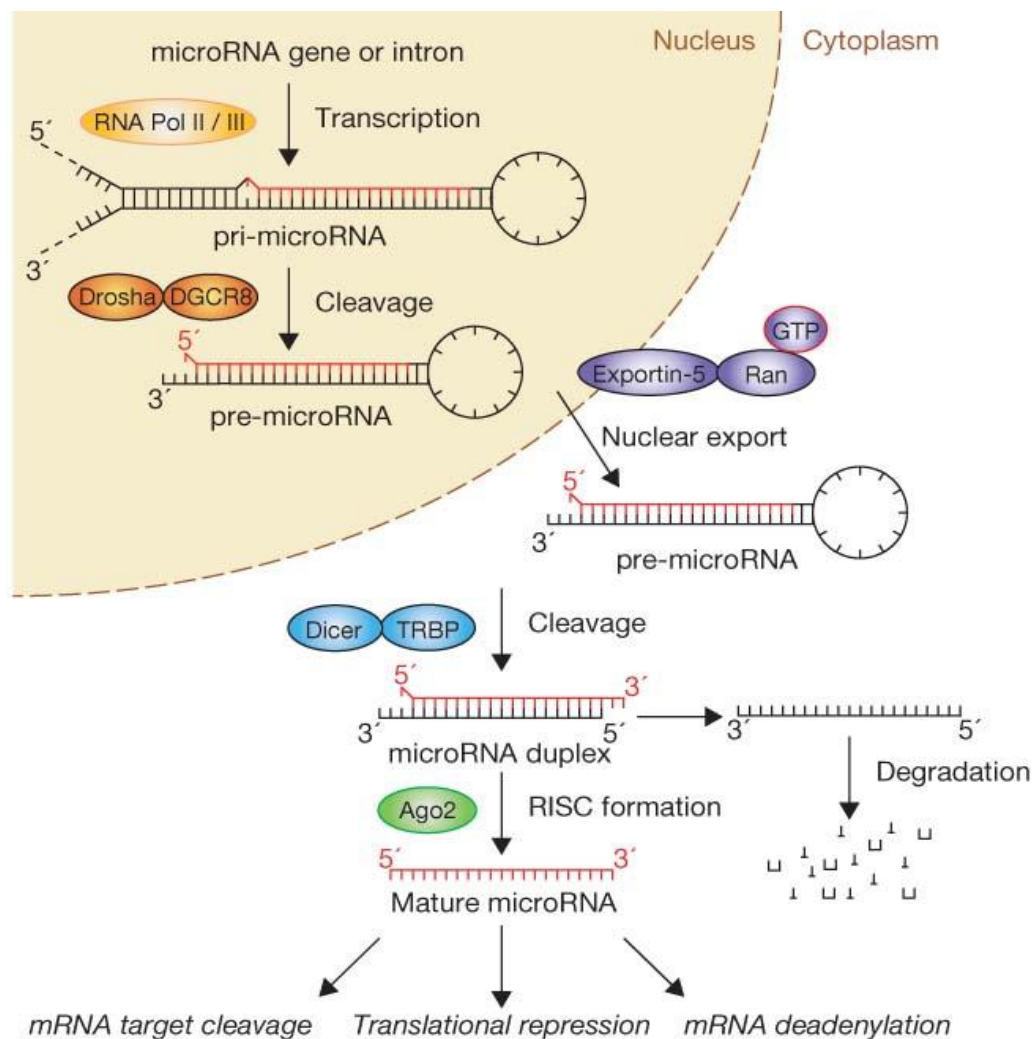
The Wnt family of growth factors are evolutionarily conserved signaling ligands critical for many developmental processes during embryogenesis, such as embryonic induction, gastrulation and axis development (Arkell et al., 2013; Herr et al., 2012; Logan and Nusse, 2004). In mammals, Wnt ligands act via the canonical Wnt signaling pathway and the non-canonical planar cell polarity (PCP) pathway. The canonical Wnt signaling pathway is activated through binding of the Wnt ligand to the Frizzled receptor/low density lipoprotein receptor-related protein (Lrp5/6) complex. The ligand-receptor interaction inhibits Glycogen synthase kinase 3 (GSK3)-mediated degradation pathway, and leads to the accumulation of  $\beta$ -catenin (Niehrs and Acebron, 2010). The accumulated  $\beta$ -catenin then translocates into the nucleus, where it interacts with the high mobility group box transcription factors of the T-cell factor (Tcf) and lymphoid enhancing factor (Lef) families to regulate the rate of transcription of a number of canonical Wnt target genes, such as the PS genes *T* and *Axin 2* (Jho et al., 2002; Lickert et al., 2002; Logan and Nusse, 2004; Moon, 2005; Seidensticker and Behrens, 2000; Tamai et al., 2000; Wray and Hartmann, 2012; Yamaguchi et al., 1999). Several Wnts (Wnt1, 2, 2b, 3, 3a, 6, 7b, 8a and 8b) have been shown to be involved in the canonical pathway; whereas other Wnts, such as Wnt5a and Wnt11, act through non-canonical pathway, which includes the activation of PCP pathway, the regulation of calcium flux and the activation of JNK and Src kinase (Kemp et al., 2005).

Genetic studies indicate the Wnt/ $\beta$ -catenin signaling pathway is a key factor of the molecular events that control the mesendoderm induction in mouse embryos (Tortelote et al., 2013). Multiple Wnt ligands (*Wnt2b*, *Wnt3*, *Wnt3a*, *Wnt5a*, *Wnt8a*, *Wnt11*) are expressed in the posterior embryo and PS domains (Kemp et al., 2005) (Figure 1.3B), and play roles in the process of formation of PS, mesoderm, and

endoderm (Lickert et al., 2002; Tam and Loebel, 2007). Deletion of the components of the pathway, such as ligand, co-receptor, inhibits PS formation (Biechele et al., 2011; Fu et al., 2009; Huelsken et al., 2000; Kelly et al., 2004; Liu et al., 1999), indicating that Wnt/ $\beta$ -catenin signaling is essential for this process. Analysis of *Wnt null* alleles indicates that *Wnt3* plays an important role in germ layer formation and expression of mesendoderm genes (Liu et al., 1999). In the absence of *Wnt3*, the embryos fail to maintain *Nodal* expression, and have defects in the formation of the anterior-posterior axis and induction of the mesoderm and endoderm germ layers (Ben-Haim et al., 2006; Liu et al., 1999).  $\beta$ -catenin null mutants also show that embryonic structures can not form properly due to the lack of mesendoderm induction (Haegel et al., 1995). During mesendoderm segregation,  $\beta$ -catenin is essential for endoderm induction and failure of Wnt/ $\beta$ -catenin signaling leads to ectopic mesoderm formation (Lickert et al., 2002).

### 1.3.2 miRNAs biogenesis and function in development

MicroRNAs (miRNAs) are 20~25 nucleotides single-stranded non-coding RNAs, which bind to the 3' untranslated region (UTR) of target mRNA transcripts, and lead to either degradation of the targeted mRNAs or inhibition of translation (Ambros, 2004; Bartel et al., 2004). The first miRNA molecules, *lin-4* and *lin-7*, were identified 20 years ago in regulating the developmental timing of *Caenorhabditis elegans* (*C. elegans*) (Lee et al., 1993; Reinhart et al., 2000; Wightman et al., 1993). *Lin-4* was found as a translational repressor of *lin-14*, as it is partially complementary to the 3' UTR of the *lin-14* gene (Lee et al., 1993). Since then, hundreds of miRNAs in animals, plants, and viruses have been identified by molecular cloning and bioinformatic approaches (Berezikov et al., 2006; Ruby et al., 2006). Predictions show 30% of human genes are regulated by miRNAs, and a single miRNA may have more than hundreds of mRNA targets involved in different processes in the cell (Rajewsky, 2006). Given that the miRNAs can regulate numerous mRNAs which are involved in many pathways, therefore, miRNAs are thought to be fine tuners of molecular programs.



(Adapted from Winter et al., 2009)

**Figure 1.4 Biogenesis of canonical miRNAs****1.3.2.1 miRNA biogenesis**

miRNAs are generated through a series of posttranscriptional biogenesis steps in both the nuclear and cytoplasmic compartments (Figure 1.3). First, long primary precursor miRNAs (pri-miRNA) are generated mostly by RNA polymerase II in the nucleus (Lee et al., 2004; Lin et al., 2003; Shomron and Levy, 2009). Then the long pri-miRNAs are further processed by Drosha-like nuclear RNase III and the dsRNA binding protein DiGeorge critical region 8 (DGCR8) to form shorter hairpin precursor miRNAs (pre-miRNAs), which are about 70 nucleotides in length. Some miRNAs, such as mirtons generated from introns, are processed by splicing to generate pre-miRNA (Kim and Kim, 2007). Pre-miRNAs are subsequently transported by Exportin-5 (Exp5) and Ran-GTP into the cytoplasm, where they are further cleaved by the RNase III enzyme Dicer to produce double-stranded 20~25 nt miRNAs duplex

intermediates with 2 nt-3' overhangs at both ends (Lund et al., 2004; Yi et al., 2003). Only one strand (guide strand) of the duplex incorporates into a ribonuclear particle and forms the RNA-induced gene silencing complex (RISC). The RISC complex can recognize and repress target mRNA expression; whereas the other passenger strand is released and degraded (Du and Zamore, 2005). The miRNA seed sequence, particularly nucleotides 2–7 from the 5' end of the miRNA, guides the complex to their targets through base-pair interactions (Lewis and Tam, 2006). The overall degree of the base-pair interaction complementarity leads to either directly degradation of the mRNA (Guo et al., 2010a), or inhibition of the translation of the target mRNA (Mourelatos et al., 2002; Pratt and MacRae, 2009).

### 1.3.2.2 miRNA function in early development

Based on the role of *Dicer* and *DGCR8* in miRNAs processing, *Dicer/DGCR8* knockout mice or mESCs have been used for the analysis of miRNA function. Most of the *Dicer* KO mice die before gastrulation, and a subset of surviving embryos appear small and morphologically abnormal (Bernstein et al., 2003). The epiblast was formed in *Dicer* mutants with expression of the early mesoderm maker T in the posterior epiblast, but the PS fails to elongate and there is a drastically reduced or lost expression of the DE markers *Hhex* and *Cer11* (Spruce et al., 2010).

The functions of specific miRNAs in early development have been broadly identified. Initially, miRNAs, such as *lin-4* and *let-7*, were discovered in regulating the developmental timing by directly regulate lineage- and larval stage-specific genes in *C. elegans* (Ambros, 2000; Slack and Ruvkun, 1997). These miRNAs target specifically miRNAs coding for regulatory proteins, to switch between symmetric vs. asymmetric cell divisions or to induce terminal differentiation, thus ensuring developmental timing in the worm (Ambros, 2011). In vertebrates, miRNAs and processing enzymes control the developmental progression during oocyte maturation (Murchison et al., 2007). This suggests that vertebrate miRNAs function analogously to regulate spatio-temporal transitions in gene expression programs during cell-fate acquisition. This is supported by the role of a conserved miR-430/427/302 family, which has been shown to promote mesendoderm formation and repress the neuroectoderm differentiation by targeting of components of the Nodal pathway in

---

*Xenopus* (Rosa et al., 2009). Other miRNAs, such as miR-15 and miR-16, which target the Nodal receptor Activin receptor type II and are negatively regulated by Wnt/ $\beta$ -catenin signaling to establish a dorso-ventral Nodal signaling gradient (Martello et al., 2007). Moreover, miR-430 targets both the Nodal agonist *squint* and the Nodal/TGF- $\beta$  antagonist *Lefty* to control the availability of Nodal ligands in the extracellular space and establish morphogen gradients in the *Zebrafish* embryos (Choi et al., 2007).

### 1.3.2.3 miRNAs function in mESCs

miRNAs have been shown to play a very important role in the maintenance and differentiation of pluripotent mESCs. Both knock-outs of *Dicer* and *DGCR8* prevent the differentiation of mESCs, because they fail to efficiently down-regulate the pluripotency network and up-regulate the differentiation factors (Kanellopoulou et al., 2005; Murchison et al., 2005; Wang et al., 2007). In addition, *Dicer*-null mESCs exhibits low proliferation rates and defective differentiation. The pluripotency factor Oct4 is only partially down-regulated, and the endodermal and mesodermal markers are not detectable when those cells have been induced to embryoid body (EB) differentiation (Kanellopoulou et al., 2005; Murchison et al., 2005). The same defective differentiation problem has been found in *DGCR8*-null mESCs (Wang et al., 2007). These studies indicate the important role of mature miRNAs in mESCs self-renewal and differentiation.

miRNAs are important components of regulatory networks of mESCs, as many core pluripotency factors, such as *Oct4*, *Nanog*, and *Sox2* are targets of miRNAs. On the other hand, the transcription of miRNAs are regulated by these core factors (Cell et al., 2008; Xu et al., 2009). Therefore, these miRNAs and their targets could form a double-negative feedback loop that switches mESCs between self-renewal and differentiation. The genome-wide miRNA expression profiles of mESCs and the differentiated lineage have been performed to identify candidate miRNAs that might be important in mESCs maintenance or differentiation. For instance, the expression of the miR-290-295 cluster and miR-296 is enriched in mESCs and decreases as the mESCs differentiate. This highlights their role in maintaining pluripotency (Houbaviy et al., 2003). On the other hand, some miRNAs, such as miRNA-145 and miRNA-134,

are important in promoting the exit from the pluripotent state by targeting pluripotency genes *Oct4*, *Sox2*, and *Klf4* (Tay et al., 2008; Xu et al., 2009).

### 1.3.3 Important transcription factors (TFs) regulating development

#### 1.3.3.1 The mesoderm-specific TF Brachyury (T)

The gene *Brachyury* (Greek for 'short tail'), or *T* (tail), was first described in 1927 because of the phenotype of short and often slightly kinked tail in mutant mice (Gluecksohn-Schoenheimer, 1938). *T* is an important TF in the process of gastrulation and is required for mesoderm and notochord development (Herrmann, 1991; Kavka and Green, 1997; Schulte-Merker et al., 1994; Wilson et al., 1995).

The mouse *T* gene was identified by positional cloning and is part of the *t*-complex that spans 40 cM on chromosome 17, (Wilkinson et al., 1990). The expression pattern of *T* is conserved amongst vertebrates and marks the mesodermal lineages. Both the mRNA and protein are found in the posterior epiblast and PS from the onset of gastrulation (Burtscher and Lickert, 2009; Wilkinson et al., 1990). *T* mRNA expression increases from the early- to mid-streak stage, but decreases as cells move away from the PS (Smith, 1997). After gastrulation, *T* mRNA is expressed in the node and head process, and is then restricted to the notochord (axial mesoderm) and tail bud (Herrmann, 1991; Technau, 2001). During later stage of development mRNA transcription is silenced.

Many mutant alleles of *T* have been identified and studied by displaying conserved phenotypes within vertebrates. The homozygous *T*<sup>-/-</sup> mice embryos show several morphological defects. During gastrulation, the mutant embryos show a defect in PS, which leads to insufficient mesoderm generation (Yanagisawa et al., 1981). The deficiency in mesoderm formation results in the embryos lacking the notochord and allantois with abnormal somites. The embryos die and are resorbed in *utero* at E10.0~11.0 (Chesley, 1935; Kispert and Herrmann, 1993; Velopment et al., 1944). The heterozygous mice form normal somites and other PS-derived structures, but they do not compete axis formation and have a variable short tail. The phenotype of heterozygous mice can be rescued by incorporating a single copy of the wild type (WT) allele into the genome as a transgene (Clements et al., 1996; Stott et al., 1993).



The zebrafish phenotype of *ntl* and *bra*, which are two *T* ortholog, are similar to mouse *T* mutants, with the absence of a differentiated notochord, tail somites and defect in the most anterior trunk somites (Halpern et al., 1993; Martin and Kimelman, 2008; Schulte-Merker et al., 1994).

### 1.3.3.2 The endoderm-specific TF Foxa2

The forkhead box (Fox) gene family is named after the *Drosophila* gene *fork head*, because its mutations show defects in the formation of the anterior and posterior gut, resulting in a spiked head structure (or 'forks') (Weigel et al., 1989). The Fox gene family has more than 100 Fox genes that have been identified and classified into subfamilies, and many of them have been shown to be involved in a broad variety of biological processes (Carlsson and Mahlapuu, 2002; Kaestner et al., 2000).

The Foxa family of winged helix/forkhead box TFs is a subfamily of the Fox gene family, which is cloned from liver and shown to replace linker histones and affect chromatin structure (Cirillo et al., 2002; Lai et al., 1990, 1991). In mice, this family contains three highly related TFs: Foxa1, Foxa2, and Foxa3 (Hannenhalli and Kaestner, 2009). The Foxa family genes are very important TFs, and act as pioneer factors to open the chromatin structure and allow access to cis-regulatory elements for other TFs (Zaret and Carroll, 2011). Among the three Foxa genes, *Foxa2* is the first gene which is activated during embryogenesis. *Foxa2* expression has been found in embryonic organizing centers of the gastrulating *C.elegans* (Horner et al., 1998), *Zebrafish* (Strähle et al., 1993), *Xenopus* (Bolce et al., 1993), chick and mouse (Ang and Rossant, 1994; Sasaki and Hogan, 1993). In mouse, *Foxa2* expression is detected in the posterior epiblast at the early streak stage. At E7.5, *Foxa2* expression is maintained in the neural plate, notochord and throughout the DE (Ang et al., 1993; Lantz and Kaestner, 2005; Sasaki and Hogan, 1993). Later, *Foxa2* is expressed in endodermally-derived tissues, such as the liver, lung, pancreas, and gastrointestinal tract. *Foxa2* cooperates with Foxa1 to induce tissue specification, branching morphogenesis, and regulation of key endodermal TFs (Burtscher et al., 2013; Friedman and Kaestner, 2006; Mirosevich et al., 2005).

Embryos that lack the *Foxa2* gene show several defects. DE was not formed because the *Foxa2-null* cells fail to maintain epithelial polarity in endoderm

progenitors, which results in lack of proper cellular junctions. These cells are unable to integrate into the outside epithelium (Burtscher and Lickert, 2009). This indicates that *Foxa2* is involved in polarization and epithelialization, and plays multiple roles in endoderm formation in the mouse. Later, the *Foxa2* null embryos show defects in the specification of the foregut DE with the reduction of the expression of the foregut genes *thyrotropin-releasing hormone (Trh)* and *peptide YY (Pyy)*, however, midgut and hindgut formation are unaffected (Dufort et al., 1998; McKnight et al., 2010). The *Foxa2* null embryos do not form a distinct node and notochord, and die at later stage due to embryonic patterning defects of the primary body axes (Ang and Rossant, 1994; Tamplin et al., 2008). In addition, *Foxa2* knock-out analysis reveals that *Foxa2* is required for terminal differentiation and maturation of many endoderm-derived cells. It is involved in multiple stages of pancreatic development (Friedman and Kaestner, 2006), initiation of liver development (Lee et al., 2005), lung alveolarization and regulates airway epithelial cell differentiation during postnatal life (Wan et al., 2004).

### 1.3.3.3 The endoderm-specific TF Sox17

*Sox17* belongs to the Sry (Sex determining region Y)-related HMG (high mobility group) box (Sox) TF family, which was first identified by discovery of the mammalian testis-determining factor Sry (Gubbay et al., 1990). The Sox family consists of 20 different members in mice and humans and encodes a diverse group of TFs that regulate cell-fate decisions, tissue differentiation and regeneration (Pevny and Lovell-Badge, 1997; Sarkar and Hochedlinger, 2013).

In vertebrates, *Sox17* together with *Sox7* and *Sox18* belongs to the Sox group F and was first identified as a stage-specific transcription activator during spermatogenesis (Kanai et al., 1996). *Sox17* shows an endodermally restricted expression pattern in *Xenopus* (Hudson et al., 1997) and *zebrafish* (Alexander and Stainier, 1999) from onset of gastrulation. In the mouse, *Sox17* was shown to be expressed in ExVE and AVE from E6.0 onwards (Kanai-Azuma et al., 2002; Pfister et al., 2007). In the late PS stage, *Sox17* expression is detected in DE, allantois and the blood islands together with the other endoderm TF *Foxa2* (Engert et al., 2009). After gastrulation, *Sox17* is mainly expressed in the posterior endoderm at 3 - 5 somite stage, and then restricted to the mid- and hindgut endoderm; The mRNA expression is no longer

detectable in the endoderm from E9.5 onwards, but is then observed in vascular endothelial cells and blood vessels (Burtscher et al., 2012; Engert et al., 2009, 2013; Kanai-Azuma et al., 2002; Matsui et al., 2006).

In the mouse, after canonical Wnt stimulation  $\beta$ -catenin translocates into the nucleus and interacts with Tcf4 to activate the *Sox17* gene. Associates with  *$\beta$ -catenin*, *Sox17* could activate the transcription of its targets, such as *Foxa1*, *Foxa2*, *Edd* and *Hnf1 $\beta$* , and regulate endodermal transcription (Sinner et al., 2004). In *Xenopus*, ectopic expression of *Sox17* induces the expression of endoderm markers including endodermin, hepatocyte nuclear factor-1 $\beta$  (HNF-1 $\beta$ ), and intestinal fatty acid binding protein (IFABP). Moreover, *in vitro* experiments have shown that *Sox17* overexpression leads ESCs differentiation towards various endoderm lineages including PrE, parietal endoderm, and DE. It also promotes the differentiation of ESC-derived PrE cells into parietal endoderm and VE lineages (Qu et al., 2008; Shirai et al., 2005). Conversely, block of the activity of *Sox17* repressed endodermal marker gene expression in the *Xenopus* embryo (Clements and Woodland, 2000; Hudson et al., 1997). In mouse, *Sox17*-null mutant embryos have defect in DE formation resulting in abnormal morphogenesis of the mid- and hindgut. Another finding is that the loss of *Sox17* function leads to lack of *Pdx1* expression in the residual endodermal cells indicating the pancreas formation is not induced (Kanai-Azuma et al., 2002). All of these observations point to the essential role of *Sox17* during endoderm and organ formation.

### **1.3.4 Epithelial-to-mesenchymal transition (EMT) during gastrulation**

In any given animal tissue, the cell types can be normally categorized into two cell types, an epithelial cell type and mesenchymal cell type. The typical epithelial cells are attached to the basement membrane, display apical-basal polarity and contact with neighbouring cells through adherens and tight junctions (Acloque et al., 2009; Chen et al., 2012; Nakaya and Sheng, 2008). In contrast, mesenchymal cells are often migratory, loosely connected and regarded as cells that do not display epithelial morphology. EMT and its reverse mesenchymal-to-epithelial transition (MET) are important mechanisms that occur during development as well as the tumor metastasis and invasion (Lim and Thiery, 2012; Shook and Keller, 2003).

EMT is essential for gastrulation. During gastrulation, a subset of epithelial cells from the posterior epiblast undergo EMT and migrate through the PS to generate mesoderm and endoderm (Beddington and Robertson, 1999; Tam and Loebel, 2007). During this process, the epiblast cells lose cell-cell adhesion and cell polarity, downregulate E-Cadherin (E-Cad) expression, and breakthrough the basement membrane (BM) to invade into the PS region, in which the mesodermal progenitor acquires a mesenchymal cellular fate to form mesoderm. The formation of mesoderm, thereby, is regarded as a typical EMT process (Nakaya and Sheng, 2008; Williams et al., 2012). In contrast, the endodermal lineage acquires an epithelial fate and intercalates into the overlying VE to form the DE (Acloque et al., 2009; Bartscher and Lickert, 2009). The formation of DE involves a migration from mesenchymal PS to an epithelial endoderm layer, therefore, it is commonly regarded as a MET process.

The complex networks of EMT/MET are initiated and regulated by multiple factors, such as Nodal, Wnt3, Fgf, as well as some EMT related TFs, such as Snail1. Among these factors, Nodal and Wnt3 appear to induce EMT, as both Nodal and Wnt3 mutants fail to induce PS and mesendoderm formation (Brennan et al., 2001; Liu et al., 1999). Fibroblast growth factor (Fgf) receptors are also important to induce the EMT program. In *Fgf8* and *Fgfr1* mutant embryos, Snail1 is not induced to downregulate *E-Cad* expression, thereby, the mesoderm cells fail to migrate away from the PS (Deng et al., 1994; Sun et al., 1999). Snail1 belongs to the Snail family of zinc finger TFs, which have been shown to play crucial roles in the EMT process (Carver et al., 2001; Nakaya and Sheng, 2008; Nieto, 2002), and function as transcriptional repressor of some specific adhesion gene (*E-Cad*), tight junction genes (*Claudins* and *Occludins*) and cell polarity genes (*Crumbs3* and *Discs large*) (Cano et al., 2000; Ikenouchi et al., 2003; Moreno-Bueno et al., 2008). *Snail1* knock-out embryos show abnormal mesoderm morphology with maintained cell polarity and adherens junctions in the mesoderm germ layer (Carver et al. 2001). Other TFs including Twist1, Zeb1, and Zeb2 have also been shown to act as repressors of *E-Cad* and are important to induce mesenchymal genes (Lamouille et al., 2013; Peinado et al., 2007), but are not important for the gastrulation process.

Although the mechanisms involved in the EMT processes during mesoderm formation have been extensively studied, it is currently not clear if also endoderm

undergoes EMT during gastrulation. It is still questionable if DE cells undergo a rapid EMT and then MET transition, or just shortly loosen their epithelial characteristic to migrate out of the epiblast and re-epithelialize in the outside VE. This remains unknown due to the difficulties of continuous single cell tracking *in vivo*. Elucidation of the molecular mechanisms that trigger and promote endoderm by EMT using an *in vitro* system is necessary for a better understanding of embryogenesis.

#### **1.4 Maintenance and differentiation of mESCs *in vitro***

mESCs are isolated from the ICM of blastocyst stage embryos (Evans and Kaufman, 1981; Martin, 1981). They have the capability to self-renew and differentiate into all cell types of an organism. The first mESCs were identified from mice in 1981 (Evans and Kaufman, 1981; Martin, 1981), following by the isolation of human ESCs two decades later (Thomson, 1998). Since then, ESCs have been successfully used to differentiate to desired cell types by appropriate culture conditions and protocols, such as neurons (Cho et al., 2008; Di Giorgio et al., 2008), cardiomyocytes (Laflamme et al., 2007; Vidarsson et al., 2010), hepatocytes (Agarwal et al., 2008; Fagoonee et al., 2010; Hay et al., 2008), and pancreatic endocrine cells (Bruin et al., 2013; D'Amour et al., 2005; Kroon et al., 2008; Xu et al., 2011). The differentiated cell types, such as  $\beta$  cells are used for cell-replacement therapy, disease modeling, drug discovery and toxicity testing (Keller, 2005; Sui et al., 2013).

##### **1.4.1 Maintenance of mESCs pluripotency**

mESCs were originally established and maintained in media containing serum by co-culture with mitotically inactivated mouse embryo fibroblasts (feeder layer, MEF) (Evans and Kaufman, 1981; Martin, 1981), which produce the leukemia inhibitory factor (Lif) and Bmp4 to maintain mESCs pluripotency (Chambers and Smith, 2004; Qi et al., 2004). With the improvement of the culture systems, it is now possible to grow mESCs with defined factors in the absence of feeder cells or serum. Lif, a member of the interleukin-6 family of cytokines, has been used to replace the feeder cell function in the presence of appropriate batches of fetal calf serum (FCS),

indicating its role in supporting the self-renewal of mESCs (Smith et al., 1988). LIF signals through gp130 and activates the Janus kinase (Jak)-signal transducer and activator of transcription 3 (Stat3) pathway to support self-renewal of mESCs. Inhibition of *Stat3* or Jak activity results in the induction of differentiation of mESCs (Niwa et al., 1998). In the presence of LIF, serum could be replaced by Bmp4 to maintain mESCs (Ying et al., 2003). In the mESCs system, the function of Bmp4 is similar to that observed in the mouse embryo. Bmp4 has been implicated in maintaining pluripotency by inducing the expression of the inhibitor of differentiation genes, which can inhibit neuronal differentiation (Ng et al., 2005; Ying et al., 2003). Besides LIF and Bmp4, Wnt ligands are also crucial for preventing mESCs differentiation (ten Berge et al., 2011; Sato et al., 2004; Wray and Hartmann, 2012). In the absence of other defined factors, Wnt proteins in combination with LIF are sufficient to support mESCs self-renewal (ten Berge et al., 2011).

Multiple TFs have been shown to be involved in the maintenance of self-renewal and pluripotency of mESCs. Among them, Oct4, Sox2, and Nanog are core components of pluripotency TF network. These TFs form a regulatory feedback circuit to maintain pluripotency in mESCs by binding efficiently to their own promoters. Therefore, they can regulate themselves and maintain their own expression with autoregulatory and cross-regulatory interactions (Orkin et al., 2008). On the other hand, these TFs also prevent the expression of differentiation-promoting genes (Jiang et al., 2008; Pan and Thomson, 2007). Oct4, a POU domain-containing TF encoded by *Pou5f1*, serves as a master transcriptional regulator of naive pluripotency maintenance. In the absence of Oct4, pluripotent cells lost pluripotency and revert to the trophoblast-like state both *in vivo* and *in vitro*. In contrast, overexpression of Oct4 in mESCs results in the differentiation towards extraembryonic endoderm (Niwa et al., 2000). These divergent effects of Oct4 in mESCs indicate that Oct4 could target both pluripotency and differentiation factors and therefore Oct4 has multiple functions during development.

---

### 1.4.2 Differentiation of mESCs to recapitulate gastrulation

ESCs have the potential to differentiate into all embryonic lineages under appropriate conditions in culture. This enables us to get large quantity of lineage-specific cells for molecular or functional studies, which is difficult to acquire in *in vivo* studies due to the limited size of the embryo or small amount of differentiated cells in organs.

In the literature, three approaches are commonly used to initiate mESCs differentiation. Some protocols use the formation of embryonic bodies (EBs) to initiate mESCs differentiation (Kubo et al., 2004a). In EBs, mESCs form a three dimensional embryo-like structure, which can enhance cell-cell interaction and signaling to form mesoderm, endoderm and ectoderm lineages (Haque et al., 2010). The EB protocol is regarded as the best to mimic embryonic development. The second differentiation protocol is culturing mESCs directly on feeder cells, which provide specific growth factors of the particular used cell line. The fated population is more pure in this case. Recently, mESCs are cultured on a monolayer and extracellular matrices, which can minimize the influence of neighboring cells and feeder cells and better control the culture conditions. However, due to complex and heterogeneous culture conditions, controlling mESCs proliferation and differentiation is still challenging (Haque et al., 2010).

mESCs can differentiate towards mesoderm, endoderm, and ectoderm under certain defined conditions, which are established from the *in vivo* knock-out studies. As described before, the mouse embryos null for Nodal or Wnt3 fail to form the PS, while knock-out of Bmp4 leads to mesoderm defects (Conlon et al., 1994; Liu et al., 1999; Winnier et al., 1995); This points to a role of these signals in PS formation and subsequent lineage specification. Based on this, studies have used Nodal/ActA and Wnt3 to induce APS and posterior PS in a concentration-dependent manner in adherent differentiation of mESCs and in EB cultures (Kubo et al., 2004a; Nakanishi et al., 2009; Yasunaga et al., 2005). After PS formation *in vitro*, higher levels of Nodal/ActA are in support of endoderm induction; whereas lower level of Nodal/ActA in combination with Bmp4 favors mesoderm induction (Johansson and Wiles, 1995; Zorn and Wells, 2009). Thus Wnt3 supplementary with ActA could promote Sox17-positive DE induction, but this effect is minimal after the initiation of DE differentiation (Hansson et al., 2009).

Given that the defined cytokines and growth factors can be supplied in a more controlled manner in mESCs culture, many *in vitro* differentiation model systems have been established (Fehling, 2003; Gouon-Evans et al., 2006; Kubo et al., 2004a; Tada et al., 2005; Yasunaga et al., 2005). Using a T knock-in ES line, Fehling and his colleagues could follow mesoderm induction and its specification to the hemangioblast lineage. They could also separate mesodermal progenitors from those with neuroectoderm potential (Fehling, 2003). This cell line was then used to investigate endoderm formation, and they found endodermal cells were developed from a T positive bi-potent population (Kubo et al., 2004b). This population was further identified by a gooseoid (Gsc) knock-in ES line, whereby the Gsc<sup>+</sup>/E-cad<sup>+</sup> population represents mesendoderm. Subsequently, the Gsc<sup>+</sup>/E-cad<sup>+</sup>/Pdgfra<sup>-</sup> population will give rise to endoderm, while the Gsc<sup>+</sup>/E-cad<sup>-</sup>/Pdgfra<sup>+</sup> population contribute to mesoderm (Tada et al., 2005). The establishment of the *in vitro* differentiation system, thereby, provides us with a model system to investigate elucidate the molecular mechanisms that are involved in development of the germ layers. schon heute im zivilen Luftverkehr nachhaltig hergestellte, alternative Kraftstoffe zu verwenden.



## 1.5 Aims of the study

Pluripotent epiblast cells acquire different signals and migrate through the PS to form either endoderm or mesoderm (Beddington and Robertson, 1999; Zorn and Wells, 2009). However, the temporal and spatial segregations of these endodermal and mesodermal progenitors as well as their molecular control during gastrulation remain unclear due to the lack of continuous single cell analysis at this stage of mouse development. The overall goal of this study was to uncover the progenitor populations and the mechanisms involved in mesendoderm lineage specification.

Therefore, the first aim was to generate a dual fluorescent knock-in  $TGFP^{+/-}$ ;  $Foxa2^{tagRFP/+}$  reporter mESCs line to investigate the lineage specification during endoderm and mesoderm formation on single cell level. Using this dual reporter mESC line and an *in vitro* differentiation system we traced endoderm and mesoderm formation on a continuous single cell level to discover progenitor relationships.

The second aim was to uncover the mechanisms involved in the formation of endoderm and mesoderm. Therefore, progenitor cells using the dual reporter mESC were sorted by FACS and analyzed by whole genome mRNA profiling. Data mining and temporal-spatial expression analysis of these progenitors was analysed to uncover novel mechanisms of progenitor specification and determination and to investigate whether these *in vitro*-derived progenitors are equivalent to their embryonic counterpart *in vivo*. Some morphogenetic processes during gastrulation, such as EMT were also investigated by analyzing the expression of specific marker genes.

Our final goal was to identify novel miRNAs that regulate the mesendoderm lineage specification. To investigate this, we performed a time-resolved miRNA profile to discover differentially expressed miRNAs during differentiation. To identify miRNAs that are involved in the regulation of mesendoderm lineage specification, we selected miRNAs that were predicted to target known signaling cascades (Nodal, Wnt, and Bmp) or TFs (Oct4, Foxa2, and Sox17) by miRNA target prediction analysis. The impact of candidate miRNAs during this process was functionally tested by miRNA overexpression and knock-out approaches *in vivo* or *in vitro*.



## **2. Results**

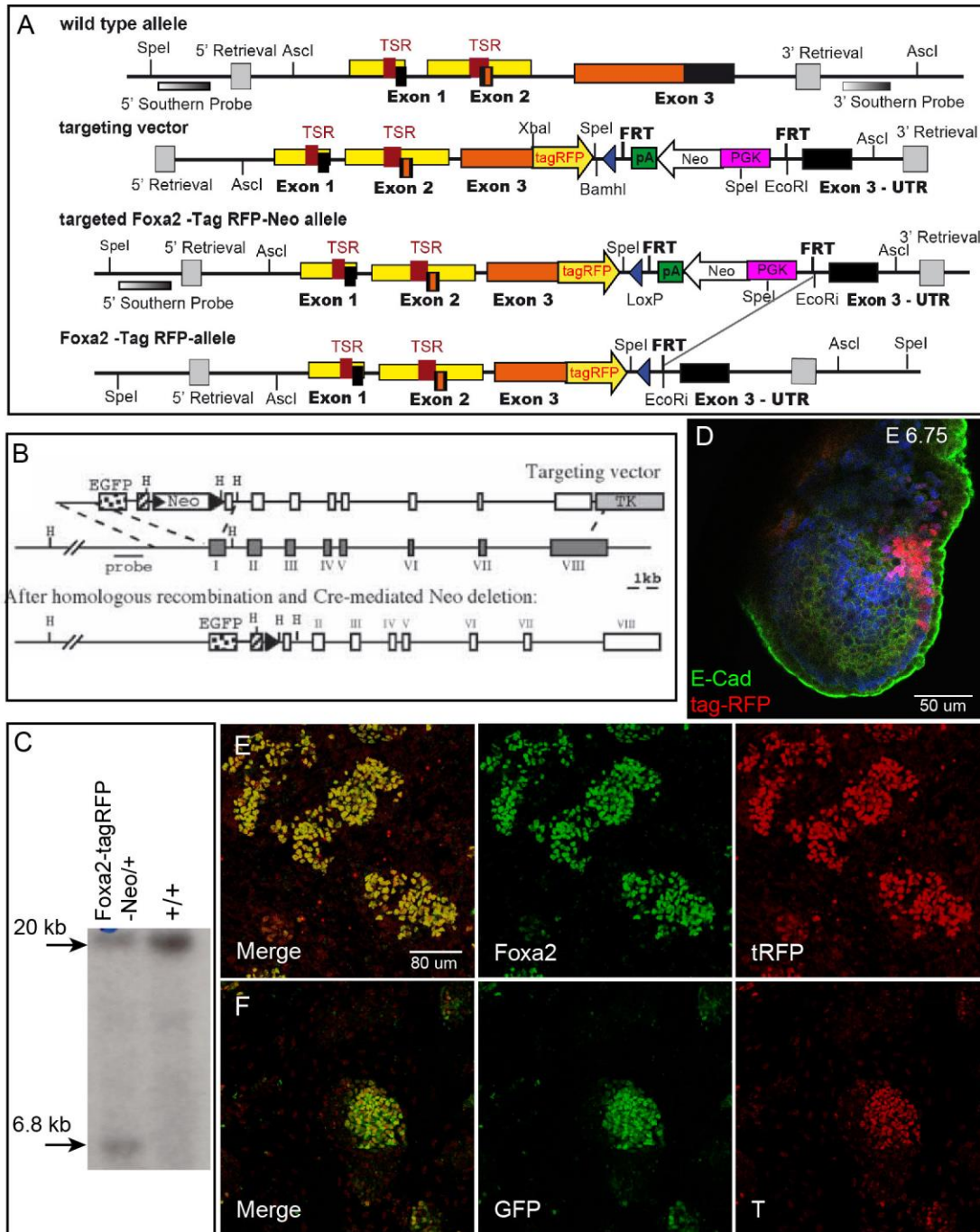
### **2.1 Single cell continuous lineage analysis reveals a novel mechanism of endoderm and mesoderm formation**

### 2.1.1 Generation of a dual knock-in reporter mESC line

The mesendoderm lineage specification as well as the subsequent endoderm or mesoderm formation has been well studied by different lineage tracing approaches (Lawson et al., 1991; Lawson and Pedersen, 1992; Tam and Beddington, 1992; Lawson and Hage, 1994). Since most of the studies are not based on single-cell measurements, it is still unclear when and how these progenitor cells are separated and further differentiate during gastrulation.

Previous studies have shown that the mesodermal TF T and endodermal TF Foxa2 are expressed at various stages of gastrulation within different regions (Beddington and Robertson, 1999; Burtscher and Lickert, 2009; Herrmann, 1991; Tam and Loebel, 2007). In pre/mid-streak stage embryos, T and Foxa2 are synthesized and differentially expressed in proximal and distal domains of the posterior epiblast. The proximal T<sup>+</sup> epiblast cells acquire mesenchymal fate, up-regulate T expression, and migrate through PS to form mesoderm; whereas distal Foxa2<sup>+</sup> epiblast cells up-regulate Foxa2 expression and become endoderm (Figure 1.2). Therefore, the expression of Foxa2 and T could mark the endodermal and mesodermal progenitors of different stages of gastrulation. For this purpose, we generated a *T<sup>GFP/+</sup>; Foxa2<sup>tagRFP/+</sup>* dual knock-in reporter mESCs line to investigate the differentiation process on a single cell level.

First, a *Foxa2-tagRFP* targeting construct was generated by Dr. Ingo Burtscher using conventional cloning techniques and bacterial recombination (Figure 2.1 A). The translational stop codon of *Foxa2* in exon 3 was removed and replaced by the open reading frame (ORF) coding for Red Fluorescent Protein tagRFP. A *loxP-PGK-Neo-loxP* cassette was inserted as a positive selection marker. After electroporation into an available *T-GFP* knock-in mESCs line in which one allele of the ORF coding for T was replaced by GFP (Figure 2.1 B Fehling et al., 2003), *Neo* resistant clones were isolated as homologous events and further confirmed by Southern blotting (Figure 2.1 C). To test whether this mESC line remain pluripotency and has the ability to give rise to three germ layers, an ESC-derived chimera was generated using tetraploid complementation (Nagy et al., 1993). IHC of the ESC-derived mouse embryo showed that the expression of tagRFP was restricted in the posterior epiblast and PS region at E6.75 (Figure 2.1 D).



**Figure 2.1 Generation of dual knock-in reporter  $T^{GFP/+}; Foxa2^{tagRFP/+}$  mESCs line**

(A, B) Schematic representation of Foxa2-tagRFP (A, for details see material and methods) and T-GFP (B, adapted from Fehling et al., 2003) targeting strategy.

(C) Southern blot confirms the correct insertion of tagRFP into the Foxa2 allele. The calculated sized bands of Foxa2 wild type (WT) is 20012 kb, while Foxa2-tagRFP-Neo allele with 20012 bp and 6800 bp.

(D) IHC of ESC-derived mouse embryo shows that Foxa2-tagRFP is expressed at posterior epiblast and PS region at E6.75. Scale bar: 50  $\mu$ m.

(E, F) IHC of Foxa2 and tagRFP as well as T and GFP showed colocalization in endoderm and mesoderm differentiated cells. Scale bars E and F: 80  $\mu$ m.

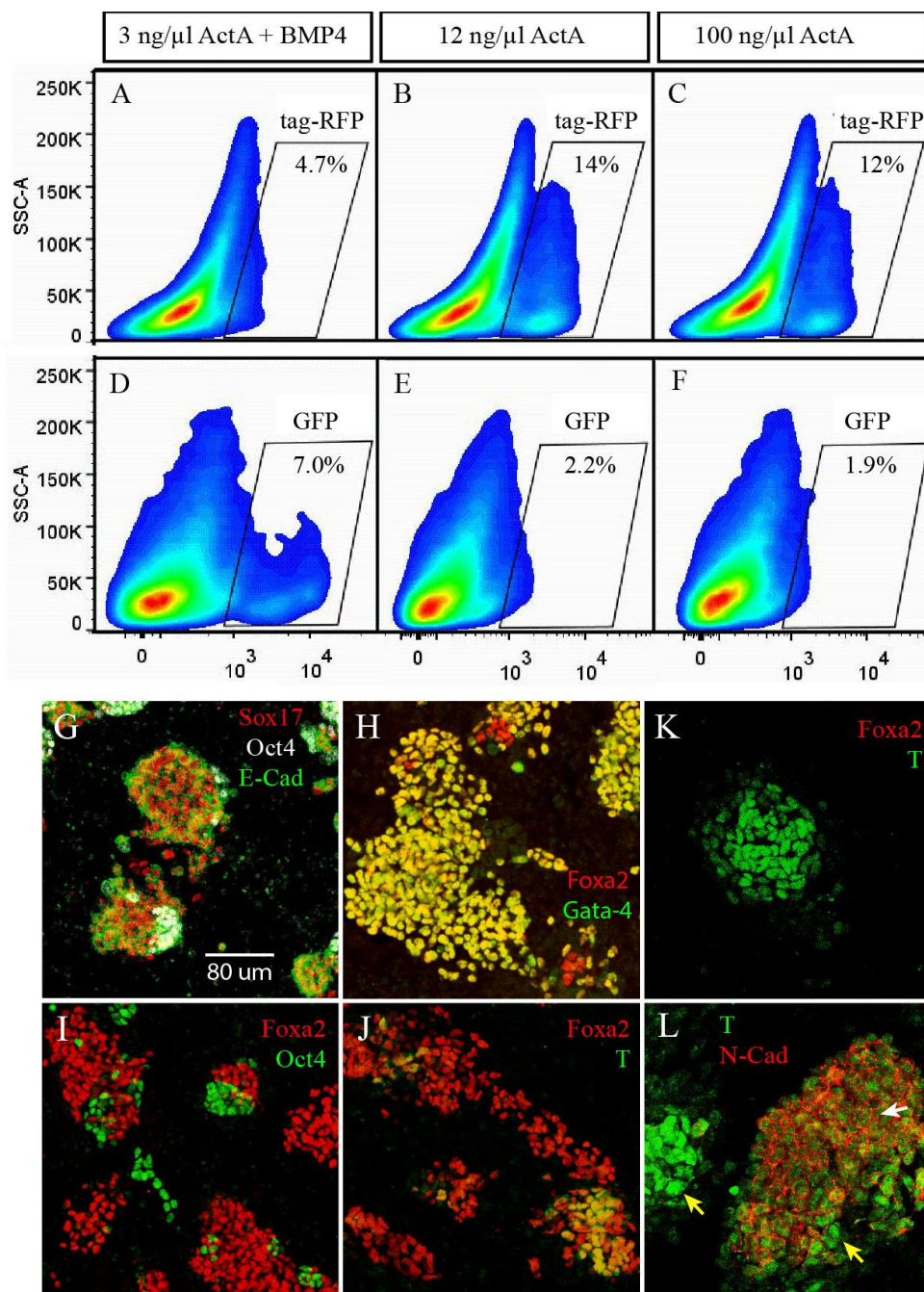
To investigate whether the  $T^{GFP/+}$ ;  $Foxa2^{tagRFP/+}$  dual reporter ESCs line can be used to monitor the differentiated  $Foxa2^+$  and  $T^+$  cells, we differentiated the mESCs (for details see material and methods). After 96 hour (hr) endoderm/mesoderm induction, IHC of  $Foxa2$  and tagRFP as well as T and GFP demonstrated that the expression of tagRFP and GFP correlated with that of endogenous  $Foxa2$  and T, respectively (Figure 2.1 E and F).

Taken together, these results suggested that the dual reporter mESC line remain a pluripotent state and can be used to monitor  $Foxa2$  and T expression.

