Stem cell–like reprogramming is required for leukemia-initiating activity in B-ALL

Vincent Fregona, Manon Bayet, Mathieu Bouttier, Laetitia Largeaud, Camille Hamelle, Laura A Jamrog, Naïs Prade, Stéphanie Lagarde, Sylvie Hebrard, Isabelle Luquet, et al.

To cite this version:

Vincent Fregona, Manon Bayet, Mathieu Bouttier, Laetitia Largeaud, Camille Hamelle, et al.. Stem cell–like reprogramming is required for leukemia-initiating activity in B-ALL. Journal of Experimental Medicine, 2023, 221 (1), 10.1084/jem.20230279 . hal-04283055

HAL Id: hal-04283055

https://hal.science/hal-04283055

Submitted on 13 Nov 2023

JEM Journal of Experimental Medicine

ARTICLE

Stem cell–like reprogramming is required for leukemia-initiating activity in B-ALL

Vincent Fregona[1,2,3], Manon Bayet[1,2,3], Mathieu Bouttier[1,2,3], Laetitia Largeaud[1,2,3,4,5], Camille Hamelle[1,2,3], Laura A. Jamrog[1,2,3], Naïs Prade[1,2,3,4,5], Stéphanie Lagarde[1,2,3,4,5], Sylvie Hebrard[1,2,3], Isabelle Luquet[1,2,3,4,5], Véronique Mansat-De Mas[1,4,5], Marie Nolla[1,2,3,5], Marlène Pasquet[1,2,3,5], Christine Didier[1,2,3], Ahmed Amine Khamlichi[6], Cyril Broccardo[1,2,3], Éric Delabesse[1,2,3,4,5*], Stéphane J.C. Mancini[7*], and Bastien Gerby[1,2,3]

B cell acute lymphoblastic leukemia (B-ALL) is a multistep disease characterized by the hierarchical acquisition of genetic alterations. However, the question of how a primary oncogene reprograms stem cell–like properties in committed B cells and leads to a preneoplastic population remains unclear. Here, we used the *PAX5::ELN* oncogenic model to demonstrate a causal link between the differentiation blockade, the self-renewal, and the emergence of preleukemic stem cells (pre-LSCs). We show that PAX5::ELN disrupts the differentiation of preleukemic cells by enforcing the IL7r/JAK-STAT pathway. This disruption is associated with the induction of rare and quiescent pre-LSCs that sustain the leukemia-initiating activity, as assessed using the *H2B-GFP* model. Integration of transcriptomic and chromatin accessibility data reveals that those quiescent pre-LSCs lose B cell identity and reactivate an immature molecular program, reminiscent of human B-ALL chemo-resistant cells. Finally, our transcriptional regulatory network reveals the transcription factor EGR1 as a strong candidate to control quiescence/resistance of PAX5::ELN pre-LSCs as well as of blasts from human B-ALL.

Introduction

Cell quiescence and self-renewal activity are distinctive characteristics of normal stem cells that are lost throughout the differentiation process. Nevertheless, it is established that molecular reprogramming occurring in cancer cells frequently leads to tumor dedifferentiation and the acquisition of stemness features (Bradner et al., 2017). This notion has been exploited to predict the clinical outcome of patients with solid cancers and hematological malignancies (Malta et al., 2018; Ng et al., 2016; Yan et al., 2020) and supports the idea that cellular plasticity, reprogramming, and cancer resistance are tightly intertwined. Among hematological malignancies, B cell acute lymphoblastic leukemia (B-ALL) is defined as the most frequent pediatric cancer. Although current chemotherapy is efficient at reducing the tumor load by targeting proliferating and metabolically active leukemic cells, the disease relapse points to the presence of resistant cells that escape treatment (Inaba and Mullighan, 2020). Thus, the biological properties of the cell-of-origin of leukemia, including sustained self-renewal activity, cell quiescence, and drug resistance, can significantly affect leukemia treatment and should be considered in the search for new targeted therapies (Fregona et al., 2021).

The evolution and the dynamic of subclones in B-ALL cells have been fully explored in several works through the comparison of the genetic landscape of paired diagnostic and relapse samples (Dobson et al., 2020; Mullighan et al., 2008; van Delft et al., 2011; Waanders et al., 2020). These studies predicted the existence of an ancestral clone of preleukemic stem cells (pre-LSCs) harboring a restricted number of genetic alterations, such as a founding chromosomal translocation, which is not yet transformed but at the apex of the clonal hierarchy. The notion of a multistep process of B-ALL development has also been gained from studies exploring leukemia initiation and progression in monochorionic twins (Greaves, 2018; Hong et al., 2008). These approaches demonstrated that a primary chromosomal translocation by itself does not have the capacity to induce the disease but establishes a preleukemic subclonal compartment in blood cells many years before the leukemia onset. Recurrent chromosomal translocations have been described in B-ALL and

---

[1]Université de Toulouse, Inserm, Centre Nationale de la Recherche Scientifique, Université Toulouse III-Paul Sabatier, Centre de Recherches en Cancérologie de Toulouse, Toulouse, France; [2]Equipe Labellisée Ligue Contre le Cancer 2023, Toulouse, France; [3]Équipe Labellisée Institut Carnot Opale, Toulouse, France; [4]Institut Universitaire du Cancer de Toulouse-Oncopole, Toulouse, France; [5]Centre Hospitalier Universitaire de Toulouse, Toulouse, France; [6]Institut de Pharmacologie et de Biologie Structurale, Université de Toulouse, Centre Nationale de la Recherche Scientifique, Université Toulouse III—Paul Sabatier (UT3), Toulouse, France; [7]Université de Rennes, Etablissement Français du Sang, Inserm, MOBIDIC—UMR_S 1236, Rennes, France.

*É. Delabesse and S.J.C. Mancini contributed equally to this paper. Correspondence to Bastien Gerby: bastien.gerby@inserm.fr.

Rockefeller University Press
J. Exp. Med. 2024 Vol. 221 No. 1 e20230279
https://doi.org/10.1084/jem.20230279

lead to the expression of chimeric fusion proteins. These oncoproteins often harbor domains of transcription factors (TFs) (e.g., ETV6::RUNX1, TCF3::PBX1, PAX5::ETV6, and PAX5::ELN) (Coyaud et al., 2010; Familiades et al., 2009; Mullighan et al., 2007). For instance, the TF PAX5, known to play an important role in the transcriptional regulatory networks (TNRs) of early B cells, the definitive B cell commitment, and the plasticity of committed B cells (Cobaleda et al., 2007a, 2007b), was also shown to promote leukemogenesis once translocated in mouse models (Jamrog et al., 2018; Jurado et al., 2022; Smeenk et al., 2017). Thus, it is widely thought that recurrent chromosomal translocations can act as a first oncogenic event in the early steps of B-ALL initiation, though the precise mechanisms at play are yet ill-known.

The development of transgenic mouse models with activated primary oncogenes not only helped to identify the oncogenic collaborators required for leukemia development but also allowed to identify pre-LSCs in the early steps of the disease. Various studies using in vivo models demonstrated that, in both lymphoid and myeloid malignancies, a primary genetic alteration can confer stem cell–like properties to normally committed progenitors and convert them into self-renewing pre-LSCs (Cozzio et al., 2003; Gerby et al., 2014; Huntly et al., 2004; Krivtsov et al., 2006; McCormack et al., 2010; Wojiski et al., 2009). Other studies using T-acute lymphoblastic leukemia (T-ALL) mouse models showed that pre-LSCs are resistant to chemotherapeutic agents because of their distinctive slow-division rate (Gerby et al., 2016; Tremblay et al., 2018). These findings support the view that self-renewal and cell quiescence are early and obligatory events in leukemia initiation, two specific features of the cell-of-origin, and differ from the propagating activity of fully transformed leukemic blasts. However, how a primary oncogene reprograms committed B cell progenitors into pre-LSCs with altered self-renewal and survival properties remains to be fully explored (Fregona et al., 2021).

To this end, we took advantage of a genetically engineered mouse model expressing the human PAX5::ELN fusion oncoprotein (Jamrog et al., 2018), recurrently identified in B-ALL patients (Bousquet et al., 2007). PAX5::ELN-expressing mice efficiently develop B-ALL, reproduce accurately the key features of leukemia development including secondary mutations acquired during the multistep process of human B-ALL, and provide an unrestrictive and reproducible access to a pre-LSC population (Jamrog et al., 2018). Here, by combining genetic and functional assays, together with gene expression and chromatin accessibility profiling, we deciphered the biological mechanisms by which PAX5::ELN reprograms normal B cell progenitors and leads to the emergence of quiescent pre-LSCs. Our data demonstrate that B-lineage identity is perturbed in quiescent pre-LSCs, associated with the reactivation of an immature molecular program. Moreover, our TNR predicted a list of TFs to control pre-LSC reprogramming. Among them, *Early growth response 1* (Egr1) represents a prime candidate TF to regulate pre-LSC quiescence. Finally, our analyses suggest that the pre-LSC signature mimics the resistance/quiescence of human B-ALL cells. In particular, we show that *EGR1* expression is associated with poor clinical outcomes and is activated in quiescent/resistant leukemic blasts from patients.

Results

PAX5::ELN oncoprotein induces aberrant phenotype and function in preleukemic B cells

Adult B cell development is initiated in the bone marrow (BM) by the entry of hematopoietic progenitors into the B cell lineage transcription program and the sequential rearrangement of immunoglobulin heavy (IgH) and light (IgL) chain loci through V(D)J recombination (Hardy et al., 2000). Phenotypic characterization of the different B cell subsets in mice using a combination of cell surface markers was established (Fig. S1 A) (Aurrand-Lions and Mancini, 2018). We previously developed a transgenic mouse model in which the human cDNA encoding the PAX5::ELN oncoprotein was inserted within the *IgH* locus to ensure an early and B cell–restricted expression (Jamrog et al., 2018). To determine whether PAX5::ELN oncoprotein perturbs the normal B cell development at the preleukemic stage, we performed a multiparametric staining by FACS covering exhaustively the steps of B cell differentiation (Fig. S1 A and Table S1). Unsupervised clustering allowed the visualization of the natural phenotypic progression of B cells throughout differentiation and showed that all the normal B cell subpopulations are present in preleukemic PAX5::ELN (PE^{tg}) mice as compared with wild type (wt) littermates (Fig. 1, A and B; and Fig. S1, B and C). Indeed, we distinguished in both lines the presence of pre-pro-B ($B220^+CD19^-CD2^-$), pro-B ($B220^+CD19^+Kit^+$), pre-BI ($B220^+CD19^+BP1^+$), early pre-BII ($B220^+CD19^+CD2^+IL7r^+$), late pre-BII ($B220^+CD19^+CD2^+IL7r^-$), immature B ($B220^+CD19^+Ig\kappa\lambda^+IgM^+$), and circulating/mature-B ($B220^+CD19^+Ig\kappa\lambda^+IgM^+CD23^+$) cell populations, corresponding to the different stages of differentiation (Fig. 1, A–C; and Fig. S1, B and C). However, our clustering approach revealed an aberrant population of B cell progenitors (aberrant B) in the PE^{tg} BMs (Fig. 1, A and B), featuring low expression of B220 ($B220^{low}$) and inappropriate expression of Kit (CD117), IL7r (CD127), IgM, and λ5 (CD179b) (Fig. 1 C), which is directly associated with the transgene expression (Fig. S1 D). Indeed, ELN expression is restricted to aberrant B cells ($CD19^+B220^{low}$) and is negative in all the normal B cell subpopulations ($CD19^+B220^+$) as well as in hematopoietic progenitors ($CD19^-B220^-Kit^{+/low}$) from PE^{tg} mice (Fig. S1 D). Interestingly, we noticed through the exploration of BP1, CD2, CD25 (IL2rα), and Igκλ markers, a phenotypic progression within the aberrant B population that is reminiscent of normal B cell differentiation (Fig. 1 A and Fig. S1 E). Moreover, in addition to the strong expression of Kit, we also observed a concomitant expression of the stem cell surface marker Sca-1 (Morcos et al., 2017) in about 30% of the aberrant B cells from the preleukemic PE^{tg} BM (Fig. S1 F). Together, these results indicate that PAX5::ELN oncoprotein induces a partial blockade of differentiation associated with the emergence of an aberrant B cell population in a steady-state condition.

To explore the functional impact of PAX5::ELN oncoprotein on preleukemic B cells, we performed serial transplantations of wt and PE^{tg} preleukemic total B cells (Fig. 1 D). Short-term

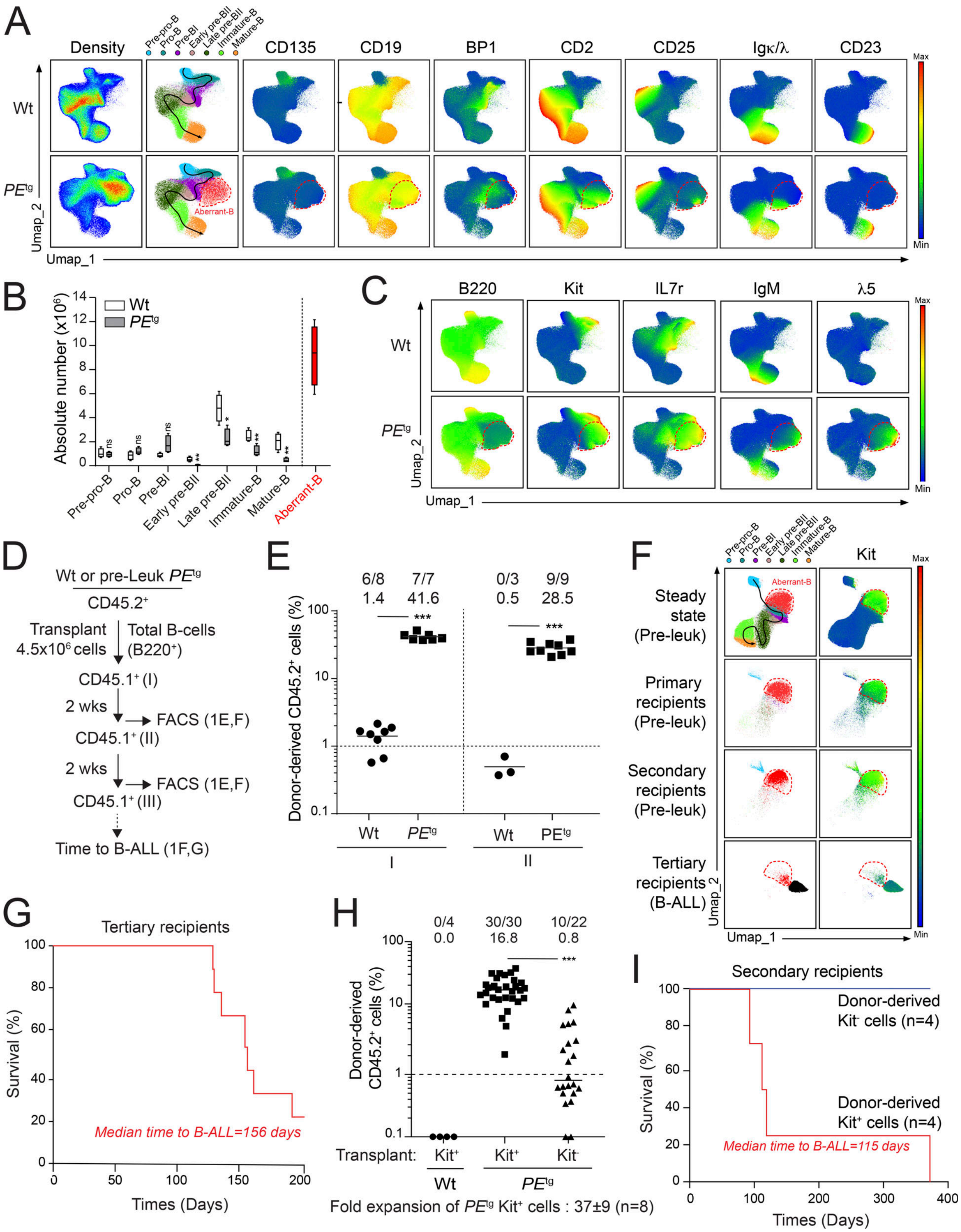

Figure 1. **PAX5::ELN induces aberrant B cells at the preleukemic stage. (A–C)** Phenotypic characterization of the B cell lineage from the BM of wt and PAX5::ELN (*PE*^{tg}) preleukemic mice (30 days old, $n = 4$). UMAP of the cell density (A, left panels), of the clustering analysis of the B cell subpopulations (A, central panels) based on the gating strategy shown in Fig. S1, B and C. Each subset is represented by one color. Black arrow indicates the physiological phenotypic progression of B cells in the differentiation. Red dotted line delimits the aberrant preleukemic B cell population induced by PAX5-ELN (A, central panels). UMAP of the expression level of CD135, CD19, BP1, CD2, CD25, Igκ/λ, and CD23 (A, right panels). Absolute numbers of B cell subpopulations were calculated. The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values (B, $^*P < 0.05$, $^{**}P < 0.01$). UMAP of the expression level of B220, Kit, IL7r, IgM and λ5 (C). Data were compiled

from four mice per condition and are representative at least of two independent experiments ($^{*}P < 0.05$, $^{**}P < 0.01$). **(D)** Experimental procedure to study the functional impact of PAX5::ELN on B cells. **(E)** Engraftment efficiency (CD45.2^{+}) in the recipient BM of primary (I) and secondary (II) mice was analyzed each 2 wk after transplantation. The number of positive mice and the median of engraftment are indicated ($n = 3–9$, $^{***}P < 0.001$). **(F)** UMAP of the CD45.2^{+} PE^{tg} B cell subpopulations (left panel) before (steady state) and after transplantation in primary, secondary, and tertiary recipients, associated with the expression level of Kit (right panels). **(G)** Kaplan–Meier survival curve of tertiary recipients transplanted with donor-derived PE^{tg} B cells ($n = 9$). **(H)** Engraftment efficiency in the recipient BM of primary mice was analyzed 3 wk after transplantation. The number of positive mice, the median of engraftment and the in vivo fold expansion of PE^{tg} Kit^{+} cells are indicated. Data were compiled from three independent experiments ($n = 4–30$, $^{***}P < 0.001$). **(I)** Kaplan–Meier survival curve of secondary recipient mice respectively transplanted with donor-derived Kit^{+} cells ($n = 4$, red) or Kit^{-} cells ($n = 4$, blue).

---

reconstitution potential was assessed by analyzing the BM engraftment 2 wk after transplantations in primary and secondary recipient mice. In contrast to normal B cells that are devoid of self-renewal potential, preleukemic B cells efficiently engrafted over primary and secondary recipients (Fig. 1 E). The phenotypic characterization of engrafted cells confirmed that this short-term self-renewal activity is associated with the selection and expansion of the aberrant B cell population (Fig. 1 F). Indeed, this phenotype was strongly stable over primary and secondary recipients (Fig. S2 A), as exemplified by the high level of Kit expression (Fig. 1 F). Moreover, tertiary transplantations of preleukemic B cells eventually led to long-term B-ALL development in about 5 mo (Fig. 1 G), which is the delay required to acquire the full pattern of transforming secondary mutations (Jamrog et al., 2018). Strikingly, fully transformed leukemic blasts from tertiary recipients exhibited a distinct phenotype (Fig. 1 F and Fig. S2 A) and were able to efficiently propagate B-ALL a few days only after quaternary transplantations, even with low doses of injected cells (Fig. S2 B). Finally, Kit^{+} and Kit^{-} B cells from PE^{tg} BMs were purified to compare their short-term engraftment capacity after transplantation in primary mice (Fig. S2 C). We controlled that purified Kit^{+} cells from wt mice were devoid of engraftment potential, and that the aberrant self-renewal potential of PE^{tg} B cells was well enriched in the Kit^{+} compartment (Fig. 1 H). Moreover, we confirmed that only Kit^{+} preleukemic B cells sustained long-term B-ALL development after secondary transplantation (Fig. S2 C and Fig. 1 I). Together, our results suggest that PAX5::ELN oncoprotein converts Kit^{+} B cell progenitors into pre-LSCs that have acquired de novo self-renewing activity in the first steps of the disease.

Preleukemic cells restricted in the cell cycle are chemoresistant and support leukemia-initiating activity

Cell-cycle restriction is a critical feature of normal hematopoietic stem cells (HSCs) to preserve their self-renewal function (Wilson et al., 2008). We therefore hypothesized that cell quiescence could be a biological mark of pre-LSCs in the PE^{tg} model. To test this hypothesis, we analyzed Ki67 expression in the different B cell subsets of wt and PE^{tg} preleukemic mice in steady-state condition (Fig. 2 A and Table S1). As expected, Ki67 expression followed the physiological proliferation rate of normal B cell differentiation (Tomura et al., 2013) in wt and PE^{tg} mice (Fig. 2, A and B; and Fig. S2 D). Furthermore, non-cycling Ki67^{-} cells were detected within the aberrant PE^{tg} B cell population (Fig. 2, A and B; and Fig. S2 D). Since cell-cycle restriction is also an important characteristic of treatment-resistant cells in B-ALL (Ebinger et al., 2016; Turati et al., 2021), we assessed in vitro dose responses of doxorubicin (DOXO), methotrexate (MTX), and vincristine (VCR) on the aberrant PE^{tg} B cell population and compared the sensitivity of negative and positive Ki67 cells (Fig. 2 C and Fig. S2 E). We observed that Ki67^{-} cells were still detected after 48 h of coculture and were strongly selected upon the treatment (Fig. S2 E). Moreover, the analysis of the absolute numbers of each fraction after the treatment showed that Ki67^{-} cells were less sensitive to the three chemotherapeutic agents than Ki67^{+} cells (Fig. 2 C).

