Supplementary Figures

Figure S1: Robust expression of induced Sox2 at E16.5

IHC using an antibody against Sox2 on intestinal cross-sections at E14.5, E16.5 and E18.5 of double transgenic animals shows that robust Sox2 expression is observed from E14.5 onwards (A). In order to evaluate the expression level of induced Sox2, we compared Sox2 mRNA expression levels of Dox-induced double transgenic embryos with the stomach and intestine of control animals. We show an approximate 6-fold increase of expression of induced Sox2 compared to the endogenous level in the stomach. As expected Sox2 expression in the control intestine was negligible (B). Additionally, we analysed the unaffected stomach and the intestinal tract of the induced double transgenic animals with immunohistochemistry using limiting amounts of Sox2 antibody. Because the titration of the antibody showed diminishing staining, we could perfectly compare the intensity of the staining in de stomach and intestine and quantify the expression level of the ectopically expressed Sox2 in the intestinal epithelium using ImageJ. This revealed an average increase in intensity of 32% in the intestine, compared to endogenous Sox2 expression in the stomach (C). Scale bars: 50 μm.

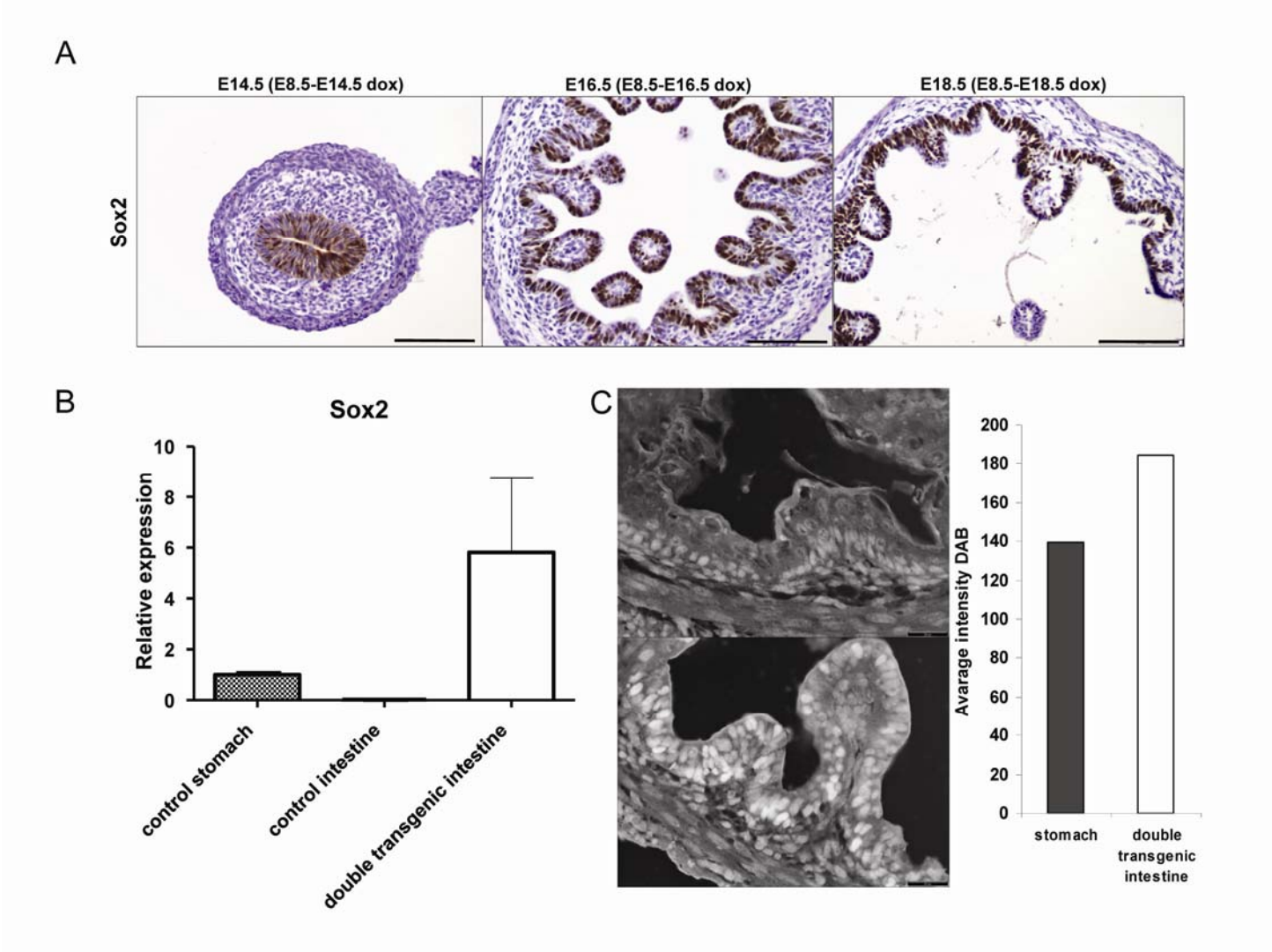


Figure S2: Ectopic Sox2 induces significant enlargement of the intestine

The average intestinal lumen size was measured in at least three controls and double transgenic animals. The double transgenic animals show a two-fold increase in lumen size.

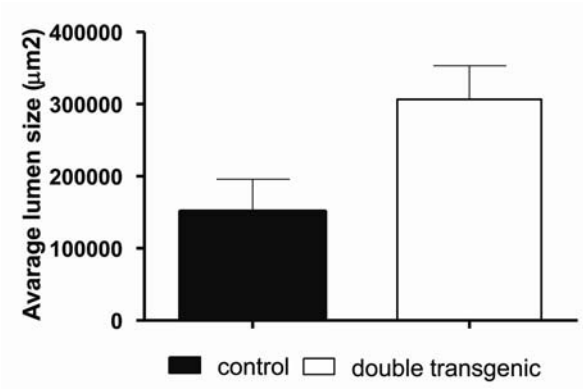


Figure S3: Overview of IHC on control and double transgenic intestines

Overview of IHC results with antibodies against Ki67 (A), Muc2 (B), Syp2 (C), p63 (D), H+/K+ Atpase4β (E) and Hnf4a (F) on cross-sections of duodenum of control (left panels) and double transgenic (center panels) animals at E18.5. Each panel consists of three images, representing 100, 200 and 400 times magnification (scale bars: 200, 100 and 20 μm, respectively). Quantification is represented as the number of positive cells per total number of epithelial cells in at least 3 microscopic fields (right panels). Black bars represent the control and white bars the double transgenic animals.

