

# RAG1 and RAG2 non-core regions are implicated in leukemogenesis and off-target V(D)J recombination in BCR-ABL1-driven B-cell lineage lymphoblastic leukemia

Xiaozhuo Yu<sup>1†</sup>, Wen Zhou<sup>1†</sup>, Xiaodong Chen<sup>1</sup>, Shunyu He<sup>1</sup>, Mengting Qin<sup>1</sup>, Meng Yuan<sup>1</sup>, Yang Wang<sup>1</sup>, Woodvine Otieno Odhiambo<sup>1</sup>, Yinsha Miao<sup>2\*</sup>, Yanhong Ji<sup>1,2\*</sup>

<sup>1</sup>Department of Pathogenic Biology and Immunology, School of Basic Medical Sciences, Xi'an Jiaotong University Health Science Center, Xi'an, China; <sup>2</sup>Department of Clinical Laboratory, Xi'an No. 3 Hospital, the Affiliated Hospital of Northwest University, Xian, China

**\*For correspondence:**

miaoyinsha@med.nwu.edu.cn (YM);

jiyanhong@xjtu.edu.cn (YJ)

<sup>†</sup>These authors contributed equally to this work

**Competing interest:** The authors declare that no competing interests exist.

**Funding:** See page 16

**Sent for Review**

08 August 2023

**Preprint posted**

02 October 2023

**Reviewed preprint posted**

01 November 2023

**Reviewed preprint revised**

09 May 2024

**Version of Record published**

26 July 2024

**Reviewing Editor:** Amit K

Verma, Albert Einstein College of Medicine, United States

© Copyright Yu, Zhou et al. This article is distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use and redistribution provided that the original author and source are credited.

**Abstract** The evolutionary conservation of non-core RAG regions suggests significant roles that might involve quantitative or qualitative alterations in RAG activity. Off-target V(D)J recombination contributes to lymphomagenesis and is exacerbated by RAG2' C-terminus absence in *Tp53*<sup>-/-</sup> mice thymic lymphomas. However, the genomic stability effects of non-core regions from both *Rag1*<sup>cl/c</sup> and *Rag2*<sup>cl/c</sup> in *BCR-ABL1*<sup>+</sup> B-lymphoblastic leukemia (*BCR-ABL1*<sup>+</sup> B-ALL), the characteristics, and mechanisms of non-core regions in suppressing off-target V(D)J recombination remain unclear. Here, we established three mouse models of *BCR-ABL1*<sup>+</sup> B-ALL in mice expressing full-length RAG (*Rag*<sup>fl/fl</sup>), core RAG1 (*Rag1*<sup>cl/c</sup>), and core RAG2 (*Rag2*<sup>cl/c</sup>). The *Rag*<sup>cl/c</sup> (*Rag1*<sup>cl/c</sup> and *Rag2*<sup>cl/c</sup>) leukemia cells exhibited greater malignant tumor characteristics compared to *Rag*<sup>fl/fl</sup> cells. Additionally, *Rag*<sup>cl/c</sup> cells showed higher frequency of off-target V(D)J recombination and oncogenic mutations than *Rag*<sup>fl/fl</sup>. We also revealed decreased RAG cleavage accuracy in *Rag*<sup>cl/c</sup> cells and a smaller recombinant size in *Rag1*<sup>cl/c</sup> cells, which could potentially exacerbate off-target V(D)J recombination in *Rag*<sup>cl/c</sup> cells. In conclusion, these findings indicate that the non-core RAG regions, particularly the non-core region of RAG1, play a significant role in preserving V(D)J recombination precision and genomic stability in *BCR-ABL1*<sup>+</sup> B-ALL.

## eLife assessment

Using a set of animal models, this **valuable** paper shows tumor suppressive function of the non-core regions of RAG1/2 recombinases. The conclusions are supported by **solid** evidence.

## Introduction

V(D)J recombination serves as the central process for early lymphocyte development and generates diversity in antigen receptors. This process involves the double-strand DNA breaks of gene segments by the V(D)J recombinase, including Rag1 and Rag2 (hereinafter referred to as RAG). RAG recognizes conserved recombination signal sequences (RSSs) positioned adjacent to V, D, and J gene segments. A bona fide RSS contains a conserved palindromic heptamer (consensus 5'-CACAGTG) and A-rich

nonamer (consensus 5'-ACAAAACC) separated by a degenerate spacer of either 12 or 23 base pairs (*Hirokawa et al., 2020; Schatz and Ji, 2011a*). The process of efficient recombination is contingent upon the presence of RSSs with differing spacer lengths, as dictated by the '12/23 rule' (*Banerjee and Schatz, 2014; Eastman et al., 1996; Shi et al., 2020*). Following cleavage, the DNA ends are joined via non-homologous end-joining (NHEJ), resulting in the precise alignment of the two coding ends and the signal ends (*Rooney et al., 2004*). V(D)J recombination promotes B-cell development, but aberrant V(D)J recombination can lead to precursor B-cell malignancies through RAG-mediated off-target effects (*Mendes et al., 2014; Onozawa and Aplan, 2012; Thomson et al., 2020*).

The regulation of RAG expression and activity is multifactorial, serving to ensure V(D)J recombination and B-cell development (*Gan et al., 2021; Kumari et al., 2021*). The RAGs consist of core and non-core region. Although non-core regions of Rag1/2 are not strictly required for V(D)J recombination, the evolutionarily conserved non-core RAG regions indicate their potential significance in vivo that may involve quantitative or qualitative modifications in RAG activity and expression (*Braams et al., 2023; Curry and Schlissel, 2008; Liu et al., 2022; Liu et al., 2022; Sekiguchi et al., 2001*). Specifically, the non-core Rag2 region (amino acids 384–527 of 527 residues) contains a plant homeodomain (PHD) that can recognize histone H3K4 trimethylation, as well as a T490 locus that mediates a cell cycle-regulated protein degradation signal in proliferated pre-B-cell stage (*Liu et al., 2007; Matthews et al., 2007*). Failure to degrade Rag2 during the S stage poses a threat to the genome (*Zhang et al., 2011*). Moreover, the off-target V(D)J recombination frequency is significantly higher when Rag2 is C-terminally truncated, thereby establishing a mechanistic connection between the PHD domain, H3K4me3-modified chromatin, and the suppression of off-target V(D)J recombination (*Lu et al., 2015; Mijušković et al., 2015*). The Rag1' non-core region (amino acids 1–383 of 1040 residues) has been identified as a Rag1 regulator. While the core Rag1 maintains its catalytic activity, its in vivo recombination efficiency and fidelity are reduced in comparison to the full-length Rag1 (*Silver et al., 1993*). In addition, the Rag1 binding to the genome is more indiscriminate (*Beilinson et al., 2021; Sadofsky et al., 1993*). The N-terminal domain (NTD), which is evolutionarily conserved, is predicted to contain multiple zinc-binding motifs, including a Really Interesting New Gene (RING) domain (aa 287–351) that can ubiquitylate various targets, including Rag1 itself (*Deng et al., 2015*). Furthermore, NTD contains a specific region (amino acids 1–215) that facilitates interaction with DCAF1, leading to the degradation of Rag1 in a CRL4-dependent manner (*Schabla et al., 2018*). Additionally, the NTD plays a role in chromatin binding and the genomic targeting of the RAG complex (*Schatz and Swanson, 2011b*). Despite increased evidence emphasizing the significance of non-core RAG regions, particularly Rag1's non-core region, the function of non-core RAG regions in off-target V(D)J recombination and the underlying mechanistic basis have not been fully clarified.

Typically, genomic DNA is safeguarded against inappropriate RAG cleavage by the inaccessibility of cryptic RSSs (cRSSs), which are estimated to occur once per 600 base pairs (*Teng et al., 2015*). However, recent research has demonstrated that epigenetic reprogramming in cancer can result in heritable alterations in gene expression, including the accessibility of cRSSs (*Becker et al., 2020; Fatma et al., 2022; Goel et al., 2022; Khoshchereh et al., 2019*). We selected the *BCR-ABL1*<sup>+</sup> B-ALL model, which is characterized by ongoing V(D)J recombinase activity and *BCR-ABL1* gene rearrangement in pre-B leukemic cells (*Schjerven et al., 2017; Wong and Witte, 2004*). The genome structural variations (SVs) analysis was conducted on leukemic cells from *Rag1<sup>fl/fl</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>*, *BCR-ABL1*<sup>+</sup> B-ALL mice to examine the involvement of non-core RAG regions in off-target V(D)J recombination events. The non-core domain deletion in both Rag1 and RAG2 led to accelerated leukemia onset and progression, as well as an increased off-target V(D)J recombination. Our analysis showed a reduction in RAG cleavage accuracy in *RAG<sup>c/c</sup>* cells and a decrease in recombinant size in *RAG1<sup>c/c</sup>* cells, which may be responsible for the increased off-target V(D)J recombination in *RAG<sup>c/c</sup>* leukemia cells. In conclusion, our results highlight the potential importance of the non-core RAG region, particularly RAG1's non-core region, in maintaining accuracy of V(D)J recombination and genomic stability in *BCR-ABL1*<sup>+</sup> B-ALL.

