

Sox17-dependent gene expression and early heart and gut development in Sox17-deficient mouse embryos

SABINE PFISTER¹, VANESSA J. JONES¹, MELINDA POWER¹, GERMAINE L. TRUISI¹,
POH-LYNN KHOO¹, KIRSTEN A. STEINER¹, MASAMI KANAI-AZUMA², YOSHIKIRA KANA³,
PATRICK P. L. TAM^{1,4} and DAVID A. F. LOEBEL^{*,1,4,5}

¹Embryology Unit, Children's Medical Research Institute, New South Wales, Australia,

²Center for Experimental Animals, Tokyo Medical and Dental University, Tokyo, Japan,

³Department of Veterinary Medicine, University of Tokyo, Tokyo, Japan and

⁴Sydney Medical School, University of Sydney, New South Wales, Australia

ABSTRACT *Sox17* is a transcription factor that is required for maintenance of the definitive endoderm in mouse embryos. By expression profiling of wild-type and mutant embryos and *Sox17*-overexpressing hepatoma cells, we identified genes with *Sox17*-dependent expression. Among the genes that were up-regulated in *Sox17*-null embryos and down-regulated by *Sox17*-expressing HepG2 cells is a set of genes that are expressed in the developing liver, suggesting that one function of *Sox17* is the repression of liver gene expression, which is compatible with a role for *Sox17* in maintaining the definitive endoderm in a progenitor state. Consistent with these findings, *Sox17*^{-/-} cells display a diminished capacity to contribute to the definitive endoderm when transplanted into wild-type hosts. Analysis of gene ontology further revealed that many genes related to heart development were downregulated in *Sox17*-null embryos. This is associated with the defective development of the heart in the mutant embryos, which is accompanied by localised loss of *Myocd*-expressing cardiogenic progenitors and the malformation of the anterior intestinal portal.

KEY WORDS: *Sox17*, downstream gene, endoderm, heart morphogenesis

Introduction

Endoderm specification and differentiation during embryogenesis in frogs and fish involves transcriptional activation of zinc finger GATA factors, Mix-homeodomain proteins and High Mobility Group SOX proteins by TGF β /Nodal signalling and T-box and POU-domain transcription factors (Kiefer, 2003; Lunde *et al.*, 2004; Reim *et al.*, 2004; Shivdasani, 2002; Stainier, 2002; Woodland and Zorn, 2008; Zorn and Wells, 2007; Zorn and Wells, 2009). A similar set of genes including *Mix/Mixer*, *Sox17* and *Foxa* is associated with endoderm formation in mouse, frogs and fish, but their function and regulation differ in mouse development.

The mouse *Mix1* gene is expressed in the primitive streak where the endoderm progenitors reside but not in the definitive endoderm (Pearce and Evans, 1999), unlike in *Xenopus* where *Mix/Mixer* expression is endoderm-specific. *Mix1* function is critical for the generation of definitive endoderm in the mouse.

Mix1-null embryos contain a reduced population of endoderm that lacks the expression of molecular markers of definitive endoderm (Hart *et al.*, 2002) and *Mix1*-null embryonic stem cells and primitive streak cells are less efficient in populating the definitive endoderm (Hart *et al.*, 2002; Tam *et al.*, 2007). Whereas *Sox17* is essential for endoderm formation in *Xenopus* and Zebrafish embryos, loss of *Sox17* in the mouse does not appear to affect lineage allocation but impairs the viability of the endoderm of the foregut and the proliferation of endoderm in the posterior gut (Kanai-Azuma *et al.*, 2002; Tam *et al.*, 2003). In *Xenopus*, the endodermal genes *HNF1 β* , *Foxa1*, *Foxa2* and *Endodermin* are direct transcriptional targets of *Sox17* (Clements *et al.*, 2003; Sinner *et al.*, 2004). The expression of *Foxa1* and

Abbreviations used in this paper: APS, anterior primitive streak; GO, gene ontology.

*Address correspondence to: David Loebel, Embryology Unit, Children's Medical Research Institute, Locked Bag 23 Wentworthville, NSW Australia 2145.
Fax: +61-2-9687-2120. e-mail: Dloebel@cmri.org.au

Supplementary Material (7 tables) for this paper is available at: <http://dx.doi.org/10.1387/ijdb.103158sp>

Accepted: 11 January 2011. Final author corrected PDF published online: 1 February 2011.

ISSN: Online 1696-3547, Print 0214-6282

© 2011 UBC Press
Printed in Spain

Foxa2 is severely reduced in *Sox17*^{-/-} mouse embryos (Kanai-Azuma *et al.*, 2002), but it is not clear whether they are direct transcriptional targets of Sox17.

At present, the transcriptional targets of *Sox17* in the mouse embryo at early-organogenesis when the null-mutant phenotype first manifests have not been identified. Previous studies dealing with Sox17-dependent gene expression in mammalian cells have used mouse or human embryonic stem (ES) cells, in which *Sox17* was over-expressed. Results conflicted between the two stem cell types, with *Sox17* inducing molecular characteristics of definitive endoderm in human ES cells (Seguin *et al.*, 2008) but those of extraembryonic endoderm in mouse ES cells (Niakan *et al.*, 2010). However, in mouse ES cells, Sox17 can induce both definitive and visceral endoderm differentiation under some circumstances (Qu *et al.*, 2008). The variation in results between *in vitro* models suggests that Sox17-dependent changes in the transcriptome and the identification of potential Sox17 transcriptional targets is influenced by the inherent properties of the cell model used.

In this study, we examined gene expression profiles in the mouse embryo, in which *Sox17* has been shown to be required for development of the definitive but not visceral endoderm, to identify *Sox17*-dependent changes in gene transcription. To

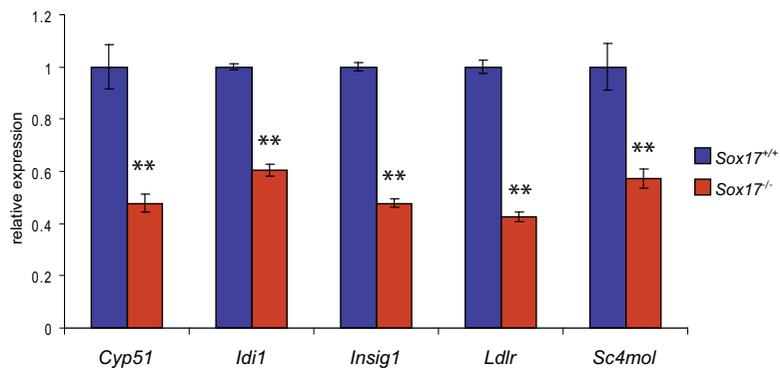


Fig. 1. Validation of the down-regulation of genes associated with cholesterol metabolism in *Sox17*^{-/-} embryos. Real-time quantitative RT-PCR analysis was performed on cDNA generated from RNA samples of E8.25 *Sox17*^{+/+} and *Sox17*^{-/-} embryos. Expression levels are shown relative to wild-type. Gapd was the reference gene. Error bars indicate standard error of the mean. N=3 for each assay. **Significant difference at $p < 0.01$ by two-tailed t-test from *Sox17*^{+/+} embryos.

identify possible Sox17 target genes, we compared the changes in gene expression profiles of mutant embryos with those resulting from *Sox17*-overexpression in HepG2 cells. The findings show that changes in *Sox17* activity impact on the genetic activity associated with endoderm development. Embryological analysis

TABLE 1

TOP 50 GENES SIGNIFICANTLY UP-REGULATED IN *SOX17*^{-/-} EMBRYOS ORDERED BY LOG₂ FOLD-CHANGE IN EXPRESSION

Symbol	Description	GenBank	GO Biological functions summary	Probe	Log ₂ FoldChange
<i>Tsix</i>	X (inactive)-specific transcript, antisense	BG806300	dosage compensation;	1436936_s_at	3.27
<i>Afp</i>	alpha fetoprotein	NM_007423		1416645_a_at 1416646_at 1436879_x_at	3.14 2.96 2.63
<i>Slc2a2</i>	solute carrier family 2 (facilitated glucose transporter), member 2	NM_031197	carbohydrate transport	1449067_at	2.83
<i>Tspan8</i>	tetraspanin 8	BC025461		1424649_a_at	2.77
<i>D7Erd715e</i>	DNA segment, Chr 7, ERATO Doi 715, expressed	BB314814		1436964_at	2.72
<i>Timd2</i>	T-cell immunoglobulin and mucin domain containing 2	BC028829		1418766_s_at	2.37
<i>Apo2</i>	apolipoprotein C-II	NM_009695	lipid transport	1418069_at	2.35
<i>Xlr4b</i>	X-linked lymphocyte-regulated 4B	NM_021365		1449347_a_at	2.35
<i>Apoa1</i>	apolipoprotein A-I	NM_009692	endothelial cell proliferation; lipid transport	1438840_x_at 1419232_a_at 1455201_x_at 1419233_x_at	2.6 2.55 2.34 2.28
<i>Ccnb1ip1</i>	cyclin B1 interacting protein 1	BG066504	blastocyst formation; apoptosis	1435998_at	2.19
<i>Fgg</i>	fibrinogen, gamma polypeptide	NM_133862	platelet activation	1416025_at	1.99
<i>Lgals2</i>	lectin, galactose-binding, soluble 2	NM_025622		1417079_s_at	1.89
<i>Pdzk1</i>	PDZ domain containing 1	AK006269		1431701_a_at	1.85
<i>Cldn2</i>	claudin 2	NM_016675	calcium-independent cell-cell adhesion	1417231_at	1.82
<i>Cubn</i>	cubilin (intrinsic factor-cobalamin receptor)	AF197159	receptor-mediated endocytosis	1426990_at 1452270_s_at	1.94 1.8
<i>Rbp4</i>	retinol binding protein 4, plasma	U63146	eye development; spermatogenesis; lung development; cardiac muscle development female genitalia morphogenesis	1426225_at	1.8
<i>2410003J06Rik</i>	RIKEN cDNA 2410003J06 gene	AK010362		1429701_at	1.75
<i>Sfmbt2</i>	Scm-like with four mbt domains 2	BM200222	regulation of transcription	1434353_at	1.72
<i>Trap1a</i>	tumor rejection antigen P1A	NM_011635		1460226_at	1.7
<i>Apoa4</i>	apolipoprotein A-IV	BC010769	lipid transport; regulation of cholesterol absorption	1417761_at	1.68
<i>Nrk</i>	Nik related kinase	AK012873	protein amino acid phosphorylation	1450078_at 1450079_at	1.94 1.65
<i>Ttr</i>	transthyretin	BG141874	thyroid hormone generation	1455913_x_at 1459737_s_at 1451580_a_at 1454608_x_at	1.77 1.71 1.65 1.64