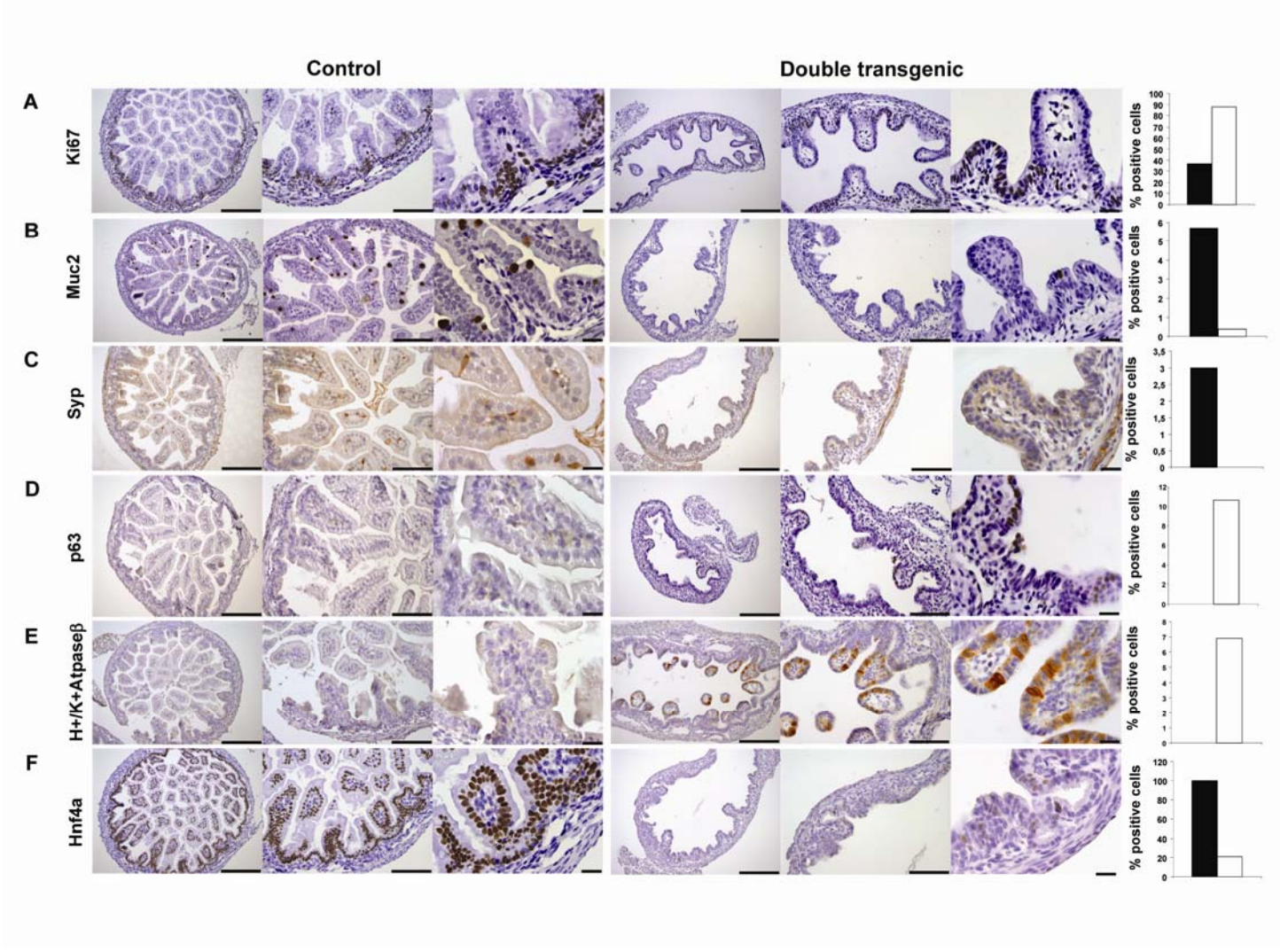


Figure S4: Quantification of the expression of aquaporin family members

Analysis of the expression of Aquaporin family members Aqp1, Aqp2, Aqp4, Aqp7, Aqp8, Aqp11 showed a downregulation in the double transgenic animals, whereas Aqp 12 was upregulated compared to the controls.

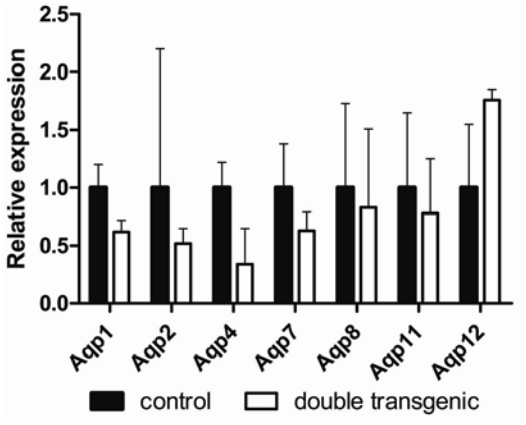


Figure S5: GO enrichment analysis of control and double transgenic intestines

Enriched GO terms within the set of 1,354-regulated probesets (both up- and downregulated genes) for GO biological processes (A), GO molecular function (B) and GO cellular component (C). Enriched categories are those identified as significantly enriched (P<0.05) after multiple testing. *P < 0.05; **P < 0.01; ***P < 0.001. (C) Enriched GO terms within the set of 906-downregulated probesets for GO biological processes (D). Enriched categories are those identified as significantly enriched (P<0.05) after multiple testing. *P < 0.05; **P < 0.01; ***P < 0.001.