### **2.1.2 Establishing a mESCs-derived endoderm and mesoderm differentiation system**

Triggering Wnt/ $\beta$ -catenin and Nodal/ActA signaling that induce gastrulation in the mouse embryo leads to differentiation of adherent ESC colonies into the mesendoderm lineage under serum-free conditions (Yasunaga et al, 2005). The subsequent mesendoderm lineage specification is controlled by gradients of signals, where higher levels of Nodal/ActA induce endoderm, and lower doses of Nodal/ActA and Bmp4 signaling promote mesoderm formation (Kubo et al., 2004b; Tada et al., 2005; Winnier et al., 1995).

To test the effect of different signals on cell fate decisions, we differentiated mESCs using different concentrations of cytokines; high concentration (100 ng/ml of ActA), intermediate concentration (12 ng/ml of ActA), and low concentration (3 ng/ml of ActA and 7.5 ng/ml of Bmp4) in a chemically defined medium in the absence of serum (for details see material and methods). Using the  $T^{GFP/+}$ ;  $Foxa2^{tagRFP/+}$  cell line, the efficiency of endoderm ( $Foxa2^+$ ) or mesoderm ( $T^+$ ) differentiation could be easily measured and quantified by expression of tag-RFP or GFP using flow cytometry. Within 2 days of differentiation,  $Foxa2$ -tagRFP was significantly induced with high and intermediate ActA concentrations, reaching a number of 12% and 14% of the total cells (including Wnt3a feeders), respectively; whereas hardly any  $Foxa2^+$  cells were found with low ActA concentrations induction (Figure 2.2 A-C). Since there was no significant difference of  $Foxa2^+$  cells by high and intermediate ActA concentrations, we used 12ng/ml ActA for endoderm differentiation. In contrast, low concentrations of ActA and Bmp4 induced the highest amount of T-GFP expression and mesoderm



**Figure 2.2 Stepwise mESCs differentiation towards endoderm and mesoderm**

(A-F) FACS analysis of Foxa2-tagRFP and T-GFP induction using the following combinations of cytokines: 3 ng/μl ActA + Bmp4, 12 ng/μl ActA, and 100 ng/μl ActA. The percentage of Foxa2 tagRFP<sup>+</sup> cells is significantly increased with high and intermediate ActA induction when compared to low concentrations (A-C). The percentage of T-GFP<sup>+</sup> cells is significantly increased with the induction paradigm using low ActA and Bmp4 in combination (D) when compared to high ActA induction (E and F).

(G-L) IHC of specific pluripotency, endoderm and mesoderm markers in endoderm (G-J) and mesoderm (K and L) differentiated cells at day 4. N-Cad<sup>+</sup> and N-Cad<sup>-</sup> cells are marked by white arrow and yellow arrow, respectively. Scale bars G - L: 80 μm.

induction (Figure 2.2 D-2F). Time-course fluorescence activated cell sorting (FACS) experiments revealed that peak *Foxa2* induction occurred at day 4 of differentiation (data not shown). The FACS data was confirmed by IHC showing the majority of differentiated mESCs give rise to endoderm at day 4 with the expression of a broad range of endoderm markers *Foxa2*, *Sox17*, *Gata4*, and *E-Cad*; whereas only few cells expressed the pluripotency marker *Oct4* (Figure 2.2 G-I). Expression of mesendodermal marker *T* after 4 days of endoderm differentiation was significantly decreased (Figure 2.2 J). These data indicates a high efficiency of endoderm induction at day 4. In parallel, we observed that *T* was highly expressed under mesoderm differentiation conditions and there was a lack of expression of the endoderm marker *Foxa2* (Figure 2.2 K). Interestingly, we observed that the mesoderm derived *T*<sup>+</sup> cells showed heterogeneity in *N-Cadherin* (*N-Cad*) expression, suggesting the existence of different subpopulations under mesoderm induction conditions (Figure 2.2 L). These results confirm that modulating *ActA* and *Bmp4* concentrations lead to the mESCs towards either endoderm or mesoderm *in vitro*.

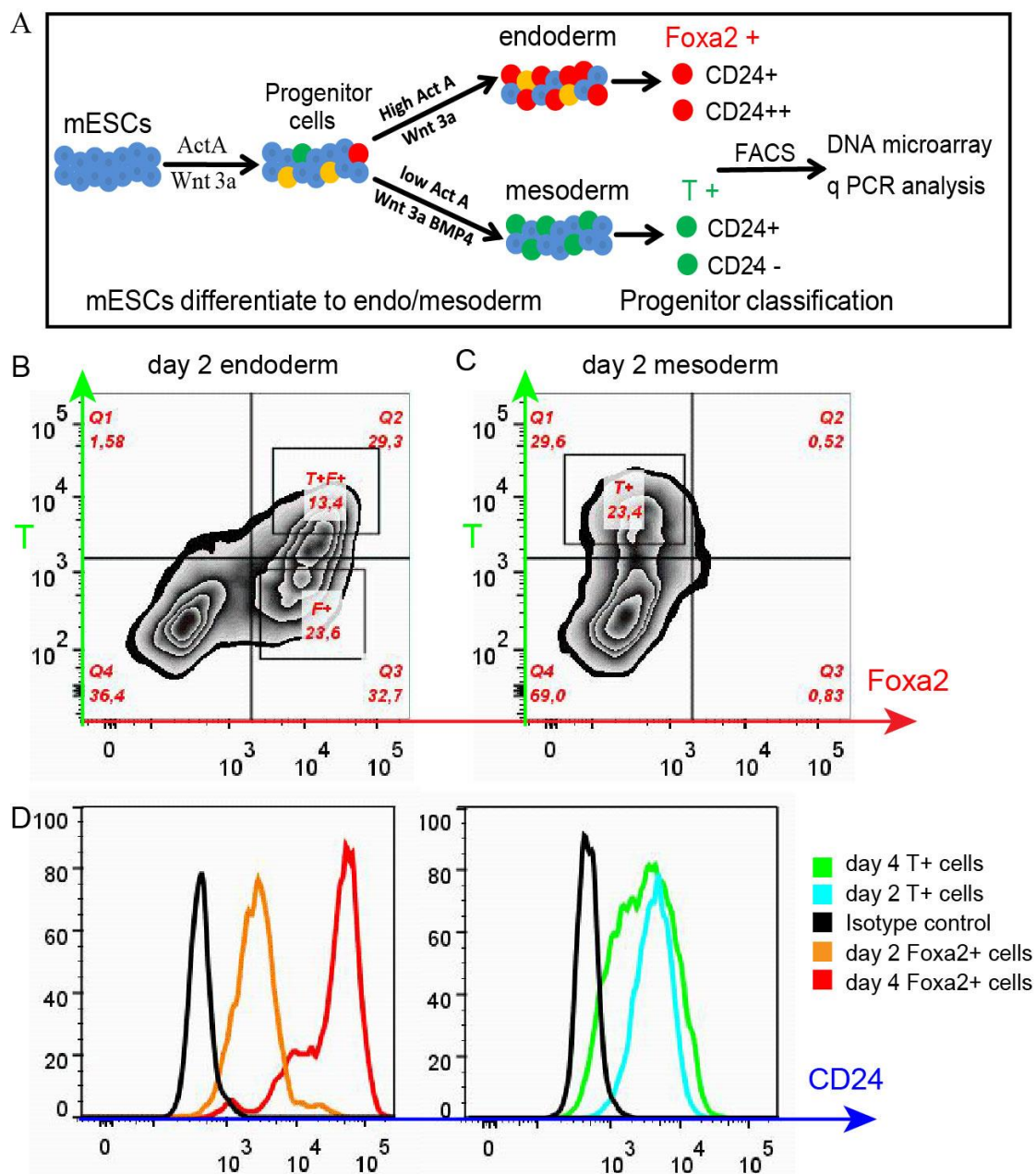
### 2.1.3 Isolation of mESCs-derived endodermal and mesodermal progenitors

The formation of mesoderm or endoderm included changes of adherens junctions, tight junctions, apical-basal polarity, as well as an ingression process, this is considered to be an EMT process for mesoderm formation and an EMT followed by MET process for the DE (Nakaya and Sheng, 2008; Viotti et al., 2014). For these reasons, we assumed that *Foxa2*<sup>+</sup> and *T*<sup>+</sup> epithelial epiblast progenitor population could be further separated from the mesoderm and endoderm lineage by an additional epithelial or mesenchymal marker.

To test this hypothesis, *Foxa2*-tagRFP<sup>+</sup> and *T*-GFP<sup>+</sup> cells were sorted using FACS after 2 or 4 days of differentiation and then further stained with some epithelial and mesenchymal markers, such as *E-Cad*, *N-Cad*, *CD24*, and *CD133* (Figure 2.3 A-D). The epithelial marker *CD24* served to be the best candidate to further separate the *Foxa2*<sup>+</sup> and *T*<sup>+</sup> progenitors. Previous studies have reported that *CD24* was enriched in DE but not in VE (Jiang et al., 2011). *CD24* combined with *Sox17* permitted isolation and separation of naïve DE and VE from tissues comprising the rest of the embryo (Wang et al, 2012). Based on these studies, we analyzed *CD24* expression in



sorted tagRFP<sup>+</sup> cells and GFP<sup>+</sup> cells. Interestingly, we observed that the endoderm differentiated Foxa2 cells could be further separated into two distinct subpopulations by low or high expression of CD24 (Figure 2.3 D). Similarly, the mesoderm differentiated T<sup>+</sup> cells represented two distinct subpopulations, with low or no expression of CD24 (Figure 2.3 D). These initial findings indicate that CD24 can be used as a surface marker to further separate the endoderm and mesoderm population.



**Figure 2.3 Isolation of mESCs-derived endodermal and mesodermal progenitors**

(A) A graphical representation of the approach for systematic differentiation, isolation and molecular characterization of mESCs derived progenitors and lineage.

## 2.Results

---

(B, C) FACS blots show the efficiency of endoderm (B) and mesoderm (C) differentiation at day 2 as well as the approach to isolate Foxa2<sup>+</sup>, Foxa2<sup>+</sup>T<sup>+</sup> and T<sup>+</sup> cells. The efficiency is shown by percentages of number.

(D) FACS analysis of CD24 expression in sorted Foxa2<sup>+</sup> and T<sup>+</sup> cells. The expression of CD24 is increased in endoderm differentiated Foxa2<sup>+</sup> cells from day 2 to day 4; whereas the CD24 level is decreased in mesoderm differentiated T<sup>+</sup> cells.

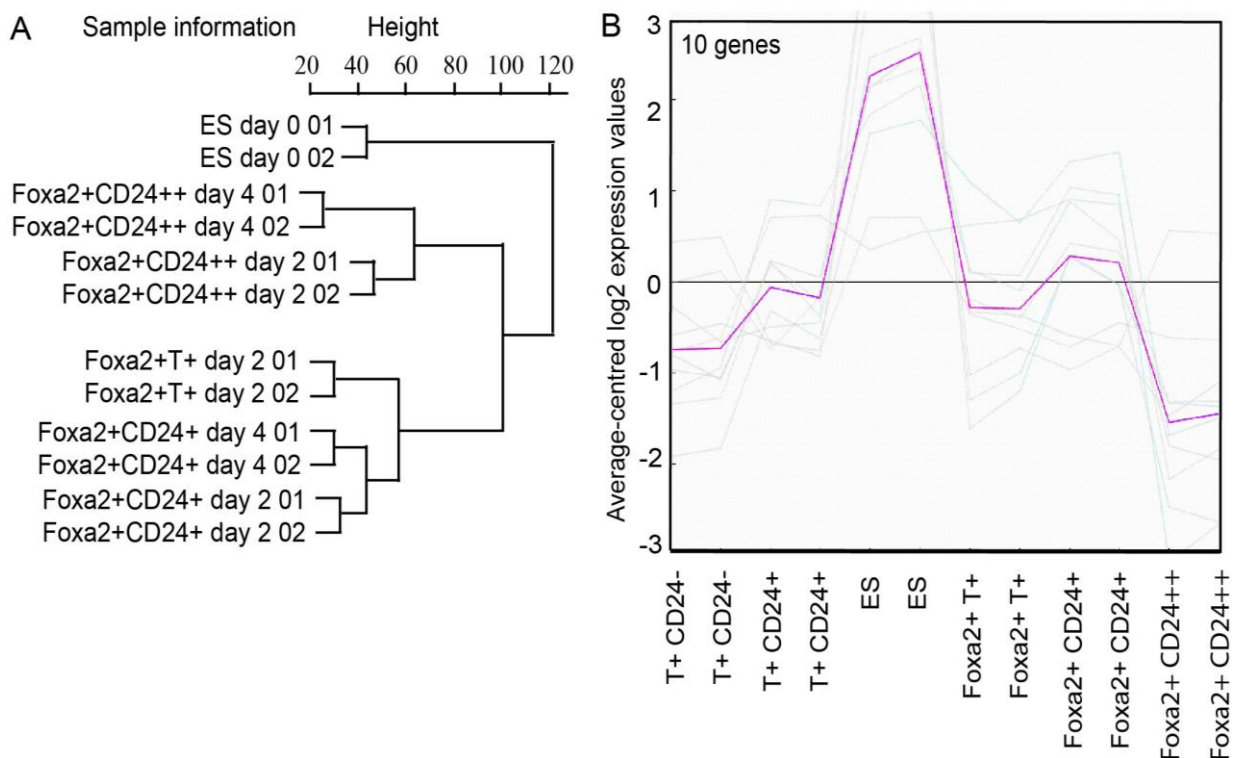
### **2.1.4 CD-24 is a valuable marker to separate early and late endodermal/mesodermal progenitors**

As shown above, the epithelial surface marker CD24 serves as a good candidate to further separate Foxa2<sup>+</sup> and T<sup>+</sup> cells progenitors and lineage. This let us to hypothesize that CD24 can be a novel molecule marker to resolve the progression of differentiation *in vitro* by separating these early and late endoderm/mesodermal cells.

To prove this hypothesis and better understand the molecular characteristics of these *in vitro* differentiated endodermal and mesodermal progenitors, we isolated these Foxa2<sup>+</sup>CD24<sup>+</sup> and Foxa2<sup>+</sup>CD24<sup>++</sup> cells, T<sup>+</sup>CD24<sup>+</sup> and T<sup>+</sup>CD24<sup>-</sup> cells, Foxa2<sup>+</sup>T<sup>+</sup> cells from day 2 to day 4 by FACS, and performed microarray gene expression analysis of these populations (Figure 2.3 A). Undifferentiated mESCs were also included at the pluripotent state. The analysis of mRNA profile revealed that a total of 8858 and 4902 probe sets (FDR ≤ 10% and ratio ≥ 1.5x) were differentially expressed in Foxa2<sup>+</sup> lineage and T<sup>+</sup> lineage when compared to mESCs, respectively. Hierarchical gene clustering of these differentially expressed probe sets revealed that all differentiated cells populations clustered closely and were distinct from the pluripotent mESCs (Figure 2.4 A). Interestingly, Foxa2 single positive cells sorted from either day 2 or 4 did not cluster together. Instead, they were clustered rather in relation to the expression of CD24 level (Figure 2.4 A), which indicates that CD24 may serve as a valuable marker to show progression of endoderm differentiation.

Before gastrulation, the epiblast cells keep a pluripotent state and show high expression of core pluripotent genes. The formation of the primary germ layers via pluripotent epiblast represents a process that needs exit from the pluripotent state. Therefore, we investigated the developmental changes of these sorted populations by multi-class analysis of 10 pluripotent marker genes (Table 2.1). As shown in

Figure 2.4 B, the expression of pluripotency genes was high in mESCs, and decreased in  $\text{Foxa2}^+\text{CD24}^+$  and  $\text{T}^+\text{CD24}^+$  cells, as well as in  $\text{Foxa2}^+\text{T}^+$  cells, but still remained relatively higher; whereas it was barely expressed or absent in  $\text{Foxa2}^+\text{CD24}^{++}$  and  $\text{T}^+\text{CD24}^-$  cells. The reduction of pluripotent gene expression in either  $\text{Foxa2}^+$  or  $\text{T}^+$  cells is related to the CD24 expression level, which suggests that CD24 accurately reflects the differentiation progression. The sustained expression of these pluripotent genes in  $\text{Foxa2}^+\text{CD24}^+$ ,  $\text{T}^+\text{CD24}^+$  and  $\text{Foxa2}^+\text{T}^+$  cells indicates that these cells are early endodermal and mesodermal progenitors, which are formed in intermediate stages of endoderm/mesoderm differentiation; whereas,  $\text{Foxa2}^+\text{CD24}^{++}$  and  $\text{T}^+\text{CD24}^-$  cells represent late endodermal and mesodermal progenitors due to their low expression of pluripotent genes.



**Figure 2.4 CD-24 is a valuable marker to show the differentiation progress**

(A) Cluster dendrogram showing the relationships between the *in vitro* endoderm differentiated cell populations and mESCs. The differentiated cells are clustered together and are distinct from mESCs. The cluster of differentiated  $\text{Foxa2}^+$  cells relies on CD24 expression.

(B) Multi-class analysis of 10 pluripotent genes in the sorted populations show that mESCs keep the highest expression of pluripotent markers; whereas the  $\text{T}^+\text{CD24}^-$  and  $\text{Foxa2}^+\text{CD24}^{++}$  populations express the lowest levels of the pluripotency markers genes.

<b>Gene</b>	<b>Gene Title</b>	<b>Probe set</b>
Sox2	SRY-box containing gene 2	10491477
Dppa5a	developmental pluripotency associated 5A	10595156
Dnmt3l	DNA (cytosine-5-)-methyltransferase 3-like	10364346
Utf1	undifferentiated embryonic cell transcription factor 1	10558580
Dppa4	developmental pluripotency associated 4	10436050
Nanog	Nanog homeobox	10541524
Stat3	signal transducer and activator of transcription 3	10391301
Gdf3	growth differentiation factor 3	10547633
Fbxo15	F-box protein 15	10457077
Dppa4	developmental pluripotency associated 4	10486954

**Table 2.1 List of pluripotency genes used in the multi-class analysis**

<b>Gene</b>	<b>Gene Title</b>	<b>Probe set</b>
Sox2	SRY-box containing gene 2	10491477
Nanog	Nanog homeobox	10541524
T	Brachyury	10441669
Tdgf1	teratocarcinoma-derived growth factor 1	10597268
Otx2	orthodenticle homolog 2 (Drosophila)	10419356
Fgf5	fibroblast growth factor 5	10523490
Foxa2	forkhead box A2	10488374

**Table 2.2 List of epiblast genes used in the multi-class analysis**

### 2.1.5 Distinct *in vitro*-derived populations reflect *in vivo* lineage progenitors

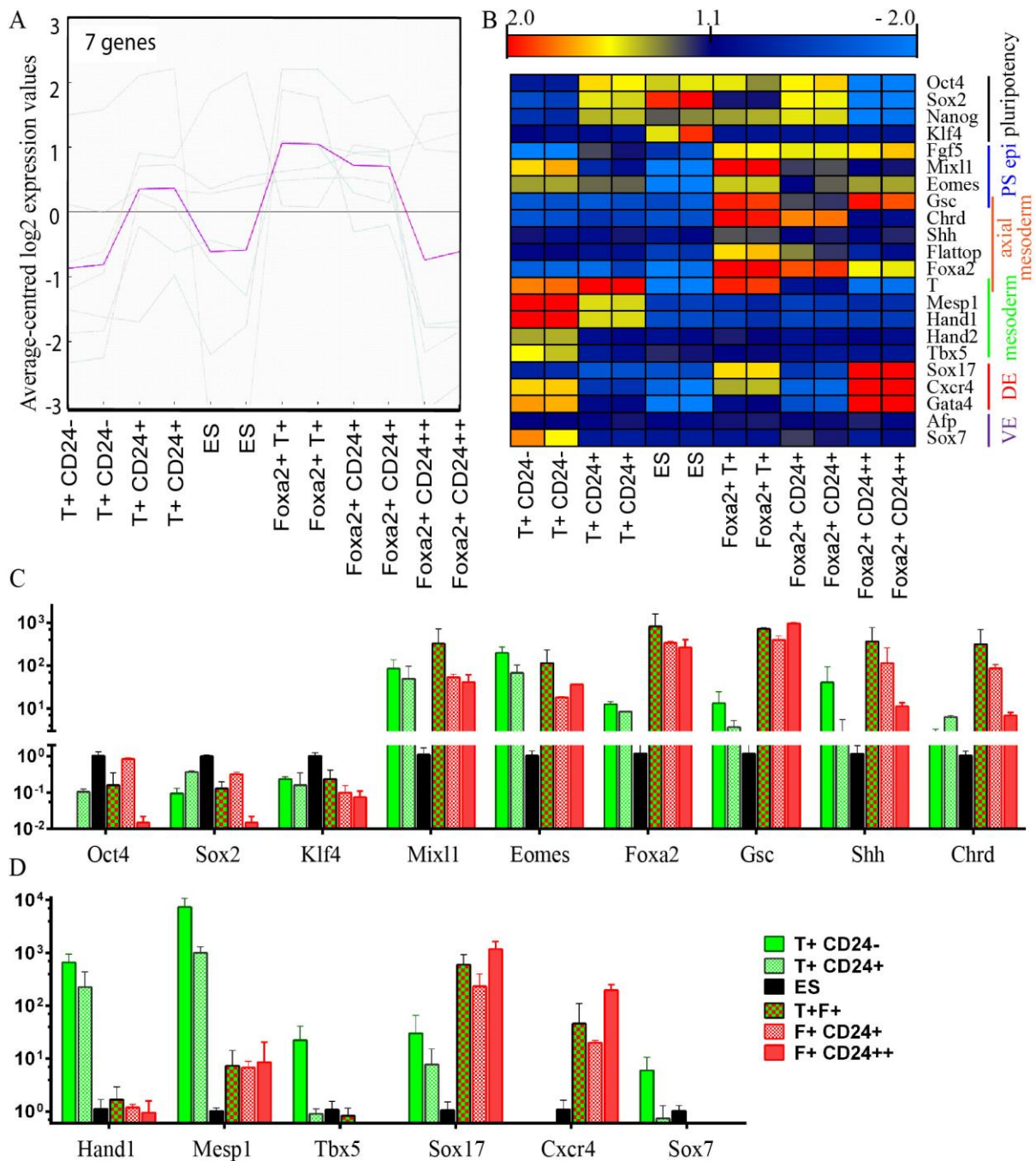
At pre-streak-stage, the pluripotent Oct4<sup>+</sup> epiblast cells were segregated into two distinct populations with the expression of either T or Foxa2. These populations up-regulate either T or Foxa2 and migrate through PS region to give rise to three different populations including T<sup>+</sup> mesoderm, Foxa2<sup>+</sup> DE, and Foxa2<sup>+</sup>T<sup>+</sup> axial mesoderm population (Burtscher and Lickert, 2009). As we have already isolated and classified these *in vitro*-derived populations, we next asked whether these early and late endodermal/mesodermal progenitors are functionally equivalent to their embryonic counterpart *in vivo*.

The sustained expression of the pluripotent genes in Foxa2<sup>+</sup>CD24<sup>+</sup>, T<sup>+</sup> CD24<sup>+</sup> and Foxa2<sup>+</sup>T<sup>+</sup> cells raised the possibility that these early endodermal and mesodermal progenitors might closely related to the posterior epiblast cells (Figure 2.4 B), which are pluripotent and specific signal gradients and response allocates them to either endoderm or mesoderm lineage. To follow up this observation, another series of multi-class analysis was performed by a list of epiblast markers (Table 2.2). This analysis revealed that epiblast genes were expressed higher in these early endodermal and mesodermal progenitors than mESCs and late endodermal/mesodermal cells (Figure 2.5 A). To further investigate the molecular characteristic of these sorted populations, we compared the gene expression patterns among them with specific pluripotency, epiblast, PS, mesoderm and DE markers. Gene expression profiling revealed that these early endodermal and mesodermal progenitors were enriched for pluripotency, epiblast and PS genes, whereas expression of mature endoderm and mesoderm markers was more pronounced in the late endodermal and mesodermal progenitors (Figure 2.5 B). Of note, *Mixer-like 1 (Mixl1)*, *Eomesodermin (Eomes)*, and *Gsc*, which define the epiblast and streak region, were higher expressed in Foxa2<sup>+</sup>T<sup>+</sup> population than the other populations (Figure 2.5 A). Quantitative PCR (qPCR) analysis reinforced these observations, showing a relative higher expression of core pluripotency markers *Sox2* and *Oct4* in these earlier endodermal/mesodermal progenitors. Among these early progenitors, Foxa2<sup>+</sup>T<sup>+</sup> cells showed highest level of *Mixl1* and *Eomes* expression (Figure 2.5 B). Collectively, the high expression of pluripotent and epiblast genes as well as low expression of PS genes in Foxa2<sup>+</sup>CD24<sup>+</sup> and T<sup>+</sup>CD24<sup>+</sup> cells, indicating that these Foxa2<sup>+</sup>CD24<sup>+</sup> and T<sup>+</sup>CD24<sup>+</sup> cells have similar properties

with the cells formed in distal or proximal domains of posterior epiblast. Conversely, the Foxa2<sup>+</sup>T<sup>+</sup> cells behave more like the cells identified in PS region due to their co-expression of PS genes and pluripotent genes.

In the mouse embryos, the Foxa2<sup>+</sup>T<sup>+</sup> cells are found in anterior PS and restrict to axial mesoderm as gastrulation ends (Burtscher and Lickert, 2009; Kinder et al., 2001). For this purpose, we examined the expression of specific axial mesoderm markers among the sorted populations. Interestingly, we found that *Flattop (Fltp)* and *Sonic hedgehog (Shh)*, which were first expressed in the embryonic node at E7.5 (Echelard et al., 1993; Gegg et al., 2014), as well as *Gsc* and *Chrd*, were expressed higher in the Foxa2<sup>+</sup>T<sup>+</sup> cells than the other populations (Figure 2.5 B and C). The high expression of these axial mesoderm related genes in Foxa2<sup>+</sup>T<sup>+</sup> cells indicating these cells might represent progenitors of axial mesoderm and node cells *in vitro*.

As described above, we observed high efficiency of endoderm and mesoderm induction at day 4 differentiation (Figure 2.2 G-L). To investigate the characteristics of sorted day 4 populations, some more specific mature mesoderm and endoderm markers were used. Our microarray data show the mesoderm differentiation towards more paraxial and cardiac mesoderm, with the up-regulation of paraxial (*Mesp1*, *Mesp2* and *Pdgfra*) and cardiac (*Hand1*, *Hand2*, and *Tbx5*) mesoderm marker expression only in T<sup>+</sup> lineage mesodermal cells. The peak mesoderm expression was in T<sup>+</sup>CD24<sup>-</sup> mesoderm cells (Figure 2.5 B). qPCR confirmed these mRNA profile data and showed an increase of paraxial and cardiac mesoderm markers expression in T<sup>+</sup> lineages (Figure 2.5 D). Our data indicates that paraxial and cardiac mesoderm share similar progenitors *in vitro*, which is in consistence with previous report that paraxial mesoderm and heart cells were derived from the similar region in the gastrulation stage embryos (Kinder et al., 1999). By contrast, the Foxa2<sup>+</sup>CD24<sup>++</sup> population expressed the highest levels of *Sox17*, *Gata4*, as well as *chemokine receptor 4 (Cxcr4)*, all of which are expressed in DE (Figure 2.5 B and D). Importantly, the expression of *Afp* and *Sox7* that are associated with the VE, but not DE, were barely detectable in the Foxa2<sup>+</sup>CD24<sup>++</sup> population as well as other Foxa2 lineage cells (Figure 2.5 B and D). This reveals that the Foxa2<sup>+</sup>CD24<sup>++</sup> population is more closely related-to the embryonic mature DE.



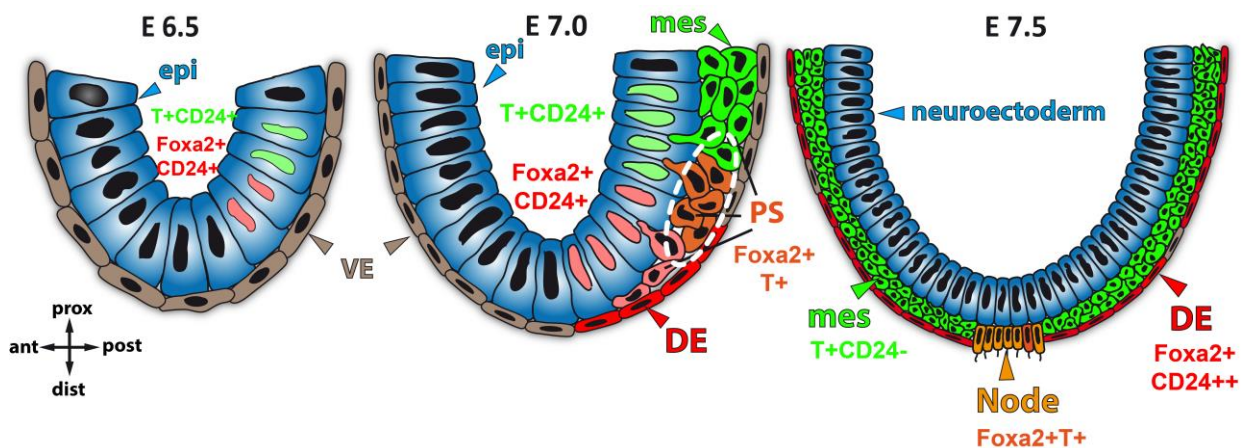
**Figure 2.5 Characterization of mESCs-derived endodermal and mesodermal progenitors**

(A) Multi-class analysis of 7 epiblast genes in the sorted populations show that Foxa2<sup>+</sup>CD24<sup>+</sup>, T<sup>+</sup>CD24 and Foxa2<sup>+</sup>T<sup>+</sup> cells have relative higher expression of epiblast markers when compared to the other populations.

(B-D) Microarray gene expression analysis of pluripotent, epiblast, PS, axial mesoderm, mesoderm, DE and VE genes in sorted populations (B). These data are further confirmed by qPCR (C, D). Data represented by mean+Standard deviation (SD).



In summary, the above results indicate that the early endodermal and mesodermal progenitors  $\text{Foxa2}^+\text{CD24}^+$  and  $\text{T}^+\text{CD24}^+$  cells have similar properties with the cells found in posterior epiblast; where the PS-like  $\text{Foxa2}^+\text{T}^+$  cells represent the progenitors of axial mesoderm. On the other hand, the late endodermal and mesodermal progenitors  $\text{Foxa2}^+\text{CD24}^{++}$  and  $\text{T}^+\text{CD24}^-$  represent more embryonic DE and mesoderm (Figure 2.6).



**Figure 2.6** *In vitro* derived populations correlate with the embryonic cell populations formed during gastrulation

### 2.1.6 Delineating the mesendoderm lineage hierarchy in culture

The mesendoderm lineage specification and formation of endoderm and mesoderm have been well studied both *in vivo* and *in vitro*. In mouse embryos it has been shown that endodermal and mesodermal cells emerge from a similar region in the posterior PS region during gastrulation (Burtscher and Lickert, 2009; Herrmann, 1991; Lawson et al., 1991; Lickert et al., 2002; Sasaki and Hogan, 1993; Tam and Beddington, 1987). *In vitro* experiments have suggested that mesoderm and endoderm are derived from a  $\text{T}^+$  bi-potent mesendoderm population, but this hypothesis has not been proven due to the lack of single cell continuous live cell analysis.

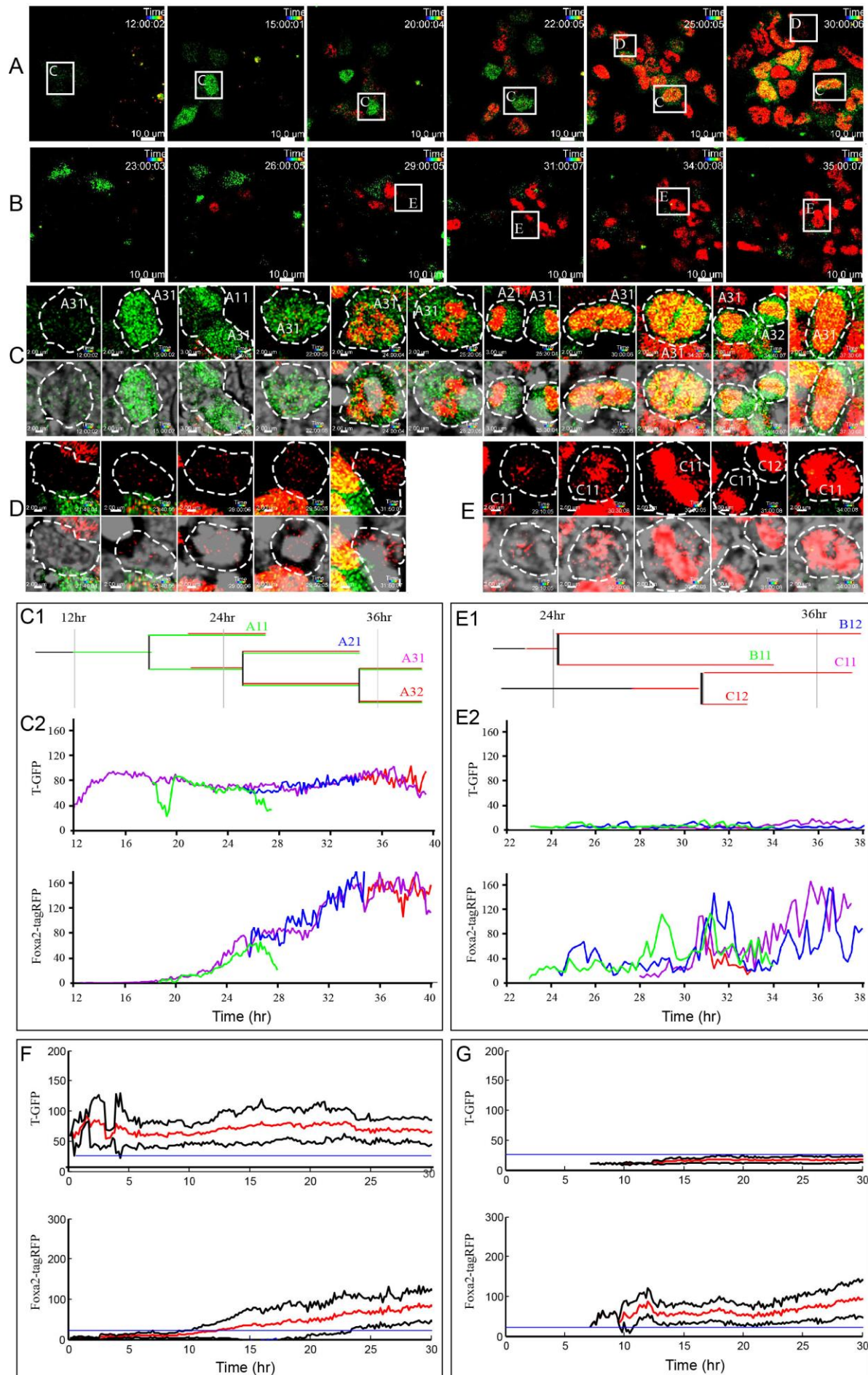


---

To investigate the mesendoderm lineage specification as well as endoderm and mesoderm formation in detail, we generated a live-cell imaging system using time-lapse confocal microscopy. Surprisingly, two distinct processes were discovered upon endoderm induction (Figure 2.7 A and B). We observed that T was first induced in differentiated mESCs, and then these T<sup>+</sup> cells up-regulate Foxa2 expression and became Foxa2<sup>+</sup>T<sup>+</sup> cells. Hardly any Foxa2<sup>+</sup>T<sup>+</sup> cells could give rise to Foxa2 single positive cells (Figure 2.7 A); whereas, the majority of the Foxa2<sup>+</sup> cells were derived directly from mESCs (Figure 2.7 B). To get a closer view, we developed an approach to visualize the differentiation on single cell level using image quantification software Imaris (Figure 2.7 C-E). This enabled us to quantify the expression of tag-RFP and GFP within single cell during the time course of differentiation. The background/noise in undifferentiated mESCs was used as a negative control for the GFP and tag-RFP signal. Consistent with the above observation (Figure 2.7 A), we showed an up-regulation of GFP signal as well as the subsequent up-regulation of tag-RFP signal in a single cell; and both signals were maintained (Figure 2.7 C and F). In parallel, we showed an up-regulation of Foxa2 expression, but not T expression, in the formation of Foxa2<sup>+</sup> single cell during the time course of differentiation (Figure 2.7 D, E, and G).

During gastrulation, T<sup>+</sup> progenitors and mesoderm are mainly formed in the proximal epiblast adjacent to the ExE. The ExE is the source of Bmp signaling, which was shown to be important for mesoderm induction. To monitor the effect of mesoderm inducing conditions (low ActA+Bmp4) on mesendoderm lineage formation, we used continuous live cell imaging as outlined above. T expression was induced quickly after induction of differentiation and maintained during the differentiation time course. Hardly any mesoderm differentiated cells expressed Foxa2 (Figure 2.8 A). In line with the previous observation, we showed an up-regulation of GFP signal at the earlier points of the movie. The signal was maintained with a relative high level from 5 hr until the end of the movie; tag-RFP signal was not increased at any time points of differentiation (Figure 2.8 B-E).

## 2. Results



---

**Figure 2.7 Single cell time-lapse imaging of endoderm differentiation *in vitro***

(A, B) Time-lapse confocal imaging of  $T^{GFP/+}$ ;  $Foxa2^{tagRFP/+}$  fluorescent reporter mESCs line under endoderm differentiation reveals two distinct lineages: 1, mESCs  $\rightarrow$  T<sup>+</sup> cells  $\rightarrow$  Foxa2<sup>+</sup>T<sup>+</sup> cells (A) and 2, mESCs  $\rightarrow$  Foxa2<sup>+</sup> cells  $\rightarrow$  Foxa2<sup>+</sup> cells (B). Time point indicates in hr:min:sec from the start of the movie. Scale bars: 10  $\mu$ m.

(C) Single cell time-lapse confocal imaging of endoderm differentiated Foxa2<sup>+</sup>T<sup>+</sup> cell. T-GFP was first induced at the beginning of the movie, followed by the up-regulation of Foxa2-tagRFP expression in these tracked cells. Single cell is depicted by dashed line. Scale bars: 2  $\mu$ m.

(C1) Single cell lineage tree analysis of endoderm differentiated Foxa2<sup>+</sup>T<sup>+</sup> cell. Green lines represent the expression of T-GFP, while the red lines show Foxa2-tagRFP expression in the time course of differentiation. A11, A21, A31 and A32 represent daughter cells that were tracked over time.

(C2) Quantification of the GFP and tagRFP intensity levels of Foxa2<sup>+</sup>T<sup>+</sup> cells at the single cell level in the time course of endoderm differentiation. The color lines represent signal intensity of the individual cells.

(D-E) Single cell time-lapse confocal imaging of endoderm differentiated Foxa2<sup>+</sup> cells. Foxa2-tagRFP was up-regulated directly from differentiated mESCs without undergoing T stage. Single cell is circled by dashed line. Scale bars: 2  $\mu$ m.

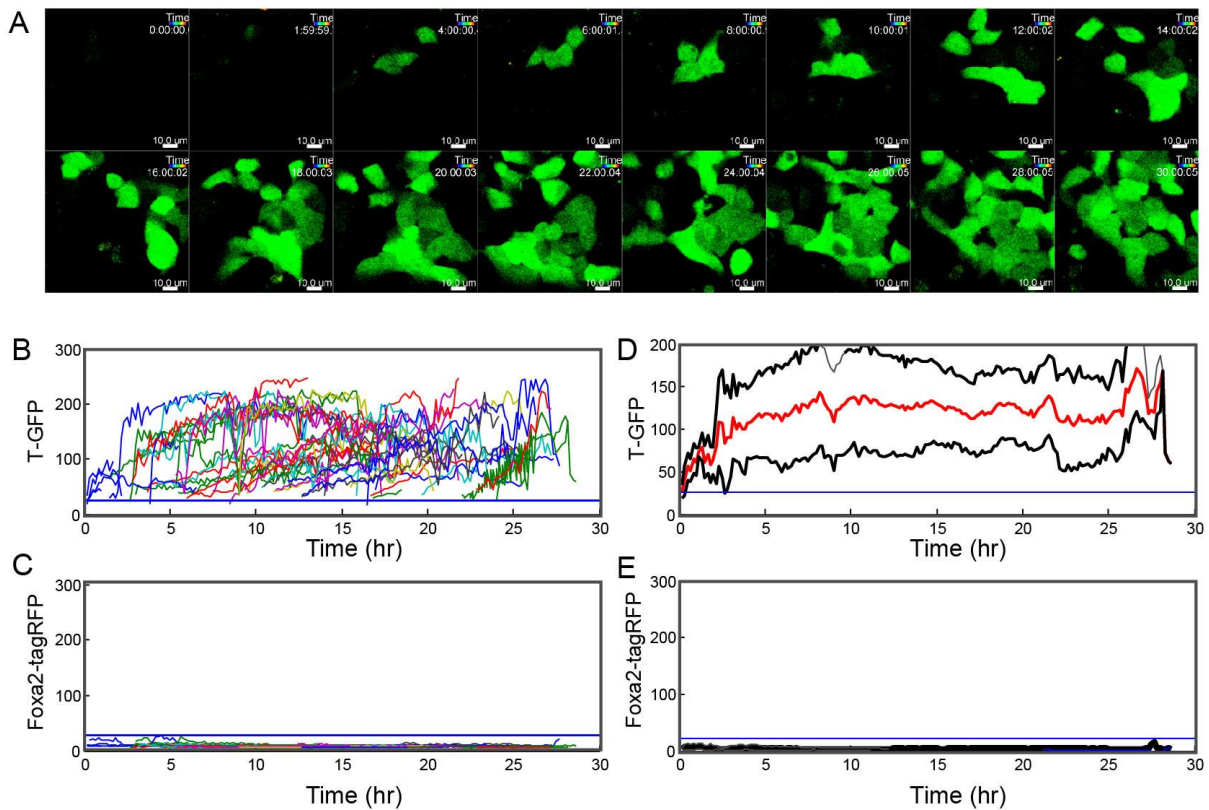
(E1) Single cell lineage tree analysis of endoderm differentiated Foxa2<sup>+</sup> cell. Green lines represent the expression of T, while the red lines stand for Foxa2 expression in the time course of differentiation.

(E2) Quantification of the GFP and tagRFP intensity levels of Foxa2<sup>+</sup> cells at the single cell level in the time course of endoderm differentiation. The color lines represent signal intensity of individual cells.

(F) Quantification of the GFP and tagRFP overall intensity levels of Foxa2<sup>+</sup>T<sup>+</sup> cells in the time course of endoderm differentiation. Data represented by **mean**  $\pm$  Standard deviation (SD). The blue line represents signal intensity of the background/noise in undifferentiated mESCs.

(G) Quantification of the GFP and tagRFP overall intensity levels of Foxa2<sup>+</sup> cells in the time course of endoderm differentiation. Data represented by **mean**  $\pm$  Standard deviation (SD). The blue line represents signal intensity of the background/noise in undifferentiated mESCs.

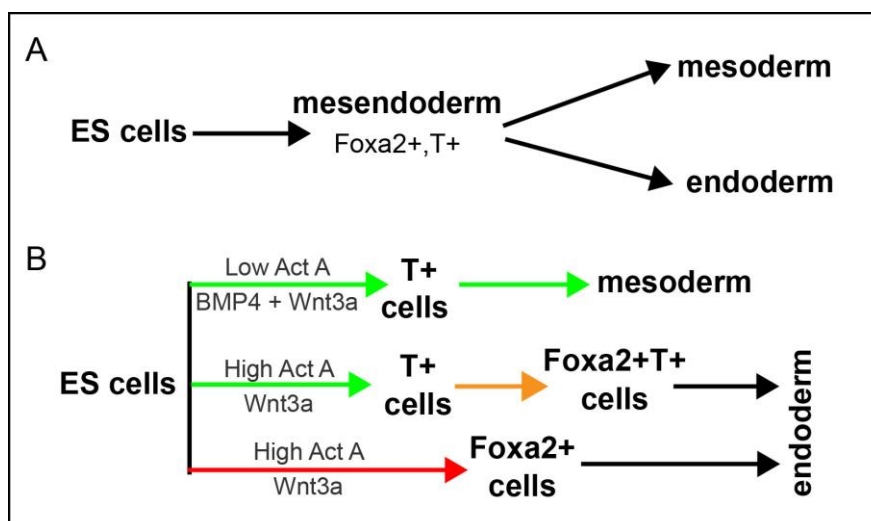
## 2.Results



**Figure 2.8 Analysis of *in vitro* mesoderm differentiation dynamics**

(A) Time-lapse confocal imaging of  $T^{GFP/+}$ ,  $Foxa2^{tagRFP/+}$  fluorescent reporter ES line under mesoderm differentiation. Time point indicated in hr:min:sec from start of the movie. Scale bars: 10  $\mu$ m.

(B-E) Plots of GFP and tagRFP intensity levels in the time course of mesoderm differentiation, presented by multi-single cell level (B and C) and average level (D and E). The intensity levels of GFP are induced from the beginning of the movie, and maintained till the end.  $Foxa2$ -tagRFP is not induced under mesoderm conditions. The blue line represents signal intensity of the background/noise in undifferentiated mESCs. Data represented by  $mean \pm$  Standard deviation (SD).



**Figure 2.9 Graphical models of *in vitro* endoderm and mesoderm formation**

---

Taken together, we identified three distinct mesendoderm lineages: mESCs→T<sup>+</sup> cells; mESCs→T<sup>+</sup>cells→Foxa2<sup>+</sup>T<sup>+</sup> cells; and mESCs→Foxa2<sup>+</sup> cells, that form directly from pluripotent ESCs under endoderm and mesoderm promoting differentiation conditions (Figure 2.9); these mesendoderm lineages could contribute to three distinct cell types during differentiation. These findings contradict the classic concept that both endoderm and mesoderm are derived from a common progenitor that co-expresses Foxa2 and T (Figure 2.9 A) and reveal that distinct progenitors give rise to the endoderm and mesoderm lineages (Fig. 2.9 B).

### 2.1.7 ADE and PDE are derived from distinct progenitor cells

In the mouse embryo, DE is patterned along the anterior-posterior into distinct A-P domains by the end of gastrulation. The foregut gives rise to lung, thyroid, pancreas and liver, and the mid/hindgut contributes to small and large intestine (Zorn and Wells, 2009). However, the origins of ADE and PDE as well as the mechanisms involved in endoderm patterning are still not fully understood.

