Supplementary Information

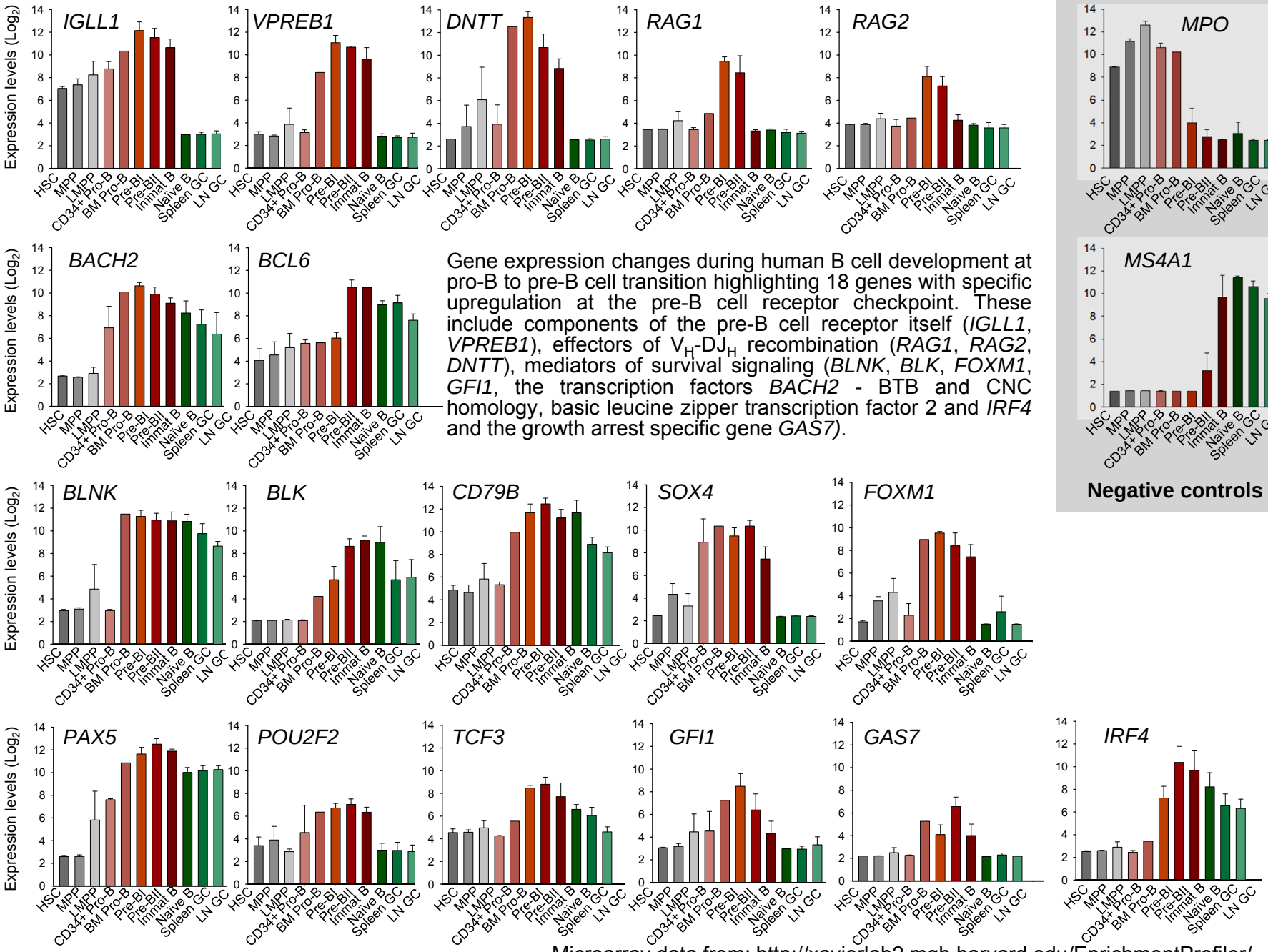
BACH2 mediates negative selection and p53-dependent tumor suppression at the pre-B cell receptor checkpoint

Srividya Swaminathan[1], Chuanxin Huang[2], Huimin Geng[1], Zhengshan Chen[1], Richard Harvey[3], Huining Kang[3] , Carina Ng[1], Björn Titz[4], Christian Hurtz[1] , Mohammed Firas Sadiyah[1], Daniel Nowak[5], Gabriela B. Thoennissen[5,10], Vikki Rand[6], Thomas G. Graeber[4], H. Phillip Koeffler[5], William L Carroll[7], Cheryl L Willman[3], Andrew G. Hall[6], Kazuhiko Igarashi[8,9], Ari Melnick[2], Markus Müschen[1]

Supplementary Figures and Legends 1-22

Supplementary Tables 1-8

Supplementary Figure 1: Identification of genes with pre-B cell receptor checkpoint-specific expression pattern

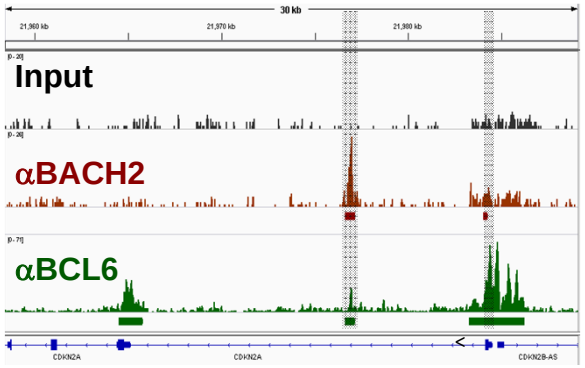


Gene expression changes during human B cell development at pro-B to pre-B cell transition highlighting 18 genes with specific upregulation at the pre-B cell receptor checkpoint. These include components of the pre-B cell receptor itself (*IGLL1*, *VPREB1*), effectors of V_H-DJ_H recombination (*RAG1*, *RAG2*, *DNTT*), mediators of survival signaling (*BLNK*, *BLK*, *FOXM1*, *GFI1*, the transcription factors *BACH2* - BTB and CNC homology, basic leucine zipper transcription factor 2 and *IRF4* and the growth arrest specific gene *GAS7)*.

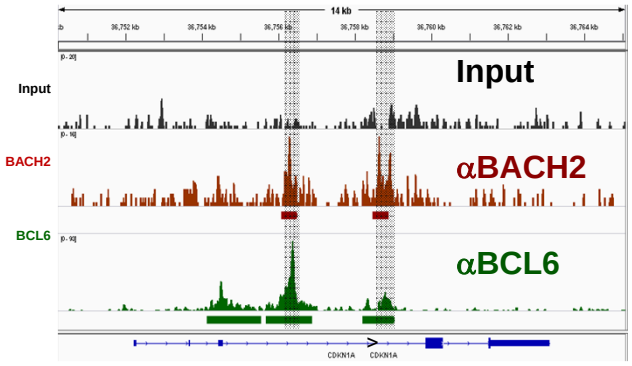
Negative controls

Microarray data from: http://xavierlab2.mgh.harvard.edu/EnrichmentProfiler/

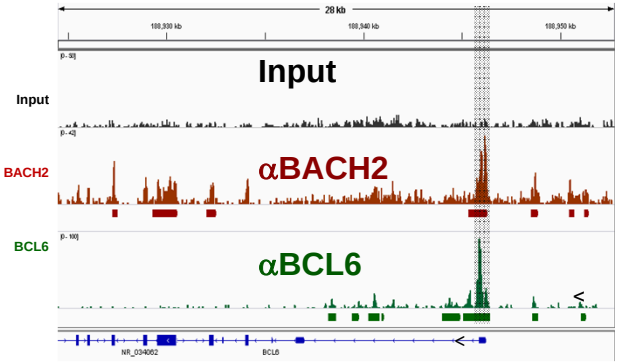
Supplementary Figure 2: BACH2 and BCL6 bind to promoters of identical tumor suppressor genes but have opposing roles on their transcriptional regulation



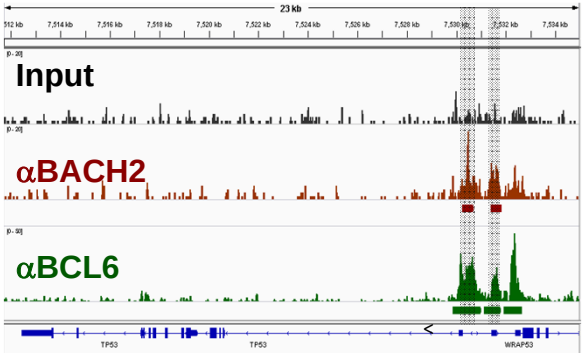
CDKN2A



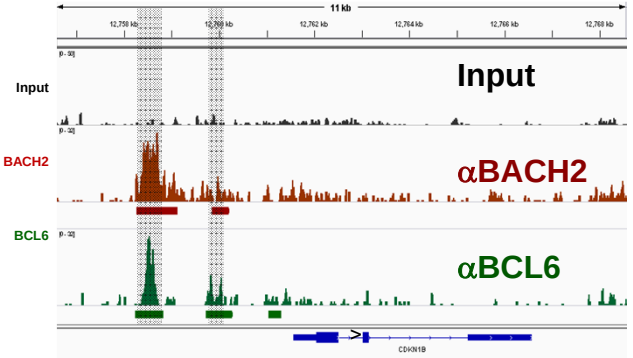
CDKN1A



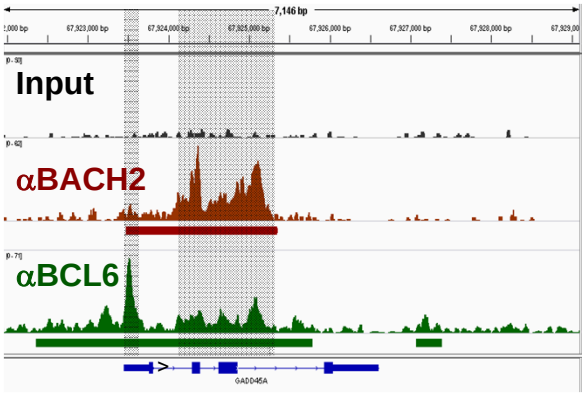
BCL6



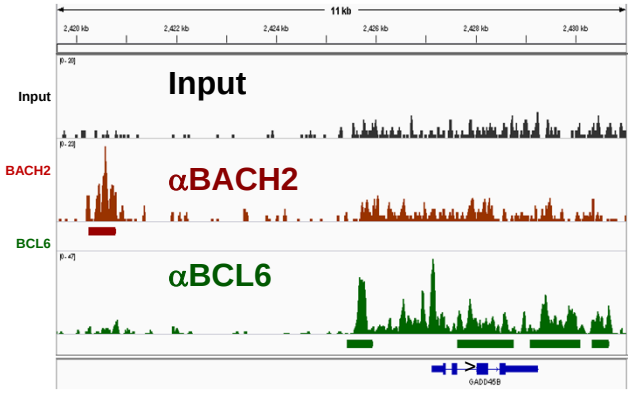
TP53



CDKN1B



GADD45A

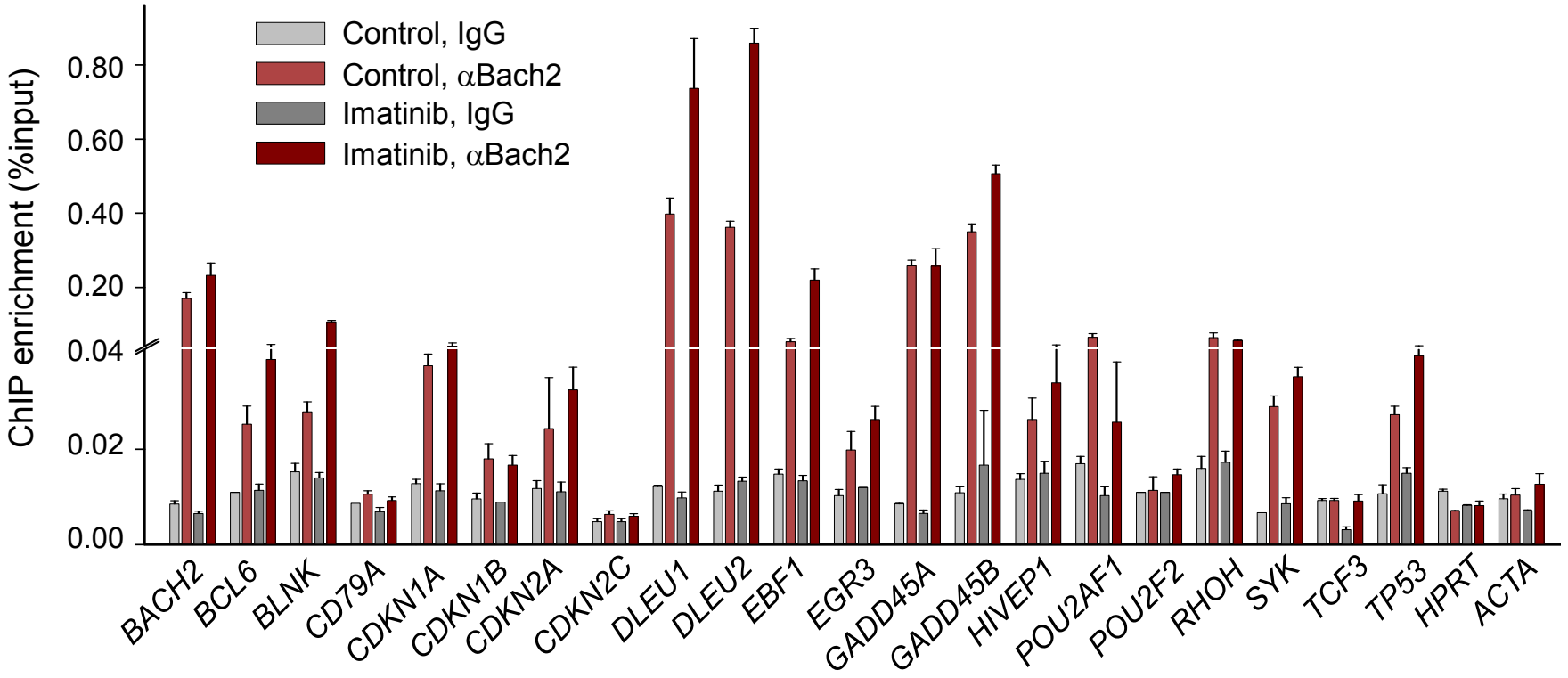


GADD45B

Bach2-regulated genes were identified by microarray analysis of *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ cells. Bcl6-regulated genes were identified by microarray analysis of *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ cells. Among the genes shared by both signatures, only those that show opposing transcriptional fates upon Bach2 and Bcl6 binding were selected.

