TECHNISCHE UNIVERSITÄT MÜNCHEN
Lehrstuhl für Entwicklungsgenetik

Endoderm lineage labelling using BAC reporter constructs in ES cells and mouse embryos

SASCHA IMHOF

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. S. Scherer
Prüfer der Dissertation:	1. apl. Prof. Dr. J. Graw
	2. Univ.-Prof. A. Schnieke, Ph.D.

Die Dissertation wurde am 11.05.2009 bei der Technischen Universität München eingereicht und durch die Fakultät Wissenschaftszentrum Weihenstephan für Ernährung, Landnutzung und Umwelt am 31.07.2009 angenommen.

Strength does not come from physical capacity. It comes from an indomitable will.

Mahatma Gandhi

Indian political and spiritual leader (1869 – 1948)

1 INTRODUCTION

1.1 The early development of the mouse embryo with a focus on endoderm development

Embryogenesis is the process between the fertilized oocyte and birth. As schematized in Figure 1, the fertilized oocyte, the zygote is surrounded by a protecting layer named *zona pellucida*. The zygote starts to divide and forms blastomeres, duplicating their number at each cleavage division. The first divisions form the zygote-blastocyst do not increase the size of the embryo. Then, the 8-cell morula becomes compacted, the blastomeres become polarized and subsequent symmetric or asymmetric cell divisions generate outside polar cells and inside apolar cells (Figure 1) (Rossant and Tam, 2009). The outside polar cells will subsequent

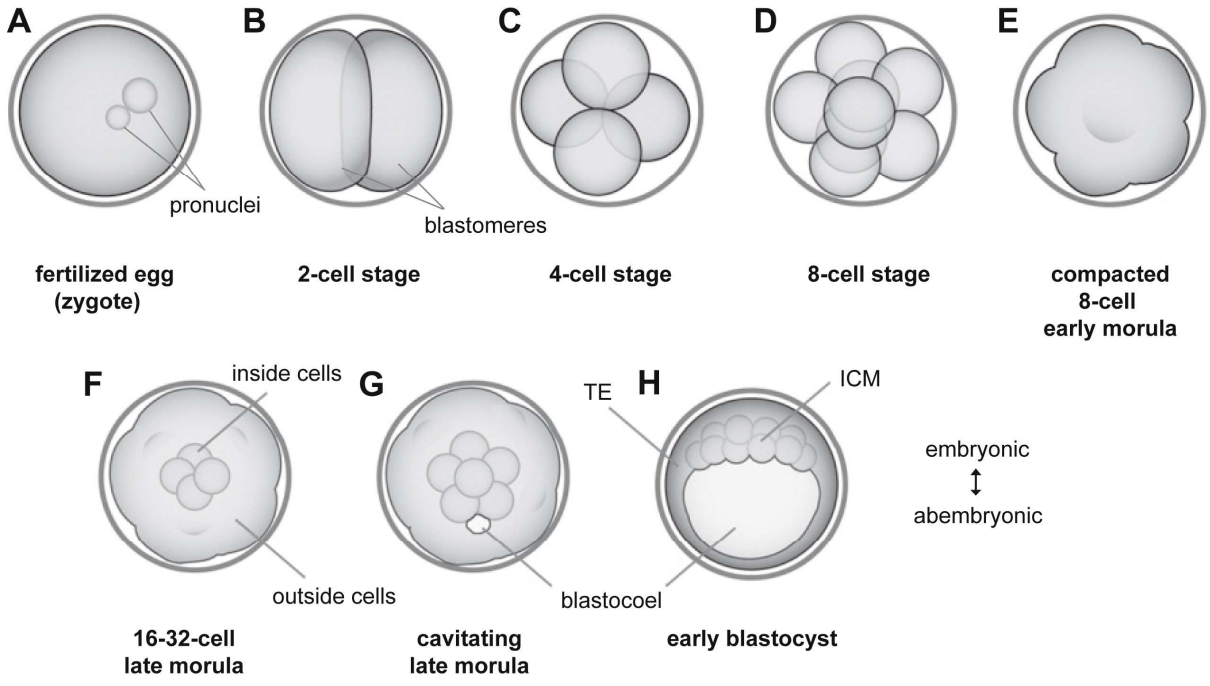


Figure 1 Stages of preimplantation mouse development
(a) The fertilized egg (zygote) containing the two pronuclei and surrounded by the protecting *zona pellucida*. (b)-(d) Three rounds of cleavage division of the blastomeres lead to the 8-cell stage early morula. (e) The early morula cells compact so that cell outlines are no longer visible. (f) At late morula stage, the inside and outside cell populations are established. (g) At late morula stage the nascent blastocoel cavity forms. (h) The early blastocyst with the blastocoel cavity and two distinct lineages, the trophectoderm (TE) and the inner cell mass (ICM).

form the trophectoderm (TE), that gives rise to all the trophoblast cell types, as for example the majority of the fetal part of the placenta. The inside apolar cells become enclosed by the outer epithelium and give rise to the inner cell mass (ICM) from which embryonic stem (ES) cells can be isolated (Martin, 1981).

The first lineagedecision of the developing embryo is the formation of the primitive endoderm (PrE) from the ICM. The PrE gives later rise to the extraembryonic endoderm layers of the visceral and parietal yolk sac, which form the membranes surrounding the development embryo. From embryonic day (E) 3.25 the blastocyst is apparent; and at E3.5 it is a hollow structure with an outer layer of trophectoderm cells, which encloses the fluid filled blastocoel cavity and the asymmetrically positioned ICM. This asymmetry leads to the first embryonic axis which distinguishes between the embryonic side, with ICM and polar trophectoderm (pTE), and the abembryonic side, where the mural trophectoderm (mTE) is localized (Figure 2 a) (Chazaud *et al*., 2006). The segregation of the ICM is accompanied by expression of the key transcription factors (TFs) *Octamer-binding transcription factor 4* (*Oct4*), to establish ICM cell fate, and the *caudal type homeobox 2* (*Cdx2*; James and Kazenwadel, 1991) for Trophectodermal (TE) fate (Nichols *et al*., 1998; Chawengsaksophak *et al*., 1997; Strumpf *et al*., 2005).

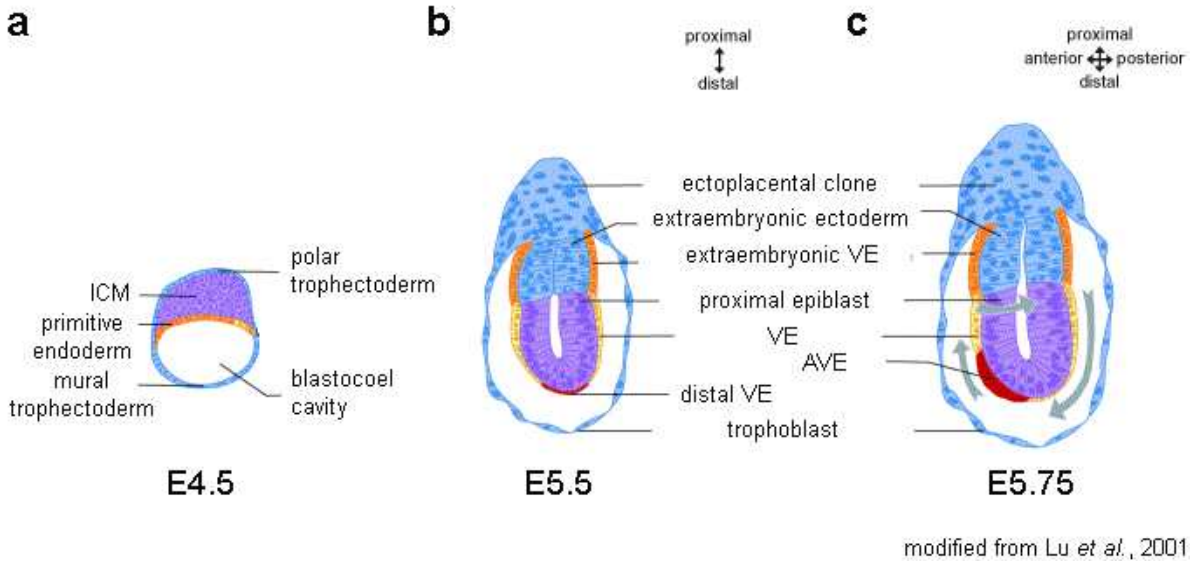


Figure 2 Development of the mouse embryo from blastocyst stage E4.5 to E5.75
(a) An E4.5 blastocyst stage mouse embryo with hatched out *zona pellicula* and the polar trophectoderm (blue) which implants in the uterus. (b) An implanted embryo at E5.5. The VE (red) has formed at the distal tip of the embryo from epiblast cells (violet). (c) At E5.75 the distal VE (red) moves toward the future anterior side and becomes AVE. Abbreviations: ICM = inner cell mass; VE = visceral endoderm; AVE = anterior visceral endoderm.

The ICM is later characterized by the presence of TFs that function as determinants to maintain pluripotency as *Oct4*, *Nanog homeobox gene* (*Nanog*), *SRY-related HMG box transcription factor 2* (*Sox2*) and *Forkhead box transcription factor d3* (*Foxd3*) (Nichols *et al*., 1998; Hanna *et al*., 2002; Avilion *et al*., 2003; Mitsui *et al*., 2003).

At E4.5, the blastocyst hatches out of the *zona pellucida*, implants in the uterus and the ICM differentiates into two distinct tissues: the epiblast and the overlaying PrE (Figure 2 a), which express the two TFs *haematopoetically expressed homeobox gene* (*Hhex*) (Thomas *et el*., 1998; Mesnard *et al*., 2006) and *GATA binding protein 6* (*Gata6*) (Rossant *et al*., 2003). After the implantation, cells of the uterus mucosa interact with trophoblast cells and form the placenta, which is a combination of both extraembryonic and maternal tissue (Wang and Dey, 2006).

Subsequently, cells of the ICM undergo a burst of proliferation and elongate into the blastocoel cavity, generating on one side a distal embryonic part and on the other side a proximal extraembryonic part (Figure 2 b). The embryonic part has a cup shaped structure and consists of the thick one-cell-layer columnar epithelium of the epiblast, surrounded by the one-cell-layer planar epithelium of the visceral endoderm (VE) (Figure 2 b). Epiblast cells that originate from the ICM will give rise to the embryo proper, while the VE that originate from the PrE will form the visceral yolk sac (Wells and Melton, 1999; Beddington and Robertson, 1999; Ang and Constam, 2004). The extraembryonic part consists of the extra-embryonic ectoderm (ExE), which originates from outer TE cells and is also surrounded by the VE, here named abembryonic visceral endoderm (Srinivas, 2006).

At E5.5, as the proximo-distal axis of the embryo has been determined, VE cells at the distal tip of the egg cylinder differentiate into a morphologically distinct tissue, the distal visceral endoderm (DVE) (Figure 2 b). These cells are also characterized by the expression of marker genes such as *Hhex* (Thomas *et al*., 1998; Rivera-Perez *et al*., 2003; Srinivas *et al*., 2004), *Cerberus-like 1* (*Cer1*) (Belo *et al*., 1997) and *left right determination factor 1* (*Lefty1*) (Meno *et al*., 1999). As early as stage E6.0, an anterior-posterior (A-P) embryonic polarity is established, after cells of the DVE start expressing *Cer1* and *Lefty1* that are inhibitors of posterior fate, and move towards one side, which becomes then “anterior”; and the DVE becomes then anterior visceral endoderm (AVE) (Figure 2 c) (Thomas and Beddington, 1996; Thomas *et al*., 1998; Kimura *et al*., 2000; Perea-Gomez *et al*., 2001; Rivera-Perez *et al*., 2003; Srinivas *et al*., 2004; Chazaud and Rossant, 2006). At this point, genes such as

Brachyury (*T*) appear in the proximal epiblast and the ExE (Thomas *et al*., 1998; Liu *et al*., 1999; Brennan *et al*., 2001; Kimura *et al*., 2001; Beck *et al*., 2002; Perea-Gomez *et al*., 2004).

At E6.5, gastrulation starts at the posterior side of the embryo with the formation of the primitive streak. During gastrulation the three principal germ layers, ectoderm, mesoderm and endoderm form (Beddington and Robertson, 1999; Wells and Melton, 1999; Tam and Loebel, 2007). Before gastrulation starts (pre-streak stage, PS; E6.25), the embryonic cup-shaped epiblast is surrounded by the VE (Figure 3 a). With the beginning of gastrulation, pluripotent epiblast cells at the posterior side undergo epithelial-mesenchymal transition (EMT) and ingress in between epiblast and VE to form a new embryonic structure: the primitive streak, along the embryonic-extraembryonic border. The formation of the primitive streak at E6.5 (early-streak stage; ES) on the posterior side is the first visible indication of the introduced anterior-posterior (A-P) axis in the embryo. At this time, the AVE secrets inhibitors of

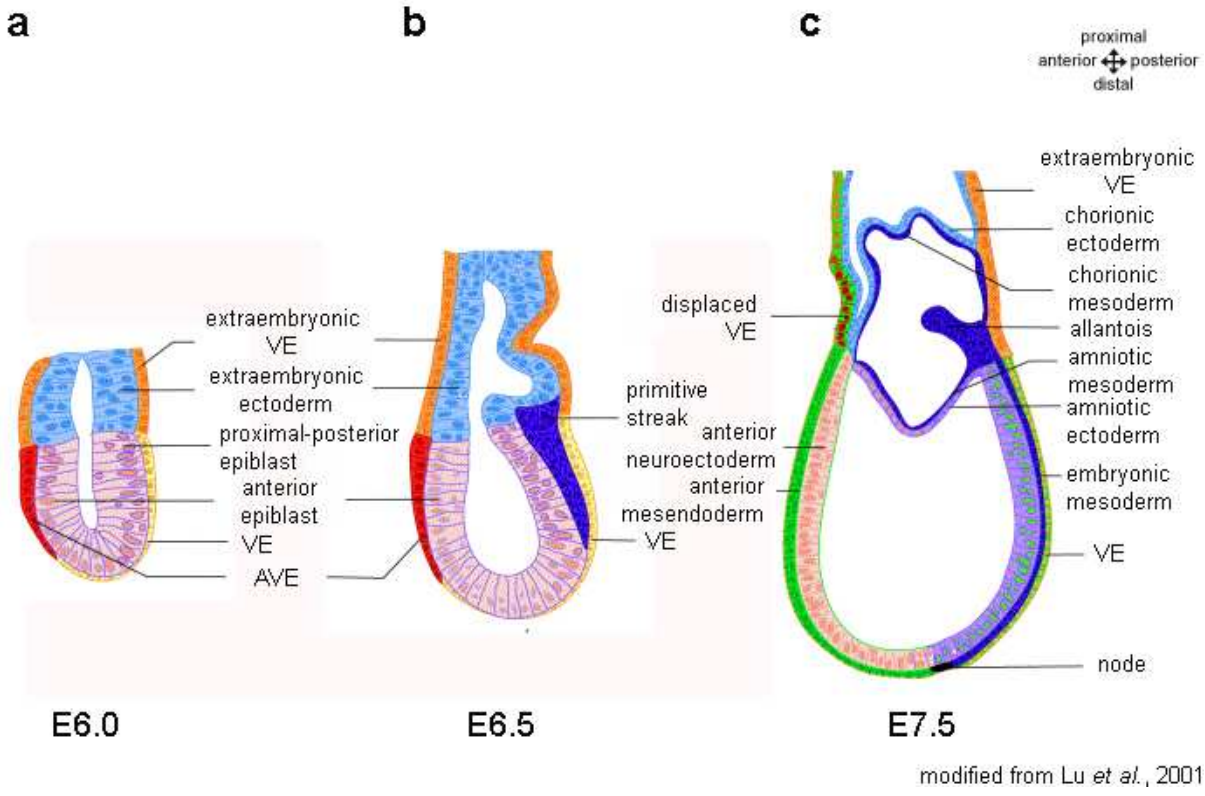


Figure 3 Development of the mouse embryo from pre-gastrulation to late gastrulation stage
(a) Embryo at E6.0 prior to gastrulation with the AVE (red) which establishes the anterior-posterior axis. (b) At E6.5 (early-streak (ES) stage) with the onset of gastrulation the primitive streak (blue) forms at the posterior side of the embryo. The primitive streak forms from epiblast cells (pink) which undergo epithelial-mesenchymal transition and migrate out. The cells give rise to definitive endoderm (DE; green/yellow), which replaces the overlying VE (yellow) or to mesoderm (blue) which forms under the DE. (c) At E7.5 (late-streak (LS) stage) the primitive streak has expanded to the distal tip of the embryo with the node (black) on its end. Abbreviations: VE = visceral endoderm; AVE = anterior visceral endoderm.

Wnt/β-catenin (Wingless/Int) and TGFβ/Nodal (transforming growth factor) signalling to prevent primitive streak formation on the anterior side; while Wnt/β-catenin and TGFβ/Nodal are expressed posteriorly to induce primitive streak formation (Conlon et al., 1994; Liu *et al*., 1999; Kimura *et al*., 2000; Perea-Gomez *et al*., 2002; Tam *et al*., 2006).

Epiblast cells undergo EMT and the cell fate decision between mesoderm and endoderm. Then endodermal cells undergo mesenchymal-epithelial transition (MET), intercalate into VE to form the definitive endoderm (DE) and displace the VE (E7.5, Fig 2.1.3). At the same time, mesodermal cells express the *T* gene and stay between the newly formed ectoderm (from the underlying epiblast) and the endoderm; while endodermal cells express the *Forkhead box transcription factor a2* (*Foxa2*) and form the DE on the surface of the embryo (Figure 3 c) (Beddington and Robertson 1998, Wells and Melton, 1999; Tam *et al*., 2003; Lewis and Tam, 2006). As more DE is recruited into the endodermal layer during gastrulation, the incoming population expands anteriorly and laterally from the site of integration and extends the primitive streak distally (Lewis and Tam, 2006). The existence of a bipotential mesendodermal progenitor cell population, which can give rise to mesodermal and endodermal cell fate, was shown for *C. elegans*, *Xenopus* and zebrafish (Rodaway *et al*., 1999; Kimelman and Griffin, 2000; Rodaway and Patient, 2001).

In vitro differentiation experiments using mouse ES cells (Tada *et al*.; 2005), as well as results from conditional gene targeting experiments (Lickert *et al*., 2002), strongly indicates that such a bipotential mes-endodermal progenitor cell population exists also in the mouse. Recent results from our lab (Burtscher and Lickert, 2008; in revision) showed that already fate-specified *Foxa2* or *T* expressing epiblast cells loose their polarity, undergo EMT, intercalate in the primitive streak and give rise to *Foxa2*-expressing endoderm and *T*-expressing mesoderm. The first markers of the endoderm germ layer are *Foxa2* and *SRY-related HMG box transcription factor 17* (*Sox17*); they are shown to be key factors and essential for early and late endoderm development (Dufort *et al*., 1998; Kanai-Azuma *et al*., 2002; Kinder *et al*., 2001; Lewis and Tam, 2006).

On the distal tip of the primitive streak the anatomical structure of the node forms. This is an analogous structure of the Spemann/Mangold organizer and important for the A-P, dorsal-ventral and left-right patterning of the gastrulation-stage embryo (Beddington, 1994; Tam *et al*., 1997; Kinder *et al*., 2001). The node produces

numerous growth factors, such as members of the fibroblast growth factor (FGF), transforming growth factor (TGFβ) a part of the BMP signalling and Wnt growth factor families, morphogenic factors, such as retinoic acid (RA) and inhibits Chordin, Noggin and Follistatin which are inhibitors of BMP signalling (Hogan *et al*., 1992; Conlon *et al*., 1994; Tam and Behringer, 1997; Yamaguchi and Rossant, 1995; Wells and Melton, 1999).

The node is also able to induce a second embryonic axis, as demonstrated by the famous transplantation experiment of the blastopore lip the organizer region in the newt, by Spemann and Mangold in 1924 from which this organizer region is named "Spemann organizer" (Harland and Gerhart, 1997). Similar organizer regions could be also identified in several vertebrates, as the Hensen´s node in the chick, in the rabbit and the fish (Hensen, 1876; Wetzel, 1925; Waddington, 1933; Shih and Fraser, 1996, Viehbahn, 2001). At successive stages of gastrulation, the organizer cells of the primitive streak are composed of a dynamic population of precursor cells of different mesodermal and endodermal lineages that are destined into different parts of the forming body (Kinder *et al*., 1999; Kinder *et al*., 2001).

Cells with organizer potential at the early-streak stage (ES, E6.5) in the early-streak embryo are called early-gastrula organizer (EGO) and contribute to the most anterior definitive endoderm (ADE) and the precordal plate mesoderm. At mid-streak stage (MS; E7.0) these cells are called mid-gastrula organizer (MGO) and contribute later to the axial mesoderm, ADE and parts of the notochord. At late-streak stage (LS, E7.5), when the primitive streak has elongated to the distal tip of the embryo, the organizer cells on the anteriormostpart are called late-gastrula organizer (LGO) and form the node, which consists of condensed cells and contributes to the floor plate and notochord (Kinder *et al*., 2001).

The progenitor cell population of mesodermal fate is marked by the expression of *T* which is expressed in the posterior epiblast of an ES embryo and in the nascent mesoderm of the primitive streak at MS stage. *T* can be detected in the node and notochord from LS stage onwards (Inman and Downs, 2006). The transcription factor *Foxa2* is expressed at ES stage in the posterior epiblast and is required for the formation of the ADE and axial mesoderm which consists of prechordal plate, notochord and node and later on the heard process (Monaghan et al., 1993; Sasaki and Hogan, 1993). The majority of DE cells ingress through the anterior end of the

primitive streak at the MS stage and intercalate into the overlying VE to give rise to the foregut (Tam and Beddington, 1992; Burtscher and Lickert, 2009).

By the end of gastrulation at E7.5 the VE is completely replaced by mesoderm and covered by the overlying endodermal cell layer which covers the whole embryo (Figure 4 a). The most anterior endoderm of the E7.5 embryo contributes to the ventral foregut adjacent to the developing heart and organs such as liver and ventral pancreas (Figure 4 b and Figure 4 a-c). The anterior-distal and distal endoderm gives rise to dorsal foregut and midgut, which contributes later in development to stomach, pancreas, duodenum and parts of the intestine. The posterior endoderm forms the hindgut, which generates the large intestine and the colon.

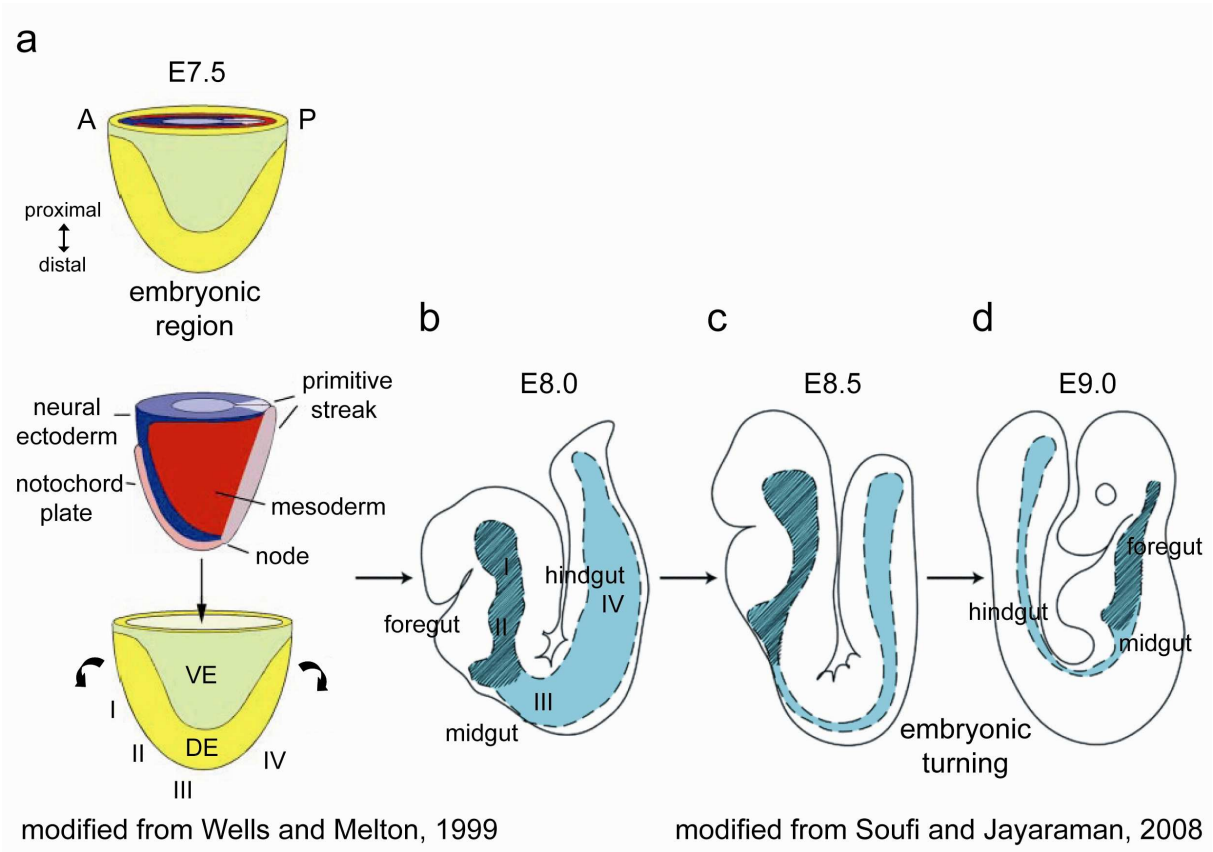


Figure 4 Endoderm formation from gastrulation (E7.5) to early organogenesis (E9.0)
(a) Embryonic region at the end of gastrulation at E7.5 with the three germ layers; endoderm = yellow; mesoderm = red; ectoderm = blue. The notochord (pink) is mesodermally derived, and the primitive streak is derived from the ectoderm (formerly the epiblast) but contains precursors of the endoderm and mesoderm. The outside layer of the embryo is covered by VE (yolk sac endoderm) and DE (gut tube endoderm). The roman numerals I-IV represent regions of the E7.5 endoderm that fate map to regions I-IV of the E8.5 gut tube. The foregut tube forms as region I folds over region II (arrow) and migrates in a posterior direction. The hindgut tube forms when region IV folds over and migrates in anterior direction. (b, c) The positions of the formed gut tube (blue) with the gut regions in the embryo at E8.0 and E8.5 prior to embryonic turning. (d) The embryo at E9.0 after turning with the gut toward the inner (ventral) side of the embryo. Abbreviations: A = anterior; P = posterior; VE = visceral endoderm; DE = definitive endoderm.

At the end of gastrulation the endoderm germ layer is not determined by its regional fate, in contrast to mesoderm and ectoderm. It was in fact demonstrated by Wells and Melton (2000) that the anterior endoderm respecifies after association with posterior tissues. Thus, the endoderm gets specified by its localization along the A-P axis (Lawson *et al*., 1991; Wells and Melton, 2000) and the influence of signalling pathways. In line with this, the restriction of Wnt and Fgf signalling to the posterior side of the epiblast is necessary for the establishment of hindgut versus foregut endoderm. In addition, the proper development of the foregut is dependent on the inhibition of posteriorizing factors (Dessimoz *et al*., 2006; McLin *et al*., 2007).

Also the influence of adjacent germ layers and their signalling activities, which act from close proximity, has been underestimated. Indeed, the influence of BMP and Fgf signalling activities from the adjacent mesoderm is necessary for the induction of liver, pancreas and lung development in the embryo in dose-dependent manner (Rossi *et al*., 2001; Zaret, 2001, Zaret, 2002; Lemaigre and Zaret, 2004; Serls *et al*., 2005; Dessimoz *et al*., 2006, Calmont *et al*., 2006).

An influence on foregut formation and patterning has also been demonstrated for retinoic acid (RA) and Wnt signalling (Wodarz and Nusse, 1998; Ross *et al*., 2000; McLin *et al*., 2007). At the end of gastrulation, four groups of mesoderm can be distinguished: 1. the axial mesoderm located along the midline, that will give rise to Chorda dorsalis; 2. the paraxial mesoderm, that segregates into the somites, which further subdivide into ventral sclerotome (precursor of bones and cartilage) and dorsal dermomyotome (precursor of skeletal muscles, dermis of the back and connective tissues); 3. the intermediate mesoderm, which gives rise to the urogenital apparatus with gonads and cells of the kidneys; 4. the lateral mesoderm gives rise to the heart, the circulatory system, walls of the gut, ventral and lateral embryo. Together with the paraxial mesoderm it gives rise to the lymphatic system, which originates from the veins (Figure 6).

After gastrulation, development proceeds with the metamorphosis of the embryo from a two-dimensional body structure of the epiblast into the three-dimensional architecture of the embryo (Figure 4 a-d; Figure 5 a, b). The morphogenesis begins at E7.75 (early headfold stage, EHF) with the formation of two intestinal pockets: the anterior intestinal port (AIP) that forms on the anterior side by the move of the most anterior DE cells posteriorly; and a second pocket, the caudal intestinal port (CIP), that forms slightly later on the posterior part of the embryo by the move of the most

posterior DE cells anteriorly (Figure 5 b). During this process the first somites (from E8.0) are derived from paraxial mesoderm. As the two gut pockets elongate and close, the lateral walls of the embryonic gut elongate (Figure 4 b), all the tissue folds meet at the yolk stalk and form a closed gut tube (Figure 5 b) (Tam and Behringer, 1997; Wells and Melton 1999; Tremblay and Zaret, 2005; Lewis and Tam, 2006). At the same time, the neuroectoderm forms when dorsal ectodermal tissue invaginates and therefore originates the neural folds; these will move then dorsally and fuse, forming a dorsal tube. All neuronal structures, like brain, spinal cord, nerves, sensory organs and all pigmented cells, and also skin, oral cavity and anus will be formed from the ectoderm (Figure 6).

Between E8.5 and E9.0 the embryo rotates (“turns”) around its A-P axis, so that the dorsal side with the somites will lay outside and the ventral side with the forming

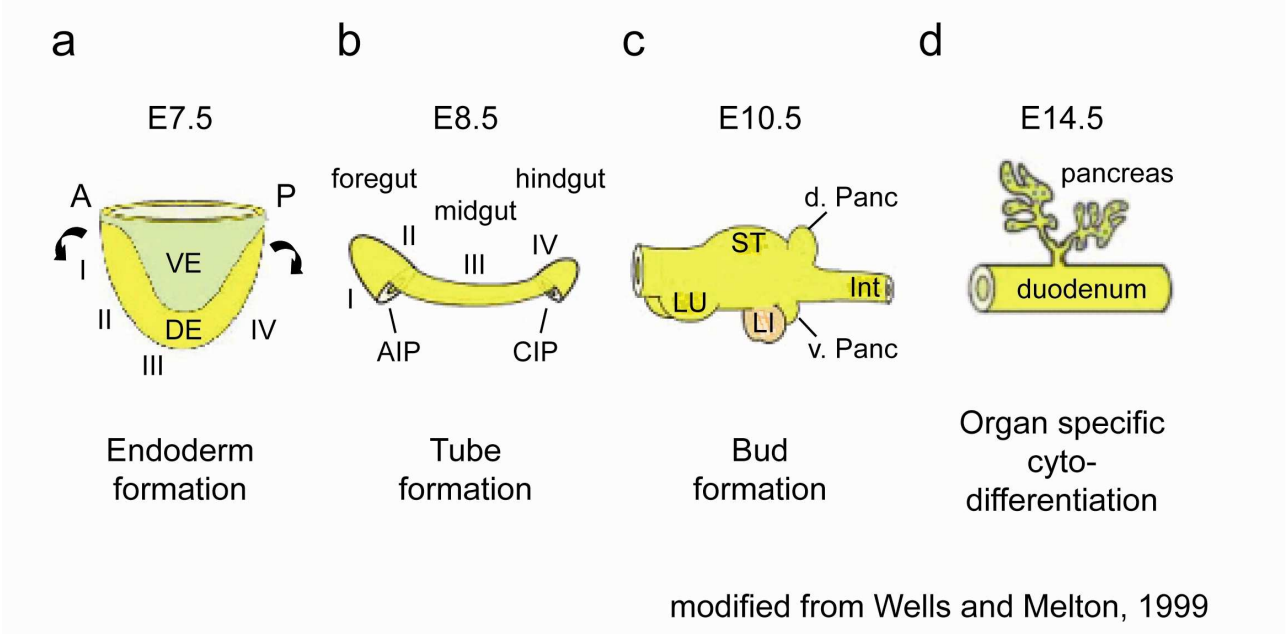


Figure 5 Stages of endoderm development
(a) The embryo at the end of gastrulation with the DE, the gut tube endoderm on the outside layer. The roman numerals I-IV represent regions of the E7.5 endoderm that fate map to regions I-IV of the E8.5 gut tube. The foregut tube forms as region I folds over region II (arrow) and migrates in a posterior direction. The hindgut tube forms when region IV folds over and migrates in anterior direction. (b) The gut tube has formed with the two intestinal pockets, the anterior intestinal port (AIP) and the caudal intestinal port (CIP). The region I give rise to liver, ventral pancreas, lungs and stomach while the region II gives rise to stomach, dorsal pancreas and duodenum. The region III gives rise to the small intestine and region IV gives rise to the large intestine. (c) At E10.5 the organ buds have formed out of the gut tube (LU = lung; ST = stomach; LI = liver; d. Panc = dorsal pancreas; v. Panc = ventral pancreas; Int = duodenum/intestine). (d) The specific branching and differentiation of the organ cell types have occurred. Abbreviations: A = anterior; P = posterior; VE = visceral endoderm; DE = definitive endoderm; AIP = anterior intestinal port; CIP = caudal intestinal port; LU = lung; ST = stomach; LI = liver; d. Panc = dorsal pancreas; v. Panc = ventral pancreas; Int = duodenum/intestine.

organs will lie inside (Figure 4 c, d). At E9.5 the formation of organs of endodermal origin takes place along the A-P axis, which has the morphology of a tube, with an open end on the anterior side. The expression of several TFs and their specific regionalized distribution along the gut tube contribute to differentiate the fore-, mid- and hindgut. Moreover, their particular overlaps determine the different endodermal derived organs (Wells and Melton, 1999).

In particular, the foregut contributes to lung, thyroid, thymus and liver; the midgut contributes to stomach, pancreas, duodenum and parts of the intestine; while large intestine and colon are originated from the hindgut (Figure 4 a, b; Figure 5 a-c). During organogenesis, the endoderm-derived organs bud off from different regions of the primitive gut tube and their specific morphological differences becomes evident at E10.5 (Figure 5 c). Organ development proceeds with specific cyto-differentiation to form functionally different cells to warrant the metabolic and physiological facts of the various organs (Figure 5 d) (Wells and Melton, 1999). The primordium of the lung emerges from the ventral wall of the anterior foregut, as such as the laryngo-tracheal diverticulum (Ten Have-Opbroek, 1981); instead, the dorsal wall of the ventral foregut serves as progenitor of the oesophagus. With ongoing development, the tracheal and oesophageal primordia separate, with the distal end of the tracheal primordium (the lung buds) invading the surrounding splanchic mesenchyme.

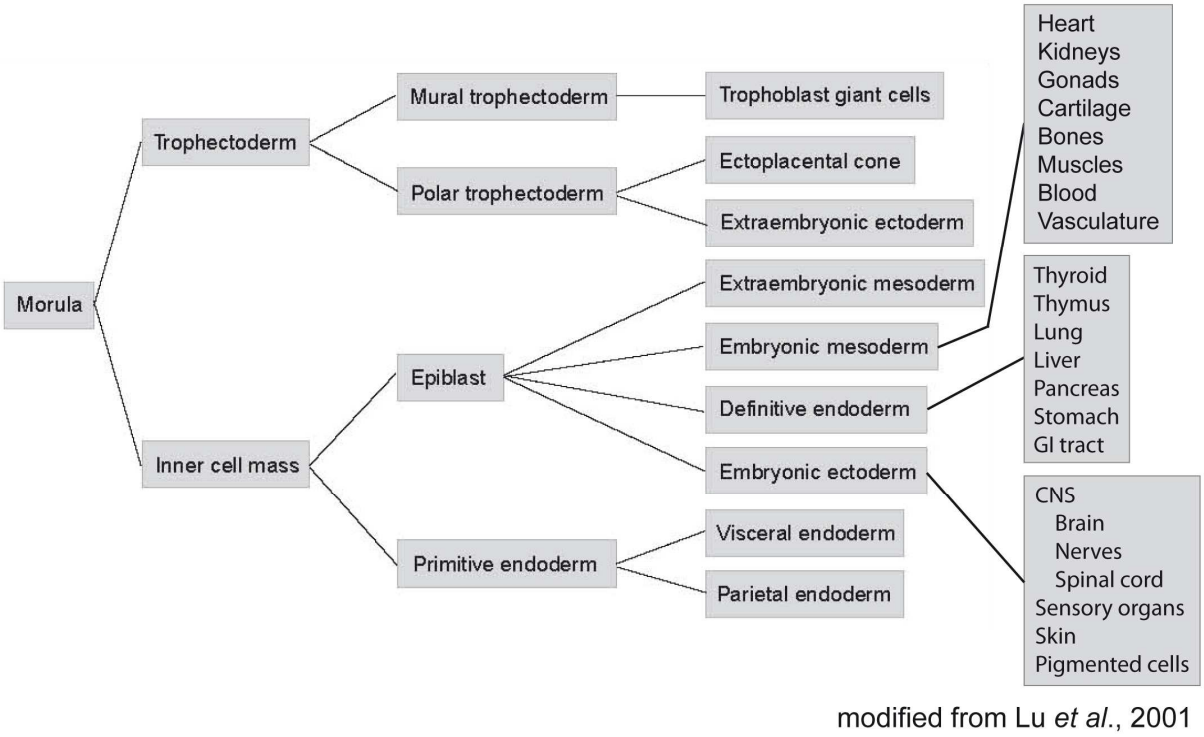
The expression of TFs such as *NK homeodomain transcription factor 2.1* (*Nkx2.1*) or *Gata6* in epithelial cells of the developing lung induces the expression of lung specific genes for further differentiation of the lung epithelium (Liu *et al*., 2002; Nishida *et al*., 2002; Weidenfeld *et al*., 2002; Zhang *et al*., 2005). The uniformly columnar cells of the lung buds undergo branching morphogenesis (E10-E15) and determine lung patterning, giving rise to the formation of conducting and respiratory components of the pulmonary system.

The development of the liver begins with the induction of early liver genes in a particular region of the ventral foregut endoderm, called the hepatic endoderm. A columnar epithelium rapidly forms and by E9.5 the hepatic endoderm cells will delaminate, budding out from the gut lumen into the stroma and form the liver bud.

These cells are now referred to as hepatoblasts (Wilson *et al*., 1963; Lemaigre and Zaret, 2004; Bort *et al*., 2005). The first known effectors for liver bud morphogenesis and of the developing liver are the TFs *Hhex* and *prospero-related homeobox 1* (*Prox1*) (Oliver *et al*., 1993). Once the liver bud emerges, hematopoietic cells migrate

there and start proliferating. The hepatoblasts soon differentiate into their lineage fate and give rise into definitive hepatocytes and bile duct cells (Shiojiri, 1984; Shiojiri *et al*., 2001). Meanwhile, the hepatic epithelium polarizes, therefore creating a small apical domain that lines channels the canaliculi between the cells and connect to bile ducts to further drain into the intestine. The liver as such becomes functional when the basal layer becomes juxtaposed to a fenestrated endothelium that lines sinusoids and forms with infloating blood from the arterial and intestinal portal blood circulation (Zaret, 2001; Zaret, 2002; Lemaigre and Zaret, 2004; Bort *et al*., 2006).

The pancreas development starts with the first expression of pancreas specific genes as *pancreatic-duodenal homeobox 1* (*Pdx1*) in ventral foregut endodermal cells, that line neighbouring hepatic endodermal cells. At E9.5, the outcropping of a dorsal and a ventral pancreatic bud is visible in the gut tube; these then develop into branching ducts and undifferentiated epithelium (E12.5). The buds start to differentiate into endocrine and exocrine cellular lineages, while expanding (E14.0). For the formation of a functional organ, dorsal and ventral pancreas must rotate and fuse, before final maturation of the endocrine cells (E19.0) (Habener *et al*., 2005). Exocrine cells produce digestive enzymes endocrine cells produce insulin and glucagon to regulate blood sugar levels.



modified from Lu *et al*., 2001

Figure 6 Cell lineage decisions in the early mouse embryo
Lineage decision from the beginning of development to the germ layer derived organs and tissues.

1.1.1 The BMP signalling pathway

The secreted bone morphogenic proteins (BMPs) play pleiotrophic roles during embryonic development and in adult life. BMPs were originally identified 40 years ago by their ability to cause bone differentiation (Urist, 1965). The BMP family of proteins possesses potent bone inducing properties (Sampath and Reddi, 1981; Reddi, 1994). Beyond bone formation, an additional role of BMPs for anterior-posterior and dorso-ventral axis formation has been demonstrated (Dale *et al*., 1992; Jones *et al*., 1992).

The TGF-β superfamily can be subdivided into two branches: the BMP/GDF (growth and differentiation factor) and the TGF-β/Activin/Nodal subfamily. Members of the TGF-β/Activin/Nodal subfamily bind to the hetero multimeric transmembrane receptors (serine/threonine kinases) of type II, which in turn phosphorylates and

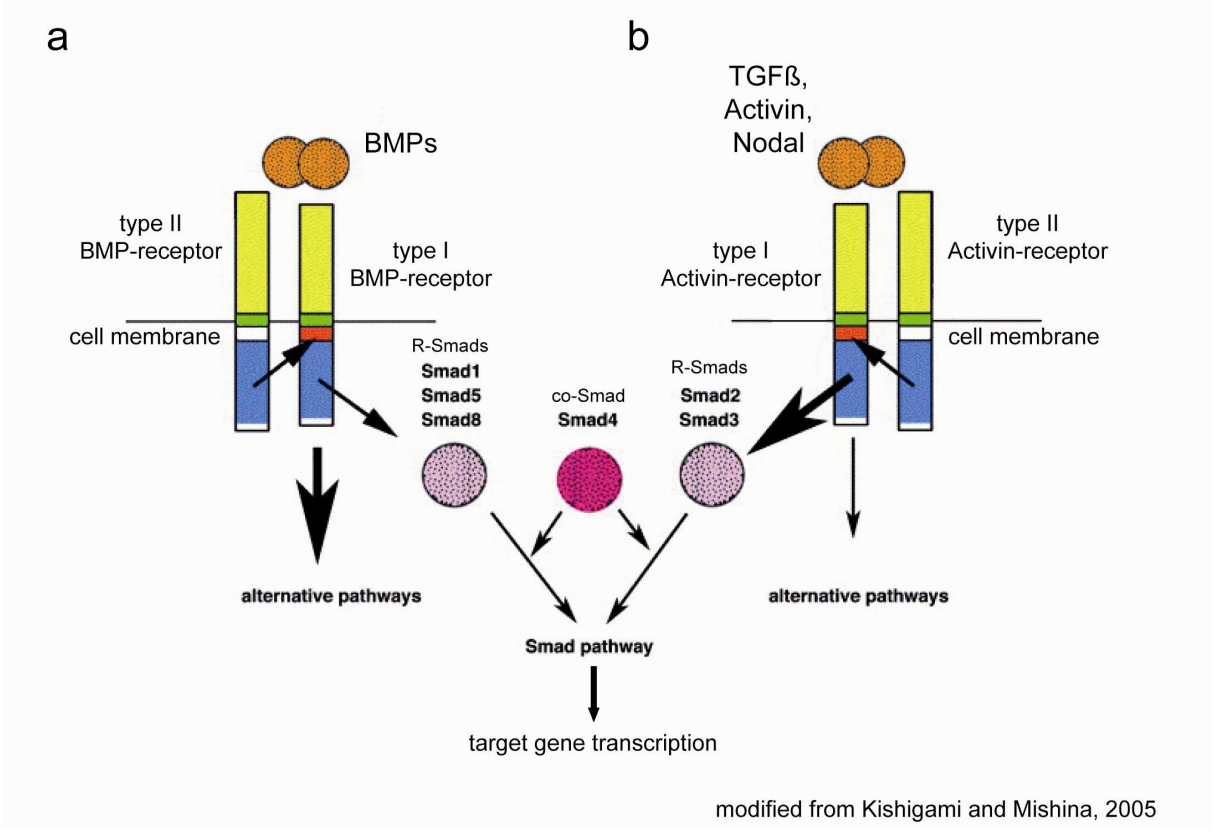


Figure 7 Regulation of BMP signalling
(a) BMPs bind to specific type II / type I transmembrane receptors. The type II receptor activates the type I receptor by phosphorylation which activates the R-Smads Smad1, -5, -8 by phosphorylation. The R-Smads bind to the co-Smad, Smad4 which translocates in the nucleus and activates the transcription of target genes. b) TGFβ, Activin, Nodal bind to specific type II / type I transmembrane receptors. The type II receptor activates the type I receptor by phosphorylation which activates the R-Smads Smad2, -3 by phosphorylation. The R-Smads bind to the co-Smad, Smad4 which translocates in the nucleus and activates the transcription of target genes.

activates a TGF-β/Activin/Nodal specific type I receptor kinase (Figure 7 b). Members of the BMP/GDF subfamily bind to receptors of type II, which then phosphorylates and activates a BMP/GDF specific type I receptor kinase (Figure 7 a). Among the five type II receptors identified in vertebrates, only three act as BMP specific; while among the seven known type I receptors, only three behave as BMP receptors. After ligand binding to TGF-β/Activin/Nodal specific type I receptor kinase, the receptor-activated Smad proteins (R-Smads) Smad2 and Smad3 will be subsequently phosphorylated; whereas the activation of BMP/GDF specific type I receptor will lead to phosphorylation of the R-Smads, Smad1, -5, -8 (Derynck *et al*., 1996; Heldin *et al*., 1997; Wrana and Pawson, 1997; Massague, 1998).

The phosphorylated-activated R-Smads act as signal transducers and bind to the common mediator Smad4 or Smad4β (co-Smad), form heterodimers and translocate into the nucleus, where they activate the transcription of specific target genes (Figure 7) (Derynck and Feng, 1997; Massague, 1998; Zwijsen *et al*., 2003).Notably, the TGF-β receptors can activate different signalling cascades, among which are both Smad-dependent and Smad-independent ones (Derynck and Zhang, 2003).

1.1.2 The transcription factor *Foxa2*: essential for endoderm development

The transcription factor *Foxa2* (*forkhead box transcription factor a2*), formerly known as *hepatocyte nuclear factor 3β* (*HNF3β*) (Lai *et al*., 1991) is part of the group of Fox transcription factors together with *Foxa1* and *Foxa3*, which contain the DNA binding domain forkhead. The gene consists of 3 exons, is located on chromosome 2 and was originally identified by its ability to bind to regulatory elements of liver-specific genes (Lai *et al*., 1991).

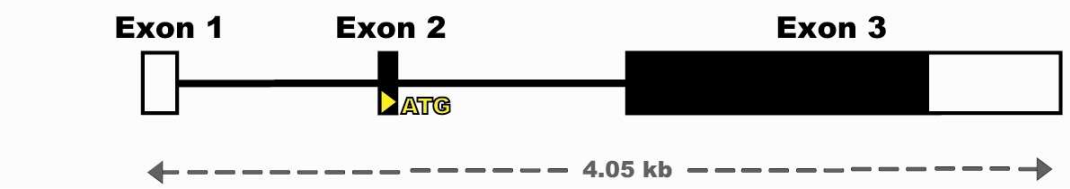


Figure 8 The *Foxa2* locus
The *Foxa2* gene consists of 3 exons (black frames). The ORF (black boxes) starts within exon 2 and ends in exon 3. Abbreviation: ORF = open reading frame.

Homologues were found in human (Ang *et al*., 1993; Monaghan *et al*., 1993; Sasaki and Hogan, 1993), in the fruit fly D. melanogaster (Weigel and Jaeckle, 1990), in the frog *Xenopus laevis* (Strähle *et al*., 1993) and in hydra (Martinez *et al*., 1997).

The first *Foxa2* expression can be detected at E6.0 in the AVE where it regulates the expression of *Orthodenticle homolog 2* (*Otx2*); there it plays a role for the establishment of the A-P axis by regulation of AVE-specific genes and of Nodal (Zhou *et al*., 1993) and Wnt antagonists *dickkopf 1* (*Dkk1*; Glinka *et al*., 1998) and *Cerberus-like* (*Cerl*; Kimura-Yoshida *et al*., 2007). At the beginning of gastrulation (E6.5), *Foxa2* is expressed in the first DE cells which appear on the posterior side of the embryo (Sasaki and Hogan, 1993; Monaghan *et al*., 1993; Ang *et al*., 1993; Ang *et al*., 1994).

Foxa2 is known to bind to specific enhancers and open chromatin, therefore facilitating DNA interaction with other transcription-factors for consequent activation of gene expression. Thereby, *Foxa2* acts for example as pioneer factor for the induction of endodermal and hepatic genes expression (Gualdi *et al*., 1996; Bossard and Zaret; 1998; Cirillo *et al*., 2001; Zaret, 2002). The expression of *Foxa2* during gastrulation is restricted to VE, DE, node and axial mesoderm. At headfold stage *Foxa2* is present in the foregut endoderm, neural plate and notochord. As the initiation of liver development is dependent on *Foxa2* expression, this gene can be detected in the liver primordium at E9.5 (Overdier *et al*., 1994; Kaufmann and Knochel, 1996; Lee *et al*., 2005).

The influence of *Foxa2* on branching morphogenesis of the developing lung was demonstrated in mutant animals lacking *Foxa2*, that therefore fail the transition to air breathing at birth due to missed inactivation of surfactant proteins and lipid synthesis (Wan *et al*., 2004; Wan *et al*., 2005). At E16.5, *Foxa2* expression is detectable in endoderm-derived organs like liver, lung and pancreas, with an especially high expression in the latter (Sund *et al*., 2001; Lantz *et al*., 2004). In the adult, *Foxa2* is expressed predominantly in the liver but is also present in lungs, intestine and stomach (Lai *et al*., 1991; Kaestner *et al*., 1994). Moreover, its expression is detectable in pancreas, in liver metabolism and in type II pneumocytes of the lung (Mason *et al*., 1997; Lee *et al*., 2002; Zhang *et al*., 2005; Friedman and Kaestner, 2006).

The continuous expression of *Foxa2* from early embryonic stages until the adulthood supports its role in both specification and maintenance of endoderm-derived tissues

(Sasaki and Hogan, 1993; Monaghan *et al*., 1993; Ang *et al*., 1993; Ruiz i Altaba, 1994; Ang and Rossant, 1994). To deeper clarify the role of *Foxa2* in development, knockout mice were generated. Embryos lacking *Foxa2* (-/-) die at the age of E9-E10 and are smaller in size as compared to wild-type littermates. Among their phenotypes, they show absence of node and notochord; they also manifest disorganized somites, neural tube, fore- and midgut defects and a failure of proper primitive streak elongation (Ang *et al*., 1994; Weinstein *et al*., 1994; Levinson-Dushnik *et al*., 1997; Dufort *et al*., 1998). As they lack a notochord, null mutant embryos do also not develop floorplate and motor neurons.