To further address the question of whether PAX5::ELN perturbs the cell division kinetic of preleukemic B cells, wt and PE^{tg} Kit^{+} B cells were labeled with the cell trace violet (CTV) and cultured for 3 days (Fig. 2 D). This in vitro cell division assay indicated that PAX5::ELN reduced the global proliferation rate of preleukemic cells and revealed the persistence of an undivided (D0) population (Fig. 2, E and F) that maintained the expression of Kit (Fig. S2 F). Next, we purified the slow-cycling (D0–2) PE^{tg} B cells population, which represents <1% of total generated cells and are undetectable in CTV-labeled wt counterparts cultured in the same experimental conditions (Fig. 2, E and F). High-cycling (D8–9) cells from both wt and PE^{tg} conditions were also purified as controls (Fig. 2 E). RNA sequencing (RNA-seq) demonstrated the presence of *PAX5::ELN* fusion transcript in both D0–2 and D8–9 preleukemic cells (Fig. S2 G) and their polyclonality of $V_H(D_H)J_H$ and V_LJ_L rearrangements (Fig. 2 G), confirming their preleukemic state. These cells were then treated in vitro separately by DOXO, MTX, or VCR (Fig. 2, D and H) with a sublethal concentration that we previously determined by dose-response experiments on total wt and PE^{tg} Kit^{+} B cells (Fig. S2 H). This assay revealed a significant resistance of slow-cycling PE^{tg} D0–2 cells in contrast to wt and PE^{tg} D8–9 cells (Fig. 2 H). To address at the functional level the role of cell-cycle restriction in pre-LSC activity, equal numbers of D0–2 and D8–9 preleukemic cells were transplanted (Fig. 2 D). Both D0–2 and D8–9 preleukemic PE^{tg} B cells were able to engraft the BM of recipient mice 3 wk after transplantation (Fig. 2 I). However, only donor-derived D0–2 PE^{tg} cells maintained high-level expression of Kit in vivo (Fig. S2 I), sustained the reconstitution 16 wk after transplantation (Fig. 2 I), and the long-term B-ALL development (Fig. 2 J). Together, our results indicate that cell-cycle restriction of preleukemic cells is associated with chemoresistance and leukemia initiation.

PAX5::ELN oncoprotein disrupts the differentiation of preleukemic cells by enforcing the IL7r/JAK-STAT signaling pathway

PAX5::ELN perturbs B cell differentiation (Fig. 1) and the cell cycle of preleukemic cells (Fig. 2). Thus, we asked whether the main B cell differentiation molecular programs are affected by

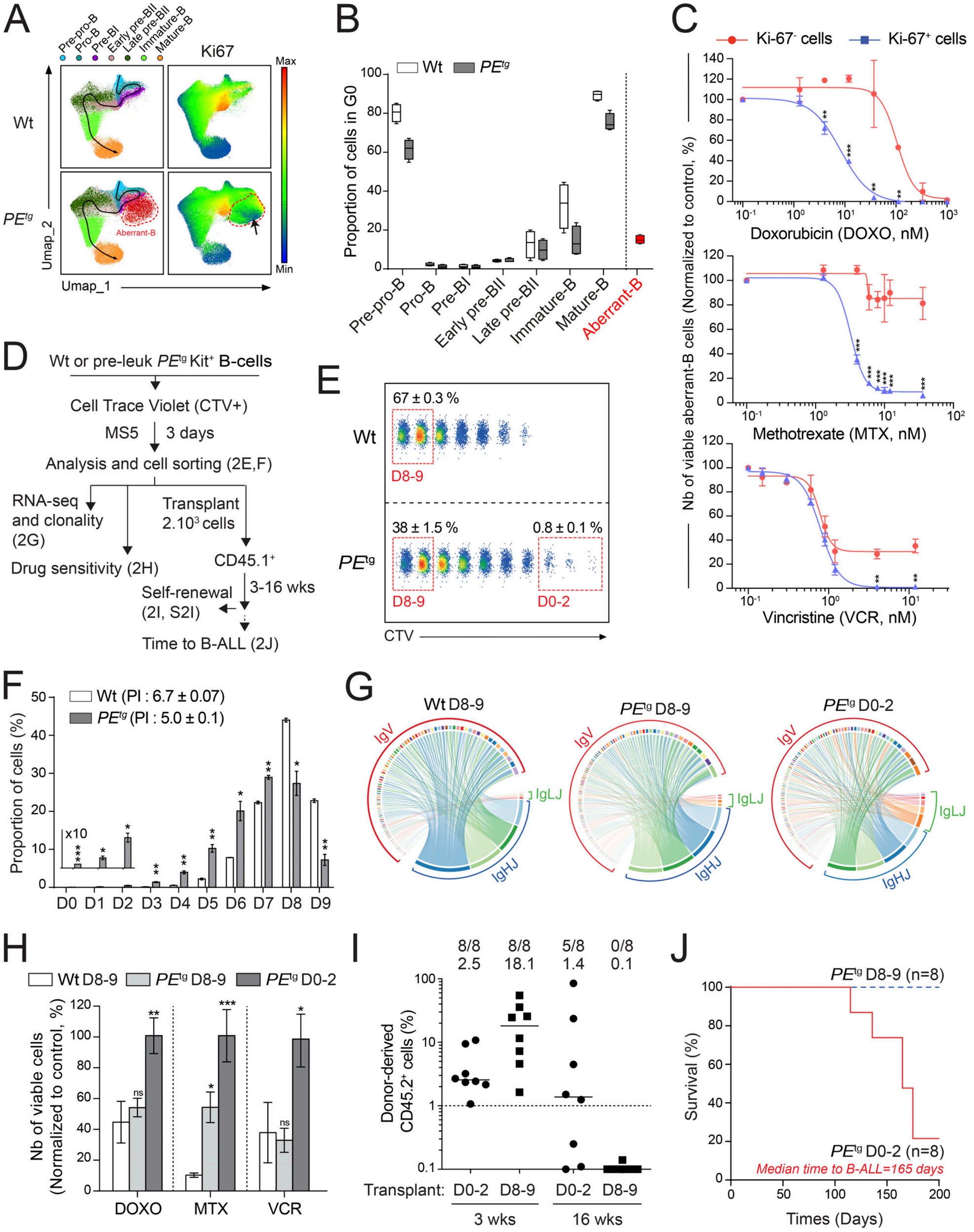

Figure 2. **Cell quiescence is a feature of PAX5::ELN pre-LSCs. (A and B)** UMAP representation of the wt and *PE*^{tg} B cell subpopulations (A, left panel), associated with the Ki67 expression level (A, right panel). The arrow indicates the population in G0 phase (Ki67^{-} cells in blue) within the aberrant B cell population. Quantification of the proportion of cells in G0 phase (Ki67^{-} cells) within each wt and *PE*^{tg} B cell subpopulation (B). Data were compiled from four mice per condition and are representative at least of two independent experiments .The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. **(C)** Preleukemic *PE*^{tg} cells were treated in vitro on MS5 stromal cells with a dose-response of DOXO, MTX, or VCR for 48 h. Absolute numbers of Ki67^{+} and Ki67^{-} cells within the aberrant B cell

population were then analyzed by FACS and normalized to the number of untreated cells ($n = 3$, one experiment, $^{**}P < 0.01$, $^{***}P < 0.001$). **(D)** Experimental procedure to study cell division of Kit^+ B-progenitors from wt and preleukemic PE^{tg} mice. **(E and F)** The number of cell divisions (D0 to D9) after the coculture was then analyzed (E), the proportion of cells in each division was quantified and the PI (PI ± SD) was calculated for each condition (F, $n = 2$, mean ± SD, representative of two independent experiments, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). Red dotted gates were used to purify the D0–2 and D8–9 populations (E). **(G)** Cell clonality analysis of PE^{tg} D0–2, PE^{tg} D8–9 and wt D8–9 populations after RNA-seq. The proportions of $V_H(D_H)J_H$ and V_LJ_L rearranged transcripts were represented by Circos diagrams. **(H)** Drug sensitivity of PE^{tg} D0–2 and D8–9 cells and of wt D8–9 cells upon the treatment with an IC50 of DOXO, MTX, and VCR during 48 h. Cell numbers of treated cells were analyzed by FACS and normalized to the number of untreated cells ($n = 3$, one experiment, mean ± SD, $^{*}P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). **(I and J)** Engraftment efficiency 3 and 16 wk after transplantation (I) and Kaplan–Meier survival curve of recipient mice transplanted with D0–2 and D8–9 PE^{tg} cells (J). The number of positive mice and the median of engraftment are indicated (I, $n = 8$). The median of time required to develop the B-ALL is indicated (J).

PAX5::ELN in slow- and high-cycling preleukemic cells. Early B cell development is controlled by the IL7 receptor and the pre-B cell receptor (pre-BCR), two interacting signaling systems that tightly coordinate alternative phases of cell survival and differentiation through the JAK-STAT and pre-BCR/BLNK pathways, respectively (Clark et al., 2014; Reth and Nielsen, 2014). Each pathway has antagonistic and balanced functions to coordinate the proliferation and differentiation switch at the pre-B stage (Fig. 3 A). Interestingly, IL7r is strongly expressed in the aberrant PE^{tg} B cell population, together with the accumulation of the surrogate light chain λ5 at their surface (Fig. 1 C). Thus, we analyzed IL7r and pre-BCR/BLNK pathways by using a phosphoflow cytometry approach after ligand and chemical ex vivo stimulations, respectively (Fig. S3 A and Fig. 3 B). Ligand stimulation of IL7r efficiently induced the phosphorylation of its two downstream effectors, STAT5 and FOXO1, in the aberrant B ($B220^{low}$) PE^{tg} population, in a similar way to normal ($B220^{high}$) pro-B cells (Fig. S3 A and Fig. 3 B, left panels). Accordingly, the presence of IL-7 protected aberrant PE^{tg} B cells from apoptosis (Fig. 3 C) and efficiently promoted their proliferation (Fig. 3 D) in vitro. In addition, this process was abrogated in the presence of tofacitinib, a well-known JAK-STAT inhibitor (Fig. 3, C and D), showing that IL-7r signaling is critical to maintaining the survival of preleukemic cells. We also assessed the (pre-)BCR signaling response by mimicking its downstream activation ex vivo with hydrogen peroxide (H_2O_2) (Fig. S3 A and Fig. 3 B, right panels), a well-known inducer of BLNK/PLCγ2 phosphorylation cascade (Patterson et al., 2015; Reth, 2002). As expected, H_2O_2 induced the phosphorylation of BLNK and PLCγ2 along the normal B cell differentiation pathway. In contrast, the aberrant PE^{tg} B cell population was unresponsive to BLNK/PLCγ2 stimulation (Fig. S3 A and Fig. 3 B, right panels).

Interestingly, ligand stimulation of IL7r efficiently induced the phosphorylation of STAT5 in both $Ki67^+$ and $Ki67^-$ cells within the aberrant PE^{tg} B cell population, which is abrogated in the presence of tofacitinib (Fig. 3 E, right panels). In contrast, IL7r activation is concentrated in cycling $Ki67^+$ cells from normal B cells (Fig. 3 E, left panels), corresponding to cycling pro-B cells (Fig. 2 B and Fig. S2 D). Moreover, gene set enrichment analysis (GSEA) from the RNA-seq of purified PE^{tg} D0–2, PE^{tg} D8–9, and wt D8–9 cells (Fig. 2 D and Table S2) confirmed that the IL7r/JAK-STAT axis was significantly upregulated by PAX5::ELN oncoprotein in slow- and high-cycling preleukemic cells (Fig. S3 B). This observation was reinforced by the comparison of our PAX5::ELN-modified genes with the Stat5-binding genes arising from the ChIP sequencing of murine *Stat5b*-induced B-ALL cells (Katerndahl et al., 2017), showing a substantial enrichment of Stat5 targets in PAX5::ELN upregulated genes in slow- (D0–2) and high- (D8–9) cycling preleukemic cells (Fig. 3 F and Table S2). Accordingly, the downstream targets of IL7r/JAK-STAT axis, including the prosurvival genes *Cdkn2a*, *Il2rα*, *Tnfrsf13b*, and *Socs3*, were significantly activated in both D0–2 and D8–9 cells (Fig. 3 G, left panels). In contrast, several components of the pre-BCR/BLNK signaling including *CD79b*, *CD79a*, *Lyn*, *Syk*, *Blnk*, *Ikzf3*, and *Irf8* were downregulated in both D0–2 and D8–9 cells (Fig. 3 G, right panels). Among them, the pre-B cell differentiation genes *Blnk*, *Ikzf3*, and *Syk* also belong to Stat5 targets (Fig. 3 F), reinforcing the notion that IL7r/JAK-STAT axis negatively regulates the B cell differentiation (Fig. 3 A).

Collectively, combined with our phosphoflow cytometry results, our RNA-seq data suggest that the activation of IL7r/JAK-STAT pathway favors the maintenance and survival of preleukemic cells, in slow- and high-cycling fractions.

B cell molecular identity is impaired in quiescent pre-LSCs

While the above results demonstrate that the leukemia-initiating potential is enriched in slow-cycling population, the imbalance between IL7r/JAK-STAT and pre-BCR/BLNK molecular circuits is a shared feature of all *PAX5::ELN* preleukemic cells. Thus, to further identify and characterize quiescent pre-LSCs, we generated PE^{tg}::$H2B$-GFP^{tg} mice, derived from doxycycline-inducible $H2B$-GFP^{tg} mice, an in vivo model allowing the study of cell quiescence in normal HSCs (Foudi et al., 2009; Wilson et al., 2008). After a doxycycline pulse, we compared the cell division kinetics in vivo between wt and preleukemic PE^{tg} B cells based on the loss of GFP expression during a time course of doxycycline chase (Fig. 4 A). At the end of the doxycycline pulse, almost all B cells in both wt and preleukemic BM expressed the H2B-GFP marker (Fig. 4 B). However, the decrease of GFP after the withdrawal (chase period) of doxycycline revealed a substantial delay of the global GFP loss in preleukemic PE^{tg} B cells, with the persistence of GFP^{high} PE^{tg} cells after 1-wk chase of doxycycline (Fig. 4 B). Based on the loss of GFP after 2- and 3-wk chase of doxycycline, we defined GFP^{high} cells as quiescent rather than dormant cells, such as reported for normal stem cells (Foudi et al., 2009; Wilson et al., 2008). Therefore, we adapted the multiparametric staining to identify the quiescent GFP^{high} B cell subset. Quiescent cells were exclusively found in the aberrant B cell population of PE^{tg} BM (Fig. S3 C and Fig. 4 C). Additionally, we observed that GFP loss was proportional to the acquisition of the Ki67 marker, almost all GFP^{high} cells did not express Ki67 (Fig. 4 D), and only this non-

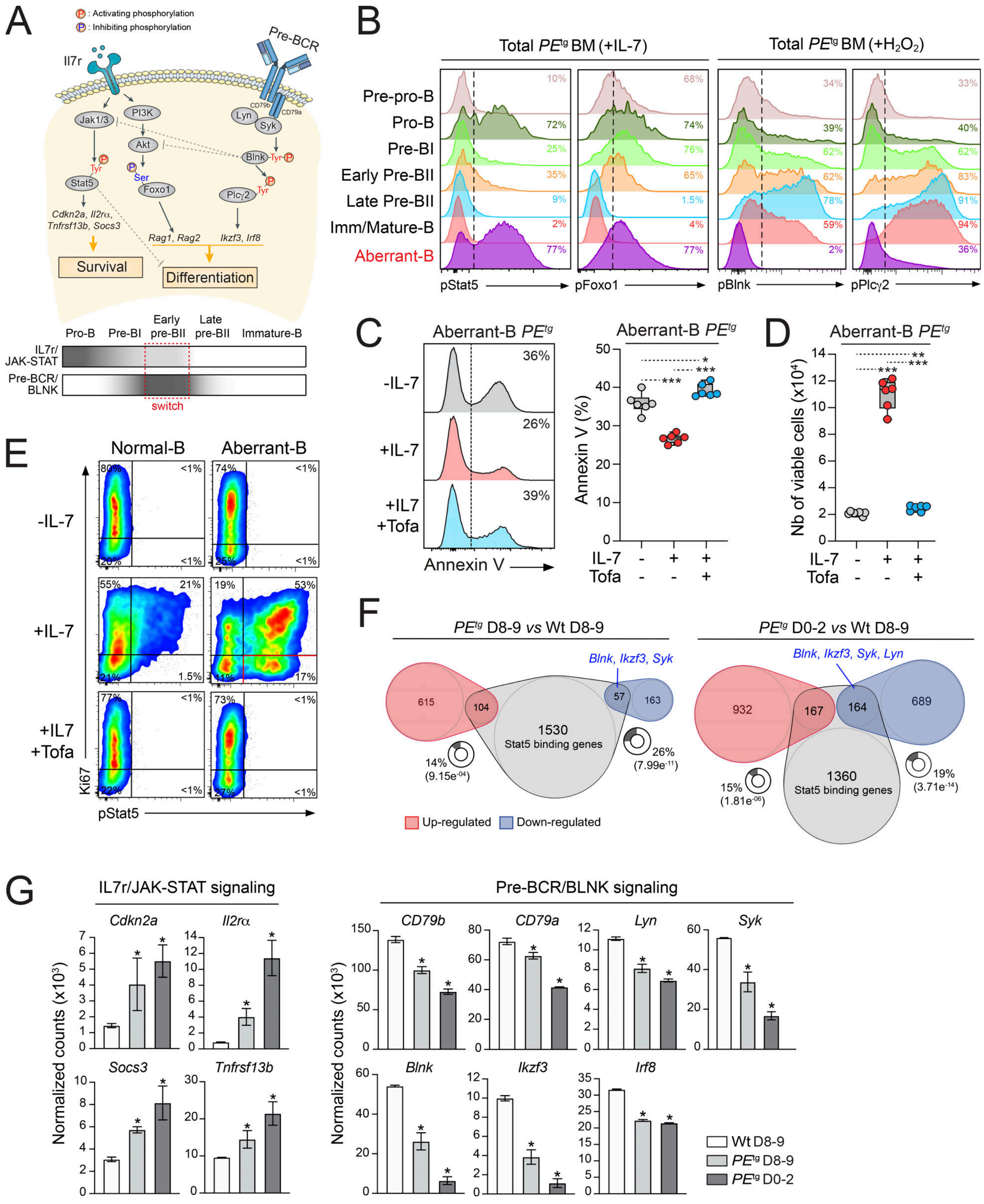

Figure 3. **Imbalance of IL7r/JAK-STAT and pre-BCR/BLNK pathways in PAX5::ELN preleukemic cells. (A)** Simplified schematic representation of the IL7r and pre-BCR signaling pathways adapted from Clark et al. (2014). **(B)** FACS analysis of pStat5 and pFoxo1 on the aberrant B cell population and on each normal B cell subsets from preleukemic *PE*^{tg} mice after ex vivo IL-7 stimulation (+IL-7) (left panels). FACS analysis of pBlnk and pPlcγ2 after ex vivo H_2O_2 stimulation (+H_2O_2) is shown in right panels. **(C and D)** Preleukemic *PE*^{tg} cells were treated in vitro on MS5 stromal cells with IL-7 (+IL-7) or not (−IL-7) for 48 h. The proportion of Annexin V$^+$ cells within the aberrant B *PE*^{tg} population was measured (left panel) and quantified (right panel) by FACS and the absolute numbers of viable aberrant B *PE*^{tg} cells were calculated (D). Cells treated with 1 µM tofacitinib (Tofa) were used as controls (*n* = 6, representative of two independent experiments, $^*P < 0.05$, $^{**}P < 0.01$, $^{***}P < 0.001$). The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. **(E)** Representative FACS analysis of pStat5 after IL-7 stimulation (+IL-7)

in presence or not (vehicle) of 1 μM tofacitinib associated with Ki67 expression in normal B cells and in aberrant B PE^{tg} cells. Unstimulated (−IL-7) cells were used as controls. **(F)** Venn diagram of PAX5::ELN-modified genes overlapped with a list of Stat5-binding genes arising from the ChIP sequencing of murine *Stat5b*-induced B-ALL cells (Katerndahl et al., 2017). Numbers in the plot indicate quantification of genes in each group. Indicated P value was calculated by hypergeometric test. **(G)** Expression level of selected genes involved in IL7r/JAK-STAT (left panel, *Cdkn2a*, *Il2rα*, *Tnfrsf13b*, and *Socs3*) and in pre-BCR/BLNK (right panel, CD79b, CD79a, *Lyn*, *Syk*, *Blnk*, *Ikzf3*, and *Irf8*) signaling pathways extracted from the RNA-seq of PE^{tg} D0–2, PE^{tg} D8–9 and wt D8–9 cells ($n = 2$, mean ± SD, $^{*}P < 0.05$).

cycling population sustained engraftment capacity after transplantation (Fig. 4 E) and the long-term B-ALL development (Fig. S3 D).

To refine the molecular characterization of quiescent pre-LSCs in vivo, GFP^{high} and GFP^{low} cells were purified from aberrant B cells of $PE^{tg}H2B\text{-}GFP^{tg}$ mice after 1-wk chase of doxycycline, and an integrative molecular approach was performed comparing both gene expression profiles by RNA-seq and chromatin accessibility by assay for transposase-accessible chromatin sequencing (ATAC-seq) (Fig. 4 F) (Buenrostro et al., 2013). Transcriptome analysis identified 836 upregulated and 789 downregulated genes with an expression difference of >1.5-fold and an adjusted P value of <0.05 in quiescent GFP^{high} pre-LSCs (Fig. 4 F and Table S3). Combined with an expected downregulation of cell cycle–related pathways, the analysis revealed that cytokine signaling pathways were upregulated in quiescent GFP^{high} pre-LSCs (Fig. S3 E). Indeed, these modifications, involving the downregulation of E2f/Myc targets and the upregulation of TNFα signaling (Fig. S3 E), are crucial for HSC and cancer stem cell survival (Cabezas-Wallscheid et al., 2017; Vazquez-Santillan et al., 2015; Yamashita and Passegué, 2019; Zhang et al., 2022). Differentially accessible chromatin regions were revealed by ATAC-seq analysis in preleukemic GFP^{high} (7,439 peaks) and GFP^{low} (5,360 peaks) cells (Fig. 4 F and Table S4). Their associated genes were enriched for molecular pathways similar to those identified at the mRNA level (Fig. S3, E and F). These observations highlight the direct relationship between the expression of genes and their surrounding accessibilities, allowing the identification of specific regulators.

