



OPEN **Transcriptional repression of the oncofetal *LIN28B* gene by the transcription factor SOX6**

Valentina Pastori^{1,6}, Gianluca Zambanini^{2,3,4,6}, Elisabetta Citterio¹, Tamina Weiss^{2,3}, Yukio Nakamura⁵, Claudio Cantù^{2,3} & Antonella Ellena Ronchi¹✉

The identification of regulatory networks contributing to fetal/adult gene expression switches is a major challenge in developmental biology and key to understand the aberrant proliferation of cancer cells, which often reactivate fetal oncogenes. One key example is represented by the developmental gene *LIN28B*, whose aberrant reactivation in adult tissues promotes tumor initiation and progression. Despite the prominent role of *LIN28B* in development and cancer, the mechanisms of its transcriptional regulation are largely unknown. Here, by using quantitative RT-PCR and single cell RNA sequencing data, we show that in erythropoiesis the expression of the transcription factor *SOX6* matched a sharp decline of *LIN28B* mRNA during human embryo/fetal to adult globin switching. *SOX6* overexpression repressed *LIN28B* not only in a panel of fetal-like erythroid cells (K562, HEL and HUDEP1; $\approx 92\% p < 0.0001$, $54\% p = 0.0009$ and $\approx 60\% p < 0.0001$ reduction, respectively), but also in hepatoblastoma HepG2 and neuroblastoma SH-SY5H cells ($\approx 99\% p < 0.0001$ and $\approx 59\% p < 0.0001$ reduction, respectively). *SOX6*-mediated repression caused downregulation of the *LIN28B*/Let-7 targets, including *MYC* and *IGF2BP1*, and rapidly blocks cell proliferation. Mechanistically, *Lin28B* repression is accompanied by *SOX6* physical binding within its locus, suggesting a direct mechanism of *LIN28B* downregulation that might contribute to the fetal/adult erythropoietic transition and restrict cancer proliferation.

Development is dictated by the timely integration of a complex network of genetic regulatory circuits. The identification of these fetal/adult switches is key to understand not only the developmental process but also the aberrant proliferation of cancer cells, which very often rely on the expression of fetal oncoproteins to sustain their proliferation¹.

LIN28B is an oncofetal gene whose expression is preferentially confined to early developmental stages in various tissues. Originally identified in *C. elegans*, where it controls the timing of larval development², *LIN28* is a RNA-binding protein encoded in mammals by two genes, *Lin28A* and *Lin28B*^{3,4}. In humans, *LIN28* has emerged as a Quantitative Trait Locus (QTL), influencing the timing of developmental processes^{5–8} and has been implicated in cell reprogramming⁹, tissue repair¹⁰ and cancer, where it is upregulated in a large number of human primary cancers and tumor cell lines^{11–18}.

A major role of *LIN28B* is to block the processing from precursor to mature miRNAs of Let-7 family, which promote differentiation by downregulating growth-promoting genes, including *c-MYC*, *RAS*, *IGF2BP1-3* and *HMGA2*¹⁹. This evidence suggests that the oncogenic role of *LIN28B* involves the de-repression of pro-proliferative Let-7 targets.

In hematopoiesis, *LIN28* expression is confined to fetal HSPCs cells. Its ectopic expression in adult cells promotes a fetal-like phenotype, characterized by active self-renewal and fetal-like lymphoid and myelo-erythroid differentiation^{20–23}. *LIN28B* expression in hematopoietic adult cells represses Let-7²⁴, the most upregulated miRNAs family in adult versus fetal erythroblasts²⁵. Interestingly, *LIN28B* overexpression is observed in specific subtypes of pediatric leukemia²⁶, including juvenile myelomonocytic leukemias of fetal origin (JMML), characterized by high fetal hemoglobin levels and considered to arise from a stem-progenitor cell²⁷, possibly of fetal origin^{24,28,29}.

¹Dipartimento di Biotecnologie e Bioscienze, Università degli Studi di Milano-Bicocca, Milan, Italy. ²Wallenberg Centre for Molecular Medicine, Linköping University, Linköping, Sweden. ³Division of Molecular Medicine and Virology, Department of Biomedical and Clinical Sciences, Faculty of Medicine and Health Sciences, Linköping University, Linköping, Sweden. ⁴Max-Planck-Institut für molekulare Genetik, Berlin, Germany. ⁵RIKEN BioResource Research Center, Tsukuba, Ibaraki, Japan. ⁶These authors contributed equally: Valentina Pastori and Gianluca Zambanini. ✉email: antonella.ronchi@unimib.it

SOX6 is a member of the HMG box family of Transcription Factors, preferentially expressed in adult cells³⁰. In hematopoiesis, SOX6 is expressed in quiescent LT-HSC³¹ and it acts as a tumor suppressor in chronic myeloid leukemia stem cells (CML-LSC)³². In erythropoiesis, it is expressed in adult cells, where it promotes differentiation^{33,34} and cooperates with the BCL11A-XL in the silencing of fetal γ -globin³⁵.

SOX6 downregulation is observed in different human cancers, including acute myeloid leukemias³⁶. Its low expression often correlates with poor prognosis. Restoring SOX6 expression mitigates the hyper-proliferative phenotype, indicating that it can act as tumor suppressor^{37–41}.

Here we show that SOX6 represses *LIN28B* in different human fetal-like erythroid cell types and in myeloid leukemia cell lines. The repression of *LIN28B* upon SOX6 transduction results in the repression of Let-7 targets, including *c-MYC*, and in arrested cell growth.

We find that the repression of *LIN28B* by SOX6 also occurs in tumor cell types representative of other tissues, such as SH-SY5Y neuroblastoma and HepG2 hepatoblastoma cells, where *LIN28B* is highly expressed^{13,14}. In all the contexts tested, the repression of *LIN28B* upon SOX6 overexpression invariably results in decreased cell growth and proliferation. Finally, we identify SOX6 consensus binding sites within the *LIN28B* genomic locus. ChIP-qPCR and CUT&RUN experimentally validate SOX6 physical occupancy of these sites, supporting a direct transcriptional repression of *LIN28B* by SOX6.

Overall, our data reveal a mechanism by which SOX6 can act as a tumor suppressor in *LIN28B*-positive cancer cells.

Materials and methods

Cell and cell cultures

Cell lines, culture media, supplements and conditions are listed in Supplementary table 1.

CD1 mice were housed in our animal facility with 12 h light and dark cycles and free access to water and chow. Females 8–14 week-old were mated with males from the same strain and inspected daily for vaginal plugs. The appearance of the vaginal plug was designated as day 0.5 of pregnancy. At days 11.5, 12.5 and 13.5 pregnant females were sacrificed by cervical dislocation. N = 3–5 females were sacrificed for each time point and embryos from the same litter were pooled for subsequent analysis. Bone Marrow (BM) was harvested from the sacrificed pregnant females. Mouse experiments were conducted under the approval of the Italian Ministry of Health (protocol number 358-2016/PR), in accordance with European Union (86/609/EEC) and ARRIVE guidelines.

Data analysis from publicly available single-cell RNA sequencing datasets

The single-cell RNA sequencing (scRNAseq) datasets, encompassing human fetal liver and yolk sac from ref Popescu et al.⁴², human neuroblastoma from Kildisiute et al.⁴³, and human hepatocellular carcinoma from Lu et al.⁴⁴, were publicly available and underwent preprocessing as outlined in the Table 1. Subsequently, Anndata objects were analyzed using Python packages Scanpy⁴⁵ (v.1.9.5) and Anndata (<https://github.com/scverse/anndata>) (v.0.9.2), inspired by Seurat⁴⁶, to visualize the expression of genes of interest. Cluster assignment of cell types was achieved utilizing cell labels and marker genes sourced from the respective papers. For the yolk sac dataset, neighborhood graphs were constructed with 10 neighbors and 40 principal components (PCs), followed by dimension reduction to two dimensions using UMAP as recommended in the Scanpy clustering pipeline.

