Published in final edited form as:
Stem Cells. 2012 October ; 30(10): 2297–2308. doi:10.1002/stem.1192.

Dual lineage-specific expression of Sox17 during mouse embryogenesis

Eunyoung Choi[1], **Marine R-C. Kraus**[2], **Laurence A. Lemaire**[2,3], **Momoko Yoshimoto**[4], **Sasidhar Vemula**[4], **Leah A. Potter**[1], **Elisabetta Manduchi**[5], **Christian J. Stoeckert Jr.**[5], **Anne Grapin-Botton**[2,3], and **Mark A. Magnuson**[1,*]

[1]Center for Stem Cell Biology and Department of Molecular Physiology and Biophysics, Vanderbilt University, Nashville, TN [2]Swiss Institute for Experimental Cancer Research, School of Life Sciences, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland [3]DanStem, University of Copenhagen, 3B Blegdamsvej, DK-2200 Copenhagen N, Denmark [4]Department of Pediatrics, Wells Center for Pediatric Research, Indiana University School of Medicine, Indianapolis, IN [5]Penn Center for Bioinformatics and Department of Genetics, University of Pennsylvania School of Medicine, Philadelphia, PA

Abstract

Sox17 is essential for both endoderm development and fetal hematopoietic stem cell (HSC) maintenance. While endoderm-derived organs are well known to originate from Sox17-expressing cells it is less certain whether fetal HSCs also originate from Sox17-expressing cells. By generating a *Sox17*^{GFPCre} allele and using it to assess the fate of Sox17-expressing cells during embryogenesis we confirmed that both endodermal and a part of definitive hematopoietic cells are derived from Sox17-positive cells. Prior to E9.5 the expression of Sox17 is restricted to the endoderm lineage. However, at E9.5 Sox17 is expressed in the endothelial cells (ECs) at the para-aortic splanchnopleural (P-Sp) region that contribute to the formation of HSCs at a later stage. The identification of two distinct progenitor cell populations that express Sox17 at E9.5 was confirmed using FACS together with RNA-Seq to determine the gene expression profiles of the two cell populations. Interestingly, this analysis revealed differences in the RNA processing of the *Sox17* mRNA during embryogenesis. Taken together, these results indicate that Sox17 is expressed in progenitor cells derived from two different germ layers, further demonstrating the complex expression pattern of this gene and suggesting caution when using Sox17 as a lineage-specific marker.

Introduction

Sox17, a member of the Sry-related high mobility group box (Sox) transcription factors, plays an essential role in the differentiation of many types cells[1–4]. During mouse embryogenesis, Sox17 is first detected in extraembryonic visceral endoderm at embryonic day (E) 6.0 and in the endoderm of mid- to late-gastrula stage embryos (e.g. around E7.5)

*Corresponding author, 9465 MRB-IV, 2213 Garland Avenue, Vanderbilt University School of Medicine, Nashville, TN 37232-0494, Tel: 615.322.7006, Fax: 615.332.6645, mark.magnuson@vanderbilt.edu.

Author contributions: E.C.: conception and design, collection and assembly of data, data analysis and interpretation, manuscript writing; M.R-C.K.,L.A.L., M.Y. and S.V.: collection and assembly of data; L.A.P.: data analysis and manuscript writing; E.M. and C.J.S.: data analysis and interpretation, manuscript evaluation; A.G-B.: provision of study material, financial support, manuscript evaluation; M.A.M: conception and design, data analysis, financial support, manuscript writing, final approval of manuscript.

Disclosure of potential conflicts of interest
No potential conflicts of interest.

where it plays an essential role in organ formation[5]. Epithelial cells of the gut tube endoderm maintain Sox17 expression until approximately E8.5 as they undergo specification into distinct endoderm-derived organs[5–8]. Expression of Sox17 in endoderm, albeit transient, has led to this gene being widely used as a marker for definitive endoderm in studies using embryonic stem (ES) cells[9–12].

Within the developing endoderm, Sox17 is critical for specifying pancreatic progenitors in the ventral foregut endoderm. Mice that are globally deficient in *Sox17* fail to undergo axis rotation at E8.5 and exhibit a severe defect of the posterior region of the embryo[5, 13]. Moreover, while *Sox17*-null embryos express Hnf3a/b and other endoderm derived organ-specific markers such as Hhex and Cdx2, they do not express Pdx1, an essential factor for pancreatic outgrowth and development[5, 14]. Interestingly, Sox17 is expressed in a ventral foregut endoderm region from which the ventral pancreas and extrahepatobiliary ducts emerge where it appears to be essential for the segregation between liver and pancreatibiliary systems[15].

Sox17 also plays a crucial role in the maintenance of fetal and neonatal hematopoietic stem cell (HSC)identity[13, 16]. During vascular system development, *Sox17* expression begins at approximately E8.5[17]. However, the precise time at which *Sox17* is expressed during embryonic hematopoiesis has not been as thoroughly investigated[18]. During embryogenesis, HSCs originate from the hemogenic endothelial cells (ECs) located on the aortic floor at E10.5, and migrate to the liver to expand in number[19–24]. Utilizing Tie2-Cre to conditionally eliminate *Sox17* expression, Kim et al. investigated the role of Sox17 in HSC development post-migration and found a marked impairment in the number of HSCs in the liver at E11.5[13]. A role for Sox17 in fetal HSC maintenance was further supported by He et al. who reported that the transient expression of Sox17 in adult bone marrow (BM) caused adult hematopoietic cells to adopt characteristics of fetal HSCs[16]. However, while both of these studies demonstrated the importance of Sox17 in the maintenance of the fetal HSCs, neither explored whether Sox17-expressing ECs can give rise to hematopoietic cells or not.

In this study, we generated mice with a $Sox17^{GFPCre}$ allele and used it to identify Sox17-expressing cells and their progeny. At E9.5, we identified two distinct progenitor populations of Sox17-expressing cells, both an endoderm-derived ventral pancreatic epithelial cell population and a mesoderm-derived endothelial cell population that has hemogenic potential. Furthermore, we show that the two populations exhibit distinct gene expression signatures as assessed by whole transcriptome profiling.

Materials and Methods

Gene targeting and RMCE

Mice containing a $Sox17^{GFPCre}$ allele were derived using both gene targeting and recombinase-mediated cassette exchange (RMCE). First, a *Sox17* loxed cassette acceptor (LCA) allele was made by gene targeting. The targeting vector made by BAC recombineering replaced a 3.793 kb region of the gene containing exons 3–5 with a *puromycin resistance-Δthymidine kinase* (*puΔTK*) fusion gene driven by the mouse *phosphoglycerol kinase* (*PGK*) promoter and a *kanamycin resistance* gene driven by the bacterial EM7 promoter. The selection cassettes were flanked by tandemly-oriented lox71 and lox2272 sites, two homology arms, and a *PGK-diphtheria toxin A* (*DT-A*) cassette for negative selection. After linearization using NotI, a 129S6-derived mouse ES cell line was electroporated with the targeting vector, and puromycin-resistant clones were isolated. Homologous recombination was verified by Southern hybridization with 5’ and 3’ probes following digestion with either SphI or SpeI. Second, an exchange cassette flanked by tandem lox66 and lox2272 sites was made with an enhanced green fluorescent protein (GFP)

and Cre fusion gene in place of the coding sequences of *Sox17* and replace the selectable markers in the $Sox17^{LCA}$ allele. In addition, the vector contained a *PGK-hygromycin resistance* ($Hygro^{R}$) cassette flanked by tandem flippase recognition target (FRT) sites. Following co-electroporation of the exchange vector and a Cre-expression plasmid, positively exchanged clones were identified by a staggered positive-negative selection strategy using both hygromycin and gancyclovir[25]. To identify correctly exchanged clones, PCR was performed using the following primers: 5'-ACAGTCTTACACGCTACGGAT and 5'-CAAGACCTCTTGGGGAAATAG on the 5' end (a); 5'-CAGAGGTATGCAGATCTCTGT and 5'-CATTCTGGTCAACATGTAAGGT on the 3' end (b).

Mouse strains

Chimeric mice containing the $Sox17^{GFPCre}$ allele were derived by the microinjection of clone 1G3:1C10 ES cells into blastocysts of C57BL/6J mice. After germline transmission, the $Hygro^{R}$ cassette was removed by cross breeding with *ACTB:FLPe* mice[26]. The $Sox17^{GFPCre}$ allele was maintained within a CD1 background for experiments. Mice with the $R26R^{eYFP}$ and $R26R^{LacZ}$ (*Rosa26:LacZ* (*Gt(ROSA)26Sor*^{tm1Sor})) alleles were previously described[27]. Embryos were considered to be E0.5 at noon on the day a vaginal plug was detected. All experimental protocols were approved by the Vanderbilt Institutional Animal Care and Use Committee.

Immunohistochemistry

Five micron frozen sections were preincubated with 5% normal donkey serum (NDS) for 30 mins at room temperature, incubated with primary antibodies at 4°C overnight, and washed in phosphate buffered saline (PBS) with 0.1% Tween 20. Secondary antibodies were incubated at room temperature for 1 hour then washed in PBS containing 0.1% Tween 20. All antibodies were diluted in PBS containing 5% NDS and 0.1% Tween 20. The sources of the antibodies used are listed in Table S3. Images were acquired using either a Zeiss LSM510 or LSM710 inverted confocal microscope at the Vanderbilt Cell Imaging Shared Resource.

β-D-Galactosidase staining

$Sox17^{GFPCre/+}$;$R26R^{LacZ/+}$ embryos were harvested at E7.5, E9.5, or E12.5 and fixed at room temperature for 10 mins in 4% paraformaldehyde. 5-bromo-4-chloro-3-indolyl β-D-Galactosidase staining was performed on 10 μm-thick serial transversal sections for 4 hours at 37°C (Thompson et al., 2012). Images were acquired with a Leica DM5500B microscope equipped with a DFC 320 color camera.