BACH2 and BCL6 chIP sequencing in the human lymphoma cell line OCI-Ly7 depicting the common tumor suppressor genes which are transcriptional targets of both proteins. Significance of binding peaks was determined by statistical validation using the ChIPseeqer algorithm (underlined with red or green lines; Olivier Elemento *et al.* (http://physiology.med.cornell.edu/faculty/elemento/lab/chipseq.shtml).

Supplementary Figure 3: Single-locus BACH2 qChIP for promoter-binding of genes that show opposite regulation by BACH2 and BCL6 in a primary human Ph⁺ ALL sample



Bach2-regulated genes were identified by microarray analysis of *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ cells. Bcl6-regulated genes were identified by microarray analysis of *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ cells. From the common Bach2- and Bcl6-regulated genes, those were selected that show opposing transcriptional fates upon Bach2 and Bcl6 binding. Quantitative chromatin immunoprecipitation (qChIP) was performed using BACH2-specific antibodies or an IgG control. Enrichment of BACH2 at promoters of these genes in human Ph⁺ ALL cells before (light red) and after (dark red) imatinib treatment is depicted in the bar chart. Target genes exhibiting stronger BACH2 binding are shown in the left panel and the loci with weaker overall BACH2 binding are shown in the right panel.

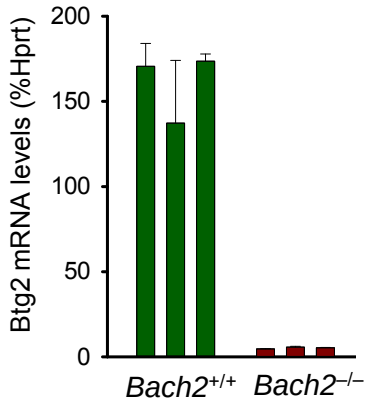
Supplementary Figure 4: Genetic modulation of relative Bach2 vs Bcl6 expression

Right: Relative Bach2 and Bcl6 expression levels were genetically modulated by expressing an inducible vector system for Bach2 in Bcl6$^{+/+}$ and *Bcl6*$^{-/-}$ pre-B ALL cells (*BCR-ABL1*). *Bach2*$^{+/+}$, *Bach2*$^{-/-}$, and Bach2-overexpressing *Bcl6*$^{+/+}$ and Bach2-overexpressing *Bcl6*$^{-/-}$ pre-B ALL cells (*BCR-ABL1)* were studied for phenotypic differences using microarray analysis. Cloning of the tamoxifen-inducible Bach2-ER^{T2} vector is described in the methods section. A schematic for the relative expression levels of Bach2 (red) and Bcl6 (green) is depicted at the top of the heatmap. Genes shown in the heatmap are selected based on three criteria:

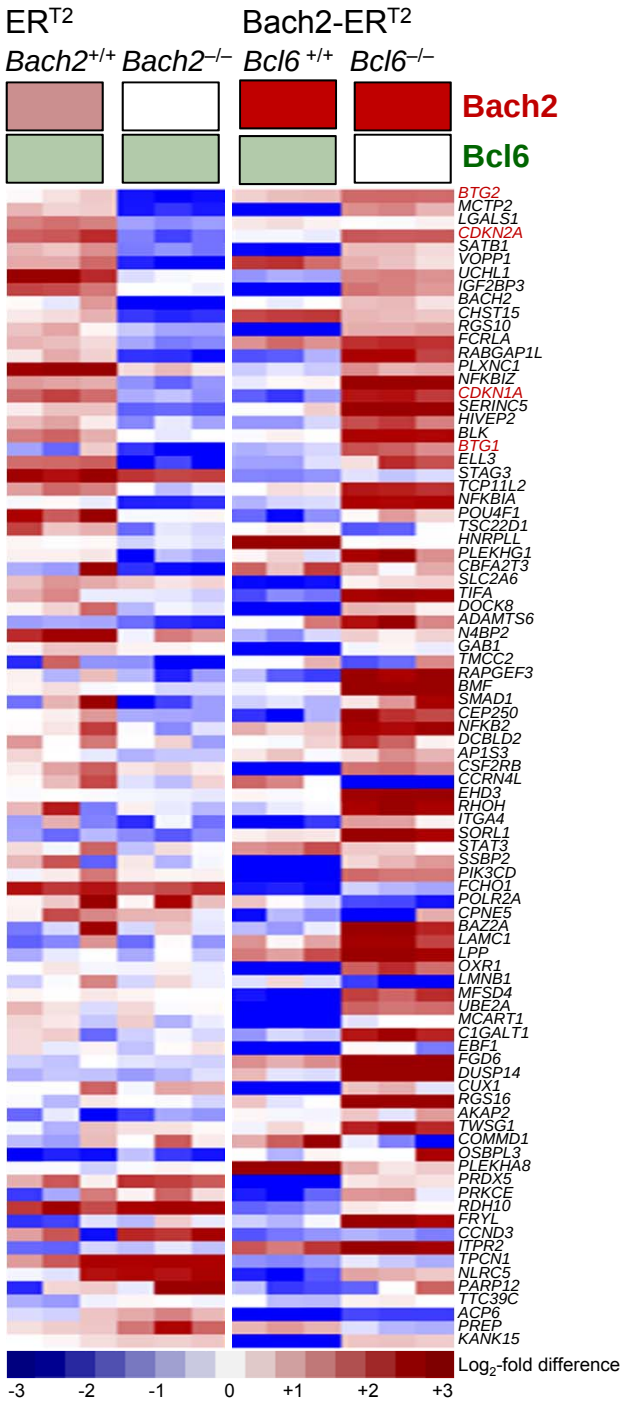
(1) Direct targets of BACH2 and BCL6 as identified by the BACH2 and BCL6 ChIPseq.

(2) Downregulated by Bcl6 (fold change >1.5 and *P*<0.05 by t-test; GEO accession no. GSE20987: *BCR-ABL1*-transformed B cell precursors from *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ mice)

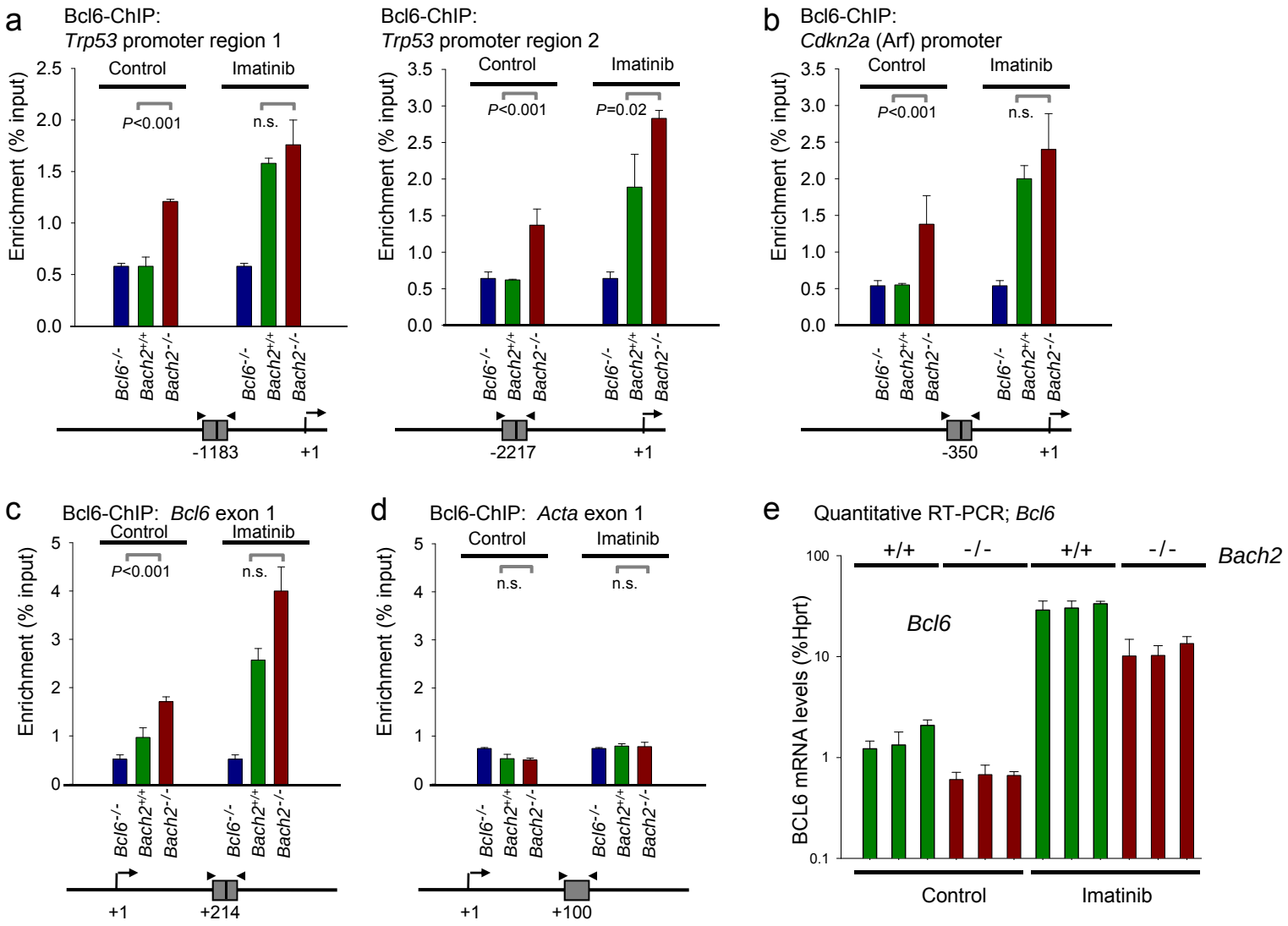
(3) Upregulated by Bach2 (fold change >1.5, *P*<0.05) or no changes by Bach2 (fold change <1.5; GEO accession no. GSE30883: *BCR-ABL1*-transformed B cell precursors from Bach2$^{+/+}$ and Bach2$^{-/-}$ mice).



Above: Expression of Btg2 mRNA in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ pre-B ALL cells *(BCR-ABL1).* Btg2 mRNA levels are plotted as percentage of the reference gene Hprt.



Supplementary Figure 5: Bach2 inhibits recruitment of Bcl6 to *Cdkn2a* (*Arf*) and *Tp53* promoters

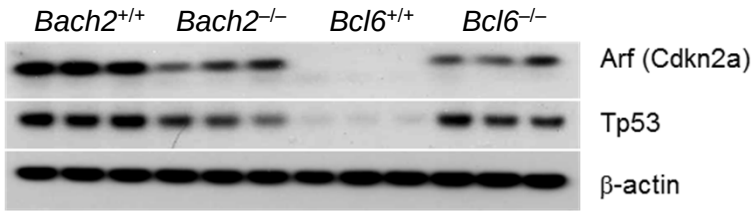


Supplementary Figure 5: Bach2 inhibits recruitment of Bcl6 to *Cdkn2a* (*Arf*) and *Tp53* promoters, Legend

(**a**) Bcl6-ChIP was performed at two regions of the *Tp53* promoter, (**b**) one region of the *Cdkn2a* (Arf) promoter, (**c**) at *Bcl6* exon 1 (positive control for Bcl6 autoregulation, and (**d**) Acta exon 1 (negative control for non-specific binding. In all conditions a-d, *BCR-ABL1*-transformed pre-B ALL cells were studied in the presence or absence of Imatinib treatment (2 μmol l^{-1}, 24 hours). ChIP was performed using anti-Bcl6 IgG (Santa Cruz, Clone N3) or IgG controls. *BCR-ABL1*-transformed pre-B ALL cells were derived from *Bcl6*$^{-/-}$ (additional negative control) mice and *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ mice.

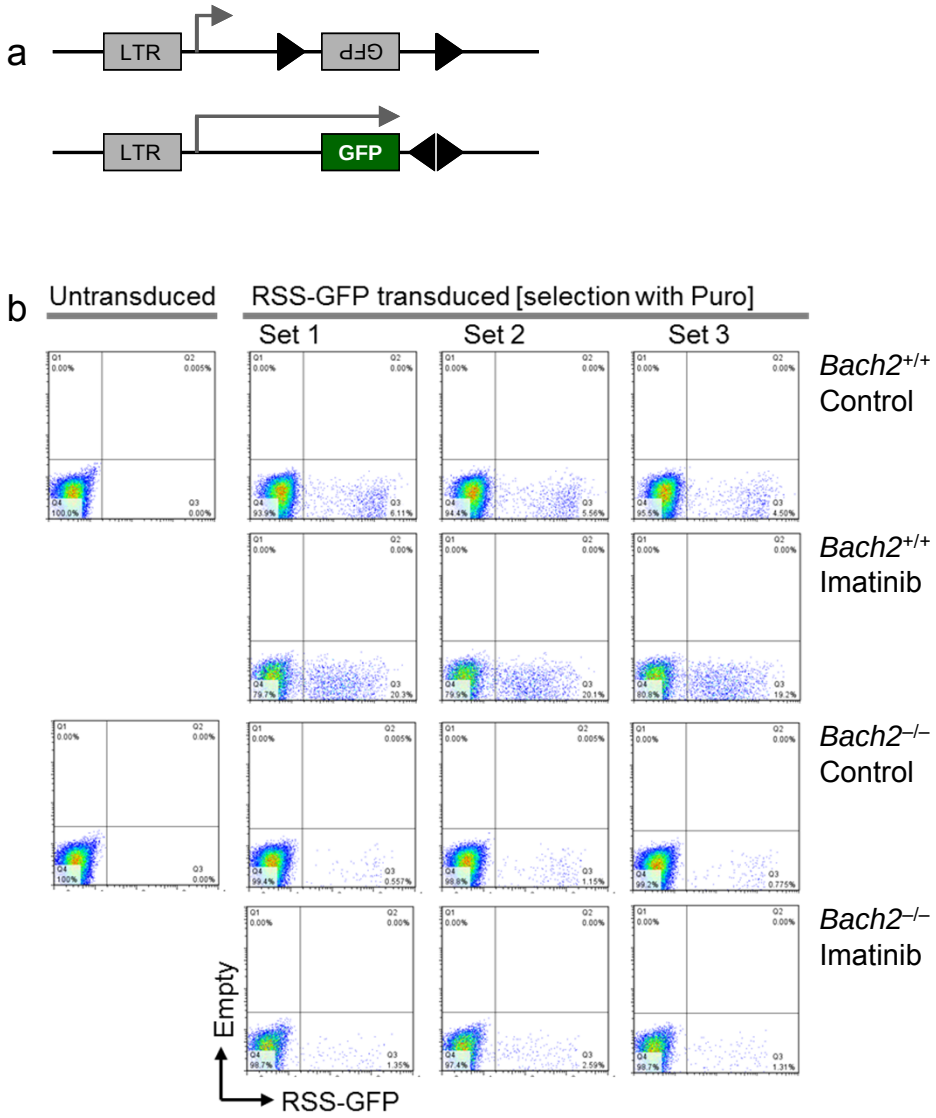
For *Tp53*, *Cdkn2a* (Arf) promoters, specific recruitment of Bcl6 was compared in the presence (*Bach2*$^{+/+}$) and absence (*Bach2*$^{-/-}$) of Bach2 and indicated on the y-axis as enrichment (% of input). Treatment with Imatinib results in high expression levels of Bcl6. In this case, Bcl6 was highly enriched at *Tp53* and *Cdkn2a* promoters regardless of Bach2. In the absence of Imatinib, expression levels of Bach2 are low. In this case, Bcl6 was only strongly enriched in the absence, but not in the presence of Bach2. These findings indicate that Bach2 negatively regulates binding of Bcl6 to *Tp53* and *Cdkn2a* promoters unless promoters are saturated in the presence of very high Bcl6 expression levels. (**e**) To address the possibility that higher enrichment of Bcl6 at *Tp53* and *Cdkn2a* promoters in *Bach2*$^{-/-}$ leukemia cells may reflect higher expression levels of Bcl6 in these cells, we performed qRT-PCR analysis of Bcl6 in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ leukemia cells for Bcl6 mRNA.

