

OPEN

# Clinicopathological features and prognostic value of SOX11 in childhood acute lymphoblastic leukemia

Toni Grönroos<sup>1,8\*</sup>, Artturi Mäkinen<sup>1,2,8</sup>, Saara Laukkanen<sup>1</sup>, Juha Mehtonen<sup>3</sup>, Atte Nikkilä<sup>1</sup>, Laura Oksa<sup>1</sup>, Samuli Rounioja<sup>4</sup>, Yanara Marincecic-Zuniga<sup>5</sup>, Jessica Nordlund<sup>5</sup>, Virva Pohjolainen<sup>2</sup>, Timo Paavonen<sup>2,6</sup>, Merja Heinäniemi<sup>3</sup> & Olli Lohi<sup>1,7</sup>

Acute lymphoblastic leukemia is marked by aberrant transcriptional features that alter cell differentiation, self-renewal, and proliferative features. We sought to identify the transcription factors exhibiting altered and subtype-specific expression patterns in B-ALL and report here that SOX11, a developmental and neuronal transcription factor, is aberrantly expressed in the ETV6-RUNX1 and TCF3-PBX1 subtypes of acute B-cell leukemias. We show that a high expression of SOX11 leads to alterations of gene expression that are typically associated with cell adhesion, migration, and differentiation. A high expression is associated with DNA hypomethylation at the SOX11 locus and a favorable outcome. The results indicate that SOX11 expression marks a group of patients with good outcomes and thereby prompts further study of its use as a biomarker.

SOX11 is a transcription factor (TF) encoded by the *SOX11* gene located in chromosome 2p25<sup>1</sup>. It is a member of the *SoxC* (sex-determining region Y-related HMG box) group of genes and consists of two functional domains—the N-terminal DNA-binding and the C-terminal transactivation domains<sup>2,3</sup>. Other *SoxC* family members include SOX4 and SOX12. Of these, SOX4 is a crucial TF in B lymphopoiesis and is expressed in the B- and T-cell lineages<sup>2,4,5</sup>. SOX11 is normally expressed in the developing central nervous system of the embryo, in keratinocytes, and in some other epithelial tissues<sup>1,6–8</sup>. It is also expressed in ovarian and breast cancer, in which both tumor suppressor and oncogenic functions have been suggested<sup>9,10</sup>. A knockout mouse model revealed the vital role of *SOX11* during embryonic development, as *SOX11*-deficient mice died after birth likely from multiple heart defects, asplenia, and organ hypoplasia in the lungs, stomach, and pancreas<sup>11</sup>. *SOX11* deletions and mutations are associated with neurodevelopmental disorders<sup>12</sup>.

Previous studies have shown that SOX11 mRNA and nuclear protein expression is a specific marker for conventional but not indolent mantle cell lymphoma (MCL)<sup>8,13–15</sup>. In MCL, SOX11 has been associated with either increased or reduced cell proliferation<sup>16–23</sup> and either good or bad prognosis<sup>24–27</sup>. In a cohort of 50 adult acute myeloid leukemia (AML) patients, a high SOX11 expression was associated with FLT/ITD and NPM1 mutations and a shortened disease-free survival<sup>28</sup>.

There is other evidence linking SOX11 with B-lineage malignancies. Dictor *et al.* (2009) reported that the nuclear expression of SOX11 was found in eight cases of B-lymphoblastic neoplasias<sup>29</sup>. Another study reported the strong nuclear expression of SOX11 in a single B-cell and five T-cell lymphoblastic lymphoma/leukemias<sup>13</sup>. Vegliante *et al.* (2011) demonstrated the increased expression of *SOX11* mRNA in ETV6-RUNX1 (E/R) and

<sup>1</sup>Tampere Center for Child Health Research, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. <sup>2</sup>Fimlab Laboratories, Department of Pathology, Tampere University Hospital, Tampere, Finland. <sup>3</sup>Institute of Biomedicine, School of Medicine, University of Eastern Finland, Kuopio, Finland. <sup>4</sup>Fimlab Laboratories, Department of Hematology, Tampere University Hospital, Tampere, Finland. <sup>5</sup>Department of Medical Sciences, Molecular Medicine and Science for Life Laboratory, Uppsala University, Uppsala, Sweden. <sup>6</sup>Department of Pathology, Faculty of Medicine and Health Technology, Tampere University, Tampere, Finland. <sup>7</sup>Department of Pediatrics, Tampere University Hospital, Tampere, Finland. <sup>8</sup>These authors contributed equally: Toni Grönroos and Artturi Mäkinen. \*email: [toni.gronroos@tuni.fi](mailto:toni.gronroos@tuni.fi)

TCF3-rearranged B-cell precursor acute lymphoblastic leukemia (BCP-ALL)<sup>30</sup>, whereas Nordlund *et al.* (2012) and Busche *et al.* (2013) observed the prominent expression of *SOX11* in the *E/R* subtype of ALL<sup>31,32</sup>.

We investigated the expression of *SOX11* across lymphoid malignancies, focusing on B-lymphoblastic leukemias. The function of *SOX11* in leukemias and its clinical significance as a biomarker were further explored.

## Materials and Methods

**Microarray datasets.** We used three independent datasets to study *SOX11* expression—a combined microarray dataset (“Hemap”) retrieved from Gene Expression Omnibus (GEO)<sup>33,34</sup>, the GEO series GSE47051<sup>35</sup>, and the publicly available BCP-ALL data from the recent PanALL study<sup>36</sup>. The sample sizes for each dataset are shown in Supplementary Table 1.

**Cell lines, cell culture, and drug treatments.** NALM-6, REH, 697, RCH-ACV, KOPN-8, KASUMI-2, JURKAT, MOLT-16, P12-ICHIKAWA, HPB-ALL, and CCRF-CEM were cultured in RPMI Medium 1640 (Gibco, Thermo Fisher Scientific, Waltham, MA, USA) with 2 mM L-glut, 100 U penicillin, 100 µg/ml streptomycin with 10% FBS (Gibco), and MOLT-4, PEER, and MHH-CALL3 with 20% FBS (Gibco) at 37 °C in 5% CO<sub>2</sub>. An inducible *E/R* fusion in the NALM-6 cell line and a knockdown of *E/R* by a short hairpin RNA (shRNA) in the REH cell line have been previously described<sup>37</sup>. *E/R* expression was induced with 500 ng/ml doxycycline (Clontech). *E/R* expression changes were confirmed with RT-qPCR, with fusion gene-specific primers (Table S2). Mycoplasma tests were done regularly for the cell lines, and Eurofins Genomics (Ebersberg, Germany) services were used to authenticate the cell lines by STR genotyping. All cell lines used in this study were purchased from the Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures (Braunschweig, Germany).

For the methyltransferase inhibition experiments, the cultured cells were treated for 72 h with decitabine, 5-Aza-2'-deoxycytidine solved in DMSO (A3656, Sigma-Aldrich, St. Louis, MO, USA) at 0, 0.1, and 1 µM concentrations. The media were changed at a 24 h interval to compensate for decitabine instability under cell culture conditions. After the treatment cells were collected, RNA was extracted for RT-qPCR analyses. Corticosteroid and chemotherapy treatments were conducted with the indicated concentration ranges, and cell viabilities were measured after either 72 (697 and RCH-ACV) or 96 h (REH). The corticosteroids included prednisolone (P6004, Sigma-Aldrich) and dexamethasone (D8893, Sigma-Aldrich), and the chemotherapy agents used were asparaginase (A3809, Sigma-Aldrich) and vincristine (V8879, Sigma-Aldrich). The applied concentrations for each cell line are indicated in Table S3.

**Quantitative real-time PCR.** Total RNA was extracted using the PureLink™ RNA Mini Kit, and the On-Column PureLink® DNase Treatment Protocol was used for DNA removal (Ambion® by Life Technologies and Invitrogen, Thermo Fisher Scientific, Waltham, MA, USA); 100–500 ng of the extracted RNA was used as a starting material for cDNA synthesis, which was performed with iScript (Bio-Rad, Hercules, CA, USA). RT-qPCR reactions were conducted according to the manufacturer's instructions with SsoFast EvaGreen® Supermix (Bio-Rad). The following program was performed with the Bio-Rad CFX96™ Real-Time System (Bio-Rad): initial denaturation at 96 °C for 30 s, 39 cycles of denaturation at 96 °C for 2 s, annealing at 60 °C for 5 s, and plate read. Independent experiments performed in triplicate were used as the starting material for the RT-qPCR measurements, and the relative 2<sup>-ΔΔC<sub>T</sub></sup> method was used for quantification<sup>38</sup>. The primer sequences used in RT-qPCR are listed in Table S2.