1.1.3 The transcription factor *Sox17*: necessary for proper endoderm development

The transcription factor *Sox17* (*SRY related HMG box transcription factor 17*) (Kanai *et al*., 1996) belongs to the group of Sry-related HMG-box transcription factors. The gene consists of 5 exons, is located on chromosome 1 and was identified in a mouse testis cDNA screen (Kanai *et al*., 1996). 33 Sox genes are until now classified and divided in the subgroups named from A to J. *Sox17* belongs together with *Sox7* and *Sox18* to subgroup F, which contains TFs involved in early embryonic development (Bowles *et al*., 2000; Lefebvre *et al*., 2007). Orthologous genes have been identified in the zebrafish, *Danio rerio* (Alexander and Stainier, 1999) and in the frog *Xenopus laevis* (Hudson *et al*., 1997; Clements and Woodland, 2000).

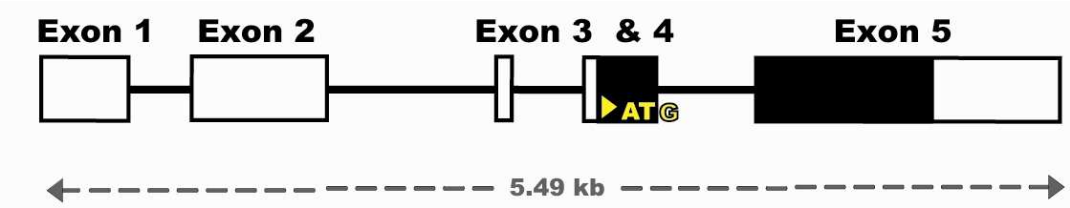


Figure 9 The *Sox17* locus
The *Sox17* gene consists of 5 exons (black frames). The ORF (black boxes) starts within exon 4 and ends in exon 5. Abbreviation: ORF = open reading frame.

The earliest *Sox17* expression can detected in the VE at E6.0. At E7.0, it is expressed in the DE and embryonic hindgut at the anterior end of the primitive streak (Kanai *et al*., 1996; Kanai-Azuma, 2002). At HF stage the expression is restricted to the foregut endoderm and the VE. From E9.5 the expression is restricted to the vascular endothelial cells of the dorsal aortae and the intersomitic blood vessels as

well as in fetal hematopoietic stem cells (Matsui *et al*., 2006; Kim *et al*., 2007). *Sox17* is important for cardiac specification in primitive mesoderm and cardiac myogenesis (Kitajima *et al*., 2000). Later in development, *Sox17* is expressed in testis and lung (Kanai *et al*., 1996), where it becomes restricted to respiratory epithelial cells and primary ciliated cells (Park *et al*., 2006).

Sox17 -/- mutant mice show a reduced invasion of posterior and lateral region of the prospective mid- and hindgut by the DE in HF stage embryos (Kanai-Azuma *et al*., 2002); on the contrary, the prospective foregut develops properly. *Sox17* absence also results in early lethality by E9.5 and the embryos are not able to turn, due to defects in gut endoderm. The essential role of *Sox17* in endoderm formation was also shown in its direct interaction with *Foxa2* (Sinner *et al*., 2004).

1.1.4 *Brachyury* (*T*): important for development of the mesoderm

A mutation (short tail) of the T-box gene *Brachyury* (*T*) was described since the twenties of the last century (Dobrovolskaia-Zavadskaia, 1927; Chesley, 1935; Glücksohn-Schönheimer, 1938; Grüneberg, 1958). The gene was cloned from Herrmann *et al*. in 1990. *T* is a member of the T-box genes which play roles during gastrulation, embryogenesis and organogenesis. Homologues were found in the zebrafish, *Danio rerio* (Schulte-Merker *et al*., 1994), the frog, *Xenopus laevis* (Smith *et al*., 1991) and in the chick, *Gallus gallus* (Kispert *et al*., 1995; Liu *et al*., 2003). The gene consists of 8 exons and is localized on chromosome 17.

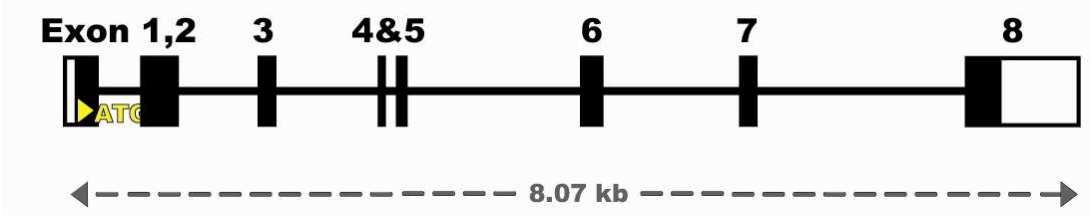


Figure 10 The *Brachyury* (*T*) locus
The *Brachyury* (*T*) gene consists of 8 exons (black frames). The ORF (black boxes) starts within exon 1 and ends in exon 8. Abbreviation: ORF = open reading frame.

The first expression of *T* can be detected at E5.5 in the distal extra-embryonic ectoderm and later on the posterior side of the epiblast (Rivera-Perez and Magnuson, 2005). Then, *T* can be detected in the primitive streak at the onset of gastrulation (E6.5); later, it is expressed in the expanding primitive streak, as it

extends distally to its anterior limit, and additionally in the nascent mesoderm. In line with its expression in the streak, *T* was also observed in derivatives of the streak such as node, notochord, midline and nascent embryonic mesoderm, as well as in notochord-associated endoderm and posterior neuroectoderm (Mac Murray and Shin, 1988).

At E9.0, *T* expression is restricted to posterior mesoderm and notochord, posterior neuroectoderm and the hindgut endoderm associated with the notochord (Yanagisawa, 1990; Inman and Downs, 2006). *T* homozygote null mutant embryos die at E10.0 due to a deficiency in mesoderm formation that blocks primitive streak elongation and results in an incomplete axial development and truncation of the posterior body axis (Wilkinson *et al*., 1990; Herrmann, 1991). Mice which are heterozygote for *T* show defects in completing axial development and thereby have short tails.

1.1.5 The role of *Hhex* during endoderm development and organogenesis

The transcription factor *Hhex* (*haematopoietically expressed homeobox*, also *Hex* or *HHEX* in human) (Bedford et al., 1993), *prospero-related homeobox 1* (*Prhx1*) or *prolin-rich homeodomain* (*Prh*) (Crompton *et al*., 1992; Oliver *et al*., 1993; Hromas *et al*., 1993) is part of the group of homeodomain transcription factors. The gene consists of 4 exons, is located on chromosome 19 and was identified in a screen for homeobox genes. Orthologous of *Hhex* have been found in human (Crompton *et al*., 1992; Hromas *et al*., 1993), rat (Tanaka *et al*., 1999), chick (Crompton *et al*., 1992; Yatskievych *et al*., 1999), zebrafish (Ho *et al*., 1999) and frog (Newman *et al*., 1997), thus highly indicating that it is conserved throughout evolution.

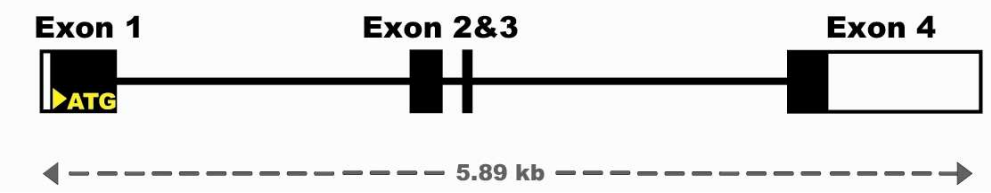


Figure 11 The *Hhex* locus
The *Hhex* gene consists of 4 exons (black frames). The ORF (black boxes) starts within exon 1 and ends in exon 4. Abbreviation: ORF = open reading frame.

First *Hhex* expression in mice can be detected in the PrE of blastocyst stage at E4.5, in the epithelial layer lining the top of the ICM (Thomas *et al*., 1998; Mesnard *et al*., 2006; Chazaud *et al*., 2006). At stage E5.0-E5.25, *Hhex* becomes downregulated and at stage E5.5, its expression is confined to a small patch of VE cells at the distal tip of the egg cylinder that are destined to give rise to AVE. As these cells move to an anterior position, they still express *Hhex*. Therefore, *Hhex* expression is considered to mark the earliest molecular A-P asymmetry in the mouse embryo.

At E5.75 *Hhex* expressing cells are localized to the anterior side of the epiblast in the AVE. At early-streak stage (ES; E6.5) *Hhex* positive cells can be found on the anterior tip of the primitive streak in the earliest DE; and from mid-streak stage (MS, E7.0) to late-streak stage (LS, E7. 5) they can be seen in the AVE, axial mesendoderm and blood islands. The expression of *Hhex* at E8.5 is localized to the endothelial cell precursors, notochord and the foregut diverticulum. At E9.5, this is restricted to the primordia of liver and thymus, where it represents one of their earliest markers (Crompton *et al*., 1992; Keng *et al*., 1998; Thomas *et al*., 1998; Rodriguez *et al*., 2001; Rivera-Perez *et al*., 2003).

Within the mesoderm, *Hhex* is transiently expressed in the nascent blood islands of the visceral yolk sac, where primitive erythrocyte and blood vessel formation is initiated; and at later stages it can be detected in embryonic angioblasts and endocardium. *Hhex* positive cells give rise to several foregut endoderm-derived organs. At E10.0, *Hhex* mRNA is detectable in the forming thymus, liver, thyroid, gallbladder and dorsal pancreatic bud. From E13.5 to E16.5, the expression persists in thymus, liver and epithelial cells of the pancreas and lung. At E18.5, *Hhex* mRNA expression is present in thymus and lung, where it is maintained until adulthood, when it can also be found in the liver (Bogue *et al*., 2000).

In *Hhex*-null embryos at E7.5, in particular the DE appears to be compromised in its anterior displacement. These embryos show also defects in liver, thyroid and hematopoietic lineages. Moreover, the lack of *Hhex* results in a complete failure of ventral pancreatic specification, while the specification of the liver proceeds, although the proliferation of cells in the diverticulum is reduced (Bort *et al*., 2004). The migration of early hepatoblasts into the surrounding mesenchyme also fails to occur (Bort *et al*., 2006), thus resulting in the absence of the liver bud (Keng *et al*., 2000; Martinez-Barbera *et al*., 2000). The embryos can still properly turn and the gut tube closure occurs normally. As they grow until E11.5, it is possible to further phenotype

them and observe a lack of liver, thyroid and parts of the forebrain (Martinez-Barbera *et al*., 2000; Keng *et al*., 2000). Furthermore, it has been shown that *Hhex* plays a crucial role in several cell-lineage differentiation pathways including haematopoiesis, lymphopoiesis, cardiogenesis and the vascular smooth muscle cell differentiation (Bogue *et al*., 2003; Oyama *et al*., 2004; Foley and Mercola, 2005; Kubo *et al*., 2005). Recently it was shown that polymorphisms of the human *HHEX* gene may confer risk of type 2 diabetes (Grarup *et al*., 2007; Furukawa *et al*., 2007; Staiger *et al*., 2008; Herder *et al*., 2008).

1.1.6 The role of *Nkx2.1* for endodermal organ development

The transcription factor Nkx2.1 (NK homeodomain transcription factor 2.1) also known as thyroid transcription factor 1 (Titf1 or TTF1 in human; Guazzi et al., 1990) or thyroid-specific enhancer-binding protein (T/ebp; Mizuno et al., 1991) belongs to the group of NK-2 class of homeobox genes. The gene consists of 3 exons, is located on chromosome 12 and was first identified in the fruit fly, *Drosophila melanogaster* (Kim and Nirenberg, 1989).

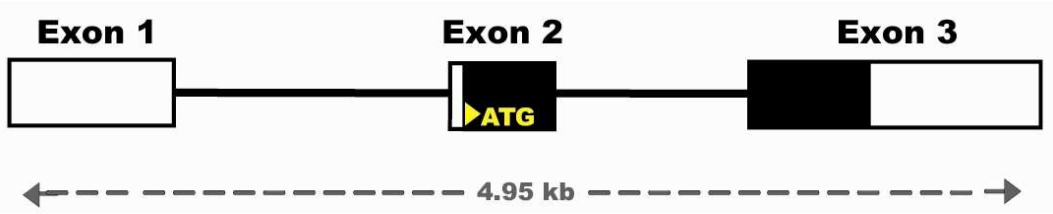


Figure 12 The *Nkx2.1* locus
The *Nkx2.1* gene consists of 3 exons (black frames). The ORF (black boxes) starts within exon 2 and ends in exon 3. Abbreviation: ORF = open reading frame.

In the mouse the earliest expression of *Nkx2.1* appears at E9.0 in a group of cells surrounding the ventral wall of the anterior foregut. At E9.5 *Nkx2.1* is expressed in the developing lung fields, thyroid and forebrain which originated from the gut endoderm (Lazzaro et *al*., 1991; Kimura *at el*., 1999; Minoo *et al*., 1999). Thus, *Nkx2.1* is considered as the earliest marker of lung endodermal cells (Desai *et al*., 2004). In the lung, *Nkx2.1* is expressed in the two main bronchi of the primitive lung bud, in all epithelial cells at the early pulmonary morphogenesis and it marks all pulmonary epithelial derivatives. In the adult lung *Nkx2.1* expression is confined to distal alveolar and proximal bronchiolar epithelial cells. *Nkx2.*1-null mice show an

early and severe interruption of branching morphogenesis of the lung, as well as a severe loss of epithelial cell differentiation (Kimura *et al*., 1996; Minoo *et al*., 1999; Yuan *et al*, 2000). The targeted disruption of the *Nkx2.1* locus results in immediate postnatal death and exhibit a hypoplastic lung phenotype, namely the lungs contain only a rudimentary bronchial tree, show abnormal bronchial epithelium and lack lung parenchyme (Kimura *et al*., 1996). The importance of *Nkx2.1* for lung development was also shown in *in vitro* experiments in which *Nkx2.1* induces the transcription of lung-specific genes (Sawaya *et al*., 1993; Bohinski *et al*., 1994; Bruno *et al*, 1995; Kelly *et al*., 1996). In humans, dominant mutations of *Nkx2.1* lead to congenital pulmonary and cardiac defects (Krude *et al*., 2002).

1.1.7 The role of *Pdx1* for endodermal organ development

The transcription factor *Pdx1* (*pancreatic duodenal homeobox 1*) was identified contemporaneously by several laboratories and has therefore multiple names, as *insulin-promotor binding protein* (*Ipf1*; Ohlsson *et al*., 1993), *somatostatin transactivating factor 1* (*Stf1*) (Leonard *et al*., 1993), and *islet/duodenum homeobox-1* (*Idx1*; Miller *et al*., 1994). *Pdx1* belongs to a para-*Hox* gene cluster which lies outside the major Hox cluster of homeodomain proteins (Brooke *et al*., 1998) consisting of *Pdx* , *Gsh* (Singh *et al*., 1991) and *Cdx* (Duprey *et al*., 1998), all expressed in anterior, middle, and posterior locations in the vertebrate axis. The gene consists of 2 exons and is located on chromosome 5.

Figure 13 The *Pdx1* locus
The *Pdx1* gene consists of 2 exons (black frames). The ORF (black boxes) starts within exon 1 and ends in exon 2. Abbreviation: ORF = open reading frame.

The first expression of *Pdx1* in mouse embryos can be detected at E8.5 in cells of the foregut endoderm, which later give rise to ventral and dorsal pancreatic buds (E9.5), where Pdx1 is also expressed. From E11.5 to E13.5 the *Pdx1* expression is localized to the ductal tree and shifts at E14-E15 to the endocrine compartment when

the exocrine pancreas forms and the islets begin to develop. From E13.5 to E16.5, the expression of *Pdx1* colocalizes with amylase, an exocrine cell-specific enzyme, and by E18.5 it becomes mostly restricted to the mature β-cells of the endocrine pancreas (Guz *et al*., 1995; Offield *et al*., 1996).

In the adult pancreas *Pdx1* expression can be observed at highest levels within the β-cells. *Pdx1* is a master regulator of pancreatic endocrine cell development, involved in both the differentiation of progenitor cells into the β-cell phenotype and for adult islet β-cell function (Stoffers *et al*., 1997; Habener and Stoffers, 1998; Gu *et al*., 2002). In the differentiated β-cell, *Pdx1* acts as glucose-responsive regulator of the insulin gene expression (Petersen *et al*., 1994; Petersen *et al*., 1998). It was shown that *Pdx1* is a β- and δ-cell specific regulatory transcription factor for transcriptional expression of insulin and somatostatin genes, as well as regulator of the expression of islet-specific genes (Waeber *et al*., 1996; Watada *et al*., 1996).

In response to glucose, Pdx1 function is regulated by its phosphorylation and nuclear translocation (McFarlane *et al*., 1994; McFarlane *et al*., 1997; Rafiq *et al*., 1998). *Pdx1* expression in the developing pancreas is maintained during the whole development and provides spatial and temporal contributions to the commitment of endoderm into the pancreatic organ. Thus, as expected for a gene responsible for pancreas development, its disruption results in lack of this organ (Jonsson *et al*., 1994; Offield *et al*., 1996).

Pdx1 null mice die a few days after birth due to defective pancreatic epithelium formation. The visibly normal pancreatic epithelium is not able to differentiate due to its inability to respond to mesenchymal signals which are required for its further differentiation (Ahlgren *et al*., 1996). These animals also show defects in development of the rostral duodenal epithelium (Offield *et al*., 1996). Interestingly, a few insulin-expressing cells could be detected in *Pdx1* null mice, suggesting that a population of insulin-positive, but *Pdx1*-negative, cells may occur distinct from the mature *Pdx1*-expressing β-cells of a developed pancreas. In humans, the analysis of a person born without a pancreas, who was found homozygous for an inactivating mutation in *Pdx1* (Stoffers *et al*., 1997), revealed the importance of *Pdx1* for pancreas formation in vertebrates.

The role of *Pdx1* for the expression of genes which regulate the sensing and regulation of the blood glucose level is also shown by the phenotyping of *Pdx1* heterozygous mice mutants, which show symptoms of glucose intolerance; the fact

that already heterozygote mice show a defect in glucose metabolism suggests a dosage effect on insulin expression and β-cell function (Ahlgren *et al*., 1998; Dutta *et al*., 1998). In humans, a heterozygous mutation leads to susceptibility to a Type II diabetes form (Stoffers *et al*., 1997).

1.2 BAC transgenes: A technology which allows expression of reporter genes

The construction of the first bacterial artificial chromosome (BAC) was done in 1992 during the Human Genome Project (Shizuya *et al*., 1992). The system based on the construction of vectors which bases on the well-characterized *Escherichia coli* (*E. coli*) F-factor, a low copy plasmid that exists in a supercoiled circular form in host cells. The replication of the F-factor is strictly controlled by the regulatory functions of *E. coli*, and as a result the F-factor is maintained in a copy number of one or two copies per cell. The structural features of the F-factor allow stable maintenance of individual DNA clones as well as easy manipulation of the cloned DNA (Shizuya and Kouros-Mehr, 2001).

Practically, for generation of a BAC library the genomic DNA of an organism is partially digested, selected by size (e.g. between 100-250 kb) and subcloned in a BAC vector. An advantage of BACs is the high stability of the inserted DNA, which is shown to be maintained after more than 100 bacterial generations (Cai *et al*., 1995; Nakamura *et al*., 1997). BAC technology, developed to enable the faster and more convenient genome sequencing and analysis of the human genome has developed rapidly and is used in a wide array of applications. Beginning from the generation of BAC libraries which containing the whole genome of organisms BACs had been used for genome sequencing and the formation of physical maps of the genome, BACs have been widely used to identify disease loci as well as individual genes e.g. by positional cloning or the library screening for known sequences. After this, the generation of transgenic mice by using BACs has become fashionable (Nielsen *et al*., 1997).

Traditionally produced transgenic mice, created by microinjection of copies of the gene of interest into mice fertilized eggs, and then introduced as embryos into a mouse, have some disadvantages. Often the upstream and downstream regulatory elements of the gene are not included in the transgene due to their restricted size,

which results in an incomplete transgene expression (Nielsen *et al*., 1997). Due to the large insert size of BACs of ~100-300 kb in most of the cases all of the regulatory elements can be found on a single BAC (Yang *et al*., 1997; Antoch *et al*., 1997; Gong *et al*., 2003; Heintz, 2004). BAC transgenes can emulate the native environment of a gene resulting in a tightly regulated pattern even if randomly inserted in the genome. It was shown that the large genomic clones of BACs often contain DNA sequences which are able to organize the chromatinstructure at the local level and so become immune to position effects which affect expression of smaller transgenes (Nistala and Sigmund, 2002).

During the last years the possibility of BAC modification in a different manner has made the BAC system more attractive for different approaches. Several methods are now available for manipulations of BACs by homologous recombination in E. coli (Yang *et al*., 1997; Muyrers *et al*., 1999; Lee *et al*., 2001; Gong *et al*., 2002) or Cre / *loxP* site specific recombination from bacteriophage P1 (Mejia and Larin, 2000; Wang *et al*., 2001). BAC transgene expression is even proportional to the inserted number of copies (Gebhard *et al*., 2007). The possibility of multi-copy number insertion has even become more important for the introduction of fluorescent reporters in the genome since the diversity of fluorescent reporters has been increased in number and quality in terms of reduced toxicicity and increased emission strength (Campbell *et al*., 2002; Vintersten *et al*., 2004; Shaner *et al*., 2004).

Together with the progression in live imaging techniques during the last years. BAC transgene technology using fluorescent reporters can be used to visualize active processes during embryonic development *in vivo*, live and on cellular level, which was never possible before (Burtscher and Lickert, 2009). The temporal and spatial expression of endoderm-specific transcription factors under the influence of signalling pathways determines endoderm development. Therefore, cell-lineage labelling using endoderm-specific factors enables a better understanding of the process regulating endoderm induction and organ formation.

The earliest endodermal progenitor cells are characterized by expression of the Forkhead transcription factor *Foxa*2. Later on, definitive endoderm express the *SRY related HMG box transcription factor Sox17*. The transcription factor *Hhex* is expressed in anterior definitive endoderm and marks the endoderm population which gives later rise to the liver. *Nkx2.1* and *Pdx1* are expressed during endoderm patterning in organ precursors of the forming lung and pancreas, respectively.

The use of fluorescent reporters under the control of these specific promoters can thus allow following cell fate determination during endodermal development. In particular, the use of BAC transgene technology, using large genomic fragments can help to overcome the problem of other transgenic technologies, as for example the lack of regulatory elements which results in an incomplete expression pattern. Moreover, BAC transgenes enable multiple copy insertions of the transgenic construct in the genome, which enhances the reporter signal and enables live imaging. Another advantage is that both alleles of the endogenous gene are maintained, contrarily to strategies as in knock-ins. Moreover, the generation of BAC transgenic ES cells opens more experimental options, instead of the direct generation of transgenic mice by microinjection. The BAC transgenic ES cells can be used for differentiation in endoderm, *in vitro* characterization studies and pre-screening of the transgenic ES clones.

From the Ensembl database (www.ensembl.org) BACs were chosen on the basis of the central position of the gene of interest and highest content of up- and downstream endogenous genomic sequence. This increased the probability that all the important regulatory sequences required for normal gene expression could be found on the same BAC (Yang *et al*., 1997; Antoch *et al*., 1997; Heintz, 2004). The technology of BAC recombination-mediated genetic engineering (recombineering; Warming *et al*., 2005) via homologous recombination in *Escherichia coli* (*E. coli*) enables the convenient modification of BACs (Copeland *et al*., 2001 and Court *et al*., 2002).

2 SCOPE OF THE WORK

The forkhead box transcription factor *Foxa2* and the SRY-related HMG box transcription factor
Sox17 are among the earliest marker genes for endoderm differentiation. These progenitor cells which form in the mouse embryo at embryonic (E) day 7.5 give over the next two days rise to the endoderm-derived organs liver, lung and pancreas.

The earliest marker gene for the liver is the haematopoetically expressed homeobox gene *Hhex*. The earliest marker gene for the lung is the NK homeodomain transcription factor 2.1 *Nkx2.1*. The earliest marker gene for pancreas is the pancreatic duodenal homeobox 1 gene *Pdx1*.
Open questions of endoderm development are:
1. Which signals and factors influence early lineage decisions in the endoderm?
2. Do the *Foxa2* and *Sox17* progenitors give rise to all endoderm derived organs or do these progenitors constitute subpopulations?
3. Which morphogenetic events lead to the formation of the organ primordia?

To answer this questions BAC transgenic fluorescent reporter lines of *Foxa2*, *Sox17*, *Hhex*, *Nkx2.1* and *Pdx1* were produced using BAC recombineering techniques.

Together with the possibilities of embryonic live imaging approaches and whole mount immunochemistry this would give new insights in the lineage decisions of the early mouse embryo and the organogenesis of endodermal organs.

3 RESULTS

3.1 Generation of BAC transgenic embryonic stem (ES) cells and mouse lines

At the beginning mini-targeting vectors that contained the fluorescent reporters were cloned (Methods 6.2.12-6.2.15; Figure 14 a). Then the BAC was transformed in bacteria. The recombineering started with the homologous recombination of the mini-targeting vector in the ATG – containing exon or in an upstream exon (Figure 14 b). The resistance cassette was removed by Flp recombinase-mediated excision (Figure 14 d). A new selection marker for subsequent ES cell screening was introduced in the BAC vector by retrofitting (Figure 14 e). Stable BAC reporter ES cell lines were produced and the endoderm-specific expression was tested by *in vitro* differentiation into endoderm. The *in vivo* expression was tested by generation of completely ES-cell derived embryos using the tetraploid complementation technique and the generation of BAC transgenic mouse lines.

The mini-targeting vectors consisted of a 5´ and 3´ homology arms corresponding to the target gene, flanking a fluorescent reporter gene (Tomato or Venus), a localization sequence to the membrane (Lyn) or the nucleus (H2B), and for selection, a *Flippase* (*Flp*)-*recombinase target* (*FRT*) flanked resistance cassette (Dymecki, 1996). They included the PGK (phosphor-glycerate kinase) promoter/EM7 (synthetic bacterial promoter)-driven *Neomycin* (*Neo*) resistance gene and an artificial intron to avoid non-sense-mediated decay (Maquat, 2002). For recombineering, the BAC was transformed in the bacterial strain EL250 which contains a defective λ prophage stably inserted into its bacterial genome (Yu *et al*., 2000; Lee *et al*., 2001).

When heat shocked, the EL250 cells express recombination proteins which mediate homologous recombination of the mini-targeting vector via the 5´and 3´ homology arms in the BAC (Figure 14 b). The next step was the removal of the *FRT*-flanked PGK-*Neo* resistance cassette, as several previous data showed that the presence of a strong promoter as PGK near a transgene can cause interference with its regulation. The *FRT*-flanked *Neo* resistance cassette was deleted by arabinose-induced *Flp* recombinase-mediated excision. To be able to later select for the

successful integration of the BAC in ES cells, a new selection cassette was needed; this was inserted in the BAC vector sequence pBACe3.6 (Frengen *et al*., 1999) located approximately 29-123 kb upstream of the reporter gene. This modification of BACs to introduce selection markers and other functional sequences is called retrofitting. We used an *in vivo* retrofitting approach with a single vector (pRetro-ES;

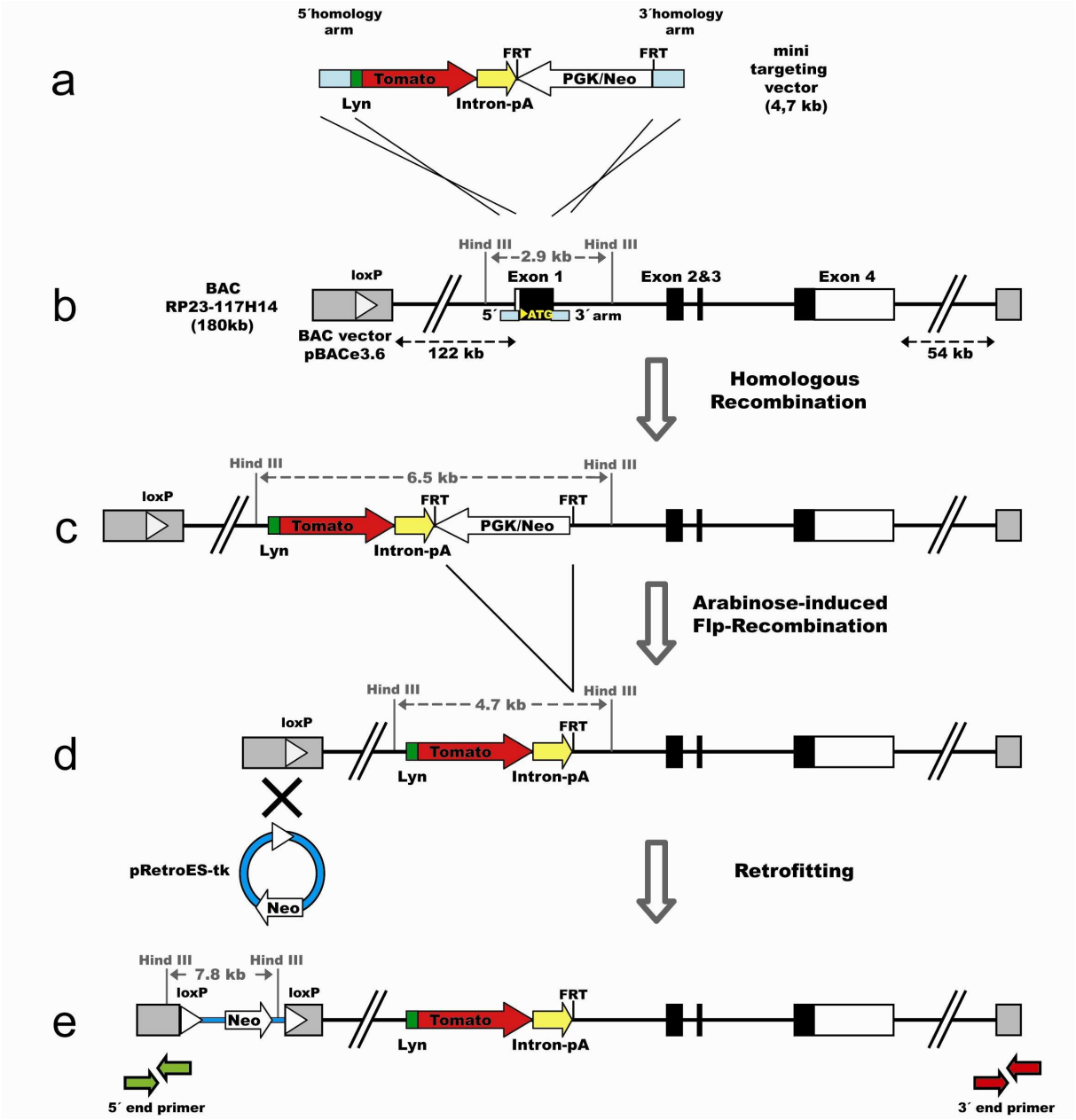


Figure 14 Method to generate BAC fluorescent reporter transgenes
(a, b) The mini-targeting vector consisting out of 5´and 3´ homology arms flanking the fluorescent gene and *FRT*-flanked selection cassette is introduced by BAC recombination into the start-ATG of *Hhex* exon1 (c, d) Using Arabinose induced Flp-mediated recombination the FRT-flanked PGK/*Neo* resistance cassette is removed in bacteria. (d, e) Insertion of a new *Neo* resistance cassette in the BAC vector pBACe3.6 via retrofitting, using the vector pRetroES-tk.

Wang *et al*., 2001) to introduce a new Neo resistance gene conferring resistance in bacteria and in ES cells in the BAC vector pBACe3.6 sequence (Frengen *et al*., 1999).

Furthermore, to enhance selection for multiple copy integration of the BAC transgene in ES cells, the stronger phosphor-glycerate kinase (PGK) promoter of the pRetroES was exchanged with the weaker thymidin kinase (TK) promoter and then pRetroES was named pRetroES-tk. The vector pRetroES-tk contains a promoter which drives the expression of a *causes recombination* (*Cre*) recombinase (Sternberg and Hamilton, 1981) and a *locus of X* over Pi (*lox*P; Hoess *et al*., 1982) site, other than the already mentioned *Neo* resistance gene.

After transformation, the *Cre* recombinase promotes the recombination between the *lox*P sites in pRetroES-tk and in the BAC vector pBACe3.6, therefore leading to the integration of pRetroES-tk into the BAC. After integration the promoter is separated from the fusion gene and stops the production of *Cre* recombinase. DNA fingerprinting was used, to follow the steps of bacterial recombineering, to check for successful recombineering and to test the integrity of the BAC after each modification step. The DNA fingerprinting profile was performed after: 1. the transformation of the wild-type (*wt*) BAC in EL250 cells; 2. the homologous recombination of the mini-targeting vectors in the BACs; 3. the removal of the Neo resistance cassette; 4. the integration of the pRetroES-tk vector in the retrofitting process. At least three different restriction enzyme digestions were used after each modification step to check and verify the BAC integrity using the wt BAC as control. DNA preparations were done from BAC clones, digested overnight and visualized on agarose gel (details are given in the Methods section).

3.2 Generation of a *Hhex::Lyn-Tomato* BAC transgene

The homeobox gene *Hhex* is expressed at blastocyst stage in the PrE and plays a pivotal role in the establishment of the A-P axis prior to gastrulation, when *Hhex* is expressed in the DVE and the forming AVE. During gastrulation *Hhex* is then expressed in the AVE and the DE. *Hhex* is essential for the formation of the liver and can be found in the earliest progenitors of the developing liver as well as in the adult organ. The lineage labelling of *Hhex* gene allows the analysis of its temporal expression from very early embryo stages on, its expression in the endoderm germ

layer and its important role during the formation of endodermal organs. The lineage labelling allows insights over various stages endoderm development until the adult organism. The selected BAC RPCI-23-117H14 (BAC library RPCI-23; Osoegawa *et al*., 2000) had a size of 180 kb, contained the *Hhex* open reading frame (ORF) and contained also the ORF of 4 other genes. The *exocyst complex component 6* (*Exoc6*; Scheufler, 1969), is involved in protein transport, and the *kinesin family member 11*

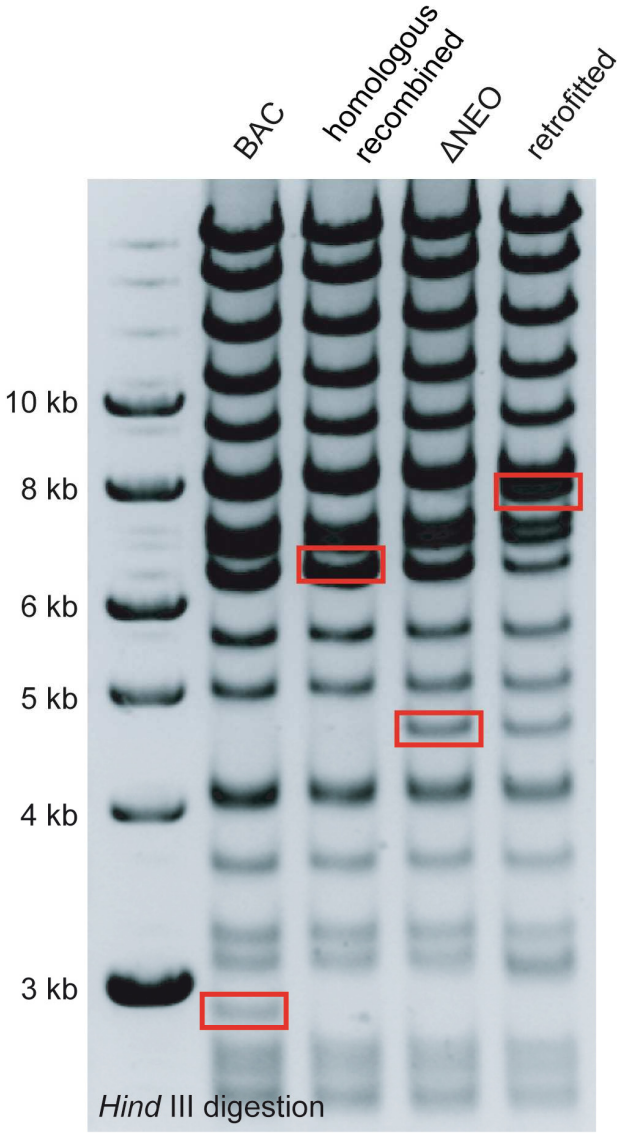


Figure 15 Generation and verification of a *Hhex::Lyn-Tomato* reporter transgene
Agarose gel of *Hind* III digested *Hhex* BAC DNA from colonies of the unmodified BAC (BAC), the BAC after homolog recombination of the mini-targeting construct (homologous recombined), the BAC after deletion of the *Neo* cassette by Flp recombination (ΔNEO) and the BAC after the retrofitting process using pRetroES-tk (retrofitted). The continuous restriction pattern shows the BAC integrity while the restriction fragment length polymorphism of some fragments (RFLP; indicated by red boxes) shows the BAC modification steps (see also Figure 14).

(*Kif11*; Winter *et al*., 1994), which is involved in microtubule motor activity. The other two genes are the *insulin-degrading enzyme* (*Ide*; Mirra *et al*., 1991), a metalloprotease which is involved in proteolyse; and the *membrane associated ring finger 5* (*March5*) which is involved in ubiquitin cycle. Using the strategy outlined above, the *Hhex* BAC was modified (Figure 14).

The DNA fingerprinting gel of the *Hhex* BAC, recombined with the Lyn-Tomato mini-targeting vector is shown in Figure 15. The maintained restriction pattern demonstrated the preserved integrity of the BAC, while the restriction fragment length polymorphism of some fragments (RFLP; indicated by red boxes) indicated the BAC modification steps (see also Figure 14). The *Hind* III digestion showed RFLP from 2.9 kb to 6.5 kb after integration of the mini-targeting vector in the BAC by homologous recombination (Figure 15). The deletion of the Neo resistance cassette caused a RFLP from 6.5 kb to 4.7 kb ("ΔNEO"). The integration of the new selection marker by retrofitting created a new restriction fragment of 7.8 kb length. The results of the RFLP confirmed the alterations which were expected from *in silico* data of the fingerprinting. The used steps of BAC recombineering in bacteria appeared as very efficient, as indicated by the high number of bacterial colonies. About 95% (n=20) of the tested *Hhex* BAC clones showed a maintained integrity during the recombineering steps.

The efficiency of the retrofitting step was 20% (n=30). Positive retrofitted bacterial colonies were identified by PCR screening (Yang and Seed, 2003; Methods 6.2.16). The modification of the *Hhex* BAC was successful and the single steps could be followed precisely by RFLP.

3.2.1 Generation of *Hhex* BAC transgenic reporter ES cells

BAC transgenic mice can be generated by pronuclear injection of BAC DNA or from BAC transgenic ES cells by diploid complementation technique. Pronuclear injections are laborious, expensive and technically demanding. On the other hand the generation of BAC transgenic ES cells has some advantages. They can be used for *in vitro* studies in culture by differentiation in endoderm, and they can be checked for transgene expression. The transgenic clones can be pre-selected for multi-copy insertions in the genome by Southern analysis and for complete integration of the BAC in the genome by PCR. Additionally, the *in vivo* expression can be tested for

accurate reporter expression in completely ES cell-derived embryos before generating a mouse line. *Hhex::Lyn-tomato* transgenic BAC clones were produced from *Hhex* BAC DNA, which was linearized with the PI-*Sce*I enzyme (Christ *et al*., 1999) that cleaves a unique site in the BAC vector backbone pBACe3.6. The linearized *Hhex* BAC DNA was electroporated in IDG3.2 ES cells (Hitz *et al*., 2007). *Neomycin* resistant ES cell colonies were tested by Southern analysis using a probe which detected the fluorescent protein Tomato. DNA from a clone with single copy integration was used to determine the copy numbers of the integrated BAC.

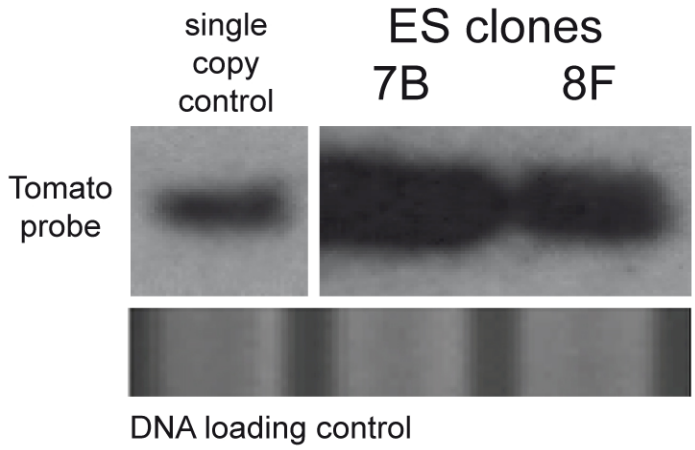


Figure 16 Southern analysis of individual Hhex BAC transgenic ES clones showing different numbers of transgene integrations
The used Southern probe was Tomato (size 6.7 kb) for Hhex (4.7 kb; see Figure 14). The BAC DNA was digested with *Hind* III. DNA from single copy Tomato knock-in ES cell integration was used as control. The DNA amount is shown in the loading control to estimate the number of integrated BACs in the genome. The clone 7B shows integration of 8, clone 8F of 6 BACs in the genome. Not shown are the Southern positive clones Hhex 4C, Hhex G2 and Hhex 1A which show single copy insertions.

Aim was, to identify clones with multiple insertions which would later give an amplification of the fluorescent signal above the background level for live imaging applications. One electroporation in ES cells using 30 µg linearized BAC DNA resulted in 19 ES cell clones. The Southern analysis showed 5 out of 19 positive clones for the Tomato probe, which corresponded to a ratio of 26 % positive ES cell clones. The number of integrated BACs in the genome was estimated by comparison between the signal strength of the single copy control with that of the positive clones, after normalization of the amount of loaded DNA (DNA loading control). All BAC clones showed a single integration in the genome, except the Hhex clones 7B and 8F. The number of integrated Hhex BACs into genome was estimated as 8 for the

Hhex clone 7B and about 6 for the Hhex clone 8F (Figure 16). Overall, the percentage of Southern positive clones obtained with our strategy was very high.

3.2.2 Integrity check of BAC transgenic ES cells by BAC end PCR

Although the Southern blot analysis is a convincing method to demonstrate the genomic integration of the fluorescent Tomato transgene, still it does not demonstrate the integration of the complete BAC sequence in the genome. The advantage of BACs to contain all regulatory elements of a gene requires the complete insertion of the BAC. Therefore, to show the complete integration of the intact BAC, a PCR check of the BAC transgenic ES clones was done.

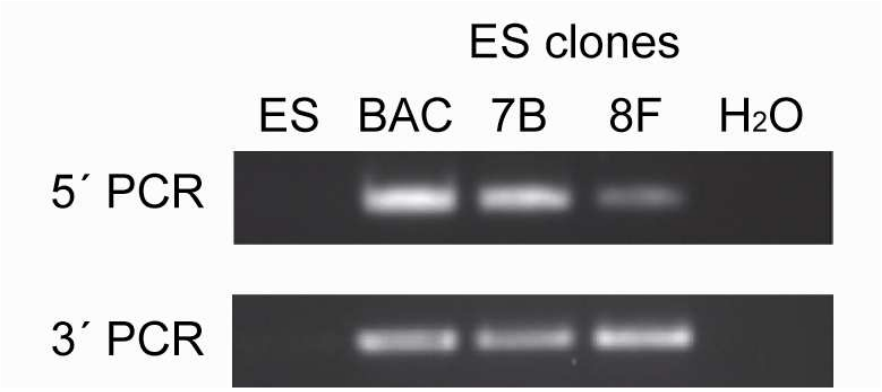


Figure 17 BAC-end PCR of Hhex BAC transgenic ES cells demonstrates the complete integration of the BAC in the genome
BAC clones of Hhex were tested for complete integration using genomic ES DNA as template for the PCR on the 5´ and the 3´ BAC end. The PCR shows the complete integration of the Hhex clones 7B and 8F. ES DNA from unmodified IDG3.2 cells was used as negative control (ES) while BAC DNA of the corresponding BAC was used as positive control (BAC).

The linearization of the BAC at the PI-*Sce*I site in the BAC vector backbone pBACe3.6 before the electroporation in the ES cells generates two “BAC ends” named “5’” and “3’”, which could be independently verified by PCR on the genomic DNA (Figure 14.2 e). PCR primer pairs on both sides were designed to demonstrate the integration of the complete BAC as described by Yang and Seed (2003).
The *Hhex* BAC transgenic clones 7B and 8F were selected for the PCR check of complete integration as they showed multi-copy insertions of the BAC transgene in the genome by Southern blot (Figure 16). Both tested transgenic clones showed amplification products for the 5´and the 3´ BAC-end (Figure 17). This demonstrated the complete integration of the BAC in the genome of clones 7B and 8F. After the verification of the integration of the Tomato reporter by Southern analysis and

demonstration of complete insertion of the BAC by BAC-end PCR, ES clones were further characterized by *in vivo* analysis.

3.2.3 *In vitro* expression of Hhex BAC transgenic ES cells in ES cells and *in vitro* differentiated ES cells into endoderm

Recently it was shown that ES cells can be efficiently differentiated in endodermal cells in culture (Tada *et al*., 2005; Kubo *et al*., 2004). This offered a good opportunity to test the BAC transgenic ES cells in *in vitro* culture for reporter gene expression. The procedure allowed the direct selection of positive functional clones for BAC transgene expression before continuing with the time consuming test of *in vivo* expression. The *in vitro* differentiation allows the analysis in the BAC transgenic ES cells of the endoderm progenitor *Foxa2*, the endoderm precursor *Sox17* and of *Hhex* which are expressed at early stages of endoderm development.

Transgene expression could be observed in undifferentiated ES cells of the *Hhex::Lyn-Tomato* clones 7B and 8F (Figure 18). The localisation of the expression was detectable in the membrane (Figure 18 a´), as expected from a transgene driven by the N-terminal sequence of the Lyn kinase gene. The *Hhex::Lyn-Tomato* transgenic BAC clones 7B and 8F were also differentiated in endoderm. This enabled later the comparison of the transgenic clones of *Hhex*, *Foxa2* and *Sox17* in terms of signal localization. The Hhex clones 7B and 8F showed already transgene expression at day 0 of differentiation (in the undifferentiated ES cells).

The expression level, showed a peak of intensity on day 4 (Figure 19 e, f) and a following slight decrease of intensity until day 8 of differentiation. The transgene expression was more spread over the cells from day 5 to day 8 and the cell-walls are hardly visible. At these time point cell morphology appeared to be irregular due to the formation of both multiple layers of differentiating ES cells and three dimensional structures. Both *Hhex* BAC transgenic ES clones showed transgene expression in undifferentiated as in *in vitro* differentiated ES cells.

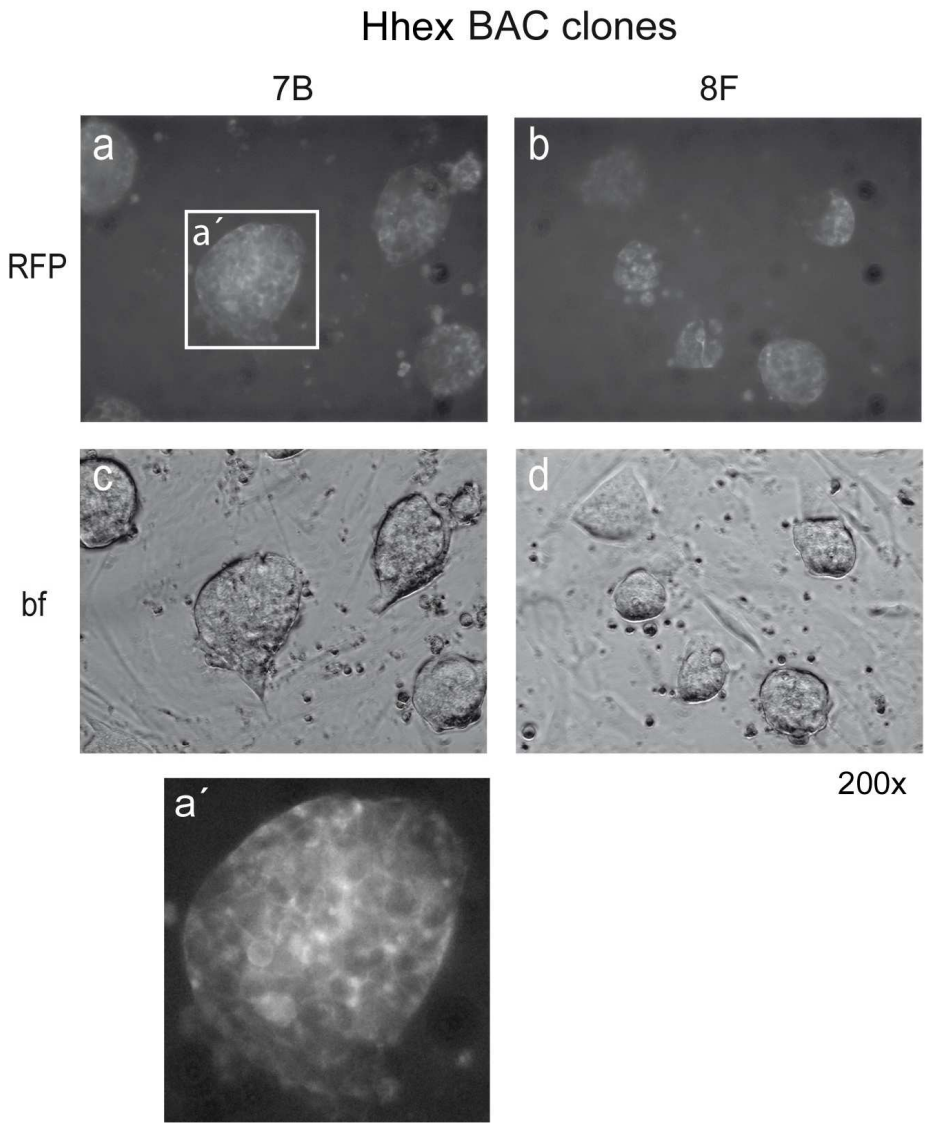


Figure 18 *Hhex::Lyn-Tomato* transgene expression in undifferentiated ES cells
Transgene expression of the Tomato reporter in Hhex BAC transgenic ES cell clones 7B (a, c) and 8F (b, d). The expression is localized to the cell membrane (a´; higher magnification) as expected from the transgene driven by promoters of the Lyn kinase.