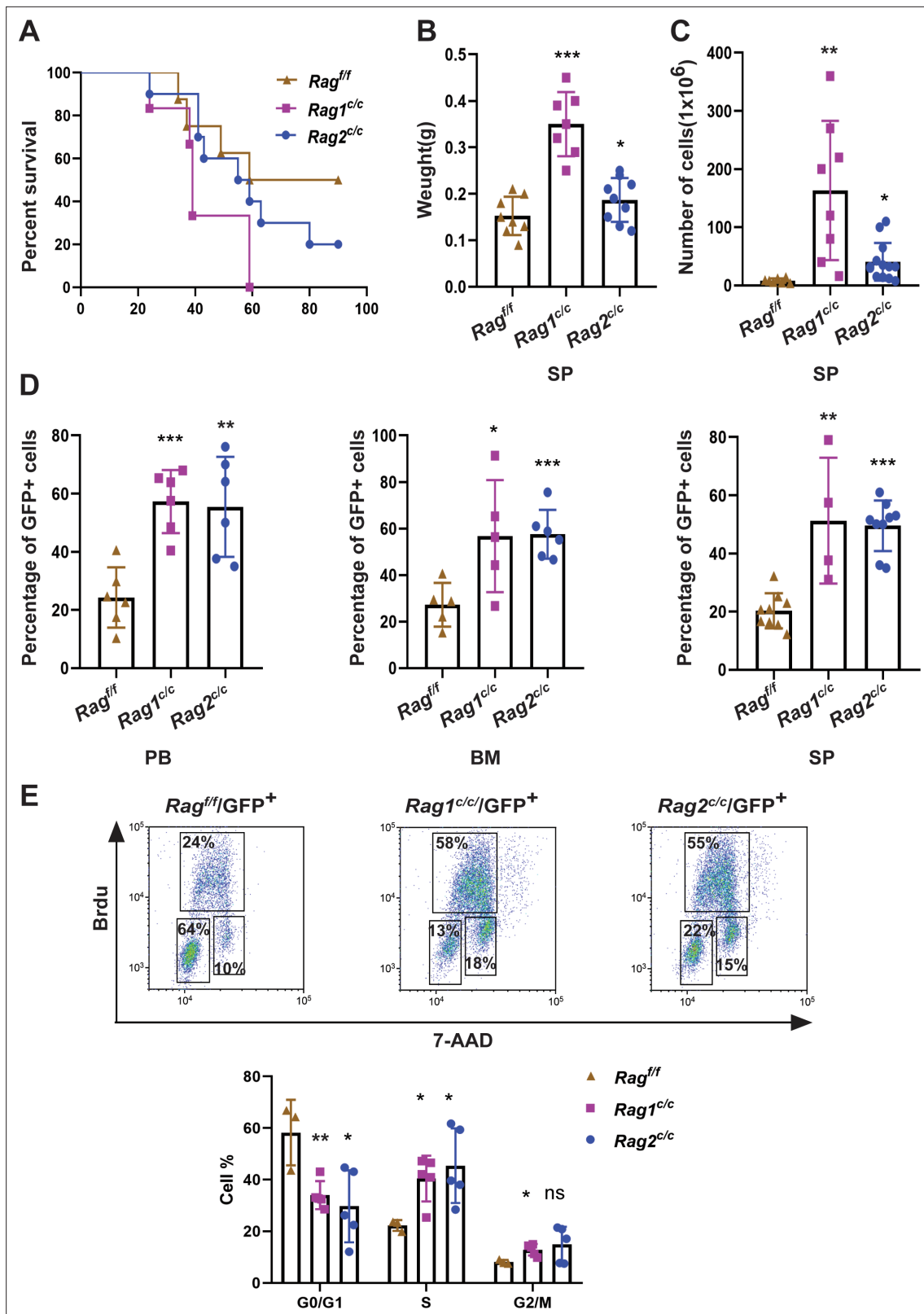
## Results

### **Rag<sup>c/c</sup> give more aggressive leukemia in a mouse model of BCR-ABL1<sup>+</sup> B-ALL**

In order to assess the impact of RAG activity on the clonal evolution of BCR-ABL1<sup>+</sup> B-ALL through a genetic experiment, we utilized bone marrow transplantation to compare disease progression in Rag<sup>fl/fl</sup>, Rag1<sup>c/c</sup>, and Rag2<sup>c/c</sup> BCR-ABL1<sup>+</sup> B-ALL (Schjerven et al., 2017; Wong and Witte, 2004). Bone marrow cells transduced with a BCR-ABL1/green fluorescence protein (GFP) retrovirus were administered into syngeneic lethally irradiated mice, and CD19<sup>+</sup> B-cell leukemia developed within 30–80 days (Figure 1A, Figure 1—figure supplement 1). Western blotting results confirmed equivalent transduction efficiencies of the retroviral BCR-ABL1 in all three cohorts (Figure 1—figure supplement 2A). To investigate potential variances in leukemia outcome across different genomic backgrounds, we employed Mantel–Cox analysis to evaluate survival rates in Rag<sup>fl/fl</sup>, Rag1<sup>c/c</sup>, or Rag2<sup>c/c</sup> mice transplanted with BCR-ABL1-transformed bone marrow cells. Our results show that, during the primary transplant phase, BCR-ABL1<sup>+</sup> B-ALL mice expressing Rag1<sup>c/c</sup> or Rag2<sup>c/c</sup> demonstrated lower survival rates compared to their counterparts with Rag<sup>fl/fl</sup> (median 74.5 days vs 39 or 57 days,  $p < 0.0425$ , Figure 1A). This survival rates discrepancy was also observed during the secondary transplant phase, wherein leukemic cells were extracted from the spleens of primary recipients and subsequently purified via GFP<sup>+</sup> cell sorting. A total of 10<sup>5</sup>, 10<sup>4</sup>, and 10<sup>3</sup> GFP<sup>+</sup> leukemic cells that originated from Rag<sup>fl/fl</sup>, Rag1<sup>c/c</sup>, or Rag2<sup>c/c</sup> leukemic mice were transplanted into corresponding non-irradiated immunocompetent syngeneic recipients (survival days Rag<sup>fl/fl</sup>, 11–26 days, Rag1<sup>c/c</sup>, 10–16 days, Rag2<sup>c/c</sup>, 11–21 days, Figure 1—figure supplement 2B). Additionally, the Rag<sup>c/c</sup> mice exhibited significantly higher leukemia burdens in the peripheral blood, bone marrow, and spleen compared to the Rag<sup>fl/fl</sup> mice (Figure 1B, D). To elucidate the cellular mechanisms driving the accelerated proliferation observed in Rag<sup>c/c</sup> BCR-ABL1<sup>+</sup> B-ALL, flow cytometry analyses were conducted to evaluate cell cycle dynamics and apoptotic activity. Results revealed a higher fraction of Rag<sup>c/c</sup> BCR-ABL1<sup>+</sup> B-ALL cells residing in the S/G2-M phase of the cell cycle compared to their Rag<sup>fl/fl</sup> counterparts (Figure 1E). Additionally, the enhanced proliferation in Rag<sup>c/c</sup> leukemic cells was attributed to a reduction in apoptosis rates (Figure 1—figure supplement 2C). RNA-seq analysis demonstrated the changes of cell differentiation and proliferation/apoptotic pathways (Figure 1—figure supplement 3). These findings indicate that the absence of non-core RAG regions accelerates malignant transformation and leukemic proliferation, leading to a more aggressive disease phenotype in the Rag<sup>c/c</sup> BCR-ABL1<sup>+</sup> B-ALL mouse model.

### **The loss of non-core RAG regions corresponds to a less mature cell surface phenotype but does not impede IgH VDJ recombination**

To delineate the developmental stages of B cells from which the leukemic cells originated, we performed flow cytometry on single cells stained with B-cell-specific surface markers. Analysis revealed that 91–98% of GFP<sup>+</sup> cells in Rag<sup>c/c</sup> mice were CD19<sup>+</sup>BP-1<sup>+</sup>B220<sup>+</sup>CD43<sup>+</sup>, indicating that most leukemic cells were at the large pre-B-cell stage (Figure 2A; Hardy and Hayakawa, 2001). Conversely, in Rag<sup>fl/fl</sup> leukemic mice, the distribution was 65% large pre-B cells (GFP<sup>+</sup>CD19<sup>+</sup>BP-1<sup>+</sup>B220<sup>+</sup>CD43<sup>+</sup>) and 35% small pre-B cells (GFP<sup>+</sup>CD19<sup>+</sup>BP-1<sup>+</sup>B220<sup>+</sup>CD43<sup>-</sup>) (Figure 2A). Moreover, approximately 5% of leukemic cells in Rag<sup>fl/fl</sup> mice expressed  $\mu$ HC, in contrast to minimal expression in Rag<sup>c/c</sup> leukemic cells. This suggests that Rag<sup>fl/fl</sup> leukemic cells may differentiate further, associated with an immune phenotype (Figure 2B). IgH rearrangement initiates with D<sub>H</sub>–J<sub>H</sub> joining in pro-B cells, followed by V<sub>H</sub>–DJ<sub>H</sub> joining in pro-B cells, and ultimately, V<sub>L</sub>–J<sub>L</sub> rearrangements occur at the IgL loci in small pre-B cells (Schatz and Ji, 2011a). Genomic polymerase chain reaction analysis of DNA from GFP<sup>+</sup>CD19<sup>+</sup> cells was utilized to assess V<sub>H</sub>DJ<sub>H</sub> rearrangement. The results showed a pronounced oligoclonality in Rag<sup>c/c</sup> leukemic cells, with tumors consistently demonstrating rearrangements involving a restricted set of V<sub>H</sub> family members. In contrast, Rag<sup>fl/fl</sup> leukemias displayed significant polyclonality, evidenced by the widespread rearrangement of various V<sub>H</sub> family members to all potential J<sub>H</sub>1–3 segments, indicative of a broader clonal diversity (Figure 2—figure supplement 1). This observation aligns with the more aggressive leukemia phenotype seen in Rag<sup>c/c</sup> BCR-ABL1<sup>+</sup> B-ALL mice. Such oligoclonality in Rag<sup>c/c</sup> leukemic cells suggests a selection process driven by BCR-ABL1-induced leukemia, favoring the emergence of a limited number of dominant leukemic clones. The absence of non-core RAG regions appears to restrict the diversity of leukemic clones, leading to the formation of oligoclonal tumors.



**Figure 1.** RAG<sup>c/c</sup>s give more aggressive leukemia in mice model of BCR-ABL1<sup>+</sup> B-ALL. (A) Kaplan–Meier survival curve for *Rag<sup>fl/fl</sup>* (n = 8), *Rag1<sup>c/c</sup>* (n = 6), and *Rag2<sup>c/c</sup>* (n = 10) recipient mice. The survival was calculated by Mantel–Cox test (*Rag<sup>fl/fl</sup>* vs *Rag1<sup>c/c</sup>*, p < 0.0381, *Rag<sup>fl/fl</sup>* vs *Rag2<sup>c/c</sup>*, p < 0.0412). (B) The spleen weights of *Rag<sup>fl/fl</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>* leukemic mice (*Rag<sup>fl/fl</sup>*, n = 8, *Rag1<sup>c/c</sup>*, n = 7, *Rag2<sup>c/c</sup>*, n = 9; *Rag<sup>fl/fl</sup>* vs *Rag1<sup>c/c</sup>*, p < 0.0001, *Rag<sup>fl/fl</sup>* vs *Rag2<sup>c/c</sup>*, p = 0.1352). (C) The numbers of spleen cell in *Rag<sup>fl/fl</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>* leukemic mice (*Rag<sup>fl/fl</sup>*, n = 7, *Rag1<sup>c/c</sup>*, n = 8, *Rag2<sup>c/c</sup>*, n = 13; *Rag<sup>fl/fl</sup>* vs *Rag1<sup>c/c</sup>*, p =

Figure 1 continued on next page

Figure 1 continued

0.0047,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0180$ ). (D) The percentage of GFP<sup>+</sup> cells in peripheral blood (PB) ( $Rag^{ff}$ ,  $n = 6$ ,  $Rag1^{c/c}$ ,  $n = 6$ ,  $Rag2^{c/c}$ ,  $n = 6$ ;  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0003$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0035$ ), bone marrow (BM,  $Rag^{ff}$ ,  $n = 5$ ,  $Rag1^{c/c}$ ,  $n = 5$ ,  $Rag2^{c/c}$ ,  $n = 6$ ;  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0341$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0008$ ), and spleen (SP,  $Rag^{ff}$ ,  $n = 9$ ,  $Rag1^{c/c}$ ,  $n = 4$ , cRAG2,  $n = 9$ ;  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0016$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p < 0.0001$ ) of  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  leukemic mice. (E) Representative flow cytometry plots of cell cycle arrest of leukemic cells in  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  mice. In the graph, the percentages of each phase of the cell cycle are summarized below ( $Rag^{ff}$ ,  $n = 3$ ,  $Rag1^{c/c}$ ,  $n = 5$ ,  $Rag2^{c/c}$ ,  $n = 5$ ; G0/G1,  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0082$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0279$ ; S,  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0146$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0370$ ; G2/M,  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0134$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.1507$ ). In figures B, C, D, and J, error bars represent the mean  $\pm$  standard deviation (s.d.),  $p$  values were calculated by Student's  $t$  test and \* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\* $p < 0.001$ .