**Western blot.** Protein extraction was performed using M-PER reagent (Thermo Fisher Scientific) according to the manufacturer's instructions to lyse the cells, and protein concentrations were measured with DC Protein Assay (Bio-Rad); 15–20 µg of a protein sample was loaded into the precast 10% Mini-PROTEAN® TGX Stain-Free™ Gels (Bio-Rad). After the electrophoresis run, Trans-Blot® Turbo™ Pack (Bio-Rad) was used to transfer the proteins from the gel to the nitrocellulose membrane. Transfer was done with the Trans-Blot® Turbo™ Transfer system according to the manufacturer's instructions (Bio-Rad). A prestained protein ladder, PageRuler Plus (#26619, Thermo Fischer Scientific), was used as a protein size marker. We utilized antibodies against *SOX11* (1:1,000) (HPA000536, Sigma-Aldrich, Lot # BB107024) and Histone H3 (1:75,000) (ab4729, Abcam, Cambridge, UK, Lot # GR167613-1). Horseradish peroxidase conjugated anti-rabbit (1:2,000) (P0217, Lot # 00069121) was used as a secondary antibody (Agilent Technologies, Santa Clara, CA, USA). Chemiluminescence reaction by Amersham ECL reagent was detected with ChemiDoc™ XRS+ using Image Lab™ Software (Bio-Rad).

**Gene silencing with nucleofection.** Knockdowns for suspension cells were performed using 4D-Nucleofector™ (Lonza, Basel, Switzerland) for transfections. *SOX11* knockdown was done using gene sequence-specific small interfering RNA (siRNA) (Sigma-Aldrich), and a non-specific siRNA was used as a control (Table S4). Before nucleofection, 20 µM stock solutions were diluted in Resuspension Buffer (SR30005, OriGene, Rockville, MD, USA) so that the final concentration per reaction was 300 nM. One million cells were used for each nucleofection reaction. Nucleofection reactions were conducted in proper solutions and programs according to the manufacturer's instructions in single nucleocuvettes (Table S5). Then, the cells were transferred to 12-well plates with prewarmed fresh media. The transfected cells were used in cell viability assays and RNA sequencing (RNA-seq). Western blot and RT-qPCR were used to assess knockdown levels.

**RNA sequencing of cell lines.** *SOX11*-specific and control siRNAs were used in the nucleofection for the 697, RCH-ACV, and REH cell lines to create *SOX11* knockdown and control samples. After 48 h of nucleofection, three million cells per sample were collected for the total RNA extraction. RNA extraction was performed with the PureLink™ RNA Mini Kit, and the On-Column PureLink® DNase Treatment Protocol was used to avoid contamination by genomic DNA (Ambion® by Life Technologies and Invitrogen). Protein samples were also

collected, and the *SOX11* knockdown level was verified by both RT-qPCR and Western blotting. Three independent biological replicates were collected for each transfected cell line, and each sample had 25 ng of RNA in 40  $\mu$ l. Library preparation and RNA-seq (GSE123943) were performed in the Finnish Functional Genomics Centre (Turku, Finland). See more details in Supplementary Information.

The quality of the raw sequencing reads was ensured with FastQC (v0.10.1). Based on the FastQC results, reads were trimmed and their quality was filtered using the FASTX Toolkit (0.0.14). The reads were mapped to the human reference genome version hg19 using STAR aligner software (2.5.3a modified); reads aligning to more than two locations were discarded<sup>39</sup>. The alignment file was turned into tag directories, and read counts were calculated using the HOMER tool kit (v4.8).

Differential gene expression was analyzed using the quasi-likelihood F-test from edgeR, an R package<sup>40</sup>. Differentially expressed genes from all cell lines were filtered using the adjusted p-value < 0.05 (Benjamini–Hochberg method) as a cut-off, and the resulting gene lists from three different cell lines were compared by drawing a Venn diagram with an interactive Venn online tool (<http://www.interactivenn.net/>). Heatmaps, presenting all biological replicates, were drawn by using the z-scores of reads per kilobase of transcript, the per million mapped reads (RPKM) normalized count matrix, and the ComplexHeatmap R-package<sup>41</sup>. Gene set enrichment was analyzed by Gene Set Enrichment Analysis (GSEA) 3.0 software using logFC ranked lists of genes from differential gene expression analysis, and the results presented had an adjusted p-value < 0.02 (Benjamini–Hochberg method)<sup>42,43</sup>. Gene ontology (GO) term enrichment was studied with two approaches—with GSEA software using f-statistics ranked lists of genes and with the GOrilla online tool using unranked target and background lists<sup>44,45</sup>. The target lists were created with a threshold of adjusted p-value < 0.05 and logFC > 0.5 or < –0.5.

**RNA sequencing and methylation analysis of the patient samples.** Previous transcriptome sequencing of pediatric ALL cohort by Marincevic-Zuniga *et al.*<sup>46</sup> included 116 BCP-ALL cases, of which 115 cases were used to assess the *SOX11* mRNA expression level in this study. DNA methylation data (GSE49031) were available for 112 of the 115 BCP-ALL cases<sup>35</sup>. See more details in Supplementary Information.

**Cell viability and proliferation assays.** Fresh media were replaced on the transfected cells after 24 h of transfection. For the cell viability assay, the cells were counted, and 10,000 cells per well were used in a 96-well plate. The cells were allowed to grow up to 72 (697 and RCH-ACV) and 120 h (REH), and cell viability was measured every 24 h with 10  $\mu$ l of Alamar Blue reagent per well with a 2 h incubation before fluorescence measurement with excitation of 560 nm and emission of 590 nm using the Tecan fluorometer Infinite 200 (Tecan, Männedorf, Switzerland). For each time point, we used four technical replicates per sample to calculate the mean.

Cell proliferation assays were performed by counting the cells every 24 h after the transfection up to 96 (697 and RCH-ACV) and 120 h (REH).

**Clinical data and the patient samples.** The data on pediatric ALL cases below 18 years of age and treated in the Tampere University Hospital were retrieved from the clinical registry from years 1990 to 2017. Essential clinical information, such as age, leukocyte count, gender, relapse, death, central nervous system leukemia, immunophenotype based on flow cytometry, and clinical genetic information, was collected. The study obtained permission from the local ethical committee, and informed consent was sought, as needed (Pirkanmaa Hospital District Ethical Committee, R16054 and R13109). The use of old biopsy samples was approved by the National Supervisory Authority for Welfare and Health (Valvira), and the samples were handled in accordance with relevant guidelines and regulations.

A total of 126 representative diagnostic formalin-fixed and paraffin-embedded (FFPE) decalcified bone marrow trephine biopsy samples were collected from the pathology department archives on the basis of the primary sample reports. For a proportion of cases, remission and relapse samples were also retrieved. Plastic-embedded and inadequate samples were excluded. The cases were classified based on the WHO 2017 Classification of Tumors and Haematopoietic and Lymphoid Tissues<sup>47</sup>.

**Immunohistochemistry.** Four micron-thick whole tissue sections were used for immunohistochemistry. All cases were stained with anti-*SOX11* antibody (clone MRQ-58, Cell Marque, Sigma-Aldrich, Lot # 1331005 A and 1430213 C) using Ventana Benchmark Classic at a dilution of 1:50. FFPE human MCL was used as a positive control material, and remission bone marrow samples were used as negative controls.

Staining intensity was graded, and cases with less than 20% of positivity of leukemic blast cell nuclei were interpreted as negative. Cases with immunohistochemical positivity ranging from 20% to 50% in the blast cell nuclei were graded positive, and cases with immunohistochemical nuclear positivity of over 50% were considered strongly positive. *SOX11* expression was independently analyzed with a light microscope by two experienced pathologists without knowledge of the clinical data. Cases with discrepant scores were further analyzed by a third pathologist.

**Fluorescence *in situ* hybridization.** For cases lacking the genetic subtype information, fluorescence *in situ* hybridization analysis was performed on either the bone marrow aspiration samples or the FFPE samples. The following probes were used: Metasystems E2A Break Apart Probe 19p13 (Lot # 18216), Metasystems XL MLL plus Break Apart Probe 11q23 (Lot # 18451), Metasystems XL BCR/ABL1 plus (Lot # 19082), and Metasystems XL t(12;21) (Lot # 19133).

**Flow cytometry analysis.** For flow cytometry analysis, the cells were permeabilized using Fix&Perm reagent according to the manufacturer's instructions (GAS003, Invitrogen). The cells were then stained with Alexa Fluor 647<sup>®</sup> conjugated rabbit anti-human *SOX11* antibody [EPR8191 (2)] (ab198540, Abcam), while CD3-APC antibody (345767) (Becton Dickinson, Franklin Lakes, NJ, USA) served as a negative control. The data were

acquired with the Beckman Coulter Navios cytometer (Beckman Coulter, Brea, CA, USA) using Red laser (638 nm) and 660/20 bandpass filter. Data analysis was conducted using Kaluza software (Beckman Coulter).