Supplementary Figure 6: Bach2 reverses Bcl6-mediated repression of Arf and Tp53



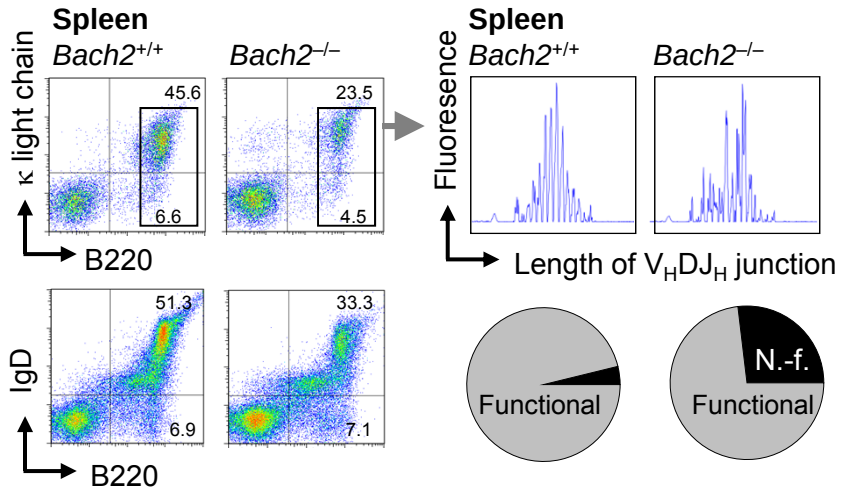
Western blot for Arf (Cdkn2a) and Tp53 in *Bach2*$^{+/+}$, *Bach2*$^{-/-}$ and *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ pre-B cells transformed with *BCR-ABL1*. β-actin was used as loading control. It should be noted that *Bach2*$^{+/+}$ (C57BL6/SV129) and *Bcl6*$^{+/+}$ (C57BL6/SV129) pre-B ALL cells have significantly different baseline levels of Arf and Tp53 expression, which likely reflects the different admixture of C57BL6 and SV129 in these mice. For this reason, exposure of Western blots was adjusted for comparable Arf and Tp53 baseline expression levels in *Bach2*$^{+/+}$ and *Bcl6*$^{+/+}$ pre-B ALL cells.

Supplementary Figure 7: Measurement of V_H-DJ_H recombinase activity in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ pre-B cells



(**a**) A retroviral vector that reports recombinase activity in transduced cells. The vector carries an inverted GFP flanked by recombination signal sequences (RSS; depicted as black triangles) and a puromycin resistance cassette. (**b**) Recombinase activity in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ pre-B ALL cells (*BCR-ABL1*) transduced with the vector shown in (a), before and after imatinib treatment.

Supplementary Figure 8: Sequence analysis of V_H-DJ_H junctions in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ B cells



Bone marrow

Bach2$^{+/+}$	Bach2$^{-/-}$
CARSHGSKGYDMDYW	CARGEGSN#YFDYW
CARGLRSHYAMDYW	CARGYYGSRRGLYYFDYW
CARSECDGYYAYAMDYW	CARGG#FAYW
CARGGGRYLLDYW	CARSGAYNSKG#AMDYW
CARSYGSSYDYFDYW	CARGYYGSRRGLYYFDYW
CARGGAYW	CARSNTQAT#FAYW
CARRG*#YAMDYW	CARDYYGSSWYFDVW
CARSITGTSRFAYW	CARPFNYYGKGSG#YFDYW
CARFYYDYDDYFDYW	CARFDDGYLAW#FAYW
CARGGYW	CARRSSYLYWYFDVW
CARGQLGLREGAYW	CAREGEQLRLR#FAYW
CARKSGYLLFAYW	CVITTVRYFDYW
CARGG*GG#DYW	CASGGITTY#YYYAMDYW
CARSDYDAFAYW	CARGGFAYW
CARGDYGYGFAYW	CALGNYTGDY#W
CAREGITTVVAERYFDVW	CASGGITTVVAT#FAYW
CAAVDGYYFDYW	CARRLRYYFDYW
CARDSSGYFDYW	CALPLYAMDYW#
CAREGDVGPAWFAYW	CARFYYYGSSYWYFDVW
CARDRVFDYW	CARGLR##YFDYW
CARGDYDDYYAMDYW	CAREGRGLRAWFAYW
CARWNYYGSSYGYFDVW	CARFYGYYFDYW
CAREGDDYDGYAMDYW	CVKGG*LRL#WFAYW
CARFEEYFDYW	CARHSLRLRGFAYW
CASPRITYTN#YAMDYW	CASPYYG**LW#YW
CTRNWDGYAMDYW	CARGIYYARQKYFDVW

P=0.00189

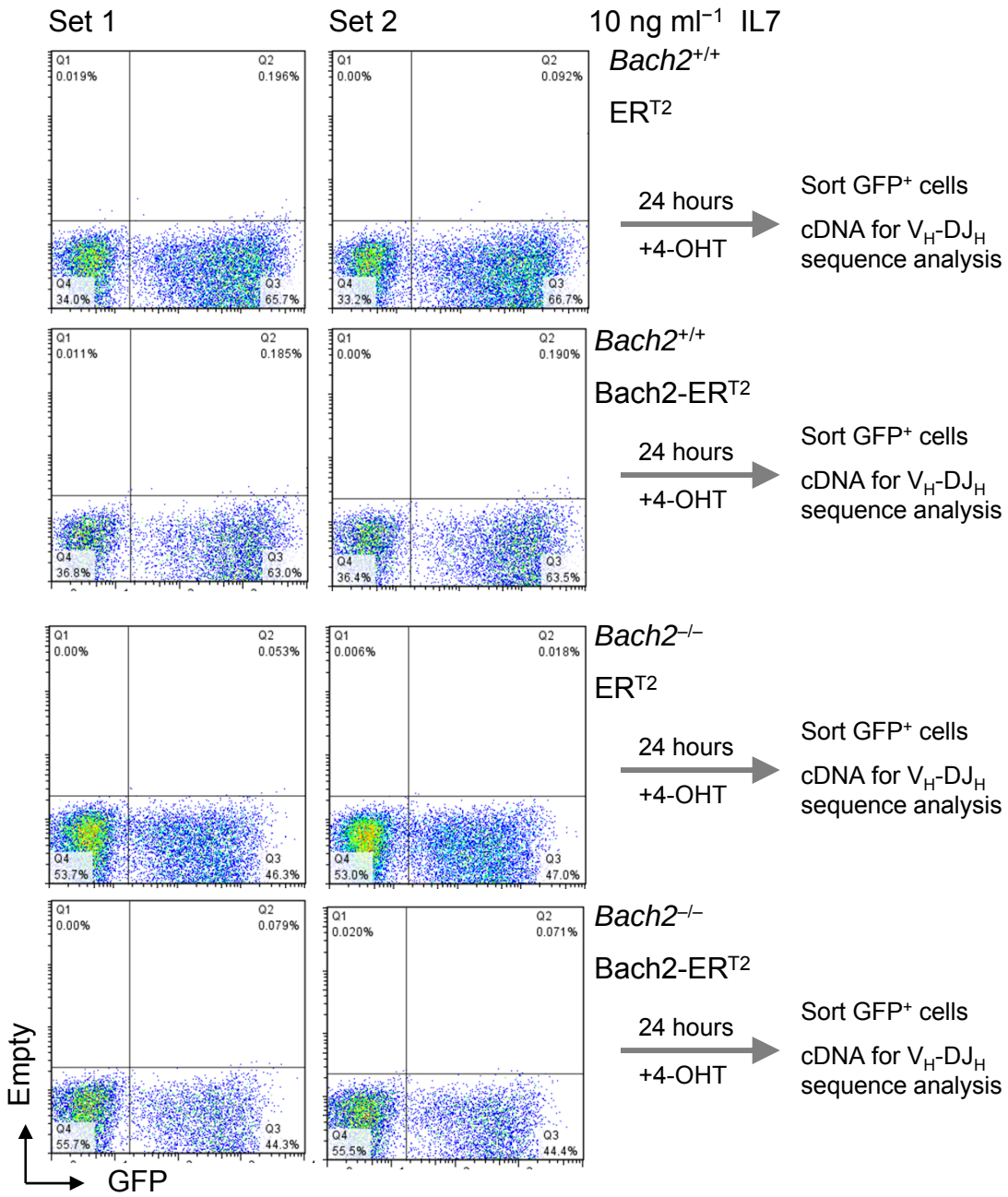
Spleen

Bach2$^{+/+}$	Bach2$^{-/-}$
CTRAYGSLFDYW	CARFGSWYFDVW
CARRTGYYFDYW	CARGLITTVPNYYAMDYW
CARWITTEGYFDVW	CARGG#FAYW
CARGGNYGYFDVW	CARAFYYAMDYW
CARSGTYW	CARKDIPRAAMDYW
CARGDYW	CARDIPYYGIDYW
CAKAFTTVVFDYW	CARGYYGK*##DYW
CTRGGGYW	CARGELFDYW
CAHSSGSPYFDYW	CVLYDG#YEYFDVW
STR*GSSTMITG#AMDYW	CARSSYGNWGYFDVW
CARGLDYDGGYAMDYW	CARRVYYDYDG#YFDYW
CARGDYW	CAKITTVVATDFDVW
CARDGNWYFDVW	CARNLLHEPRDYW
CARRDYYYGSSYLGAMDYW	CAGITTVGKPFDYW
CARSNWDYFDYW	CARYPLLR**#FFDYW
CARSGSYYGSSFAYW	CARPSPLLR**L##YAMDYW
CARSLYDYDPSWFAYW	CARPDGYYRGFAYW
CARKGQGDPRFAYW	CARDYDYFFFDYW
CARWGTVPAMDYW	CARSGLDYDGRGYYFDYW
CARSRGSITTVVATDYFDYW	CARGGFAYW
CACYYDYDVGFAYW	CARSKHYYGSSSYYFDYW
CARTRDYYGSSSYAMDYW	CARRGTTYAMDYW
CARSRLQLGYYYAMDYW	CARYSNYAMDYW
CAQGGLLLRSFAYW	CARDGSSYFDYW
CTCSWNYYAMDYW	CARDFITTVVAT#FDYW
CAREGYGSSYGGYW	CTIYYGNLGFAYW