	Fetal liver	Yolk sac	Neuroblastoma	Hepatocellular carcinoma
Species	Human	Human	Human	Human
Number of subjects	14	3	16	10
Cell count	113,063	10,071	13,281	16,498
Tissue / design	Fetal liver 7–17 PCW	4–6 PCW	Mainly pre-treated tumors, viable tumor areas	Tumor and adjacent liver, primary and relapsed tumor
scRNAseq Kit	10 × 3' v2		CEL-seq2	10 × 3' v2
Preprocessing	Cell ranger alignment on GRCh38 (STAR), Cells filtered for > 200 detected genes and total mtCount < 20%; Genes filtered for expressed in > 3 cells		Cell ranger alignment on GRCh38 (STAR), Cells filtered for > 200 detected genes and > 500 UMIs and total mtCount < 20%; Genes filtered for not mtGenes and not hspGenes	Cell Ranger alignment on GRCh38, Cells filtered for > 200 detected UMIs and > 200 / < 8000 detected genes and total mtCount < 10%
Normalization	By sequencing depth scale to 10,000 counts (NormalizeData, LogNormalize method), data feature scaling, variable gene detection, PCA, Louvain graph-based clustering with a resolution of 30 with standard parameters (fetal liver only, Seurat)		By sequencing depth scaled to 10 000 counts (NormalizeData, LogNormalize method), Scaling, variable gene detection (most 2000), PCA of 2000 most variable genes, 50 first PCs were used to calculate a UMAP (resolution parameter 1) (Seurat)	Seurat standard normalization pipeline, Shared-nearest neighbor clustering obtained a final of 53 clusters that were used to calculate a UMAP (Resolution parameter 3) (Seurat)
Download source	https://developmental.cellatlas.io/fetal-liver		http://neuroblastomacellatlas.org	http://omic.tech/scrna-hcc/
Download format	Anndata .h5ad format containing count matrix and meta-data		Anndata .h5ad format count matrix and metadata	Anndata .h5ad format count matrix and metadata
Reference	Popescu et al. ⁴² (Nature)		Kildisiute et al. ⁴³ (Science Advances)	Lu et al. ⁴⁴ (Nature Communications)
Raw data	E-MTAB-7407 (Array Express)		EGAD00001008345	EGAC00001001616

Table 1. Single-cell RNA sequencing data sets and processing.

RNA isolation and real time PCR

Total RNA from $\geq 10^5$ cells were purified with TRIzol Reagent (Euroclon) and retrotranscribed (High Capacity cDNA Reverse Transcription Kit, Applied Biosystem). RT-PCR analysis was performed using StepOne (ThermoFisher). Specific PCR product accumulation was monitored by using SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad) fluorescent dye in 12 μ l reaction volume. Dissociation curves confirmed the homogeneity of PCR products. Primers, designed to amplify 150 to 300bp amplicons, are listed in the Supplementary table 3. Each experiment was done in three biological replicates and cDNAs from each replicate were analysed in technical triplicates.

Protein extracts

For whole protein extracts, cells were resuspended in RIPA buffer (20 mM Tris HCl pH 7.4, 137 mM NaCl, 10% glycerol, 0.1% SDS, 0.5% deoxycolate, 1% Triton X-100, 2mM EDTA and proteases inhibitor cocktail). Lysis was performed for 30 min in ice and, after centrifugation (15 min at 16,000 \times g at 4 °C) the supernatant was collected and analysed.

Immunoblotting

Protein extracts were resolved by SDS/PAGE in a 10% acrylamide gel and blotted onto Hybond-ECL Nitrocellulose membrane (GE healthcare) at 400 mA for 90 min at 4 °C (Biorad Transblot apparatus). Membranes were blocked for 1 h at RT with milk 5% in TBS-T (Tris Buffered Saline, pH 7.6 and 0,1% Tween20) and incubated with the appropriate primary antibody diluted in milk 5% TBS-T overnight at 4 °C. Membranes were washed in TBS-T and incubated with the appropriate HRP-conjugated secondary antibody (in milk 5% TBS-T) for 1 h at room temperature. Antibodies binding was detected by ECL (Millipore). The blots were cut prior to hybridisation with the indicated antibodies. Full size gels are shown in Supplementary Fig. 10.

Lentiviral particles production

The Sox6 murine cDNA was cloned in frame with a FLAG epitope immediately upstream to the IRES-eGFP cassette into the CSI vector, as in³⁴. Packaging HEK-293T cells were transfected with the SOX6 expression vector (or the corresponding empty control Vector, EV) and with the psPAX2 and pMD-VSVG packaging vectors. 72 h upon transfection, viral pseudoparticles were concentrated by centrifugation at 4 °C. The viral pellet was resuspended in 1X PBS and aliquoted at –80 °C. 72h upon lentiviral transduction (MOI \geq 25) GFP + cells were scored by Flow Cytometry (FC) (CytoFlex, Beckman Coulter). For HUDEP1 and HepG2 cells where GFP positivity was about 60–70%, cells were sorted (BD FACSMelody, BD Biosciences) before further analysis. All experiments were performed on cells \geq 90% GFP + (Supplementary Fig. 1).

CUT&RUN analysis

The detailed protocol for CUT&RUN on HUDEP1 overexpressing SOX6-Flag and for data analysis are detailed in the Supplementary file material. Sox6 peaks are listed in suppl. table 4. The CUT&RUN datasets (raw and processed files) have been deposited at ArrayExpress (<https://www.ebi.ac.uk/arrayexpress/>) under accession number E-MTAB-12800.

Chromatin IP

HEL cells overexpressing SOX6 were fixed with 0.4% formaldehyde for 10 min at RT. Chromatin was sonicated (Bioruptor®, Diagenode) to the size of \approx 500bp. DNA immunoprecipitation was obtained by incubation with the appropriate antibodies and subsequent isolation with protein A-agarose beads (Upstate). Immunoprecipitated DNA fragments were analysed by qPCR for the regions of interest (primers are listed in Supplementary Table 3). Full size gels are shown in Supplementary Fig. 9.

Primers, antibodies and reagents

Antibodies, reagents and primers are listed in the Supplementary Tables 2 and 3.

Statistical analysis

All experiments were performed on a least of $n \geq 3$ biological replicates, each analysed in technical triplicates. For Western Blots, statistics was calculated on densitometry values from 3 independent experiments obtained by using ImageJ software. Statistical analysis was performed by using GraphPad Prism version 8.0.0 for Windows, GraphPad Software (www.graphpad.com). Data were analysed by using a two-tailed unpaired t-test and are expressed as mean \pm SEM (standard error of mean). In Figures, error bars represent SEM and p-values are represented as follows: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. The numerical values are in Figure legends.

Figures were prepared in compliance with the Scientific Reports digital image and integrity policies.

Results

Sox6 and Lin28B show opposite expression profiles during the fetal to adult transition in erythroid cells

Erythropoiesis represents a model system of developmental dynamics. In particular, in the mouse fetal liver (FL), the E11.5–E13.5 timeframe corresponds to the transition from embryo/fetal to adult erythropoiesis, marked by the switching from embryo/fetal to adult globin genes expression (Fig. 1a). RT-qPCR shows that, within this developmental window, the expression of Sox6 and Lin28B inversely correlate: the highest expression of Sox6 at E13.5 coincides with the drop in Lin28B and with the establishment of the adult pattern of globin expression.