To this end, we combined footprinting analysis using hint-ATAC and “diffTF” computational tools (Berest et al., 2019) to identify and classify TF motif accessibilities that were differentially regulated in quiescent GFP^{high} pre-LSCs (Fig. 4 G). We observed that the motif accessibility of critical TFs involved in the B cell identity, such as Ebf and Pax, or in the cell cycle, such as E2f, was significantly downregulated in quiescent GFP^{high} pre-LSCs (Fig. 4 G and Fig. S4 A). While the accessibility footprint (Tn5 insertion site) of Ebf and Pax was decreased in quiescent GFP^{high} pre-LSCs (Fig. S4 A), we observed that the expression of *Ebf1* and total *Pax5* was tightly increased at the transcriptional level (significantly for *Ebf1* but not for total *Pax5*, Fig. S4 B). Of note, the expressions of *Pax5* WT and *PAX5-ELN* were quietly similar between $H2B\text{-}GFP^{high}$ and $H2B\text{-}GFP^{low}$ cells (Fig. S4 B). This observation indicates that the binding activity of these two TF is affected in $H2B\text{-}GFP^{high}$ pre-LSCs, but not their expression. Interestingly, we observed that Pax5-repressed genes, but not Pax5-activated genes (Revilla-I-Domingo et al., 2012), were drastically enriched in GFP^{high} pre-LSCs (Fig. 4 H and Fig. S4 C). Together, these results strongly suggest that Pax5 has lost its repressive activity in quiescent pre-LSCs, allowing the upregulation of a non-B cell molecular program. Indeed, we observed the upregulation of the chromatin accessibility to Spi and Runx motifs in quiescent GFP^{high} pre-LSCs (Fig. 4 G and Fig. S4 A). This was associated with increased expression of Spi1 (PU.1) and Runx1 (Fig. S4 B), two critical TFs regulating the quiescence and the myeloid differentiation process of normal HSCs (Chavez et al., 2021; Imperato et al., 2015; Staber et al., 2014). This observation was also associated with a downstream activation of a myeloid program in quiescent GFP^{high} pre-LSCs at the transcriptional level. Indeed, the analysis mainly revealed the upregulation of genes involved in TNFα signaling and in myeloid differentiation in GFP^{high} pre-LSCs (Fig. 4 I). Finally, we observed a strong enrichment of both murine (Fig. 4 J and Table S3) (Venezia et al., 2004) and human (Fig. S4 D and Table S3) (Garcia-Prat et al., 2021) dormant/quiescent HSC signature in the GFP^{high} pre-LSCs, suggesting redundant molecular mechanisms between pre-LSCs and HSCs. Collectively, our data indicate that quiescent pre-LSCs partially lose B-lineage molecular identity and concomitantly reactivate an inappropriate and immature transcriptional program.

Multi-lymphoid progenitor (MLP)/hematopoietic stem and progenitor cell (HSPC) molecular signatures are partially reprogrammed in quiescent pre-LSCs

The above findings raise the question as to what extent stem cell–like molecular features are reprogrammed in quiescent pre-LSCs. Thus, we integrated our results within ATAC-seq and RNA-seq data from the Immunological Genome Project (ImmGen) consortium (https://www.immgen.org/). First, we extracted both chromatin accessibility and gene expression profiles from normal HSPCs (LT-HSCs, MPP1, ST-HSCs, MPP4), MLPs, common lymphoid progenitor (CLP), Fr.A (i.e., Pre-pro-B), and B cells (Fr.B/Pro-B to Fr.E/immature-B) (Fig. 5 A). Then, differentially accessible peaks and expressed genes from preleukemic GFP^{high} and GFP^{low} cells were clustered by a k-means approach according to their signal in the different steps of normal differentiation (Fig. 5, A–C and Table S5). Strikingly, ATAC-seq data revealed that open chromatin regions from HSPCs, including LT-HSCs and other multipotent immature progenitors (clusters C1 to C4), were enriched in quiescent GFP^{high} pre-LSCs, contrasting in this regard with GFP^{low} cycling cells that were rather enriched in accessible regions from committed B cells (clusters C7, C8) (Fig. 5 B). Furthermore, RNA-seq data identified four clusters of genes in preleukemic GFP^{high} and three clusters of genes in GFP^{low} cells. This analysis notably revealed a unique “$HSPC^{like}$ cluster” in quiescent GFP^{high} cells, composed of genes mostly upregulated in LT-HSCs, MPP1, ST-HSCs, and MPP4 fractions (Fig. 5 C). An “MLP^{like} cluster” was

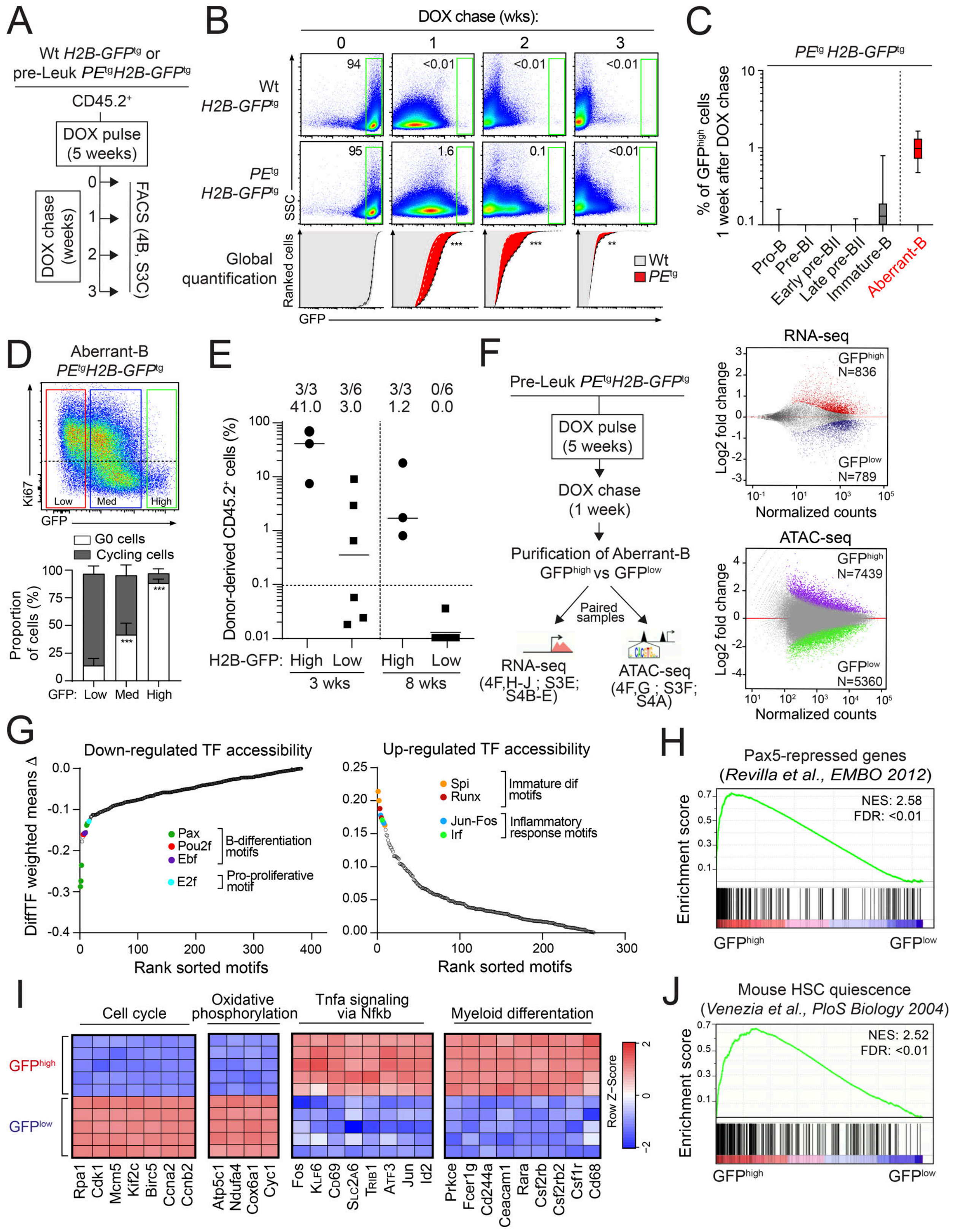

Figure 4. **Loss of B cell molecular identity in quiescent pre-LSCs. (A)** Experimental procedure to study the cell division kinetics in vivo of wt and preleukemic B cells. TFBS, TF binding site. **(B)** Representative FACS analysis of GFP expression in B cells ($CD19^+CD23^-$) from the BM of wt *H2B-GFP*^{tg} ($n = 14$, upper panels) and preleukemic *PE*^{tg}*H2B-GFP*^{tg} ($n = 12$, central panels) mice after 0, 1, 2, and 3 wk of DOX chase. AUC quantification of the GFP was compared between the two conditions (lower panels, $^{**}P < 0.01$, $^{***}P < 0.001$). **(C)** Quantification of the proportion of GFP^{high} within preleukemic *PE*^{tg}*H2B-GFP*^{tg} ($n = 12$) B cell subpopulations after 1 wk of DOX chase. The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. **(D)** Representative FACS analysis (upper panel) and quantification (lower panel, $n =$

8, mean ± SD, ***P < 0.001) of the proportion of G0 (Ki67⁻) and cycling (Ki67⁺) cells in GFP^{low}, GFP^{med} and GFP^{high} populations. Data from B–D were representative at least of two independent experiments. **(E)** Engraftment efficiency of purified GFP^{high} (*n* = 3) and GFP^{low} (*n* = 6) cells from *PE*^{tg}*H2B-GFP*^{tg} mice was analyzed by FACS 3 and 8 wk after transplantation. The number of positive mice and the median of engraftment are indicated. **(F)** Experimental procedure for the molecular characterization of quiescent pre-LSCs in vivo (left panel). Scatter plots showing differentially expressed genes (RNA-seq, *n* = 5) and differentially chromatin accessibility (ATAC-seq signal, *n* = 3) between purified GFP^{high} and GFP^{low} cells from *PE*^{tg}*H2B-GFP*^{tg} mice (right panels). **(G)** TF motifs with differential chromatin accessibility between GFP^{high} and GFP^{low} cells identified and sorted using the DiffTF computational tool (Berest et al., 2019). Down- (left panel) and up- (right panel) regulated TF accessible regions in quiescent GFP^{high} cells were represented. **(H)** GSEA of the well-established list of in vivo PAX5-repressed genes (Revilla-I-Domingo et al., 2012) between GFP^{high} and GFP^{low} cells. **(I)** Heatmap of genes involved in indicated pathways between GFP^{high} (*n* = 5) and GFP^{low} (*n* = 5) cells. **(J)** GSEA of the established list of in vivo mouse HSC quiescence genes (Venezia et al., 2004) between GFP^{high} and GFP^{low} cells. The normalized enrichment score (NES) and the false discovery rate (FDR) are indicated in H and J.

---

also defined in quiescent GFP^{high} cells, composed of genes upregulated in CLP and pre-pro-B precursors, two populations that retain multilymphoid potential (Fig. 5 C). This MLP signature, which is strongly activated in quiescent GFP^{high} cells, was also observed in human B-ALL chemoresistant cells (Turati et al., 2021).

These observations revealed at the transcriptional level the plasticity and reprogramming of quiescent pre-LSCs that can partially reactivate a non-B cell molecular program.

TRN reveals Egr1 to control pre-LSC quiescence

According to the global upregulation of the cell-cycle program in GFP^{low} cells (Fig. 4 I and Fig. S3 E), we identified a cluster of genes named “cell-cycle^{like} cluster,” whose expression pattern followed the normal proliferation rate of each subset in the differentiation (Fig. 5 C). Conversely, a “quiescence^{like} cluster” corresponding to the mirror image of the “cell-cycle^{like} cluster” was detected within GFP^{high} cells (Fig. 5 C).

To predict the “core” TFs defining pre-LSC identity, and in particular pre-LSC quiescence, we built a TRN from ATAC-seq (Fig. 5 B) and RNA-seq (Fig. 5 C) data using LASSO-StARS algorithm (Miraldi et al., 2019; Pokrovskii et al., 2019). This algorithm integrates (i) motif analyses, (ii) differential accessibility motifs, and (iii) gene expression to construct a set of TF–gene interactions. This analysis predicted TFs that positively (activating) or negatively (repressive) regulate their target genes. Thus, we contrasted five specific TRNs (#1 to #5) to one shared TRN involved in the pre-LSC signature (Fig. 5 D). In particular, a quiescence module (TRN #1) was identified in GFP^{high} pre-LSCs composed of four TFs (Egr1, Mnt, Klf6, and Atf3) (Fig. 5 D). Among them, *Early growth response 1* (Egr1) regulates the higher number of target genes (121 activated and 242 repressed genes). Moreover, Egr1 was the top predicted TF to regulate positively the quiescence and negatively the cell cycle (TRN#1, Fig. 5 D). Accordingly, we observed a strong downregulation of the quiescence^{like} cluster and upregulation of the cell-cycle^{like} cluster in human HSPCs in which EGR1 has been silenced by RNA interference (Stoddart et al., 2022) (Fig. S4 E). Consistent with the critical role of Egr1 in controlling normal HSC dormancy and functions (Cabezas-Wallscheid et al., 2017; Min et al., 2008; Scheicher et al., 2015), our result suggests that it may be also relevant in pre-LSCs.

To address this question, we used a loss-of-function approach by CRISPR/Cas9 strategy (Fig. S4 F). We generated *PE*^{tg}*Cas9*^{tg} mice that constitutively express the Cas9 endonuclease (Tzelepis et al., 2016) and designed two single-guide RNAs (sgRNAs) (sgEgr1 T5 and sgEgr1 T7) that target the genomic region encoding the exon 1 of *Egr1* gene (Table S1). Aberrant B cells from preleukemic *PE*^{tg}*Cas9*^{tg} mice were purified and transduced with lentiviral vectors encoding sgEgr1 T5 or sgEgr1 T7 or a non-targeted control sgRNA (sgCTL). Transduced *PE*^{tg}*Cas9*^{tg} GFP⁺ cells were then purified (Fig. S4 G) and genome editing efficiency was determined by analyzing the rate of insertions/deletions (InDels) induced in the targeted region. Although sgEgr1 T5 did not induce InDels in the *Egr1* locus, sgEgr1 T7 allowed for almost a complete (93%) genome editing of the *Egr1* locus (Fig. S4 H). Subsequent in vitro cell division assay indicated that Egr1 genomic editing by sgEgr1 T7 slightly but significantly increases the global proliferation rate of aberrant B cells (Fig. S4, I and J). Interestingly, this observation was associated with a reduction of the proportion of slow-cycling (D0–2) cells (Fig. 5 E), with a decrease of Kit expression and an increase of IgM expression at the surface of aberrant B cells (Fig. 5 F). Thus, our results show that Egr1 genomic editing reverses, at least partially, the quiescence and the aberrant expression of Kit of preleukemic cells.

Finally, we built the molecular network of genes involved in TRN#1 from Fig. 5 D (Fig. S4 K). Strikingly, several members of the network such as the TFs Atf3 (Liu et al., 2020) and Klf6 (Adelman et al., 2019), the GTPase Gimap5 (Chen et al., 2011), and the transcriptional coactivator Taz (Althoff et al., 2020) have been also documented to control HSC functions.

Collectively, the data arising from our integrative molecular approach predict a list of TFs and their target genes to control molecular reprogramming in pre-LSCs and highlights Egr1 as a strong candidate to regulate pre-LSC quiescence.

EGR1 is activated in quiescent and therapy-resistant leukemic blasts from human B-ALL

Finally, we addressed the relevance of the expression modifications of murine pre-LSCs in human-resistant and quiescent leukemic cells. Interestingly, the downregulation of E2f/Myc targets (Fig. S3 E) and cell cycle/oxidative phosphorylation program (Fig. 4 I), combined with the upregulation of TNFα signaling (Fig. S3 E and Fig. 4 I) and HSPC/MLP signatures (Fig. 5 C) that we identified in GFP^{high} pre-LSCs from the *PAX5::ELN* model, have been also observed in treatment-resistant (minimal residual disease [MRD]) cells and in CFSE label-retaining cells (LRC) from B-ALL patient-derived xenografts (PDXs) in two independent studies (Ebinger et al., 2016; Turati et al., 2021). These molecular similarities (Fig. 6, A and B; and Fig. S5 A) suggest that PAX5::ELN pre-LSC signature mimics in some aspect the resistance/quiescence of human B-ALL cells.

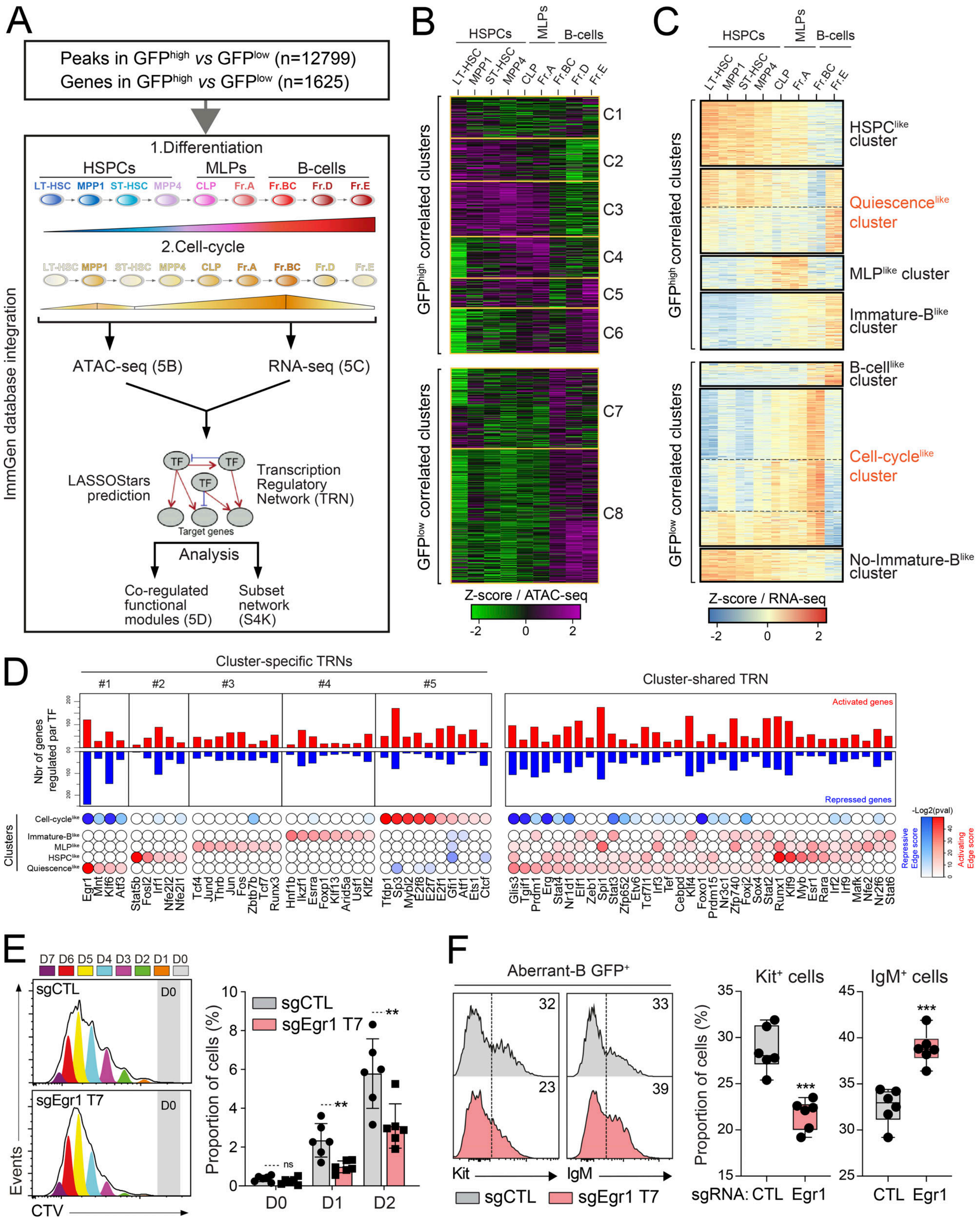

Figure 5. **Transcriptional molecular network reprogrammed in quiescent pre-LSCs. (A)** Study design for the generation of the pre-LSC transcription regulatory network. **(B and C)** Heatmap of k-means clustering for the differentially accessible region (B) or differentially expressed genes (C) integrating the different subsets of the normal B-lymphopoiesis. Clusters are separated depending on the correlation with GFP^{high} or GFP^{low}. **(D)** Core TRNs are defined according to enriched positive regulation of upregulated gene (red) or negative regulation of downregulated genes (blue) in a given cluster identified in Fig. 5 C (lower panels). The significance of enrichment is displayed on the right. The upper panels displayed the number of positive (red) and negative (blue) regulatory interactions per TF. The analysis defined five cluster-specific TRNs and one cluster-shared TRN of the pre-LSC signature. **(E and F)** Cell division assay (E) and

cell differentiation assay (F) of $PE^{tg}Cas9^{tg}$ aberrant B cells expressing sgCTL or sgEgr1 T7 (GFP^+) according to the experimental procedure detailed in Fig. S4 F. The number of cell divisions (D0 to D7) after 6 days of coculture was then analyzed (E, left panel) and the proportion of cells in the divisions D0, D1, and D2 was quantified (E, right panel, $n = 6$, one experiment, mean ± SD, $^{**}P < 0.01$). The proportion of Kit^+ and IgM^+ cells was evaluated by FACS (F, left panel) and quantified (F, right panel, $n = 6$, one experiment, $^{***}P < 0.001$). The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values.