P=0.02032

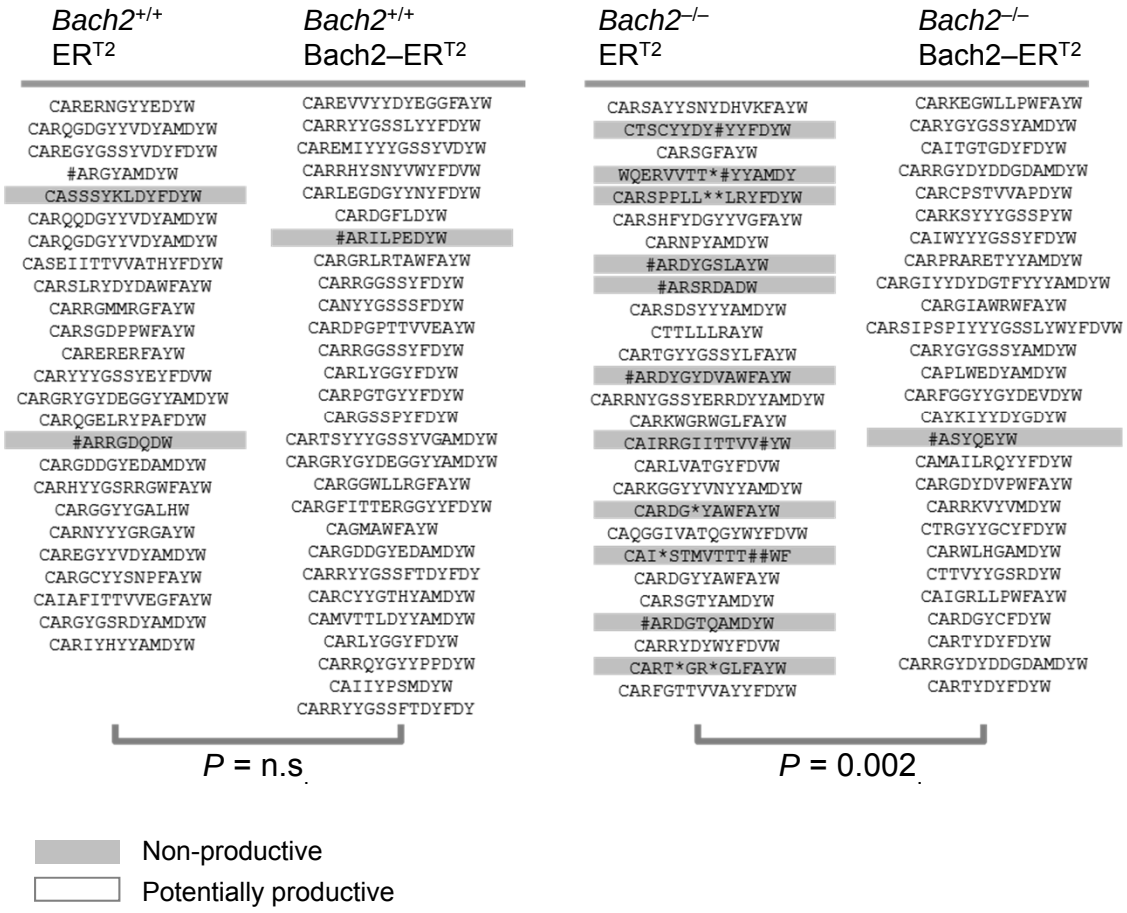
Non-productive
Potentially productive

Supplementary Figure 9: Inducible overexpression of Bach2



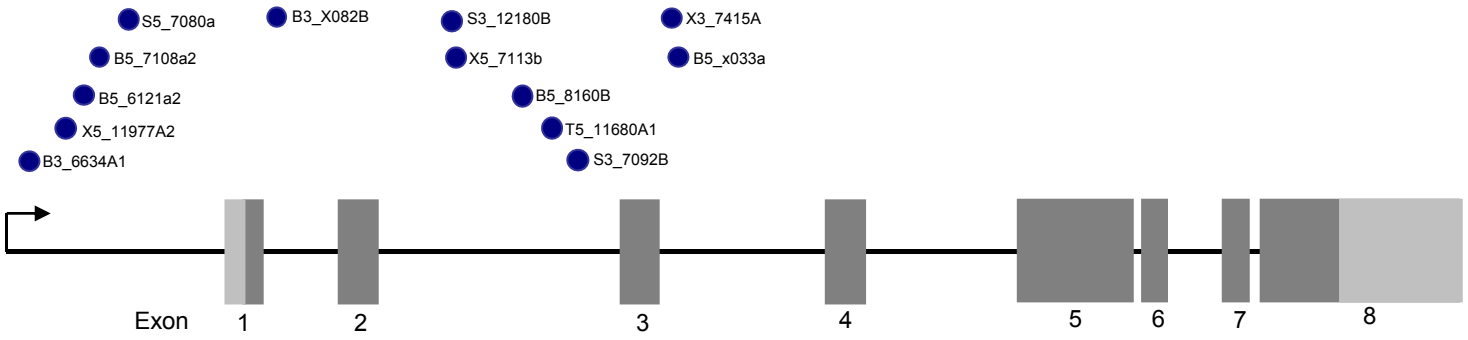
Bach2$^{+/+}$ and *Bach2*$^{-/-}$ IL-7-dependent pre-B cells were transduced with empty vector (EV-ER^{T2}) or with Bach2-ER^{T2}. Induction with tamoxifen for 24 hours resulted in Bach2 translocation to the nucleus in the groups carrying the Bach2-ER^{T2} construct. Following this, V_H-DJ_H junctions were sequenced.

Supplementary Figure 10: Inducible overexpression of Bach2 rescues the negative selection defect seen in Bach2$^{-/-}$ pre-B cells



The sequences of the V_H-DJ_H junctions obtained after carrying out the rescue experiment as described in Supplementary Figure 9. Non-productive rearrangements are shaded gray while productive ones are indicated by white background.

Supplementary Figure 11: Common retrovirus integration sites in the *Bach2* locus in B lymphoid malignancies

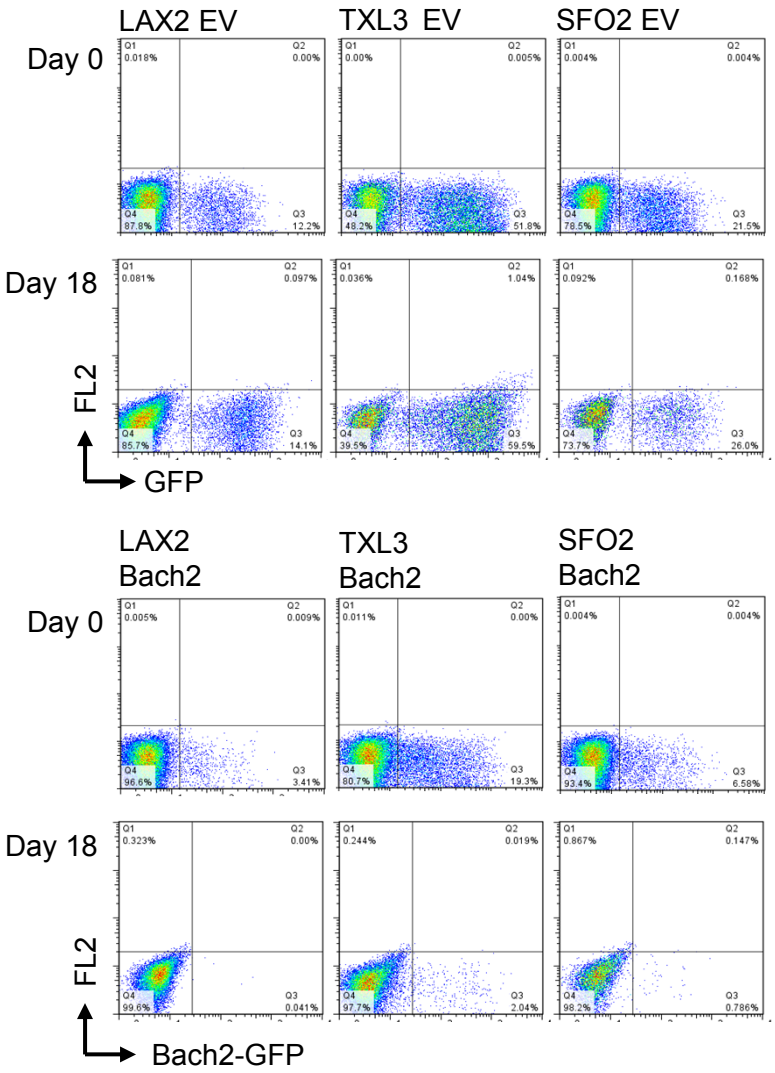


Sequence of 50 bp regions flanking the common integration sites of Bach2

	5' to CIS	3' to CIS
B3_6634A1	TTAGAATACTAAACGTGACACTCCTGGCTGACTGGCATTAATGCACCTTT	CAAAAGTGCCCAGTTTTTCCCTTGGAAGTAATTTCAGCAGCAATCGGCTA
X5_11977A2	TCAGGCTGCCAGGCTTTTGTGGCCACTGTTCTTACCCACCAAGCCACCTC	ATTGTCCCTAATATTCATGGATTTTTTTTGAAGATTTAATAATGGTTGAA
B5_6121a2	GGCTACTACCACACACACACACACACACACACACACACACACACACACAC	ACACACACACACACACACTGCATTCCTCTGGAGGCATAATAGGAGTGTGG
B5_7108a2	AAAATGTGTTAAAATGTTAGTCATCACACTATATGACTAAGGACCTATGG	TTGTTGTATCTTTTGCTCTTCCAGCATTGCTGAACAGCTTTCCACATTAT
S5_7080a	GACATTAAAGGTTGGTCCAGGCCTTGAGACTGCCTCATCTGAAAGCCTGC	TGGTACCCACCCCCCACTCCACCCCTGGGCTCTCTTGCCTGCTACTCTCC
S3_12180B	TGGCCCTACATGTACCTGAATGTATTCCTCTGTACCTTAATACTTTTAAT	AGAACTGTTTCCTTCCCTGACTTCTAAATTGTCAAGAAAAAAAAATACTG
X5_7113b	TATGTCTGAGAATGAACAGACAAAAGTTCTAACTCATGAAGCTTATACGC	AGCCAGGGAGGAAAAAAGATAAGCAAGAAGCAAACGGATGGTGCTGTGTT
B5_8160B	TTATTTCCAGGAGCCCTCAGGGAAGACCCCAAACACTCACTGGTGACATG	GGTCATATGTGCCAGATATGATCTTGTGACTCTGGGACATGACAGGATTG
T5_11680A1	GCAGGTACCAGTGTCACTTCAGAACGCCTCTTATCAAGCAACAACAATTG	TATTATTTTGGTTAAATTTCCACCCTTGGACTGATGTCTTTTTATATTCA
S3_7092B	GGTTGTTGGGGAAGAGAGCAGCTCGGCAGAGATCAGTTAAGAGTACAGTT	TCTCTTGGAAGGAAACAGAAGTCCTTTTACTTTGAGAATAGTAAGTCCAG
X3_7415A	AGCCCTGCAGGGTGTGCGTCCCCTACCAGAGAGTGATGGGCCAGTTCCTC	TAGTTTCTGGATAATGAGCCAGTGCACAGGTAACGATGCCTTTCTCACGG
B5_x033a	GATAAGTCAACCTTGTTTTCCATGAAAGTCCTGAGACAGAAAAAGAATGA	AAGTACAAACAGTGACTGGAAAAAGTGTCACACACTTGAGCCACTTTGCA

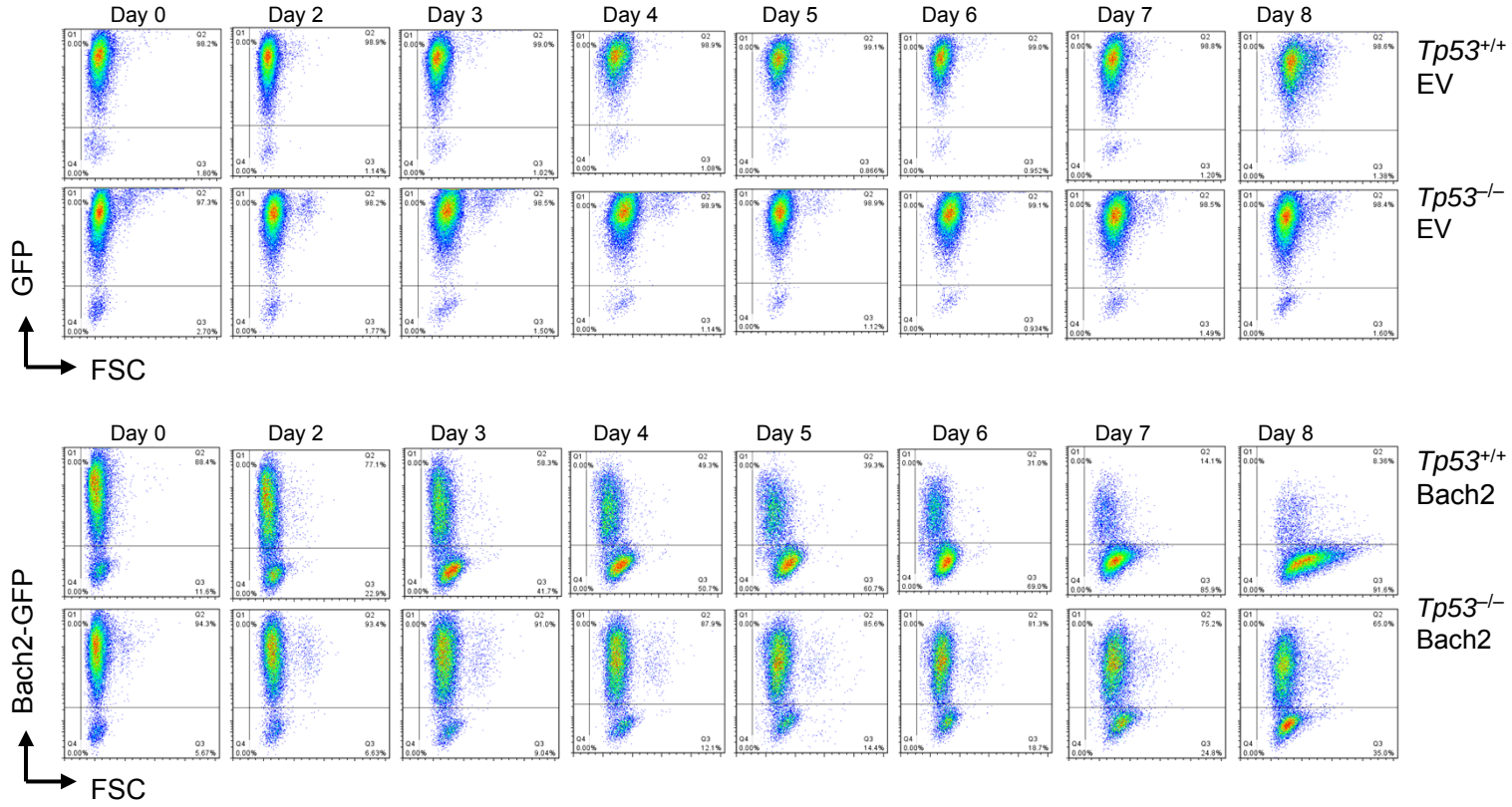
The data of Retrovirus Integration Sites (RISs) in the RTCGD is obtained from using a high-throughput inverse PCR method or splinkerette method. Independent RISs were cloned and sequenced from tumor and then added to RTCGD. Then, by using public UCSC mouse genome server, these RIS sequences were positioned in the mouse genome and candidate genes were identified.
http://variation.osu.edu/cgi-bin/rtcgd/hits_finder.cgi?dataset=retrovirus&assembly=mm9&mouse_symbol=Bach2&cischeck=1&mouse_chr=chr4

Supplementary Figure 12: Overexpression of Bach2 in primary Ph⁺ ALL cells



Primary human Ph⁺ ALL cells were transduced with Bach2–GFP or GFP empty vector (EV) controls. The ratio of transduced (GFP⁺) and non-transduced (GFP⁻) cells was monitored over time (Days 0-18). Representative FACS analysis of days 0 and 18 showing reduction in percentage of GFP⁺ cells in three patient-derived Ph⁺ ALL cases upon overexpression of Bach2–GFP.