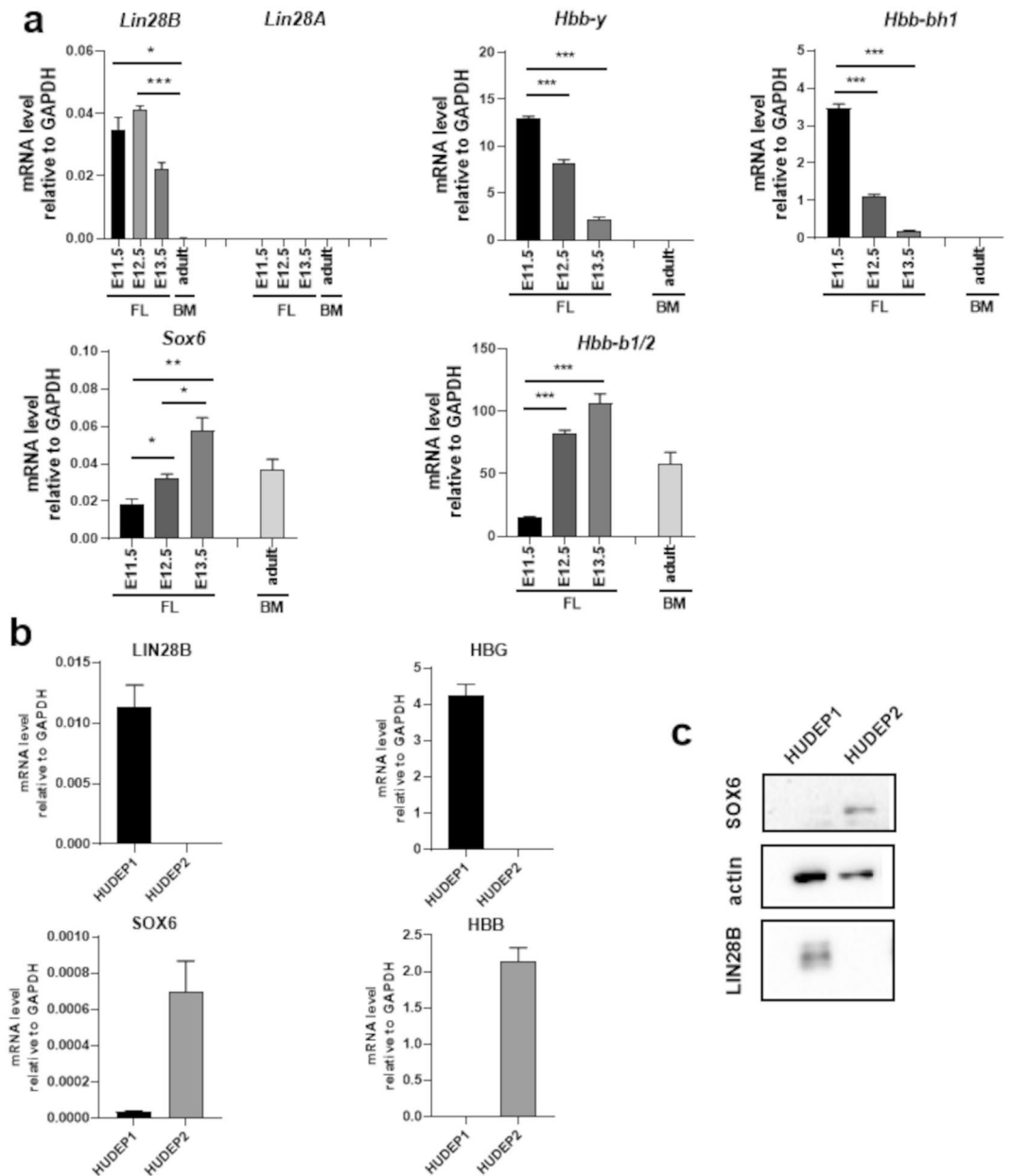


Figure 1. Expression levels of *LIN28* and *SOX6* in erythropoiesis during development. **(a)** RT-qPCR: *Sox6*, *Lin28B* and globins expression relative to *GAPDH* in E11.5-E13.5 mouse fetal liver (FL) and in adult bone marrow (BM). The decline in *Lin28B* and the opposite increase in *Sox6* parallel the transition (hemoglobin switching) from embryo/fetal (*Hbb-y* and *Hbb-bh1*) to adult (*Hbb-b1/2*) mouse globins. In the Figure, * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$. *Lin28B* and *Sox6* fold changes calculated on mean values: *Lin28B* fold reduction E12.5-E13.5 $\approx 45\%$ $p = 0.0003$, *Sox6* fold increase E12.5-E13.5 $\approx 45\%$ $p = 0.012$. **(b)** RT-qPCR and Western Blot showing the opposite expression of *LIN28B* and *SOX6* and human fetal γ (*HBG*) and adult β (*HBB*) globins in human fetal-like HUDEP1 and adult HUDEP2 cells. **(c)** Western blot showing the relative protein levels.

Consistently, in adult bone marrow (BM) erythropoiesis where *Sox6* expression persists throughout adult life, *Lin28B* is absent (Fig. 1a). *Lin28B* and *Sox6* expression profiles, both at the mRNA and protein level, also inversely correlate in human HUDEP1 and HUDEP2 cells⁴⁷, representative of embryo/fetal and adult erythropoiesis, respectively (Fig. 1b and c). The *Lin28B* paralog, *Lin28A*, is not expressed (Fig. 1a).

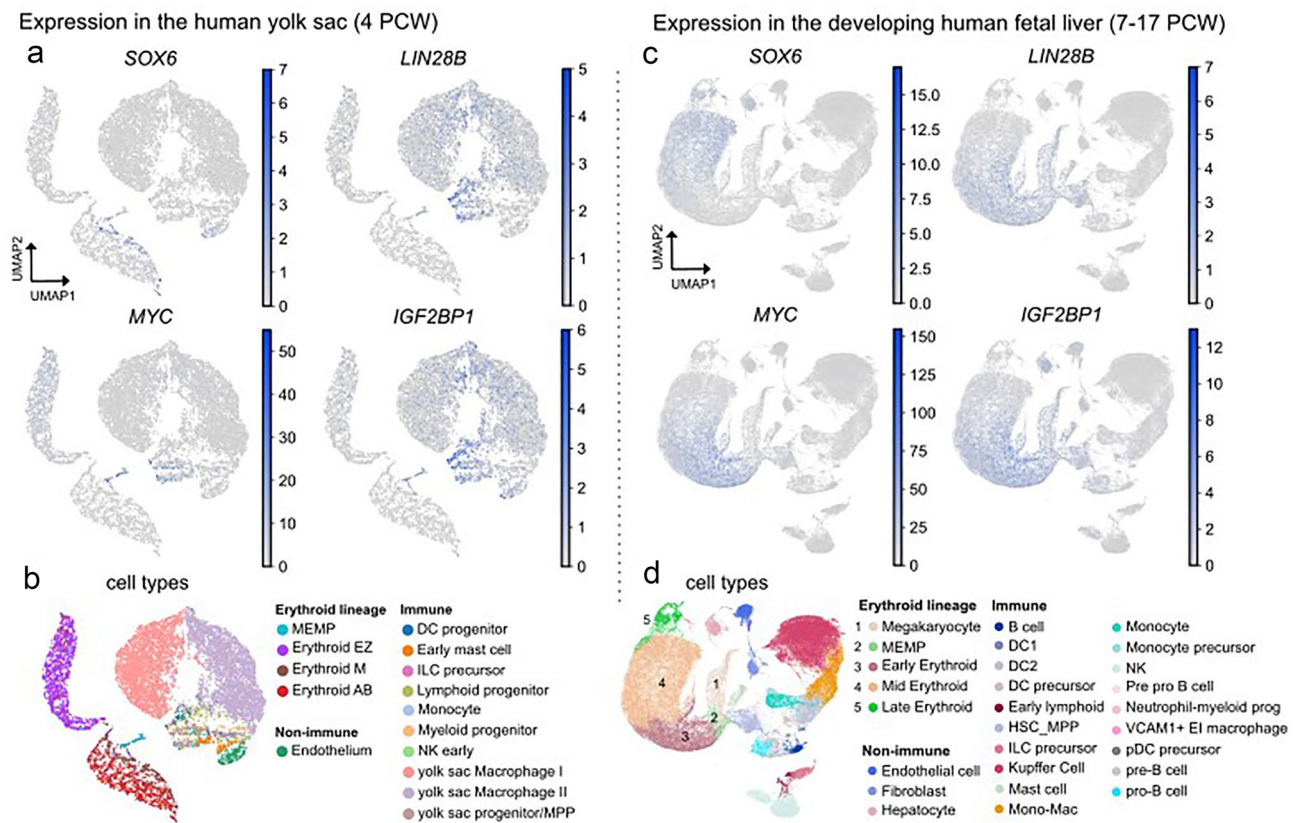


Figure 2. Single-cell RNA expression of the SOX6-LIN28B/Let-7 axis in human fetal hematopoiesis. (a) Normalized expression of *SOX6*, *LIN28B*, as well as the downstream genes *MYC*, and *IGF2BP1* in human yolk sac cells shown as UMAP. (b) 15 UMAP clusters are associated with cell types according to the expression of distinct marker genes from Popescu et al.⁴². Dataset contains 10,071 yolk sac cells derived from 3 individuals at the age of 4–6 PCW. (c) Normalized expression of *SOX6*, *LIN28B*, as well as the downstream genes *MYC*, and *IGF2BP1* in human fetal liver cells shown as UMAP (Uniform Manifold Approximation and Projection) plots. (d) 27 UMAP clusters are associated with cell types according to the expression of distinct marker genes. Dataset contains 113,063 fetal liver cells derived from 16 individuals at the age of 7–17 PCW, post conception week; *MEMP*, megakaryocyte-erythroid progenitors; *DC*, dendritic cells; *HSC-MPP*, hematopoietic stem cells and multipotent progenitors; *ILS*, induced leucocyte stem; *Erythroid EZ*, early (HGE expressing) erythroid; *Erythroid M*, mid erythroid; *Erythroid AB*, late (HGA and HGB expressing) erythroid.