3.2.4 *In vivo* analysis of *Hhex::Lyn-Tomato* BAC transgenic ES cells by tetraploid complementation technique

To analyze the *in vivo* expression of the Hhex BAC transgenic ES clones which were successfully tested *in vitro*, completely ES cell-derived embryos were generated using tetraploid complementation technique (Nagy *et al*., 1993). Tetraploid embryos (4n) were aggregated with small clumps of 6-8 BAC transgenic ES cells and transferred in the uterus of E2.5 pseudopregnant recipient female mice. In the tetraploid chimeras the BAC transgenic ES cells can only contribute to the

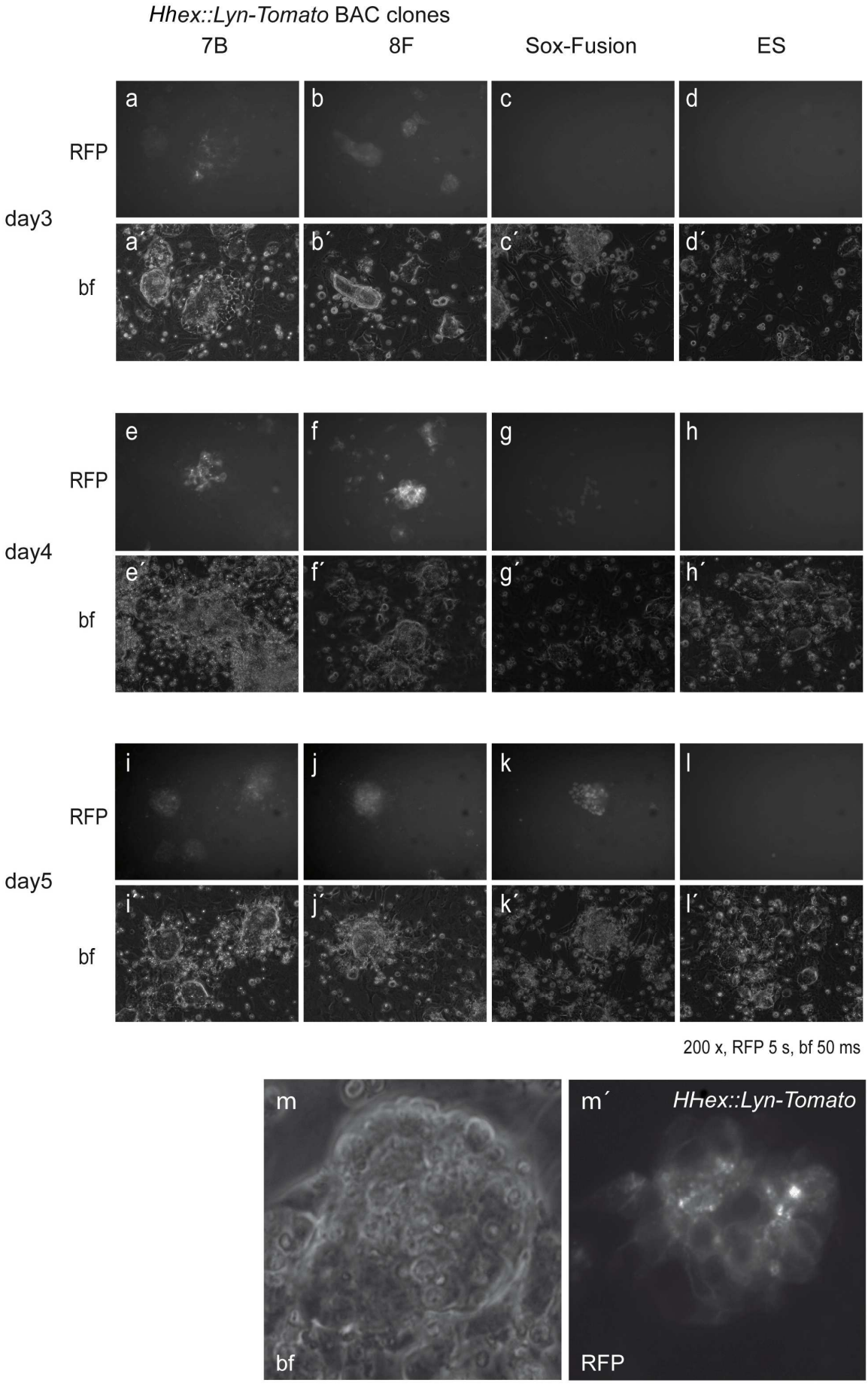

Hhex::Lyn-Tomato BAC clones
7B
8F
Sox-Fusion
ES
day3
RFP
bf
a
b
c
d
a´
b´
c´
d´
day4
RFP
bf
e
f
g
h
e´
f´
g´
h´
day5
RFP
bf
i
j
k
l
i´
j´
k´
l´
200 x, RFP 5 s, bf 50 ms
m
bf
m´
Hhex::Lyn-Tomato
RFP

Figure 19 *Hhex::Lyn-Tomato* transgene expression during ES cell *in vitro* differentiation in endoderm
Transgene expression of the Tomato reporter in Hhex BAC transgenic ES cell clones 7B (a, a´, e, e´, i, i´) and 8F (b, b´, f, f´, j, j´) from day 3 to day 5 of in vitro differentiation in endoderm. The expression is localized to the cell membrane (m, m´; clone 8F at day4 of differentiation in higher magnification). ES cells from a Sox-Fusion (knock-in) were used as positive control for the in vitro differentiation. Wild-type ES cells (ES) were used as negative control.

embryonic lineages (e.g. definitive endoderm) whereas the cells of the 4n embryo contribute only to the extraembryonic lineages (e.g. visceral endoderm). This method allowed testing accurately several BAC clones for *in vivo* expression in the endoderm. *Hhex::Lyn-Tomato* BAC transgenic ES clones 7B and 8F were aggregated with tetraploid YVI embryos (Hadjantonakis *et al*., 2002) which express ubiquitously the yellow fluorescent protein, thus enabling to distinguish between the Tomato expressing transgenic ES cells (red) and the 4n ES cells (green). Transgene expression in the DE could be observed in both clones (as an example for both, clone 7B is shown in Figure 20).

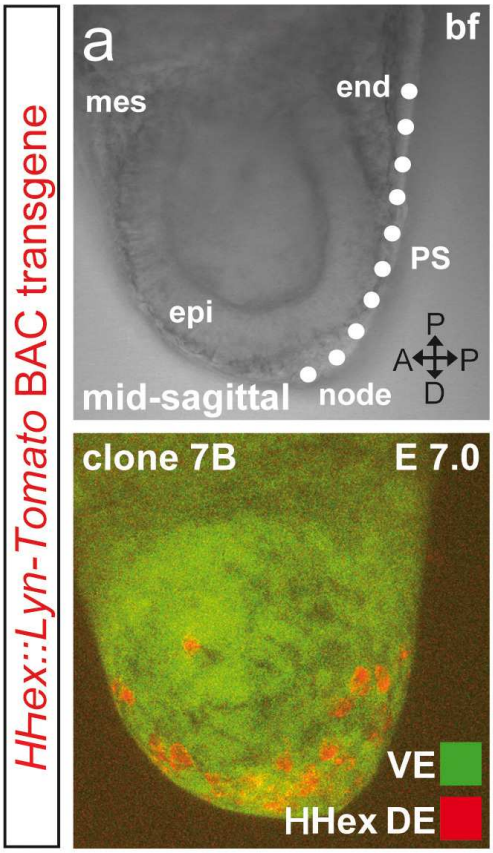


Figure 20 *Hhex* BAC reporter transgene expression *in vivo* in tetraploid-derived embryos
(a, b) *In vivo* expression of the reporter transgene in the definitive endoderm (DE; red) appearing in the visceral endoderm (VE; green) of a *Hhex::Lyn-Tomato* BAC transgenic ES cell-derived embryo at E7.0 (Hhex BAC transgenic ES cells <-> 4n YFP embryo). Abbreviations: Mesoderm (mes), epiblast (epi), primitive streak (PS), endoderm (end).

Hhex::Lyn-Tomato transgenic DE cells emerged in the VE epithelium at the anterior end of the primitive streak (PS), in the node region and along the midline of analyzed embryos between E7.0 and E7.5, consistent with *Hhex* mRNA expression (Thomas *et al*., 1998).

The *in vivo* test of BAC transgenic ES cells by tetraploid aggregation technique showed a specific transgene expression in 2 of 2 tested clones. Taken together, the complete pre-screening scheme for BAC transgenic ES cells allowed selection for clones that expressed the transgene accurately and at levels which are high enough for endoderm lineage labelling in live imaging approaches.

3.2.5 Expression analysis of the BAC transgenic *Hhex::Lyn-Tomato* mouse line

To study *Hhex* gene expression and lineage differentiation in the developing mouse embryo, *Hhex::Lyn-Tomato* transgenic mouse lines were generated from the pre-selected and tested ES cell clones (as described in Results 3.2 - 3.2.4). Namely, BAC transgenic ES cells of clone 7B and 8F were used to generate chimeras by diploid aggregation technique. Then, these were crossed to CD1 outbread mice for germline transmission of the transgene and generation of a *Hhex::Lyn-Tomato* transgenic mouse line.

The following results were obtained from the mouse line generated from *Hhex::Lyn-Tomato* ES clone 7B. At blastocyst stage between E3.5 and E4.5, the PrE forms by sorting out from all progenitor cells of the ICM, those of the PrE; at E4.5 the latter forms an epithelial layer lining the top of the ICM. With the help of a confocal fluorescent microscope, the first *Hhex::Lyn-Tomato* transgene expression was detected at blastocyst stage E3.5 - E3.75 in few cells of the ICM (Figure 21 a, a´). Therefore, the Lyn-Tomato fluorescent signal detected at E3.5 – E3.75 might correspond to PrE progenitors in the ICM.

The earliest reported detection of *Hhex* mRNA expression was found in the PrE at E4.5 (Thomas *et al*., 1998; Mesnard *et al*., 2006). During pre-gastrulation stage (E5.5) the earliest transgene expression was detected in the DVE on the distal tip of the embryo, as reported, previously (Thomas *et al*., 1998; Rodriguez *et al*.; 2001). During gastrulation stage, transgene expression

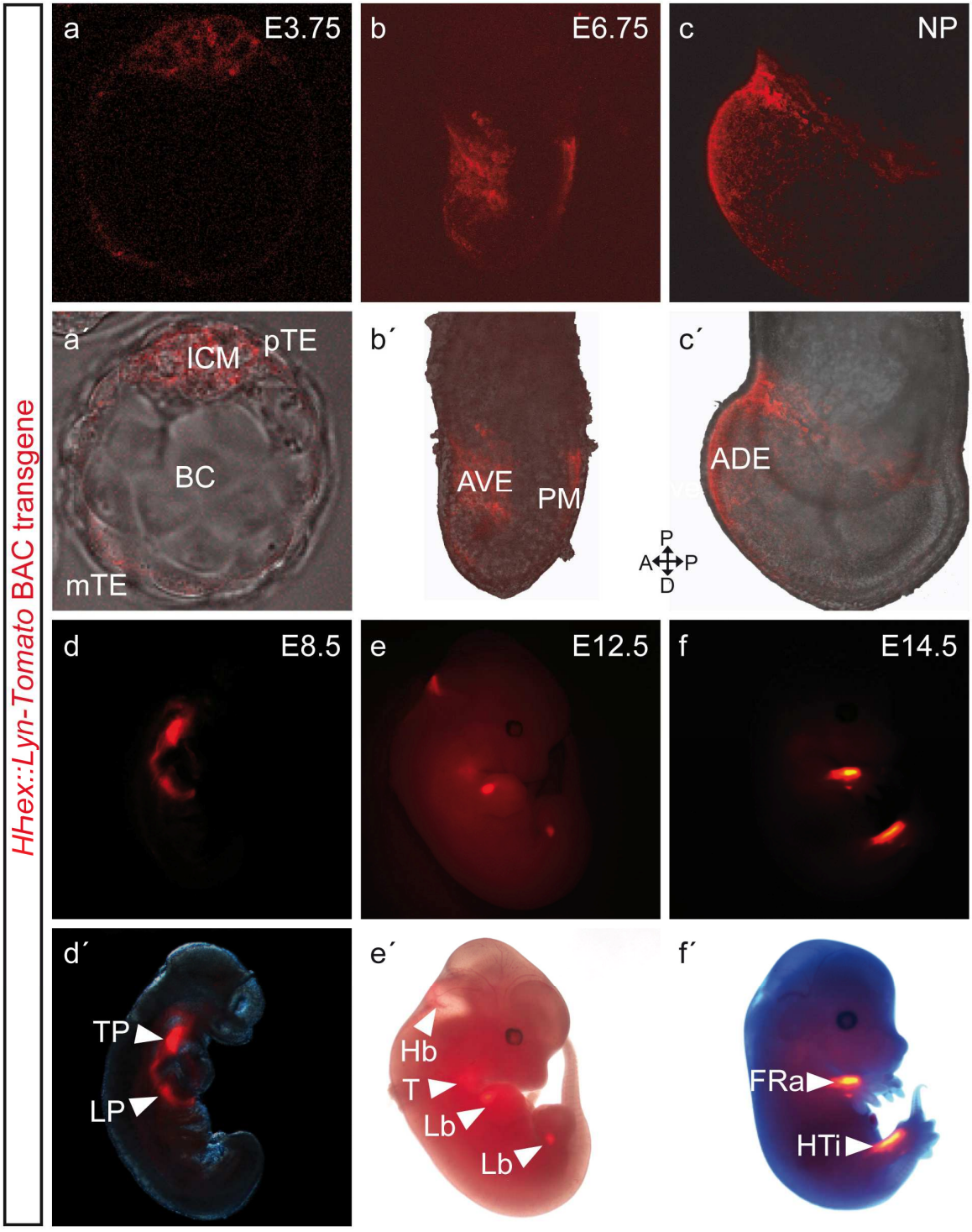


Figure 21 Analysis of *Hhex::Lyn-Tomato* transgene expression in mouse embryos
(a, a´) *Hhex::Lyn-Tomato* expression in the inner cell mass (ICM) at blastocyst stage. (b, b´) At E6.75 (early/mid-streak stage), *Hhex::Lyn-Tomato* is expressed in the anterior visceral endoderm (AVE) and posterior mesoderm (PM). (c, c´) At neural-plate stage, *Hhex::Lyn-Tomato* is localized to the anterior definitive endoderm (ADE). (d, d´) At E8.5 fluorescent signal is detected in the thyroid primordium (TP) and the liver primordium (LP). (e, e´) At E12.5, expression is restricted to the hindbrain (Hb), Thymus (T) and the fore- and hindlimb (Lb) (f, f´) At E14.5 expression in the radius of the forelimb (FRa) and the tibia of the hindlimb (HTi). Image acquisition was done using laser scanning confocal imaging until stage E7.75 (NP) and by fluorescence inverse microscopy of later stage embryos. Abbreviations: Polar trophectoderm (pTE), blastocoel cavity (BC), mural trophectoderm (mTE).

could be detected at E6.75 (pre-streak stage) in the AVE and the posterior mesoderm (PM) of the embryo (Figure 21 b, b´); at E7.25 (early/mid-streak stage) the transgene was still expressed in the visceral endoderm and in the posterior mesoderm. At neural plate (NP) stage, the transgene expression was detected in the ADE (Figure 21 c, c´) as shown before from Thomas *et al*. (1998) and Rodriguez *et al*. (2001).

At E8.5 when the gut tube closes and the foregut pocket forms, the *Hhex::Lyn-Tomato* reporter was detected in the region of the forming liver and thymus (Figure 21 d, d´). A strong transgene activity could be detected at E10.0 and E12.5 in the thymus (Figure 21 e, e´). At E12.5 transgene expression was also evident in the rostral hindbrain region (Hb) and in the fore- and hindlimbs (Lb, Figure 21 e, e´). In whole-mount imaging at stage E 14.5 and E16.5 a strong transgene expression was detectable in specific regions of the fore- and hindlimbs; these clearly corresponded to radius of the forelimb and tibia of the hindlimb and these signals were maintained until stage E18.5 (Figure 21 f, f´ and Figure 22 a-b´).

These results indicated that Hhex::Lyn-Tomato expression could correlate with limb bud patterning and radius and tibia bone formation. Whole-mount organ preparations from E16.5 and E18.5 showed transgene expression in the liver and pancreas (Figure 22. c-f´; as shown by Bogue *et al*., 2000), but no expression was detectable in lung, kidney, heart, gut, thymus, stomach and thyroids.

The analysis of the generated *Hhex::Lyn-Tomato* mouse line showed that transgene expression accurately reflected the endogenous mRNA expression patterns. Thus, enabling the study of *Hhex* regulation and lineage allocation in ES cell culture and mouse embryos. It was not possible to generate a homozygous *Hhex::Lyn-Tomato* mouse line by crossing heterozygote animals. To this end, 9 matings with an offspring of 96 heterozygous pups were examined and tested for transgene expression.

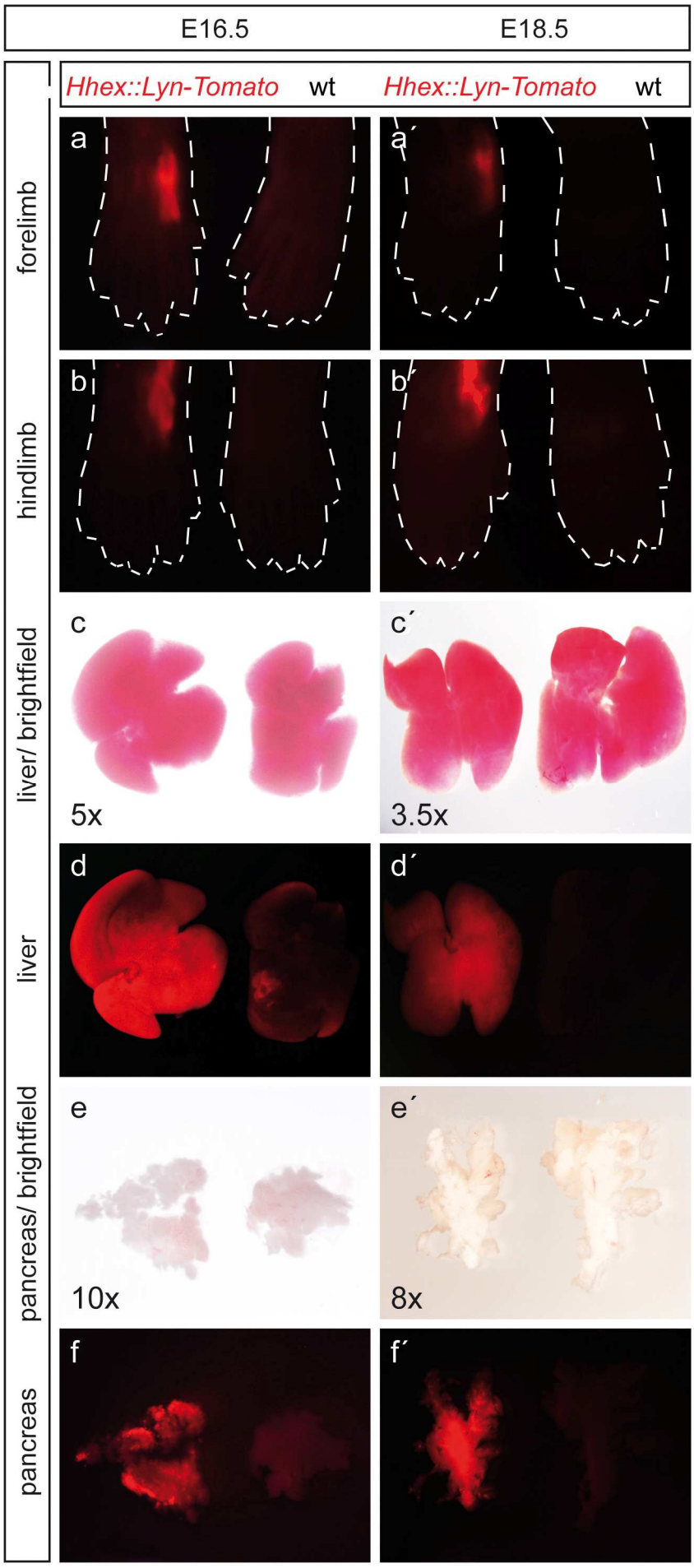

E16.5
E18.5
Hhex::Lyn-Tomato
wt
Hhex::Lyn-Tomato
wt
forelimb
hindlimb
liver/ brightfield
liver
pancreas/ brightfield
pancreas
a
a´
b
b´
c
c´
5x
3.5x
d
d´
e
e´
10x
8x
f
f´

Figure 22 Analysis of *Hhex::Lyn-Tomato* transgene expression in limbs and different organs at E16.5 and E18.5
(a - b´) Transgene expression in the radius of forelimb E16.5 to E18.5 and the tibia of the hindlimb is detected in *Hhex::Lyn-Tomato* but not in wt embryos. (c-f´) *Hhex* transgene expression is detected in the liver and pancreas at E16.5 (c, d, e, f) and at E18.5 (c´,d´,e´,f´). Transgene expression in the pancreas at E16.5 and E18.5. Brightfield and corresponding *Hhex::Lyn-Tomato* is shown in c-f´. Image acquisition was done using fluorescence inverse microscopy.

3.3 Comparison of *Hhex::Lyn-Tomato* BAC transgenic mouse line with a Hhex-GFP transgene mouse line

To enable a direct comparison between the *Hhex* expression pattern in our newly generated mouse line and other existing *Hhex*-reporter lines, *Hhex::Lyn-Tomato* BAC transgenic mice were crossed to a *Hhex-GFP* mouse line (Rodriguez *et al*., 2001). Double heterozygous embryos of these intercrosses were analyzed by live imaging at gastrulation stage, where the expression patterns of the transgene started showing a different expression pattern. In general, the BAC *Hhex::Lyn-Tomato* transgene showed a more homogenous pattern in comparison with the *Hhex-GFP* transgene;

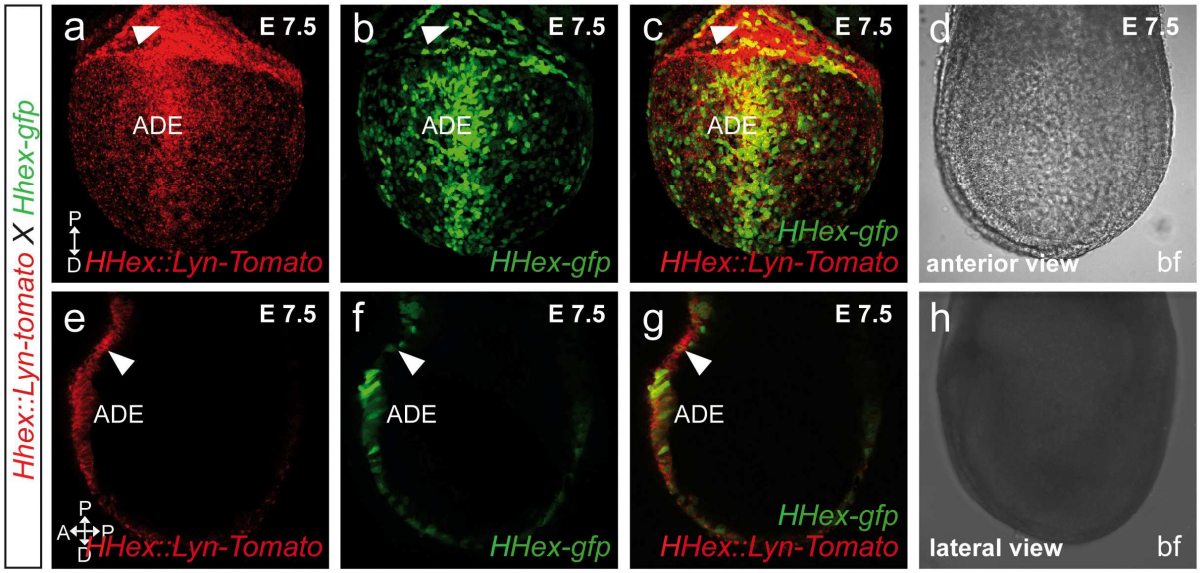


Figure 23 *Hhex::Lyn-Tomato* BAC transgene expression and *Hhex-GFP* transgene expression in the anterior definitive endoderm (ADE) at stage E7.5
(a) Transgene expression of *Hhex::Lyn-Tomato* at E7.5 in the anterior definitive endoderm (ADE; merge of planes). (b) *Hhex-GFP* transgene expression at E7.5 in the ADE (merge of planes). (c) Overlay of *Hhex::Lyn-Tomato* and *Hhex-GFP* transgene expression pattern (merge of planes). (d) Brightfield picture of the in a-c shown embryo at E7.5 (single plane). (e) *Hhex::Lyn-tomato* transgene expression in the ADE and the anterior extraembryonic VE (lateral view, single plane, mid-sagittal section). (f) *Hhex-GFP* transgene expression in the ADE and the anterior proximal region. (g) Overlay of *Hhex:: Lyn-Tomato* and *Hhex-GFP* transgene expression pattern at E7.5 in mid-sagittal, lateral view (single plane) (h) Brightfield picture of the in e-g shown embryo at E7.5 (merge of planes). Anterior definitive endoderm (ADE), anterior extraembryonic VE = white arrowhead.

in fact the *GFP* fluorescent signal was more “patchy” and presented “holes” in the expression, as for example in the ADE and the anterior extraembryonic VE; (see Figure 23 a-c; white arrowhead). These differences in transgene expression were more prominent in the anterior proximal region of the embryo, where the BAC transgene showed the strongest expression domain, while the *Hhex-GFP* transgene was only expressed in some parts of the region. These differences appeared even more obvious in the overlay of both signals (Figure 23 g, white arrowhead): in this lateral view of an embryo at stage E7.5 single plane confocal sections highlighted the homogenous expression pattern of the *Hhex::Lyn-Tomato* transgene and the spotty and patchy pattern of the *Hhex-GFP* transgene (Figure 23 g).

The results of this comparison are a confirmation of the advantage of BAC transgenes as they contain most or all regulatory elements of a gene, as shown by Heintz (2004) with the generation of 250 BAC transgenic mouse lines. In fact, the “spotty” and “patchy” expression pattern of the *Hhex-GFP* transgene could be due to the lack of regulator elements, which are instead contained in the *Hhex::Lyn-Tomato* BAC transgenic line.

3.4 Generation of a *Sox17::H2B-Tomato* BAC transgenic mouse line

Sox17 belongs to the group of Sry-related HMG-box transcription factors and is expressed at early stages of endoderm development. It is expressed at E6.0 in the visceral endoderm (Kanai *et al*., 1996) and from E7.0 it can be detected in the anterior definitive endoderm (ADE) and the hindgut, where it is essential for the formation of the endodermal hindgut (Kanai-Azuma, 2002). *Sox17* is, together with *Foxa2*, one of the first markers of the endoderm germ layer and plays a pivotal role in early endoderm formation (Kanai *et al*., 1996).

At later stages of development *Sox17* influences early vasculogenesis. The study of gene expression and lineage differentiation of *Sox17* became an important issue for its role as key factor in formation of the endoderm germ layer. The selected BAC RPCI-23-285G23 has a size of 179 kb, and contains only the *Sox17* ORF. We used the same strategy as outlined above (Figure 14) to introduce H2B-Tomato (with nuclear restricted expression) under the control of the *Sox17* regulatory elements.

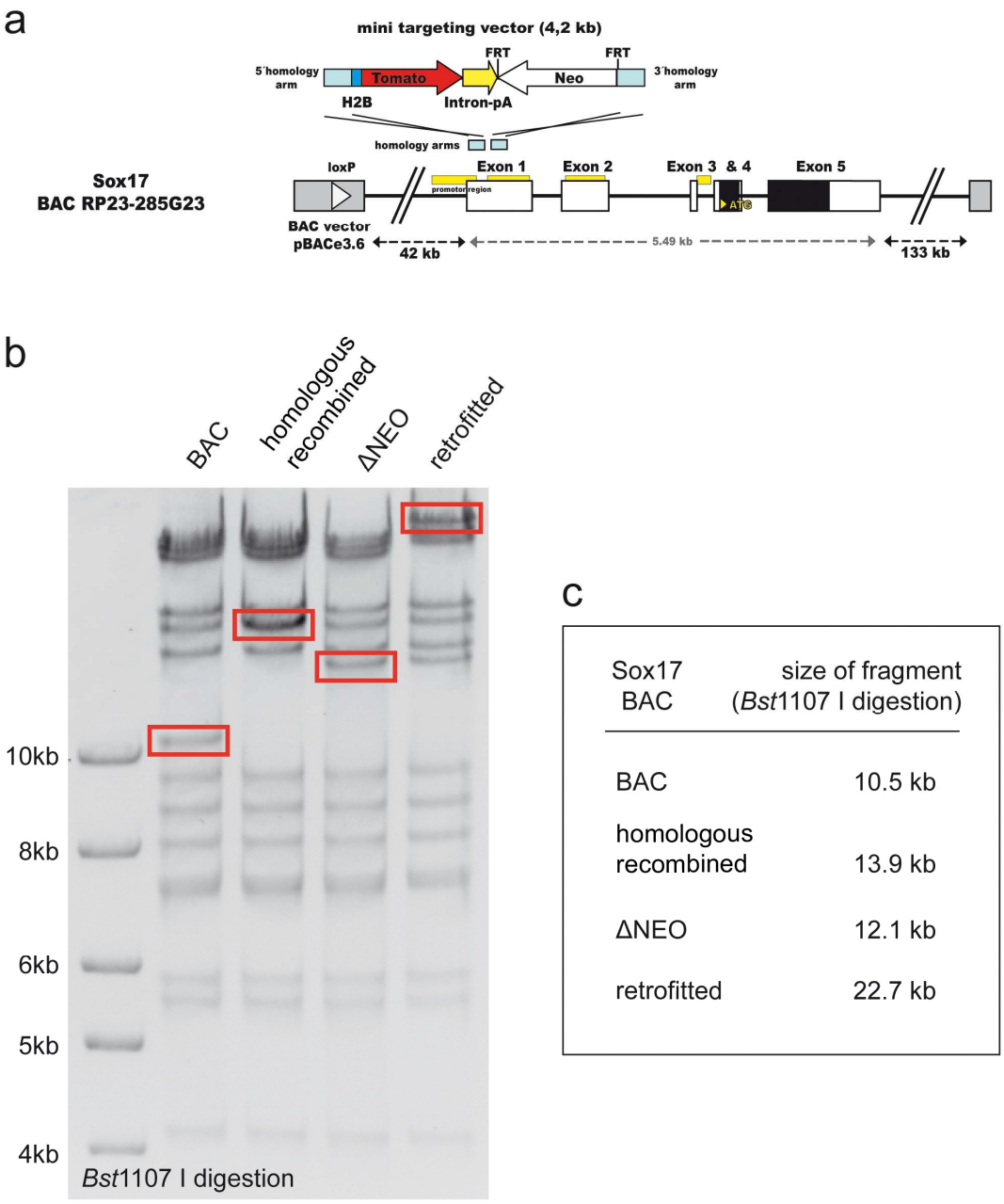


Figure 24 Generation and verification of a *Sox17::H2B-Tomato* reporter transgene
(a) Mini targeting vector containing 5´and 3´ homology arm, H2B-Tomato reporter gene, Intron-polyA and FRT-flanked *Neo* resistance gene. This vector was used to knock-in the reported gene into exon1 of *Sox17* as indicated. (b) Agarose gel of *Bst*1107 I digested *Sox17* BAC DNA isolated from colonies of the unmodified BAC (BAC), the BAC after homologous recombination of the mini-targeting construct (homologous recombined), the BAC after deletion of the *Neo* cassette by Flp recombination (ΔNEO) and the BAC after the retrofitting process using pRetroES-tk (retrofitted). The continuous restriction pattern shows the BAC integrity while the restriction fragment length polymorphism (RFLP; indicated by red boxes) shows the BAC modification steps. (c) The box shows the indicated RFLP fragment sizes of the modification steps.

Therefore the mini-targeting sequence was inserted in *Sox17* Exon1.The DNA fingerprinting gel of the *Sox17* BAC, recombined with the H2B-Tomato mini-targeting vector is shown in Figure 24.The maintained restriction pattern demonstrated the

preserved integrity of the *Sox17* BAC, while the restriction fragment length polymorphism (RFLP; Figure 24 b) indicated the BAC modification steps. The *Bst1107* I digestion showed RFLP from 10.5 kb to 13.9 kb after integration of the mini-targeting vector in the BAC by homologous recombination ("homologous recombined" in Figure 24 a-c).The deletion of the *Neo* resistance cassette caused a RFLP from 13.9 kb to 12.1 kb ("ΔNEO"). The integration of the new selection marker by retrofitting creates a new restriction fragment of 22.7 kb length ("retrofitted") (Figure 24 c).The results of the RFLP confirmed the alterations which were expected from *in silico* data of the fingerprinting (Figure 24 c). The used steps of BAC recombineering in bacteria appeared very efficient as indicated by the high number of growing bacteria colonies. About 90% (n=20) of the tested *Sox17* BAC clones showed maintained integrity during the recombineering steps. The efficiency of the retrofitting step was 17% (n=30). Positive retrofitted bacterial colonies were identified by PCR screening (Yang and Seed, 2003; Methods 6.2.16). The modification of the *Sox17* BAC was successful and the single steps could be followed precisely by RFLP.

3.4.1 Generation of *Sox17* BAC transgenic reporter ES cells

Sox17 BAC transgenic ES cells were produced as described in Results 3.2.1. for the Hhex transgene (Results 3.2.1). The linearized *Sox17* BAC DNA was electroporated in IDG3.2 ES cells (Hitz *et al*., 2007). Neomycin resistant ES cell colonies were tested by Southern analysis using a probe which detected the fluorescent protein Tomato. One electroporation in ES cells using 30 µg linearized BAC DNA resulted in 10 ES cell clones. The Southern analysis showed 2 out of 10 positive clones for the Tomato probe which corresponded to a ratio of 20 %.

The number of integrated BACs in the genome was estimated by comparison between the signal strength of the single copy control with that of the positive clones, after normalization of the amount of loaded DNA (DNA loading control). The Southern positive *Sox17* BAC transgenic clones 4B and F3 showed a single integration in the genome (Figure 25). The number of generated clones was very low but the percentage of Southern positive clones obtained with our strategy was very high as compared to the rate of Southern positive clones generated with classical knock-in strategies.

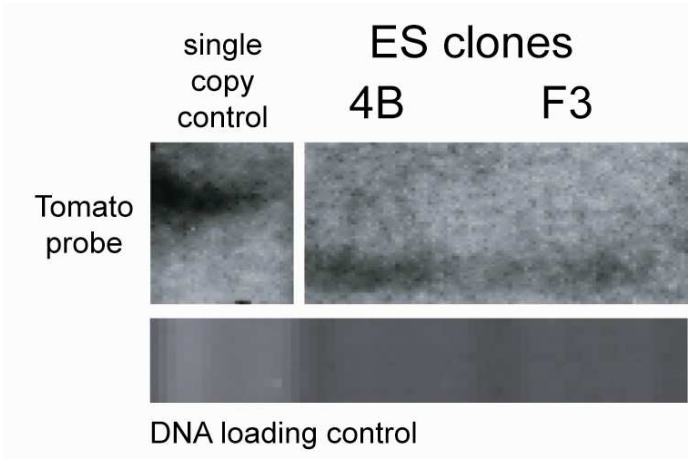


Figure 25 Southern analysis of individual *Sox17* BAC transgenic ES clones showing integration of one transgene in the genome
The used Southern probe was Tomato (size 6.7 kb) for Sox17 (4.3 kb). The BAC DNA was digested with *Hind* III. DNA from single copy integration was used as control. The DNA amount is shown in the loading control to estimate the number of integrated BACs in the genome. The clones 4B and F3 show integration of one BAC in the genome.

The fact that just 2 out of 10 clones were positive in the Southern analysis indicated that the ratio of complete integrated BACs in the genome was very low. This was very likely due to damages of the BAC in the DNA preparation process.

3.4.2 Integrity check of *Sox17* BAC transgenic ES cells by BAC end PCR

A PCR check of the *Sox17* BAC transgenic ES clones was done, to show the complete integration of the intact BAC. The linearization of the BAC at the PI-*Sce*I site in the BAC vector backbone pBACe3.6 before the electroporation in the ES cells

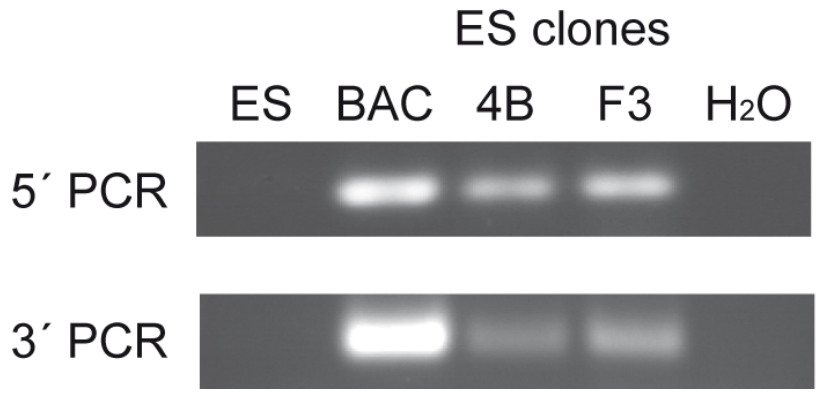


Figure 26 BAC-end PCR of *Sox17* BAC transgenic ES cells demonstrates the complete integration of the BAC in the genome
BAC clones of *Sox17* were tested for complete integration using genomic ES DNA as template for the PCR on the 5´ and the 3´ BAC end. The PCR shows the complete integration of the *Sox17* clones 4B and F3. ES DNA from unmodified IDG3.2 cells was used as negative control (ES) while BAC DNA of the corresponding BAC was used as positive control (BAC).

caused the generation of two “BAC ends” named “5 PCR” and “3 PCR”, which could be independently verified by PCR on the genomic DNA (as described in Results 3.2.2 and in Methods 6.2.8).

PCR primer pairs on both sides were designed to demonstrate the integration of the complete BAC as described by Yang and Seed (2003).Both *Sox17* BAC transgenic clones 4B and F3 were selected for the PCR check of complete integration and showed amplificates of the 5´and the 3´ BAC-end (Figure 26).

This demonstrated the complete integration of the BAC transgenic clones 4B and F3 in the genome. The Southern analysis used to verify of the integration the Tomato reporter; while the BAC-end PCR demonstrated the complete insertion of the BAC.

3.4.3 *In vitro* expression of *Sox::H2B-Tomato* BAC transgenic ES cells in *in vitro* differentiated ES cells into endoderm

Sox17 BAC transgenic ES cells were tested in *in vitro* culture for reporter gene expression by differentiation in endodermal cells, as described for *Hhex* BAC transgenic cells in Results 3.2.3. This allowed the direct selection of positive functional clones for BAC transgene expression before continuing with the time consuming test of *in vivo* expression.

The *Sox17::H2B-Tomato* BAC clones 4B and F3 showed transgene expression in differentiated ES cells from day 3 on (Figure 27). The expression started at day 3, began decreasing at day 4 and day 5, until the end of the differentiation at day 6. The expression was clearly localized to the nucleus (Figure 27 m, m´) as expected from a transgene containing the promoter of the H2B gene.

The expression of both *Sox17* BAC transgenic ES clones in *in vitro* differentiated ES cells displayed the effectivity of the confirmation steps done before.

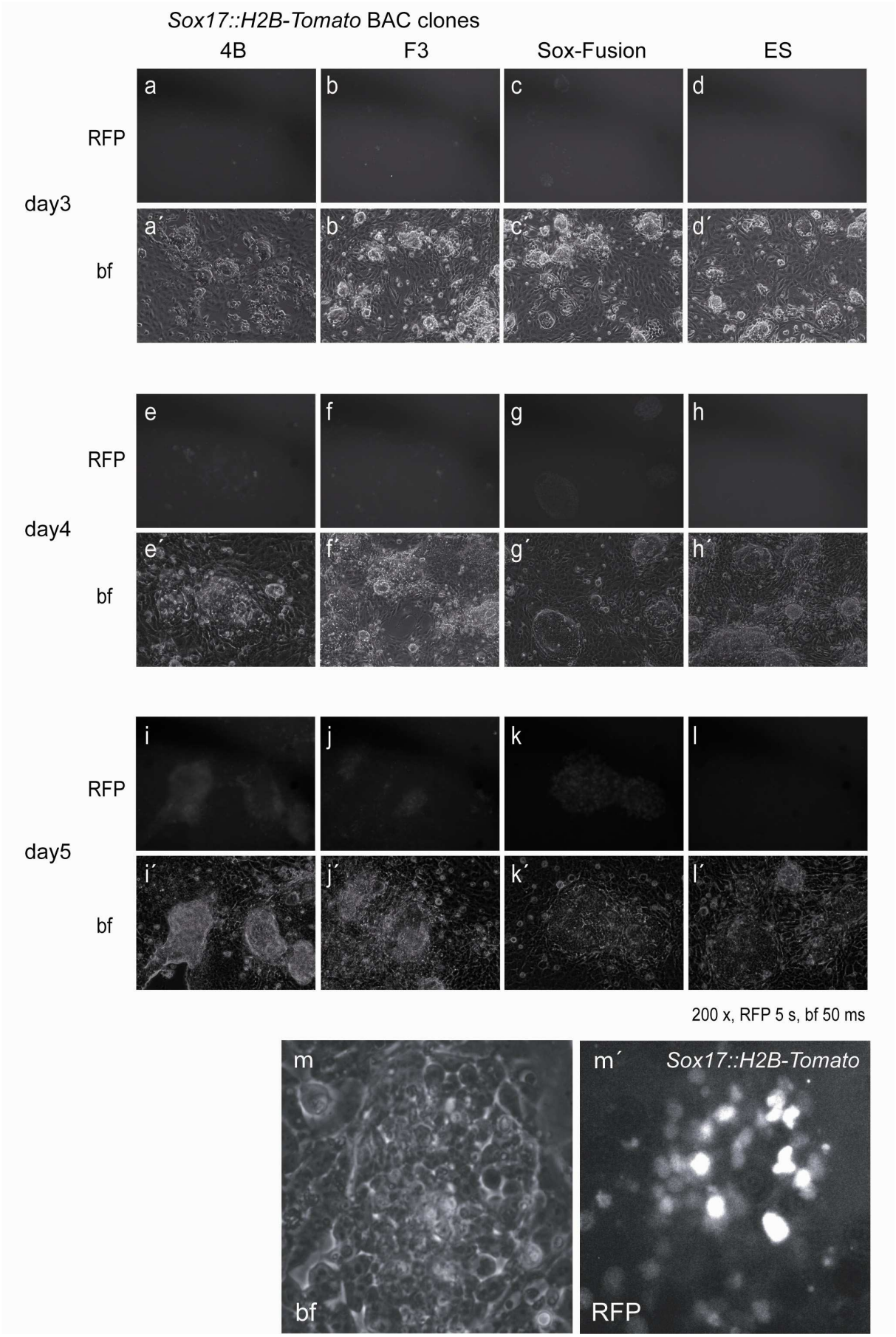

Sox17::H2B-Tomato BAC clones
4B
F3
Sox-Fusion
ES
day3
RFP
bf
a
b
c
d
a´
b´
c´
d´
day4
RFP
bf
e
f
g
h
e´
f´
g´
h´
day5
RFP
bf
i
j
k
l
i´
j´
k´
l´
200 x, RFP 5 s, bf 50 ms
m
bf
m´
Sox17::H2B-Tomato
RFP

Figure 27 *Sox17::H2B-Tomato* transgene expression during ES cell *in vitro* differentiation in endoderm
Transgene expression of the Tomato reporter in Sox17 BAC transgenic ES cell clones 4B (a, a´, e, e´, i, i´) and F3 (b, b´, f, f´, j, j´) from day 3 to day 5 of in vitro differentiation in endoderm. The expression is localized to the nucleus of the cells (m, m´; clone 4B at day4 of differentiation in higher magnification). ES cells from a Sox-Fusion (knock-in) were used as positive control for the in vitro differentiation. Wild-type ES cells (ES) were used as negative control.

3.4.4 Analysis of a *Sox17::H2B-Tomato* BAC transgenic mouse line

To study *Sox17* gene expression and lineage differentiation in the developing mouse embryo, we generated a *Sox17::H2B-Tomato* transgenic mouse lines from our pre-selected and tested ES cell clones (described in Results 3.3 - 3.3.3). Therefore, BAC transgenic ES cells of clone 4B and F3 were used to generate chimeras by diploid aggregation technique. As for the *Hhex* transgenic line, these were crossed to CD1 mice for germline transmission of the transgene.

After the establishment of the mouse line from *Sox17::H2B-Tomato* ES clone 4B, heterozygote animals were breed to generate embryos and collect several developmental stages to analyse the transgene expression pattern. The expression of *Sox17* from E7.0 on is restricted to the embryonic hindgut at the anterior end of the primitive streak and the DE (Kanai *et al*., 1996; Kanai-Azuma, 2002). *Sox17* expression at HF stage is restricted to the VE and the foregut endoderm and from E9.5 on to the intersomitic blood vessels, fetal hematopoietic stem cells and the vascular endothelial cells of the dorsal aortae (Matsui *et al*., 2006; Kim *et al*., 2007). Only a weak expression of the *Sox17::H2B-Tomato* transgene in the DE of mouse embryos at E7.5 could be detected (see Figure 28 a). But the very weak signal intensity, almost at the limit of detection, prevented a clear detection of the fluorescent signal above autofluorescent background levels (see Figure 28 a, b). This could be due to the single integration of the transgene in the genome, as shown in Figure 25.

It was anyway possible to observe that the signal was localized into the nucleus (Figure 28, a), as expected. *Sox::H2B-Tomato* transgene expression could not be observed at earlier stages of gastrulation (E7.0, E7.25) using confocal imaging technique. Between embryonic stages E9.5 and E12.5 transgene expression could not be detected by inverse fluorescent microscope imaging; although, it was already shown that from E9.5 on, Sox17 is expressed in intersomitic blood vessels and vascular endothelial cells of the dorsal aorta (Matsui *et al*., 2006).

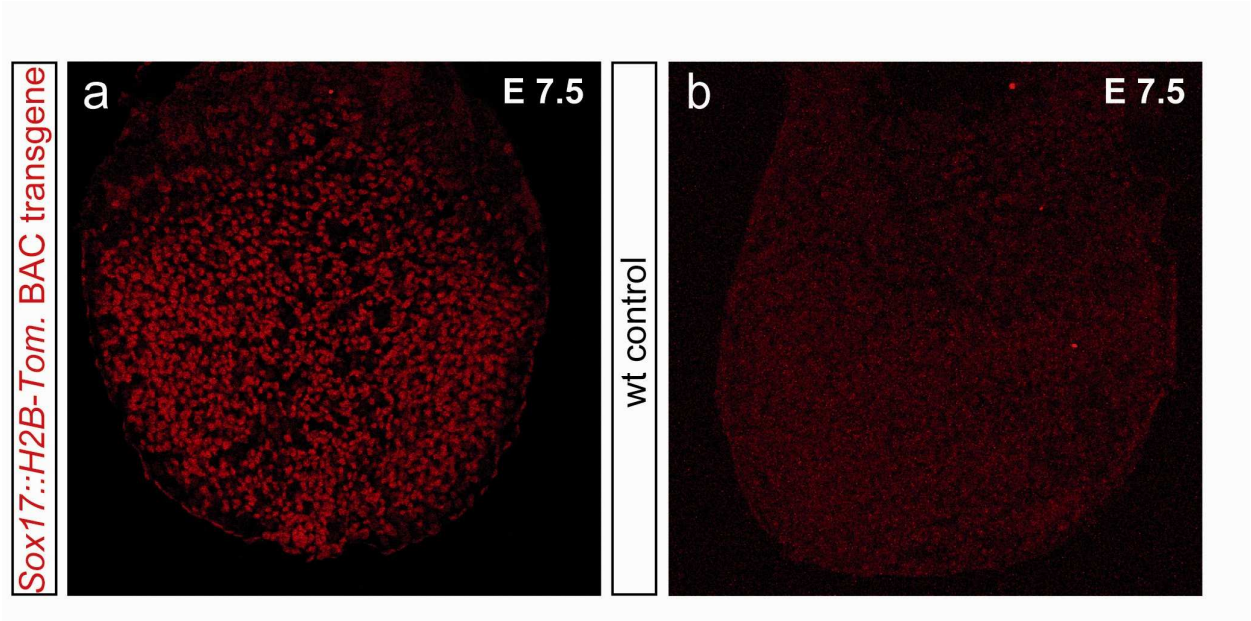


Figure 28 Expression of *Sox::H2B-Tomato* transgene at stage E7.5 in the definitive endoderm
(a) *Sox::H2B-Tomato* transgene expression at E7.5 in the definitive endoderm (DE). (b) Wild-type embryo at E7.5 shows autofluorescent background. Pictures a, b; merge of planes.

3.5 Generation of *Foxa2::Lyn-Venus* BAC transgenic ES cells and expression analysis in tetraploid-derived mouse embryos

The transcription factor *Foxa2* is a pioneer factor for endoderm formation. *Foxa2* defines the first endodermal cells in the primitive streak and is essential for the induction of expression of other endodermal genes. It is expressed in the VE and DE during gastrulation and in the foregut endoderm at later stages of development. In the adult mouse *Foxa2* is expressed in lung, liver and pancreas. *Foxa2* was used to study its expression and lineage differentiation in the endoderm germ layer by lineage labelling.

The selected BAC RPCI-23-254G2 has a size of 201 kb, and contains only the *Foxa2* ORF. The DNA fingerprinting gel of the *Foxa2* BAC, recombined with the Lyn-Venus mini-targeting vector is shown in Figure 21. The maintained restriction pattern demonstrated the preserved integrity of the *Sox17* BAC, while the restriction fragment length polymorphism (RFLP; Figure 29 b) indicated the BAC modification steps. The *Kpn* I digestion showed RFLP from 6.8 kb to 10.0 kb after integration of the mini-targeting vector in the BAC by homologous recombination ("homologous recombined" in Figure 29 a-c).

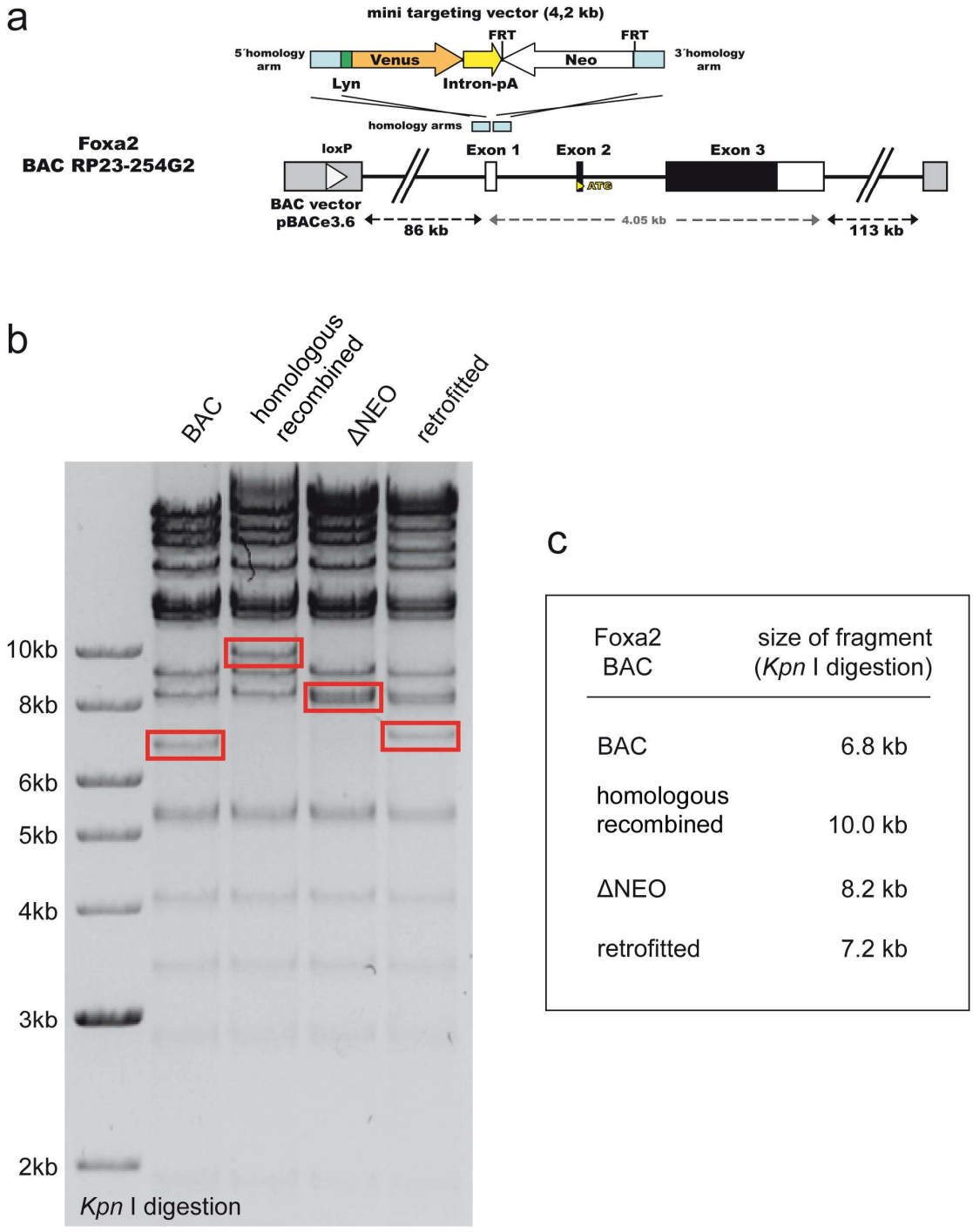


Figure 29 Generation and verification of a *Foxa2::Lyn-Venus* reporter transgene
(a) Mini targeting vector containing 5´and 3´ homology arm, Lyn-Venus reporter gene, Intron-polyA and FRT-flanked *Neo* resistance gene. This vector was used to knock-in the reported gene into exon1 of *Foxa2* as indicated. (b) Agarose gel of *Kpn* I digested *Foxa2* BAC DNA isolated from colonies of the unmodified BAC (BAC), the BAC after homologous recombination of the mini-targeting construct (homologous recombined), the BAC after deletion of the *Neo* cassette by Flp recombination (ΔNEO) and the BAC after the retrofitting process using pRetroES-tk (retrofitted). The continuous restriction pattern demonstrates the integrity of the BAC while the restriction fragment length polymorphism (RFLP; indicated by red boxes) shows the BAC modification steps. (c) The box shows the indicated RFLP fragment sizes of the modification steps.

