

Infection Exposure Promotes *ETV6-RUNX1* Precursor B-cell Leukemia via Impaired H3K4 Demethylases



Guillermo Rodríguez-Hernández^{1,2}, Julia Hauer³, Alberto Martín-Lorenzo^{1,2}, Daniel Schäfer³, Christoph Bartenhagen⁴, Idoia García-Ramírez^{1,2}, Franziska Auer³, Inés González-Herrero^{1,2}, Lucia Ruiz-Roca^{1,2}, Michael Gombert³, Vera Okpanyi³, Ute Fischer³, Cai Chen³, Martin Dugas⁴, Sanil Bhatia³, René Martin Linka³, Marta Garcia-Suquia⁵, María Victoria Rascón-Trincado⁵, Angel Garcia-Sanchez^{5,6}, Oscar Blanco^{2,7}, Maria Begoña García-Cenador^{2,8}, Francisco Javier García-Criado^{2,8}, César Cobaleda⁹, Diego Alonso-López¹⁰, Javier De Las Rivas^{2,10,11}, Markus Müschen¹², Carolina Vicente-Dueñas², Isidro Sánchez-García^{1,2}, and Arndt Borkhardt³

Abstract

ETV6-RUNX1 is associated with the most common subtype of childhood leukemia. As few *ETV6-RUNX1* carriers develop precursor B-cell acute lymphocytic leukemia (pB-ALL), the underlying genetic basis for development of full-blown leukemia remains to be identified, but the appearance of leukemia cases in time-space clusters keeps infection as a potential causal factor. Here, we present *in vivo* genetic evidence mechanistically connecting preleukemic *ETV6-RUNX1* expression in hematopoietic stem cells/precursor cells (HSC/PC) and postnatal infections for human-like pB-ALL. In our model, *ETV6-RUNX1* conferred a low risk of developing pB-ALL after exposure to common pathogens, corroborating the low incidence observed in humans. Murine preleukemic *ETV6-RUNX1* pro/preB cells showed high

Rag1/2 expression, known for human *ETV6-RUNX1* pB-ALL. Murine and human *ETV6-RUNX1* pB-ALL revealed recurrent genomic alterations, with a relevant proportion affecting genes of the lysine demethylase (KDM) family. KDM5C loss of function resulted in increased levels of H3K4me3, which coprecipitated with RAG2 in a human cell line model, laying the molecular basis for recombination activity. We conclude that alterations of KDM family members represent a disease-driving mechanism and an explanation for RAG off-target cleavage observed in humans. Our results explain the genetic basis for clonal evolution of an *ETV6-RUNX1* preleukemic clone to pB-ALL after infection exposure and offer the possibility of novel therapeutic approaches. *Cancer Res*; 77(16): 4365–77. ©2017 AACR.

Introduction

Childhood B-cell precursor-acute lymphoblastic leukemia (pB-ALL) arises from interactions between genetic (inherited) susceptibility and exogenous exposures. The *ETV6-RUNX1* fusion gene is the most common chromosomal alteration in pediatric cancer and occurs in approximately 25% of childhood pB-ALL (1, 2). This fusion gene is one of the most extreme examples of

genotype–phenotype association. Its presence is only associated with pB-ALL and it seems to confer a low risk of developing this leukemia. Secondary genetic events are required for a pB-ALL evolving from an *ETV6-RUNX1* preleukemic clone. Studies on twins with concordant ALL have revealed that *ETV6-RUNX1* gene fusion represents the first event ("hit") in the process of leukemogenesis, creating a preleukemic clone, which requires secondary postnatal genetic aberrations (3–5). This is further supported

¹Experimental Therapeutics and Translational Oncology Program, Instituto de Biología Molecular y Celular del Cáncer, CSIC/Universidad de Salamanca, Campus M. de Unamuno s/n, Salamanca, Spain. ²Institute of Biomedical Research of Salamanca (IBSAL), Salamanca, Spain. ³Department of Pediatric Oncology, Hematology and Clinical Immunology, Heinrich-Heine University Düsseldorf, Medical Faculty, Düsseldorf, Germany. ⁴Department of Computer Science, Bonn-Rhein-Sieg University of Applied Sciences, Sankt Augustin, Germany. ⁵Departamento de Ciencias Biomédicas y del Diagnóstico, Área de Obstetricia y Ginecología, HUS-Universidad de Salamanca, Salamanca, Spain. ⁶IBSAL, Facultad de Medicina, Universidad de Salamanca, Salamanca, Spain. ⁷Departamento de Anatomía Patológica, Universidad de Salamanca, Salamanca, Spain. ⁸Departamento de Cirugía, Universidad de Salamanca, Salamanca, Spain. ⁹Centro de Biología Molecular Severo Ochoa, CSIC/Universidad Autónoma de Madrid, Campus de Cantoblanco, Madrid, Spain. ¹⁰Bioinformatics Unit, Cancer Research Center (CSIC-USAL), Salamanca, Spain. ¹¹Bioinformatics and Functional Genomics Research Group,

Cancer Research Center (CSIC-USAL), Salamanca, Spain. ¹²Department of Laboratory Medicine, University of California San Francisco, San Francisco, California.

Note: Supplementary data for this article are available at Cancer Research Online (<http://cancerres.aacrjournals.org/>).

G. Rodríguez-Hernández, J. Hauer, and A. Martín-Lorenzo share first authorship of this article.

C. Vicente-Dueñas, I. Sánchez-García, and A. Borkhardt share senior authorship of this article.

Corresponding Author: Arndt Borkhardt, Moorenstr. 5, 40225 Düsseldorf, Germany. Phone: 49-211-81-17680; Fax: 49-211-81-16707. E-mail: Arndt.Borkhardt@med.uni-duesseldorf.de

doi: 10.1158/0008-5472.CAN-17-0701

©2017 American Association for Cancer Research.

by the fact that the fusion gene may be present for up to 10 years before leukemia diagnosis (3–5). Correspondingly, high-resolution genome-wide analysis revealed multiple additional genetic lesions in overt *ETV6-RUNX1*-associated leukemia (6). These critical secondary events are frequently driven by genomic alterations mediated by RAG recombinase activity, and only infrequently by point mutations (7). This suggests that high RAG expression and/or facilitated recruitment of RAG to cryptic recombination signal sequence (RSS) might be a common path in the acquisition of secondary alterations in *ETV6-RUNX1* pB-ALL. The pending challenge is to identify the relevant exposures and decipher how and when these factors contribute to the multistep natural history of *ETV6-RUNX1* pB-ALL, because it can offer prophylactic interventions. The lack of genetically engineered human-like *ETV6-RUNX1* pB-ALL models has hampered our better understanding of the pathogenesis of this disease (8, 9). Thus far, experimental models of *ETV6-RUNX1*-associated leukemia have recapitulated some aspects of disease evolution but not the full disease phenotype, irrespective of the mode and timing of expression or the source of target cells (10–15).

To address these issues, we have used a novel experimental approach to model the *ETV6-RUNX1* preleukemic state by expressing the human *ETV6-RUNX1* fusion gene under the control of the Sca1 promoter in the mouse hematopoietic system. Sca1-*ETV6-RUNX1* mice developed exclusively pB-ALL at a low disease penetrance only when they were exposed to common pathogens.

Materials and Methods

Primary human samples

Primary cases (Supplementary Fig. S1 and Supplementary Table S1) were obtained with informed consent in compliance with the institutional review board of the University of Düsseldorf, Germany.

Mouse leukemia model for *ETV6-RUNX1* pB-ALL

All animal work has been conducted according to relevant national and international guidelines, and it has been approved by the Bioethics Committee of University of Salamanca and by the Bioethics Subcommittee of Consejo Superior de Investigaciones Científicas (CSIC). The *ETV6-RUNX1* vector was generated by inserting the human *ETV6-RUNX1* cDNA into the ClaI site of the pLy6 vector. The transgene fragment was excised from its vector by restriction digestion with NotI, purified and injected (2 ng/mL) into CBAx57BL/6J-fertilized eggs. Transgenic mice were identified by Southern blot analysis of tail snip DNA after EcoRI digestion, using *ETV6-RUNX1* cDNA to detect the transgene. Two independent transgenic lines were generated and analyzed. The two independent Sca1-*ETV6-RUNX1* founder lines had normal gestation and were viable and developed normally and were used to examine the phenotype further. Upon signs of disease, Sca1-*ETV6-RUNX1* mice were sacrificed and subjected to standard necropsy procedures. All major organs were examined under the dissecting microscope. Tissue samples were taken from homogeneous portions of the resected organ and fixed immediately after excision. Differences in Kaplan–Meier survival plots of transgenic and WT mice were analyzed using the log-rank (Mantel–Cox) test.

FACS analysis

Nucleated cells were obtained from total mouse bone marrow (BM; flushing from the long bones), peripheral blood (PB), thymus, or spleen. In order to prepare cells for flow cytometry,

contaminating red blood cells were lysed with RCLB lysis buffer and the remaining cells were then washed in PBS with 1% FCS. After staining, all cells were washed once in PBS and were resuspended in PBS with 1% FCS containing 2 mg/mL propidium iodide (PI) to allow dead cells to be excluded from both analyses and sorting procedures. The samples and the data were acquired in an AccuriC6 Flow Cytometer and analyzed using FlowJo software. Specific fluorescence of FITC, PE, PI, and APC excited at 488 nm (0.4 W) and 633 nm (30 mW), respectively, as well as known forward and orthogonal light scattering properties of mouse cells were used to establish gates. Nonspecific antibody binding was suppressed by preincubation of cells with CD16/CD32 Fc-block solution (BD Biosciences). For each analysis, a total of at least 50,000 viable (PI[−]) cells were assessed.