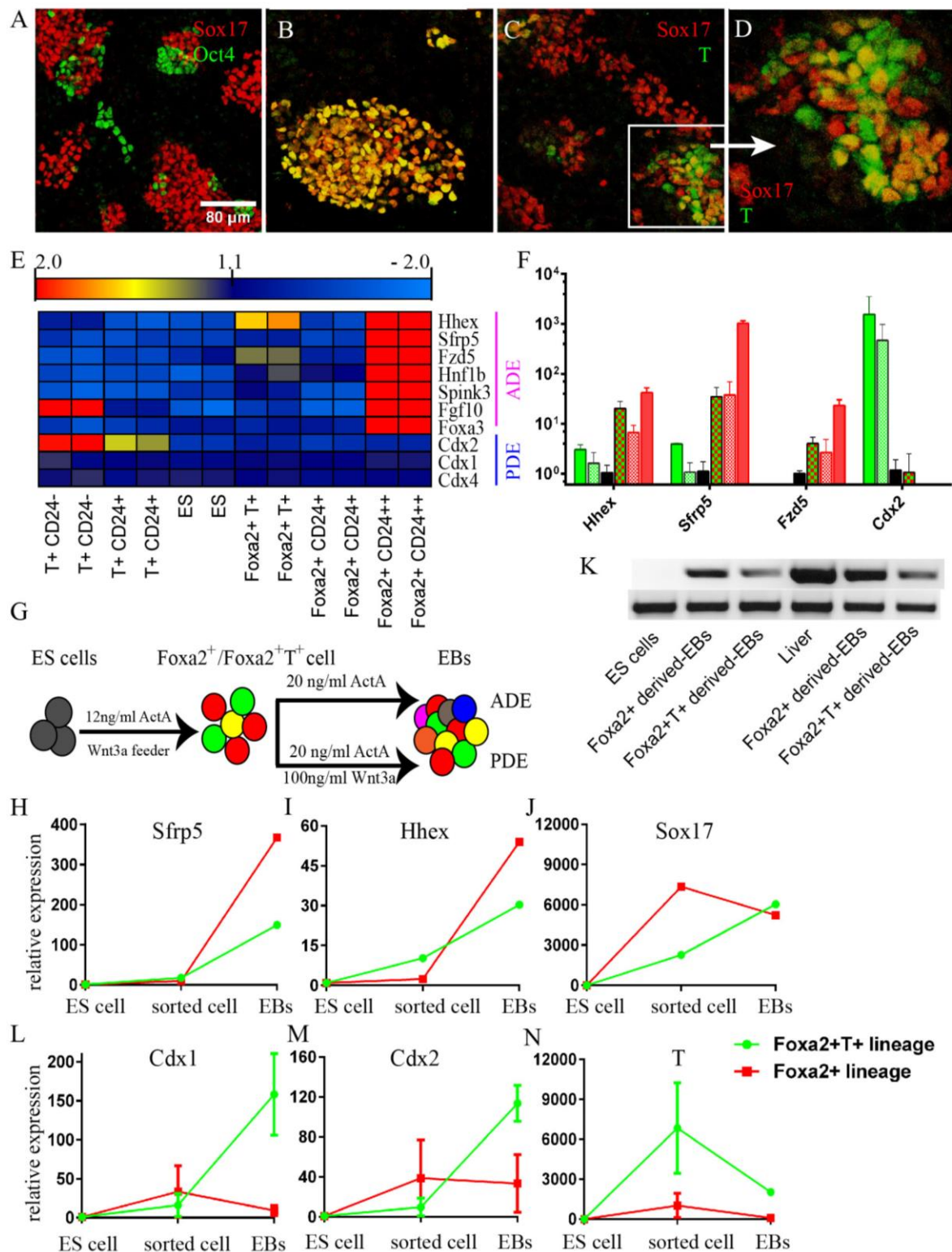
To address these questions, we first checked the expression pattern of the late endoderm differentiated cells with well-known endoderm markers. In line with previous data, we observed a high efficiency of endoderm induction at day 4 with the high expression of endoderm markers Foxa2 and Sox17 and low expression of pluripotent gene Oct4 (2.10 A and B). IHC showed that all Sox17<sup>+</sup> cells expressed Foxa2, but only few of them expressed T (Figure 2.10 C and D). In combination with our observations from the time-lapse imaging analysis described in Figure 2.6, where two main cell types Foxa2<sup>+</sup> cells and Foxa2<sup>+</sup>T<sup>+</sup> cells were observed, we assumed that both of Foxa2<sup>+</sup> cells and Foxa2<sup>+</sup>T<sup>+</sup> cells could give rise to Sox17<sup>+</sup> DE cells. This was supported by the mRNA profile data showing relative higher Sox17 mRNA level in Foxa2<sup>+</sup> cells and Foxa2<sup>+</sup>T<sup>+</sup> cells compared to the other populations (Figure 2.5 B).

Next we sought to analyze whether these distinct endoderm lineages may have different differentiation potential, as both Foxa2<sup>+</sup> cells and Foxa2<sup>+</sup>T<sup>+</sup> cells potentially could give rise to DE. Microarray analysis showed a unique signature of the Foxa2<sup>+</sup>CD24<sup>++</sup> subpopulation with the expression of a significant number of ADE genes, such as *Hex*, *Sfrp5*, *Fzd5*, *Hnf1b*, *Spink3*, *Fgf10*, and *Foxa3* (Figure 2.10 E). On the other hand, there is a minimal to no expression of the PDE markers *Cdx1*,

*Cdx2*, and *Cdx4* (Figure 2.10 E). Consistent with the microarray data, qPCR analysis showed that the *Foxa2*<sup>+</sup>*CD24*<sup>++</sup> population expressed the highest mRNA levels of ADE markers *Hhex* and *Sfrp5* and lowest levels of *Cdx2* transcript (Figure 2.10 F). These data indicate that the gene expression profiles of *Foxa2*<sup>+</sup>*CD24*<sup>++</sup> subpopulation more closely resembled ADE than PDE.

The data above suggests that these endodermal *Foxa2*<sup>+</sup> cells and *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells could form DE, but the *Foxa2*<sup>+</sup> cells favour the ADE fate. To better understand the fate of the *Foxa2*<sup>+</sup> cells and *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells, both ADE and PDE cultures were used to further differentiate these cells (Figure 2.10 G). In mice, the high Nodal environment of the PS promotes ADE fate; whereas Wnt signaling is essential for PDE induction (Sherwood et al., 2011; Thomas et al., 1998). Based on these previous studies, we sorted these *Foxa2*<sup>+</sup> and *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells at day 2, and re-cultured them aggregated as EBs in endoderm conditions supplemented with only ActA (20ng/ml) for another 3 days. We found that the ADE markers *Sfrp5* and *Hhex* were up-regulated in both lineages. Notably, the ADE markers were expressed much higher in *Foxa2*<sup>+</sup> cell-derived EBs than *Foxa2*<sup>+</sup>*T*<sup>+</sup> cell-derived EBs (Figure 2.10 H-J). The higher capability of *Foxa2*<sup>+</sup> cell to form ADE led us to hypothesize that these *Foxa2*<sup>+</sup> cell are specified for the ADE lineages. To investigate this, the culture conditions were switched to a previously published hepatic culture system (Gouon-Evans et al., 2006). The early hepatic marker AFP was expressed in both induced population, but it was significantly higher in the hepatoblast-like cells derived from *Foxa2*<sup>+</sup> population (Figure 2.10 K). These results indicate that the *Foxa2*<sup>+</sup> cells exhibit greater competence to differentiate into liver progenitors. Next, we also re-cultured these *Foxa2*<sup>+</sup> and *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells in a PDE condition supplemented with ActA (20ng/ml) and Wnt3a (100ng/ml) for another 3 days. *Foxa2*<sup>+</sup>*T*<sup>+</sup> cell-derived EBs exhibited significant higher levels of PDE marker expression when compared to *Foxa2*<sup>+</sup> cell-derived EBs (Figure 2.10 L-N). Taken together, these results suggest that *Foxa2*<sup>+</sup> cells are likely to contribute towards ADE and subsequent hepatic fate better than *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells, whereas *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells rather favour the PDE fate.





**Figure 2.10** cell fate analyses of mESCs–derived early endoderm progenitor cells

(A-D) Immunostaining of specific pluripotency (Oct4), endoderm (Foxa2, Sox17) and mesoderm (T) markers in 96 hr endoderm differentiated cells. Scale bars A - C: 80  $\mu$ m.

## 2.Results

---

(E-F) Microarray gene expression analysis of ADE and PDE genes in sorted populations. All the ADE genes are highly expressed in  $Foxa2^+CD24^{++}$  cells when come to the other sorted populations. These data are further confirmed by qPCR (B, C). Data represented by mean+SD.

(G) Schematic representation of the stepwise differentiation towards either ADE or PDE.

(H-K) Comparison of  $Foxa2^{+T^+}$  and  $Foxa2^+$  derives under ADE induction by q PCR analysis (H-J) and in hepatic culture conditions (K) by semi-PCR. The expression of ADE genes including *Sfrp5* (H) and *Hhex* (I) as well as the early hepatic marker *AFP* (K) are relatively higher expressed in  $Foxa2^+$  cells-derived EBs compared to the expression in  $Foxa2^{+T^+}$  cells-derived EBs.

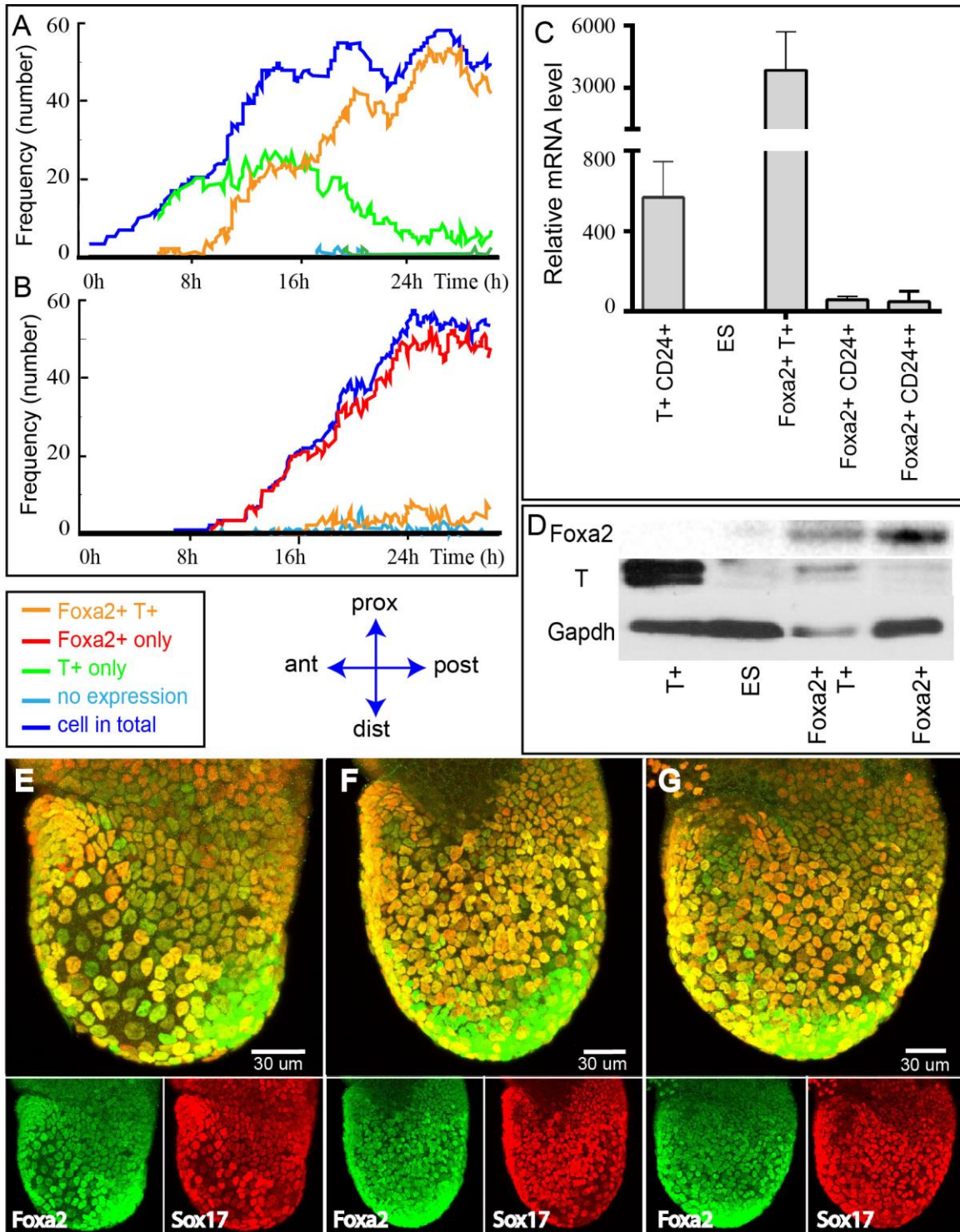
(L-N) Comparison of  $Foxa2^{+T^+}$  and  $Foxa2^+$  derives under PDE induction by q PCR analysis. The expression of PDE genes including *Cdx1* (L) and *Cdx2* (M) are relatively higher expressed in  $Foxa2^{+T^+}$ cells-derived EBs compared to the expression in  $Foxa2^+$  cells-derived EBs. Data represented by mean plus standard error of the mean (mean +SEM).



### 2.1.8 The Brachyury (T) function is not required for endoderm formation

Previous studies in mouse have revealed that T plays a very important role in mesoderm formation. The homozygous  $T^{-/-}$  embryos show defects in formation of PS and mesoderm, disruption in morphogenesis of mesoderm-derived structures (Gruneberg, 1958; Wilkinson, Bhatt, 1990; Wilson et al., 1995; Yanagisawa et al., 1981). However, no clear defects in endoderm formation were reported. Moreover, it is commonly thought that DE cells derive from a  $T^{+}$  progenitor, as T expression is switched on first in ESC endoderm differentiation cultures (Fehling, 2003; Kubo et al., 2004b). In contrast, in our single cell time-lapse imaging experiment, the formation of  $Foxa2^{+}$  cells is independent from  $T^{+}/Foxa2^{+}T^{+}$  cells (Figure 2.7). To analyse if the TF T is necessary for DE formation, we compared the onset of  $T^{+}$  cell,  $Foxa2^{+}$  cell, and  $Foxa2^{+}T^{+}$  cell by the frequency of positive cells produced in the time course of endoderm differentiation. Similar with our observation described above, the onset of T expression occurred from start of mESCs differentiation, and was switched on earlier than in other cell types (Figure 2.11 A and B). Interestingly, we found that the onset of  $Foxa2$  expression is similar in  $Foxa2^{+}$  cells and  $Foxa2^{+}T^{+}$  cells, suggesting that the up-regulation of  $Foxa2$  expression is not regulated by T (Figure 2.11 A and B). These findings lead us to hypothesize that T is not required for endoderm formation.

To functionally prove this hypothesis, we first investigated the T expression in the endoderm differentiated  $Foxa2^{+}$  cells at different time points. As expected, both RNA and protein levels of T were significantly expressed higher in  $T^{+}$  and  $Foxa2^{+}T^{+}$  subpopulations when compared to mESCs negative control, while they were barely expressed in  $Foxa2^{+}$  cells (Figure 2.11 C and D). To examine whether T is required or not in the formation of endoderm *in vivo*, homozygous and heterozygous T knock-out mouse lines were generated (Figure 2.1 B Fehling, 2003) and analyzed the expression of  $Foxa2$  and  $Sox17$  by IHC. As shown in Fig 2.11 E-G, no obvious defects were found in the formation of DE cells in homozygous and heterozygous mutant embryos (Figure 2.11 E-G), when compared to in WT littermates. Our data clearly show that T is not expressed in the  $Foxa2^{+}$  lineage that gives rise to DE and knock-out analysis reveals that T is not necessary for DE formation.



**Figure 2.11 T function is not required during endoderm formation**

(A, B) Representative traces show the frequency of Foxa2<sup>+</sup>, T<sup>+</sup>, Foxa2<sup>+</sup>T<sup>+</sup> cells emerge in the time course of endoderm differentiation. A represents the formation process of Foxa2<sup>+</sup>T<sup>+</sup> lineage cells, and B represents the formation of Foxa2<sup>+</sup> lineage cells.

(C, D) T is enriched in mesoderm differentiated T<sup>+</sup> cells and endoderm differentiated Foxa2<sup>+</sup>T<sup>+</sup> cells, which are quantified by q PCR(C) and western blot analysis (D). q PCR analysis was done in triplicate, and error bars represent the standard deviation.

---

(E-G) IHC of endoderm markers Foxa2 and Sox17 in WT (E), heterozygous (F), and homozygous (G) T knock-out mouse embryos. Compared to the Foxa2 and Sox17 expression in WT embryos, no defects is found in T knock-out mouse embryos. Scale bars E - G: 30  $\mu$ m.

### 2.1.9 Key regulators of classical EMT are not involved in endoderm formation

EMT occurs during gastrulation allowing cells to enter the PS region of the embryo to form either mesoderm or endoderm. The formation of mesoderm is considered to be a typical EMT process with a clear up-regulation of the expression of mesenchymal markers, as well as down-regulation of epithelial markers (Nakaya and Sheng, 2008). E-Cad was down-regulated when epiblast cells acquire a mesenchymal fate, migrate through the PS and form the mesodermal lineage. The reduction of E-Cad was mainly due to the activation of the Snail TF, which is induced by TGF- $\beta$  signaling and serves as repressor of E-Cad to induce the EMT process (Barrallo-Gimeno and Nieto, 2005; Peinado et al., 2007). In contrast, when cells acquire an epithelial fate and intercalate into the outside epithelium becoming DE cells, E-Cad was maintain during the epiblast-endoderm transition. Relatively little is known about the genetic regulatory network that controls EMT in the endoderm cells, especially in terms of the molecular mechanisms of initiation and execution of EMT in endoderm cells, and how these molecular mechanisms control cell adhesion, motility, invasion, survival and differentiation.

A variety of factors have been identified in EMT in mouse gastrulation. To examine that we compared the EMT characteristic expression profile among these sorted subpopulations. As expected, the well-known EMT TFs Zeb1, Zeb2, Snail1, Snail2, and Twist1 were markedly up-regulated in T<sup>+</sup> lineage subpopulations, accompanied with the up-regulation of mesenchymal markers Fibronectin1 (Fn1), Vimentin (Vim), and N-Cad. There was also a considerable decrease in the expression of epithelial markers, such as E-Cad, CD24, and Cldn6. The most notable mesenchymal feature was found in T<sup>+</sup>CD24<sup>-</sup> subpopulation, which represent differentiated mesoderm (Figure 2.12 A and B). In contrast, the Foxa2<sup>+</sup> lineage exhibited a strong induction of epithelial marker (Figure 2.12 A and B). Of note, the expression of CD24a in the sorted subpopulations fits our FACS data, and shows clear low and high expression in mesenchymal mesoderm cells and epithelial endoderm cells (Figure 2.11 A and B).

Interestingly, we found all the EMT TFs were not up-regulated in the Foxa2<sup>+</sup> lineage (Figure 2.12 A and B).

*Snail1* is one of the key inducers of EMT in gastrulating mouse embryos (Barrallo-Gimeno and Nieto, 2005; Lim and Thiery, 2012). Embryos null of *Snail1* exhibit defects in gastrulation and in the EMT required for generation of the mesoderm cell layer (Carver et al., 2001). To investigate the cellular distribution of *Snail1* during gastrulation, we used whole-mount IHC with antibodies to *Snail1*, T and Foxa2. Interestingly, we observed that *Snail1* was only expressed in proximal domain of the posterior epiblast, where it was co-expressed with T; whereas, the distal epiblast cells up-regulate Foxa2 expression, but were absent in *Snail1* expression (Figure 2.12 C). To confirm these *in vivo* observations, we closely examined the expression of *Snail1* in the sorted populations. As shown in Fig 11E, *Snail1* was only enriched in mesoderm T<sup>+</sup> lineage, but was not detectable in the Foxa2<sup>+</sup> lineage (Figure 2.12 D).

These results provide first evidence that mesoderm formation involves a typical EMT process. In contrast, the lack of EMT TFs expression in Foxa2 positive lineage indicates that the key regulators of classical EMT are not involved in endoderm formation. Thus we hypothesize that epithelial columnarepiblast cells directly give rise to squamous epithelial endoderm cells in a process that should be called epithelial-to-epithelial transition (EET).

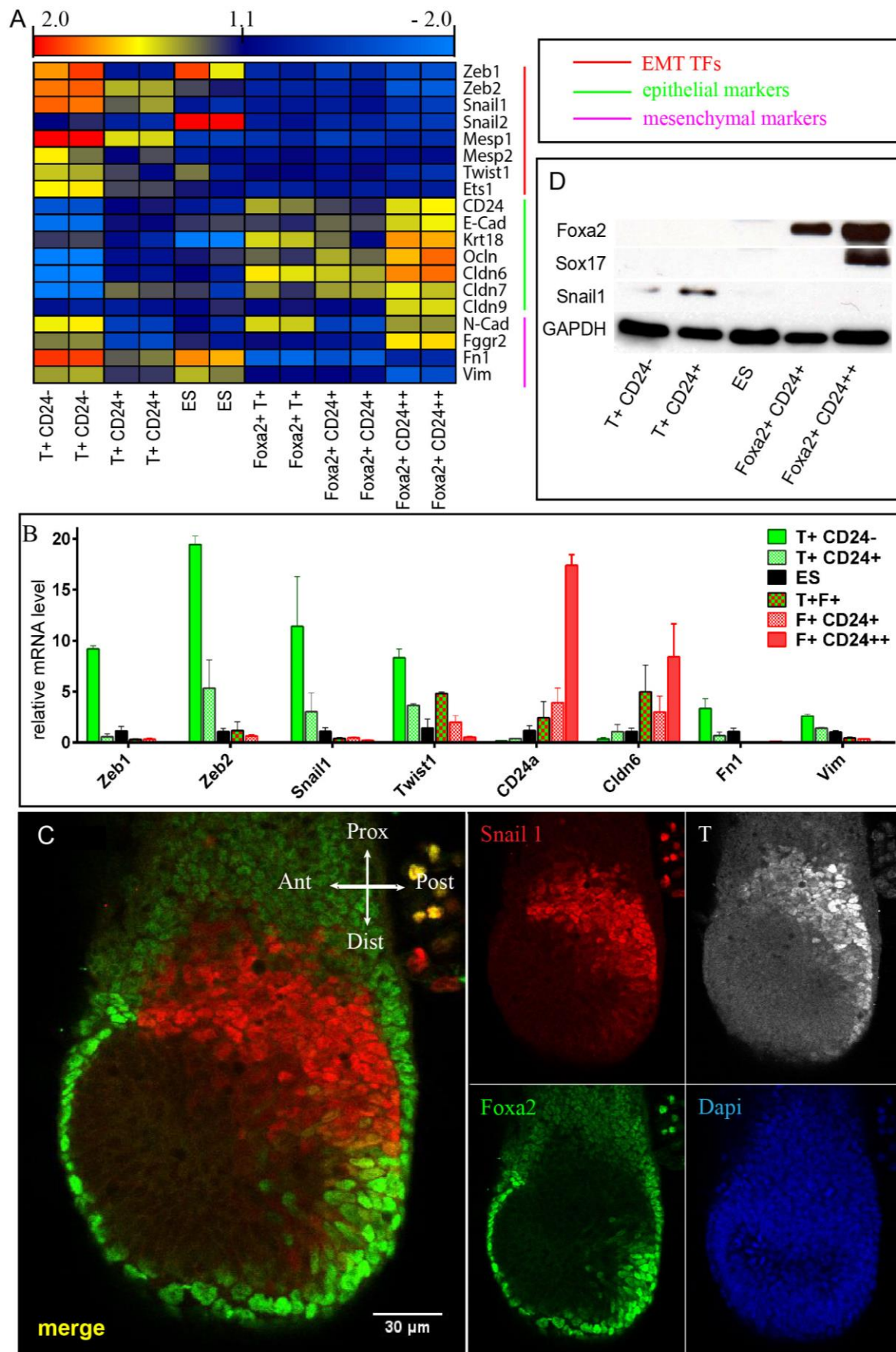


Figure 2.12 Analysis of EMT associated genes in endoderm and mesoderm progenitors

## 2.Results

---

(A, B) mRNA profile shows that the expression of EMT associated TFs, such as Zeb1, Zeb2, Snail and Snail2, as well as mesenchymal markers Fn1 and Vim, is up-regulated in T<sup>+</sup> lineage cells; whereas the expression of epithelial markers, such as E-Cad, Ocln, Cldn6 and Cldn7, is up-regulated in Foxa2<sup>+</sup> lineage cells (A). These data are further confirmed by quantitative real time RT-PCR (B). Two biological replicates were performed to generate error bars (standard deviation). (C, D) The expression pattern of Snail1 is correlated with the expression of T, which is confirmed by IHC in gastru-stage mouse embryos (C) and western blot analysis in sorted populations (D). Scale barC: 30  $\mu$ m.

## **2.2 miR-335 promotes mesendodermal lineage segregation and shapes a TF gradient in the endoderm**

Parts of this dissertation were recently published in the peer-reviewed journal *Development*. Incorporation of the publication is in agreement with the journal and the supervisors of the dissertation.

**Yang D**, Lutter D, Burtscher I, Uetzmann L, Theis F, Lickert H (2014). miR-335 promotes mesendodermal lineage segregation and shapes a transcription factor gradient in the endoderm, *Development* 141(3) 514-25.

### 2.2.1 Identification of miRNAs regulating mESCs differentiation

miRNAs have been shown to play important roles during embryonic development and in stem cell biology (Song and Tuan, 2006). *Dicer*-null mESCs exhibit low proliferation rates, loss of pluripotency, and defective differentiation (Kanellopoulou et al., 2005; Murchison et al., 2005), suggesting that miRNAs are required in maintenance of ESC populations and pluripotency. On the other hand, by targeting some key pluripotency genes, such as *Oct4*, *Sox2*, and *Nanog*, miRNAs have also been shown to promote ESCs differentiation (Tay et al., 2008; Xu et al., 2009). These studies indicate the essential role of miRNAs in mESCs maintenance and differentiation. However, relatively little is known about how miRNAs regulate lineage specification, especially endoderm formation during gastrulation.

To identify differentially expressed miRNAs in pluripotent mESCs and differentiated endoderm, which might be important for endoderm development, we used a *Foxa2*-Venus fusion (FVF) and *Sox17*-mCherry fusion (SCF) mESCs line (Burtscher et al., 2012, 2013). We then isolated RNA from fifteen samples using five different time points of endoderm differentiation, and profiled mRNA and miRNA expression by microarray analysis. The quality of the mESCs differentiation was confirmed by IHC with specific markers. Qualitative (Figure 2.13 A) and quantitative analyses (Figure 2.13 B) revealed that the endoderm TF *Foxa2* was up-regulated from onset of differentiation and reached peak levels between day 2 and day 3 of differentiation. The other endoderm TF *Sox17* was observed at day 2 and reached peak levels the following days of differentiation; whereas, the expression of pluripotent marker *Oct4* was down-regulated during differentiation.

In order to discover factors that regulate the differentiation from one stage to the next, the comparison of mRNA profile was done by Limma t-test and Benjamini-Hochberg multiple testing correction (FDR<10%). We obtained lists of genes with at least 1.5-fold changes in expression level. As shown in Table 2.3, there were 255 and 74 differentially expressed mRNA probes in the comparisons of 12 hr VS 0 hr, 24hr VS 12 hr of differentiated cells, separately (Table 2.3). Mesendoderm genes including *T*, *Foxa2*, *Gsc*, and *Eomes*, were up-regulated shortly after starts of endoderm induction (Figure 2.13 C). The peaks of differentially expressed mRNA probes were among 48 hr to 72 hr of differentiation accompanied by an up-regulation of critical endoderm



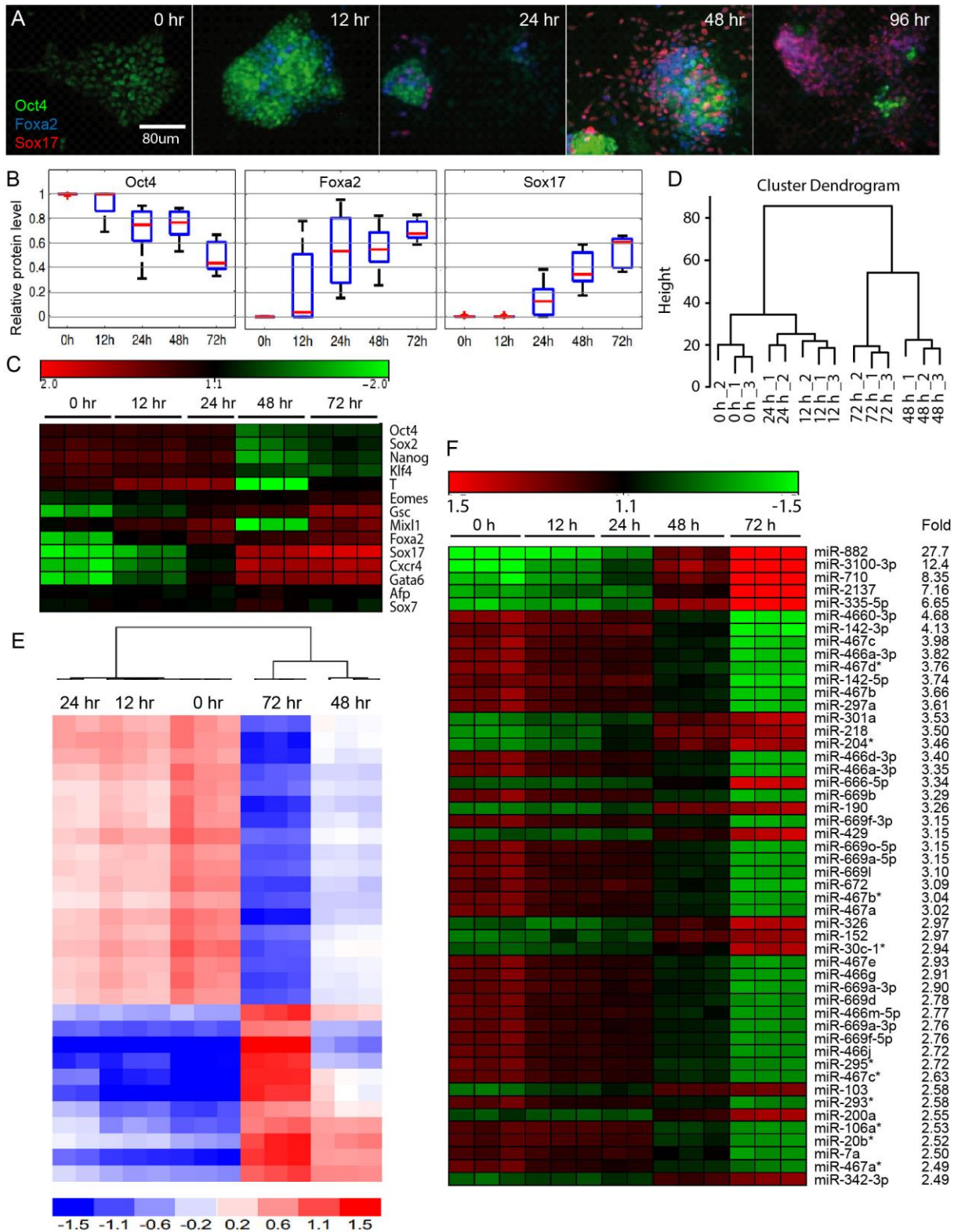
---

genes *Sox17* and *Cxcr4*, and a significant reduction of pluripotent genes expression including *Oct4*, *Sox2* and *Nanog* (Table 2.3 and Figure 2.13 C). In line with the heat map analysis, two main groups were separated by hierarchical cluster analysis. The first comprising samples of 12 hr and 24 hr differentiated cells, which display early stages of differentiation and clustered together with undifferentiated mESCs; whereas, the samples from 48 hr and 72 hr differentiated cells were clustered together and displayed late stage of differentiation (Figure 2.13D).

As we aimed to identify miRNAs potentially involved in differentiation of mESCs, the miRNA probes from each differentiation stage to the following stage were then analyzed by ANOVA analysis. Interestingly, we observed that most of the differentially expressed miRNA probes were up-/down-regulated at the late stage of differentiation (Table 2.3). Consistent with the hierarchical cluster analysis above based on the mRNA expression, the miRNA analysis also separated the samples into two main clusters; early and late progenitors (Figure 2.13 E). To identify miRNAs that display temporal expression regulation, we compared miRNA profile among different stages. Figure 2.13 F shows the top 50 differentially expressed miRNAs in mESCs during the differentiation time course (Figure 2.13 F). Among the top regulated miRNAs were miR-882, miR-3100-3p, miR-710, miR-2137, and miR-335-5p, which indicate they may have important function in the mESCs maintenance and differentiation process.

To identify specific miRNAs that involved in mesendoderm lineage specification and endoderm formation, the differentially expressed miRNAs that potentially target mesendoderm TF *Foxa2* and endoderm TF *Sox17*, were selected. Among them, miR-335 was especially interesting, because it is an intronic miRNA embedded in intron 2 of the mesoderm-specific transcription factor (*Mest*) mRNA transcript (Figure 2.14 A), and is predicted to target both TFs; *Foxa2* and *Sox17* (Figure 2.14 B-E). These findings let us to hypothesize that the mesoderm-specific miR-335 could regulate mesendoderm lineage specification and segregation by targeting endoderm-specific TFs, *Foxa2* and *Sox17* (Figure 2.14 D).

## 2.Results



**Figure 2.13 Analysis of mRNA and microRNA profiling during endoderm differentiation**

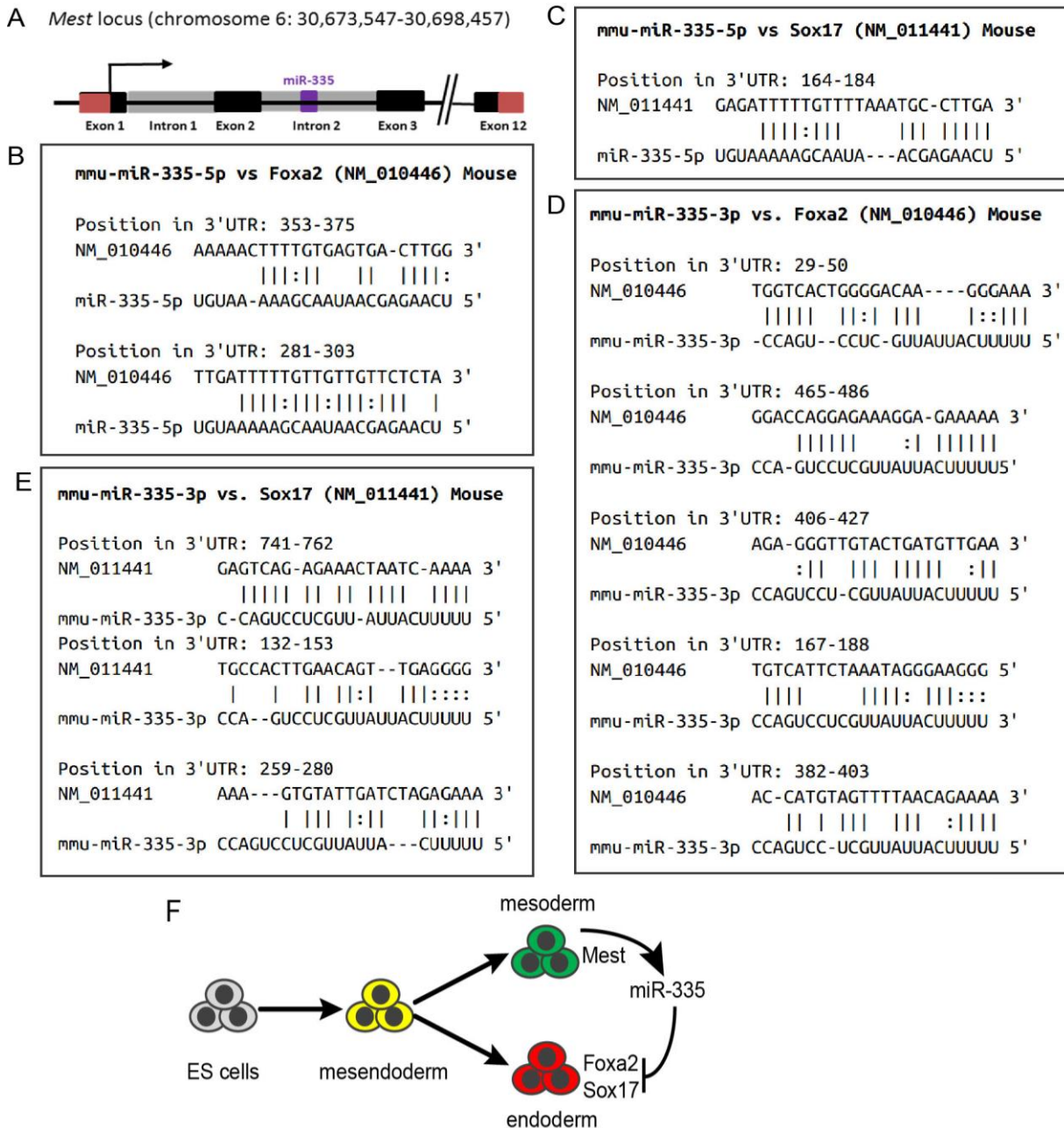
(A-B) Qualitative and quantitative analyses of time course of mesendoderm and endoderm differentiation by immunostaining with stage specific markers. Scale bar A: 80  $\mu$ m.

(C) Microarray gene expression heat map of mESCs endoderm differentiation by stage specific markers.

(D) Cluster analysis of mESCs and differentiated endodermal cells.

(E) Differentially expressed miRNAs (fold changes  $> \pm 3.0$ ) in all five samples were presented by hierarchical clustering analysis.

(F) Top 50 differentially expressed miRNAs comparing 72 h differentiated endodermal cells with mESCs.



**Figure 2.14 miR-335 is a potential regulator of mesendoderm development**

(A) *Mest* locus showing the genomic sequence encoding miR-335, which is located in the second intron of *Mest*.

(B-E) All predicted miR-335-5p and miR-335-3p binding sites for *Foxa2* and *Sox17*. Predictions were calculated using RNA22 allowing for 2 unpaired bases in seed sequence.

(F) Proposed model of miR-335 in mesendoderm lineage specification.

Significant probe sets	mRNA	miRNA
12 hr VS 0 hr	255	1
24 hr VS 12 hr	74	1
48 hr VS 24 hr	3104	34
72 hr VS 48 hr	1386	67

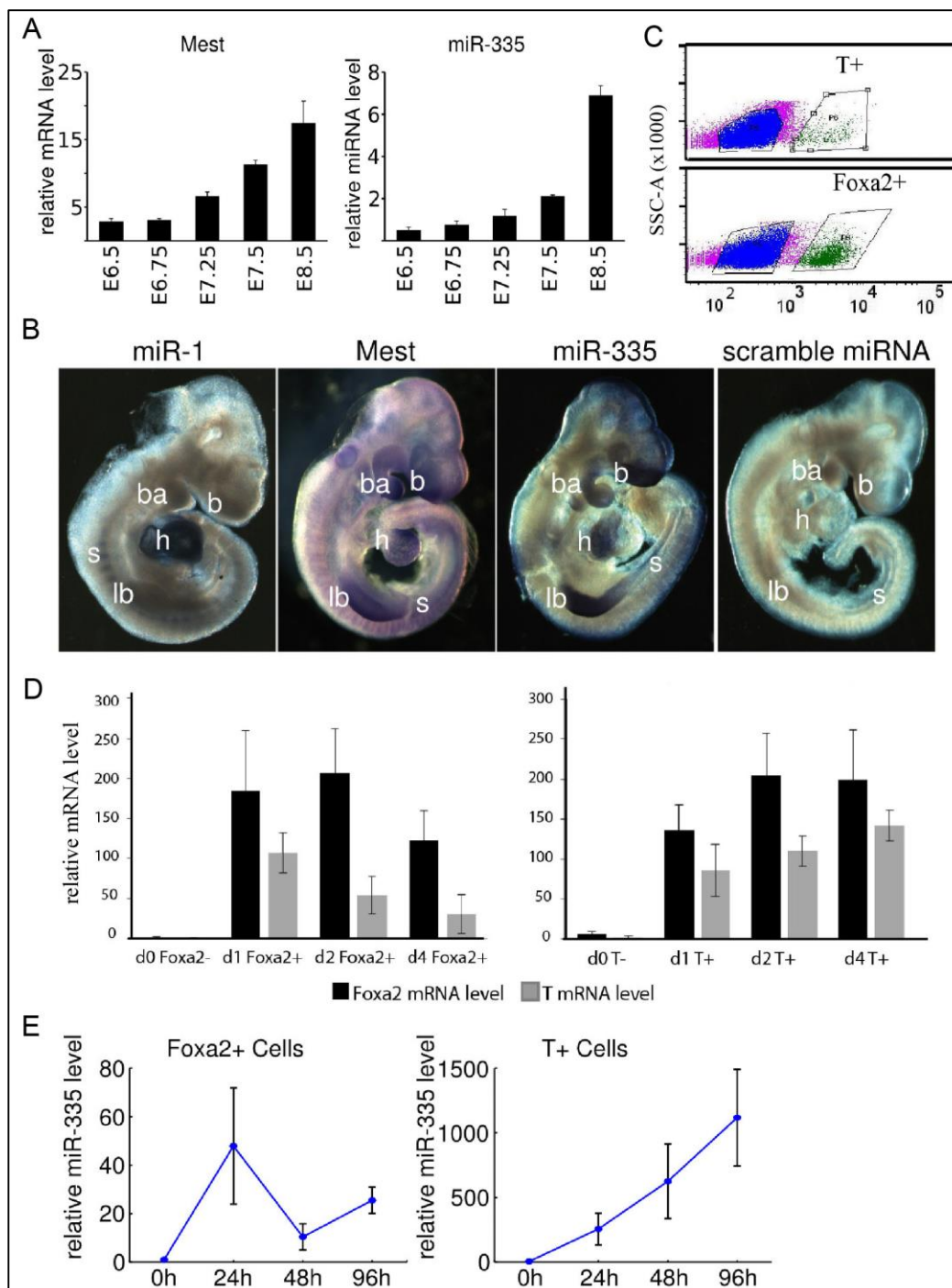
**Table 2.3 Number of mRNA and miRNA differentially expressed during mesendoderm and endoderm differentiation (fold change  $\geq 1.5$ ).**

### 2.2.2 Spatio-temporal expression of miR-335 during early development

miR-335 was identified in mammalian neurons (Kim et al., 2003) and has been well studied in cancer and identified as a candidate tumour suppressor (Gao et al., 2014; Png et al., 2011; Xiong et al., 2013; Xu et al., 2012). To date, relatively few studies have investigated miR-335 function in early mouse development and mESCs differentiation.

miR-335 is an intronic miRNA and according to previous findings should be co-expressed with its host gene (Ronchetti et al., 2008). Using qPCR we confirmed that miR-335 expression level correlates with *Mest* expression during mouse development at E6.5-8.5 (Figure 2.15 A). Furthermore, we confirmed that miR-335 and *Mest* mRNA are co-expressed in a tissue-specific manner in the brain, branchial arches, heart, limb buds, and somites at E9.5 using whole-mount *in situ* hybridization (*ISH*) with anti-sense probes against *Mest* and miR-335 (Figure 2.15 B). As previously shown that miRNAs regulate pathways in a tissue-specific manner (Kowarsch et al., 2011), we further analyzed the cell type-specific *Mest* and miR-335 expression during ESC differentiation. Using two different ESC differentiation paradigms, we isolated the *Foxa2*<sup>+</sup> mesendoderm and endoderm lineage using the FVF mESCs line under endoderm conditions, as well as the T<sup>+</sup> mesendoderm and mesoderm populations using the T-GFP mESCs line under mesoderm conditions in a time-course experiment using FACS (Figure 2.15 C) and analyzed the miR-335 expression levels by qPCR (Figure 2.15 E). This revealed that miR-335 is transiently up-regulated approximately 40-fold in the *Foxa2*<sup>+</sup> mesendoderm progenitors after 24 h of differentiation. In contrast, miR-335 steadily increases in a four day time-course up to over a 1000-fold in the T<sup>+</sup> mesendoderm and mesoderm lineage (Figure 2.15 E),

consistent with the strong mesoderm-specific expression in heart and somites, but not in the gut endoderm *in vivo* (Figure 2.15 B). Together the spatio-temporal expression suggests that miR-335 functions transiently in the Foxa2<sup>+</sup> endoderm progenitors and later during mesoderm formation.



**Figure 2.15 Co-expression of Mest/miR-335 in mESCs and mouse embryos**

(A) The temporal expression levels of miR-335 correlate with those of its host coding-gene Mest in mouse embryos at E6.5-8.5.

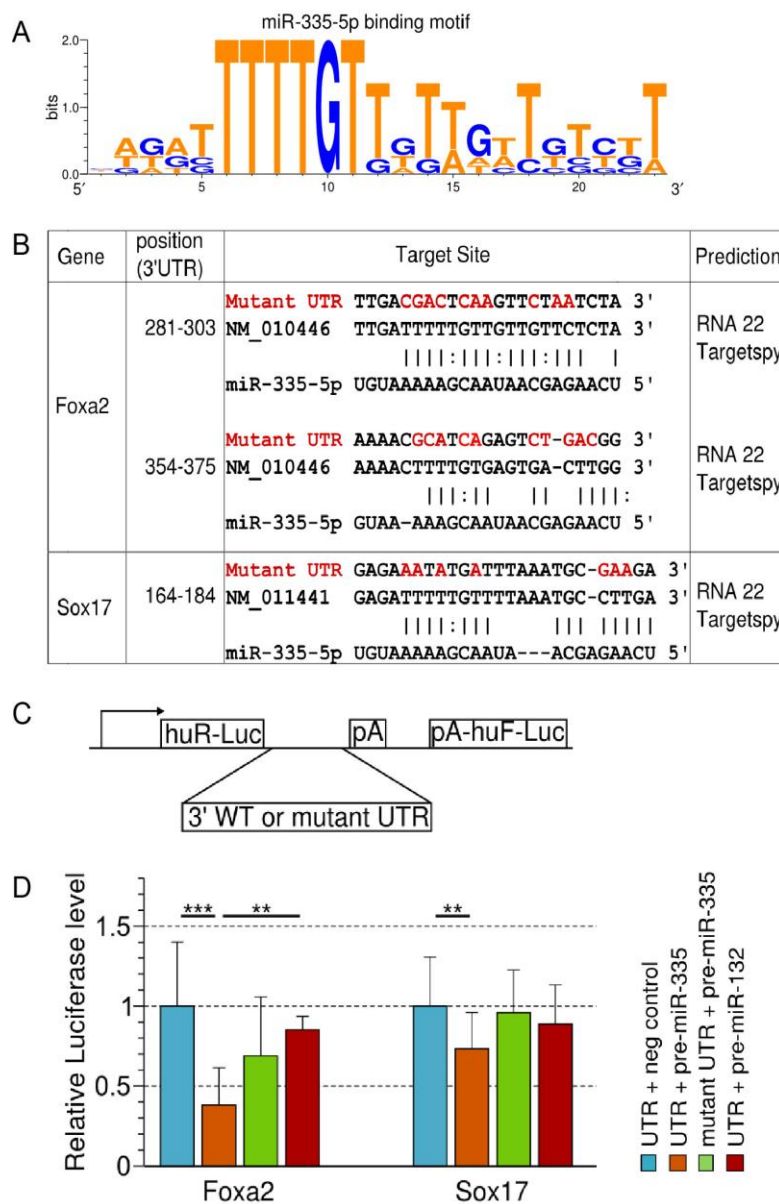


## 2.Results

(B) Whole-mount *ISH* shows co-expression of *Mest* and miR-335 in the brain (b), heart (h), limb bud (lb), and branchial arches (ba) at E9.5. The heart and somites (s) specific expressed miR-1 served as a positive control (Kloosterman et al., 2006), whereas a scrambled miR-335 served as a negative control.