FACS analysis

$Sox17^{GFPCre/+}$ embryos were identified by direct fluorescence using a Leica MZ 16 FA stereoscope. The pooled embryos, containing five to seven embryos, were dissociated with AccuMax (Sigma) and 5 μg/ml DNase I (Sigma) and passed through a 35 μm mesh filter into FACS tubes (BD). After centrifugation, the cells were resuspended with FACS staining buffer (R&D Systems), blocked using 1 μg/ml of mouse IgG at room temperature for 15 min, then immunolabeled with PE conjugated-EpCAM (G8.8) antibody (Santa Cruz) at room temperature for 30 mins. The cells were washed three times in FACS staining buffer, centrifuged, and resuspended with FACS staining buffer with 10mM HEPES, 1mg/ml BSA and 1% penicillin/streptomycin. 7-aminoactinomycin D (7-AAD, Molecular Probes) was added at a dilution of 1:1000 to assess cell viability, and cells were analyzed using an LSRII (BD Biosciences). To isolate the Sox17-expressing cell populations, the midgut regions from 26 to 29 $Sox17^{GFPCre/+}$ embryos at E9.5 were dissected (Fig. S5C) and pooled prior to

cellular dissociation. Samples were prepared as described above, and cells were isolated using either an Aria II or III (BD Biosciences).

Fetal liver cells were obtained from dissected E12.5 embryo livers then dissociated into a single cell suspension using a 200 μl large orifice tip pipette. Bone marrow was collected from the tibias and femurs of 6 week old mice. The cells were flushed using 1X PBS through a 21G needle more than 10 times then filtered with a 100 μm cell strainer. Red blood cells were lysed using ACK lysis buffer (0.15 M NH_4Cl, 1 mM $KHCO_3$, 0.1 mM Na_2EDTA, pH 7.4, sterilized using 0.2 mm filter). The following antibodies were used for immunolabeling: anti-CD41 (MWReg30), anti-VE-Cad (BV13), anti-CD45 (30-F11), anti-CD19 (1D3), anti-CD11b/Mac1 (M1/70), anti-CD3 (17A2) and anti-Ter119. DAPI (1 μg/ml) was added to assess cell viability.

Hemogenic endothelial cell culture

$GFPCre^+$ or $GFPCre^-$ endothelial cells ($CD45^-Ter119^-CD41^-VE\text{-}cad^+CD31^+$ cells) from the YS and P-Sp region of E9.5 *Sox17*$^{GFPCre/+}$ embryos were sorted and 3,000 cells were plated onto an OP9 stromal cell layer in 6 well plates with SCF (10 ng/ml), Flit3-ligand (10 ng/ml), IL3 (10 ng/ml), IL7 (10 ng/ml), and TPO (10 ng/ml). Endothelial cells from wild-type YS and P-Sp were also plated as a positive control. Floating cells in the co-culture were collected every 3–4 days and analyzed by flow cytometry using anti-TER119 and Mac1 antibodies.

Total RNA isolation, amplification, library construction, RNA-Seq and analysis

Total RNA was isolated using TRIzol LS (Invitrogen) containing 40 μg/ml mussel glycogen (Sigma). Following treatment with DNase (Ambion), total RNA was column-purified using a DNA-Free RNA kit (Zymo Research). RNA integrity was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, CA), and RNAs with an RNA Integrity Number (RIN) > 7.0 were processed for RNA-Seq. Total RNA (10 ng) was amplified using the Ovation RNA-Seq system (NuGen, CA), and amplified cDNA (750 ng) was sheared using Covaris S2 system (Covaris, MA) and used to construct the cDNA library using the Illumina TruSeq DNA prep kit (Illumina, CA). The libraries were used to generate ≥110 base reads as single-end tags using an Illumina HiSeq 2000, and the HiSeq Control Software (HCS 1.4.8/RTA 1.12) was used for image analysis. Reads were mapped to mm9 genome with RNA-Seq unified mapper (RUM, v1.09, http://www.cbil.upenn.edu/RUM/)[28]. Expression was quantified as reads per kilobase of exon model per million mapped reads (RPKM). A summary of the mapped reads is inTable S4, and confidence values were determined by PaGE (http://www.cbil.upenn.edu/PaGE/)[29]. Differentially expressed genes were selected as those displaying at least a 4-fold change in RPKM, ≥ 4 RPKM in one sample, and ≥ 80% confidence value, and these genes were clustered using Cluster program[30]. Gene Ontology analysis was performed using PANTHER (v7.0)[31].

Meta-data and processed data for the RNA-Seq are available at http://genomics.betacell.org. The RNA-Seq data has been deposited at ArrayExpress (accession number E-MTAB-970) and the Sequence Read Archive (accession number ERP001235).

Semiquantitative RT-PCR

RT-PCR was performed using 50 ng of amplified cDNA template used for RNA-Seq. *Sox17* transcript variants were detected using the following primers: 5’ -GGATACGCCAGTGACGACCA with 5’ -CGTTCGTCTTTGGCCCACAC (a) or 5’ -TCATGCGCTTCACCTGCTTG (b); 5’-ATGGCCCACTCACACTGCTG with 5’ -ATGTAGCTCTCCTGCCTCTC (c) or 5’ -CGTTCGTCTTTGGCCCACAC (d).

Results

Derivation of mice with a *Sox17^{GFPCre}* allele

To explore the stage-specific expression and lineage of Sox17-expressing cells, we derived mice that express a GFPCre fusion protein under control of the endogenous *Sox17* gene locus through gene targeting and recombinase-mediated cassette exchange (RMCE)[32] (Fig. 1A). First, we performed gene targeting to generate mouse ES cells containing a *Sox17^{LCA}* allele. Two of 84 mouse ES cell clones that survived puromycin selection were correctly targeted as confirmed by Southern hybridization (Fig. 1B). Second, we generated an exchange cassette that replaced coding sequences in the *Sox17* gene with a GFPCre fusion protein. Following RMCE into the *Sox17^{LCA}* allele, the exchange was confirmed using PCR across both the lox66/71 and lox2272 sites, and one clone (1G3:1C10) was used to generate mice containing the *Sox17$^{GFPCre(+HygroR)}$* allele (Fig. 1C). After germline transmission, mice containing the *Sox17$^{GFPCre(+HygroR)}$* allele were bred with FLPe-expressing mice to delete the FRT-flanked hygromycin resistance cassette, thereby establishing the *Sox17^{GFPCre}* line.

To determine whether expression of the *Sox17^{GFPCre}* reporter allele faithfully recapitulates that of the wild type *Sox17* allele, we first performed whole mount fluorescence microscopy of *Sox17$^{GFPCre/+}$* mouse embryos. In agreement with previous reports[5, 15], we observed fluorescence in the extra-embryonic region beginning at E6.5 (Fig. 2A–a,b) and in the intra-embryonic region beginning at E7.5 (Fig. 2A–c,d). A three-dimensional reconstruction of E7.5 embryo images revealed that GFPCre-expressing cells were highly abundant and clustered together in the definitive endoderm rather than the visceral endoderm (Fig. S1). At E8.5, between the fourth and ninth somite (S) stage, fluorescence was localized to the foregut endoderm (Fig. 2A–e,f), and histological analysis revealed co-expression with Sox17 (Fig. 2B–a). At E9.5 (>20S), fluorescence throughout the endoderm was greatly diminished, but localized expression was evident in the ventral pancreas region (Fig. 2A–g,h and 2B–b) At E9.5, we also observed cells expressing both GFPCre and Sox17 in the dorsal aorta (Fig. 2B–c). This finding indicates that Sox17 is not only expressed in the ventral pancreatic epithelium at E9.5 but is also present in cells in the dorsal aorta.

Two types of *Sox17^{GFPCre}*-expressing cells during organogenesis in the mouse embryo

At E9.5, Sox17 was detected primarily in the ventral pancreatic epithelium but not in the liver epithelium (Fig. 3A). GFPCre was strongly expressed in the tip area of the ventral pancreatic epithelium (Fig. 3A, left arrow) and at a lower level in the Pdx1-high area of ventral pancreatic epithelium (Fig. 3A, right arrow)[15]. Several GFPCre-expressing cells were also located in the region of the liver epithelium; however, they were not co-localized with the liver marker Hnf4a, indicating that these GFPCre-expressing cells were separate but intermingled with the developing liver epithelium at this stage (Fig. 3A). We also observed GFPCre-expressing cells localized within the dorsal aorta (Fig. 3B)[17] and in the vicinity of the neural tube area (Fig. 3C). While these GFPCre-expressing cells were not co-localized with Sox2, an ectoderm marker, Foxa2, a floor plate marker, or Sox10, a marker of neural crest cells, they were co-localized with PECAM, an endothelial cell marker (Fig. 3E, arrow)[17]. Conversely, the GFPCre-expressing cells in the ventral pancreatobiliary epithelium co-localized with epithelial cell adhesion molecule (EpCAM), an epithelial cell marker (Fig. 3D, arrow). These data indicate that Sox17 is expressed in two distinct progenitor cell populations at E9.5: 1) an epithelial cell population found in the ventral foregut epithelium that gives rise to the ventral pancreas, extrahepatic ducts and gall bladder and 2) an endothelial cell population found in the P-Sp area, which includes the dorsal aorta.

Two distinct cell populations expressing Sox17 are alternatively derived from two different origins

To trace the lineage of Sox17-expressing cells the *Sox17^{GFPCre}* mice were crossed with either a *Gt(ROSA)26Sor^{tm1Sor} (R26R^{LacZ})* or *ROSA26-eYFP (R26R^{eYFP})* cre-reporter line[27, 33] and the locations of either LacZ or YFP, repectively, were determined (Fig. 4 and S2). Consistent with previous studies, LacZ was detected as early as E7.5 in extraembryonic and embryonic visceral endoderm, as well as the definitive endoderm where it was detected primarily in the proximal endoderm (data not shown)[34]. Previous studies have also shown that Sox17 plays a key role in determining endoderm and pancreas fates[5, 35]. Consistent with this[15, 34, 36], we found that at E9.5 all endoderm-derived epithelial cells, from the branchial arch endoderm to the hindgut, were derived from cells that once expressed Sox17 (Fig. S2A–D). At E9.5, YFP was co-expressed with EpCAM or Foxa2 within the gut endoderm and the pancreatic epithelium (Figs. 4A and S3A). At E12.5, we confirmed that endoderm-derived organs, including the thyroid, thymus, parathyroid, esophagus, trachea, lungs, stomach, liver epithelium and endothelium (but not blood cells), dorsal and ventral pancreas, small intestine, cecum, and large intestine were labeled by LacZ (Fig. S2 E–G). We also observed YFP labeling in Ptf1a-expressing pancreatic multipotent progenitor cells (MPCs) at E11.5. Additionally, as observed at E15.5, acinar, ductal, endocrine progenitor, and insulin-expressing cells were derived from a Sox17-expressing lineage (Fig. S3B).