<i>Amn</i>	amniotless	NM_033603	excretion	1417920_at	1.6
<i>Fgb</i>	fibrinogen, B beta polypeptide	AK011118	platelet activation	1428079_at	1.53
<i>Spp2</i>	secreted phosphoprotein 2	NM_029269	bone remodeling	1418916_a_at	1.53
<i>Xlr3a</i>	X-linked lymphocyte-regulated 3A	NM_011726		1420357_s_at	1.5
<i>Paip1</i>	polyadenylate binding protein-interacting protein 1	BB381990	regulation of translation	1425521_att	1.49
				1441955_s_a	1.61
<i>Vil1</i>	villin 1	NM_009509	cytoskeleton organization	1448837_at	1.41
<i>Mfap3l</i>	microfibrillar-associated protein 3-like	AV262974		1428904_at	1.46
				1441481_at	0.92
<i>Rab11fip5</i>	RAB11 family interacting protein 5 (class I)	BF682225	protein transport	1434314_s_at	1.4
				1427405_s_at	0.62
<i>Ctsh</i>	cathepsin H	NM_007801	proteolysis	1443814_x_at	1.39
				1418365_at	1.46
<i>2610528J11Rik</i>	RIKEN cDNA 2610528J11 gene	AK012175		1450947_at	1.39
<i>Mbl2</i>	mannose-binding lectin (protein C) 2	NM_010776	innate immune response	1418787_at	1.38
<i>Npl</i>	N-acetylneuraminidase 7	BC022734	carbohydrate metabolic process	1424265_at	1.37
<i>Slc7a9</i>	solute carrier family 7 (cationic amino acid transporter, y+ system), member 9	NM_021291	amino acid transport	1448783_at	1.36
<i>Tdh</i>	L-threonine dehydrogenase	NM_021480	cellular metabolic process	1449064_at	1.36
<i>Car7</i>	carbonic anhydrase 7	BB193643		1443824_s_at	1.35
<i>Soat2</i>	sterol O-acyltransferase 2	BC025931	lipid; steroid; cholesterol metabolic process	1460722_at	1.35
<i>Trf</i>	transferrin	AF440692	iron ion transport	1425546_a_at	1.33
<i>Fga</i>	fibrinogen, alpha polypeptide	BC005467	blood coagulation	1424279_at	1.31
<i>Rhox5</i>	reproductive homeobox 5	BM210473	sperm motility; germ cell programmed cell death	1423429_at	1.3
<i>Slco4c1</i>	solute carrier organic anion transporter family, member 4C1	AV024403	spermatogenesis; organic anion transport c	1437870_at	1.27
				1460616_at	1.29
<i>Reep6</i>	receptor accessory protein 6	AK002562		1430128_a_at	1.24
<i>Aass</i>	aminoadipate-semialdehyde synthase	BF687395	generation of precursor metabolites	1423523_at	1.22
<i>Morc4</i>	microorchidia 4	AV036158		1434436_at	1.18
<i>Lgmn</i>	legumain	NM_011175	negative regulation of growth	1448883_at	1.16
<i>Adora2b</i>	adenosine A2b receptor	BB709140	mast cell degranulation; relaxation of vascular smooth muscle	1434430_s_att	1.16
				1434431_x_a	0.78
<i>Cfi</i>	complement component factor i	NM_007686	proteolysis immune response complement activation; classical pathway	1418724_at	1.15
<i>Gpr155</i>	G protein-coupled receptor 155	BB762731		1452353_at	1.14
<i>Dab2</i>	disabled homolog 2 (<i>Drosophila</i>)	NM_023118	cell morphogenesis	1420498_a_at	1.14

further shows that *Sox17*-null cells are less efficient in contributing to the definitive endoderm. Loss of *Sox17* in the endoderm also disrupts form-shaping activity in the anterior intestinal portal and the morphogenesis of the heart tube. *Sox17*-null mutant embryos also showed defective differentiation of the cardiogenic mesoderm. These findings are consistent with the definitive endoderm being a source of morphogenetic cues that patterns the tissues associated with the gut (Lewis and Tam, 2006),

Results

Identification of Sox17-dependent changes in gene expression

To identify Sox17-dependent gene expression we compared the expression profile of *Sox17*-null mutant and wild type embryos at the 4-5 somite stage, in order to capture the changes in the transcriptome before the *Sox17*-null phenotype becomes morphologically evident (Kanai-Azuma *et al.*, 2002). This was expected to minimize the confounding effects caused by secondary changes in gene expression due to aberrant development. Among the up-regulated transcripts in the mutant embryos, there is a preponderance of those encoding apolipoproteins and solute transporters (Table 1, Supplementary Table S1). This is reflected in the over-representation of genes defined by Gene Ontology (GO) biological function terms relating to homeostasis, absorption, transport as well as cholesterol, steroid and lipid metabolism, catabolism and biosynthesis (Table 2A). These genes include *Apoa1*, *Apob*, *Apoc2*, *Apom* and *Slc2a2* that are known to be

expressed primarily in the extraembryonic visceral endoderm during post-implantation development. Other genes that are expressed in the visceral endoderm were also upregulated in mutant embryos (including *Afp*, *Cubn*, *Dab*, *Rbp4*, *Rhox5* and *Ttr*; Table 1, Supplementary Table S1). Some of these putative visceral endoderm genes (e.g.: *Afp*, *Ttr* and *Apo* family members) are also expressed in the liver later in development but they are normally only expressed at low levels in the foregut endoderm of wild-type early-somite stage embryos. The profiling results are consistent with the observation that definitive endoderm is replaced by cells displaying phenotypes characteristic of the extraembryonic visceral endoderm in the embryonic gut region of the *Sox17*-null mutant (Kanai-Azuma *et al.*, 2002). Alternatively, it may be that one function of Sox17 is to maintain liver-specific gene expression at a low level in the early definitive endoderm.

Among the genes that were down-regulated in the mutant embryos (Table 3, Supplementary Table S1), there is an over-representation of genes involved in pathways for biosynthesis, metabolism and transport of lipids and cholesterol (Table 2B). Transcripts of genes involved in cholesterol biosynthesis (*Cyp51*, *Idi1*, *Insig1* and *Sc4mol*) and transport (*Ldlh*) showed reduced expression in the absence of *Sox17*, and downregulation was validated by qRT-PCR (Fig. 1). Cholesterol is a key component of the Hedgehog signalling pathway that is essential for patterning the primitive gut tube (Harmon *et al.*, 2002). Notably, the concerted changes of the cholesterol pathway genes are consistent with the concept that cholesterologenic enzymes constitute a synexpression group in embryos and cultured fibroblasts (Iyer *et al.*

al., 1999; Laubner *et al.*, 2003; Marijanovic *et al.*, 2003). In the *Sox17*-null embryo, expression of *lhh* in the gut endoderm and *Ptch1* in the lateral plate mesoderm is down-regulated, indicating that loss of *Sox17* function is associated with reduced Hedgehog signalling activity (Kanai-Azuma *et al.*, 2002). Also of note is the reduced expression of *Neprn* (Table 3, Supplementary Table S1), which is normally expressed in the lateral midgut endoderm at this stage (Hou *et al.*, 2007). Its down-regulation is consistent with the depletion of definitive endoderm in the mutant embryos.

Analysis of networks of transcriptional targets using Metacore (www.genego.com) revealed that 75 of the differentially expressed genes were also potentially regulated by *Hnf4*, with either a known effect on transcription or evidence for binding of *Hnf4* to

putative regulatory regions (Supplementary Fig. S1). *Hnf4* is expressed in liver and visceral endoderm suggesting a change in cell composition of the gut or dys-regulation of genes associated with liver development. *Sox17* may also be regulated by *Hnf4* via a cross-acting *Hnf4* and *Sox17* transcriptional networks.

To identify changes in *Sox17*-dependent gene expression without the complicating factor of altered tissue contents due to the loss of gene function, we over-expressed *Sox17* in HepG2 hepatoma cells, an endoderm cell line of epithelial morphology which originates from the liver tissue. HepG2 cells were transfected with either a GFP expressing construct or a bi-cistronic construct expressing *Sox17* along with a GFP reporter. Transfected cells were sorted for GFP expression, and expression

profiles analysed using Illumina human Sentrix-6 chips. Results were filtered for significantly different expression levels between GFP-only and *Sox17*-IRES-GFP transfected cells (Supplementary Table S2). The list of differentially expressed genes was compared to the list of genes that were differentially expressed between wild-type and *Sox17*-mutant embryos, and we identified 32 genes that were differentially expressed in both sets of data (Table 4). Of these, 21 showed concordance between the two data sets; that is, these genes were up-regulated in the HepG2 over-expressing *Sox17* but down-regulated in *Sox17*-null embryos, and vice versa. These 21 genes are therefore potential candidates for positive or negative regulation by *Sox17*. Results of qRT-PCR have validated in the majority of cases the differential expression of these genes in transfected HepG2 cells (Fig. 2A) and in embryos (Fig. 2B).

Two previous studies have addressed the question of *Sox17*-dependent gene expression either in mouse (Niakan *et al.*, 2010) or human (Seguin *et al.*, 2008) ES cells in which *Sox17* was over-expressed. To gain additional insights into the genes that potentially act downstream of *Sox17* we compared our embryo and HepG2 microarray data to these studies. Comparing the list of genes that were up- or down-regulated in *Sox17*-null embryos with those that were changed in the opposite direction in *Sox17*-over-expressing mouse (Niakan *et al.*, 2010) and human (GEO accession GSE10809) ES cells revealed an overlap of 6 genes each in the embryo-mouse ES and embryo-human ES sets. These genes are involved in diverse processes including metabolism (folic acid, TCA, sorbitol, carbohydrate), cytoskeletal arrangement, transcription and cell proliferation (Supplementary Table S3A, B). There

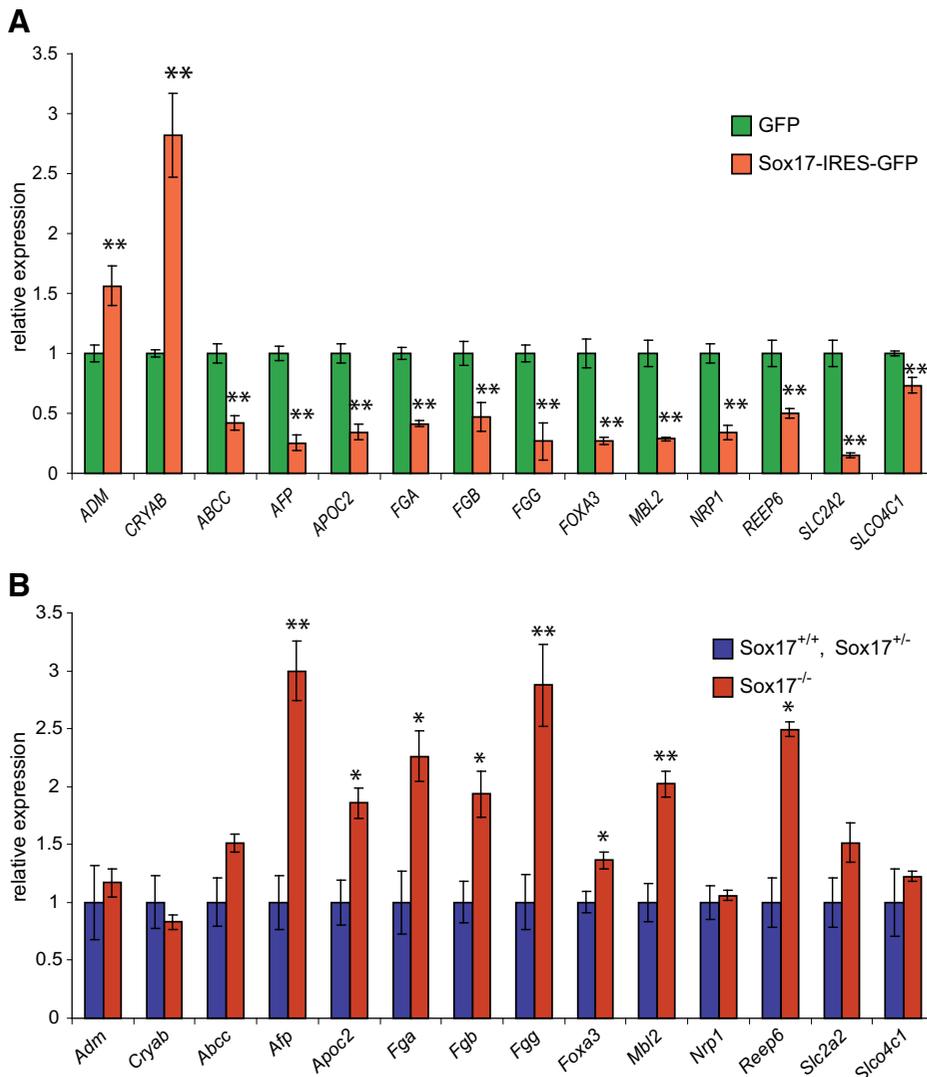


Fig. 2. Validation of the microarray results of differential gene expression in HepG2 cells over-expressing *Sox17*-IRES-GFP and in *Sox17*^{-/-} embryos. Real-time quantitative RT-PCR analysis was performed on cDNA generated from: (A) RNA samples of HepG2 cells transfected with GFP only or *Sox17*-IRES-GFP expression vector; and (B) RNA samples of 6-8 somite *Sox17*^{+/+} and *Sox17*^{+/-} or *Sox17*^{-/-} embryos. Expression levels are shown relative to GFP only (A) or *Sox17*^{+/+}/*Sox17*^{+/-} embryos (B). POLR2A was the reference gene for (A) and *Gapd* was the reference gene for (B). Error bars indicate standard error of the mean. $N=4$ for *Sox17*^{+/+}, *Sox17*^{+/-} and $N=3$ for each all others. **Significant difference at $p<0.01$ or * at $p<0.05$ by two-tailed t-test from GFP-only expressing cells or *Sox17*^{+/+}/*Sox17*^{+/-} embryos.

were no genes common to both sets. Most likely, this reflects the different effects of *Sox17* over-expression in the two cell models. In mouse ES cells, *Sox17* appears to direct extraembryonic endoderm differentiation, whereas in human ES cells, *Sox17* overexpression results in up-regulation of genes that are characteristic of definitive endoderm. Among the genes of the embryo-mouse ES and embryo-human ES sets, *Adm*, *Fgb*, *Folr1* were also detected by ChIP from mouse ES or extraembryonic endoderm stem (XEN) cells (Niakan *et al.*, 2010) suggesting that *Sox17* may act directly on these genes.