(C, D) *Foxa2*<sup>+</sup> or *T*<sup>+</sup> cells were sorted at day 1, day 2 and day 4 under endoderm or mesoderm differentiation conditions and analysed by qPCR, day 0 mESCs used as a control.

(E) The level of miR-335 in the *Foxa2*<sup>+</sup> lineage is transient upregulated during endoderm differentiation (left graph), whereas miR-335 highly accumulates in the *T*<sup>+</sup> lineage during mesoderm differentiation (right graph).



**Figure 2.16** *Foxa2* and *Sox17* are targets of miR-335

(A) Target site consensus sequences based on previously described miR-335 target sites. To generate conserved miR335 target site consensus motifs we combined a number of available validated and predicted miR-335-3p and miR-335-5p target sequences from mouse and human 3'UTR sequences.

(B) Summary of miR-335-5p target sites in the 3'-UTR of Foxa2 and Sox17. Nucleotides shown in red indicate changes in the mutant 3'-UTR.

(C) Scheme of the reporter constructs used for target validation. huR-Luc, humanized Renilla luciferase; huF-Luc, humanized firefly luciferase; pA, polyadenylation signal; WT, wild type.

(D) miR-335 directly represses its targets in a luciferase assay in HEK293T cells. Renilla luciferase activity was assayed 40 hr after transfection and the values were normalized to the activity of firefly luciferase encoded in the same vector.

### 2.2.3 miR-335 directly targets Foxa2 and Sox17

miRNAs play important regulatory roles in many developmental processes by binding to 3'-UTRs of target mRNAs and leading either to mRNA degradation or repression of protein translation (Bartel et al., 2004; Guo et al., 2010b). As miR-335 has predicted target recognition sites in the 3'-UTRs of Foxa2 and Sox17 (Figure 2.14 B-E), we first investigated whether it directly targets these TFs.

A position weight matrix for miR-335 binding motif using previously described miR-335 target mRNA binding sites (Miranda et al., 2006; Tavazoie et al., 2008) was generated by Dr. Domink Lutter (Table 2.4). Scanning the 3'-UTRs identified one and two miR-335-5p binding motives for the Sox17 and Foxa2 mRNA; respectively (Figure 2.14 B and C); whereas miR-335-3p shows five and three binding motives for the Sox17 and Foxa2 mRNA, respectively (Figure 2.14 D and E). To investigate whether Foxa2 and Sox17 can be directly targeted by miR-335, we generated dual Renilla-firefly luciferase reporters that contain either the WT 3'-UTRs or mutant 3'-UTRs with a modification at the highly conserved positions (Figure 2.16 A - C). The luciferase reporters were co-transfected with miR-335-5p into Hek 293T cells. This revealed that both Foxa2 and Sox17 3'-UTR are significantly down-regulated ( $p$ -value < 0.01) by miR-335, which can be rescued to different degrees following mutation of the miR-335-5p binding site (Figure 2.16 D). In contrast, miRNA-132, which has no target site of Foxa2 and Sox17 mRNA, did not affect the reporter activities. In summary, full rescue of the Sox17 3'UTR reporter activity by mutation of the miR-335-5p binding motif suggests direct regulation, whereas additional miR-

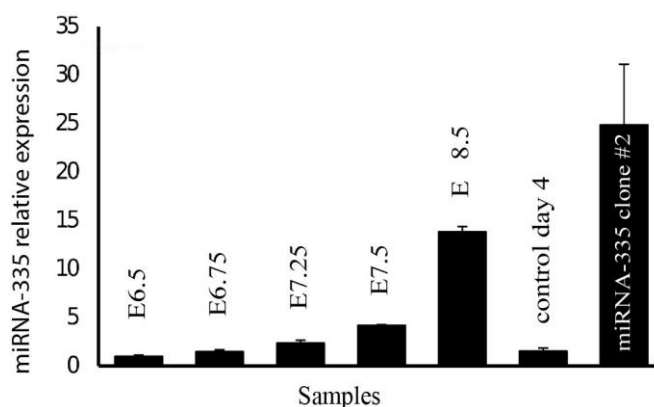
## 2. Results

335-3p or non-consensus binding motives in the Foxa2 3'UTR might be targeted by miR-335-3p.

Consensus Sequence for miR-335-5p:	Consensus Sequences for miR-335-3p:
>utr 3MMUR058050 Foxa2 TTTGATTTTTGTTGTTGTTCTCTA	>NM 010446 Foxa2 GGACCAGGAGAAAGGAGAAAAA
>utr 3MMUR058050 Foxa2 AAAAACTTTTGTGAGTGACTTGGT	>NM 010446 Foxa2 AGAGGGTTGTA CTGATGTTGAA
>utr 3MMUR052152 Sox17 CGAGATTTTTGTTTTAAATGCCTT	>NM 009238 Sox4 GAGGCAGGAGAGGAGAGAGGGA
>utr 3HSAR026431 SOX4 CCTTGGTTTTGTTTTATTTTGCTT	>NM 009238 Sox4 AGGGGAGCATTGGCATGGAGAA
>utr 3HSAR022836 SOX17 CTGGGTTTTGTTGTTGCTGTTGT	>NM 003107 SOX4 TACAGGGGCAGTCAGTGGAGGG
>utr 3HSAR023247 FOXA2 TGATTTTTTTGTTGTTGTTGTTCT	>NM 003107 SOX4 GGGCCGGGGGGGTAGGAGAGG
	>NM 011441 Sox17 GAGTCAGAGAACTAATCAAAA

**Table 2.4**  
**Target Sequences**

(Binding motifs were generated using WebLogo).



**Figure 2.17**  
**Comparison of embryonic and overexpressed miR-335 levels measured by qPCR**

### 2.2.4 miR-335 overexpression represses endoderm formation

Cell fate decisions of mESCs differentiation are controlled by signaling pathways via the activation or repression of lineage-specific genes. Give the role of miRNAs in fine tuning of gene expression; miRNAs are important for the regulation of cell fate decisions by targeting lineage-specific genes. According to our previous results described above (Figure 2.15 and 2.16), we assumed that mR-335 could serve as a mesendoderm cell fate regulator by inhibiting Foxa2- and Sox17-mediated endoderm formation and cause a compensatory shift in fate towards mesoderm.

Next we studied the function role of miR-335 in modulating endoderm differentiation by a gain-of-function (GOF) approach. To assess the long-term of miR-335, several stable ESC lines that constitutively express miR-335 were generated (Uetzmann,



2009). Followed with the RNA polymerase II-driven human ubiquitin C promoter to co-express miR-335 from a modified intronic miRNA-155 precursor, an exon coding for histone 2B-cyan fluorescent reporter protein (H2B-CFP) was inserted (Chung et al., 2006). Using this bicistronic vector expression system, miR-335 expression was correlated to H2B-CFP fluorescent reporter activity (Yang et al., 2014a).

We used two independent mESCs lines with medium (miR-335 #1) and high (miR-335 #2) H2B-CFP reporter activity to analyze a dose-dependent effect of miR-335 on mESCs differentiation. We compared the miR-335 overexpression (miR-335 #2) to endogenous miR-335 levels measured in embryos to assure that the overexpression is in a physiological range (Figure 2.17). Qualitative (Figure 2.18 A) and quantitative analyses (Figure 2.18 B) revealed that endoderm differentiation was blocked by miR-335 overexpression in a concentration-dependent manner. Upon miR-335 GOF, Oct4<sup>+</sup> pluripotent ESC colonies remained round in shape and only a few flattened Foxa2<sup>+</sup>Sox17<sup>+</sup> DE cells appeared at the edge of the colonies (Figure 2.18 A). Western blot analysis confirmed these results and revealed that Foxa2-mediated mesendoderm induction at 48 hr occurred normal, while further differentiation into the Foxa2<sup>+</sup>Sox17<sup>+</sup> DE lineage was strongly reduced (Figure 2.18 C). Even after 96 hr Oct4 protein levels remained high, suggesting that differentiation was blocked at the Oct4<sup>+</sup>Foxa2<sup>+</sup> mesendoderm progenitor cell stage (Figure 2.18 C). Next, the effect of miR-335 GOF in completely ESC-derived mouse embryos *in vivo* was analyzed (Nagy et al., 1993; Tam and Rossant, 2003). As shown in Figure 2.18 D, the Foxa2<sup>+</sup>Sox17<sup>+</sup> DE cells were nicely formed at E7.5 in wt control mouse embryos. In contrast, tetraploid (4n) embryo  $\leftrightarrow$  miR-335#2 ESC aggregation chimera implanted normally and proceeded to the mesendoderm stage (Figure 2.18 D); however, hardly any DE cells were formed upon miR-335 GOF. These *in vivo* results directly reflect our mESCs differentiation results (Figure 2.18 A-C).

To find whether mesoderm or ectoderm formation is affected by miR-335 GOF, we performed EB differentiation experiments. Using the ESC clone #2 that expresses high levels of miR-335, we observed a marked decrease of the endoderm-specific marker Sox17 expression, but the expression of mesoderm-specific marker T and ectoderm-specific marker Sox1 seems unaffected (Figure 2.18 E-H).

In summary, these data suggest that miR-335 GOF blocks DE differentiation at the Foxa2<sup>+</sup> mesendoderm progenitor stage *in vitro* and *in vivo*.

2.Results

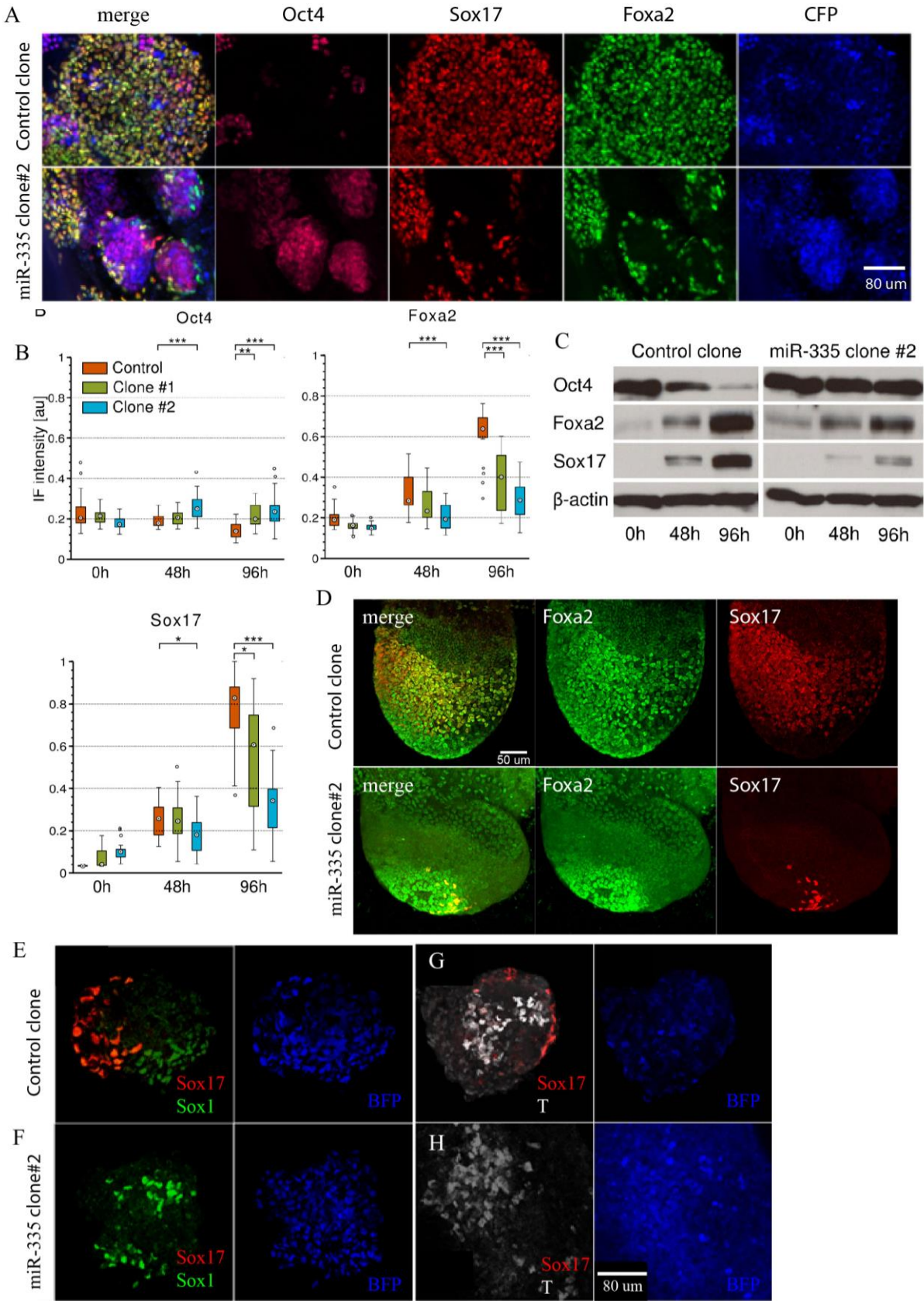


Figure 2.18 miR-335 overexpression represses endoderm differentiation

(A-C) During endoderm differentiation, both Foxa2 and Sox17 were significantly down-regulated by over-expression of miR-335 at medium and high levels, when compared to a control clone, as

shown by immunostaining (A), quantitative analysis of immunofluorescence (B) and Western blot analysis of Foxa2, Sox17 and Oct4 in differentiating mESCs (C). Scale barA: 80  $\mu$ m.

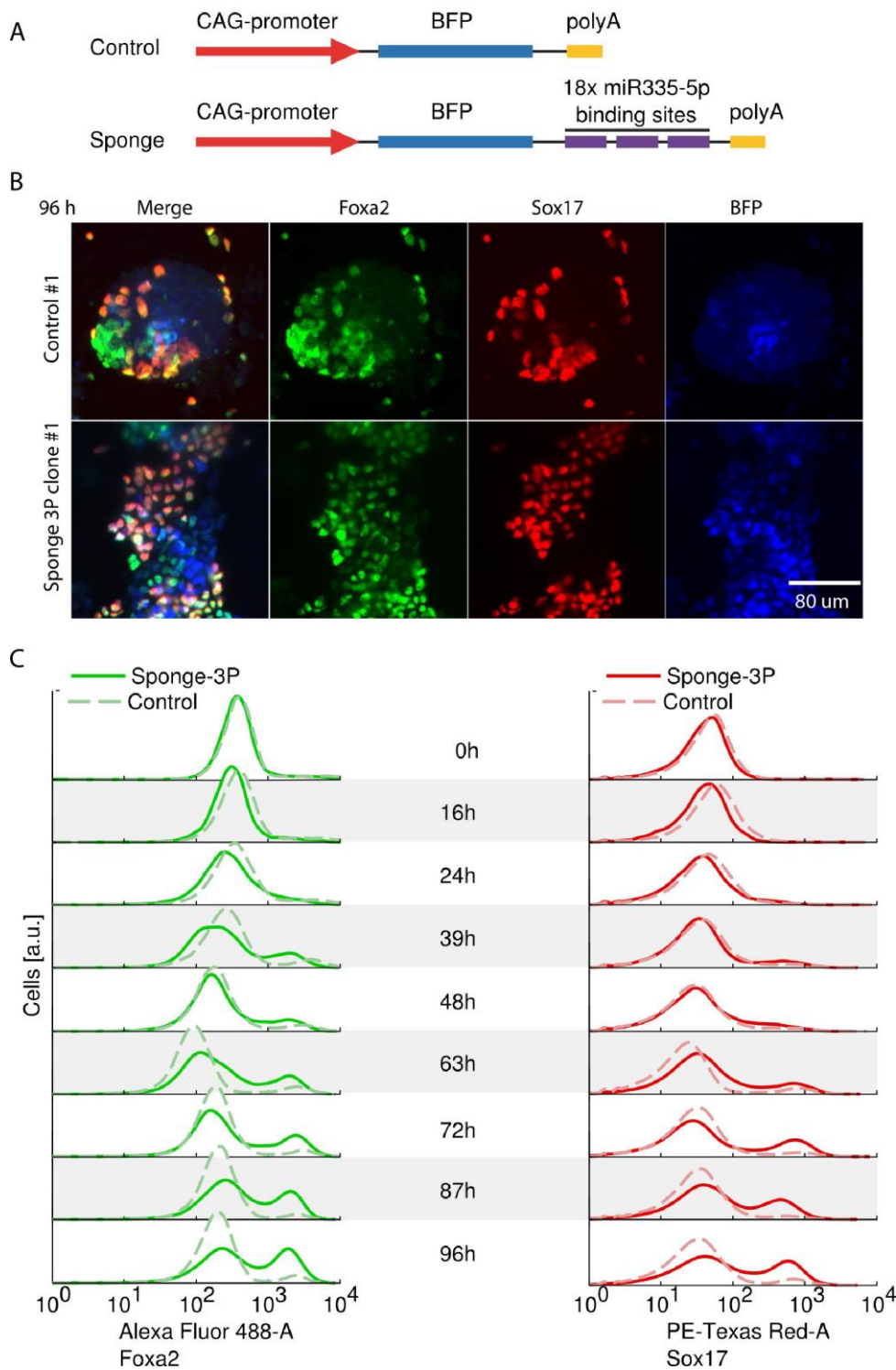
(D) IHC of completely mESCs-derived mouse embryos at gastrulation stage at E7.5. Foxa2 and Sox17 expression was only found in newly formed endoderm cells in the anterior primitive streak region, whereas Foxa2 and Sox17 were strongly suppressed in older endoderm cells in the anterior and lateral regions of miR-335 overexpressing embryo when compared to the wt control. Anterior to the left, distal to the bottom. Scale bar: 50  $\mu$ m.

(E-H) IHC of mESCs-derived EBs with specific mesoderm marker T, ectoderm marker Sox1, and endoderm marker Sox17. BFP was used to monitor the expression of miR-335. Sox17 expression was decreased, while T and Sox1 were not affected in miR335 clone #2 when compared to the control clone. Scale barE-H: 80  $\mu$ m.

### 2.2.5 Knock-down of miR-335 leads to an increase of endoderm differentiation

To gain further insight into the physiological function of miR-335 during mesendoderm formation, we examined the effect of miR-335 loss-of-function (LOF) on differentiation progression using competitive inhibition by overexpression of a sponge construct. For this purpose, a sponge-3P construct was generated by using a CMV-enhancer  $\beta$ -Actin promoter (CAG) to express H2B-blue fluorescent protein (H2B-BFP) followed by a 3'-UTR with 18 miR-335-5p binding motives. In contrast, the control construct was generated by the CAG to express H2B-BFP alone (Figure 2.19 A). Three different ESC sub-clones to constitutively express the control or sponge-3P construct from the FVF, SCF mESCs line were generated for live-cell analysis on single cell level. Both, FVF and SCF knock-in reporter mESC lines utilize the endogenous 3'-UTR of the Foxa2 and Sox17 mRNA and act as miRNA sensors for direct analysis of TF levels.

The analyses of FVF and SCF reporter activity by IHC (Figure 2.19 B) and quantitative single cell FACS analysis (Figure 2.19 C) using three independent mESCs clones expressing either sponge-3P construct or corresponding control in a time-course experiment suggested that miR-335 LOF led to an increase in Foxa2<sup>+</sup>Sox17<sup>+</sup> cells after 96 hr of DE induction. Together, these results suggest that endogenous miR-335-5p activity blocks Foxa2 and Sox17 translation, which can be specifically released by competitive inhibition for enhanced endoderm formation.



**Figure 2.19 Loss of miR-335 leads to an increase in endoderm formation**

(A) Schematic representation of the control and sponge-3P expression vectors.

(B) IHC analysis of differentiated ESCs shows significantly increased number of Foxa2<sup>+</sup> and Sox17<sup>+</sup> in sponge vector expressing clones compared to control clones. Scale bar: 80  $\mu$ m.

(C) FACS analysis of a time course experiment over 96 h endoderm differentiation shows increase of Foxa2 and Sox17 double positive population in sponge-3P expressing clones. Pooled FACS data of three control and three sponge clones was analyzed.

---

### 2.2.6 Mathematical modeling predicts miR-335 function in TF gradient formation

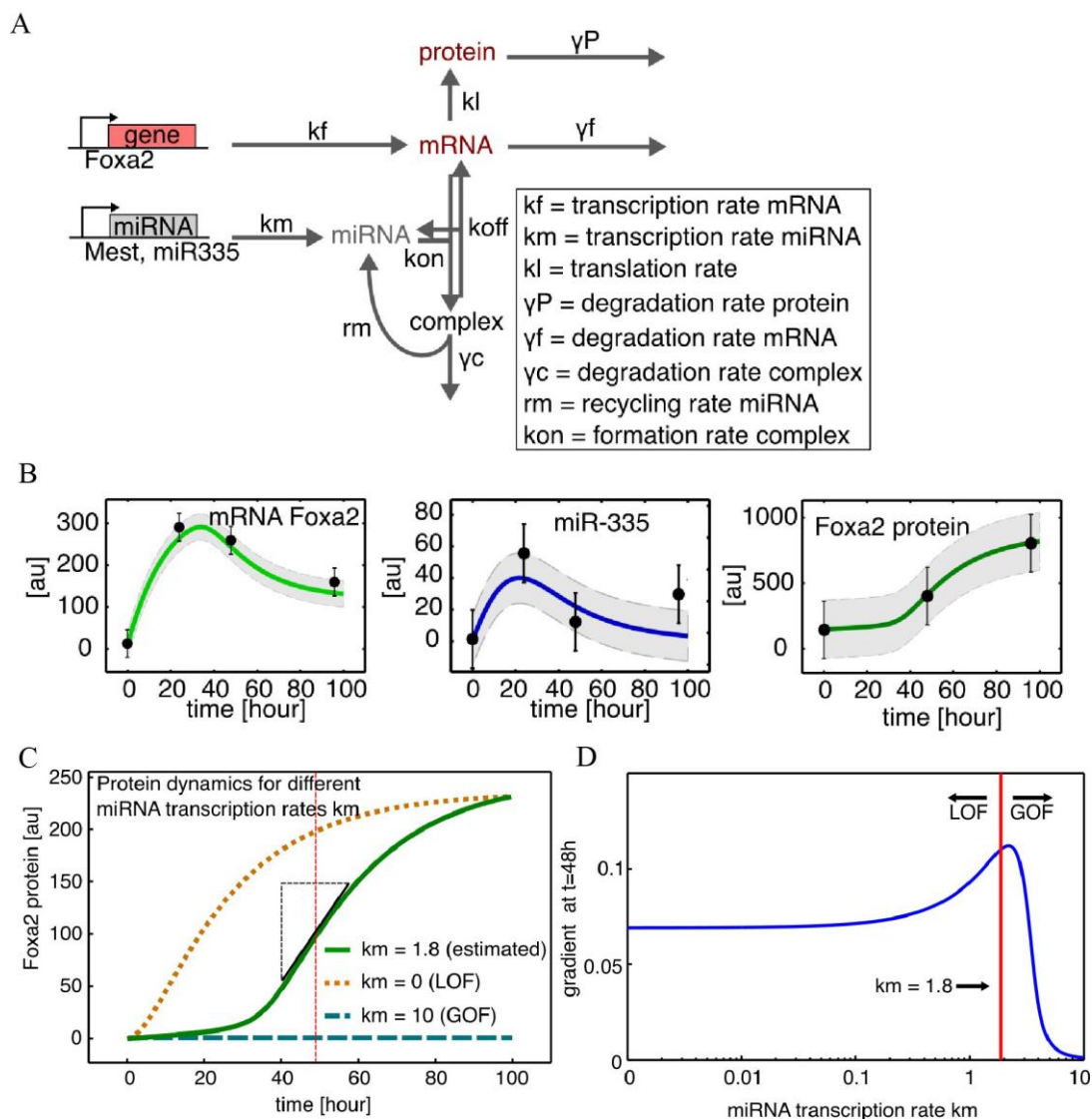
To analyze and describe the effect of miR-335 on mesendoderm and DE differentiation in a continuously quantitative resolved fashion, Dr. Dominik Lutter generated an *in silico* molecular mathematical model based on ordinary differential equations (ODEs; Figure 2.20 A). The model is based on the generic models previously used to analyze miRNA-mediated effects on protein expression (Levine et al., 2007; Mukherji et al., 2011). In contrast to these models, where miRNA concentrations were assumed to be constant, dynamic miRNA turnover rates as deduced from qPCR data were introduced. For all RNA molecules distinct transcription rates as well as for the free mRNA and protein degradation rates were unknown and therefore estimated from the data. To test for parameter identifiability and to estimate confidence intervals, the profile likelihood estimation (PLE) was exploited (Raue et al., 2009). We assumed that degradation of the complex led to a partly degradation of the miRNA, thus a fraction of the active miRNA was recycled. Given that miRNA half-lives are much longer compared to mRNA half-lives (Krol et al., 2010), thus, the model does not include a separate degradation rate for the free miRNA. Referring to the *Foxa2* expression we assumed a constant transcription  $k_f$  rate for the target mRNA and - since miRNA expression decreases after 24 hr - we model the miRNA transcription using a bell-shaped function with an estimated maximum at time = 0 hr. The transcription rate  $k_m$  thus refers to the maximal miRNA transcription rate.

To calibrate the model parameter, qPCR measured mRNA and miRNA data from FACS isolated *Foxa2*<sup>+</sup> cells (Figure 2.15 C-E) and FVF and SCF reporter activities measured at 0, 48 and 96 hr of endoderm differentiation were used (Figure 2.20 B). To study the model dynamics, we started simulations with very low RNA and protein concentrations, as predicted from the experimental data. After calibrating the model parameters, we studied model dynamics for the protein readout depending on different miRNA transcription rates. To measure miRNA dependent effects on the protein dynamics, the time dependent protein gradient at time = 48 hr, which corresponds to half of simulated differentiation time, was calculated (Figure 2.20 C). For the estimated physiological miRNA transcription rate ( $k_m$ ) of 9.8 [miRNA/time a.u.], a sigmoid-like expression curve was observed. We then analyzed miRNA LOF



## 2.Results

and GOF by simulating the model with  $k_m$  between 0 [miRNA/time a.u.] and 25 [miRNA/time a.u.] (Figure 2.20 C). As shown miR-335 GOF completely blocks protein expression, whereas miR-335 LOF allows for a faster accumulation of the target protein, thus decreasing the gradient at time = 48 hr. We proved the model's confidence by comparing GOF and LOF trajectories for estimated parameter below the 95% PLE threshold. The continuous miR-335 dependent protein (Foxa2) gradient for time = 48 hr is displayed in Figure 2.19 D.



**Figure 2.20 Modeling miRNA mediated protein expression dynamics**

(A) Model describing the dynamical behavior of the free mRNA, the free miRNA as well as the complex (miR binds to mRNA) and the translated protein.

(B) Model parameter estimation for Foxa2<sup>+</sup> cells. Experimental data (black dots) were measured from FACS sorted Foxa2<sup>+</sup> cells. Foxa2 mRNA and miR-335 was measured using qPCR, Foxa2 protein expression was measured by Western blot. Pulsed miR-335 expression was modeled using a Gaussian expression function. Gray areas denote the confidence intervals for the parameter estimation.

(C) Model prediction for different miR expression rates  $k_m$ . Protein expression was predicted using different miR-335 expression rates. The green solid line displays protein dynamics for the estimated miR-335 expression rate ( $k_m = 1.8$ ) for pulsed expression as estimated from experimental data. Protein dynamics for a simulated complete miR-335 knockdown ( $k_m = 0$ ) is shown by the dashed orange line. Linear miR-335 ( $k_m = 10$ ) expression as estimated for the mesodermal T<sup>+</sup> cells is shown by the dashed dark blue line. The dashed red line indicates  $t = 48\text{hr}$ .

(D) Predicted protein gradient for different miR-335 transcription rates ( $k_m$ ). The simulated protein gradient at time = 48hr is shown as a function of miR-335 transcription rates  $k_m$  ( $\log_{10}$ ). The gradient increases for increasing transcription rates until it drops when a miR-mediated knockdown is reached. The red line denotes the estimated transcription rate for Foxa2<sup>+</sup> cells. The curve left of the red line shows simulated LOF behavior, right of the red line shows GOF behavior.

### 2.2.7 miR-335 shapes a TF gradient in the endoderm

Our previous study showed that the endoderm-specific TFs, Foxa2 and Sox17, accumulated in an A-P gradients in the endoderm germ layer (Burtscher et al., 2012). How these spatio-temporal TF gradients in the endoderm are formed remains unclear. Recently, miRNAs, such as miR-15 and miR-16, were shown to be involved in the establishment of morphogen gradients, including TGF $\beta$ , Wnt and other growth factors by acting on their secretion, distribution and clearance (Inui et al., 2012; Martello et al., 2007). The presence of miRNAs regulating morphogen gradients, therefore, may leads to the TF gradient.

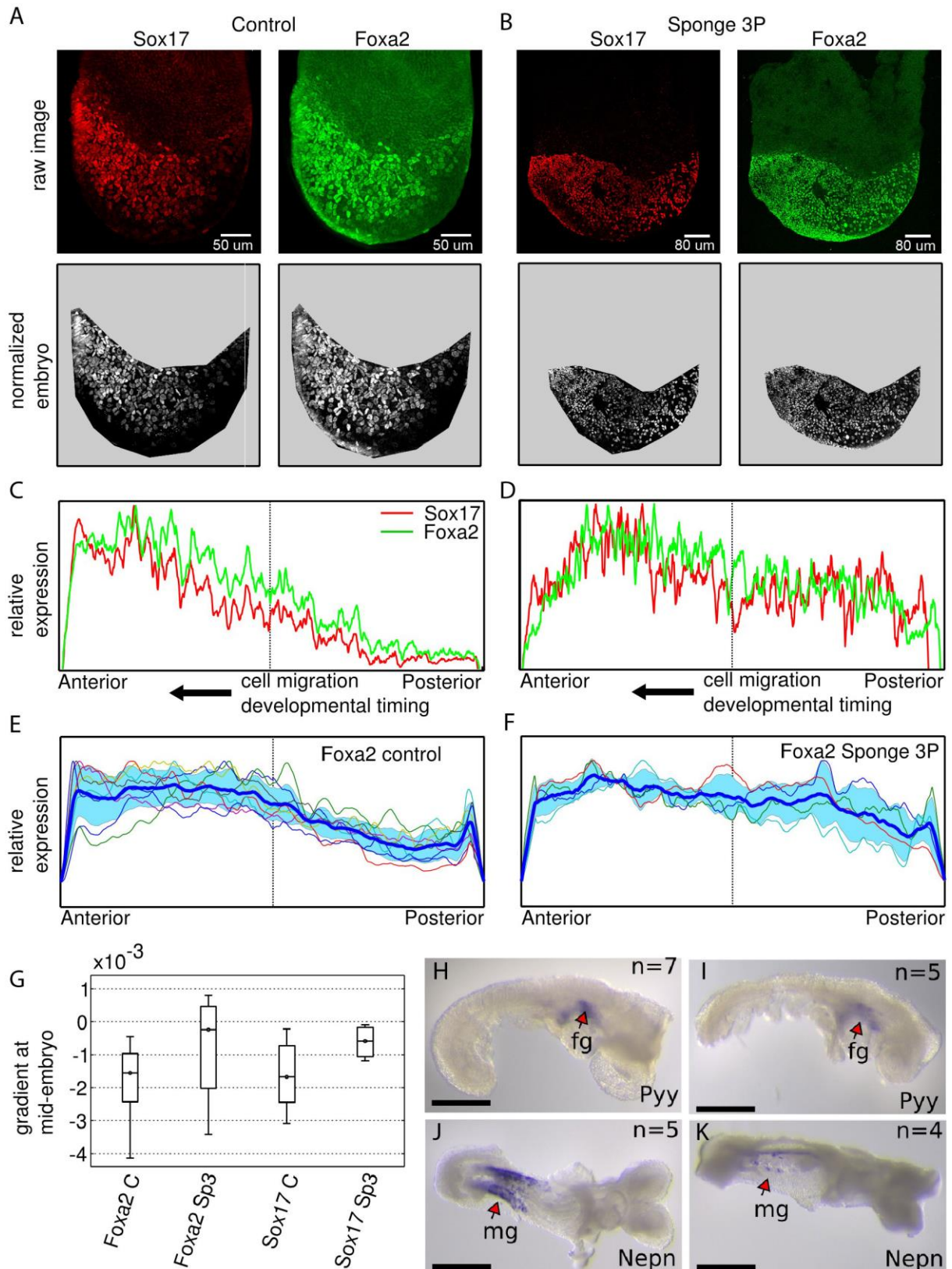
As shown above, miR-335 is transiently expressed in the Foxa2-mesendoderm progenitors, but quickly down-regulated in the DE (Figure 2.15 E). Moreover, miR-335 GOF blocks DE formation after mesendoderm induction (Figure 2.18), whereas miR-335 LOF increases DE formation and Foxa2/Sox17 protein accumulation (Figure 2.19). Therefore, we assume that miR-335 acts at the level of the

Foxa2<sup>+</sup>mesendoderm progenitor to shape a gradient of Foxa2 and Sox17 in the anterior-posterior patterned DE, as predicted by the mathematical model.

To validate this model experimentally, we analyzed TF gradient formation in the gastrula-stage mouse embryo, where different morphogen activities translate into an A-P gradient of Foxa2 and Sox17 (Figure 2.21 A) (Burtscher et al., 2012). Foxa2<sup>+</sup> mesendoderm epiblast progenitor cells are recruited at the posterior side of the embryo, where miR-335 levels are still high, and differentiating Foxa2<sup>+</sup>/Sox17<sup>+</sup> DE cells intercalate and migrate to the anterior side of the embryo (Burtscher et al., 2012), where according to our ESC data the miR-335 levels decreases (Figure 2.15 E). This is reflected in the accumulation of Foxa2 and Sox17 protein in an A-P gradient as revealed by IHC and LSM analysis of wild type embryos at E7.5 (Figure 2.21 A).

To confirm that, an automated image quantification method to determine the protein amounts in the DE along the A-P axis from fluorescent images was used. Quantification of the Foxa2 and Sox17 protein levels revealed a spatial and temporal protein gradient along the A-P axis in single embryos (Figure 2.21 C and D) and pooled embryo groups at gastrulation (Figure 2.21 E and F), which we quantified explicitly (Figure 2.21 C and D). Since DE cells migrate over time along the A-P axis, we quantify a spatial gradient that reflects the time-dependent protein accumulation. From these data we predicted that the miR-335 LOF should lead to an increase of Foxa2 and Sox17 protein levels in developmentally younger cells and therefore should accumulate faster at the posterior side of the embryo. To confirm this, the completely ESC-derived sponge-3P expressing embryos were generated and the TF gradient was analyzed (Figure 2.21 B). Analysis of Foxa2 and Sox17 protein accumulation (Figure 2.21 B-F) and gradient formation (Figure 2.21 G) confirmed our model predictions and revealed that miR-335 functions to shape a TF gradient in the endoderm in vivo. Finally, we tested whether the change in the TF gradient at gastrulation stage leads to patterning defects in the gut tube at E8.5. Whole-mount *ISH* revealed that the foregut gene *Pyy* was normally expressed (Figure 2.21 H and I), whereas the midgut gene *Nepn* was strongly reduced in miR-335 sponge expressing embryos when compared to controls (Figure 2.21 J and K). Taken together, these results suggest that miR-335 functions to shape a TF gradient in the endoderm that translates into gut tube patterning.





**Figure 2.21 miR-335 shapes TF gradients in the endoderm**

(A, B) Upper row: Whole-mount immunofluorescence stainings for FVF and SCF fusion proteins in control and sponge-3P ESC-derived embryos showing levels of proteins at E7.5. Lower row: Normalized intensities of TF gradient.

## 2.Results

---

(C, D) Plots display the normalized mean TF levels along the A-P axis for FVF and SCF in control and sponge-3P embryos. Loss of miR-335 function leads to increased FVF and SCF protein at the posterior, developmental younger side.

(E,F) Summary plot of all TF gradients in multiple embryos (colorful lines). The mean (thick blue line) and standard deviation (blue area) is shown for FVF of control embryos (upper) and miR-335 LOF (Sponge-3P) embryos.

(G) Estimated protein gradient for control (n = 7) and sponge-3P (n = 4) embryos. The *Foxa2* and *Sox17* protein gradient was estimated in the medial range (dashed black line in C, D, E) along the A-P axis. For both proteins a less distinctive gradient was observed for loss of miR-335 function consistent with the modeling results in Figure 6.

(H-K) Whole mount *ISH* with indicated probes at E8.5. CD1 control (H, J) and sponge-3P embryos (I, K) showed similar foregut (fg) expression of *Pyy*, whereas the midgut (mg) expression of *Nepn* is strongly reduced in sponge-3P embryos. Scale bars G-J: 500  $\mu$ m.

---

## 3. Discussion

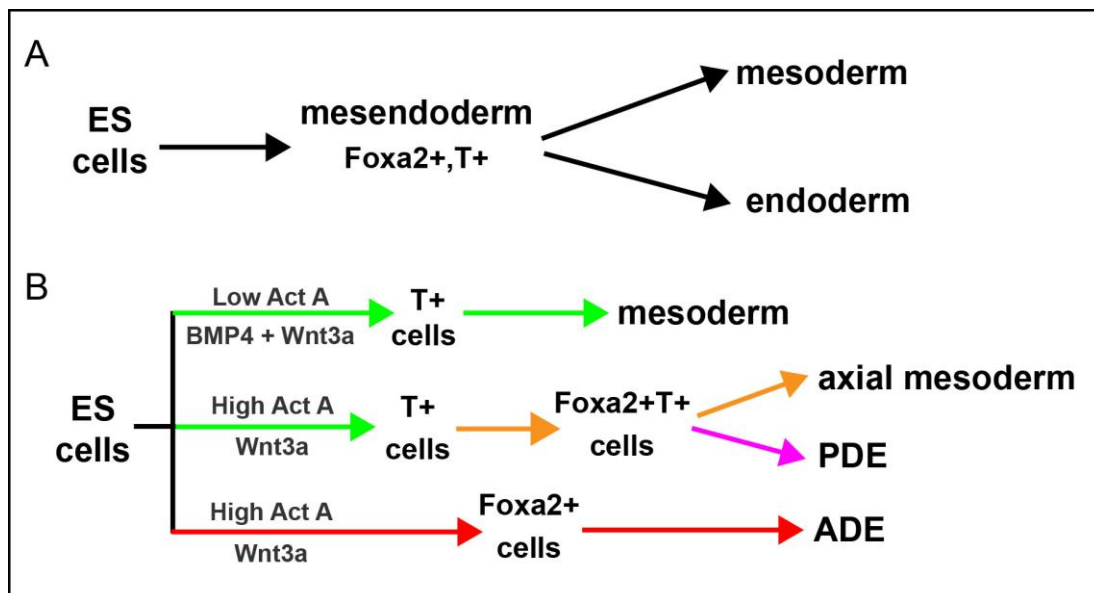
### 3.1 Summary of results

Despite many efforts to reveal the mechanisms of either mesendoderm or subsequent endoderm and mesoderm differentiation, the lineage specification during gastrulation is not fully understood. The aim of this work was to investigate the mechanisms that are involved in lineage specification during gastrulation.

In this study, we report an effective strategy for the stepwise differentiation of mESCs into mesendoderm, and later endoderm/mesoderm lineages by triggering Wnt/ $\beta$ -catenin, Nodal/ActA and Bmp4 signaling. We have generated a dual fluorescent knock-in  $T^{GFP/+}; Foxa2^{tagRFP/+}$  reporter ESC line for lineage specification analyses during gastrulation. Combined with a surface marker CD24, our *in vitro* differentiation system enabled us to isolate early and late endodermal/mesodermal progenitors. We have characterized these isolated populations by microarray analysis. The mRNA expression profile from isolated populations nicely correlate with the cell populations formed in gastrulating embryos, suggesting that our *in vitro* differentiation system serves as a relevant model system to investigate lineage segregation during gastrulation.

Second, we have conducted a single-cell time-resolved lineage tree analysis of mESCs differentiation using  $T^{GFP/+}; Foxa2^{tagRFP/+}$  reporter ESC line. We demonstrated that three different mesendoderm lineages are directly formed from pluripotent ESCs under endoderm and mesoderm promoting differentiation conditions in culture. The time-resolved isolation and molecular profiling of these distinct lineages further uncovered that mesoderm, axial mesendoderm/PDE as well as ADE are formed from three distinct  $T^+$ ,  $Foxa2^+T^+$ , and  $Foxa2^+$  progenitor populations, respectively. In contrast to the common belief (Figure 3.1A), we demonstrated that ADE is derived from  $Foxa2^+$  progenitor and not from a  $T^+$  progenitor, which instead give rise to axial mesoderm and PDE (Figure 3.1B). Moreover, both homozygous and heterozygous  $T$  knock-out mouse embryos show no defect in the formation of DE cells, which further confirms that DE formation does not require  $T$  function. These data demonstrate an

unprecedented progenitor relationship and lineage hierarchy during mesendoderm differentiation, which are crucially important for the differentiation of the correct progenitor population into functional cell types. Furthermore, using the time-resolved molecular profiling we confirmed that the process of mesoderm formation is a typical EMT due to the up-regulation of EMT related key TFs and down-regulation of epithelial markers in the T<sup>+</sup> lineage. In contrast, there was no up-regulation of EMT TFs in the Foxa2<sup>+</sup> lineage. This reveals that the endoderm is formed by a process independent of classical EMT. As columnar epithelial epiblast progenitors directly give rise to squamous epithelial endoderm, we named this process EET.



**Figure 3.1 Graphical representation of the hypothetical *in vitro* differentiation model**

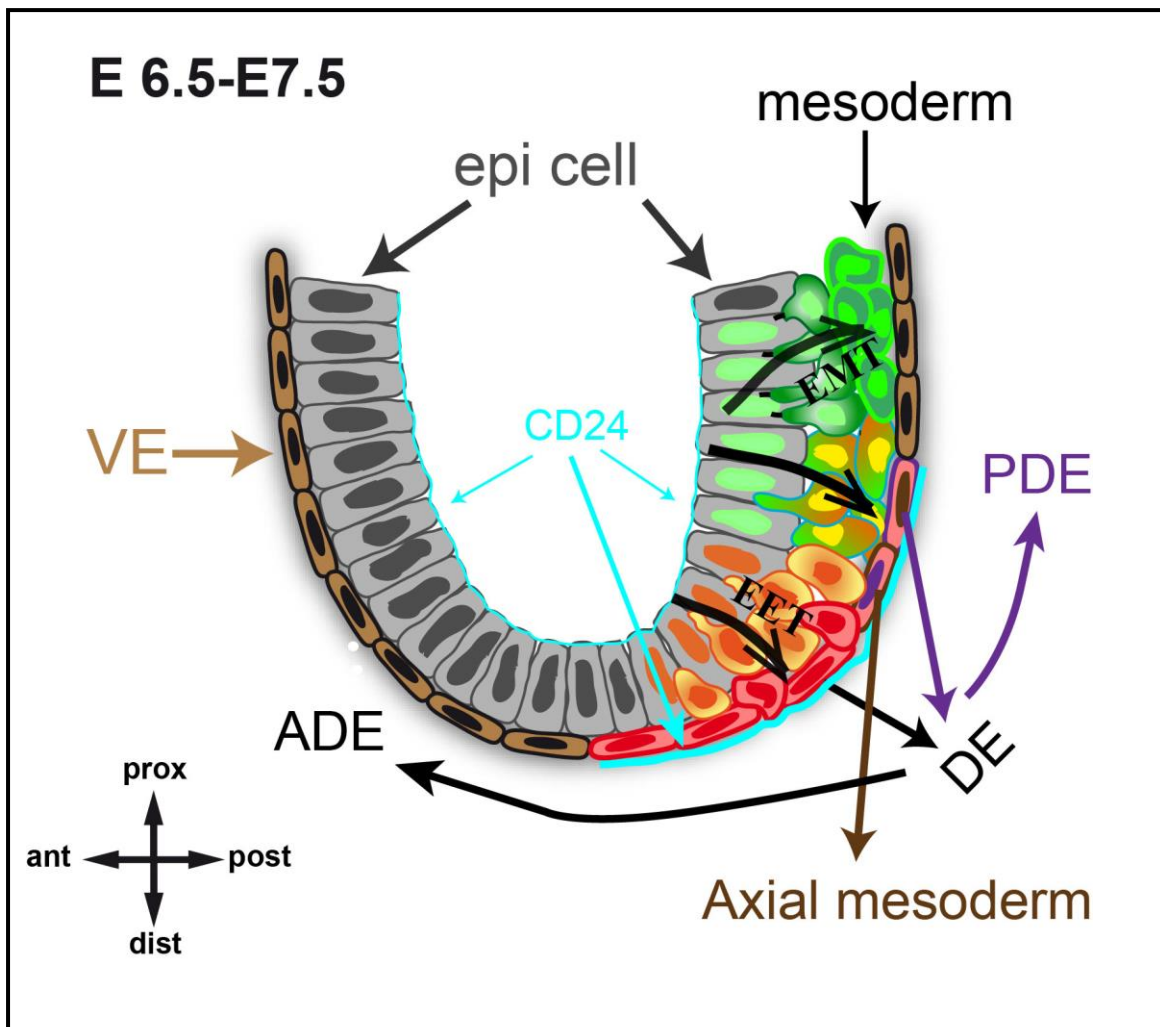
(A) Mesoderm and endoderm derive from the Foxa2<sup>+</sup>T<sup>+</sup> mesendoderm population *in vitro* (previous common model).