In addition to contributing to endoderm-derived lineages, we observed YFP-expressing cells in the endothelia, such as dorsal aorta and veins at E9.5 (Fig. 4B and S4A). An analysis of YFP-positive cells in the heart, which is comprised of both endothelial cells and mesenchymal cells, showed that YFP was selectively observed in the PECAM-positive endothelial cells (Fig. 4C, arrow) but not in the VCAM-positive cardiac myocytes, representing cells of the mesenchymal lineage[37] (Fig. 4C, arrowheads). These findings indicate that Sox17 is not expressed in mesoderm progenitor cells that gives rise to both heart endothelium and mesenchyme[38], but is expressed after the mesoderm is specified into either the mesenchymal or endothelial lineage.

Sox17 is expressed in hemogenic endothelial cells at E9.5

While it has been previously shown that Sox17 has a role in the maintenance of fetal and neonatal HSCs[13], it is not clear from these studies whether Sox17 is expressed in the hemogenic ECs in the YS and in the P-Sp region. Given that Sox17 was detected in the ECs of P-Sp region, we sought to determine whether the Sox17-expressing cells were hematopoietic in nature. At E9.5, many GFPCre-expressing cells along the aortic wall co-expressed c-Kit (Fig. 5A, arrows). However, GFP was not co-localized in CD41^{+}c-Kit^{+} HPCs[39] (Fig. 5A, arrowheads), indicating that Sox17 expression occurs only in the endothelial stage and is not sustained as the cells differentiate into HPCs. In contrast, some Sox17-derived cells, as indicated by YFP expression, were detected in CD41^{+}c-Kit^{+} HPCs (Fig. 5B, arrowheads), indicating that CD41^{+}c-Kit^{+} HPCs were derived from Sox17 expressing cells. Consistent with this, a few CD41^{+}c-Kit^{+} cells (~1%) were produced when GFPCre^{+} EC cells were cultured on OP9 cells (data not shown). To determine whether Sox17-expressing ECs have hemogenic potential we sorted GFPCre^{+} EC cells (CD45^{-}CD41^{-}Ter119^{-}CD31^{+} or VEcad^{+} cells), which accounted for around 40–50% of the EC cells, from E9.5 YS and P-Sp and plated them on an OP9 stromal cell layer along with stimulatory cytokines. After 8 days of co-culture, large cobblestone-appearing areas were observed that were proven to be Ter119^{+} or Mac1^{+} erythro-myeloid colonies by FACS (Fig. 5C). Of note, GFPCre^{-} EC also produced erythro-myeloid cells in the OP9 co-culture (data not shown). Taken together, these data indicates that Sox17 is expressed in the hemogenic ECs.

Use of Sox17 lineage tracing to show the development of definitive hematopoiesis *in vivo*

We next sought to determine how many hematopoietic cells are derived from Sox17-expressing cells *in vivo*. To do so, we determined the number of Sox17-lineaged cells in E9.5 embryos, E12.0 fetal liver (FL) and in adult bone marrow (BM) (Fig. 5D). While only 25.1 ± 12.3% of $CD41^+$ cells, 9.5±2% of $Ter119^+$ cells, and 21.4±11.1% of $Mac1^+$ cells were YFP^+ in the E9.5 embryos (Fig. 5D upper panel), the number of YFP^+ cells was dramatically increased in the E12.5 FL. Indeed, up to 81.9 ± 2.2 of the $CD41^+$ cells, 39.5 ± 5.5% of the $Ter119^+$ cells and 81.6 ± 4.1% the $Mac1^+$ cells were lineage-labeled (Fig. 5D middle panel). Furthermore, YFP^+ cells were also detected among $CD19^+$ cells (55.8% ± 10.4%) and $CD3^+$ cells (58.9 ± 1.1%), representing lymphocytes. YFP^+ cells were also detected in the VE-cad^+ cells (81.9 ± 2.2%). These results suggest that Sox17 is expressed in VE-cad^+ endothelial cells and is linked to the production of definitive hematopoietic cells, perhaps through hemogenic ECs, based on the previous reports[40–42]. In adult BM, all $CD45^+$ hematopoietic cells were YFP^+, clearly demonstrating that all the blood cells in the adult BM are derived from Sox17-expressing cells. This is consistent with the previous report that HSCs in the FL express Sox17[13], and that FL HSCs migrate into the BM and support hematopoiesis, as is commonly considered. However, the low percentage of YFP^+ cells among $Ter119^+$ cells were detected in adult BM, likewise in E12.5 FL. Since definitive erythroid cells are enucleated, these cells may lose the expression of YFP as they mature. Also, about 17% of the $CD3^+$ cells were YFP^-. While this may suggest that T cell lymphopoiesis occurs in $Sox17^-$ hemogenic EC cells in the YS and P-Sp before emergence of HSCs[42], this finding could also be due to a low amount of GFPCre in the hemogenic ECs resulting in only partial recombination of the $R26R^{eYFP}$ reporter allele. In either case, Sox17 marks a part of hemogenic ECs and Sox17 lineage tracing shows the progression of definitive hematopoiesis in the embryo.

Sox17-expressing endothelial and epithelial cells exhibit transcriptional differences in distinct protein classes

To further demonstrate differences in the two Sox17-expressing populations, we isolated both cell populations using fluorescence-activated cells sorting (FACS) and performed whole transcriptome sequencing (RNA-Seq). Using E9.5 $Sox17^{GFPCre/+}$ mouse embryos, the epithelial and endothelial cells were separated based on expression of EpCAM, an epithelial cell marker (Fig. S5A). From whole embryos, we obtained less than one percent GFPCre-expressing cells (Fig. S5A) of which only 7.0% ± 3.2% were Sox17-expressing epithelial cells based on the expression of both GFP and EpCAM (Fig. S5B). Since endodermal expression becomes restricted to the ventral posterior foregut at E9.5, we dissected the mid region of the embryo to increase the yield of Sox17-expressing cells (Figs. S5C and 6A). Even so, the total number of cells was low, thereby requiring RNA amplification prior to sequencing. As summarized in Table S1, 62.5% – 75.6% of the RNA-Seq reads obtained were aligned to the mouse genome (mm9)[28]. Each profile from three biological replicates showed high reproducibility (Table S2). Interestingly, the profiles between $EpCAM^+$ and $EpCAM^-$ Sox17-expressing cells were also highly correlated, simply implying that the global gene expression profiles for the two populations are similar.

To identify differentially expressed transcripts within the two populations, we examined the transcriptional levels of 28,683 genes. While the overall gene expression profiles were similar (Table S2), the two populations were distinguished by many differentially-expressed genes with high confidence values (Table S4). To more clearly illustrate the differences we selected 321 genes that displayed a four-fold or greater change in expression (Table S5) and performed a cluster analysis. By doing so, differences between the epithelial and endothelial cells became readily apparent (Fig. 6B). Furthermore, many of the differentially-expressed

genes are important for ventral foregut development, hematopoiesis and/or the regulation of cell fate and signal transduction based on their gene ontology annotations in PANTHER[43].

In the transcription factor cluster (PC00218), there were 11 up- and 26 down-regulated genes in Sox17-expressing epithelial cells compared to the endothelial cells (Fig. 6C). The up-regulated transcripts primarily consisted of endoderm and pancreas development-related genes, such as *Onecut1, Foxa3, Foxa2, Hnf1b, Sox9*, and *Prox1*. However, other transcription factors, including *Cited1, Tcea3, ID2, Lin28a* and *Aes*, were also highly expressed in the Sox17-expressing epithelial cells. Conversely, most of the down-regulated genes, including *Gata2, Ets1, Elk3, Lmo2, Kef2c, Klf2, Sox7, Tal1* and *Sox18,* are known to play a role in endothelial and HSC development. In the receptor (PC00197) and cell adhesion molecule (PC00070) clusters, we found 11 up- and 19 down-regulated genes in Sox17-expressing epithelial cells as compared to the endothelial cells (Fig. 6C). *Tacstd1* (*EpCAM*), which was used to isolate the two Sox17-expressing cell populations, was one of the most abundant genes in epithelial cells, and several epithelial cell-related genes, such as *Cdh1, Tmprss2, Emb* and *Bcam,* were also abundant. *Paqr9*, a member of the progestin and adipoQ receptor family, and *Dlk1*, a member of the EGF-like family, were also highly expressed; however, their role in pancreas development has not been elucidated. Conversely, in the Sox17-expressing endothelial cells, many genes important for endothelial cell development or hematopoiesis, such as *Kdr, Eltd1, Eng, Cd97, Cd34, Esam, Nid1*, were highly expressed. In the signaling molecule cluster (PC00207), we found 13 up- and 13 down-regulated genes in Sox17-expressing epithelial cells as compared to the endothelial cells (Fig. 6C). In agreement with studies highlighting the roles of Wnt and TGFβ signaling in pancreas development[44–45], we found that Wnt signaling-related genes, *Fzd7* and *Sfrp5*, and TGFβ signaling-related genes, such as *Npnt* and *Bmp7*, were abundant in epithelial cells. Similarly, *Habp2*, a hepatocyte growth factor activator-like protein and endoderm-enriched gene, and *Sdc4*, which is involved in organogenesis of the kidney, were highly expressed.These data not only confirm the existence of two distinct Sox17-expressing progenitor populations at E9.5, but they also reveal numerous transcription factors and cell surface molecules that may be useful for distinguishing the two populations.

Alternative splicing variant of *Sox17* gene is expressed in two populations

Previously, tissue-specific splicing variants of *Sox17* were identified in mouse adult testis and lung[46], and the presence of alternative transcription start sites has been suggested during embryonic development[34, 36]. An analysis of the RNA-Seq reads mapped to the *Sox17* locus revealed differences in exon abundance (Fig. S6). Thus, to further explore whether the Sox17-expressing epithelial and endothelial cells express unique *Sox17* mRNA transcripts (Fig. 7A), we performed exon-specific RT-PCR (Fig. 7B). We found that the fourth and fifth exons, which contain the coding sequences, were amplified from cDNA from both Sox17-expressing epithelial and endothelial cells (Fig. 7B–a,b). However, sequences in exons one, two and three were amplified only in the Sox17-expressing endothelial cDNA (Fig. 7B–c,d). This data indicates that the *Sox17* pre-mRNA is alternatively processed in the two different cell types.

Discussion

Our study indicates the existence of two distinct Sox17-expressing cell lineages in the P-Sp region of the mouse at E9.5. The first population of Sox17-expressing cells, DE (Fig. 2), goes on to form endoderm-derived organs, such as lung, stomach, pancreas, liver, and intestine. The second population is mesoderm-derived endothelium, which go on to contribute to the derivation of HPC. While the expression of Sox17 in the ECs and fetal HSCs has been previously reported using reporter mouse lines[13, 34, 47], it has not been clear when Sox17-expressing ECs and fetal HSCs begin to emerge.