Comparison of genes with changed expression levels in *Sox17* mutant embryos and *Sox17*-overexpressing human ES cells (Seguin *et al.*, 2008) revealed a set of genes that encode proteins that affect diverse processes including actin cytoskeleton organization (CDC42, RDX), proliferation and differentiation (FST, GNG4, SERPINE2) and responses to signalling (ACSL4, ADM, FST, GNG4, STC2; Supplementary Table S4). The best candidates *Sox17*-regulated genes in the definitive endoderm are those that are differentially expressed in *Sox17*-null embryos and in the over-expression HepG2 cells and to which *Sox17* binds. Comparison with published data (Niakan *et al.*, 2010), (Supplementary Table S5) revealed a set of six genes that fulfilled this criterion amongst those identified by ChIP of ES cells only (*Fgb*, *Slco4c1*), XEN cells only (*Mbl2*) or both ES and XEN cells (*Adm*, *Nrp1*, *Fbm4b*).

***Sox17*-deficient cells contribute less efficiently to gut endoderm**

Our microarray studies revealed a higher representation of transcripts of genes that are characteristic of extraembryonic visceral endoderm in the *Sox17*-null embryo, suggesting a preponderance of this type of endoderm in the early-somite stage *Sox17*-null embryo. This is consistent with findings that *Sox17*^{-/-} embryo is depleted of definitive endoderm, which is partly replaced by cells that display properties of extraembryonic visceral endoderm (Kanai-Azuma *et al.*, 2002). To test directly whether *Sox17*-deficient progenitor cells are compromised in their ability to populate the gut endoderm, the fates of cells of the anterior segment of the primitive streak (APS) of gastrula-stage embryos, where the progenitors of definitive endoderm are localised (Kinder *et al.*, 2000), were analysed by orthotopic transplantation to wild type recipient embryos (Fig. 3 A, H). Fluorescence imaging of the recipient embryos cultured for 29 hours revealed a widespread distribution of the EGFP and *lacZ*-expressing APS-derived cells from *Sox17*^{+/+} and *Sox17*^{+/-} donor in the host embryos (Fig. 3 B, B', E, E') and to tissues derived from different germ layers (Fig. 3 C, D, F, G). Although a similar average number of cells were generated by grafts of *Sox17*^{-/-} APS cells, a smaller proportion of graft-derived cells were found in the gut endoderm and the presomitic mesoderm (Fig. 3 I-N, Table 5). Instead, *Sox17*^{-/-} cells were found more frequently in the ectoderm of the posterior neural tube and the extraembryonic mesoderm (Table 5). *Sox17*^{-/-} APS cells therefore are impaired in their ability to contribute to gut endoderm.

Morphogenesis of the anterior intestinal portal is disrupted in Sox17^{-/-} embryos

Cells of the anterior definitive endoderm cells are fated to become the endoderm of the anterior intestinal portal (the foregut

invagination) that forms the upper digestive tract and associated organs (Tremblay and Zaret, 2005). In wild type and heterozygous mutant embryos, formation of the anterior intestinal portal proceeded first by invagination, followed by anterior extension and narrowing of the foregut pocket (Fig. 4 A, B, 0, 5 and 23 hrs)

TABLE 2

SIGNIFICANTLY OVER-REPRESENTED GENE ONTOLOGY BIOLOGICAL FUNCTION TERMS AMONGST GENES THAT WERE UP- OR DOWNREGULATED IN SOX17^{-/-} EMBRYOS

(A) Upregulated in Sox17^{-/-} embryos

P value	Odds Ratio	Expected Count	Count	Term
0.000	69.201	0	3	cellular zinc ion homeostasis
0.000	3.533	5	15	response to wounding
0.000	51.896	0	3	lipoprotein transport
0.000	6.084	1	7	regulation of body fluid levels
0.000	5.866	1	7	lipid transport
0.001	68.800	0	2	regulation of cholesterol absorption
0.001	68.800	0	2	cholesterol efflux
0.002	45.863	0	2	negative regulation of blood coagulation
0.002	45.863	0	2	lipid digestion
0.002	13.826	0	3	regeneration
0.002	13.826	0	3	vasodilation
0.003	34.394	0	2	lipoprotein catabolic process
0.003	11.519	0	3	cholesterol homeostasis
0.003	5.358	1	5	cholesterol metabolic process
0.004	10.912	0	3	blastocyst formation
0.004	10.912	0	3	zinc ion transport
0.005	4.796	1	5	negative regulation of multicellular organismal process
0.005	23.371	0	2	lipoprotein metabolic process
0.006	9.011	0	3	Gastrulation in deuterostomes
0.006	9.011	0	3	iron ion transport
0.006	9.011	0	3	lipid homeostasis
0.006	3.428	2	7	steroid metabolic process
0.006	3.879	2	6	di-, tri-valent inorganic cation homeostasis
0.007	3.808	2	6	cellular cation homeostasis
0.007	2.235	7	14	reproduction
0.007	19.649	0	2	axon regeneration
0.007	5.660	1	4	lipoprotein biosynthetic process

(B) Downregulated in Sox17^{-/-} embryos

P value	Odds Ratio	Expected Count	Count	Term
0.000	38.762	0	6	cholesterol biosynthetic process
0.000	6.713	2	14	cellular alcohol metabolic process
0.000	12.039	1	7	steroid biosynthetic process
0.000	3.184	5	15	oxidation reduction
0.001	11.490	0	4	cholesterol metabolic process
0.001	18.822	0	3	cholesterol homeostasis
0.002	14.724	0	3	lipid homeostasis
0.004	6.822	1	4	steroid metabolic process
0.004	4.325	1	6	lipid biosynthetic process
0.006	20.544	0	2	isoprenoid biosynthetic process
0.009	3.580	2	6	heart development
0.009	Infinite	0	1	S-adenosylmethionine biosynthetic process
0.009	Infinite	0	1	isoprenoid catabolic process
0.009	Infinite	0	1	low-density lipoprotein receptor metabolic process
0.009	Infinite	0	1	regulation of low-density lipoprotein receptor catabolic process
0.009	Infinite	0	1	muscle cell fate specification
0.009	Infinite	0	1	arachidonic acid secretion

TABLE 3

**TOP 50 GENES THAT WERE SIGNIFICANTLY DOWN-REGULATED IN *SOX17*^{-/-} EMBRYOS
ORDERED BY LOG₂ FOLD CHANGE IN EXPRESSION**

Symbol	Description	GenBank	GO Biological functions summary	Probe	Log ₂ FoldChange
<i>Sox17</i>	SRY-box containing gene 17	AK004781	angiogenesis; vasculogenesis; negative regulation of Wnt receptor signaling pathway	1429177_x_at 1421657_a_at	-4.88 -1.24
<i>Eif2s3y</i>	eukaryotic translation initiation factor 2, subunit 3, structural gene Y-linked	NM_012011	translation	1417210_at	-3.7
<i>Ddx3y</i>	DEAD (Asp-Glu-Ala-Asp) box polypeptide 3, Y-linked	AA210261		1426438_at	-3.65
<i>Prl8a2</i>	prolactin family 8, subfamily a, member 2	NM_010088	response to hypoxia	1448608_at	-2.72
<i>Jarid1d</i>	jumonji, AT rich interactive domain 1D (Rbp2 like)	AF127244	chromatin modification	1424903_at	-2.39
<i>Uty</i>	ubiquitously transcribed tetratricopeptide repeat gene, Y chromosome	BB742957	chromatin modification	1426598_at 1422247_a_at	-2.33 -1.35
<i>Mid1</i>	midline 1	BG073178	negative regulation of microtubule depolymerization	1438239_at	-1.99
<i>Nepr</i>	nephrocan	NM_025684	negative regulation of transforming growth factor beta receptor signaling pathway	1419065_at	-1.92
<i>Myl3</i>	myosin, light polypeptide 3	X67685		1427768_s_at	-1.88
<i>Sec23ip</i>	Sec23 interacting protein	BE685845		1439882_at	-1.8
<i>Actc1</i>	actin, alpha, cardiac muscle 1	NM_009608	apoptosis; muscle thin filament assembly; cardiac muscle tissue morphogenesis	1415927_at	-1.58
<i>A230083H22Rik</i>	RIKEN cDNA A230083H22 gene	AK018172		1432198_at	-1.53
<i>Ctla2a</i>	cytotoxic T lymphocyte-associated protein 2 alpha	NM_007796		1416811_s_at	-1.51 -1.05
<i>Adm</i>	adrenomedullin	AV378441	heart development; cell proliferation	1447839_x_at	-1.43
<i>Myl7</i>	myosin, light polypeptide 7, regulatory	NM_022879		1449071_at	-1.38
<i>3110040N11Rik</i>	RIKEN cDNA 3110040N11 gene	AK019261		1450972_at	-1.33
<i>Myl4</i>	myosin, light polypeptide 4	NM_010858		1422580_at	-1.33
<i>Rpl17</i>	ribosomal protein L17	BF453369	translation	1453752_at	-1.28
<i>Sc4mol</i>	sterol-C4-methyl oxidase-like	AK005441	steroid biosynthesis	1423078_a_at	-1.24
<i>6720422M22Rik</i>	RIKEN cDNA 6720422M22 gene	BB051012		1437798_at	-1.21
<i>Strp1</i>	secreted frizzled-related protein 1	BI658627	somitogenesis; Wnt receptor signaling pathway	1460187_at 1428136_at	-1.21 -0.7
<i>Fosb</i>	FBJ osteosarcoma oncogene B	NM_008036	Regulation of transcription	1422134_at	-1.17
<i>Idi1</i>	isopentenyl-diphosphate delta isomerase	BC004801	cholesterol biosynthesis	1423804_a_at 1451122_at	-1.09 -1.02
<i>Insig1</i>	insulin induced gene 1	BB005488	cholesterol metabolic process	1454671_at	-1.08
<i>Ldlr</i>	low density lipoprotein receptor	AF425607	cholesterol metabolic process;	1421821_at 1459403_at	-1.07 -0.63
<i>A2m</i>	alpha-2-macroglobulin	BB185854	pregnancy	1434719_at	-1.06
<i>Ctla2b</i>	cytotoxic T lymphocyte-associated protein 2 beta	BG064656		1452352_at	-1.06
<i>Cryab</i>	crystallin, alpha B	AV016515	muscle development; eye development	1434369_a_at	-1.04
<i>Myocd</i>	myocardin	AF384055	vasculogenesis; heart development; regulation of myoblast differentiation	1425978_at	-1.02
<i>Rassf5</i>	Ras association (RalGDS/AF-6) domain family member 5	NM_018750	apoptosis; cell cycle regulation	1422638_s_at	-1.01
<i>Trub1</i>	TruB pseudouridine (psi) synthase homolog 1 (E. coli)	AK011362	tRNA processing	1428281_at	-1
<i>Myl9</i>	myosin, light polypeptide 9, regulatory	AK007972		1452670_at	-0.98
<i>Shisa4</i>	shisa homolog 4 (<i>Xenopus laevis</i>)	BF468228		1438426_at	-0.98
<i>Cyp51</i>	cytochrome P450, family 51	NM_020010	cholesterol biosynthetic process	1422533_at 1450646_at	-0.96 -0.85
<i>Fabp4</i>	fatty acid binding protein 4, adipocyte	BC002148	cholesterol homeostasis	1451263_a_at 1417023_a_at	-0.96 -0.54
<i>Tnni1</i>	troponin I, skeletal, slow 1	NM_021467	ventricular cardiac muscle morphogenesis	1450813_a_at	-0.95
<i>ENSMUSG00000068790</i>	predicted gene, ENSMUSG00000068790	BM195235		1452731_x_at 1428301_at	-0.9 -0.83
<i>Gprc5b</i>	G protein-coupled receptor, family C, group 5, member B	BC020004	signal transduction	1451411_at	-0.9
<i>Idh1</i>	isocitrate dehydrogenase 1 (NADP+), soluble	NM_010497	glyoxylate cycle tricarboxylic acid cycle	1422433_s_at	-0.89
<i>Slc30a1</i>	solute carrier family 30 (zinc transporter), member 1	BE685959	zinc ion transport	1436164_at	-0.89
		AU042527		1447096_at	-0.89
<i>2610528B01Rik</i>	RIKEN cDNA 2610528B01 gene	AK012160		1429232_at	-0.88
<i>Scd1</i>	stearoyl-Coenzyme A desaturase 1	NM_009127	fatty acid biosynthetic process	1415964_at 1415965_at	-0.87 -0.48
		BI134319		1442257_at	-0.86
		BB235490		1441050_at	-0.83
<i>D14Erd449e</i>	DNA segment, Chr 14, ERATO Doi 449, expressed	BG072279		1428738_a_at	-0.83
<i>Rabif</i>	RAB interacting factor	AI482417	protein transport	1457969_at	-0.83
<i>LOC100040592</i>	similar to Hmgcs1 protein	BB705380		1433445_x_at 1433444_at 1433443_a_at	-0.83 -0.78 -0.66
<i>Duxbl</i>	double homeobox B-like	AV321065		1445710_x_x_at	-0.82
<i>Mef2c</i>	myocyte enhancer factor 2C	BB280300	blood vessel development; osteoblast differentiation; heart development transcription activator activity smooth muscle cell differentiation	1424852_at 1451507_at 1421028_a_at	-0.79 -0.71 -0.44