Thus, we explored *EGR1* expression in resistant and/or quiescent human B-ALL cells. We transplanted leukemic blasts from two "de novo" *ETV6-RUNX1* B-ALL patients (HB#010 and HB#007) into immunodeficient NOD.*Cg-Prkdc*^{scid}*Il2rg*^{tmWjl} (referred to as NSG) mice and monitored for engraftment before treatment with a chemotherapeutic cocktail (dexamethasone [DEXA] + VCR) or vehicle control (Fig. 6, C and D; and Fig. S5, B and C). Strikingly, B-ALL progression was prevented in chemotherapy-treated mice characterized by a significant reduction of the leukemic burden in the BM and the spleen (Fig. 6, E and F; and Fig. S5, D and E). This drastic reduction of human B-ALL cells in vivo was associated with an enrichment of quiescent $Ki67^-$ leukemic blasts (Fig. 6 G and Fig. S5 F). This resistance was also associated with a significant upregulation of *EGR1* transcript in purified residual leukemic blasts (Fig. 6 H and Fig. S5 G). Interestingly, we observed that residual leukemic blasts developed leukemia in secondary recipients with a significant delay as compared with untreated cells (Fig. S5 H). This result reinforces the notion that quiescent/resistant cells from PDX models exhibit a long-term leukemia-initiating activity and mimics, in some aspects, relapse-initiating cells from B-ALL patients.

Based on our previous observations, we hypothesized that *EGR1* expression was already upregulated in quiescent cells within the tumor burden of B-ALL. Thus, leukemic blasts from two other *ETV6::RUNX1* B-ALL patients (HB#002 and HB#008) were labeled with CTV and cocultured on MS5 stromal cells (Fig. 6 I and Fig. S5 I). Strikingly, our data showed the feasibility of identifying undividing/slow-dividing cells from "de novo" B-ALL samples using this in vitro cell division approach (Fig. 6 J and Fig. S5 J). Of note, we observed that these cells were enriched in the $CD45^{neg/low}$ phenotype (Fig. 6 J and Fig. S5 J), which has been associated with drug-tolerant leukemic clones leading to B-ALL relapse in patients (Dobson et al., 2020). Therefore, we confirmed that *EGR1* expression was higher in purified undivided (CTV^{high}) B-ALL cells from the two patients as compared with cycling (CTV^{low}) leukemic cells (Fig. 6 K and Fig. S5 K). Together, our observations suggest that EGR1 overexpression is a "de novo" molecular characteristic of quiescent/resistant cells within the tumor bulk.

Finally, we took advantage of the RNA-seq and clinical data of a large cohort of B-ALL patients (Gu et al., 2019) to explore *EGR1* expression and its clinical outcome. The analysis of the event-free survival (EFS) curves showed that the clinical outcome was drastically affected for all B-ALL patients with high levels of *EGR1* (Fig. S5 L). This is partially explained by the observation that *EGR1* expression was higher in high-risk genetic subgroups (i.e., Ph-like and *KMT2A*-rearranged subgroups) than in B-ALL subgroups associated with a better prognosis (i.e., *ETV6::RUNX1*, *TCF3::PBX1*, hyperdiploid, and *PAX5*alt) (Inaba and Pui, 2021) (Fig. S5 M). Furthermore, the poor clinical outcome was also strikingly associated with EGR1 expression within the different B-ALL oncogenic subgroups (Fig. 6 L). Collectively, our observations strongly suggest that cell quiescence and *EGR1* activation are common and oncogene-independent features of treatment-resistant B-ALL cells.

Discussion

In the present study, we found that a single primary oncogenic event of B-ALL is sufficient to confer stem cell–like features to a subset of committed B cell progenitors, converting them to pre-LSCs. Combined with an unsupervised clustering representation of FACS data, our multiparametric immunophenotypic approach allowed for a precise visualization of the normal progression of B cells throughout differentiation. In addition, this strategy revealed concomitantly the presence of an aberrant B cell population in preleukemic PE^{tg} mice, characterized by the abnormal co-occurrence of phenotypic markers. To date, this exhaustive coverage of the B cell differentiation has not yet been described in leukemic mouse models, mainly limited by the number of simultaneously assayed markers. However, the identification of aberrant antigen coexpression is commonly used in clinics to define the MRD in B-ALL (Cherian and Soma, 2021; Hassanein et al., 2009) based on the principle that leukemic cells express phenotypic features that can be used to distinguish them from normal hematopoietic cells. Moreover, the development of single-cell mass cytometry (CyTOF) approach, quantifying simultaneously a large number of surface markers and intracellular phosphorylated proteins, allowed for the identification of several features involved in normal and pathological B cell development in humans with unprecedented resolution (Bendall et al., 2014; Good et al., 2018). Thus, our approach reinforces the notion that exploring abnormal coexpression using a large number of phenotypic markers is a valuable approach to predict the aberrant differentiation state occurring in cells.

Combined with RNA-seq, the integration of intracellular phosphorylated proteins established that the ectopic activation of IL7r/JAK-STAT pathway, which is normally shut down by the pre-BCR/BLNK axis, favors the differentiation blockade of preleukemic cells. Thus, the imbalance between the IL7r/JAK-STAT pathway and pre-BCR/BLNK axis appears as a critical feature of B-ALL initiation. Interestingly, this molecular imbalance, including unresponsive pre-BCR signaling, has been shown to predict patient relapse in B-ALL (Good et al., 2018). Moreover, secondary alterations in human B-ALL frequently mimic cytokine-receptor signaling through the constitutive phosphorylation of STAT5 (*IL7R*, *CRLF2*, and *JAK* mutations) or mimic the pre-BCR downstream pathway through the constitutive activation of the RAS/MAPK signaling (*NRAS*, *KRAS*, and *PTPN11*

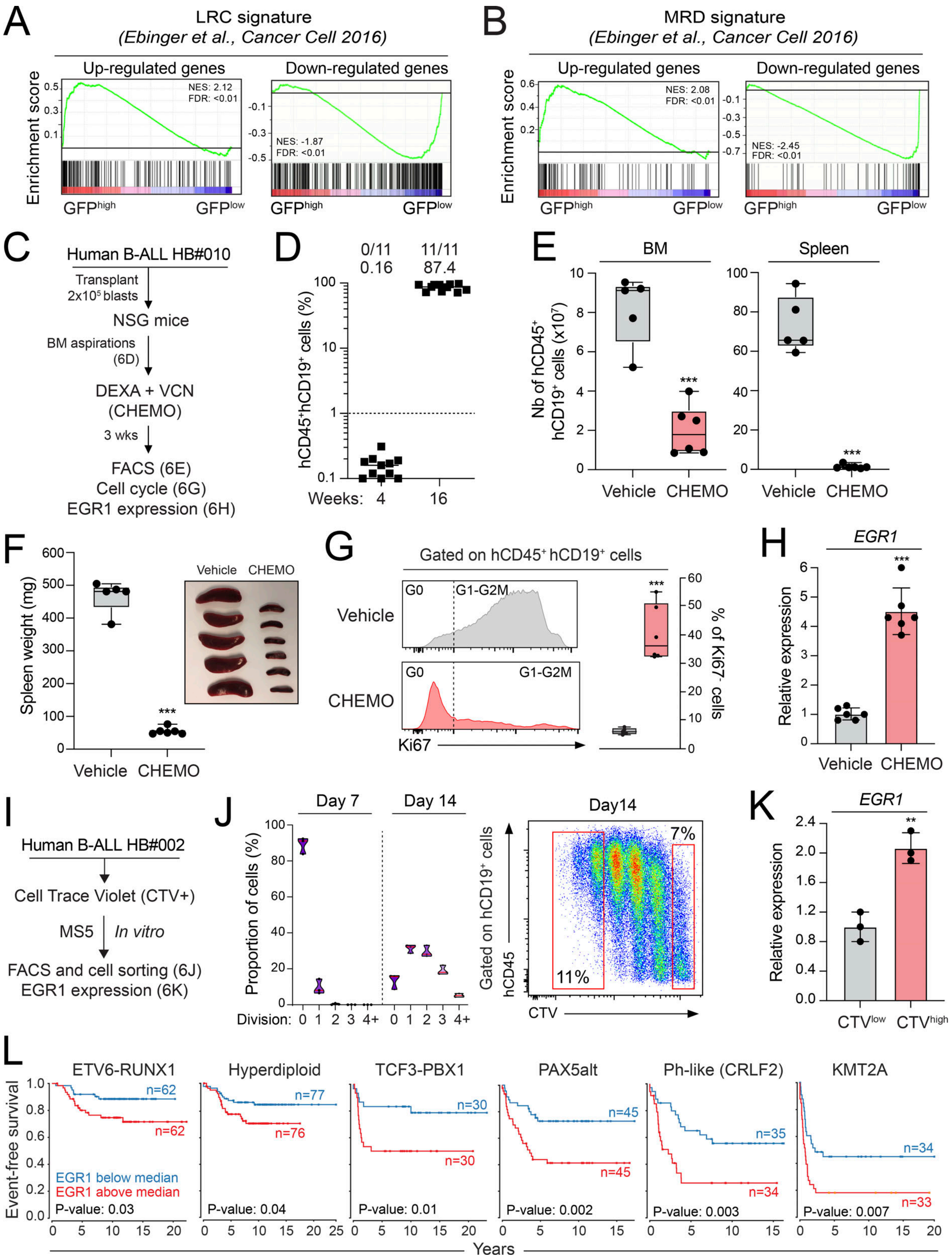

Figure 6. **EGR1 activation in quiescent and therapy-resistant human B-ALL. (A and B)** GSEA of the up- (left panels) and down- (right panels) regulated genes from CFSE label-retaining human B-ALL cells (LRC signature, A) or from residual human B-ALL cells after chemotherapies (MRD signature, B) (Ebinger et al., 2016) between GFP^{high} and GFP^{low} cells. NES and FDR values are indicated. **(C–H)** Experimental procedure to explore EGR1 expression in resistant human B-ALL cells (C). Leukemic blasts from the "de novo" B-ALL patient HB#010 were transplanted into NSG mice ($n = 11$). The proportion of hCD45$^+$hCD19$^+$ B-ALL blasts in BM aspirations was assessed 4 and 16 wk after transplantation (D). Engrafted mice were randomly selected for treatment (CHEMO) or not (vehicle) with a chemotherapeutic cocktail (DEXA+VCR) during 3 wk. Human B-ALL reconstitution (number of hCD45$^+$hCD19$^+$ B-ALL blasts) was monitored by FACS in the BM and the spleen after the treatment (E). Spleen weight was measured, and a picture of the spleen is shown (F). Ki67 expression was determined by FACS

in blasts from the BM of treated and untreated mice (G). Data obtained with B-ALL patient HB#010 were compiled from five (vehicle) and six (CHEMO) mice per condition (***P < 0.001). The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. Engrafted blasts were purified and mRNA levels of the human *EGR1* were determined by quantitative RT-PCR and normalized to *ABL1* gene. Error bars indicate mean ± SD (*n* = 6, ***P < 0.001) from two independent experiments (H). **(I–K)** Experimental procedure to explore EGR1 expression in quiescent human B-ALL cells (I). Leukemic blasts from the "de novo" B-ALL patient HB#002 were stained with CTV and the number of cell divisions (D0 to D4–5) after the coculture was then analyzed (J). The proportion of cells in each division was quantified (J, left panel). Red dotted gates were used to purify the D0 (CTV^{high}) and D4–5 (CTV^{low}) populations (J, right panel) and the expression of human *EGR1* were determined by quantitative RT-PCR (K, mean ± SD, *n* = 3, one experiment, **P < 0.01). **(L)** *EGR1* expression and its associated clinical outcome in B-ALL patient extracted from Gu et al. (2019). EFS curves of B-ALL patients with *EGR1* expression below (blue) and above (red) to the median. Patients are classified within the different B-ALL oncogenic subgroups.

---

mutations) (Jamrog et al., 2018; Mullighan et al., 2007). It has been recently demonstrated that mutations in these two pathways are mutually exclusive in human B-ALL (Chan et al., 2020). In addition, the concurrent reactivation of one pathway affects the activity of the other, reciprocally, leading to the subversion of B-ALL transformation (Chan et al., 2020). B-ALL development in the PAX5::ELN model is also associated with the acquisition of secondary mutations in genes of either JAK/STAT (*Jak3* mutations) or RAS/MAPK (*Kras* and *Ptpn11* mutations) signaling pathways (Jamrog et al., 2018). These observations, therefore, suggest that secondary mutations either reinforce the aberrant activation of the JAK/STAT pathway already primed in preleukemic cells or allow for a bypass of pre-BCR signaling deficiency to trigger B-ALL transformation.

Various studies suggest that the restricted cell cycle is an important mechanism of therapeutic resistance in both human and murine leukemia models (Ebinger et al., 2016; Gerby et al., 2016; Prost et al., 2015; Tremblay et al., 2018; Trumpp et al., 2010; Turati et al., 2021). In parallel with an in vitro cell division assay, we used the doxycycline-inducible *H2B-GFP*^{tg} mouse model to demonstrate in vivo the existence of a slow-cycling and drug-resistant population of pre-LSCs in the PAX5::ELN-induced B-ALL model. In particular, our work corroborates that self-renewal and clonal evolution could be restricted to a rare and slow-cycling population of preleukemic cells, as previously observed in the *Lmo2*-induced T-ALL model (Tremblay et al., 2018). This not only reinforces the importance of cell-cycle restriction in pre-LSC activities in two distinct lineages but also makes the *H2B-GFP* system a powerful in vivo tool to develop strategies targeting the quiescence and the resistance of relapse-inducing clones in different models.

Additionally, the present study brings new insights regarding the molecular mechanisms of the cell-of-origin of B-ALL demonstrating that quiescent pre-LSCs are characterized by a partial loss of B cell lineage identity and activate an HSPC-like molecular program. This finding is in agreement with pioneer works in acute myeloid leukemia and T-ALL demonstrating that a primary oncogene can activate an aberrant stem cell–like program in committed myeloid progenitors (Krivtsov et al., 2006) or thymocytes (McCormack et al., 2010). While the presence of PAX5::ELN induces the emergence of quiescent cells in committed B cells, its expression level in quiescent and cycling cells is quite similar. Notwithstanding the fact that functional and molecular stem cell–like features were focused in quiescent cells, this observation raises at least two questions: (i) whether cell-cycle restriction is a common aberrant characteristic of pre-LSCs or is a critical property to prime leukemia initiation, and (ii) whether stem cell properties of pre-LSCs such as self-renewal, drug resistance, and leukemia-initiating activity are reversible processes, as demonstrated on human B-ALL blasts (Ebinger et al., 2016).

The question of whether a primary genetic alteration induces sufficient molecular and functional plasticity to cause lineage subversion remains poorly explored. A previous investigation on the cell-of-origin in mixed phenotype acute leukemia supports the notion that a founding alteration, rather than the secondary events, induces lineage plasticity in preleukemic clones (Alexander et al., 2018). Another study using lineage-specific oncogene activation also shows that *ETV6::RUNX1* fusion protein can induce both B-ALL and T-ALL, and that leukemia transformation is dependent on the nature of the secondary mutations (Rodríguez-Hernández et al., 2021). While *ETV6::RUNX1* is associated with B-ALL in humans, this work suggests that preleukemic clones can exhibit a T cell potential when expressed in the appropriate lineage. Our results revealed that the TF activity of PAX5 and EBF1, considered master factors to maintain B cell identity by suppressing alternative lineage choices (Busslinger, 2004; Nutt et al., 1999; Ramamoorthy et al., 2020), is downregulated in quiescent pre-LSCs. The downregulation of PAX5 and EBF1 activity is associated with the upregulation of PU-1 and RUNX1 activity, and with the activation of a non-B cell molecular program. Integrating chromatin accessibility and gene expression data recently brought major insights and facilitated the identification of putative regulatory regions and their candidate binding TFs, thus allowing a more precise definition of the regulatory networks in immune cells (Miraldi et al., 2019; Pokrovskii et al., 2019). Here, we applied this computational method to build cluster-specific TRNs that define the molecular identity and plasticity of quiescent pre-LSCs in our model. In particular, this approach allowed us to identify Egr1 as a strong TF candidate regulating positively pre-LSC quiescence. In addition, our loss-of-function approach demonstrated that Egr1 inhibition reverses, at least partially, the cell quiescence of preleukemic PAX5-ELN cells. This observation is consistent with its critical role in controlling normal HSC dormancy, localization, and functions that have been described using both *Egr1*$^{-/-}$ mouse model (Min et al., 2008) and RNA interference in human CD34^{+} cord blood–derived HSPCs (Stoddart et al., 2022). Finally, molecular similarities with resistant clones in B-ALL patients after chemotherapy (Ebinger et al., 2016; Turati et al., 2021), including the upregulation of *EGR1*, strongly make our PAX5::ELN quiescent pre-LSCs a relevant drug-resistant cellular platform to develop new targeted therapies.

Materials and methods

Mice

PAX5::ELN (PE^{tg}) transgenic mouse model was developed by our team as previously described (Jamrog et al., 2018). Briefly, cDNA encoding the human PAX5::ELN fusion protein was inserted at the *IgH* locus under the control of a VH promoter (PVH) and the endogenous Eμ enhancer to trigger PAX5::ELN expression in the early phase of B cell development. The previously developed inducible *TetOP-H2B-GFP* ($H2B\text{-}GFP^{tg}$) mouse model (Foudi et al., 2009) was purchased from Jackson Laboratory. *Rosa26-Cas9* ($Cas9^{tg}$) mice (Tzelepis et al., 2016) were generously provided by the Wellcome Sanger Institute of Cambridge (UK). Homozygous PE^{tg}, $H2B\text{-}GFP^{tg}$, and $Cas9^{tg}$ mice were maintained by crossbreeding, their genotypes were verified by PCR, and all the experiments were performed in heterozygous mice. C57BL6 (Ly6.2, CD45.2) mice were purchased from Charles River Laboratories and Pep3b (Ly6.1, CD45.1) B6.SJL congenic mice were initially obtained from The Jackson Laboratory. All animals were housed in pathogen-free conditions (Anexplo US006 CREFRE) in accordance with the European Directive 2010/63/EU and the French Institutional Guidelines for animal handling. Mice were handled according to protocols approved by the Regional Ethical Committee (agreement #A31555010). Mice were backcrossed into the C57BL6 background for >10 generations.

FACS analysis and cell sorting

Single-cell suspensions were prepared from BM of wt and PE^{tg} mice of the indicated ages in Iscove's modified Dulbecco's medium (IMDM; Gibco) supplemented with 2% fetal bovine serum (FBS) (Stemcell Technologies). Immunostainings were performed using antibodies for flow cytometry obtained from Pharmingen (BD Biosciences), Invitrogen, and Miltenyi are listed in Table S1. FACS analysis was performed on a Fortessa cytometer (BD Biosciences) using FlowJo (BD Biosciences) software. For surface staining, cells were incubated with the antibodies for 20 min in IMDM 2% FBS at 4°C and washed twice with PBS1X before analysis. For cell-cycle analysis, cells were fixed and permeabilized (CytofixCytoperm Plus; BD Bioscience) for 30 min before staining with the anti-Ki67 antibody, in accordance with the manufacturer's instructions. For cell sorting, BM cells from wt or preleukemic PE^{tg} mice were flushed, and B cell fraction were enriched by B220 magnetics microbead sorting (Miltenyi Biotech). B cell fraction was subsequently incubated with anti-CD19, -CD23, and -Kit antibodies listed in Table S1. Cell sorting of Kit^+ and Kit^- ($CD19^+CD23^-$) B cells was performed on a MoFlo Astrios sorter (Beckman Coulter).

Uniform manifold approximation and projection (UMAP) analysis

UMAP analysis was performed using FlowJo software. At least four samples per condition were pregated on single cells of the population of interest. Then, informatic cleaning and normalization were performed using FlowAI package (V2.1) (Monaco et al., 2016) and CytoNorm package (V1.0) (Van Gassen et al., 2020) for each sample and the same number of cells were then concatenated. Arc sin transformation was performed manually for each marker to discriminate the different B cell populations. The dimension reduction algorithm UMAP (umap-learn Python package v2.4.0) (Becht et al., 2018) was run using Euclidian distance with 15 nearest neighbors and 0.5 distance parameters. Thus, the UMAP analysis was performed using a multiparametric staining by FACS integrating the 12 markers shown in Fig. S1 A covering all the steps of the B cell differentiation. UMAP representations of the cell density and the clustering analysis of the different B cell subpopulations are based on the gating strategy shown in Fig. S1, B and C. In the UMAP analyses, each subset is represented by one color. Black arrow indicates the physiological phenotypic progression of B cells in the differentiation, and red dotted line delimits the aberrant preleukemic B cell population induced by PAX5::ELN.

In vitro coculture assays

B cell populations were cocultured on MS5 stromal cells in a B cell medium composed of IMDM 10% FBS, 0.05 mM β-mercaptoethanol (Sigma-Aldrich), 2 mM *L*-glutamine (Invitrogen), penicillin (100 U/ml), and streptomycin (100 U/ml), and supplemented with 5 ng/ml of murine Il7, 10 ng/ml of murine Flt3L, and 10 ng/ml of murine stem cell factor (SCF) (Peprotech).

In vivo transplantation assays

Donor and recipient mice were on C57BL6 (Ly6.2, CD45.2) and Pep3b (Ly6.1, CD45.1) backgrounds, respectively, allowing for the discrimination of host- and donor-derived cells on the basis of CD45 alleles. For the serial transplantation assay, total B cells ($B220^+$) from the BM of wt or preleukemic PE^{tg} mice (30 days old; $CD45.2^+$) were purified and transplanted intravenously in equal numbers (4.5×10^6 cells per mouse) into primary (I) recipient mice (6–8 wk-old; $CD45.1^+$). Then, one fifth of the total BM cells from primary mice were transplanted in equal numbers (~20×10^6 cells per mouse) into secondary (II) recipients. To measure the short-term reconstitution potential of preleukemic B cells, the chimerism was analyzed by flow cytometry (FACS) 2 wk after each transplantation and was illustrated by the percentage of donor-derived cells (% $CD45.2^+$) found in the BM of recipient mice. Finally, one fifth of the total BM cells from secondary mice were transplanted in equal numbers (~20×10^6 cells per mouse) into tertiary (III) recipients to assess the long-term B-ALL development, represented by Kaplan–Meier survival curves. All the recipient mice ($CD45.1^+$) were pretreated, 24 h before transplantation, with 30 mg/kg of Busulfan (Busilvex; Pierre Fabre) (related to Fig. 1, D–G).