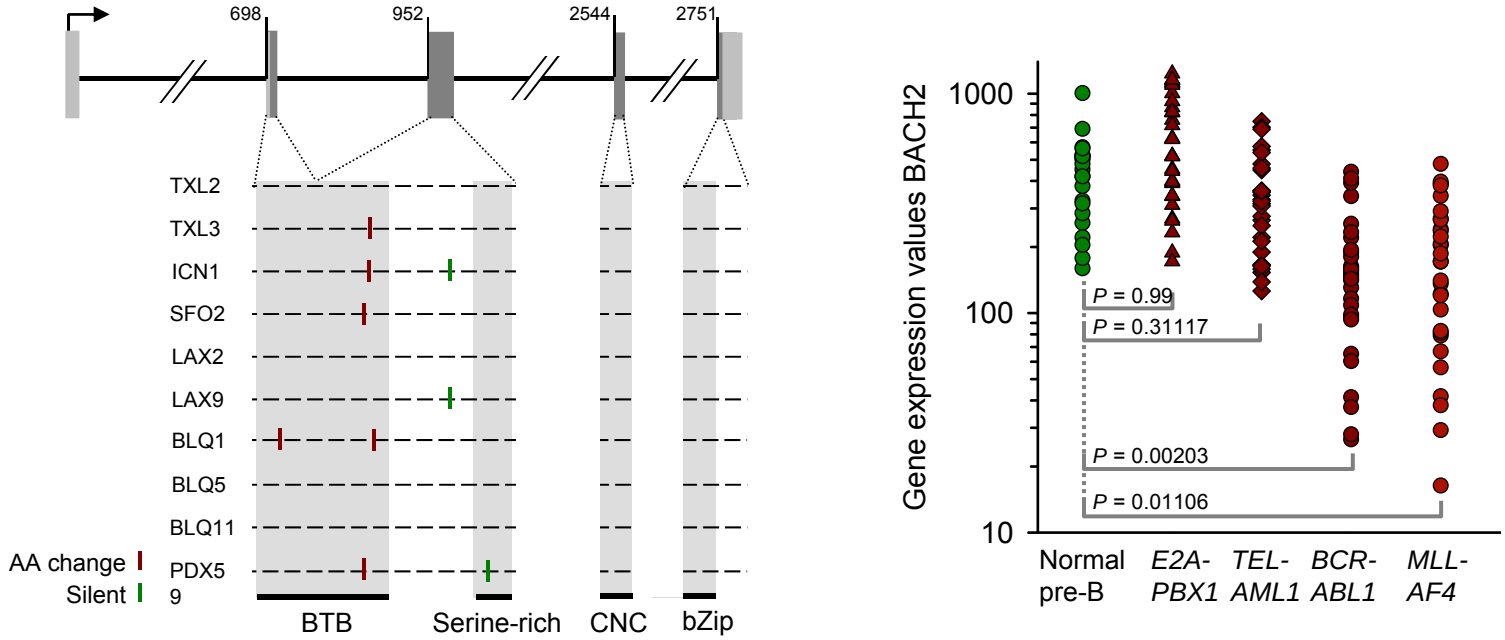
Supplementary Figure 13: Tp53 is required for Bach2–mediated tumor suppression



Tp53$^{+/+}$ and *Tp53*$^{-/-}$ leukemia cells were transduced with Bach2–GFP or GFP empty vector controls. The ratio of transduced (GFP^{+}) and non-transduced (GFP^{-}) cells was monitored over time (Days 0-8).

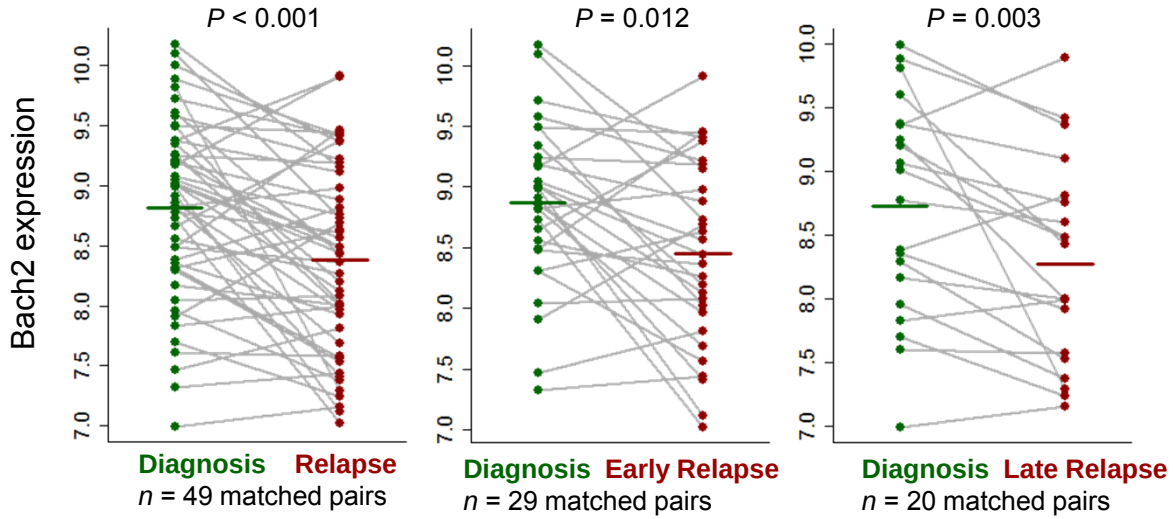
Supplementary Figure 14: Somatic mutations of the *BACH2* gene in primary Ph⁺ ALL samples

Ph⁺ ALL case	N	Sequence	Aa change	Clinical	Karyotype
TXL2	0			at diagnosis	t(9;22)(q34;q11)
TXL3	1	C1039T	R111C	at diagnosis	t(9;22)(q34;q11)
ICN1	2	C1779T C1039T	Silent R111C	at diagnosis	t(9;22)(q34;q11)
SFO2	1	C1039T	R111C	at diagnosis	t(9;22)(q34;q11)
LAX2	0			T315I, Relapse	t(9;22)(q34;q11)
LAX9	1	C1779T	Silent	at diagnosis	t(9;22)(q34;q11)
BLQ1	2	G938T C1039T	S77I R111C	T315I, Relapse	FISH der(9), der(22)
BLQ5	0			T315I, Relapse	FISH der(9), der(22)
BLQ11	0			T315I, Relapse	FISH der(9), der(22)
PDX59	2	A2214G C1039T	Silent R111C		46,XY,del(9)(p13),t(9;22)(q34;q11.2)[12]/46,XY,del(9)(p13),der(9)t(9;22)(q34;q11.2),ider(22)(q10)t(9;22)(q34;q11.2)[4]/46,XY[7]



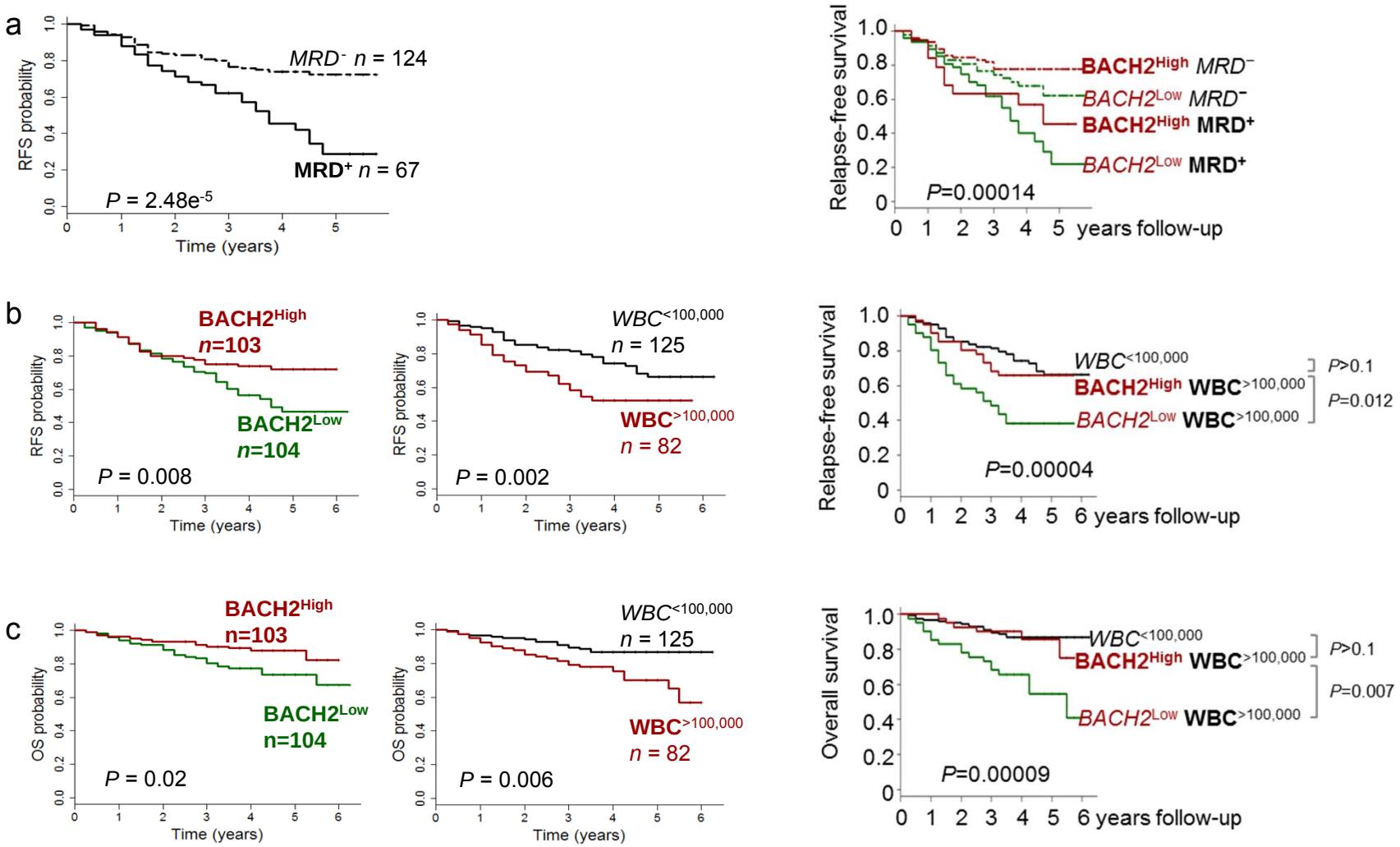
The coding sequence of *BACH2* was amplified and sequenced from 10 primary cases of Ph⁺ ALL. Details of these leukemia cases are summarized in the Table (top). In five cases, point mutations encoding amino acid changes in the *BACH2* BTB domain were found (diagram, bottom left). A metaanalysis of publicly available gene expression data was performed to compare mRNA levels of *BACH2* in normal pre-B cells (green) to pre-B ALL subtypes with different cytogentics (red) (bottom right).

Supplementary Figure 15: Bach2 expression levels are lower at relapse in pediatric ALL patients than that observed at the time of diagnosis



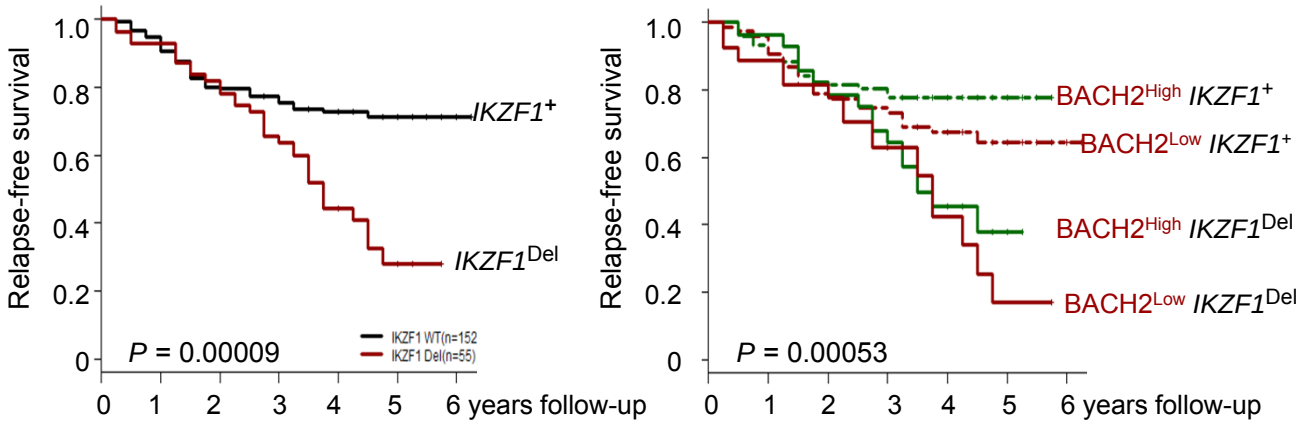
A comparison of the expression levels of Bach2 mRNA at diagnosis and relapse in matched sample pairs of pediatric ALL patients. The left panel shows all 49 pairs of Diagnosis-Relapse samples, the middle and right panels show early (*n* = 29) and late relapses (*n* = 20) respectively. Bach2 mRNA levels are significantly decreased in most patients at relapse demonstrating that low levels of Bach2 expression are associated with relapse of leukemia.

Supplementary Figure 16: Low level of BACH2 is an independent predictor of poor outcome in patients with ALL



Positive minimal residual disease (MRD on day 29 post treatment) and white blood cell count (WBC >100,000) were used as indicators of high risk ALL. (**a**) Kaplan-Meier analysis of relapse free survival (RFS) for patients with and without positive MRD status on day 29 is shown. A multivariate analysis of Bach2 mRNA levels versus MRD status on day 29 is shown in pediatric ALL to compare RFS in patients (P9906) with higher and lower than median mRNA levels of Bach2 (a, right). Multivariate analysis of Bach2 mRNA levels versus WBC count to compare (**b**) relapse free survival (RFS) and (**c**) overall survival (OS) of pediatric patients.