The following antibodies were used for flow cytometry: anti-B220 (RA3-6B2), CD3ε (145-2C11), CD4 (RM4-5, 1:500), CD8α (53-6.7, 1:500), CD11b/Mac1 (M1/70, 1:200), CD19 (1D3), CD117/c-Kit (2B8, 1:200), CD127/IL-7Rα (A7R34, 1:50), Ly-6G/Gr1 (RB6-8C5), IgM (R6-60.2), Sca1/Ly-6A/E (E13-161.7, 1:50), CD25 (PC61), CD48 (HM48-1) and CD150 (TC15-12F12.2) antibodies. Unspecific antibody binding was suppressed by preincubation with CD16/CD32 (2.4G2) Fc-block solution (PharMingen). FACS definition of B-cell developmental stages was performed according to Kwon and colleagues (16) with minor modifications. Briefly, BM proB cells (CD19⁺c-Kit⁺), BM pre-B cells (B220⁺CD25⁺IgM[−]), BM immature B cells (B220⁺IgM^{hi}IgD[−]), BM recirculating B cells (B220⁺IgD^{hi}), peripheral transitional B cells (B220⁺IgM^{hi}IgD^{hi}), peripheral mature B cells (B220⁺IgM^{lo}IgD^{hi}), marginal zone (MZ) B cells (B220⁺CD21^{hi}CD23^{lo}), and follicular (FO) B cells (B220⁺CD21^{int}CD23^{hi}). All antibodies were purchased from BD Biosciences. All antibodies were used at a 1:100 dilution unless otherwise indicated.

Histology

Animals were sacrificed by cervical dislocation; tissue samples were formalin-fixed and included in paraffin. Pathology assessment was performed on hematoxylin–eosin-stained sections under the supervision of Dr. Oscar Blanco, an expert pathologist at the Salamanca University Hospital.

V(D)J recombination

Immunoglobulin rearrangements were amplified by PCR using the primers in Table 1. Cycling conditions consisted of an initial heat-activation at 95°C followed by 31 to 37 cycles of denaturation for 1 minute at 95°C, annealing for 1 minute at 65°C, and elongation for 1 minute 45 seconds at 72°C. This was followed by a final elongation for 10 minutes at 72°C. To determine the DNA sequences of individual V(D)J rearrangements, the PCR fragments were isolated from the agarose gel and cloned into the pGEM-T Easy vector (Promega).

Mouse exome library preparation and next-generation sequencing

Sample acquisition. The AllPrep DNA/RNA Mini Kit (Qiagen) was used to purify DNA according to the manufacturer's instructions.

Exome library preparation and next-generation sequencing. Exome library preparation was performed using the Agilent SureSelectXT Mouse All Exon kit with modifications adapted from Fisher and colleagues (17). Briefly, we added SPRI beads to the original

Table 1. Primers used to determine the V(D)J recombination status (first 10 primers) and to validate mutations by Sanger sequencing (last eight primers)

Primer	Sequence [5'-3']
V _H J558-f	CGAGCTCTCCARACAGCCTWCATGCARCTCARC
V _H J558-r	GTCTAGATTCTACAAGAGTCCGATAGACCTGG
V _H 7183-f	CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC
V _H 7183-r	GTCTAGATTCTACAAGAGTCCGATAGACCTGG
V _H Q52-f	CGGTACCAGACTGARCATCASCAAGGACAAYTCC
V _H Q52-r	GTCTAGATTCTACAAGAGTCCGATAGACCTGG
DH-f	TTCAAAGCACAATGCCTGGCT
DH-r	GTCTAGATTCTACAAGAGTCCGATAGACCTGG
C μ -f	TGGCCATGGGCTGCCTAGCCCGGGACTT
C μ -r	GCCTGACTGAGCTCACACAAGGAGGA
KDM5C-f	CTTCTCCCTCCCTACCCCTTAT
KDM5C-r	ATTACACGCTCCAGAACTCC
Ebf1-f	GTCGTGGTGTCTACCACAG
Ebf1-r	ATGATTGCTCTACCATGTTC
Kdm5c-f	AGGTTAGAGGGACTCTTCAG
Kdm5c-r	CCTCACAACATACCCAG
Kdm2b-f	CCTGCCAAGCGGAGAAGTGAGTG
Kdm2b-r	CCAGGCCTTCCAGTTCGTA

protocol and reduced the size of the reaction to 0.5 μ L in order to be able to use PCR tubes for subsequent steps. Furthermore, we reduced the volume for washing. We minimized sample loss and optimized sample processing by reducing sample handling. We therefore just added freshly prepared 20% PEG/2.5 mol/L NaCl (Sigma) instead of elution of samples from the SPRI beads for library preparation. Targeted capture by hybridization to an RNA library was performed according to the manufacturer's protocol. Purification and enrichment of the captured library was achieved by binding to MyOne Streptavidin T1 Dynabeads (Life Technologies) and off-bead PCR amplification in the linear range. Sequencing (2 \times 100 bp) with a 6 bp index read was performed using the TruSeq SBS Kit v3 on the HiSeq 2500 (Illumina).

Data analysis. Fastq files were generated by using BcltoFastq 1.8.4 (Illumina). BWA version 0.7.4 (18) was used to align sequence data to the mouse reference genome (GRCm38.71). Conversion steps were carried out using Samtools (19, 20) followed by removal of duplicate reads (<http://broadinstitute.github.io/picard>). Local realignment around indels, SNP-calling, annotation, and recalibration was facilitated by GATK 2.4.9 (21). Mouse dbSNP138 and dbSNP for the used mouse strains were used as training datasets for recalibration. Resulting variation calls were annotated by Variant Effect Predictor (22) using the Ensembl database (v70) and imported into an in-house MySQL database to facilitate automatic and manual annotation, reconciliation, and data analysis by complex database queries. Loss-of-function prediction scores for PolyPhen2 (23) and SIFT (24) were extracted from this Ensembl release.

Only entries with at least 9% difference in allele frequency between tumor and normal were kept for further analysis. Cancer-related genes were determined by translating the cancer gene consensus from COSMIC (25) using ENSEMBL's biomaRT.

Sequencing

Mutations were validated using Sanger sequencing and specific primers (Table 1) on a 3130 Genetic Analyzer (Applied Biosystems).

ProB cell culture

Iscove's modified Dulbecco's medium supplemented with 50 μ mol/L β -mercaptoethanol, 1 mmol/L L-glutamine, 2% heat-

inactivated fetal calf serum, 1 mmol/L penicillin-streptomycin (BioWhittaker), and 0.03% (w/v) primatone RL (Sigma) was used for proB cell-culture experiments. ProB cells isolated by MACS sorting for B220⁺ (Milteny Biotec) from BM were cultured on Mitomycin C-treated ST2 cells in IMDM medium containing IL7 (R&D Systems). IL7-independent tumor proB cells were grown in the same medium without IL7.

Apoptosis assays

The extent of IL7-withdrawal apoptosis was evaluated by Annexin V-FITC/PI staining. Aliquots (200 μ L) containing 10⁶ cells in 10 mmol/L Hepes/NaOH pH 7.4, 140 mmol/L NaCl, and 2.5 mmol/L CaCl₂ were incubated with Annexin V-FITC (BD Biosciences) to a final concentration of 1 μ g/mL for 10 minutes, at room temperature in the dark, and then labeled with propidium iodide to a final concentration of 2 μ g/mL (PI in binding buffer). Flow cytometry was performed within an hour of labeling, and data were analyzed using FlowJo software. The difference between experimental variables was determined using the Student *t* test.

Microarray data analysis

Total RNA was isolated in two steps using TRIzol (Life Technologies) followed by RNeasy Mini-Kit (Qiagen) purification following the manufacturer's RNA Clean-up protocol with the optional On-column DNase treatment. The integrity and the quality of the RNA were verified by electrophoresis and its concentration measured. Samples were analyzed using Affymetrix Mouse Gene 1.0 ST arrays.

Briefly, the robust microarray analysis (RMA) algorithm was used for background correction, intra- and inter-microarray normalization, and expression signal calculation (26–28). Once the absolute expression signal for each gene (i.e., the signal value for each probe set) was calculated in each microarray, a method called significance analysis of microarray (SAM; ref. 29) was applied to calculate significant differential expression and find the gene probe sets that characterized the BM pro/preB cells from preleukemic *ETV6-RUNX1* mice compared BM pro/preB cells from WT mice. The method uses permutations to provide robust statistical inference of the most significant genes and provides *P* values adjusted to multiple testing using false discovery rate (FDR; ref. 30). A cutoff of FDR < 0.1 was used for the differential expression calculations. We applied all these methods using R (31) and Bioconductor (32).