which led to the medial convergence and longitudinal extension of the lateral endoderm cell populations. Subsequently, these cells were found primarily in the lateral wall of the anterior intestinal portal (Fig 4 A, B left and right). A unique morphogenetic feature of the anterior intestinal portal is the asymmetric displacement of the endoderm cells in the plane of the epithelium (Franklin *et al.*, 2008). This was observed in the wild type and *Sox17*^{-/-} mutant embryos in which cells from the left side of the portal were found on the floor (ventral wall) whereas those from the right side were found in the roof (dorsal wall) (Fig. 4 A, B left and right; Table 6). The *Sox17*^{-/-} embryos characteristically showed flattened neural folds and a shallow depression in the prospective foregut area at the initial phase of portal formation (Fig. 4C, 0hr). The lateral regions of the portal were wider apart, resulting in a short portal and with a broad entrance (Fig. 4C, 23hr). In addition, the lateral

cell populations were mostly confined to the lateral walls and infrequently found asymmetrically on the dorsal or ventral wall of the portal (Fig. 4C, left and right; Table 6). Morphogenesis of the anterior intestinal portal in *Sox17*^{-/-} embryos is therefore defective with less convergence-extension and planar rotation of the anterior definitive endoderm.

Loss of Sox17 leads to down-regulation of cardiac genes and abnormal morphogenesis of the heart tube

Among the genes that were down-regulated in *Sox17*-null embryos, there was an over-representation of genes associated with heart development (Table 2). In particular, *Myocd*, a transcription factor expressed in the heart tube (Wang *et al.*, 2001), is down-regulated significantly in early-somite stage *Sox17*^{-/-} embryos (Fig. 5A). Whole mount *in situ* hybridisation revealed that

TABLE 4

GENES THAT WERE SIGNIFICANTLY DIFFERENTIALLY EXPRESSED IN SOX17^{-/-} EMBRYOS COMPARED TO WILD TYPE EMBRYOS AND IN SOX17-IRES-GFP TRANSFECTED HEPG2 CELLS COMPARED WITH GFP TRANSFECTED HEPG2 CELLS

Symbol (mouse/human)	Name	Affymetrix probe ID	<i>Sox17</i> ^{-/-} mouse embryos (Log ₂ fold change)	Illumina target ID	<i>Sox17</i> -IRES-GFP HepG2 cells (Log ₂ fold change)
<i>Abcc2/ABCC2*</i>	ATP-binding cassette, sub-family C (CFTR/MRP), member 2	1450109_s_at	1	ILMN_9691	-0.6043
<i>Adm/ADM*</i>	adrenomedullin	1447839_x_at	-1.43	ILMN_29514	1.009719
<i>Afp/AFP*</i>	alpha fetoprotein	1416646_at 1436879_x_at 1416645_a_at	3.14 2.96 2.63	ILMN_19039	-0.20155
<i>Agpat5/AGPAT5</i>	1-acylglycerol-3-phosphate O-acyltransferase 5 (lysophosphatidic acid acyltransferase, epsilon)	1453257_at	-0.44	ILMN_9737	-0.72299
<i>Apoc2/APOC2*</i>	apolipoprotein C-II	1418069_at	2.35	ILMN_26723	-0.35034
<i>Bmp4/BMP4</i>	bone morphogenetic protein 4	1422912_at	0.59	ILMN_27187	-0.46856
<i>Cryab/CRYAB*</i>	crystallin, alpha B	1434369_a_at	-1.04	ILMN_6827	1.910507
<i>Eno3/ENO3</i>	enolase 3, beta muscle	1417951_at	-0.46	ILMN_16651	-0.36165
<i>Eps8/EPS8</i>	epidermal growth factor receptor pathway substrate 8	1422823_at 1422824_s_at	0.52 0.79	ILMN_17717	0.556275
<i>F10/F10*</i>	coagulation factor X	1449305_at	0.65	ILMN_138620	-0.71965
<i>Fga/FGA*</i>	fibrinogen, alpha polypeptide	1424279_at	1.31	ILMN_11182	-0.47527
<i>Fgb/FGB*</i>	fibrinogen, B beta polypeptide	1428079_at	1.53	ILMN_13882	-0.32784
<i>Fgg/FGG*</i>	fibrinogen, gamma polypeptide	1416025_at	1.99	ILMN_26176	-0.20352
<i>Foxa3/FOXA3*</i>	forkhead box A3	1431900_a_at	0.42	ILMN_22171	-1.193
<i>Ldlr/LDLR*</i>	low density lipoprotein receptor	1421821_at 1459403_at	-1.07 -0.63	ILMN_10126	0.282144
<i>Lgmn/LGMN</i>	legumain	1448883_at	1.16	ILMN_20242	0.620039
<i>Lmna/LMNA*</i>	lamin A	1421654_a_at	0.39	ILMN_12442	-0.38805
<i>Lrp8/LRP8</i>	low density lipoprotein receptor-related protein 8, apolipoprotein e receptor	1440882_at 1442347_at	-0.54 -0.46	ILMN_2030	-0.42656
<i>Mbl2/MBL2*</i>	mannose-binding lectin (protein C) 2	1418787_at	1.38	ILMN_6942	-1.53863
<i>Nrp1/NRP1*</i>	neuropilin 1	1457198_at 1448944_at	0.53 0.54	ILMN_17483	-1.10432
<i>Nsdhl/NSDHL</i>	NAD(P) dependent steroid dehydrogenase-like	1416222_at	-0.63	ILMN_13529	-0.42729
<i>Rbm4b/RBM4B*</i>	RNA binding motif protein 4B	1430032_at	0.41	ILMN_29996	-0.28508
<i>Reep6/REEP6*</i>	receptor accessory protein 6	1430128_a_at	1.24	ILMN_15192	-0.64812
<i>Sc4mol/SC4MOL</i>	sterol-C4-methyl oxidase-like	1423078_a_at	-1.24	ILMN_2770	-0.41272
<i>Sec23ip/SEC23IP</i>	Sec23 interacting protein	1439882_at	-1.8	ILMN_7522	-0.4024
<i>Slc16a10/SLC16A10</i>	solute carrier family 16 (monocarboxylic acid transporters), member 10	1434592_at	0.31	ILMN_10556	0.425748
<i>Slc2a2/SLC2A2*</i>	solute carrier family 2 (facilitated glucose transporter), member 2	1449067_at	2.83	ILMN_28285	-1.20611
<i>Slc30a1/SLC30A1*</i>	solute carrier family 30 (zinc transporter), member 1	1436164_at	-0.89	ILMN_5933	0.439786
<i>Slc39a5/SLC39A5*</i>	solute carrier family 39 (metal ion transporter), member 5	1429523_a_at	0.78	ILMN_14803	-0.64595
<i>Slco4c1/SLCO4C1*</i>	solute carrier organic anion transporter family, member 4C1	1460616_at 1437870_at	1.29 1.27	ILMN_3183	-0.37455
<i>Slu7/SLU7</i>	SLU7 splicing factor homolog (S. cerevisiae)	1425488_at	0.46	ILMN_12938	0.927178
<i>Zcchc14/ZCCHC14*</i>	zinc finger, CCHC domain containing 14	1418170_a_at	0.6	ILMN_138708	-0.63098

Asterisk indicates genes that were up-regulated when *Sox17* was over-expressed and down-regulated in its absence, and vice versa