The online version of this article includes the following source data and figure supplement(s) for figure 1:

**Figure supplement 1.** Construction of  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$ , BCR-ABL1<sup>+</sup> B-ALL mice models using bone marrow transplantation.

**Figure supplement 2.** Biological behavior of leukemia in  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  BCR-ABL1<sup>+</sup> B-ALL mouse model.

**Figure supplement 2—source data 1.** Original file for the western blot analysis in **Figure 1—figure supplement 2A**.

**Figure supplement 3.** The genetic pathways in  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  BCR-ABL1<sup>+</sup> lymphocytes.

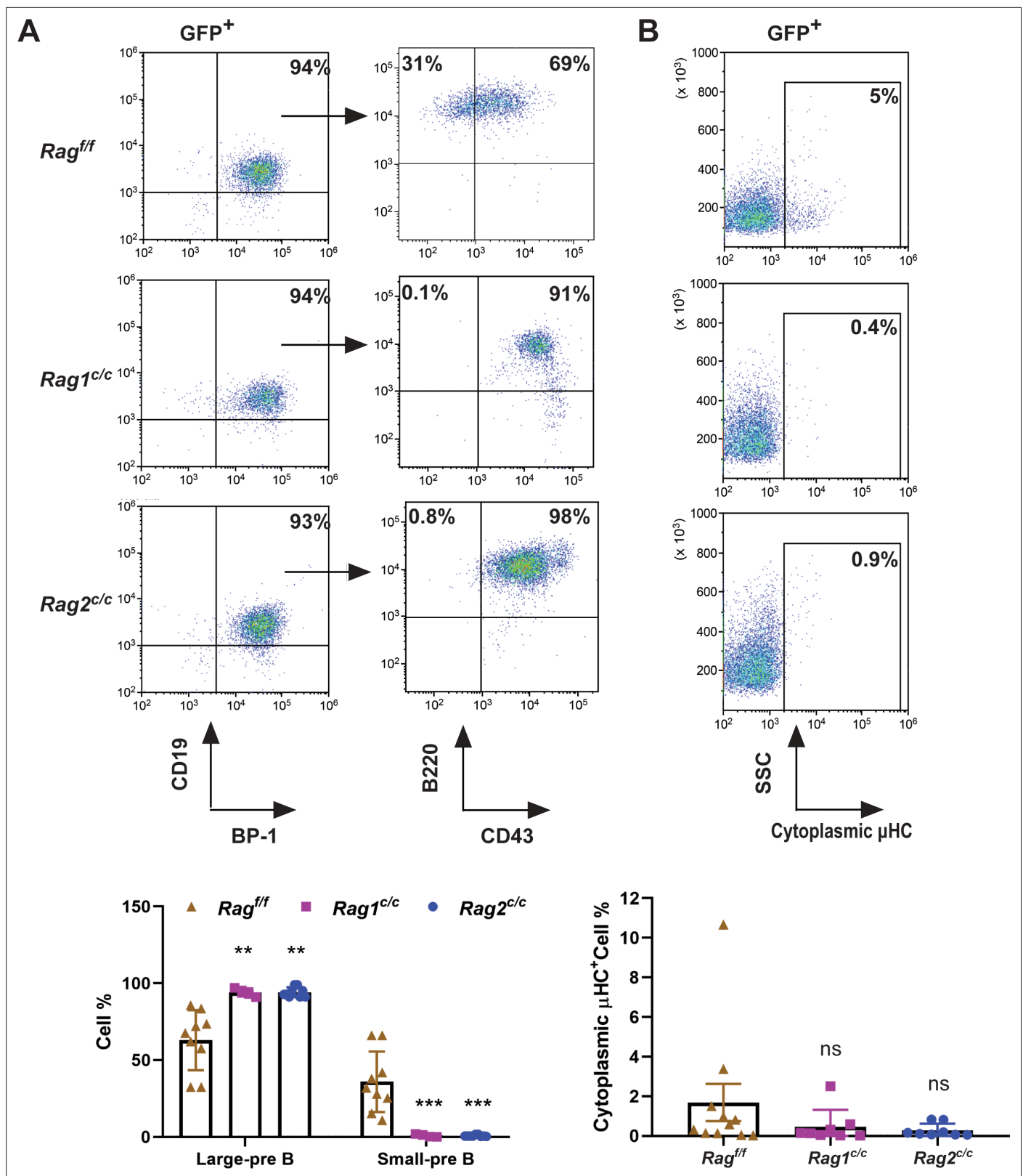
## The loss of non-core RAG regions highlights genomic DNA damage

The findings indicate that leukemic cells from three types of mice exhibited variable arrests at the large pre-B-cell stage, deviating from normal B-cell developmental trajectory. Typically, at this juncture, B cells initiate the degradation of Rag2 via the cyclin-dependent kinase cyclinA/Cdk2, leading to a downregulation of RAG activity. It is therefore crucial to explore the impact of deletions in non-core regions on the expression and functionality of RAG in these leukemic cells. Western blotting analysis showed that expression of Rag1<sup>ff</sup>, Rag1<sup>c/c</sup>, Rag2<sup>ff</sup>, and Rag2<sup>c/c</sup> proteins were present in GFP<sup>+</sup>CD19<sup>+</sup> splenic leukemic cells with the indicated Rag1 and Rag2 genotypes (**Figure 3A**). Notably, we observed an upregulation expression of Rag1<sup>c/c</sup> and Rag2<sup>c/c</sup> in leukemic cells from  $Rag1^{c/c}$  or  $Rag2^{c/c}$  mice compared to those from  $Rag^{ff}$  mice (**Figure 3A**, **Figure 3—figure supplement 1A**). The in vitro V(D)J recombination assay confirmed that different forms of RAG exhibited cleavage activity in BCR-ABL1<sup>+</sup> B-ALL (**Figure 3B**, **Figure 3—figure supplement 1B**). Moreover, we analyzed pediatric acute lymphoid leukemia patients from the TARGET cohort and found that high RAG expression was related to low survival in this disease (**Figure 3—figure supplement 2**).

To examine the potential correlation between aberrant RAG activity and increased DNA double-strand breaks (DSBs), we assessed levels of phosphorylated H2AX ( $\gamma$ -H2AX), a marker of DSB response, in leukemic cells from  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  mice (gated on GFP<sup>+</sup> cells). This evaluation aimed to gauge DNA DSBs and overall genomic instability. Flow cytometry analysis revealed elevated  $\gamma$ -H2AX levels in  $Rag1^{c/c}$  and  $Rag2^{c/c}$  leukemic cells compared to those from  $Rag^{ff}$  (**Figure 3C**), indicating a more pronounced role of Rag<sup>c/c</sup> in mediating somatic structural variants (SVs) in BCR-ABL1<sup>+</sup> B cells. These findings suggest enhanced expression of Rag1<sup>c/c</sup> endonucleases in  $Rag1^{c/c}$  leukemic cells and increased DNA damage in cells lacking core RAG regions.

## Off-target recombination mediated by RAG in BCR-ABL1<sup>+</sup> B cells

Genome-wide sequencing and analysis were performed to compare somatic SVs in BCR-ABL1<sup>+</sup> B cells derived from  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  mice. The leukemic cells were sequenced with an average coverage of 25 $\times$  (**Supplementary file 1**). The SVs generated by RAG were screened based on two criteria: the presence of a CAC to the right (or GTG to the left) of both breakpoints, and its occurrence within 21 bp from the breakpoint (**Mijušković et al., 2015**). Further elaboration on these criteria can be found in **Figure 4—figure supplement 1**. Consequently, aberrant V-to-V junctions and V to intergenic regions were encompassed in five validated abnormal rearrangements at *Ig* loci in  $Rag^{c/c}$  leukemic mice (**Supplementary file 2**). Additionally, seven samples had 24 somatic SVs, with an average of 3.4 coding region mutations per sample (range of 0–9), which is consistent with the limited number of acquired somatic mutations observed in hematological cancers. The results of the study demonstrate that  $Rag^{ff}$  cells had low SVs (0–1 per sample),  $Rag1^{c/c}$  cells exhibited higher SVs (6–9 per sample) while  $Rag2^{c/c}$  cells had moderate SVs incidence (1–4 per sample) (**Figure 4** and **Supplementary file 3**). These findings suggest that Rag<sup>c/c</sup> may lead to an elevated off-target recombination, eventually posing a threat to the BCR-ABL1<sup>+</sup> B-cell genome.



**Figure 2.** The non-core RAG region loss corresponds to a less mature cell surface phenotype. (A) Flow cytometry analysis of the B-cell markers CD19, BP-1, B220, and CD43 on BCR-ABL1-transformed *Rag1<sup>ff</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>* leukemic bone marrow cells. The percentages of each phase of the B-cell stage are summarized in the bottom graph (*Rag1<sup>ff</sup>*, *n* = 9, *Rag1<sup>c/c</sup>*, *n* = 4, *Rag2<sup>c/c</sup>*, *n* = 9; Large-preB, *Rag1<sup>ff</sup>*, vs *Rag1<sup>c/c</sup>*, *p* = 0.0349, *Rag1<sup>ff</sup>*, vs *Rag2<sup>c/c</sup>*, *p* = 0.0017; Small-pre-B, *Rag1<sup>ff</sup>*, vs *Rag1<sup>c/c</sup>*, *p* = 0.0141, *Rag1<sup>ff</sup>*, vs *Rag2<sup>c/c</sup>*, *p* = 0.0005). The expression of the cytoplasmic  $\mu$  chain was analyzed by flow

Figure 2 continued on next page

Figure 2 continued

cytometry. Representative samples are shown in (B), and the results from multiple samples analyzed in independent experiments are summarized in the bottom graph as the fraction of cells expressing cytoplasmic factors ( $Rag^{ff}$ ,  $n = 11$ ,  $Rag1^{c/c}$ ,  $n = 9$ ,  $Rag2^{c/c}$ ,  $n = 8$ ;  $Rag^{ff}$  vs  $Rag1^{c/c}$ ,  $p = 0.0312$ ,  $Rag^{ff}$  vs  $Rag2^{c/c}$ ,  $p = 0.0441$ ). Error bars represent the mean  $\pm$  standard deviation (s.d.),  $p$  values were calculated by Student's  $t$  test and  $**p < 0.01$ ,  $***p < 0.001$ .