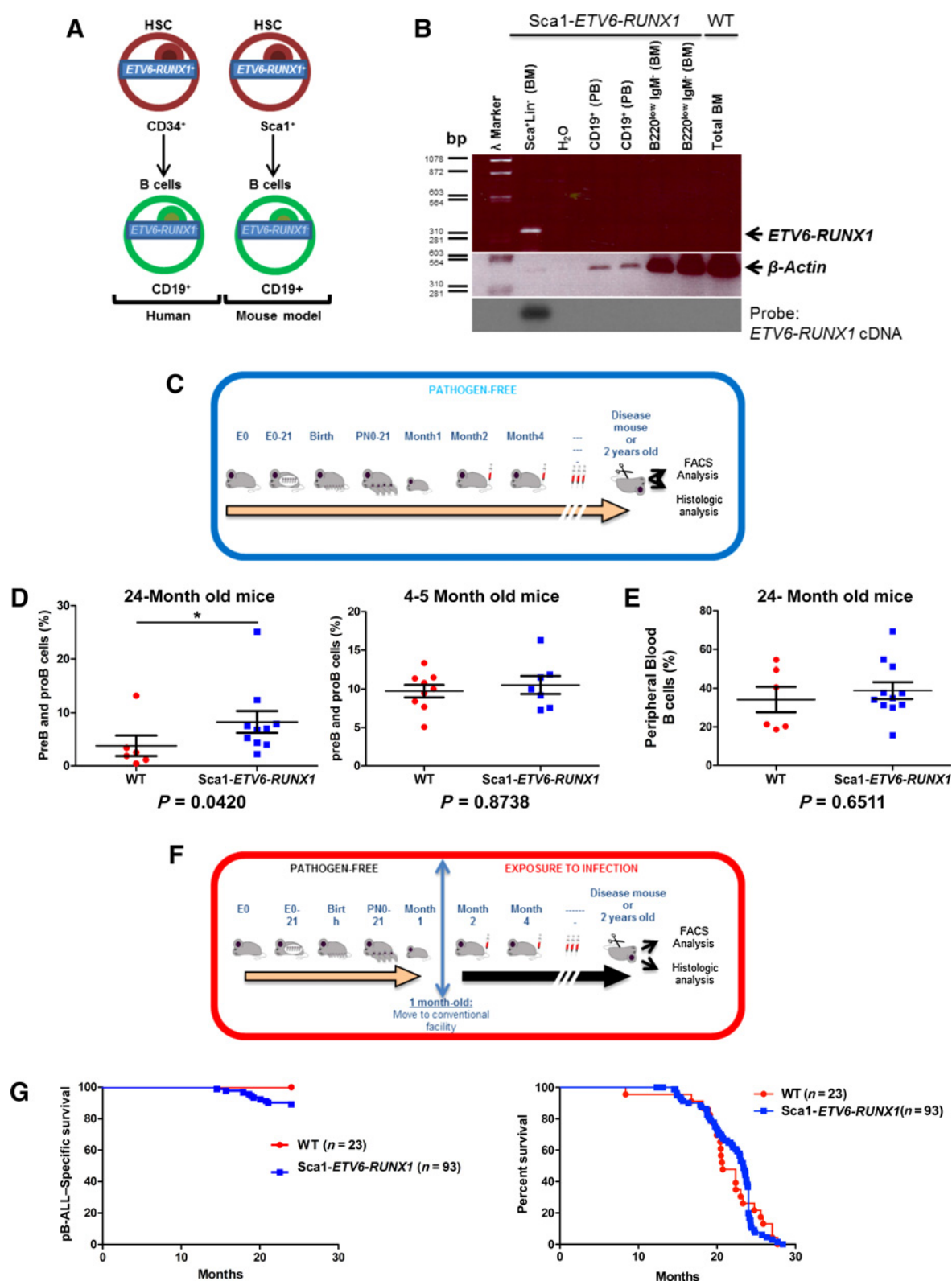
The data discussed in this publication have been deposited in NCBI's Gene Expression Omnibus (33) and are accessible through GEO Series accession number GSE70492 (<http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE70492>).

Enrichment analysis

Differentially expressed genes were tested for enrichment using GSEA enrichment analysis from MSigDB (ref. 34; <http://www.broad.mit.edu/gsea/>). We carried out one enrichment analysis for the *Lysine(K)-specific demethylases* gene signature described in HUGO gene nomenclature committee.

Transplantation experiments

In order to determine the nature of the leukemogenic cell, BM transplantation experiments were performed. Sca1-*ETV6-RUNX1* proB cells were injected intravenously into sublethally irradiated (4 Gy) secondary recipient 12-week-old male syngenic mice (C57BL/6 \times CBA). Disease development in the recipient mice

**Figure 1.**

Modeling the first hit function of *ETV6-RUNX1* in mice. **A**, Mouse model mimicking human disease. *ETV6-RUNX1* is expressed in HSC but not in terminal differentiated B cells in the *Sca1-ETV6-RUNX1* mouse model mimicking the preleukemic stage of human *ETV6-RUNX1* pB-ALL disease. **B**, *ETV6-RUNX1* pB-ALL expression in *Sca1-ETV6-RUNX1* mice. (Continued on the following page.)

was monitored by periodic PB analysis until blast cells were detected. Then, they were sacrificed and assessed for pB-ALL development.

CRISPR-mediated gene editing

Knockout of the *KDM5C* gene was done by CRISPR/Cas9-mediated knock-in of a selection cassette following the double nicking strategy by Ran and colleagues (35). A pair of guide RNAs targeting exon 1 of *KDM5C* adjacent to its transcription start site was designed with the help of the CRISPR design tool (<http://crispr.mit.edu/>; ref. 36). Expression cassettes for Cas9^{D10A} (Casn) and the guide RNAs were taken from pX335 (Addgene #42335; ref. 37) and transferred to a smaller pUC19 backbone via PCR-based cloning. The knock-in was achieved by offering a repair template addressing the homologous recombination repair pathway: the coding sequence for the truncated human low-affinity nerve growth factor receptor (LNGFR, taken from pMACS LNGFR, Miltenyi Biotech) followed by a Poly A from SV40 was bedded between 0.9 and 1 kb homologous sequence for *KDM5C* flanking the expected nicking sites. All four plasmids (Casn, 2 × gRNA, and repair template) were nucleofected in NALM-6 cells (Kit V, T-001, Lonza)—the repair template was also circularized to diminish random integration. Selection of positive cells was done by three sequential rounds of enrichment by using the MACSelect LNGFR System (Miltenyi).

Coimmunoprecipitation and immunoblotting analyses

Wild-type (WT) and *KDM5C*[−] NALM-6 cells were nucleofected (Lonza) with c-myc-tagged-RAG2 plasmid. After 48 hours, coimmunoprecipitation was performed with Anti-c-myc microbeads following the manufacturer's guidelines (Miltenyi), and later precipitated proteins were immunoblotted with anti-RAG2 (Abcam) and anti-H3K4me3 (Merck Millipore) antibodies.

Results

Development of a mouse model for preleukemic hematopoietic cell precursors expressing *ETV6-RUNX1*

The *ETV6-RUNX1* transcript expression is not detected in progenitor B cells of healthy children carrying the preleukemic clone (4, 38). We generated a murine strain that mimicked this scenario by expressing *ETV6-RUNX1* specifically within hematopoietic stem/progenitor cells (HSPC) through placing a human *ETV6-RUNX1* cDNA under control of the Sca1 promoter (Fig. 1A and Supplementary Fig. S2A and S2B). This system ensures the expression of *ETV6-RUNX1* into HSPC (39–41). In accord with the known capacity of Sca1 regulatory elements to drive transgene expression in Sca1⁺ cells, Sca1-*ETV6-RUNX1* transgene expression was detected in HSPCs but not in proB cells or later stages of B-cell development (Fig. 1B). Therefore, this limited expression

of *ETV6-RUNX1* during early hematopoietic development mimicked human *ETV6-RUNX1* preleukemic biology (4, 38) and provided a preclinical model to evaluate environmental factors, which act synergistically in *ETV6-RUNX1* pB-ALL development.

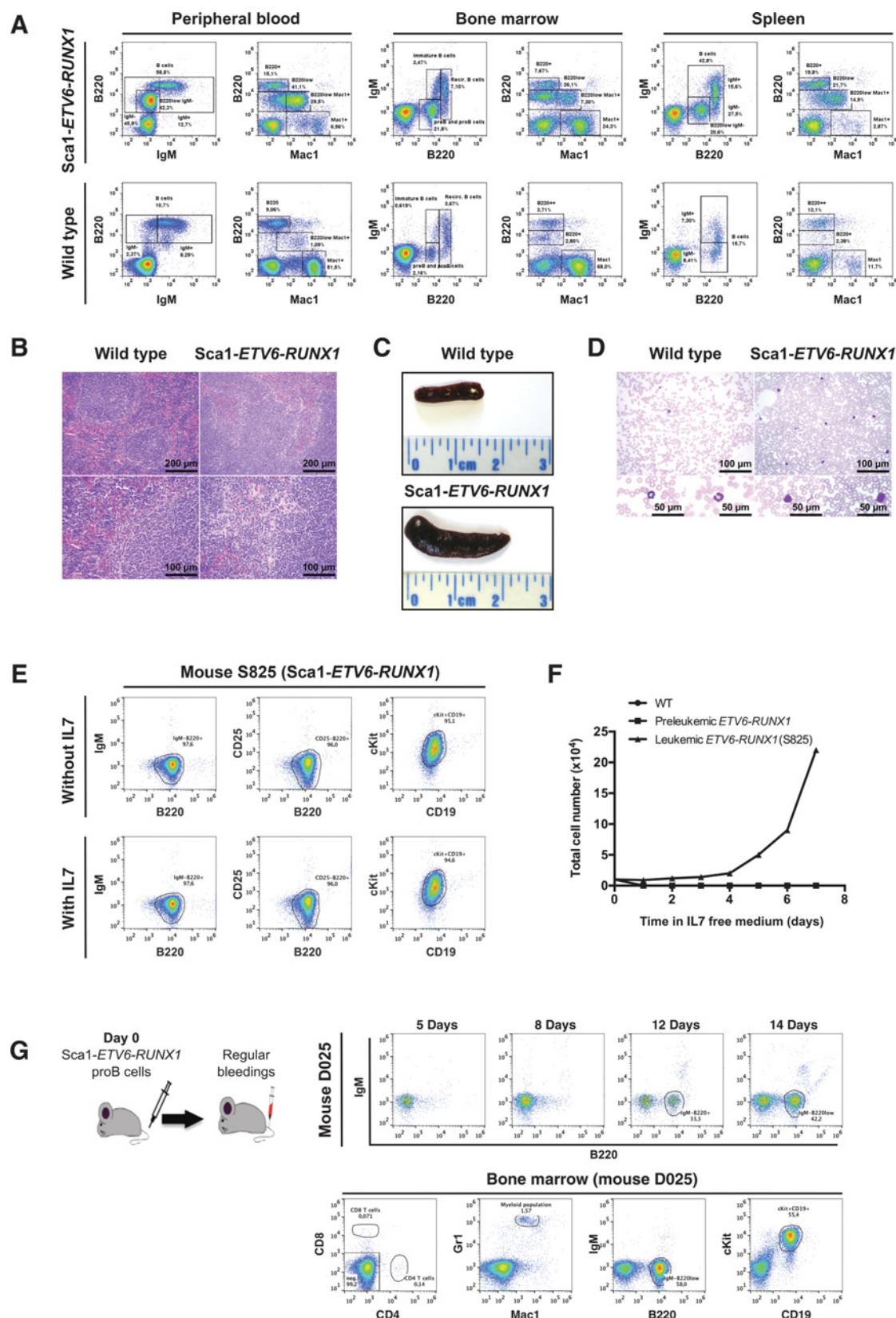
Sca1-*ETV6-RUNX1* mice do not develop pB-ALL under specific pathogen-free conditions