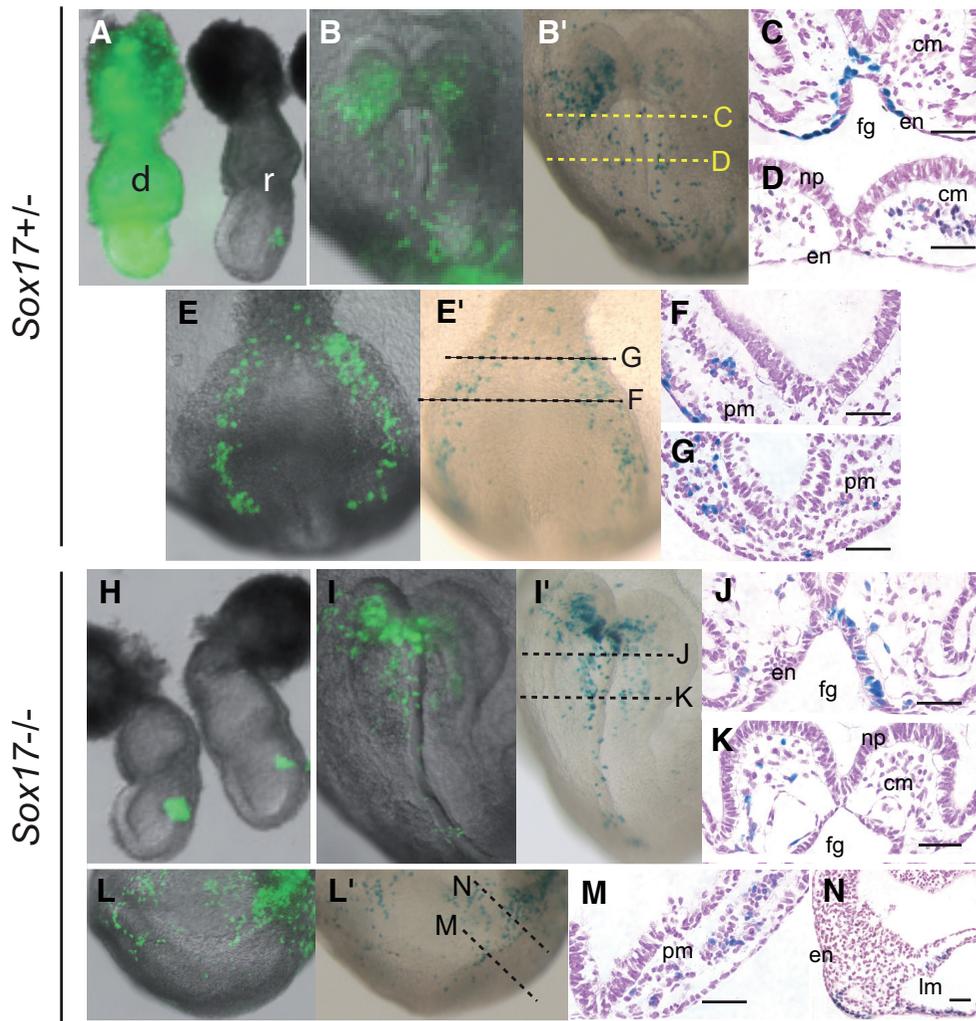


Fig. 3. Contribution of graft-derived cells in the recipient embryos. Transplantation of cells from the EGFP-expressing (A) Sox17^{+/-} donor (d) embryo and (H) Sox17^{-/-} donor embryos to the anterior region of the primitive streak of the wild type mid-streak stage recipient (r) embryo followed by visualization of the graft-derived cells by fluorescence imaging (for EGFP: B, E, I, L) and X-gal staining (for lacZ: B', E', I', L') in (B, B', I, I') the anterior and (E, E', L, L') posterior regions of the recipient embryo after 24 hours of in vitro development. Histology was performed on recipient embryos (in the same series of the transplantation but not necessarily of the specimens shown in B', E', I' and L'). For orientation, planes of the histological sections are shown in B', E', I' and L'. Graft-derived (lacZ-expressing, visualized by blue X-gal staining reaction) (C, D, F, G) Sox17^{+/-} cells and (J, K, M, N) Sox17^{-/-} cells populate the endoderm (en) of the foregut (fg), and the cranial mesoderm (cm) underneath the neural plate (np) in the anterior region; and the presomitic mesoderm (pm) and lateral plate mesoderm (lm) in the posterior region of the embryo. Scale bars, 100 μ m.

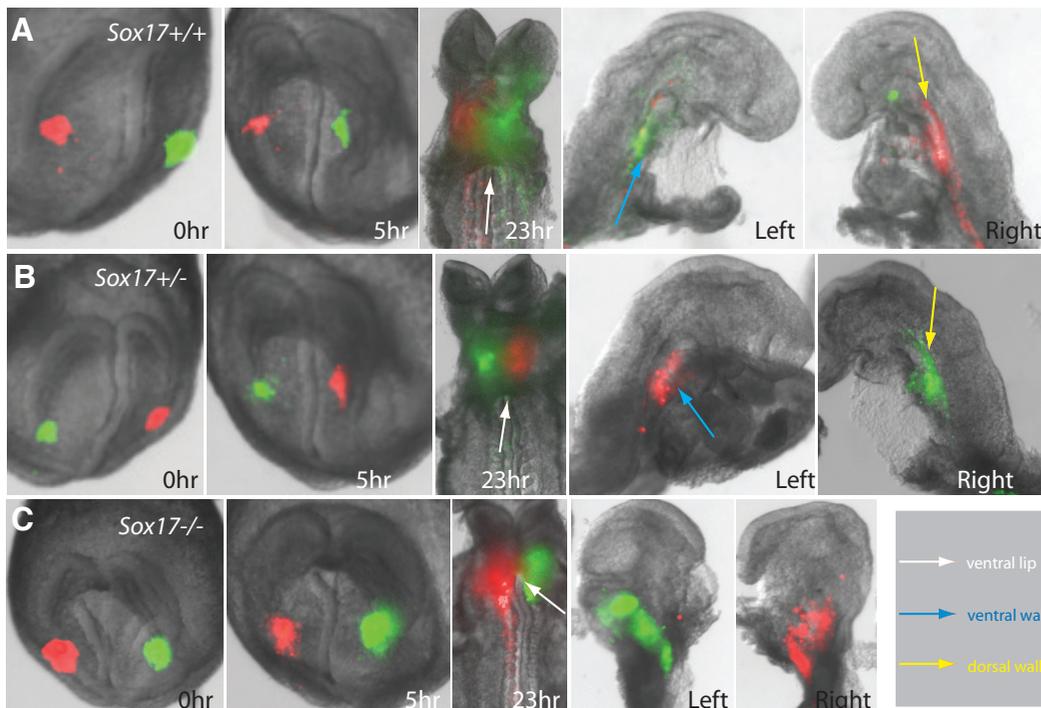


Fig. 4. Formation of the anterior intestinal portal. Distribution of contra-lateral cell population of the anterior definitive endoderm in the anterior intestinal portal at different time points (0, 5 and 23 hours) during its formation in (A) Sox17^{+/+}, (B) Sox17^{+/-} and (C) Sox17^{-/-} embryos. Cells were labeled with DiO (green) and Dil (red) carbocyanine dye. Embryos were imaged intact during in vitro culture and were imaged after bisected into left and right halves at the end of culture to reveal the localization of the labelled cells. Arrows (see legend for colour code) point to the ventral lip, ventral wall and dorsal wall of the anterior intestinal portal.

TABLE 5

CONTRIBUTION OF THE ANTERIOR PRIMITIVE STREAK CELLS TO THE GERM LAYER DERIVATIVES DURING GASTRULATION AND EARLY ORGANOGENESIS

Donor genotypes (no. of donor embryos)	<i>Sox17^{+/+}</i> or <i>Sox17^{+/-}</i> (n=15)		<i>Sox17^{-/-}</i> (n=15)	
	Cell count	% of total	Cell count	% of total
Cellular contribution in recipient embryos to:				
Endoderm*	290	8.1%	163	4.1%
Posterior neurectoderm*	17	0.4%	51	1.3%
Cranial mesoderm	361	10.1%	545	13.7%
Paraxial mesoderm	460	12.8%	444	11.2%
Presomitic mesoderm*	777	21.7%	511	12.9%
Lateral mesoderm	624	17.4%	652	16.4%
Heart mesoderm	312	8.7%	392	9.9%
Extraembryonic mesoderm*	716	20.8%	1207	30.4%
TOTAL	3587		3965	
Average	239.1		264.3	

*Significant difference between genotypes by Chi-squared test at $P < 0.05$.

markers of the primary heart field (*Myocd* and *Tbx5*) and the secondary heart field (*Isl1*) are expressed in the cardiogenic mesoderm of *Sox17^{-/-}* embryos at the early head-fold stage (Fig. 5 B, D, F), suggesting that the heart precursors were present in the embryo and that the initial distribution of the cardiogenic mesoderm to the heart field is not affected by the loss of *Sox17* in the endoderm. However, *Myocd* and *Tbx5* expression was absent from the midline tissues in the primary heart field of the *Sox17^{-/-}* embryos (3/4), resulting in a hiatus in the cardiac crescent where the lateral population of the heart progenitors failed to unite medially (Fig. 5 C-G).

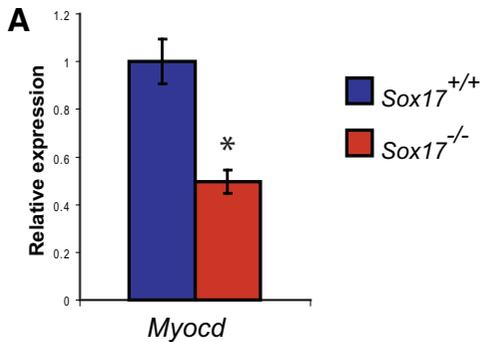


Fig. 5. Expression of heart genes in *Sox17^{-/-}* early head fold stage (pre-somite, E8.25) embryos revealed by whole mount *in situ* hybridization. (A) Down-regulation of *Myocd* in *Sox17^{-/-}* embryos revealed by real time quantitative RT-PCR. Error bars indicate standard error of the mean. $N=3$ for each assay (* $P < 0.05$ by *t* test). Whole mount *in situ* hybridization of embryos showing the expression pattern of (B,C) *Myocd* and (D,E) *Tbx5* in the cardiac crescent of (B) *Sox17^{+/+}*, (D) *Sox17^{+/-}* and (C,E) *Sox17^{-/-}* embryos. In *Sox17*-null embryos, expression of *Tbx5* and *Myocd* is absent in the midline region of the cardiac crescent (arrows). (F, G) Expression of *Isl1* is intact in the tissues in the secondary heart field of the *Sox17^{-/-}* embryo.

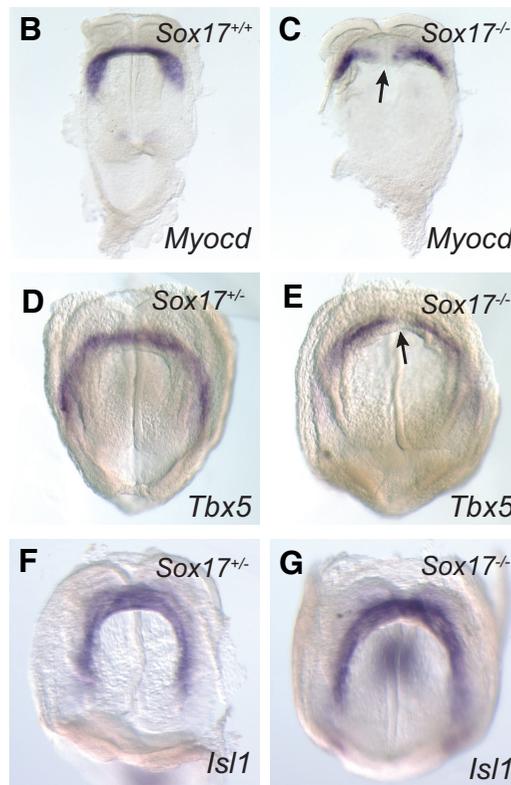


TABLE 6

PATTERNS OF DISTRIBUTION OF CELLS ORIGINATING FROM THE LATERAL REGIONS OF THE ANTERIOR DEFINITIVE ENDODERM IN THE ANTERIOR INTESTINAL PORTAL

Sox17 Genotype of embryo +/+ or +/-	Distribution of labeled cells in the anterior intestinal portal*						
	No of embryos	Left side			Right side		
		Ventral	Lateral	Dorsal	Ventral	Lateral	Dorsal
+/+ or +/-	2	■	■			■	
	6	■	■			■	
	1	■	■		■	■	
	1*	■	■		■	■	
Total	10	4	10	1	8	9	
-/-	1	■			■	■	
	1*		■		■	■	
	2		■		■	■	
	1		■		■	■	
Total	5	1	4	1	5	1	

■ DiO labeled cells; ■ DiI labeled cells. * Three out of five null-mutant embryos showed localization of the labeled cells only to the lateral wall of the anterior intestinal portal. Only 1/5 null-mutant as contrasted to 9/10 wild type or heterozygous embryos showed proper asymmetrical distribution of labeled cells: significant difference at $P < 0.01$ by χ^2 test. * One each of null-mutant and wild type/heterozygous showed atypical localization of labeled cells from right side to the ventral wall (floor) instead of dorsal wall (roof).

Examination the heart morphology of E8.5 *Sox17*-null embryos revealed abnormal heart tube phenotypes ranging from single heart tube that had not initiated looping (1/5) (Fig. 6B), incomplete union of the heart tube (1/5) (Fig. 6C) and the formation of two completely separated heart tubes (cardia bifida, 3/5) (Fig. 6D). The heart of *Sox17^{+/-}* embryos (not shown) was indistinguishable from that of the wild type (Fig. 6A). In the *Sox17* mutant embryos the anterior intestinal portal was irregularly shaped and the floor crinkled extensively (Fig. 6 F, G, J), in contrast to the uniform contour in the wild type embryo (Fig. 6E). In *Sox17^{-/-}* mutant embryo with incomplete fusion or separated heart tubes, the more caudal segment of the portal was open, resulting a wide ventral gap (Fig. 6 K, L), which in the wild type embryo was closed by this stage (Fig. 6I). The defects of heart tube formation in the *Sox17*-null mutant embryos are therefore accompanied by the abnormal morphogenesis of the anterior intestinal portal in addition to the defects in cardiac mesoderm differentiation.