(B) Mesoderm, axial mesoderm/PDE, and ADE are generated from distinct T<sup>+</sup>, Foxa2<sup>+</sup>T<sup>+</sup>, and Foxa2<sup>+</sup> progenitor populations, respectively (new model).

---

Third, we have analyzed the physiological functions of miR-335 in the regulation of mesendoderm lineage specification. We identified miR-335 as an intragenic miRNA embedded in the second intron of the mesoderm-specific transcript *Mest*. MiR-335 is transiently expressed in *Foxa2*<sup>+</sup> endoderm progenitors and accumulates highly in the T<sup>+</sup> mesoderm lineage. It specifically targets the 3'-UTR of *Foxa2* and *Sox17* and GOF blocks DE differentiation, whereas LOF enhances DE formation. Quantitative mathematical modeling predicted a miR-335 function in endoderm TF gradient formation, which we confirmed experimentally in developing embryos. Taken together our results suggest two physiological functions of miR-335: First, low miR-335 expression levels dampen *Foxa2* and *Sox17* protein levels in nascent *Foxa2*<sup>+</sup> DE progenitors to establish a TF gradient along the A-P axis in the endoderm germ layer. Second, high miR-335 expression levels promote mesendoderm lineage segregation and prevent lineage inappropriate expression of endoderm TFs in the T<sup>+</sup> mesoderm lineage by default repression.

Based on these findings, we hypothesized a developmental model during gastrulation *in vivo* (Figure 3.2). The lineage specification occurs from onset of gastrulation due to the induction of different signals. At pre-streak-stage, the posterior epithelial epiblast cells (CD24<sup>+</sup>) are segregated into two progenitor populations marked by the expression of T and *Foxa2*. When gastrulation proceeds, these proximal T<sup>+</sup> cells acquire more *Bmp4* signal and undergo a typical EMT process to give rise to mesoderm cells; the intermediate T<sup>+</sup> epiblast cells receive *Bmp4*, *Wnt/β-Catenin* and *Nodal* signals produced from proximal/distal regions of the embryos, and up-regulate *Foxa2* expression. These *Foxa2*<sup>+</sup>T<sup>+</sup> population likely contribute to both, axial mesoderm and PDE. In the meantime, the T<sup>+</sup> cells produce miR-335, which represses the endoderm differentiation program in more differentiated mesendoderm cells. In contrast, the distal *Foxa2*<sup>+</sup> epiblast cells (CD24<sup>+</sup>) acquire more *Nodal* signal produced from the anterior end of the PS, up-regulate epithelial polarity, and allocate to the ADE (CD24<sup>++</sup>). This process is independent of the classical EMT process and represents an EET.



**Figure 3.2 Graphical representation of lineage specification during gastrulation**

During gastrulation, the proximal posterior  $T^+$  epiblast cells ( $CD24^+$ ) acquire more mesenchymal cellular fate, and break through the basement membrane to form mesoderm ( $T^+CD24^-$ ); these intermediate  $T^+$  epiblast cells upregulate *Foxa2* expression and become PS like cells, these cells subsequently give rise to axial mesoderm and PDE. During the migration, these  $T^+$  cells upregulate *miR-335* expression and repress the expression of *Foxa2* and *Sox17*. The distal posterior *Foxa2*<sup>+</sup> epiblast cells acquire more epithelial cellular fate and contribute to ADE.

---

## **3.2 *In vitro* mESCs differentiation system: A valuable tool to investigate mouse gastrulation**

The aim of stem cell research is to model development *in vitro* and ultimately to generate functional cell types for tissue replacement and regenerative medicine. Translation of embryonic principles to ESC differentiation might eventually lead to the generation of cell types for tissue replacement and regenerative medicine. Therefore developmental design principals have to be translated to the culture dish for the generation of functional cell types, such as cardiomyocyte or pancreatic insulin-producing  $\beta$  cells. Lineage specification during gastrulation is a critical step in generating progenitor cells. Here we established an *in vitro* differentiation system, which enabled us to purify and characterize endodermal and mesodermal progenitors and monitor their formation process live over time on the cellular and molecular level; this provides novel practical solutions for demonstrating lineage specification *in vitro*.

### **3.2.1 Establishment of stepwise mESCs differentiation towards endoderm and mesoderm**

ESCs were broadly used recently due to their remarkable developmental potential. Pluripotent ESCs can give rise to all the germ layers and subsequent lineages upon specific induction, raising prospects for biomedical research and for regenerative medicine (Murry and Keller, 2008). In this study, we took the advantage of the mESCs and generated a 2-D *in vitro* differentiation system to mimic early embryonic development. Unlike the previous reported 3-D EB differentiation system, which generates all three germ layers, our differentiation protocol induces the differentiation of mESCs towards either the endoderm or mesoderm fate. Therefore, the differentiation is more restricted and directed towards two populations, which allows the dissection of the mesendoderm lineage hierarchy and growth factor requirements in detail, because the complexity in the 2-D mESCs differentiation system is reduced and the adherent morphology is easier to track at single cell level.

One of the advantages of mESCs differentiation is the recapitulation of the signaling pathways of the early mouse development in the differentiation system, in which defined cytokines and growth factors can be supplied in a more controlled manner to mimic *in vivo* development. In this study, we dissected the cooperative interactions of ActA, Wnt3a and Bmp4 for the induction of mesendoderm and subsequent DE and mesoderm formation. Our data revealed that both high and intermediate concentrations of ActA could induce mESCs differentiation towards an endodermal fate, while lower level of ActA with Bmp4 leads to mesoderm differentiation. These data are in line with previous *in vivo* studies that shown DE cells derived from PS region, which is in close to Nodal source, whereas the mesoderm cells are induced in the adjacent cells, which produced more Bmp4 signal (Ben-Haim et al., 2006; Hagos and Dougan, 2007; Rodriguez et al., 2005; Zorn and Wells, 2009). During gastrulation, Nodal and Wnt3 are expressed in posterior PS region and essential for PS formation. The inhibitors of these pathways are expressed in the AVE that restricts their activity to the posterior site of the embryo. As such, the posterior epiblast acquires high levels of Nodal and Wnt3 and progenitors are fated towards the mesendoderm lineage (Arnold and Robertson, 2009; Tam et al., 2006). Similarly, Bmp4 and its inhibitors are expressed in proximal ExE and distal node region, respectively (Bachiller et al., 2000; Beppu et al., 2000; Klingensmith et al., 1999; Winnier et al., 1995). The gradients formed by these signaling molecules along the proximal-distal region are important for mesendoderm specification. The proximal posterior epiblast cells, which are in close contact with the ExE acquire more Bmp4 signal and relative lower ActA to form mesoderm, while the distal posterior epiblast cells acquire high levels of Nodal and give rise to DE. Our data nicely recapitulate the *in vivo* situation and support the idea that a balance of Nodal/Activin, Wnt3 and Bmp4 specifies the cells towards the mesoderm and endoderm fate. Consistent with previous *in vitro* mESCs differentiation protocols (Burtscher et al., 2013; D'Amour et al., 2005; Kubo et al., 2004; Yasunaga et al., 2005), we used ActA to induce differentiation, as both ActA and Nodal belong to TGF- $\beta$  superfamily and signaling through the same receptor to activate Smad signaling (Schier, 2003). However, in our study 100 ng/ml ActA concentration did not significantly induce endoderm differentiation compare to that of 12 ng/ml ActA concentration, suggesting that concentrations of at least 12 ng/ml ActA are required for DE induction.



---

Advances have been made in inducing both human and mouse ESCs to differentiate towards endoderm or mesoderm (Borowiak et al., 2009; Burtscher et al., 2013; Fu et al., 2011; Gadue et al., 2006; Kroon et al., 2008; Kubo et al., 2004b; Yang et al., 2014b; Yasunaga et al., 2005). However, none of these protocols could really recapitulate all the development processes of the actual embryo due to yet undefined factors and environmental conditions. Similar to the previous published *in vitro* differentiation protocols (Morrison et al., 2008; Villegas et al., 2013), our endoderm differentiation is directed more towards ADE than PDE, which should allow to differentiate ADE-derived cell types of the lung, pancreas and liver. This is surprising, as it is generally assumed that Nodal/Activin and Wnt/ $\beta$ -catenin signaling is important for the induction of mesendoderm and the formation of endoderm, but the patterning of ADE is thought to rely on the inhibition of these pathways by antagonists secreted from the AVE. One possible explanation for the successful differentiation of ADE is that these cells synthesize these antagonists during differentiation, which then shield and dampen signaling for correct tissue patterning.

### **3.2.2 The dual knock-in reporter mESCs line: $T^{GFP/+}$ ; $Foxa2^{tagRFP/+}$ , a useful tool to investigate lineage specification during gastrulation**

Cell fate analysis has been used to track the differentiation of progenitors during the time course of development. Fate maps of pre- and early streak embryos revealed which regions contained the progenitors of ectoderm, mesoderm and endoderm. The proximal posterior epiblast, which is adjacent to the ExE, contains the precursors of the extra-embryonic mesoderm and the primordial germ cells. By contrast, the intermediate posterior epiblast gives rise to the most anterior mesoderm and DE (Lawson et al., 1991; Lawson and Pedersen, 1992; Tam and Beddington, 1992; Lawson and Hage, 1994). Classical approaches for lineage tracing included the injection of either horseradish peroxidase (HRP) or dextran linked fluorescent dyes into the mouse embryos to track the descendants of labelled cell over time. These provided insight for tissue formation and regionalization with potential clone boundaries, but not correspond to distinct cell lineages (Buckingham and Meilhac, 2011). In the context of gastrulation, several distinct lineages have been identified. How to distinguish and track these distinct lineages which give rise to either

endoderm or mesoderm in a continuous single cell level still remains a challenge *in vivo*.

The major questions in germ layer specification are when and how endodermal and mesodermal progenitors are specified and determined. To address these questions in detail, we generated a dual knock-in reporter approach whereby endogenous locus with *T* and *Foxa2* were targeted by the expression of two reporter molecules, GFP and tag-RFP, respectively. The use of fluorescent proteins of single-cell labelling has had a remarkable impact on lineage analysis, as they are genetically encoded. In combination with the time-lapse imaging analysis, the targeted fusion proteins enabled us track the developmental process on the single cell lineage to resolve binary lineage decisions. *Foxa2* and *T* were used as their expression marks the endodermal and mesodermal progenitors and progeny. *T* and *Foxa2* are expressed from onset of gastrulation in the posterior epiblast and PS as well as in endoderm and mesoderm until gastrulation ends (Burtscher and Lickert, 2009; Herrmann, 1991; Monaghan et al., 1993; Sasaki and Hogan, 1993). For the molecular analysis we further separated the progenitors from the progeny with help of the epithelial surface marker CD24. In combination with CD24 the early and late *Foxa2*<sup>+</sup> and *T*<sup>+</sup> cells could be further separated.

The other advantage of the *T*<sup>GFP/+</sup>; *Foxa2*<sup>tagRFP/+</sup> dual knock-in reporter mESCs line, is the targeted fluorescent proteins enable us to isolate and characterise the specific cell population for further analysis. The time-resolved isolation and molecular profiling data shows that the gene expression signature of early progenitors, *Foxa2*<sup>+</sup>CD24<sup>+</sup>, *T*<sup>+</sup>CD24<sup>+</sup>, and *Foxa2*<sup>+</sup>*T*<sup>+</sup> cells, are similar with epiblast stem cells (epiSCs). EpiSCs are derived from postimplantation epiblast at E5.5-7.5 and show co-expression of core pluripotency markers and mesendodermal markers, such as *Foxa2* and *T* (Bernemann et al., 2011; Brons et al., 2007; Tesar et al., 2007). EpiSCs are considered to be pluripotent and can generate teratomas, but do not effectively form blastocysts chimeras (Bernemann et al., 2011; Nichols and Smith, 2015). Similar to epiblast cells, the *Foxa2*<sup>+</sup>CD24<sup>+</sup> or *T*<sup>+</sup>CD24<sup>+</sup> cells still keep expression of core pluripotency genes. In mouse, although cell fate is acquired during gastrulation, epiblast cells maintain pluripotent until the start of somitogenesis (Osorno et al., 2012). Pluripotent genes, such as *Oct4* and *Nanog* are expressed in epiblast and PS cells, and have been shown to regulate germ layer specification through

---

Eomesodermin (Teo et al., 2011). Besides that, we observed an up-regulation of *FGF5* expression in  $Foxa2^+CD24^+$  and  $T^+CD24^+$  populations. *FGF5* is regarded as a marker of primitive ectoderm/epiblast cells, because it is up-regulated in epiblast cells but not in ICM or mESCs (Hébert et al., 1991). However, in contrast to epiblast cells showing heterogeneous PS-like and neural-like characteristics, the  $Foxa2^+CD24^+$  and  $T^+CD24^+$  cells are more specified towards endoderm and mesoderm, without neural fate-like characteristics (Tsakiridis et al., 2014). This is due to the induction of Wnt, Activin and Bmp4 signaling in culture. These signals induce the PS and repress neural differentiation. In contrast, the characteristics of  $Foxa2^+T^+$  cells resemble more close the cells formed in the PS region, due to the high expression of PS as well as pluripotency genes *Oct4* and *Nanog*, but not *Sox2*. *Sox2* expression was restricted to presumptive neuroectoderm but excluded from PS by mid-late-streak stages (Avilion et al., 2003).

In contrast, the pluripotent genes are low expressed in the  $Foxa2^+CD24^{++}$  and  $T^+CD24^-$  subpopulations, which instead show high expression of endodermal and mesodermal markers, respectively. The  $Foxa2^+CD24^{++}$  subpopulation expresses much higher levels of *Cxcr4* and *Sox17* when compared to the others. Although *Sox17* is expressed not only in DE but also in mouse visceral endoderm and parietal endoderm (Kanai-Azuma et al., 2002), the expression of *Cxcr4* is restricted to DE (McGrath et al., 1999) and permits the isolation of DE (D'Amour et al., 2005). Given that the sorting is based on the expression level of CD24, which has been reported to be enriched in DE but not in VE (Jiang et al., 2011; Wang et al., 2012), it is strongly suggested that the  $Foxa2^+CD24^{++}$  represents mature DE.

Overall, the establishment of the dual knock-in reporter mESCs line as well as the *in vitro* culture system provide the basis for modelling lineage specification during gastrulation.

### 3.3 Novel mechanisms of endoderm and mesoderm formation

Many endoderm and mesoderm derivatives have been successfully induced from ESCs under certain conditions induction (Kehat et al., 2001; Pagliuca et al., 2014; Shiba et al., 2012). However, the achievement in delivering terminally appropriate cell populations is limited due to the heterogeneous cultures of the ESCs differentiation, in which the appropriate cells represent a minority of the entire population. In this study, we have identified three distinct mesendoderm lineages ( $T^+$ ,  $Foxa2^+$ , and  $Foxa2^+T^+$ ) that are likely fated towards the mesoderm, ADE, and axial mesoderm/PDE lineage, respectively. A better understanding of the lineage specification and cell fates of these distinct lineages during gastrulation is crucial, as this will help to find a more efficient way to differentiate functional cells for future tissue replacement.

#### 3.3.1 Mesendoderm lineage segregation and cell fate decisions

Mesendoderm was identified at the end of last century because of its potential to give rise to both endoderm and mesoderm. Labelling of single cells in *Xenopus*, *Zebrafish* and frog showed that endoderm and mesoderm derived from common marginal cells (Nieuwkoop, 1997; Warga and Nüsslein-Volhard, 1999), which are then referred as bi-potent mesendoderm (Rodaway and Patient, 2001). In mice, the mesendoderm cells were existed in PS and show potential to give rise to both endoderm and mesoderm (Kimelman and Griffin, 2000; Rodaway and Patient, 2001). In support of this notion are *in vitro* mESCs models showing endoderm cells were derived from  $T^+$  cells, which supposed to be mesendoderm population that co-expressed  $T$  and  $Foxa2$  (Kubo et al., 2004b). Mesendoderm was further characterized by  $Gsc^+/E-cad^+$  or  $Cxcr4^+/Pdgfra^+$  cells with the possibility to differentiate towards both DE and mesoderm (Kopper and Benvenisty, 2012; Tada et al., 2005). These mesendoderm populations were shown to have similar properties with cells in the anterior PS (Kubo et al., 2004b; Tada et al., 2005). These results support the idea that the bi-potent mesendoderm population exists in mammals. However, not all mesoderm and endoderm cells were derived from these bi-potent progenitors. Indeed only a subset of mesoderm cells that fated to cardiac mesoderm were derived from these bi-potent

---

progenitors (Rodaway and Patient, 2001), suggesting the existence of non-shared mesendoderm progenitors. Despite these efforts, the mesendoderm lineage hierarchy has not been fully established in any of the animal models and *in vitro* studies due to the lack of continuous single cell lineage tracking approaches.

In this study, we provide another model of mesendoderm lineage segregation upon time-lapse imaging analysis and molecular profiling. We demonstrate that three distinct progenitors were purified and characterized in mesendoderm stage *in vitro*. The  $Foxa2^+CD24^+$  and  $T^+CD24^+$  cells showed similar properties with posteriorepiblast cells, while  $Foxa2^+T^+$  cells display more PS and axial mesoderm properties. But what are the fates of these distinct mesendoderm lineages? We have indicated that these endoderm differentiated  $Foxa2^+CD24^+$  cells mostly give rise to  $Foxa2^+CD24^{++}$  cells, which have a gene expression profile similar to ADE. The  $Foxa2^+$  lineage is thereby fated to ADE, which is further supported by following criteria: first, these  $Foxa2^+CD24^{++}$  cells express high levels of ADE related genes including *Cxcr4*, *Hhex*, *Sfrp5* and *Fzd5*, but low levels of VE, PDE, and ESCs related markers. Second, the isolated early  $Foxa2^+$  cells show more capacity to differentiate further towards ADE and ADE-derived hepatocytes when compared with  $Foxa2^+T^+$  cells; whereas,  $T^+$  cells derived under mesoderm condition show paraxial and cardiac mesoderm characteristics. However, the formation of  $Foxa2^+$  cells and  $T^+$  cells in either endoderm or mesoderm conditions did not differentiate via the  $Foxa2^+T^+$  stage; In contrast, these endoderm differentiated  $Foxa2^+T^+$  cells could form  $Sox17^+$  endoderm cells, but only with a low percentage. It has been reported that triggering Wnt signaling concentration is essential for PDE differentiation (Sherwood et al., 2011), therefore, we optimized our differentiation protocol and further cultured these sorted early  $Foxa2^+$  and  $Foxa2^+T^+$  cells with a EB system under high Wnt3a concentration conditions. Interestingly, we found that the  $Foxa2^+T^+$  cell-derived EBs showed higher expression of *Cdx1* and *Cdx2* when compared to  $Foxa2^+$  cell-derived EBs. *Cdx1* and *Cdx2* are posterior DE markers and have been shown to be important for intestinal specification (Gao et al., 2009). Our data is further supported by lineage tracing approach with broad T-Cre activity in mesoderm and posterior gut endoderm (Kumar et al., 2008; Perantoni et al., 2005). Likewise, using a kidney capsule transplantation experiment, Kubo and his colleges found that *Ifabp*, which is indicative of intestinal development was expressed in the endodermal GFP-Bry<sup>+</sup> cells

derivatives (Kubo et al., 2004a). These findings support that the endoderm differentiated Foxa2<sup>+</sup>T<sup>+</sup> cells could give rise to PDE-derived organs.

Notably, we do find these endoderm-derived Foxa2<sup>+</sup>T<sup>+</sup> cells appear to be axial mesoderm due to the high expression of node cell related genes *Gsc*, *Chrd*, *Shh* and *Fltp* in this population. In the mouse embryo, Foxa2 and T are restricted to axial mesoderm, which is derived from the anterior segment of the PS and then gives rise to notochord and node (Kinder et al., 1999, 2001). The high expression of node cell related genes in Foxa2<sup>+</sup>T<sup>+</sup> cells indicate that they have the potential to form the node upon certain inductive conditions. The establishment of mesendoderm lineage hierarchy provide fundamental aspects of endoderm and mesoderm formation, as well as later axis patterning, tissue specification and differentiation. Given that the self-renewing endodermal progenitor lines has already been generated in culture and could differentiate into numerous endodermal lineages, such as pancreatic  $\beta$ -cells, hepatocytes, and intestinal epithelia (Cheng et al., 2012). Therefore, the identification of mesendoderm progenitors and establishment of the self-renewing progenitor cell lines offer a better starting point of endodermal- and mesodermal-derived tissues for cell-replacement therapies.

The heterogeneity of these *in vitro* endoderm-differentiated early progenitors indicates that the epiblast cells might be already specified before entering PS with different fates due to the gradients of signals. These signals including Nodal, Wnt3, and Bmp4, were produced from PS and the surrounding VE (Conlon et al., 1994; Lawson et al., 1999; Liu et al., 1999; Varlet et al., 1997). Epiblast cells within different region acquire different level of signals, therefore showing different fates. These proximal posterior T<sup>+</sup> epiblast cells acquire Bmp4 and Nodal signals, up-regulate T protein and form the mature mesenchymal mesoderm, whereas these intermediate posterior T<sup>+</sup> epiblast cells acquire relative higher level of Nodal, up-regulate Foxa2 protein, and contribute to axial mesoderm and PDE. These Foxa2<sup>+</sup>T<sup>+</sup> cells did not acquire a mesenchymal fate along with the proximal posterior T<sup>+</sup> cells, instead they up-regulate epithelial markers (E-Cad, CD24, and Cldn6) expression, reflecting that they are highly polarized and connected through cell-cell adhesion. Similarly, the distal posterior epiblast Foxa2<sup>+</sup> cells acquire the highest level of Nodal, up-regulate epithelial markers expression and form ADE. The differential expression of epithelial/

---

mesenchymal in Foxa2<sup>+</sup> and T<sup>+</sup> lineages indicates that mesoderm and endoderm formation might display different EMT regulation process.

### 3.3.2 The formation of endoderm does not require EMT

EMT is an evolutionarily conserved cell-biological program that is important for the formation of the body plan and mesenchymal cell types (Lim and Thiery, 2012). It has been described over the past decade as a key event in regulating the cell morphological transition, especially in mesoderm formation or tumor progression and metastasis. In mouse embryogenesis, the first EMT event occurs at gastrulation, where it is responsible for the formation of mesoderm and definitive endoderm (Thiery and Sleeman, 2006). The formation of mesoderm is a typical EMT process with a clear cellular ingression movement and morphology change. However, the effects of EMT on endoderm formation are somewhat less clear. One predicted model is that these specified endodermal epiblast cells first undergo an EMT process to ingress into the PS and there they undergo another mesenchymal epithelial transition (MET) process to intercalate into the outside epithelium and to form the mature endoderm. Another model is, instead of undergoing EMT/MET processes, these endodermal cells rather maintain epithelial polarity and up/down-regulate dynamic cellular adhesions to leave the epiblast epithelium, or are pushed passively into the PS by specified mesodermal epiblast cells that lose polarity cell-cell adhesion.

As we were able to obtain and characterize all the subpopulations involved in the endoderm and mesoderm differentiation process, our *in vitro* differentiation system enables us to get a clear view of the mechanistic regulation of EMT process during gastrulation. In this study, we clearly demonstrated an up-regulation of mesoderm associated TFs (T, Mesp1, and so on), and some well-known EMT-related TFs, such as Snail1, Snail2, Zeb1, Zeb2, and Twist, in mesoderm differentiated cells, suggesting a role for these EMT TFs in early mesoderm development. It has been shown that Nodal, Wnt and FGF signaling play an important role in activation and maintenance of the EMT process (Ciruna and Rossant, 2001; Liu et al., 1999; Mathieu et al., 2004). These signaling induce PS formation and activate the expression of some mesoderm TFs, such as T and Mesp1. Mesp1, as well as these signaling themselves, could induce the expression of EMT-related TFs (Ciruna and

Rossant, 2001; Kemler et al., 2004; Lindsley et al., 2008), thereby in turn induce EMT. The up-regulation of these EMT-related TFs are all responsible for EMT regulation process, where they serve as repressors of E-Cad through direct or intermediate pathways (Kalluri and Weinberg, 2009). In agreement with the previous findings, we did find a clear reduction of E-Cad expression in mesoderm cells. Besides that, we also observed an up-regulation of mesenchymal genes, such as *N-Cad*, *Fibronectin*, and *Vimentin*. These results suggest that the process of mesoderm formation includes a clear global transition from epithelial-to-mesenchymal gene expression, representing a classical EMT process. In contrast, we found that almost all the EMT-related TFs, such as Snail1, Snail2, Zeb1, and Zeb2, were not up-regulated in the endoderm lineages. This observation is in line with the previous study in Snail1 deficient embryos, where mutants exhibit defects in mesoderm; As a consequence, E-Cad expression was not down-regulated in the mesoderm of mutant embryos (Carver et al., 2001; Grau et al., 1984). Of note, no clear endoderm defects have been reported so far. The lack of defects in endoderm formation indicates that this process is independent of a Snail-activated EMT process. An essential feature of EMT is down-regulation of junctional proteins and cell adhesion molecules, such as E-Cadherin,  $\beta$ -Catenin, Cldl4, and Cldl6. Interestingly, we did not observe a down-regulation of these epithelial markers, instead all of these markers were up-regulated in the time course of endoderm differentiation. These data suggest that the formation of endoderm is regulated by a novel mechanism and represents an EET event.

#### **3.3.3 miR-335 regulates TF gradients**

There is growing evidence that miRNAs have a critical role in stem cell maintenance and differentiation. Most miRNAs have been found involved in the regulation of ESCs pluripotency and differentiation networks, in which they function as TFs that regulate entire programs of stem cell pluripotency and differentiation. However, relatively little is known about the function of miRNAs in mesendoderm lineage specification.

Our data demonstrate that miR-335 is transiently expressed in Foxa2<sup>+</sup> endoderm progenitors and highly accumulating in the T<sup>+</sup> mesoderm lineage, where it represses DE differentiation by targeting endoderm TFs Foxa2 and Sox17. The up-regulation of miR-335 in Foxa2<sup>+</sup>T<sup>+</sup> mesendoderm as well as in T<sup>+</sup> mesoderm lineages might result



---

in guiding mESCs differentiation toward mesendoderm and subsequent mesoderm, where it represses the initiation of the endoderm program. In contrast the low levels of miR-335 in Foxa2<sup>+</sup> late endoderm progenitors rather results in an endoderm TFs gradient. How morphogen gradients are established and how they pattern developing tissues in a dose-dependent manner is a long-standing question in developmental biology (Rogers and Schier, 2011). It was long thought that the expression of miRNAs dampen rather than silence the expression of mRNA targets, which makes them prime candidates to fine-tune morphogen and TF gradients. This idea is based on the fact that genome-wide computational and transcriptome analyses showed that miRNA-mRNA target pairs correlate more positive than negative in their tissue expression (Martinez et al., 2008). Prime examples are the regulation of Nodal signaling by the Wnt/ $\beta$ -catenin inhibited miRNA-15 and miRNA-16 during dorso-ventral patterning in *Xenopus laevis* (Martello et al., 2007) and control of extracellular Nodal morphogen availability by miRNA-430 (Choi et al., 2007). Different with the other miRNAs establishing a gradient by targeting morphogens, we provide a new model that rather directly targets the mRNAs of TFs to form a gradient. As TF gradients were often viewed as the integration and net result of several synergistic and antagonistic acting morphogen gradients, this example suggests that miRNAs dose-dependently regulate TF accumulation on the post-transcriptional level by dampening rather than silencing the target mRNAs, which adds an additional layer of regulation to fine-tune morphogen gradients. The effects of miR-335 in either preventing endoderm differentiation or controlling TF gradients reflect novel mechanisms of miRNAs in lineage specification. Increasing miR-335 levels before DE induction leads to repression of DE formation, while modulating miR-335 levels after DE induction may help to generate anterior vs posterior endoderm populations that can give rise to lung, liver, pancreas and gastro-intestinal tract along the A-P axis. Specified cell types, such as insulin-producing  $\beta$ -cells or GLP-1-producing L cells of the gut are of great therapeutic interest to treat metabolic disease.

But how does miR-335 generate the spatio-temporal gradient? In the mouse embryo, the first endoderm recruited from the epiblast progenitors migrates in anterior direction to overlie the forming headfold and is by definition older than PDE (Lawson and Pedersen, 1987; Thomas et al., 1998). In the meantime, the mesoderm T<sup>+</sup> cells are also formed, but are restricted to the posterior side of the embryo. These

posterior T<sup>+</sup> cells produce miR-335, and thereby repress the endoderm differentiation program of the neighboring PDE cells. Another explanation is, through quantitative mathematical modeling based on ordinary differential equations we predicted that the Mest-miR-335 expression level directly correlates with the dose-dependent degree of TF accumulation in the DE. These predictions were confirmed by GOF and LOF experiments in ESC differentiation and developing mouse embryos. The exact concentration and duration of miR-335 levels along the spatio-temporal axis *in vivo* is up-to-date difficult to determine. However, our GOF and LOF results demonstrate how different thresholds of miRNA expression either completely block or increase Foxa2 and Sox17 protein accumulation, thus explaining how low levels of miR-335 in endoderm progenitors lead to a delayed accumulation of TFs along the spatio-temporal axis for gradient formation.

miR-335 is encoded in the second intron of Mest. Our data demonstrate that under the control of mesoderm-specific enhancers, miR-335 is highly expressed in the heart, somites, limb bud and branchial arch mesenchyme. Interestingly, the same tissues are Foxa2<sup>+</sup> lineage positive (Horn et al., 2012; Uetzmann et al., 2008), but neither express Foxa2 nor Sox17 mRNA at E9.5 (Kanai-Azuma et al., 2002; Sasaki and Hogan, 1993). Pioneering work in the fruit fly suggested that miRNAs function to repress their mRNA targets in tissues where they should not be expressed to confer robustness of gene expression (Stark et al., 2005). This indicates that high levels of miR-335 suppress an endoderm TF program by default repression in mesoderm-derived tissues.

Furthermore, several cancers that happened in breast, brain, lung, and pancreas have been associated with differential miR-335 expression (Dohi et al., 2012; Lynch et al., 2012; Polytarchou et al., 2012; Vickers et al., 2012; Xu et al., 2012; Yan et al., 2012). The function of miR-335 in tumor initiation and progression is controversial. For instance, high levels of hsa-miR-335 associated with a high frequency of recurrence and poor survival in gastric cancer patients (Yan et al., 2012) and epigenetic silencing of miR-335 and MEST is correlated with hepatocellular carcinoma (Dohi et al., 2012). Our investigation of the miR-335 function during normal development indicates that miR-335 targets key endoderm TFs downstream of Nodal/TGFβ- and Wnt/β-catenin-signaling. Interestingly, both signaling pathways are implicated in tumor formation and progression and miR-335 is associated

---

differentially regulated during both processes (Lynch et al., 2012; Polytaichou et al., 2012). Our results suggest that overexpression of miR-335 in tumors might silence the epithelialization factor Foxa2 (Burtscher and Lickert, 2009) and contribute to EMT and metastasis formation. Furthermore, the loss of endodermal organ-specific expression of Foxa2 might lead to de-differentiation of mature cell types and acquiring of a more naïve proliferative cellular status. In contrast, silencing miR-335 and its host gene MEST by methylation might cause inappropriate up-regulation of its target genes, such as Foxa2, Sox17, and Sox4 and promote tumor growth in several cancer forms (Dohi et al., 2012; Huang et al., 2012). Overall, our findings allow for new perspectives in several disease related aspects, both for cell-replacement and cancer therapy.



## 4. Materials and methods

### 4.1 Materials

#### 4.1.1 Instruments

Agarose gel chamber	Harnischmacher, Midi 450
Autoclave	Systemec, HX-320
Balances	KERN, ABS KERN, EWB
Cell culture Cabinets	Thermo Scientific, MSC-Advantage™ Class II Biological Safety Cabinets
Centrifuges	Eppendorf, 5417 R Eppendorf, 5430 Thermo, Haereus Rotanta 460R Hettich, Universal 30F Hettich, MIKRO 220R
Counting chamber (cells)	Neubauer (LO - Laboroptik GmbH, Friedrichsdorf) BioRad, TC20 Automated Cell Counter
Developing machine	AGFA, CURIX 60
Digital camera	Zeiss, AxioCam MRc5 Zeiss, AxioCam HRm
Electroporation system	BioRad, Gene Pulser Xcell
FACS	BD, FACS AriaIII
Film cassettes	Amersham, Hypercassette AGFA, CEA RP NEW
Freezer	-20°C, Liebherr -80°C, Heraeus, HFU 686 Top
Fridge	4°C, Liebherr
Gel documentation system	BioRad, UV-Transilluminator

#### 4. Materials and methods

---

Glassware	Schott-Duran
Hybridisation tubes	Thermo, Hybridizer HB 100
Ice machine	Scotsman, AF 30
Incubation system	Shel Lab, Sheldon Manufacturing, Cornelius) Thermo Scientific, BBD 6220 CO2Incubator
Microscopes	Zeiss, Zeiss Stereo Lumar V12 Zeiss, Axiovert 200M Leica, DFC 450C Leica, DMIL LED Leica, TCS SP5II
NanoDrop	Thermo, 2000
Oven	65°C Southern Blot, Thermo Electron Thermo Fisher, HBSNSRS220 6240
PCR machine	Thermo Fisher, Px2 ThermoHybaid Thermo Fisher, PXE0.2 Thermo Cycler Applied Biosystems, 7900HT Applied Biosystems, VIIA 7
pH meter	InoLab, pH Level 1
Pipettes	Eppendorf, 1000 µl / 100 µl / 20 µl / 10 µl
Radiation Monitor	Berthold, LB122
Robot	Intavis, In situ Pro Robot
Shaker	neoLab, DOS-10L
Plastic ware	VITLAB UVsolo TS
Vortexer	Scientific Industries, Vortex Genie 2
Water bath	VWR
Western Blot semi-dry system	BioRad, Trans-Blot® SD, Semi-Dry Transfer cell BioRad, Trans-Blot Turbo

### 4.1.2 Consumables

0.2 ml PCR tubes	Eppendorf
1.5 mL/2 mL reaction tubes	Eppendorf
MicroAmp Fast optical 96-well plate	Life Technologies
Optical Adhesive Covers	Life Technologies
15 mL/50 mL tubes	BD Falcon
15 cm/10 cm/ 6 cm dishes	Nunc
Multiwell plates (6, 12, 48, 96 wells)	Nunc
Corning® Costar® cell culture plates	Sigma-Aldrich
DNA ladder (100bp)	NEB
Filter paper	Whatman 3MM
Pipettes (1ml/2 ml/ 5ml/ 10ml/ 25ml)	Greiner Bio One
PVDF membrane (protein)	BioRad, Immun-Blot PVDF-Membrane
Blotting paper	Whatman paper
Films	Amersham, Hyperfilm ECL

### 4.1.3 Kits

Dual luciferase assay kit	Promega
ECL Detection Kit	Millipore
High Capacity RNA-to-cDNA Kit	Invitrogen
iQ™ SYBR® Green PCR supermix	BioRad
miRNeasy Mini Kit	Qiagen
QIAquick PCR Purification Kit	Qiagen
QIAquick Gel Extraction Kit	Qiagen
QIAgen Maxi Kit	Qiagen
QIAgen Mini Kit	Qiagen
QuantiTect Reverse Transcription Kit	Qiagen
Quick Start™ Bradford Protein Assay	BIO-RAD
RNeasy Mini Kit	Qiagen
TaqMan® Universal PCR Master Mix	Life technologies
Transfection kit	Invitrogen, Lipfectamine 2000

#### 4.1.4 Software

Adobe Design standard CS6	Creative Suite
AxioVision Rel 4.8	Zeiss
CARMAweb	Bioinformatics Graz
Clone manager professional 9	Sci-Ed Software
Flowjo_v10	FLOWJO
GraphPad Prism 6	GraphPad Software
Imaris x64 7.6.5	Bitplane
Leica LAS AF	Leica
Microsoft Office 2010	Microsoft

#### 4.1.5 Chemicals

A	Acetic acid	Sigma
	Activin A, human	R&D Systems
	Acrylamide/bisacrylamide (Rotiphorese Gel)	Roche
	Agarose	Biozym Scientific
	Ampicillin	Sigma
	Ammonium peroxodisulfate (APS)	Roche
	Anti-Digoxigenin-AP Fab fragments	Roche
B	BCA	Sigma
	Blocking Reagent	Roche
	BM purple AP Substrate	Roche
	BSA	BioRad
	Bradford reagent	Sigma
C	Chloroform	Sigma
	Citric acid monohydrate	Roche
	Collagenase P	Roche
D	Dextranulphate	Roche
	Diethylpyrocarbonate (DEPC)	Sigma
	Dimethylsulfoxide (DMSO)	Sigma
	dNTPs	Fermentas
E	EDTA	Sigma



---

	Ethanol	Merck
F	Foetal Bovine Serum (FBS)	Invitrogen
	Formaldehyde	Sigma
	Formamide	Sigma
	Gelatine	Sigma
	Glutamine	Sigma
	Glutaraldehyde	Sigma
	Glycerol	Sigma
H	H <sub>2</sub> O <sub>2</sub>	Roche
	HCl	Sigma
I	Isopropanol, 100%	Sigma
	Levamisol Hydrochloride	AppliChem
	Mek1	Cell Signaling
M	Methanol, 100%	Sigma
	Milk powder	Rothe
	Mineral oil	Sigma
	Mitomycin C	Sigma
N	Nitrogen	Linde
	NaOH	Sigma
	NaCl	Sigma
O	Orange G	AppliChem
P	Paraformaldehyde	Sigma-Aldrich
	Polyethylenimine (PEI)	Polysciences
	Protein inhibitor	Sigma
	Protein marker	Fermentas, SM1811
	Puromycin	Sigma
R	Roti-Phenol / C / I	Rothe
	Sheep serum	Sigma-Aldrich, S-2263
S	Sodiumdodecylsulphate (SDS)	Sigma
T	TEMED	Sigma
	TWEEN20	Sigma
	Tris	Sigma
	Triton X-100	Sigma
	Trizol Reagent	Invitrogen

### 4.1.6 Cell culture reagents

Dulbecco's Modified Eagle Medium (DMEM)	Gibco
Dulbecco's Phosphate Buffered Saline (DPBS)	Gibco
HEPES Buffer Solution	Gibco
PAA	Gibco
PAN	PAN
SFO-3 Medium	Sanko Junyaku
0.05% Trypsin/EDTA (1x)	Gibco
2-Mercaptoethanol	Gibco
Penicillin/Streptomycin (Pen/Strep)	Gibco
Leukemia Inhibitory Factor (LIF)	ESGRO-Millipore
L-glutamine	Invitrogen
Wnt3a	R&D Systems
Bmp4	R&D Systems
bFGF	R&D Systems
Phosphate Buffered Saline (PBS)	Invitrogen
Non-Essential Amino Acids (NEAA)	Gibco

### 4.1.7 Buffer and solutions

#### 4.1.7.1 Common buffers

PBS (10x)	75.97 g Sodium Chloride 1.46 g Disodium Phosphate 4.80 g Sodium Dihydrogen Phosphate H <sub>2</sub> O fill up to 1000 ml, adjust pH to 7.0 with Sodium hydroxide
Tris-HCl	1Mm EDTA 1M Tris base, pH 7.5
TAE buffer (50x stock)	2 M Tris 50 mM Glacial acetic acid 50 mM EDTA
TE buffer	10 mM Tris HCl, pH 8.0 0.1 mM EDTA

---

Loading buffer DNA	100mM EDTA 2% SDS 60% Glycerol 0.2% Bromophenol blue
4 % PFA	4 % Paraformaldehyde solution (PFA) w/v in PBS

#### 4.1.7.2 Plasmid preparation

P1 buffer	50 mM Tris HCl pH 8.0 10 mM EDTA 100 µg/ml RNase A
P2 buffer	200 mM Sodium hydroxide 1% SDS
P3 buffer	3 M Potassium acetate, pH 5.5
QBT buffer	750 mM Sodium chloride 50 mM MOPS, pH 7.0 15% Isopropanol (v/v) 0.15% Triton X-100 (v/v)
QC buffer	1 M Sodium chloride 50 mM MOPS, pH 7.0 15% Isopropanol (v/v)
QF buffer	1.25 M Sodium chloride 50 mM Tris HCl, pH 8.5 15% Isopropanol
EB buffer	10 mM Tris HCl, pH 8.0

#### 4.1.7.3 Isolation of genomic DNA

Proteinase K lysis buffer	100 mM Tris pH8-8.5, 5 mM EDTA pH 8.0 0.2% SDS, 200mM Sodium chloride
Lysis buffer 96-well	2.5 ml 1M Tris pH7.5, 5 ml 0.5EDTA 0.5ml 5M NaCl, set pH 7.0 and autoclave, add 0.5% N-laurolysarcosine sodium salt and 0.4mg/ml Proteinase K before use

#### 4.1.7.4 Southern blot

Depurination (fragments $\geq 10$ kb)	11ml HCl in 989ml H <sub>2</sub> O MilliQ
Denaturation (all gels)	87.66 g Sodium chloride, 20.00 g NaOH 1000 ml H <sub>2</sub> O
Neutralization (all gels)	87.66 g Sodium chloride, 60.50 g Tris 1000 ml H <sub>2</sub> O, pH7.5
Transfer, 20x SSC (all gels)	88.23 g Tri-sodium-citrat, 175.32 g Sodium chloride fill up to 1000 ml with H <sub>2</sub> O, pH7-8
Hybridisation solution	1 M Sodium chloride, 50 mM Tris (pH7.5), 1% SDS, 10% Dextran sulfate, 250 $\mu$ g/ml Salmon Sperm (SS)- DNA sonicated, store at -20°C
20x SSC	175.3 g Sodium chloride, 88.2 g Sodium citrate, pH7.0 1000ml DEPC-H <sub>2</sub> O
20% SDS	200 g SDS in 1000 ml H <sub>2</sub> O

#### 4.1.7.5 Western blot

RIPA lysisbuffer	50 mM Tris/HCl, pH 7.4, 150 mM Sodium chloride 0.1% SDS, 0.5% sodium deoxycholate, 1% Nonidet P-40, filtrate sterile Add Proteinase Inhibitor before use (1:200)
4x Tris/SDS pH6.8	0.5 M Tris (pH6.8), 0.4% SDS
4x Tris/SDS pH8.8	1.5 M Tris (pH8.8), 0.4% SDS
4x SDS-loading dye	200 mM Tris/HCl, pH6.8, 8% SDS, 40% Glycerol 0.4% Bromophenol blue, store at -20°C (2 M DTT add freshly: 40 $\mu$ l DTT to 160 $\mu$ l buffer)
2 M DTT	Dissolve 3.085 g DTT powder 10 ml H <sub>2</sub> O (-20°C)
10x Running buffer	100 ml 20% SDS, 60.6 g Tris, 288.2 g Glycine H <sub>2</sub> O up to 2000ml
10% APS	10g Ammonium peroxodisulfate powder 90 ml H <sub>2</sub> O, store at -20°C

---

10x TBS	171.4 g NaCl, 150 g Tris in 2000ml H <sub>2</sub> O, pH 7.5
KP-Buffer	3 g Tris-HCl, 3 g Glycine, 100 ml Methanol H <sub>2</sub> O up to 1000ml
AP I-Buffer	36.3 g Tris-HCl, 100 ml Methanol H <sub>2</sub> O up to 1000ml
AP II-Buffer	3 g Tris-HCl, 100 ml Methanol H <sub>2</sub> O up to 1000ml
Blocking solution	milk powder 1:10 in 1x TBST
ECL solution	Solution A and B mix: 1:1
Protein ladder	SeeBlueR Plus2 Pre-Stained Standard (Gibco, Invitrogen TM Cooperation, Carlsbad, CA)

**4.1.7.6 IHC**

Blocking solution	1% BSA, 10% FCS 3% donkey serum in PBST (0.1% Tween20 in 1x PBS)
Permeabilisation	0.1% TritonX-100 100 mM Glycine, in H <sub>2</sub> O (mouse embryo) 100% ice cold methanol (cell culture)

**4.1.7.7 ISH**

DEPC-H <sub>2</sub> O	Add DEPC 1:1000 to H <sub>2</sub> O stir overnight, autoclave
Heparin	50 mg/ml in DEPC-H <sub>2</sub> O
tRNA	20 mg/ml in DEPC-H <sub>2</sub> O
SDS	10 g SDS in DEPC-H <sub>2</sub> O
20xSSC	175 g Sodium Chloride, 88.2 g Sodium citrate in 1000ml DEPC-H <sub>2</sub> O
Prehybridisation buffer	50% Formamide, 5 x SSC, p5.4 1% SDS, 50 µg/ml yeast tRNA 50 µg/ml Hepairne, store at -20°C
Hybridisation buffer	1 µg RNA probe in 1 ml prehybridisation buffer

#### 4. Materials and methods

---

Solution I	50% Formamide, 5 x SSC, p5.4 1% SDS in MiliQ H <sub>2</sub> O, -20°C
Solution II	50% Formamide, 2 x SSC, p5.4 0.2% SDS in MiliQ H <sub>2</sub> O, -20°C
MAB	100 mM Maleic acid, 150 mM Sodium Chloride 2 mM Levamisole 0.1% Tween-20 in MiliQ H <sub>2</sub> O, pH 7.5
MAB block	2% Boehringer Mannheim blocking reagent in MAB
Antibody solution	10 mg embryo powder in 5 ml MAB block, vortex and incubate 30 min at 70°C, cool down on ice. Then add 50 µl sheep serum, 4 µl a-Dig Alkaline Phosphatase (1:5000), incubate 1 hour at 4°C. Centrifuge with 5000 rpm for 10 min at 4°C. Transfer supernatant to new tube, add 154 µl sheep serum and dilute to 20 ml MAB block.
TNT	10 mM Tris pH7.5, 0.5 mM Sodium Chloride, 0.1% Tween-20 in MiliQ H <sub>2</sub> O
NTMT	100 mM Tris pH9.5, 50mM MgCl <sub>2</sub> 100mM Sodium Chloride, 0.1% Tween-20 100 µl Levamisolein in MiliQ H <sub>2</sub> O

#### 4.1.7.8 FACS

FACS buffer (50 ml)	FCS 5ml 0.5M EDTA 100µl PBS 45 ml
Collagenase Buffer	2mM EDTA 2 mg/ml collagenase in PBS

### 4.1.8 Antibodies

**Table 4.1 Primary antibody:**

Name	Species	Dilution	Producer
Actin Ab-5	Mouse	1:5000	BD
Anti-tRFP	Rabbit	1:2500	Biocat/Evrogen
beta-Catenin	Mouse	1:2000	BD
Brachyury	Goat	1:1000	Santa Cruz
E-Cadherin	Mouse	1:2000	NEB
Foxa2	Goat	1:1000	Santa Cruz
Foxa2	Rabbit	1:1000	abcam
GAPDH	Mouse	1:6000	Merck Biosciences
GFP	Chicken	1:1000	Aves Labs
GFP	Rabbit	1:2000	Invitrogen
GFP	Rabbit	1:5000	abcam
N-Cadherin	Mouse	1:1000	BD
Oct-4	Mouse	1:1000	Santa Cruz
RFP	Rabbit	1:1000	Biotrend
Snail 1	Rabbit	1:1000	Cell signaling
Sox17	Goat	1:800	Acris/Novus
Troma-1	Rat	1:100	DSHB Hybridoma
Twist 1	Rabbit	1:400	abcam
ZEB1	Rabbit	1:2000	Acris/Novus
CD-24	Mouse	1:100	Biolegend
CD-133	Mouse	1:250	Biolegend
Isotype control	Mouse	1:100/250	Biolegend

**Table 4.2 Secondary antibody:**

Name	Conjugated	Dilution	Producer
Donkey-anti-chicken IgY	Cy2	1:800	Dianova
Donkey-anti-rabbit IgG 488	Fluorescent	1:800	Invitrogen
Donkey-anti-rabbit IgG 555	Fluorescent	1:800	Invitrogen
Donkey-anti-rabbit IgG 649	Fluorescent	1:800	Dianova
Donkey anti-mouse IgG 488	Fluorescent	1:800	Dianova
Donkey anti-mouse IgG 647	Fluorescent	1:800	Dianova
Donkey anti-goat IgG 488	Fluorescent	1:800	Invitrogen
Donkey anti-goat IgG 555	Fluorescent	1:800	Invitrogen
Donkey anti-goat IgG 633	Fluorescent	1:800	Invitrogen
Donkey-anti-rat IgG 549	Cy3	1:800	Dianova
Donkey-anti-rat IgG 649	Fluorescent	1:800	Dianova
Goat-anti-mouse IgG	HRP	1:20000	Dianova
Goat-anti-rabbit IgG	HRP	1:20000	Dianova
Rabbit-anti-goat IgG	HRP	1:20000	Dianova
Rabbit-anti-rat IgG	HRP	1:20000	abcam



## **4.2 Methods**

### **4.2.1 Molecular biology**

#### **4.2.1.1 Isolation of DNA**

##### **Plasmid preparations according to QIAGEN plasmid Kits**

To isolate plasmid DNA out of bacterial culture, a Mini or Maxi Prep was done using the QIAGEN Mini/Maxi Kit. Bacteria suspension was added to the LB medium with antibiotic and cultured at 37°C overnight with shaking. Bacteria were collected by centrifuging and further processed according to the Mini/Maxi Kit manufacturer's protocol. DNA was resuspended in TE buffer and stored at -20°C for long term storage. The concentration of DNA is calculated by measuring the extinction at 260 nm with a photometer (NanoDrop).