The deletion of the *Neo* resistance cassette caused a RFLP from 10.0 kb to 8.2 kb ("ΔNEO"). The integration of the new selection marker by retrofitting created a new restriction fragment of 7.2 kb length ("retrofitted") (Figure 29 c).

The results of the RFLP confirmed the alterations which were expected from *in silico* data of the fingerprinting (Figure 29 c). The used steps of BAC recombineering in bacteria appeared as very efficient which was indicated by the high number of bacteria colonies. About 95% (n=20) of the tested *Foxa2* BAC clones showed maintained integrity during the recombineering steps. The efficiency of the retrofitting step was 23% (n=30). Positive retrofitted bacterial colonies were identified by PCR screening (Yang and Seed, 2003; Methods 6.2.16). The modification of the *Foxa2* BAC was successful and the single steps could be followed precisely by RFLP.

3.5.1 Generation of *Foxa2* BAC transgenic reporter ES cells

Foxa2 BAC transgenic ES cells were produced as described above for *Hhex* transgene (Results 3.2.1). The linearized *Foxa2* BAC DNA was electroporated in IDG3.2 ES cells (Hitz *et al*., 2007). *Neomycin* resistant ES cell colonies were tested by Southern analysis using a probe which detected the fluorescent protein Tomato. One electroporation in ES cells using 30 µg linearized BAC DNA resulted in 14 ES cell clones.

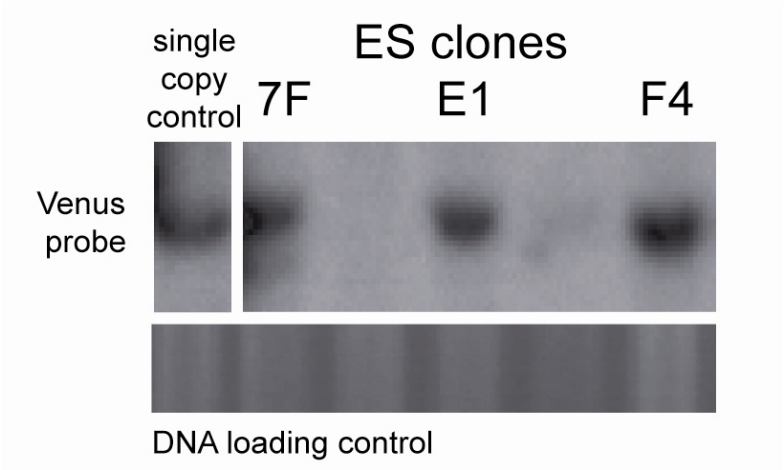


Figure 30 Southern analysis of individual *Foxa2* BAC transgenic ES clones showing integration of one transgene in the genome
The used Southern probe was Venus (size 4.6 kb) for Foxa2 (9.3 kb). The BAC DNA was digested with *EcoR* V. DNA from single copy integration was used as control. The DNA amount is shown in the loading control to estimate the number of integrated BACs in the genome. The *Foxa2* clones 7F, E1 and F4 show integration of one BAC in the genome. Not shown *Foxa2* clone E4.

The Southern analysis showed 4 out of 14 positive clones for the Tomato probe which corresponds to a ratio of 29 %. The number of integrated BACs in the genome was estimated by comparison between the signal strength of the single copy control with that of the positive clones, after normalization of the amount of loaded DNA (DNA loading control). The Southern positive *Foxa2* BAC transgenic clones 7F, E1, E4 and F4 showed a single integration in the genome (Figure 30). The number of Southern positive clones was with 29% very high.

3.5.2 Integrity check of *Foxa2* BAC transgenic ES cells by BAC end PCR

The integrity of the BAC was demonstrated by several digestions in Results 3.5. To show the complete integration of the intact BAC, a PCR check of the *Foxa2* BAC transgenic ES clones was done. The linearization of the BAC at the PI-*Sce*I site in the BAC vector backbone pBACe3.6 before the electroporation in the ES cells caused the generation of two “BAC ends” named “5’” and “3’”, which could be independently verified by PCR on the genomic DNA. PCR primer pairs on both sides were designed to demonstrate the integration of the complete BAC as described by Yang and Seed (2003).

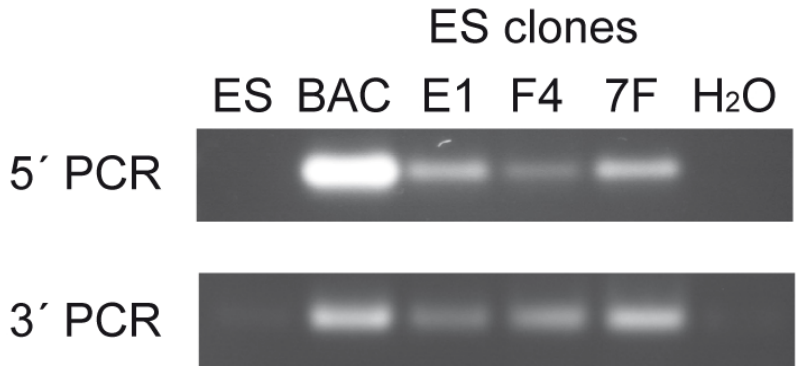


Figure 31 BAC-end PCR of *Foxa2* BAC transgenic ES cells demonstrates the complete integration of the BAC in the genome
BAC clones of Foxa2 were tested for complete integration using genomic ES DNA as template for the PCR on the 5´ and the 3´ BAC end. The PCR shows the complete integration of all three clones E1, F4 and 7F. ES DNA from unmodified IDG3.2 cells was used as negative control (ES) while BAC DNA of the corresponding BAC was used as positive control (BAC).

All three *Foxa2* BAC transgenic clones (E1, F4, 7F) were selected for the PCR check of complete integration of the BAC-ends. All clones showed amplified bands for of the 5´and the 3´ BAC-end (Figure 31). This demonstrated the complete integration of all BAC clones in the genome. After the verification of the integration of the Venus

reporter by Southern analysis and demonstration of complete insertion of the BAC by BAC-end PCR. ES clones were further characterized by *in vitro* analysis.

3.5.3 *In vitro* expression of *Foxa2* BAC transgenic ES cells in *in vitro* differentiated ES cells into endoderm

The three *Foxa2* BAC transgenic ES cells were tested in *in vitro* culture for reporter gene expression by differentiation in endodermal cells as described for *Hhex* BAC transgenic cells (Results 3.2.3). This allowed the direct selection of positive functional clones for BAC transgene expression before continuing with the test of *in vivo* expression. The *in vitro* differentiation with *Foxa2::Lyn-Venus* BAC transgenic ES cells showed transgene expression in all three tested BAC clones E1, F4, 7F (Figure 32) and clone E4 (not shown).

The transgene expression was weak in comparison with the Sox-Fusion control or with the *in vitro* differentiation of *Hhex* and *Sox17* transgenes (Figure 19 and Figure 27, respectively). Also in this case, the maximum intensity was observed on day 4. The differentiation procedure showed a difference in intensity between the clones. The transgene expression was higher in clones F4 and 7F than in clone E1 (Figure 32 g, h, f). Unfortunately, the weak signal intensity did not allow a clear localisation of the transgene in the cells (see Figure 32 p, p´).

3.5.4 *In vivo* analysis of *Foxa2::Lyn-Venus* BAC transgenic ES cells by tetraploid complementation technique

To check the *in vivo* expression of the *Foxa2* BAC transgenic ES clones which were successfully tested *in vitro*, completely ES cell-derived embryos were generated using tetraploid complementation technique (Nagy *et al*., 1993). Tetraploid embryos (4n) were generated using the *Foxa2::Lyn-Venus* BAC transgenic clones E1, E4 and 7F were used to generate tetraploid-derived embryos using dsRed tetraploid embryos (Vintersten *et al*., 2004) to be able to distinguish between the cells. The expression of *Foxa2* during gastrulation is restricted to the VE, DE, node and axial mesoderm (Sasaki and Hogan, 1993; Monaghan *et al*., 1993; Ang *et al*., 1993; Ang *et al*., 1994). *Foxa2::Lyn-Venus* transgenic cells (green) could be observed in definitive endodermal cells which emerged between the visceral endodermal cells at

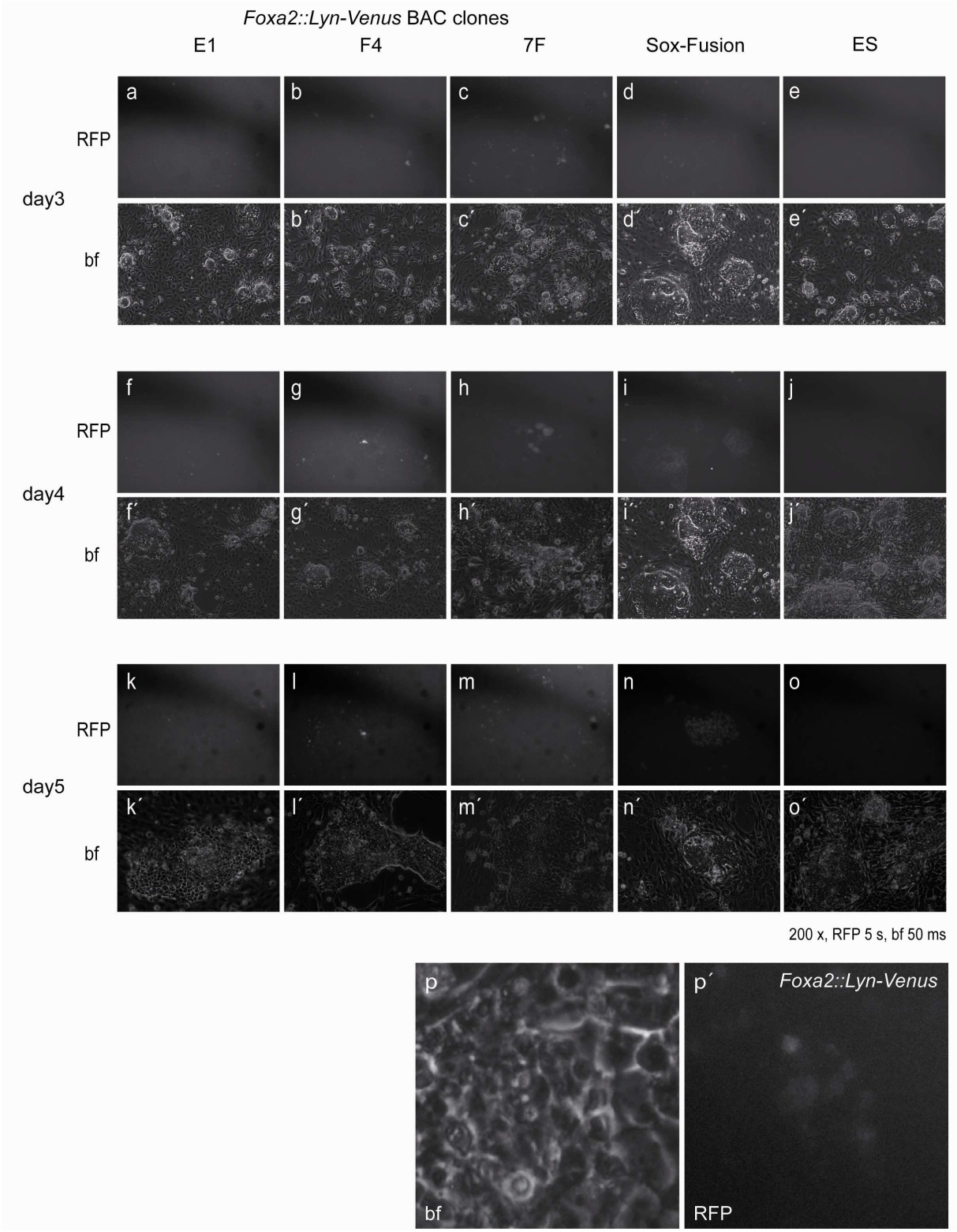


Figure 32 *Foxa2::Lyn-Venus* transgene expression during ES cell *in vitro* differentiation in endoderm

Transgene expression of the Venus reporter in *Foxa2* BAC transgenic ES cell clones E1, F4 and 7F from day 3 to day 5 of *in vitro* differentiation in endoderm. The expression is localized to the membrane of the cells (p, p´; clone 7F at day5 of differentiation in higher magnification). ES cells from a Sox-Fusion (knock-in) were used as positive control for the *in vitro* differentiation. Wild-type ES cells (ES) were used as negative control.

the anterior end of the primitive streak (PS) in the node region and along the midline of tetraploid embryos between E7.0 and E7.5 (as an example for all, clone 7F is shown in Figure 33).

These results are consistent with the published mRNA expression of *Foxa2* (Ang *et al*., 1993; Monaghan *et al*., 1993; Sasaki and Hogan, 1993).

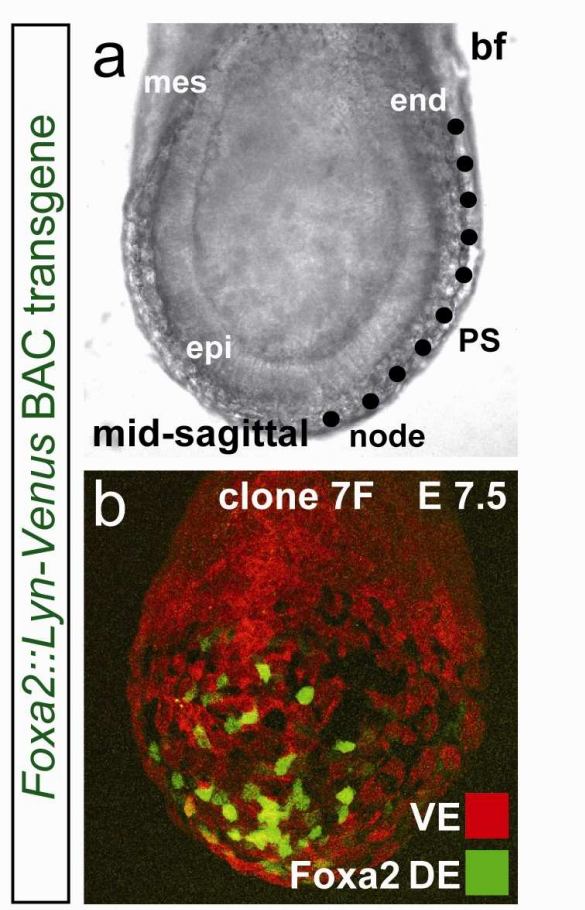


Figure 33 BAC reporter transgene expression *in vivo* in tetraploid-derived embryos
(a, b) In vivo expression of *Foxa2::Lyn-Venus* transgene in the definitive endoderm of a *Foxa2::Lyn-Venus* BAC transgenic ES cell-derived embryo at E7.5 (*Foxa2* BAC transgenic ES cells <-> 4n dsRed embryo). Image aquisition was done using laser scanning confocal imaging. Abbreviations: Mesoderm (mes), epiblast (epi), primitive streak (PS), endoderm (end.

The *in vivo* test of BAC transgenic ES cells by tetraploid aggregation technique showed a specific transgene expression in 3 of 3 tested clones. Taken together, the complete pre-screening scheme for BAC transgenic ES cells allowed selection for clones which expressed the transgene accurately in the endoderm germ layer.

3.6 Generation of *Nkx2.1::H2B-Venus* BAC transgenic ES cells and expression analysis in tetraploid-derived mouse embryos

The gene was selected for lineage labelling as an essential marker gene for the formation of an endodermal organ which is expressed from the earliest stages of organ formation on. The selected BAC RPCI-23-3H18 has a size of 216 kb, and contains only the *Nkx2.1* ORF.

The DNA fingerprinting gel of the *Nkx2.1* BAC, recombined with the H2B-Venus mini-targeting vector is shown in Figure 34. The maintained restriction pattern demonstrated the preserved integrity of the *Nkx2.*1 BAC, while the restriction fragment length polymorphism (RFLP; Figure 34 b) indicated the BAC modification steps. The *EcoR* V digestion showed RFLP from 13.1 kb to 16.1 kb after integration of the mini-targeting vector in the BAC by homologous recombination ("homologous recombined" in Figure 34 a-c). The deletion of the *Neo* resistance cassette caused a RFLP from 16.1 kb to 14.3 kb ("ΔNEO"). The integration of the new selection marker by retrofitting created a new restriction fragment of 10.7 kb length ("retrofitted") (Figure 34 c).

The results of the RFLP confirmed the alterations which were expected from *in silico* data of the fingerprinting (Figure 34 c). The used steps of BAC recombineering in bacteria appeared as very efficient which was indicated by the high number of bacteria colonies. About 85% (n=20) of the tested *Nkx2.1* BAC clones showed maintained integrity during the recombineering steps. The efficiency of the retrofitting step was 13% (n=30). Positive retrofitted bacterial colonies were identified by PCR screening (Yang and Seed, 2003; Methods 6.2.16). The modification of the *Nkx2.1* BAC was successful and the single steps could be followed precisely by RFLP.

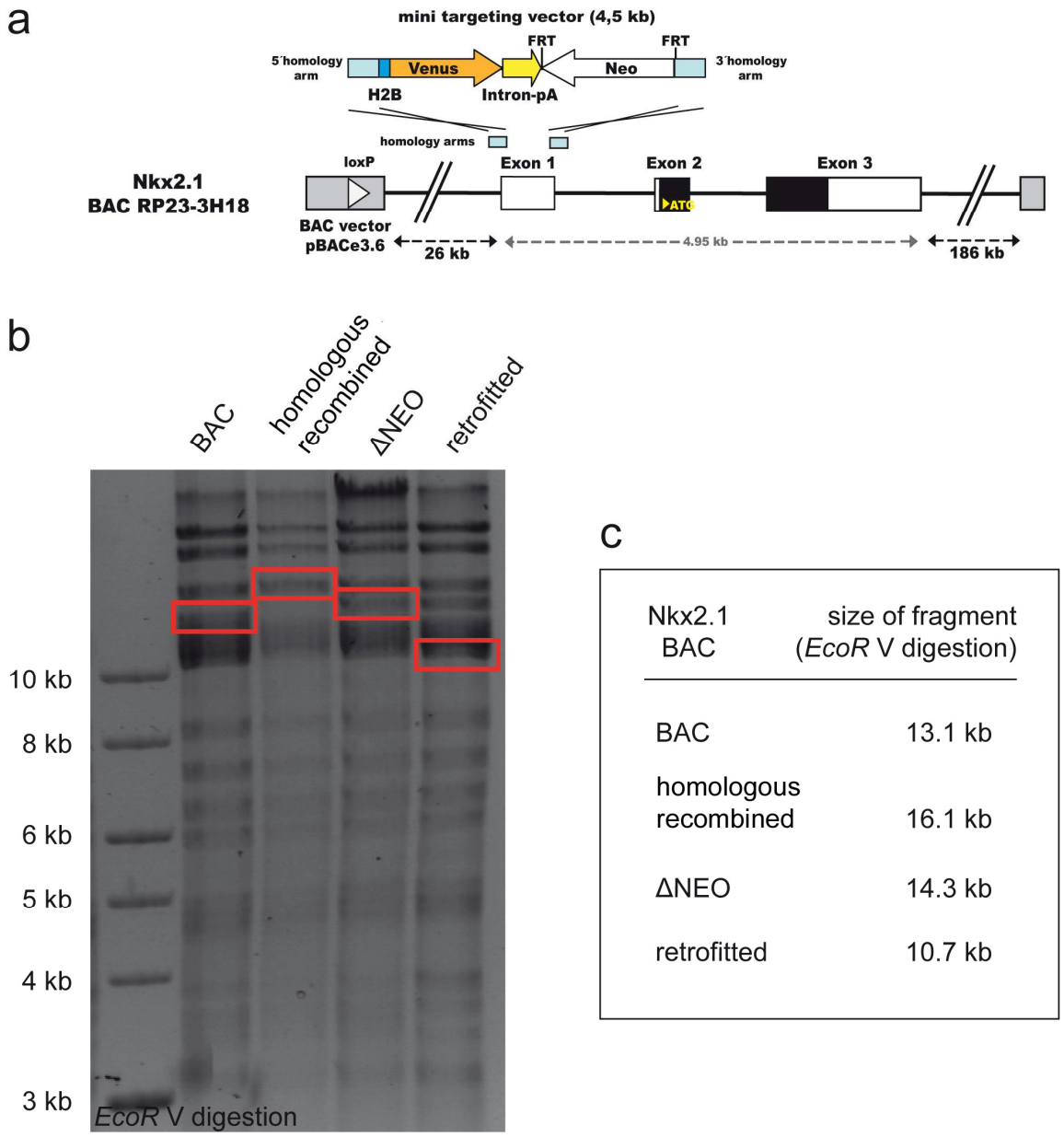


Figure 34 Generation and verification of a *Nkx2.1::H2B-Venus* reporter transgene
(a) Mini targeting vector containing 5´and 3´ homology arm, H2B-Venus reporter gene, Intron-polyA and FRT-flanked *Neo* resistance gene. This vector was used to knock-in the reported gene into exon1 of *Nkx2.1* as indicated. (b) Agarose gel of *EcoR* V digested *Nkx2.1* BAC DNA isolated from colonies of the unmodified BAC (BAC), the BAC after homolog recombination of the mini-targeting construct (homolog recombinated), the BAC after deletion of the *Neo* cassette by Flp recombination (ΔNEO) and the BAC after the retrofitting process using pRetroES-tk (retrofitted). The continuous restriction pattern shows the BAC integrity while the restriction fragment length polymorphism (RFLP; indicated by red boxes) shows the BAC modification steps. (c) The box shows the indicated RFLP fragment sizes of the modification steps.

3.6.1 Generation of *Nkx2.1* BAC transgenic reporter ES cells

Nkx2.1 BAC transgenic ES cells were produced as described above for *Hhex* transgenes (Results 3.2.1). The linearized *Nkx2.*1 BAC DNA was electroporated in IDG3.2 ES cells (Hitz *et al*., 2007). *Neomycin* resistant ES cell colonies were tested by Southern analysis using a probe which detected the fluorescent protein Tomato. One electroporation in ES cells using 30 µg linearized BAC DNA resulted in 10 ES cell clones. The Southern analysis showed 4 out of 14 positive clones for the Venus probe which corresponds to a ratio of 29 %.

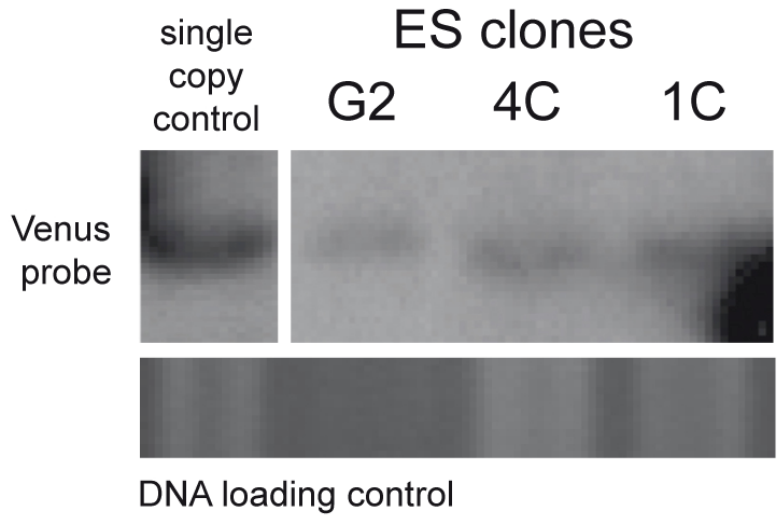


Figure 35 Southern analysis of individual *Nkx2.1* BAC transgenic ES clones showing integration of one transgene in the genome
The used Southern probe was Venus (size 4.6 kb) for Nkx2.1 (3.7 kb). The BAC DNA was digested with *Hind* III. DNA from single copy integration was used as control. The DNA amount is shown in the loading control to estimate the number of integrated BACs in the genome. The clones G2, 4C, 1C and D1 (not shown) show the integration of one BAC in the genome.

The number of integrated BACs in the genome was estimated by comparison between the signal strength of the single copy control with that of the positive clones, after normalization of the amount of loaded DNA (DNA loading control). The Southern positive *Nkx2.1* BAC transgenic clones G2, 4C, 1C and D1 showed a single integration in the genome (Figure 35).

The number of generated clones was very low but the percentage of Southern positive clones obtained with our strategy was with 29% very high compared to the rate of Southern positive clones generated in a knock-in strategy.

3.6.2 Integrity check of *Nkx2.1* BAC transgenic ES cells by BAC end PCR

To show the complete integration of the intact BAC, a PCR check of 2 of 4 *Nkx2.1* BAC transgenic ES clones was done. The linearization of the BAC at the PI-SceI site in the BAC vector backbone pBACe3.6 before the electroporation in the ES cells caused the generation of two “BAC ends” named “5’” and “3’”, which could be independently verified by PCR on the genomic DNA. PCR primer pairs on both sides were designed to demonstrate the integration of the complete BAC as described by Yang and Seed (2003).

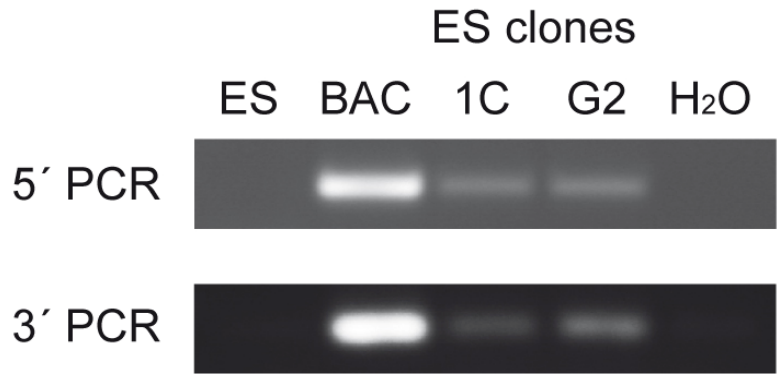


Figure 36 BAC-end PCR of *Nkx2.1* BAC transgenic ES cells demonstrates the complete integration of the BAC in the genome
BAC clones of Nkx2.1 were tested for complete integration using genomic ES DNA as template for the PCR on the 5´ and the 3´ BAC end. The PCR shows the complete integration of the *Nkx2.1* clones 1C and G2. ES DNA from unmodified IDG3.2 cells was used as negative control (ES) while BAC DNA of the corresponding BAC was used as positive control (BAC).

The tested *Nkx2.1* BAC transgenic clones 1C and G2 were selected for the PCR check of complete integration. Both tested transgenic clones showed amplificates of the 5´and the 3´ BAC-end (Figure 36). This demonstrated the complete integration of the BAC clones in the genome. After the verification of the integration of the Venus reporter by Southern analysis and demonstration of complete insertion of the BAC by BAC-end PCR, ES clones were further characterized by *in vivo* analysis. The *in vitro* differentiation of *Nkx2.1* BAC transgenic ES cells in endoderm to test for *in vitro* transgene expression of the ES cell clones was not possible. The differentiated ES cells do not reach the developmental stage of endodermal organ progenitors where *Nkx2.1* expression begins (Tada *et al*., 2005; Kubo *et al*., 2004). The expression test of the *Nkx2.1* BAC transgenic ES cells was done direct *in vivo* by generation of tetraploid-derived embryos.

3.6.3 *In vivo* analysis of *Nkx2.1::H2B-Venus* BAC transgenic ES cells by tetraploid complementation technique

To check the *in vivo* expression of the *Nkx2.1* BAC transgenic ES clones completely ES cell-derived embryos were generated using tetraploid complementation technique (Nagy *et al*., 1993). Tetraploid embryos (4n) were generated as described. The *Nkx2.1* clones 1C and G2 were analysed with tetraploid complementation technique. The expression of the transgene could be observed in the lung buds of E11.5 derived tetraploid embryos (Figure 37). *Nkx2.1* is expressed in the lung endodermal cells which originated from the gut endoderm and in the two main bronchi of the primitive lung bud and in all epithelial cells at the early pulmonary morphogenesis (Lazzaro et *al*., 1991; Kimura *at el*., 1999; Minoo *et al*., 1999, Desai *et al*., 2004). Using fluorescent microscopy we were not able to detect a fluorescent signal above

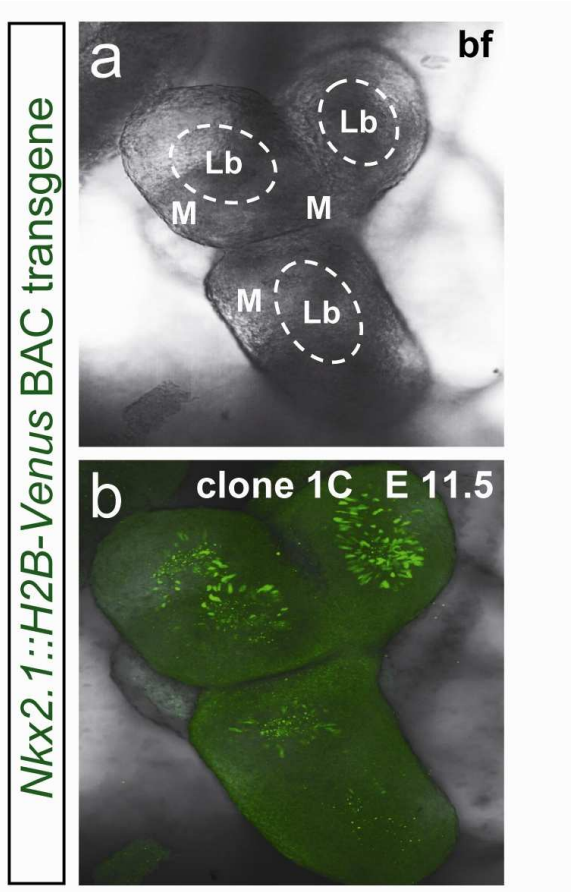


Figure 37 BAC reporter transgene expression *in vivo* in tetraploid-derived embryos
(a, b) In vivo expression of *Nkx2.1::H2B-Venus* transgene in the lung buds of a E 11.5 ES cell-derived embryo (*Nkx2.1* BAC transgenic ES cells <-> 4n CD1 embryo). *gfp*-antibody staining on the Venus reporter. Image acquisition was done using laser scanning confocal imaging. Abbreviations: lung-bud (Lb), mesenchyme (M).

background level in embryonic lung, most likely due to the single copy integration of the BAC reporter in the genome (Figure 35). Using whole mount immunohistochemistry, we could readily detect the Venus protein in epithelial cells of the lung, but not in the surrounding mesenchyme (as an example for both, *Nkx2.1* clone 1C is shown in Figure 37), corresponding to what reported in literature for the endogenous *Nkx2.1* mRNA (Lazzaro *et al*., 1991).

The *in vivo* test of BAC transgenic ES cells by tetraploid aggregation technique showed a specific transgene expression in 2 of 2 tested clones. Taken together, the complete pre-screening scheme for BAC transgenic ES cells allowed selection for clones which expressed the transgene accurately. The one copy insertion of the reporter was too weak to be detected in the lung bud at E11.5 under the surrounding mesenchyme. The *Nkx2.1* expressing cells in the lung-bud were detectable after whole-mount antibody staining using a gfp-antibody against the Venus reporter.

3.7 Generation of a *Pdx1::H2B-Venus* BAC

The transcription factor *Pdx1* is the master regulator of pancreatic endocrine cell development and essential for the development of the pancreas. *Pdx1* is first expressed at E8.5 in endodermal foregut cells which give rise to the pancreatic bud. It is expressed during the whole pancreatic development and in the adult organ. The gene was selected for lineage labelling as an essential marker gene for the formation of an endodermal organ which is expressed from the earliest stages of organ formation on.

The selected BAC RPCI-23-355P12 has a size of 246 kb, is located on chromosome 5, and contains the *Pdx1* ORF and additionally the ORF of *GS homeobox 1* (Gsh1; Singh *et al*., 1991), *Cdx2* (James and Kazenwadel, 1991), *parahox cluster neighbour* (*Prhoxnb*; Ramazzina *et al*., 2006), *FMS-like tyrosine kinase 3* (*Flt3*; Rosnet *et al*., 1991). The DNA fingerprinting gel of the *Pdx1* BAC, recombined with the H2B-Venus mini-targeting vector is shown in Figure 38. The maintained restriction pattern demonstrated the preserved integrity of the *Pdx1* BAC, while the restriction fragment length polymorphism (RFLP; Figure 38 b) indicated the BAC modification steps. The *Xho* I digestion showed RFLP from 4.1 kb to 7.2 kb after integration of the mini-targeting vector in the BAC by homologous recombination (“homologous recombined” in Figure 38 a-c).

The deletion of the *Neo* resistance cassette caused a RFLP from 7.2 kb to 5.4 kb ("ΔNEO"). The integration of the new selection marker by retrofitting created a new restriction fragment of 4.5 kb length ("retrofitted") (Figure 38 c).

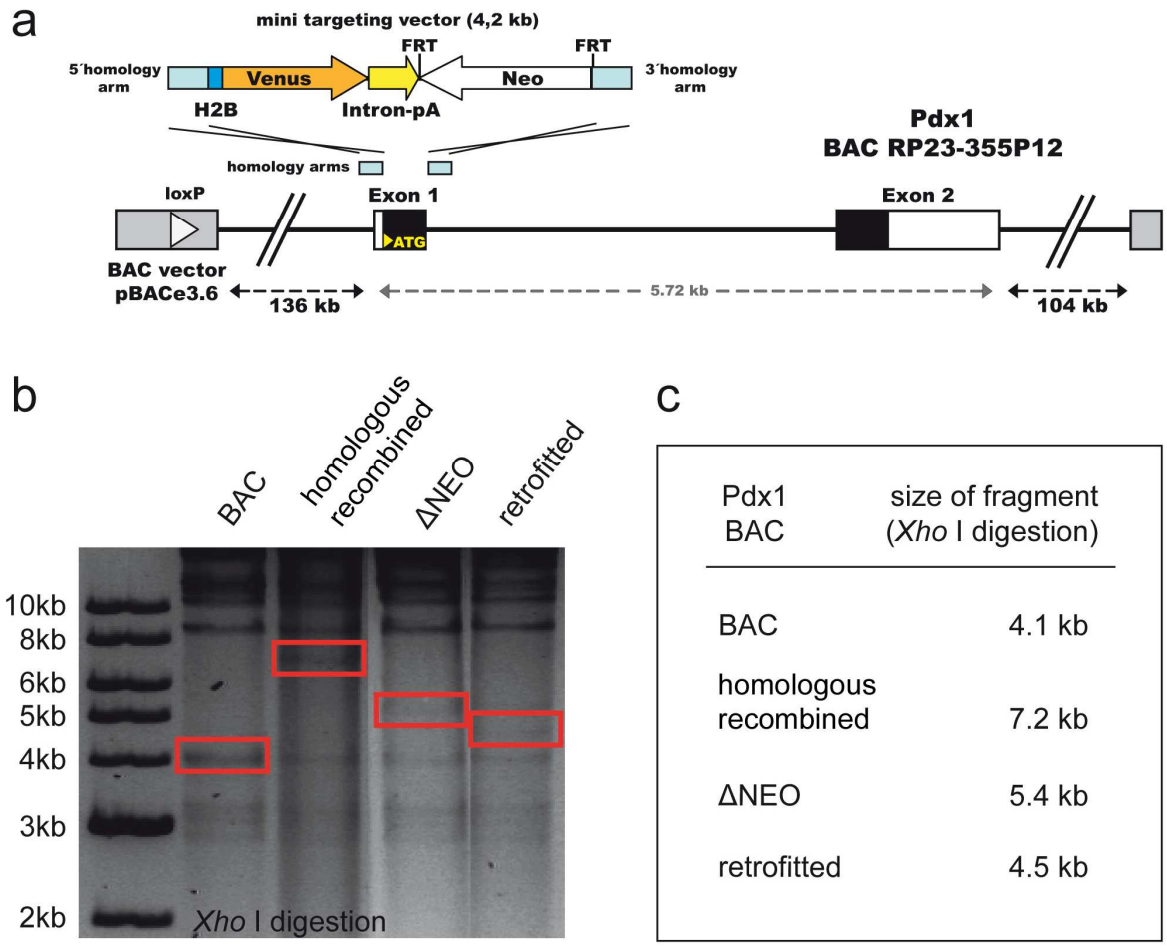


Figure 38 Generation and verification of a *Pdx1::H2B-Venus* reporter transgene
(a) Mini targeting vector containing 5´and 3´ homology arm, Lyn-Venus reporter gene, Intron-polyA and FRT-flanked *Neo* resistance gene. This vector was used to knock-in the reported gene into exon1 of *Pdx1* as indicated. (b) Agarose gel of *Xho* I digested *Pdx1* BAC DNA isolated from colonies of the unmodified BAC (BAC), the BAC after homolog recombination of the mini-targeting construct (homolog recombinated), the BAC after deletion of the *Neo* cassette by Flp recombination (ΔNEO) and the BAC after the retrofitting process using pRetroES-tk (retrofitted). The continuous restriction pattern shows the BAC integrity while the restriction fragment length polymorphism (RFLP; indicated by red boxes) shows the BAC modification steps. (c) The box shows the indicated RFLP fragment sizes of the modification steps.

The results of the RFLP confirm the alterations which were expected from *in silico* data of the fingerprinting (Figure 38 c). The used steps of BAC recombineering in bacteria appeared as very efficient which was indicated by the high number of bacteria colonies. About 85% (n=20) of the tested BAC clones showed maintained integrity during the recombineering steps. The efficiency of the retrofitting step was 20% (n=30).

Positive retrofitted bacterial colonies were identified by PCR screening (Yang and Seed, 2003; Methods 6.2.16). The modification of the *Pdx1* BAC was successful and the single steps could be followed precisely by RFLP.
The modified *Pdx1* BAC was used for generation of *Pdx1* BAC transgenic ES cells as described above for the *Hhex* transgenes (Results 3.2.1). Unfortunately, no positive clone could be detected out of 19 tested *Pdx1* clones.

3.8 Whole mount Foxa2 antibody staining on *Hhex::Lyn-Tomato* transgenic mouse embryos

Whole mount antibody stainings were done to compare *Hhex::Lyn-Tomato* transgene expression and lineage differentiation to *Foxa2* expression in the developing mouse embryo. It was recently shown that forkhead box transcription factor family member *Foxa2* is a key factor for the formation of endoderm development, as the loss of the *Foxa2* gene leads to absence of anterior definitive endoderm (ADE) and axial mesoderm as well as node and notochord (Ang and Rossant, 1994; Weinstein *et al*., 1994).
The whole mount staining of mouse embryos allows the visualisation of gene expression on single cell level. The method used is based on a protocol from Nakaya *et. al*. (2005) (see Methods 6.22): basically, it avoids steps which are chemically too harsh, the use of dehydrating and rehydrating processes, and it includes just a short-fixation step. This allows also the use of the sensitive anti-phosphoproteins antibodies which bind to active molecules. As previously described above, the *Hhex::Lyn-Tomato* transgene expression could be detected in the anterior visceral endoderm (AVE) and definitive endoderm (DE) at mid-streak stage (Figure 39 b, c). At this stage the *Foxa2* protein expression was also detectable in the same endodermal regions, although to less extent in the AVE (Figure 39 a, c, d). Therefore, *Foxa2*/ *Hhex::Lyn-Tomato* double positive cells could be detected in the anterior visceral endoderm (Figure 39 c´); but also *Foxa2* negative/ *Hhex::Lyn-Tomato* positive cells could be found in the same region.
The superficial single layer of visceral endoderm cells can be distinguished from the definitive endodermal cells in sagittal sections (Figure 39. d). These sections also highlighted the membrane localisation of the *Hhex::Lyn-Tomato* transgene in the surface cell layer of the visceral endoderm and definitive endoderm, but not in

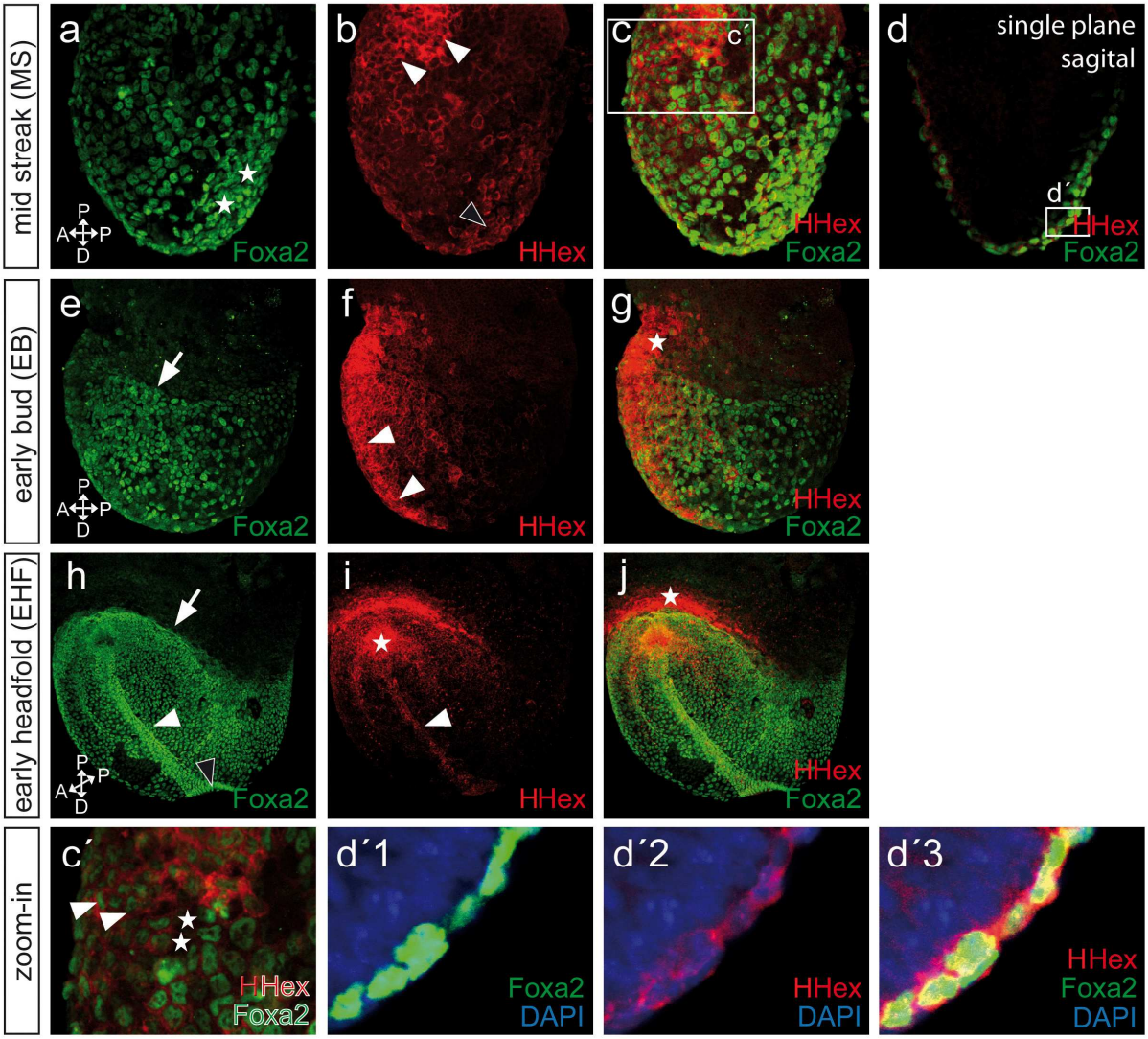


Figure 39 *Foxa2* whole mount staining on *Hhex::Lyn-Tomato* transgenic mouse embryos
(a) *Foxa2* antibody staining on mid-streak (MS) stage embryo shows *Foxa2* expression in the definitive endoderm (DE, white asterisks) and the visceral endoderm. (b) *Hhex::Lyn-Tomato* transgene expression in the anterior visceral endoderm (AVE, white arrowheads) and the definitive endoderm (DE, black arrowhead). (c) Overlay of *Foxa2* antibody staining and *Hhex::Lyn-Tomato* transgene expression. (d) Sagittal section (single plane) shows Foxa2 expression in the one layer thick visceral endoderm (anterior) and multi layer *Foxa2* expressing cells in the definitive endoderm (posterior) and *Foxa2* positive epiblast cells. (e) *Foxa2* expressing definitive endoderm cells at early bud (EB) stage in the embryonic part with a distinct border to the extraembryonic part (white arrow) of the embryo. (f) *Hhex::Lyn-Tomato* transgene expression in the midline region (white arrowheads), lateral endoderm and anterior extraembryonic VE. (g) The overlay of both expression profiles shows the strong *Hhex::Lyn-Tomato* transgene expression in the anterior extraembryonic VE (asterisk). (h) *Foxa2* expression at early headfold stage in the complete embryonic part of the embryo (border, indicated by white arrow). Dense cells indicating the primitive streak (white arrowhead) and node (black arrowhead). (i) *Hhex::Lyn-Tomato* transgene expression in the midline, anterior visceral endoderm (AVE) anterior extraembryonic VE and strong expression at the proximal end of the midline (asterisk). (j) Overlay shows the strong extraembryonic expression domain of the Hhex transgene. (c´) Zoom-in at mid-streak stage shows *Foxa2*/*Hhex::Lyn-Tomato* double positive cells in the anterior visceral endoderm (white arrowheads) and Foxa2 negative/ *Hhex::Lyn-Tomato* positive cells (asterisk). (d´1-3) Cellular localisation of *Foxa2* and the *Hhex::Lyn-Tomato* transgene shows double positive definitive endodermal (DE) and visceral endodermal (VE) cells and the membrane localisation of the *Hhex::Lyn-Tomato* transgene (d´2) in the surface cellular layer but not the epiblast. Pictures: d, d´1-d´3; single planes); a-c, e-j; merge of planes.

epiblast cells. Epiblast cells leave the epithelium in the APS region. These cells express weak *Foxa2* and start to intercalate into the overlaying DE/VE layer. Then, cells up-regulate *Foxa2*/*Hhex* expression and form endoderm epithelium surrounding the gastrulation stage embryo (Figure 39 d´1-d´3). At early bud stage, *Foxa2* positive definitive endodermal cells were detectable in the embryonic part of the embryo (Figure 39 e), marking a sharp border with the extraembryonic part (Figure 39 e and h, white arrows). At this stage, the *Hhex::Lyn-Tomato* transgene was detectable in the ADE with a strong expression in the midline region (Figure 39 f, white arrowheads) and in the anterior extraembryonic VE of the embryo (Figure 39 g, asterisk). The higher cellular density among the *Foxa2* positive cell population at the early headfold stage (EHF) indicates the midline and the node (Figure 39 h white arrowhead and black arrowhead). At this stage, *Hhex::Lyn-Tomato* transgene

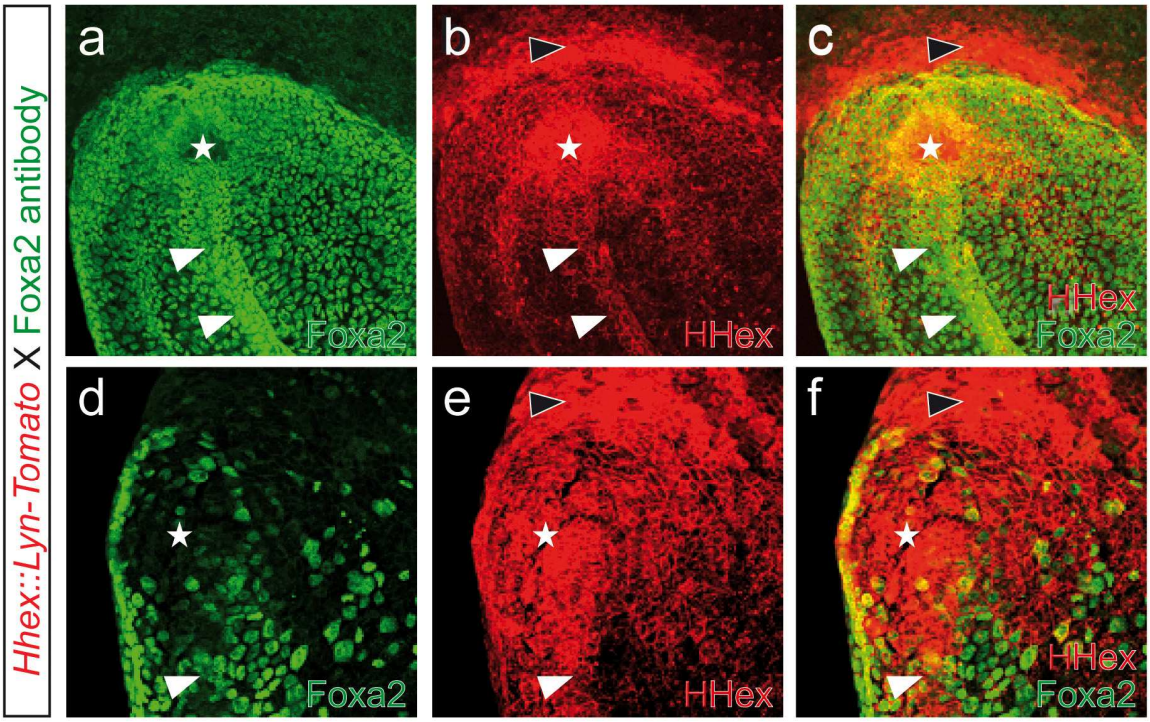


Figure 40 *Hhex::Lyn-Tomato* transgene expression and *Foxa2* antibody staining in the upper midline region of early headfold stage embryo
(a, d) *Foxa2* expressing cells in the anterior definitive endoderm (ADE) the midline region (white arrowheads) and the lack of *Foxa2* expressing cells at the proximal end of the midline. (b, e) *Hhex::Lyn-Tomato* transgene region in the ADE, anterior extraembryonic VE (black arrowhead) and at the proximal end of the midline (asterisk). (c, f) Overlay of *Foxa2* and *Hhex* expression shows the *Foxa2* expression in the midline (white arrowheads) but not at its proximal end and the *Hhex* transgene expression at the proximal end (asterisk) and in the anterior extraembryonic VE (black arrowhead). Pictures: a-f; merge of planes.

expression was restricted both to the midline (Figure 39 i, white arrowhead), in particular in a spot located at the distal tip of the midline (Figure 39 i, asterisk), and to the anterior extraembryonic VE (Figure 39 j, asterisk).