The online version of this article includes the following source data and figure supplement(s) for figure 2:

**Figure supplement 1.** VDJ recombination in leukemic cells with different genetic backgrounds.

**Figure supplement 1—source data 1.** Original file for the PCR analysis in **Figure 2—figure supplement 1A**.

**Figure supplement 1—source data 2.** Original file for the PCR analysis in **Figure 2—figure supplement 1B**.

## Off-target V(D)J recombination characteristics in $BCR-ABL1^+$ B cells

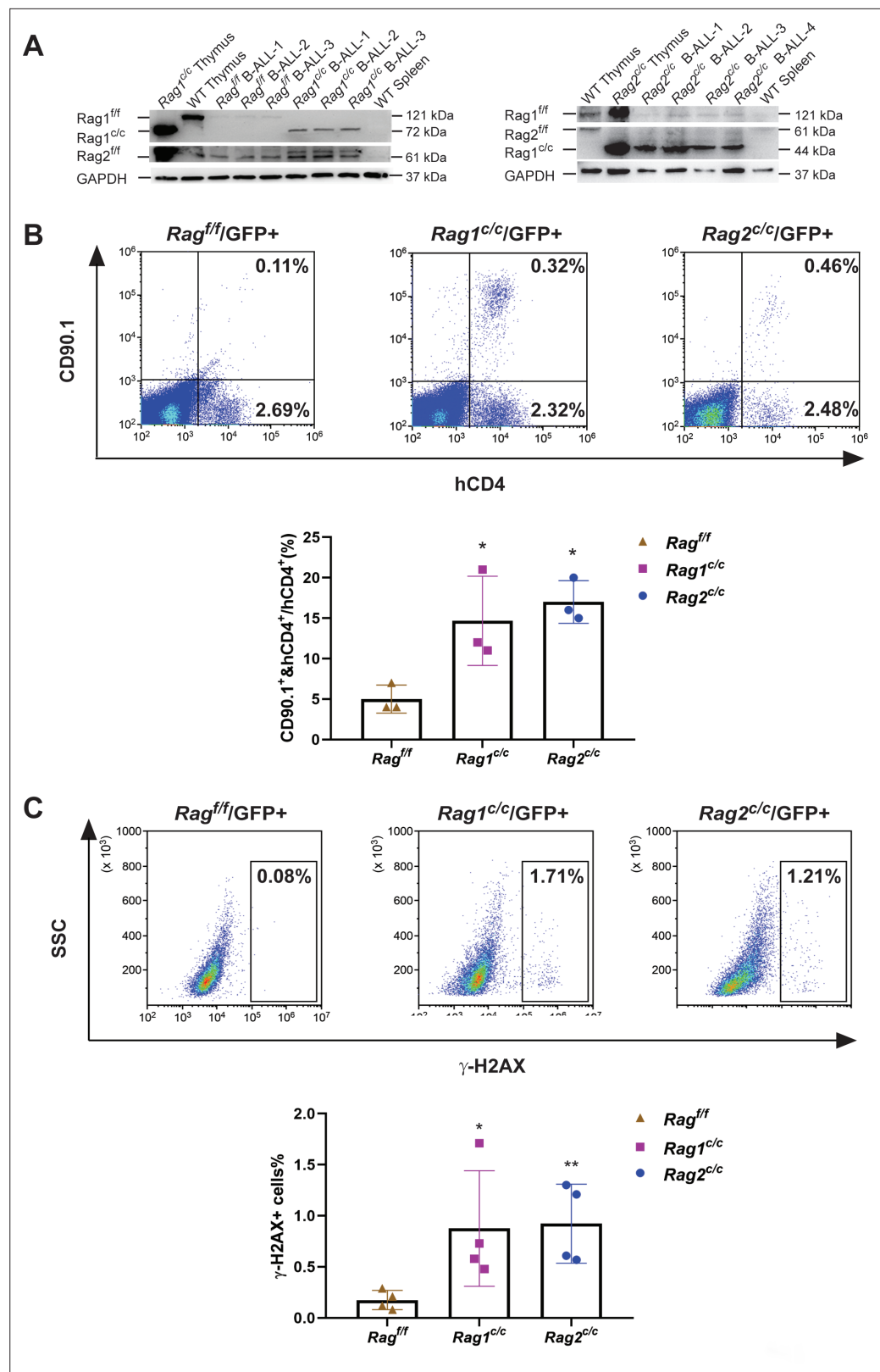
We further examined the characteristics of the identified SVs. Specifically, we analyzed the exon–intron distribution profiles of 41 breakpoints from 24 SVs through genome analysis. The results indicated that 57% of the breakpoints were located within the gene body, while 43% were enriched in the flanking sequences, the majority of which were identified as transcriptional regulatory sequences (**Figure 5A**). P and N nucleotides are recognized as distinctive characteristics of V(D)J recombination (**Repasky et al., 2004**). RSS-to-RSS and cRSS-to-cRSS recombination have P nucleotide lengths of 7 and 9, respectively, and N lengths of 5, so nucleotide lengths are basically the same during RSS-to-RSS and cRSS-to-cRSS recombination (**Figure 5B**). However, the frequency of P and N sequences in RSS-to-RSS recombination was 50%/50% (P/N), compared to 4%/8% (P/N) in cRSS-to-cRSS recombination (**Figure 5B**). This significant reduction in the frequency of P and N sequences suggests that DNA repair at off-target sites in  $BCR-ABL1^+$  B cells diverges from the classical V(D)J recombination repair process.

The hybrid joints were notably prevalent in  $Rag1^{c/c}$  and  $Rag2^{c/c}$  leukemic cells (93% and 100%, respectively), suggesting that the non-core regions of RAG may play a role in inhibiting harmful transposition events (**Figure 5C**). To evaluate the effect of deleting non-core RAG regions on the emergence of oncogenic mutations, we performed a comparative analysis of cancer-related genes across three types of leukemic cells. We found that  $Rag1^{c/c}$  leukemic cells harbored a significantly higher number of cancer genes compared to the other groups. This finding corresponds with the most aggressive leukemia phenotype observed in  $Rag1^{c/c}$   $BCR-ABL1^+$  B-ALL mice and associated changes in their transcription profiles.

## The non-core regions have effects on RAG cleavage and off-target recombination size in $BCR-ABL1^+$ B cells

Sequence logos were employed to visually contrast RSS and cRSSs within *Ig* and non-*Ig* loci, respectively. Notably, the RSS elements in *Ig* loci displayed a higher similarity to the canonical RSS, especially at critical functional sites. The initial four nucleotides (highlighted) of the canonical heptamer sequence CACAGTG were recognized as the cleavage site for  $Rag^{ff}$ . Conversely, in leukemic cells, the cleavage site for  $Rag^{c/c}$  was pinpointed to the first three nucleotides, the CAC trinucleotide, of the heptamer sequence (**Figure 6A**). While both motifs (CAC and CACA) align with the highly conserved segment of the RSS heptamer sequence, differences in the cRSS sequences across off-target genes in both  $Rag^{ff}$  and  $Rag^{c/c}$  mice suggest that deletion of RAG's non-core regions broadens the spectrum of off-target substrates in  $BCR-ABL1^+$  B cells.

Antigen receptor genes are assembled by large-scale deletions and inversions (**Teng et al., 2015**). The off-target recombination size was determined as the DNA fragment size spanning the two breakpoints. Our analysis demonstrated that both  $Rag^{ff}$  and  $Rag2^{c/c}$  leukemic cells produced off-target recombinations with 100% and 92% of events, respectively, spanning over 10,000 bp in length. In contrast,  $Rag1^{c/c}$  leukemic cells showed only 6% of off-target recombinations exceeding 10,000 bp, with 48% under 1000 bp and 46% ranging between 1000 and 10,000 bp (**Figure 6B, C**). These findings suggest that the  $Rag1^{c/c}$  variant primarily facilitates smaller-scale off-target recombinations in  $BCR-ABL1^+$  B cells, highlighting the role of the non-core RAG1 region in influencing the extent of off-target recombination. The deletion of the non-core RAG1 region appears to constrict the size of off-target recombination, potentially contributing to the elevated frequency of off-target V(D)J recombination observed in  $Rag1^{c/c}$  leukemic cells (**Figure 6D**).



**Figure 3.** The non-core RAG region loss highlights genomic DNA damage. (A) Western blotting analysis showed RAG1 and RAG2 expression in GFP<sup>+</sup>CD19<sup>+</sup> leukemic cells originating from BCR-ABL1<sup>+</sup> B-ALL in different genetic backgrounds. The experiment was repeated under the same conditions three times. (B) Rearrangement substrate retrovirus was transduced into leukemic cells. Flow cytometry was used to analyze the percentage of CD90.1- and hCD4<sup>+</sup> cells. (C) Flow cytometry was used to analyze the percentage of  $\gamma$ -H2AX<sup>+</sup> cells. \*p < 0.05, \*\*p < 0.01. Figure 3 continued on next page



Figure 3 continued

hCD4-positive cells, and the percentage populations are shown in the bottom graph ( $Rag^{ff}$ ,  $n = 3$ ,  $Rag1^{c/c}$ ,  $n = 3$ ,  $Rag2^{c/c}$ ,  $n = 3$ ;  $Rag^{ff}$ , vs  $Rag1^{c/c}$ ,  $p = 0.0002$ ,  $Rag^{ff}$ , vs  $Rag2^{c/c}$ ,  $p = 0.5865$ ). (C) Flow cytometry analysis of  $\gamma$ -H2AX levels in  $Rag^{ff}$ ,  $Rag1^{c/c}$ , and  $Rag2^{c/c}$  leukemic cells and the percentage of  $\gamma$ -H2AX-positive cell populations shown in the bottom graph ( $Rag^{ff}$ ,  $n = 11$ ,  $Rag1^{c/c}$ ,  $n = 8$ ,  $Rag2^{c/c}$ ,  $n = 8$ ;  $Rag^{ff}$ , vs  $Rag1^{c/c}$ ,  $p = 0.0505$ ,  $Rag^{ff}$ , vs  $Rag2^{c/c}$ ,  $p = 0.0094$ ). Error bars represent the mean  $\pm$  standard deviation (s.d.),  $p$  values were calculated by Student's  $t$  test and  $*p < 0.05$ ,  $**p < 0.01$ .

The online version of this article includes the following source data and figure supplement(s) for figure 3:

**Source data 1.** Original file for the western blot analysis in **Figure 3A**.

**Figure supplement 1.** RAG protein expression levels and schematic diagram of the recombinant substrate vector.

**Figure supplement 2.** The relationship of RAG1 mRNA levels and survival of pediatric acute lymphoid leukemia.

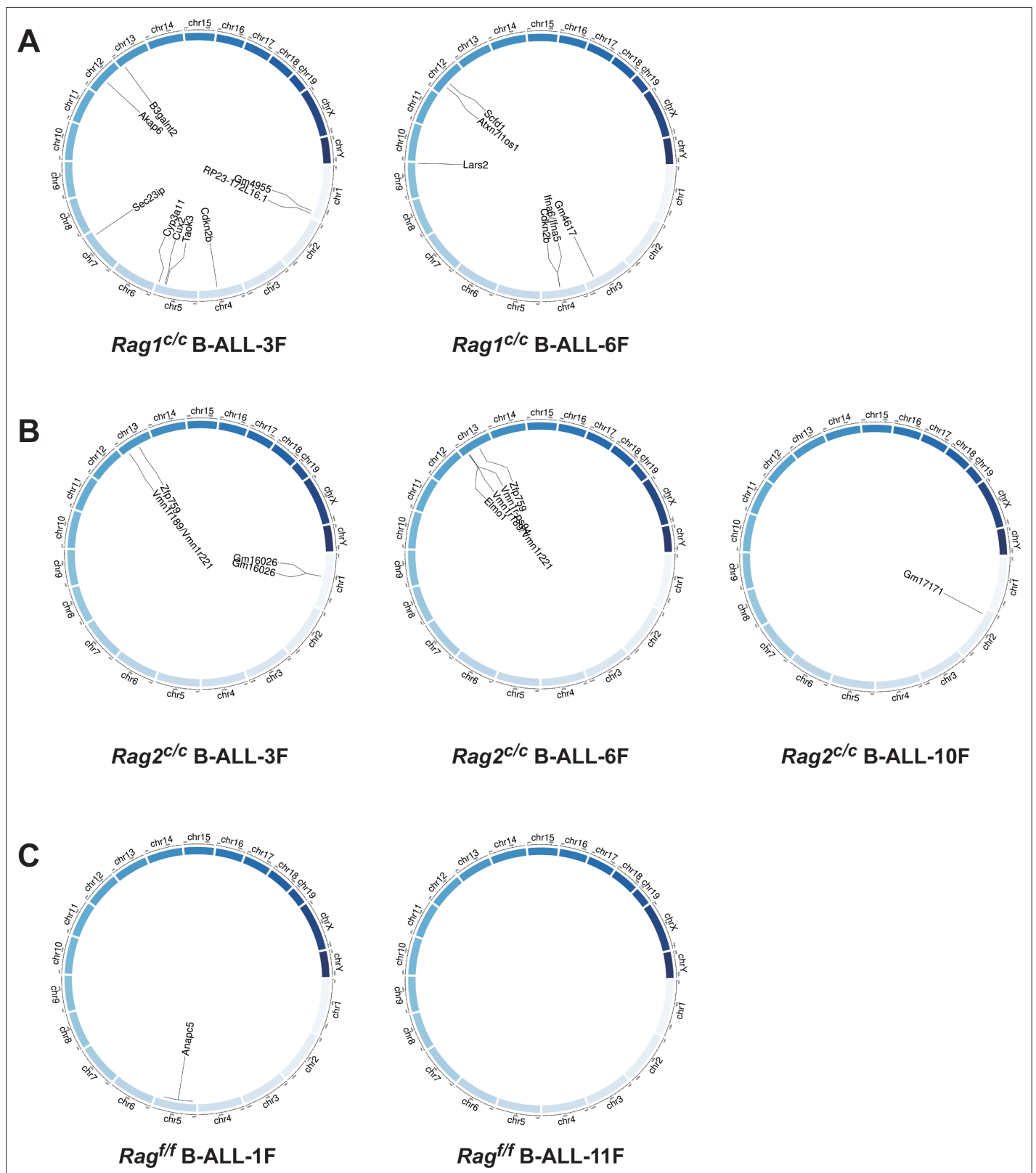
## Discussion

In this study, we have demonstrated that non-core region deletion of both Rag1 and Rag2 leads to accelerated development of leukemia and increased off-target V(D)J recombination in mouse models of  $BCR-ABL1^+$  B cells. Furthermore, we report reduced RAG<sup>c/c</sup> cleavage accuracy and off-target recombination size in  $Rag1^{c/c}$  and  $Rag2^{c/c}$  leukemia cells, which might contribute to exacerbated off-target V(D)J recombination. These findings suggest that the non-core regions, particularly the non-core region of Rag1, play a crucial role in maintaining accuracy of V(D)J recombination and genomic stability in  $BCR-ABL1^+$  B cells.

Our findings suggest that leukemic cells with RAG<sup>c/c</sup> regions exhibit increased production of hybrid joints, implying that non-core RAG regions might suppress the formation of these hybrid joints in vivo. Post-cleavage synaptic complexes (PSCs), comprising RAG proteins, coding ends, and RSS ends, are believed to have evolved to form with optimal conformation and/or stability for conventional coding and RSS end-joining (Fugmann et al., 2000; Libri et al., 2021). In contrast, RAG<sup>c/c</sup> PSCs could promote RAG-mediated hybrid joints by facilitating closer proximity of coding and RSS ends or by increasing PSC stability. It is also conceivable that RAG recruit disassembly/remodeling factors to PSCs, a process that could allow NHEJ factors to complete the normal reaction (Fugmann et al., 2000). Conversely, RAG<sup>c/c</sup> may have a diminished recruitment capacity due to changes in overall conformation or the absence of specific motifs, leading to more unstable PSCs and a heightened risk of accumulating incomplete hybrid joints (Raghavan et al., 2006; Talukder et al., 2004). Our data reveal that over 90% of junctions were hybrid joints in  $Rag^{c/c}$  leukemic cells, a frequency exceeding that reported in previous studies. This suggests that deficiencies in the NHEJ pathway could contribute to chromosomal instability and lymphomagenesis (Gaymes et al., 2002; Rassool, 2003; Scully et al., 2019; Wiegmanns et al., 2021). Significantly, our analysis uncovered variations in the NHEJ repair pathway among leukemic cells from different genetic backgrounds, suggesting a potential aberrant expression of DNA repair pathways in  $Rag^{c/c}$  leukemic cells (Figure 1—figure supplement 3B). These findings highlight the potential of RAG<sup>c/c</sup> to foster increased hybrid joint formation, especially when a normal pathway for efficient coding and RSS joining is compromised in an NHEJ-aberrant context.

Our data demonstrate that the deletion of the RAG1 non-core region results in more severe off-target V(D)J recombination compared to the deletion of the RAG2 non-core region. This observation is supported by the fact that the RAG1 terminus contains multiple zinc-binding motifs and ubiquitin ligase activity, which are known to enhance the efficiency of the rearrangement reaction (Beilinson et al., 2021; Burn et al., 2022). Furthermore, our research reveals that Rag1 expression persists in  $BCR-ABL1^+$  progenitor B cells, and deletion of the non-core region of Rag1 results in elevated expression of RAG in comparison to  $Rag^{ff}$ . Consequently, as demonstrated in this study,  $Rag1^{c/c}$  from  $BCR-ABL1^+$  B leukemic mice is prone to generating off-target V(D)J recombination. The distinct function of Rag1's non-core region in thymic lymphomas of  $Tp53^{-/-}$  mice and  $BCR-ABL1^+$  B leukemic mice leads to dissimilar off-target activity of  $Rag1^{c/c}$  (Mijušković et al., 2015). Therefore, it would be intriguing to replicate these analyses across various subtypes of ALL to further investigate this phenomenon.

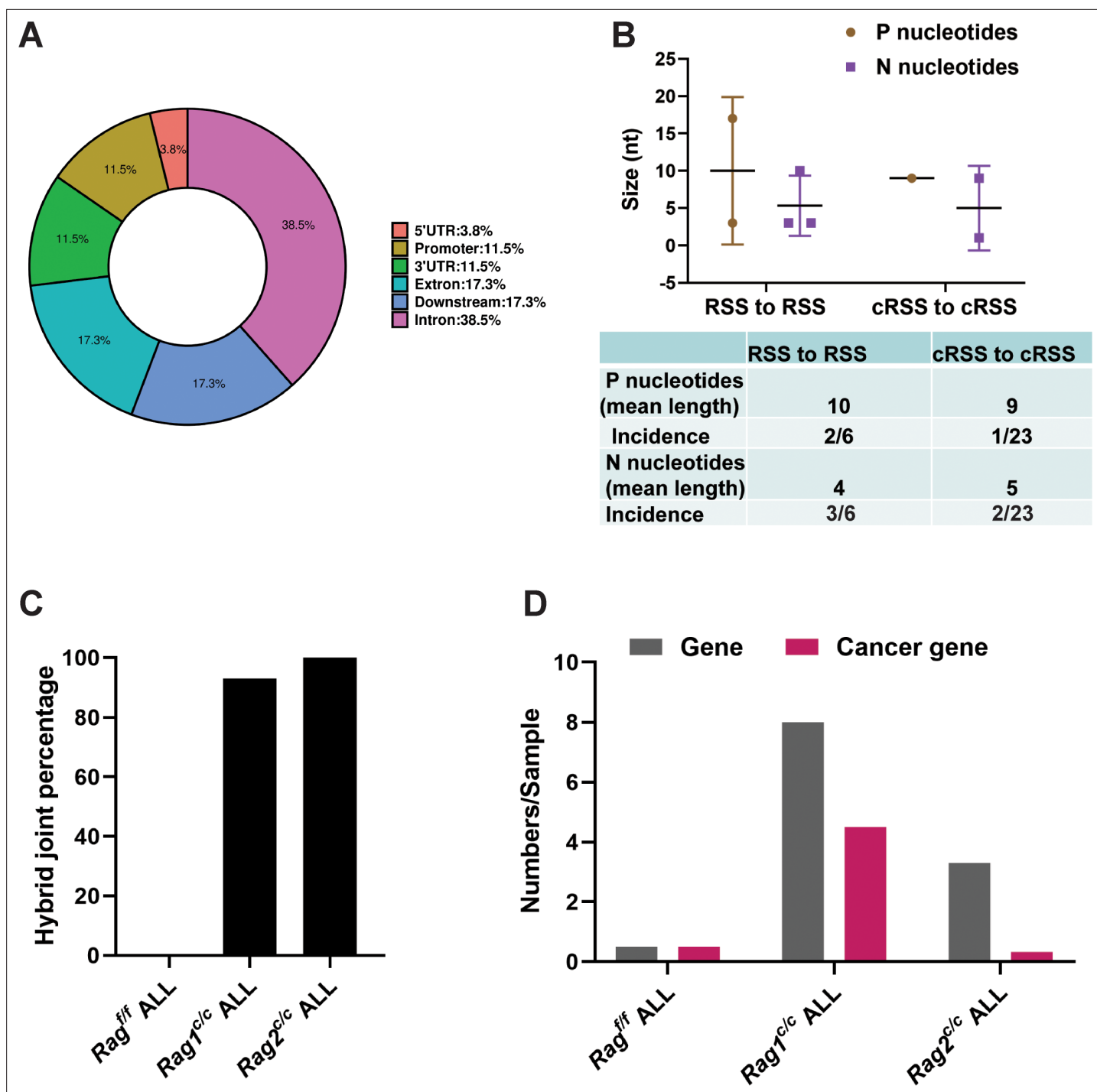
In human *ETV6-RUNX1* ALL, the *ETV6-RUNX1* fusion gene is believed to initiate prenatally, yet the disease remains clinically latent until critical secondary events occur, leading to leukemic transformation – 'pre-leukemia to leukemia' (Mijušković et al., 2015; Bateman et al., 2010). Genomic rearrangement, mediated by aberrant RAG recombinase activity, is a frequent driver of these secondary