Most mature blood cells at E9.5 are considered to be primitive erythro-myeloid cells derived directly from mesoderm, not ECs[48–49]. Thus, the YFP^+ hematopoietic cells detected in this study may reflect definitive hematopoiesis from hemogenic ECs, although there is a possibility that primitive hematopoietic cells are derived from $Sox17^-$ hemogenic ECs. At E12.5, when definitive hematopoiesis is increased, YFP^+ cells were not only detected in erythro-myeloid cells, but also in both $CD19^+$ and $CD3^+$ cells. This indicates that Sox17 also gives rise to lymphocyte cells. However, the fact that half of the $CD19^+$ and $CD3^+$ cells were YFP^- in the lineage tracing experiment suggest the need for further, even more detailed studies of the biological potential of Sox17-expressing ECs. Nonetheless, when taken together, our findings indicate that Sox17-expressing ECs have, at a minimum, the potential to form erythro-myeloid cells.

To better understand the characteristics of the two populations, we utilized EpCAM immunolabeling to isolate Sox17-expressing epithelial and endothelial cells from dissected mouse embryos. Using whole transcriptome profiling, we determined that the gene signatures for the epithelial and endothelial cells exhibit specific differences and that the *Sox17* mRNA in these two populations is alternatively processed (Fig. 7). However, the coding sequences for Sox17 in the two variants are unaffected. While it is well known that alternative splicing occurs in many genes during embryonic development[50] and is an important means of regulating gene expression, it is not known whether the two different *Sox17* mRNAs that distinguish the ventral foregut progenitor and hemogenic endothelial cell populations have any functional significance.

Previous studies have reported the failure of normal gut tube and blood cell formation due to *Sox17*-deficiency[5, 13]. While some endoderm-derived organs such as the liver and thyroid are not dependent on *Sox17* for development, the specification of other organs, such as the pancreas is dependent on *Sox17*[5]. The transcriptome profiling we performed revealed the expression of many different genes, in addition to *Sox17*, known to be involved in endoderm and foregut development in the ventral pancreatobiliary epithelium at E9.5 (Fig. 6C). In endothelial development, previous studies have shown that the vascular system is unaffected in *Sox17*-null embryos at E8.25[17]; however, at E11.5 there is a significant impairment in the number of HSCs in *Sox17*-null embryos[13].

In our analysis of differentially expressed genes, we found that Sox7 and Sox18, other members of Sox-F family, are more abundant in the Sox17-expressing endothelial cells at E9.5 (Fig. 6C). While these Sox-F members have a redundant role in vascular development[17], it has been suggested that Sox7 and Sox18 have no impact in the generation of HSCs by zebrafish experiments performed with morpholinos to inhibit translation of *Sox7* and *Sox18*[51]. However, little is known about the role of Sox-F family members in HSC generation in mice. Our lineage tracing study shows that Sox17 is expressed in hemogenic ECs that give rise to definitive HPCs. To date, no in-depth studies have been performed to assess the role of the other Sox-F members in intra- or extra-embryonic hematopoiesis. Here, our study reveals that Sox17, one Sox-F member, contributes to embryonic HPC generation.

In spite of the highly similar global expression profiles for the two Sox17-expressing populations, the transcriptional profiles clearly revealed distinct differences as would be expected for two distinct cellular lineages. Many pancreatic and hematopoietic cell-type specific genes distinguish the two populations, suggesting that combinatorial expression of tissue-specific transcription factors is crucial for the fate decision of Sox17-expressing cells. First, numerous pancreas-related genes are highly expressed in Sox17-expressing epithelial cells (Fig. 6C). Specifically, these cells express not only *Onecut1* and *Sox9,* genes important for the continuance of pancreatic progenitor cell fate, but also *Lin28a*, a gene important for

cell pluripotency[52–54]. The combinatorial expression of these transcription factors with Sox17 in the epithelial cells may be crucial for the specification of pancreatic progenitor cells. Furthermore, we identified high expression of *Fzd7* (*Frizzled-7*), a Wnt receptor, in the Sox17-expressing epithelial cells. Given that Sox17 regulates transcription of endodermal target genes through the beta-catenin pathway and Wnt/beta-catenin signaling is required for the proliferation of pancreatic progenitor cells[55–56], Fzd7 may mediate the Wnt/beta-catenin signaling pathway during the pancreas segregation from the gut and biliary system within the Sox17-expressing epithelium. Second, a different set of genes is expressed in the Sox17-expressing endothelial cells. Specifically, we observed the expression of transcription factors such as *Gata2, Lmo2, Tal1* and *Mef2c*, all of which have roles in both vascular and HSC development[57–60], further supporting our identification of a novel Sox17-expressing hemogenic ECs during embryogenesis. The expression of *CD34* and *Kdr* confirms endothelial phenotype. Indeed, given that multiple markers currently need to be used in combination to identify early HSC populations[61], the *Sox17*^{GFPCre} mice we have derived may be useful in future studies that seek to identify and isolate HSC-producing hemogenic ECs.

In conclusion, our study indicates the existence of two lineage-specific, Sox17-expressing progenitor cell populations during early mouse development. Given that Sox17 is widely used to identify endoderm-like cells during embryonic stem cell directed differentiation, our results suggests that Sox17 is not solely an endoderm-specific marker but that it identifies both ventral pancreatic MPCs and a part of hemogenic ECs at E9.5.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Drs. Stacey S. Huppert and Mervin C. Yoder for critically reviewing the manuscript and other helpful suggestions; Susan B. Hipkens and Rama Gangula for establishing and maintaining the mouse line; Travis Clark for RNA amplification and library construction; Anna Osipovich, Judsen D. Schneider and Anil Laxman for helpful suggestions; the Vanderbilt Transgenic Mouse/Embryonic Stem Cell Shared Resource for performing blastocyst injections; the Vanderbilt Flow Cytometry Shared Resource for assistance with FACS and the Vanderbilt Genome Sciences Resource for performing the RNA-Seq.

These studies were supported by NIH grants DK72473 and DK89523 to MAM and DK072495 to AGB; CA68485 and DK58404 to the Vanderbilt Flow Cytometry Shared Resource; CA68485 and DK20593 to the Vanderbilt Transgenic Mouse/ES Cell Shared Resource; and CA68485 to the Vanderbilt Genome Sciences Resource.

References

1. Foster JW, Dominguez-Steglich MA, Guioli S, et al. Campomelic dysplasia and autosomal sex reversal caused by mutations in an SRY-related gene. Nature. 1994; 372:525–530. [PubMed: 7990924]
2. Kamachi Y, Uchikawa M, Collignon J, et al. Involvement of Sox1, 2 and 3 in the early and subsequent molecular events of lens induction. Development. 1998; 125:2521–2532. [PubMed: 9609835]
3. Schilham MW, Oosterwegel MA, Moerer P, et al. Defects in cardiac outflow tract formation and pro-B-lymphocyte expansion in mice lacking Sox-4. Nature. 1996; 380:711–714. [PubMed: 8614465]
4. Pingault V, Bondurand N, Kuhlbrodt K, et al. SOX10 mutations in patients with Waardenburg-Hirschsprung disease. Nat Genet. 1998; 18:171–173. [PubMed: 9462749]
5. Kanai-Azuma M, Kanai Y, Gad JM, et al. Depletion of definitive gut endoderm in Sox17-null mutant mice. Development. 2002; 129:2367–2379. [PubMed: 11973269]

6. Zorn AM, Wells JM. Vertebrate endoderm development and organ formation. Annu Rev Cell Dev Biol. 2009; 25:221–251. [PubMed: 19575677]
7. Beck F, Erler T, Russell A, et al. Expression of Cdx-2 in the mouse embryo and placenta: possible role in patterning of the extra-embryonic membranes. Dev Dyn. 1995; 204:219–227. [PubMed: 8573715]
8. Bogue CW, Ganea GR, Sturm E, et al. Hex expression suggests a role in the development and function of organs derived from foregut endoderm. Dev Dyn. 2000; 219:84–89. [PubMed: 10974674]
9. Peterslund JM, Serup P. Activation of FGFR(IIIc) isoforms promotes activin-induced mesendoderm development in mouse embryonic stem cells and reduces Sox17 coexpression in EpCAM+ cells. Stem Cell Res. 2011; 6:262–275. [PubMed: 21513905]
10. D'Amour KA, Bang AG, Eliazer S, et al. Production of pancreatic hormone-expressing endocrine cells from human embryonic stem cells. Nat Biotechnol. 2006; 24:1392–1401. [PubMed: 17053790]
11. D'Amour KA, Agulnick AD, Eliazer S, et al. Efficient differentiation of human embryonic stem cells to definitive endoderm. Nat Biotechnol. 2005; 23:1534–1541. [PubMed: 16258519]
12. Borowiak M, Maehr R, Chen S, et al. Small molecules efficiently direct endodermal differentiation of mouse and human embryonic stem cells. Cell Stem Cell. 2009; 4:348–358. [PubMed: 19341624]
13. Kim I, Saunders TL, Morrison SJ. Sox17 dependence distinguishes the transcriptional regulation of fetal from adult hematopoietic stem cells. Cell. 2007; 130:470–483. [PubMed: 17655922]
14. Fujitani Y, Fujitani S, Boyer DF, et al. Targeted deletion of a cis-regulatory region reveals differential gene dosage requirements for Pdx1 in foregut organ differentiation and pancreas formation. Genes Dev. 2006; 20:253–266. [PubMed: 16418487]
15. Spence JR, Lange AW, Lin SC, et al. Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. Dev Cell. 2009; 17:62–74. [PubMed: 19619492]
16. He S, Kim I, Lim MS, et al. Sox17 expression confers self-renewal potential and fetal stem cell characteristics upon adult hematopoietic progenitors. Genes Dev. 2011; 25:1613–1627. [PubMed: 21828271]
17. Sakamoto Y, Hara K, Kanai-Azuma M, et al. Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. Biochem Biophys Res Commun. 2007; 360:539–544. [PubMed: 17610846]
18. Jang YY, Sharkis SJ. Fetal to adult stem cell transition: knocking Sox17 off. Cell. 2007; 130:403–404. [PubMed: 17693249]
19. Jaffredo T, Gautier R, Eichmann A, et al. Intraaortic hemopoietic cells are derived from endothelial cells during ontogeny. Development. 1998; 125:4575–4583. [PubMed: 9778515]
20. Zovein AC, Hofmann JJ, Lynch M, et al. Fate tracing reveals the endothelial origin of hematopoietic stem cells. Cell Stem Cell. 2008; 3:625–636. [PubMed: 19041779]
21. Kissa K, Herbomel P. Blood stem cells emerge from aortic endothelium by a novel type of cell transition. Nature. 2010; 464:112–115. [PubMed: 20154732]
22. Bertrand JY, Chi NC, Santoso B, et al. Haematopoietic stem cells derive directly from aortic endothelium during development. Nature. 2010; 464:108–111. [PubMed: 20154733]
23. Boisset JC, van Cappellen W, Andrieu-Soler C, et al. In vivo imaging of haematopoietic cells emerging from the mouse aortic endothelium. Nature. 2010; 464:116–120. [PubMed: 20154729]
24. Mikkola HK, Orkin SH. The journey of developing hematopoietic stem cells. Development. 2006; 133:3733–3744. [PubMed: 16968814]
25. Long Q, Shelton KD, Lindner J, et al. Efficient DNA cassette exchange in mouse embryonic stem cells by staggered positive-negative selection. Genesis. 2004; 39:256–262. [PubMed: 15286998]
26. Rodriguez CI, Buchholz F, Galloway J, et al. High-efficiency deleter mice show that FLPe is an alternative to Cre-loxP. Nat Genet. 2000; 25:139–140. [PubMed: 10835623]
27. Srinivas S, Watanabe T, Lin CS, et al. Cre reporter strains produced by targeted insertion of EYFP and ECFP into the ROSA26 locus. BMC Dev Biol. 2001; 1:4. [PubMed: 11299042]