Supplementary Figure 17: Low level of BACH2 is an independent predictor of poor outcome in patients with high risk ALL

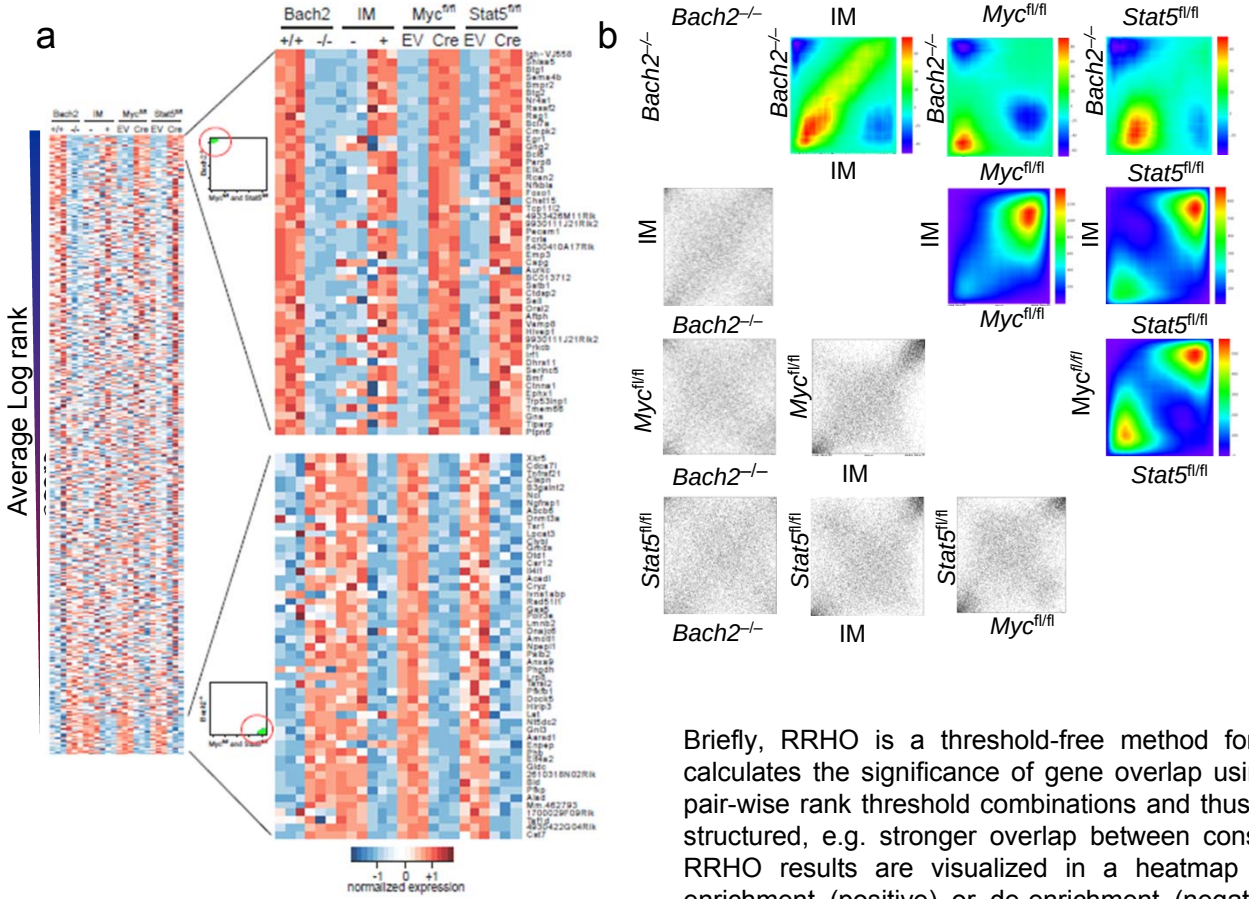


Multivariate analysis showing *BACH2* mRNA levels as an independent predictor of poor clinical outcome in high risk ALL patients with *IKZF1*-deletion (*P* = 0.00053).

Reference:

Mullighan CG, Su X, Zhang J, Radtke I, Phillips LA, Miller CB, Ma J, Liu W, Cheng C, Schulman BA, Harvey RC, Chen IM, Clifford RJ, Carroll WL, Reaman G, Bowman WP, Devidas M, Gerhard DS, Yang W, Relling MV, Shurtleff SA, Campana D, Borowitz MJ, Pui CH, Smith M, Hunger SP, Willman CL, Downing JR; Children's Oncology Group. Deletion of IKZF1 and prognosis in acute lymphoblastic leukemia. *N Engl J Med.* 2009; 360: 470-80.

Supplementary Figure 18: Bach2 recapitulates a common gene expression signature of *BCR-ABL1* inhibition and *Myc/Stat5*-deletion

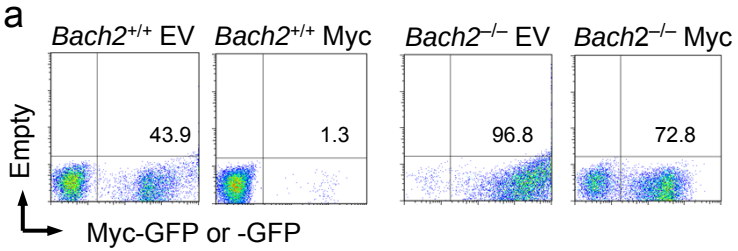


(**a**) The *Bach2*$^{-/-}$ gene expression signature demonstrates a significant overlap with the signatures of *BCR-ABL1* inhibition by imatinib treatment (IM), Myc-deletion (*Myc*$^{fl/fl}$), and Stat5-deletion (*Stat5*$^{fl/fl}$). In addition to enrichment of commonly upregulated genes (lower-left corners in scatter plots and RRHO maps), the *Bach2*$^{-/-}$ signature comparisons also prominently display an enrichment of genes anti-correlated between the *Bach2*$^{-/-}$ and the other gene signatures (blue significance values in left-upper and right-lower corners in scatter plots and heat maps). For this analysis, the overlap between the gene expression signatures was visualized with perturbation rank-based scatter plots (lower left plots) and statistically analyzed with rank-rank hypergeometric overlap (RRHO) heat maps (upper right plots). See Plaisier *et al.* for a detailed explanation of the RRHO procedure.

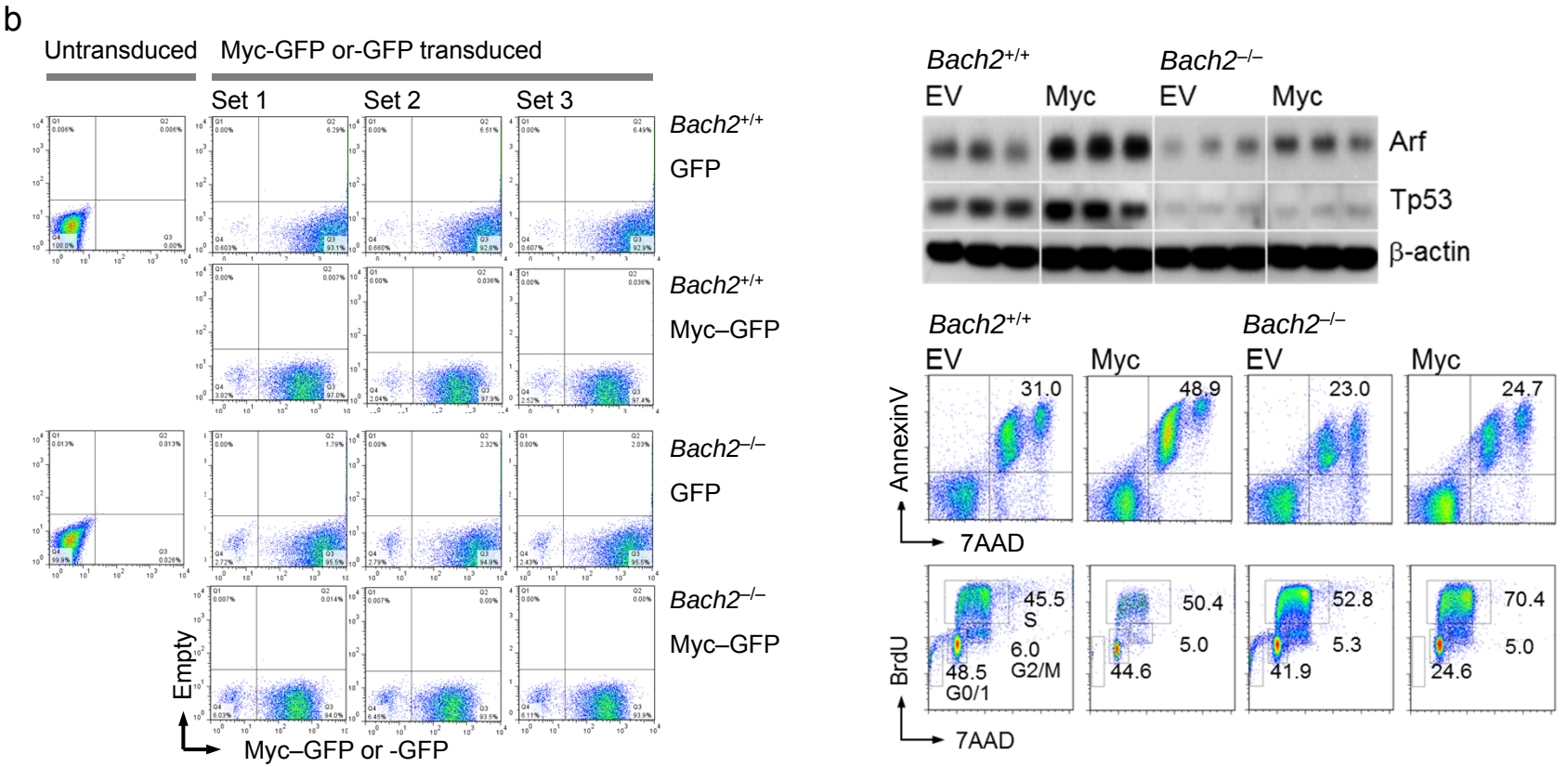
Briefly, RRHO is a threshold-free method for the comparison of ranked gene lists, which calculates the significance of gene overlap using the hypergeometric distribution at all possible pair-wise rank threshold combinations and thus allows for a detailed analysis how the overlap is structured, e.g. stronger overlap between consistently down-regulated vs. up-regulated genes. RRHO results are visualized in a heatmap as the signed log hypergeometric p-value for enrichment (positive) or de-enrichment (negative) compared to random expectations. In the heatmap, the axes coordinates correspond to the rank thresholds used in the hypergeometric calculation. RRHO can be thought of as a two-dimensional analog of the Gene Set Enrichment Analysis (GSEA).

(**b**) Heatmap representation of genes anti-correlated between the *Bach2*$^{-/-}$ and the other gene expression signatures. Each row of the heatmap represents a gene, each column a sample. For this representation, each gene value was mean centered and scaled by its standard deviation within each experimental batch of control and perturbation matched samples. These normalized expression values are color-coded according to the scale bar shown below. To focus on the anti-correlated genes, the genes in the heatmap were ranked by their differential perturbation score rank between the *Bach2*$^{-/-}$ and the *Myc*$^{fl/fl}$ and *Stat5*$^{fl/fl}$ signatures (see methods section for an explanation of the perturbation score). With this, the 50 genes ranked at the top (upper zoom-in) represent anti-correlated genes enriched in the upper-left corners in the RRHO maps (see (a), blue values) and the 50 genes ranked at the bottom (lower zoom-in) represent anti-correlated genes enriched in the lower-right corners in the RRHO maps.

Supplementary Figure 19: Bach2-deficient pre-B cells are permissive to overexpression of Myc

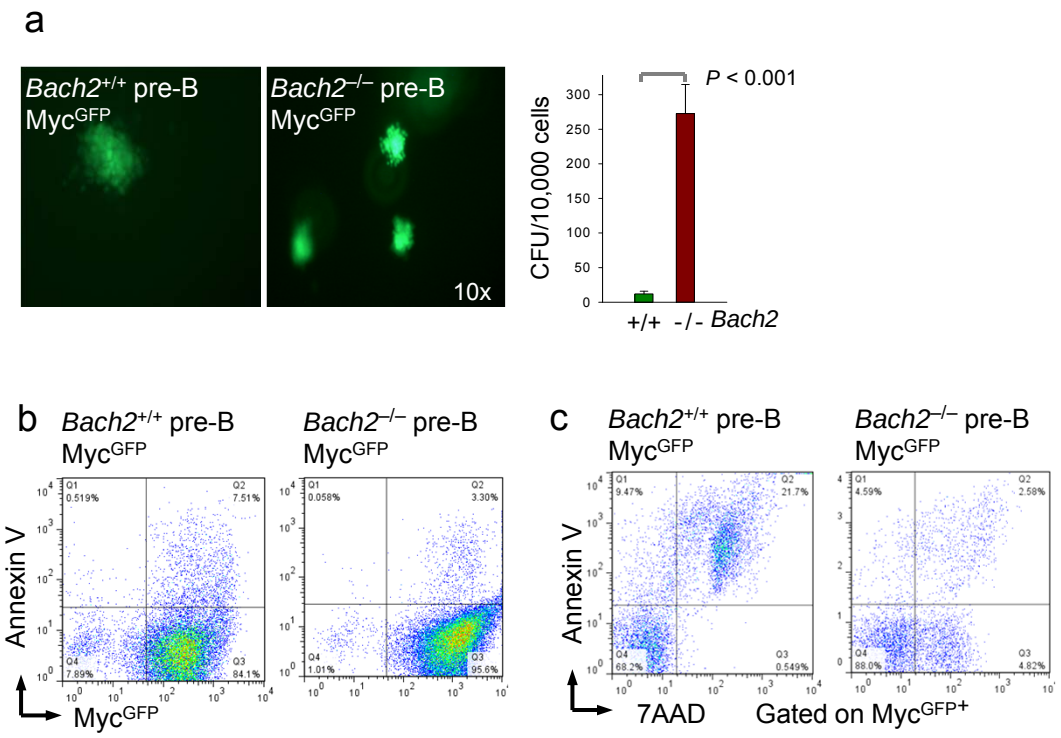


(**a**) Flow cytometry depicting the percentages of GFP+ cells after transduction of *Bach2*+/+ and *Bach2*–/– pre-B ALL cells (*BCR-ABL1*) with Myc–GFP or an empty GFP vector control (EV).