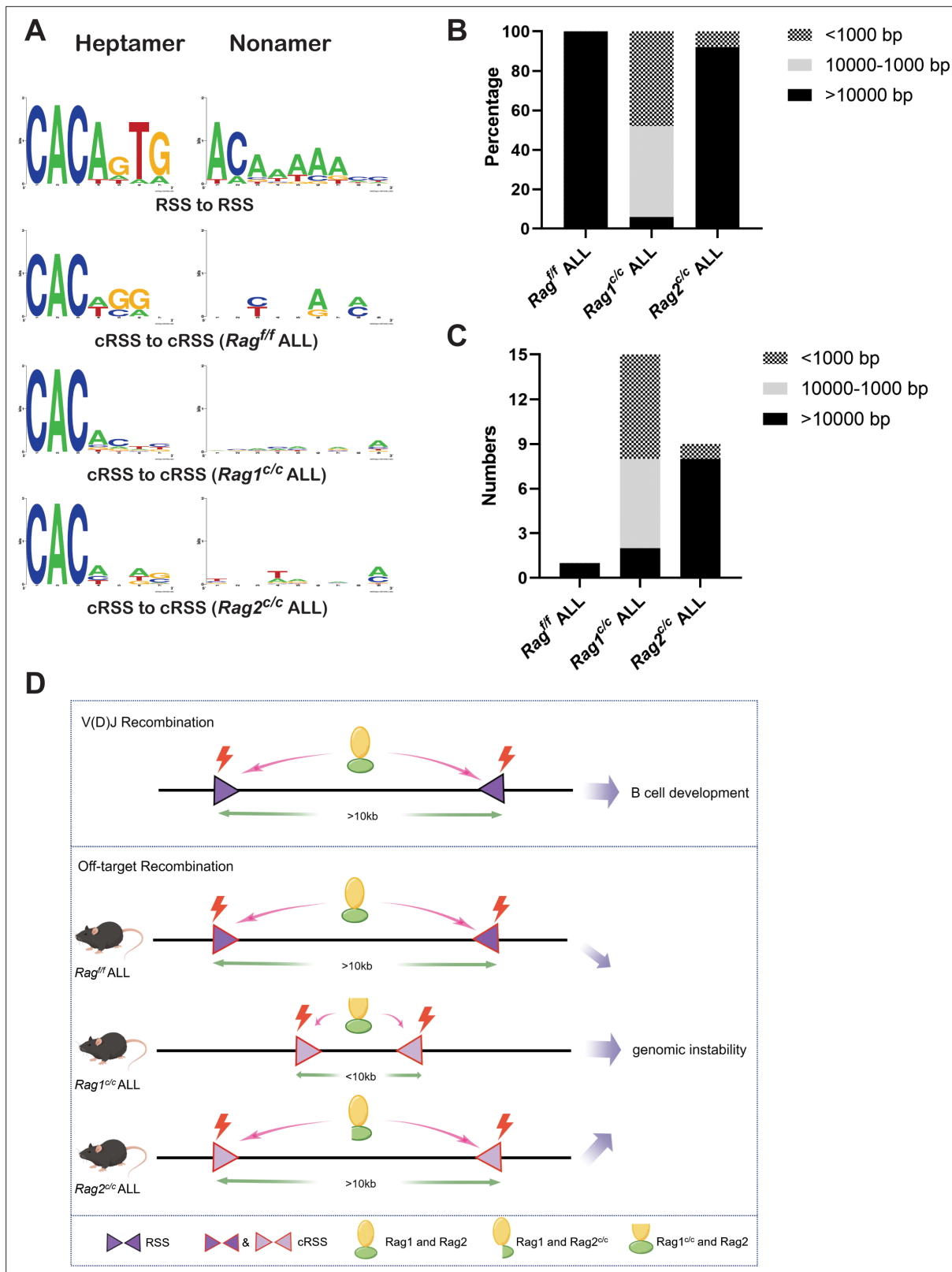
**Figure 4.** Structural alterations in *BCR-ABL1*<sup>+</sup> B lymphocytes. (A–C) Circos plot representation of all off-target recombination detected in the genome-wide analyses of *Rag<sup>f/f</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>* leukemic cells.

The online version of this article includes the following figure supplement(s) for figure 4:

**Figure supplement 1.** The criteria for identifying off-target recombination.



**Figure 5.** Overview and characteristics of off-target recombination in *BCR-ABL1*<sup>+</sup> B-ALL leukemic cells from *Rag*<sup>fl/fl</sup> and *Rag*<sup>c/c</sup> mice. (A) Exon–intron distribution profiles of 41 breakpoints generated by 24 structural variations (SVs). Gene body includes exon ( $n = 9$ ; 17.3%) and intron ( $n = 20$ ; 38.5%). Flanking sequence includes 3'UTR ( $n = 6$ ; 11.5%), 5'UTR ( $n = 2$ ; 3.8%), promoter ( $n = 6$ ; 11.5%), and downstream ( $n = 9$ ; 17.3%). (B) The off-target recombination was filtered and verified by whole genomic sequence and PCR, respectively. P and N nucleotides of recombination signal sequence (RSS) to RSS and cryptic RSS (cRSS) to cRSS were calculated in *BCR-ABL1*<sup>+</sup> B-ALL. (C) Hybrid joint percentage generated by either *Rag*<sup>fl/fl</sup>, *Rag1*<sup>c/c</sup>, or *Rag2*<sup>c/c</sup> in *BCR-ABL1*<sup>+</sup> B-ALL. It was 0, 100%, and 93% in *Rag*<sup>fl/fl</sup>, *Rag1*<sup>c/c</sup>, or *Rag2*<sup>c/c</sup> leukemic cells, respectively. (D) The 24 off-target recombination genes were retrieved by COSMIC Cancer Gene Census (<http://cancer.sanger.ac.uk/census/>). 0.5 genes and 0.5 cancer genes average sample in *Rag*<sup>fl/fl</sup> leukemic cells; 8 genes and 4.5 cancer genes average sample in *Rag1*<sup>c/c</sup> leukemic cells; 3.3 genes and 0.3 cancer genes average sample in *Rag2*<sup>c/c</sup> leukemic cells.



**Figure 6.** The non-core regions have effects on RAG-binding accuracy and recombinant size in *BCR-ABL1<sup>+</sup>* B lymphocyte. **(A)** Sequence logos were used to compare the recombination signal sequence (RSS) and cryptic RSS (cRSS) in Ig loci and non-Ig loci. Top panel: V(D)J recombination at Ig locus; the next three panels: RAG-mediated off-target recombination at non-Ig locus from *Rag1<sup>ff</sup>*, *Rag1<sup>c/c</sup>*, and *Rag2<sup>c/c</sup>* leukemic cells, respectively. The scale of recombinant size was categorized into three ranges: <1000, 1000–10,000, and >10,000 bp. The distribution of different recombinant sizes in *Rag1<sup>ff</sup>*,  
 Figure 6 continued on next page

Figure 6 continued

*Rag1<sup>cl</sup>*, and *Rag2<sup>cl</sup>* leukemic cells was presented in (B), while the number of different recombinant sizes in *Rag<sup>fl</sup>*, *Rag1<sup>cl</sup>*, and *Rag2<sup>cl</sup>* leukemic cells is displayed in (C). (D) A schematic depiction of the mechanism of cRAG-accelerated off-target V(D)J recombination was provided. Both RAG1 and RAG2's non-core region deletion decreases RAG-binding accuracy in *Rag1<sup>cl</sup>* and *Rag2<sup>cl</sup>*, *BCR-ABL1<sup>+</sup>* B-ALL. Additionally, RAG1's non-core region deletion significantly reduces the size and scale of off-target V(D)J recombination in *Rag1<sup>cl</sup>* *BCR-ABL1<sup>+</sup>* B-ALL.

events in *ETV6-RUNX1* ALL (Chen et al., 2021; Wong et al., 2014). In contrast, RAG-mediated off-target V(D)J recombination is also observed in *BCR-ABL1<sup>+</sup>* B-ALL. These oncogenic SVs can also be considered as secondary events that promote the transition – 'leukemia to aggressive leukemia'. The enhancement of *BCR-ABL1<sup>+</sup>* B-ALL deterioration and progression by RAG<sup>cl</sup> in mouse model was consistent with our previous study that RAG enhances *BCR-ABL1*-positive leukemic cell growth through its endonuclease activity (Yuan et al., 2021). Additionally, we showed that non-core RAG1 region deletion leads to increased *Rag1<sup>cl</sup>* expression and high RAG expression related to low survival in pediatric acute lymphoid leukemia. Therefore, more attention should be paid to the non-core RAG region mutation in *BCR-ABL1<sup>+</sup>* B-ALL for the role of non-core region in leukemia suppression and off-target V(D)J recombination.

## Methods

### Mice

The C57BL/6 mice were purchased from the Experimental Animal Center of Xi'an Jiaotong University, while *Rag1<sup>cl</sup>* mice (amino acids 384–1040) and *Rag2<sup>cl</sup>* mice (amino acids 1–383) were obtained from Dr. David G. Schatz (Yale University, New Haven, CT, USA). The mice were bred and maintained in a specific pathogen-free environment at the Experimental Animal Center of Xi'an Jiaotong University. All animal-related procedures were in accordance with the guidelines approved by the Xi'an Jiaotong University Ethics Committee for Animal Experiments.

### Generation of retrovirus stocks

The pMSCV-BCR-BAL1-IRES-GFP vector is capable of co-expressing the human *BCR-ABL1* fusion protein and GFP, while the pMSCV-GFP vector serves as a negative control by solely expressing GFP. To produce viral particles, 293T cells were transfected with either the MSCV-BCR-BAL1-IRES-GFP or MSCV-GFP vector, along with the packaging vector PKAT2, utilizing the X-tremeGENE HP DNA Transfection Reagent from Roche (Basel, Switzerland). After 48 hr, the viral supernatants were collected, filtered, and stored at  $-80^{\circ}\text{C}$ .

### Bone marrow transduction and transplantation

Experiments were conducted using mice aged between 6 and 10 weeks. *BCR-ABL1<sup>+</sup>* B-ALL murine model was induced by utilizing marrow from donor mice who had not undergone 5-fluorouracil treatment. The donor mice were euthanized through  $\text{CO}_2$  asphyxiation, and the bone marrow was harvested by flushing the femur and tibia with a syringe and 26-gauge needle. Erythrocytes were not removed, and  $1 \times 10^6$  cells per well were plated in 6-well plates. A single round of co-sedimentation with retroviral stock was performed in medium containing 5% WEHI-3B-conditioned medium and 10 ng/ml IL-7 (Peprotech, USA). After transduction, cells were either transplanted into syngeneic female recipient mice ( $1 \times 10^6$  cells each) that had been lethally irradiated ( $2 \times 450$  cGy), or cultured in RPMI-1640 (Hyclone, Logan, UT) medium supplemented with 10% fetal calf serum (Hyclone), 200 mmol/l L-glutamine, 50 mmol/l 2-mercaptoethanol (Sigma, St Louis, MO), and 1.0 mg/ml penicillin/streptomycin (Hyclone). Subsequently, recipient mice were monitored daily for indications of morbidity, weight loss, failure to thrive, and splenomegaly. Weekly assessment of peripheral blood GFP percentage was done using flow cytometry analysis of tail vein blood. Hematopoietic tissues and cells were utilized for histopathology, in vitro culture, FACS analysis, secondary transplantation, genomic DNA preparation, protein lysate preparation, or lineage analysis, contingent upon the unique characteristics of mice under study.

## Secondary transplants

Thawed bone marrow cells were sorted using a BD FACS Aria II (Becton Dickinson, San Jose, CA, USA). GFP-positive leukemic cells ( $1 \times 10^6$ ,  $1 \times 10^5$ ,  $1 \times 10^4$ , and  $1 \times 10^3$ ) were then resuspended in 0.4 ml Hank's Balanced Salt Solution and intravenously administered to unirradiated syngeneic mice.