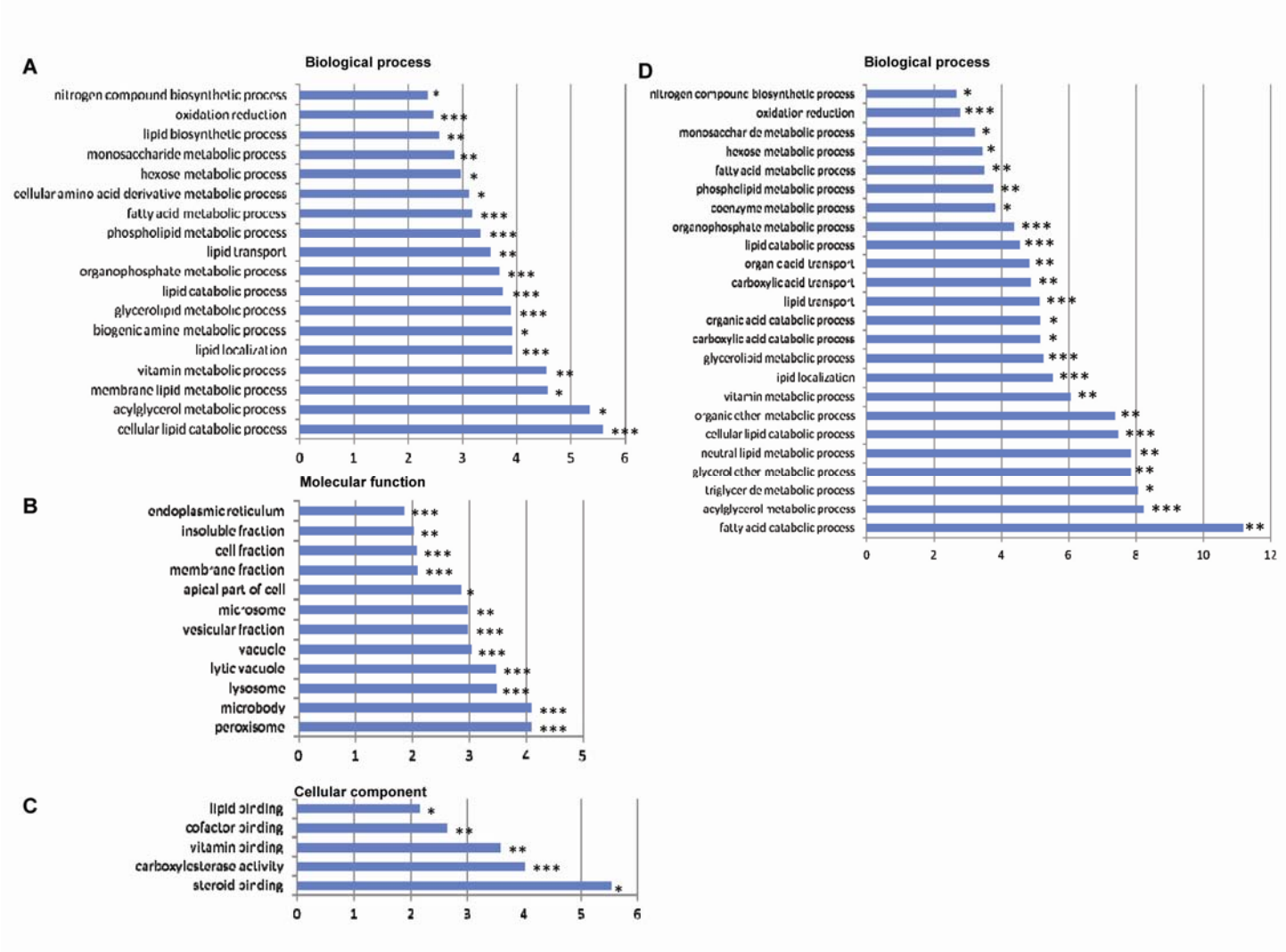


Figure S6: Sox2 affects the intestinal brush border

Close up electron microscopy images of the intestinal brush border of control and double