28. Grant GR, Farkas MH, Pizarro AD, et al. Comparative analysis of RNA-Seq alignment algorithms and the RNA-Seq unified mapper (RUM). Bioinformatics. 2011; 27:2518–2528. [PubMed: 21775302]
29. Grant GR, Liu J, Stoeckert CJ Jr. A practical false discovery rate approach to identifying patterns of differential expression in microarray data. Bioinformatics. 2005; 21:2684–2690. [PubMed: 15797908]
30. Eisen MB, Spellman PT, Brown PO, et al. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A. 1998; 95:14863–14868. [PubMed: 9843981]
31. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet. 2000; 25:25–29. [PubMed: 10802651]
32. Chen SX, Osipovich AB, Ustione A, et al. Quantification of factors influencing fluorescent protein expression using RMCE to generate an allelic series in the ROSA26 locus in mice. Dis Model Mech. 2011; 4:537–547. [PubMed: 21324933]
33. Soriano P. Generalized lacZ expression with the ROSA26 Cre reporter strain. Nat Genet. 1999; 21:70–71. [PubMed: 9916792]
34. Engert S, Liao WP, Burtscher I, et al. Sox17-2A-iCre: a knock-in mouse line expressing Cre recombinase in endoderm and vascular endothelial cells. Genesis. 2009; 47:603–610. [PubMed: 19548312]
35. Hudson C, Clements D, Friday RV, et al. Xsox17alpha and -beta mediate endoderm formation in Xenopus. Cell. 1997; 91:397–405. [PubMed: 9363948]
36. Liao WP, Uetzmann L, Burtscher I, et al. Generation of a mouse line expressing Sox17-driven Cre recombinase with specific activity in arteries. Genesis. 2009; 47:476–483. [PubMed: 19415628]
37. Grogan SP, Miyaki S, Asahara H, et al. Mesenchymal progenitor cell markers in human articular cartilage: normal distribution and changes in osteoarthritis. Arthritis Res Ther. 2009; 11:R85. [PubMed: 19500336]
38. Garry DJ, Olson EN. A common progenitor at the heart of development. Cell. 2006; 127:1101–1104. [PubMed: 17174889]
39. Ferkowicz MJ, Starr M, Xie X, et al. CD41 expression defines the onset of primitive and definitive hematopoiesis in the murine embryo. Development. 2003; 130:4393–4403. [PubMed: 12900455]
40. Nishikawa SI, Nishikawa S, Kawamoto H, et al. In vitro generation of lymphohematopoietic cells from endothelial cells purified from murine embryos. Immunity. 1998; 8:761–769. [PubMed: 9655490]
41. Yoshimoto M, Montecino-Rodriguez E, Ferkowicz MJ, et al. Embryonic day 9 yolk sac and intra-embryonic hemogenic endothelium independently generate a B-1 and marginal zone progenitor lacking B-2 potential. Proc Natl Acad Sci U S A. 2011; 108:1468–1473. [PubMed: 21209332]
42. Yoshimoto M, Porayette P, Glosson NL, et al. Autonomous murine T cell progenitor production in the extra-embryonic yolk sac prior to HSC emergence. Blood. 2012
43. Mi H, Dong Q, Muruganujan A, et al. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res. 2010; 38:D204–D210. [PubMed: 20015972]
44. Heller RS, Dichmann DS, Jensen J, et al. Expression patterns of Wnts, Frizzleds, sFRPs, and misexpression in transgenic mice suggesting a role for Wnts in pancreas and foregut pattern formation. Dev Dyn. 2002; 225:260–270. [PubMed: 12412008]
45. Wandzioch E, Zaret KS. Dynamic signaling network for the specification of embryonic pancreas and liver progenitors. Science. 2009; 324:1707–1710. [PubMed: 19556507]
46. Kanai Y, Kanai-Azuma M, Noce T, et al. 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. 1996; 133:667–681. [PubMed: 8636240]
47. Burtscher I, Barkey W, Schwarzfischer M, et al. The Sox17-mCherry fusion mouse line allows visualization of endoderm and vascular endothelial development. Genesis. 2011
48. Baron MH, Isern J, Fraser ST. The embryonic origins of erythropoiesis in mammals. Blood. 2012; 119:4828–4837. [PubMed: 22337720]

49. Fraser ST, Isern J, Baron MH. Maturation and enucleation of primitive erythroblasts during mouse embryogenesis is accompanied by changes in cell-surface antigen expression. Blood. 2007; 109:343–352. [PubMed: 16940424]
50. Kalsotra A, Cooper TA. Functional consequences of developmentally regulated alternative splicing. Nat Rev Genet. 2011; 12:715–729. [PubMed: 21921927]
51. Pendeville H, Winandy M, Manfroid I, et al. Zebrafish Sox7 and Sox18 function together to control arterial-venous identity. Dev Biol. 2008; 317:405–416. [PubMed: 18377889]
52. Patterson M, Chan DN, Ha I, et al. Defining the nature of human pluripotent stem cell progeny. Cell Res. 2011
53. Seymour PA, Freude KK, Tran MN, et al. SOX9 is required for maintenance of the pancreatic progenitor cell pool. Proc Natl Acad Sci U S A. 2007; 104:1865–1870. [PubMed: 17267606]
54. Furuyama K, Kawaguchi Y, Akiyama H, et al. Continuous cell supply from a Sox9-expressing progenitor zone in adult liver, exocrine pancreas and intestine. Nat Genet. 2011; 43:34–41. [PubMed: 21113154]
55. Dessimoz J, Bonnard C, Huelsken J, et al. Pancreas-specific deletion of beta-catenin reveals Wnt-dependent and Wnt-independent functions during development. Curr Biol. 2005; 15:1677–1683. [PubMed: 16169491]
56. Sinner D, Rankin S, Lee M, et al. Sox17 and beta-catenin cooperate to regulate the transcription of endodermal genes. Development. 2004; 131:3069–3080. [PubMed: 15163629]
57. De Val S, Black BL. Transcriptional control of endothelial cell development. Dev Cell. 2009; 16:180–195. [PubMed: 19217421]
58. Reynaud D, Ravet E, Titeux M, et al. SCL/TAL1 expression level regulates human hematopoietic stem cell self-renewal and engraftment. Blood. 2005; 106:2318–2328. [PubMed: 15961517]
59. Gerstein RM. Deciding the decider: Mef2c in hematopoiesis. Nat Immunol. 2009; 10:235–236. [PubMed: 19221552]
60. Ling KW, Ottersbach K, van Hamburg JP, et al. GATA-2 plays two functionally distinct roles during the ontogeny of hematopoietic stem cells. J Exp Med. 2004; 200:871–882. [PubMed: 15466621]
61. Medvinsky A, Rybtsov S, Taoudi S. Embryonic origin of the adult hematopoietic system: advances and questions. Development. 2011; 138:1017–1031. [PubMed: 21343360]
62. Zhang H, Ables ET, Pope CF, et al. Multiple, temporal-specific roles for HNF6 in pancreatic endocrine and ductal differentiation. Mech Dev. 2009; 126:958–973. [PubMed: 19766716]