Kit^+ and Kit^- ($CD19^+CD23^-$) B cells from the BM of preleukemic PE^{tg} mice (CD45.2) were purified and intravenously transplanted in equal numbers (5×10^4 per mouse) in primary (I) recipient CD45.1 mice. Kit^+ ($CD19^+CD23^-$) B cells from the BM of wt mice were purified and transplanted (5×10^4 per mouse) as control. BM cells from the primary transplantation with preleukemic PE^{tg} Kit^+ et Kit^- cells were transplanted in secondary (II) recipients (related to Fig. S2 C and Fig. 1, H and I).

Purified D0–2 and D8–9 cells (CD45.2$^+$) from CTV-labeled preleukemic *PE*^{tg} cells were transplanted in equal numbers (2 × 10^3 per mouse) in recipient CD45.1 mice (related to Fig. 2, D, I, and J).

Purified GFP^{high} and GFP^{low} cells from preleukemic *PE*^{tg}*H2B-GFP*^{tg} mice (CD45.2$^+$) were transplanted in equal numbers (1.5 × 10^4 per mouse) in recipient CD45.1 mice. Engraftment was monitored by FACS and was illustrated by the percentage of donor-derived cells (% CD45.2$^+$) found in recipient mice after BM punctions. For all transplantation experiments, the number of positive mice and the median of engraftment are indicated. Statistical significance was calculated using an unpaired *t* test. For Kaplan–Meyer survival curves, the median of time required to develop the B-ALL is indicated on the curve (related to Fig. 4 E and Fig. S3 D).

In vitro cell division assay

Cell division of Kit$^+$ (CD19$^+$CD23$^-$) wt and *PE*^{tg} B cells were analyzed by using CTV dilution. Briefly, total B cells (10^6 cells/ml) from wt and preleukemic *PE*^{tg} mice (30 days old) were labeled with 5 µM CTV (C34557; Thermo Fisher Scientific) in PBS1X during 20 min at 37°C. CTV-labeled B cells were then washed twice in PBS1X 1% FBS according to the manufacturer's recommendations and plated on MS5 stromal cells for 24 h to stabilize the staining. Then, Kit$^+$ (CD19$^+$CD23$^-$) CTV-labeled B cells were purified by cell sorting and cocultured for 3 days on an MS5 cell line in a supplemented B cell medium. Since dividing cells equally distribute the dye to daughter cells, the number of cell divisions can be followed by a decrease in the fluorescence intensity. Cell division was analyzed using Winlist (Verity Software House), and proliferation index (PI) was calculated as described (Roederer, 2011). Slow- (D0–2) and high- (D8–9) cycling *PE*^{tg} cells were purified by cell sorting. High- (D8–9) cycling wt cells were also purified as a control to perform the subsequent functional and molecular experiments as indicated in the experimental procedure detailed in Fig. 2 D.

In vivo cell division assay

Wt or homozygous *PE*^{tg} mice were bred with homozygous *H2B-GFP*^{tg} mice to obtain heterozygous *H2B-GFP*^{tg} and *PE*^{tg}*H2B-GFP*^{tg} mice, respectively. To induce *H2B-GFP* transgene expression, 20-day-old mice were intraperitoneally treated with 40 mg/kg of doxycyxline (DOX, D9891; Sigma-Aldrich) and 2 mg/ml of DOX, supplemented with sucrose (S9378; Sigma-Aldrich) at 10 mg/ml, was added to drinking water for 5 wk (DOX pulse). GFP expression in all B cell subpopulations was then examined following withdrawal of doxycycline (DOX chase) for 0, 1, 2 and 3 wk. Quantification of GFP was determined for each time point by ranking 10^4 wt or *PE*^{tg} cells from the lowest to the highest value of their GFP fluorescence intensity and the area under the curve (AUC) was then calculated. White and black dotted lines indicate the standard error of the mean (SEM) of wt and *PE*^{tg} AUC, respectively. The proportion of persisting GFP-retained (GFP^{high}) cells is indicated for each time point (related to Fig. 4, A–C). Purification of GFP^{high}, GFP^{med}, and GFP^{low} B cells was performed after 1 wk of DOX chase using a MoFlo Astrios cell sorter (Beckman Coulter).

Phosphoflow cytometry

Phosphoflow cytometry experiments are related to Fig. 3, B and E; and Fig. S3 A to explore IL7r and pre-BCR signaling pathways in preleukemic B cells. As shown in Fig. 3 A, the IL7r signal induces Stat5 phosphorylation that activates the transcription of *Cdkn2a*, *IL2rα*, *Tnfrsf13b*, and *Socs3* to promote cell survival and proliferation. In parallel, IL7r activates the PI3K signaling pathway that phosphorylates and represses Foxo1, the inducer of *Rag1* and *Rag2* gene transcription, therefore inhibiting *Igκ* gene recombination. In contrast, pre-BCR signal induces Syk/Blnk/Plcγ2 phosphorylation cascade that converges to the transcription of *Ikzf3* and *Irf8* that inhibit pre-B cell proliferation and initiate *Igκ* gene recombination, respectively. Together, each receptor has antagonistic and balanced functions to coordinate the proliferation and differentiation switch in pre-B cells.

1.5 × 10^6 cells from each wt and preleukemic *PE*^{tg} BM were mixed, and surface markers were stained in RPMI (Gibco) supplemented with 2% FBS for 10 min at 37°C and washed with RPMI (centrifugation, 5 min; 1,200 rpm).

IL7r-ligand stimulation was assessed by phosphoflow cytometry after the ex vivo activation of preleukemic *PE*^{tg} BM cells with IL-7 (+IL-7). Unstimulated (−IL-7) cells were used as control. Normal B cells and aberrant B *PE*^{tg} cells were distinguished according to the level of B220 expression, and pStat5 and pFoxo1 expression were detected by FACS. IL7r-ligand stimulation was performed by incubating the cells with RPMI 2% FBS supplemented or not with 30 ng/ml IL-7 (Peprotech) for 15 min at 37°C for the detection of pStat5 and for 16 h at 37°C for the detection of pFoxo1. For the staining of pStat5, cells were then fixed with prewarmed Lyse/Fix buffer 1× for 12 min at 37°C, washed twice (centrifugation, 7 min; 500 rpm) in RPMI 2% FBS, and permeabilized with cold Perm Buffer III for 30 min at 4°C, according to the manufacturer's instructions (BD Phosphoflow). Cells were washed twice in PermWash buffer and the intracellular immunostainings with the anti-pStat5 antibody combined with the anti-Ki67 antibody were performed for 30 min at room temperature, followed by two washes before FACS analysis. For the staining of pFoxo1, cells were fixed and permeabilized with cold CytofixCytoperm Plus buffer (BD Bioscience) for 24 h at 4°C and washed twice (centrifugation, 7 min; 500 rpm) in PermWash buffer (BD Bioscience). The intracellular immunostainings with the anti-pFoxo1 antibody combined with the anti-Ki67 antibody were performed in PermWash buffer for 1 h at room temperature, followed by two washes (centrifugation, 7 min; 500 rpm) before FACS analysis.

(Pre-)BCR stimulation was assessed by phosphoflow cytometry after the ex vivo activation of preleukemic *PE*^{tg} BM cells with H_2O_2 ($H_2O_2^+$). Unstimulated ($H_2O_2^-$) cells were used as control. Normal B cells and aberrant B *PE*^{tg} cells were distinguished according to the level of B220 expression, and pBlnk and pPlcγ2 expression was detected by FACS. (Pre-)BCR stimulation was performed by incubating the cells with RPMI 2% FBS supplemented or not with 1 mM H_2O_2 for 5 min at 37°C. Cells were then fixed and permeabilized with cold CytofixCytoperm Plus buffer (BD Bioscience) for 45 min at 4°C and washed twice (centrifugation, 7 min; 500 rpm) in PermWash buffer (BD Bioscience). The intracellular immunostainings with the anti-pBlnk or the anti-pPlcγ2 antibodies combined with the anti-Ki67 antibody were performed in PermWash buffer for 30 min at room temperature, followed by two washes (centrifugation, 7 min; 500 rpm) before FACS analysis. Immunostaining combinations used for the phospflow cytometry are listed in Table S1.

Lentiviral production and transduction for gene editing by CRISPR/Cas9

sgRNAs targeting the exon 1 of *Egr1* (Table S1) were designed using CCTop-CRISPR/Cas9 target online predictor software (Center for Organismal Studies Heidelberg) and subcloned in pLKO5-sgRNA-EFS-GFP, in which the sgRNA and GFP coding sequences are controlled by the U6 and the EF1α promoters, respectively (Heckl et al., 2014). Concentrated VSV/G pseudotyped lentiviral vectors pLKO5-sgRNA-EFS-GFP encoding sgRNA *Egr1* (sgEgr1 T5 and T7) or sgCTL were produced by the vectorology facility Vect'UB (TBMCore—CNRS UAR3427, INSERM US005, Universtiy of Bordeaux, Bordeaux, France). Titrations of produced lentiviral vectors were performed on the HEK293T cell line. Homozygous *PE*^{tg} mice were bred with homozygous *Cas9*^{tg} mice to obtain heterozygous *PE*^{tg}*Cas9*^{tg} mice. Aberrant B cells from the BM of preleukemic *PE*^{tg}*Cas9*^{tg} mice were purified by cell sorting, plated in suspension culture in B cell medium and transduced with lentiviral vectors expressing sgCTL or sgEgr1 T5 and sgEgr1 T7 at a multiplicity of infection of 30 in the presence of 4 µg/ml of polybrene (Sigma-Aldrich) for 24 h. Transduced cells were then washed in PBS1X and cocultured on MS5 stromal cells in a fresh supplemented B cell medium for 3 days. Transduced GFP^{+} cells were then purified by cell sorting to perform in vitro cell division assay and cell differentiation assay in coculture on MS5 stromal cells in a B cell medium supplemented with a low concentration of Il-7 (0.5 ng/ml) and with Flt3L (10 ng/ml) and SCF (10 ng/ml) (related to Fig. 5, E and F; and Fig. S4, F–J).

Next-generation sequencing (NGS) of targeted gene regions

The targeted region (exon 1) of *Egr1* was performed by NGS, and the genomic editing efficiency was analyzed by the proportion (%) of insertions and deletions (InDel) induced by the sgRNA Egr1 T5 and T7. Genomic DNA was extracted and purified using QIAamp DNA mini kit (Qiagen). Sequencing was performed using a MiSeq sequencer (Illumina) using a Miseq Reagent kit V2 (paired-end sequencing 2 × 150 cycles). Genome editing efficiency of the targeted *Egr1* region (exon 1) was analyzed using the online CRISPResso software (https://crispresso.pinellolab.partners.org/submission; Canver et al., 2018). Targets and primers are listed in Table S1.

Human B-ALL samples

Human B-ALL samples were obtained from the Toulouse University Hospital (Toulouse, France) after signed written informed consent for research use in accordance with the Declaration of Helsinki and stored at the Hémopathies malignes de l'Inserm en Midi-Pyrénées (HIMIP) collection (BB-0033-00060). According to the French law, HIMIP biobank collection has been declared to the Ministry of Higher Education and Research (DC 2008-307, collection 1) and obtained a transfer agreement for research applications (AC 2008-129) after approbation by our institutional review board and ethics committee (Comité de Protection des Personnes Sud-Oust et Outremer II). Biological annotations of the samples have been declared to the Comité National Informatique et Libertés (i.e., Data Processing and Liberties National Committee). For in vitro cell division assay using CTV dilution, human B-ALL blasts (HB#002 and HB#008) were stained as previously described, and CTV-labeled cells were grown in cocultured on MS5 cell line in supplemented B cell medium and analyzed at the indicated time points in the Fig. 6 J and Fig. S5 J.

PDX model and in vivo treatment with chemotherapy

NSG mice were produced at the Génotoul-Anexplo platform in Toulouse, France, using breeders obtained from Charles River Laboratories. Mice were housed in sterile conditions using high-efficiency particulate arrestance-filtered microisolators and fed with irradiated food and sterile water. Mice (6–9 wk old) were sublethally treated with busulfan 30 mg/kg 24 h before injection of leukemic cells. Leukemic blasts from B-ALL patients were thawed at room temperature, washed twice in PBS, and transplanted by intravenous injection (2×10^5 cells per mouse). At the indicated time points in Fig. 6 D and Fig. S5 C, leukemia engraftment was monitored by FACS with hCD45 (557748; BD Bioscience) and hCD19 (562294; BD Bioscience) staining after BM punctions. For combination therapy, a chemotherapeutic cocktail composed of DEXA (10 mg/kg) and VCR (0.5 mg/kg) was administered weekly for 3 wk (HB#010; Fig. 6 C) and for 1 wk (HB#007; Fig. S5 B) via intraperitoneal injection. A daily administration of 10 mg/kg DEXA was supplemented during the period of treatment. Mice were euthanized and analyzed after the treatment.

Quantitative RT-PCR

RNA was isolated using the Trizol method, and cDNA was synthesized using SuperScript VILO cDNA Synthesis Kit (Invitrogen) according to the manufacturer's instructions. Quantitative SYBR Green PCR was performed on a LightCycler480 II System (Roche) to quantify human *EGR1* cDNAs expression in Fig. 6, H and K; and Fig. S5, G and K using LightCycler480 SYBR Green I Master (Roche Diagnostics GmbH) according to the manufacturer's instructions. All PCR were carried out as follows in a 20 µl volume: 5 min at 95°C, followed by 45 cycles of 10 s at 95°C, 10 s at annealing temperature of 60°C and 10 s at 72°C. Quantification was performed using the Δ*Ct* method with normalization to the human *ABL1* gene expression levels. Data were analyzed using the LC480 software (Roche Diagnostics). The forward (Fw) and reverse (Rev) primers used were as follows: for EGR1 (Fw: 5′-CAGCCCTACGAGCACCTGAC-3′ and Rev: 5′-GTGGTTTGGCTGGGGTAACT-3′); for ABL1 (Fw: 5′-TGGAGATAACACTCTAAGCATAACTAAAGGT-3′ and Rev: 5′-GATGTAGTTGCTTGGGACCCA-3′).

Statistical analysis

The number of biological replicates is indicated in the relevant figure legends. Error bars for pooled replicates represent standard deviation (SD). Statistical differences were determined using a two-tailed unpaired Student's *t* test for comparison of quantitative variables, assuming normality and equal distribution of variance between the different groups analyzed. All statistical analyses were performed using GraphPad Prism software, version 7 (GraphPad). A P value of <0.05 was considered statistically significant (*P < 0.05, **P < 0.01, ***P < 0.001). Survival in mouse experiments was represented with Kaplan–Meier curves (GraphPad Prism).

RNA-seq

RNA from purified *PE*$^{\mathrm{tg}}$ D0–2 (*n* = 2) and D8–9 (*n* = 2) cells and from wt D8–9 (*n* = 2) cells was isolated with the RNeasy Plus Mini Kit (Qiagen). RNA-seq libraries were prepared using TruSeq Stranded mRNA Low Sample kit (Illumina) according to the manufacturer's protocol, starting with 300 ng mRNA. Cluster generation and sequencing were carried out by using the Illumina NEXTSEQ 550 system and NEXTSEQ 500/550 HIGH OUTPUT KIT v2.5 (150 cycles) with a read length of 2 × 75 nucleotides according to the manufacturer's guidelines. The expression of mRNA transcripts rearranged for the *IgH* and *IgL* loci was analyzed using Mixcr software and represented by Circos diagrams using VDJ tool software (related to Fig. 2 G).

ULI RNA-seq

GFP$^{\mathrm{high}}$ and GFP$^{\mathrm{low}}$ cells from the aberrant B population of preleukemic *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ mice after 1 wk of DOX chase were purified by cell sorting (Melody, BD) directly in RLT buffer, and the RNA was isolated with the RNeasy plus micro kit (Qiagen). Smart-seqv4 libraries were prepared as previously described (Picelli et al., 2014) using the Takara SMART-Seq v4 full-length transcriptome analysis kit according to the manufacture's guidelines. Paired-end sequencing was performed on an Illumina NextSeq 500 using 2 × 75 bp reads. Low-quality reads were trimmed, and adaptor sequence was removed using fastp (v). Short reads were then mapped to mm10 genome using RNA-STAR (v2.7.8a) with doublepass parameter. Low-quality mapping and duplicated reads were removed and the remaining reads was count using featurecount (v) with default paired-end parameter.

A specific region in the exon 7 of Pax5 (Chr4: 44,609,844–44,609,786), which is enriched in SNPs, was for the quantification of reads from murine *Pax5* (Pax5 WT) and from human *PAX5* (PAX5-ELN) (related to Fig. S4 B).

GSEA of top 500 differentially expressed genes (bulk RNA-seq EGAD00001006133) of residual human B-ALL cells after chemotherapies published in Turati et al. (2021) (corresponding to acutely treated versus untreated PDX from patient 2) was performed in preleukemic *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ (Fig. S5 A).

For the integration of the ImmGen database, the same preprocessing pipeline was used for all the ImmGen fastq (GSE109125) except that the reads of the *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ samples were trimmed to 25 bp to match the ImmGen reads length and avoid mappability bias. The ComBat function in the sva R package was used to correct the batch effect between our and ImmGen samples. Regularized log-transformed (rlog) values were calculated by DESeq2, used to calculate k-means, and perform our integrative approach.

ATAC-seq libraries generation and processing

Sorted *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ populations were prepared according to Buenrostro et al. (2013) using ATAC-Seq kit (#53150; Activemotif). The size distribution and concentration of the libraries were assessed on Tapestation with a DNA High Sensitivity kit (Agilent Technologies). Paired-end 37.5 bp sequences were generated from samples on an NextSeq 500 (Illumina), generating an average of 125 millions of reads per sample.

First, FastQC was used to assess the sequence quality. Foreign sequences removal and trimming are realized with Sickle (qual threshold 20 and length threshold 20). Sequences were mapped to the murine genome with Bowtie2 (2.4.2) (Langmead and Salzberg, 2012) with -X 2,000 (maximal fragment length), very sensitive, and against mm10. Next, we performed various cleaning steps according to Berest et al. (2019): in brief, removing mitochondrial reads, filtering reads with mapping quality <20, removing duplicate reads, and adjusting read start sites as described previously (Buenrostro et al., 2013) (+4 −5). Lastly, a GC bias diagnosis and correction using deepTools was run for each sample. The output of this preprocessing pipeline was used to peak calling using MACS2 with -SE -200 -100 -lambdafix parameter and removal of blacklisted regions.

diffTF and HINT-atac analysis

The complete diffTF pipeline (Berest et al., 2019) was run using TF binding sites generated by PWMscan analysis (cutoff P val: 0.00001; background base composition: 0.29;0.21;0.21;0.29) of each JASPAR2020 PWM motif to obtain the position of 644 nonredundant and specific motifs. The analytical approach was preferred due to the small size of the sample, and paired design was used for DeSeq2 parameter. For the plotting of individual activity for each TF, we used Hint-atac pipeline (Li et al., 2019) with a standard parameter on the combination of the bam for each condition to scan the JASPAR2020 database.

Fastq ATAC-seq data of interest from the ImmGen database was extracted from GSE100738 preprocessed as described previously, and the totality of the reads from all samples was normalized together using local loess normalization as described in Reske et al. (2020) with csaw package. Finally, to represent pile-up traces of the integrated data, Becorrect (Gontarz et al., 2020) was used on each bedgraph file with their corresponding normalized counts.

TNR inference

Peaks were associated with putative TF binding events and target genes to generate a prior network: $P\varepsilon R^{|\mathrm{genes}| \times |\mathrm{TFs}|}$ of TF–gene interactions. We used vertebrate motifs from JASPAR2020 to scan the differential *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ peaks, running construct_atac_prior.R script (P val cutoff = 10^{-5}; window_feature = 25 kb; background set as A = 0.29, C = 0.21, G = 0.21, T = 0.29). We built gene expression models according to Miraldi et al. (2019) and included in our study a target gene expression matrix containing the 1,625 core *PE*$^{\mathrm{tg}}$*H2B-GFP*$^{\mathrm{tg}}$ genes for a total of 26 samples: LT-HSC (*n* = 2), MPP1 (*n* = 2), ST-HSC (*n* = 2), MPP4 (*n* = 2), CLP (*n* = 2), FrA (*n* = 2), FrBC (*n* = 2), FrE (*n* = 2), H2B$^{\mathrm{high}}$ (*n* = 5), and H2B$^{\mathrm{low}}$ (*n* = 5). We considered 275 potential TF regulators corresponding to most variable TFs across the samples. The parameters used were moderate prior reinforcement (λbias = 0.05), 50 subsamples of size 0.90 × |samples|, and an instability cutoff of

0.05 to solve the TF–gene interactions. The $PE^{tg}H2B\text{-}GFP^{tg}$ TRN size was set to an average of 15 TFs per gene and resulted in a total of 24,275 TF–gene interactions.