SOX6 and LIN28B/Let-7 downstream genes are expressed in non-overlapping cell populations during human embryo/fetal hematopoiesis

To investigate the relation between *SOX6* and *LIN28B* expression in human early development at a single-cell resolution, we analyzed single-cell RNA sequencing (scRNA-seq) datasets obtained from yolk-sac (4 PCW, post conception weeks) and fetal liver (7 to 17 PCW) cells⁴². The integrated data analysis at these stages shows that the expression of *SOX6* and *LIN28B* is confined to distinct cell populations (Fig. 2 and Supplementary Fig. 2): whereas *LIN28B* and the LIN28B/Let-7 downstream targets *MYC* and *IGF2BP1* are prominently expressed in early erythroid progenitors (*MEMP*), *SOX6* raises in mid-late erythroid cells and this is paralleled by loss of *LIN28B* (together with *MYC* and *IGF2BP1*). Complementing the data in Fig. 1, the expression pattern of *SOX6* and *LIN28B* during human fetal erythropoiesis at a single cell level strongly suggests a potential switch-off mechanism of *LIN28B* transcription by *SOX6*.

SOX6 represses Lin28B blocking the LIN28B/Let-7 axis and cell proliferation in different erythroid cellular models

The LIN28B/Let-7 axis, which represents the main LIN28B downstream pathway, ultimately results in the control of proliferative genes, including *MYC* and *IGF2BP1*¹⁹. Our observations (Figs. 1 and 2) suggest that *SOX6* may interfere on the LIN28B/Let-7 axis. To test if *SOX6* is causally involved in the transcriptional regulation of *LIN28B*, we overexpressed it and monitor how *LIN28B* mRNA expression is affected as a consequence. To this aim, we selected three different “fetal-like” human erythroid cells lines: K562 and HEL erythroleukemias and HUDEP1, which model human embryonic erythropoiesis (Supplementary Fig. 3).

We infected the above cell lines with viral pseudoparticles carrying a *SOX6*Flag-IRES-GFP overexpressing vector (*SOX6*) and, in parallel, with the corresponding empty control vector (EV), (Supplementary Fig. 1). Upon *SOX6* overexpression, *LIN28B* is significantly downregulated in all the tested cell lines, both at the mRNA and

ERYTHROID CELLS

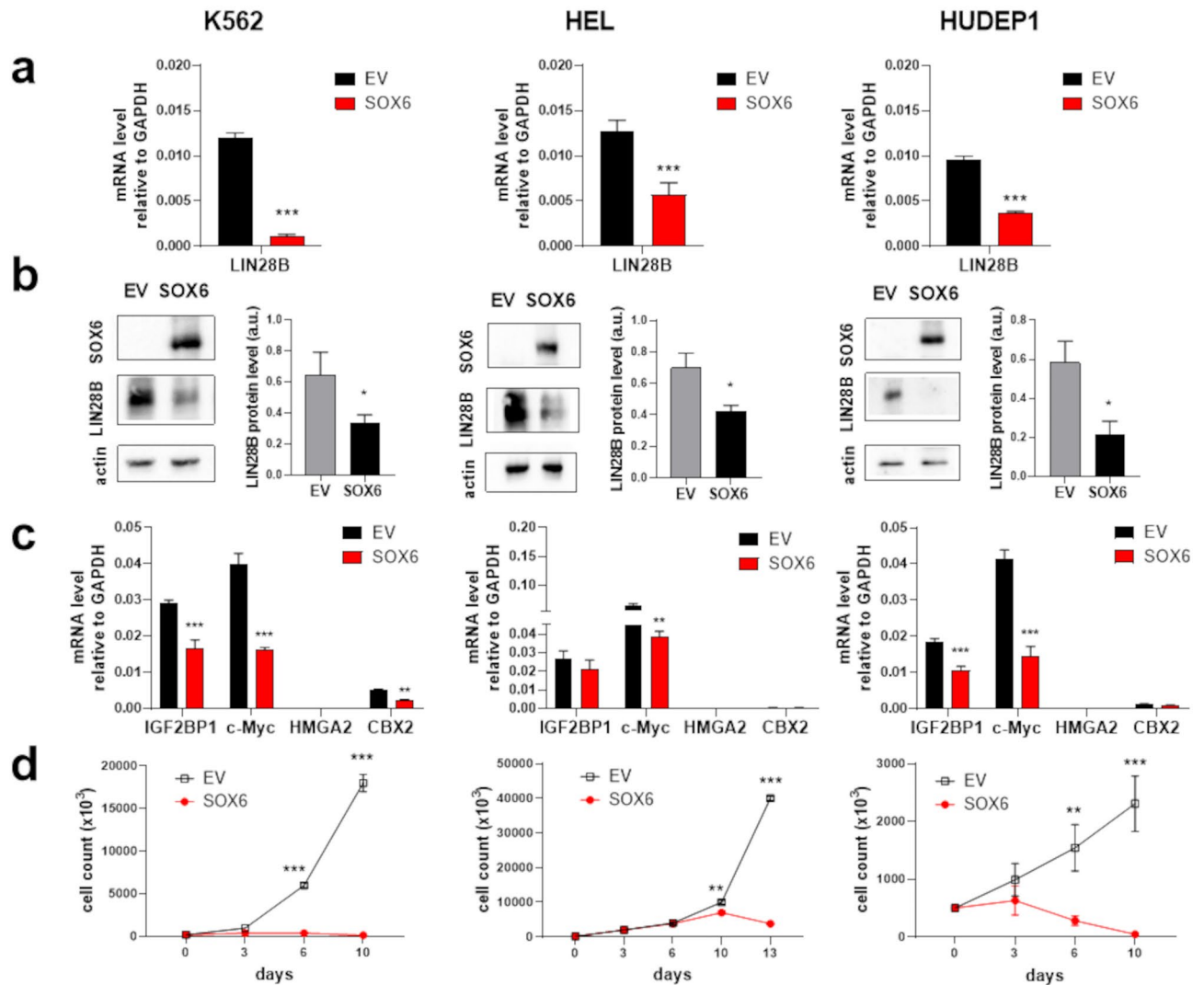


Figure 3. SOX6 represses *LIN28B* and blocks proliferation in erythroid cells. **(a)** Effect of SOX6 expression on *LIN28B* mRNA level. SOX6: cells transduced with the SOX6 overexpressing vector; Empty Vector (EV): cells transduced with the corresponding empty vector. Fold reduction calculated on the mean values: K562 $\approx 92\%$ $p < 0.0001$, HEL $\approx 54\%$ $p = 0.0009$, HUDEP1 $\approx 60\%$ $p < 0.0001$. **(b)** Representative Western blots and quantification of the *LIN28B* reduction upon SOX6 overexpression at the protein level (densitometry analysis from 3 independent Western blot experiments was performed by using the ImageJ software). Fold reduction calculated on the mean values: K562 $\approx 48\%$ $p = 0.03$, HEL $\approx 40\%$ $p = 0.0281$, HUDEP1 $\approx 63\%$ $p = 0.046$. **(c and d)** Effects of SOX6 expression on representative Let-7 downstream targets and on cell proliferation. Empty Vector (EV), in black; cells transduced with the empty vector. SOX6, in red; cells transduced with the SOX6 expressing vector. Fold reduction of targets calculated on the mean values: K562 *IGF2BP1* $\approx 42\%$ $p = 0.0005$, K562 *c-MYC* $\approx 60\%$ $p < 0.0001$, K562 *CBX2* $\approx 60\%$ $p = 0.0058$, HEL, *c-MYC* $\approx 40\%$ $p = 0.0045$. HUDEP1 *IGF2BP1* $\approx 39\%$ $p = 0.0006$, HUDEP1 *c-MYC* $\approx 64\%$ $p = 0.0008$. For all RT-PCR shown, $n \geq 4$, error bars: standard error of mean. * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