In this study, the authors identified this region as contributing to liver bud progenitors and identified it as anterior intestinal portal (AIP). The high expression of *Hhex::Lyn-Tomato* that was observed in this area would support the proposed role of *Hhex* as the earliest known marker of liver primordium (Zaret, 2002; Bort *et al*., 2004; Bort *et al*., 2005; Calmont *et al*., 2006; Martinez-Barbera *et al*., 2000).

In the same line, the lack of *Foxa2* positive cells in this region and at this stage of development supports the idea that in early liver development *Foxa2* has to be downregulated to drive endodermal cells towards an hepatic cell fate (Zaret, 2002).

3.9 Whole mount embryo stainings using antibodies against active Smad1/5/8 and *Foxa2*

To study endoderm differentiation and lineage segregation processes in the developing embryo under the influence of signalling pathways, whole mount immunostainings were used. The use of antibodies against active signalling molecules is a direct method to visualize the signalling activity in the developing embryo on a single cell level.

With this approach, the activity of bone morphogenic protein receptors (BMPR) by using an antibody against phosphorylated Smad1, Smad5 and Smad8 (Smad1/5/8) was checked, as these are known to be BMPR downstream effectors activated via phosphorylation; then, with use of co-immunostainings as before, the results were compared directly with the *Foxa2* expression in the embryo.

At early-streak stage, active Smad1/5/8 expressing cells were detected in epiblast cells of the posterior and posterior-to-distal region of the embryo as well as in VE cells on the surface of the embryo (Figure 41 a-d). Later, at mid- to late-streak stage, phospho-Smad1/5/8 (p-Smad1/5/8) expressing cells were found in the surface layer of visceral endoderm in the whole embryo. P-Smad1/5/8

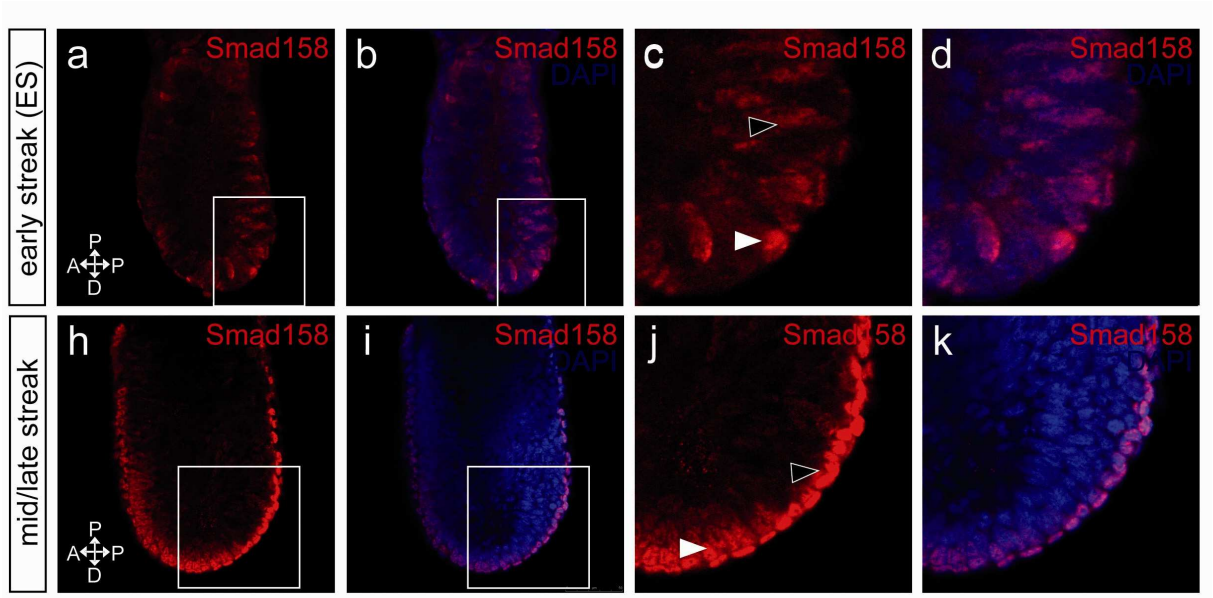


Figure 41 Anti active-Smad1/5/8 staining on early and mid/late-streak stage embryo
(a, b) Anti active-Smad1/5/8 staining on early-streak embryo shows active Smad1/5/8 expressing cells in the posterior and posterior-distal region of the embryo. (c, d) Active Smad1/5/8 expression is localized to epiblast cells (black arrowhead) and cells of the outer layer of visceral endoderm (white arrowhead). (h, i) Anti active-Smad1/5/8 staining on mid/late-streak stage embryo shown Smad expressing cells in the outer cell layer of the whole embryo. (j, k) In the proximal region the active Smad1/5/8 expression is restricted to cells of the outer cell layer (black arrowhead), while in the distal and anterior region underlying epiblast cells are also active Smad1/5/8 positive (white arrowhead) . Pictures: a-k; single planes.

positive epiblast cells could be detected in the distal and anterior, but not in the posterior, regions of the primitive streak (Figure 41 h, j).

For a more detailed analysis, co-immunostainings with an anti-*Foxa2* antibody were performed. *Foxa2*, as endodermal marker for definitive endoderm, allows the exact cellular localisation of the active Smad1/5/8 expression in the embryo. As described before, at mid-streak stage, *Foxa2* was expressed in the VE and DE localized at the posterior side of the embryo (Figure 42 a; white arrowhead). A single plane confocal allowed to distinguish between the thick one cell layer epithelium of visceral endodermal cells the squamous epithelium of the definitive endodermal cells (Figure 42 e´; white arrowhead, black arrowhead).

Higher magnifications of the same plane showed visceral endodermal cells and epiblast cells positive for p-Smad1/5/8 labelling both in distal and anterior sides of the embryo (Figure 42 e´´-j, white arrowhead, black arrowhead); on the contrary, epiblast cells expressing p-Smad1/5/8 were not detectable on the posterior side of the embryo, below the *Foxa2* positive definitive endodermal cells (Figure 42 f´, g´; white arrowhead, black arrowhead).

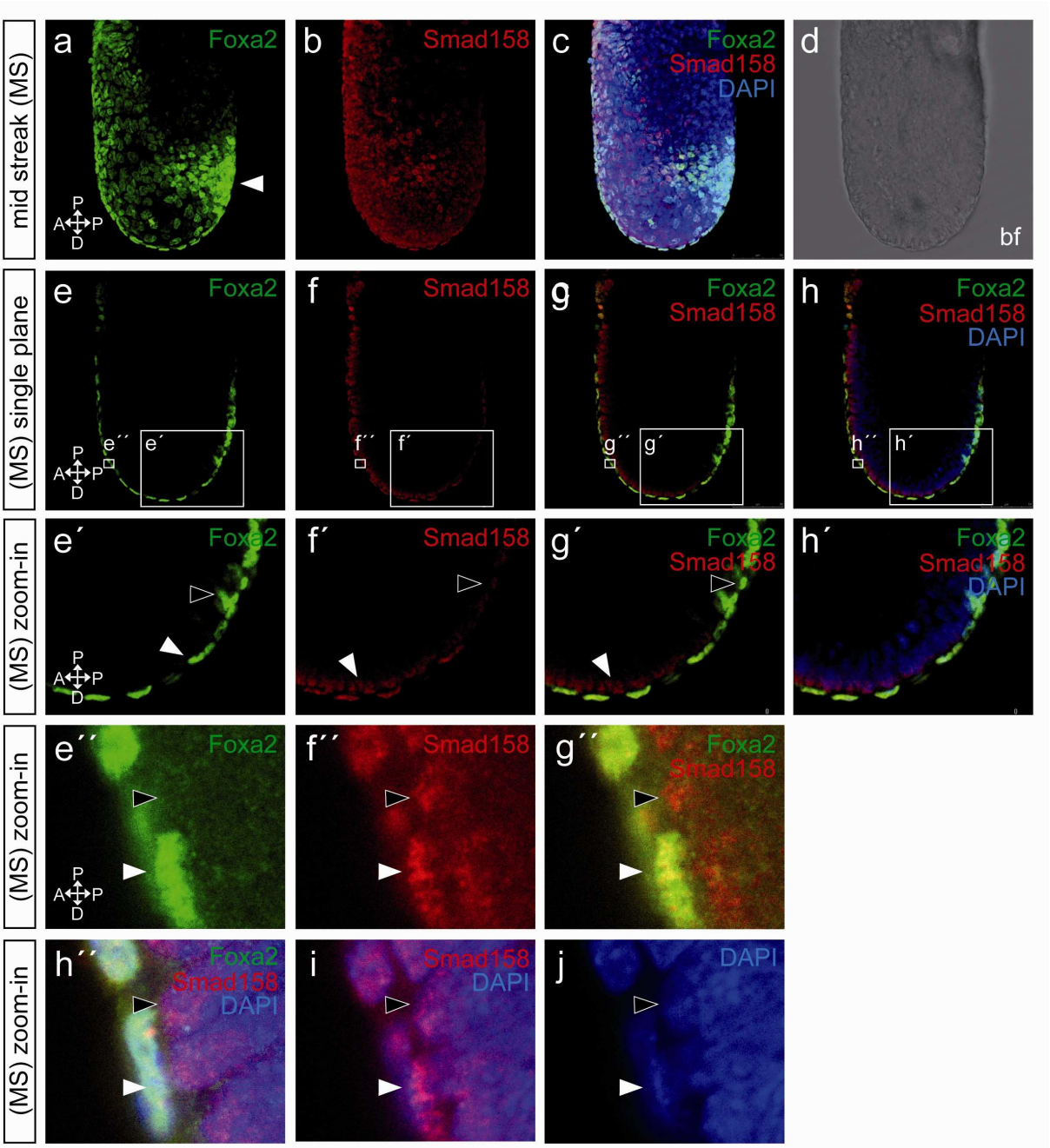


Figure 42 *Foxa2* and active Smad 1/5/8 antibody staining on mid-streak stage embryo
(a) Staining of a mid-streak stage embryo with a *Foxa2* antibody shows *Foxa2* expression in definitive endodermal cells (white arrowhead) in the primitive streak region and cells of the visceral endoderm. (b) Staining against active Smad1/5/8 shows active Smad1/5/8 shows expressing cells in the posterior-distal, distal and anterior region of the embryo. (c) Overlay of stainings. (d) Brightfield picture. (e, e´) Single plane of *Foxa2* staining shows *Foxa2* expressing epiblast cells (black arrowhead) overlaying definitive endodermal cells in the posterior region as well as *Foxa*2 expressing visceral endodermal cells in the distal and anterior region (white arrowhead). (f, f´) Single plane of active Smad1/5/8 expression shows expression just in the outer cell layer in the posterior region (black arrowhead) and expression in outer layer and epiblast cells in the distal and anterior region (white arrowhead). (g, g´) Overlay shows that in the primitive streak region, characterized by *Foxa2* expressing epiblast cells (black arrowhead) active Smad1/5/8 is only expressed in the outer cell layer of definitive endodermal cells. In other parts of the embryo, active Smad1/5/8 is also expressed in epiblast cells. (e´´-g´´) Higher magnification of anterior region shows localisation of active Smad1/5/8 in the *Foxa2* expressing cells of the visceral endoderm (white arrowhead) and in epiblast cells (black arrowhead). Pictures: a-c; merge of planes; d-j; single planes.

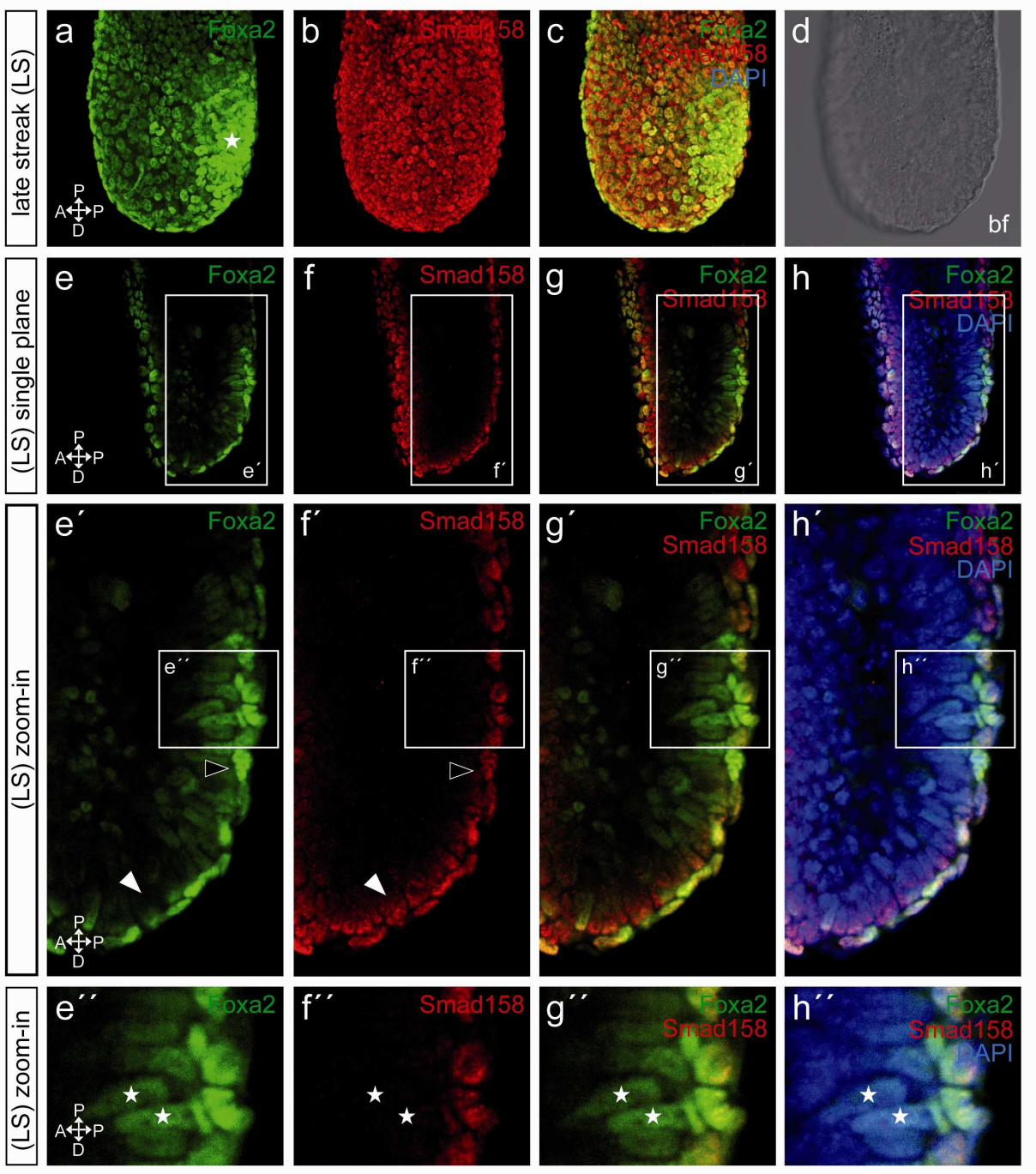


Figure 43 *Foxa2* and active Smad 1/5/8 antibody staining on late-streak stage embryo
(a) *Foxa2* staining on late-streak stage embryo shows *Foxa2* expressing definitive endoderm (DE) cells and visceral endoderm (VE) cells. (b) Active Smad1/5/8 staining shows expression in the whole embryo. (c) Overlay of stainings. (d) Brightfield picture. (e, e´, e´´) *Foxa2* expressing epiblast cells in the primitive streak region (e´´, asterisk) but not in distal region (white arrowhead). (f, f´, f´´) Smad1/5/8 staining shows expression in the surface cell layer of definitive endodermal cells in the primitive streak (f´´) but not in the underlying epiblast cells (f´´; asterisk). Expression of active Smad1/5/8 in epiblast cells in the distal region (f´; white arrowhead). (g, g´, g´´, h, h´, h´´) Overlays show absence of active Smad1/5/8 expressing epiblast cells but presence of *Foxa2* expressing epiblast cells in the primitive streak region. Pictures: a-c; merge of planes; d-h´´; single planes.

Specifically, in this region, where *Foxa2* positive cells could be found in the epiblast, p-Smad1/5/8 signal was absent and even showed a reduced expression in the definitive endoderm (Figure 42 f´, g´; white arrowhead, black arrowhead).A late-streak stage embryo showed a similar expression pattern (Figure 43). *Foxa2* positive definitive endodermal cells (Figure 43 a, asterisk; e, e´, black arrowhead; e´´, asterisk) in the epiblast did not show active Smad1/5/8 expression, which was restricted to the outer cell layer in this region (Figure 43. f´´, white arrowheads). In contrast, in distal regions of the embryo, active Smad1/5/8 in inner layers of the epiblast, under visceral endodermal cells were found (Figure 43 f´, white arrowhead).

At early bud stage *Foxa2* as well as active Smad1/5/8 expressing cells were restricted to the embryonic part of the embryo (Figure 44 a-c). The expression of active Smad1/5/8 and *Foxa2* in epiblast cells was not any more detectable (Figure 44 e´- h´).

The co-immunostainings for active Smad1/5/8 together with *Foxa2* showed the absence of Smad1/5/8 in epiblast cells which undergo the EMT to endodermal cell fate. During this step of development BMP signalling seems to be downregulated.

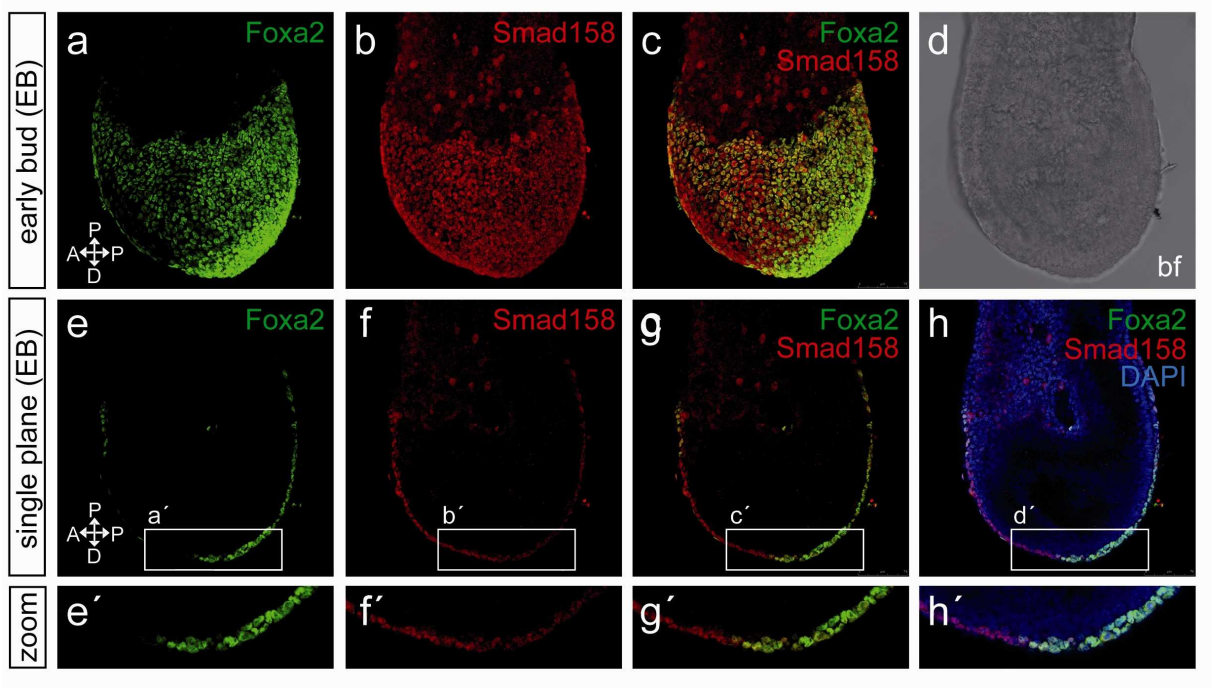


Figure 44 *Foxa2* and active Smad1/5/8 antibody staining on early bud stage embryo
(a) *Foxa2* staining on early bud stage embryo shows *Foxa2* expressing definitive endoderm (DE) cells and visceral endoderm (VE) cells. (b) Active Smad1/5/8 staining shows expression in the whole embryonic part of the embryo. (c) Overlay of stainings. (d) Brightfield picture. (e, e´) *Foxa2* expressing definitive endoderm (DE) cells. (f, f´) Active Smad1/5/8 expression visible only in the outer cell layer of the embryo (g, g´, h, h´). Overlay of *Foxa2* and active Smad1/5/8 expression.

3.10 Whole mount embryo stainings using antibodies against active Smad1/5/8 and *Brachyury* (*T*)

After the study of co-immunolabelling of active Smad1/5/8 with *Foxa2* in the embryo, whole mount antibody stainings were used to show the possible colocalization of active Smad1/5/8 with *Brachyury* (*T*). *T* is necessary for posterior, but not anterior mesoderm formation (Wilkinson *et al*., 1990). Gene functional analysis showed that *T* null mutants lack posterior mesoderm and notochord (Wilkinson *et al*., 1990; Kispert and Herrmann, 1994).

An embryo at late-streak stage, stained for active Smad1/5/8 and Brachyury showed still *T* expression in the distal and posterior region (Figure 45 a, e, e´) in both the surface cell layer and in epiblast cells. The expression of active Smad1/5/8 was restricted to the distal and anterior side of the embryo in the surface

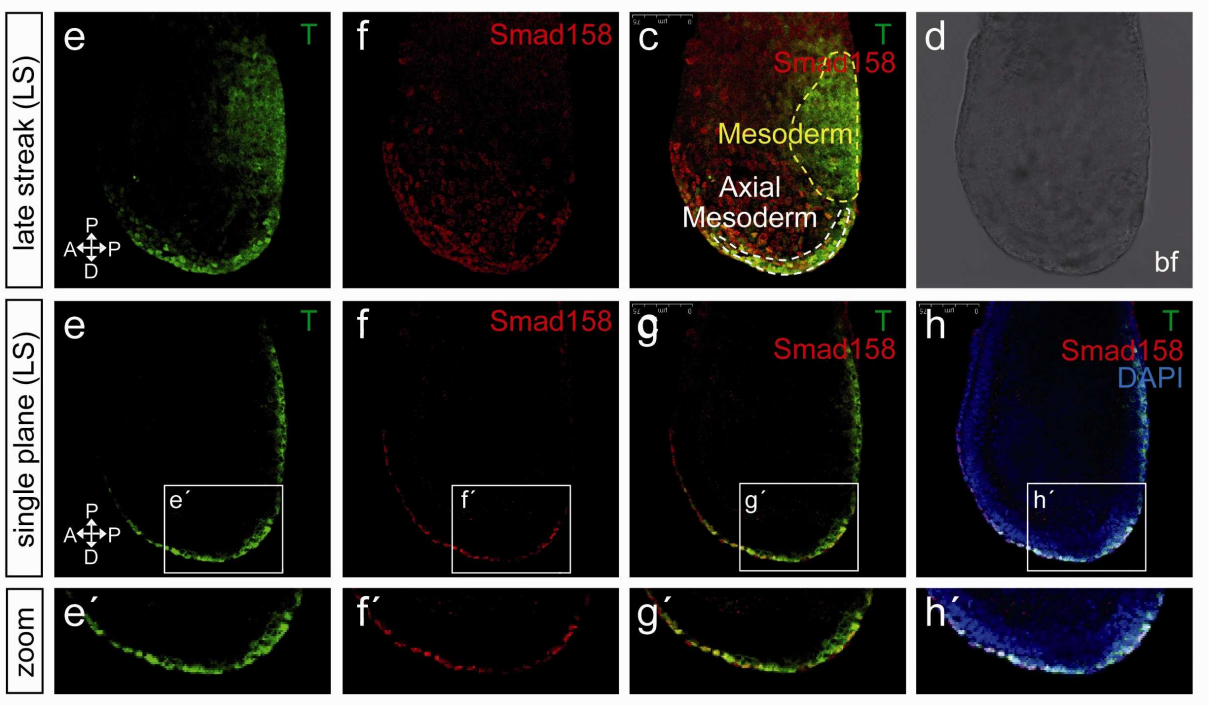


Figure 45 *Brachyury* (*T*) and active Smad1/5/8 antibody staining on late-streak stage embryo
(a) *T* staining on late-streak stage embryo shows *T* expressing cells on the posterior side and distal posterior side of the embryo in the axial mesoderm. (b) Active Smad1/5/8 staining shows expression in the distal and anterior side of the embryo. (c) Overlay of stainings. (d) Brightfield picture. (e, e´) Single plane pictures show *T* expression on the posterior and distal side of the embryo. (f, f´) Active Smad1/5/8 expression on the distal and anterior distal side of the embryo. (g, g´, h, h´) Overlay pictures showing colocalization of *T* and active Smad1/5/8 expression in the distal part of the embryo with an expression of active Smad1/5/8 in the surface cell layer and *T* expression in the surface cells layer and in epiblast cells..

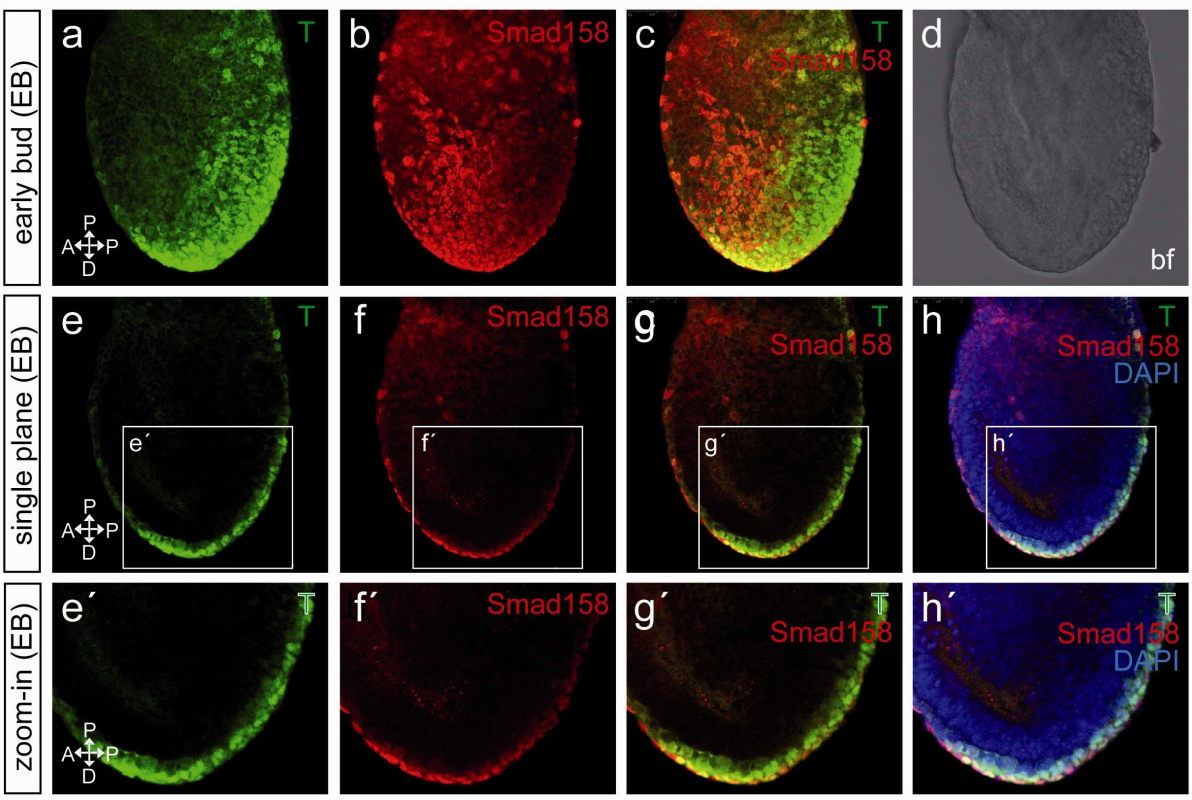


Figure 46 *Brachyury* (*T*) and active Smad1/5/8 antibody staining on early bud stage embryo
(a) *T* staining on early bud stage embryo shows *T* expressing cells on the posterior and distal posterior side of the embryo. (b) Active Smad1/5/8 staining shows expression in the whole embryo with a strong expression on the posterior and distal side. (c) Overlay of stainings. (d) Brightfield picture. (e, e´) Single plane picture shows *T* expression on the posterior and distal posterior side of the embryo. (f, f´) Active Smad1/5/8 expression at the posterior and distal side of the embryo. (g, g´, h, h´) Overlay pictures showing colocalization of *T* and active Smad1/5/8 expression in the posterior and distal posterior side of the embryo with an expression of active Smad1/5/8 in the surface cell layer and *T* expression in the surface cells layer and in epiblast cells.

cell layer (Figure 45 b, f, f´). Therefore, overlay pictures could show coexpression only in the distal region of the embryo (Figure 45 c- h´). Later, at early bud stage, *T* expression could be observed in the distal and posterior-distal region of the embryo (Figure 46 a, e, e´).

The expression of active Smad1/5/8 at this stage was detectable in the whole embryo, with a stronger expression in the distal tip of it (Figure 46 b, f, f´). Thus, the overlay of the two signals showed an overlap only in the cells of the outer layer in the distal region (Figure 46 c-h´).

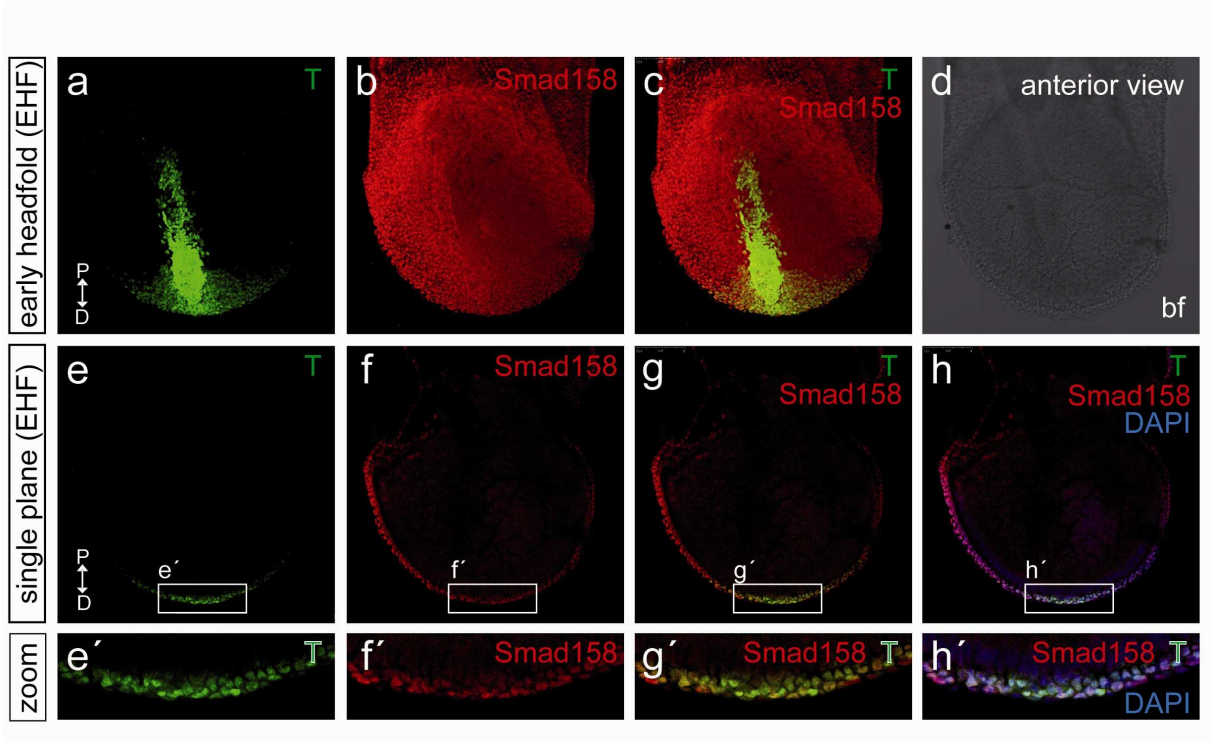


Figure 47 *Brachyury* (*T*) and active Smad1/5/8 antibody staining on early headfold stage embryo
(a) *T* staining on early headfold stage embryo shows *T* expressing cells in the midline and at the distal tip of the embryo. (b) Active Smad1/5/8 expression in the whole embryo. (c) Overlay of stainings. (d) Brightfield picture. (e, e´) Single plane picture shows *T* expression in the distal tip of the embryo. (f, f´) Active Smad 1/5/8 expression in the distal and lateral region of the embryo. (g, g´, h, h´) Overlay picture shows *T* expression in the surface cell layer and epiblast cells and the expression of active Smad1/5/8 in the surface cell layer.

The *T* expression at early headfold stage was restricted to the primitive streak and the node (Figure 47 a, e, e´), while active Smad1/5/8 was expressed in the whole embryo (Figure 47 b, f, f´); so the overlay of both could be observed in all *T* expressing cells (Figure 47 c-h´).

4 DISCUSSION

4.1 The efficiency of BAC transgene technology by bacterial recombineering and the following analysis of BAC transgenic ES cells *in vitro* and *in vivo* demonstrates advantages of BAC transgenes

Since the generation of the first bacterial artificial chromosome (BAC) in 1992 during the Human Genome Project (Shizuya *et al*., 1992), BAC technology, developed rapidly and is now used in a wide array of applications. The lack of upstream and downstream regulatory elements of genes in generated by small promoter-reporter constructs transgenic lines results in an incomplete transgene expression which can be overcome by BAC transgenes as the average transcription unit for mammalian genes is less than 100 kb (Nielsen *et al*., 1997; Yang *et al*., 1997; Antoch *et al*., 1997; Gong *et al*., 2003; Heintz, 2004).

BAC transgenes can emulate the native environment of a gene resulting in a tightly regulated pattern even if randomly inserted in the genome. It was shown that the large genomic clones of BACs (of ~100-300 kb) often contain DNA sequences which are able to mantipulate chromatin at the local level and so become immune to position effects which affect expression of smaller transgenes (Nistala and Sigmund, 2002). The first publication showing the mapping of gene expression at the cellular level using modified BACs was from Yang *et al*. (1997). The possibilities of multi-copy number insertions of BACs has become more important for the introduction of fluorescent reporters in the genome since this amplifies the signal of the reporter and enables live imaging. This enables for the first time to visualize active processes during embryonic development *in vivo*, live and on cellular level, which was not possible before (Burtscher and Lickert, 2009).

The results of the bacterial recombineering in this work demonstrate that BAC recombineering is an elegant and very efficient method to modify BACs. 90% of the BACs were successfully processed at the single modification steps and their integrity was maintained (Figure 49). The further retrofitting step of introducing a new selection marker was with 19% not as effectice as the other modification steps. The

pre-sreening of clones by PCR for positives overcomes this and confirms the published data of this step (Yang and Seed, 2003).

Overall the BAC modification in bacteria in the described way is very effective and can be done in large scale for many genes in parallel and / or different loci of one gene (Gong *et al*., 2003; Valenzuela *et al*., 2003). The required time to produce a higher number of BAC transgenic genes and/or to target different loci of the genes is short if different mini-targeting vectors are already availiable and just the homology arms have to be exchanged. As expected, the establishing of BAC technology e.g. BAC handling in the lab was time consuming as BACs are not yet part of the molecular repertoire of most of the labs. If the details and important points of handling and protocol procedures are known the used protocol can be nicely reproduced. The Group of Prof. Dr. Götz, Dr. Schröder and a group at the TU München are successfully reproducing the used techniques in total or in parts for their approaches and verifies the here described results.

The generation of BAC transgenic ES cells by preparation of modified BAC DNA, linearization and electroporation in ES cells was found to be delicate. Unil now, the majority of transgenic mice were produced by labourious, expensive and technically demanding pronuclear injections of BAC DNA in fertilized oocytes.

To enable a pre-screening of the BAC transgenic ES cells *in vitro* and *in vivo* before establishing a mouse line we choose the way of generating BAC transgenic ES cells. For electroporation of BAC DNA in ES cells it is necessary to have a great amount of high purified and intact BAC DNA and the right conditions for the electroporation. The evaluation and establishment of the ideal preparation method with a high yield and an acceptable effort of time and ressources was time consuming.

Finally an upscaling of the BAC mini protocol with a few modifications was used. The number of produced ES clones was unless very low due to the size of the DNA vector which is up to 40 times larger than in knock-ins and other (non BAC-) transgenes. The effort for ES cell culture work should also not be underestimated due to the time consuming passage of each single ES cell clone. The Southern analysis of the BAC transgenic ES clones showed that 18% of the clones were positive for the fluorescent transporter which is a very high rate of positives compared to knock-ins (Results 3.2; Figure 48). This although demonstrated that around 80% of the integrated BACs were not completely integrated in the genome

and damaged during the DNA preparation procedure or the electroporation in ES cells.

The results showed just in two cases (both *Hhex* clones; Figure 15) a multi-copy instertion of the BACs in the genome which indicated that these incidents are not occuring so frequently. Although we changed the (PGK) promoter which drives the resistsance gene to give selection to G418 in ES cells to a weaker (tk) promoter to select for multiple copy insertion of BACs. The chosen strategy did not perform as well as planned. The complete integration of the BAC in the ES cell genome was successfully demonstrated for all 9 tested clones by BAC end PCR (Results 3.3, Figure 48).

step \ construct	*Hex::Lyn-Tomato*	*Sox17::H2B-Tomato*	*Foxa2::Lyn-Venus*	*Nkx2.1::H2B-Venus*	*Pdx1::H2B-Venus*	percentage	numbers (n=)
recombineering	95%	90%	95%	85%	85%	90%	n=100
retrofitting	20%	17%	23%	13%	20%	19%	n=150
ES clones	5/19	2/10	4/14	4/14	0/19	18%	n=76
BAC end PCR	2/2	2/2	3/3	2/2		100%	n=9
tetraploid clones	2/2	2/2	3/3	2/2		100%	n=9
mouse line	1/1	1/1				100%	n=2

Figure 48 Statistic of BAC transgenes
The numbers of intact BAC clones for the steps of recombineering (homologous recombination step and *Neo* deletion step) and retrofitting are given in percent. The total numbers of Southern positive ES clones/ Southern tested ES clones are given in percent. The total numbers of tested clones/ BAC end PCR positive clones are given. The total numbers of BAC clones tested in tetraploid complementation/ BAC clones showed correct transgene expression in tetraploid-derived embryos. Numbers of generated mouse lines/ correct transgene expression in the generated mouse line.

The BAC end PCR verification was as fast and efficient method which guaranteed that all following in vitro and in vivo expression analysis steps were done with completely and intact integrated BACs in the ES genome and approves the result of Yang and Seed (2003). The step of *in vitro* expression analysis of BAC transgenic

ES cells by differentiation of BAC transgenic ES cells in endoderm was performed as pre-check before *in vivo* expression analysis in tetraploid-derived embryos or a BAC transgenic mouse line. This step of differentiation of BAC transgenic ES cells in endoderm was never done before.

Nevertheless during the production of this work, the ES cell differentiation of BAC transgenic ES cells to neural cell fate was done by Tomishima *et al*. (2007) and supports the chosen procedure. Recently established techniques of differentiating of ES cells in endoderm (Kubo *et al*., 2004; Tada *et al*., 2005) allowed the *in vivo* pre-screening of the BAC transgenic ES cells for transgene expression in endoderm, a rapid and time- and resource-saving test for *in vivo* expression. As *Hhex* is already expressed in undifferentiated ES cells in culture (Zamparini *et al*., 2006), two transgene expressing *Hhex* clones with multi-copy insertions of the transgene could be easily identified, as known that the transgene expression level corresponds to the number of inserted transgenes (Gebhard *et al*., 2006).

As expected, the two *Hhex* clones showed also expression in ES cells, differentiated in endoderm cell fate (Results 3.4.2). All three tested *Foxa2* BAC transgenic ES clones showed a weak transgene expression *in vitro* even with only one copy insertion in the genome (Results 3.2.1 b and 3.4.3). Slightly differences in the expression intensity of the clones might be due to different insertion loci in the ES cell genome. Both tested *Sox17* BAC transgenic ES clones showed transgene expression in ES cells, differentiated in endoderm (Results 3.4.4). In total, the *in vitro* analysis showed correct transgene expression from 7 of 10 tested BAC transgenic ES cell clones and enabled beyond a selection of very promising candidated clones for *in vivo* expression analysis.

Unfortunately, the system was not useful to analyse BAC transgenic ES cells of *Nkx2.1* and *Pdx1* as they are expressed only from endodermal organ formation on, and the differentiated ES cells could not reach this developmental stage with the used protocol (Tada *et al*., 2005; Kubo *et al*., 2004).

The *in vivo* expression check of BAC transgenic ES cells was performed by tetraploid complementation technique which allowed a fast *in vivo* testing of the candidate clones for expression in the endoderm. The two *in vitro* tested *Hhex::Lyn-Tomato* BAC transgenic ES cell clones showed in tetraploid-derived embryos transgene expression in the definitive endoderm which was consistent with the reported *Hhex* mRNA expression and expression in *Hhex* transgenic lines (Results 3.5; Thomas *et*

al., 1998; Bogue *et al*., 2000; Rodriguez *et al*., 2001; Mesnard *et al*., 2006; Morrison *et al*., 2008). The establishment of a BAC libary of the most important CNS-expressed genes of the mammalian brain using BAC transgenes with fluorescent reporter genes was performed in the Heintz lab (Gong *et al*., 2003; Heintz *et al*., 2004). This was the first large scale effort to combine *in situ* hybridization and BAC transgenic methods for expression analysis of 250 genes. There was shown that the increased sensitivity of the reporter genes particulary in lines carrying multiple copies of the BAC transgene may allow the detection of sites of expression that are not detectable in *in situ* hybridizazion. These observations are consistent with results, shown in this work.

The *Hhex::Lyn-Tomato* transgene expression in the ICM at blastocyst stage and the expression in the limbs may be due to this reason which causes differences between the BAC expression data sets and the *in situ* data. Three of three in vitro tested *Foxa2::Lyn-Venus* BAC transgenic ES clones showed transgene expression in definitive endodermal cells of tetraploid-derived embryos, consistent with the published mRNA expression of *Foxa2* (Ang *et al*., 1993; Monaghan *et al*., 1993; Sasaki and Hogan, 1993).

The *in vivo* expression of two *Nkx2.1::H2B-Venus* BAC *transgenic* ES cell clones was also tested in tetraploid complementation technique derived embryos. Due to the single copy integration of the BAC reporter in the genome (Results 3.2.1 d) it was not possible to detect a fluorescent signal by the mean of confocal microscopy. The visualization was also aggravated due to the late timepoint of *Nkx2.1::H2B-Venus* transgene expression in the embryo which complicated the detection in the thicker tissue of the lung buds at E11.5, as it is surrounded by mesenchyme. Using whole mount immunohistochemistry the *Nkx2.1::H2B-Venus* transgene could be detected in epithelial cells of the lung (Results 3.5.1 c) as reported for the endogenous Nkx2.1 mRNA expression from Lazzaro *et al*. (1991). The *in vivo* expression analysis using tetraploid complementation technique reflected in 7 of 7 tested BAC transgenic ES cell clones the published gene expression. This *in vivo* expression test demonstrated in a very convincing way that the BAC integrity checks done before and the *in vitro* tests of the candidate clones had been the right strategy to select for accurate reporter expressing BAC clones.

In the large scale BAC transgene libary screen (Gong *et al*., 2003; Heintz, 2004) were BAC transgenic lines for 250 genes produced. It was demonstrated that ~85%

of the produced BACs expressed reproducibly in multiple transgenic lines while ~15% of BACs did not show reproducibly caused by position effects on the integration site. The generated BAC transgenic lines of this work showed reproducibly in the *in vitro* and *in vivo* experiments. The teraploid derived embryos showed that a one copy insertion of the transgene in the genome can be enough for detection of the reporter in gastrulation stage embryos as shown for *Foxa2*, depending on the gene and its endogenous regulation level. For *Hhex*, the multi-copy insertion of the transgene showed expression, strong enough for live imaging for several hours (shown by Dr. Burtscher).

In the literature (Gong *et al*., 2003) was reported that in many cases insertions of at least 5 copies of the fluorescent reporter genes are required for a sufficient signal strength for the detection of the fluorescent reporter. In this work only the *Hhex::Lyn-Tomato* transgene consisted of multi copy insertions in the genome while the other transgenic lines consisted of only one copy insertions. In this work could be shown that fluorescent transporter cells can be easily detected in BAC transgenic lines with one copy insertions. This may be due to improvements of the fluorescent proteins.

Another reason could be the timepoint of analysis at this early stage of development which might facilitate the detection. But one copy insertion of the transgene seems to be too weak for detection at later stages of development (older than E8.0) due to physical imaging limitations caused by the increased size of the specimen, the limited depth of the focus level and the increased thickness of the surrounding tissue. This demonstrated that BAC reporter ES cell lines can be used for studies to test accurate reporter expression *in vitro* and *in vivo* and for live imaging approaches.

Multiple purposes to generate genetically modified mice by knock-ins or BAC transgenes demonstrated another not to neglect advantage of BAC transgenes. The living offspring of chimeras after germline transmission can be directly used to analyse transgene expression instead of time-consuming crossings of knock-ins for deletion of the *Neo* resistance. Knock-ins, constructed in the lab showed different expression of the knock-in before and after deletion of the *Neo* cassette and hypomorphic effects. One of the biggest advantages of BAC transgenes, the ability to visualize a low number of gene expressing cells and to follow them during the development clould be shown in this work and is coincident with Gong *et al*. (2003).

4.2 A *Hhex::Lyn-Tomato* BAC transgenic mouse line reflects the accurate reporter expression and shows some new aspects of *Hhex* expression in the mouse

The *Hhex::Lyn-tomato* transgenic mouse line was generated for studying *Hhex* gene expression and lineage differentiation in the developing mouse embryo (Results 3.6). The found transgene expression at blastocyst stage embryos between E3.5 and E3.75 in the ICM which might correspond to PrE progenitors in the ICM was never reported before and may be due to the regulatory elements contained in the BAC transgenic line which are not present in existing transgenic lines (Results 3.6). As previously reported, undifferentiated ES cell cultures contain subpopulations of cells corresponding to ICM, PrE and epiblast (Toyooka *et al.*, 2008); this might indicate that the *Hhex::Lyn-Tomato* expression in undifferentiated ES cells could correspond to the PrE subpopulation.

Instead, the expression at the following stages of development, at E5.5 in the distal tip of the embryo in the VE and during gastrulation in the VE, posterior mesoderm and ADE corresponded to the *Hhex* expression reported before (Thomas *et al*.,1998; Rodriguez *et al*.,2001). Therefore *Hhex* transgene expression corresponded to the already reported *Hhex* expression. From E14.5 to E18.5 *Hhex::Lyn-tomato* transgene expression could be detected in the radius of the forelimb and the tibia of the hindlimb. This was also never reported before and indicated that *Hhex* could play a role in limb bud patterning and radius and tibia formation. This has to be confirmed by mRNA *in situ* hybridization which did not show utilizable results until now. The expression analysis of the *Hhex* transgene in organs at E16.5 and E18.5 confirmed the reported expression in liver and pancreas (Bogue *et al*., 2000).

The results indicated that the *Hhex::Lyn-Tomato* reporter line accurately reflected the endogenous mRNA expression pattern. As shown, the *Hhex::Lyn-Tomato* mouse line is working nicely to study *Hhex* transgene expression in static analysis and in live imaging approaches for lineage allocation, cell population analysis and cell migration approaches.

Compared to the reported *Hhex* mRNA expression in ISH, the new generated *Hhex::Lyn-Tomato* mouse line shows a more homogenous expression pattern (Thomas *et al*., 1998; Bogue *et al*., 2000; Rodriguez *et al*., 2001).

It was not possible to generate a homozygous *Hhex::Lyn-Tomato* mouse line by crossing heterozygote animals. This might be due to the multi-copy insertion of the *Hhex* BAC in the genome (see Figure 16). The doubled copy number and its corresponding transgene expression might be too high in homozygote animals and therefore interfere with cellular viability due to cytotoxic effects. That is also supported by the failure to generate a ubiquitous expressing H2B-Tomato mouse line in the lab. The overexpression of the other genes, which are localized on the *Hhex* BAC might also have some negative effects on cell viability.

4.3 The *Hhex* BAC transgenic mouse line shows a homogenous and detailed expression pattern in comparison to a Hhex-GFP transgenic line

The direct comparison of the Hhex-GFP transgenic mouse line and the *Hhex::Lyn-Tomato* BAC transgenic mouse line by intercrosses displayed the advantages of BAC transgenes compared to minimal transgenic reporters. The Hhex-GFP transgenic line was constructed of a 8 kb genomic fragment with 4.2 kb upstream and 3.8 kb downstream sequence of the ATG while the *Hhex::Lyn-Tomato* BAC transgenic line based on a BAC with a size of 180 kb and 54 kb upstream and 122 kb downstream genomic sequence of the ATG.

As reported, a slightly differerent and more detailed and homogenous expression pattern from the BAC transgenic mouse line was observed, due to the presence of most or all regulatory elements on the BAC as it was whown in the large scale BAC transgene libary screen (Gong *et al*., 2003; Heintz, 2004) in which BAC transgenic lines for 250 genes were produced, that the average transcription unit for mammalian genes is less than 100 kb so that ~85% of the produced BACs expressed reproducibly in multiple transgenic lines (Yang *et al*., 1997; Antoch *et al*., 1997; Gong *et al*., 2003; Heintz, 2004).

The differences of expression of the *Hhex* lines could be directly visualized by co-expression of both transgenes in the same embryo (Results 3.7.1). There the minimal transgene Hhex-GFP showed, a spotty and patchy expression pattern, when compared to the *Hhex* BAC transgenic line. The strong expression domain in the anterior proximal region of the *Hhex* BAC transgenic line was in the Hhex-GFP line slightly decreased and very patchy, due to the lack of some regulator elements.

So the direct comparison strongly showed the advantages of BAC transgenes if an accurate transgene expression is required and expected form a transgenic line for characterizing the gene expression.