transgenic E18.5 pups, and of the embryonic stomach, which does not have a brush border. Scale bars: 1.4μm

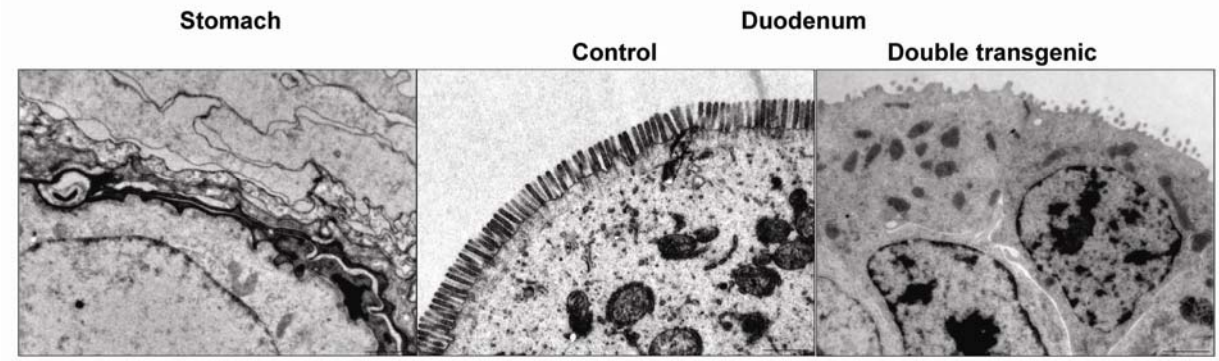
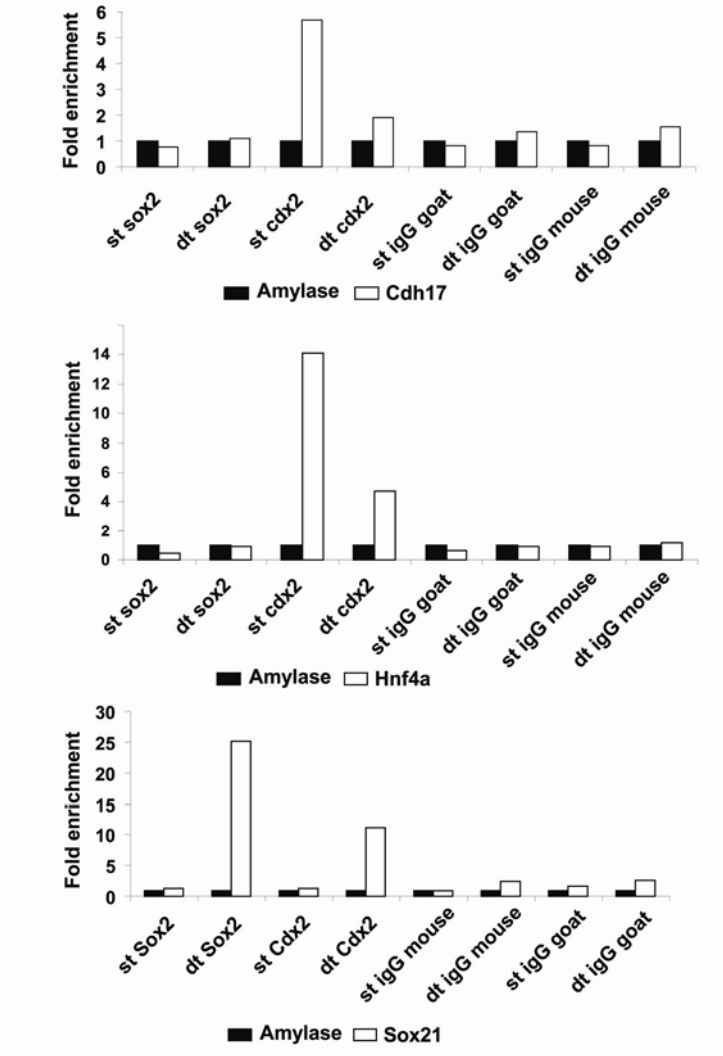