To determine if *ETV6-RUNX1* expression in HSPCs predisposes to leukemia as implicit in its assignment as the initiating hit in childhood pB-ALL, we monitored cohorts of Sca1-*ETV6-RUNX1* mice and control WT mice born and kept in the specific pathogen-free (SPF) environment during their lifespan ($n = 37$; observed for 2 years; Fig. 1C). To comprehensively study the long-term impact of *ETV6-RUNX1* on BM lymphopoiesis, we characterized the different stages of B-cell development. The different cell populations were analyzed by flow cytometry in BM, spleen, and PB of 4-month-old Sca1-*ETV6-RUNX1* transgenic mice compared with an age-matched WT control. At 4 to 5 months of age, an increase of the BM pro/preB cells and immature B cells could not be detected (Supplementary Fig. S2C). In contrast, BM pro/preB cells were not gradually lost in Sca1-*ETV6-RUNX1* mice (Fig. 1D and Supplementary Fig. S2D). These data suggest that *ETV6-RUNX1* fusion favors the maintenance of a B-cell precursor compartment in the BM, although the differentiation to mature PB B cells is not impaired (Fig. 1E). However, in the absence of oncogenic environments, this is not sufficient to produce leukemia, because Sca1-*ETV6-RUNX1* mice remained without evidence of disease.

Infection exposure induces pB-ALL in Sca1-*ETV6-RUNX1* mice

Thus, we next asked if we could provoke pB-ALL development in the Sca1-*ETV6-RUNX1* mice by exposing them to oncogenic environments relevant to human disease. Infection exposure was the first suggested cause for childhood *ETV6-RUNX1* pB-ALL and remains one of the strongest candidates (3). Moreover, *in vitro* inflammatory and infection stimuli are known to increase RAG expression, which is relevant for the clonal evolution of human *ETV6-RUNX1* pB-ALL (42–44). However, the role of infection in the conversion of the *ETV6-RUNX1* preleukemic clone into pB-ALL remains unknown. To test this hypothesis, we monitored cohorts of Sca1-*ETV6-RUNX1* mice and control WT mice born and kept in the SPF environment until exposed to common infectious environment (Fig. 1F; murine norovirus, murine hepatitis virus, helicobacter species, and *Trichomonas muris*; Supplementary Table S2). Under this scenario, specific pB-ALL development was observed only in Sca1-*ETV6-RUNX1* mice exposed to common infectious environment (10.75%; 10 out of 93). The low penetrance of pB-ALL in the murine model closely resembled the low penetrance of pB-ALL development in children carrying the preleukemic *ETV6-RUNX1* clone (3, 5). Due to the low incidence of the pB-ALL disease, the overall survival of Sca1-*ETV6-RUNX1*

(Continued.) BM HSC, BM proB, and preB cells and PB CD19⁺ cells from Sca1-*ETV6-RUNX1* mice and total BM from WT mice were tested for *ETV6-RUNX1* expression. *ETV6-RUNX1* is only expressed in BM HSC. Shown is one representative experiment out of two biological replicates. **C**, Experimental set-up. Mice were born and followed up in SPF conditions until they were sacrificed. **D**, Percentage of BM proB and preB cells (B220^{low} IgM[−]) in 24-month-old Sca1-*ETV6-RUNX1* mice ($n = 10$) versus young (4–5 months old) mice compared with age-matched control littermate WT mice ($n = 6$). Error bars, SD. The Mann-Whitney *U* test was used and *P* value is indicated. All the mice were housed in SPF facility. **E**, Percentage of PB B cells in 24-month-old Sca1-*ETV6-RUNX1* mice ($n = 10$) compared with age-matched control littermate WT mice ($n = 6$). Error bars, SD. The Mann-Whitney *U* test was used and *P* value is indicated. All the mice were housed in SPF facility. **F**, Experimental set-up. Mice were born under SPF conditions and were moved to a non-SPF facility where they were exposed to common infectious pathogens 1 month after birth. **G**, Kaplan-Meier survival curve of Sca1-*ETV6-RUNX1* mice (right) and specific pB-ALL survival curve of Sca1-*ETV6-RUNX1* mice (left). Overall survival of Sca1-*ETV6-RUNX1* mice (blue; $n = 93$) compared with WT control mice (red; $n = 23$). Log-rank (Mantel-Cox) *P* = 0.9352 (right). Leukemia incidence of Sca1-*ETV6-RUNX1* mice (blue; $n = 93$) compared with WT control mice (red; $n = 23$). Log-rank (Mantel-Cox) *P* = 0.0855 (left).



mice was not significantly reduced compared with WT mice ($P = 0.9352$; Fig. 1G). The appearance of leukemia in Sca1-*ETV6-RUNX1* mice manifested with splenomegaly, disruption of splenic architecture due to blast infiltration, and appearance of blast cells in PB (Fig. 2A–D). Sca1-*ETV6-RUNX1* pB-ALLs displayed clonal immature BCR rearrangement (Supplementary Fig. S3A). FACS analysis revealed a cell surface phenotype $CD19^+B220^+IgM^-$ for tumor cells that extended through BM, PB, and spleen (Supplementary Fig. S3B–S3D). Consistently, tumor proB cells grew independent of IL7 (Fig. 2E and F), and these cells were able to propagate the disease when transplanted into sublethally irradiated syngeneic recipient mice (Fig. 2G). These results provide evidence that the Sca1-*ETV6-RUNX1* model closely reproduces the human disease, as the presence of the fusion gene is only associated with pB-ALL, and it confers a low risk of developing pB-ALL (3, 5). Overall, these results represent the first proof that infection exposure can induce human-like pB-ALL in mice carrying an *ETV6-RUNX1* preleukemic clone.

ETV6-RUNX1 impairs B-cell development under exposure to infection

We next aimed to explain the mechanism why the susceptible preleukemic *ETV6-RUNX1* clone progresses to pB-ALL under exposure to infections. To this end, the different hematopoietic compartments were studied using flow cytometry in young WT and preleukemic Sca1-*ETV6-RUNX1* littermates of the same breeding after delayed exposure to infection. No abnormalities were detected, neither in the myeloid nor in the T-lymphoid compartments (as determined by Gr1, Mac1 or CD4, and CD8 staining) in the BM, PB, thymus, and spleen (Supplementary Fig. S3E–S3G). However, the analysis of the different B-cell developmental stages showed a significant and specific increase of pro/preB cells ($B220^{low}IgM^-$) and immature B cells ($B220^{low}IgM^+$) in Sca1-*ETV6-RUNX1* mice compared with age-matched WT littermates but not in recirculating B cells ($B220^{++}IgM^+$; Fig. 3A; Supplementary Table S3). However, this increase of pro/preB cells ($B220^{low}IgM^-$) was not observed in 24-month-old leukemia-free Sca1-*ETV6-RUNX1* mice compared with age-matched WT mice under exposure to infection conditions (Supplementary Fig. S3H). These data suggest that the specific and temporal increase of BM pro/preB cells was induced by the exposure to common pathogens as a similar increase was not observed in Sca1-*ETV6-RUNX1* mice housed in an SPF environment (Supplementary Fig. S2C). In contrast, *ETV6-RUNX1* expression did not result in striking alterations in $Lin^-Sca1^+cKit^+$ (LSK) cells,