4.4 A *Sox17* BAC transgenic mouse line showns restricted transgene expression in the definitive endoderm

The *Sox17::H2B-Tomato* transgenic mouse line was generated from BAC transgenic ES cells which were checked for BAC integrity and differentiated in endoderm were the cells showed successful transgene expression *in vitro* (Results 3.4.4). The generated mouse line showed only a very weak expression in the DE at E7.5 (Results 3.8.1) but not at earlier and later stages of development, as reported for endogenous *Sox17* expression (Kanai *et al*., 1996; Matsui *et al*., 2006). To explain this, it was useful to compare the results of the BAC transgenic *Sox17* mouse line with a $Sox17^{iCre}$ (iCre: improved causes recombination, Shimshek *et al*., 2002) mouse line which was produced from Perry Liao by knock-in of the iCre in the same locus of exon 1 as in the *Sox17* BAC transgenic line (Liao *et al*., Genesis, in press).

This *Sox17* iCre line showed expression in the endoderm vessel endothelium restricted to arteries from E9.5, but hardly showed any recombination activity at earlier stages in the endoderm germ layer. The analysis of this result revealed that a previously unknown promoter region also exists which lead to an alternative mRNA of *Sox17* coding for a transcript expressed in the endoderm. The mRNA which codes for the transcript, which is expressed by promoters of exon 1 leads to *Sox17* expression mostly restricted to the arterial endothelium (Liao *et al*., 2008, Genesis, in press).
Considering these new aspects the arteries of the mouse brain between E9.5 and E12.5 were tested for transgene expression of the *Sox17::H2B-Tomato* as shown in the *Sox17* iCre line. Nevertheless, it was not possible to detect transgene expression in bood vessles due to limitided possibilities of fluorescent microscope imaging and the fact of a one copy insertion of the BAC in the genome (Results 3.2.1). The expression pattern in the iCre line showed just single positive stained cells in the blood vessels of the brain. The visualization of single positive tomato reporter cells in

the blood vessles of the brain under the covering skin with laser scanning confocal imaging was not possible.

The *Sox17::H2B-Tomato* transgenic mouse line with the described attributes appeard to be not useful for characterisation of the *Sox17* endoderm population and lineage allocation studies during embryonic development. The occurred expression effects are caused by the unpredicted and unknown alternative promoter. These arguments are in favour of the BAC transgenic strategy which allows targeting of different loci of a gene to avoid the described problems. Using BAC technology, this results not in extensively more work because just the homology arms of the mini-targeting vector have to be exchanged.

4.5 Whole mount *Foxa2* staining on *Hhex::Lyn Tomato* BAC transgenic mouse embryos enable the mapping of populations on cellular level

Whole mount antibody *Foxa2* stainings on *Hhex::Lyn-tomato* transgenic mouse embryos enable the visualization and mapping of the populations in cellular resolution and the determination of subpopulations. The forkhead transcription factor *Foxa2* is required for anterior axial mesoderm and definitive endoderm formation (Ang and Rossan, 1994; Weinstein *et al*., 1994).

Using whole mount immunohistochemistry (IHC) *Foxa2* expression was detected in the one cell layer overlaying VE and the multi layer DE which was easy to distinguish by morphology and localization in the embryo (Results 3.9.1). The overlay with the *Hhex* transgene expression enabled the identification of *Foxa2* expressing AVE cells that were also positive for *Hhex* transgene expression. The *Hhex* transgene expression could be clearly localized to the membrane and just be found in the outer cell layer of the embryo. By using the *Foxa2* antibody staining it was nicely possible to determine the *Hhex* transgene expression in the ADE as well as the strong *Hhex* transgene expression in the anterior extraembryonic VE of the embryo.

This expression domain, clearly visible in the early bud stage and early headfold stage (Results 3.9.1) appeared in the BAC transgenic line more evident and with a more homogenous expression pattern than in the *Hhex*-GFP line (Results 3.7.1, 3.7.2 and Discussion 4.3). Based on this detailed visualization of expression a region could be identified on the proximal end of the midline at early headfold stage where

Foxa2 expression was absent, but the *Hhex* transgene was highly expressed. This was a completely new observation, which was never reported before in the literature. This region was previously described as AIP in a fate mapping study using dye labelling by Tremblay and Zaret (2005) and suggested to contribute for liver bud progenitors.

This indicates that *Foxa2* has to be downregulated for driving endodermal cells in hepatic cell fate. The result coincides also with the verified role of *Hhex* as earliest known effector of the liver (Zaret, 2002; Bort *et al*., 2004; Bort *et al*., 2005; Calmont *et al*., 2006; Martinez-Barbera *et al*., 2000).

The observed expression of *Foxa2* in whole mount stainings coincides with the reported *Foxa2* expression in the mouse embryo (Ang *et al*., 1993; Monaghan *et al*., 1993; Sasaki and Hogan, 1993). The *Hhex::Lyn-Tomato* transgene localization to the membrane was in whole mount stainings more evident than in the live imaging scans (shown in Results 3.6). This might be due to the multiple washing steps performed during the immunohistochemical procedure, which could have reduced the amount of Tomato reporter protein in the cytoplasm. The expression pattern of the *Hhex::Lyn-Tomato* transgene allowed a much more detailed analysis of expression under the *Hhex* promoter than described before (Thomas *et al*., 1998; Bogue *et al*., 2000; Rodriguez *et al*., 2001).

4.6 Whole mount *Foxa2, T* and active Smad1/5/8 staining in gastrulation stage embryos

The intention was to perform whole mount antibody staining of gastrulation stage embryos to generate a map of active signalling pathways. For this approach antibodys against active β-catenin (as key component of the Wnt-Signallimg pathway), active Smad 2/3 (for TGF β/Nodal signalling) and active Erk1/2 (for FGF signalling) were tested in whole mount stainings. Unfortunately none of the tested antibody performed positive in the whole mount staining procedure.

Studies in *Xenopus* and chick suggest that BMP signalling is a “right determinant” (Capdevila *et al*., 2000). Recently was shown that the type I receptor ACVRI (ALK2) for BMPs play important roles in left-right patterning of the mouse (Gu *et al*., 1999; Mishina *et al*., 1999). As there is no evidence that *Nodal* is required for left-right patterning it seems that BMP signalling through ACVRI in the node regulates *Nodal*

expression in the node as an upstream factor. The visualization of active BMP signalling during gastrulation in the embryo was performed using phosphorylated antibodies against the active receptor substrates Smad1, Smad5 and Smad8 (Smad1/5/8). Using this approach active Smad1/5/8 expressing cells could be observed in the posterior distal and distal region of the embryo at early-streak stage in the VE and in epiblast cells (Results 3.10.1).

Surprisingly, at later stages of development, active Smad1/5/8 expressing visceral endodermal cells and underlying epiblast cells could be found in the distal and anterior region but just a weak expression of active Smad1/5/8 expressing cells could be found in *Foxa2* expressing DE cells and no expression could be found in the underlying epiblast cells (Results 3.10.2, 3.10.3). In this way, the downregualtion of active Smad1/5/8, active BMP-signalling in the region of endoderm formation could be clearly demonstrated.

The *Foxa2* shown here expression corresponds with the reported *Foxa2* expression at early-streak stage in the posterior epiblast, anterior definitive endoderm (ADE) and later in the axial mesoderm, prechordal plate, notochord and node (Sasaki and Hogan, 1993; Monaghan *et al*., 1993).

The expression of Smad1 from E6.5 in the VE and during gastrulation in the VE and DE on was shown by Tremblay *et al*. (2001). The data shown in this work show the cells of active Smad1/5/8 signalling in the VE and DE during gastrulation with cellular resolution. The model from Kishigami and Mishina (2005) of BMP function prior to and during gastrulation shows that the BMP receptor 1A signalling in the ExE and VE is critical for the initiation of gastrulation. Later on this is important for the proper development of the AVE. The T-box transcription factor *Brachyury* (*T*) is necessary for posterior, but not anterior mesoderm formation (Wilkinson *et al*., 1990). *T* was used as marker for mesoderm in whole mount antibody stainings to show the active Smad 1/5/8 expression in the embryo towards the mesoderm. The stainings showed an overlap of expression in early stages of development as *T* and active Smad1/5/8 were expressed in the posterior and distal regions of the embryo. *T* expression was found in mesodermal cells and epiblast cells while the expression of active Smad1/5/8 was mostly restricted to the outer cell layer. The shown results are consistent with the published expression of *T* and the conditional allel of BMP receptor BMPR1A during gastrulation (Miura *et al*., 2006). In contrast to Miura *et al*. (2006) where the *T* and BMPR expression is shown by *in situ* hybridization, the here

presented data show *T* and the active BMP receptors through their activated downstream targets.

The observed expression of *T* during gastrulation and at later stages of development corresponded very well with the already published expression (Inman and Downs, 2006).

5 MATERIALS

5.1 Instruments

autoradiography cassettes	Amersham, Hypercassette (Amersham, GE Healthcare GmbH, München)
Balance	Scout Pro (Ohaus, Pine Brook (NJ), USA)
centrifuges	Eppendorf 5417R (Rotor F45-30-11) Eppendorf 5804R (Rotor F34-6-38) (Eppendorf AG, Hamburg)
confocal microscopes	Olympus FV1000 (Olympus, Tokio, Japan) Leica DMI 6000 (Leica Microsystems GmbH, Wetzlar)
crosslinker	Spectrolinker XL-1500 (Spectronics Corporation, Westbury (NY), USA)
developing machine	Agfa, Curix 60 (Agfa AG, München)
digital camera	AxioCam MRc, AxioCam ARc (Carl Zeiss AG, Göttingen AG, Göttingen)
dissecting microcope	Leica MS5 (Leica Microsystems GmbH, Wetzlar, Wetzlar, Germany)
DNA sequencer	Applied Biotech, DNA Analyzer 3730 (Applied Biotech, San Diego (CA), USA)
electroporator	Gene-Pulser Xcell BioRad (BioRad, München)
freezer (-20℃)	(Liebherr GmbH, Ochsenhausen)
freezer (-80℃) (Sanyo MDF)	(Sanyo, Osaka, Japan)
fridge (4℃)	(Liebherr GmbH, Ochsenhausen)
gel chambers	Gelelektrophorese chambers Mini, Midi, Large (Harnischmacher, Kassel)
gel documentation system	Gel Flash BioRad (BioRad, München)

gel power supply	Thermo EC250 (Thermo, Waltham (MA), USA)
glassware	(Schott AG, Mainz, Germany)
hybridization oven	Hybridization oven SN (Thermo, Waltham (MA), USA)
hybridization bottles	ThermoHybaid (Thermo, Waltham (MA), USA)
incubator (32℃, bacteria)	Incubator hood TH30 (Edm und Bühler, Hechingen)
incubator (37℃, bacteria)	Shaking incubator S16 Sh el Lab (Cornelius (OR), USA)
incubator (37℃, cells)	(Heareus, Hanau, Germany)
ice machine	Scotsman, AF 30 (Scotsman, Vernon Hills (IL), USA)
inverse microscope	Zeiss Axiovert 200M (Carl Zeiss AG, Göttingen)
laminar flow	Flow Microflow 514 (Nunc, Roskilde, Denmark)
magnetic stirrer / heater	Heidolph, MR3001 (Heidolph, Kelheim)
microwave	(Severin, Sundern)
nanodrop	ND-1000 Spectrophotometer (Nanodrop, Wilmington (DE), USA)
Neubauer counting chamber	(Brand, Gießen)
PCR cycler	Thermo Hybaid Px2, Thermo Thermal cycler PxE0.2 (Thermo, Waltham (MA), USA)
pH-meter	Hanna pH211 (Hanna, Leighton Buzzard, UK)
photometer	BioPhotometer (Eppendorf AG, Hamburg)
pipetteboy	Easypet (Eppendorf AG, Hamburg,)
pipettes	Research fix (Eppendorf AG, Hamburg)
roller incubator	Stuart STR1 (Stuart, Staffordshire, UK)
rotator	Grant Boeckel VSR23 (Grant Boeckel,

	Cambridge, UK)
shaker	Eppendorf Thermomixer comfort (Eppendorf AG, Hamburg)
stereomicroscope	Zeiss Lumar V12 (Carl Zeiss AG, Göttingen)
szintillation counter	Hidex, Triathler (Hidex, Turku, Finland)
UV-tansilluminator	(Biocom, Berlin)
vortex	(VWR, Darmstadt)
water barth	Elmeco PB WB-10 (Elmeco, Lorton (VA), USA)

5.2 Chemicals

All used chemicals were delivered from AppiChem GmbH, Darmstadt; Beckton Dickinson GmbH, Heidelberg; Biozym GmbH, Hessisch Oldendorf; Carl Roth GmbH & Co. KG, Karlsruhe; Merck KaGaA, Darmstadt and Sigma-Aldrich GmbH, Hamburg.

5.3 Kits

Kit	Company
DNA Maxi Prep Kit	Qiagen, Hilden
Gel extraction Kit	Qiagen, Hilden
Large construct Kit	Qiagen, Hilden
PCR Purification Kit	Qiagen, Hilden
Prime-It Random Prime Labeling Kit	Stratagene

5.4 Commonly used stock solutions

Solution	Composition
loading buffer for agarose gels	15% Ficoll 400
	200 mM EDTA
	2% orange G
PBS (1x)	4.3 mM Sodium Phosphate

	137 mM Sodium Chloride
	2.7 mM Potassium Chloride
	1.4 mM Potassium Phosphate, monobasic
	adjust to pH 7.4 with HCl
P1: Resuspension buffer	50 mM Tris (pH 8.0)
	10 mM EDTA (pH 8.0)
	0.5 mg/ml RNaseA
P2: Lysis buffer	200 mM NaOH
	1% SDS
P3: Neutralization buffer	3 M Potassium acetate (pH 5.5),
	adjust with glacial acetic acid
TAE buffer (10x)	0.4 M Tris base
	0.1 M acetate
	0.01 M EDTA
TBE buffer (10x)	0.89 M Tris base
	0.89 M boric acid
	0.02 M EDTA
TE buffer (Tris-EDTA)	10 mM Tris-HCl pH 8.0
	1 mM EDTA
Lysis buffer for mouse tails	100 mM Tris (pH8.0)
	5 mM EDTA (pH8.0)
	2% SDS
	200 mM NaCl
	10 mg/ml Proteinase K

5.5 Solutions for the work with bacteria

Solution	Composition	
LB agar	98.5 % LB medium	
	1.5 % bacto agar	
	pH 7,5	
LB (Luria-Bertani) medium	10 g	Bacto peptone
	5 g	yeast extract

	5 g NaCl ad to 1l with H_2O
LB^{amp} medium / agar	LB medium / agar with 100 µg/ml ampicillin
LB^{cam} medium / agar	LB medium / agar with 20 µg/ml chloramphenicol
LB^{kan} medium / agar	LB medium / agar with 20 µg/ml kanamycin
$LB^{amp/kan}$ medium / agar	LB medium / agar with 50 µg/ml ampicillin and 20 µg/ml kanamycin

5.6 Solutions for cell culture

Solution	Composition
Differentiation medium in endodermal cells	50 ml SFO3 serum free media (Sanko Junyaku, Tokyo, Japan) 50 µl human recombinant Activin A (Sigma-Aldrich GmbH, Steinheim) 50 µl ß-mercaptoethanol (50mM, Gibco, Carlsbad (CA), USA) 500 µl Pen/Strep (100x, Gibco, Carlsbad (CA), USA)
ES cell freezing medium (2x)	4 ml ES cell medium 4 ml FCS 2 ml DMSO
ES cell lysisbuffer for 96-well plate	10 mM Tris (pH 7.5) 10 mM EDTA 10 mM NaCl
ES cell medium	DMEM (Dulbecco´s Modified Eagle Medium, Invitrogen, Carlsbad (CA), USA) 15% fetal calf serum (FCS, PAN, Aidenbach) 20 mM Hepes (100x, Invitrogen,

	Carlsbad (CA), USA) 2 mM L-glutamine (Invitrogen (200mM), Carlsbad (CA), USA) 1x nonessential amino acids (Invitrogen (100x), Carlsbad (CA), USA) 100µM β-mercaptoethanol (Invitrogen (50 mM), Carlsbad (CA), USA) 1500 U/ml LIF (Leukaemia inhibitory factor, (Chemicon, Billerica (MA), USA)
Feeder Medium (for MEFs)	DMEM (Dulbecco´s Modified Eagle Medium, Invitrogen, Carlsbad (CA), USA) 15% fetal calf serum (FCS, PAA, Pasching, Austria) 2 mM L-glutamine (Invitrogen (200mM), Carlsbad (CA), USA) 1x nonessential amino acids (Invitrogen (100x), Carlsbad (CA), USA) 100µM β-mercaptoethanol (Invitrogen (50 mM), Carlsbad (CA), USA)

5.7 Solutions for embryo work

Solution	Composition
Dissecting media (for live imaging)	500 ml DMEM (Dulbecco´s Modified Eagle Medium), without phenol red (Invitrogen 31053-028, Carlsbad (CA), USA)) 15% foetal calf serum (FCS) (PAA A15-102, Pasching, Austria) (20 minutes at 65℃ heat inactivated) 20 mM 1M Hepes (Gibco 15630, Carlsbad (CA), USA) 2 mM L-Glutamin 100x (Gibco 25030,

	Carlsbad (CA), USA) 50.000 Units/ml (50.000 µg/ml) Penicillin-Streptomycin (Gibco 15140, Carlsbad (CA), USA) 100 µM β-Mercaptoethanol (Gibco 31350-010, Carlsbad (CA), USA) 1mM Sodium pyruvate (Sigma-Aldrich GmbH, Steinheim)
Dissecting PBS	Solution A (20x) 348 mM Na_2HPO_4 70 mM NaH_2PO_4
	Solution B (20x) 18 mM $CaCl_2$ 70 mM KCl 18 mM $MgCl_2$ 2740 mM NaCl 1:1 mix of Solution A and B in H_2O
embryo culture medium	50% rat serum 40% DMEM without phenol red (Invitrogen 31053-028) 2 mM glutamine 1 mM sodium pyruvate 100 µM β-mecaptoethanol
Paraformaldehyd solutuion (PFA, 4%)	4% PFA in PBS

5.8 Solutions for Southern blot analysis

Solution	Composition
Hybridisation solution	1 M NaCl 50 mM Tris (pH 7.5) 10% dextransulfat 1% SDS

	250 µg/ml salmon sperm DNA (sonified)
SSC buffer (20x)	3M NaCl
	0.3 M sodium citrate – 2 H_2O
	pH 7.0
Washing solution I	2 x SSC
	0.1 % SDS
Washing solution II	0.2 x SSC
	0.1 % SDS

5.9 Solutions for immunochemistry

Solution	Composition
Permeabilization buffer	0.1% Triton X-100
	100 mM glycerin
	in PBS
TBST	20 mM Tris-HCl (pH8.0)
	150 mM NaCl
	0.05% Tween-20

5.10 Enzymes

All used Restriction endonucleases were delivered from New England Biolabs GmbH, Frankfurt am Main and Fermentas GmbH, St. Leon-Rot.

5.11 Antibodies

Antibody (ID)	**Organism**	**Dilution**	**Company (Number)**
Anti-Brachyury (T) (N-19) (25)	goat polyclonal	IC 1:250	Santa Cruz (sc17743), Santa Cruz (CA), USA
anti-Active-ß-Catenin (3)	mouse monocl. IgG	IC 1:250	Millipore (05-665), Billerica (MA), USA
anti-HNf-3ß (Foxa2) M-20 (23)	goat polyclonal	IC 1:1000	Santa Cruz (sc-6554), Canta Cruz (CA), USA

anti-*gfp* (7)	rabbit IgG	IC 1:1000	Invitrogen (A11122), Carlsbad (CA), USA
anti-phospho-Smad 2/3 (58)	goat polyclonal	IC 1:250	Santa Cruz (sc-11769), Canta Cruz (CA), USA
Secondary Antibody	**Conjugated**	**Dilution**	**Company (Number)**
rabbit anti-goat IgG 488 (4)	fluorescent	1:800	Invitrogen (A11078), Carlsbad (CA), USA
rabbit anti-goat IgG 555 (5)	fluorescent	1:800	Invitrogen (A21431), Carlsbad (CA), USA
goat anti-mouse 488 (7)	fluorescent	1:800	Invitrogen (A11029), Carlsbad (CA), USA
goat anti-rabbit 594 (8)	fluorescent	1:800	Invitrogen (A11037), Carlsbad (CA), USA
goat anti-mouse IgG 546 (12)	fluorescent	1:800	Invitrogen (A11030), Carlsbad (CA), USA
goat anti-rabbit IgG 488 (13)	fluorescent	1:800	Invitrogen (A11234), Carlsbad (CA), USA
goat anti-mouse IgG Cy3 (14)	fluorescent	1:800	Dianova (115-165-003), Hamburg, Germany
goat anti-mouse IgG 405 (17)	fluorescent	1:800	Invitrogen (A31553), Carlsbad (CA), USA
donkey anti-goat IgG 633 (18)	fluorescent	1:800	Invitrogen (A21082), Carlsbad (CA), USA
donkey anti-mouse IgG 488 (23)	fluorescent	1:800	Invitrogen (A21202), Carlsbad (CA), USA
donkey anti-rabbit IgG 555 (24)	fluorescent	1:800	Invitrogen (A31572), Carlsbad (CA), USA
donkey anti-goat IgG 488 (25)	fluorescent	1:800	Invitrogen (A11055), Carlsbad (CA), USA
donkey anti-rabbit IgG 488 (26)	fluorescent	1:800	Invitrogen (A21206), Carlsbad (CA), USA

5.12 Bacterial strains

Strain (description; genotype)	Source
E. coli K12 DH5α F- φ80lacZΔM15 Δ(□lacZYA-argF)U169 recA1 endA1 hsdR17(r_k^-, m_k^+) phoA supE44 thi-1 gyrA96 relA1 λ^-	Hanahan *et al*., 1985
E. coli K12 DH10B F^- mcrA Δ (mrr–hsdRMS–mcrBC) Ø80dlacZΔ M15 Δ lacX74 deoR recA1 endA1 araD139 Δ (ara, leu)7649 ga/U ga/K rspL nupG	Grant *et al*., 1990
E. coli K12 EL250 Derived from DY380 (DH10B-derived strain). The bacteria contain a defective λ prophage with recombination proteins exo, bet, and gam being controlled by the temperature-sensitive repressor cI857. Genotype DY380: F- mcrA Δ(mrr-hsdRMS-mcrBC) Φ80dlacZ M15 ΔlacX74 deoR recA1 endA1 araD139 Δ(ara, leu) 7649 galU galK rspL nupG [λcI857 (cro-bioA) <> tet]. EL250 also contains a tightly controlled arabinose-inducible flpe gene. flpe will mediate recombination between two identical frt sites. Genotype EL250: DY380 [(cro-bioA) <> araC-PBADflpe]	Lee *et al*., 2001

5.13 BACs

All used BACs are from the RPCI-23 BAC library (Osoegawa et al. 2000) and derived from the DNA of a female C57BL/6J mouse. The DNA of this libary was digested with a combination of EcoRI and EcoRI methylase and cloned between the pBACe3.6 vector (Frengen et al. 1999) at the EcoRI sites. BACs were obtained from RZPD (Heidelberg, Germany).

BAC ID	Included gene
RPCI-23-117H14	Hhex
RPCI-23-285G23	Sox17
RPCI-23-3H18	Nkx2.1
RPCI-23-254G2	Foxa2
RPCI-23-355P12	Pdx1

5.14 Vectors and plasmids

Vector	Source
PL-451 FRT-PGK-EM7-NeobpA-Frt-loxP	Lui *et al*., 2003
pBluescript KS(-)(pKS$^-$) vector	Invitrogen, Carlsbad (CA), USA
CG8967, CT25769, DTrk, Dtkr, IGGtyk, OFT/TRK, Tk48D, Trk48D, anon-WO2004063362.81	

5.15 DNA Ladders, modifying enzymes

Ladder, modifying enzyme	Company
100 bp ladder	NEB, Frankfurt am Main
1 kb ladder	NEB, Frankfurt am Main
T4 DNA Ligase	NEB, Frankfurt am Main

5.16 Oligonucleotides

The Oligonucleotides were done by Mr. Linzner (Helmholtz Zentrum München) and at Metabion (Martinsried).

Purpose	Primer	Sequence	Product	Conditions
Hhex 5´ hom. arm	Hhex Ex1A fw	5´- NNNGAGCTCATGCATCGGTC GCAGTCGCTCGCTAGGTGCC – 3´ (76℃)		95℃ 2min 95℃ 30s 67℃ 45s 72℃ 45s
	Hhex Ex1A rev	5´- NNNGCGGCCGCTCCGCAGCC CCTCGCAGAGAGC – 3´ (77℃)	322 bp	30 cycles 72℃ 5min
Hhex 3´ hom. arm	Hhex Ex1B fw	5´-NNNGCGGCCGCGAATTC GAACGACTACACGCACGCC CTACTCC - 3´ (72℃)		95℃ 2min 95℃ 30s 67℃ 45s 72℃ 45s
	Hhex Ex1B rev	5´- NNNCAAGCTTGTTGAACCA AGCCCAGCCCAG – 3´ (73℃)	463 bp	30 cycles 72℃ 5min
Nkx2.1 5´ hom. arm	Nkx2.1 Ex1A fw	5´- NNNCCGCGGATGCAT GTCGTCGGCTTTGTGTTG AGGGAGTAG - 3´ (70℃)		95℃ 2min 95℃ 30s 65℃ 45s 72℃ 45s
	Nkx2.1 Ex1A rev	5´- NNNGCGGCCGCCTCATTGCTG TTACCAGGTCAGGATC – 3´ (70℃)	466 bp	30 cycles 72℃ 5min
Nkx2.1 3´ hom. arm	Nkx2.1 Ex1B fw	5´ -NNNGCGGCCGCGAATTC GCCGAAGTATCCAGGCCATG GGCAAGG - 3´ (74℃)		95℃ 2min 95℃ 30s 69℃ 45s 72℃ 45s
	Nkx2.1 Ex1B rev	5´- NNNAAGCTTGACCACATCGGG CTTCGCTGCGCTGAG – 3´ (76℃)	477bp	30 cycles 72℃ 5min
Sox17 5´ hom. arm	Sox17 Exon1A fw	5´- NNNGAGCTCGATCTG CTTG AGTGCCCACGGATCCTGTGC - 3´ (75℃)		95℃ 2min 95℃ 30s 68℃ 45s 72℃ 45s
	Sox17 Exon1A	5´- NNNGCGGCCGCCGC CAGCAGTGTGAGAGGGC CATATTTCAG - 3´ (73℃)	322 bp	30 cycles 72℃ 5min
Sox17 3´ hom. arm	rev Sox17 Exon1B fw	5´- NNNGAATTCCACTCCTCCCA AAGTATCTATCAAGAGAATG – 3´ (66℃)		95℃ 2min 95℃ 30s 61℃ 45s

				72℃ 45s
	Sox17	5´-NNNAAGGTTGCATTTTCTCT		30 cycles
	Exon1B	GTCTTCCCTGTCTTGGTTG – 3´		
	rev	(69℃)	463 bp	72℃ 5min
Foxa2	Foxa2	5´- NNNGAGCTCATGCATT		95℃ 2min
5´hom.	Exon1A	GTAGCCCGGATGTCTTGAA		95℃ 30s
arm	fw	ACTCACTC - 3´ (69℃)		66℃ 45s
				72℃ 45s
	Foxa2	5´-NNNGCGGCCGCGGTCGTC		30 cycles
	Exon1A	AGTTACCTCAGTCCTCTTTC - 3´		
	rev	(66℃)	322 bp	72℃ 5min
Foxa2	Foxa2	5´- NNNGAATTCCACTCCTCCCA		95℃ 2min
3´hom.	Exon1B	AAGTATCTATCAAGAGAATG – 3´		95℃ 30s
arm	fw	(66℃)		61℃ 45s
				72℃ 45s
	Foxa2	5´- NNNAAGCTTCGGGCAGC		30 cycles
	Exon1B	CCATTTGAATAATCAGC – 3´		
	rev	(68℃)	463 bp	72℃ 5min
Pdx1	Pdx	5´- NNNGAGCTCCACAGCAGCAA		95℃ 2min
5´hom.	Ex 1A	GCAGGGATCAGGCGACTG – 3´		95℃ 30s
arm	fw	(71℃)		66℃ 45s
				72℃ 45s
	Pdx			30 cycles
	Ex 1A			
	rev	5´- NNNGCGGCCGCGGTGGCAGC	328 bp	72℃ 5min
		CGGCACTTGGGGGCCAGCAG – 3´		
		(79℃)		
Pdx1	Pdx	5´- NNNGAATTCGTAAGCC		95℃ 2min
3´hom.	Ex 1B	TGGCTCTTTCT – 3´ (58℃)		95℃ 30s
arm	fw			53℃ 45s
				72℃ 45s
	Pdx	5´ - NNNCTCGAGCCGTTTCTCT		30 cycles
	Ex 1B			
	rev	AATAAGGCCACTA – 3´ (61℃)	447 bp	72℃ 5min
Retro-	Retro-R	5´-ATCGACCGGTAAT	328 bp	94℃ 3min
fitting		GCAGGCA - 3´ (66℃)	(integrati	
check			on in	94℃ 45s
	F1	5´- TCAGCGTGAGACT	*lox511*)	56℃ 45s
		ACGATTC – 3´ (61℃)	367 bp	72℃ 45s
			(integrati	35 cycles
	F2	5´ - GTTGCTACGCCTG	on in	
		AATAAGTG – 3´ (61℃)	*loxP*)	72℃ 5min

5´BAC-end PCR	5´fw pBAC e3.6	5´- GAGTGAATACCACG ACGATTTCC – 3´ (63℃)		95℃ 3min 95℃ 45s 58℃ 45s 72℃ 45s
	5´rev pBAC e3.6	5´- TCACAAACGGCAT GATGAACC – 3´ (63℃)	272 bp	35 cycles 72℃ 5min
3´BAC-end PCR	3´Sce fw	5´– GTAGACTTAATTAAG GATCGATCC – 3´ (58℃)		95℃ 3min 95℃ 30s 54℃ 45s 72℃ 60s
	3´Sce rev	5´- CTCTCCCTATAG TGAGTC – 3´ (53℃)	167 bp	35 cycles 72℃ 5min
Hhex::Lyn-Tomato geno-typing	T3 for	5´– CACTGCATTCTAGTTG TGGTTTGTCC – 3´ (65℃)		95℃ 5min 95℃ 30s 59℃ 60s 72℃ 60s
	H5 rev	5´- GATAAAATAATCAGCAG CGTGCACTACTCC – 3´ (64℃)	713bp	35 cycles 72℃ 5min
Sox17::H2B-Tomato geno-typing	Neo rev EP271	5´–CTTCGCCCAATAGC AGCCAGTCC – 3´ (69℃)		95℃ 5min 95℃ 30s 64℃ 60s 72℃ 60s
	Neo fw EP272	5´-GGCGCGGTCCCA GGTCCAC – 3´ (70℃)	320bp	35 cycles 72℃ 5min

5.17 Computer programs

Software	Company
Leica Microsystems GmbH, Wetzlar	Leica Microsystems GmbH, Wetzlar
LAS AF	
Olympus FW10-ASW 1.6	Olympus, Tokio, Japan
Carl Zeiss AG, Göttingen Axiovision 4.6	Carl Zeiss AG, Göttingen

5.18 Plastic ware and other material

Material	Company
cell culture dishes	Nunc, Roskilde, Denmark
centrifugation tubes (15 ml, 50 ml)	BD Falcon, Becton Dickinson GmbH,

	Heidelberg
cleaning columns MicroSpin S-300 (Southern blot)	Amersham, GE Healthcare GmbH, München
coverslips	Menzel Gläser
cuvettes for electroporation (0.4 cm)	BioRad GmbH, München
Eppendorf reaction tubes	Eppendorf AG, Hamburg
films for autoradiography Kodak, Biomax MS, MR	Sigma-Aldrich GmbH, Hamburg
filter paper Whatman 3 mm	GE Healthcare Buchler GmbH & Co. KG
filter tips (10 µl, 200 µl, 1000 µl)	Starlab, Hamburg
gloves	Kimberly-Clark, Hamburg
Nylon membrane for Southern transfer Amersham Hybond N Plus	Amersham, GE Healthcare GmbH, München
Pasteur pipettes	Brand, Gießen
PCR tubes	Biozym GmbH, St. Leon-Rot

6 METHODS

6.1 Isolation and Purification of Nucleic Acids

6.1.1 DNA isolation small scale (plasmid mini-prep protocol)

A 2 ml overnight LB culture was pelleted for 1 minute at 14000 r.p.m., the supernatant removed and the pellet was dissolved in 350 µl buffer P1 (Qiagen, Miniprep Kit) and transferred to an Eppendorf tube. An aliquot of 350 µl P2 was added followed by mixing and incubated for up to 5 minutes at room temperature. An aliquot of 350 µl P3 was added and mixed by inversion. The supernatant was cleared by centrifugation at 14000 r.p.m. for 10 minutes in a tabletop centrifuge. The DNA precipitation was performed by addition of 630 µl isopropanol (1:0.7) to 900 µl supernatant and centrifugation for 10 minute at 14000 r.p.m.. The pellet was washed with 500 µl 70% ethanol in a 5 minutes 14000 r.p.m. centrifugation step. Residual ethanol was removed and the DNA pellet air dried up to 10 minutes. The DNA was dissolved in 100 µl TE buffer on a shaker at 750 r.p.m. for 1 hour.

6.1.2 DNA isolation large scale (maxis)

Large scale DNA plasmid preparations were done using the Qiagen Maxi Kit (Qiagen AG, Hilden) following manufacturers instructions.

6.1.3 Isolation of BAC DNA small scale (minis)

A 5 ml overnight LB culture was pelleted for 5 minutes at 5000 r.p.m., the supernatant removed and the pellet was dissolved in 250 µl buffer P1 (Miniprep Kit, Qiagen AG, Hilden) and transferred to an Eppendorf tube. An aliquot of 250 µl P2 was added followed by mixing and incubated for up to 5 minutes at room temperature. An aliquot of 250 µl P3 was added, mixed by inversion and incubated on ice for 5 minutes. The supernatant was cleared two times by centrifugation at 14000 r.p.m. for 5 minutes in a tabletop centrifuge. The DNA precipitation was

performed by addition of 750 µl isopropanol (1:1), mixing by inversion, incubation on ice for 10 minutes and centrifugatuion for 10 minutes at 14000 r.p.m.. The pellet was washed with 500 µl 70% ethanol in a 5 minutes 14000 r.p.m. centrifugation step. Residual ethanol was removed and the DNA pellet air dried up to 5 minutes. The DNA was dissolved in 50 µl TE buffer on a shaker at 750 r.p.m. for 1 hour (protocol modified from Warming *et al*., 2005).

6.1.4 Isolation of BAC DNA large scale (maxis)

The BAC Miniprep protocol was upscaled 40 times so that the volume of the overnight LB culture was 200 ml. The protocol was similar as described for BAC Minipreps until the drying of the DNA pellet. But the centrifugation steps were done in 50 ml Falcon tubes in Eppendorf 5804R centrifuge at 9500 rpm; and the centrifugation times were extended of about 50% of the time for miniprep.
To prepare BAC DNA for electroporation all steps from the ethanol removal step on were done under steril conditions to maintain sterile conditions. The pellet was scratched out of the Falcon tube with a spatula and transferred in an Eppendorf tube which contained 500 µl of cell culture grade wate (Invitrogen GmbH, Karlsruhe). The tube was incubated for 2 hours at 55 ℃ in an Eppen dorf Thermomixer, it was then gently shaking with 500 rpm for resuspension of the DNA pellet and stored at 7℃ till use.
Alternatively, BAC DNA large scale preparation was performed using the Large Construct Kit (Qiagen AG, Hilden) following manufacturer´s instructions.

6.1.5 Phenolextraction-chloroform extraction

The purification of DNA was performed in 15 ml Falcon tubes (BD Biosciences, San Jose (CA), USA) by addition of two times 1 volume phenol-chloroform-isoamylalcohol (ratio 25:25:1; Carl Roth GmbH & Co. KG, Karlsruhe) to the DNA (soluted in TE or H_2O) followed by a 30 seconds vortexing and 15 minutes centrifugation step at 14000 rpm. The upper phase was discarded and 1 volume chloroform (Carl Roth GmbH & Co. KG, Karlsruhe) was added, mixed and centrifugated as described above. After the removing of the upper phase 2.5 volumes of 100% Ethanol 1/10 volume (3M) Sodium Acetat were added and incubated for 30 minutes at -20℃. For sedimentation

the mixture was spun 15 minutes at 14000 rpm at 4℃. The DNA was washed two times using 70% Ethanol in centrifugation steps of 5 minutes at 14000 rpm followed by a 5 minutes air dry step and resuspension in PBS, TE or H_2O.

6.1.6 Isolation of DNA from mouse tails

Mouse tails were collected stored in autoclaved Eppendorf tubes and stored at –20℃ until use. For DNA isolation 500 µl of Lysisbuffer supplemented with 10mg/ml ProteinaseK (Invitrogen GmbH, Karlsruhe) are added and each tube incubated overnight at 55℃. The day after, tubes were vortex ed and then centrifugated for 10 minutes at 14.000 r.p.m. to sediment tail hair. The supernatant was transferred in new Eppendorf tube prepared with 500 µl isopropanol. The DNA was precipitated by inverting the tube 2-3 times and transferred in an Eppendorf tube with 500 µl dd$H_2$0 using a 200 µl pipet tip.

6.1.7 DNA purification from Agarose gels

For the isolation of DNA fragments of a specific size from agarose gels, the bands were visualized on long wave UV radiation (366 nm) to prevent damage of the DNA. The band of interest was cut out with a clean scalpel and elution of DNA was performed using the Qiagen Gel Extraction Kit (Qiagen AG, Hilden) following manufacturer´s instructions.

6.1.8 Purification of DNA fragments from PCR

For the isolation of PCR amplified DNA fragments, the Qiagen PCR purification Kit (Qiagen AG, Hilden) was used following manufacturer´s instructions.

6.1.9 Quantification of Nucleic acids

The concentrations of nucleic acids were determined by measuring the optical densitiy (OD) in a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Waltham (MA), USA) using 1 µl sample normalized to 1 µl TE or H_2O following manufacturer´s instructions.

6.2 Molecular biological methods on nucleic acid

6.2.1 Gel electrophoresis (Agarose gel, BAC agarose gel)

DNA Agarose (Biozym Scientific GmbH, Hessisch Oldendorf) was soluted in 1xTAE buffer (add of Ethidiumbromid 50µg/ 50 ml gel, Carl Roth GmbH & Co. KG, Karlsruhe). Gels were done with 1%, PCR-gels 1.5% and BAC gels 0.7% agarose. The volume was 50ml or 100ml, BAC gels 500ml and they were run at 80 – 100V, BAC gels at 40V overnight.

6.2.2 Restriction digestion of Plasmid DNA

1 µg DNA was digested using 1 unit restriction endonuclease with the recommended 10 x reaction buffer at the recommended temperature for 1 – 2 hours. BAC DNA was digested overnight.

6.2.3 Restriction digestion of BAC DNA

BAC Mini DNA (50 µl volume) for the check of BAC integrity and the recombeneering steps was digested overnight using 20 units of the corresponding restriction enzyme in a total volume of 60 µl.

6.2.4 Digest of BAC DNA for electroporation

The BAC DNA was prepared for electroporation by using the Large Construct Kit (Qiuagen AG, Hilden) or by upscaling of the BAC mini prep protocol (40x) and resuspended in 500 µl TE buffer under a cell culture hood. The digestion was done by addition of 40U PI-*Sce*I (NEB), 60 µl PI-SceI buffer and 6 µl BSA and incubation overnight at 37℃.

6.2.5 Ligations

The ligations were done with a ration of 1:2.5 of vector (200 ng DNA) to insert using 1 µl T4 DNA ligase (NEB) and 1 µl ligase buffer in total volume of 10 µl. The ligations were performed overnight at 16℃.

6.2.6 Tansformation of Plasmid DNA

For transformation of DNA 50 µl chemical compentent *E.coli* DH5α were incubated with 1 µl of the ligation mix for 30 minutes on ice, 2 minutes incubated on 42℃ and transferred for 30 – 60 minutes in LB medium at 37°C. Then cells were spun for 1 minute at 5000 rpm the supernatant removed and the cells plated out.

6.2.7 Polymerase-chain reaction

The PCR mix contained 2.5 U T7 DNA polymerase (NEB), 200 µM dNTPs, 1µM Primer (each), 100 – 1000 ng genomic DNA, 10x buffer, 2 mM $MgCl_2$ and was filled with H_2O up to 20 µl.
The standard settings were: denaturation for 2 - 5 minutes at 95℃ (1 cycle), denaturation for 30 – 60 seconds at 95℃, annealing for 30 – 60 seconds at melting temperature of primers – 5℃, polymerisation for 30 – 60 seconds at 72℃ (30 – 40 cycles), polymerisation 2 – 5 minutes at 72℃ (1 cy cle).

6.2.8 BAC-end PCR

The PCR at the BAC ends for the check of the complete integration of the BAC in the genome was done the following way. The 5´BAC-end PCR was performed using primers 3´fw pBACe3.6, 3´rev pBACe3.6, NEB Taq polymerase (M0320L) and the program: 95℃ for 3 minutes, 1 cycle; 95℃ for 45 s econds, 58℃ for 45 seconds, 72℃ for 45 seconds, 35 cycles; 72℃ for 5 minutes. The size of the PCR product is 272 bp. The 3´BAC-end PCR was done using the primers A: 3´ Sce fw, D: Sce rev and the program: 95℃ for 3 minutes, 1 cycle; 95℃ for 30 seconds, 54℃ for 45 seconds, 72℃ for 60 seconds, 35 cycles; 72℃ for 5 minutes. The size of the PCR product is 167 bp.

6.2.9 *Hhex::Lyn-Tomato* and *Sox17::H2B-Tomato* genotyping

The genotyping of the *Hhex::Lyn-Tomato* mice was done using the primers T3 for and
H5 rev and the following conditions: 95℃ 5min; 95° C 30s, 59℃ 60s, 72℃ 60s for 35 cycles 72℃ 5minutes. Genotyping of the *Sox::H2B-Tomato* mice was done using the primers Neo fw and Neo rev and the following conditions 95℃ 5min; 95℃ 30s, 64℃ 60s, 72℃ 60s for 35 cycles 72℃ 5minutes.

6.2.10 Southern blot

Southern blot gel
The digested genomic DNA was loaded on a 0.8% agarose gel together with 5 µl loading buffer and run over night at 40 V. The gel was photographed (GenFlash, BioRad, München) with a fluorescent ruler to be able to relocate DNA bands. The gel was placed on a rotator for 15 minutes in depurination solution (989 ml H_2O + 11 ml HCl conc.), 15 minutes in denaturation solution (0.4 M NaOH, 0.6 M NaCl) and 15 minutes in neutralization solution (0.5 M Tris, 1 M NaCl, pH7.2).

Blotting
The blot was assembled in a plastic bowl on a glas plate and all membranes and paper were pre-wetted with 20x SSC transfer buffer. The glas plate was placed over the plastic bowl containing the SSC. Blotting paper (Whatman Gel blotting paper 3 mm, GE Healthcare Buchler GmbH & Co. KG) was placed air-bubble free in two layers from the glas plate in the transfer buffer and additionally 4 layers on it (size of the gel). The gel was placed upside down on the paper and covered on the sides and over the slots with plastic foil. The transfer membrane (Hybond-N+, Amersham, GE Healthcare GmbH, München) was placed on the gel and the gel was covered by 4 more layers of blotting paper, two staples of cellulose paper and a thin catalogue and incubated over night. At the next day the gel was turned together with the transfermembrane and the position of the slots was marked with a ballpen. The DNA on the membrane was crosslinked (Spectrolinker XL-1500, Spectronics Corporation, Westbury (NY), USA). The membrane was put in glas tubes and placed in the

hybridisation over at 65℃ together with 25 ml pre- warmed hybridisation solution for 2 – 3 hours.

Labelling and preparation of the Southern probe

For the labelling the Prime-It Random Prime Labeling Kit (Stratagene) was used. As DNA-sonde 25 ng DNA were used (conferring to 5 µl tomato or venus probe), 10 µl H_2O was added and together with 10 µl random 9-er primer boiled for 5 minutes. The mixture was spun put on ice and 10 µl 5x buffer, 1 µl Klenow and 50 µCi (5 µl) radioactive dCTP ^{32}P (Amersham, GE Healthcare GmbH, München) was added and incubated for 10 minutes at 37℃. The reaction was stopped with 2 µl stop mix put back on ice and 49 µl TE was added. Micro-Spin columns (MicroSpin S-300HR, Amersham, GE Healthcare GmbH, München) were prepared by removal of the cap and centrifugation at 760g for 2 minutes. Then the labelling mix was loaded, spun for 1 minute at 760g put back on ice and 1 µl of the probe was counted with the scintillator. One million counts/ml hybridisation solution was used.

Chemical denaturation of the Southern probe

500 µl salmon sperm DNA was denaturated for 10 minutes in a boiling water bath put on ice and transferred in a falcon tube. The hot probe was added, 50 µl 10N NaOH, 300 µl 2M Tris (pH 8.0) and 475 µl 1M HCl drop per drop were added, while shaking. The probe was given in the glas tube with the hybridisation solution and the membranes and incubated at 65℃ over night.

Membrane wash and development

The membranes were washed in a 65℃ water bath short time using 2xSSC/0.5% SDS and for 15 -30 minutes till the membrane had 30 – 60 counts/second. Then the membrane was placed on a polytrap-paper (Benchcoat), covered with foil (Saran) and put in a film cassette (Hypercassette, Amersham, GE Healthcare GmbH, München) using a Hyperfilm MP (Amersham, GE Healthcare GmbH, München) and placed at -80℃ and exposed for 1 - 3 days till development.

6.2.11 Sequencing

The DNA sequencing was done at the Genome Analysis Centre (GAC) of the Helmholtz Zentrum München using ABI Prism 3700 DNA Analyzer (Applied Biosystems, Foster City (CA), USA). The DNA was amplified by PCR and purified before sequencing. For the PCR were used for each 10 ng DNA for 1000 bp´s of a Plasmid and mixed with 10 pM primer DNA, 1 µl Big Dye Terminator (Version 3.1, Applied Biosystems, Foster City (CA), USA) and 1 µl 5x sequencing buffer (Applied Biosystems, Foster City (CA), USA) and filled with H_2O to a total volume of 5 µl. The PCR programm was 96℃ for 1 minute, one cycle, then 35 cycles of 96℃ denaturation for 10 seconds, 50℃ for 5 seconds, 60 ℃ elongation for 4 minutes and 4℃ till end. The PCR product was purified by Ethanol-precipitation by adding 0.5 µl of 125 mM EDTA, 2 µl of 3 M sodiumacetat, 50 µl of 100% ethanol, mixing and incubated for 15 minutes at room temperature. The DNA was sedimented in 30 minutes 11000 rpm centrifugation step at 4℃; the supernatant was removed and the DNA washed in 70 µl of 70% ethanol by a 10 minutes 11000 rpm centrifugation step; and the supernatant air dried. The DNA was resuspended in 25 µl HPLC-H_2O on a shaker and filled in a 96-well sequencing plate.

6.2.12 Cloning of targeting vectors for homolog recombination

The mini-targeting vectors for the BACs recombineering were cloned in a pBS^- Vector. The targeting sequence consisted of either the superior fluorescent protein Tomato (tdTomato; Campbell *et al*., 2002; Shaner *et al*., 2004) or Venus (Nagai *et al*. 2002). The tandem dimer Tomato (tdTomato) is a fusion of two copies of dimer (d)Tomato which is the product of a directed evolution of red fluorescent protein from Discosoma sp. (dsRed; Baird *et al*. 2000). This fluorescent protein shows a brighter fluorescence and a higher tolerance of N- and C-terminal fusions than the wild-type dsRed.

The protein Venus (Nagai *et al*., 2002) is the improved version of the yellow fluorescent protein (YFP; Miyawaki *et al*., 1997), a variant of the jellyfish (Aequorea victoria) derived green fluorescent protein (GFP; Shimomura et al., 1962; reviewed by Tsien, 1998). For localisation of the fluorescent proteins to the membrane NH_2-terminal residues of the myristoylation-palmitoylation sequence of the kinase Lyn

(MGCIKSKRKDNLNDDE; Zacharias *et al*., 2002) were fused to the N terminus of the fluorescent protein. For localisation to the nucleus the amino acids of human histone 2B (H2B) were fused to the fluorescent protein. The mini-targeting vectors contained also an artificial intron construct to avoid nonsense mediated decay (shown by Maquat 2005) and a poly(A) tail (pA) to protect the mRNA molecule from degradation and enhances its translation.

A *Neomycin* resistance gene (*Neo*) which enables selection for successful integration is also part of the targeting vector and its expression is driven by a PGK (phospho-glycerate kinase promoter)/EM7 (synthetic bacterial) promoter. The PGK-*Neo* cassette is flanked with FRT sequences (Meyers *et al*., 1998) which allows Flp-mediated excision. The targeting cassette is flanked on the 5´ and 3´ end by homology arms which allow the homologous integration in the BAC (see also Figure 6.2.12.1.). Details about the cloning of the mini-targeting vectors for *Hhex*, *Nkx2.1* and *Pdx1* are given in Chapters 6.2.13, 6.2.14 and 6.2.15.

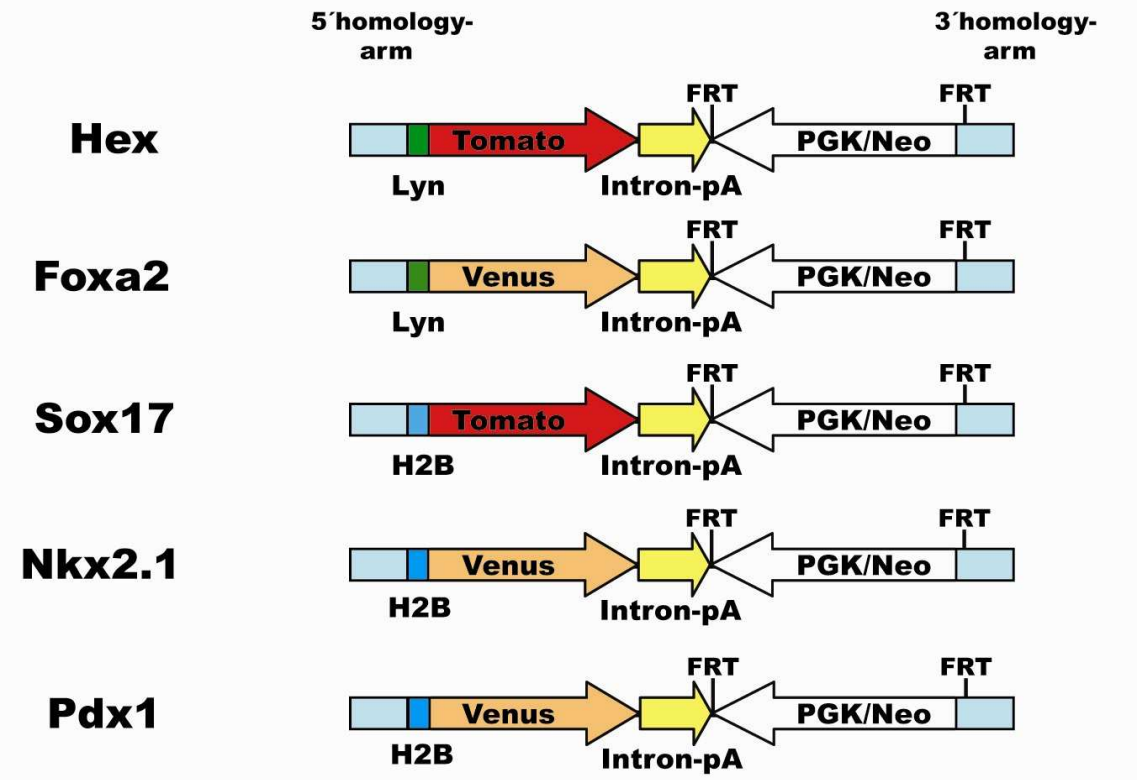


Figure 49 Mini-targeting vectors for BAC recombineering in bacteria
The targeting vectors containing 3´ and 5´ homology arms for targeting by homolog recombination to the translation start (ATG) codon or an upstream exon of the corresponding gene. The fluorescent reporters Tomato and Venus are fused either to the localisation sequence Lyn (myristoylation-palmitoylation sequence of the Lyn kinase) for localisation to the membrane or H2B (histone 2 B) for localisation to the nucleus. The artificial intron and the poly-A sequence is to avoid nonsense mediated decay and mRNA stabilization. The FRT flanked PGK/*Neo* cassette is to allow selection in ES cells.