Discussion

Sox17 downstream genes and molecular pathways

Microarray analysis of wild-type and *Sox17^{-/-}* embryos revealed changes in expression of genes that are associated primarily with the definitive endoderm (*Nepn*, *Foxa3*, and *Hnf4a*), as well as those that are expressed in both the extraembryonic visceral endoderm and the definitive endoderm (e.g.: Apolipoprotein genes, *Ttr*, *Afp*). Furthermore, there are

changes in the expression of genes (e.g. heart development genes, cholesterol pathway genes) that are not associated with the endoderm.

To identify transcripts that are candidates for direct dependence on *Sox17* expression, we over-expressed *Sox17* in HepG2 cells. As the liver is derived from the definitive endoderm, we reasoned that HepG2 cells would provide a cellular environment comparable to the gut endoderm of the mouse embryo, where *Sox17* is normally expressed. Comparison of the results of the two array screens revealed a set of 21 genes whose changes in expression are consistent with either negative or positive regulation by *Sox17*. Analysis of publicly available gene expression data (<http://biogps.gnf.org>) revealed that 9 of these genes were specifically or predominantly expressed in adult and/or fetal liver (*Afp*, *Apoc2*, *F10*, *Fga*, *Fgb*, *Fgg*, *Mbl2*, *Reep6* and *Slc2a2*). All of these genes were found in our microarray study to be downregulated in the *Sox17*-overexpressing HepG2 cells and upregulated in *Sox17*^{-/-} embryos. These data, along with the significant overlap with *Hnf4a* regulated genes, suggest the possibility that *Sox17* plays a role in repression of liver-specific gene expression. This is consistent with a role for *Sox17* in maintaining definitive endoderm in the progenitor state, and the reported requirement for *Sox17* in the delineation of specific foregut endoderm-derived lineages, including the gallbladder and bile duct (Spence *et al.*, 2009; Uemura *et al.*, 2010).

By comparing our mouse embryo and HepG2 microarray expression data with ChIP data from mouse ES or XEN cells, we have identified a shortlist of six genes (*Adm*, *Fgb*, *Mbl2*, *Nrp1*, *Rbm4b*, *Slco4c1*) for which *Sox17* influences their expression

and there is evidence for binding of *Sox17* protein to regulatory regions. These genes are diverse in function and expression and represent strong candidates for *Sox17* transcriptional target genes.

Loss of *Sox17* function impairs the allocation of endoderm cells

Although the *Sox17*-mutant phenotype indicates a role in the development of the definitive endoderm, several lines of evidence also point to a requirement of *Sox17* in the differentiation of extraembryonic endoderm. In mouse ESC, enforced expression of *Sox17* promotes the differentiation of extraembryonic endoderm, whereas *Sox17*-null ES cells maintain the expression of pluripotency genes and fail to differentiate into extraembryonic endoderm. Stem cells with extraembryonic endoderm properties (XEN cells) cannot be generated from *Sox17*-null embryos (Niakan *et al.*, 2010). In contrast, in human ESC, progenitors with stable characteristics of definitive endoderm can be generated by constitutive expression of *Sox17*, whereas expression of another Group F Sox factor, *Sox7*, produces cells with more restricted potential reminiscent of the extraembryonic endoderm (Seguin *et al.*, 2008). Transcriptional profiling of *SOX17* and *SOX7* overexpressing human ESCs revealed that the sets of genes that were specifically upregulated by either *SOX17* or *SOX7* were far greater than the set of genes that was upregulated by both, indicating unique, tissue-specific roles for the two transcription factors. In agreement with these findings, our microarray data do not support a role for *Sox17* in promoting extraembryonic endoderm differentiation in mouse embryos. Instead, our data show up-regulation of extraembryonic gene expression in *Sox17*-null embryos, consistent with apparent replacement of definitive endoderm with visceral endoderm-like cells (Kanai-Azuma *et al.*, 2002).

Sox17-null ES cells are unable to contribute significantly to the definitive endoderm of the gut in the presence of wild type cells in the chimera (Kanai-Azuma *et al.*, 2002). Cell transplantation experiments performed in the present study further show that cells of the *Sox17*^{-/-} anterior primitive streak, where the endoderm progenitors are localised (Lawson *et al.*, 1991), are less efficient in contributing descendants to the host embryonic gut. The chimera and transplantation studies seem to show that, in addition to the maintenance of the lineage, *Sox17* function may also play a role in, but not be essential for, the allocation of progenitors of the endoderm lineage. The elevated level of expression of many visceral endoderm genes in the *Sox17*-null embryos might result from the increased presence of extraembryonic endoderm which takes

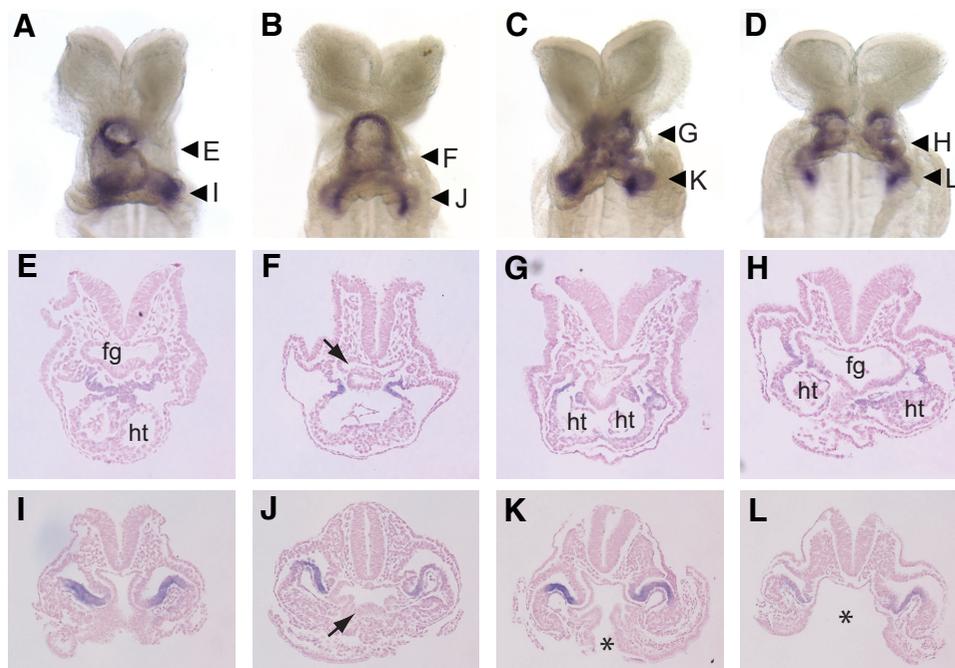


Fig. 6. Heart phenotype of E8.5 (7-8 somites) *Sox17*^{-/-} embryos. Wild type embryo shows a looping heart tube (A), whereas *Sox17*^{-/-} embryos display (B) un-looped heart tube, (C) partially united and (D) separate heart tube. Heart tissues are highlighted by *Myocd* expression by whole mount in situ hybridization (A-D) and in histological preparations (E-L, planes of sectioning indicated by arrowheads in A-D). The foregut of *Sox17*^{-/-} embryo has an abnormal shape (F, arrow) and an expanded corrugated floor (J, arrow) and opens ventrally (K, L: asterisk) in the posterior region of the foregut of embryos with bifid heart tube. ht: heart tube, fg: foregut.

the place of the gut endoderm cells. This may be the result of not only the prevalence of apoptosis and reduced proliferative activity but also of the compromised ability of the mutant cells to contribute to the definitive endoderm.

Morphogenesis of the embryonic foregut and the heart tube is dependent on Sox17 function

Development of the embryonic foregut is initiated by the formation of the anterior intestinal portal beginning with the invagination of the anterior definitive endoderm underneath the head folds, accompanied by the longitudinal extension of the medial regions and the convergence of the lateral parts of the endoderm layer (Franklin *et al.*, 2008). An intriguing feature of the morphogenesis of the portal is the pattern of movement of the contra-lateral endoderm populations of the prospective foregut. While these populations are undergoing longitudinal extension, they are also displaced asymmetrically along the wall of the portal: cells on the right populate the right and dorsal (roof) sides while cells on the left populate the left and ventral (floor) of the portal. The asymmetrical distribution of contra-lateral endoderm cells is accomplished by the directional relocation of cells within the epithelium that lines the portal and not due to the "rotation" of the foregut. However, the asymmetry of cell displacement corresponds with the laterality of subsequent rotation (left to ventral and right to dorsal) of the lower foregut. *Sox17*-null embryos form an anterior intestinal portal which is shorter and wider than the wild type counterpart, with excess folding of the floor and a split posterior ventral lip of the portal. Tracking the morphogenetic movement of anterior definitive endoderm in the null mutant embryo further revealed a much reduced longitudinal extension and medial convergence of the lateral endoderm, culminating in the distribution of these cells to wide lateral domain in the portal and very restricted rotational displacement.

Heart defects were not reported in the initial analysis of the *Sox17*-null phenotype (Kanai-Azuma *et al.*, 2002). A subsequent study revealed that the heart of the mutant embryo displayed defects in the looping of the heart tube (Sakamoto *et al.*, 2007). Our study further shows that *Sox17*-null mutant embryos display additional and more severe defects of the heart tube: lack of union of the cardiac precursors from the contra-lateral part of the cardiac crescent and bifid heart tubes. One possible cause of the difference in the manifestation of heart defects in our study and that of Sakamoto *et al.* (2007) is the difference in background. Whereas in the latter study, the mice were maintained on a mixed 129/Sv X C57BL/6 background, our mice have been maintained on a predominantly 129/Sv background. A precedent for the effect of genetic background on the heart defects is seen in *Fibronectin 1* mutants which display less severe heart phenotype with an increasing contribution of C57/BL6 over the 129/Sv background (George *et al.*, 1997).

The formation of a wide portal with a ventral hiatus and the lack of asymmetrical displacement of the cells in the portal could be the morphogenetic factors underlying the incomplete fusion and lack of looping of the heart tube. Similar cardia bifida phenotypes found in other mouse and Zebrafish mutants are also accompanied by defects in foregut development, such as absence or reduction of foregut endoderm and incomplete closure of the foregut pocket (Alexander *et al.*, 1999; Kuo *et al.*, 1997; Li *et al.*, 2004; Molkenstein *et al.*, 1997; Reiter *et al.*, 1999; Roebroek *et al.*,

1998). It is not known whether the heart tube defects are due to the abnormal morphogenesis of the anterior intestinal portal and whether, in a broader context, the laterality in tissue movement in the anterior intestinal portal may predispose the direction of heart looping and the left-right asymmetry of the heart and the digestive tract.

Sox17 activity in the endoderm influences cardiac cell differentiation

In the mouse embryo, the cardiac cells are derived from two sources of progenitor cells: the primary and the secondary heart field. The primary heart field, consisting of *Tbx5* and *Myocd*-expressing cells contributes to the left ventricle and atria. The secondary heart field, populating by *Isl1*-expressing cells, contributes to the outflow tract, right ventricle and atria and the pharyngeal mesoderm (Buckingham *et al.*, 2005). Expression profiling analysis reveals that loss of *Sox17* is associated with the down-regulation of *Myocd* expression. The heart-specific transcription factors *Myocd*, *Tbx5* and *Isl1* are expressed in the cardiac crescent. However, *Myocd* and *Tbx5* are specifically down-regulated in the tissues at the vertex of the crescent. The domain where *Myocd* and *Tbx5* expression is lacking corresponds to the region of defective union of the lateral cardiac progenitors in the midline. The cardia bifida phenotype in the absence of *Sox17* is therefore most likely to be initiated by a failure in the fusion of the cardiac crescent at the midline leading to the formation of two separate lateral heart tubes. Our findings therefore suggest that *Sox17* activity in the definitive endoderm is required for the induction *Myocd*-expressing cardiogenic mesoderm in a specific population of cells in the primary heart field that form the atria and part of the ventricle.