##### **Isolation of genomic DNA from 96-well plate cultured cells**

Let the cells grow until they are confluent. Before adding lysis buffer, these cells were washed twice with PBS-Mg<sup>2+</sup>/Ca<sup>2+</sup>. 50 µl lysis buffer mixed with 100 µg/ml proteinase K was added per well. The plate was incubated with a humid atmosphere at 60°C for overnight. The next day, 150 µl NaCl was added to 10 ml icecold ethanol. For DNA precipitation, 100 µl NaCl/Ethanol mixtures were added to every well and incubated for 30 mins at RT. Then carefully invert the plate to remove the liquid, and out it on a paper towel to remove the rest liquid. The DNA was washed three times with 150 µl 70% icecold ethanol. The DNA could be stored in 70% ethanol at -20°C. After removing ethanol, invert the dish and allow it to dry for 10-15 min at RT. The DNA was then dissolved in 25 µl TE buffer or H<sub>2</sub>O for 1 h at 37° C or 4 °C overnight with shaking.

#### **4.2.1.2 Restriction analysis of DNA**

##### **Restriction digestion of genomic DNA from ESCs in 96-well plates for Southern blotting**

The mixture (see below) except DNA was added to each well of 96-well plate, and incubated overnight at 37° C. The plate was sealed with parafilm to prevent evaporation and incubated in a humid chamber.

#### 4. Materials and methods

---

25 µl	DNA
0.40 µl	100 x BSA
0.40 µl	100 mM spermidine
0.25 µl	RNaseA 1 mg/ml
4.00 µl	10 x enzyme buffer
2.50 µl	Enzyme 10 U/ µl
7.45 µl	H <sub>2</sub> O
<hr/>	
40 µl	

#### **Restriction digestion of plasmid DNA for electroporation (10 cm plate)**

30 µl	DNA (30 µg)
4 µl	100 x BSA
10 µl	10 x buffer
1.00 µl	Enzyme 10 U/µl
55 µl	H <sub>2</sub> O
<hr/>	
100 µl	

The digest was incubated at 37°C for minimal 1.5 h or overnight.

#### **4.2.1.3 DNA purification**

Same volume of phenol was added to the dissolved DNA, vortexed for 1 min, and centrifuged with full speed for 1 min. The aqueous upper phase was transferred to a new Eppendorf tube, and the same step was repeated. After the upper phase transferring to the new tube, the same volume of chloroform was added, vortexed for 1 min, and centrifuged with full speed for 1 min. Then the upper phase was transferred to a new Eppendorf tube, and mixed with 0.7 volume of Isopropanol as well as 0.1 volume of NaAcetate (3M). The mixture was stored at -20 ° C for 20 min, and then centrifuged for 10 min with full speed. The pellet was washed with 500 µl 70% ethanol, and centrifuged again with full speed for 5 min. Then let the pellet dry for 5 min. The DNA pellet was resuspended in 50 µl H<sub>2</sub>O at 37° C for 30 min. DNA concentration was measured by Nanodrop.

#### **4.2.1.4 Agarose gel electrophoresis**

Different concentrations (0.8 - 2%) of agarose gels were prepared according to the size of DNA. The agarose gels were prepared by melting agarose in 1 x TAE (Tris-acetate, EDTA) buffer. 5 µl Ethidium Bromid (EtBr) which intercalates into double-

stranded DNA was added to per 100 ml melting agarose. DNA samples mixed with loading buffer (5:1) as well as one DNA marker were loaded into the wells. The gel was run at approximately 120 V in 1x TAE buffer for 40 minutes. The fragment pattern of DNA was visualized under UV light and photographed using Gel documentation system (BioRad, UV-Transilluminator).

#### 4.2.1.5 Extraction of DNA fragments from agarose gel

After gel electrophoresis, the agarose gels containing the correct DNA fragments were cut out under UV light, and transferred to an Eppendorf tube. The DNA was isolated according to QIAquick Gel Extraction Kit.

#### 4.2.1.6 Dephosphorylation of linearized DNA

A typical mix for dephosphorylation of linearized plasmid DNA is as followed:

1.0 µg	DNA
1.0 µl	10 x enzyme buffer
1.0 µl	Alkaline phosphatase (1U/ µl)
7.0 µl	H <sub>2</sub> O
<hr/>	
10.0 µl	

#### 4.2.1.7 Ligation

A typical mix for ligation of vector and insert DNA is as followed:

1.0 µl	10x T4 ligation buffer
1.0 µl	T4 ligase (10U/ µl; NEB)
1.0 µl	vector DNA
0.5 µl	insert DNA
6.5 µl	H <sub>2</sub> O

The calculation of the volume of vector and insert DNA is as followed:

Ratio: vector/insert = 1/3

vector concentration:  $c(v)$  [ng/µl] = x ng/µl      size (v) [bp] = y bp

insert concentration:  $c(i)$  [ng/µl] = a ng/µl      size (i) [bp] = b bp

used amount of vector-DNA: 100-400 ng      > used for ligation: v ng / w µl

used amount of insert-DNA: intron [ng] = 3 \* (b bp / y bp) \* w µl > intron DNA [µl]

The ligation is incubated overnight at 14° C.

#### 4.2.1.8 Isolation of RNA

RNA extraction of either cultured cell or mouse tissue was performed according to QIAGEN miRNeasy kit. For all related work, RNase-free solutions, pipette tips, and Eppendorf tube were used. Isolated RNA was dissolved in RNase-free water. RNA was stored immediately at -80°C for long-term storage. RNA concentration was measured with a NanoDrop photometer. The ratio of 260/280 and 230/260 between 1.9 and 2.1 was considered to be a good RNA preparation.

#### 4.2.1.9 Reverse transcription

Reverse transcription was done according to QIAGEN QuantiTect Reverse Transcription Kit, or High Capacity RNA-to-cDNA Kit. The procedure for QuantiTect Reverse Transcription Kit is as following:

##### Reaction components I

2 µl gDNA Wipeout Buffer

RNA sample

NNase-free water

---

Total volume 14.00 µl

##### Reaction components II

1 µl Reverse transcription master mix

4 µl RT buffer, 5x

1 µl RT Primer Mix

14.00 µl Template RNA

---

Total volume 20 µl

Reaction components I was incubated on ice for 2 min at 42°C. This step is to eliminate genomic DNA. The reverse transcription was done by incubating reaction components I and II mixture at 42°C for 15 min, and inactivated by 95°C for 3 min. The cDNA samples were store at -20°C for longer use.

Reverse transcription for miRNA was performed according to TaqMan Small RNA Assays with following procedures:

---

0.15 µl 100mM dNTPs (with dTTP)
1.00 µl 50 U/µL MultiScribe™ Reverse Transcriptase
1.50 µl reverse transcription buffer (10 x)
0.19 µl RNase inhibitor, 20 U/µL
4.16 µl Nuclease-free water
<hr/> Total volume 7.00 µl

Reverse transcription was performed with the following program:

16° C	30 min
42° C	30 min
85° C	5 min
4° C	∞

#### 4.2.1.10 PCR

##### I. Semi-quantitative PCR

The typical reaction scale and program for PCR is as followed:

##### PCR reaction mix: Program for PCR:

1 µl cDNA	94°C	5min	
2 µl 10 x Taq buffer with (NH <sub>4</sub> ) <sub>2</sub> SO <sub>4</sub>	94°C	1min	
2 µl 25 mM MgCl <sub>2</sub>	X°C	30 s	35 cycles
1 µl 10 mM dNTPs	72°C	45 s	
1 µl forward primer			
1 µl reverse primer	72°C	10min	
0.2 µl Taq polymerase	4°C	∞	
11.8 µl Nuclease-free water			
<hr/> 20 µl			

**Table 4.3 Primer pairs for RT-PCR:**

Gene	Primer pairs (forward & reverse)	T <sub>m</sub> (°C)	Cycles
<i>Afp</i>	5'-GCTCACACCAAAGCGTCAAC -3' 5'-CCTGTGAACTCTGGTATCAG -3'	56	40
<i>Alb1</i>	5'-GCTACGGCACAGTGCTTG -3' 5'-CAGGATTGCAGACAGATAGTC -3'	56	40
<i>GAPDH</i>	5'-TGGATGCAGGGATGATGT-3' 5'-ATTCAACGGCACAGTCAA-3'	56	28

## II. Real-Time PCR

The mouse miR-335, *Foxa2*, and *Brachyury* transcripts were quantified by real-time RT-PCR using the corresponding TaqMan Gene Expression Assays (Applied Biosystems). U6 and  $\beta$ -actin were used as endogenous normalization controls for miRNA and protein-coding genes, respectively (Applied Biosystems). Endogenous mRNA levels of *Gapdh*, *Hhex*, *Mest*, *Sfrp5*, *Cdx1*, *Cdx2*, *Sox17*, and *T* were measured using SYBR Green PCR Master Mix (Applied Biosystems). The mRNA levels of the other genes were quantified by array card assay.

### qPCR with SYBR® Green assay

#### qPCR reaction mix:

cDNA template	1 $\mu$ l	1 $\mu$ l
2x SYBR® Green PCR Master Mix	12.5 $\mu$ l	5 $\mu$ l
Forward primer	1 $\mu$ l	0.8 $\mu$ l
Reverse primer	1 $\mu$ l	0.8 $\mu$ l
Nuclease-free water	6 $\mu$ l	2.4 $\mu$ l
	25 $\mu$ l	10 $\mu$ l

**Table 4.4 Primer pairs for qPCR assays with SYBR® Green system:**

Gene	Primer pairs (forward & reverse)
<i>Cdx1</i>	5'-GCTCTGCACTCATGGAAGAC-3' 5'-GATCTTTACCTGCCGCTCTG-3'
<i>Cdx2</i>	5'-GGAAGCCAAGTGAAAACCAG-3' 5'-CCAGCTCACTTTTCCTCCTG-3'
<i>GAPDH</i>	5'-TGGATGCAGGGATGATGT-3' 5'-ATTCAACGGCACAGTCAA-3'
<i>Hhex</i>	5'-GAGGTTCTCCAACGACCAGA-3' 5'-GTCCAACGCATCCTTTTTGT-3'
<i>Mest</i>	5'-GCTCTGCACTCATGGAAGAC-3' 5'-GGCGATCACTCGATGGAAC-3'
<i>Sfrp5</i>	5'-CTGAGCCCTAGTCATTGCATACTG-3' 5'-TTAATGCGCATCTTGACCAC-3'
<i>Sox17</i>	5'-GGTCTGAAGTGCGGTTGG-3' 5'-TGTCTTCCCTGTCTTGGTTGA-3'
<i>T</i>	5'-CCACCGCTGGAAATATGTG-3' 5'-CAGCTATGAACTGGGTCTCG-3'

### qPCR with TaqMan Assays

#### qPCR reaction mix:

1.00 $\mu$ l TaqMan Assay (20 x)
1.33 $\mu$ l RT reaction product
10.00 $\mu$ l TaqMan <sup>®</sup> Universal PCR Master Mix II (2 x)
7.67 $\mu$ l Nuclease-free water
20.00 $\mu$ l

#### qPCR program:

95° C	10 min	
95° C	15 s	40 cycles
60° C	60 s	
95° C	15s	
60° C	60s	
95° C	15s	

**Table 4.5 Genes and assay IDs used in qPCR assays:**

Gene	assay IDs
<i>Foxa2</i>	Mm01976556_s1
<i>T</i>	Mm01318252_m1
<i><math>\beta</math>-Actin</i>	Mm00607939_s1
<i>miR-335</i>	000546
<i>U6</i>	001973

### PCRs on TaqMan Array Cards

48 PCR Assay IDs were ordered and placed within the TaqMan Array Card (Applied Biosystems) architecture by the manufacturer. Array card was loaded by mixing 50  $\mu$ L cDNA samples (minimal 200ng) with 50  $\mu$ L TaqMan Universal Master Mix (Applied Biosystems). After 2x1 min 1200 g centrifugation, the card was sealed and the inlet ports were removed following the manufacturer's instructions.

**Table 4.6 Gene lists and assay IDs in the array card:**

genes	Assay IDs
<i>Sox2</i>	Mm03053810_s1
<i>Pou5f1</i>	Mm03053917_g1
<i>Nanog</i>	Mm02384862_g1
<i>Klf4</i>	Mm00516104_m1
<i>T</i>	Mm01318252_m1
<i>Foxa2</i>	Mm01976556_s1
<i>Mixl1</i>	Mm00489085_m1
<i>Eomes</i>	Mm01351985_m1
<i>Gsc</i>	Mm00650681_g1
<i>Sox17</i>	Mm00488363_m1
<i>Cxcr4</i>	Mm01996749_s1
<i>Hhex</i>	Mm00433954_m1
<i>Lefty1</i>	Mm03053915_s1
<i>Lefty2</i>	Mm00774547_m1
<i>Sfrp5</i>	Mm01194236_m1
<i>Fzd5</i>	Mm00445623_s1
<i>Cer1</i>	Mm00515474_m1
<i>Cdx1</i>	Mm00438172_m1
<i>Cdx2</i>	Mm01212280_m1
<i>Cdx4</i>	Mm00432451_m1
<i>Sox7</i>	Mm00776876_m1
<i>Afp</i>	Mm00431715_m1
<i>Pdgfra</i>	Mm00440701_m1
<i>Shh</i>	Mm00436528_m1
<i>Nog</i>	Mm01297833_s1
<i>Chrd</i>	Mm00438203_m1
<i>Hand1</i>	Mm00433931_m1
<i>Hand2</i>	Mm00439247_m1
<i>Tbx5</i>	Mm00803518_m1
<i>Gata4</i>	Mm00484689_m1
<i>Gata6</i>	Mm00802636_m1
<i>Mesp1</i>	Mm00801883_g1
<i>Mesp2</i>	Mm00655937_m1
<i>Tbx6</i>	Hs00365539_m1
<i>Zeb1</i>	Mm00495564_m1
<i>Zeb2</i>	Mm00497193_m1
<i>snail1</i>	Mm00441533_g1



<i>snail2</i>	Mm00441531_m1
<i>Twist1</i>	Mm04208233_g1
<i>FoxC2</i>	Mm00546194_s1
<i>cdh1</i>	Mm01247357_m1
<i>CD24</i>	Mm00782538_sH
<i>Cldn6</i>	Mm00490040_s1
<i>FN-1</i>	Mm01256744_m1
<i>Vim</i>	Mm01333430_m1
<i>N-cadherin</i>	Mm01162497_m1
<i>Gapdh</i>	Mm99999915_g1
<i>18s</i>	Hs99999901_s1

All the qPCR reactions were performed either in ABI PRISM 7900HT sequence detection System (Applied Biosystems) or ViiA™ 7 real-time PCR system (Applied Biosystems) according to the manufacturer's instructions.

#### 4.2.1.11 Generation of the Foxa2-tagRFP targeting construct

The targeting construct was designed as shown in Figure 2.1. The targeting strategy of Foxa2-tagRFP involved the fusion of the ORF of Foxa2 (orange boxes) to fluorescent reporter gene tagRFP (yellow arrow). A previously established pBKs-Foxa2Ex3-Venus construct (Burtscher et al., 2013) was used and digested via SpeI and XbaI to generate pBKs-Foxa2Ex3. The tagRFP Sequence with translational stop codon was amplified subcloned between homology region of pBKs-Foxa2Ex3 via SpeI and XbaI resulting in pBKs-Foxa2Ex3-tagRFP. Next, the phospho-glycerate kinase (PGK) promoter-driven Neomycin (Neo) resistance gene cassette flanked by two flippase recognition target (FRT) sites was cloned from the PL452 via EcoRI and BamHI of the tagRFP sequence into the upstream product resulting in the pBKs-Foxa2Ex3-tagRFP-Neo. In the last step, the mini-targeting cassette was released by digest with Eco47III and KpnI and introduced into the pL254-Foxa2 (Burtscher et al., 2013) via bacterial homologous recombination in EL250 bacteria, resulting in the final targeting construct (Figure 2.1).

#### Primers for tagRFP cloning

EP-696: 5'-NNNGCGGCCGCGCCACCATGTCTAGAATGGTGTCTAAGGGCGAA GAGC-3'

EP-644: 5'-NNNACTAGTTCAATTAAGTTTGTGCCCCAGTTTGCTAGGGAG-3'

Foxa2 5' and 3' UTRs (black boxes) and coding region (orange boxes) as well as the predicted promoter sites (yellow boxes) and transcriptional start region (TSR, red boxes) are indicated. The location of the 3' external probe and appendent restriction sites EcoRV and SpeI are indicated. Homology regions to generate the targeting construct are indicated as 5' and 3' retrieval.

The targeting construct was electroporated into an available Brachyury (T)-GFP mESCs line. The positive clones, which were *Neo* resistant, were isolated as homologous events and further confirmed by Southern blotting. Flp recombinase-mediated excision was used to eliminate the PGK driven *Neo*-resistant cassette flanked by FRT sites and a 5' loxP site.

### 4.1.1.12 Transformation of bacteria

#### I. Transformation of bacteria using electroporation

The competent bacteria were thawed on ice and then mixed with 1  $\mu$ l DNA. The mixture was incubated on ice for 5 min. The mixture was then carefully transferred to a pre-cooled electroporation cuvette (0.1 cm) without bubbles. Electroporation was performed with a tension of  $U=2.5$  kV. The bacteria were immediately transferred into a prepared Eppendorf tube with 1 ml LB medium inside. The mixture was then incubated at 37 °C for 60 min with shaking (850 rpm) to allow the bacteria to regenerate. After centrifuge, the bacteria were put on LB plates supplemented with the appropriate antibiotics, and incubated overnight at 37°C.

#### II. Transformation of bacteria using heat shock

1  $\mu$ l of vector DNA was mixed with the bacteria and incubated on ice for 20-30 min. Then the mixture was incubated at 42°C for 90 s, and then incubated on ice for 5 min. The mixture was then transferred to an Eppendorf tube with 900  $\mu$ l LB medium, and incubated at 37 °C for 60 min with shaking (850 rpm) to allow the bacteria to regenerate.

#### **4.2.1.13 Southern blot**

##### **I. Gel electrophoresis:**

The DNA was isolated and digested in 96-well plate. The digestion was stopped by adding 5 µl Orange G. The mixture was loaded in 0.7-0.8 % agarose gel, and applied with a maximal V for 10 min. The gel was cut by three parts, separated with some space, and run for overnight with 35-40 V. The next day, the gel was photographed under UV-light next to a ruler, which serves as reference length. Repetitive sequences can be easily detected by a distinct band under UV-light if the digest is complete.

##### **II. Blot:**

The gel was depurinated by incubating it in depurination solution for 10-15 min with shaking. After washed with MilliQ water, the gel was incubated in denaturation solution for 30-60 min while shaking. For neutralization the gel was incubated in a solution for 30-60 min while shaking. After washing with H<sub>2</sub>O, the blot was built up in 20 x SSC and the transfer was performed overnight at RT. The next day, the blot was taken apart and the slots of the gel were marked on the membrane with a pen. Afterwards, the membrane was dried between Whatman paper at RT, and then were incubated in oven at 80°C for 30 min to cross-link the DNA to the membrane.

##### **III. Hybridisation:**

###### **Prehybridisation**

Hybridization buffer 30 ml per membrane was pre-heated to 65°C. The membrane was rolled and put into the big glass flask and the pre-heated hybridization buffer was poured into the tube. The membrane was pre-hybridized on the rotor at 65°C for at least 1.5 to 3 h.

###### **Radioactive labelling of the probe and hybridisation**

Foxa2 3' probe Southern blot, length: 732bp

##### Sequence:

5'-CTGGATATGCTCTAGAAAGGCAGAAGTTTACAGTTTTTTTTAATATCAGGCCTCCTTTC  
TAGTCAGTGAACCTTAGACTGGGTTTACCAATTTTGGTGCATGGCTCTTCCAGCTACTTGA  
AGCATTGCCCCCTAGACCTTCCTGTGCCATTGAGACTACCTGGCTCTAGGTTGTGCC  
GGGAGGGCAGCCTGTCTCAGTCTCACAGGTGTTATCCAGGTATTGGGAAACCTTGCTA  
GGCTAGGAACGATGAGCCACCTAATCTGGGGAAACATTTTAACATTGGGAATTGGGTAT  
AATTGCATAGTTAAGGGTAACCCCAAATCTTTTATTAAGAAGTTATTCTGTGGGTGGGG  
AGATAGGGAGGGATGGAAGGGTGCCCTGAGCAGCTTAGCAAATGACTCCCAAAGTAGT  
GAAATCCCAGTGTCTCAGGAATGGTGTCTCCCTTCTACCAGCCAGGGCAAAGCTGTTTG  
TTAGCTTAGGAAGCTCCTATAGGCCAAACCACACTTGAGGCCCAGGGACTGAATGGGTAT  
TTTGTGAGCCTCCAGGAAAATACAAAGACCCCAAATAAAACCTCACCAATCATTTCCACC  
ACTCTGCAGATTTTCAAATTGACGGGTAAGTGTAGAGGAGGTTCGTGTTTTGCAAAGG  
AGCCTCCTCACGCTGACCTGCATCTCCTGCCCTTGAAGCTGTCCCTCCCGCCCGCCCC  
CAGTCTGACTTTCCATAGGCCATTC-3'

The radioactive labelling of the probe was performed as following steps. Approximately 25 ng of linearized DNA probe was diluted with dist. H<sub>2</sub>O to a final volume of 23 µl in an Eppendorf tube. 10 µl of random oligonucleotides were added and the mixture was denatured in a water bath for 5 min at 100°C. Subsequently, the mixture was put on ice, shortly centrifuged, and put it back on ice. 10 µl pre-cooled 5x dCTP buffer was added. Radioactive labelled dCTP (50 µCi) and 1 µl Klenow-enzyme (pink, 5 U) were added, the mixture carefully mixed, shortly centrifuged, and labelled at 37°C for 30-60 min. To stop the reaction, 2 µl 0.2M EDTA (pH 8.0) were added, and kept on ice. Then 49 µl TE were added into the mixture. Afterwards, the micro-spin columns were prepared, and the seal from the columns were removed. Then the columns were placed into 1.5 ml Eppendorf tube, and centrifuged at 760 G (3000rpm) for 2 min. The columns were placed into fresh Eppendorf tube. The 100 µl labelling reaction mixture was carefully transferred to prepared micro-spin columns, and centrifuged at 760 G (3000rpm) for 1 min. Non-incorporated nucleotides remained on the column. The radioactive probe was kept on ice, and 1 µl of the samples was used to measure the activity. For denaturation of the probe, 500 µl SS-DNA (10 mg/ml) were denatured at 100°C in a water bath for 10 min. Then the SS-DNA was put in a 50 ml Falcon tube, and stored on ice. The labelled probe was added with a final concentration of 1x10<sup>6</sup> counts/ml hybridization buffer. While carefully swirling, 50 µl of 10 N sodium hydroxyl were added to denature the samples, and then 300 µl 2 M Tris (pH 8.0) as well as 475 µl 1 M HCl were pipetted drop wise into the tubes for neutralization.

---

After the pre-hybridization step the hybridization buffer were discarded and the radioactive-labelled DNA probe was pipetted into the hybridization tube. The tubes were put back into the oven at 65°C for overnight while rotating.

### **Washing membrane**

The following day the solution was removed and discarded into P<sup>32</sup> liquid waste. The membranes were washed with preheated 2xSSC/0.5% SDS buffer (65°C) for 30 min while shaking. If membranes are still hot, the membranes were washed with 1xSSC/0.5% SDS buffer. The membranes were wrapped tightly in Saran wrap and fixed in a film cassette, and stored at -80°C for 1-3 days depending how hot the membranes are.

#### **4.2.1.15 Generation of the Luciferase Reporter vector for Foxa2 and Sox17**

3' UTR of Foxa2 and Sox17 were amplified by PCR using mouse genomic DNA. After gel electrophoresis, DNA fragments of correct size were extracted from the gel and ligated with pKS vector. After transformation, single positive colony was picked and further cultured into LB medium containing Amp. Then the plasmid DNA was extracted from the bacteria using mini preparation kit (Qiagen) according to the manufacturer's protocol. The plasmid DNA of interest was identified by restriction digestion and gel electrophoresis. The 3' UTR of Foxa2 and Sox17 were then cut out from pKS vector with restriction endonuclease NotI and XhoI, purified using the gel extraction kit, and ligated with NotI and XhoI-digested psiCHECK<sup>TM</sup>-3 (Promega) using T4 DNA ligase (Roche) according to the manufacturer's protocol. The psiCHECK<sup>TM</sup>-3-Foxa2 or Sox17 3'UTR vector was amplified and isolated in bacteria later.

We mutated Foxa2 UTR according to miR-335 binding sites. The mutated Foxa2 UTR was then amplified, and ligated with pKS vector with restriction endonuclease NotI and PstI. After transformation, the mutated Foxa2 UTR was cut out from the pKS vector with restriction endonuclease NotI and XhoI, and ligated with NotI and XhoI-digested psiCHECK<sup>TM</sup>-3 (Promega). We use the same approach for generation of Sox17 mutated UTR.

**Table 4.7 Primers for 3'UTR cloning:**

Gene	Forward primer and Reverse primer
<i>Foxa2</i> 3'UTR	5'-NNN CTCGAGGAAGATGGCTTTCAGGCCCTGCTAGCTC-3' 5'-NNN GCGGCCGCATTCTAGCCAGAACACACATTTATAAGC-3'
<i>Foxa2</i> 3'UTR mutant	5'-NNNCTGCAGAGTTTGACGACTCAAGTTCTAATCTATTGCTGTTGTTGCA GAAAAGTCTGACTTTAAAAACAAACAAACAAAAACGCATCAGAGT CTGACGGTGTAACCATGTAGTTTTAACAG-3' 5'-NNN GCGGCCGCATTCTAGCCAGAACACACATTTATAAGC-3'
<i>Sox17</i> 3'UTR	5'-NNN CTCGAGCGGTTGCCGACCCGACCTGAGGGCCAGAA-3' 5'-NNN GCGGCCGCCACTAACAGTCACAACACAACTTTATTTTG-3'
<i>Sox17</i> 3'UTR mutant	5'-NNN CTCGAGCGGTTGCCGACCCGACCTGAGGGCCAGAA-3' 5'-GTGATTGTGGGAGCAAGTCCCTCTTCGCATTTAAATCATATTTCTCGT GTAGCCCCTCAACTGTTCAAGTGGCAGAC-3'

#### 4.2.1.16 Luciferase Reporter Transfection and Dual Luciferase Assay

The full-length 3'-UTRs and mutated 3'-UTRs of miR-335-5p target genes *Foxa2* and *Sox17* were amplified and individually cloned into the psiCHECK™-3 (Promega) dual luciferase reporter vector. One day before transfection, the HEK293T cells were split to 24-well plate with the density of  $8 \times 10^4$  cells /24-well. The following day, HEK293T cells were co-transfected with each reporter construct and synthetic pre-miR-335 (Ambion) at the following concentration: 100 ng of the UTR reporter, 30 pmol miRNA precursor molecules and 25  $\mu$ l of PEI (polyethylenimine, Polysciences) per 24-well. Then the solution was added to the feeder medium, up to 500  $\mu$ l per 24-well plate. Incubate at least 6 hour before changing the medium. Cells were lysed 40 h after transfection and the ratio of Renilla to firefly luciferase was measured with the dual luciferase assay (Promega). Pre-miR negative control#2 (Ambion) and Pre-miR-132 (Ambion) were used as control. Significance was estimated by performing a multiple comparison test using One-Way ANOVA.

#### 4.2.1.17 Generation of the miR-335 and miR-335 sponge-3P overexpression vector

The generation of miR-335 expression vector has been done previously (Uetzmann, 2009). The ORF of green fluorescent protein (GFP) in the published vector (Chung et

al., 2006) was exchanged by the complex of ORF of cyan fluorescent protein linked to the histon 2B localization sequence (H2B-CFP) and a puromycin resistance gene.

For the generation of sponge vector, a pCAG-H2B-BFP-2A-IRES-Puro polyA vector was used. The oligos contained each three miR-335-3p binding motives were aligned and phosphorylated, ligated into the dephosphorylated pCAG-H2B-BFP-2A-IRES-Puro polyA vector using the *Nsi*I site. HindIII digestion was used to check the correct orientation. Sub-cloning of this oligos into the *Nsi*I site was repeated six times resulting in a total of 18x mir335-5p binding sites in the 3'-UTR of the sponge-3P vector.

**Table 4.8 Primers for cloning and sequencing:**

name	Sequence
5' sponge fwd	5'- TGCACGACATTTTTTGTTCCTGCTCTTGAACCTAAGCTTGATACATTTTT CGTTAATTTTGCTCTTGACCGTGATACACATTTTTTGTTGCTCTTGAC AGACTGCA-3'
5' sponge rev	5'- GTCTGTCAAGAGCAACACAAAAAATGTGTATCACGGTCAAGAGCAAAT TAACGAAAAATGTATCAAGCTTAGTTCAAGAGCAGGAAACAAAAAATGT CGTGCAATGCA-3'
3' sponge fwd	5'TGCACGGGTCAGGAGCTGAATGAAAAACTAAGCTTGATGGTCAGGA GCTCTGAATGAAAAACCGTGATACGGTCGGGAGUTC GGATGAAAAACA GACTGCA-3'
3' sponge rev	5'- GTCTGTTTTTCATCCGAACCTCCCGACCGTATCACGGTTTTTCATTCAGA GCTCCTGACCATCAAGCTTAGTTTTTCATTCAGCTCCTGACCCGTGCA TGCA-3'
Sequencing sense	5'-GCTCCGCACAGATTTGGGAC-3'
Sequencing reverse	5'-CCCACGGTGGCCATTTGTTC-3'

#### **4.2.1.18 Protein biochemistry**

##### **I. Protein extraction**

###### **Protein extraction from whole lysate**

Before lysing the cells, the cells were washed once with cold PBS. After removing PBS, the cells were lysed in 100 - 200  $\mu$ l (for 6-well plate cultured cells) lysis buffer supplemented with protein inhibitor (1:200) on ice. A cell scraper was used to abrade the cells. The mixture was transferred to an Eppendorf tube, and centrifuged at 14000 rpm at 4 °C for 10 min to pellet insoluble constituents, such as nuclei and parts of the cell membrane. The supernatant was transferred to a new Eppendorf tube and stored at -20°C for long storage.

###### **Protein extraction from mouse tissue**

Mouse embryos were dissected quickly, and immediately kept in cold PBS. The mixture was centrifuged, and the supernatant was removed carefully. The remaining tissue was further homogenized and sonicated by 200  $\mu$ l the mixture of 4x SDS and 2M dithiothreitol (DTT). After centrifugation, the supernatant was transferred to a new Eppendorf tube and stored at -20°C for long storage.

##### **II. Determination of protein concentrations by Bradford assay**

1  $\mu$ l protein sample was diluted 1:10 with MilliQ water, and mixed with 990  $\mu$ l Bradford. The mixture was incubated at RT for 10 min. The mixture was measured against a blank value, and the relative value was calculated based on the directly proportional absorbance measured.

##### **III. Western blot**

###### **Denaturing SDS-polyacrylamide gel electrophoresis**

For 10% separating gels were made by the following composition:



	2 gels	4 gels
Acrylamide	5 ml	10 ml
4x Tris/SDS buffer, pH8.8	3.75 ml	7.50 ml
H <sub>2</sub> O	6.25 ml	12.5 ml
TEMED	20 $\mu$ l	40 $\mu$ l
10% APS	150 $\mu$ l	300 $\mu$ l

The mixture was immediately filled between two glass plates and covered with isopropanol to achieve a sharp and straight border without bubbles. The isopropanol was decanted after polymerization. The gels were then washed with MilliQ water, and the water was completely sucked off with a paper towel.

Four collecting gels were prepared by the following mixture:

	2 gels	4 gels
Acrylamide	0.65 ml	1.3 ml
4x Tris/SDS buffer, pH6.8	1.25 ml	2.5 ml
H <sub>2</sub> O	3.1 ml	6.2 ml
TEMED	10 $\mu$ l	20 $\mu$ l
10% APS	50 $\mu$ l	100 $\mu$ l

The mixture was immediately filled between the two glass plates till the upper rim was reached then the comb was carefully inserted to avoid bubbles. The gels were put into the gel chamber filled with 1x Tris glycine running buffer after removing the combs. The gel pockets were rinsed with running buffer to remove bubbles and remaining gel filaments.

### Western blot

About 20-30  $\mu$ g proteins from each sample were mixed 1:3 with the mixture of 4x SDS loading buffer and 2 M DTT. The mixture was denatured at 95°C for 4 min and chilled on ice and then loaded on the prepared gels. A protein marker was loaded to estimate the size of the protein bands. Electrophoresis was performed at

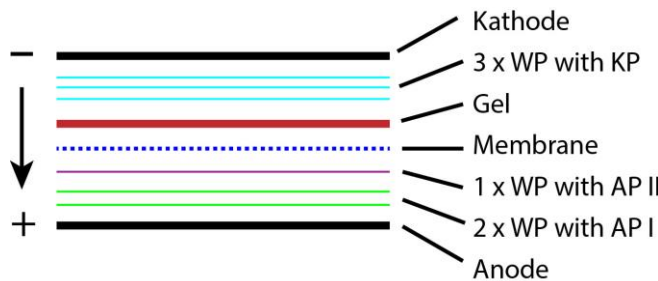
#### 4. Materials and methods

approximately 120V for about 1.5 hr until the proteins were well separated according to their size.

After gel electrophoresis, the gels and the PVDF membranes were incubated with the following steps:

Gel	20 min in KP
Membrane	15 s in Methanol 2 min in H <sub>2</sub> O 5 min in AP II-Buffer

The blot was set up as followed:



**Figure 4.1 Western blot setup**

Blotting was performed at 0.22A for 30 min with the transfer system (Bio-RAD). The Whatman filterpapers (WP) containing different solutions were used to build the blot.

After blotting, the membrane was blocked with blocking solution (4 %

skim milk in 1x TBS-T) for 1 hr at RT. Afterwards, the membrane was incubated with the primary antibody (diluted in blocking solution) at 4 °C overnight while shaking. The next day, the membrane was washed 3 x TBST for 10 min, and then incubated with the secondary antibody in blocking solution for at least 1 hr at RT, and washed again three times. Afterwards, the membrane was covered and incubated with ECL detection reagent for 1 minute. The membrane was transferred on fresh foil and placed in a film cassette, and then exposed to a BIOMAX film for several seconds to several minutes in a dark environment, depending on the intensity of the signal. Exposed films were developed using a developing machine.

#### IV. Immunohistochemistry (IHC)

**IHC of cultured cells**

Cells were washed with 2 x PBS, fixed with 4% paraformaldehyde (PFA) for 5 min, and then permeabilized in cold methanol for 5 min. Afterwards, cells were blocked with blocking solution (3% donkey serum, 10 % fetal bovine serum (FBS), 0.1% BSA in PBS containing 0.2 % Tween-20 (PBST)) for at least 1hr, and incubated with primary antibodies at 4°C overnight while shaking. The next day, the cells was washed 3 x PBST for 10 min, and then incubated with secondary antibodies in blocking solution for 1 hr at RT, and washed again three times. Images were acquired with a Zeiss Axiovision inverted microscope with a 20xobjective.

**IHC of whole mount embryos**

Embryos were isolated in PBS<sup>+</sup> (PBS containing Mg<sup>2+</sup> and Ca<sup>2+</sup>), fixed for 20 min at RT in 2% PFA in PBS<sup>+</sup>, and then permeabilized in 0.1 M glycine /0.1% Triton X-100. After 1-2 hrs incubating in blocking solution, embryos were incubated with the primary antibody at 4°C overnight while shaking, and incubate another 3 hrs at RT the next day. After 3 x PBST<sup>+</sup> washes for 15 min, embryos were incubated with secondary antibodies in blocking solution for minimum 3 hrs, and washed again three times. Afterwards, embryos were equilibrated in 15% and 30% glycerol in PBST<sup>+</sup> for 5 min, separately. Image acquisition was performed on a Leica SP5 confocal microscope.

## 4.2.2 Cell culture

### 4.2.2.1 Cell lines

IDG3.2	murine ES cell line (F1); genetic background 129S6/SvEvTac x C57BL/6J (Hitz et al., 2007)
FVF; SCF	(Burtscher et al., 2012; 2013)
T <sup>GFP/+</sup> ; Foxa2 <sup>tagRFP/+</sup>	ES line generated in this thesis
MEF	primary murine embryonic fibroblasts isolated on E13.5
Wnt3a feeders	Wnt3a over-expressing NIH3T3 cells, which is a murine embryonic fibroblast cell line (Todaro and Green, 1963)
HEK293T	human embryonic kidney cells stably expressing the T-large antigen of SV40 (Graham et al., 1977)

### 4.2.2.2 Cell culture media

Feeder cell medium	Dulbecco's Modified Eagle Medium (D-MEM) supplemented with 10 % PAA, 1x MEMs non-essential amino acids (NEAA), 1x L-glutamine, 10 mM Pen/Strep, 20 mM HEPES and 1% 2-mercaptoethanol.
ES cell medium	Dulbecco's Modified Eagle Medium (D-MEM) supplemented with 15 % PAN/PAA, LIF, 1x NEAA, 1x L-glutamine, 10 mM Pen/Strep, 20 mM HEPES and 1% 2-mercaptoethanol.
ES cell freeze medium (1x)	50% PAA, 40% ES cell medium, 10% DMSO
Feeder cell freeze medium (1x)	50% PAA, 40% feeder cell medium, 10% DMSO
Differentiation medium	SFO-3 medium/Endoderm differentiation medium (EDM) supplemented with 3 ng/ml or 12ng/ml human ActA, 10 mM Pen/Strep. Bmp4 (7.5 ng/ml) was added for the mesoderm differentiation.

---

#### 4.2.2.3 Maintenance of mouse ES cells

Mouse embryonic stem cells were maintained and passaged on 0.1% gelatine coated plates with murine embryonic fibroblasts (MEF) feeders in ESC medium. MEF feeders were expanded to passage 4 and treated with mitomycin C before using. ESCs were cultured at 37°C with 5-7% CO<sub>2</sub> in a humid incubator. Medium was changed every other day. For passaging 10 cm plate ES cells, the cells were first washed with 10 ml PBS and then incubated with 1 ml 0.05% trypsin-EDTA and 1 ml PBS at 37° C for ~4 min depending on the cell line. Trypsinization was stopped by adding 6 ml ES medium. The cells were resuspended by pipetting gently, and transferred into a 50 ml falcon tube and centrifuged at 1200 rpm for 5 minutes. The supernatant was removed carefully and the cell pellet was resuspended with fresh ES medium and added to new plates.

#### 4.2.2.4 Generate of new ES cell lines

ESCs were cultured and trypsinized as described before. Single cell suspension was washed with 10 ml PBS, and centrifuged at 1200 rpm for 5 minutes. The cell pellet was resuspended in 700 µl cold PBS and mixed with 100 µl DNA plasmids (25 µg). The suspension was transferred to a cuvette, and the electroporation was performed by GenePulser Xcell with 220V and 500 µF. The cells were kept on ice for 10 min and transferred to MMC-treated MEF dish with pre-warmed ES medium. The ES cells were selected with 300µg/ml Geneticin or 1µg/ml puromycin according to different vector constructs for 7 days. Colonies that looks compact and round were picked under the stereo in 20 µl PBS. Single colony was transferred to a coated conical 96-well with 55 µl PBS inside, and treated with 30 µl of trypsin/EDTA, then incubated for 15 min in cell culture incubator. After trypsinization, the cell suspension was resuspended and transferred to two prepared 96-well plate with 50 µl for each. One plate with MEF feeder layer is the master plate for freezing. After 2-4 days culture in the incubator, those cells were treated with 40 µl trypsin/EDTA at 37° C for 5 min. 60 µl of ESC medium was added to stop the trypsinization. Then the single cell suspension was mixed with 100 µl of ES freezing medium (2x), and stored at -80 °C. The ES cells in the template plate were grown until the medium turned to yellow from orange within one day. Afterwards, DNA was isolated for the southern blot analysis.

After southern blot, the positive ES cells were expanded from 96 well to 10 cm plates, and frozen with 1x ES freezing medium for future experiment.

##### **4.2.2.5 Differentiation of mESCs**

Prior to differentiation, wnt-3a feeders were seeded at a density of  $\sim 5 \times 10^4$  cells /24-well. ESCs were passaged onto gelatine-coated plates for 30 min twice to remove feeders. After washing with PBS, ESCs were transferred to the Wnt3a-expressing feeder plate with a seeding density of  $2 \times 10^5$  cells /24-well. SFO-3/EDM supplemented with high ActA (12ng/ml) was used to induce the endoderm differentiation. Different with endoderm differentiation, low ActA (3 ng/ml) and Bmp4 (7.5 ng/ml) were used to induce the differentiation towards to the mesoderm lineage. The differentiation medium was changed every day.

For the EB differentiation, two different approaches were used. First, the hanging drop culture system was used to induce the differentiation towards to three germ layers. The cells are grown in the ES medium lacking LIF with a density of 5000 cells/drop. Second, for the differentiation towards to ADE or PDE, we sorted these Foxa2<sup>+</sup> and Foxa2<sup>+</sup>T<sup>+</sup> cells at day 2, and re-cultured them aggregated as EBs in a low attachment plate (Sigma-Aldrich) for another 3 days using endoderm conditions supplemented with only ActA (20ng/ml, ADE condition) or ActA (20ng/ml) and Wnt3a (100ng/ml, PDE condition). The hepatic culture system was used according to a previously published protocol (Gouon-Evans et al., 2006). The Foxa2<sup>+</sup> and Foxa2<sup>+</sup>T<sup>+</sup> cells were sorted at day 2 and re-cultured them aggregated as EBs in the ADE condition for another two days. Then the culture medium was switched to EDM supplemented with ActA (50ng/ml), Bmp4 (50ng/ml), and bFGF (10ng/ml) for another 3 days. The EBs then were plated and cultured with the EDM medium supplemented with EGF (10 ng/ml), bFGF (10 ng/ml), HGF (20 ng/ml), TGF $\alpha$  (20 ng/ml), and VEGF (10 ng/ml) (R&D Systems) onto gelatine-coated plates for another 3 days.

##### **4.2.2.6 HEK293T cells culture**

HEK293T cells were cultured with feeder cell medium. After 2 days culture, HEK293T cells were split into 24-well plate with the density of  $8 \times 10^4$  cells /24-well for transfection.

### 4.2.3 Embryology

#### 4.2.3.1 Generation of chimeras

Tetraploid chimeras were generated according to standard protocols (Nagy et al., 2003). Tetraploid embryos were generated by electro-fusion of two-cell embryos isolated from superovulated wild type CD1 females. Embryos were incubated overnight in KSOM (Specialty Media) in a 37°C, 5% CO<sub>2</sub> incubator. The following day, the zona pellucida was removed using Tyrode's Solution (Sigma). Two embryos and an approximately 6-8 cell ESC clump were placed into small wells and allowed to aggregate overnight in KSOM in a 37°C, 5%CO<sub>2</sub> incubator. The following day, aggregated embryos were transferred to the oviduct or uterus of a pseudo-pregnant CD1 female.