Homology arms were generated by PCR amplification using the corresponding wild-type BACs as templates (for details see Methods 5.18.). The targeting vectors for *Foxa2* and *Sox17* were cloned and nicely provided by Dr. Ingo Burtscher from my Lab. The sequences of all mini-targeting constructs were verified by sequencing before using them for BAC recombineering.

6.2.13 The *Hhex* mini-targeting Vector

The targeting vector for *Hhex* was cloned by modification of a pKS$^-$ vector (p bluescript KS$^-$; 4312 bp) containing Lyn-Venus-Intron-pA (Figure 60). The Vector was digested with *Xba*I and *Spe*I to excise Venus and integrate Tomato (1454 bp). The resulting vector was digested with *BamH*I and *EcoR*I to introduce the PGK/*Neo* cassette (1841 bp) which was obtained from a PL-451 vector by *BamH*I and *EcoR*I digestion. The *Hhex* homology arms were done by PCR using the primers *Hhex* Ex 1A fw (EP-55), *Hhex* Ex 1A rev (EP-56) for the 5´ homology arm and the primers Hhex Ex 1B fw (EP-57),*Hhex* Ex 1B rev (EP-58) for the 3´ homology arm. Amplification was done using DNA from BAC RPCI-23-117H14 (*Hhex*) as template. The exact conditions are given in methods 5.18. The PCR primers introduce *Sac*I, *Nsi*I (5´ end) and *Not*I (3´ end) restriction sites in the 5´prime homology arm (322 bp) and *EcoR*I (5´ end) and *Hind*III (3´ end) restriction sites in the 3´ homology arm (463 bp).

One µl of the PCR product was checked on agarose gel, the PCR product was digested to trim the fragment ends and gel purified. First, the 3´ homology arm was ligated in a pBKS$^-$ vector using *EcoR*I and *Hind*III then the 5´ homology arm was introduced using *Sac*I and *Not*I. The Lyn-Tomato-Intron-pA-*Neo* cassette (6818 bp) was excised using *Not*I and *EcoR*I and cloned in between the prepared 5´ and 3´ homology arms. The final vector was verified by sequencing. For BAC recombineering the targeting sequence (4682 bp) was excised by *Nsi*I / *Hind*III digestion of the vector (7584 bp) and gel purified. In the homolog recombination step, the *Hhex* mini-targeting vector replaces in the BAC the sequence of Exon1 of *Hhex* from the start ATG to the beginning of Intron 1.

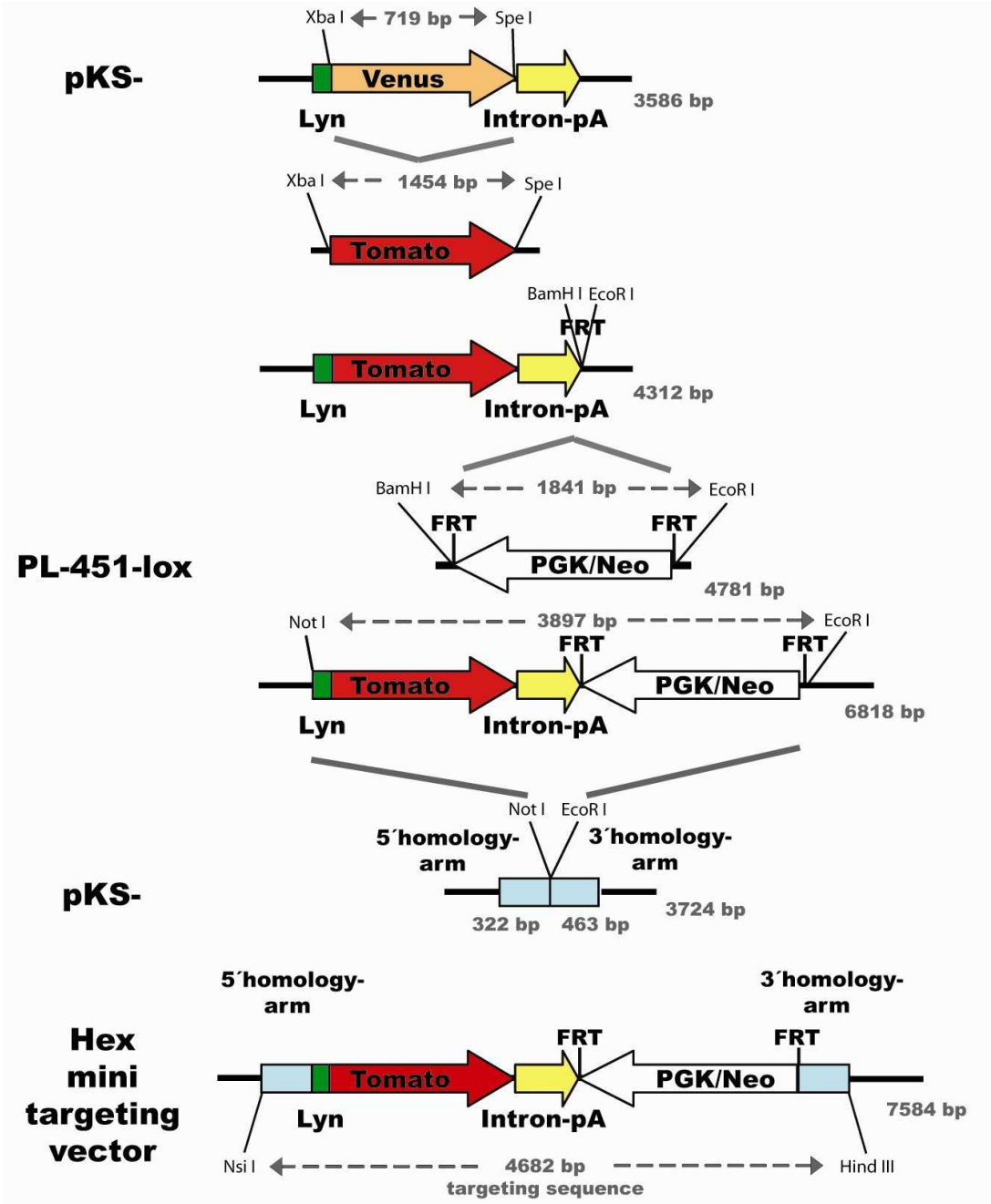


Figure 50 Cloning of the *Hhex* mini-targeting vector
Cloning strategy for the generation of the *Hhex* mini-targeting vector. The names of the used vectors for subcloning steps are given in the left side. The oligonucleotide primers used for amplification of the homology arms and the amplification conditions are given in the methods (5.18).

6.2.14 The *Nkx2.1* mini-targeting vector

The targeting vector for *Nkx2.1* was cloned by modification of a pKS$^-$ vector (p bluescript KS$^-$; 4312 bp) containing the Lyn-Venus-Intron-pA sequence (1373 bp). The Vector was digested with *BamH*I (cuts after the Intron-pA) and *EcoR*I. The PGK/*Neo* cassette (1841 bp) was obtained from a PL-451 vector by *BamH*I and

*EcoR*I digestion and ligated in the prepared pKS¯ vector (6153 bp). The *Nkx2.1* homology arms were done by PCR using the primers *Nkx2.1* Ex 1A fw (EP-51), *Nkx2.1* Ex 1A rev (EP-52) for the 5´ homology arm and the primers *Nkx2.1* Ex 1B fw (EP-53), *Nkx2.1* Ex 1B rev (EP-54) for the ¯3´ homology arm. The amplification

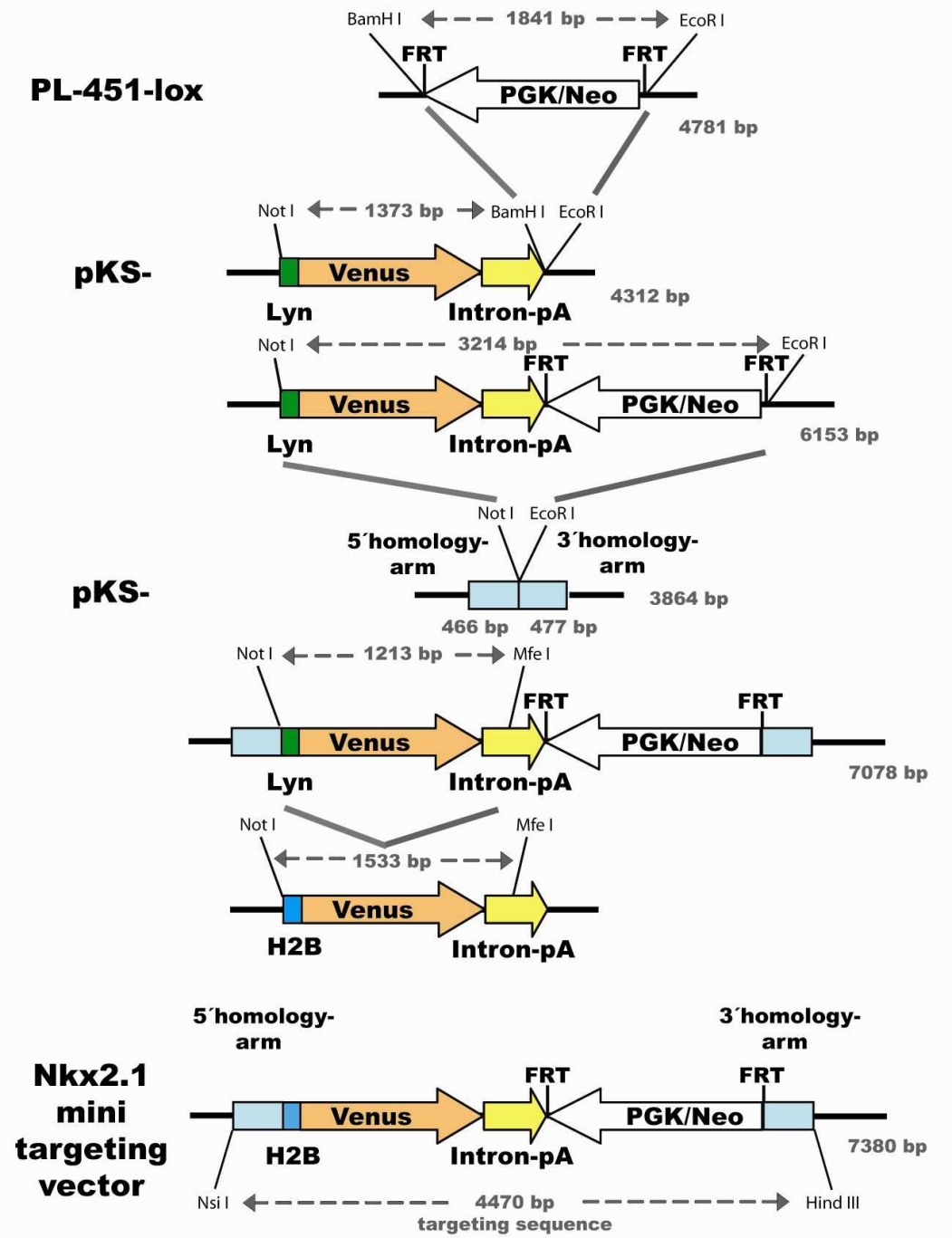


Figure 51 Cloning of the *Nkx2.1* mini-targeting vector
Cloning strategy for the generation of the *Nkx2.1* mini-targeting vector. The names of the used vectors for subcloning steps are given in the left side. The oligonucleotide primers used for amplification of the homology arms and the amplification conditions are given in the methods (5.18).

conditions are given in methods 5.18. The PCR primers introduce *Sac*II, *Nsi*I (5´ end) and *Not*I (3´ end) restriction sites in the 5´ homology arm (466 bp) and *Not*I, *EcoR*I (5´ end) and *Hind*III (3´ end) restriction sites in the 3´ homology arm (477 bp). One µl of the PCR product was checked on agarose gel then the PCR product was digested to trim the fragment ends and gel purified. Both homology arms were ligated in a three way ligation in a pBKS$^-$ vector which was digested with *Sac*II and *Hind*III. After ligation this vector was cut between the 5´ and the 3´ homology arm with *Not*I and *EcoR*I. The Lyn-Venus-Intron-pA-*Neo* cassette (3214 bp) was cut from pKS$^-$ using *Not*I and *EcoR*I and ligated in between the 5´and 3´ homology arm. To complete the targeting vector, the Lyn-Venus-Intron-pA sequence (1231 bp) was replaced by H2B-Venus-Intron-pA (1533 bp) in a *Not*I / *Mfe*I digestion and ligation step. The final vector was verified by sequencing. For homolog recombination the targeting sequence (4470 bp) was excised by *Nsi*I / *Hind*III digestion of the vector (7380 bp) and gel purified. In the homolog recombination step, the *Nkx2.1* mini-targeting vector replaces in the BAC the sequence of whole Exon1 of *Nkx2.1* from the 5´UTR to the beginning of Intron 1.

6.2.15 The *Pdx1* mini-targeting vector

The targeting vector for *Pdx1* was cloned by modification of the targeting vector of *Nkx2.1* (see 6.2.14.). The H2B-Venus-Intron-pA sequence (1675 bp) was excised in a *Not*I / *BamH*I digestion while the floxed *Neo* cassette was excised by a *BamH*I / *Xho*I digestion. The *Pdx*1 homology arms were done by PCR using the primers Pdx Ex 1A fw (EP-373), Pdx Ex 1A rev (EP-374) for the 5´ homology arm and the primers Pdx Ex 1B fw (EP-375), Pdx Ex 1B rev (EP-376) for the 3´ homology arm. The amplification conditions are given in methods 5.18. The PCR primers introduce *Sac*I (5´ end) and *Not*I (3´ end) restriction sites in the 5´ homology arm (328 bp) and *EcoR*I (5´ end) and *BamH*I / *Xho*I (3´ end) restriction sites in the 3´ homology arm (447 bp). One µl of the PCR product was checked on agarose gel then the PCR product was digested to trim the fragment ends and gel purified.

The 5´ homology arm was together with the H2B-Venus-Intron-pA sequence (1675 bp) ligated in a *Sac*I / *BamH*I digested pBKS$^-$ vector (4924 bp). The 3´ homology arm was ligated together with the *Neo* cassette (1841 bp) in a *Xho*I digested pBKS$^-$ vector (5190 bp). The final vector was created by *BamH*I / *Xho*I digestion of the 3´ homology

arm fused to the *Neo* cassette (2283 bp) and ligated behind the H2B-Venus-Intron-pA sequence of the *BamH*I / *Xho*I linearized pBKS$^-$ vector (7156 bp). The vector sequence was verified by sequencing. For homolog recombination the targeting sequence (4194 bp) was excised of the vector (7156 bp) by *Afl*II / *Xho*I digestion and gel purified. In the homolog recombination step, the Pdx1 mini-targeting vector replaces in the BAC the sequence of Exon 1 of *Pdx*1 from the start ATG to the end of Exon 1.

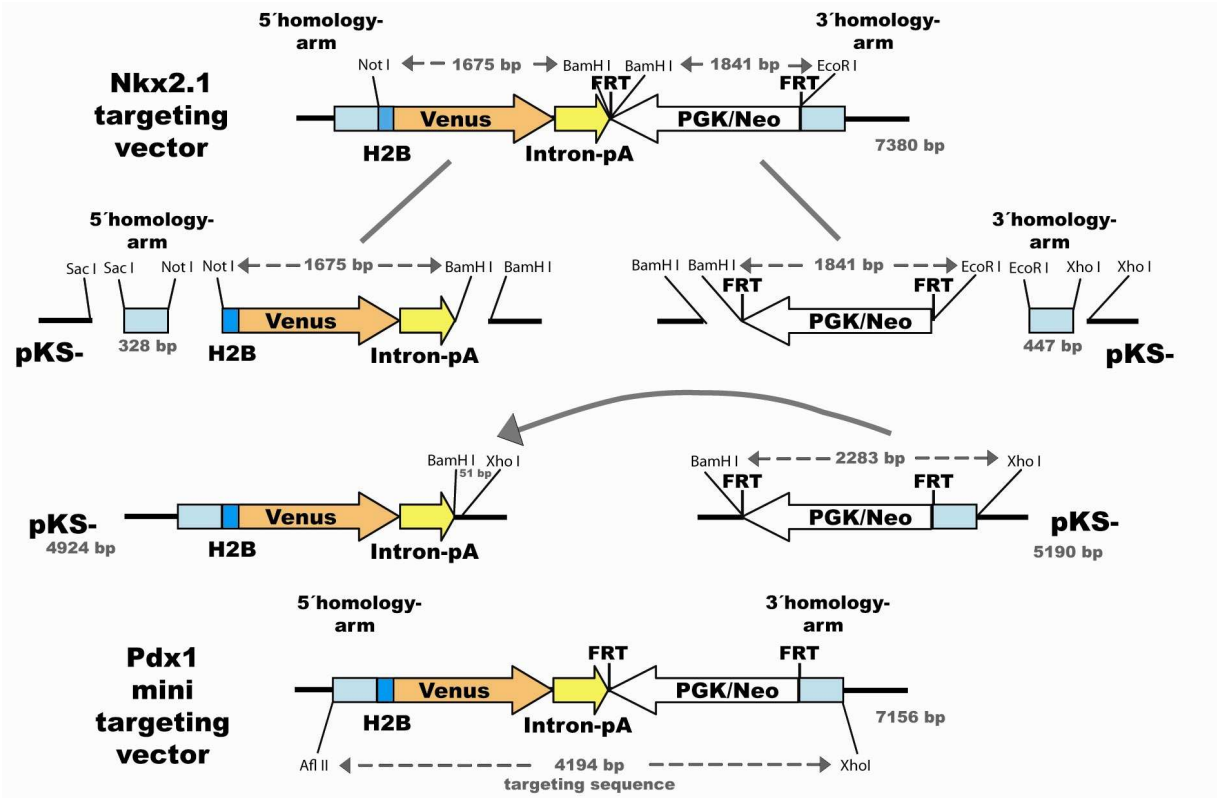


Figure 52 Cloning of the *Pdx*1 mini-targeting vector
Cloning strategy for the generation of the *Pdx*1 mini-targeting vector. The names of the used vectors for subcloning steps are given in the left and the right side. The oligonucleotide primers used for amplification of the homology arms and the amplification conditions are given in the methods (5.18).

6.2.16 Recombineering of BACs in bacteria

Transformation of BAC in bacteria

Recombineering (recombination-mediated genetic engineering) allows the modification of BAC DNA by homologous recombination in bacteria (Copeland *et al*.; 2001; Court *et al*., 2002). It is based on homologous recombination in E. coli using recombination proteins provided from the λ phage.

We used the bacterial strain EL250 (derived from DY380 a DH10B derived strain) which contains a defective λ prophage who is inserted into the bacterial genome (Lee *et al*., 2001). The phage genes exo, bet and gam are transcribed from the λpL promoter. The promoter is repressed by the temperature-sensitive repressor cI857 at 32℃ and derepressed (inactive) at 42℃. When EL250 cells are kept at 32℃ no recombination proteins are produced but after a brief (15 minutes) heat-shock at 42℃ a sufficient amount of recombination proteins are produced (Yu *et al*., 2000; Lee *et al*., 2001).

The gene exo encodes a 5' - 3' exonuclease that creates single-stranded overhangs on introduced linear DNA. The gene bet encodes a pairing protein that binds to the 3´ overhangs and mediates its annealing and homologous recombination with complementary DNA. The gene gam encodes an inhibitor of the E. coli RecBCD exonuclease and protects in this way the linear DNA-targeting cassette from degradation.

Competent EL250 cells were prepared by inoculation of 1-2 ml over-night culture in 25 ml LB medium at 32° (EL250 cells are without selection marker). Cells were grown to OD_{600}= 0.6, cooled down in a water-ice bath and centrifugated at 5000 rpm for 5 minutes. The pellet was resuspended in 1ml 10% ice-cold glycerol, transferred in an Eppendorf tube and three times resuspended in 1ml 10% ice-cold glycerol and sedimented at 14000 rpm for 20 s in a tabletop centrifuge. Finally the pellet was resuspended in 50µl 10% ice-cold glycerol which was directly used for electroporation or stored at -80℃ (modified from Lee *et al*., 2001).

The BAC DNA (chloramphenicol resistant) was transformed in competent EL250 cells by electroporation of 250 ng DNA (~1/4 BAC-mini; dissolved in up to 15µl H_2O) with 50µl competent EL250 cells in a pre-cooled 0.1 cm cuvette (BioRad Gene Pulser) using a BioRad GenePulser Xcell electroporator with the settings 1.8 kV, 200Ω, 25µF (modified from Lee *et al*., 2001).

After electroporation cells were promptly transferred in an Eppendorf tube with 1ml LB media, incubated for 60 - 90 minutes at 32℃ shaking at 750 rpm and plated out in dilutions of 1:10, 1:100 and 1:1000 on LB agar plates containing 12.5 µg/ml choramphenicol. The plates were incubated for 24h at 32℃, colonies were picked and glycerol stocks done. To check the integrity of the BAC after transformation, BAC minis were done (see protocol), digested overnight and the restriction pattern of

several enzymes were visualized on an agarose gel. The unmodified and untransformed BAC was used as control.

Homolog recombination of the targeting vector in the BAC

A 500 µl over-night culture of the clone was inoculated from the glycerol stock for the homologous recombination of the targeting construct in the BAC. With this starter culture a 20ml LB culture was inoculated, grown until OD_{600}=0.6 and split in two parts. The control part was incubated at 32℃, while the second half for the heat-shock induction of the recombinant proteins was incubated for exactly15 minutes at 42℃ in a shaking water bath. Then the cells were rapidly cooled down in a water-ice bath and competent cells were done as described above. The competent heat-shocked cells and the non heat-induced control cells were electroporated with 20 ng DNA of the targeting construct (see above). Cells were plated out in dilutions of 1:10, 1:100 and 1:1000 on LB agar plates containing 12.5 µg/ml of kanamycin. Colonies were picked, glycerol stocks done and the clones were checked as described above for transformation of BAC in bacteria.

Arabinose-induced Flp recombination

The EL250 cells also contain a tightly controlled arabinose-inducible flpe gene. flpe will mediate recombination between two identical frt sites and excise the `floxed` neomycin resistance cassette which mediates resistance to kanamycin.

A 50 ml LB media day culture (12.5 µg/ml choramphenicol) of a successfully homolog recombinated and checked clone was inoculated from a 5 ml over-night culture (LB medium and 12.5 µg/ml kanamycin) and grown till OD_{600}= 0.4. Then 0.5 ml of 10% L(+) Arabinose (Sigma; final concentration 0,1%) was added and the cells were incubated for 1 hour at 32℃ on a shaker. The cells were diluted 1:5000, 1:10000 and 1:100000 and plated out on LB agar containing 12.5 µg/ml choramphenicol and incubated over-night at 32℃. To check for successful recombination 10 colonies were chosen, picked marked with numbers on the plates and either plated out on LB agar plates containing kanamycin (12.5 µg/ml) and LB agar plates containing choramphenicol (12.5 µg/ml).

Only colonies which grew on kanamycin LB plates were checked by digestion as described above.

Retrofitting

The modification of BACs to introduce transfection markers and other functional sequences is called retrofitting. We used an *in vivo* retrofitting approach with a single vector (pRetro-ES; Wang *et al*., 2001) to introduce a neo-resistance gene conferring resistance in bacteria and in ES cells in the BAC vector pBACe3.6 sequence (Frengen *et al*., 1998) to be able to select for the successful integration of the BAC in ES cells. The Vector pRetroES is a self-catalytic, suicidal cre fusion gene containing Vector (tac-GST-loxP-cre fusion gene). pRetroES contains a conditional replication origin (ori^{r6ky}) which functions only in the presence of the λpir gene product which is produced by the bacteria strain JM109λpir. So the expression of pRetroES is restricted to JM109λpir cells. The fusion gene provides enough Cre recombinase to promote site-specific recombination for integration of the construct into the BAC vector when transformed in the EL250 cells which containing the modified BAC.

Once integrated, the cre fusion gene is broken at the loxP site and becomes non-functional, preventing the production of any excess Cre recombinase. For selection of multiple integrations of the BAC in the ES cell genome the strong PGK promoter of pRetroES was replaced by a weak thymidine kinase (TK) promoter. The modified Vector was named pRetroES-tk.

Competent cells of a successful Flp recombinated clone were done (using LB medium containing 12.5 µg/ml choramphenicol) and electroporated with 200 ng DNA of pRetroES-tk (Wang *et al*., 2001). After electroporation cells were directly transferred in 1 ml LB medium containing 15 µg/ml choramphenicol and incubated for 90 min at 32℃ on a shaker. Then cells were plated out on LB plates (1/10 and 9/10) containing 15 µg/ml choramphenicol and 50 µg/ml ampicillin as the integrated vector mediates additionally ampicillin resistance. Expected clones were 5 - 40 clones (1/10) and 50 - 300 (9/10) clones.

About 30 clones were picked and BAC minis were done. The clones were screened by PCR for successful retrofitting in the *lox*P site. The used primers were EP-65 (Retro-R), EP-67 (F1) and EP-68 (F2) which lead to a product of 367 bp for integration of pRetroES-tk in loxP and 328 bp for integration in *lox*511. The program was 94℃ for 3 minutes, 1 cycle, 94℃ for 45 second s, 56℃ for 45 seconds and 72℃ for 45 seconds, 35 cycles, 72℃ for 5 minutes. Clon es which had only a integration in the loxP site were selected and checked by multiple digests as described above.

Positive clones were verified for BAC integrity by PCR on the 5´ prime and 3´ prime end of the BAC vector sequence (pBACe3.6) as second check for the complete integration of the BAC into the genome (Yang and Seed, 2003).

6.3 Microbiological methods

6.3.1 Transformation of plasmid DNA in bacteria

Transformation of plasmid DNA in bacteria was done using chemically competent E. cloli (Invitrogen, Carlsbad (CA), USA). The cells were thawed on ice and 20 ng plasmid DNA was added. The mixture was kept on ice for 5 minutes and incubated for 2 minutes at 42 ℃. Then the cells were transfe rred in 1 ml LB medium, incubated for 60 minutes at 37 ℃ and plated on LB plates.

6.3.2 Generation of glycerin stocks

Glycerol stocks (10%) were done by adding of 110 µl 86 % Glycerol to 860 µl LB culture. The stocks were kept at - 80 ℃.

6.4 ES cell culture

6.4.1 Murine ES cells

For all experiments mouse IDG3.2 embryonic stem (ES) cells from F_1 (C57/BL/6Jx129S6/SvEvTac) were used (Hitz *et al*., 2007).

6.4.2 Preparation of feeder cells

Feeder cells are mouse embryonic fibroblast cells, which have been mitotically inactivated by mitomycin c treatment. Primary neomycin resistant fibroblasts were obtained from embryos of the transgenic mouse strain C57Bl/6J-Tg(*pPGKneobpA*)3Ems/J at the age of E14.5 to E16.5 (Jackson Laboratories) under sterile conditions. Cultivated primary fibroblasts were expanded for two passages

and grown on 10 cm cell culture dishes until confluence. Cells were incubated with medium containing 10 µg/ml mitomycin c for 2 hours at 37°C. After intensive washing with PBS, feeder cells were trypsinized and plated on a fresh culture dish or frozen at -80°C (in 1x freezing medium) for later use. Feeder cells were plated at a density of 2 - 2.5 x 10^4 cells/cm^2 at least for several hours or one day prior to plating of ES cells.

6.4.3 Splitting of ES cells

For expansion, ES cells were splitted every 36 – 45 hours when colonies have grown nicely but not yet to a confluence of 70-80%. The medium was sucked off, cells were washed twice with PBS and trypsinized for 5 minutes at 37°C until cells detached from the surface. The reaction was stopped by adding an equal amount of medium to the cells and the suspension was resuspended carefully by pipetting 10 times up and down and splitted on several fresh culture dishes depending on the desired amount of cells per dish. For determination of cell number, cells were pipetted in a Neubauer counting chamber and ES cells were counted.

6.4.4 Freezing and thawing of ES cells

ES cells were stored -80°C for short term storage (up to 2-3 months) or in liquid nitrogen for long term storage. Cells were trypsinized as described, centrifugated, resuspended in ice cold 1x freezing medium, and pipetted into a 2 ml cryovial. Vials were frozen in a freezing container at -80°C. Due to the isopropanol in the freezingcontainer the temperature is lowered for one degree every minute inside until it reaches the final temperature of -80°C. For long term storage, the vials were then transferred into liquid nitrogen. For thawing of ES cells the cryovials were placed in the water bath at 37°C, rapidly transferrd in 10 ml ES medium and centrifugated to get rid of the DMSO. Then the cells were plated on dishes with feeder cells.

6.4.5 *In vitro* differentiation of ES cells endoderm

NIH3T3 fibroblasts expressing murine Wnt3a (Kispert *et al*., 1998) were plated on 6-well dish at a densitiy of 3x10^5/ well in Dulbecco´s Modified Eagle Medium (DMEM, Invitrogen) containing 10% fetal calf serum (FCS, PAN). The medium was

exchanged after 3h when the feeders attached by 2 ml SF-O3 serum free culture medium (Takakura *et al*., 1996; Sanko Junyaku) supplemented with 50 µM β-mercaptoethanol and recombinant human activin A (10 ng/ml, R&D Systems) for the induction of ES cell differentiation. Two times pre-plated ES cells were plated at a density of $6x10^5$ cells per well on the Wnt3a expressing feeders. During differentiation the medium was changed every second day. Protocol according to Tada *et al*., 2005.

6.4.6 Electroporation of BAC DNA in ES cells

ES cells were grown till 70 – 80% confluency in a 10 cm tissue culture dish trypsinized using 1 ml trypsin and 2 ml PBS incubated for 5 minutes at 37℃ and resuspended. Then ES cells were transferred in a Falcon tube centrifugated, the supernatant was sucked off and the cells were washed with 10 ml PBS and resuspended in 1ml ice-cold PBS. The cells were counted in a Neubauer chamber and $1x10^7$ cells were resuspended in a volume of 200 µl $PBS^{-/-}$.
The digested BAC DNA (total volume 600 µl) was mixed with the ES cells transferred in a pre-cooled 0.4 cm cuvette (BioRad) and incubated on ice. Electroporation was done using a BioRad GenePulser Xcell electroporator with the settings 220 V, 500 µF, resistance ∞. After electroporation the cuvette was kept on ice for additionally 5 minutes till the ES cells were transferred on two 10 cm tissue culture dishes with MEF feeder cells and ES medium. After 2 days the medium was exchanged by selection ES medium containing G418 (300 µg/ml). The medium was replaced every day by fresh medium.

6.4.7 Selection and picking of ES cell clones

Plates were incubated for 24 – 48 hours at 37℃ till the medium was exchanged by ES medium containing 300 µg/ml G418 (Gibco). Then the medium was exchanged daily by fresh ES medium containing G418. The selection for ES cell clones after electroporation with BAC DNA took 6 to 10 days. The medium was removed and replaced by $PBS^{-/-}$ and the clones got picked from the plate with a 200 µl pipet and transferred in a conical 96 well plate containing 60 µl $PBS^{-/-}$. After adding of 30 µl trypsin the clones got resuspended and two times 45 µl medium was transferred in

96 well plates containing 45 µl ES medium and MEF feeder cells. One plate was used as “master” to maintain the clones while the other was used for DNA preparation for the Southern blot screen.

6.4.8 Freezing of ES cell clones in 96 well format

The medium was aspirate off and the ES cells were washed with 200 µl $PBS^{-/-}$. Trypsin was added (40 µl per well) and the plates were incubated for 5 minutes at 37℃ till rapidly 60 µl of cold ES cell media was added and the cells were resuspended by pipetting. Finally 200 µl of cold 2 x ES freezing medium was added and mixed by pipetting. Then the lid was drawn the plate closed with Parafilm and enwrapped using a napkin and placed at –80℃.

6.4.9 Preparation of genomic ES cell DNA in 96 well format

The ES cells on 96-well plates grew till they are 80% confluent. The medium got aspirate off and the cells were washed two times with 100 µl PBS. Then 50 µl lysisbuffer was added to each well and the plate was closed and incubated in a humid chamber over night at 55℃. At the next day 100 µl of a NaCl/EtOH mixture was added (150 µl 5M NaCl + 10 ml ice-cold 100% EtOH) to each well and incubated for 30 minutes. Then the plate was turned very, very slow to remove the NaCl/EtOH mixture without removing the precipitated DNA. The DNA was washed three times with 150 µl 70% EtOH dried for 10 – 15 minutes at room temperature and dissolved in 25 µl TE (pH 7.5) on a rotator for one hour.

6.4.10 Digestion of ES cell DNA in 96 well format

The dissolved ES cell DNA (25 µl) in the 96 wl plate was incubated overnight with 40 µl of a mastermix containing 25 units enzyme/reaction, 1x buffer, 6.25 µg RNase A, 1 mM spermidin and 1 x BSA.

6.5 Mouse embryo technique

6.5.1 Mouse strains

All used mice were kept and used in the animal house (A) of the Helmholtz Zentrum München. The BAC transgenic mouse lines *Hhex::Lyn-Tomato* and *Sox17::H2B-Tomato* were generated by crossing of a chimera generated by diploid aggregation technique with CD1 females and intercrosses of the offspring. Mice for the experiments were generated by intercroses of heterozygote transgenic mice or crossings of them to CD1 mice.

The mouse lines for the generation of diploid or tetraploid chimera embryos were collected from YFP mice (Hadjantonakis *et al*., 2002) or ds Red mice (Vintersten *et al*., 2004) which were both maintained on a mixed genetic background (CD1/129Sv/C57/Bl6). The *Hhex-GFP* mouse line (Rodriguez *et al*., 2001) was nicely provided from Tristan A.Rodriguez. The line was generated by the insertion of a 8 kb promoter-reporter construct that includes a 4.2 kb sequence upstream of the *Hhex* ATG (containing the *Hhex* endogenous promoter) and a 3.8 kb sequence downstream of the ATG (containing exons 1-3, introns 1-2, and 800 bp of Intron 3) which drive the *enhancedGFP* (*eGFP*) reporter gene and some regulatory elements of the *Hhex* gene.

6.5.2 Isolation and preparation of blastocyst stage embryos

Embryos at stages E5.5 to E8.5 were dissected in dissecting media (live imaging) containing DMEM 10% FCS, 20mM HEPES (Sigma). Then the Embryos were cultured on glass bottom dishes (MAT TEK, P35GC-1.5-14-C) using 100-200 µl embryo culture medium (50% rat serum, 40% DMEM without phenol red (Invitrogen 31053-028), 2 mM glutamine, 1 mM sodium pyruvate and 100 µM β-mercaptoethaol and placed in a incubator at 37℃ containing 5% CO^2 and 5% O^2. To avoid evaporation the medium was covered with mineral oil. Image aquisiution was done using an Olympus FV1000 or Leica DMI 6000 confocal microscope. Embryos older than E8.5 were dissected in dissecting PBS and kept in dissecting media (live

imaging). Image acquisition was done using a Zeiss Axiovert 200M inverse microscope with AxioCam MRc or ARc.

6.5.3 Tetraploid complementation technique (tetraploid aggregation)

The embryos were recovered at two cell stage from the oviducts of 1.5-dpc pregnant CD1 mice obtained from Jackson Laboratories. The embryos were washed in 0.3 M mannitol solution (Sigma M9546) and placed on the plastic dish to the side of electrode slide, lined up in a row between the electrodes (BLS, CF-150B) using 0.5 – 1 V and pulsed with 35 V for 35 microseconds to initiate the blastomere fusion. After pulsing the embryos were washed through sequential drops of M2 media (Chemicon International, MR-015P-5F) to remove mannitol and then placed in a dish containing drops of KSMO media (Chemicon International, MR-020P-5F) covered with mineral oil. For fusion of the blastomeres they were incubated for 10 to 15 minutes in the incubator. The fused blastomeres were cultured overnight to develop to the four-cell stage. Then the zona pellucida was dissolved in acide tyrode solution (Sigma, P0930) and the embryos were placed into depression wells of aggregation plates. The ES cells for the fusion were lightly trypsinized to release small clumps of 8 to 15 cells and transferred into the depression well containing the four-cell stage embryos that they fuse together. The fused cells were transferred after overnight incubation in the uterus of E 2.5 pseudopregnant recipient female mice.

6.6 Immunohistochemistry

6.6.1 Fixation and staining of blastocyst stage embryos (pre- and early implanted embryos)

The blastocysts were flushed from oviduct or uterus with M2 media (Sigma), the zona pellucida was removed using tyrodes solution (Sigma) and washed 3 times with PBT. The fixation was done with 2% PFA for 10 minutes followed by 2 washing steps and the bocking in 2% serum for 30 minutes on a rotator. The 1st antibody was incubated for 30 minutes in blocking solution then washed 3 times for 10 minutes and the 2nd antibody was incubated in blocking solution for 20 minutes. The blastocysts were

washed 3 times in blocking solution for 10 minutes, in the last step additional with DAPI (0.2 µg/ml) and placed on a glass bottom dish (MatTek), overlayed with mineral oil (partially based on Vestweber and Kemler, 1984).

6.6.2 Fixation and staining of gastrulation stage embryos (whole mount staining protocol; E6.5-E7.75)

The mouse urterus was directly placed in ice-cold dissecting PBS, the embryos rapidly dissected (Leica MS5) under cold conditions and fixed with 2% ice-cold PFA for 13 minutes at 7℃ and 13 minutes at room temperature on a roller incubator (Stuart SRT1). The embryos were rinsed two times in TBST, permeabilized for 10 minutes at room temperature (0.1% Triton X-100 and 100 mM glycine in PBS) and again rinsed two times in TBST.

The blocking was done in10% FCS, 0.1% BSA (Sigma) and 3% serum in TBST (20 mM Tris-HCl, pH 8.0, 150 mM NaCl. 0.05% Tween-20) for 1-2 hours. The blocking serum species (and also for the first and secondary antibody) was different to the origin species of the first antibody. The primary antibody was diluted in TBST containing 0.1% BSA and 1.5% serum and incubated at 7℃ overnight followed by multiple TBST washing steps for 30 - 45 minutes. The secondary antibody was diluted in TBST containing 0.1% BSA and 1.5% serum and incubated for 3 hours at room temperature followed by multiple TBST washing steps for 45 minutes. The embryos were rinsed two times in dissecting PBS and incubated in dissecting PBS with DAPI (0.2 µg/ml) for 20 - 30 minutes, rinsed two times in TBST, transferred for 5 minutes in 30% glycerol and 5 minutes in 60% glycerol (all steps on Rotator, Grant Boekel, VSR23).

For embedding the embryo was placed in the pre-wetted (ProLong Gold antifade reagent, Invitrogen, P36931) well of a secure-seal spacer (Invitrogen, 9mm diameter, 0,12 mm deep, S24737) which was attached to a coverslip (24x60mm, Roth).

Residual glycerol was removed, the embryo embedded in the antifade reagent and the well closed with a second coverslip. The specimen was dried over night at room temperature in the dark and long term stored at 7℃ in the dark (protocol partially based on Nakaya *et al*.; 2005).

7 APPENDIX

7.1 Publications

Roessler, S., Györy, I., Imhof, S., Spivakov, M., Williams, R.R., Busslinger, M., Fisher, A.G. and Grosschedl, R. (2007). Distinct Promoters Mediate the Regulation of *Ebf1* Gene Expression by Interleukin-7 and Pax5. Mol Cell Biol27, 579-94.

Imhof, S., Burtscher, I., Lickert, H. Endoderm lineage labelling using BAC reporter constructs in ES cells and mouse embryos. Genesis, under review.

7.2 Abbreviations

A

ADE	anterior definitive endoderm
AIP	anterior intestinal port
A-P	anterior-posterior
AVE	anterior visceral endoderm

B

BAC	bacterial artificial chromosome
bf	brightfield

C

Cer1	Cerberus-like 1
Cdx2	caudal type homeobox 2
CIP	caudal intestinal port
Cre	causes recombination

D

DE	definitive endoderm
Dkk1	dickkopf 1

DVE	distal visceral endoderm
E	
E	embryonic day
EGO	early gastrula organizer
EHF	early headfold stage
EMT	epithelial-mesenchymal transition
ES	embryonic stem
ExE	extra-embryonic endoderm
F	
FGF	fibroblast growth factor
Flp	Flippase
Foxa2	Forkheadbox transcription factor a2
Foxd3	Forkhead box transcription factor d3
FRT	Flp-recombinase target
G	
Gata6	GATA binding protein 6
GFP	green fluorescent protein
H	
H2B	Histon 2BHhex Hematopoietically
Hhex	Hematopoietically expressed homeobox gene
HF	headfold
HMG	high mobility group
I	
ICM	inner cell mass
iCre	improved Cre
IRES	internal ribosomal entry site
L	
Lefty1	left right determination factor 1
LGO	late gastrula organizer

loxP	locus of X over Pi
M	
MET	mesenchymal-epithelial transition
MGO	mid-gastrula organizer
MS	mid-streak stage
mTE	mural trophectoderm
N	
Nanog	Nanog homeobox gene
Neo	neomycin resistance
Nkx2.1	NK homeodomain transcription factor 2.1
O	
Oct4	Octamere-binding protein 4
O/N	over night
ORF	open reading frame
Otx2	Orthodenticle homolog 2
P	
Pdx1	pancreatic-duodenal homeobox 1
PGK	phosphoglycerate-kinase
PM	posterior mesoderm
PrE	primitive endoderm
Prox1	prospero-related homeobox 1
pTE	polar trophectoderm
R	
RA	retinoic acid
rpm	rounds per minute
S	
Sox2	Sry-related HMG-box transcription factor 2

Sox17	Sry-related HMG-box transcription factor 17
Sry	sex determing region y
T	
T	Brachyury
TE	trophectoderm
TF	transcription factor
TGFβ	transforming growth factor
V	
VE	visceral endoderm
W	
Wnt	Wingless/Int

7.3 List of Figures

7.4 Curriculum vitae

Personal Data

Name:	Sascha Imhof
Date of Birth:	13th of March 1977
Place of Birth:	Marburg
Nationality:	German

Education

1983-1987	Elementary school (Grundschule Biedenkopf)
1987-1993	Junior high school (Realschule Biedenkopf)
1993-1996	Secondary school (Städtisches Gymnasium Bad Laasphe) Certificate: (Abitur)
1997-2000	Academic studies in biology (Philipps-Universität in Marburg)
2000-2004	Academic studies in biology (Ludwig-Maximilians-Universität in München)
2004-2005	Diploma thesis in biology (Max-Planck-Institut für Immunbiologie in Freiburg) Certificate: Diplom (Dipl. Biol. Univ.)
2005-2008	PhD studies in biology (Technische Universität in München) Aspired certificate: Dr. *rer. nat.*

8 REFERENCES

Ahlgren, J.D. (1996). Chemotherapy for pancreatic carcinoma. Cancer *78*, 654-663.

Alexander, J., and Stainier, D.Y. (1999). A molecular pathway leading to endoderm formation in zebrafish. Curr Biol *9*, 1147-1157.

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., and Constam, D.B. (2004). A gene network establishing polarity in the early mouse embryo. Semin Cell Dev Biol *15*, 555-561.

Antoch, M.P., Song, E.J., Chang, A.M., Vitaterna, M.H., Zhao, Y., Wilsbacher, L.D., Sangoram, A.M., King, D.P., Pinto, L.H., and Takahashi, J.S. (1997). Functional identification of the mouse circadian Clock gene by transgenic BAC rescue. Cell *89*, 655-667.

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.

Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2000). Biochemistry, mutagenesis, and oligomerization of DsRed, a red fluorescent protein from coral. Proc Natl Acad Sci U S A *97*, 11984-11989.

Beck, S., Le Good, J.A., Guzman, M., Ben Haim, N., Roy, K., Beermann, F., and Constam, D.B. (2002). Extraembryonic proteases regulate Nodal signalling during gastrulation. Nat Cell Biol *4*, 981-985.

Beddington, R.S. (1994). Induction of a second neural axis by the mouse node. Development *120*, 613-620.

Beddington, R.S., and Robertson, E.J. (1998). Anterior patterning in mouse. Trends Genet *14*, 277-284.

Beddington, R.S., and Robertson, E.J. (1999). Axis development and early asymmetry in mammals. Cell *96*, 195-209.

Bedford, F.K., Ashworth, A., Enver, T., and Wiedemann, L.M. (1993). HHEX: a novel homeobox gene expressed during haematopoiesis and conserved between mouse and human. Nucleic Acids Res *21*, 1245-1249.

Belo, J.A., Bouwmeester, T., Leyns, L., Kertesz, N., Gallo, M., Follettie, M., and De Robertis, E.M. (1997). Cerberus-like is a secreted factor with neutralizing activity expressed in the anterior primitive endoderm of the mouse gastrula. Mech Dev *68*, 45-57.

Bogue, C.W., Ganea, G.R., Sturm, E., Ianucci, R., and Jacobs, H.C. (2000). Hhex expression suggests a role in the development and function of organs derived from foregut endoderm. Dev Dyn *219*, 84-89.

Bogue, C.W., Zhang, P.X., McGrath, J., Jacobs, H.C., and Fuleihan, R.L. (2003). Impaired B cell development and function in mice with a targeted disruption of the homeobox gene Hhex. Proc Natl Acad Sci U S A *100*, 556-561.

Bohinski, R.J., Di Lauro, R., and Whitsett, J.A. (1994). The lung-specific surfactant protein B gene promoter is a target for thyroid transcription factor 1 and hepatocyte nuclear factor 3, indicating common factors for organ-specific gene expression along the foregut axis. Mol Cell Biol *14*, 5671-5681.

Bort, R., Martinez-Barbera, J.P., Beddington, R.S., and Zaret, K.S. (2004). Hhex homeobox gene-dependent tissue positioning is required for organogenesis of the ventral pancreas. Development *131*, 797-806.

Bort, R., Signore, M., Tremblay, K., Martinez Barbera, J.P., and Zaret, K.S. (2006). Hhex homeobox gene controls the transition of the endoderm to a pseudostratified, cell emergent epithelium for liver bud development. Dev Biol *290*, 44-56.

Bossard, P., and Zaret, K.S. (1998). GATA transcription factors as potentiators of gut endoderm differentiation. Development *125*, 4909-4917.

Bowles, J., Schepers, G., and Koopman, P. (2000). Phylogeny of the SOX family of developmental transcription factors based on sequence and structural indicators. Dev Biol *227*, 239-255.

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.

Bruno, M.D., Bohinski, R.J., Huelsman, K.M., Whitsett, J.A., and Korfhagen, T.R. (1995). Lung cell-specific expression of the murine surfactant protein A (SP-A) gene is mediated by interactions between the SP-A promoter and thyroid transcription factor-1. J Biol Chem *270*, 6531-6536.

Burtscher I. and Lickert H. (2009). Foxa2 regulates polarity and epithelialization in the endoderm germ layer of the mouse embryo. Development *136*, 1029-38.

Cai, L., Taylor, J.F., Wing, R.A., Gallagher, D.S., Woo, S.S., and Davis, S.K. (1995). Construction and characterization of a bovine bacterial artificial chromosome library. Genomics *29*, 413-425.

Calmont, A., Wandzioch, E., Tremblay, K.D., Minowada, G., Kaestner, K.H., Martin, G.R., and Zaret, K.S. (2006). An FGF response pathway that mediates hepatic gene induction in embryonic endoderm cells. Dev Cell *11*, 339-348.

Campbell, R.E., Tour, O., Palmer, A.E., Steinbach, P.A., Baird, G.S., Zacharias, D.A., and Tsien, R.Y. (2002). A monomeric red fluorescent protein. Proc Natl Acad Sci U S A *99*, 7877-7882.

Capdevila, J., Vogan, K.J., Tabin, C.J., Izpisue Belmonte, J.C. (2000). Mechanisms of left-right determination in vertebrates. Cell, 101, 9-21.

Chawengsaksophak, K., James, R., Hammond, V.E., Kontgen, F., and Beck, F. (1997). Homeosis and intestinal tumours in Cdx2 mutant mice. Nature *386*, 84-87.

Chazaud, C., and Rossant, J. (2006). Disruption of early proximodistal patterning and AVE formation in Apc mutants. Development *133*, 3379-3387.

Chazaud, C., Yamanaka, Y., Pawson, T., and Rossant, J. (2006). Early lineage segregation between epiblast and primitive endoderm in mouse blastocysts through the Grb2-MAPK pathway. Dev Cell *10*, 615-624.

Christ, F., Schoettler, S., Wende, W., Steuer, S., Pingoud, A., and Pingoud, V. (1999). The monomeric homing endonuclease PI-SceI has two catalytic centres for cleavage of the two strands of its DNA substrate. EMBO J *18*, 6908-6916.

Cirillo, A., Chifflet, S., and Villar, B. (2001). Neural retina of chick embryo in organ culture: effects of blockade of growth factors by suramin. Cell Tissue Res *304*, 323-331.

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.

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.

Copeland, N.G., Jenkins, N.A., and Court, D.L. (2001). Recombineering: a powerful new tool for mouse functional genomics. Nat Rev Genet *2*, 769-779.

Court, D.L., Sawitzke, J.A., and Thomason, L.C. (2002). Genetic engineering using homologous recombination. Annu Rev Genet *36*, 361-388.

Crompton, M.R., Bartlett, T.J., MacGregor, A.D., Manfioletti, G., Buratti, E., Giancotti, V., and Goodwin, G.H. (1992). Identification of a novel vertebrate homeobox gene expressed in haematopoietic cells. Nucleic Acids Res *20*, 5661-5667.

Dale, L., Howes, G., Price, B.M., and Smith, J.C. (1992). Bone morphogenetic protein 4: a ventralizing factor in early Xenopus development. Development *115*, 573-585.

Derynck, R., Gelbart, W.M., Harland, R.M., Heldin, C.H., Kern, S.E., Massague, J., Melton, D.A., Mlodzik, M., Padgett, R.W., Roberts, A.B.*, et al.* (1996). Nomenclature: vertebrate mediators of TGFbeta family signals. Cell *87*, 173.

Derynck, R., and Feng, X.H. (1997). TGF-beta receptor signaling. Biochim Biophys Acta *1333*, F105-150.

Derynck, R., and Zhang, Y.E. (2003). Smad-dependent and Smad-independent

pathways in TGF-beta family signalling. Nature *425*, 577-584.