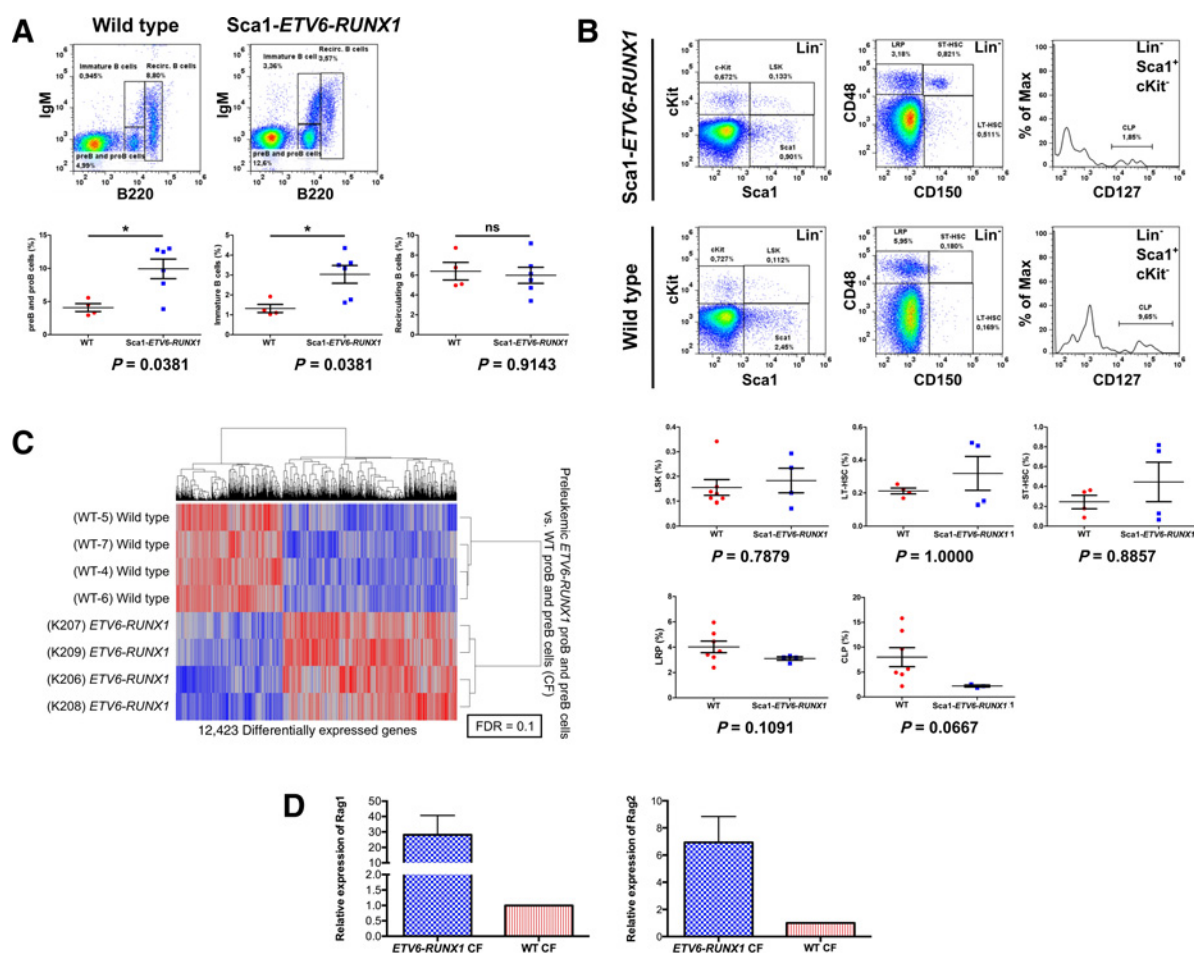
long-term hematopoietic stem cells (LT-HSC), short-term hematopoietic stem cells (ST-HSC), lineage-restricted progenitor (LRP) cells, and common lymphoid progenitors (CLP) within Sca1-*ETV6-RUNX1* mice compared to WT mice (Fig. 3B). These data therefore suggest that *ETV6-RUNX1* fusion favors the appearance of an aberrant B-cell precursor compartment in the BM of young Sca1-*ETV6-RUNX1* mice only under exposure to common pathogens, although the differentiation to mature PB B cells is not impaired. These observations are in agreement with reports of preleukemic activity in mice bearing pB-ALL-associated oncogenes and/or inherited susceptibility genes (3, 45). In agreement with these findings, expression array analysis of preleukemic *ETV6-RUNX1* pro/preB cells derived from mice housed in conventional facility (CF; $n = 4$, ID: K206-K209) shows a distinct expression pattern compared with healthy age-matched WT mice ($n = 4$, ID: WT-4–WT-7; Fig. 3C; Supplementary Table S4). The analysis of the differential expressed genes revealed significant higher *Rag1* and *Rag2* expression (Fig. 3D), which is a hallmark in human translocation t(12;21)-positive leukemia (MILE study; <http://r2.amc.nl>). Preleukemic *ETV6-RUNX1* pro/preB cells showed differential regulation of epigenetic regulator genes of the KDM family, including *KDM2A*, *KDM3A/B*, *KDM4A/C*, *KDM5A/B/C*, and *KDM6A/B* (Table 2), but not of other epigenetic regulator gene families, including HDAC or PMT. However, we did not find this high proportion of differentially regulated epigenetic regulator genes in a related *Pax5*^{+/−} model of pB-ALL (Table 2 and Supplementary Table S5; ref. 46). These data indicate that *ETV6-RUNX1* specifically regulates transcription of histone modifying genes of the KDM family and support a specific role of histone modification in preleukemic *ETV6-RUNX1* cells (Supplementary Table S5).

Genomic analysis identified the mechanism of clonal evolution in infection-driving murine pB-ALL

In order to elucidate how infection exposure causes pB-ALL in the context of aberrant histone modification and elevated *Rag1* expression in a preleukemic *ETV6-RUNX1* clone, we next performed whole exome and whole genome sequencing of six Sca1-*ETV6-RUNX1* tumors and corresponding germline on a HiSeq 2500 (Illumina) platform for the detection of structural aberrations ($n = 3$) as well as of SNVs ($n = 6$; Supplementary Fig. S4A). Sca1-*ETV6-RUNX1* tumor cells were derived from BM of diseased mice. We did not observe aberrations of the *Etv6* locus, which are described in up to 70% of human *ETV6-RUNX1* patients (6), most likely because both *Etv6* alleles are intact in the transgenic mice.

Figure 2.

pB-ALL development in Sca1-*ETV6-RUNX1*. **A**, Flow cytometric analysis of hematopoietic subsets in diseased Sca1-*ETV6-RUNX1* mice. Representative plots of cell subsets from the PB, BM, and spleen are shown. These show accumulation of blast B cells in Sca1-*ETV6-RUNX1* mice ($n = 10$) compared with age-matched control littermate WT mice ($n = 23$). **B**, Hematoxylin and eosin staining of spleen sections from WT and tumor-bearing Sca1-*ETV6-RUNX1* mice show disruption of normal spleen architecture due to tumor-infiltrating cells. Scale bar, 100 μm ($\times 400$) and 200 μm ($\times 200$; $n = 10$). **C**, Splenomegaly observed in diseased Sca1-*ETV6-RUNX1* mice. WT mouse spleen is shown for reference (representative of $n = 10$). **D**, Blood smear from WT and tumor-bearing Sca1-*ETV6-RUNX1* mice showing the presence of blast cells in the PB of diseased mice. Scale bar, 100 μm ($\times 400$) and 50 μm ($\times 400$; $n = 4$). **E**, Immunophenotype of tumor Sca1-*ETV6-RUNX1* proB cells. Sorted $B220^+$ cells from BM of the diseased Sca1-*ETV6-RUNX1* mouse were cultured under conditions to allow the isolation and expansion of a pure population of proB cells. These tumor proB cells grow in the absence of IL7. **F**, WT, preleukemic Sca1-*ETV6-RUNX1* and leukemic Sca1-*ETV6-RUNX1* primary proB cells were cultured in media without IL7 and their proliferation measured every day using Trypan blue. Values represent the mean out of three replicates with essentially identical datasets. **G**, Sca1-*ETV6-RUNX1* pB-ALL is transplantable to secondary recipients. Experimental set-up. One hundred thousand leukemic Sca1-*ETV6-RUNX1* proB cells were injected into sublethally irradiated WT syngeneic mice. Regular bleedings were performed in order to monitor the development of the pB-ALL. Representative flow cytometric analysis of four mice injected with leukemic Sca1-*ETV6-RUNX1* proB cells. Cytometric analyses at different time points show that leukemic pB-ALL cells ($B220^{low}IgM^-cKit^+CD19^+$) were able to grow in secondary recipients. Representative flow cytometric analysis of four mice injected with leukemic Sca1-*ETV6-RUNX1* proB cells shows the accumulation of pB-ALL cells ($B220^{low}IgM^-cKit^+CD19^+$) in BM.

**Figure 3.**

B-cell development in young Sca1-ETV6-RUNX1 mice housed in conventional facility. **A**, Percentage of BM B cell compartments in 7-month-old leukemia-free Sca1-ETV6-RUNX1 mice. Sca1-ETV6-RUNX1 mice ($n = 6$) show a significant increase in preB and proB cells (B220^{low} IgM⁺) and immature B cells (B220^{low} IgM⁺) compared with age-matched control littermate WT mice ($n = 4$) but not in recirculating B cells (B220⁺ IgM⁺). **B**, Flow cytometric analysis of hematopoietic precursor subsets in 4-month-old tumor-free mice. No differences can be seen in Lin⁺ Sca1⁺ cKit⁺ (LSK) cells, long-term hematopoietic stem cells (LT-HSC), short-term hematopoietic stem cells (ST-HSC), lineage-restricted progenitor (LRP) cells, and common lymphoid progenitors (CLP) within Sca1-ETV6-RUNX1 mice ($n = 4$) compared with age-matched control littermate WT mice ($n = 4-7$). The Mann-Whitney U test was used throughout **A-B** and P value is indicated in each case. Data represent the means \pm SD. **C**, Differential gene expression analysis of preleukemic proB and preB cells of four ETV6-RUNX1 mice compared with proB and preB cells of four age-matched WT mice from CF shows significant differences. **D**, *Rag* expression. Relative expression of *Rag1* and *Rag2* from BM pro/preB cells of ETV6-RUNX1 mice under infection exposure. BM proB/preB cells of WT mice also under infection exposure were used as a reference. Error bars represent the mean of relative expression \pm SD of three replicates.

However, we identified secondary genomic alterations in genes known to be relevant for human ETV6-RUNX1 pathogenesis, but with a very heterogeneous pattern in the single tumors, which is a common feature of human ETV6-RUNX1 pB-ALL (Table 3). Among others, we identified a deletion in the B-cell differentiation factor gene *Ebf1* leading to the loss of three amino acids, which is associated with human ETV6-RUNX1 leukemia (Supplementary Fig. S4B; ref. 6).