Figure S7:Sox2 interferes with Cdx2 binding to target genes

ChiP assay for Cdx2 shows loss of binding to Cdh17 (top panel) and Hnf4a (middle panel) in the double transgenic animals, compared to the control. Sox2 does not bind to Cdh17 or Hnf4a, but does bind to its downstream target Sox21 (bottom). Mouse IgG or goat IgG serve as negative antibody controls Cdx2 and Sox2, respectively. Amylase served as negative control for the qPCR.



Supplementary Tables

Supplementary Table S1

Antigen	Clone	Target	Source	Concentration	Dilution	Amplificat Kit
Myc-epitope	9E10	Myc-epitope	Roche	5mg/ml	1:800	None
Sox2	401196	Sox2	Immune systems	1mg/ml	1:500	ABC-kit
Cdx2	Cdx2-88	Cdx2	Biogenex	10-15mg/ml	1:20	ABC-kit
Phospho-histone H3	32219	mitosis	Upstate, cell sinaling solutions	1mg/ml	1:800	ABC-kit
Ki67	TEC-3	proliferation	Dakocytomation		1:50	ABC-kit
Cleaved Caspase 3	5A1E	apoptosis	Cell Signaling	100µg/ml	1:100	ABC-kit
Mucin 2	H-300 Sc-15334	goblet cells	Santa Cruz Biotechnology, INC	1mg/ml	1:400	ABC-kit
Synaptophysin	A0010	entero-endocrine cells	DakoCytomation	300µg/ml	1:250	Envision
p63	4A4 Sc-8431	basal cells	Santa Cruz Biotechnology, INC	200µg/ml	1:200	ABC-kit
H/K+ ATPase 4B	2G11	parietal cells	thermo scientific	10µg/ml	1:2000	ABC-kit
HNF4	K9218	HNF4	Abcam	1mg/ml	1:100	ABC-kit
GSII Lectin	L21415	mucous neck cells	Molecular probes	1mg/ml	1:2000	None
Mist1		zymogenic chief cells	Gift of Jason Mills		1:500	None
ZO-1	61-7300	Apical membranes	Invitrogen	250µg /ml	1:50	None
Laminin	L9393	Basement membranes	Sigma		1:400	None
E-cadherin	24E10	Lateral membranes	Cell SignalingTechnology		1:20	ABC-kit

* StreptABCcomplex/HRP (Dako)
** Envision+systems (Dako)

Supplementary Table S2

Genomic region	Sequence (forward)	Sequence (reverse)
Sox21	GCAGGCGCATAAATAAATAA	ATATCCATTCAAAGGGCATT
Cdh17	TTAAAACAACACCACCACCAC	CCCCAGTCAAACATTAACCAC
Hnf4a	AGGCTGAGGCTATGAGAAC	AACTCTCCCCTGACTCCTTGC
Amylase	CTCCTTGTACGGGTTGGT	AATGATGTGCACAGCTGAA