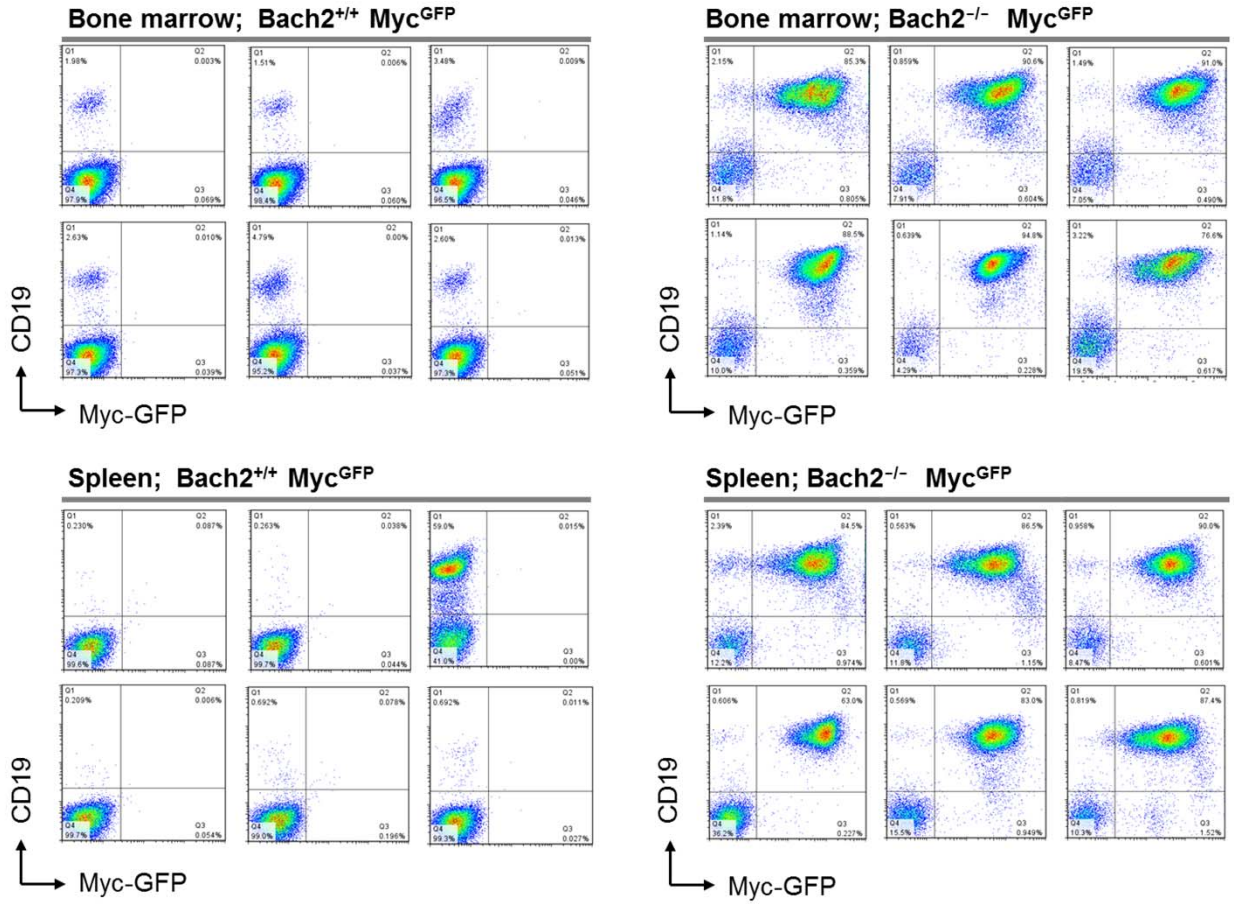
(**b**) For Western blot experiments (top right) and flow cytometry experiments (bottom right), *Bach2*+/+ and *Bach2*–/– pre-B ALL cells (*BCR-ABL1*) were transduced with Myc–GFP or an empty vector (EV) control (GFP), and then sorted based on GFP expression. Western blot analysis of sorted Myc–GFP- and EV–GFP-transduced *Bach2*+/+ and *Bach2*–/– pre-B leukemia cells for expression of Arf and Tp53 was performed using β-actin as loading control (top right). Viability and cell cycle progression of Myc–GFP- and EV–GFP-transduced *Bach2*+/+ and *Bach2*–/– pre-B ALL cells (*BCR-ABL1*) was measured by Annexin V/7AAD staining and BrdU staining (bottom right).

Supplementary Figure 20: Bach2 limits Myc-mediated normal pre-B cell transformation



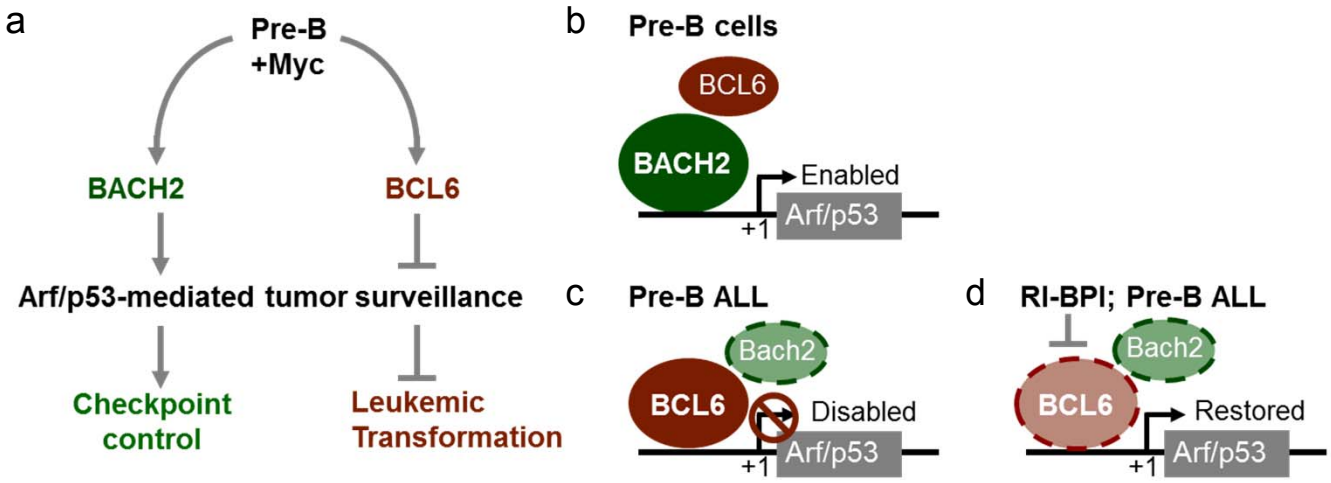
(**a**) Comparison of the abilities of Myc^{GFP}-transduced *Bach2*$^{+/+}$ and *Bach2*$^{–/–}$ IL7-dependent pre-B cells to form colonies on semi-solid methylcellulose medium. Comparison of percentages of Myc-GFP^{+} cells in (**b**) early and (**c**) late apoptotic stages after Myc overexpression in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ IL7-dependent pre-B cells.

Supplementary Figure 21: Flow cytometry of bone marrow and spleen isolated from mice injected with Myc^{GFP} -transduced *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ pre-B cells



Bone marrows (top) and spleens (bottom) of NOG mice injected with Myc^{GFP}-transduced *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ IL7-dependent pre-B cells were studied for CD19^{+} GFP^{+} cells (i.e. leukemia) by flow cytometry.

Supplementary Figure 22: The balance between BACH2 and BCL6 determines pre-B cell receptor checkpoint control and propensity to leukemic transformation



(**a**) Activation of protooncogenes (e.g. Myc) in pre-B cells causes either Arf/p53-dependent failsafe control (BACH2) or leukemic transformation (BCL6). The divergent outcomes of oncogene activation are influenced by BACH2 and its competitive inhibitor, BCL6. (**b**) In normal pre-B cells, BACH2 mediates pre-B cell receptor checkpoint control and tumor suppression through activation of ARF/TP53. (**c**) In pre-B ALL, *BACH2* is frequently inactivated. In the absence of BACH2, the ARF/TP53 pathway is disabled through transcriptional repression by BCL6 and the outcome of Myc-overexpression is leukemic transformation. (**d**) As potential therapeutic intervention, we propose to use BCL6 inhibitors (e.g. RI-BPI) to restore the balance between BACH2 and BCL6 and to relieve BCL6-mediated transcriptional repression of ARF/TP53.

Supplementary Table 1: Overview of clinically-derived human samples of Ph$^+$ ALL used in the study

Case	Cytogenetics	*Oncogene*		Clinical course	Gender/Age
Primary cases					
LAX2	t(9;22)(q34;q11)	*BCR-ABL1;* p210,	T315I	Relapse (Imatinib)	m/38
LAX9	t(9;22)(q34;q11) del(12)(p12;p13)	*BCR-ABL1;* p190,	unmutated	At diganosis	m
SFO2	t(9;22)(q34;q11)	*BCR-ABL1;* p210,	unmutated	at diagnosis	m/7
BLQ1	FISH der(9), der(22)	*BCR-ABL1;* p210,	T315I	Relapse (Imatinib)	
BLQ4	FISH der(9), der(22)	*BCR-ABL1;* p210,	unmutated	Relapse (Imatinib)	f
BLQ5	FISH der(9), der(22)	*BCR-ABL1;* p190,	T315I	Relapse (Imatinib)	f
BLQ6	FISH der(9), der(22)	*BCR-ABL1;* n.d.		Relapse (Imatinib)	m
BLQ11	FISH der(9), der(22)	*BCR-ABL1;* p210,	T315I	Relapse (Imatinib)	m
TXL1	t(9;22)(q34;q11)	*BCR-ABL1;* n.d.,	unmutated	at diagnosis	m/19
TXL2	t(9;22)(q34;q11)	*BCR-ABL1;* p210,	unmutated	at diagnosis	
TXL3	t(9;22)(q34;q11)	*BCR-ABL1;* p210,	unmutated	at diagnosis	
TXL4	t(9;22)(q34;q11)	*BCR-ABL1;* p190,	unmutated	at diagnosis	f/56
ICN1	t(9;22)(q34;q11)	*BCR-ABL1;* p210,	unmutated	at diagnosis	
ICN10	der(9)(q10)t(9;22)(q34;q11)	*BCR-ABL1;* n.d.		at diagnosis	f
NCL37	46,XY,add(9)(q34), del(11)(q23)[3]/46,XY[6]	*ABL1*		Relapse	m/5.3
NCL45	55-59,XY,+X,+1,del(1)(p3) +2,+5,+6,+8,+10,+11,+12	n.d.		Relapse	m/13.7
NCL64	47,XX,t(1;11)(p21.3q22.1)	n.d.		Relapse	f/2.5
NCL68	Fail	n.d.		Relapse	m/11.3
NCL169	45,XY,-20[46]/46,XY[19]	n.d.		Relapse	m/3.5
NCL173	46,XY,t(8;14)(q24;q11)	*MYC*		Relapse	m/2.3
NCL296	Fail	n.d.		Relapse	f/14.1
NCL405	46,XY,t(11;19)(q23;p13.3) [6]/47,idem,+X[3]	*MLL-ENL*	.	Relapse	m/3.4
NCL554	43-46,XY,der(3;17)(q10;q10), +8,16,add(19)(q13) [cp5]/46,	n.d.		Relapse	m/7
NCL578	46,XX[20]	n.d.		Relapse	f/3.7
NCL625	46,XY,t(17;19)(q21;p13)	*TCF3-HLF*		Relapse	m/14

Notes: All primary samples are bone marrow biopsies, blast content >80%; LAX, Los Angeles; BLQ, Bologna; TXL, Berlin; SFO, San Francisco; ICN, Seoul; n.d., not done; f, female; m, male

Supplementary Table 2: Overview of ALL cell lines used in the study

Case	Cytogenetics	*Oncogene*	Clinical course	Gender/Age
BEL-1	46, XX, t (4; 11)(q21; q23), del (6)(q11q21), der (7) I (7)(q10) add (7)(q36), +13, −15	*MLL-AF4*	Relapse	f/41
HPB-Null	n.d.	n.d.	n.d.	m/47
EB2	n.d.	*MYC-IGH*	n.d.	f/7
LC4-1	n.d.	n.d.	n.d.	f/13
MN-60	46(45-47)<2n>XY, dup(1)(q21q41), del(6)(q21), t(8;14)(q24;q32), i(13q)	*MYC-IGH*	Partial Remission	m/20
U-698-M	49(44-50)<2n>XY, +3, +7, -14, +mar, dup(1)(q43q21.2), der(2)t(2;3)(p16;p11), add(3)(p11), del(6)(q15q22), add(3)(p11), del(6)(q15q22), del(9)(p22), dup(11)(q23q13),	n.d.	at diagnosis	m/7
	add(13)(p12), add(16)(q24), carries large submetacentric dup(1) marker			
BE-13	81-89<4n>XXXX, +1, +4, -9, -9, -10, -16, -21, del(1)(q13.2), add(1)(p11-21), del(4)(q11q31.2)/i(4p)x2, i(4q)del(4q31.2)/i(4q)add(4)(q31)x2, del(5)(q14q21)x2, del(6)(q22)x2, del(9)(p22)x2, der(17)t(17;?21)(p11;q11)x2 - 5q- and apparent 9p22-pter nullisomy	n.d	Relapse	f/11
697	46(45-48)<2n>XY, t(1;19)(q23;p13), del(6)(q21)	*E2A-PBX1*	Relapse	m/12
MHH-CALL-3	46(45-46)<2n>XX, del(6)(q15), der(9)t(9;9)(p21;q11), der(19)t(1;19)(q23;p13) - carries t(1;19) primary and 6q- secondary rearrangements associated with pre B-ALL	*E2A-PBX1*	at diagnosis	f/11

Supplementary Table 3: Overview of mouse strains used in this study

Mouse strain	Source	Purpose
Bach2$^{-/-}$	Kazuhiko Igarashi, Tohoku University	Genetic loss-of-function experiments
[a]*Bcl6*$^{-/-}$	Riccardo Dalla-Favera, Columbia University	Genetic loss-of-function experiments
[b]*Myc*$^{fl/fl}$	Ignacio Moreno de Alborán, CINES	Inducible deletion of *Myc*
[c]*Stat5*$^{fl/fl}$	Lothar Hennighausen, NIDDK	Inducible deletion of *Stat5a* and *Stat5b*
NOD/SCID IL2rg$^{-/-}$(NSG)	Jackson Laboratories	Xenograft recipient mice
Tp53$^{-/-}$	Jackson Laboratories	Analysis of *Tp53* as Bach2 target gene
Arf$^{-/-}$	Jackson Laboratories	Analysis of *Arf* as Bach2 target gene

Notes:

a Ye BH, Cattoretti G, Shen Q, Zhang J, Hawe N, de Waard R, Leung C, Nouri-Shirazi M, Orazi A, Chaganti RS, Rothman P, Stall AM, Pandolfi PP, Dalla-Favera R. The BCL-6 proto-oncogene controls germinal-centre formation and Th2-type inflammation. *Nat Genet.* 1997; 16: 161-70.

b de Alboran IM, O'Hagan RC, Gartner F, Malynn B, Davidson L, Rickert R, Rajewsky K. DePinho RA, Alt FW. Analysis of c-Myc function in normal cells via conditional gene-targeted mutation. Immunity. 2001; 14:45-55.

c Liu X, Robinson GW, Wagner KU, Garrett L, Wynshaw-Boris A, Hennighausen L. Stat5a is mandatory for adult mammary gland development and lactogenesis. *Genes Dev*. 1997; 11: 179-86.