at the protein level (Fig. 3a, b). In order to evaluate the secondary impact of *LIN28B* repression on the *LIN28B*/Let-7 axis, we selected different known Let-7 target genes (Fig. 3c), including *c-MYC*, *HMG2*, *IGF2BP1* and *CBX2*^{10,14,17}. Among them, *HMG2* is not expressed in all the three cell lines, *CBX2* is expressed at low level in K562 and HUDEP1 cells and *c-MYC* and *IGF2BP1* are highly expressed in all three cell lines. *IGF2BP1*, which belongs to the family of the oncofetal IGF2 mRNA binding proteins, is highly expressed during embryogenesis and overexpressed in various tumors^{48,49}. In erythropoiesis, *IGF2BP1* is a major target of the Let-7 miRNA family^{19,50}, expressed in fetal erythroblasts²⁵ and capable of inducing γ -globin expression^{51,52}. As shown in Fig. 3c, *c-MYC*, *IGF2BP1* and *CBX2* (in K562) expression is reduced by SOX6 overexpression. The overexpression of SOX6 results in a marked decline in cell proliferation (Fig. 3d). We previously showed that SOX6 blocks cell proliferation in Set-2 and Uke1 myeloid cancer cells³⁶. Here we show that SOX6 overexpression downregulates the *LIN28B*/Let-7 pathway also in these cells (Supplementary Fig. 4). Together, these data confirm the ability

of SOX6 to repress *LIN28B* and the pro-proliferative genes downstream to the LIN28/Let-7 axis, pointing to a wide tumor suppressor role of SOX6.

***LIN28B* repression by SOX6 results in the repression of Let-7 targets in non-hematopoietic *LIN28B*-positive cancer cells**

LIN28B is broadly expressed in early development and is reactivated in cancers¹⁷. We analyzed *LIN28B* expression in the large panel of comprehensively characterized Cancer Cell Line Encyclopedia (CCLE, Broad Institute, <https://sites.broadinstitute.org/ccle/>). Consistently with a broad role in cancer, *LIN28B* is found highly expressed in a wide spectrum of cancer cell lines, including myeloid, liver, and central nervous system cells, compared to other tissues (Supplementary Fig. 5a). Moreover, *LIN28B* is overexpressed in both primary as well as metastatic cell lines (Supplementary Fig. 5c).

To test whether high *LIN28B* gene expression correlates with genetic dependency, we analyzed dependencies scores calculated by Chronos for *LIN28B* CRISPR in CCLE cell lines⁵³. According to DepMap (<https://depmap.org/portal/>), *LIN28B* expression and dependency varied across cell lines. In particular, neuroblastoma cell lines, including SH-SY5Y, show high expression and significant dependency on *LIN28B* (Supplementary Fig. 5e, f). This is in line with the known involvement of *LIN28B* in high-risk neuroblastoma^{13,54}. In addition, in neuroblastoma cell lines, SOX6 and *LIN28B* expression negatively correlate (Supplementary Fig. 5e, f). We therefore assessed the effect of SOX6 overexpression in human neuroblastoma SH-SY5Y cells (Fig. 4). To expand our observation, we performed the same analysis on human hepatic cancer HepG2 cells (Fig. 4) in which *LIN28B* was demonstrated by functional experiments to be essential for proliferation^{11,14}.

In both cell types, we observed that overexpression of SOX6 represses *LIN28B* and ultimately results in a marked reduction of proliferation, although with slightly different kinetics. In these cells, Let-7 targets are repressed downstream to SOX6 overexpression.

Finally, we examined single-cell RNA-seq datasets obtained from neuroblastoma and hepatocellular carcinoma human biopsies (Supplementary Fig. 6). In neuroblastoma, the subset of cells expressing SOX6 shows minimal overlap with those expressing *LIN28B*, supporting a SOX6 mediated *LIN28B* downregulation in this context (Supplementary Fig. 6a, b). In hepatocellular carcinomas, the relationship between SOX6 and *LIN28B* expression is less clear indicating more intricate regulatory dynamics (Supplementary Fig. 6c, d).

***LIN28B* is a direct transcriptional target of SOX6**

Our data points to *LIN28B* as a candidate gene transcriptionally repressed by SOX6. To test this hypothesis, we first searched for genomic binding sites of SOX6 across the *LIN28B* genomic locus. We found several SOX6 binding motifs within the *LIN28B* promoter and third intron (Fig. 5a and b, Supplementary Fig. 7). To test the ability of SOX6 to bind these sequences in vivo we performed a CUT&RUN analysis targeting the exogenously expressed SOX6-Flag in HUDEP1 (Supplementary Fig. 8). Untransduced cells were used as negative control. As shown in Fig. 5a, we identified reproducible signal enrichment in the vicinity of *LIN28B* transcriptional start site (TSS), in an evolutionarily conserved region enriched by H3K27Ac, a marker of active regulatory regions. This result constitutes robust evidence that SOX6 physically associates with this region (Fig. 5a).

Targeted chromatin immunoprecipitation (ChIP) followed by qPCR in HEL cells (Fig. 5c and d) validates the CUT&RUN findings and provides evidence for multiple binding anchors for SOX6 within the *LIN28B* locus in different cell types, suggesting robust, potentially redundant regulatory mechanisms of *LIN28B* transcription repression by SOX6.

Discussion

LIN28B is a stemness-related *oncofetal* gene whose repression is required to acquire adult-specific fate in different systems^{20–23,53,54}. Its reactivation is observed in several tumors¹⁷ and often coincide with the acquisition of fetal-like phenotypes^{23,24}. Despite its broad role in development and cancer, the molecular basis of its transcriptional repression is largely unknown. In this work we identify the pro-differentiative, growth-restriction SOX6 “adult” transcription factor^{31,54} as a direct repressor of the *LIN28B* gene.

First we show that *LIN28B* and *Sox6* expression is mutually exclusive in mouse and human erythropoiesis, which represents a well-defined model of development (Figs. 1 and 2) as well in several cancer cells (Figs. 3, 4 and Supplementary Figs. 5 and 6), in particular neuroblastoma and hepatocarcinoma cell lines. SOX6 binds in vivo to several consensus binding sites that we identified within the *LIN28B* locus (Fig. 5 and Supplementary Fig. 7), represses *LIN28B* transcription and cell proliferation (Figs. 3, 4).

At the molecular level, the *LIN28B* oncogenic potential largely relies in its ability to repress the processing of miRNAs of the Let-7 family, which in adult cells promote differentiation and repress pro-proliferative genes¹⁷. We demonstrate that SOX6 impacts on the expression of genes downstream to the LIN28B/Let-7 axis (Figs. 2, 3, 4), such as *c-MYC* and *IGF2BP1*, confirming the involvement of this pathway. Based on our observations, we propose a general model in which SOX6 acts as a tumor suppressor in cells of different tissue origin acting through the LIN28B/Let-7 pathway (Fig. 6).

We cannot exclude that additional molecular mechanisms/pathways downstream to SOX6, independent from *LIN28B* downregulation concur to the observed tumor suppressive outcome and it is also likely that reduction of *LIN28B* expression does not exclusively impact the Let-7 axis. Nevertheless, we believe that the genetic circuit involving the direct repression of *LIN28B* by SOX6 represents a novel layer of regulation with relevant therapeutic implications for those tumors relying on *LIN28B* for their survival, often associated with high-risk malignancy.

The identification of the mechanistic link between SOX6 and *LIN28B* may provide novel opportunities for tumors stratification and novel therapeutic options.