Sox17 has been shown to act in a non-cell autonomous manner in eliciting cardiac differentiation in mouse ES cells. Inhibition of *Sox17* in mouse ES cells leads to a suppression of cardiac differentiation (Liu *et al.*, 2007). The expression of *Sox17* in the foregut endoderm could therefore be required for the induction of heart-specific genes in the embryo. This may occur either by direct regulation of genes encoding signalling molecules or antagonists by *Sox17*, or indirectly with *Sox17* playing a role in the maintenance or differentiation of the definitive endoderm, which then expresses the morphogenetic signals for heart development. Consistent with the non-cell autonomous role of *Sox17* in heart development, over-expressed *Sox17* in the pluripotent fibroblast cell line C3H10T1/2 does not induce *Myocd* expression (our unpublished data). On the other hand, co-culture of C3H10T1/2 with HepG2 cells does induce *Myocd* expression in the C3H10T1/2 cells, suggesting that the endoderm derived HepG2 cells are secreting a cardiac-inducing factor into the media. This is consistent with *Sox17* playing an indirect role in heart induction. However, transfection of a *Sox17* expression construct into the HepG2 cells does not result in increased *Myocd* expression in co-cultured C3H10T1/2 cells (F. Kruiswijk, DAFL and PPLT unpublished data). In this case, the HepG2 cells are already primed to express the cardiac-inducing factor independently of *Sox17* expression.

Materials and Methods

Mouse strains and genotyping

Sox17 mutant mice (Kanai-Azuma *et al.*, 2002) were maintained on a 129/Sv background. Heterozygous *Sox17*^{+/-} mice were inter-crossed to

generate wild type embryos, and heterozygous and homozygous mutant embryos for phenotypic analysis, expression studies and embryological experimentation. For testing the lineage potential of *Sox17*^{-/-} embryonic cells in the cell grafting experiment, *Sox17*^{-/-} mice were crossed with transgenic mice (H253) that co-express two transgenes (*Hmgcr-nls-lacZ* and pCAGG-EGFP) widely in embryonic tissues (Hadjantonakis et al., 1998; Kinder et al., 2001; Tam and Tan, 1992) to derive a line of *Sox17*^{-/-};H253 mice (SGX). The SGX mice were inter-crossed to generate embryos of different *Sox17* genotypes that express the transgenic reporters, so that the cells harvested from these embryos can be tracked during the course of the cell grafting experiment.

Microarray analyses

Total RNA from embryonic specimens was isolated using the RNeasy Micro Kit (QIAGEN) according to the manufacturer's instructions. RNA was isolated from two pairs of somite-number matched (4-5 somites) *Sox17*^{-/-} mutant and wild type embryos. The quality and quantity of the isolated RNAs were assessed on an Agilent Bioanalyser. RNA probes were amplified and labelled using the Affymetrix two-cycle process. Each RNA probe was hybridised to two or three Affymetrix GeneChip mouse genome 430 2.0 arrays, comprising all the known and predicted genes in the mouse genome. The arrays were read with an Affymetrix GS3000 scanner and raw data imported into Chipster v 1.4.4 (<http://chipster.csc.fi>) for normalization (RMA) and statistical analysis by linear modelling using LIMMA (significance at $P < 0.05$), taking into account biological and technical replicates. Analysis of gene ontology was carried out within Chipster using a hypergeometric test for over-representation of terms.

HepG2 (Human hepatocellular carcinoma) cells were maintained in DMEM + 10% FCS. Prior to transfection, approximately 2×10^6 cells were seeded into T75-flasks. One day later, the HepG2 cultures were transiently transfected with 8 µg of pCMV-*Sox17*-EGFP or CMV-EGFP control and 24 µl FuGENE6 (Roche). After 48 hours the EGFP expressing cells were isolated using the flow cytometry (FACSVantage cell sorter). Total RNA was extracted from the cell pellets using the RNeasy Mini Kit (QIAGEN) according to the manufacturer's protocol. RNA samples from three independent groups of transfected cells were amplified, labelled and hybridised to Illumina Sentrix® Human-6 v2 Expression BeadChips that cover the human genome. The arrays were analysed with the Illumina BeadStation 500 reader. Beadstudio was used to normalize the data (cubic spline) and identify transcripts whose expression level had significantly changed (Illumina Custom test, $p < 0.01$).

Quantitative real-time RT-PCR

RNA samples from *Sox17* mutant and wild type embryos (4-5 somites) were reverse transcribed into cDNA using the SuperScript III First-Strand Synthesis System (Invitrogen). Real-time quantitative RT-PCR was performed on Rotor-Gene thermocyclers (Corbett Research) using QuantiTect SYBR Green (QIAGEN) or Platinum Taq (Invitrogen). Details of PCR primers are given in Supplementary Tables S6 and S7.

Generation of riboprobes

To generate a plasmid from which to make a *Myocd* whole mount *in situ* hybridisation probe, a fragment was amplified from mouse E8.5 cDNA using *Pfu* DNA polymerase (Roche). Primer sequences for generating the 750bp probe were forward: 5'-TGGGCTAGACTCTGAGAAGGAC-3' and reverse: 5'-TGGGTGATATCTGAAACTGCTG-3'. The amplified fragments were subsequently cloned into pGEM-T (Promega) downstream of the T7 promoter. The *Tbx5* antisense probe has been described previously (Chapman et al., 1996). The *Is1* probe was kindly provided by Gerhard Przemek (Helmholtz Zentrum München). For the generation of antisense RNA probes from linearized cDNA clones the DIG RNA labelling kit (Roche) has been used according to the manufacturer's instructions.

Whole mount in situ hybridisation

For whole mount *in situ* hybridisations embryos were fixed in 4% PFA

in PBS over night, dehydrated in methanol and stored at -20°C until usage. Whole mount *in situ* hybridisations were performed as described previously (Chapman et al., 1996; Davidson et al., 1999). In brief, embryos were rehydrated in PBT (PBS plus 0.1% Tween-20), bleached with 6% hydrogen peroxide in PBT for 30min, washed in PBT, refixed in 4% PFA in PBS containing 0.2% glutaraldehyde for 20min and washed in PBT. Hybridisations with DIG-labelled antisense probes (1:200) were performed in hybridisation buffer (50% formamide, 5 x SSC pH4.5, 1% SDS, 50 µg/ml Heparin) at 70°C over night. After two washes each in 2 x SSC, 0.1% SDS and 0.2 x SSC, 0.1% SDS at 70°C, embryos were washed 3 times in MABT (100mM maleic acid pH 7.5, 150mM NaCl, 1% Tween-20, 2mM Levamisole). Embryos were blocked in 2% blocking reagent (Roche) in MABT for 1h, in 2% blocking reagent with 20% FCS in MABT for 1-3 h and incubated with alkaline phosphatase coupled anti-DIG antibody (Roche) (1:2000) in 2% blocking reagent with 20% FCS in MABT at 4°C over night. After extensive washing in MABT for 1-3 days the embryos were washed twice in 2mM Levamisole, 0.1% Tween-20. BM purple AP substrate (Roche) was used for subsequent alkaline phosphatase staining. The embryos were kept in 4% PFA. After whole mount *in situ* hybridisations the embryos were dehydrated, embedded in paraffin wax and sectioned at 7 µm. The sections were counterstained with nuclear fast red and mounted using Canada balsam.

X-Gal staining of embryos

Embryos were washed twice in PBS, fixed in 4% PFA in PBS for 5min and washed again twice in PBS. In order to detect *lacZ* the embryos were stained in X-gal [5-bromo-5-chloro-3-indoyl-β-D-galactopyranoside] staining solution (1mg/ml X-gal, 5mM K₃Fe(CN)₆, 5mM K₄Fe(CN)₆, 2mM MgCl₂, 0.01% Tween-20, 0.2% PFA in PBS) for 40 min at 37°C. Afterwards the embryos were rinsed twice in PBS and fixed in 4% PFA in PBS over night. The stained embryos were processed for wax histology and the relative contribution of the graft-derived cells to various types of germ layer derivatives evaluated by scoring the number of X-gal stained cells in serial sections of the recipient embryos.

Cell transplantation

Embryos from SGX mating were collected at the E7.0 mid-streak stage. GFP positive embryos were selected as donors for transplantation. Tissue fragments from the anterior primitive streak of these embryos were isolated and further dissociated into small clumps of cells. The cell clumps were then transplanted to the anterior primitive streak of stage-matched wild type (ARC/s) embryos using a Leica micromanipulator. The remaining part of the embryo was collected for PCR genotyping. After grafting, the recipient embryos were checked by fluorescence microscopy for the correct placement of the transplanted cells. Embryos were cultured for 24-28h in a medium made up of 75% heat-inactivated rat serum and 25% Dulbecco's modified Eagle medium at 37°C in glass bottles rotating at 30 RPM with a continuously replenished gas phase of 5% CO₂, 20% O₂ and 75% N₂ (Sturm and Tam, 1993). At the end of culture, recipient embryos were imaged by fluorescence photomicroscopy to visualise and record the distribution of the EGFP-expressing cells. Embryos containing EGFP-expressing graft-derived cells were then fixed in 4% PFA and stained in X-gal staining solution to detect *lacZ* expression.

Dye labelling of the foregut endoderm

E7.0 mid-streak-stage embryos were harvested from pregnant mice generated by inter-crossing *Sox17*^{-/-} mice. Embryos were dissected from the decidua and the Reichert's membrane was removed. Embryos were selected by somite numbers to ensure stage-matching (1-3 somites) between groups of different *Sox17* genotypes. They were kept in 100% rat serum in a 5% CO₂ incubator at 37°C prior to micro-manipulation. For dye labelling, cells in the anterior definitive endoderm were painted with carbocyanine dyes: DiO (D275, Molecular Probes) and CM-Dil (C-7001, Molecular Probes) (Bildsoe et al., 2007). In each embryo, around 100-150 cells each were labelled with one of the two dyes in order to reveal the

relative spatial distribution of different cell populations and to track their morphogenetic movement during foregut morphogenesis. Following dye labelling, embryos were cultured for 12 hours under the same conditions as the cell grafting studies. Labelled embryos were imaged by fluorescence microscopy within 1 hour after labelling to ascertain the site of labelling. The embryos were re-imaged at the end of a 12-hour culture period to visualise the distribution of labelled cells in the embryonic gut. Photographs were taken using a Leica MZ16 microscope with a SPOT Advanced digital camera and fluorescent and bright field images were digitally edited and merged with the SPOT 4.0 software and Adobe Photoshop 7.0. The yolk sac of the embryo was collected to prepare the DNA for genotyping.

Acknowledgements

We acknowledge the support of a Deutscher Akademischer Austauschdienst Fellowship (SP) and studentship from University of Applied Sciences, Mannheim (GT). DL is a Kimberly-Clark Research Fellow and PT is a NMHRC Senior Principal Research Fellow. This work is supported by NHMRC project grant 321704. Our thanks also to Mr. James Fairfax.