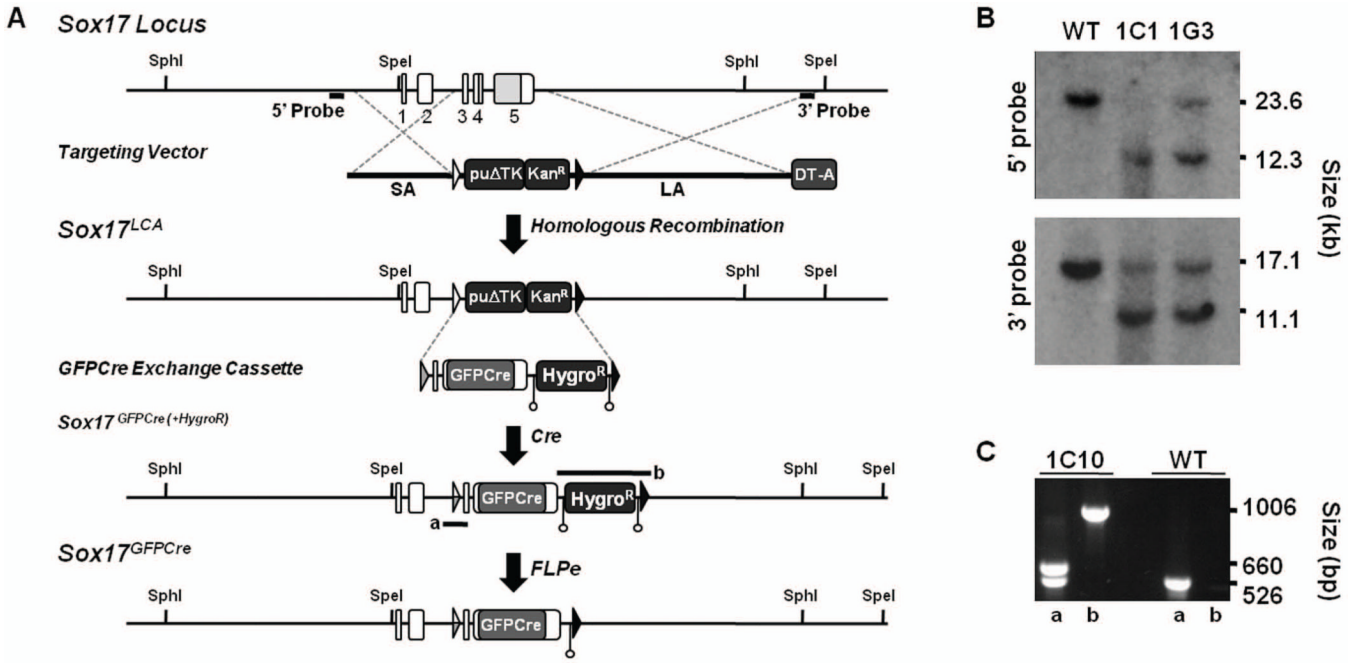

Figure 1. Generation the *Sox17^{GFPCre}* allele

A) Diagram of the *Sox17* locus, targeting vector, *Sox17^{LCA}* allele, GFPCre exchange cassette, *Sox17$^{GFPCre(+HygroR)}$*, and *Sox17^{GFPCre}* allele. A targeting vector for the mouse *Sox17* gene was constructed where sequence including exons 3 to 5, which contain the coding region of *Sox17*, was replaced with a *puromycin resistance-Δ-thymidine kinase* fusion gene (*puΔTK*) and an EM7-driven *kanamycin resistance* gene (*Kan^R*) flanked by lox66 (open triangle) and lox2272 (black triangle) sites. The GFPCre exchange cassette was flanked by lox71 (grey triangle) and lox2272 sites and contained a *PGK*-driven *hygromycin resistance* gene (*Hygro^R*) flanked by FRT sites (open circles). Following exchange into *Sox17^{LCA}*-containing mouse ES cells by RMCE, mice containing the *Sox17$^{GFPCre(+HygroR)}$* allele were bred with FLPe-expressing transgenic mice, thereby generating the final *Sox17^{GFPCre}* allele. PCR amplifications for 5' and 3' screening of *Sox17$^{GFPCre(+HygroR)}$* allele depicted as a and b. SA, short arm. LA, long arm. B) Southern blot analysis of genomic DNA from puromycin-resistant *Sox17^{LCA}* ES cells. DNA was digested with SphI or SpeI and hybridized with a 5' or 3' probe as indicated in panel A. Clones 1C1 and 1G3 were correctly targeted by presence of a 12.3 kb and 11.1 kb band on the 5' and 3' ends, respectively. Clone 1G3 was used for RMCE. C) PCR screening of *Sox17$^{GFPCre(+HygroR)}$* exchanged clones. The proper exchange of clone 1C10 was identified by 660 and 1,006 basepair (bp) bands on the 5' and 3' ends, respectively.

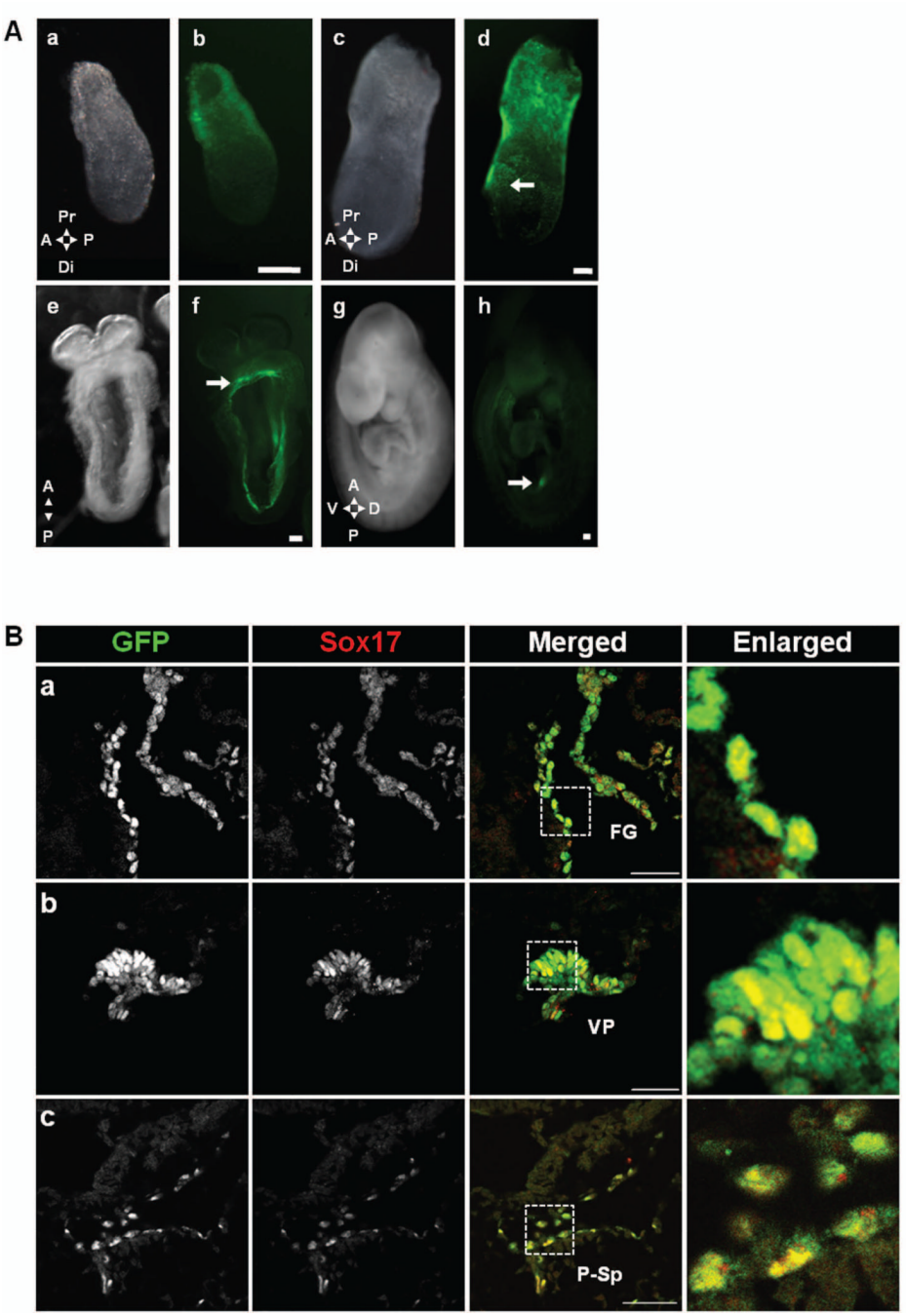

Figure 2. Expression pattern of *Sox17*$^{GFPCre/+}$ during development

A) GFPCre expression was observed from E6.5 to E9.5 in *Sox17*$^{GFPCre/+}$ embryos. At E6.5, GFPCre was observed only in the extra-embryonic region (a&b). Conversely, at E7.5, it was seen in the embryonic region (c&d), specially the definitive endoderm area (arrow). At E8.5, GFPCre was observed in the gut tube area (e&f) with expression in the foregut region marked (arrow). At E9.5, the expression was diminished in the gut (g&h) but was still seen in the ventral pancreatic bud (arrow). Scale bar = 100 μm. Anterior (A), posterior (P), proximal (Pr), distal (Di), dorsal (D), ventral (V). B) Immunolabeling revealed co-localization of both GFPCre and Sox17 in E8.5 and E9.5 *Sox17*$^{GFPCre/+}$ mouse embryos. At E8.5, GFPCre and Sox17 were co-expressed in the foregut endoderm (a) and at E9.5 in the

ventral pancreatic bud (b) and the para-aortic splanchnopleural area (c). White boxed areas depict regions enlarged. Scale bar = 50 μm. Foregut (FG), ventral pancreatic bud (VP), para-aortic splanchnopleural (P-Sp).