NON HEMATOPOIETIC CANCER CELLS

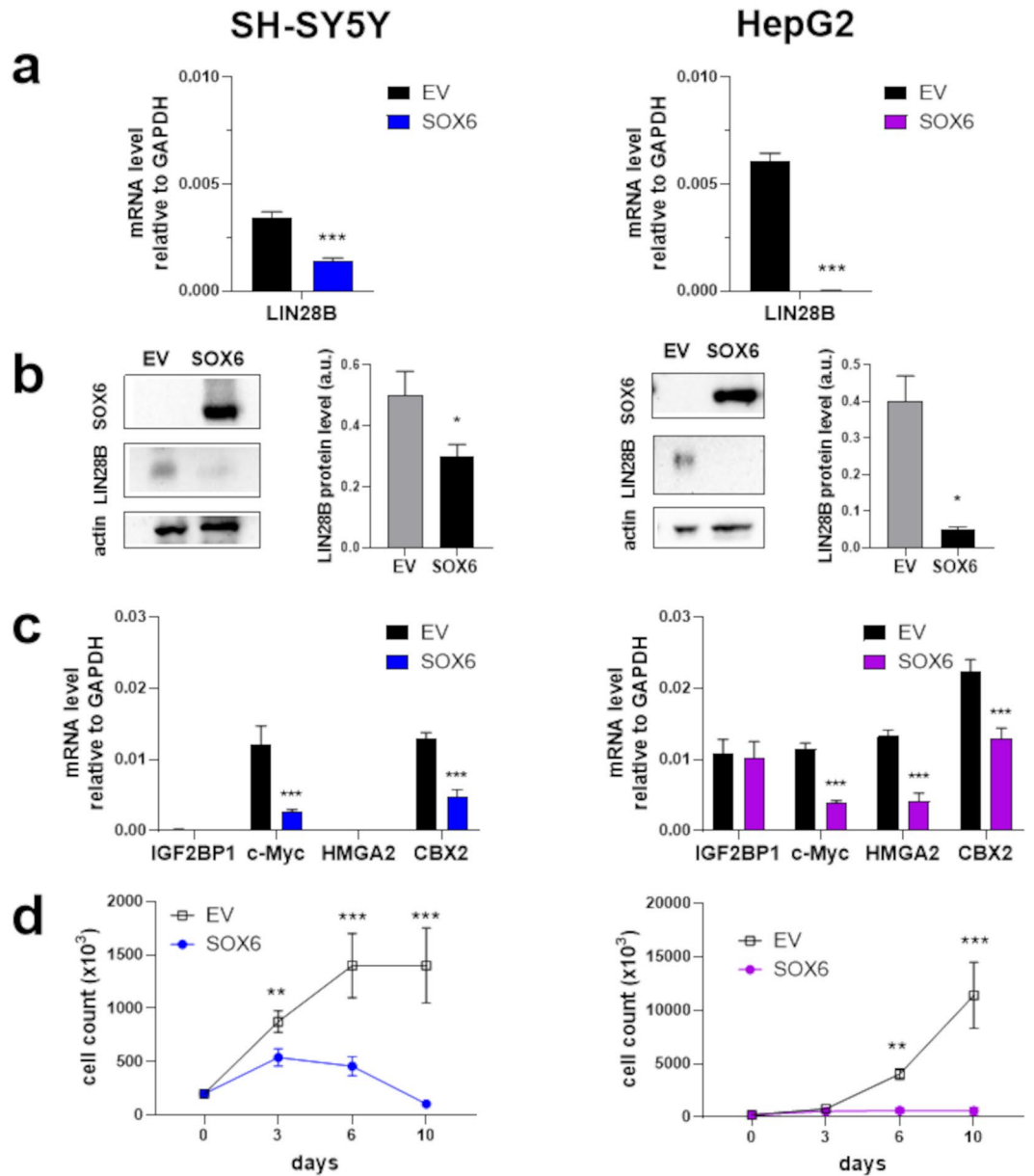


Figure 4. SOX6 represses *LIN28B* and blocks proliferation in neuroblastoma SH-SY5Y and hepatoblastoma HepG2 cells. The same analysis as in Fig. 3 was carried out in two non-hematopoietic cancer cells whose growth is sustained by *LIN28B*. SOX6: cells transduced with the SOX6 overexpressing vector; Empty Vector (EV): cells transduced with the corresponding empty vector. (a) mRNA fold reduction calculated on the mean values: SH-SY5Y $\approx 59\%$ $p < 0.0001$, HepG2 $\approx 99.7\%$ $p < 0.0001$. (b) Protein fold reduction calculated on the mean values: SH-SY5Y $\approx 40\%$ $p < 0.03$ HepG2 $\approx 88\%$ $p < 0.02$. (c) Fold reduction of targets calculated on the mean values: SH-SY5Y *c-MYC* $\approx 75\%$ $p < 0.0001$, SH-SY5Y *CBX2* $\approx 62\%$ $p = 0.0004$, HepG2 *c-MYC* $\approx 63\%$ $p < 0.0001$, HepG2 *HMGA2* $\approx 70\%$ $p = 0.0008$, HepG2 *CBX2* $\approx 41\%$ $p = 0.0007$. (d) Cell proliferation curves.