Online supplemental material

Fig. S1 describes the gating strategy of the different B cell subpopulations from the BM of wt and PE^{tg} preleukemic mice, and the expression of intracytoplasmic ELN and of Sca-1 surface marker are explored. Fig. S2 displays the immunophenotype of PE^{tg} B cells after serial transplantations; the survival curve of quaternary recipient mice after transplantation of PE^{tg} B-ALL cells; the cell-cycle status (Ki67) of wt and PE^{tg} preleukemic B cells in steady state and in response to chemotherapy; the expression of Kit in CTV-labeled wt and preleukemic PE^{tg} cells; the expression of PAX5-ELN in PE^{tg} D0–2, PE^{tg} D8–9, and wt D8–9 cells; the dose-response of chemotherapy on wt and preleukemic PE^{tg} Kit^{+} cells; and the immunophenotype of engrafted cells in recipient mice transplanted with D0–2 and D8–9 preleukemic PE^{tg} cells. Fig. S3 shows FACS analysis of pStat5/pFoxo1 and of pBlnk/pPlcγ2 on PE^{tg} B cells after IL-7 and H_2O_2 stimulation respectively; signaling pathways that are up- and down-regulated in PE^{tg} D0–2 and PE^{tg} D8–9 cells; the kinetic of GFP expression in the different B cell subsets from wt $H2B\text{-}GFP^{tg}$; and preleukemic $PE^{tg}H2B\text{-}GFP^{tg}$ mice after DOX chase, the survival curve of recipient mice transplanted with purified GFP^{high} and GFP^{low} cells from $PE^{tg}H2B\text{-}GFP^{tg}$ mice and the heatmaps of signaling pathways up- and downregulated in GFP^{high} from RNA-seq and ATAC-seq data. Fig. S4 shows the accessibility footprints (ATAC-seq) and the expression levels (RNA-seq) of *Ebf1*, *Spi1*, *Runx1*, and *Pax5* in GFP^{high} and GFP^{low} cells from $PE^{tg}H2B\text{-}GFP^{tg}$ mice; GSEA highlighting the quiescent state of GFP^{high} cells; the transduction efficiency of preleukemic $PE^{tg}Cas9^{tg}$ B cells with sgCTL, sgEgr1 T5, and sgEgr1 T7 lentiviral vectors combined to the genomic editing efficiency of the Egr1 locus; the impact of Egr1 editing on the cell division and proliferation of preleukemic cells; and the molecular network of genes involving Egr1 from TRN#1. Fig. S5 displays the molecular similarities between GFP^{high} cells and residual human B-ALL cells after chemotherapies and shows EGR1 activation in quiescent and therapy-resistant human B-ALL. Table S1 contains the list of FACS antibodies and the sequence of oligonucleotide primers used in the study. Table S2 contains the RNA-seq data of PE^{tg} D0–2, PE^{tg} D8–9, and Wt D8–9 cells. Tables S3 and S4 contain the RNA-seq data and the ATAC-seq data of GFP^{high} and GFP^{low} cells, respectively. Table S5 presents the clustering data of modified genes and peaks from GFP^{high} and GFP^{low} cells according to their signal in the different steps of the normal differentiation.

Data availability

The RNA-seq and ATAC-seq data reported in this study (Tables S2, S3, and S4) are available at the at the Gene Expression Omnibus repository under the accession number GSE210926.

Acknowledgments

We acknowledge Pr. Tariq Enver and Pr. Javier Herrero for having shared their molecular data published in Turati et al. (2021). We acknowledge Manon Farcé from the cytometry and cell sorting facility of the Cancer Research Center of Toulouse (Institut National de la Santé et de la Recherche Médicale [INSERM] U1037) and the Anexplo/Genotoul platforms (UMS006) for technical assistance. We thank the vectorology facility Vect'UB for providing lentiviral particles and for technical support, TBMCore (Centre National de la Recherche Scientifique [CNRS] UAR3427, INSERM US005, University of Bordeaux, Bordeaux, France).

This study was supported by institutional grants from INSERM, from CNRS, the Institut National du Cancer (INCa-2020-096), the Agence Nationale de la Recherche (ANR-18-CE13-0002-01), and the Fondation ARC pour la Recherche sur le Cancer (PJA-20181207977). The team is supported by the Ligue Contre le Cancer and the associations "Laurette Fugain," "111 des arts," "Cassandra," and "Constance la petite guerrière astronaute."

Author contributions: V. Fregona designed the study, performed and analyzed the experiments, and wrote the manuscript. M. Bayet, M. Bouttier, L. Largeaud, C. Hamelle, L.A. Jamrog, N. Prade, S. Lagarde, S. Hebrard, I. Luquet, V. Mansat-De Mas, and M. Nolla performed experiments. M. Pasquet, C. Didier, A.A. Khamlichi, C. Broccardo, É. Delabesse, and S.J.C. Mancini reviewed the data and the manuscript. B. Gerby performed and analyzed experiments, conceived and supervised the project, and wrote the manuscript. All authors contributed to the final draft.

Disclosures: The authors declare no competing interests exist.

Submitted: 14 February 2023
Revised: 31 August 2023
Accepted: 3 October 2023

References

Adelman, E.R., H.T. Huang, A. Roisman, A. Olsson, A. Colaprico, T. Qin, R.C. Lindsley, R. Bejar, N. Salomonis, H.L. Grimes, and M.E. Figueroa. 2019. Aging human hematopoietic stem cells manifest profound epigenetic reprogramming of enhancers that may predispose to leukemia. *Cancer Discov.* 9:1080–1101. https://doi.org/10.1158/2159-8290.CD-18-1474

Alexander, T.B., Z. Gu, I. Iacobucci, K. Dickerson, J.K. Choi, B. Xu, D. Payne-Turner, H. Yoshihara, M.L. Loh, J. Horan, et al. 2018. The genetic basis and cell of origin of mixed phenotype acute leukaemia. *Nature.* 562:373–379. https://doi.org/10.1038/s41586-018-0436-0

Althoff, M.J., R.C. Nayak, S. Hegde, A.M. Wellendorf, B. Bohan, M.D. Filippi, M. Xin, Q.R. Lu, H. Geiger, Y. Zheng, et al. 2020. Yap1-scribble polarization is required for hematopoietic stem cell division and fate. *Blood.* 136:1824–1836. https://doi.org/10.1182/blood.2019004113

Aurrand-Lions, M., and S.J.C. Mancini. 2018. Murine bone marrow niches from hematopoietic stem cells to B cells. *Int. J. Mol. Sci.* 19:2353. https://doi.org/10.3390/ijms19082353

Becht, E., L. McInnes, J. Healy, C.A. Dutertre, I.W.H. Kwok, L.G. Ng, F. Ginhoux, and E.W. Newell. 2018. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat. Biotechnol.* https://doi.org/10.1038/nbt.4314

Bendall, S.C., K.L. Davis, E.A.D. Amir, M.D. Tadmor, E.F. Simonds, T.J. Chen, D.K. Shenfeld, G.P. Nolan, and D. Pe'er. 2014. Single-cell trajectory detection uncovers progression and regulatory coordination in human B cell development. *Cell.* 157:714–725. https://doi.org/10.1016/j.cell.2014.04.005

Berest, I., C. Arnold, A. Reyes-Palomares, G. Palla, K.D. Rasmussen, H. Giles, P.M. Bruch, W. Huber, S. Dietrich, K. Helin, and J.B. Zaugg. 2019. Quantification of differential transcription factor activity and multiomics-based classification into activators and repressors: diffTF. *Cell Rep.* 29:3147–3159.e12. https://doi.org/10.1016/j.celrep.2019.10.106

Bousquet, M., C. Broccardo, C. Quelen, F. Meggetto, E. Kuhlein, G. Delsol, N. Dastugue, and P. Brousset. 2007. A novel PAX5-ELN fusion protein identified in B-cell acute lymphoblastic leukemia acts as a dominant negative on wild-type PAX5. *Blood*. 109:3417–3423. https://doi.org/10.1182/blood-2006-05-025221

Bradner, J.E., D. Hnisz, and R.A. Young. 2017. Transcriptional addiction in cancer. *Cell*. 168:629–643. https://doi.org/10.1016/j.cell.2016.12.013

Buenrostro, J.D., P.G. Giresi, L.C. Zaba, H.Y. Chang, and W.J. Greenleaf. 2013. Transposition of native chromatin for fast and sensitive epigenomic profiling of open chromatin, DNA-binding proteins and nucleosome position. *Nat. Methods*. 10:1213–1218. https://doi.org/10.1038/nmeth.2688

Busslinger, M. 2004. Transcriptional control of early B cell development. *Annu. Rev. Immunol.* 22:55–79. https://doi.org/10.1146/annurev.immunol.22.012703.104807

Cabezas-Wallscheid, N., F. Buettner, P. Sommerkamp, D. Klimmeck, L. Ladel, F.B. Thalheimer, D. Pastor-Flores, L.P. Roma, S. Renders, P. Zeisberger, et al. 2017. Vitamin A-retinoic acid signaling regulates hematopoietic stem cell dormancy. *Cell*. 169:807–823.e19. https://doi.org/10.1016/j.cell.2017.04.018

Canver, M.C., M. Haeussler, D.E. Bauer, S.H. Orkin, N.E. Sanjana, O. Shalem, G.C. Yuan, F. Zhang, J.P. Concordet, and L. Pinello. 2018. Integrated design, execution, and analysis of arrayed and pooled CRISPR genome-editing experiments. *Nat. Protoc.* 13:946–986. https://doi.org/10.1038/nprot.2018.005

Chan, L.N., M.A. Murakami, M.E. Robinson, R. Caeser, T. Sadras, J. Lee, K.N. Cosgun, K. Kume, V. Khairnar, G. Xiao, et al. 2020. Signalling input from divergent pathways subverts B cell transformation. *Nature*. 583:845–851. https://doi.org/10.1038/s41586-020-2513-4

Chavez, J.S., J.L. Rabe, D. Loeffler, K.C. Higa, G. Hernandez, T.S. Mills, N. Ahmed, R.L. Gessner, Z. Ke, B.M. Idler, et al. 2021. PU.1 enforces quiescence and limits hematopoietic stem cell expansion during inflammatory stress. *J. Exp. Med.* 218:e20201169. https://doi.org/10.1084/jem.20201169

Chen, Y., M. Yu, X. Dai, M. Zogg, R. Wen, H. Weiler, and D. Wang. 2011. Critical role for Gimap5 in the survival of mouse hematopoietic stem and progenitor cells. *J. Exp. Med.* 208:923–935. https://doi.org/10.1084/jem.20101192

Cherian, S., and L.A. Soma. 2021. How I diagnose minimal/measurable residual disease in B lymphoblastic leukemia/lymphoma by flow cytometry. *Am. J. Clin. Pathol.* 155:38–54. https://doi.org/10.1093/ajcp/aqaa242

Clark, M.R., M. Mandal, K. Ochiai, and H. Singh. 2014. Orchestrating B cell lymphopoiesis through interplay of IL-7 receptor and pre-B cell receptor signalling. *Nat. Rev. Immunol.* 14:69–80. https://doi.org/10.1038/nri3570

Cobaleda, C., W. Jochum, and M. Busslinger. 2007a. Conversion of mature B cells into T cells by dedifferentiation to uncommitted progenitors. *Nature*. 449:473–477. https://doi.org/10.1038/nature06159

Cobaleda, C., A. Schebesta, A. Delogu, and M. Busslinger. 2007b. Pax5: The guardian of B cell identity and function. *Nat. Immunol.* 8:463–470. https://doi.org/10.1038/ni1454

Coyaud, E., S. Struski, N. Prade, J. Familiades, R. Eichner, C. Quelen, M. Bousquet, F. Mugneret, P. Talmant, M.P. Pages, et al. 2010. Wide diversity of PAX5 alterations in B-ALL: A Groupe Francophone de Cytogenetique Hematologique study. *Blood*. 115:3089–3097. https://doi.org/10.1182/blood-2009-07-234229

Cozzio, A., E. Passegué, P.M. Ayton, H. Karsunky, M.L. Cleary, and I.L. Weissman. 2003. Similar MLL-associated leukemias arising from self-renewing stem cells and short-lived myeloid progenitors. *Genes Dev.* 17:3029–3035. https://doi.org/10.1101/gad.1143403

Dobson, S.M., L. García-Prat, R.J. Vanner, J. Wintersinger, E. Waanders, Z. Gu, J. McLeod, O.I. Gan, I. Grandal, D. Payne-Turner, et al. 2020. Relapse-fated latent diagnosis subclones in acute B lineage leukemia are drug tolerant and possess distinct metabolic programs. *Cancer Discov.* 10:568–587. https://doi.org/10.1158/2159-8290.CD-19-1059

Ebinger, S., E.Z. Özdemir, C. Ziegenhain, S. Tiedt, C. Castro Alves, M. Grunert, M. Dworzak, C. Lutz, V.A. Turati, T. Enver, et al. 2016. Characterization of rare, dormant, and therapy-resistant cells in acute lymphoblastic leukemia. *Cancer Cell*. 30:849–862. https://doi.org/10.1016/j.ccell.2016.11.002

Familiades, J., M. Bousquet, M. Lafage-Pochitaloff, M.C. Béné, K. Beldjord, J. De Vos, N. Dastugue, E. Coyaud, S. Struski, C. Quelen, et al. 2009. PAX5 mutations occur frequently in adult B-cell progenitor acute lymphoblastic leukemia and PAX5 haploinsufficiency is associated with BCR-ABL1 and TCF3-PBX1 fusion genes: A GRAALL study. *Leukemia*. 23:1989–1998. https://doi.org/10.1038/leu.2009.135

Foudi, A., K. Hochedlinger, D. Van Buren, J.W. Schindler, R. Jaenisch, V. Carey, and H. Hock. 2009. Analysis of histone 2B-GFP retention reveals slowly cycling hematopoietic stem cells. *Nat. Biotechnol.* 27:84–90. https://doi.org/10.1038/nbt.1517

Fregona, V., M. Bayet, and B. Gerby. 2021. Oncogene-induced reprogramming in acute lymphoblastic leukemia: Towards targeted therapy of leukemia-initiating cells. *Cancers*. 13:5511. https://doi.org/10.3390/cancers13215511

Garcia-Prat, L., K.B. Kaufmann, F. Schneiter, V. Voisin, A. Murison, J. Chen, M. Chan-Seng-Yue, O.I. Gan, J.L. McLeod, S.A. Smith, et al. 2021. TFEB-mediated endolysosomal activity controls human hematopoietic stem cell fate. *Cell Stem Cell*. 28:1838–1850.e10. https://doi.org/10.1016/j.stem.2021.07.003

Gerby, B., C.S. Tremblay, M. Tremblay, S. Rojas-Sutterlin, S. Herblot, J. Hébert, G. Sauvageau, S. Lemieux, E. Lécuyer, D.F.T. Veiga, and T. Hoang. 2014. SCL, LMO1 and Notch1 reprogram thymocytes into self-renewing cells. *PLoS Genet.* 10:e1004768. https://doi.org/10.1371/journal.pgen.1004768

Gerby, B., D.F.T. Veiga, J. Krosl, S. Nourreddine, J. Ouellette, A. Haman, G. Lavoie, I. Fares, M. Tremblay, V. Litalien, et al. 2016. High-throughput screening in niche-based assay identifies compounds to target preleukemic stem cells. *J. Clin. Invest.* 126:4569–4584. https://doi.org/10.1172/JCI86489

Gontarz, P., S. Fu, X. Xing, S. Liu, B. Miao, V. Bazylianska, A. Sharma, P. Madden, K. Cates, A. Yoo, et al. 2020. Comparison of differential accessibility analysis strategies for ATAC-seq data. *Sci. Rep.* 10:10150. https://doi.org/10.1038/s41598-020-66998-4

Good, Z., J. Sarno, A. Jager, N. Samusik, N. Aghaeepour, E.F. Simonds, L. White, N.J. Lacayo, W.J. Fantl, G. Fazio, et al. 2018. Single-cell developmental classification of B cell precursor acute lymphoblastic leukemia at diagnosis reveals predictors of relapse. *Nat. Med.* 24:474–483. https://doi.org/10.1038/nm.4505

Greaves, M. 2018. A causal mechanism for childhood acute lymphoblastic leukaemia. *Nat. Rev. Cancer*. 18:471–484. https://doi.org/10.1038/s41568-018-0015-6

Gu, Z., M.L. Churchman, K.G. Roberts, I. Moore, X. Zhou, J. Nakitandwe, K. Hagiwara, S. Pelletier, S. Gingras, H. Berns, et al. 2019. PAX5-driven subtypes of B-progenitor acute lymphoblastic leukemia. *Nat. Genet.* 51:296–307. https://doi.org/10.1038/s41588-018-0315-5

Hardy, R.R., and K. Hayakawa. 2001. B cell development pathways. *Annu. Rev. Immunol.* 19:595–621. https://doi.org/10.1146/annurev.immunol.19.1.595

Hardy, R.R., Y.S. Li, D. Allman, M. Asano, M. Gui, and K. Hayakawa. 2000. B-cell commitment, development and selection. *Immunol. Rev.* 175:23–32. https://doi.org/10.1111/j.1600-065X.2000.imr017517.x

Hassanein, N.M., F. Alcancia, K.R. Perkinson, P.J. Buckley, and A.S. Lagoo. 2009. Distinct expression patterns of CD123 and CD34 on normal bone marrow B-cell precursors ("hematogones") and B lymphoblastic leukemia blasts. *Am. J. Clin. Pathol.* 132:573–580. https://doi.org/10.1309/AJCPO4DS0GTLSOEI

Heckl, D., M.S. Kowalczyk, D. Yudovich, R. Belizaire, R.V. Puram, M.E. McConkey, A. Thielke, J.C. Aster, A. Regev, and B.L. Ebert. 2014. Generation of mouse models of myeloid malignancy with combinatorial genetic lesions using CRISPR-Cas9 genome editing. *Nat. Biotechnol.* 32:941–946. https://doi.org/10.1038/nbt.2951

Hong, D., R. Gupta, P. Ancliff, A. Atzberger, J. Brown, S. Soneji, J. Green, S. Colman, W. Piacibello, V. Buckle, et al. 2008. Initiating and cancer-propagating cells in TEL-AML1-associated childhood leukemia. *Science*. 319:336–339. https://doi.org/10.1126/science.1150648

Huntly, B.J.P., H. Shigematsu, K. Deguchi, B.H. Lee, S. Mizuno, N. Duclos, R. Rowan, S. Amaral, D. Curley, I.R. Williams, et al. 2004. MOZ-TIF2, but not BCR-ABL, confers properties of leukemic stem cells to committed murine hematopoietic progenitors. *Cancer Cell*. 6:587–596. https://doi.org/10.1016/j.ccr.2004.10.015

Imperato, M.R., P. Cauchy, N. Obier, and C. Bonifer. 2015. The RUNX1-PU.1 axis in the control of hematopoiesis. *Int. J. Hematol.* 101:319–329. https://doi.org/10.1007/s12185-015-1762-8

Inaba, H., and C.G. Mullighan. 2020. Pediatric acute lymphoblastic leukemia. *Haematologica*. 105:2524–2539. https://doi.org/10.3324/haematol.2020.247031

Inaba, H., and C.H. Pui. 2021. Advances in the diagnosis and treatment of pediatric acute lymphoblastic leukemia. *J. Clin. Med.* 10:1926. https://doi.org/10.3390/jcm10091926

Jamrog, L., G. Chemin, V. Fregona, L. Coster, M. Pasquet, C. Oudinet, N. Rouquie, N. Prade, S. Lagarde, C. Cresson, et al. 2018. PAX5-ELN oncoprotein promotes multistep B-cell acute lymphoblastic leukemia in mice. *Proc. Natl. Acad. Sci. USA*. 115:10357–10362. https://doi.org/10.1073/pnas.1721678115

Jurado, S., A.S. Fedl, M. Jaritz, D. Kostanova-Poliakova, S.G. Malin, C.G. Mullighan, S. Strehl, M. Fischer, and M. Busslinger. 2022. The PAX5-JAK2 translocation acts as dual-hit mutation that promotes aggressive B-cell leukemia via nuclear STAT5 activation. *EMBO J*. 41:e108397. https://doi.org/10.15252/embj.2021108397

Katerndahl, C.D.S., L.M. Heltemes-Harris, M.J.L. Willette, C.M. Henzler, S. Frietze, R. Yang, H. Schjerven, K.A.T. Silverstein, L.B. Ramsey, G. Hubbard, et al. 2017. Antagonism of B cell enhancer networks by STAT5 drives leukemia and poor patient survival. *Nat. Immunol*. 18:694–704. https://doi.org/10.1038/ni.3716

Krivtsov, A.V., D. Twomey, Z. Feng, M.C. Stubbs, Y. Wang, J. Faber, J.E. Levine, J. Wang, W.C. Hahn, D.G. Gilliland, et al. 2006. Transformation from committed progenitor to leukaemia stem cell initiated by MLL-AF9. *Nature*. 442:818–822. https://doi.org/10.1038/nature04980

Langmead, B., and S.L. Salzberg. 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods*. 9:357–359. https://doi.org/10.1038/nmeth.1923

Li, Z., M.H. Schulz, T. Look, M. Begemann, M. Zenke, and I.G. Costa. 2019. Identification of transcription factor binding sites using ATAC-seq. *Genome Biol*. 20:45. https://doi.org/10.1186/s13059-019-1642-2

Liu, Y., Y. Chen, X. Deng, and J. Zhou. 2020. ATF3 prevents stress-induced hematopoietic stem cell exhaustion. *Front. Cell Dev. Biol*. 8:585771. https://doi.org/10.3389/fcell.2020.585771

Malta, T.M., A. Sokolov, A.J. Gentles, T. Burzykowski, L. Poisson, J.N. Weinstein, B. Kamińska, J. Huelsken, L. Omberg, O. Gevaert, et al. 2018. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell*. 173:338–354.e15. https://doi.org/10.1016/j.cell.2018.03.034

McCormack, M.P., L.F. Young, S. Vasudevan, C.A. de Graaf, R. Codrington, T.H. Rabbitts, S.M. Jane, and D.J. Curtis. 2010. The Lmo2 oncogene initiates leukemia in mice by inducing thymocyte self-renewal. *Science*. 327:879–883. https://doi.org/10.1126/science.1182378

Min, I.M., G. Pietramaggiori, F.S. Kim, E. Passegué, K.E. Stevenson, and A.J. Wagers. 2008. The transcription factor EGR1 controls both the proliferation and localization of hematopoietic stem cells. *Cell Stem Cell*. 2:380–391. https://doi.org/10.1016/j.stem.2008.01.015

Miraldi, E.R., M. Pokrovskii, A. Watters, D.M. Castro, N. De Veaux, J.A. Hall, J.Y. Lee, M. Ciofani, A. Madar, N. Carriero, et al. 2019. Leveraging chromatin accessibility for transcriptional regulatory network inference in T Helper 17 Cells. *Genome Res*. 29:449–463. https://doi.org/10.1101/gr.238253.118

Monaco, G., H. Chen, M. Poidinger, J. Chen, J.P. de Magalhães, and A. Larbi. 2016. flowAI: Automatic and interactive anomaly discerning tools for flow cytometry data. *Bioinformatics*. 32:2473–2480. https://doi.org/10.1093/bioinformatics/btw191