**Statistical analysis.** IBM SPSS Statistics (v. 22) and R (v. 3.40) were used for the statistical analysis. Kaplan–Meier survival analysis was performed, and Log-rank test was used, with a  $p$ -value  $< 0.05$  considered statistically significant. Event-free survival (EFS) was analyzed using death, relapse, resistant disease (blast count  $> 25\%$  at the end of induction), and secondary malignancy as events. Cox proportional hazards models were fitted to estimate the effects of potential risk factors on survival. Chi-squared test was performed on *SOX11* expression and clinicopathological prognostic variables. Kruskal–Wallis H and Mann–Whitney U tests were used to evaluate the differential expression of *SOX11* in distinct leukemia subgroups (Table S6). All performed statistical tests were two tailed, and no corrections for multiple testing were used.

## Results

**SOX11 is overexpressed in acute lymphoblastic leukemias.** Derailed differentiation and abnormal proliferation of B-cells are thought to underlie the genesis of BCP-ALL. As TFs are key cell differentiation drivers, we sought to identify aberrantly expressed TFs in BCP-ALL. We utilized a large, curated dataset of microarray-based gene expression profiles retrieved from GEO<sup>34</sup>. This dataset comprises a total of 9,544 hematological gene expression profiles, including 4,418 leukemias, 428 healthy controls, and 862 cell lines. Among the genes with altered expression, TF *SOX11* showed prominent expression in MCL and in the E/R and T/P subtypes of BCP-ALL (Fig. 1a). Compared with healthy hematopoietic cells, *SOX11* had a 4.7- and 4-fold higher expression in the E/R and T/P subtypes, respectively, with an expression comparable to that in MCL (Fig. 1a). A similar subtype-specific expression of *SOX11* was replicated in two additional ALL gene expression datasets (Fig. 1a). Interestingly, the PanALL study revealed that *SOX11* expression is also elevated in novel subtypes, such as E/R-like, IKZF1 N159Y (*IKZF1* missense alteration encoding p.Asn159Tyr), MEF2D rearrangement, and DUX4 rearrangement.

As *SOX11* belongs to the SoxC family of TFs, we investigated the expression of the two other members of this family. *SOX4*, which has been reported to affect survival, progression, and proliferation in BCP-ALL<sup>48</sup>, was highly expressed in BCP-ALL and T-ALL, but it showed no subtype-specific expression pattern. The expression of *SOX12* did not vary markedly between the studied subtypes (Fig. S1a).

A subtype-specific expression pattern was also present in BCP-ALL cell lines, as REH cells, representing the E/R subtype, and RCH-ACV and 697 cells, representing the T/P subtype, showed an increased expression of *SOX11* by RT-qPCR (Fig. 1b). Neither knockdown of *E/R* in REH cells nor their overexpression in NALM-6 cells had any effect on *SOX11* expression levels (Fig. S2a).

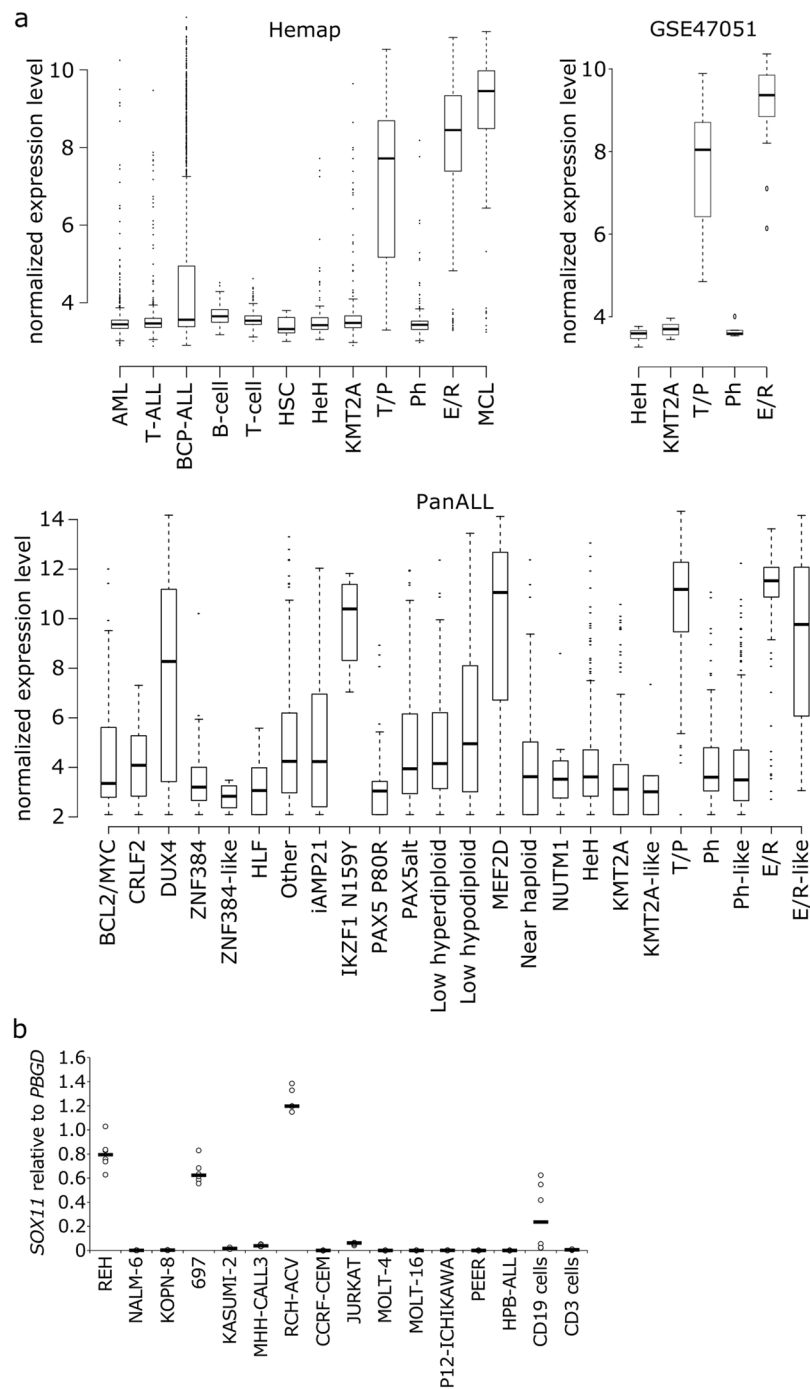
**Immunohistochemical analysis of SOX11 expression in ALL.** In order to confirm the *SOX11* expression at the protein level and relate it to clinical features, we collected a retrospective cohort of 119 B-ALL cases with available bone marrow biopsies and associated clinical data (Table 1). We performed immunohistochemical staining of the biopsy samples by using a *SOX11*-specific antibody. Staining intensity was graded from 0 to 2, with 0 marking negative, 1 positive, and 2 strongly positive samples. Eighty out of 119 primary B-ALL bone marrow samples stained negative for *SOX11*, whereas 34 cases were positive and 5 were strongly positive (Fig. 2a and Table 2). A total of 29/39 (74.4%) of the positively staining B-ALL cases represented either the E/R or T/P subtype. A statistically significant association was observed with the E/R subtype ( $p$ -value  $< 0.001$ ) but not with the T/P subtype (likely because of the low number of cases:  $N = 4$ , 3 positive cases). Cases with either *KMT2A* rearrangement or the Philadelphia chromosome (*BCR-ABL1*-translocation) did not express the *SOX11* protein, and the majority of hyperdiploid cases were also *SOX11* negative. Additionally, seven T-ALL cases included in the whole ALL patient cohort were all negative.

**SOX11 expression is associated with a favorable outcome.** We next sought to evaluate the clinical significance of *SOX11* expression in B-ALL ( $N = 119$ ). The overall survival (OS) was better in the *SOX11*-positive group, and no deaths occurred among *SOX11*-positive cases ( $p = 0.039$ ) (Fig. 2b). The EFS and relapse-free survival (RFS) adjusted for the competing event (death) showed similar trends but did not reach statistical significance (Fig. 2b). *SOX11* positivity was not associated with good early therapy response, as measured by a minimal residual disease below 0.1% at the end of induction therapy (OR = 0.54, 95% CI 0.22, 1.28,  $p = 0.17$ ). In multivariate analysis of EFS with covariates (age, WBC, MRD and subtypes), a positive immunohistochemical staining for *SOX11* protein showed a favorable trend (Table 3).