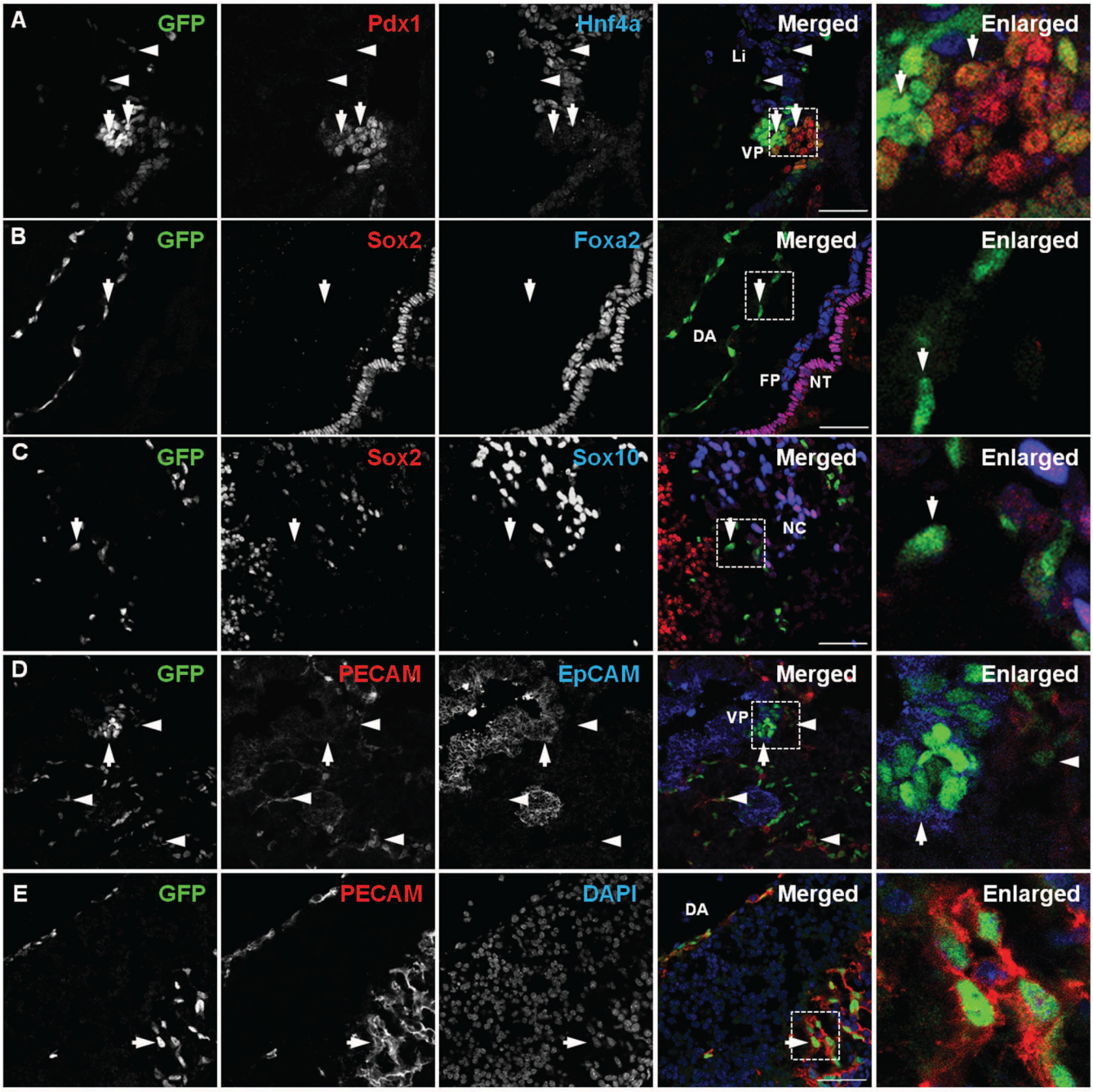

Figure 3. Two distinct types of *Sox17*^{GFPCre}-expressing cells at E9.5

A) Immunolabeling revealed that GFP expression at E9.5 in the ventral pancreatic bud (VP) was co-localized with Pdx1 (arrows). Several GFPCre-expressing cells were found in the liver bud region (Li) but did not co-localize with Hnf4α (arrowheads). B) GFP was detected in the dorsal aorta (DA); however, it did not co-localize with Sox2, an early ectoderm marker, or Foxa2, a floor plate marker, (arrow). C) GFP was detected in the neural tube (NT); however, it did not co-localize either Sox2 or Sox10, a neural crest cell marker (arrow). D) GFP in the ventral pancreatic bud co-localized with EpCAM (arrow) but not with PECAM (arrowheads). E) GFP in the neural tube co-localized with PECAM (arrow).

White boxed areas depict regions enlarged. Scale bar = 50 μm. Floor plate (FP), neural crest cells (NC).

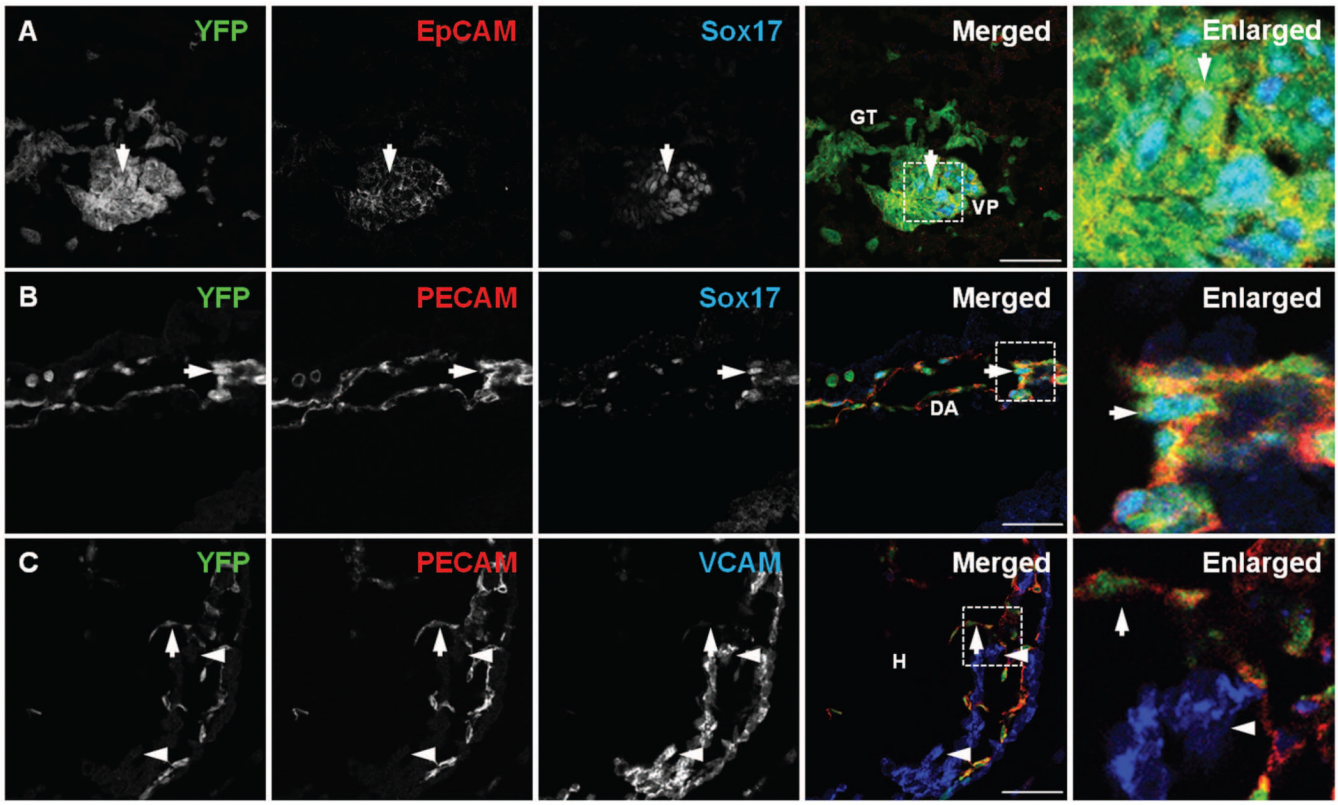

Figure 4. Fate tracing of Sox17-expressing cells at E9.5

A) In E9.5 *R26R*^{eYFP};*Sox17*^{GFPCre} mouse embryos, YFP was detected in the ventral pancreatic bud (VP) and gut tube[62] and displayed co-localization with EpCAM. YFP also co-localized with *Sox17*^{GFPCre}-expressing cells in the ventral pancreatic bud (arrow). B) YFP was detected in the dorsal aorta (DA) and co-localized with PECAM. Some YFP-positive cells also co-localized with Sox17 (arrow). C) YFP co-localized with PECAM in the endocardium of heart (H) (arrow); however, it did not co-localize with VCAM in the myocardium (arrowheads). White boxed areas depict regions enlarged. Scale bar = 50 μm.

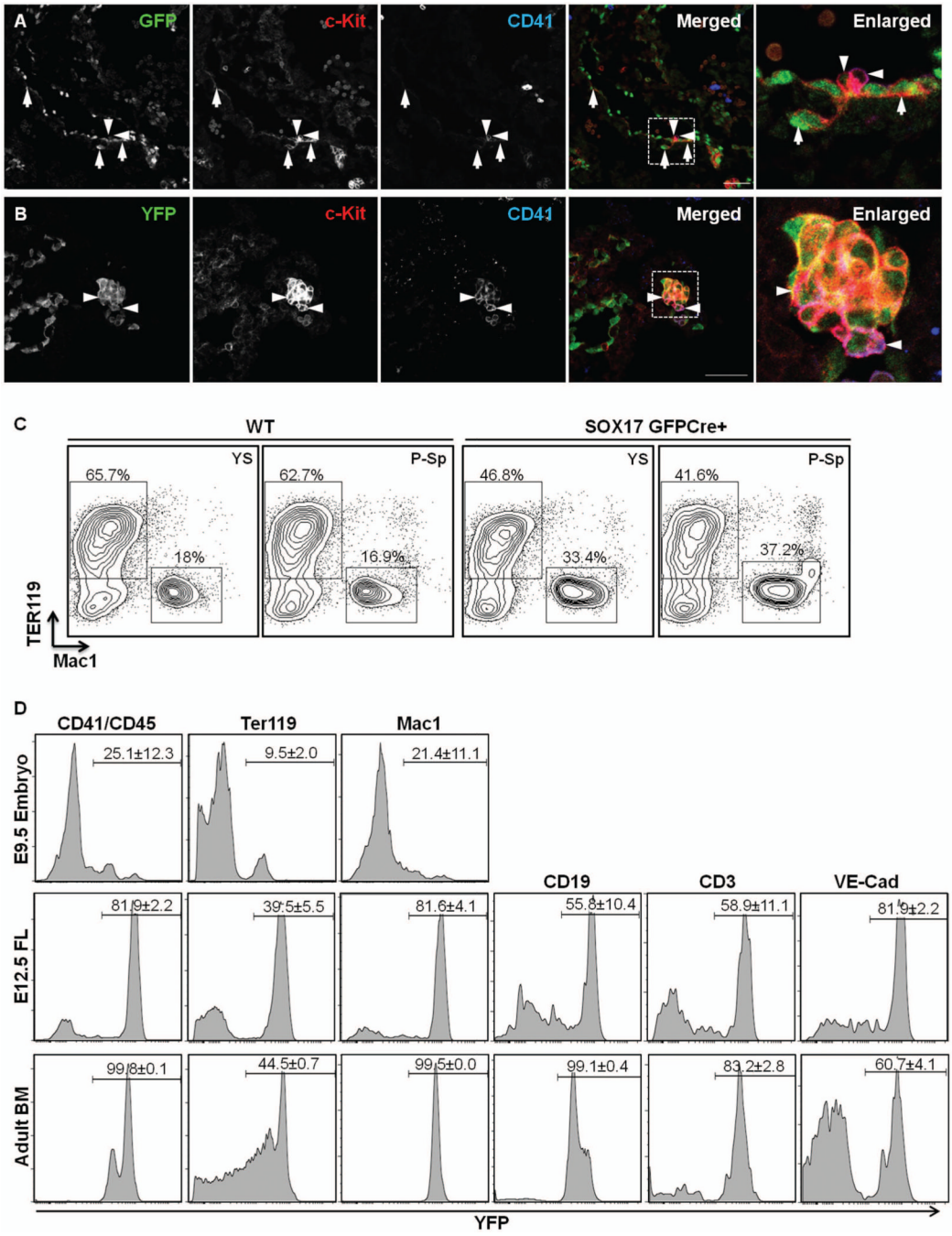

Figure 5. Sox17-expressing endothelial cells exhibit hemogenic potential

A) In E9.5 *Sox17*$^{GFPCre/+}$ embryos, GFP was detected in the para-aortic splanchnopleural (P-Sp) area. GFP co-localized with c-Kit-positive cells in the aortic floor (arrows); however, GFP was diminished or not detected in CD41-positive and/or c-Kit positive hematopoietic cells (arrowheads). White boxed areas depict regions enlarged. Scale bar = 50 μm. B) In E9.5 *R26R*^{eYFP};*Sox17*^{GFPCre} embryos, YFP was detected in the para-aortic splanchnopleural (P-Sp) area. YFP co-localized with c-Kit- and CD41-positive cells (arrowheads). C) Both erythrocytes and myeloid (Ter119+ or Mac1+) cells were differentiated from wild type and Sox17-expressing ECs obtained from E9.5 YS and P-Sp. D) YFP expression in hematopoietic cells from E9.5 embryos, E12.5 fetal liver (FL), and

adult mouse BM of *R26R*^{eYFP};*Sox17*^{GFPCre} mice (n≥3). YFP-expression was analyzed with hematopoietic cell marker-gated cells.