Morcos, M.N.F., K.B. Schoedel, A. Hoppe, R. Behrendt, O. Basak, H.C. Clevers, A. Roers, and A. Gerbaulet. 2017. SCA-1 expression level identifies quiescent hematopoietic stem and progenitor cells. *Stem Cell Rep*. 8:1472–1478. https://doi.org/10.1016/j.stemcr.2017.04.012

Mullighan, C.G., S. Goorha, I. Radtke, C.B. Miller, E. Coustan-Smith, J.D. Dalton, K. Girtman, S. Mathew, J. Ma, S.B. Pounds, et al. 2007. Genome-wide analysis of genetic alterations in acute lymphoblastic leukaemia. *Nature*. 446:758–764. https://doi.org/10.1038/nature05690

Mullighan, C.G., L.A. Phillips, X. Su, J. Ma, C.B. Miller, S.A. Shurtleff, and J.R. Downing. 2008. Genomic analysis of the clonal origins of relapsed acute lymphoblastic leukemia. *Science*. 322:1377–1380. https://doi.org/10.1126/science.1164266

Ng, S.W.K., A. Mitchell, J.A. Kennedy, W.C. Chen, J. McLeod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A.D. Schimmer, et al. 2016. A 17-gene stemness score for rapid determination of risk in acute leukaemia. *Nature*. 540:433–437. https://doi.org/10.1038/nature20598

Nutt, S.L., B. Heavey, A.G. Rolink, and M. Busslinger. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature*. 401:556–562. https://doi.org/10.1038/44076

Patterson, H.C., C. Gerbeth, P. Thiru, N.F. Vögtle, M. Knoll, A. Shahsafaei, K.E. Samocha, C.X. Huang, M.M. Harden, R. Song, et al. 2015. A respiratory chain controll signal transduction cascade in the mitochondrial intermembrane space mediates hydrogen peroxide signaling. *Proc. Natl. Acad. Sci. USA*. 112:E5679–E5688. https://doi.org/10.1073/pnas.1517932112

Picelli, S., O.R. Faridani, A.K. Björklund, G. Winberg, S. Sagasser, and R. Sandberg. 2014. Full-length RNA-seq from single cells using Smart-seq2. *Nat. Protoc*. 9:171–181. https://doi.org/10.1038/nprot.2014.006

Pokrovskii, M., J.A. Hall, D.E. Ochayon, R. Yi, N.S. Chaimowitz, H. Seelamneni, N. Carriero, A. Watters, S.N. Waggoner, D.R. Littman, et al. 2019. Characterization of transcriptional regulatory networks that promote and restrict identities and functions of intestinal innate lymphoid cells. *Immunity*. 51:185–197.e6. https://doi.org/10.1016/j.immuni.2019.06.001

Prost, S., F. Relouzat, M. Spentchian, Y. Ouzegdouh, J. Saliba, G. Massonnet, J.P. Beressi, E. Verhoeyen, V. Raggueneau, B. Maneglier, et al. 2015. Erosion of the chronic myeloid leukaemia stem cell pool by PPARγ agonists. *Nature*. 525:380–383. https://doi.org/10.1038/nature15248

Ramamoorthy, S., K. Kometani, J.S. Herman, M. Bayer, S. Boller, J. Edwards-Hicks, H. Ramachandran, R. Li, R. Klein-Geltink, E.L. Pearce, et al. 2020. EBF1 and Pax5 safeguard leukemic transformation by limiting IL-7 signaling, Myc expression, and folate metabolism. *Genes Dev*. 34:1503–1519. https://doi.org/10.1101/gad.340216.120

Reske, J.J., M.R. Wilson, and R.L. Chandler. 2020. ATAC-seq normalization method can significantly affect differential accessibility analysis and interpretation. *Epigenetics Chromatin*. 13:22. https://doi.org/10.1186/s13072-020-00342-y

Reth, M. 2002. Hydrogen peroxide as second messenger in lymphocyte activation. *Nat. Immunol*. 3:1129–1134. https://doi.org/10.1038/ni1202-1129

Reth, M., and P. Nielsen. 2014. Signaling circuits in early B-cell development. *Adv. Immunol*. 122:129–175. https://doi.org/10.1016/B978-0-12-800267-4.00004-3

Revilla-I-Domingo, R., I. Bilic, B. Vilagos, H. Tagoh, A. Ebert, I.M. Tamir, L. Smeenk, J. Trupke, A. Sommer, M. Jaritz, and M. Busslinger. 2012. The B-cell identity factor Pax5 regulates distinct transcriptional programmes in early and late B lymphopoiesis. *EMBO J*. 31:3130–3146. https://doi.org/10.1038/emboj.2012.155

Rodríguez-Hernández, G., A. Casado-García, M. Isidro-Hernández, D. Picard, J. Raboso-Gallego, S. Alemán-Arteaga, A. Orfao, O. Blanco, S. Riesco, P. Prieto-Matos, et al. 2021. The second oncogenic hit determines the cell fate of ETV6-RUNX1 positive leukemia. *Front. Cell Dev. Biol*. 9:704591. https://doi.org/10.3389/fcell.2021.704591

Roederer, M. 2011. Interpretation of cellular proliferation data: Avoid the panglossian. *Cytometry A*. 79:95–101. https://doi.org/10.1002/cyto.a.21010

Rolink, A., and F. Melchers. 1996. B-cell development in the mouse. *Immunol. Lett*. 54:157–161. https://doi.org/10.1016/S0165-2478(96)02666-1

Scheicher, R., A. Hoelbl-Kovacic, F. Bellutti, A.S. Tigan, M. Prchal-Murphy, G. Heller, C. Schneckenleithner, M. Salazar-Roa, S. Zöchbauer-Müller, J. Zuber, et al. 2015. CDK6 as a key regulator of hematopoietic and leukemic stem cell activation. *Blood*. 125:90–101. https://doi.org/10.1182/blood-2014-06-584417

Smeenk, L., M. Fischer, S. Jurado, M. Jaritz, A. Azaryan, B. Werner, M. Roth, J. Zuber, M. Stanulla, M.L. den Boer, et al. 2017. Molecular role of the PAX5-ETV6 oncoprotein in promoting B-cell acute lymphoblastic leukemia. *EMBO J*. 36:718–735. https://doi.org/10.15252/embj.201695495

Staber, P.B., P. Zhang, M. Ye, R.S. Welner, E. Levantini, A. Di Ruscio, A.K. Ebralidze, C. Bach, H. Zhang, J. Zhang, et al. 2014. The Runx-PU.1 pathway preserves normal and AML/ETO9a leukemic stem cells. *Blood*. 124:2391–2399. https://doi.org/10.1182/blood-2014-01-550855

Stoddart, A., A.A. Fernald, E.M. Davis, M.E. McNerney, and M.M. Le Beau. 2022. EGR1 haploinsufficiency confers a Fitness advantage to hematopoietic stem cells following chemotherapy. *Exp. Hematol*. 115:54–67. https://doi.org/10.1016/j.exphem.2022.08.003

Tomura, M., A. Sakaue-Sawano, Y. Mori, M. Takase-Utsugi, A. Hata, K. Ohtawa, O. Kanagawa, and A. Miyawaki. 2013. Contrasting quiescent G0 phase with mitotic cell cycling in the mouse immune system. *PLoS One*. 8:e73801. https://doi.org/10.1371/journal.pone.0073801

Tremblay, C.S., J. Saw, S.K. Chiu, N.C. Wong, K. Tsyganov, S. Ghotb, A.N. Graham, F. Yan, A.A. Guirguis, S.E. Sonderegger, et al. 2018. Restricted cell cycle is essential for clonal evolution and therapeutic resistance of pre-leukemic stem cells. *Nat. Commun*. 9:3535. https://doi.org/10.1038/s41467-018-06021-7

Trumpp, A., M. Essers, and A. Wilson. 2010. Awakening dormant haematopoietic stem cells. *Nat. Rev. Immunol*. 10:201–209. https://doi.org/10.1038/nri2726

Turati, V.A., J.A. Guerra-Assunção, N.E. Potter, R. Gupta, S. Ecker, A. Daneviciute, M. Tarabichi, A.P. Webster, C. Ding, G. May, et al. 2021. Chemotherapy induces canalization of cell state in childhood B-cell precursor acute lymphoblastic leukemia. *Nat. Cancer*. 2:835–852. https://doi.org/10.1038/s43018-021-00219-3

Tzelepis, K., H. Koike-Yusa, E. De Braekeleer, Y. Li, E. Metzakopian, O.M. Dovey, A. Mupo, V. Grinkevich, M. Li, M. Mazan, et al. 2016. A CRISPR dropout screen identifies genetic vulnerabilities and therapeutic targets in acute myeloid leukemia. *Cell Rep*. 17:1193–1205. https://doi.org/10.1016/j.celrep.2016.09.079

van Delft, F.W., S. Horsley, S. Colman, K. Anderson, C. Bateman, H. Kempski, J. Zuna, C. Eckert, V. Saha, L. Kearney, et al. 2011. Clonal origins of relapse in ETV6-RUNX1 acute lymphoblastic leukemia. *Blood*. 117: 6247–6254. https://doi.org/10.1182/blood-2010-10-314674

Van Gassen, S., B. Gaudilliere, M.S. Angst, Y. Saeys, and N. Aghaeepour. 2020. CytoNorm: A normalization algorithm for cytometry data. *Cytometry A*. 97:268–278. https://doi.org/10.1002/cyto.a.23904

Vazquez-Santillan, K., J. Melendez-Zajgla, L. Jimenez-Hernandez, G. Martínez-Ruiz, and V. Maldonado. 2015. NF-κB signaling in cancer stem cells: A promising therapeutic target? *Cell Oncol*. 38:327–339. https://doi.org/10.1007/s13402-015-0236-6

Venezia, T.A., A.A. Merchant, C.A. Ramos, N.L. Whitehouse, A.S. Young, C.A. Shaw, and M.A. Goodell. 2004. Molecular signatures of proliferation and quiescence in hematopoietic stem cells. *PLoS Biol*. 2:e301. https://doi.org/10.1371/journal.pbio.0020301

Waanders, E., Z. Gu, S.M. Dobson, Ž. Antić, J.C. Crawford, X. Ma, M.N. Edmonson, D. Payne-Turner, M. van de Vorst, M.C.J. Jongmans, et al. 2020. Mutational landscape and patterns of clonal evolution in relapsed pediatric acute lymphoblastic leukemia. *Blood Cancer Discov*. 1:96–111. https://doi.org/10.1158/0008-5472.BCD-19-0041

Wilson, A., E. Laurenti, G. Oser, R.C. van der Wath, W. Blanco-Bose, M. Jaworski, S. Offner, C.F. Dunant, L. Eshkind, E. Bockamp, et al. 2008. Hematopoietic stem cells reversibly switch from dormancy to self-renewal during homeostasis and repair. *Cell*. 135:1118–1129. https://doi.org/10.1016/j.cell.2008.10.048

Wojiski, S., F.C. Guibal, T. Kindler, B.H. Lee, J.L. Jesneck, A. Fabian, D.G. Tenen, and D.G. Gilliland. 2009. PML-RARalpha initiates leukemia by conferring properties of self-renewal to committed promyelocytic progenitors. *Leukemia*. 23:1462–1471. https://doi.org/10.1038/leu.2009.63

Yamashita, M., and E. Passegué. 2019. TNF-A coordinates hematopoietic stem cell survival and myeloid regeneration. *Cell Stem Cell*. 25:357–372.e7. https://doi.org/10.1016/j.stem.2019.05.019

Yan, F., N.C. Wong, D.R. Powell, and D.J. Curtis. 2020. A 9-gene score for risk stratification in B-cell acute lymphoblastic leukemia. *Leukemia*. 34: 3070–3074. https://doi.org/10.1038/s41375-020-0888-8

Zhang, Y.W., J. Mess, N. Aizarani, P. Mishra, C. Johnson, M.C. Romero-Mulero, J. Rettkowski, K. Schönberger, N. Obier, K. Jäcklein, et al. 2022. Hyaluronic acid-GPRC5C signalling promotes dormancy in haematopoietic stem cells. *Nat. Cell Biol*. 24:1038–1048. https://doi.org/10.1038/s41556-022-00931-x

Supplemental material

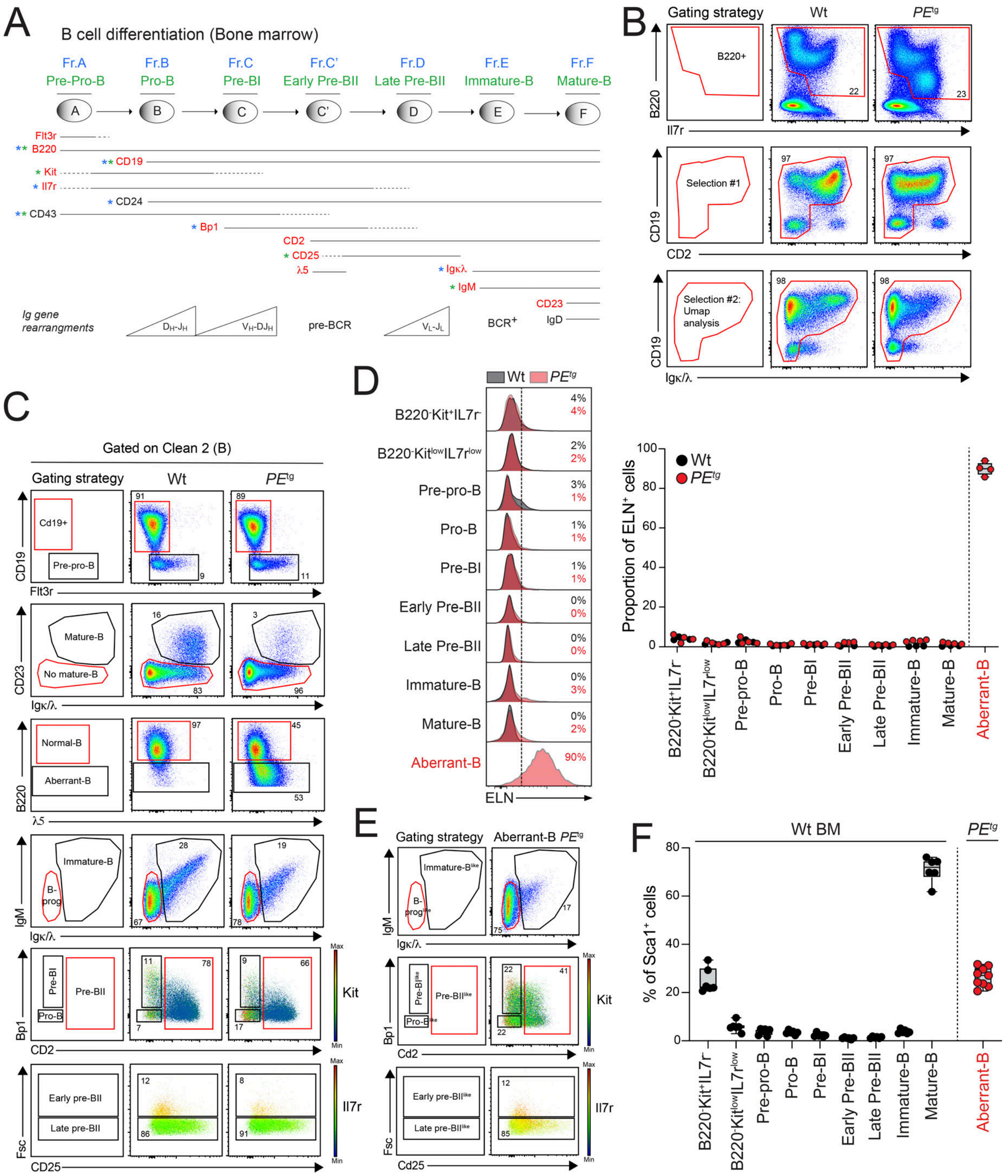

Figure S1. **Phenotypic characterization of preleukemic mice. (A)** Schematic representation of the murine BM B cell differentiation. Representation of the different B cell differentiation stages (pre-pro-B, pro-B, pre-BI, early pre-BII, late pre-BII, immature-B, and mature-B) and pattern of expression of the main markers used to characterize each subset. Both Philadelphia (Hardy and Hayakawa, 2001) and Basel (Rolink and Melchers, 1996) nomenclatures are indicated in blue and green, respectively. Blue and green stars indicate the markers used in the respective nomenclatures. The thickness of the line is representative of the expression level. The dotted lines indicate subsets where expression is progressively acquired or lost. The markers used to perform the multiparametric staining by FACS, and the UMAP representation are indicated in red. $V_HD_HJ_H$ and V_LJ_L rearrangements are indicated. **(B and C)** Gating strategy of the B cell compartment (B) and of the different B cell subpopulations (C) from the BM of wt and *PE*^{tg} preleukemic mice to perform the clustering analysis by UMAP shown in Fig. 1 A. Each red gate represents the parent population of the following below set of FACS plots. Kit expression level in pro-B, pre-BI, and pre-BII subsets and IL7r expression level in early and late pre-BII subsets are shown. **(D)** Expression of intracytoplasmic ELN in aberrant B cells (CD19$^+$B220^{low}), in all the steps of B cell differentiation (CD19$^+$B220$^+$), in HSPC-enriched (CD19$^-$B220$^-$Kit$^+$IL7r$^-$) and CLP-enriched (CD19$^-$B220$^-$Kit^{low}IL7r^{low}) populations from preleukemic PE^{tg} mice (*n* = 4). Wt mice (*n* = 4) were used as controls. **(E)** Representative immunophenotype of the aberrant B population from *PE*^{tg} preleukemic BM identified in Fig. S1 C. The aberrant B *PE*^{tg} population was subdivided into the different phenotypic-like subsets pro-B^{like}, pre-BI^{like}, early pre-BII^{like}, late pre-BII^{like}, immature-B^{like}. Kit expression levels in pro-B^{like}, pre-BI^{like}, and pre-BII^{like} subsets and Il7r expression levels in early and late pre-BII^{like} subsets are shown. **(F)** Expression of Sca-1 in aberrant B cells (CD19$^+$B220^{low}), in all the steps of B cell differentiation (CD19$^+$B220$^+$), in HSPC-enriched (CD19$^-$B220$^-$Kit$^+$IL7r$^-$) and CLP-enriched (CD19$^-$B220$^-$Kit^{low}IL7r^{low}) populations from PE^{tg} mice (*n* = 9). Wt mice (*n* = 6) were used as controls.

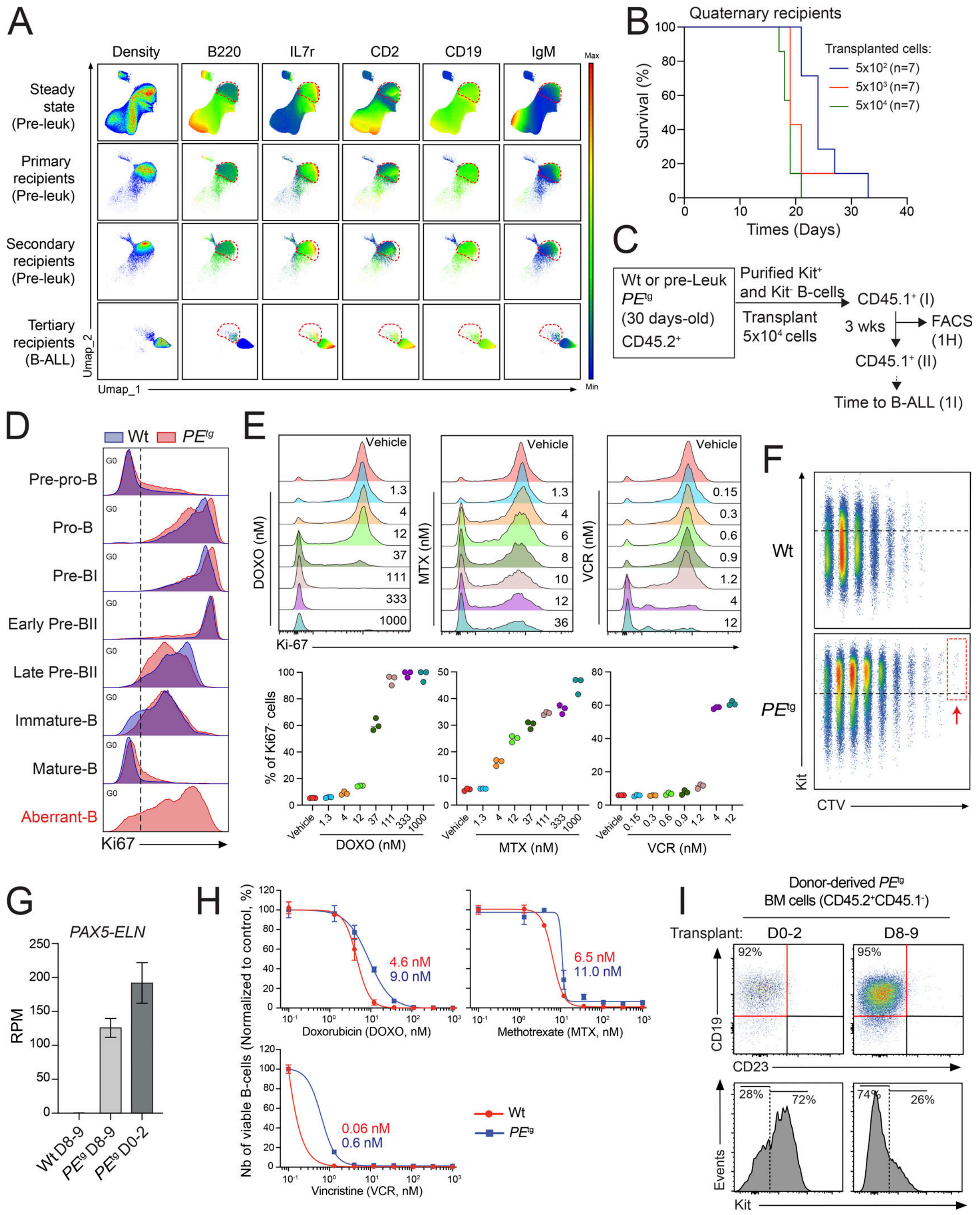

Figure S2. **Cell-cycle status of preleukemic B cells. (A)** UMAP of the cell density (left panel) before (steady state) and after transplantation in primary, secondary, and tertiary recipients, associated with the expression level of B220, IL7r, CD2, CD19, and IgM (right panels). **(B)** Kaplan–Meier survival curve of quaternary recipient mice transplanted with 5 × 10² (*n* = 7), 5 × 10³ (*n* = 7), and 5 × 10⁴ (*n* = 7) donor-derived blasts from tertiary B-ALL *PE*tg mice. **(C)** Experimental procedure to study the engraftment potential of kit⁺ aberrant B cells. **(D)** Ki67 staining was performed on B cells from the BM of wt and preleukemic *PE*tg mice according to Fig. 2 A (*n* = 4). Representative FACS analysis of the proportion of cells in G0 phase (Ki67⁻ cells) within each wt and *PE*tg B cell subpopulation. **(E)** Preleukemic *PE*tg cells were plated on MS5 stromal cells and treated with increasing doses (*n* = 3 for each dose, one experiment) of DOXO, MTX, or VCR for 48 h. Ki67 staining was then performed, and the proportion of Ki67⁺ and Ki67⁻ within the aberrant B cell population was shown (upper panels) and quantified (lower panels) for each dose. **(F)** FACS analysis (Winlist software) of the cell divisions (CTV) in function of Kit expression level after the coculture of CTV-labeled wt and preleukemic *PE*tg cells. Red arrow indicates the undivided (D0) population. **(G)** Expression level of the *PAX5::ELN* fusion transcript in *PE*tg D0–2 (*n* = 2), *PE*tg D8–9 (*n* = 2), and wt D8–9 (*n* = 2) cells (mean ± SD). RNA-seq data were aligned independently on the sequence of *PAX5::ELN* human transgene. Expression levels are represented in "reads" per million (RPM). **(H)** Purified wt and preleukemic *PE*tg Kit⁺ cells were plated on MS5 stromal cells and treated with increasing doses (*n* = 3 for each dose, one experiment) of DOXO, MTX, or VCR for 48 h. Cell numbers of treated cells were then analyzed by FACS and normalized to the number of untreated cells. IC50 of each compound on wt and preleukemic *PE*tg cells were determined according to the respective dose-response curves. **(I)** Representative immunophenotype of donor-derived (CD45.2⁺CD45.1⁻) cells in engrafted recipient BM 3 wk after the transplantation of slow- (D0–2) and high- (D8–9) cycling preleukemic *PE*tg cells.