Dessimoz, J., Opoka, R., Kordich, J.J., Grapin-Botton, A., and Wells, J.M. (2006). FGF signaling is necessary for establishing gut tube domains along the anterior-posterior axis in vivo. Mech Dev *123*, 42-55.

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.

Duprey, P., Chowdhury, K., Dressler, G.R., Balling, R., Simon, D., Guenet, J.L., and Gruss, P. (1988). A mouse gene homologous to the Drosophila gene caudal is expressed in epithelial cells from the embryonic intestine. Genes Dev *2*, 1647-1654.

Dutta, S., Bonner-Weir, S., Montminy, M., and Wright, C. (1998). Regulatory factor linked to late-onset diabetes? Nature *392*, 560.

Foley, A.C., and Mercola, M. (2005). Heart induction by Wnt antagonists depends on the homeodomain transcription factor Hhex. Genes Dev *19*, 387-396.

Frengen, E., Weichenhan, D., Zhao, B., Osoegawa, K., van Geel, M., and de Jong, P.J. (1999). A modular, positive selection bacterial artificial chromosome vector with multiple cloning sites. Genomics *58*, 250-253.

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.

Furukawa, Y., Shimada, T., Furuta, H., Matsuno, S., Kusuyama, A., Doi, A., Nishi, M., Sasaki, H., Sanke, T., and Nanjo, K. (2008). Polymorphisms in the IDE-KIF11-HHEX gene locus are reproducibly associated with type 2 diabetes in a Japanese population. J Clin Endocrinol Metab *93*, 310-314.

Gebhard, S., Hattori, T., Bauer, E., Bosl, M.R., Schlund, B., Poschl, E., Adam, N., de Crombrugghe, B., and von der Mark, K. (2007). BAC constructs in transgenic reporter mouse lines control efficient and specific LacZ expression in hypertrophic chondrocytes under the complete Col10a1 promoter. Histochem Cell Biol *127*, 183-194.

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.

Gong, S., Yang, X.W., Li, C., and Heintz, N. (2002). Highly efficient modification of bacterial artificial chromosomes (BACs) using novel shuttle vectors containing the R6Kgamma origin of replication. Genome Res *12*, 1992-1998.

Gong, S., Zheng, C., Doughty, M.L., Losos, K., Didkovsky, N., Schambra, U.B., Nowak, N.J., Joyner, A., Leblanc, G., Hatten, M.E.*, et al.* (2003). A gene expression atlas of the central nervous system based on bacterial artificial chromosomes. Nature *425*, 917-925.

Grarup, N., Rose, C.S., Andersson, E.A., Andersen, G., Nielsen, A.L., Albrechtsen, A., Clausen, J.O., Rasmussen, S.S., Jorgensen, T., Sandbaek, A.*, et al.* (2007). Studies of association of variants near the HHEX, CDKN2A/B, and IGF2BP2 genes with type 2 diabetes and impaired insulin release in 10,705 Danish subjects: validation and extension of genome-wide association studies. Diabetes *56*, 3105-3111.

Gu, Z., Reynolds, E.M., Song, J., Lei, H., Feijen, A., Yu, L., He, W., MacLaughlin D.T., van den Eijnden-van Raaij J., Donahoe, P.K., Li, E (1999). The type I serine/threonine kinase receptor ActRIA (ALK2) is required for gastrulation of the mouse embryo. Development *126*, 2551-61.

Gu, G., Dubauskaite, J., and Melton, D.A. (2002). Direct evidence for the pancreatic lineage: NGN3+ cells are islet progenitors and are distinct from duct progenitors. Development *129*, 2447-2457.

Gualdi, R., Bossard, P., Zheng, M., Hamada, Y., Coleman, J.R., and Zaret, K.S. (1996). Hepatic specification of the gut endoderm in vitro: cell signaling and transcriptional control. Genes Dev *10*, 1670-1682.

Guazzi, S., Price, M., De Felice, M., Damante, G., Mattei, M.G., and Di Lauro, R. (1990). Thyroid nuclear factor 1 (TTF-1) contains a homeodomain and displays a novel DNA binding specificity. EMBO J *9*, 3631-3639.

Guz, Y., Montminy, M.R., Stein, R., Leonard, J., Gamer, L.W., Wright, C.V., and Teitelman, G. (1995). Expression of murine STF-1, a putative insulin gene transcription factor, in beta cells of pancreas, duodenal epithelium and pancreatic exocrine and endocrine progenitors during ontogeny. Development *121*, 11-18.

Habener, J.F., and Stoffers, D.A. (1998). A newly discovered role of transcription factors involved in pancreas development and the pathogenesis of diabetes mellitus. Proc Assoc Am Physicians *110*, 12-21.

Hadjantonakis, A.K., Macmaster, S., and Nagy, A. (2002). Embryonic stem cells and mice expressing different GFP variants for multiple non-invasive reporter usage within a single animal. BMC Biotechnol *2*, 11.

Habener, J.F., Kemp, D.M., and Thomas, M.K. (2005). Minireview: transcriptional regulation in pancreatic development. Endocrinology *146*, 1025-1034.

Hanna, L.A., Foreman, R.K., Tarasenko, I.A., Kessler, D.S., and Labosky, P.A. (2002). Requirement for Foxd3 in maintaining pluripotent cells of the early mouse embryo. Genes Dev *16*, 2650-2661.

Harland, R., and Gerhart, J. (1997). Formation and function of Spemann's organizer. Annu Rev Cell Dev Biol *13*, 611-667.

Heintz, D., Wurtz, V., High, A.A., Van Dorsselaer, A., Reski, R., and Sarnighausen, E. (2004). An efficient protocol for the identification of protein phosphorylation in a seedless plant, sensitive enough to detect members of signalling cascades. Electrophoresis *25*, 1149-1159.

Heintz, D. (2004). Gene Expression Nervous System Atlas (GENSAT). Nature Neuroscience *7*, 483.

Heldin, C.H., Miyazono, K., and ten Dijke, P. (1997). TGF-beta signalling from cell membrane to nucleus through SMAD proteins. Nature *390*, 465-471.

Herder, C., Rathmann, W., Strassburger, K., Finner, H., Grallert, H., Huth, C., Meisinger, C., Gieger, C., Martin, S., Giani, G.*, et al.* (2008). Variants of the PPARG, IGF2BP2, CDKAL1, HHEX, and TCF7L2 genes confer risk of type 2 diabetes independently of BMI in the German KORA studies. Horm Metab Res *40*, 722-726.

Herrmann, B.G., Labeit, S., Poustka, A., King, T.R., and Lehrach, H. (1990). Cloning of the T gene required in mesoderm formation in the mouse. Nature *343*, 617-622.

Herrmann, B.G. (1991). Expression pattern of the Brachyury gene in whole-mount TWis/TWis mutant embryos. Development *113*, 913-917.

Hitz, C., Wurst, W., and Kuhn, R. (2007). Conditional brain-specific knockdown of MAPK using Cre/loxP regulated RNA interference. Nucleic Acids Res *35*, e90.

Ho, C.Y., Houart, C., Wilson, S.W., and Stainier, D.Y. (1999). A role for the extraembryonic yolk syncytial layer in patterning the zebrafish embryo suggested by properties of the Hhex gene. Curr Biol *9*, 1131-1134.

Hoess, R.H., Ziese, M., and Sternberg, N. (1982). P1 site-specific recombination: nucleotide sequence of the recombining sites. Proc Natl Acad Sci U S A *79*, 3398-3402.

Hogan, B.L., Thaller, C., and Eichele, G. (1992). Evidence that Hensen's node is a site of retinoic acid synthesis. Nature *359*, 237-241.

Hromas, R., Radich, J., and Collins, S. (1993). PCR cloning of an orphan homeobox gene (PRH) preferentially expressed in myeloid and liver cells. Biochem Biophys Res Commun *195*, 976-983.

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.

Inman, K.E., and Downs, K.M. (2006). Brachyury is required for elongation and vasculogenesis in the murine allantois. Development *133*, 2947-2959.

James, R., and Kazenwadel, J. (1991). Homeobox gene expression in the intestinal epithelium of adult mice. J Biol Chem *266*, 3246-3251.

Johnson, M.H., Maro, B., and Takeichi, M. (1986). The role of cell adhesion in the synchronization and orientation of polarization in 8-cell mouse blastomeres. J Embryol Exp Morphol *93*, 239-255.

Johnson, M.H., and McConnell, J.M. (2004). Lineage allocation and cell polarity during mouse embryogenesis. Semin Cell Dev Biol *15*, 583-597.

Jones, C.M., Lyons, K.M., Lapan, P.M., Wright, C.V., and Hogan, B.L. (1992). DVR-4

(bone morphogenetic protein-4) as a posterior-ventralizing factor in Xenopus mesoderm induction. Development *115*, 639-647.

Jonsson, J., Carlsson, L., Edlund, T., and Edlund, H. (1994). Insulin-promoter-factor 1 is required for pancreas development in mice. Nature *371*, 606-609.

Kaestner, K.H., Hiemisch, H., Luckow, B., and Schutz, G. (1994). The HNF-3 gene family of transcription factors in mice: gene structure, cDNA sequence, and mRNA distribution. Genomics *20*, 377-385.

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.*, et al.* (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. Development *129*, 2367-2379.

Kaufmann, E., and Knochel, W. (1996). Five years on the wings of fork head. Mech Dev *57*, 3-20.

Keng, V.W., Fujimori, K.E., Myint, Z., Tamamaki, N., Nojyo, Y., and Noguchi, T. (1998). Expression of Hhex mRNA in early murine postimplantation embryo development. FEBS Lett *426*, 183-186.

Keng, V.W., Yagi, H., Ikawa, M., Nagano, T., Myint, Z., Yamada, K., Tanaka, T., Sato, A., Muramatsu, I., Okabe, M.*, et al.* (2000). Homeobox gene Hhex is essential for onset of mouse embryonic liver development and differentiation of the monocyte lineage. Biochem Biophys Res Commun *276*, 1155-1161.

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

Kim, Y., and Nirenberg, M. (1989). Drosophila NK-homeobox genes. Proc Natl Acad Sci U S A *86*, 7716-7720.

Kimelman, D., and Griffin, K.J. (2000). Vertebrate mesendoderm induction and patterning. Curr Opin Genet Dev *10*, 350-356.

Kimura, S., Hara, Y., Pineau, T., Fernandez-Salguero, P., Fox, C.H., Ward, J.M., and Gonzalez, F.J. (1996). The T/ebp null mouse: thyroid-specific enhancer-binding protein is essential for the organogenesis of the thyroid, lung, ventral forebrain, and pituitary. Genes Dev *10*, 60-69.

Kimura, C., Yoshinaga, K., Tian, E., Suzuki, M., Aizawa, S., and Matsuo, I. (2000). Visceral endoderm mediates forebrain development by suppressing posteriorizing signals. Dev Biol *225*, 304-321.

Kimura, C., Shen, M.M., Takeda, N., Aizawa, S., and Matsuo, I. (2001). Complementary functions of Otx2 and Cripto in initial patterning of mouse epiblast. Dev Biol *235*, 12-32.

Kinder, S.J., Tsang, T.E., Quinlan, G.A., Hadjantonakis, A.K., 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., Ang, S.L., Behringer, R.R., and Tam, P.P. (2001). Defects of the body plan of mutant embryos lacking Lim1, Otx2 or Hnf3beta activity. Int J Dev Biol *45*, 347-355.

Kishigami, S. and Mishina, Y. (2005). BMP signaling and early embryonic patterning. Cytokine & Growth Factor Rev. *16*, 265-278.

Kispert, A., Koschorz, B., and Herrmann, B.G. (1995). The T protein encoded by Brachyury is a tissue-specific transcription factor. EMBO J *14*, 4763-4772.

Kitajima, S., Takagi, A., Inoue, T., and Saga, Y. (2000). MesP1 and MesP2 are essential for the development of cardiac mesoderm. Development *127*, 3215-3226.

Krude, H., Schutz, B., Biebermann, H., von Moers, A., Schnabel, D., Neitzel, H., Tonnies, H., Weise, D., Lafferty, A., Schwarz, S.*, et al.* (2002). Choreoathetosis, hypothyroidism, and pulmonary alterations due to human NKX2-1 haploinsufficiency. J Clin Invest *109*, 475-480.

Kubo, A., Shinozaki, K., Shannon, J.M., Kouskoff, V., Kennedy, M., Woo, S., Fehling, H.J., and Keller, G. (2004). Development of definitive endoderm from embryonic stem cells in culture. Development *131*, 1651-1662.

Kubo, A., Chen, V., Kennedy, M., Zahradka, E., Daley, G.Q., and Keller, G. (2005). The homeobox gene HHEX regulates proliferation and differentiation of hemangioblasts and endothelial cells during ES cell differentiation. Blood *105*, 4590-4597.

Lai, E., Prezioso, V.R., Tao, W.F., Chen, W.S., and Darnell, J.E., Jr. (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.

Lantz, K.A., Vatamaniuk, M.Z., Brestelli, J.E., Friedman, J.R., Matschinsky, F.M., and Kaestner, K.H. (2004). Foxa2 regulates multiple pathways of insulin secretion. J Clin Invest *114*, 512-520.

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.

Lazzaro, D., Price, M., de Felice, M., and Di Lauro, R. (1991). The transcription factor TTF-1 is expressed at the onset of thyroid and lung morphogenesis and in restricted regions of the foetal brain. Development *113*, 1093-1104.

Lee, C.S., Sund, N.J., Vatamaniuk, M.Z., Matschinsky, F.M., Stoffers, D.A., and Kaestner, K.H. (2002). Foxa2 controls Pdx1 gene expression in pancreatic beta-cells in vivo. Diabetes *51*, 2546-2551.

Lee, C.S., Sund, N.J., Behr, R., Herrera, P.L., and Kaestner, K.H. (2005). Foxa2 is required for the differentiation of pancreatic alpha-cells. Dev Biol *278*, 484-495.

Lee, E.C., Yu, D., Martinez de Velasco, J., Tessarollo, L., Swing, D.A., Court, D.L., Jenkins, N.A., and Copeland, N.G. (2001). A highly efficient Escherichia coli-based chromosome engineering system adapted for recombinogenic targeting and subcloning of BAC DNA. Genomics *73*, 56-65.

Lefebvre, V., Dumitriu, B., Penzo-Mendez, A., Han, Y., and Pallavi, B. (2007). Control of cell fate and differentiation by Sry-related high-mobility-group box (Sox) transcription factors. Int J Biochem Cell Biol *39*, 2195-2214.

Lemaigre, F., and Zaret, K.S. (2004). Liver development update: new embryo models, cell lineage control, and morphogenesis. Curr Opin Genet Dev *14*, 582-590.

Levinson-Dushnik, M., and Benvenisty, N. (1997). Involvement of hepatocyte nuclear factor 3 in endoderm differentiation of embryonic stem cells. Mol Cell Biol *17*, 3817-3822.

Lewis, S.L., and Tam, P.P. (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.

Liu, C., Glasser, S.W., Wan, H., and Whitsett, J.A. (2002). GATA-6 and thyroid transcription factor-1 directly interact and regulate surfactant protein-C gene expression. J Biol Chem *277*, 4519-4525.

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, P., Jenkins, N.A., and Copeland, N.G. (2003). A highly efficient recombineering-based method for generating conditional knockout mutations. Genome Res *13*, 476-484.

MacFarlane, W.M., Read, M.L., Gilligan, M., Bujalska, I., and Docherty, K. (1994). Glucose modulates the binding activity of the beta-cell transcription factor IUF1 in a phosphorylation-dependent manner. Biochem J *303 (Pt 2)*, 625-631.

MacFarlane, W.M., Smith, S.B., James, R.F., Clifton, A.D., Doza, Y.N., Cohen, P., and Docherty, K. (1997). The p38/reactivating kinase mitogen-activated protein kinase cascade mediates the activation of the transcription factor insulin upstream factor 1 and insulin gene transcription by high glucose in pancreatic beta-cells. J Biol Chem *272*, 20936-20944.

MacMurray, A., and Shin, H.S. (1988). The antimorphic nature of the Tc allele at the mouse T locus. Genetics *120*, 545-550.

Maquat, L.E. (2002). Nonsense-mediated mRNA decay. Curr Biol *12*, R196-197.

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.

Martinez Barbera, J.P., Clements, M., Thomas, P., Rodriguez, T., Meloy, D., Kioussis, D., and Beddington, R.S. (2000). The homeobox gene Hhex is required in definitive endodermal tissues for normal forebrain, liver and thyroid formation. Development *127*, 2433-2445.

Martinez, D.E., Dirksen, M.L., Bode, P.M., Jamrich, M., Steele, R.E., and Bode, H.R. (1997). Budhead, a fork head/HNF-3 homologue, is expressed during axis formation and head specification in hydra. Dev Biol *192*, 523-536.

Mason, R.J., Williams, M.C., Moses, H.L., Mohla, S., and Berberich, M.A. (1997). Stem cells in lung development, disease, and therapy. Am J Respir Cell Mol Biol *16*, 355-363.

Massague, J. (1998). TGF-beta signal transduction. Annu Rev Biochem *67*, 753-791.

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.

McLin, V.A., Rankin, S.A., and Zorn, A.M. (2007). Repression of Wnt/beta-catenin signaling in the anterior endoderm is essential for liver and pancreas development. Development *134*, 2207-2217.

Mejia, J.E., and Larin, Z. (2000). The assembly of large BACs by in vivo recombination. Genomics *70*, 165-170.

Meno, C., Gritsman, K., Ohishi, S., Ohfuji, Y., Heckscher, E., Mochida, K., Shimono, A., Kondoh, H., Talbot, W.S., Robertson, E.J.*, et al.* (1999). Mouse Lefty2 and zebrafish antivin are feedback inhibitors of nodal signaling during vertebrate gastrulation. Mol Cell *4*, 287-298.

Mesnard, D., Guzman-Ayala, M., and Constam, D.B. (2006). Nodal specifies embryonic visceral endoderm and sustains pluripotent cells in the epiblast before overt axial patterning. Development *133*, 2497-2505.

Minoo, P., Su, G., Drum, H., Bringas, P., and Kimura, S. (1999). Defects in tracheoesophageal and lung morphogenesis in Nkx2.1(-/-) mouse embryos. Dev Biol *209*, 60-71.

Mishina, Y., Crombie, R., Bradley, A., Behringer, R.R. (1999). Multiple roles for activin-like kinase-2 signaling during mouse embryogenesis. Dev Biol *213*, 314-26.

Mitsui, K., Tokuzawa, Y., Itoh, H., Segawa, K., Murakami, M., Takahashi, K., Maruyama, M., Maeda, M., and Yamanaka, S. (2003). The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. Cell *113*,

631-642.

Miyawaki, A., Llopis, J., Heim, R., McCaffery, J.M., Adams, J.A., Ikura, M., and Tsien, R.Y. (1997). Fluorescent indicators for Ca2+ based on green fluorescent proteins and calmodulin. Nature *388*, 882-887.

Mizuno, K., Gonzalez, F.J., and Kimura, S. (1991). Thyroid-specific enhancer-binding protein (T/EBP): cDNA cloning, functional characterization, and structural identity with thyroid transcription factor TTF-1. Mol Cell Biol *11*, 4927-4933.

Monaghan, A.P., Kaestner, K.H., Grau, E., and Schütz, 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.

Morrison, G.M., Oikonomopoulou, I., Migueles, R.P., Soneji, S., Livigni, A., Enver, T., Brickman, J.M. (2008). Anterior definitive endoderm from ESCs reveals a role for FGF signaling. Cell Stem Cell *3*, 355-6.

Muyrers, J.P., Zhang, Y., Testa, G., and Stewart, A.F. (1999). Rapid modification of bacterial artificial chromosomes by ET-recombination. Nucleic Acids Res *27*, 1555-1557.

Nagai, T., Ibata, K., Park, E.S., Kubota, M., Mikoshiba, K., and Miyawaki, A. (2002). A variant of yellow fluorescent protein with fast and efficient maturation for cell-biological applications. Nat Biotechnol *20*, 87-90.

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.

Nakamura, S., Asakawa, S., Ohmido, N., Fukui, K., Shimizu, N., and Kawasaki, S. (1997). Construction of an 800-kb contig in the near-centromeric region of the rice blast resistance gene Pi-ta2 using a highly representative rice BAC library. Mol Gen Genet *254*, 611-620.

Newman, C.S., Chia, F., and Krieg, P.A. (1997). The XHhex homeobox gene is expressed during development of the vascular endothelium: overexpression leads to an increase in vascular endothelial cell number. Mech Dev *66*, 83-93.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. Cell *95*, 379-391.

Nielsen, L.B., McCormick, S.P., Pierotti, V., Tam, C., Gunn, M.D., Shizuya, H., and Young, S.G. (1997). Human apolipoprotein B transgenic mice generated with 207- and 145-kilobase pair bacterial artificial chromosomes. Evidence that a distant 5'-element confers appropriate transgene expression in the intestine. J Biol Chem *272*, 29752-29758.

Nishida, W., Nakamura, M., Mori, S., Takahashi, M., Ohkawa, Y., Tadokoro, S.,

Yoshida, K., Hiwada, K., Hayashi, K., and Sobue, K. (2002). A triad of serum response factor and the GATA and NK families governs the transcription of smooth and cardiac muscle genes. J Biol Chem *277*, 7308-7317.

Nistala, R., and Sigmund, C.D. (2002). A reliable and efficient method for deleting operational sequences in PACs and BACs. Nucleic Acids Res *30*, e41.

Offield, M.F., Jetton, T.L., Labosky, P.A., Ray, M., Stein, R.W., Magnuson, M.A., Hogan, B.L., and Wright, C.V. (1996). PDX-1 is required for pancreatic outgrowth and differentiation of the rostral duodenum. Development *122*, 983-995.

Oliver, G., Sosa-Pineda, B., Geisendorf, S., Spana, E.P., Doe, C.Q., and Gruss, P. (1993). Prox 1, a prospero-related homeobox gene expressed during mouse development. Mech Dev *44*, 3-16.

Overdier, D.G., Porcella, A., and Costa, R.H. (1994). The DNA-binding specificity of the hepatocyte nuclear factor 3/forkhead domain is influenced by amino-acid residues adjacent to the recognition helix. Mol Cell Biol *14*, 2755-2766.

Oyama, Y., Kawai-Kowase, K., Sekiguchi, K., Sato, M., Sato, H., Yamazaki, M., Ohyama, Y., Aihara, Y., Iso, T., Okamaoto, E.*, et al.* (2004). Homeobox protein Hhex facilitates serum responsive factor-mediated activation of the SM22alpha gene transcription in embryonic fibroblasts. Arterioscler Thromb Vasc Biol *24*, 1602-1607.

Park, K.S., Wells, J.M., Zorn, A.M., Wert, S.E., and Whitsett, J.A. (2006). Sox17 influences the differentiation of respiratory epithelial cells. Dev Biol *294*, 192-202.

Perea-Gomez, A., Rhinn, M., and Ang, S.L. (2001). Role of the anterior visceral endoderm in restricting posterior signals in the mouse embryo. Int J Dev Biol *45*, 311-320.

Perea-Gomez, A., Vella, F.D., 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.

Petersen, H.V., Serup, P., Leonard, J., Michelsen, B.K., and Madsen, O.D. (1994). Transcriptional regulation of the human insulin gene is dependent on the homeodomain protein STF1/IPF1 acting through the CT boxes. Proc Natl Acad Sci U S A *91*, 10465-10469.

Petersen, H.V., Peshavaria, M., Pedersen, A.A., Philippe, J., Stein, R., Madsen, O.D., and Serup, P. (1998). Glucose stimulates the activation domain potential of the PDX-1 homeodomain transcription factor. FEBS Lett *431*, 362-366.

Rafiq, M., Suen, C.K., Choudhury, N., Joannou, C.L., White, K.N., and Evans, R.W. (1997). Expression of recombinant human ceruloplasmin--an absolute requirement for splicing signals in the expression cassette. FEBS Lett *407*, 132-136.

Ramazzina, I., Folli, C., Secchi, A., Berni, R., and Percudani, R. (2006). Completing the uric acid degradation pathway through phylogenetic comparison of whole

genomes. Nat Chem Biol *2*, 144-148.

Reddi, A.H. (1994). Bone and cartilage differentiation. Curr Opin Genet Dev *4*, 737-744.

Rivera-Perez, J.A., Mager, J., and Magnuson, T. (2003). Dynamic morphogenetic events characterize the mouse visceral endoderm. Dev Biol *261*, 470-487.

Rivera-Perez, J.A., and Magnuson, T. (2005). Primitive streak formation in mice is preceded by localized activation of Brachyury and Wnt3. Dev Biol *288*, 363-371.

Rodaway, A., Takeda, H., Koshida, S., Broadbent, J., Price, B., Smith, J.C., Patient, R., and Holder, N. (1999). Induction of the mesendoderm in the zebrafish germ ring by yolk cell-derived TGF-beta family signals and discrimination of mesoderm and endoderm by FGF. Development *126*, 3067-3078.

Rodaway, A., and Patient, R. (2001). Mesendoderm. an ancient germ layer? Cell *105*, 169-172.

Rodriguez, T.A., Casey, E.S., Harland, R.M., Smith, J.C., and Beddington, R.S. (2001). Distinct enhancer elements control Hhex expression during gastrulation and early organogenesis. Dev Biol *234*, 304-316.

Rosnet, O., Marchetto, S., deLapeyriere, O., and Birnbaum, D. (1991). Murine Flt3, a gene encoding a novel tyrosine kinase receptor of the PDGFR/CSF1R family. Oncogene *6*, 1641-1650.

Ross, S.A., McCaffery, P.J., Drager, U.C., and De Luca, L.M. (2000). Retinoids in embryonal development. Physiol Rev *80*, 1021-1054.

Rossant, J., Chazaud, C., and Yamanaka, Y. (2003). Lineage allocation and asymmetries in the early mouse embryo. In Philos Trans R Soc Lond B Biol Sci, pp. 1341-1348; discussion 1349.

Rossant, J. and Tam, P.P. (2009). Blastocyst lineage formation, early embryonic asymmetries and axis patterning in the mouse. Development *136*, 701-13.

Rossi, J.M., Dunn, N.R., Hogan, B.L., and Zaret, K.S. (2001). Distinct mesodermal signals, including BMPs from the septum transversum mesenchyme, are required in combination for hepatogenesis from the endoderm. Genes Dev *15*, 1998-2009.

Ruiz i Altaba, A. (1994). Pattern formation in the vertebrate neural plate. Trends Neurosci *17*, 233-243.

Sampath, T.K., and Reddi, A.H. (1981). Dissociative extraction and reconstitution of extracellular matrix components involved in local bone differentiation. Proc Natl Acad Sci U S A *78*, 7599-7603.

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.

Sawaya, P.L., Stripp, B.R., Whitsett, J.A., and Luse, D.S. (1993). The lung-specific CC10 gene is regulated by transcription factors from the AP-1, octamer, and hepatocyte nuclear factor 3 families. Mol Cell Biol *13*, 3860-3871.

Scheufler, H. (1969). [An additional house mouse mutant with anemia (hemoglobin deficiency)]. Z Versuchstierkd *11*, 348-353.

Schulte-Merker, S., van Eeden, F.J., Halpern, M.E., Kimmel, C.B., and Nusslein-Volhard, C. (1994). no tail (ntl) is the zebrafish homologue of the mouse T (Brachyury) gene. Development *120*, 1009-1015.

Serls, A.E., Doherty, S., Parvatiyar, P., Wells, J.M., and Deutsch, G.H. (2005). Different thresholds of fibroblast growth factors pattern the ventral foregut into liver and lung. Development *132*, 35-47.

Shaner, N.C., Campbell, R.E., Steinbach, P.A., Giepmans, B.N., Palmer, A.E., and Tsien, R.Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. Nat Biotechnol *22*, 1567-1572.

Shih, J., and Fraser, S.E. (1996). Characterizing the zebrafish organizer: microsurgical analysis at the early-shield stage. Development *122*, 1313-1322.

Shimomura, O., Johnson, F.H., and Saiga, Y. (1962). Extraction, purification and properties of aequorin, a bioluminescent protein from the luminous hydromedusan, Aequorea. J Cell Comp Physiol *59*, 223-239.

Shiojiri, N. (1984). The origin of intrahepatic bile duct cells in the mouse. J Embryol Exp Morphol *79*, 25-39.

Shiojiri, N., Inujima, S., Ishikawa, K., Terada, K., and Mori, M. (2001). Cell lineage analysis during liver development using the spf(ash)-heterozygous mouse. Lab Invest *81*, 17-25.

Shizuya, H., Birren, B., Kim, U.J., Mancino, V., Slepak, T., Tachiiri, Y., and Simon, M. (1992). Cloning and stable maintenance of 300-kilobase-pair fragments of human DNA in Escherichia coli using an F-factor-based vector. Proc Natl Acad Sci U S A *89*, 8794-8797.

Singh, G., Kaur, S., Stock, J.L., Jenkins, N.A., Gilbert, D.J., Copeland, N.G., and Potter, S.S. (1991). Identification of 10 murine homeobox genes. Proc Natl Acad Sci U S A *88*, 10706-10710.

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.

Smith, J.C., Price, B.M., Green, J.B., Weigel, D., and Herrmann, B.G. (1991). Expression of a Xenopus homolog of Brachyury (T) is an immediate-early response to mesoderm induction. Cell *67*, 79-87.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J Mol Biol *98*, 503-517.

Srinivas, S., Rodriguez, T., Clements, M., Smith, J.C., and Beddington, R.S. (2004). Active cell migration drives the unilateral movements of the anterior visceral endoderm. Development *131*, 1157-1164.

Srinivas, S. (2006). The anterior visceral endoderm-turning heads. Genesis *44*, 565-572.

Staiger, H., Stancakova, A., Zilinskaite, J., Vanttinen, M., Hansen, T., Marini, M.A., Hammarstedt, A., Jansson, P.A., Sesti, G., Smith, U*., et al.* (2008). A candidate type 2 diabetes polymorphism near the HHEX locus affects acute glucose-stimulated insulin release in European populations: results from the EUGENE2 study. Diabetes *57*, 514-517.

Sternberg, N., and Hamilton, D. (1981). Bacteriophage P1 site-specific recombination. I. Recombination between loxP sites. J Mol Biol *150*, 467-486.

Sternberg, N., Hamilton, D., Austin, S., Yarmolinsky, M., and Hoess, R. (1981). Site-specific recombination and its role in the life cycle of bacteriophage P1. Cold Spring Harb Symp Quant Biol *45 Pt 1*, 297-309.

Sternberg, N., Hamilton, D., and Hoess, R. (1981). Bacteriophage P1 site-specific recombination. II. Recombination between loxP and the bacterial chromosome. J Mol Biol *150*, 487-507.

Stoffers, D.A., Thomas, M.K., and Habener, J.F. (1997). Homeodomain Protein IDX-1 A Master Regulator of Pancreas Development and Insulin Gene Expression. Trends Endocrinol Metab *8*, 145-151.

Strahle, U., Blader, P., Henrique, D., and Ingham, P.W. (1993). Axial, a zebrafish gene expressed along the developing body axis, shows altered expression in cyclops mutant embryos. Genes Dev *7*, 1436-1446.

Strumpf, D., Mao, C.A., Yamanaka, Y., Ralston, A., Chawengsaksophak, K., Beck, F., and Rossant, J. (2005). Cdx2 is required for correct cell fate specification and differentiation of trophectoderm in the mouse blastocyst. Development *132*, 2093-2102.

Sund, N.J., Vatamaniuk, M.Z., Casey, M., Ang, S.L., Magnuson, M.A., Stoffers, D.A., Matschinsky, F.M., and Kaestner, K.H. (2001). Tissue-specific deletion of Foxa2 in pancreatic beta cells results in hyperinsulinemic hypoglycemia. Genes Dev *15*, 1706-1715.

Tada, S., Era, T., Furusawa, C., Sakurai, H., Nishikawa, S., Kinoshita, M., Nakao, K., Chiba, T., and Nishikawa, S. (2005). Characterization of mesendoderm: a diverging point of the definitive endoderm and mesoderm in embryonic stem cell differentiation culture. Development *132*, 4363-4374.

Takakura, N., Kodama, H., Nishikawa, S., and Nishikawa, S. (1996). Preferential

proliferation of murine colony-forming units in culture in a chemically defined condition with a macrophage colony-stimulating factor-negative stromal cell clone. J Exp Med *184*, 2301-2309.

Tam, P.P., and Beddington, R.S. (1992). Establishment and organization of germ layers in the gastrulating mouse embryo. Ciba Found Symp *165*, 27-41; discussion 42-29.

Tam, P.P., and Behringer, R.R. (1997). Mouse gastrulation: the formation of a mammalian body plan. Mech Dev *68*, 3-25.

Tam, P.P., Steiner, K.A., Zhou, S.X., and Quinlan, G.A. (1997). Lineage and functional analyses of the mouse organizer. Cold Spring Harb Symp Quant Biol *62*, 135-144.

Tam, P.P., Kanai-Azuma, M., and Kanai, Y. (2003). Early endoderm development in vertebrates: lineage differentiation and morphogenetic function. Curr Opin Genet Dev *13*, 393-400.

Tam, P.P., Loebel, D.A., and Tanaka, S.S. (2006). Building the mouse gastrula: signals, asymmetry and lineages. Curr Opin Genet Dev *16*, 419-425.

Tam, P.P., and Loebel, D.A. (2007). Gene function in mouse embryogenesis: get set for gastrulation. Nat Rev Genet *8*, 368-381.

Tanaka, T., Inazu, T., Yamada, K., Myint, Z., Keng, V.W., Inoue, Y., Taniguchi, N., and Noguchi, T. (1999). cDNA cloning and expression of rat homeobox gene, Hhex, and functional characterization of the protein. Biochem J *339 (Pt 1)*, 111-117.

Ten Have-Opbroek, A.A. (1981). The development of the lung in mammals: an analysis of concepts and findings. Am J Anat *162*, 201-219.

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). Hhex: a homeobox gene revealing peri-implantation asymmetry in the mouse embryo and an early transient marker of endothelial cell precursors. Development *125*, 85-94.

Tomishima, M.J., Hadjantonakis, A.K., Gong, S., Studer, L. (2007). Production of Green Fluorescent Protein Transgenic Embryonic Stem Cells Using the GENSAT Bacterial Artificial Chromosome Library. Stem Cell 25, 39-45.

Toyooka, Y., Shimosato, D., Murakami, K., Takahashi, K., and Niwa, H. (2008). Identification and characterization of subpopulations in undifferentiated ES cell culture. Development *135*, 909-918.

Tremblay, K.D., Dunn, N.R. and Robertson, E.J. (2001) Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. Development *128*, 3601-21.

Tremblay, K.D., and Zaret, K.S. (2005). Distinct populations of endoderm cells converge to generate the embryonic liver bud and ventral foregut tissues. Dev Biol *280*, 87-99.

Tsien, R.Y. (1998). The green fluorescent protein. Annu Rev Biochem *67*, 509-544.

Urist, M.R. (1965). Bone: formation by autoinduction. Science *150*, 893-899.

Valenzuela, D.M., Murphy, A.J., Frendewey, D., Gale, N.W., Economides, A.N., Auerbach, W., Poueymirou, W.T., Adams, N.C., Rojas, J., Yasenchak, J.*, et al.* (2003). High-throughput engineering of the mouse genome coupled with high-resolution expression analysis. Nat Biotechnol *21*, 652-659.

Vintersten, K., Monetti, C., Gertsenstein, M., Zhang, P., Laszlo, L., Biechele, S., and Nagy, A. (2004). Mouse in red: red fluorescent protein expression in mouse ES cells, embryos, and adult animals. Genesis *40*, 241-246.

Waddington, C.H., and Waterman, A.J. (1933). The Development in vitro of Young Rabbit Embryos. J Anat *67*, 355-370.

Waeber, G., Thompson, N., Nicod, P., and Bonny, C. (1996). Transcriptional activation of the GLUT2 gene by the IPF-1/STF-1/IDX-1 homeobox factor. Mol Endocrinol *10*, 1327-1334.

Wan, H., Xu, Y., Ikegami, M., Stahlman, M.T., Kaestner, K.H., Ang, S.L., and Whitsett, J.A. (2004). Foxa2 is required for transition to air breathing at birth. Proc Natl Acad Sci U S A *101*, 14449-14454.

Wan, H., Dingle, S., Xu, Y., Besnard, V., Kaestner, K.H., Ang, S.L., Wert, S., Stahlman, M.T., and Whitsett, J.A. (2005). Compensatory roles of Foxa1 and Foxa2 during lung morphogenesis. J Biol Chem *280*, 13809-13816.

Wang, H., and Dey, S.K. (2006). Roadmap to embryo implantation: clues from mouse models. Nat Rev Genet *7*, 185-199.

Wang, Z., Engler, P., Longacre, A., and Storb, U. (2001). An efficient method for high-fidelity BAC/PAC retrofitting with a selectable marker for mammalian cell transfection. Genome Res *11*, 137-142.

Watada, H., Kajimoto, Y., Miyagawa, J., Hanafusa, T., Hamaguchi, K., Matsuoka, T., Yamamoto, K., Matsuzawa, Y., Kawamori, R., and Yamasaki, Y. (1996). PDX-1 induces insulin and glucokinase gene expressions in alphaTC1 clone 6 cells in the presence of betacellulin. Diabetes *45*, 1826-1831.

Weidenfeld, J., Shu, W., Zhang, L., Millar, S.E., and Morrisey, E.E. (2002). The WNT7b promoter is regulated by TTF-1, GATA6, and Foxa2 in lung epithelium. J Biol Chem *277*, 21061-21070.

Weigel, D., and Jackle, H. (1990). The fork head domain: a novel DNA binding motif of eukaryotic transcription factors? Cell *63*, 455-456.

Weinstein, D.C., Ruiz i Altaba, A., Chen, W.S., Hoodless, P., Prezioso, V.R., Jessell, T.M., and Darnell, J.E., Jr. (1994). The winged-helix transcription factor HNF-3 beta is required for notochord development in the mouse embryo. Cell *78*, 575-588.

Wells, J.M., and Melton, D.A. (1999). Vertebrate endoderm development. Annu Rev Cell Dev Biol *15*, 393-410.

Wells, J.M., and Melton, D.A. (2000). Early mouse endoderm is patterned by soluble factors from adjacent germ layers. Development *127*, 1563-1572.

Wilkinson, D.G., Bhatt, S., and Herrmann, B.G. (1990). Expression pattern of the mouse T gene and its role in mesoderm formation. Nature *343*, 657-659.

Winter, H., Siry, P., Tobiasch, E., and Schweizer, J. (1994). Sequence and expression of murine type I hair keratins mHa2 and mHa3. Exp Cell Res *212*, 190-200.

Wodarz, A., Stewart, D.B., Nelson, W.J., and Nusse, R. (2006). Wingless signaling modulates cadherin-mediated cell adhesion in Drosophila imaginal disc cells. J Cell Sci *119*, 2425-2434.

Wrana, J., and Pawson, T. (1997). Signal transduction. Mad about SMADs. Nature *388*, 28-29.

Yamaguchi, T.P., and Rossant, J. (1995). Fibroblast growth factors in mammalian development. Curr Opin Genet Dev *5*, 485-491.

Yang, X.W., Model, P., and Heintz, N. (1997). Homologous recombination based modification in Escherichia coli and germline transmission in transgenic mice of a bacterial artificial chromosome. Nat Biotechnol *15*, 859-865.

Yang, Y., and Seed, B. (2003). Site-specific gene targeting in mouse embryonic stem cells with intact bacterial artificial chromosomes. Nat Biotechnol *21*, 447-451.

Yatskievych, T.A., Pascoe, S., and Antin, P.B. (1999). Expression of the homebox gene Hhex during early stages of chick embryo development. Mech Dev *80*, 107-109.

Yu, D., Ellis, H.M., Lee, E.C., Jenkins, N.A., Copeland, N.G., and Court, D.L. (2000). An efficient recombination system for chromosome engineering in Escherichia coli. Proc Natl Acad Sci U S A *97*, 5978-5983.

Yuan, B., Li, C., Kimura, S., Engelhardt, R.T., Smith, B.R., and Minoo, P. (2000). Inhibition of distal lung morphogenesis in Nkx2.1(-/-) embryos. Dev Dyn *217*, 180-190.

Zacharias, D.A., Violin, J.D., Newton, A.C., and Tsien, R.Y. (2002). Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science *296*, 913-916.

Zamparini, A.L., Watts, T., Gardner, C.E., Tomlinson, S.R., Johnston, G.I., and

Brickman, J.M. (2006). Hhex acts with beta-catenin to regulate anteroposterior patterning via a Groucho-related co-repressor and Nodal. Development *133*, 3709-3722.

Zaret, K.S. (2001). Hepatocyte differentiation: from the endoderm and beyond. Curr Opin Genet Dev *11*, 568-574.

Zaret, K.S. (2002). Regulatory phases of early liver development: paradigms of organogenesis. Nat Rev Genet *3*, 499-512.

Zaret, K.S. (2004). Regenerative medicine: self-help for insulin cells. Nature *429*, 30-31.

Zhang, L., Rubins, N.E., Ahima, R.S., Greenbaum, L.E., and Kaestner, K.H. (2005). Foxa2 integrates the transcriptional response of the hepatocyte to fasting. Cell Metab *2*, 141-148.

Zhou, X., Sasaki, H., Lowe, L., Hogan, B.L., and Kuehn, M.R. (1993). Nodal is a novel TGF-beta-like gene expressed in the mouse node during gastrulation. Nature *361*, 543-547.

Zwijsen, A., Verschueren, K., and Huylebroeck, D. (2003). New intracellular components of bone morphogenetic protein/Smad signaling cascades. FEBS Lett *546*, 133-139.

9 ACKNOWLEDGMENTS

Ich danke allen, die mir während der Promotion und während des Studiums seelisch und moralisch sowie ganz praktisch mit Rat und Tat geholfen haben. Dank an meinen Eltern ohne die das Studium nicht möglich gewesen wäre und an meine Schwester. Danke an Carsten, Wally, Nils, Dunja, Nadine, Claudio, Steffi und Grisi die mich während des Studiums unterstützt haben. Vielen Dank an Ingo, Claude, Patrizia, Wenke und das ganze Lickert Labor das mir während der Promotion mit Tipps, Tricks und allerlei Diskussionen geholfen hat. Ganz besonders möchte ich Barbara für die vergangenen zwei Jahre danken.

Vielen Dank an Dr. Heiko Lickert, der mir ermöglicht hat dieses Projekt zu übernehmen und mir immer wieder mit Anregungen und Vorschlägen durch die Promotion geholfen hat.

Herrn Prof. Jochen Graw danke ich für die Übernahme dieser Arbeit als „Doktorvater“, für seine gewissenhafte Betreuung während der gesamten Zeit sowie seine Vorschläge und Anregungen.

Frau Prof. Angelika Schnieke sowie Herrn Prof. Siegfried Scherer danke ich für die unkomplitzierte Übernahme der 2. Prüferin sowie des Vorsitzenden der Prüfungskommission.

10 SUMMARY / ZUSAMMENFASSUNG

10.1 Summary

The scope of this work was to study endoderm development in mouse embryonic stem (ES) cells and in mouse embryos to get new insights about endoderm formation. The aim was to reveal where endodermal cell populations appear and migrate from the beginning of gastrulation until early organogenesis. Therefore, the BAC recombineering method was improved to stably introduce fluorescent proteins in ES cells under the control of specific endodermal genes.

BAC transgenes were generated that contained either the earliest marker genes of endoderm precursor cell populations , *Foxa2* and *Sox17,* or the endodermal-derived organ primordial of lung, liver and pancreas, *Nkx2.1*, *Hhex* and *Pdx1*, respectively. The generated BAC reporter transgenic ES cells were analyzed for the expected cell-population-specific and time-specific expression of the fluorescent protein by *in vitro* differentiation.

Complete ES cell-derived embryos and BAC transgenic mouse lines were further generated using tetraploid aggregation technique to show how the reporter activity reflected endogenous gene expression pattern in *in vivo* live imaging under a confocal microscope. In this way, it was demonstrated that a one copy insertion of a fluorescent reporter could be sufficient for static analysis in live imaging approaches during early embryogenesis.

In particular, a new *Hhex::Lyn-tomato* reporter line was working nicely to study *Hhex* transgene expression in static analysis and in live imaging approaches for lineage allocation, cell population analysis and cell migration studies. Furthermore, *Hhex* expression during blastocyst stage in the ICM and between E14.5 to E18.5 in the fore- and hindlimb was revealed, that was never reported before. Moreover, the question whether *Foxa2* and *Sox17* progenitors do give rise to all endoderm derived organs or if they constitute only endodermal subpopulations could also be answered. Namely, it was shown that *Foxa2*, like also *Sox17*, progenitors do not give rise to all endoderm derived organs; in fact, for example, *Foxa2* has to be first downregulated

to drive endodermal cells towards hepatic cell fate in the anterior definitive endoderm.

A whole mount *Foxa2* immunostaining on the *Hhex::Lyn-tomato* BAC transgenic mouse line was also established, that enabled the mapping of endoderm populations on the cellular level. In this way, it was shown that epiblast cells are already expressing *Foxa2* before intercalating in the definitive endoderm. This demonstrated that endodermal cells are determined before they intercalate into the visceral endoderm and form the definitive endoderm.

Whole mount antibody stainings were also used for the analysis of signalling factors that influence early lineage decisions in the endoderm. Antibodies against several signaling pathways were tested, but unfortunately only the antibody against BMP-signalling performed well and could be used for further analysis.

The possible role of BMP-signalling during gastrulation and organogenesis was therefore clarified using whole mount stainings with antibodies against the activated forms of BMP´s downstream effectors Smad1/5/8. The results showed that BMP-signalling is actually not active in the region of endoderm formation.

In conclusion, this work showed the advantages of BAC transgenes in the study of endoderm formation during gastrulation stage. BAC transgenes can be used to deep analysis of gene expression and to follow this over time at the cellular level in the living mouse embryo. The generated tools gave some new detailed insights about the formation of the endoderm germ layer and can be used for further experiments and live imaging approaches to understand several other aspects of endoderm development. Moreover, the used whole mount antibody staining methodology could allow the complete mapping of signaling pathways during gastrulation and could be therefore highly useful to show which signaling pathways influence early decisions in the endoderm at the cellular level. Thus, allowing to eventually generate an atlas of the gastrulation stage embryo showing cell fate decisions at the cellular level.

10.2 Zusammenfassung

Der Zweck dieser Arbeit war die Erforschung der Entwicklung des Entoderms in embryonalen Stammzellen der Maus und Mausembryonen um neue Erkenntnisse

über die Entstehung des Entoderms zu erhalten. Ziel war es zu zeigen, wo Zellpopulationen des Entoderms während der Gastrulation erscheinen und wohin sie bis zum Zeitpunkt der Organogenese migrieren.

Hierzu wurde die Methode der Rekombination von BAC´s verbessert um fluoreszierende Proteine unter der Kontrolle spezifischer Gene des Entoderms in embryonalen Stammzellen stabil zu integrieren. Es wurden BAC-transgene der frühesten exprimierten Gene des Entoderms, *Foxa2* und *Sox17* erzeugt. Außerdem wurden BAC-transgene der, in den entodermal-abgeleiteten Organen Lunge, Leber und Bauchspeicheldrüse exprimierten Gene *Nkx2.1*, *Hhex* sowie *Pdx1* erzeugt. Die generierten BAC-transgenen embryonalen Stammzellen wurden anschließend *in vitro* in Entoderm differenziert und die erwartete orts- und zeitspezifische Expression der fluoreszierenden Reportergene wurde analysiert.

Mittels tetraploider Komplementation wurden Embryonen generiert, die komplett aus den BAC-transgenen embryonalen Stammzellen gebildet wurden. Im direkten „live imaging“ unter dem Konfokalen Lasermikroskop konnte gezeigt werden, dass die Expression der Reportengene der endogenen Genexpression entspricht. Auf diese Weise konnte gezeigt werden dass bereits eine Kopie des Reportergens für statische Aufnahmen im „live imaging“ während der frühen Embryogenese ausreichend ist.

Besonders die *Hhex-Lyn-Tomato* Reporterlinie erwies sich als sehr gut geeignet in statischen Analysen und im „live imaging“ der Population der Reportergen exprimierenden Zellen. Es konnte eine Reportergenexpression in der inneren Zellmasse der Blastozyste sowie zwischen E14.5 und E18.5 in den vorderen und hinteren Gliedmaßen beobachtet werden die bisher unbekannt war. Die Frage konnte beantwortet werden ob alle *Foxa2* und *Sox17* exprimierenden Vorläuferzellen zu allen entodermal-abgeleiteten Organen beitragen oder nur zu einem Teil der Organe. Es konnte gezeigt werden dass *Foxa2*, genauso wie *Sox17*, nicht zu allen entodermal-abgeleiteten Organen beiträgt. So wird zum Beispiel *Foxa2* in den Zellen des Anterioren Definitiven Entoderms herunterreguliert aus denen die Leber hervorgeht. Es wurde das Verfahren der Komplettfärbung von *Hhex::Lyn-Tomato* BAC-transgenen Embryonen mit Antikörpern gegen *Foxa2* etabliert, das die Darstellung der Populationen des Entoderms auf zellulärer Ebene ermöglicht. Auf diesem Wege konnte gezeigt werden, dass Zellen des Epiblasts bereits *Foxa2* exprimieren bevor sie in das Definitive Entoderm integrieren. Dies zeigt, dass Zellen

des Entoderms bereits vor der Integration in das Viszerale Entoderm und der Bildung des Definitiven Entoderms determiniert sind.
Die Komplettfärbung von Embryonen wurde zudem verwendet bei der Analyse von Signaltransduktionswegen, die die frühe Bildung des Entoderms beeinflussen. Es wurden Antikörper gegen verschiedene Schlüsselfaktoren von Signaltransduktionswegen getestet. Es funktionierte nur der Antikörper gegen den BMP-Signaltransduktionsweg, der dann in weiteren Experimenten eingesetzt wurde. Die mögliche Rolle des BMP-Signaltransduktionswegs während der Gastrulation und der Bildung der entodermal-abgeleiteten Organe wurde durch Antikörperfärbungen gegen die aktivierte Form der BMP-Effektoren Smad1/5/8 untersucht. Die Ergebnisse zeigten, dass der BMP-Signaltransduktionsweg in der Region, in der das Entoderm gebildet wird inaktiv ist.

Im Ergebnis zeigt diese Arbeit die Vorteile von BAC-transgenen bei der Erforschung der Bildung des Entoderms während der Gastrulation. BAC-transgene können zur vertieften Analyse der Expression von Reportergenen genutzt werden. Die Expression kann auf zellulärer Ebene über einen Zeitraum statisch und live im Embryo der Maus visualisiert werden. Mit den geschaffenen BAC-transgenen konnten neue Einblicke über die Bildung des Entoderms gewonnen werden. Sie können in „live imaging“ Experimenten zukünftig eingesetzt werden um weitere Aspekte der Bildung des Entoderms zu untersuchen.
Die Technik der Komplettfärbung von Embryonen ermöglicht die komplette Kartierung von Signaltransduktionswegen während der Gastrulation der Maus. Sie erlaubt es auf zellulärer Ebene zu zeigen welche Signaltransduktionswege die frühen Entscheidungen bei der Bildung des Entoderms beeinflussen. Damit wäre es möglich einen „Atlas“ der Maus während der Gastrulation zu erstellen auf dem die Zellschicksalsentscheidungen auf zellulärer Ebene dargestellt werden können.