When *SOX11* immunostaining positivity was analyzed separately within the E/R subtype, *SOX11*-positive cases had a better OS (Log-rank test  $p = 0.004$ ), but EFS did not show a statistically significant difference with a hazard ratio of 0.67 (95% CI 0.07, 6.43).

We replicated the survival findings in another dataset. Transcriptomic data from a patient cohort comprising 115 BCP-ALL cases were analyzed for *SOX11* expression and patients' survival status<sup>46</sup>. Figure 3a shows that a high *SOX11* mRNA expression was associated with a favorable trend in EFS analysis.

**SOX11 overexpression is associated with DNA hypomethylation.** We next investigated the biology behind the increased *SOX11* expression in leukemia. Neither primary transcription (GRO-seq,  $N = 8$ ) nor whole-genome sequencing (WGS,  $N = 8$ ) of the *SOX11* gene in BCP-ALL cases revealed any aberrant enhancer activity or somatic mutations, respectively, in the *SOX11* gene or nearby regions (data not shown), prompting us to look for other mechanisms. In MCL, hypomethylation (partly) drives the increased *SOX11* expression<sup>30</sup>. We utilized the above-mentioned BCP-ALL patient cohort with readily available genome-wide CpG methylation data<sup>35,46</sup>. Altogether, 23 CpG sites were located within the *SOX11* gene locus, and a strong pattern of DNA hypomethylation was seen in patients with a high *SOX11* mRNA expression (Fig. 3b,c). Nevertheless,



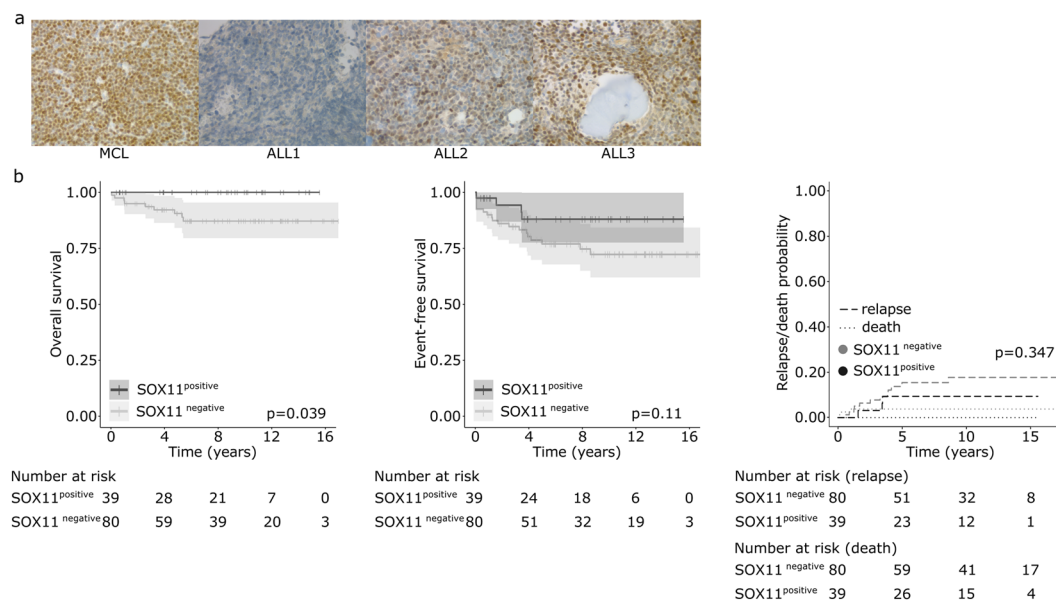
**Figure 1.** *SOX11* expression in hematological malignancies. **(a)** Expression boxplots of *SOX11* in healthy cells, leukemias, and mantle cell lymphoma. Data sources: Hemap microarray dataset<sup>34</sup>, GSE47051<sup>35</sup>, and the PanALL study<sup>36</sup>. See Supplementary Information for more details. **(b)** *SOX11* expression in ALL cell lines and healthy B- and T-cells, as measured by RT-qPCR (N = 2, black lines indicate the median). ALL, acute lymphoblastic leukemia; AML, acute myeloid leukemia; BCL2/MYC, BCL2/MYC rearranged; BCP-ALL, B-cell precursor ALL; CRLF2, CRLF2 (non-Ph-like); DUX4, DUX4 rearranged; E/R, ETV6-RUNX1; HeH, high hyperdiploid; HLF, TCF3/TCF4-HLF; HSC, hematopoietic stem cell; iAMP21, intrachromosomal amplification of chromosome 21; IKZF1 N159Y, IKZF1 missense alteration encoding p.Asn159Tyr; KMT2A, KMT2A rearranged; MCL, mantle cell lymphoma; MEF2D, MEF2D rearranged; NUTM1, NUTM1 rearranged; PAX5alt, PAX5 alterations; PAX5 P80R, PAX5 p.Pro80Arg (P80R) alteration; Ph, Philadelphia chromosome; T/P, TCF3-PBX1; ZNF384, ZNF384 rearranged. Reproduced with permission<sup>66</sup>.

	WHO Subtype								Total/Combined
	Burkitt	NOS	Ph	KMT2A	ETV6-RUNX1	Hyper-diploid	Hypo-diploid	TCF3-PBX1	
N of cases	2	42	2	5	33	30	1	4	119
Age, Md (min/max)	12.7	8.1	10.1	1.3	4	3.7	4.1	7.4	4.3 (0.9/17.6)
WBC count (10 <sup>9</sup> /l), Md (min/max)	18.9	6.7	156.3	109.7	7	7.1	1.9	43.4	7.1 (1/311)
Deceased	2	6	1	0	1	1	0	0	9
Relapse	0	7	1	0	3	3	0	1	15
CNS disease	0	2	1	0	1	2	0	0	6
Resistant disease	0	3	0	0	1	0	0	0	4
MRD (%) at EOI, Md (min/max)	0	0.05	1.75	0	0.06	0	0	0	0.02 (0/44)
Follow-up (years), Md (min/max)	7.5	9.7	8.4	4	7.1	9.2	11.5	7.6	8.2 (0.1/17)

**Table 1.** Summary of clinical characteristics of B-ALL patients. CNS, central nervous system; EOI, end of induction therapy; KMT2A, KMT2A rearrangement; Max, maximum; Md, median; Min, minimum; MRD, minimal residual disease; NOS, not otherwise specified; Ph, Philadelphia chromosome; WBC, white blood cell.

SOX11 IHC	WHO Subtype								All cases
	Burkitt	NOS	Ph	KMT2A	ETV6-RUNX1	Hyper-diploid	Hypo-diploid	TCF3-PBX1	
Negative	1	35	2	5	7	29	0	1	80
Positive	1	6	0	0	24	0	1	2	34
Strong positive	0	1	0	0	2	1	0	1	5
Total	2	42	2	5	33	30	1	4	119

**Table 2.** Summary of SOX11 protein expression by immunohistochemical staining in B-ALL. IHC, immunohistochemistry; KMT2A, KMT2A rearrangement; NOS, not otherwise specified; Ph, Philadelphia chromosome.