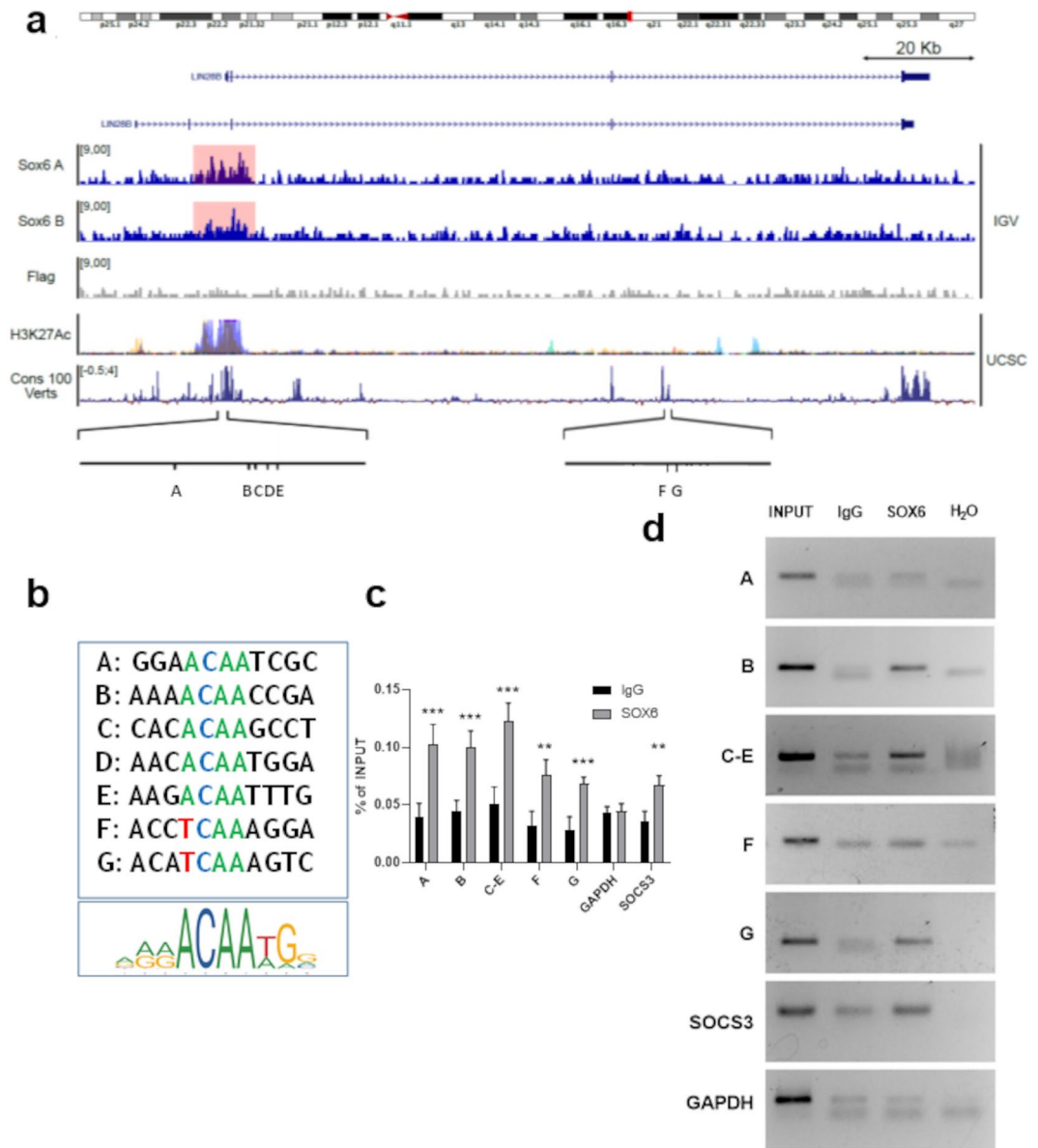


Figure 5. Identification of the SOX6 binding sites within the *LIN28B* locus. **(a)** *Lin28B* genomic locus visualized in Integrative Genomic Viewer (IGV) and with the UCSC genome browser (<https://genome.ucsc.edu/>). From the top: chromosome representation (IGV), gene transcripts (UCSC), two SOX6 CUT&RUN replicates and the respective negative Flag control in HUDEP1 cells (IGV), H3K27ac overlap (UCSC), vertebrate conservation (UCSC) and positions of the SOX6 binding sites within the promoter (A–E) and the third intron (F–G). **(b)** SOX6 sites compared with the SOX6 consensus Jaspar matrix 515.1. **(c)** Chromatin IP results obtained in HEL cells expressing SOX6. *GAPDH* locus: negative control; *SOCS3* promoter: positive control³⁴. Histograms represent the results of three biological replicates (n = 3, each of them in three technical replicates, as analyzed by RTqPCR (Error bars: standard error of mean; *p < 0.05; **p < 0.01; ***p < 0.001). **(d)** representative ChIP result. Antibodies are listed in the Supplementary table 2.

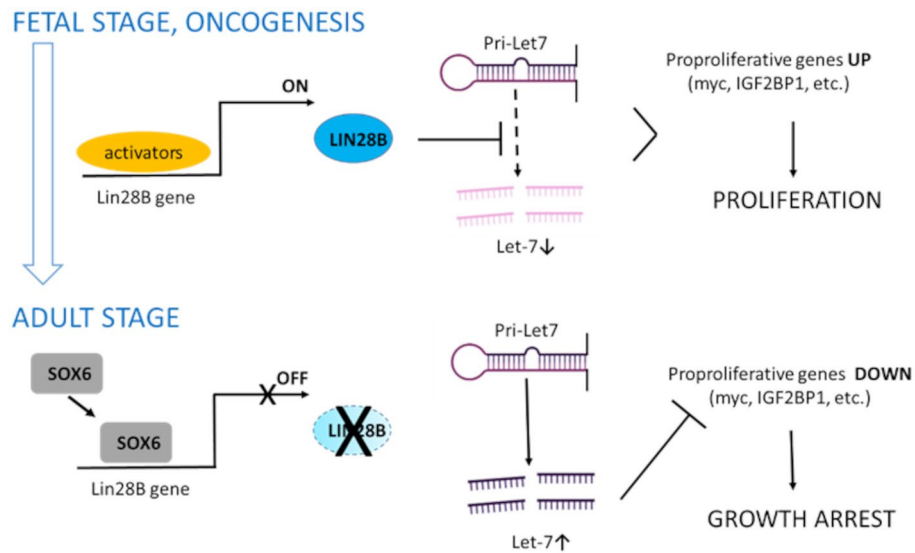


Figure 6. Proposed model of *LIN28B* repression by SOX6. In fetal cells and in cancers with illegitimate *LIN28B* reactivation, LIN28B blocks the processing of Let-7 precursors to Let-7 miRNA. Let-7 reduction leads to the failure of silencing of proliferative genes, including *c-MYC* and *IGF2BP1*, thus promoting cell proliferation. SOX6, by repressing *LIN28B* transcription, allows the maturation of Let-7 miRNAs, among the most abundant miRNAs in adult cells¹⁷. Let-7 miRNAs post-transcriptionally repress proliferative genes inducing growth arrest (created with Biorender).

Data availability

All data have been deposited in ArrayExpress under accession number E-MTAB-12800.

Received: 19 October 2023; Accepted: 23 April 2024

Published online: 04 May 2024

References

- Marshall, G. M. *et al.* The prenatal origins of cancer. *Nat. Rev. Cancer* **14**(4), 277–289 (2014).
- Moss, E. G., Lee, R. C. & Ambros, V. The cold shock domain protein LIN-28 controls developmental timing in *C. elegans* and is regulated by the lin-4 RNA. *Cell* **88**(5), 637–646 (1997).
- Moss, E. G. & Tang, L. Conservation of the heterochronic regulator Lin-28, its developmental expression and microRNA complementary sites. *Dev. Biol.* **258**(2), 432–442 (2003).
- Guo, Y. *et al.* Identification and characterization of lin-28 homolog B (*LIN28B*) in human hepatocellular carcinoma. *Gene* **384**, 51–61 (2006).
- Zhu, H. *et al.* Lin28a transgenic mice manifest size and puberty phenotypes identified in human genetic association studies. *Nat. Genet.* **42**(7), 626–630 (2010).
- He, C. *et al.* Genome-wide association studies identify loci associated with age at menarche and age at natural menopause. *Nat. Genet.* **41**(6), 724–728 (2009).
- Ong, K. K. *et al.* Genetic variation in *LIN28B* is associated with the timing of puberty. *Nat. Genet.* **41**(6), 729–733 (2009).
- Sulem, P. *et al.* Genome-wide association study identifies sequence variants on 6q21 associated with age at menarche. *Nat. Genet.* **41**(6), 734–738 (2009).
- Yu, J. *et al.* Induced pluripotent stem cell lines derived from human somatic cells. *Science* **318**(5858), 1917–1920 (2007).
- Shyh-Chang, N. *et al.* Lin8 enhances tissue repair by reprogramming cellular metabolism. *Cell* **155**(4), 778–792 (2013).
- Nguyen, L. H. *et al.* Lin28b is sufficient to drive liver cancer and necessary for its maintenance in murine models. *Cancer Cell* **26**(2), 248–261 (2014).
- Zhou, J., Ng, S. B. & Chng, W. J. *LIN28/LIN28B*: an emerging oncogenic driver in cancer stem cells. *Int. J. Biochem. Cell Biol.* **45**(5), 973–978 (2013).
- Molenaar, J. J. *et al.* *LIN28B* induces neuroblastoma and enhances *MYCN* levels via *let-7* suppression. *Nat. Genet.* **44**(11), 1199–1206 (2012).
- Viswanathan, S. R. *et al.* Lin28 promotes transformation and is associated with advanced human malignancies. *Nat. Genet.* **41**(7), 843–848 (2009).
- Zhou, J. *et al.* *LIN28B* activation by PRL-3 Promotes leukemogenesis and a stem cell-like transcriptional program in AML. *Mol. Cancer Res.* **15**(3), 294–303 (2017).
- Manier, S. *et al.* The *LIN28B/let-7* axis is a novel therapeutic pathway in multiple myeloma. *Leukemia* **31**(4), 853–860 (2017).
- Balzeau, J., Menezes, M. R., Cao, S. & Hagan, J. P. The *LIN28/let-7* pathway in cancer. *Front. Genet.* **8**, 31 (2017).
- Gewalt, T., Noh, K. W. & Meder, L. The role of *LIN28B* in tumor progression and metastasis in solid tumor entities. *Oncol. Res.* **31**(2), 101–115 (2023).
- Busch, B. *et al.* The oncogenic triangle of *HMGA2*, *LIN28B* and *IGF2BP1* antagonizes tumor-suppressive actions of the *let-7* family. *Nucl. Acids Res.* **44**(8), 3845–3864 (2016).
- Copley, M. R. *et al.* The *Lin28b-let-7-Hmga2* axis determines the higher self-renewal potential of fetal haematopoietic stem cells. *Nat. Cell Biol.* **15**(8), 916–925 (2013).
- Rowe, R. G. *et al.* Developmental regulation of myeloerythroid progenitor function by the *Lin28b-let-7-Hmga2* axis. *J. Exp. Med.* **213**(8), 1497–1512 (2016).