## Flow cytometry analysis and sorting

Bone marrow, spleen cells, and peripheral blood were harvested from leukemic mice. Red blood cells were eliminated using  $\text{NH}_4\text{Cl}$  red blood cell lysis buffer, and the remaining nucleated cells were washed with cold phosphate-buffered saline (PBS). In order to conduct in vitro cell surface receptor staining,  $1 \times 10^6$  cells were subjected to antibody staining for 20 min at  $4^\circ\text{C}$  in  $1 \times$  PBS containing 3% BSA. Cells were then washed with  $1 \times$  PBS and analyzed using a CytoFLEX Flow Cytometer (Beckman Coulter, Miami, FL) or sorted on a BD FACS Aria II. Apoptosis was analyzed by resuspending the cells in Binding Buffer (BD Biosciences, Baltimore, MD, USA), and subsequent labeling with anti-annexin V-AF647 antibody (BD Biosciences) and propidium iodide (BD Biosciences) for 15 min at room temperature. The lineage analysis was performed using the following antibodies, which were purchased from BD Biosciences: anti-BP-1-PerCP-Cy7, anti-CD19-PerCP-CyTM<sup>5.5</sup>, anti-CD43-PE, anti-B220-APC, and anti- $\mu\text{HC}$ -APC.

## BrdU incorporation and analysis

Cells obtained from primary leukemic mice were cultured in 6-well plates containing RPMI-1640 medium supplemented with 10% fetal bovine serum and 50  $\mu\text{g}/\text{ml}$  BrdU. After incubation at  $37^\circ\text{C}$  for 30 min, the cells were harvested and intranuclearly stained with anti-BrdU and 7-AAD antibodies, following the manufacturer's instructions.

## The in vitro V(D)J recombination assay

The retroviral recombination substrate pINV-12/23 was introduced into primary leukemic cells utilizing X-treme GENE HP DNA Transfection Reagent (Roche). Recombination efficiency of pINV-12/23 was evaluated through flow cytometry analysis for mouse CD90 (mCD90) and hCD4 expression (Yuan *et al.*, 2021).

## Western blotting analysis

Over  $1 \times 10^6$  leukemic cells were centrifuged and washed with ice-cold PBS. The cells were then treated with ice-cold radioimmune precipitation assay buffer, consisting of 50 mM Tris-HCl (pH 7.4), 0.15 M NaCl, 1% Triton X-100, 0.5% NaDoc, 0.1% sodium dodecyl sulphide, 1 mM ethylene diamine tetraacetic acid, 1 mM phenylmethanesulphonyl fluoride (Amresco), and fresh protease inhibitor cocktail Pepstain A (Sigma). After sonication using a Bioruptor TMUCD-200 (Diagenode, Seraing, Belgium), the suspension was spined at  $14,000 \times g$  for 3 min at  $4^\circ\text{C}$ . The total cell lysate was either utilized immediately or stored at  $-80^\circ\text{C}$ . Protein concentrations were determined using DC Protein Assay (Bio-Rad Laboratories, Hercules, California, USA). Subsequently, the protein samples (20  $\mu\text{g}$ ) were incubated with  $\alpha$ -RAG1 (mAb 23) and  $\alpha$ -RAG2 (mAb 39) antibodies (Teng *et al.*, 2015), with glyceraldehyde-3-phosphate dehydrogenase (GAPDH) serving as the loading control. The signal was further detected using secondary antibody of goat anti-rabbit IgG conjugated with horseradish peroxidase (Thermo Scientific, Waltham, MA). The band signal was developed with Immobilon Western Chemiluminescent HRP substrate (Millipore, Billerica, MA). The band development was analyzed using GEL-PRO ANALYZER software (Media Cybernetics, Bethesda, MD).

## Genomic PCR

Genomic PCR was performed in a 20- $\mu\text{l}$  reaction containing 50 ng of genomic DNA, 0.2  $\mu\text{M}$  of forward and reverse primer, and 10  $\mu\text{l}$  Premix Ex Taq (TaKaRa, Shiga, Japan). Amplification conditions were as follows:  $94^\circ\text{C}$  for 5 min; 35 cycles of 30 s at  $94^\circ\text{C}$ , 30 s at  $60^\circ\text{C}$  and 1 min at  $72^\circ\text{C}$ ;  $72^\circ\text{C}$  for 5 min (Bio-Rad, Hercules, CA). Genomic PCR was performed using the following primers:

DHL-5'-GGAATTCGTTTTTGTSAAGGGATCTACTACTGTG-3'; JH3-5'-GTCTAGATTCTCAC  
-AAGAGTCCGATAGACCCCTGG-3'; VHQ52-5'-CGGTACCAGACTGARCATCASCAAG

-GACAAYTCC-3'; VH558-5'-CGAGCTCTCCARCACAGCCTWCATGCARCTCARC-3'; VH7183-5'-CGGTACCAAGAASAMCCTGTWCCTGCAAATGASC-3' (*Schlissel et al., 1991*).

## RNA-seq library preparation and sequencing

GFP<sup>+</sup>CD19<sup>+</sup> cells were sorted from the spleen of *Rag1<sup>c/c</sup>* ( $n = 3$ ,  $1 \times 10^6$  cells/sample), *Rag2<sup>c/c</sup>* ( $n = 3$ ,  $1 \times 10^6$  cells/sample), and *Rag1<sup>ff</sup>* ( $n = 3$ ,  $1 \times 10^6$  cells/sample) *BCR-ABL1<sup>+</sup>* B-ALL mice. Total RNA was extracted using Trizol reagent (Invitrogen, CA, USA) following the manufacturer's guidelines. RNA quantity and purity analysis were done using Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent, CA, USA) with RNA integrity number (RIN) >7.0. RNA-seq libraries were prepared by using 200 ng total RNA with TruSeq RNA sample prep kit (Illumina). Oligo(dT)-enriched mRNAs were fragmented randomly with fragmentation buffer, followed by first- and second-strand cDNA synthesis. After a series of terminal repair, the double-stranded cDNA library was obtained through PCR enrichment and size selection. cDNA libraries were sequenced with the Illumina HiSeq 2000 sequencer (Illumina HiSeq 2000 v4 Single-Read 50 bp) after pooling according to its expected data volume and effective concentration. Two biological replicates were performed in the RNA-seq analysis. Raw reads were then aligned to the mouse genome (GRCm38) using Tophat2 RNA-seq alignment software, and unique reads were retained to quantify gene expression counts from Tophat2 alignment files. The differentially expressed mRNAs and genes were selected with  $\log_2$  (fold change) >1 or  $\log_2$  (fold change) <-1 and with statistical significance ( $p$  value <0.05) by R package. Bioinformatic analysis was performed using the OmicStudio tools at <https://www.omicstudio.cn/tool>.

## Preparation of tumor DNA samples

GFP<sup>+</sup>CD19<sup>+</sup> splenic cells, tail and kidney tissue were obtained from *Rag1<sup>c/c</sup>*, *Rag2<sup>c/c</sup>*, and *Rag1<sup>ff</sup>* *BCR-ABL1<sup>+</sup>* B-ALL mice, and genomic DNA was extracted using a TIANamp Genomic DNA Kit (TIANGEN-DP304). Subsequently, paired-end libraries were constructed from 1  $\mu$ g of the initial genomic material using the TruSeq DNA v2 Sample Prep Kit (Illumina, #FC-121-2001) as per the manufacturer's instructions. The size distribution of the libraries was assessed using an Agilent 2100 Bioanalyzer (Agilent Technologies, #5067-4626), and the DNA concentration was quantified using a Qubit dsDNA HS Assay Kit (Life Technologies, #Q32851). The Illumina HiSeq 4000 was utilized to sequence the samples, with two to four lanes allocated for sequencing the tumor and one lane for the control DNA library of the kidney or liver, each with 150 bp paired end reads.

## Read alignment and SV calling

Fastq files were generated using Casava 1.8 (Illumina), and BWA 37 was employed to align the reads to mm9. PCR duplicates were eliminated using Picard's Mark Duplicates tool ([sourceforge.net/apps/mediawiki/picard](http://sourceforge.net/apps/mediawiki/picard)). Our custom scripts (<http://sourceforge.net/projects/svddetection>) were utilized to eliminate BWA-designated concordant and read pairs with low BWA mapping quality scores. Intra- and inter-chromosomal rearrangements were identified using SV Detect from discordant, quality prefiltered read pairs. The mean insertion size and standard deviation for this analysis were obtained through Picard's InsertSizeMetrics tool (<https://sourceforge.net/projects/picard/>). Tumor-specific SVs were identified using the manta software (<https://github.com/Illumina/manta>; *Saunders, 2016*).

## Validation of high-confidence off-target candidates

The elimination of non-specific structural mutations from the kidney or tail was necessary for tumor-specific SVs identification. Subsequently, the method involving 21 bp CAC-to-breakpoint was employed to filter RAG-mediated off-target gene. The validation of high-confidence off-target candidates was carried out through PCR. Oligonucleotide primers were designed to hybridize within the 'linking' regions of SV Detect, in the appropriate orientation. The PCR product was subjected to Sanger sequencing and aligned to the mouse mm9 reference genome using BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

## Statistics

Statistical analysis was conducted using SPSS 20.0 (IBM Corp) and GraphPad Prism 6.0 (GraphPad Software). Descriptive statistics were reported as means  $\pm$  standard deviation for continuous variables. Statistical analyses were applied to biologically independent mice or technical replicates for

each experiment which was independently repeated at least three times. The equality of variances was assessed using Levene's test. Two-group comparisons, multiple group comparisons, and survival comparisons were performed using independent-samples *t* test, one-way analyses of variance with post hoc Fisher's least significant difference (LSD) test, and log-rank Mantel–Cox analysis, respectively. Kaplan–Meier survival curves were utilized to depict the changes in survival rate over time. Statistical significance was set at  $p < 0.05$ .

## Acknowledgements

This study was supported by grants (no. 31170821, 31370874, and 81670157 to YJ) from the National Natural Science Foundation of China and by a grant (no. 2022JZ-44 to YJ) from the Natural Science Foundation of Shaanxi Province. The authors would like to thank Professor Shaoguang Li from the Division of Hematology/Oncology, University of Massachusetts Medical School, for providing the MSCV-BCR-BAL1-IRES-GFP construct. The authors would also like to thank Mr. Xiaofei Wang (Xi'an Jiaotong University Health Science Centre) for providing expert technical assistance with cell sorting.

---

## Additional information

### Funding

Funder	Grant reference number	Author
National Natural Science Foundation of China	31170821	Yanhong Ji
Natural Science Foundation of Shaanxi Province	2022JZ-44	Yanhong Ji
National Natural Science Foundation of China	31370874	Yanhong Ji
National Natural Science Foundation of China	81670157	Yanhong Ji
Xi 'an Jiaotong University Medical "Basic - Clinical" Fusion Innovation Project	YXJLRH2022012	Yanhong Ji
Xi 'an Jiaotong University Medical "Basic - Clinical" Fusion Innovation Project	YJSCX-2024-007	Yanhong Ji

The funders had no role in study design, data collection, and interpretation, or the decision to submit the work for publication.

### Author contributions

Xiaozhuo Yu, Conceptualization, Validation, Visualization, Methodology, Writing – original draft; Wen Zhou, Validation, Visualization, Methodology; Xiaodong Chen, Validation, Methodology, Writing – review and editing; Shunyu He, Data curation; Mengting Qin, Data curation, Methodology; Meng Yuan, Validation, Writing – review and editing; Yang Wang, Software, Writing – review and editing; Woodvine Otieno Odhiambo, Writing – review and editing; Yinsha Miao, Writing – original draft; Yanhong Ji, Conceptualization, Formal analysis, Investigation, Methodology, Writing – original draft, Writing – review and editing

### Author ORCIDs

Xiaozhuo Yu  <http://orcid.org/0009-0006-7565-8941>

Yanhong Ji  <http://orcid.org/0000-0003-4144-4786>

### Ethics

All animal experiments were performed with the approval of the Institutional Animal Care and Use Committee of Xi'an Jiaotong University (under reference XJTU#20202272).