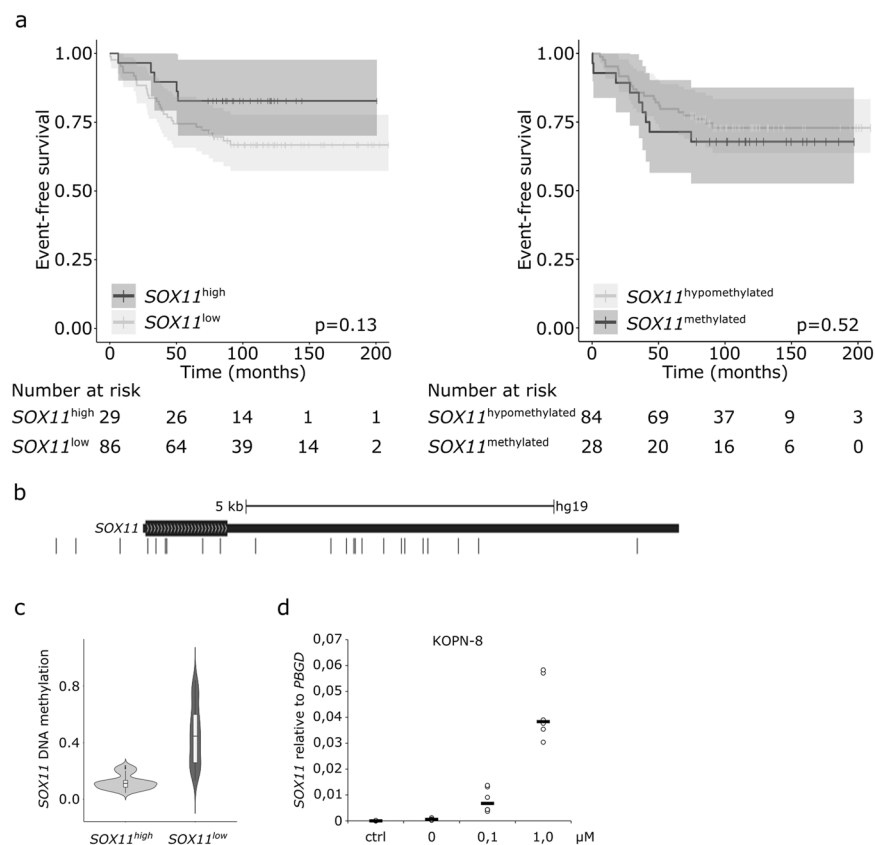


**Figure 2.** SOX11 protein expression and impact on the outcome in BCP-ALL. (a) Expression of SOX11 protein by immunohistochemistry. Mantle cell lymphoma (MCL), a strongly positive case for SOX11 (400 $\times$ ); ALL1, a negative B-ALL case (400 $\times$ ); ALL2, a positive B-ALL case (400 $\times$ ); ALL3, a strongly positive B-ALL case (400 $\times$ ). (b) Kaplan–Meier survival curves and Log-rank p-values for OS, EFS, and RFS in the SOX11-positive (high) and -negative (low) groups. Reproduced with permission<sup>66</sup>.

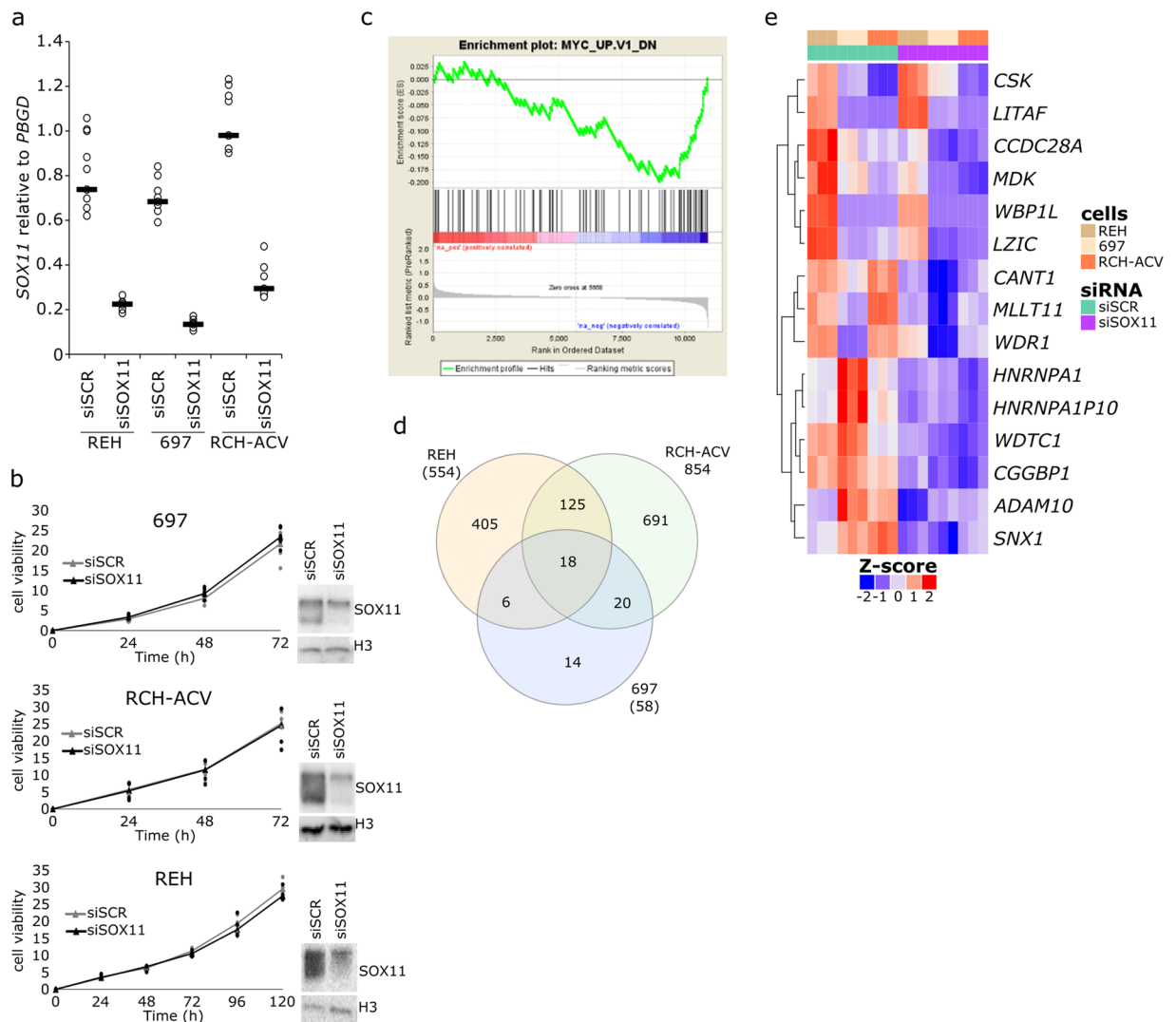
DNA hypomethylation of the CpG sites at the *SOX11* locus was not associated with a better EFS (Fig. 3a, right panel). We also tested whether a methyltransferase inhibitor, decitabine, could reverse *SOX11* expression in leukemia cell lines. After 72 h of decitabine treatment, a marked increase in *SOX11* expression was observed in a concentration-dependent manner in KOPN-8 and REH cells (Figs. 3d and S2b).

	N	MULTIVARIATE			UNIVARIATE		
		HR	95% CI	p-value	HR	95% CI	p-value
<b>Age (years)</b>							
≥1 and ≤10	93	1.00*			1.00*		
<1 and >10	26	0.87	0.31–2.38	0.79	1.27	0.50–3.23	0.61
<b>WBC count (10<sup>9</sup>/l)</b>							
<50	101	1.00*			1.00*		
≥50	18	0.55	0.16–1.92	0.35	0.90	0.27–3.04	0.87
<b>Subtype</b>							
Other B-ALL	56	1.00*			1.00*		
High hyperdiploidy	30	0.27	0.07–1.00	0.05	0.45	0.15–1.36	0.16
ETV6-RUNX1	33	0.72	0.18–2.83	0.64	0.50	0.17–1.51	0.22
<b>SOX11 expression</b>							
negative	80	1.00*			1.00*		
positive	39	0.37	0.10–1.43	0.15	0.43	0.15–1.26	0.12
<b>MRD at EOI</b>							
<0.1%	78	1.00*			1.00*		
≥0.1%	37	1.70	0.68–4.24	0.26	2.18	0.94–5.06	0.07

**Table 3.** Multivariate and univariate analyses of the event-free survival based on expression level of SOX11 protein in pediatric B-ALL. Cox proportional hazards regression calculated for known risk factors. SOX11 expression was treated as a binary variable. 95% CI, 95% confidence interval; B-ALL, B-cell acute lymphoblastic leukemia; HR, hazard ratio; WBC, white blood cell. \*Marks reference groups of each categorical variable.



**Figure 3.** Survival analysis and methylation status of cases with either a high or low SOX11 mRNA expression. **(a)** Kaplan–Meier survival curves and p-values of Log-rank test for EFS in patients with a low or high expression of SOX11 and in patients with a low or high methylation of the SOX11 gene locus. **(b)** CpG sites at the SOX11 locus in chromosome 2<sup>35</sup>. **(c)** DNA methylation at the SOX11 gene locus among patients with either a low or high expression of SOX11. **(d)** Effect of decitabine treatment on SOX11 mRNA expression in the KOPN-8 cell line. Reproduced with permission<sup>66</sup>.



