SUPPORTING MATERIALS

Index:

Materials and Methods

Retroviral transduction and transfections.

Retroviral gene transfer was performed using high-titer retroviral stocks generated by transient transfection of HEK293T cell line with different retroviral constructs. In particular, we used pLPC-puro either empty or carrying *H-RasV12*, or *E1a-ires-H-RasV12*. Viral supernatants were added to 10^6 MEFs for 8 hours, and 24 hr later the same procedure was repeated. Infected cells were selected for 3 days in the presence of 2.5 μg/ml of puromycin. After selection, protein extracts were prepared at days 2 post-selection.

For *SOX9* or *Bmi1* knockdown by shRNA: Cells were transfected using Lipofectamine (invitrogen), according to instructions of the reagent supplier. Cells were cultivated in the presence of puromycin for 3 weeks. Cells were selected and SOX9 expression assayed. Cells were transfected using two different *Sox9* shRNAs (Origene, *sh1* or *sh75* and *sh2* or *sh73*) and a shRNA especific for BMI1. A non-specific shRNA (*pRS*) was used as a control.

Chromatin immunoprecipitation (ChIP) assays – ChIP assays were performed according to standard procedures. Briefly, sub-confluent MEF cultures (5 x 10^6 cells) were crosslinked with 1% formaldehyde for 15 min at room temperature, and the crosslinking stopped with 0.13 M glycine. Subsequently, cells were washed with cold PBS, scrapped, lysed (1% SDS) and sonicated to obtain DNA fragments of 500-1,000 base pairs. DNA fragments were immunoprecipitated with the different antibodies coupled to protein agarose A beads previously blocked with salmon sperm DNA. The

immuno-precipitated DNA was extracted and subjected to PCR amplification with primers directed against the proximal promoter of *Sox9* or *Bmi1*. CHIP primers were:
Sox9, 5'-GATCTACAGGCCTCTACCAG and 5'-GATCTACAGGCCTCTACCAG-3';
Bmi1 5'-ACGGGCCTGACTACACCGACACT-3' and 5'-CTGAAGGCAGAGTGGAAACTGACAC-3';

FACS Analysis

Early passage MEFs were trypsinized, washed by PBS, and fixed with ice-cold 100% ethanol. DNA content was measured after stained by PI solution. Percentages of G0/G1 and S phase cells are shown.

Primers information.

The primers used were: for mice
Sox9, 5'-AGCTCACCAGACCCTGAGAA and 5'-CTCCAGCAATCGTTACCTTC;
Ink4a, 5'-AACTCT TTCGGTCGTACCCC-3' and 5'-GCGTGCTTGAGCTGAAGCTA-3';
Arf 5'-GCCGCACCGGAATCCT-3' and 5'-TTGAGCAGAAGAGCTGCTACGT-3';
Bmi1 5'-TGCAACTTCTCCTCGGTCTT-3' and 5'-TGTCCAGGTTCACAAAACCA-3';
Sox2, 5'-CACAACTCGGAGATCAGCAA and 5'-CTCCGGGAAGCGTGTACTTA;
Sox8, 5'-AGGCGAAGGAAGAGTGTGAA and 5'-CCCTTCCAGCCTTAGCTCTT;
Sox10, 5'-AGGCCTCACTGCTCCTGTTA, and 5'-TTGACCAGTTCCCACATTCA;
actin, 5′-GGCACCACACCTTCTACAATG-3′ and 5′-GTGGTGGTGAAGCTGTAGCC-3′.
For human:
SOX9, 5'-GTACCCGCACTTGCACAAC and 5'-TCGCTCTCGTTCAGAAGTCTC-3'
BMI1 5'-GGAGACCAGCAAGTATTGTCCTATTT-3' and 5'-CATTGCTGCTGCTGGGCATCGTAAG-3';

SUPPLEMENTARY TABLES

Table S1: Immunohistochemical staining of SOX9 in a pilot series of human cancers (n=6 per cancer).

cancer type	SOX9 positive cancers (%)
colorectal, pancreas, neuroglia, lung and prostate	more than 75
thyroid breast and sarcoma	between 30-50
kidney	less than 20
lymphoma	0%

Table S2: Immunohistochemical staining of SOX9 in human cancers.

cancer type	SOX9 positive cancers (%) under pathologist supervision	SOX9 positive cancers (%) by computerized analysis
colorectal cancers (n=79)	85	80
lung cancers (n=123)	75	75
neurofibroma cellular (n=3) neurofibroma (n=22) MPNST (n=10)	100 95 70	90
medulloblastoma (n=81)	87	95
prostatic carcinomas (n=54)	80	68
ovarian cancer (n=71)	68	73
pancreatic adenocarcinoma (n=12)	92	92

Table S3: Clinical data for two colorectal cancer series

	Dataset**	
	GSE25071	GSE24550
Tissue type		
Colorectal carcinomas	46	131
Normal mucosa	4	21
Sex		
Male	25	65
Female	21	66
Age		
Average age at	58	70

diagnosis		
Range	28-87	27-93
Tumor stage*		
I	9	33
II	13	47
III	18	36
IV	6	15
Localization		
Proximal	14	60
Distal	17	38
Rectum	15	32

*- the American Joint Committee on Cancer (AJCC) TNM stage criteria
** - transcriptome data sets

SUPPLEMENTARY FIGURES

Figure 1. High levels of SOX9 in human cancers.

A: Comparative pictures of SOX9 expression in sections derived from the following human cancers and their respective sane tissue (1) lung (2) medulloblastoma, (3) pancreas, (4) ovary and (5) prostate.