#### 4.2.3.2 Isolation of embryos

Dissections of embryos and organs were performed according to standard protocols (Nagy et al., 2003). The stages of mouse embryos were determined according to previously published literature (Downs and Davies, 1993).

#### 4.2.3.3 Whole-mount *ISH*

The ISH was used to detect mRNA and miRNA on whole mount mouse embryos using radiolabeled probes or DIG-labeled probes.

#### Preparation of mouse embryos

Embryos were dissected in fresh PBS, and fixed in 4% PFA for 1 hr at 4 °C. Embryos were washed with PBST, and dehydrated in 25% Methanol/PBST, 50% Methanol/PBST, 75% Methanol/PBST, and 100% Methanol, each for 5 min. Embryos were kept in 100% Methanol at -20°C for long storage.

#### Day 1

Embryos were rehydrated in a decreasing methanol/PBST gradient with 75%, 50%, and 25% Methanol in PBST each for 5 min, and washed with PBST for 5 min. Embryos were bleached in 3% H<sub>2</sub>O<sub>2</sub> in PBST at RT in the dark for 20 min, and followed by 3x washing in PBST for 5 min. The embryos were post-fixed with 4%

#### 4. Materials and methods

---

PFA/0.2% glutaraldehyde for 20 min followed by 3x washing for 5 min in PBST. The embryos were incubated in prehybridization buffer for 2 hr at 70°C. Afterwards, the prehybridization buffer was changed and the labeled RNA probes are applied in hybridization buffer and incubated at 70°C overnight with loose cap.

##### Day 2

The probe was carefully removed and stored at -20°C. The embryos were quickly washed with prehybridization buffer, followed by 3 x 30 min incubation with preheated (70°C) solution I at 70°C. In parallel, the MAB solution was prepared. After washing, the embryos were washed 3 x 5 min with TNT at RT, and treated with RNase (0.1 mg/ml) in TNT at 37°C for 1 hr to digest the remaining RNA probes. Afterwards, the embryos were washed with TNT/ solution II (1:1) for 5 min followed by 3 x 30 min in solution II, and afterwards 3 x 5 min with MAB at RT. Blocking was performed in MAB block solution containing 10 % sheep serum for 2-3 hr at RT. After removing the block solution, the embryos were incubated in 1-2 ml antibody solution at 4°C overnight.

##### Day 3

The embryos were washed 3 x 10 min with MAB, and then washed hourly at RT while shaking and incubated in MAB at 4°C overnight.

##### Day 4

The embryos were washed 3 x 10 min in NTMT at RT and transferred to BM-purple in darkness at RT. The staining reaction was stopped by 3 x PBST washing and fixed in 4% PFA. Embryos were documented using a Zeiss Stereo Lumar V12 microscope.

#### **4.2.3.4 miRNAISH**

miR-335ISH was performed according to EXIQON protocol for whole mount *in situ hybridisation* with LNA probes. All embryos were stained using BM Purple (Roche) according to the manufacturer's instruction. The miR-1 probe was used as a positive control and scrambled miR probe was used as a negative control. Embryos were photographed using a Zeiss Stereo Lumar V12 microscope.



## 4.2.4 Data processing and analysis

### 4.2.4.1 mRNA profiling analysis

Total RNA (150 ng) was amplified using the Ambion WT Expression Kit and the WT Terminal Labeling Kit (Affymetrix). Amplified cDNA was hybridized on Affymetrix Mouse Gene ST 1.0 arrays containing about 28,000 probe sets. Staining (Fluidics script FS450\_0007) and scanning was done according to the Affymetrix expression protocol.

Expression console (v.1.2, Affymetrix) was used for quality control and to obtain annotated normalized RMA gene-level data (standard settings including median polish and sketch-quantile normalisation). Statistical analyses were performed by utilizing the statistical programming environment R (The R Core Team and Team, 2013) implemented in CARMAweb (Rainer et al., 2006). Genewise testing for differential expression was done employing the limma t-test and Benjamini-Hochberg multiple testing correction (FDR < 10%). Heatmaps were generated with CARMAweb and cluster dendrograms with R scripts (hclust, agnes, diana).

The multiclass value is the score used by the program (MeV) to select for significant values (cut off was >4.46 for significance). The results of this analysis were 371 probe sets and the cluster pictures represent genes from this set.

### 4.2.4.2 miRNA profiling analysis

Total RNA was extracted from differentiated endoderm cells at 5 different time points (0h, 12h, 24h, 48, 72h) using miRNeasy Mini Kit (Qiagen, Hiden, Germany)

The miR profiling was performed by EXIQON. The samples were labeled using the miRCURY LNA™ microRNA Hi-Power Labeling Kit, Hy3™/Hy5™ and hybridized on the miRCURY LNA™ microRNA Array (6th gen - hsa, mmu & rno). ANOVA analysis was used for the comparison of miR profile data.

### 4.2.4.3 FACS sorting and analysis

Cells were trypsinized with 0.05 % trypsin-EDTA or collagenase as described above, and re-suspended in FACS buffer and filtered to single cells. Then these differentiated cells were incubated with CD24 (Biolegend) in FACS buffer for 30 min

on ice. Cells were then centrifuged and washed with FACS buffer. Next, these cells were analyzed and sorted using a Becton Dickinson FACS Aria IIIU equipped with 405 nm, 488 nm, 561 nm, and 633 nm lasers.

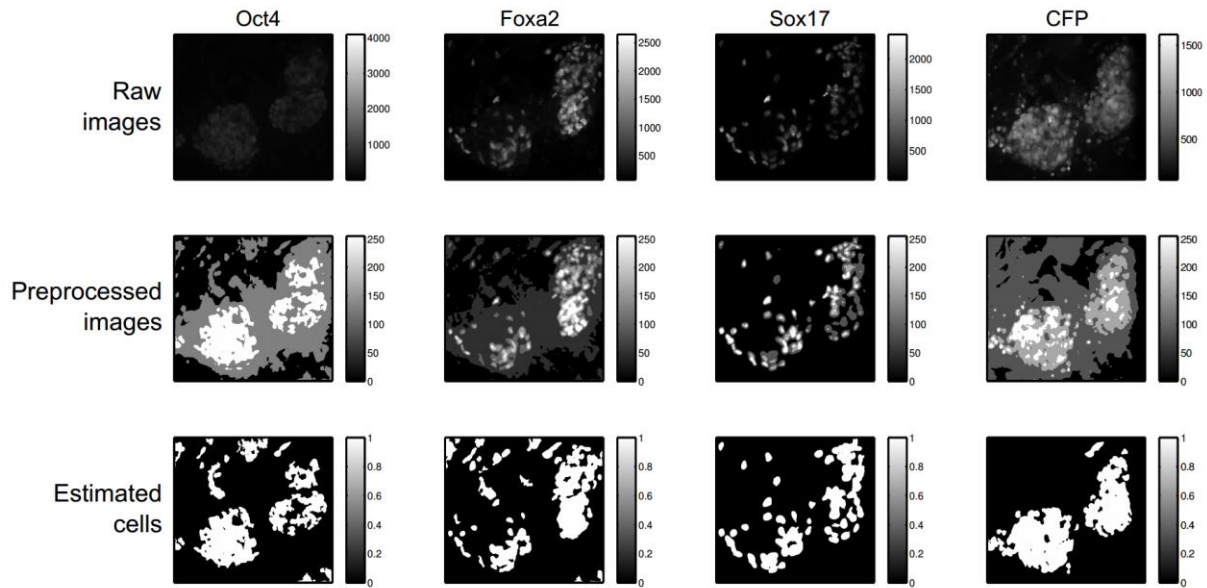
The data was analyzed either by FACS software or using FlowJo. Day 0 mESCs were used as a pluripotent stage and negative control for differentiation. The Isotype control (Biolegend) was used to set gates for positive and negative populations of CD24.

##### **4.2.4.4 Single cell tracking**

Time-lapse imaging was carried out on a Leica DMI 6000 confocal microscope equipped with an incubation system and images were made every 10 mins. Image analysis and cell tracking were performed by Imaris with a manual tracking function (Bitplane). The surpass spots function of Imaris was used to identify single cell. Cell diameter used in our tracking system was 9  $\mu\text{m}$ . For the quantification of the GFP and tagRFP signals, we used intensity mean value. The background/noise in undifferentiated ES cells was used as a negative control for GFP and tag-RFP signal.

##### **4.2.4.5 IF Image Analysis**

IF images were taken from in vitro differentiated ESCs under endoderm conditions after 0, 48 and 96 h. Images were preprocessed to improve cell identification. First, raw images were filtered using a two dimensional digital FIR filter averaging in a 10x10 window. Image contrast was increased by adjusting pixel values such that 1 % of the data is saturated at low and high intensities (Figure 4.2). Cells and boundaries were identified using image segmentation based on the maximally stable extremal regions algorithm (Matas et al., 2004). For each identified cell we calculated mean intensity by averaging over the raw image IF intensities, since adjusted intensities were not comparable between distinct images. Significance was estimated using a two-sample Kolmogorov-Smirnov test.



**Figure 4.2 Preprocessing of IHC images**

Zeiss image files were loaded into MATLAB using Bio-Formats (Linkert et al., 2010). First row: Cell culture raw intensity images for all four measured channels (Oct4, Foxa2, Sox17, CFP).

Second row: Images after preprocessing. Images were filtered using the averaging MATLAB 2D digital filter function “fspecial”. Images were normalized using the “imadjust” function such that 1% of the data is saturated at low and high intensities.

Third row: Cells were estimated using an implementation of the MSER algorithm (Buggenthin et al., 2013). Fluorescence intensities for each estimated cell (seed) were calculated by averaging all pixel intensities. Raw image (four channels), preprocessed images and estimated cells are shown.

#### 4.2.4.5 Modeling the gradient shaping effect of miR-335

This dynamic mathematical model was established by Dr. Dominik Lutter based on ordinary differential equations. The core model is derived from the miR-mRNA binding models proposed by Levine (Levine et al., 2007) and Mukherji (Mukherji et al., 2011). In contrast to these models we here allow for miR turnover to be of similar scale as gene expression. The model describes the dynamics of the free target mRNA (*foxm*) the formation and dynamics of the complex (*Cmix*) formed by the target mRNA and the miR and the free miR itself. To differentiate between mesodermal and endodermal differentiation we generated two slightly different

models. Both models were identical except for miR transcription: for the mesodermal model RNA molecules are transcribed from DNA with constant rates and functions assuming a linear expression for *Foxa2* and the miR-335 as deduced from the data. For endodermal differentiation we also assume linear mRNA expression but a temporal expression for the miR. Thus miR expression is modeled using a bell-shaped (Gaussian) expression function with an estimated maximum at time = 0. The model consists of four ordinary differential equations:

$$(1) \frac{d(foxm)}{dt} = kf - \gamma f * foxm - kon * foxm * miR + koff * complex$$

$$(2) \frac{d(Cmlx)}{dt} = kon * foxm * miR - koff * Cmlx - \gamma c * Cmlx$$

$$(3) \frac{d(miR)}{dt} = km * e^{-\frac{(\beta-t)^2}{2\sigma^2}} + \gamma c * \frac{complex}{rm+1} - kon * foxm * miR + koff * complex$$

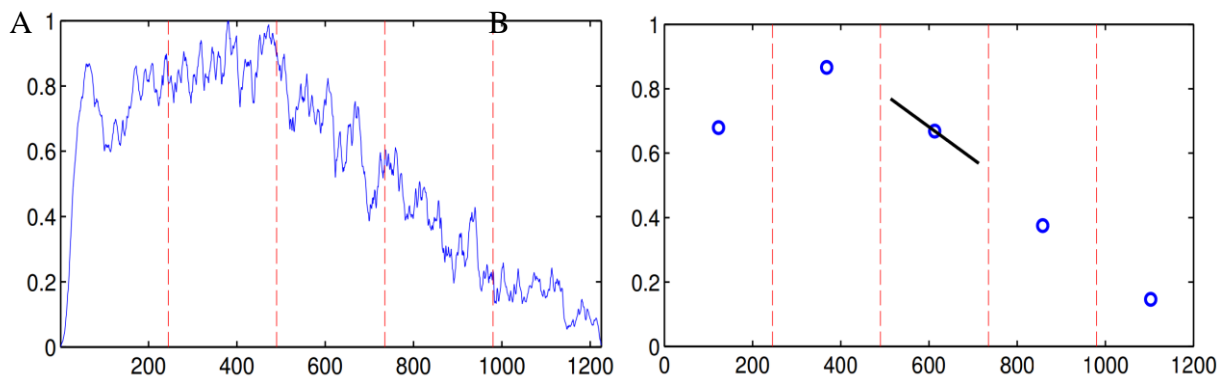
$$(3.1) \frac{d(miR)}{dt} = km + \gamma c * \frac{complex}{rm+1} - kon * foxm * miR + koff * complex$$

$$(4) \frac{d(foxP)}{dt} = kl * foxm - \gamma P * foxP$$

Further, the model assumes that only free mRNA can be translated into protein (*foxP*). Furthermore, we assume independent degradation rates for the free mRNA (*gf*) and the protein (*gC*) and miR-mediated degradation of the complex leads to a partial degradation of the miR. The system of ODEs was solved numerically by applying the CVODES solver (Hindmarsh et al., 2005) compiled as C-executable for MATLAB together with the ODE equations. Since biochemical rates are not available, we performed parameter estimation by maximum likelihood estimation for three observables (mRNA and miR concentrations measured by qPCR and protein concentrations measured by fluorescence imaging) and a total of 193 data points (data not shown). The profile likelihood was calculated applying the MATLAB implementation of the trust-region method (*lsqnonlin*). We used a log-normal error model to calculate the likelihood, including an individual noise parameter for each experimental measurement (FACS/PCR, Image Analysis).

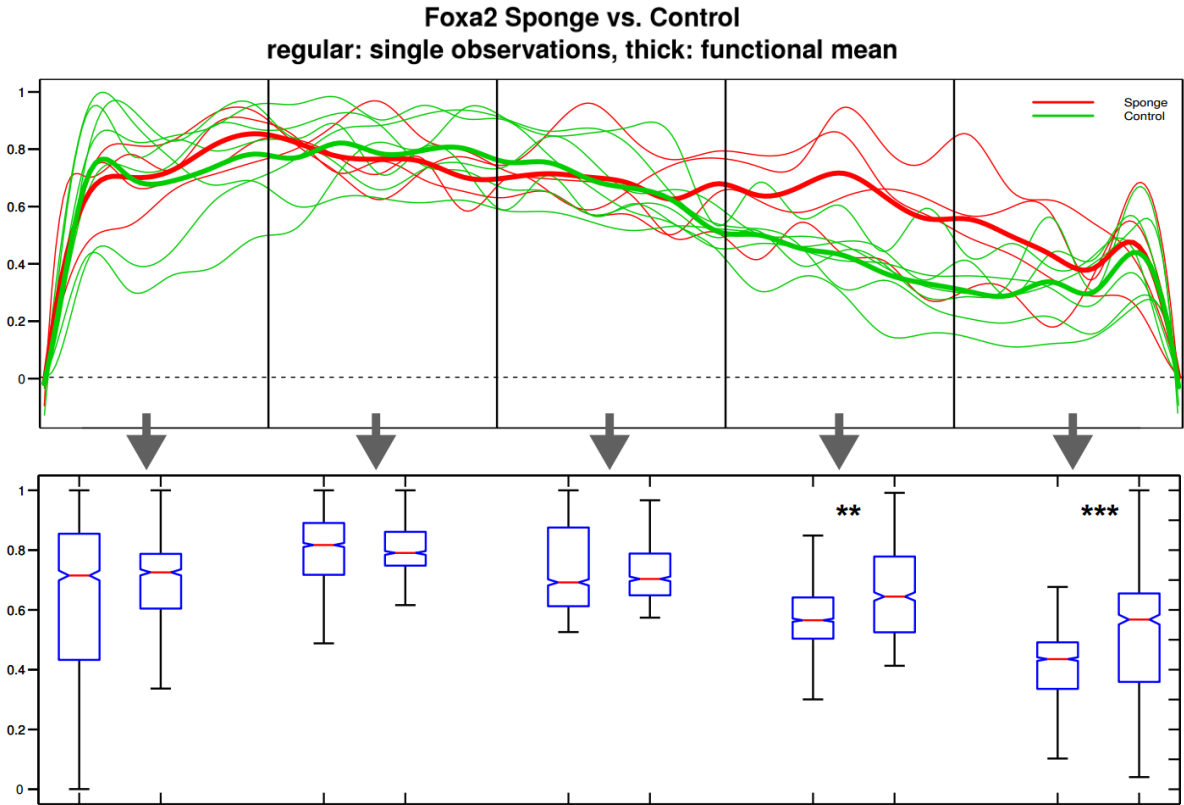
#### 4.2.4.6 Foxa2 and Sox17 gradient estimation in tetraploid embryos

TF gradients were estimated from whole-mount immunofluorescent stainings of gastrulating embryos at E7.5. For each image we manually selected a polygon mask covering the entire embryo (Figure 21A and B, lower row). Images were normalized by subtracting the overall mean and dividing each pixel by the median intensity. To estimate the gradient for each embryo we calculated the mean IF intensity along the A-P axis (right to left image axis). Gradients were normalized to the interval [0, 1] and binned into five bins along to the A-P axis. For each bin, mean intensity was calculated. We then calculated the gradient for the middle bin referring to the gradient at the middle of the A-P axis (Figure 4.3). Significant differences between the gradients were estimated using ANOVA on binned intensities (Figure 4.4).



**Figure 4.3 Foxa2 and Sox17 gradients estimation**

To estimate the protein gradient for Foxa2 and Sox17, normalized fluorescence intensities along the posterior anterior axis were used (A). Intensity levels were calculated as the mean intensity along the x-axis. Intensities were then binned and the gradient calculated at the middle bin was used to estimate miRNA mediated effect on the gradient (B).



**Figure 4.4 Gradient analysis**

To estimate differences in gradients, the sponge and control embryo images were compared. All estimated gradients were binned in 5 equidistant A-P sections. These binned intensity distributions were then pairwise compared using one-way ANOVA. Significant differences were found between sponge and control embryos in the two posterior bin

---

## 5 References

- Acloque, H., Adams, M.S., Fishwick, K., Bronner-fraser, M., and Nieto, M.A. (2009). Epithelial-mesenchymal transitions : the importance of changing cell state in development and disease. *J Clin Invest.* 119.
- Agarwal, S., Holton, K.L., and Lanza, R. (2008). Efficient differentiation of functional hepatocytes from human embryonic stem cells. *Stem Cells* 26, 1117–1127.
- Alexander, J., and Stainier, D.Y. (1999). A molecular pathway leading to endoderm formation in zebrafish. *Curr. Biol.* 9, 1147–1157.
- Ambros, V. (2000). Control of developmental timing in *Caenorhabditis elegans*. *Curr. Opin. Genet. Dev.* 10, 428–433.
- Ambros, V. (2004). The functions of animal microRNAs. *Nature* 431, 350–355.
- Ambros, V. (2011). MicroRNAs and developmental timing. *Curr. Opin. Genet. Dev.* 21, 511–517.
- Ang, S., and Rossant, J. (1994). H / W-3P Is Essential for Node and Notochord in Mouse Development. 76, 561–574.
- Ang, S.L., and Rossant, J. (1994). HNF-3 beta is essential for node and notochord formation in mouse development. *Cell* 78, 561–574.
- Ang, S.L., Wierda, a, Wong, D., Stevens, K. a, Cascio, S., Rossant, J., and Zaret, K.S. (1993). The formation and maintenance of the definitive endoderm lineage in the mouse: involvement of HNF3/forkhead proteins. *Development* 119, 1301–1315.
- Arkell, R.M., Fossat, N., and Tam, P.P. (2013). Wnt signalling in mouse gastrulation and anterior development: new players in the pathway and signal output. *Curr. Opin. Genet. Dev.* 1–7.
- Arnold, S.J., and Robertson, E.J. (2009). Making a commitment: cell lineage allocation and axis patterning in the early mouse embryo. *Nat. Rev. Mol. Cell Biol.* 10, 91–103.
- Avilion, A. a, Nicolis, S.K., Pevny, L.H., Perez, L., Vivian, N., and Lovell-Badge, R. (2003). Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev.* 17, 126–140.
- Bachiller, D., Klingensmith, J., Kemp, C., Belo, J.A., Anderson, R.M., May, S.R., McMahon, J.A., McMahon, A.P., Harland, R.M., Rossant, J., et al. (2000). The organizer factors Chordin and Noggin are required for mouse forebrain development. *Nature* 403, 658-661.
- Barrallo-Gimeno, A., and Nieto, M.A. (2005). The Snail genes as inducers of cell movement and survival: implications in development and cancer. *Development* 132, 3151–3161.
- Bartel, D.P., Lee, R., and Feinbaum, R. (2004). MicroRNAs: Genomics , Biogenesis , Mechanism , and Function. *Cell* 116, 281–297.
- Beddington, R.S., and Robertson, E.J. (1999). Axis development and early asymmetry in mammals. *Cell* 96, 195–209.

## 5. Reference

---

Belo, J. a, Bachiller, D., Agius, E., Kemp, C., Borges, a C., Marques, S., Piccolo, S., and De Robertis, E.M. (2000). Cerberus-like is a secreted BMP and nodal antagonist not essential for mouse development. *Genesis* 26, 265–270.

Ben-Haim, N., Lu, C., Guzman-Ayala, M., Pescatore, L., Mesnard, D., Bischofberger, M., Naef, F., Robertson, E.J., and Constam, D.B. (2006). The nodal precursor acting via activin receptors induces mesoderm by maintaining a source of its convertases and BMP4. *Dev. Cell* 11, 313–323.

Beppu, H., Kawabata, M., Hamamoto, T., Chytil, a, Minowa, O., Noda, T., and Miyazono, K. (2000). BMP type II receptor is required for gastrulation and early development of mouse embryos. *Dev. Biol.* 221, 249–258.

Berezikov, E., Cuppen, E., and Plasterk, R.H. a (2006). Approaches to microRNA discovery. *Nat. Genet.* 38 *Suppl*, S2–S7.

Ten Berge, D., Kurek, D., Blauwkamp, T., Koole, W., Maas, A., Eroglu, E., Siu, R.K., and Nusse, R. (2011). Embryonic stem cells require Wnt proteins to prevent differentiation to epiblast stem cells. *Nat. Cell Biol.* 13, 1070–1075.

Bernemann, C., Greber, B., Ko, K., Sternecker, J., Han, D.W., Araúzo-Bravo, M.J., and Schöler, H.R. (2011). Distinct developmental ground states of epiblast stem cell lines determine different pluripotency features. *Stem Cells* 29, 1496–1503.

Bernstein, E., Kim, S.Y., Carmell, M. a, Murchison, E.P., Alcorn, H., Li, M.Z., Mills, A. a, Elledge, S.J., Anderson, K. V, and Hannon, G.J. (2003). Dicer is essential for mouse development. *Nat. Genet.* 35, 215–217.

Biechele, S., Cox, B.J., and Rossant, J. (2011). Porcupine homolog is required for canonical Wnt signaling and gastrulation in mouse embryos. *Dev. Biol.* 355, 275–285.

Bolce, M., Hemmati-Brivanlou, and Harland, R. (1993) XFKH2, a *Xenopus* HNF-3a homologue, exhibits both Activin-inducible and autonomous phases of expression in early embryos. *Dev Biol.* 160, 413-423.

Borowiak, M., Maehr, R., Chen, S., Chen, A.E., Tang, W., Fox, J.L., Schreiber, S.L., and Melton, D. a (2009). Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. *Cell Stem Cell* 4, 348–358.

Brennan, J., Lu, C.C., Norris, D.P., Rodriguez, T. a, Beddington, R.S., and Robertson, E.J. (2001). Nodal signalling in the epiblast patterns the early mouse embryo. *Nature* 411, 965–969.

Brons, I.G.M., Smithers, L.E., Trotter, M.W.B., Rugg-Gunn, P., Sun, B., Chuva de Sousa Lopes, S.M., Howlett, S.K., Clarkson, A., Ahrlund-Richter, L., Pedersen, R. a, et al. (2007). Derivation of pluripotent epiblast stem cells from mammalian embryos. *Nature* 448, 191–195.

Bruin, J.E., Rezania, A., Xu, J., Narayan, K., Fox, J.K., O’Neil, J.J., and Kieffer, T.J. (2013). Maturation and function of human embryonic stem cell-derived pancreatic progenitors in macroencapsulation devices following transplant into mice. *Diabetologia* 56, 1987–1998.

Buckingham, M.E., and Meilhac, S.M. (2011). Tracing cells for tracking cell lineage and clonal behavior. *Dev. Cell* 21, 394–409.



- Buggenthin, F., Marr, C., Schwarzfischer, M., Hoppe, P.S., Hilsenbeck, O., Schroeder, T., and Theis, F.J. (2013). An automatic method for robust and fast cell detection in bright field images from high-throughput microscopy. *BMC Bioinformatics* 14, 297.
- Burtscher, I., and Lickert, H. (2009). *Foxa2* regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. *Development* 136, 1029–1038.
- Burtscher, I., Barkey, W., Schwarzfischer, M., Theis, F.J., and Lickert, H. (2012). The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development. *Genesis* 50, 496–505.
- Burtscher, I., Barkey, W., and Lickert, H. (2013). The *Foxa2*-Venus fusion reporter mouse line allows live-cell analysis of endoderm-derived organ formation. *Genesis* 9, 1–9.
- Cano, a, Pérez-Moreno, M. a, Rodrigo, I., Locascio, a, Blanco, M.J., del Barrio, M.G., Portillo, F., and Nieto, M. a (2000). The transcription factor snail controls epithelial-mesenchymal transitions by repressing E-cadherin expression. *Nat. Cell Biol.* 2, 76–83.
- Carlsson, P., and Mahlapuu, M. (2002). Forkhead Transcription Factors: Key Players in Development and Metabolism. *Dev. Biol.* 250, 1–23.
- Carver, E.A., Jiang, R., Lan, Y., Oram, K.F., Gridley, T., and Lan, Y.U. (2001). The Mouse Snail Gene Encodes a Key Regulator of the Epithelial-Mesenchymal Transition. *Mol Cell Biol* 21, 8184-8188.
- Cell, S., Sciences, M., Group, P.D., and Heights, Y. (2008). MicroRNA-134 modulates the differentiation of mouse embryonic stem cells, where it causes post-transcriptional attenuation of Nanog and LRH1 Stem Cells 26, 17–29.
- Chambers, I., and Smith, A. (2004). Self-renewal of teratocarcinoma and embryonic stem cells. *Oncogene* 23, 7150–7160.
- Chen, J., Han, Q., and Pei, D. (2012). Perspective EMT and MET as paradigms for cell fate switching. *J Mol Cell Biol* 4, 66–69.
- Cheng, X., Ying, L., Lu, L., Galvão, A.M., Mills, J. a, Lin, H.C., Kotton, D.N., Shen, S.S., Nostro, M.C., Choi, J.K., et al. (2012). Self-renewing endodermal progenitor lines generated from human pluripotent stem cells. *Cell Stem Cell* 10, 371–384.
- Chesley, P. (1935). Development of the short-tailed mutant in the house mouse. *J. Exp. Zool.* 70, 429–459.
- Cho, M.S., Lee, Y.-E., Kim, J.Y., Chung, S., Cho, Y.H., Kim, D.-S., Kang, S.-M., Lee, H., Kim, M.-H., Kim, J.-H., et al. (2008). Highly efficient and large-scale generation of functional dopamine neurons from human embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 105, 3392–3397.
- Choi, W.-Y., Giraldez, A.J., and Schier, A.F. (2007). Target protectors reveal dampening and balancing of Nodal agonist and antagonist by miR-430. *Science* 318, 271–274.
- Chung, K.-H., Hart, C.C., Al-Bassam, S., Avery, A., Taylor, J., Patel, P.D., Vojtek, A.B., and Turner, D.L. (2006). Polycistronic RNA polymerase II expression vectors for RNA interference based on BIC/miR-155. *Nucleic Acids Res.* 34, e53.

## 5. Reference

---

- Cirillo, L.A., Lin, F.R., Cuesta, I., Friedman, D., Jarnik, M., and Zaret, K.S. (2002). Opening of compacted chromatin by early developmental transcription factors HNF3 (FoxA) and GATA-4. *Mol. Cell* 9, 279–289.
- Ciruna, B., and Rossant, J. (2001). FGF signaling regulates Mesoderm Cell Fate Specification and Morphogenetic Movement at the Primitive Streak. *Dev Cell* 1, 37–49.
- Clements, D., and Woodland, H.R. (2000). Changes in embryonic cell fate produced by expression of an endodermal transcription factor, Xsox17. *Mech. Dev.* 99, 65–70.
- Clements, D., Taylor, H.C., Herrmann, B.G., and Stott, D. (1996). Distinct regulatory control of the Brachyury gene in axial and non-axial mesoderm suggests separation of mesoderm lineages early in mouse gastrulation. *Mech. Dev.* 56, 139–149.
- Collignon, J., Varlet, I., and Robertson, E.J. (1996). Relationship between asymmetric nodal expression and the direction of embryonic turning. *Nature* 381, 155–158.
- Conlon, F.L., Lyons, K.M., Takaesu, N., Barth, K.S., Kispert, a, Herrmann, B., and Robertson, E.J. (1994). A primary requirement for nodal in the formation and maintenance of the primitive streak in the mouse. *Development* 120, 1919–1928.
- D'Amour, K. a, Agulnick, A.D., Eliazar, S., Kelly, O.G., Kroon, E., and Baetge, E.E. (2005). Efficient differentiation of human embryonic stem cells to definitive endoderm. *Nat. Biotechnol.* 23, 1534–1541.
- Deng, C.X., Wynshaw-Boris, a, Shen, M.M., Daugherty, C., Ornitz, D.M., and Leder, P. (1994). Murine FGFR-1 is required for early postimplantation growth and axial organization. *Genes Dev.* 8, 3045–3057.
- Di-Gregorio, A., Sancho, M., Stuckey, D.W., Crompton, L. a, Godwin, J., Mishina, Y., and Rodriguez, T. a (2007). BMP signalling inhibits premature neural differentiation in the mouse embryo. *Development* 134, 3359–3369.
- Dohi, O., Yasui, K., Gen, Y., Takada, H., Endo, M., Tsuji, K., Konishi, C., Yamada, N., Mitsuyoshi, H., Yagi, N., et al. (2012). Epigenetic silencing of miR-335 and its host gene MEST in hepatocellular carcinoma. *Int. J. Oncol.* 411–418.
- Downs, K.M., and Davies, T. (1993). Staging of gastrulating mouse embryos by morphological landmarks in the dissecting microscope. *Development* 118, 1255–1266.
- Du, T., and Zamore, P.D. (2005). microPrimer: the biogenesis and function of microRNA. *Development* 132, 4645–4652.
- Dufort, D., Schwartz, L., Harpal, K., and Rossant, J. (1998). The transcription factor HNF3beta is required in visceral endoderm for normal primitive streak morphogenesis. *Development* 125, 3015–3025.
- Dunn, N.R., Vincent, S.D., Oxburgh, L., Robertson, E.J., and Bikoff, E.K. (2004). Combinatorial activities of Smad2 and Smad3 regulate mesoderm formation and patterning in the mouse embryo. *Development* 131, 1717–1728.
- Echelard, Y., Epstein, D.J., St-Jacques, B., Shen, L., Mohler, J., McMahon, J. a., and McMahon, A.P. (1993). Sonic hedgehog, a member of a family of putative signaling molecules, is implicated in the regulation of CNS polarity. *Cell* 75, 1417–1430.

- Eckert, J.J., and Fleming, T.P. (2008). Tight junction biogenesis during early development. *Biochim. Biophys. Acta* 1778, 717–728.
- Engert, S., Liao, W.P., Burtscher, I., and Lickert, H. (2009). Sox17-2A-iCre: a knock-in mouse line expressing Cre recombinase in endoderm and vascular endothelial cells. *Genesis* 47, 603–610.
- Engert, S., Burtscher, I., Liao, W.P., Dulev, S., Schotta, G., and Lickert, H. (2013). Wnt/ $\beta$ -catenin signalling regulates Sox17 expression and is essential for organizer and endoderm formation in the mouse. *Development* 140, 3128–3138.
- Evans, M.J., and Kaufman, M.H. (1981). Establishment in culture of pluripotential cells from mouse embryos. *Nature* 292, 154–156.
- Fagoonee, S., Hobbs, R.M., De Chiara, L., Cantarella, D., Piro, R.M., Tolosano, E., Medico, E., Provero, P., Pandolfi, P.P., Silengo, L., et al. (2010). Generation of functional hepatocytes from mouse germ line cell-derived pluripotent stem cells in vitro. *Stem Cells Dev.* 19, 1183–1194.
- Fehling, H.J. (2003). Tracking mesoderm induction and its specification to the hemangioblast during embryonic stem cell differentiation. *Development* 130, 4217–4227.
- Feng, X.-H., and Derynck, R. (2005). Specificity and versatility in tgf-beta signaling through Smads. *Annu. Rev. Cell Dev. Biol.* 21, 659–693.
- Friedman, J.R., and Kaestner, K.H. (2006). The Foxa family of transcription factors in development and metabolism. *Cell. Mol. Life Sci.* 63, 2317–2328.
- Fu, J., Jiang, M., Mirando, A.J., Yu, H.-M.I., and Hsu, W. (2009). Reciprocal regulation of Wnt and Gpr177/mouse Wntless is required for embryonic axis formation. *Proc. Natl. Acad. Sci. U. S. A.* 106, 18598–18603.
- Fu, S., Fei, Q., Jiang, H., Chuai, S., Shi, S., Xiong, W., Jiang, L., Lu, C., Atadja, P., Li, E., et al. (2011). Involvement of histone acetylation of Sox17 and Foxa2 promoters during mouse definitive endoderm differentiation revealed by microRNA profiling. *PLoS One* 6, e27965.
- Gadue, P., Huber, T.L., Nostro, M.C., Kattman, S., and Keller, G.M. (2005). Germ layer induction from embryonic stem cells. *Exp. Hematol.* 33, 955–964.
- Gadue, P., Huber, T.L., Paddison, P.J., and Keller, G.M. (2006). Wnt and TGF-beta signaling are required for the induction of an in vitro model of primitive streak formation using embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* 103, 16806–16811.
- Gao, L., Yang, Y., Xu, H., Liu, R., Li, D., Hong, H., Qin, M., and Wang, Y. (2014). miR-335 functions as a tumor suppressor in pancreatic cancer by targeting OCT4. *Tumour Biol.* 35, 8309–8318.
- Gao, N., White, P., and Kaestner, K.H. (2009). Establishment of intestinal identity and epithelial-mesenchymal signaling by Cdx2. *Dev. Cell* 16, 588–599.
- Gegg, M., Böttcher, A., Burtscher, I., Hasenoeder, S., Van Campenhout, C., Aichler, M., Walch, A., Grant, S.G.N., and Lickert, H. (2014). Flattop regulates basal body docking and positioning in mono- and multiciliated cells. *Elife* 3, 1–24.

Gilbert, S.F. *Developmental Biology* 9<sup>th</sup>

Di Giorgio, F.P., Boulting, G.L., Bobrowicz, S., and Eggan, K.C. (2008). Human embryonic stem cell-derived motor neurons are sensitive to the toxic effect of glial cells carrying an ALS-causing mutation. *Cell Stem Cell* 3, 637–648.

Glinka, a, Wu, W., Delius, H., Monaghan, a P., Blumenstock, C., and Niehrs, C. (1998). Dickkopf-1 is a member of a new family of secreted proteins and functions in head induction. *Nature* 391, 357–362.

Gluecksohn-schoenheimer (1938). The Development of Two Tailless Mutants in the House Mouse. *Genetics* 23, 573-84.

Gouon-Evans, V., Boussemart, L., Gadue, P., Nierhoff, D., Koehler, C.I., Kubo, A., Shafritz, D. a, and Keller, G. (2006). BMP-4 is required for hepatic specification of mouse embryonic stem cell-derived definitive endoderm. *Nat. Biotechnol.* 24, 1402–1411.

Graham, et al., 1977, and Iaa, N.W. Characteristics of a Human Cell Line Transformed by D N A from Human Adenovirus Type 5. *J Gen Virol* 36, 59-74.

Grau, Y., Carteret, C., and Simpson, P. (1984). Mutations and chromosomal rearrangements affecting the expression of snail, a gene involved in embryonic patterning in *Drosophila melanogaster*. *Genetics* 2, 347–360.

GRUNEBERG, H. (1958). Genetical Studies on the Skeleton of the Mouse. *J Embryol Exp Morphol* 6, 124–148.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010a). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835–840.

Guo, H., Ingolia, N.T., Weissman, J.S., and Bartel, D.P. (2010b). Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* 466, 835–840.

Haegel, H., Larue, L., Ohsugi, M., Fedorov, L., Herrenknecht, K., and Kemler, R. (1995). Lack of beta-catenin affects mouse development at gastrulation. *Development* 121, 3529–3537.

Hagos, E.G., and Dougan, S.T. (2007). Time-dependent patterning of the mesoderm and endoderm by Nodal signals in zebrafish. *BMC Dev. Biol.* 7, 22.

Halpern, M.E., Ho, R.K., Walker, C., and Kimmel, C.B. (1993). Induction of muscle pioneers and floor plate is distinguished by the zebrafish no tail mutation. *Cell* 75, 99–111.

Hannenhalli, S., and Kaestner, K.H. (2009). The evolution of Fox genes and their role in development and disease. *Nat. Rev. Genet.* 10, 233–240.

Hansson, M., Olesen, D.R., Peterslund, J.M., Engberg, N., Winzi, M., Klein, T., Maddox-hyttel, P., and Serup, P. (2009). A late requirement for Wnt and FGF signaling during activin-induced formation of foregut endoderm from mouse embryonic stem cells. *Dev Biol.* 330, 286–304.

Haque, M.A., Nagaoka, M., Hexig, B., and Akaike, T. (2010). Artificial extracellular matrix for embryonic stem cell cultures: a new frontier of nanobiomaterials. *Sci. Technol. Adv. Mater.* 11, 014106.

- Hay, D.C., Fletcher, J., Payne, C., Terrace, J.D., Gallagher, R.C.J., Snoeys, J., Black, J.R., Wojtacha, D., Samuel, K., Hannoun, Z., et al. (2008). Highly efficient differentiation of hESCs to functional hepatic endoderm requires ActivinA and Wnt3a signaling. *Proc. Natl. Acad. Sci. U. S. A.* *105*, 12301–12306.
- Hébert, J.M., Boyle, M., and Martin, G.R. (1991). mRNA localization studies suggest that murine FGF-5 plays a role in gastrulation. *Development* *112*, 407–415.
- Herr, P., Hausmann, G., and Basler, K. (2012). WNT secretion and signalling in human disease. *Trends Mol. Med.* *18*, 483–493.
- Herrmann, B.G. (1991). Expression pattern of the Brachyury gene in whole-mount embryos mutant. *Development* *113*, 913–917.
- Hindmarsh, A.C., Brown, P.N., Grant, K.E., Lee, S.L., Serban, R., Shumaker, D.A.N.E., and Woodward, C.S. (2005). SUNDIALS: Suite of Nonlinear and Differential / Algebraic Equation Solvers. *31*, 363–396.
- Horn, S., Kobberup, S., Jørgensen, M.C., Kalisz, M., Klein, T., Kageyama, R., Gegg, M., Lickert, H., Lindner, J., Magnuson, M. a, et al. (2012). Mind bomb 1 is required for pancreatic  $\beta$ -cell formation. *Proc. Natl. Acad. Sci. U. S. A.* *109*, 7356–7361.
- Horner, M.A., Quintin, S., Domeier, M.E., Kimble, J., Labouesse, M., and Mango, S.E. (1998). pha-4, an HNF-3 homolog, specifies pharyngeal organ identity in *Caenorhabditis elegans*. *Genes Dev.* *12*, 1947–1952.
- Houbaviy, H.B., Murray, M.F., and Sharp, P. a (2003). Embryonic stem cell-specific MicroRNAs. *Dev. Cell* *5*, 351–358.
- Huang, H.Y., Cheng, Y.Y., Liao, W.C., Tien, Y.W., Yang, C.H.J., Hsu, S.M., and Huang, P.H. (2012). SOX4 transcriptionally regulates multiple SEMA3/Plexin family members and promotes tumor growth in pancreatic cancer. *PLoS One* *7*, e48637.
- Hudson, C., Clements, D., Friday, R. V, Stott, D., and Woodland, H.R. (1997). Xsox17alpha and -beta mediate endoderm formation in *Xenopus*. *Cell* *91*, 397–405.
- Huelsken, J., Vogel, R., Brinkmann, V., Erdmann, B., Birchmeier, C., and Birchmeier, W. (2000). Requirement for beta-catenin in anterior-posterior axis formation in mice. *J. Cell Biol.* *148*, 567–578.
- Ikenouchi, J., Matsuda, M., Furuse, M., and Tsukita, S. (2003). Regulation of tight junctions during the epithelium-mesenchyme transition: direct repression of the gene expression of claudins/occludin by Snail. *J. Cell Sci.* *116*, 1959–1967.
- Inui, M., Montagner, M., and Piccolo, S. (2012). miRNAs and morphogen gradients. *Curr. Opin. Cell Biol.* *24*, 194–201.
- Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., Lovell-Badge, R.(1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryonically expressed genes. *Nature* *346*(6281), 245-50.

## 5. Reference

---

Jho, E., Zhang, T., Domon, C., Joo, C., Freund, J., and Costantini, F. (2002). Wnt/beta-catenin/Tcf signaling induces the transcription of Axin2, a negative regulator of the signaling pathway. *Mol Cell Biol.* 22, 1172-1183.

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

Jiang, W., Sui, X., Zhang, D., Liu, M., Ding, M., Shi, Y., and Deng, H. (2011). CD24: a novel surface marker for PDX1-positive pancreatic progenitors derived from human embryonic stem cells. *Stem Cells* 29, 609–617.

Johansson, B.M., and Wiles, M. V (1995). Evidence for involvement of activin A and bone morphogenetic protein 4 in mammalian mesoderm and hematopoietic development. *Mol. Cell. Biol.* 15, 141–151.

Kaestner, K.H., Knöchel, W., Martínez, D.E. (2000). Unified nomenclature for the winged helix / forkhead transcription factors. *Genes Dev.* 14, 142–146.

Kalluri, R., and Weinberg, R.A. (2009). The basics of epithelial-mesenchymal transition. *J Clin Invest.* 119, 1420-1428.

Kanai, Y., Kanai-Azuma, M., Noce, T., Saido, T.C., Shiroishi, T., Hayashi, Y., and Yazaki, K. (1996). Identification of two Sox17 messenger RNA isoforms, with and without the high mobility group box region, and their differential expression in mouse spermatogenesis. *J. Cell Biol.* 133, 667–681.

Kanai-Azuma, M., Kanai, Y., Gad, J.M., Tajima, Y., Taya, C., Kurohmaru, M., Sanai, Y., Yonekawa, H., Yazaki, K., Tam, P.P.L., et al. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* 129, 2367–2379.

Kanellopoulou, C., Muljo, S. a, Kung, A.L., Ganesan, S., Drapkin, R., Jenuwein, T., Livingston, D.M., and Rajewsky, K. (2005). Dicer-deficient mouse embryonic stem cells are defective in differentiation and centromeric silencing. *Genes Dev.* 19, 489–501.

Kavka, a I., and Green, J.B. (1997). Tales of tails: Brachyury and the T-box genes. *Biochim. Biophys. Acta* 1333, F73–F84.

Kehat, I., Kenyagin-karsenti, D., Snir, M., Segev, H., Amit, M., Gepstein, A., Livne, E., Binah, O., Itskovitz-eldor, J., and Gepstein, L. (2001). Human embryonic stem cells can differentiate into myocytes with structural and functional properties of cardiomyocytes. *J Clin Invest.* 108, 407–414.

Keller, G. (2005). Embryonic stem cell differentiation: emergence of a new era in biology and medicine. *Genes Dev.* 19, 1129–1155.

Kelly, O.G., Pinson, K.I., and Skarnes, W.C. (2004). The Wnt co-receptors Lrp5 and Lrp6 are essential for gastrulation in mice. *Development* 131, 2803–2815.

Kemler, R., Hierholzer, A., Kanzler, B., Kuppig, S., Hansen, K., Taketo, M.M., de Vries, W.N., Knowles, B.B., and Solter, D. (2004). Stabilization of beta-catenin in the mouse zygote leads to premature epithelial-mesenchymal transition in the epiblast. *Development* 131, 5817–5824.

- Kemp, C., Willems, E., Abdo, S., Lambiv, L., and Leyns, L. (2005). Expression of all Wnt genes and their secreted antagonists during mouse blastocyst and postimplantation development. *Dev. Dyn.* 233, 1064–1075.
- Kim, Y.-K., and Kim, V.N. (2007). Processing of intronic microRNAs. *EMBO J.* 26, 775–783.
- Kim, J., Krichevsky, A., Grad, Y., Hayes, G.D., Kosik, K.S., Church, G.M., and Ruvkun, G. (2003). Identification of many microRNAs that copurify with polyribosomes in mammalian neurons. *Proc Natl Acad Sci U. S. A.* 101, 360–365.
- Kimelman, D., and Griffin, K.J. (2000). Vertebrate mesendoderm induction and patterning. *Curr. Opin. Genet. Dev.* 10, 350–356.
- Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A., Nagy, A., and Tam, P.P. (1999). The orderly allocation of mesodermal cells to the extraembryonic structures and the anteroposterior axis during gastrulation of the mouse embryo. *Development* 126, 4691–4701.
- Kinder, S.J., Tsang, T.E., Wakamiya, M., Sasaki, H., Behringer, R.R., Nagy, A., and Tam, P.P.L. (2001). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* 128, 3623–3634.
- Kispert, A., and Herrmann, B.G. (1993). The Brachyury protein novel DNA binding. *EMBO J* 12, 3211–3220.
- Kitisin, K., Saha, T., Blake, T., Golestaneh, N., Deng, M., Kim, C., Tang, Y., Shetty, K., Mishra, B., and Mishra, L. (2007). Tgf-Beta signaling in development. *Sci. STKE* 2007, cm1.
- Klingensmith, J., Ang, S.L., Bachiller, D., and Rossant, J. (1999). Neural induction and patterning in the mouse in the absence of the node and its derivatives. *Dev. Biol.* 216, 535–549.
- Kloosterman, W.P., Wienholds, E., Bruijn, E. De, Kauppinen, S., and Plasterk, R.H.A. (2006). In situ detection of miRNAs in animal embryos using LNA-modified oligonucleotide probes. *Nat Methods* 3, 2005–2007.
- Kopper, O., and Benvenisty, N. (2012). Stepwise differentiation of human embryonic stem cells into early endoderm derivatives and their molecular characterization. *Stem Cell Res.* 8, 335–345.
- Kowarsch, A., Preusse, M., Marr, C., and Theis, F.J. (2011). miTALOS: Analyzing the tissue-specific regulation of signaling pathways by human and mouse microRNAs. *RNA* 17, 809–819.
- Krol, J., Loedige, I., and Filipowicz, W. (2010). The widespread regulation of microRNA biogenesis, function and decay. *Nat. Rev. Genet.* 11, 597–610.
- Kroon, E., Martinson, L. a, Kadoya, K., Bang, A.G., Kelly, O.G., Eliazer, S., Young, H., Richardson, M., Smart, N.G., Cunningham, J., et al. (2008). Pancreatic endoderm derived from human embryonic stem cells generates glucose-responsive insulin-secreting cells in vivo. *Nat. Biotechnol.* 26, 443–452.
- Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., and Keller, G. (2004a). Development of definitive endoderm from embryonic stem cells in culture. *Development* 131, 1651–1662.

- Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., and Keller, G. (2004b). Development of definitive endoderm from embryonic stem cells in culture. *Development* *131*, 1651–1662.
- Kumar, A., Lualdi, M., Lewandoski, M., and Kuehn, M.R. (2008). Broad mesodermal and endodermal deletion of Nodal at postgastrulation stages results solely in left/right axial defects. *Dev. Dyn.* *237*, 3591–3601.
- Laflamme, M. a, Chen, K.Y., Naumova, A. V, Muskheli, V., Fugate, J. a, Dupras, S.K., Reinecke, H., Xu, C., Hassanipour, M., Police, S., et al. (2007). Cardiomyocytes derived from human embryonic stem cells in pro-survival factors enhance function of infarcted rat hearts. *Nat. Biotechnol.* *25*, 1015–1024.
- Lai, E., Prezioso, V.R., Smith, E., Litvin, O., Costa, R.H., and Darnell, J.E. (1990). HNF-3A, a hepatocyte-enriched transcription factor of novel structure is regulated transcriptionally. *Genes Dev.* *4*, 1427–1436.
- Lai, E., Prezioso, V.R., Tao, W.F., Chen, W.S., and Darnell, J.E. (1991). Hepatocyte nuclear factor 3 alpha belongs to a gene family in mammals that is homologous to the Drosophila homeotic gene fork head. *Genes Dev.* *5*, 416–427.
- Lamouille, S., Subramanyam, D., Billelo, R., and Derynck, R. (2013). Regulation of epithelial-mesenchymal and mesenchymal-epithelial transitions by microRNAs. *Curr. Opin. Cell Biol.* *25*, 200–207.
- Lantz, K. a, and Kaestner, K.H. (2005). Winged-helix transcription factors and pancreatic development. *Clin. Sci. (Lond).* *108*, 195–204.
- Lawson, K. a, and Pedersen, R. a (1987). Cell fate, morphogenetic movement and population kinetics of embryonic endoderm at the time of germ layer formation in the mouse. *Development* *101*, 627–652.
- Lawson, K. a, Meneses, J.J., and Pedersen, R. a (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* *113*, 891–911.
- Lawson, K. a, Dunn, N.R., Roelen, B. a, Zeinstra, L.M., Davis, a M., Wright, C. V, Korving, J.P., and Hogan, B.L. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* *13*, 424–436.
- Lee, C.S., Friedman, J.R., Fulmer, J.T., and Kaestner, K.H. (2005). The initiation of liver development is dependent on Foxa transcription factors. *Nature* *435*, 944–947.
- Lee, R.C., Feinbaum, R.L., and Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* *75*, 843–854.
- Lee, Y., Kim, M., Han, J., Yeom, K.-H., Lee, S., Baek, S.H., and Kim, V.N. (2004). MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* *23*, 4051–4060.
- Levine, E., McHale, P., and Levine, H. (2007). Small regulatory RNAs may sharpen spatial expression patterns. *PLoS Comput. Biol.* *3*, e233.
- Lewis, S.L., and Tam, P.P.L. (2006). Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Dev. Dyn.* *235*, 2315–2329.



- Lickert, H., Kutsch, S., Kanzler, B., Tamai, Y., Taketo, M.M., and Kemler, R. (2002). Formation of multiple hearts in mice following deletion of beta-catenin in the embryonic endoderm. *Dev. Cell* 3, 171–181.
- Lim, J., and Thiery, J.P. (2012). Epithelial-mesenchymal transitions: insights from development. *Development* 139, 3471–3486.
- Lin, S.-L., Chang, D., Wu, D.-Y., and Ying, S.-Y. (2003). A novel RNA splicing-mediated gene silencing mechanism potential for genome evolution. *Biochem. Biophys. Res. Commun.* 310, 754–760.
- Lindsley, R.C., Gill, J.G., Murphy, T.L., Langer, E.M., Cai, M., Mashayekhi, M., Wang, W., Niwa, N., Nerbonne, J.M., Kyba, M., et al. (2008). *Mesp1* coordinately regulates cardiovascular fate restriction and epithelial-mesenchymal transition in differentiating ESCs. *Cell Stem Cell* 3, 55–68.
- Linkert, M., Rueden, C.T., Allan, C., Burel, J.-M., Moore, W., Patterson, A., Loranger, B., Moore, J., Neves, C., Macdonald, D., et al. (2010). Metadata matters: access to image data in the real world. *J. Cell Biol.* 189, 777–782.
- Liu, P., Wakamiya, M., Shea, M.J., Albrecht, U., Behringer, R.R., and Bradley, a (1999). Requirement for *Wnt3* in vertebrate axis formation. *Nat. Genet.* 22, 361–365.
- Liu, Y., Festing, M., Thompson, J.C., Hester, M., Rankin, S., El-Hodiri, H.M., Zorn, A.M., and Weinstein, M. (2004). *Smad2* and *Smad3* coordinately regulate craniofacial and endodermal development. *Dev. Biol.* 270, 411–426.
- Logan, C.Y., and Nusse, R. (2004). The Wnt signaling pathway in development and disease. *Annu. Rev. Cell Dev. Biol.* 20, 781–810.
- Lund, E., Güttinger, S., Calado, A., Dahlberg, J.E., and Kutay, U. (2004). Nuclear export of microRNA precursors. *Science* 303, 95–98.
- Lynch, J., Fay, J., Meehan, M., Bryan, K., Watters, K.M., Murphy, D.M., and Stallings, R.L. (2012). MiRNA-335 suppresses neuroblastoma cell invasiveness by direct targeting of multiple genes from the non-canonical TGF- $\beta$  signalling pathway. *Carcinogenesis* 33, 976–985.
- Mallanna, S.K., and Rizzino, A. (2010). Emerging roles of microRNAs in the control of embryonic stem cells and the generation of induced pluripotent stem cells. *Dev. Biol.* 344, 16–25.
- Martello, G., Zacchigna, L., Inui, M., Montagner, M., Adorno, M., Mamidi, A., Morsut, L., Soligo, S., Tran, U., Dupont, S., et al. (2007). MicroRNA control of Nodal signalling. *Nature* 449, 183–188.
- Martin, G.R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U. S. A.* 78, 7634–7638.
- Martin, B.L., and Kimelman, D. (2008). Regulation of canonical Wnt signaling by Brachyury is essential for posterior mesoderm formation. *Dev. Cell* 15, 121–133.

- Martinez, N.J., Ow, M.C., Reece-Hoyes, J.S., Barrasa, M.I., Ambros, V.R., and Walhout, A.J.M. (2008). Genome-scale spatiotemporal analysis of *Caenorhabditis elegans* microRNA promoter activity. *Genome Res.* 18, 2005–2015.
- Matas, J., Chum, O., Urban, M., and Pajdla, T. (2004). Robust wide-baseline stereo from maximally stable extremal regions. *Image Vis. Comput.* 22, 761–767.
- Mathieu, J., Griffin, K., Herbomel, P., Dickmeis, T., Strähle, U., Kimelman, D., Rosa, F.M., and Peyri eras, N. (2004). Nodal and Fgf pathways interact through a positive regulatory loop and synergize to maintain mesodermal cell populations. *Development* 131, 629–641.
- Matsui, T., Kanai-Azuma, M., Hara, K., Matoba, S., Hiramatsu, R., Kawakami, H., Kurohmaru, M., Koopman, P., and Kanai, Y. (2006). Redundant roles of Sox17 and Sox18 in postnatal angiogenesis in mice. *J. Cell Sci.* 119, 3513–3526.
- McGrath, K.E., Koniski, a D., Maltby, K.M., McGann, J.K., and Palis, J. (1999). Embryonic expression and function of the chemokine SDF-1 and its receptor, CXCR4. *Dev. Biol.* 213, 442–456.
- McKnight, K.D., Hou, J., and Hoodless, P. a (2010). Foxh1 and Foxa2 are not required for formation of the midgut and hindgut definitive endoderm. *Dev. Biol.* 337, 471–481.
- Miranda, K.C., Huynh, T., Tay, Y., Ang, Y.-S., Tam, W.-L., Thomson, A.M., Lim, B., and Rigoutsos, I. (2006). A pattern-based method for the identification of MicroRNA binding sites and their corresponding heteroduplexes. *Cell* 126, 1203–1217.
- Mirosevich, J., Gao, N., and Matusik, R.J. (2005). Expression of Foxa transcription factors in the developing and adult murine prostate. *Prostate* 62, 339–352.
- Monaghan, a P., Kaestner, K.H., Grau, E., and Sch utz, G. (1993). Postimplantation expression patterns indicate a role for the mouse forkhead/HNF-3 alpha, beta and gamma genes in determination of the definitive endoderm, chordamesoderm and neuroectoderm. *Development* 119, 567–578.
- Moon, R.T. (2005). Wnt/beta-catenin pathway. *Sci. STKE* 2005, cm1.
- Moreno-Bueno, G., Portillo, F., and Cano, a (2008). Transcriptional regulation of cell polarity in EMT and cancer. *Oncogene* 27, 6958–6969.
- Morrison, G.M., Oikonomopoulou, I., Migueles, R.P., Soneji, S., Livigni, A., Enver, T., and Brickman, J.M. (2008). Anterior definitive endoderm from ESCs reveals a role for FGF signaling. *Cell Stem Cell* 3, 402–415.
- Mourelatos, Z., Dostie, J., Paushkin, S., Sharma, A., Charroux, B., Abel, L., Rappsilber, J., Mann, M., and Dreyfuss, G. (2002). miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev.* 16, 720–728.
- Mukherji, S., Ebert, M.S., Zheng, G.X.Y., Tsang, J.S., Sharp, P. a, and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.* 43, 854–859.
- Munoz-Sanjuan, I. (2002). Gene profiling during neural induction in *Xenopus laevis*: regulation of BMP signaling by post-transcriptional mechanisms and TAB3, a novel TAK1-binding protein. *Development* 129, 5529–5540.

- 
- Murchison, E.P., Partridge, J.F., Tam, O.H., Cheloufi, S., and Hannon, G.J. (2005). Characterization of Dicer-deficient murine embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *102*, 12135–12140.
- Murchison, E.P., Stein, P., Xuan, Z., Pan, H., Zhang, M.Q., Schultz, R.M., and Hannon, G.J. (2007). Critical roles for Dicer in the female germline. *Genes Dev.* *21*, 682–693.
- Murry, C.E., and Keller, G. (2008). Differentiation of embryonic stem cells to clinically relevant populations: lessons from embryonic development. *Cell* *132*, 661–680.
- Nagy, a, Rossant, J., Nagy, R., Abramow-Newerly, W., and Roder, J.C. (1993). Derivation of completely cell culture-derived mice from early-passage embryonic stem cells. *Proc. Natl. Acad. Sci. U. S. A.* *90*, 8424–8428.
- Nakanishi, M., Kurisaki, A., Hayashi, Y., Warashina, M., Ishiura, S., Kusuda-Furue, M., and Asashima, M. (2009). Directed induction of anterior and posterior primitive streak by Wnt from embryonic stem cells cultured in a chemically defined serum-free medium. *FASEB J.* *23*, 114–122.
- Nakaya, Y., and Sheng, G. (2008). Epithelial to mesenchymal transition during gastrulation: an embryological view. *Dev. Growth Differ.* *50*, 755–766.
- Ng, E.S., Azzola, L., Sourris, K., Robb, L., Stanley, E.G., and Elefanty, A.G. (2005). The primitive streak gene *Mixl1* is required for efficient haematopoiesis and BMP4-induced ventral mesoderm patterning in differentiating ES cells. *Development* *132*, 873–884.
- Nichols, J., and Smith, A. (2015). *Pluripotency in the Embryo and in Culture*. Cold Spring Harbor Perspectives in Biology.
- Niehrs, C., and Acebron, S.P. (2010). Wnt signaling: multivesicular bodies hold GSK3 captive. *Cell* *143*, 1044–1046.
- Nieto, M.A. (2002). The snail superfamily of zinc-finger transcription factors. *Nat. Rev. Mol. Cell Biol.* *3*, 155–166.
- Nieuwkoop, P.D. (1997). Short historical survey of pattern formation in the endo-mesoderm and the neural anlage in the vertebrates: the role of vertical and planar inductive actions. *Cell. Mol. Life Sci. C.* *53*, 305–318.
- Niwa, H., Burdon, T., Chambers, I., and Smith, a. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. *Genes Dev.* *12*, 2048–2060.
- Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* *24*, 372–376.
- Ohinata, Y., Ohta, H., Shigeta, M., Yamanaka, K., Wakayama, T., and Saitou, M. (2009). A signaling principle for the specification of the germ cell lineage in mice. *Cell* *137*, 571–584.
- Orkin, S.H., Wang, J., Kim, J., Chu, J., Rao, S., Theunissen, T.W., Shen, X., and Levasseur, D.N. (2008). The transcriptional network controlling pluripotency in ES cells. *Cold Spring Harb. Symp. Quant. Biol.* *73*, 195–202.
- Osorno, R., Tsakiridis, A., Wong, F., Cambray, N., Economou, C., Wilkie, R., Blin, G., Scotting, P.J., Chambers, I., and Wilson, V. (2012). The developmental dismantling of

pluripotency is reversed by ectopic Oct4 expression. *Development* 139, 2288–2298.

Pagliuca, F.W., Millman, J.R., Gürtler, M., Segel, M., Van Dervort, A., Ryu, J.H., Peterson, Q.P., Greiner, D., and Melton, D.A. (2014). Generation of Functional Human Pancreatic  $\beta$  Cells In Vitro. *Cell* 159, 428–439.

Pan, G., and Thomson, J. a (2007). Nanog and transcriptional networks in embryonic stem cell pluripotency. *Cell Res.* 17, 42–49.

Peinado, H., Olmeda, D., and Cano, A. (2007). Snail, Zeb and bHLH factors in tumour progression: an alliance against the epithelial phenotype? *Nat. Rev. Cancer* 7, 415–428.

Perantoni, A.O., Timofeeva, O., Naillat, F., Richman, C., Pajni-Underwood, S., Wilson, C., Vainio, S., Dove, L.F., and Lewandoski, M. (2005). Inactivation of FGF8 in early mesoderm reveals an essential role in kidney development. *Development* 132, 3859–3871.

Perea-Gomez, A., Vella, F.D.J., Shawlot, W., Oulad-Abdelghani, M., Chazaud, C., Meno, C., Pfister, V., Chen, L., Robertson, E., Hamada, H., et al. (2002). Nodal antagonists in the anterior visceral endoderm prevent the formation of multiple primitive streaks. *Dev. Cell* 3, 745–756.

Pevny, L.H., and Lovell-Badge, R. (1997). Sox genes find their feet. *Curr. Opin. Genet. Dev.* 7, 338–344.

Pfister, S., Steiner, K. a, and Tam, P.P.L. (2007). Gene expression pattern and progression of embryogenesis in the immediate post-implantation period of mouse development. *Gene Expr. Patterns* 7, 558–573.

Png, K.J., Yoshida, M., Zhang, X.H.-F., Shu, W., Lee, H., Rimmer, A., Chan, T. a, Comen, E., Andrade, V.P., Kim, S.W., et al. (2011). MicroRNA-335 inhibits tumor reinitiation and is silenced through genetic and epigenetic mechanisms in human breast cancer. *Genes Dev.* 25, 226–231.

Polytarchou, C., Iliopoulos, D., and Struhl, K. (2012). An integrated transcriptional regulatory circuit that reinforces the breast cancer stem cell state. *Proc Natl Acad Sci USA* 109, 14470-14475.

Pratt, A.J., and MacRae, I.J. (2009). The RNA-induced silencing complex: a versatile gene-silencing machine. *J. Biol. Chem.* 284, 17897–17901.

Qi, X., Li, T.-G., Hao, J., Hu, J., Wang, J., Simmons, H., Miura, S., Mishina, Y., and Zhao, G.-Q. (2004). BMP4 supports self-renewal of embryonic stem cells by inhibiting mitogen-activated protein kinase pathways. *Proc. Natl. Acad. Sci. U. S. A.* 101, 6027–6032.

Qu, X.-B., Pan, J., Zhang, C., and Huang, S.-Y. (2008). Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm in vitro. *Dev. Growth Differ.* 50, 585–593.

Rainer, J., Sanchez-Cabo, F., Stocker, G., Sturn, A., and Trajanoski, Z. (2006). CARMAweb: comprehensive R- and bioconductor-based web service for microarray data analysis. *Nucleic Acids Res.* 34, W498–W503.

Rajewsky, N. (2006). microRNA target predictions in animals. *Nat. Genet.* 38 *Suppl*, S8–S13.

- 
- Raue, a, Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U., and Timmer, J. (2009). Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. *Bioinformatics* 25, 1923–1929.
- Reinhart, B.J., Slack, F.J., Basson, M., Pasquinelli, a E., Bettinger, J.C., Rougvie, a E., Horvitz, H.R., and Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403, 901–906.
- Rodaway, a, and Patient, R. (2001). Mesendoderm. an ancient germ layer? *Cell* 105, 169–172.
- Rodriguez, T. a, Srinivas, S., Clements, M.P., Smith, J.C., and Beddington, R.S.P. (2005). Induction and migration of the anterior visceral endoderm is regulated by the extra-embryonic ectoderm. *Development* 132, 2513–2520.
- Rogers, K.W., and Schier, A.F. (2011). Morphogen gradients: from generation to interpretation. *Annu. Rev. Cell Dev. Biol.* 27, 377–407.
- Ronchetti, D., Lionetti, M., Mosca, L., Agnelli, L., Andronache, A., Fabris, S., Deliliers, G.L., and Neri, A. (2008). An integrative genomic approach reveals coordinated expression of intronic miR-335, miR-342, and miR-561 with deregulated host genes in multiple myeloma. *BMC Med. Genomics* 1, 37.
- Rosa, A., Spagnoli, F.M., and Brivanlou, A.H. (2009). The miR-430/427/302 family controls mesendodermal fate specification via species-specific target selection. *Dev. Cell* 16, 517–527.
- Rossant, J., and Tam, P.P.L. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. *Development* 136, 701–713.
- Ruby, J.G., Jan, C., Player, C., Axtell, M.J., Lee, W., Nusbaum, C., Ge, H., and Bartel, D.P. (2006). Large-scale sequencing reveals 21U-RNAs and additional microRNAs and endogenous siRNAs in *C. elegans*. *Cell* 127, 1193–1207.
- Sarkar, A., and Hochedlinger, K. (2013). The sox family of transcription factors: versatile regulators of stem and progenitor cell fate. *Cell Stem Cell* 12, 15–30.
- Sasaki, H., and Hogan, B.L. (1993). Differential expression of multiple fork head related genes during gastrulation and axial pattern formation in the mouse embryo. *Development* 118, 47–59.
- Sato, N., Meijer, L., Skaltsounis, L., Greengard, P., and Brivanlou, A.H. (2004). Maintenance of pluripotency in human and mouse embryonic stem cells through activation of Wnt signaling by a pharmacological GSK-3-specific inhibitor. *Nat. Med.* 10, 55–63.
- Schier, A.F. (2003). Nodal signaling in vertebrate development. *Annu. Rev. Cell Dev. Biol.* 19, 589–621.
- Schulte-Merker, S., van Eeden, F.J., Halpern, M.E., Kimmel, C.B., and Nüsslein-Volhard, C. (1994). no tail (ntl) is the zebrafish homologue of the mouse T (Brachyury) gene. *Development* 120, 1009–1015.
- Seidensticker, M.J., and Behrens, J. (2000). Biochemical interactions in the wnt pathway. *Biochim. Biophys. Acta* 1495, 168–182.

## 5. Reference

---

- Sherwood, R.I., Maehr, R., Mazzone, E.O., and Melton, D. a (2011). Wnt signaling specifies and patterns intestinal endoderm. *Mech. Dev.* 128, 387–400.
- Shiba, Y., Fernandes, S., Zhu, W.-Z., Filice, D., Muskheli, V., Kim, J., Palpant, N.J., Gantz, J., Moyes, K.W., Reinecke, H., et al. (2012). Human ES-cell-derived cardiomyocytes electrically couple and suppress arrhythmias in injured hearts. *Nature* 489, 322–325.
- Shirai, T., Miyagi, S., Horiuchi, D., Okuda-Katayanagi, T., Nishimoto, M., Muramatsu, M., Sakamoto, Y., Nagata, M., Hagiwara, K., and Okuda, A. (2005). Identification of an enhancer that controls up-regulation of fibronectin during differentiation of embryonic stem cells into extraembryonic endoderm. *J. Biol. Chem.* 280, 7244–7252.
- Shomron, N., and Levy, C. (2009). MicroRNA-biogenesis and Pre-mRNA splicing crosstalk. *J. Biomed. Biotechnol.* 2009, 594678.
- Shook, D., and Keller, R. (2003). Mechanisms, mechanics and function of epithelial–mesenchymal transitions in early development. *Mech. Dev.* 120, 1351–1383.
- Sinner, D., Rankin, S., Lee, M., and Zorn, A.M. (2004). Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. *Development* 131, 3069–3080.
- Slack, F., and Ruvkun, G. (1997). Temporal pattern formation by heterochronic genes. *Annu. Rev. Genet.* 31, 611–634.
- Smith, J. (1997). Brachyury and the T-box genes. *Curr. Opin. Genet. Dev.* 7, 474–480.
- Smith AG, Heath JK, Donaldson DD, Wong GG, Moreau J, Stahl M, R.D. (1988). Inhibition of pluripotential embryonic stem cell differentiation by purified polypeptides. *Nature* 336, 688–690.
- Song, L., and Tuan, R.S. (2006). MicroRNAs and cell differentiation in mammalian development. *Birth Defects Res. C. Embryo Today* 78, 140–149.
- Spruce, T., Pernaute, B., Di-Gregorio, A., Cobb, B.S., Merckenschlager, M., Manzanares, M., and Rodriguez, T. a (2010). An early developmental role for miRNAs in the maintenance of extraembryonic stem cells in the mouse embryo. *Dev. Cell* 19, 207–219.
- Stark, A., Brennecke, J., Bushati, N., Russell, R.B., and Cohen, S.M. (2005). Animal MicroRNAs confer robustness to gene expression and have a significant impact on 3'UTR evolution. *Cell* 123, 1133–1146.
- Stott, D., Kispert, a, and Herrmann, B.G. (1993). Rescue of the tail defect of Brachyury mice. *Genes Dev.* 7, 197–203.
- Strähle, U., Blader, P., Henrique, D., and Unit, D.B. (1993). Axial, a zebrafish gene expressed along the the developing body axis , shows altered expression in cyclops mutant embryos. *Genes Dev.* 7, 1436–1446.
- Sui, L., Bouwens, L., and Mfopou, J.K. (2013). Signaling pathways during maintenance and definitive endoderm differentiation of embryonic stem cells. *Int. J. Dev. Biol.* 57, 1–12.
- Sun, X., Meyers, E.N., Lewandoski, M., and Martin, G.R. (1999). Targeted disruption of Fgf8 causes failure of cell migration in the gastrulating mouse embryo. *Genes Dev.* 13, 1834–1846.

- Tada, S., Era, T., Furusawa, C., Sakurai, H., Nishikawa, S., Kinoshita, M., Nakao, K., Chiba, T., and Nishikawa, S.-I. (2005). Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. *Development* 132, 4363–4374.
- Tam, P.P., and Beddington, R.S. (1987). The formation of mesodermal tissues in the mouse embryo during gastrulation and early organogenesis. *Development* 99, 109–126.
- Tam, P.P.L., and Loebel, D. a F. (2007). Gene function in mouse embryogenesis: get set for gastrulation. *Nat. Rev. Genet.* 8, 368–381.
- Tam, P.P.L., and Rossant, J. (2003). Mouse embryonic chimeras: tools for studying mammalian development. *Development* 130, 6155–6163.
- Tam, P.P.L., Loebel, D. a F., and Tanaka, S.S. (2006). Building the mouse gastrula: signals, asymmetry and lineages. *Curr. Opin. Genet. Dev.* 16, 419–425.
- Tamai, K., Semenov, M., Kato, Y., Spokony, R., Liu, C., Katsuyama, Y., Hess, F., Saint-Jeannet, J.P., and He, X. (2000). LDL-receptor-related proteins in Wnt signal transduction. *Nature* 407, 530–535.
- Tamplin, O.J., Kinzel, D., Cox, B.J., Bell, C.E., Rossant, J., and Lickert, H. (2008). 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.
- Tang, F., Barbacioru, C., Bao, S., Lee, C., Nordman, E., Wang, X., Lao, K., and Surani, M.A. (2010). Tracing the derivation of embryonic stem cells from the inner cell mass by single-cell RNA-Seq analysis. *Cell Stem Cell* 6, 468–478.
- Tavazoie, S.F., Alarcón, C., Oskarsson, T., Padua, D., Wang, Q., Bos, P.D., Gerald, W.L., and Massagué, J. (2008). Endogenous human microRNAs that suppress breast cancer metastasis. *Nature* 451, 147–152.
- Tay, Y., Zhang, J., Thomson, A.M., Lim, B., and Rigoutsos, I. (2008). MicroRNAs to *Nanog*, *Oct4* and *Sox2* coding regions modulate embryonic stem cell differentiation. *Nature* 455, 1124–1128.
- Technau, U. (2001). Brachyury, the blastopore and the evolution of the mesoderm. *Bioessays* 23, 788–794.
- Teo, A.K.K., Arnold, S.J., Trotter, M.W.B., Brown, S., Ang, L.T., Chng, Z., Robertson, E.J., Dunn, N.R., and Vallier, L. (2011). Pluripotency factors regulate definitive endoderm specification through *eomesodermin*. *Genes Dev.* 25, 238–250.
- Tesar, P.J., Chenoweth, J.G., Brook, F. a, Davies, T.J., Evans, E.P., Mack, D.L., Gardner, R.L., and McKay, R.D.G. (2007). New cell lines from mouse epiblast share defining features with human embryonic stem cells. *Nature* 448, 196–199.
- The R Core Team, 2013, and Team, T.R.C. (2013). R: A Language and Environment for Statistical Computing.
- Thiery, J.P., and Sleeman, J.P. (2006). Complex networks orchestrate epithelial-mesenchymal transitions. *Nat. Rev. Mol. Cell Biol.* 7, 131–142.

## 5. Reference

---

Thomas, P., and Beddington, R. (1996). Anterior primitive endoderm may be responsible for patterning the anterior neural plate in the mouse embryo. *Curr. Biol.* 6, 1487–1496.

Thomas, P.Q., Brown, a, and Beddington, R.S. (1998). Hex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. *Development* 125, 85–94.

Thomson, J. A. (1998). Embryonic Stem Cell Lines Derived from Human Blastocysts. *Science* 282, 1145–1147.

Todaró, G. J., and Green, H. (1963). Quantitative studies of the growth of mouse embryo cells in culture and their development into established lines. *J Cell Biol.* 17, 299–313.

Tortelote, G.G., Hernández-Hernández, J.M., Quaresma, A.J.C., Nickerson, J. a, Imbalzano, A.N., and Rivera-Pérez, J. a (2013). Wnt3 function in the epiblast is required for the maintenance but not the initiation of gastrulation in mice. *Dev. Biol.* 374, 164–173.

Tsakiridis, A., Huang, Y., Blin, G., Skylaki, S., Wymeersch, F., Osorno, R., Economou, C., Karagianni, E., Zhao, S., Lowell, S., et al. (2014). Distinct Wnt-driven primitive streak-like populations reflect in vivo lineage precursors. *Development* 141, 1209–1221.

Uetzmann, L. Differentiation of the Endoderm Lineage In the Murine System. PhD thesis

Uetzmann, L., Burtscher, I., and Lickert, H. (2008). A mouse line expressing Foxa2-driven Cre recombinase in node, notochord, floorplate, and endoderm. *Genesis* 46, 515–522.

Varlet, I., Collignon, J., and Robertson, E.J. (1997). nodal expression in the primitive endoderm is required for specification of the anterior axis during mouse gastrulation. *Development* 124, 1033–1044.

Velopment, T.H.E.D.E., Normal, O.F., Us, H., Chy, B.R.A., and The, O.F. (1944). The Development of Normal and Homozygous Brachy (TT) Mouse Embryos in the Extraembryonic Coelom of the Chick. *Proc Natl Acad Sci U. S. A.* 30, 134-140.

Vickers, M.M., Bar, J., Gorn-Hondermann, I., Yarom, N., Daneshmand, M., Hanson, J.E.L., Addison, C.L., Asmis, T.R., Jonker, D.J., Maroun, J., et al. (2012). Stage-dependent differential expression of microRNAs in colorectal cancer: potential role as markers of metastatic disease. *Clin. Exp. Metastasis* 29, 123–132.

Vidarsson, H., Hyllner, J., and Sartipy, P. (2010). Differentiation of human embryonic stem cells to cardiomyocytes for in vitro and in vivo applications. *Stem Cell Rev.* 6, 108–120.

Villegas, S.N., Rothová, M., Barrios-Llerena, M.E., Pulina, M., Hadjantonakis, A.-K., Le Bihan, T., Astrof, S., and Brickman, J.M. (2013). PI3K/Akt1 signalling specifies foregut precursors by generating regionalized extra-cellular matrix. *Elife* 2, e00806.

Vincent, S.D., Dunn, N.R., Hayashi, S., Norris, D.P., and Robertson, E.J. (2003). Cell fate decisions within the mouse organizer are governed by graded Nodal signals. *Genes Dev.* 17, 1646–1662.

Viotti, M., Foley, A.C., Hadjantonakis, A.K. (2014) Gutsy moves in mice: cellular and molecular dynamics of endoderm morphogenesis. *Philos Trans R Soc Lond B Biol Sci.* 369(1657).



- Wan, H., Kaestner, K.H., Ang, S.-L., Ikegami, M., Finkelman, F.D., Stahlman, M.T., Fulkerson, P.C., Rothenberg, M.E., and Whitsett, J. A (2004). *Foxa2* regulates alveolarization and goblet cell hyperplasia. *Development* 131, 953–964.
- Wang, P., McKnight, K.D., Wong, D.J., Rodriguez, R.T., Sugiyama, T., Gu, X., Ghodasara, A., Qu, K., Chang, H.Y., and Kim, S.K. (2012). A molecular signature for purified definitive endoderm guides differentiation and isolation of endoderm from mouse and human embryonic stem cells. *Stem Cells Dev.* 21, 2273–2287.
- Wang, Y., Medvid, R., Melton, C., Jaenisch, R., and Blalock, R. (2007). DGCR8 is essential for microRNA biogenesis and silencing of embryonic stem cell self-renewal. *Nat. Genet.* 39, 380–385.
- Warga, R.M., and Nüsslein-Volhard, C. (1999). Origin and development of the zebrafish endoderm. *Development* 126, 827–838.
- Weigel, D., Jürgens, G., Küttner, F., Seifert, E., and Jäckle, H. (1989). The homeotic gene fork head encodes a nuclear protein and is expressed in the terminal regions of the *Drosophila* embryo. *Cell* 57, 645–658.
- Wells, J.M., and Melton, D. a (1999). Vertebrate endoderm development. *Annu. Rev. Cell Dev. Biol.* 15, 393–410.
- Wightman, B., Ha, I., and Ruvkun, G. (1993). Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell* 75, 855–862.
- Wilkinson, D.G., Bhatt, S., and Herrmann, G. (1990). Expression pattern of the mouse *T* gene and its role in mesoderm formation. *Nature* 343(6259):657-659.
- Williams, M., Burdsal, C., Periasamy, A., Lewandoski, M., and Sutherland, A. (2012). Mouse primitive streak forms in situ by initiation of epithelial to mesenchymal transition without migration of a cell population. *Dev. Dyn.* 241, 270–283.
- Wilson, V., Manson, L., Skarnes, W.C., and Beddington, R.S. (1995). The *T* gene is necessary for normal mesodermal morphogenetic cell movements during gastrulation. *Development* 121, 877–886.
- Winnier, G., Blessing, M., Labosky, P. a, and Hogan, B.L. (1995). Bone morphogenetic protein-4 is required for mesoderm formation and patterning in the mouse. *Genes Dev.* 9, 2105–2116.
- Winter, J., Jung, S., Keller, S., Gregory, R.I., and Diederichs, S. (2009). Many roads to maturity: microRNA biogenesis pathways and their regulation. *Nat. Cell Biol.* 11, 228–234.
- Wray, J., and Hartmann, C. (2012). WNTing embryonic stem cells. *Trends Cell Biol.* 22, 159–168.
- Xiong, S., Lin, T., Xu, K., Dong, W., Ling, X., Jiang, F., Chen, G., Zhong, W., and Huang, J. (2013). MicroRNA-335 acts as a candidate tumor suppressor in prostate cancer. *Pathol. Oncol. Res.* 19, 529–537.

## 5. Reference

---

- Xu, N., Papagiannakopoulos, T., Pan, G., Thomson, J. a, and Kosik, K.S. (2009). MicroRNA-145 regulates OCT4, SOX2, and KLF4 and represses pluripotency in human embryonic stem cells. *Cell* 137, 647–658.
- Xu, X., Browning, V.L., and Odorico, J.S. (2011). Activin, BMP and FGF pathways cooperate to promote endoderm and pancreatic lineage cell differentiation from human embryonic stem cells. *Mech. Dev.* 128, 412–427.
- Xu, Y., Zhao, F., Wang, Z., Song, Y., Luo, Y., Zhang, X., Jiang, L., Sun, Z., Miao, Z., and Xu, H. (2012). MicroRNA-335 acts as a metastasis suppressor in gastric cancer by targeting Bcl-w and specificity protein 1. *Oncogene* 31, 1398–1407.
- Yamaguchi, T.P., Takada, S., Yoshikawa, Y., Wu, N., and McMahon, A.P. (1999). T ( Brachyury ) is a direct target of Wnt3a during paraxial mesoderm specification. *Genes Dev.* 13, 3185–3190.
- Yan, Z., Xiong, Y., Xu, W., Gao, J., Cheng, Y., Wang, Z., Chen, F., and Zheng, G. (2012). Identification of hsa-miR-335 as a prognostic signature in gastric cancer. *PLoS One* 7, e40037.
- Yanagisawa, K.O. (1990). Does the T gene determine the anteroposterior axis of a mouse embryo? *Jpn. J. Genet.* 65, 287–297.
- Yanagisawa, K.O., Fujimoto, H., and Urushihara, H. (1981). Effects of the brachyury (T) mutation on morphogenetic movement in the mouse embryo. *Dev. Biol.* 87, 242–248.
- Yang, D., Lutter, D., Burtscher, I., Uetzmann, L., Theis, F.J., and Lickert, H. (2014a). miR-335 promotes mesendodermal lineage segregation and shapes a transcription factor gradient in the endoderm. *Development* 141, 514–525.
- Yang, D., Lutter, D., Burtscher, I., Uetzmann, L., Theis, F.J., and Lickert, H. (2014b). miR-335 promotes mesendodermal lineage segregation and shapes a transcription factor gradient in the endoderm. *Development* 141, 514–525.
- Yasunaga, M., Tada, S., Torikai-Nishikawa, S., Nakano, Y., Okada, M., Jakt, L.M., Nishikawa, S., Chiba, T., Era, T., and Nishikawa, S.-I. (2005). Induction and monitoring of definitive and visceral endoderm differentiation of mouse ES cells. *Nat. Biotechnol.* 23, 1542–1550.
- Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev.* 17, 3011–3016.
- Ying, Q.L., Nichols, J., Chambers, I., and Smith, A. (2003). BMP induction of Id proteins suppresses differentiation and sustains embryonic stem cell self-renewal in collaboration with STAT3. *Cell* 115, 281–292.
- Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev.* 25, 2227–2241.
- Zorn, A.M., and Wells, J.M. (2009). Vertebrate endoderm development and organ formation. *Annu. Rev. Cell Dev. Biol.* 25, 221–251.

---

## 6 Abbreviations

Act A	activin A
ADE	anterior definitive endoderm
AVE	anterior visceral endoderm
A-P	anterior-posterior
BM	basement membrane
BMP	bone morphogenetic protein
Cer-1	Cerberus-like protein 1
Chrd	Chordin
DE	definitive endoderm
DEPC	Diethylpyrocarbonate
DGCR8	diGeorge critical region 8
Dkk1	dickkopf homologue 1
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethylsulfoxide
DTT	dithiothreitol
E	embryonic
EB	embryonic body
ECL	enhanced chemiluminescence
EDM	endoderm differentiation medium
EET	epithelial-to-epithelial transition
EMT	epithelial-to-mesenchymal transition
EGF-CFC	epidermal growth factor-Cripto-FRL1-Cryptic
Eomes	Eomesodermin
EpiSC	epiblast stem cell
ESC	embryonic stem cell
EtBr	Ethidium Bromid
ExE	extraembryonic ectoderm
Exp5	exportin-5
E-Cad	E-cadherin
FACS	fluorescent activated cell sorting

## 6. Abbreviations

---

FBS	fetal bovine serum
FCS	fetal calf serum
Fgf	fibroblast growth factor
Fltp	Flattop
Fn1	Fibronectin1
Foxa2	winged helix/forkhead box transcription factor 2
FRT	flippase recognition target
FVF	Foxa2-Venus fusion
GFP	green fluorescent protein
GOF	gain-of-function
Gsc	gooseoid
H2B-CFP	histone 2B-cyan fluorescent reporter protein
HEK	human embryonic kidney
Hex	haematopoetically expressed homeobox gene
HNF-1 $\beta$	hepatocyte nuclear factor-1 $\beta$
hr	hour
HRP	horseradish peroxidase
ICM	inner cell mass
IFABP	intestinal fatty acid binding protein
IHC	immunohistochemistry
IP	interacting protein
IRES	internal ribosomal entry site
Jak	Janus kinase
LB	lysogeny broth
LEF	lymphoid enhancing factor
Lefty	left right determination factor
LIF	leukemia inhibitory factor
LOF	loss-of-function
L-R	Left-right
MEF	murine embryonic fibroblast
Mest	mesoderm-specific transcript
MET	Mesenchymal-to-epithelial transition
miR	miRNA

---

Mixl1	Mixer-like 1
MMC	mitomycin C
NEAA	Non-Essential Amino Acids
Neo	neomycin
N-Cad	N-cadherin
ON	over night
ORF	open reading frame
pBKS	pBluescriptKS
PBS	Phosphate Buffered Saline
PBST	PBS containing 0.2 % Tween-20
PCP	planar cell polarity
PCR	polymerase chain reaction
PDE	posterior definitive endoderm
PEI	Polyethylenimine
PFA	paraformaldehyde
PGK	phospho-glycerate kinase
PI	protein inhibitor
PLE	profile likelihood estimation
PrE	primitive endoderm
pre-miR	precursor miRNA
pri-miR	primary precursor miRNA
PS	primitive streak
Pyy	peptide YY
P-D	proximal-distal
RISC	RNA-induced gene silencing complex
RT	room temperature
SCF	Sox17-mCherry fusion
SDS	Sodiumdodecylsulphate
Shh	sonic hedgehog
Stat3	Signal transducer and activator of transcription 3
T	brachyury
TAE	Tris-acetate, EDTA
TCF	T-cell factor

## 6. Abbreviations

---

TE	trophectoderm
TE	Tris, EDTA
TF	transcription factor
TGF- $\beta$	transforming growth factor-beta
Trh	thyrotropin-releasing hormone
TSR	transcriptional start region
VE	visceral endoderm
ISH	in situ hybridization
UTR	UnTranslated Region
Vimentin	Vim
WP	Whatman filterpapers
WT	wild type

## Acknowledgements

First and foremost, I would like to thank my supervisors Prof. Heiko Lickert and Dr. Ingo Burtscher. I deeply thank Heiko for giving me the opportunity to study in his lab and his scientific guidance. I am grateful for Ingo's generous technical advisory and support. Thanks to both of you for your supervision and encouragement. The scientific attitude and the knowledge I have learnt during last 5 years is my greatest achievement, and also the greatest wealth for the future.

I'd like to thank my doctorfather Prof. Johannes Beckers and my thesis committee member Prof. Fabian Theis. I am grateful for their comments and inspiring suggestions on my projects during the meetings. Many thanks to our collaborators: Prof. Fabian Theis and Dr. Dominik Lutter for the generating of miR-335 regulation model and data analysis, and their contribution on the publication; Prof. Johannes Beckers and Dr. Martin Irmeler for the performing of mRNA profile and further bio-informatics analysis; Dr. Carsten Marr and Dr. Michael Schwarzfischer for the technical support on the single cell tracking project. Thank you all for your contribution and endless help on my projects.

I am sincerely thankful to our secretary, Donna Marie Thomson for the endless help and great support in the lab as well as in daily life. Whenever I am in trouble, she is always there with kindness to help. Many thanks go to Dr. Mostafa Bakhti for his scientific guidance and for the help with correcting my thesis and manuscript. I appreciate the help and technical support from Wenke Barkey, Heide Oller, Dr. Alexander Korostylev and Dr. Silvia Engert. I thank Dr. Pallavi Mahaddalkar and Dr. Anika Böttcher for correcting my thesis and translation of abstract. I would also like to acknowledge all my colleagues in IDR and ISF at the Helmholtz center for their help and encouragement, as well as for providing a friendly scientific environment. And last, but not least, for the happiness and support from Noah Moruzzi, Aimee Bastidas Ponce, Lisann Heyner, Sabine Stöckl, as well as other colleagues.

I'd like to sincerely thank to the China Scholarship Council (CSC) for the generous financial support that allow me to conduct the Ph.D. study in Technische Universität

## 7. Acknowledgements

---

München (TUM) in Munich and develop my science career. I am appreciated of the help from Jiqiang Dai and Chongling Huang in Chinese Consulate-General in Munich.

Diving into the new fields of science and a new cultural environment, I have suffered a lot during my studies. I am very lucky because I have received so much help and encouragement from my supervisors and colleagues, as well as my friends and family. Special thanks go to my friends who supported and accompanied me during my stay in Munich, as well as the support and wishes from time to time in China.

Finally my special thanks go to my family. I appreciate their endless support and love. As farmers from a small village, my parents have sacrificed their lives for supporting my brother and me. I will never forget the tough times we have overcome together. I love you all!



## Curriculum Vitae

### Dapeng Yang

Date of birth	Feb 27, 1983
Birthplace	Tai'an city
Gender	Male
Nationality	China
Address	Willi-Graf-Str.25/0384
Telephone	+49 89 3187 2062, +49 176 92631531
E-mail	dapeng.yang@helmholtz-muenchen.de dapeng.yang2012@gmail.com

### Educational Background

PhD candidate in biology, Institute of Stem cell Research, Institute of (2009-2014)  
Diabetes and Regeneration Research, Helmholtz Zentrum München,  
Germany

Master of Agriculture in Animal Genetics and Breeding, Northwest A & F (2006-2009)  
University, China

Bachelor of Agriculture in Veterinary Medicine, Shandong Agriculture (2002-2006)  
University, China

### Courses and conferences

1. Protection and Commercialization of intellectual Property (2010), HelmholtzZentrum münchen
2. 3<sup>rd</sup> International Congress on Stem Cells and Tissue Formation (2010), Dresden, Germany(Poster Presentation)
3. Spring School on Systems Biology (2011), Kloster Seeon, Germany
4. Stem Cells in Development and Disease (2011), Berlin, Germany (Poster Presentation)
5. The Stem Cell Niche (2012), Copenhagen, Denmark
6. 4<sup>th</sup> International Congress on Stem Cells and Tissue Formation (2012), Dresden, Germany (Poster Presentation)
7. GfE Summer School: Regulatory Networks in Development (2012), Guenzburg, Germany
8. English Phonetics Training (2013), Helmholtz Zentrum, Munich, Germany
9. Interact (2013), München, Germany
10. International Joint Meeting (2013), Heidelberg, Germany (Poster Presentation)

11. Develop your Entrepreneurial Thinking (2013), Technical University Munich, Munich, Germany
12. ISSCR 12th Annual meeting (2014), Vancouver, Canada (Tear Presentation, Travel award prize)
13. Enterprising Knowledge workshop (2014), Technical University Munich, Munich, Germany

## Publications

**Yang D**, Burtscher I, Irmeler M, Schwarzfischer M, Marr C, Beckers J, Theis F, Lickert H. Delineating the mesendoderm lineage hierarchy in culture (*Manuscripts in preparation*)

**Yang D**, Lutter D, Burtscher I, Uetzmann L, Theis F, Lickert H (2014). miR-335 promotes mesendodermal lineage segregation and shapes a transcription factor gradient in the endoderm, *Development* 141(3) 514-25.

Sun Y, Zan L, Wang H, Guo H, **Yang D**, Zhao X, Gui L (2009). Insulin Inhibits the Expression of Adiponectin and AdipoR2 mRNA in Cultured Bovine Adipocytes. *Asian Australas. J. Anim. Sci* 22 (10) 1429-1436.

**Yang D**, Zan L, Wang H, Ma Y, Tian W, Zhang Y (2009). Genetic variation of calstarcin-1 gene and association with carcass traits in 3 Chinese indigenous cattle. *African Journal of Biotechnology* 8 (12).

Han R, Zan L, **Yang D**, Hao R (2008). SNPs detection of IGF2 gene and its relationship with carcass and meat quality traits in Qinchuan cattle, *Yi Chuan*. 30(12) 1579-84.