Infection-driven murine pB-ALL and human ETV6-RUNX1⁺ pB-ALL share critical secondary genetic events affecting histone modification regulators

ETV6-RUNX1 pB-ALL cases commonly exhibit a high burden of DNA copy-number alterations, but lack a known unifying genetic

second hit. Thus, we next compared the genomic alterations found in the infection-driven murine pB-ALL to the published genomic data of pediatric ETV6-RUNX1 pB-ALL (6, 7) and identified multiple CNVs that are shared between the infection-driven murine and human ETV6-RUNX1 pB-ALL (Table 3; ref. 7). In one of our six index mice (no J408), we found an SNV in *Kdm5c*, causing a premature stop, leading to loss of function (Fig. 4A). A significant number of genes implicated in histone modification are simultaneously mutated in the human ETV6-RUNX1⁺ pB-ALL cohort of Papaemmanelli and colleagues, including *KDM3B*, *KDM4C/D/E*, *KDM5A/C*, and *KDM6A/B* (Supplementary Table S6; ref. 7). We next validated the new finding of recurrent histone modification in an independent human ETV6-RUNX1 cohort of 11 patients. We subjected matched leukemic (initial

Table 2. KDM genes misregulated in *ETV6-RUNX1* human pB-ALL and murine *ETV6-RUNX1* preleukemic cells

Gene	Regulation in preleukemic <i>ETV6-RUNX1</i> mice vs. WT in SPF (pro/preB)	Regulation in preleukemic <i>ETV6-RUNX1</i> mice vs. WT in CF (pro/preB)	MILE t(12;21)	Regulation in leukemic <i>Pax5</i> ^{+/−} mice (pro/preB)
<i>KDM2A</i>	—	1.20	Up	−0.69
<i>KDM3A</i>	—	1.29	≈	−0.46
<i>KDM3B</i>	—	1.15	Up	−0.69
<i>KDM4A</i>	—	1.30	Up	−0.63
<i>KDM4C</i>	—	1.15	Up	−0.68
<i>KDM5A</i>	—	1.32	≈	−0.64
<i>KDM5B</i>	—	1.45	Up	−0.08
<i>KDM5C</i>	—	1.27	Up	−0.69
<i>KDM6A</i>	—	1.06	≈	—
<i>KDM6B</i>	—	1.54	Down	−0.42

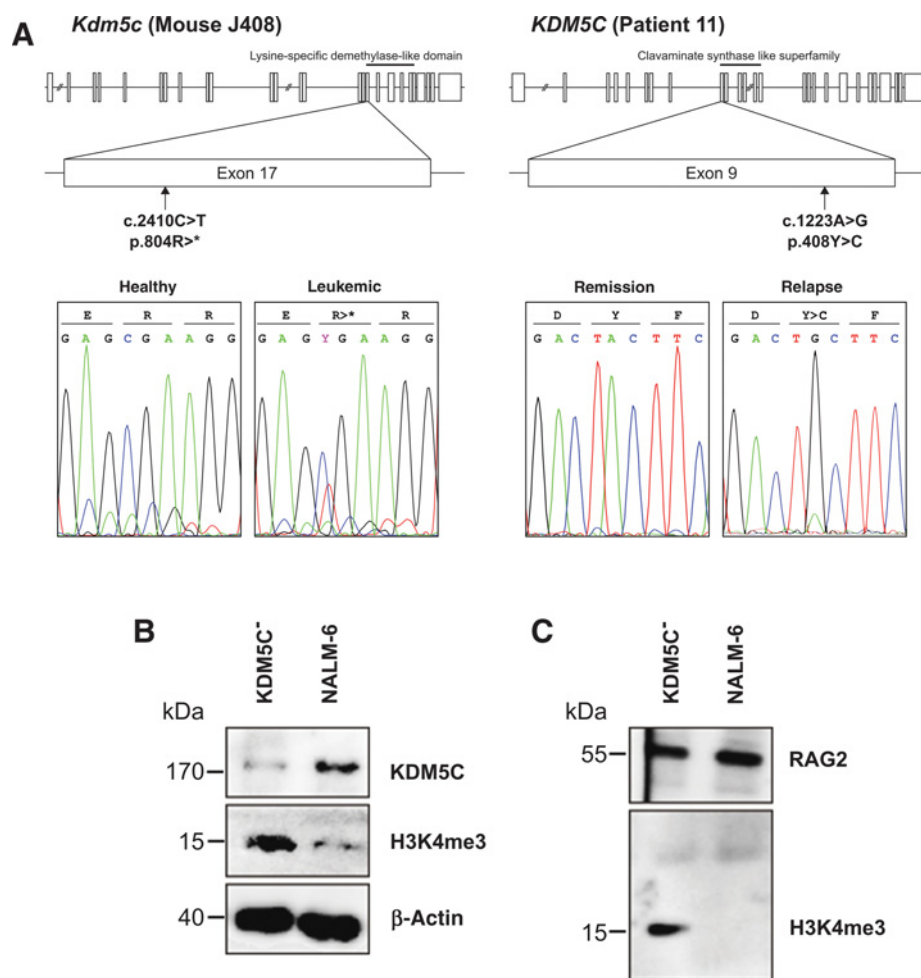
NOTE: Numbers indicated in the table correspond to the R-fold.

diagnosis and relapse) and remission DNA from 11 children with pB-ALL to WGS and WES (Supplementary Fig. S1 and Supplementary Table S1). All patients harbored the translocation t(12;21)(p13;q22) (Supplementary Fig. S1). We identified an average of 10 sequence mutations (range, 0–42) per case (Supplementary Tables S7–S9). Missense mutations (63%) were predicted to be deleterious (Supplementary Table S9), indicating

that many of these variants are involved in leukemogenesis. Our results confirmed the rarity of both recurrent coding region mutations and kinase mutations as published by Papaemmanuil and colleagues (7). However, we observed a high frequency of somatic alterations targeting histone-modifying genes in *ETV6-RUNX1* pB-ALL. Six out of 11 cases had alterations in genes encoding components of the epigenetic protein families KDM,

Table 3. Shared genomic alterations in murine and human *ETV6-RUNX1* pB-ALL

Chr.	Region	Copy number (mouse)	Affected mouse	Human samples		Genes in this region
				No. of affected samples (own cohort)	Duplicated (dup) or deleted (del) in Papaemmanuil et al. (7)	
1	1–8.130.000	3	S825	7	dup	<i>Lypla1, Tcea1, Atp6v1h, Oprk1, Rblcc1, Fam150a, Stt8, Pcmttd1, Xkr4, Rpl, Sox17, Mrpl15, Rgs20</i>
1	109.800.001–117.850.000	3	S825	1	dup	<i>Cdh7, Cntnap5a, Cdh19, Dsel</i>
2	18.110.001–18.490.000	3	J408	0	dup	<i>Mllt10, Dnajc1</i>
2	22.390.001–22.580.000	1	S825	1	del	<i>Myo3a</i>
3	8.200.001–8.290.000	4	S825	0	dup	
3	40.520.001–40.680.000	1	S825	0	del	<i>Intu, 1700017G19Rik</i>
3	131.340.001–131.500.000	1	S825	0	del	<i>Sgms2</i>
4	89.260.001–89.400.000	1	S825	0	del	<i>Cdkn2a, Cdkn2b</i>
5	124.660.001–124.710.000	3	J406	2	dup	<i>Atp6v0a2</i>
6	133.550.001–133.940.000	3	S825	0	dup	<i>Gm5886, Kap</i>
9	24.500.001–24.590.000	3	J408	1	dup	<i>Dpy19l1, Dpy19l2</i>
9	13.280.001–13.440.000	1	S825	1	del	<i>Ccdc82</i>
9	24.500.001–24.590.000	4	S825	1	dup	<i>Dpy19l1, Dpy19l2</i>
9	108.820.001–108.940.000	1	S825	0	del	<i>Celsr3, Slc26a6, Tmem89, Uqcrc1</i>
9	123.840.001–123.930.000	1	S825	0	del	<i>Fyco1, Xcr1</i>
10	7.470.001–7.610.000	1	S825	0	del	<i>Lrp11, Ulbp1</i>
10	100.290.001–100.420.000	4	S825	3	dup	
11	54.100.001–54.190.000	4	S825	3	dup	<i>P4ha2, 4933405E24Rik</i>
12	68.890.001–68.990.000	3	S825	0	dup	
14	14.090.001–14.600.000	4	S825	2	dup	<i>Atxn7, Il3ra, Psmd6</i>
14	88.130.001–97.120.000	3	S825	0	dup	<i>4930474H20Rik, 4921530L21Rik, Pcdh20, Pcdh9, Kihl1</i>
15	96.380.001–96.490.000	3	J406	0	dup	<i>Arid2, Scaf11</i>
15	6.570.001–6.960.000	3	J408	3	dup	<i>Fyb, Rictor, Osmr</i>
15	12.090.001–12.300.000	3	J408	3	dup	<i>Zfr, Mtmr12</i>
15	54.970.001–55.320.000	3	J408	0	dup	<i>Deptor, Col14a1, Taf2, Dsccl, Gm9920</i>
15	57.860.001–58.250.000	3	J408	0	dup	<i>Tbc1d31, Fam83a, Wdyh1, Der11, 9130401M01Rik, Zhx1, Atad2, Fbxo32</i>
15	96.140.001–96.500.000	3	J408	0	dup	<i>Gm4371, Arid2, 2610037D02Rik, Scaf11</i>
15	100.230.001–100.350.000	3	J408	0	dup	<i>Atf1, Tmprss12, Mettl7a1, Mettl7a3</i>
15	74.940.001–74.990.000	4	S825	0	dup	<i>Ly6e, Ly6i</i>
15	103.710.001–104.043.686	3	S825	0	dup	<i>Muc11, Spt1</i>
16	15.960.001–16.220.000	3	S825	2	dup	<i>Pkp2, Spidr</i>
16	16.380.001–16.490.000	3	S825	6	dup	<i>Fgd4</i>

**Figure 4.**

Histone Demethylase KDM5C loss of function alters the RAG2-H3K4me3 complex. **A**, Mutations in the murine and human *Kdm5c/KDM5C* gene. Left, position of the gained stop mutation in exon 17 of mouse J408 and the result of the Sanger sequencing of said mouse. Right, the same for human patient 11 and the missense mutation in exon 9. **B**, CRISPR/Cas9-mediated knockout of *KDM5C* was carried out in a pB-ALL cell line, NALM-6, and referred to as KDM5C^{-/-}. Normal NALM-6 and KDM5C^{-/-} cells were immunoblotted with anti-KDM5C and anti-H3K4me3 Ab. **C**, Coimmunoprecipitation was performed with anti-c-myc beads after nucleofection with c-myc-tagged-RAG2 plasmid to NALM-6 and KDM5C^{-/-} cells, followed by immunoblotting with anti-RAG2 and anti-H3K4me3 Ab.