22. Wang, D. *et al.* Developmental maturation of the hematopoietic system controlled by a Lin28b-let-7-Cbx2 axis. *Cell Rep.* **39**(1), 110587 (2022).
23. Yuan, J., Nguyen, C. K., Liu, X., Kanellopoulou, C. & Muljo, S. A. Lin28b reprograms adult bone marrow hematopoietic progenitors to mediate fetal-like lymphopoiesis. *Science* **335**(6073), 1195–1200 (2012).
24. Lee, Y. T. *et al.* LIN28B-mediated expression of fetal hemoglobin and production of fetal-like erythrocytes from adult human erythroblasts *ex vivo*. *Blood* **122**(6), 1034–1041 (2013).
25. Lessard, S., Beaudoin, M., Orkin, S. H., Bauer, D. E. & Lettre, G. 14q32 and let-7 microRNAs regulate transcriptional networks in fetal and adult human erythroblasts. *Hum. Mol. Genet.* **27**(8), 1411–1420 (2018).
26. Helsmoortel, H. H. *et al.* LIN28B is over-expressed in specific subtypes of pediatric leukemia and regulates lncRNA H19. *Haematologica* **101**(6), e240–e244 (2016).
27. Cooper, L. J. *et al.* Evidence that juvenile myelomonocytic leukemia can arise from a pluripotential stem cell. *Blood* **96**(6), 2310–2313 (2000).
28. Stieglitz, E. *et al.* The genomic landscape of juvenile myelomonocytic leukemia. *Nat. Genet.* **47**(11), 1326–1333 (2015).
29. Helsmoortel, H. H. *et al.* LIN28B overexpression defines a novel fetal-like subgroup of juvenile myelomonocytic leukemia. *Blood* **127**(9), 1163–1172 (2016).
30. Hagiwara, N. Sox6, jack of all trades: A versatile regulatory protein in vertebrate development. *Dev. Dyn.* **240**(6), 1311–1321 (2011).
31. Forsberg, E. C. *et al.* Differential expression of novel potential regulators in hematopoietic stem cells. *PLoS Genet.* **1**(3), e28 (2005).
32. Zhou, M. *et al.* Targeting protein lysine methyltransferase G9A impairs self-renewal of chronic myelogenous leukemia stem cells via upregulation of SOX6. *Oncogene* **40**(20), 3564–3577 (2021).
33. Dumitriu, B. *et al.* Sox6 cell-autonomously stimulates erythroid cell survival, proliferation, and terminal maturation and is thereby an important enhancer of definitive erythropoiesis during mouse development. *Blood* **108**(4), 1198–1207 (2006).
34. Cantu, C. *et al.* Sox6 enhances erythroid differentiation in human erythroid progenitors. *Blood* **117**(13), 3669–3679 (2011).
35. Xu, J. *et al.* Transcriptional silencing of $\{\gamma\}$ -globin by BCL11A involves long-range interactions and cooperation with SOX6. *Genes Dev.* **24**(8), 783–798 (2010).
36. Barbarani, G., Fugazza, C., Barabino, S. M. L. & Ronchi, A. E. SOX6 blocks the proliferation of BCR-ABL1(+) and JAK2V617F(+) leukemic cells. *Sci. Rep.* **9**(1), 3388 (2019).
37. Qin, Y. R. *et al.* Characterization of tumor-suppressive function of SOX6 in human esophageal squamous cell carcinoma. *Clin. Cancer Res.* **17**(1), 46–55 (2011).
38. Xie, Q. *et al.* Aberrant expression of microRNA 155 may accelerate cell proliferation by targeting sex-determining region Y box 6 in hepatocellular carcinoma. *Cancer* **118**(9), 2431–2442 (2012).
39. Guo, X., Yang, M., Gu, H., Zhao, J. & Zou, L. Decreased expression of SOX6 confers a poor prognosis in hepatocellular carcinoma. *Cancer Epidemiol.* **37**(5), 732–736 (2013).
40. Li, Y., Xiao, M. & Guo, F. The role of Sox6 and Netrin-1 in ovarian cancer cell growth, invasiveness, and angiogenesis. *Tumour Biol.* **39**(5), 1010428317705508 (2017).
41. Jiang, W. *et al.* Identification of Sox6 as a regulator of pancreatic cancer development. *J. Cell Mol. Med.* **22**(3), 1864–1872 (2018).
42. Popescu, D. M. *et al.* Decoding human fetal liver haematopoiesis. *Nature* **574**(7778), 365–371 (2019).
43. Kildisiute, G. *et al.* Tumor to normal single-cell mRNA comparisons reveal a pan-neuroblastoma cancer cell. *Sci. Adv.* **7**(6), eabd3311 (2021).
44. Lu, Y. *et al.* A single-cell atlas of the multicellular ecosystem of primary and metastatic hepatocellular carcinoma. *Nat. Commun.* **13**(1), 4594 (2022).
45. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: Large-scale single-cell gene expression data analysis. *Genome Biol.* **19**(1), 15 (2018).
46. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat. Biotechnol.* **36**(5), 411–420 (2018).
47. Kurita, R. *et al.* Establishment of immortalized human erythroid progenitor cell lines able to produce enucleated red blood cells. *PLoS One* **8**(3), e59890 (2013).
48. Lederer, M., Bley, N., Schleifer, C. & Huttelmaier, S. The role of the oncofetal IGF2 mRNA-binding protein 3 (IGF2BP3) in cancer. *Semin. Cancer Biol.* **29**, 3–12 (2014).
49. Hammer, N. A. *et al.* Expression of IGF-II mRNA-binding proteins (IMPs) in gonads and testicular cancer. *Reproduction* **130**(2), 203–212 (2005).
50. Zhou, J. *et al.* Inhibition of LIN28B impairs leukemia cell growth and metabolism in acute myeloid leukemia. *J. Hematol. Oncol.* **10**(1), 138 (2017).
51. Chambers, C. B. *et al.* The mRNA-binding protein IGF2BP1 restores fetal hemoglobin in cultured erythroid cells from patients with beta-hemoglobin disorders. *Mol. Ther. Methods Clin. Dev.* **17**, 429–440 (2020).
52. de Vasconcellos, J. F. *et al.* IGF2BP1 overexpression causes fetal-like hemoglobin expression patterns in cultured human adult erythroblasts. *Proc. Natl. Acad. Sci. USA* **114**(28), E5664–E5672 (2017).
53. Li, Y. *et al.* Single-cell analysis of neonatal HSC ontogeny reveals gradual and uncoordinated transcriptional reprogramming that begins before birth. *Cell Stem Cell* **27**(5), 732–747e7 (2020).
54. Li, Y. & Magee, J. A. Transcriptional reprogramming in neonatal hematopoietic stem and progenitor cells. *Exp. Hematol.* **101–102**, 25–33 (2021).

Acknowledgements

This work was supported by the European Union's Horizon 2020 Research and Innovation Program under the Marie Skłodowska Curie grant agreement No. 813091 (A.E.R.), by MUR—Ministero dell'Università e della Ricerca (PRIN: 2021-NAZ-0253) (A.E.R.) and by the European Union—NextGenerationEU through the Italian Ministry of University and Research under PNRR-M4C2-I1.3 Project PE_00000019 "HEAL ITALIA" to Antonella Ronchi (H43C22000830006), from Cancerfonden (CAN 2018/542 and 21 1572 Pj) and the Swedish Research Council (Vetenskapsrådet; 2021-03075 and 2023-01898) to C.C.. The views and opinions expressed are those of the authors only and do not necessarily reflect those of the European Union or the European Commission. Neither the European Union nor the European Commission can be HELD responsible for them. C.C. is a Wallenberg Molecular Medicine (WCMM) fellow and receives generous financial support from the Knut and Alice Wallenberg Foundation. The computations and data handling were enabled by resources provided by the National Supercomputer Centre (NSC), funded by Linköping University. Peter Münger at the National Supercomputer Centre is acknowledged for assistance concerning technical and implementational aspects in making the codes run on the Sigma resource. The sequencing was carried out at the molecular biology unit of Linköping University's core facility. We thank Gloria Barbarani and Cristina Fugazza, Agata Labeledz, Cristiana Barone and Emanuele Azzoni for experimental help and for sharing preliminary results and Silvia ML Barabino and John Strouboulis for reading, discussion and suggestions.

Author contributions

V.P., G.Z., T.W. and E.C. performed experiments and contributed to the conceptualization of this work. Y.N. provided HUDEP cells. C.C. and A.E.R. conceived the experiments and wrote the paper.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-024-60438-3>.

Correspondence and requests for materials should be addressed to A.E.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2024