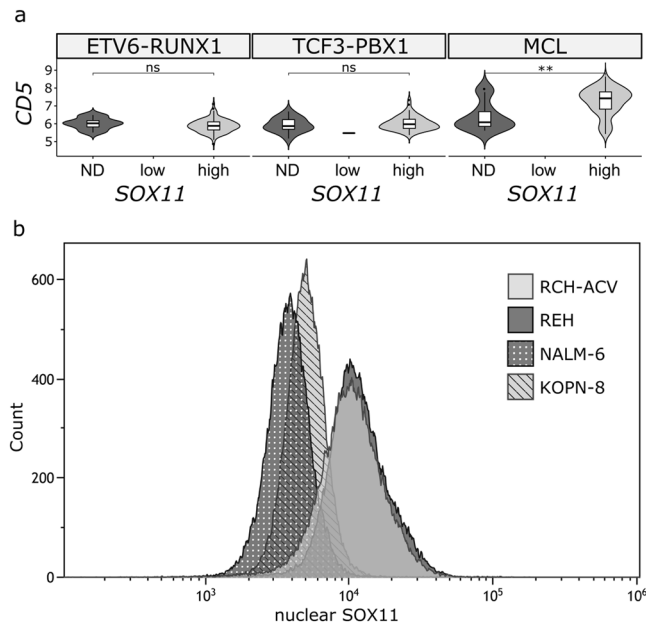
**Figure 4.** Cell viability and transcriptomic changes after knockdown of *SOX11*. **(a)** Expression level of *SOX11* after knockdown, as analyzed by RT-qPCR (N = 3, black lines indicate the median). **(b)** Knockdown of *SOX11* caused no evident changes in cell viability. Curves are drawn from the biological replicates using the median value at given time points. 697 and RCH-ACV cells represent the T/P subtype, and REH cells represent the E/R subtype. *SOX11* knockdown was confirmed by Western blotting, and cell viability assessments were conducted with the AlamarBlue assay. Measured absorbance intensities are reported as  $\times 10^3$  (697 N = 6; RCH-ACV and REH N = 4). Western blot gel figures are cropped per cell line from the original blot figures shown in Fig. S3. **(c)** *SOX11* knockdown in 697 cells caused downregulation of genes that are known to be altered after MYC upregulation. **(d)** Venn diagram of differentially expressed genes in the REH, 697, and RCH-ACV cell lines after knockdown of *SOX11* (adjusted p-value < 0.05). **(e)** Heatmap of 15 concordantly differentially expressed genes in all three cell lines after knockdown of *SOX11*. Reproduced with permission<sup>66</sup>.

### Knockdown of *SOX11* alters gene sets related to cell development, motility, and drug response pathways.

We next silenced *SOX11* expression in three cell lines that overexpress *SOX11* (REH, RCH-ACV, and 697) by using siRNA oligos. Figure 4a,b show that *SOX11* expression was decreased to 20–40% at both the mRNA and protein levels compared with scrambled siRNA-transfected cells. Cell viability and proliferation assays did not demonstrate any significant changes (Figs. 4b and S4), and *SOX11* knockdown did not have any impact on sensitivity to known leukemia drugs, such as dexamethasone, prednisolone, vincristine, and asparaginase (Fig. S5).

As *SOX11* is a TF and it regulates gene expression, we measured changes in gene expression after *SOX11* knockdown by using RNA-seq (GSE123943). Three biological replicates were used for each cell line (REH, RCH-ACV, and 697), and the data were analyzed using the R package EdgeR. In GO annotations, many of the altered terms were related to cell migration, adhesion, and differentiation (Fig. S6 and Supplementary Dataset). On the other hand, GSEA implied the altered expression of MYC and EF2 target genes (Figs. 4c and S6). *SOX11* knockdown did not have a significant effect on the expression of other Sox family members, such as *SOX4* and *SOX12* (Fig. S7).





**Figure 5.** Detection of SOX11-positive cells by flow cytometry in BCP-ALL. **(a)** Correlation of CD5 expression with SOX11 mRNA expression in the E/R and T/P subtypes of B-ALL and MCL<sup>34</sup>. SOX11 expression was categorized into three groups, not detected (ND), low and not detected (ND). **(b)** Flow cytometry analysis of nuclear SOX11 expression in cell lines with either a low (NALM-6, KOPN-8) or high (REH, RCH-ACV) expression of SOX11. ND, not detected, MCL, mantle cell lymphoma. Reproduced with permission<sup>66</sup>.

The Venn diagram in Fig. 4d shows that 18 genes were differentially expressed in all of the three cell lines after SOX11 knockdown (with an adjusted p-value < 0.05). Of these genes, 15 were concordantly down-regulated in all three cell lines and are shown in the heatmap (Fig. 4e, Supplementary Dataset). SOX11 knockdown led to the downregulation of the *WD Repeat Domain 1* (*WDR1*) gene, which is involved in the remodeling of the actin cytoskeleton, regulation of cell migration, motility of neutrophils, and maturation of megakaryocytes<sup>49,50</sup>. Another gene down-regulated by SOX11 knockdown is *Midkine* (*MDK*), a secreted growth factor that promotes cell migration and growth and is associated with an adverse prognosis in ALL possibly via increased drug resistance<sup>51,52</sup>. *WW Domain Binding Protein 1 Like* (*WBP1L*), also known as *OPAL1* (outcome predictor for acute leukemia 1), which was recently identified as a direct target of ETV6 in ALL<sup>53</sup>, was down-regulated by SOX11 silencing. Previous reports on *WBP1L*'s association with a favorable prognosis have since been refuted<sup>54</sup>. However, *WBP1L* expression had a 2.8-fold increase in the E/R subtype<sup>55</sup>, which coincides with the overexpression of SOX11 in the same subtype and may suggest co-regulation. *MLLT11* is involved in lymphoid regulation and is a known partner gene in rare leukemia translocations<sup>56,57</sup>. Similarly, *Coiled-coil domain containing 28A* (*CCDC28A*) is a fusion partner to NUP98 in AML<sup>58</sup>. *CANT1* is a calcium-dependent nucleotidase involved in pyrimidine metabolism whose regulation by SOX11 could be related with drug metabolism (cytarabine) and therapy response in ALL<sup>59</sup>. The *lipopolysaccharide-induced TNF factor* (*LITAF*) has been suggested to sensitize leukemia cells to chemotherapeutic drugs, especially in cells with a lower expression of *LITAF*<sup>60</sup>. Both *LITAF* and *Sorting nexin 1* (*SNX1*) are involved in endosomal trafficking and regulation of cell-surface receptor signaling<sup>61,62</sup>. Taken together, SOX11 knockdown leads to alterations in genes and cellular processes related to leukemia cell motility, adhesion, differentiation, and drug response.

**SOX11 protein can be detected by flow cytometry.** Finally, we searched for flow cytometry markers that could serve as surrogates of SOX11 positivity. In MCL, the surface expression of CD5 is correlated with SOX11 positivity<sup>22</sup>, but in our immunohistochemical staining and microarray dataset<sup>34</sup>, no positive correlation was observed in leukemias (Fig. 5a). Alternative surrogate markers were searched for among the routinely studied cell surface proteins, but none were associated with SOX11 positivity (data not shown). Therefore, we explored the suitability of SOX11 antibodies to discriminate SOX11-positive cases by flow cytometry. Figure 5b shows that the intensity of SOX11 antibody staining could readily separate the high expressors (REH and RCH-ACV) from the low expressors (NALM-6 and KOPN-8) in leukemia cell lines, suggesting that SOX11 antibodies could possibly be used as a biomarker in the future.

## Discussion

BCP-ALL is caused by a relatively small number of genetic mutations that impede normal B-cell differentiation, entail self-renewal capacity, and increase proliferative activity. This is evident in transcriptional programs that govern leukemic B-cells. We report here that SOX11, a developmental and neuronal TF<sup>6</sup>, is overexpressed in the E/R and T/P subtypes of BCP-ALL and also in novel E/R-like, IKZF1 N159Y, MEF2D rearrangement, and DUX4 rearrangement subtypes. A high expression is associated with DNA hypomethylation and a favorable clinical

outcome. The results suggest that a SOX11-associated transcriptional program is related with a less-aggressive disease and indicates that SOX11 warrants further study as a biomarker for low-risk ALL patients.

We observed the high expression of *SOX11* in the E/R and T/P subtypes of BCP-ALL in three separate gene expression datasets at the mRNA level<sup>34–36</sup> and confirmed this association at the protein level by immunohistochemistry, confirming previously published data<sup>8,13,29,30</sup> and adding further evidence at the protein level. We showed that the expression may be regulated epigenetically, that is, by hypomethylation of DNA at the *SOX11* locus, similar to what was earlier reported for MCL and five cases of BCP-ALL by Vegliante *et al.*<sup>30</sup>. We did not have material available for epigenetics studies, such as histone modifications, as these could also have a role in *SOX11* regulation<sup>30</sup>. Direct manipulation of the E/R fusion did not have any impact on *SOX11* expression, suggesting that regulation is indirect.