HDAC, PMT, including sequence mutations in *KDM6A*, *KDM5C*, *KDM6B*, and *KDM2B* as well as deletions, including related genes (*KDM2A/2B*, *KDM3A*, *KDM4A*, and *KDM5B*). Eight of these genes (*KDM6A/B*, *KDM5B/C*, *KDM4A*, *KDM3A*, and *KDM2A/B*) were differentially regulated in *ETV6-RUNX1* pre/proB cells (Table 2 and Supplementary Table S5). Together, 26 cases (42%) of *ETV6-RUNX1* pB-ALL, including our patients ($n = 11$) and the cohort of Papaemmanuil and colleagues ($n = 51$), had mutations affecting epigenetic regulation, which involved multiple genes in 14 cases. *In silico* prediction tools classify the mutations observed in *ETV6-RUNX1* pB-ALL as loss of function. Genes of the KDM family were specifically affected in *ETV6-RUNX1* relapsed patients (71.45%) compared with initial diagnosis and other types of pB-ALL such as relapsed hyperdiploid pB-ALL (33.33%, $n = 6$). The most affected histone modifying gene functions in both cohorts were found in H3K9me3 methylation/demethylation and H3K4me2/3 demethylation. Recurrently affected genes, which catalyze demethylation of histone 3 lysine 4 (H3K4me3), were *KDM2B* and *KDM5C*. Case UKD11 harbored a focal missense mutation of *KDM5C* (Fig. 4A), and an additional case PD4021a (7) showed deletion of the whole genomic locus (Supplementary Table S6). Additionally, patient UKD06 had a missense mutation in *KDM2B* and PD4036a and UKD09 harbored structural aberrations affecting the same gene. There-

fore, we conclude that the KDM gene family is implicated in the clonal evolution of *ETV6-RUNX1* pB-ALL and its disruption is implicated in the pathogenesis of pB-ALL.

Histone demethylase KDM5C loss of function alters the RAG2-H3K4me3 complex

It is known that KDM5C loss of function causes trimethylation of H3K4, which is the necessary requirement for Rag binding and initiation of recombination activity (47). In support of this mechanism being a rational explanation of Rag-induced clonal evolution of *ETV6-RUNX1* pB-ALL, we knocked out *KDM5C* in a human pB-ALL cell line model using the CRISPR/Cas9 technique. Western blot analysis confirmed strong diminished KDM5C expression together with elevated global levels of H3K4me3 (Fig. 4B). As predicted (47), higher levels of H3K4me3 were coprecipitated with RAG2, indicating the higher ratio of H3K4me3-bound RAG2 in the KDM5C-deficient environment, laying the basis for RAG1 binding and potential enhanced off-target cleavage activity (Fig. 4C). Taken together, these findings suggest that exposure to infection, accompanied by an increase in *Rag1/2* expression contributes to the clonal evolution of *ETV6-RUNX1* pB-ALL on the basis of misregulated histone modification, facilitating Rag recruitment to cryptic RSS.

Discussion

Understanding the biology underlying *ETV6-RUNX1* pB-ALL has been hampered by the lack of genetically modified mouse models that develop human-like disease (10–15). Although the striking uniformity of clinical features, immunophenotype transcriptional profile, and therapeutic response suggests a common underlying genetic alteration in driving *ETV6-RUNX1* pB-ALL, the disease appears to be diverse due to the presence of a remarkable diversity of genetic alterations identified in previous studies (7). Despite this heterogeneity, our results show the prevalence of loss-of-function mutations in genes involving histone modification especially of the KDM family. This suggests a common pathogenesis for the establishment of the *ETV6-RUNX1* pB-ALL clone and strongly indicates that disruption of these genes is a key event in the pathogenesis of this leukemia. The presence of the fusion gene is only associated with pB-ALL, and it seems to confer very low penetrance of leukemia (3, 5). The infection-driven pB-ALLs that developed in *Sca1-ETV6-RUNX1* mice closely resemble the human disease both in low penetrance and in pathology and genomic lesions. Therefore, the *Sca1-ETV6-RUNX1* mice mimicked human *ETV6-RUNX1* preleukemic biology (4, 38) and provided a means to evaluate the potential for oncogenic environments that contribute to pB-ALL development.

We previously proved the infectious hypothesis stated by Ward in 1917 in a murine model of *Pax5* haploinsufficiency (46), which can now be extended to the commonest subtype of childhood leukemia. It is of much broader interest for our general understanding of leukemia development to know if the very common *ETV6-RUNX1* fusion acts cooperatively with infection. In this model, exposure to infection causes the oncogenic environment conferring a selection pressure on an altered hematopoietic/B-cell precursor compartment. However, the mechanism is distinct from that observed on the basis of *Pax5* haploinsufficiency. *Sca1-ETV6-RUNX1* mice show accumulation of B cell precursors in BM only when they are exposed to infection but they never show peripheral B-cell deficiency. Interestingly, the *ETV6-RUNX1* B-cell precursors are not sensitive to IL7 withdrawal (Supplementary Fig. S5). Hence, during exposure to infection the selection pressure for activating mutations of the *Jak3/Stat5* signaling pathway is less profound in comparison to what we observed in the *Pax5* model (46). Thus, the mechanism for tumor cell selection in the *ETV6-RUNX1* model is different from that in *Pax5*^{+/-} mice. We did not find activating *Jak3* mutations in this murine model or mutations activating the *IL7/IL7R/Jak/STAT* signaling pathway. However, we observed shared genomic alterations in human *ETV6-RUNX1* pB-ALL and our *ETV6-RUNX1* mouse model involving deleterious mutations in histone modifying genes (i.e., *Kdm5c*). These loss-of-function mutations in histone-modifying genes are known to facilitate the chromatin accessibility of Rag molecules alleviating off target DNA cleavage (47). Thus, the acquisition of secondary mutations in pB-ALL with a significant enrichment in histone modifying genes of the KDM family confers the second hit for the conversion of a preleukemic clone into the clinically overt *ETV6-RUNX1* pB-ALL disease, but the mechanism linking KDM loss to oncogenesis remains to be elucidated. As the presence of a precancerous cell is a property of many cancers, our approach provides a potentially general mechanism to identify etiologic factors of cancer.

Disclosure of Potential Conflicts of Interest

C. Cobaleda has ownership interest in a patent application and is a consultant/advisory board member for German Bundesamt fuer StrahlenSchutz. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: G. Rodríguez-Hernández, J. Hauer, A. Martín-Lorenzo, M. Müschen, C. Vicente-Dueñas, I. Sánchez-García, A. Borkhardt

Development of methodology: G. Rodríguez-Hernández, A. Martín-Lorenzo, D. Schäfer, I. García-Ramírez, F. Auer, L. Ruiz-Roca, M. Gombert, V. Okpanyi, U. Fischer, C. Chen, S. Bhatia, R.M. Linka, M. García-Suquía, M.V. Rascón-Trincado, A. García-Sánchez, O. Blanco, M.B. García-Cenador, F.J. García-Criado, C. Vicente-Dueñas, I. Sánchez-García

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): G. Rodríguez-Hernández, A. Martín-Lorenzo, D. Schäfer, I. García-Ramírez, F. Auer, L. Ruiz-Roca, M. Gombert, V. Okpanyi, U. Fischer, C. Chen, S. Bhatia, R.M. Linka, M. García-Suquía, M.V. Rascón-Trincado, A. García-Sánchez, O. Blanco, M.B. García-Cenador, F.J. García-Criado, C. Vicente-Dueñas, I. Sánchez-García

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): G. Rodríguez-Hernández, J. Hauer, A. Martín-Lorenzo, D. Schäfer, C. Bartenhagen, F. Auer, M. Gombert, U. Fischer, M. Dugas, S. Bhatia, R.M. Linka, A. García-Sánchez, M.B. García-Cenador, C. Cobaleda, D. Alonso-López, J. De Las Rivas, C. Vicente-Dueñas, I. Sánchez-García, A. Borkhardt

Writing, review, and/or revision of the manuscript: G. Rodríguez-Hernández, J. Hauer, A. Martín-Lorenzo, D. Schäfer, F. Auer, M. Dugas, R.M. Linka, M.B. García-Cenador, F.J. García-Criado, C. Cobaleda, J. De Las Rivas, M. Müschen, C. Vicente-Dueñas, I. Sánchez-García, A. Borkhardt

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): D. Schäfer, F. Auer, L. Ruiz-Roca, V. Okpanyi, U. Fischer, C. Vicente-Dueñas, I. Sánchez-García, A. Borkhardt

Study supervision: J. Hauer, C. Vicente-Dueñas, I. Sánchez-García, A. Borkhardt

Acknowledgments

We are indebted to all members of our groups for useful discussions and for their critical reading of the manuscript.

Grant Support

J. Hauer has been supported by the German Children's Cancer Foundation, DJCLS 02R/2016, KKS A2016/07 and from the "Forschungskommission" of the medical faculty of the Heinrich Heine University. M. Müschen is an HHMI Faculty Scholar (HHMI-55108547) and supported by NIH/NCI through an Outstanding Investigator Award (R35CA197628) to M. Müschen, a Wellcome Trust Senior Investigator Award, the Leukemia and Lymphoma Society, the California Institute for Regenerative Medicine, the William Lawrence and Blanche Hughes Foundation. A. Borkhardt has been supported by the German Children's Cancer Foundation and the Federal Ministry of Education and Research, Bonn, Germany. Research in I. Sánchez-García's group is partially supported by FEDER and by MINECO (SAF2012-32810, SAF2015-64420-Rand Red de Excelencia Consolider OncoBIOSAF2014-57791-REDC), Instituto de Salud Carlos III (PIE14/00066), ISCIII- Plan de Ayudas IBSAL 2015 Proyectos Integrados (IBY15/00003), by Junta de Castilla y León (BIO/SA51/15, CSI001U14, UIC-017, and CSI001U16), Fundación Inocente Inocente and by the ARIMMORA project (European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 282891). I. Sánchez-García's lab is a member of the EuroSyStem and the DECIDE Network funded by the European Union under the FP7 program. A. Borkhardt and I. Sánchez-García have been supported by the German Carreras Foundation (DJCLS R13/26). Research in C. Vicente-Dueñas' group is partially supported by a "Miguel Servet" grant (CP14/00082-AES 2013-2016-FEDER) from the Instituto de Salud Carlos III (Ministerio de Economía y Competitividad) and by the Lady Tata International Award for Research in Leukaemia 2016–2017. A. Martín-Lorenzo and G. Rodríguez-Hernández were supported by FSE-Consejería de Educación de la Junta de Castilla y León (CSI001-13 and CSI001-15, respectively).