References

- ALEXANDER, J., ROTHENBERG, M., HENRY, G.L. and STAINIER, D.Y. (1999). casanova plays an early and essential role in endoderm formation in zebrafish. *Dev Biol* 215: 343-357.
- BILDSE, H., FRANKLIN, V. and TAM, P.P.L. (2007). Fate-Mapping Technique: Using Carbocyanine Dyes for Vital Labeling of Cells in Gastrula-Stage Mouse Embryos Cultured *In vitro*. *Cold Spring Harb Protoc* 2007: db.
- BUCKINGHAM, M., MEILHAC, S. and ZAFFRAN, S. (2005). Building the mammalian heart from two sources of myocardial cells. *Nat Rev Genet* 6: 826-835.
- CHAPMAN, D.L., GARVEY, N., HANCOCK, S., ALEXIOU, M., AGULNIK, S.I., GIBSON-BROWN, J.J., CEBRA-THOMAS, J., BOLLAG, R.J., SILVER, L.M. and PAPAIOANNOU, V.E. (1996). Expression of the T-box family genes, Tbx1-Tbx5, during early mouse development. *Dev Dyn* 206: 379-390.
- CLEMENTS, D., CAMELEYRE, I. and WOODLAND, H.R. (2003). Redundant early and overlapping larval roles of Xsox17 subgroup genes in *Xenopus* endoderm development. *Mech Dev* 120: 337-348.
- DAVIDSON, B.P., KINDER, S.J., STEINER, K., SCHOENWOLF, G.C. and TAM, P.P. (1999). Impact of node ablation on the morphogenesis of the body axis and the lateral asymmetry of the mouse embryo during early organogenesis. *Dev Biol* 211: 11-26.
- FRANKLIN, V., KHOO, P.L., BILDSE, H., WONG, N., LEWIS, S. and TAM, P.P. (2008). Regionalisation of the endoderm progenitors and morphogenesis of the gut portals of the mouse embryo. *Mech Dev* 125: 587-600.
- GEORGE, E.L., BALDWIN, H.S. and HYNES, R.O. (1997). Fibronectins are essential for heart and blood vessel morphogenesis but are dispensable for initial specification of precursor cells. *Blood* 90: 3073-3081.
- HADJANTONAKIS, A.K., GERTSENSTEIN, M., IKAWA, M., OKABE, M. and NAGY, A. (1998). Generating green fluorescent mice by germline transmission of green fluorescent ES cells. *Mech Dev* 76: 79-90.
- HARMON, E.B., KO, A.H. and KIM, S.K. (2002). Hedgehog signaling in gastrointestinal development and disease. *Curr Mol Med* 2: 67-82.
- HART, A.H., HARTLEY, L., SOURRIS, K., STADLER, E.S., LI, R., STANLEY, E.G., TAM, P.P., ELEFANTY, A.G. and ROBB, L. (2002). Mixl1 is required for axial mesoderm morphogenesis and patterning in the murine embryo. *Development* 129: 3597-3608.
- HOU, J., CHARTERS, A.M., LEE, S.C., ZHAO, Y., WU, M.K., JONES, S.J., MARRA, M.A. and HOODLESS, P.A. (2007). A systematic screen for genes expressed in definitive endoderm by Serial Analysis of Gene Expression (SAGE). *BMC Dev Biol* 7: 92.
- IYER, V.R., EISEN, M.B., ROSS, D.T., SCHULER, G., MOORE, T., LEE, J.C.F., TRENT, J.M., STAUDT, L.M., HUDSON, J., BOGUSKI, M.S., LASHKARI, D., SHALON, D., BOTSTEIN, D. and BROWN, P.O. (1999). The Transcriptional Program in the Response of Human Fibroblasts to Serum. *Science* 283: 83-87.
- KANAI-AZUMA, M., KANAI, Y., GAD, J.M., TAJIMA, Y., TAYA, C., KUROHMARU, M., SANAI, Y., YONEKAWA, H., YAZAKI, K., TAM, P.P. and HAYASHI, Y. (2002). Depletion of definitive gut endoderm in Sox17-null mutant mice. *Development* 129: 2367-2379.
- KIEFER, J.C. (2003). Molecular mechanisms of early gut organogenesis: a primer on development of the digestive tract. *Dev Dyn*, 228: 287-291.
- KINDER, S.J., TAN, S.S. and TAM, P.P. (2000). Cell grafting and fate mapping of the early-somite-stage mouse embryo. *Methods Mol Biol* 135: 425-437.
- KINDER, S.J., TSANG, T.E., WAKAMIYA, M., SASAKI, H., BEHRINGER, R.R., NAGY, A. and TAM, P.P. (2001). The organizer of the mouse gastrula is composed of a dynamic population of progenitor cells for the axial mesoderm. *Development* 128: 3623-3634.
- KUO, C.T., MORRISEY, E.E., ANANDAPPA, R., SIGRIST, K., LU, M.M., PARMACEK, M.S., SOUDAIS, C. and LEIDEN, J.M. (1997). GATA4 transcription factor is required for ventral morphogenesis and heart tube formation. *Genes Dev* 11: 1048-1060.
- LAUBNER, D., BREITLING, R. and ADAMSKI, J. (2003). Embryonic expression of cholesterologenic genes is restricted to distinct domains and colocalizes with apoptotic regions in mice. *Brain Res Mol Brain Res* 115: 87-92.
- 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.
- LEWIS, S.L., TAM, P.P. (2006). Definitive endoderm of the mouse embryo: formation, cell fates, and morphogenetic function. *Dev Dyn* 235: 2315-2329.
- LI, S., ZHOU, D., LU, M.M. and MORRISEY, E.E. (2004). Advanced cardiac morphogenesis does not require heart tube fusion. *Science* 305: 1619-1622.
- LIU, Y., ASAKURA, M., INOUE, H., NAKAMURA, T., SANO, M., NIU, Z., CHEN, M., SCHWARTZ, R.J. and SCHNEIDER, M.D. (2007). Sox17 is essential for the specification of cardiac mesoderm in embryonic stem cells. *Proc Natl Acad Sci USA* 104: 3859-3864.
- LUNDE, K., BELTING, H.G. and DRIEVER, W. (2004). Zebrafish pou5f1/pou2, homolog of mammalian Oct4, functions in the endoderm specification cascade. *Curr Biol* 14: 48-55.
- MARIJANOVIC, Z., LAUBNER, D., MOLLER, G., GEGER, C., HUSEN, B., ADAMSKI, J. and BREITLING, R. (2003). Closing the gap: identification of human 3-ketosteroid reductase, the last unknown enzyme of mammalian cholesterol biosynthesis. *Mol Endocrinol* 17: 1715-1725.
- MOLKENTIN, J.D., LIN, Q., DUNCAN, S.A. and OLSON, E.N. (1997). Requirement of the transcription factor GATA4 for heart tube formation and ventral morphogenesis. *Genes Dev* 11: 1061-1072.
- NIAKAN, K.K., JI, H., MAEHR, R., VOKES, S.A., RODOLFA, K.T., SHERWOOD, R.I., YAMAKI, M., DIMOS, J.T., CHEN, A.E., MELTON, D.A., MCMAHON, A.P. and EGGAN, K. (2010). Sox17 promotes differentiation in mouse embryonic stem cells by directly regulating extraembryonic gene expression and indirectly antagonizing self-renewal. *Genes Dev* 24: 312-326.
- PEARCE, J.J.H., EVANS, M.J. (1999). Mml, a mouse Mix-like gene expressed in the primitive streak. *Mech. Dev.* 87: 189-192.
- QU, X.B., PAN, J., ZHANG, C. and HUANG, S.Y. (2008). Sox17 facilitates the differentiation of mouse embryonic stem cells into primitive and definitive endoderm *in vitro*. *Dev Growth Differ* 50: 585-593.
- REIM, G., MIZOGUCHI, T., STAINIER, D.Y., KIKUCHI, Y. and BRAND, M. (2004). The POU domain protein spg (pou2/Oct4) is essential for endoderm formation in cooperation with the HMG domain protein casanova. *Dev Cell* 6: 91-101.
- REITER, J.F., ALEXANDER, J., RODAWAY, A., YELON, D., PATIENT, R., HOLDER, N. and STAINIER, D.Y. (1999). Gata5 is required for the development of the heart and endoderm in zebrafish. *Genes Dev* 13: 2983-2995.
- ROEBROEK, A.J., UMANS, L., PAULI, I.G., ROBERTSON, E.J., VAN LEUVEN, F., VAN DE VEN, W.J. and CONSTAM, D.B. (1998). Failure of ventral closure and axial rotation in embryos lacking the proprotein convertase Furin. *Development* 125: 4863-4876.
- SAKAMOTO, Y., HARA, K., KANAI-AZUMA, M., MATSUI, T., MIURA, Y., TSUNEKAWA, N., KUROHMARU, M., SAJJOH, Y., KOOPMAN, P. and KANAI, Y. (2007). Redundant roles of Sox17 and Sox18 in early cardiovascular development of mouse embryos. *Biochem Biophys Res Commun* 360: 539-544.

- SEGUIN, C.A., DRAPER, J.S., NAGY, A. and ROSSANT, J. (2008). Establishment of endoderm progenitors by SOX transcription factor expression in human embryonic stem cells. *Cell Stem Cell* 3: 182-195.
- SHIVDASANI, R.A. (2002). Molecular regulation of vertebrate early endoderm development. *Dev Biol* 249: 191-203.
- 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.
- SPENCE, J.R., LANGE, A.W., LIN, S.C., KAESTNER, K.H., LOWY, A.M., KIM, I., WHITSETT, J.A. and WELLS, J.M. (2009). Sox17 regulates organ lineage segregation of ventral foregut progenitor cells. *Dev Cell* 17: 62-74.
- STAINIER, D.Y. (2002). A glimpse into the molecular entrails of endoderm formation. *Genes Dev* 16: 893-907.
- STURM, K., TAM, P.P. (1993). Isolation and culture of whole postimplantation embryos and germ layer derivatives. *Methods Enzymol* 225: 164-190.
- 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., KHOO, P.L., LEWIS, S.L., BILDSOE, H., WONG, N., TSANG, T.E., GAD, J.M. and ROBB, L. (2007). Sequential allocation and global pattern of movement of the definitive endoderm in the mouse embryo during gastrulation. *Development* 134: 251-260.
- TAM, P.P., TAN, S.S. (1992). The somitogenetic potential of cells in the primitive streak and the tail bud of the organogenesis-stage mouse embryo. *Development* 115: 703-715.
- TREMBLAY, K.D., 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.
- UEMURA, M., HARA, K., SHITARA, H., ISHII, R., TSUNEKAWA, N., MIURA, Y., KUROMARU, M., TAYA, C., YONEKAWA, H., KANAI-AZUMA, M. and KANAI, Y. (2010). Expression and function of mouse Sox17 gene in the specification of gallbladder/bile-duct progenitors during early foregut morphogenesis. *Biochem. Biophys. Res. Commun.* 391: 357-363.
- WANG, D., CHANG, P.S., WANG, Z., SUTHERLAND, L., RICHARDSON, J.A., SMALL, E., KRIEG, P.A. and OLSON, E.N. (2001). Activation of cardiac gene expression by myocardin, a transcriptional cofactor for serum response factor. *Cell* 105: 851-862.
- WOODLAND, H.R., ZORN, A.M. (2008). The core endodermal gene network of vertebrates: combining developmental precision with evolutionary flexibility. *Bioessays* 30: 757-765.
- ZORN, A.M., WELLS, J.M. (2007). Molecular basis of vertebrate endoderm development. *Int. Rev Cytol.* 259: 49-111.
- ZORN, A.M., WELLS, J.M. (2009). Vertebrate endoderm development and organ formation. *Annu. Rev Cell Dev Biol* 25: 221-251.

Further Related Reading, published previously in the *Int. J. Dev. Biol.*

See our recent Special Issue **Placenta** edited by Joan S. Hunt and Kent L. Thornburg at:
<http://www.ijdb.ehu.es/web/contents.php?vol=54&issue=2-3>

Characterization of mouse embryonic stem cell differentiation into the pancreatic lineage in vitro by transcriptional profiling, quantitative RT-PCR and immunocytochemistry

Alexandra Rolletschek, Insa S. Schroeder, Herbert Schulz, Oliver Hummel, Norbert Huebner and Anna M. Wobus
Int. J. Dev. Biol. (2010) 54: 41-54

Movement and commitment of primitive streak precardiac cells during cardiogenesis

Carmen Lopez-Sanchez, Natividad Garcia-Masa, Carlos M. Gañan and Virginio Garcia-Martinez
Int. J. Dev. Biol. (2009) 53: 1445-1455

Building the vertebrate heart - an evolutionary approach to cardiac development

José M. Pérez-Pomares, Juan M. González-Rosa and Ramón Muñoz-Chápuli
Int. J. Dev. Biol. (2009) 53: 1427-1443

Defects of the body plan of mutant embryos lacking *Lim1*, *Otx2* or *Hnf3beta* activity

S J Kinder, T E Tsang, S L Ang, R R Behringer and P P Tam
Int. J. Dev. Biol. (2001) 45: 347-355

Expression of regulatory genes for pancreas development during murine embryonic stem cell differentiation

Josué K. Mfopou, Erik Willems, Luc Leyns and Luc Bouwens
Int. J. Dev. Biol. (2005) 49: 915-922

5 yr ISI Impact Factor (2008) = 3.271