Supplementary Table S3

Gene	Sequence (forward)	Sequence (reverse)
Aqp4	CTGTGATTCCAAACGAACTG	GGCTCCAGTATAATTGATTGCA
Aqp8	CTACTGGGACTTCCATTGGA	CCGATGAGGAGCCTAATGAG
H+/K+Atpase4a	GACCACTGATGATAATCTGTACCT	GATATTTGTGCTCTTGAACTCCTG
Cdx2	GTATGTCTGTGTTGTAAATGCC	AAACAATTCCGGTCTTCTTCAG
ChgA	CAGAAGTGTTTGAGAACCAGAG	TTCTCTTCTCCATAGTGTCCC
Krt13	CTGACTCTGGCTAAGACTGAC	AATTCCTTCATCTCCTCTTCGT
Muc2	TGCAACAACTTAACTGCTCTG	TCAGTATGGTAATAGCCAGCC
Muc5ac	CATGACCTGTTATAGCTCCGA	CTCAGTAACAACACAGCCTC
p63	CATTGTCAGTTTCTTAGCAAGG	CTCAATCTGATAGATGGTGGT
Slc2a2	CAGAAGACAAGATCACCGGA	GCATTGATCACACCGATGTC
Slc5a1	ATTGAAATAGACACAGAAGCCC	GTCATCTTTGGTCCTTTATCCT
Lct	GCTTCCTATCAGGTTGAAGGT	GTCGTCATTCCCAATCTTCAG
Hnf4a	CTTTGATCCAGATGCCAAGG	GGTCGTTGATGTAATCCTCCA
Mep1a	CATCTTCAGCTATAAATGGCTC	CTTCTGAAACAATCACAGTCCT
Heph	GCAGAAGAGATAGAGTGGGA	ATAGCTGTCTTTCTCAGATGTG
Cdx1	AAAGGAGTTTCACTACAGCC	GAACCAGATCTTTACCTGCC
Aqp1	CTCCCTAGTCGACAATTCAC	CCAATGATCTCAATGCCAG
Aqp2	ACCTCCTTGGGATCTATTTCAC	ATCATCAAACTTGCCAGTGAC
Aqp7	GAGCTACAGTTCAGTTGCAG	ATGAAGTAGGTTCTCTGAAGTG
Aqp11	TCACAGGAGCATTGTTTAACC	ATCAGCACACCTACAGAAGG
Aqp12	CACAGCCTTCTTGTCTACAG	GGATTGAAGAAGGCAGATGTG

Supplementary Materials and methods

Microarray analysis

The intestinal tract of three control and three double transgenic embryos, which received doxycycline from E8.5 onwards, were isolated at E18.5 and individually used for total RNA isolation with Trizol reagent (Invitrogen life technologies, Carlsbad, CA, USA). RNA was purified using RNeasy MinElute Cleanup kit (Qiagen, Valencia, CA, USA) and cDNA was synthesized using the GeneChip Expression 3'-Amplification Reagents One-Cycle cDNA Synthesis kit (Affymetrix, Santa Clara, CA, USA). Biotin-labelled cRNA synthesis, purification and fragmentation were performed according to standard conditions. Fragmented biotinylated cRNA was subsequently hybridised onto Affymetrix Mouse Genome 430 2.0 microarray chips.

To examine the quality of the various arrays, several R packages (including affyQCreport) were run starting from the CEL files. All created plots, including the percentage of present calls, noise, background, and ratio of GAPDH 39 to 59 (1.4) indicated a high quality of all samples and an overall comparability, except for one sample, which was excluded from further analysis. Raw intensity values of all samples were normalized by RMA normalization (Robust Multichip Analysis), background correction and quantile normalization using Partek version 6.4 (Partek Inc., St. Louis, MO).

The normalized datafile was transposed and imported into OmniViz version 6.0.1 (BioWisdom Ltd., Cambridge, UK) for further analysis. For each probe set, the geometric mean of the hybridization intensity of all samples was calculated. The level of expression of each probe set was determined relative to this geometric mean and 2log-transformed. The geometric mean of the hybridization signal of all samples was

used to ascribe equal weight to gene expression levels with similar relative distances to the geometric mean. Differentially expressed genes were identified using statistical analysis of microarrays (SAM). Cutoff values for significantly expressed genes were the number of falsely called less than 1 (FDR of 0.0006 or less) and a fold change of 2. Functional annotation of the statistical analysis of microarrays results was done using Ingenuity Pathway Analysis (Ingenuity, Mountain View, CA) and DAVID (http://david.abcc.ncifcrf.gov). DAVID calculates significant overrepresentation of gene ontology (GO)-classified biological processes. The results are shown for biological processes, which are significantly (P <0.05) enriched after multiple testing.