An interesting finding relates to the clinical significance of SOX11 expression in BCP-ALL, as we observed a better OS in the SOX11-positive cases. A similar trend was also seen in EFS, RFS, and early therapy response. This finding was supported (similar trend) in another dataset with transcriptome expression profiles<sup>46</sup>. In the subgroup analysis of E/R cases, SOX11 positivity retained its prognostic significance, suggesting that (a high) SOX11 expression could possibly be utilized as a biomarker for cases with a very good prognosis. It is noteworthy that our immunohistochemical staining series spans almost two decades, and patients have been treated using several distinct NOPHO ALL chemotherapy protocols<sup>63,64</sup>. As the most recent protocols have conferred the best survival results [63, trying to replicate these findings in the most recent protocols is necessary in the future.

To aid in the screening of SOX11 positivity at diagnosis, we also successfully tested a flow cytometry-based assay in cell lines. SOX11 inclusion into the flow cytometry panel would be convenient compared with immunohistochemical staining of bone marrow biopsies, which is slow and not routinely done in all treatment centers.

*SOX11* knockdown did not markedly influence cell viability or proliferation, nor did it affect chemotherapy sensitivity. In MCL, conflicting reports have been made about the effect of knockdown or *SOX11* overexpression on cell proliferation and tumor growth<sup>16–23</sup>. As cell viability measurement is a relatively insensitive assay, we performed transcriptional profiling of *SOX11* knockdown cells by using RNA-seq and noticed changes in the genes associated with cell migration, adhesion, oxidative phosphorylation, hypoxia, glycolysis, and differentiation, which could explain the association of SOX11 with favorable clinical outcomes. Notably, the changes observed were mostly mild to moderate. There were only few overlapping genes with previous profiling studies in both pro-B-cells<sup>18</sup> and MCL cells<sup>16,17,20,21,65</sup>. For example, we did not see marked changes in the expression of either *PAX5*, as seen in MCL<sup>21</sup>, or *Id1* and *Tal1* in pro-B-cells<sup>18</sup>. Interestingly, *MDK*, which is involved in cell migration and growth, was downregulated by *SOX11* knockdown in leukemia cell lines here and in an MCL cell line Z138<sup>16</sup>.

In conclusion, the association of SOX11 expression with a favorable prognosis invites further studies to confirm its prognostic value and applicability as a part of the diagnostic workup.

## Data availability

The datasets generated and analyzed in the current study are available in the GEO repository, GSE47051 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE47051>), GSE123943 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE123943>), and GSE49031 (<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE49031>), or are included in the article or supplementary files. RNA-seq data of the patient samples<sup>46</sup> are not publicly available, as the patient/parent consent does not cover depositing data into repositories; however, they are available from the authors upon reasonable request.

Received: 28 June 2019; Accepted: 22 January 2020;

Published online: 06 February 2020

## References

- Jay, P. *et al.* The human SOX11 gene: cloning, chromosomal assignment and tissue expression. *Genomics*. **29**, 541–545 (1995).
- Dy, P. *et al.* The three SoxC proteins—Sox4, Sox11 and Sox12—exhibit overlapping expression patterns and molecular properties. *Nucleic Acids Res.* **36**, 3101–3117 (2008).
- Penzo-Méndez, A. I. Critical roles for SoxC transcription factors in development and cancer. *Int. J. Biochem. Cell Biol.* **42**, 425–428 (2010).
- Nutt, S. L. & Kee, B. L. The transcriptional regulation of B cell lineage commitment. *Immunity*. **26**, 715–725 (2007).
- Yoshitomi, H. *et al.* Human Sox4 facilitates the development of CXCL13-producing helper T cells in inflammatory environments. *Nat. Commun.* **9**, 3762 (2018).
- Bergsland, M., Werme, M., Malewicz, M., Perlmann, T. & Muhr, J. The establishment of neuronal properties is controlled by Sox4 and Sox11. *Genes Dev.* **20**, 3475–3486 (2006).
- Bhattaram, P. *et al.* Organogenesis relies on SoxC transcription factors for the survival of neural and mesenchymal progenitors. *Nat. Commun.* **1**, 9 (2010).
- Ek, S., Dictor, M., Jerkeman, M., Jirstrom, K. & Borrebaeck, C. A. Nuclear expression of the non B-cell lineage Sox11 transcription factor identifies mantle cell lymphoma. *Blood*. **111**, 800–805 (2008).
- Shepherd, J. H. *et al.* The SOX11 transcription factor is a critical regulator of basal-like breast cancer growth, invasion, and basal-like gene expression. *Oncotarget*. **7**, 13106–13121 (2016).
- Sernbo, S. *et al.* The tumour suppressor SOX11 is associated with improved survival among high grade epithelial ovarian cancers and is regulated by reversible promoter methylation. *BMC Cancer*. **11**, 405 (2011).
- Sock, E. *et al.* Gene targeting reveals a widespread role for the high-mobility-group transcription factor Sox11 in tissue remodeling. *Mol. Cell Biol.* **24**, 6635–6644 (2004).
- Hempel, A. *et al.* Deletions and de novo mutations of SOX11 are associated with a neurodevelopmental disorder with features of Coffin-Siris syndrome. *J. Med. Genet.* **53**, 152–162 (2016).
- Mozos, A. *et al.* SOX11 expression is highly specific for mantle cell lymphoma and identifies the cyclin D1-negative subtype. *Haematologica*. **94**, 1555–1562 (2009).
- Zeng, W. *et al.* Cyclin D1-negative blastoid mantle cell lymphoma identified by SOX11 expression. *Am. J. Surg. Pathol.* **36**, 214–219 (2012).