Supplementary Table 4: Sequences of oligonucleotide primers used

Quantitative RT-PCR

*Bcl6*_F	5'-CCTGCAACTGGAAGAAGTATAAG-3'
*Bcl6*_R	5'-AGTATGGAGGCACATCTCTGTAT-3'
*Bach2*_F	5'-TGAGGTACCCACAGACACCA-3'
*Bach2*_R	5'-TGCCAGGACTGTCTTCACTG-3'
*Hprt*_F	5'-GGGGGCTATAAGTTCTTTGC-3'
*Hprt*_R	5'-TCCAACACTTCGAGAGGTCC-3'
*Btg2*_F	5'-GATGGCTCCATCTGTGTCCT-3'
*Btg2*_R	5'-TATACGGTGGCCTGTTGTCA-3'
*Rag2*_F	5'-GCAGATGGTAACAGTGGGTC-3'
*Rag2*_R	5'-ATTGCAGGCTTCAGTTTGAG-3'
*Rag1*_F	5'-TAACAACCAAGCTGCAGACA-3'
*Rag1*_R	5'-CCTCTGAGGAATCCTTCTCC-3'
*GFP*_F	5'-AGGAGCGCACCATCTTCTT-3'
*GFP*_R	5'-GCCATGATATAGACGTTGTGG-3'
*Trp53*_F	5'-TCCTTACCATCATCACACTGG-3'
*Trp53*_R	5'-CGGATCTTGAGGGTGAAATAC-3'
*Cdkn2a*_F	5'-GGACCAGGTGATGATGATG-3'
*Cdkn2a*_R	5'-ATCGCACGATGTCTTGATG-3'

Primers for cloning MSCV Bach2-ER^{T2} IRES GFP plasmid

*BACH2*_F	5'- AAAGGATCCGTCTGATCCCTTGCT -3'
*BACH2*_R	5'- AAACTCGAGGGTATAATCTTTCCT -3'

Clonality and spectratyping analysis

V_H1_F	5'- AAGGCCACACTGACTGTAGAC -3'
Cμ_R	5'- TGGCCACCAGATTCTTATCAG -3'
Cμ-FAM_R	5'- AGACGAGGGGGAAGACATTTG -3'

Supplementary Table 4 - continued: Sequences of oligonucleotide primers used

Mutation analysis in *BACH2* translated region (from BACH2 cDNA)

Primers to amplify *BACH2* coding region

Set A_F	5'-TTACATGGTGTGAACGGCATG-3'
Set A_R	5'-CCTGGCTGTGACCTCCTC-3'
Set B_F	5'-AGGAGGTCACAGCCAGG-3'
Set B_R	5'-GATGCTCTCTTCCTCATTCT-3'
Set C_F	5'-ACGCTCTGCCTGTCTGGAGA-3'
Set C_R	5'-CGGCTCAGAGAGGTCTTTGT-3'
Set D_F	5'-GTGCCAAAGGGTCTGTGGGT-3'
Set D_R	5'-CTCACACACCAATTTGCGGA-3'
Set E_F	5'-AAAGAGAAACTGTTGTCAGAG-3'
Set E_R	5'-CTAGGTATAATCTTTCCTGG-3'

Primers for sequencing *BACH2* coding region

Set A seq_F	5'-GTGTGAACGGCATGTCTGTG-3'
Set A seq_R	5'-CCTGGCTGTGACCTCCTC-3'
Set B seq_F	5'-TCACAGCCAGGGGCTTTG -3'
Set B seq_R	5'-GATGCTCTCTTCCTCATTCT-3'
Set C seq_F	5'-GCCTGTCTGGAGATGAGCC-3'
Set C seq_R	5'-CGGCTCAGAGAGGTCTTTGT-3'
Set D seq_F	5'-GGGTCTGTGGGTGGGAGC-3'
Set D seq_R	5'-CTCACACACCAATTTGCGGA-3'
Set E seq_F	5'-AAACTGTTGTCAGAGAGGAAT-3'
Set E seq_R	5'-CTAGGTATAATCTTTCCTGG-3'

***Position of primers used for mutation analysis on BACH2 cDNA**

Name of primer	Region of *BACH2* cDNA spanned (bp)
Set A	691 - 963
Set B	947 - 1575
Set C	1576 - 2091
Set D	2092 - 2760
Set E	2761 - 3234
Set A seq	698 - 963
Set B seq	953 - 1575
Set C seq	1583 - 2091
Set D seq	2100 - 2760
Set E seq	2767 - 3234

Supplementary Table 4: qChIP primers used

*Cdkn2a*_F	5'-TAGATGGACTCGGAGCAAGG-3'
*Cdkn2a*_R	5'-TTTCGCTCCGGTTAACTTTC-3'
Trp53 region1_F	5'-GCCGAGGCTAGAGTGCATTA-3'
Trp53 region1_R	5'-TCCCTGGTGATTGCTTTAGG-3'
Trp53 region2_F	5'-GAAACCCTGGGGTTGATTTT-3'
Trp53 region2_R	5'-AGTTCCAGGCAAACATGGAC-3'

Supplementary Table 4 - continued:	**qChIP primers used**
Bcl6 exon1_F	5’-CCGAGAATTGAGCTCTGTTGA-3’
Bcl6 exon1_R	5’-GGCAGCAACAGCAATAATCA-3’
*Acta1*_F	5’-AGAGTCAGAGCAGCAGGTAG-3’
*Acta1*_R	5’-CAAGGCTCAATAGCTTTCTT-3’
*BLNK*_F	5’-GCCTGGCTTCAAGTAAAAGTGT-3’
*BLNK*_R	5’-CTTCTCAGCCTGGAAATTATGG-3’
*BACH2*_F	5’-CCTACCTGGCAAAAACAAAAAC-3’
*BACH2*_R	5’-TCTTTTTGAGCAGTGGCATAGA-3’
*DLEU1*_F	5’-GGTGTTTCTTCCCACAGTCTTC-3’
*DLEU1*_R	5’-GAAATGCTGACTCACAGACACAG-3’
*DLEU2*_F	5’-CAGGCTCTAACTGCCAAATCTT-3’
*DLEU2*_R	5’-TGCGTTAGGAGAAGGGAAATAA-3’
*GADD45A*_F	5’-GGAAGAGATCCCTGTGAGTCAG-3’
*GADD45A*_R	5’-TCTGCCCTGCTAAAGGAATTAG-3’
*GADD45B*_F	5’-TCAAATGATGACTCAGCTCCAT-3’
*GADD45B*_R	5’-CTGCAAAGATGAACAAAACGAG-3’
*EBF1*_F	5’-GACTTTCTTGCTGTGTCATTCG-3’
*EBF1*_R	5’-GCCACATGTCAGCATTTTCTAA-3’
*SYK*_F	5’-TTGGTCCAATCAGTCATAGCAG-3’
*SYK*_R	5’-TTCTAGGTCAGCACATGCAAAT-3’
*CDKN1A*_F	5’-GCCACAGAACAGGACTCTGTC-3’
*CDKN1A*_R	5’-ACTGCAGCTTCCGTCTCTATTC-3’
*CDKN1B*_F	5’-AAGAATGGTGGAGTTGAGTGCT-3’
*CDKN1B*_R	5’-CCAAATGTTTCTGCGAAGGT-3’
*BCL6*_F	5’-GCAGTGGTAAAGTCCGAAGC-3’
*BCL6*_R	5’-AGCAACAGCAATAATCACCTG-3’
*RHOH*_F	5’-TGCTTAGCTGTGGTTCAGTGAT-3’
*RHOH*_R	5’-TGCTTCGGTCACAATGTTTTAC-3’
*POU2AF1*_F	5’-TCCTCTGGAAAACGTTGATCTT-3’
*POU2AF1*_R	5’-CTCCCAGTTGAGAACCAGTGAC-3’
*HIVEP1*_F	5’-TGCCTTAGAGCTGCTCCTAGAT-3’
*HIVEP1*_R	5’-TATGTGCACAGTCACGTTACCA-3’
*EGR3*_F	5’-TTCGTGGTGAAGAGGAAAGAAT-3’
*EGR3*_R	5’-TTGGAACCGTTAGGGAATTTA-3’
*RAG1*_F	5’-TTGCTCTCAATAATGGGGACT-3’
*RAG1*_R	5’-AGGAAGGTTGATGCTCCTTG -3’
*RAG2*_F	5’-TAGCAGAGCTGGCAAAGAAA -3’
*RAG2*_R	5’-AATGCAAGGCTCAGAAGGAA -3’

Supplementary Table 5: Antibodies used for Western blotting and flow cytometry

Western blot

Antigen	Clone ID	Dilution	Source
BACH2		1:1000	Ari Melnick, Weill Cornell Medical College, NY
Actb	C4	1:10,000	Santa Cruz Biotechnology
Tp53	IC12	1:1000	Cell Signaling Technology
Cdkn2a (p19 Arf)	Ab80	1:1000	Abcam
TP53	DO-7	1:1000	BD
CDKN2A (p14 ARF)	Ab470	1:1000	Abcam

Flow cytometry

Surface Antigen	Clone ID	Source
CD19	1D3	BD
Sca-1 (Ly6f)	D7	BD
Nt5e (CD73)	TY/23	BD
CD80	16-10A1	BD
K light chain	187.1	BD
IL2Rα (CD25)	7D4	BD
B220	RA3-6B2	BD
IgD	11-26c.2a	BD
IgM	R6-60.2	BD

Supplementary Table 6: Retroviral vectors used in the study

Constitutive expression	Inducible activation
MSCV *BCR-ABL1*-IRES-Neo	
MSCV Myc-IRES-GFP	MSCV-ER^{T2}-Puro
MSCV Bach2-IRES-GFP	MSCV-Cre-ER^{T2}-Puro
MSCV IRES-GFP	MSCV-ER^{T2}-GFP
EGZ Pax5	MSCV-Bach2-ER^{T2}-GFP
EGZ	
MI µHC-CD8	
MI CD8	
MOWS-RSS-eGFP-RSS (Recombination signal sequence substrate)	

Transfections of the above retroviral constructs were performed as discussed in the online methods section.

Supplementary Table 7: Accession numbers for somatic mutations in the *BACH2* gene in clinically-derived human Ph$^+$ ALL cells

Ph$^+$ ALL case	Mutation	Accession number (EMBL)
TXL3	C1039T	HE578168
ICN1	C1779T	HE578164
	C1039T	HE578169
SFO2	C1039T	HE578170
LAX9	C1779T	HE578165
BLQ1	G938T	HE578163
	C1039T	HE578173
PDX59	A2214G	HE578166
	C1039T	HE578176

Supplementary Table 8: Summary of accession numbers for gene expression and ChIP sequencing data

GEO accession	GEO description	Results	Figure
GSE30883	Role and function of Bach2 in *BCR-ABL1*- driven pre-B ALL	Differential gene expression in *Bach2*$^{+/+}$ and *Bach2*$^{-/-}$ *BCR-ABL1*-transformed ALL cells in the presence and absence of imatinib treatment	Figs. 2f and 3a, Supplementary Figs. 4 and 18
GSE30928	Role and function of Myc in *BCR-ABL1*- driven pre-B cells	Gene expression changes after *Myc* deletion in *BCR-ABL1*- driven pre-B ALL.	Supplementary Fig. 18
GSE24814	Role and function of Stat5 in *BCR-ABL1*- driven pre-B cells	Gene expression changes after *Stat5* deletion in *BCR-ABL1*- driven pre-B ALL.	Supplementary Fig. 18
GSE20987	Gene expression data of *BCR-ABL1*- transformed B cell precursors from *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ mice	Comparison of gene expression changes between *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ *BCR-ABL1*- transformed ALL cells	Fig. 2f
GSE30889	Role and function of Pax5 in *BCR-ABL1*- driven pre-B ALL	Comparison of gene expression changes before and after Pax5 overexpression in *BCR-ABL1*-driven pre-B ALL	Fig. 1a
GSE31027	Effects of pre-B cell receptor in *BCR-ABL1*- driven pre-B ALL	Comparison of gene expression changes before and after µ heavy chain overexpression in *BCR-ABL1*- driven pre-B ALL	Fig. 1a
GSE28460	Expression data from ALL diagnosis and relapse pediatric acute lymphoblastic leukemia cases	Pairwise comparison of *BACH2* expression levels in ALL patients at diagnosis and relapse	Fig. 5a, Supplementary Fig. 15
GSE41042	Effect of Bach2 overexpression in *Bcl6*$^{+/+}$ and *Bcl6*$^{-/-}$ *BCR-ABL1*-transformed pre-B cells	Gene expression changes arising after Bach2 overexpression in different *Bcl6* genotypes	Fig. 2c, Supplementary Fig. 4
GSE44420	ChIP sequencing using BACH2 and BCL6 antibodies in OCI-Ly7 cells	Comparison of binding of BACH2 and BCL6 to promoters of tumor suppressor genes	Supplementary Fig. 2, Fig. 2a
GSE34941	SuperSeries composed of GSE34861 and GSE34937; GSE34861: *BACH2* gene expression in primary ALL cases;	Gene expression of *BACH2* in different subgroups of ALL patients	Supplementary Fig. 14