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received March 9, 2017; revised June 2, 2017; accepted June 6, 2017; published OnlineFirst June 19, 2017.

References

- Borkhardt A, Cazzaniga G, Viehmann S, Valsecchi MG, Ludwig WD, Burci L, et al. Incidence and clinical relevance of TEL/AML1 fusion genes in children with acute lymphoblastic leukemia enrolled in the German and Italian multicenter therapy trials. *Associazione Italiana Ematologia Oncologia Pediatrica and the Berlin-Frankfurt-Munster Study Group*. *Blood* 1997;90:571–7.
- Pui CH, Relling MV, Downing JR. Acute lymphoblastic leukemia. *N Engl J Med* 2004;350:1535–48.
- Ford AM, Palmi C, Bueno C, Hong D, Cardus P, Knight D, et al. The TEL-AML1 leukemia fusion gene dysregulates the TGF-beta pathway in early B lineage progenitor cells. *J Clin Invest* 2009;119:826–36.
- Hong D, Gupta R, Ancliff P, Atzberger A, Brown J, Soneji S, et al. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science* 2008;319:336–9.
- Mori H, Colman SM, Xiao Z, Ford AM, Healy LE, Donaldson C, et al. Chromosome translocations and covert leukemic clones are generated during normal fetal development. *Proc Natl Acad Sci U S A* 2002;99:8242–7.
- Mullighan CG, Goorha S, Radtke I, Miller CB, Coustan-Smith E, Dalton JD, et al. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature* 2007;446:758–64.
- Papaemmanuil E, Rapado I, Li Y, Potter NE, Wedge DC, Tubio J, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet* 2014;46:116–25.
- Hauer J, Borkhardt A, Sanchez-Garcia I, Cobaleda C. Genetically engineered mouse models of human B-cell precursor leukemias. *Cell Cycle* 2014;13:2836–46.
- Sloma I, Eaves CJ. TEL me all. *Cell Stem Cell* 2009;5:5–6.
- Andreasson P, Schwaller J, Anastasiadou E, Aster J, Gilliland DG. The expression of ETV6/CBFA2 (TEL/AML1) is not sufficient for the transformation of hematopoietic cell lines in vitro or the induction of hematologic disease in vivo. *Cancer Genet Cytogenet* 2001;130:93–104.
- Bernardin F, Yang Y, Cleaves R, Zahurak M, Cheng L, Civin CI, et al. TEL-AML1, expressed from t(12;21) in human acute lymphocytic leukemia, induces acute leukemia in mice. *Cancer Res* 2002;62:3904–8.
- Fischer M, Schwieger M, Horn S, Niebuhr B, Ford A, Roscher S, et al. Defining the oncogenic function of the TEL/AML1 (ETV6/RUNX1) fusion protein in a mouse model. *Oncogene* 2005;24:7579–91.
- Kantner HP, Warsch W, Delogu A, Bauer E, Esterbauer H, Casanova E, et al. ETV6/RUNX1 induces reactive oxygen species and drives the accumulation of DNA damage in B cells. *Neoplasia* 2013;15:1292–300.
- Morrow M, Horton S, Kioussis D, Brady HJ, Williams O. TEL-AML1 promotes development of specific hematopoietic lineages consistent with preleukemic activity. *Blood* 2004;103:3890–6.
- Schindler JW, Van Buren D, Foudi A, Krejci O, Qin J, Orkin SH, et al. TEL-AML1 corrupts hematopoietic stem cells to persist in the bone marrow and initiate leukemia. *Cell Stem Cell* 2009;5:43–53.
- Kwon K, Hutter C, Sun Q, Bilic I, Cobaleda C, Malin S, et al. Instructive role of the transcription factor E2A in early B lymphopoiesis and germinal center B cell development. *Immunity* 2008;28:751–62.
- Fisher S, Barry A, Abreu J, Minie B, Nolan J, Delorey TM, et al. A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* 2011;12:R1.
- Li H, Durbin R. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 2010;26:589–95.
- Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics (Oxford, England)* 2009;25:1754–60.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, et al. The Sequence Alignment/Map format and SAMtools. *Bioinformatics (Oxford, England)* 2009;25:2078–9.
- DePristo MA, Banks E, Poplin R, Garimella KV, Maguire JR, Hartl C, et al. A framework for variation discovery and genotyping using next-generation DNA sequencing data. *Nat Genet* 2011;43:491–8.
- McLaren W, Pritchard B, Rios D, Chen Y, Flicek P, Cunningham F. Deriving the consequences of genomic variants with the Ensembl API and SNP Effect Predictor. *Bioinformatics (Oxford, England)* 2010;26:2069–70.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. *Nat Methods* 2010;7:248–9.
- Kumar P, Henikoff S, Ng PC. Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. *Nat Protoc* 2009;4:1073–81.
- Forbes SA, Beare D, Gunasekaran P, Leung K, Bindal N, Boutselakis H, et al. COSMIC: exploring the world's knowledge of somatic mutations in human cancer. *Nucleic Acids Res* 2015;43(Database issue):D805–11.
- Bolstad BM, Irizarry RA, Astrand M, Speed TP. A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. *Bioinformatics* 2003;19:185–93.
- Irizarry RA, Bolstad BM, Collin F, Cope LM, Hobbs B, Speed TP. Summaries of Affymetrix GeneChip probe level data. *Nucleic Acids Res* 2003;31:e15.
- Irizarry RA, Hobbs B, Collin F, Beazer-Barclay YD, Antonellis KJ, Scherf U, et al. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* 2003;4:249–64.
- Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci U S A* 2001;98:5116–21.
- Benjamini YH. Controlling the false discovery rate: a practical and powerful approach to multiple testing. In *J Roy Stat Soc (Ser B)* 1995:289–300.
- Team RDC. A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria ISBN 3-900051-07-0, URL <https://www.r-project.org/>. 2010.
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol* 2004;5:R80.
- Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucleic Acids Res* 2002;30:207–10.
- Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A* 2005;102:15545–50.
- Ran FA, Hsu PD, Lin CY, Gootenberg JS, Konermann S, Trevino AE, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* 2013;154:1380–9.
- Hsu PD, Scott DA, Weinstein JA, Ran FA, Konermann S, Agarwala V, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. *Nat Biotechnol* 2013;31:827–32.
- Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, et al. Multiplex genome engineering using CRISPR/Cas systems. *Science* 2013;339:819–23.
- Lausten-Thomsen U, Madsen HO, Vestergaard TR, Hjalgrim H, Nersting J, Schmiegelow K. Prevalence of t(12;21)[ETV6-RUNX1]-positive cells in healthy neonates. *Blood* 2011;117:186–9.
- Perez-Caro M, Cobaleda C, Gonzalez-Herrero I, Vicente-Duenas C, Bermejo-Rodriguez C, Sanchez-Beato M, et al. Cancer induction by restriction of oncogene expression to the stem cell compartment. *EMBO J* 2009;28:8–20.
- Vicente-Duenas C, Hauer J, Ruiz-Roca L, Ingenhag D, Rodriguez-Meira A, Auer F, et al. Tumoral stem cell reprogramming as a driver of cancer: theory, biological models, implications in cancer therapy. *Semin Cancer Biol* 2015;32:3–9.
- Vicente-Duenas C, Romero-Camarero I, Cobaleda C, Sanchez-Garcia I. Function of oncogenes in cancer development: a changing paradigm. *EMBO J* 2013;32:1502–13.
- Kumar S, Wuerffel R, Achour I, Lajoie B, Sen R, Dekker J, et al. Flexible ordering of antibody class switch and V(D)J joining during B-cell ontogeny. *Genes Dev* 2013;27:2439–44.
- Swaminathan S, Klemm L, Park E, Papaemmanuil E, Ford A, Kweon SM, et al. Mechanisms of clonal evolution in childhood acute lymphoblastic leukemia. *Nat Immunol* 2015;16:766–74.

44. Yu W, Nagaoka H, Jankovic M, Misulovin Z, Suh H, Rolink A, et al. Continued RAG expression in late stages of B cell development and no apparent re-induction after immunization. *Nature* 1999;400:682–7.
45. Tsuzuki S, Seto M, Greaves M, Enver T. Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice. *Proc Natl Acad Sci U S A* 2004;101:8443–8.
46. Martin-Lorenzo A, Hauer J, Vicente-Duenas C, Auer F, Gonzalez-Herrero I, Garcia-Ramirez I, et al. Infection exposure is a causal factor in B-precursor acute lymphoblastic leukemia as a result of Pax5 inherited susceptibility. *Cancer Discov* 2015;5:1328–43.
47. Matthews AG, Kuo AJ, Ramon-Maiques S, Han S, Champagne KS, Ivanov D, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* 2007;450:1106–10.

Cancer Research

The Journal of Cancer Research (1916–1930) | The American Journal of Cancer (1931–1940)

Infection Exposure Promotes *ETV6-RUNX1* Precursor B-cell Leukemia via Impaired H3K4 Demethylases

Guillermo Rodríguez-Hernández, Julia Hauer, Alberto Martín-Lorenzo, et al.

Cancer Res 2017;77:4365-4377. Published OnlineFirst June 19, 2017.

Updated version	Access the most recent version of this article at: doi: 10.1158/0008-5472.CAN-17-0701
Supplementary Material	Access the most recent supplemental material at: http://cancerres.aacrjournals.org/content/suppl/2017/06/17/0008-5472.CAN-17-0701.DC1

Cited articles	This article cites 45 articles, 14 of which you can access for free at: http://cancerres.aacrjournals.org/content/77/16/4365.full#ref-list-1
Citing articles	This article has been cited by 3 HighWire-hosted articles. Access the articles at: http://cancerres.aacrjournals.org/content/77/16/4365.full#related-urls

E-mail alerts	Sign up to receive free email-alerts related to this article or journal.
Reprints and Subscriptions	To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org .
Permissions	To request permission to re-use all or part of this article, use this link http://cancerres.aacrjournals.org/content/77/16/4365 . Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.