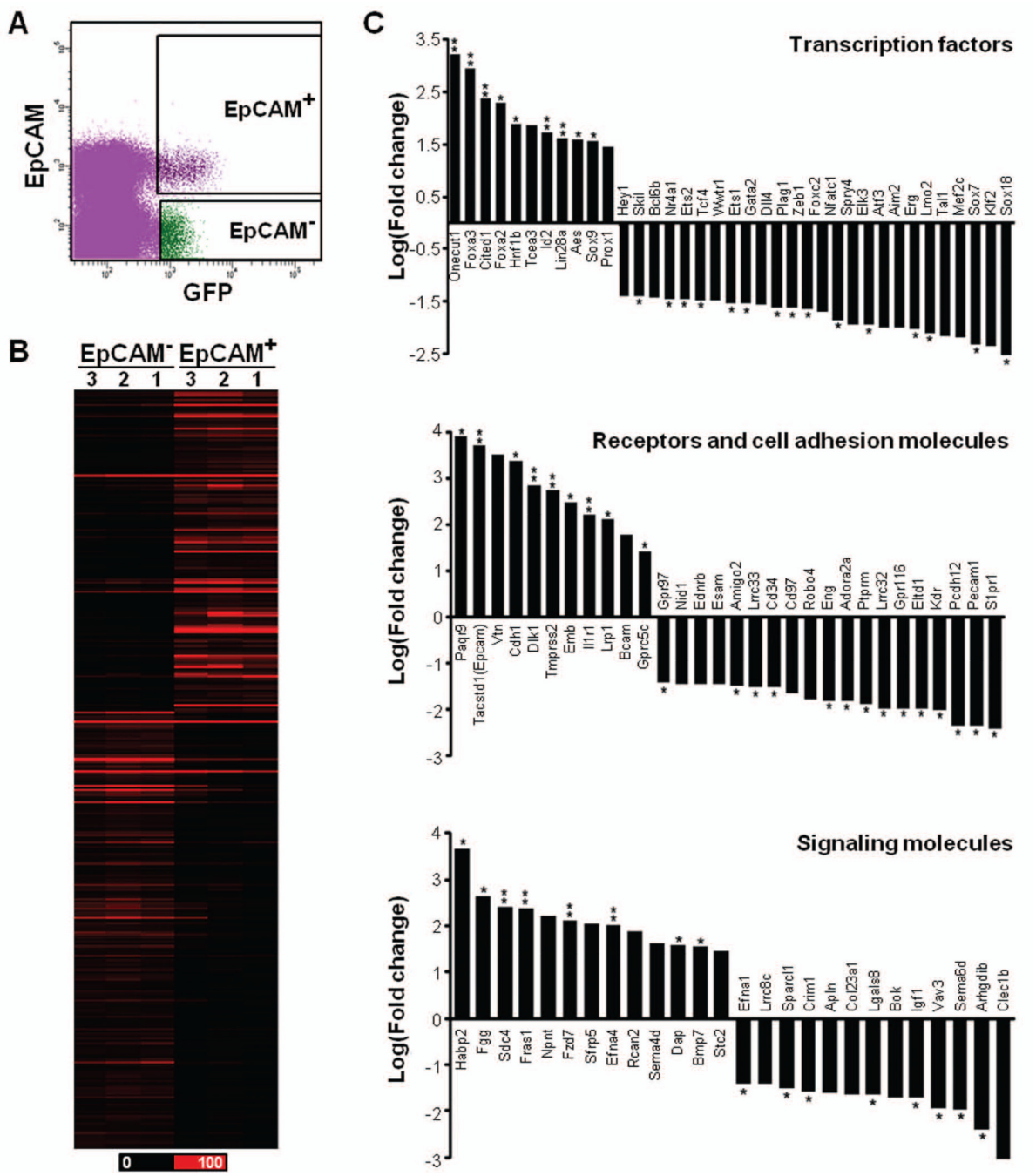

Figure 6. Comparison of differentially expressed genes in two populations

A) Fluorescence-activated cell sorting (FACS) was used to isolate GFP/EpCAM co-positive cells representing ventral pancreatic epithelial cells (EpCAM$^+$) and GFP$^+$/EpCAM$^-$ cells representing hemogenic ECs (EpCAM$^-$) from dissected E9.5 *Sox17*$^{GFPCre/+}$ embryo midguts. B) The distinct difference between the EpCAM$^+$ and EpCAM$^-$ cell populations is evident in the heat map which displays the RPKM values for 321 genes from three biological replicates for either EpCAM$^+$ or EpCAM$^-$. Black color corresponds to an RPKM value of 0, and the brightest red corresponds to ≥ 100 RPKM value. C) The selected transcripts were clustered according to protein class, and the fold change (natural log scale)

indicating gene expression in $EpCAM^+$ cells as compared to $EpCAM^-$ cells is shown. ** > 0.9 and * > 0.85 confidence value.

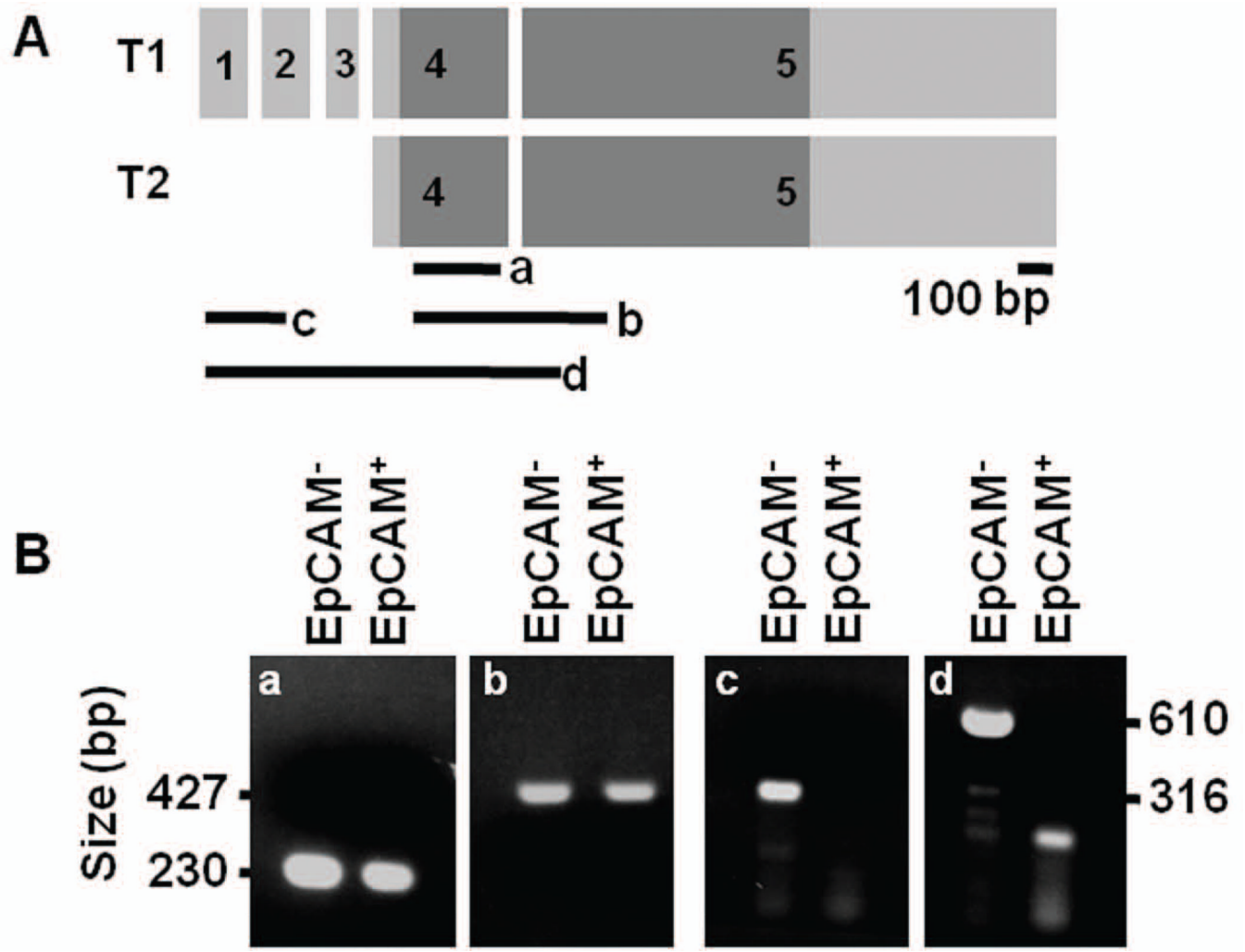

Figure 7. Isolation of Sox17-expressing cells and identification of alternative variants of *Sox17* transcript in EpCAM$^+$ and EpCAM$^-$ cells

A) Schematic of two transcript variants of *Sox17* (T1 & T2). Dark gray box shows coding regions. Black lines (a – d) indicate the amplified regions for PCR. B) Both EpCAM$^+$ and EpCAM$^-$ cells samples amplified sequence within the coding regions (a and b); however, only the EpCAM$^-$ sample amplified regions spanning the first three exons. The band in the EpCAM+ lane in d is non-specific (expected band size = 316 bp) (c and d).