15. Fernández, V. *et al.* Genomic and gene expression profiling defines indolent forms of mantle cell lymphoma. *Cancer Res.* **70**, 1408–1418 (2010).
16. Conrotto, P., Andréasson, U., Kuci, V., Borrebaeck, C. A. & Ek, S. Knock-down of SOX11 induces autotaxin-dependent increase in proliferation *in vitro* and more aggressive tumors *in vivo*. *Mol. Oncol.* **5**, 527–537 (2011).
17. Kuo, P. Y. *et al.* High-resolution chromatin immunoprecipitation (ChIP) sequencing reveals novel binding targets and prognostic role for SOX11 in mantle cell lymphoma. *Oncogene.* **34**, 1231–1240 (2015).
18. Lord, M. *et al.* Impact of Sox11 overexpression in Ba/F3 cells. *Haematologica.* **103**, e594–e597 (2018).
19. Gustavsson, E. *et al.* SOX11 expression correlates to promoter methylation and regulates tumor growth in hematopoietic malignancies. *Mol. Cancer.* **9**, 187 (2010).
20. Kuci, V., Nordström, L., Conrotto, P. & Ek, S. SOX11 and HIG-2 are cross-regulated and affect growth in mantle cell lymphoma. *Leuk. Lymphoma.* **57**, 1883–1892 (2016).
21. Vegliante, M. C. *et al.* SOX11 regulates PAX5 expression and blocks terminal B-cell differentiation in aggressive mantle cell lymphoma. *Blood.* **121**, 2175–2185 (2013).
22. Kuo, P. Y. *et al.* SOX11 augments BCR signaling to drive MCL-like tumor development. *Blood.* **131**, 2247–2255 (2018).
23. Beekman, R., Amador, V. & Campo, E. SOX11, a key oncogenic factor in mantle cell lymphoma. *Curr. Opin. Hematol.* **25**, 299–306 (2018).
24. Wang, X. *et al.* The subcellular Sox11 distribution pattern identifies subsets of mantle cell lymphoma: correlation to overall survival. *Br. J. Haematol.* **143**, 248–252 (2008).
25. Meggendorfer, M., Kern, W., Haferlach, C., Haferlach, T. & Schnittger, S. SOX11 overexpression is a specific marker for mantle cell lymphoma and correlates with t(11;14) translocation, CCND1 expression and an adverse prognosis. *Leukemia.* **27**, 2388–2391 (2013).
26. Navarro, A. *et al.* Molecular subsets of mantle cell lymphoma defined by the IGHV mutational status and SOX11 expression have distinct biologic and clinical features. *Cancer Res.* **72**, 5307–5316 (2012).
27. Nygren, L. *et al.* Prognostic role of SOX11 in a population-based cohort of mantle cell lymphoma. *Blood.* **119**, 4215–4223 (2012).
28. Totic, N. *et al.* Prognostic significance of SOX2, SOX3, SOX11, SOX14 and SOX18 gene expression in adult de novo acute myeloid leukemia. *Leuk. Res.* **67**, 32–38 (2018).
29. Dictor, M. *et al.* Strong lymphoid nuclear expression of SOX11 transcription factor defines lymphoblastic neoplasms, mantle cell lymphoma and Burkitt's lymphoma. *Haematologica.* **94**, 1563–1568 (2009).
30. Vegliante, M. C. *et al.* Epigenetic activation of SOX11 in lymphoid neoplasms by histone modifications. *PLoS One.* **6**, e21382 (2011).
31. Nordlund, J. *et al.* Digital gene expression profiling of primary acute lymphoblastic leukemia cells. *Leukemia.* **26**, 1218–1227 (2012).
32. Busche, S. *et al.* Integration of high-resolution methylome and transcriptome analyses to dissect epigenomic changes in childhood acute lymphoblastic leukemia. *Cancer Res.* **73**, 4323–4336 (2013).
33. Heinäniemi, M. *et al.* Transcription-coupled genetic instability marks acute lymphoblastic leukemia structural variation hotspots. *Elife.* **5**, e13087 (2016).
34. Pölönen, P. *et al.* Hemap: An Interactive Online Resource for Characterizing Molecular Phenotypes across Hematologic Malignancies. *Cancer Res.* **79**, 2466–2479 (2019).
35. Nordlund, J. *et al.* Genome-wide signatures of differential DNA methylation in pediatric acute lymphoblastic leukemia. *Genome Biol.* **14**, r105 (2013).
36. Gu, Z. *et al.* PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat. Genet.* **51**, 296–307 (2019).
37. Teppo, S. *et al.* Genome-wide repression of eRNA and target gene loci by the ETV6-RUNX1 fusion in acute leukemia. *Genome Res.* **26**, 1468–1477 (2016).
38. Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-ΔΔC<sub>T</sub></sup> Method. *Methods.* **25**, 402–408 (2001).
39. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* **29**, (15–21 (2013).
40. Robinson, M. D., McCarthy, D. J. & Smyth, G. K. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics.* **26**, 139–140 (2010).
41. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics.* **32**, 2847–2849 (2016).
42. Subramanian, A. *et al.* Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *PNAS.* **102**, 15545–15550 (2005).
43. Mootha, V. K. *et al.* PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nat. Genet.* **34**, 267–273 (2003).
44. Eden, E., Lipson, D., Yogeve, S. & Yakhini, Z. Discovering Motifs in Ranked Lists of DNA sequences. *PLoS Computational Biology.* **3**, e39 (2007).
45. Eden, E., Navon, R., Steinfeld, I., Lipson, D. & Yakhini, Z. GOrilla: A Tool For Discovery And Visualization of Enriched GO Terms in Ranked Gene Lists. *BMC Bioinformatics.* **10**, 48 (2009).
46. Marincevic-Zuniga, Y. *et al.* Transcriptome sequencing in pediatric acute lymphoblastic leukemia identifies fusion genes associated with distinct DNA methylation profiles. *J. Hematol. Oncol.* **10**, 148 (2017).
47. Swerdlow, S. H. *et al.* (Eds): WHO Classification of Tumours of Haematopoietic and Lymphoid Tissues (Revised 4th Edition) IARC: Lyon (2017).
48. Ramezani-Rad, P. *et al.* SOX4 enables oncogenic survival signals in acute lymphoblastic leukemia. *Blood.* **121**, 148–155 (2013).
49. Kuhns, D. B. *et al.* Cytoskeletal abnormalities and neutrophil dysfunction in WDR1 deficiency. *Blood.* **128**, 2135–2143 (2016).
50. Standing, A. S. I. *et al.* Autoinflammatory periodic fever, immunodeficiency, and thrombocytopenia (PFIT) caused by mutation in actin-regulatory gene WDR1. *J. Exp. Med.* **214**, 59–71 (2017).
51. Jia, M. *et al.* High expression of Midkine (MK) indicates poor prognosis in childhood acute lymphoblastic leukemia. *Hematology.* **21**, 69–77 (2016).
52. Hu, R. *et al.* Increased drug efflux along with midkine gene high expression in childhood B-lineage acute lymphoblastic leukemia cells. *Int. J. Hematol.* **92**, 105–110 (2010).
53. Neveu, B. *et al.* CLIC5: a novel ETV6 target gene in childhood acute lymphoblastic leukemia. *Haematologica.* **101**, 1534–1543 (2016).
54. Holleman, A. *et al.* Expression of the outcome predictor in acute leukemia 1 (OPAL1) gene is not an independent prognostic factor in patients treated according to COALL or St Jude protocols. *Blood.* **108**, 1984–1990 (2006).
55. Yeoh, E. J. *et al.* Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling. *Cancer Cell.* **1**, 133–143 (2002).
56. Tse, W., Zhu, W., Chen, H. S. & Cohen, A. A novel gene, AF1q, fused to MLL in t(11;11)(q21;q23), is specifically expressed in leukemic and immature hematopoietic cells. *Blood.* **85**, 650–656 (1995).
57. Ney Garcia, D. R. *et al.* Molecular characterization of KMT2A fusion partner genes in 13 cases of pediatric leukemia with complex or cryptic karyotypes. *Hematol. Oncol.* **35**, 760–768 (2017).
58. Petit, A. *et al.* Functional analysis of the NUP98-CCDC28A fusion protein. *Haematologica.* **97**, 379–387 (2012).
59. Fridley, B. L. *et al.* Weinshilboum RM. Gene set analysis of purine and pyrimidine antimetabolites cancer therapies. *Pharmacogenet. Genomics.* **21**, 701–712 (2011).

60. Liu, J. *et al.* PIG7 promotes leukemia cell chemosensitivity via lysosomal membrane permeabilization. *Oncotarget*. **7**, 4841–4859 (2016).
61. Chin, L. S., Raynor, M. C., Wei, X., Chen, H. Q. & Li, L. Hrs interacts with sorting nexin 1 and regulates degradation of epidermal growth factor receptor. *J. Biol. Chem.* **276**, 7069–7078 (2001).
62. Chen, Y. S. *et al.* Tiger frog virus ORF080L protein interacts with LITAF and impairs EGF-induced EGFR degradation. *Virus Res.* **217**, 133–142 (2016).
63. Toft, N. *et al.* Results of NOPHO ALL2008 treatment for patients aged 1–45 years with acute lymphoblastic leukemia. *Leukemia*. **32**, 606–615 (2018).
64. Schmiegelow, K. *et al.* Nordic Society of Paediatric Haematology and Oncology. Long-term results of NOPHO ALL-92 and ALL-2000 studies of childhood acute lymphoblastic leukemia. *Leukemia*. **24**, 345–354 (2010).
65. Wang, X. *et al.* Gene expression profiling and chromatin immunoprecipitation identify DBN1, SETMAR and HIG2 as direct targets of SOX11 in mantle cell lymphoma. *PLoS One*. **5**, e14085 (2010).
66. Grönroos, T. Transcriptional regulation and cell signaling in acute lymphoblastic leukemia and hematopoiesis. Tampere University Dissertations: 74 (Tampere University, (2019)

## Acknowledgements

We thank the Finnish Functional Genomics Centre, which is supported by the University of Turku, Åbo Akademi University, and Biocenter Finland. This work was supported by the Competitive State Research Financing of the Expert Responsibility area of Tampere University Hospital, as well as by grants from the Academy of Finland (project no. 276634 (M.H.), no. 277816 (O.L.)), the Finnish Cultural Foundation (Interdisciplinary Science Workshops (M.H.), Pirkanmaa Regional Fund, (T.G.)), the Sigrid Juselius Foundation (M.H. and O.L.), and the Cancer Society of Finland (MH, OL), the Paulo Foundation (O.L.), the Foundation for Pediatric Research (O.L.), the Jane and Aatos Erkko Foundation (O.L.), Tampere University (O.L.) and University of Eastern Finland (M.H.), Väre Foundation for Pediatric Cancer Research (T.G., A.M.), and the Finnish Hematology Association (T.G.). Grönroos Toni, Transcriptional regulation and cell signaling in acute lymphoblastic leukemia and hematopoiesis, Tampere University Dissertations: 74 (Tampere University, 2019)<sup>66</sup>.

## Author contributions

T.G., A.M.: Study design, data collection, data interpretation, figure designing, literature search, and writing (equal contribution as first authors). S.L.: Data analysis, data interpretation, figure designing, literature search, and writing. J.M.: Data collection, data analysis, figure designing, and writing. A.N.: Data collection, data analysis, data interpretation, figure designing, and writing. L.O.: Data analysis, data interpretation, figure designing, and writing. S.R.: Data collection, data analysis, data interpretation, figure designing, and writing. Y.M-Z., J.N.: Data collection, data analysis, and data interpretation. V.P.: Data collection and data interpretation. T.P.: Data analysis. M.H.: Study design, data interpretation, and writing. O.L.: Study design, data interpretation, literature search, and writing. All authors reviewed the manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary information** is available for this paper at <https://doi.org/10.1038/s41598-020-58970-z>.

**Correspondence** and requests for materials should be addressed to T.G.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020