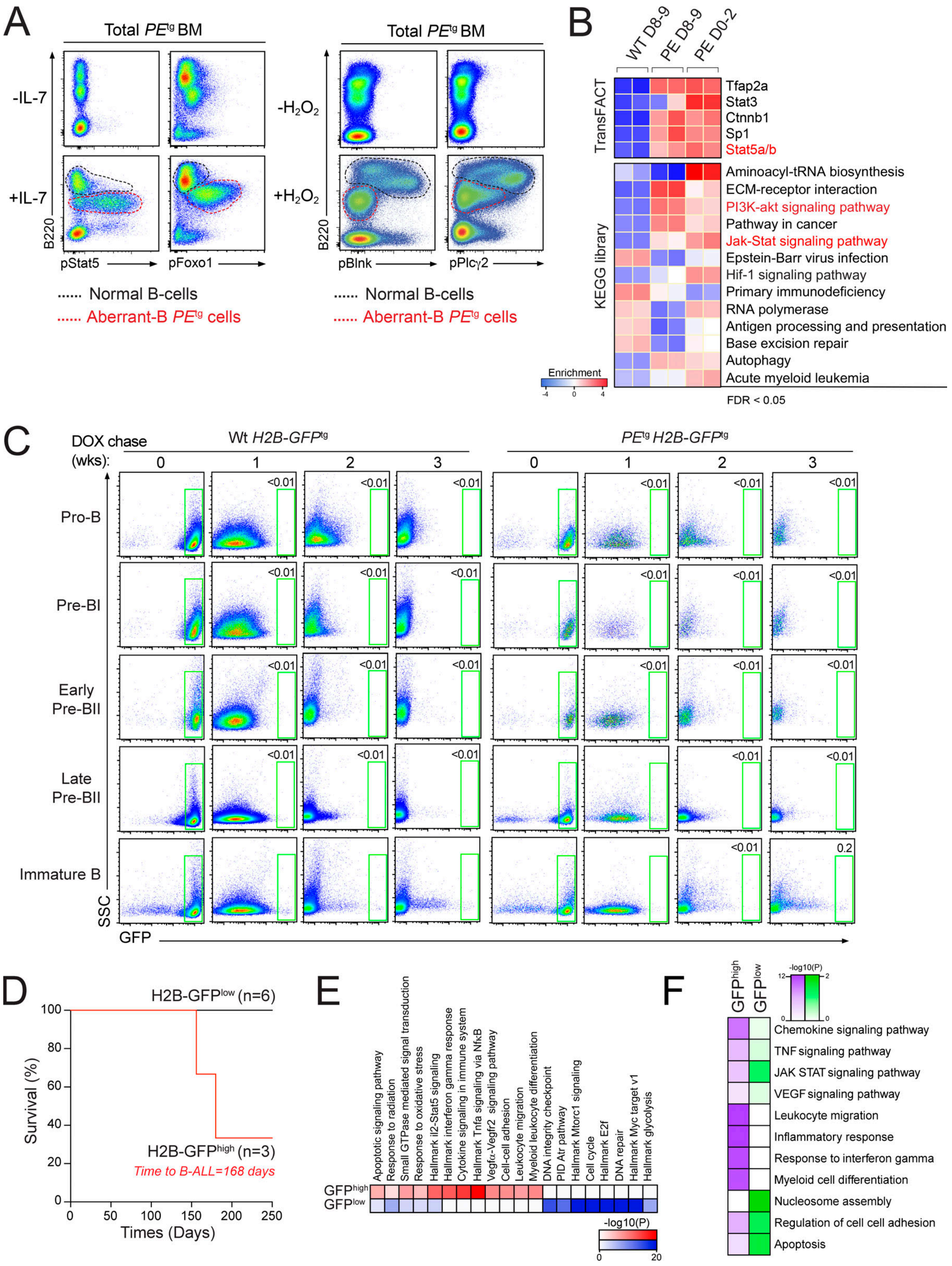

Figure S3. **Molecular characterization of quiescent preleukemic B cells. (A)** Representative FACS analysis of pStat5 and pFoxo1 on total BM from preleukemic *PE*^{tg} mice after ex vivo IL-7 stimulation (+IL-7). Unstimulated (−IL-7) cells were used as control (left panels). FACS analysis of pBlnk and pPlcγ2 on total BM from preleukemic *PE*^{tg} mice after ex vivo H_2O_2 stimulation (+H_2O_2). Unstimulated (−H_2O_2) cells were used as control (right panels). **(B)** Gene set variation analysis of *PE*^{tg} D0–2 (*n* = 2), *PE*^{tg} D8–9 (*n* = 2), and wt D8–9 (*n* = 2) cells was performed using TransFACT and KEGG libraries and top gene sets were represented in heatmap. **(C)** Representative FACS analysis of GFP expression in the different B cell subsets from wt *H2B-GFP*^{tg} (left panels) and preleukemic PE^{tg}H2B-GFP^{tg} (right panels) mice after 0, 1, 2, and 3 wk of DOX chase. The proportion of GFP^{high} cells is indicated for each time point. **(D)** Kaplan–Meier survival curve of recipient mice transplanted with purified GFP^{high} (*n* = 3) and GFP^{low} (*n* = 6) cells from *PE*^{tg}*H2B-GFP*^{tg} mice. **(E)** Heatmap of Metascape results of the top selected pathways between all the differentially expressed genes between GFP^{high} (*n* = 5) and GFP^{low} (*n* = 5) cells. **(F)** Heatmap of enrichment analysis of differentially accessible regions obtained with Cistrome-GO.

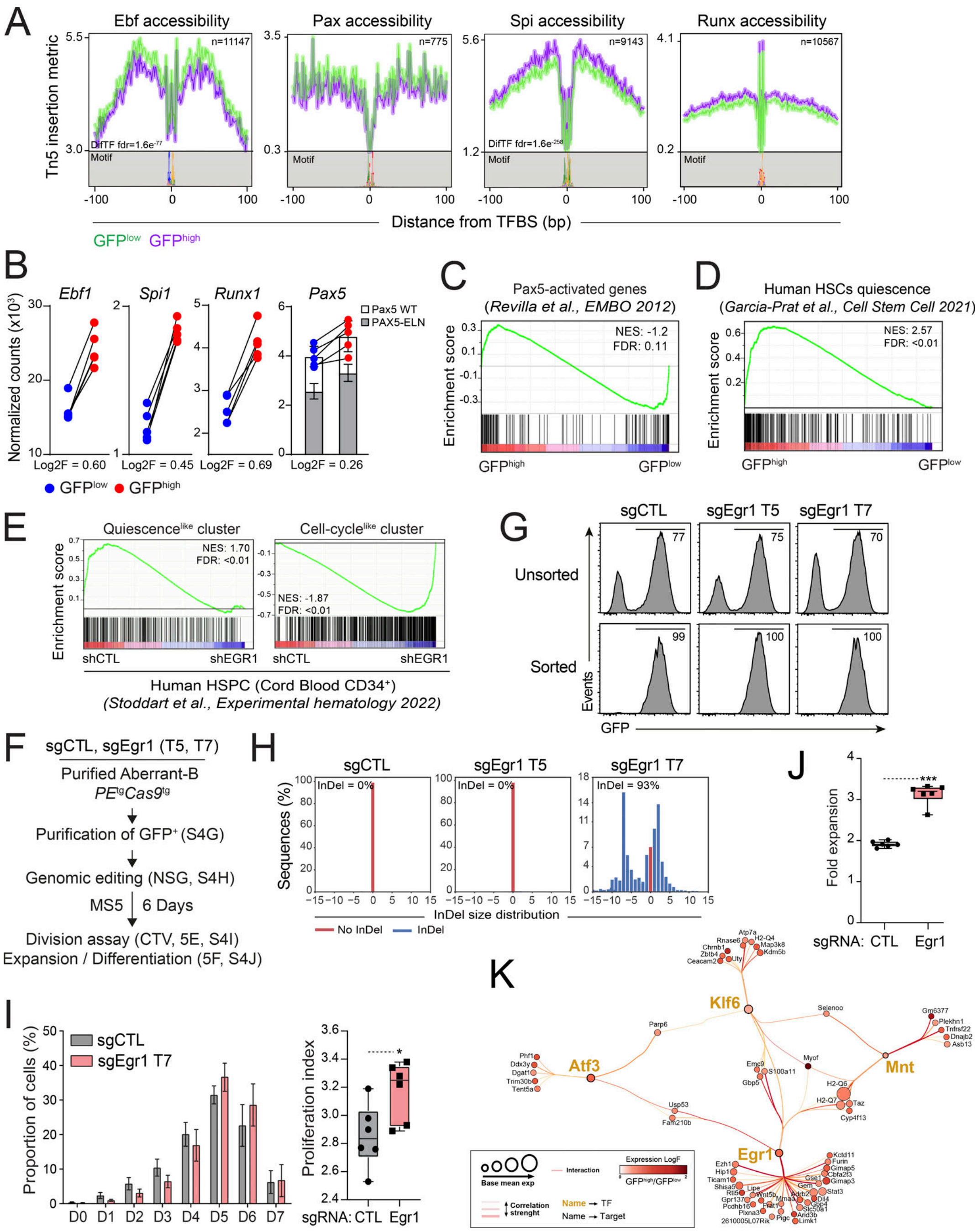

Figure S4. **Impact of Egr1 editing on cell proliferation of preleukemic B cells. (A)** Comparison of the accessibility footprint of Ebf, Pax, Spi, and Runx motifs (ATAC-seq data) between GFP^{high} (purple) and GFP^{low} (green) cells. **(B)** Comparison of *Ebf1*, *Spi1*, *Runx1*, and of total *Pax5* expression (RNA-seq data, *n* = 5, mean ± SD) between GFP^{high} (red) and GFP^{low} (blue) cells. The proportions of murine *Pax5* (*Pax5* WT) and of human *PAX5* (*PAX5-ELN*) are shown. **(C)** GSEA of the well-established list of in vivo PAX5-activated genes (Revilla-I-Domingo et al., 2012) between GFP^{high} and GFP^{low} cells. **(D)** GSEA of the established list of in vivo human HSC quiescence genes (Garcia-Prat et al., 2021) between GFP^{high} (red) and GFP^{low} (blue) cells. **(E)** GSEA of the list of genes in the quiescence^{like} cluster (left panel) and in the cell-cycle^{like} cluster (right panel) and identified in Fig. 5 C between human CD34⁺ HSPCs modified or not (shCTL) with *EGR1* shRNA (shEGR1) (Stoddart et al., 2022). **(F)** Experimental procedure to study the role of Egr1 in *PE*^{tg} pre-LSC activity. **(G)** Transduction efficiency of preleukemic *PE*^{tg}*Cas9*^{tg} aberrant B cells expressing sgCTL, sgEgr1 T5, and sgEgr1 T7 was assessed by FACS by monitoring the percentage of GFP for each condition (upper panels). Transduced *PE*^{tg}*Cas9*^{tg} GFP⁺ cells were purified by cell sorting and cell purity was then verified (lower panels). **(H)** Analysis of the targeted region (exon 1) of *Egr1* by NGS and quantification of the genomic editing efficiency. **(I)** The number of cell divisions (D0 to D7) of aberrant B GFP⁺ cells expressing sgCTL or sgEgr1 T7 after 6 days of coculture was analyzed (left panel) and the PI was calculated for each condition (right panel, *n* = 6, one experiment, $^{*}P < 0.05$). **(J)** In vitro cell growth of GFP⁺ sgCTL and sgEgr1 T7 cells. Results represent the fold expansion of aberrant B GFP⁺ cells after 6 days of coculture on MS5 stromal cells (*n* = 6, one experiment, $^{***}P < 0.001$). The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. **(K)** Molecular network of genes from TRN#1 of Fig. 5 D. Node color represents the log-fold difference between GFP^{high} and GFP^{low} subset and node size represents the base-mean expression. Target genes positively regulated by the four TFs, Egr1, Mnt, Klf6, and Atf3, were annotated.

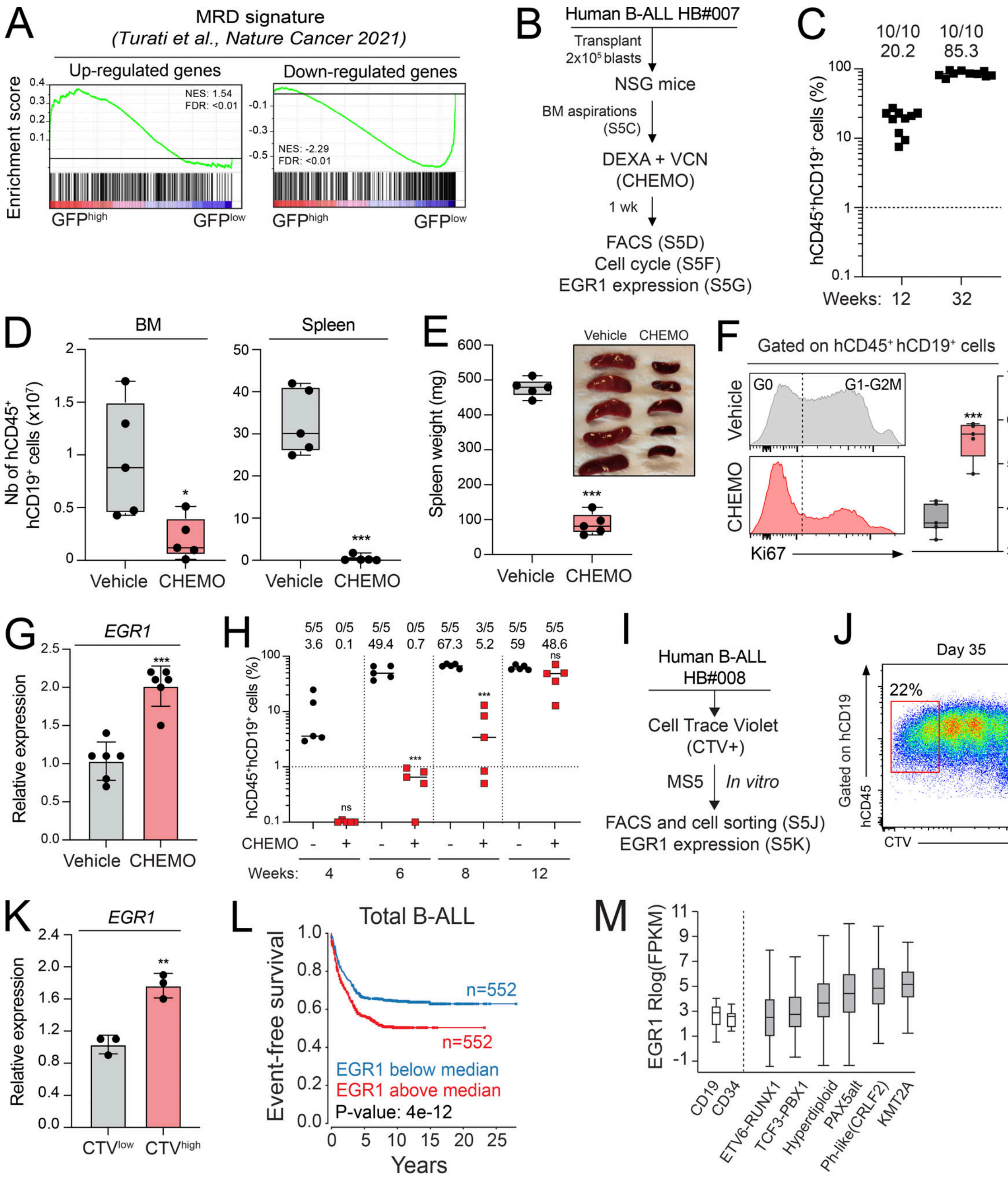

Figure S5. **Upregulation of *EGR1* expression in residual leukemic blasts. (A)** GSEA of the up- (left panel) and down- (right panel) regulated genes from residual human B-ALL cells after chemotherapies (Turati et al., 2021) between GFP^{high} and GFP^{low} cells. **(B–G)** Experimental procedure to explore EGR1 expression in resistant human B-ALL cells (B). Leukemic blasts from the “de novo” B-ALL patient HB#007 were transplanted into NSG mice ($n = 11$). The proportion of hCD45$^+$hCD19$^+$ B-ALL blasts in BM aspirations was assessed 12 and 32 wk after transplantation (C). Engrafted mice were randomly selected for treatment (CHEMO) or not (vehicle) with a chemotherapeutic cocktail (DEXA+VCR) for 1 wk. Human B-ALL reconstitution (number of hCD45$^+$hCD19$^+$ B-ALL blasts) was monitored by FACS in the BM and the spleen after the treatment (D). Spleen weight was measured, and a picture of the spleen is shown (E). Ki67 expression was determined by FACS in blasts from the BM of treated and untreated mice (F). Data obtained with B-ALL patient HB#007 were compiled from five (vehicle) and five (CHEMO) mice per condition (*$P < 0.05$, ***$P < 0.001$). The horizontal lines of the box plots indicate the median, while the boxes represent the first and the third quartiles of the data and the whiskers denote the minimum and the maximum values. Engrafted blasts were purified and mRNA levels of the human *EGR1* were determined by quantitative RT-PCR and normalized to *ABL1* gene. Error bars indicate mean ± SD ($n = 6$, ***$P < 0.001$) from two independent experiments (G). **(H)** Engrafted blasts (B-ALL patient HB#010) from the BM of treated (CHEMO) and untreated (Vehicle) mice (Fig. 6 E) were purified and transplanted in equal number (2×10^5 cells/mouse) in secondary recipients. The proportion of hCD45$^+$hCD19$^+$ B-ALL blasts in BM aspirations was assessed 4, 6, 8, and 12 wk after transplantation was monitored by FACS ($n = 5$, one experiment, ***$P < 0.001$). **(I–K)** Experimental procedure to explore EGR1 expression in quiescent human B-ALL cells (I). Leukemic blasts from the “de novo” B-ALL patient HB#008 were stained with CTV and the number of cell divisions (D0 to D4–5) after the coculture was then analyzed. Red dotted gates were used to purify the D0 (CTV^{high}) and D4–5 (CTV^{low}) populations (J) and the expression of human *EGR1* were determined by quantitative RT-PCR (K, mean ± SD, $n = 3$, one experiment, **$P < 0.01$). **(L)** *EGR1* expression and its associated clinical outcome in B-ALL patient extracted from Gu et al. (2019). EFS curves of total B-ALL patients with *EGR1* expression below (blue) and above (red) to the median. **(M)** *EGR1* expression in the different B-ALL oncogenic subgroups was shown.

Provided online are Table S1, Table S2, Table S3, Table S4, and Table S5. Table S1 is a list of FACS antibodies used for the phenotypic characterization of the B cell lineage. Table S2 shows results of differentially expressed genes of *PE*^{tg} D0–2, *PE*^{tg} D8–9, and Wt D8–9 cells analyzed with DESeq2 ($Log_2F > 0.65$; P adj <0.05). Table S3 shows results of differentially expressed genes (RNA-seq) of H2B-GFP^{high} and H2B-GFP^{low} cells analyzed with DESeq2 ($Log_2F > 0.6$; P adj <0.05). Table S4 shows results of differentially accessible regions (ATAC-seq) of GFP^{high} and GFP^{low} cells and their associated TF activity analyzed with diffTF pipeline. Table S5 shows clustering data (k-means) of significantly modified genes (related to Fig. 5 C and Table S3) and peaks (related to Fig. 5 B and Table S4) from preleukemic GFP^{high} and GFP^{low} cells according to their signal in the different steps of the normal differentiation (extracted from RNA-seq and ATAC-seq ImmGen datasets).