To analyze higher-order differentiation patterns defined by the 449 genes induced and 906 downregulated significantly by Sox2, each gene list was re-expressed in terms of the fractional representation of GO terms associated with its member genes using the GOurmet software suite.[1] The distribution of GO term fractional representations were then used as a metric to classify the Sox2 expression profiles to a database of expression profiles of mature, proliferating, and hyperplastic gastrointestinal tissues, as described before.[2]

Chromatin immunoprecipitation (ChIP)

The small intestines of, respectively, 10 double transgenic and 10 control E18.5 embryos were pooled and mechanically disrupted, after incubating for 20 minutes in ColagenaseII at 37°C. Next we filtered the cell suspension using a 70μm cell strainer. Formaldehyde was added to a final concentration of 1% for 10 min on RT to crosslink the samples and this process was stopped by adding 0.15 M glycine. Cells were washed with PBS and lysed for 10 minutes on ice in cell lysis buffer (10mM Tris pH 8.0, 10 mM Nacl, 0.2% NP-40 and complete protease inhibitors). After centrifugation the nuclei were resuspended in nuclei lysis buffer (50mM Tris pH

8.0, 10mM EDTA, 1% SDS and complete protease inhibitors) and the chromosomal DNA was fragmented by sonication (20 cycles of 15 seconds, 45 seconds in between cycles) to yield DNA fragments with an average size of 500bp.

Equal amounts of sample were diluted 1:10 with ChIP dilution Buffer (0.01% SDS, 1.1% Tx-100, 1.2 mM EDTA, 16.7mM Tris-HCL pH 8.1 and 167 mM Nacl) and from each sample 50 μl was removed to serve as input control. The samples were pre-cleared with 80μl Prot A/G agarose beads for 1 hour. Next, these samples were incubated O/N with pre-formed complexes of 10 μg antibody specific for Sox2 or Cdx2, or control IgG (goat and mouse) with 150 μl Prot A/G agarose beads.

Beads were washed with Low salt immune complex buffer (0.1% SDS, 1% Tx-100, 2mM EDTA, 20mM Tris-HCL pH 8.0 and 50 mM Nacl), High salt immune complex buffer (0.1% SDS, 1% Tx-100, 2mM EDTA, 20mM Tris-HCL pH 8.0 and 150 mM Nacl), LiCl immune complex buffer (0.25 M LiCl, 1% NP-40, 1% deoxycholate, 1mM EDTA and 10mM Tris-HCL pH 8.0) and twice with TE (10mM Tris-HCL pH 8.0, 1mM EDTA pH 8.0). The DNA was eluted by adding twice 250 μl elution buffer (1% SDS and 0.1m NaHCO3). Next the samples and the input were incubated at 65°C O/N using 20 μl 5M NaCl to de-crosslink the DNA and proteins. The eluted material was phenol-extracted and ethanol-precipitated. The DNA was resuspended in 26 μl of water and qPCRs were performed to analyze the enrichment of Cdx2 for binding to *Hnf4a* and *Cdh17*, and Sox2 for binding to *Sox21*, using Amylase as control.

1. Doherty JM, Carmichael LK, Mills JC. Gourmet: a tool for quantitative comparison and visualization of gene expression profiles based on gene ontology (GO) distributions. BMC Bioinformatics 2006;7:151.
2. Doherty JM, Geske MJ, Stappenbeck TS, Mills JC. Diverse adult stem cells share specific higher-order patterns of gene expression. Stem Cells 2008;26:2124-30.