B: Representative pictures of SOX9 expression in sections derived from the following human cancers (1) lung, (2) neuroglia, (3) breast and (4) thyroid.

Figure 2. Representative immunohistochemical staining probing for SOX9 expression in a colorectal cancer TMA.

The expression of Sox9 throughout the samples included in this tissue microarray was determined by using an automated scanning microscope and computerized image analysis system (Ariol SL-50; Genetix). The system was trained by a team of technicians and pathologists to quantify SOX9 expression. In this manner, we determined the total and positive cell number for each sample. All the samples in the TMA were represented by two cores and we checked that the results obtained for each sample were consistently reproduced in duplicate. Samples showing a low number of total cells were manually inspected and discarded due to loss of integrity of the tissue in the core. Finally, tumor samples were scored positive when the number of SOX9 positive cells was at least above 25% of all the cells in the sample.

Figure 3. High levels of SOX9 in colorectal human cancers.

A: Significantly higher expression of *SOX9* ($p = 4 \times 10^{-16}$) in another set of 131 primary carcinomas in comparison to normal colonic mucosa. These are not included in the statistics in Fig. 1 as they were carried out using the Affymetrix GeneChip Human Exon 1.0 ST arrays (Affymetrix, Santa Clara, CA, USA; HuEx). The raw data are accessible through the NCBI's Gene Expression Omnibus public repository for microarray data (accession numbers GSE25071 (1).

B: *SOX9* expression was similar in colon and rectum cancers ($p = 0.51$).

C: Variation of *SOX9* mRNA expression in CRC cell lines. *SOX9* expression is higher than in normal colonic mucosa in all cases but RKO and Colo320 cell lines

Figure 4. Gain of SOX9 copy number in colorectal human cancers.

A: Images of genomic aberrations and mRNA expression in the primary colorectal carcinomas (3/46) that showed SOX9 gain.

Figure 5. Expression of Sox8, Sox10 and Sox2 in MEFs varying Sox9 levels.

A: RNA levels of *Sox8*, *Sox10* and *Sox2* derived from 7 different wt and 9 *Z/sox9tg* independent MEF cultures. PCR data were normalized to β-actin expression and are expressed relative to gene expression levels in *wild-type* cells.

B: RNA levels of *Sox8, Sox10* and *Sox2* derived from 5 different $Sox9^{flox/flox}$*;CreERT2* independent MEF cultures treated in the absence or presence or 4OHT for 5 days. PCR data were normalized to β-actin expression and are expressed relative to gene expression levels in ethanol treated cells.

Figure 6. Sox9 inversely regulates senescence markers

A,B: Early passage *Z/Sox9tg* contains lower *Ink4a, Arf, Ink4b and DcR2* than *wild-type* cells while $Sox9^{\Delta/\Delta}$ have higher levels of these senescence markers. Expression of *Ink4a, Arf, Ink4b and DcR2* in *Z/Sox9tg* (n=7) and $Sox9^{\Delta/\Delta}$ (n=7) are shown relative to *wild-type* (n=9) and $Sox9^{flox/flox}$ (n=7) respectively.

C,D: *H-RasV12* transduced *Z/Sox9tg* contains lower *Ink4a, Arf, Ink4b and DcR2* than *wild-type* cells while $Sox9^{\Delta/\Delta}$ have higher levels of senescence markers. Senescence marker expression were obtained 3 days after selection and are representative from 3 independent cultures

Figure 7. Sox9 role in proliferation is partly dependent on *Ink4a/Arf*

A: $Ink4a/Arf^{-/-}$ MEFs were retrovirally transduced with *pCAGGS-empty vector* or *pCAGGS-Sox9-IRES-GFP* and the proliferation growth curves analyzed. 5×10^4 cells were seeded in each well (in duplicate) of six-well tissue culture plates and cell numbers were counted at the indicated times.

Figure 8. Sox9 modulates tumor growth *in vivo*.

A: Representative image of tumors developed from *Z/Sox9tg* transformed cells (left side) and *wt* cells (right side), 25 days after injection into nude mice.

B: Representative image of tumors developed from *Sox9$^{flox/flox}$;CreERT2* transformed cells treated with ethanol (*Sox9$^{flox/flox}$*) or 4OHT (*Sox9$^{\Delta/\Delta}$*) for 5 days *in vitro* before subcutaneous injection into immunodeficient mice. The tumors shown were 30 days after injection into the same mouse.

Figure 9. *SOX9* is associated to *BMI1* and *INK4a-ARF* in colorectal primary cancers.

A: Scatter plots of *SOX9* expression versus *BMI1, INK4a* and *ARF*. *SOX9* expression correlated positively with *BMI1* (Spearman's coefficient 0.21, p = 0.02) and negatively with *ARF* (Spearman's coefficient -0.32, p = 0.0002). A weak negative correlation was also observed between *SOX9* and *INK4a* (Spearman's coefficient -0.12, p = 0.19).

B: The heat map illustrates associations *SOX9*, *BMI1, INK4a* and *ARF* expression in CRC patient cancers and normal colonic mucosa specimens. mRNA expression values obtained from microarray experiments were normalized to mean zero across all samples and standard deviation of one.

Additional references:

1. Sveen A, Agesen TH, Nesbakken A, Rognum TO, Lothe RA, Skotheim RI. Transcriptome instability in colorectal cancer identified by exon microarray analyses: Associations with splicing factor expression levels and patient survival. Genome Med 2011;3(5):32.