## Peer review material

Reviewer #1 (Public Review): <https://doi.org/10.7554/eLife.91030.3.sa1>Author response <https://doi.org/10.7554/eLife.91030.3.sa2>

## Additional files

## Supplementary files

- Supplementary file 1. Sequencing statistics.
- Supplementary file 2. Abnormal rearrangements at Ig.
- Supplementary file 3. RAG-mediated off-target genes.
- MDAR checklist

## Data availability

Files containing sequencing reads were deposited in the National Institutes of Health Sequence Read Archive under PRJNA1116598.

The following dataset was generated:

Author(s)	Year	Dataset title	Dataset URL	Database and Identifier
Ji Y	2024	RAG1 and RAG2 Non-core Regions Are Implicated in Leu-kemogenesis and Off-target V(D)J Recombination in BCR-ABL1-driven B-cell Lineage Lymphoblastic Leukemia	<a href="https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1116598/">https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1116598/</a>	NCBI BioProject, PRJNA1116598

## References

- Banerjee JK**, Schatz DG. 2014. Synapsis alters RAG-mediated nicking at Tcrb recombination signal sequences: implications for the “beyond 12/23” rule. *Molecular and Cellular Biology* **34**:2566–2580. DOI: <https://doi.org/10.1128/MCB.00411-14>, PMID: 24797073
- Bateman CM**, Colman SM, Chaplin T, Young BD, Eden TO, Bhakta M, Gratias EJ, van Wering ER, Cazzaniga G, Harrison CJ, Hain R, Ancliff P, Ford AM, Kearney L, Greaves M. 2010. Acquisition of genome-wide copy number alterations in monozygotic twins with acute lymphoblastic leukemia. *Blood* **115**:3553–3558. DOI: <https://doi.org/10.1182/blood-2009-10-251413>, PMID: 20061556
- Becker LM**, O’Connell JT, Vo AP, Cain MP, Tampe D, Bizarro L, Sugimoto H, McGow AK, Asara JM, Lovisa S, McAndrews KM, Zielinski R, Lorenzi PL, Zeisberg M, Raza S, LeBleu VS, Kalluri R. 2020. Epigenetic reprogramming of cancer-associated fibroblasts deregulates glucose metabolism and facilitates progression of breast cancer. *Cell Reports* **31**:107701. DOI: <https://doi.org/10.1016/j.celrep.2020.107701>, PMID: 32492417
- Beilinson HA**, Glynn RA, Yadavalli AD, Xiao J, Corbett E, Saribasak H, Arya R, Miot C, Bhattacharyya A, Jones JM, Pongubala JMR, Bassing CH, Schatz DG. 2021. The RAG1 N-terminal region regulates the efficiency and pathways of synapsis for V(D)J recombination. *The Journal of Experimental Medicine* **218**:e20210250. DOI: <https://doi.org/10.1084/jem.20210250>, PMID: 34402853
- Braams M**, Pike-Overzet K, Staal FJT. 2023. The recombinase activating genes: architects of immune diversity during lymphocyte development. *Frontiers in Immunology* **14**:1210818. DOI: <https://doi.org/10.3389/fimmu.2023.1210818>, PMID: 37497222
- Burn TN**, Miot C, Gordon SM, Culberson EJ, Diamond T, Kreiger PA, Hayer KE, Bhattacharyya A, Jones JM, Bassing CH, Behrens EM. 2022. The RAG1 ubiquitin ligase domain stimulates recombination of tcr $\beta$  and tcr $\alpha$  genes and influences development of  $\alpha\beta$  t cell lineages. *Journal of Immunology* **209**:938–949. DOI: <https://doi.org/10.4049/jimmunol.2001441>, PMID: 35948399
- Chen D**, Camponeschi A, Nordlund J, Marincevic-Zuniga Y, Abrahamsson J, Lönnerholm G, Fogelstrand L, Mårtensson IL. 2021. RAG1 co-expression signature identifies ETV6-RUNX1-like B-cell precursor acute lymphoblastic leukemia in children. *Cancer Medicine* **10**:3997–4003. DOI: <https://doi.org/10.1002/cam4.3928>, PMID: 33987955
- Curry JD**, Schlissel MS. 2008. RAG2’s non-core domain contributes to the ordered regulation of V(D)J recombination. *Nucleic Acids Research* **36**:5750–5762. DOI: <https://doi.org/10.1093/nar/gkn553>, PMID: 18776220
- Deng Z**, Liu H, Liu X. 2015. RAG1-mediated ubiquitylation of histone H3 is required for chromosomal V(D)J recombination. *Cell Research* **25**:181–192. DOI: <https://doi.org/10.1038/cr.2015.1>, PMID: 25572281
- Eastman QM**, Leu TM, Schatz DG. 1996. Initiation of V(D)J recombination in vitro obeying the 12/23 rule. *Nature* **380**:85–88. DOI: <https://doi.org/10.1038/380085a0>, PMID: 8598914

- Fatma H**, Maurya SK, Siddique HR. 2022. Epigenetic modifications of c-MYC: role in cancer cell reprogramming, progression and chemoresistance. *Seminars in Cancer Biology* **83**:166–176. DOI: <https://doi.org/10.1016/j.semcancer.2020.11.008>, PMID: 33220458
- Fugmann SD**, Lee AI, Shockett PE, Villey IJ, Schatz DG. 2000. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annual Review of Immunology* **18**:495–527. DOI: <https://doi.org/10.1146/annurev.immunol.18.1.495>, PMID: 10837067
- Gan T**, Wang Y, Liu Y, Schatz DG, Hu J. 2021. RAG2 abolishes RAG1 aggregation to facilitate V(D)J recombination. *Cell Reports* **37**:109824. DOI: <https://doi.org/10.1016/j.celrep.2021.109824>, PMID: 34644584
- Gaymes TJ**, North PS, Brady N, Hickson ID, Mufti GJ, Rassool FV. 2002. Increased error-prone non homologous DNA end-joining—a proposed mechanism of chromosomal instability in Bloom’s syndrome. *Oncogene* **21**:2525–2533. DOI: <https://doi.org/10.1038/sj.onc.1205331>, PMID: 11971187
- Goel S**, Bhatia V, Biswas T, Ateeq B. 2022. Epigenetic reprogramming during prostate cancer progression: a perspective from development. *Seminars in Cancer Biology* **83**:136–151. DOI: <https://doi.org/10.1016/j.semcancer.2021.01.009>, PMID: 33545340
- Hardy RR**, Hayakawa K. 2001. B cell development pathways. *Annual Review of Immunology* **19**:595–621. DOI: <https://doi.org/10.1146/annurev.immunol.19.1.595>, PMID: 11244048
- Hirokawa S**, Chure G, Belliveau NM, Lovely GA, Anaya M, Schatz DG, Baltimore D, Phillips R. 2020. Sequence-dependent dynamics of synthetic and endogenous RSSs in V(D)J recombination. *Nucleic Acids Research* **48**:6726–6739. DOI: <https://doi.org/10.1093/nar/gkaa418>, PMID: 32449932
- Khoshchehreh R**, Totonchi M, Carlos Ramirez J, Torres R, Baharvand H, Aicher A, Ebrahimi M, Heeschen C. 2019. Epigenetic reprogramming of primary pancreatic cancer cells counteracts their in vivo tumorigenicity. *Oncogene* **38**:6226–6239. DOI: <https://doi.org/10.1038/s41388-019-0871-x>, PMID: 31308488
- Kumari R**, Roy U, Desai S, Nilavar NM, Van Nieuwenhuijze A, Paranjape A, Radha G, Bawa P, Srivastava M, Nambiar M, Balaji KN, Liston A, Choudhary B, Raghavan SC. 2021. MicroRNA miR-29c regulates RAG1 expression and modulates V(D)J recombination during B cell development. *Cell Reports* **36**:109390. DOI: <https://doi.org/10.1016/j.celrep.2021.109390>, PMID: 34260911
- Libri A**, Marton T, Deriano L. 2021. The (lack of) dna double-strand break repair pathway choice during v(d)j recombination. *Frontiers in Genetics* **12**:823943. DOI: <https://doi.org/10.3389/fgene.2021.823943>, PMID: 35082840
- Liu Y**, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. 2007. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity* **27**:561–571. DOI: <https://doi.org/10.1016/j.immuni.2007.09.005>, PMID: 17936034
- Liu C**, Zhang Y, Liu CC, Schatz DG. 2022. Structural insights into the evolution of the RAG recombinase. *Nature Reviews. Immunology* **22**:353–370. DOI: <https://doi.org/10.1038/s41577-021-00628-6>, PMID: 34675378
- Lu C**, Ward A, Bettridge J, Liu Y, Desiderio S. 2015. An autoregulatory mechanism imposes allosteric control on the V(D)J recombinase by histone H3 methylation. *Cell Reports* **10**:29–38. DOI: <https://doi.org/10.1016/j.celrep.2014.12.001>, PMID: 25543141
- Matthews AGW**, Kuo AJ, Ramón-Maiques S, Han S, Champagne KS, Ivanov D, Gallardo M, Carney D, Cheung P, Ciccone DN, Walter KL, Utz PJ, Shi Y, Kutateladze TG, Yang W, Gozani O, Oettinger MA. 2007. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450**:1106–1110. DOI: <https://doi.org/10.1038/nature06431>, PMID: 18033247
- Mendes RD**, Sarmiento LM, Canté-Barrett K, Zuurbier L, Buijs-Gladdines J, Póvoa V, Smits WK, Abecasis M, Yunes JA, Sonneveld E, Horstmann MA, Pieters R, Barata JT, Meijerink JPP. 2014. PTEN microdeletions in T-cell acute lymphoblastic leukemia are caused by illegitimate RAG-mediated recombination events. *Blood* **124**:567–578. DOI: <https://doi.org/10.1182/blood-2014-03-562751>, PMID: 24904117
- Mijušković M**, Chou YF, Gigi V, Lindsay CR, Shestova O, Lewis SM, Roth DB. 2015. Off-target v(d)j recombination drives lymphomagenesis and is escalated by loss of the rag2 c terminus. *Cell Reports* **12**:1842–1852. DOI: <https://doi.org/10.1016/j.celrep.2015.08.034>, PMID: 26365182
- Onozawa M**, Aplan PD. 2012. Illegitimate V(D)J recombination involving nonantigen receptor loci in lymphoid malignancy. *Genes, Chromosomes & Cancer* **51**:525–535. DOI: <https://doi.org/10.1002/gcc.21942>, PMID: 22334400
- Raghavan SC**, Tong J, Lieber MR. 2006. Hybrid joint formation in human V(D)J recombination requires nonhomologous DNA end joining. *DNA Repair* **5**:278–285. DOI: <https://doi.org/10.1016/j.dnarep.2005.09.008>, PMID: 16275127
- Rassool FV**. 2003. DNA double strand breaks (DSB) and non-homologous end joining (NHEJ) pathways in human leukemia. *Cancer Letters* **193**:1–9. DOI: [https://doi.org/10.1016/s0304-3835\(02\)00692-4](https://doi.org/10.1016/s0304-3835(02)00692-4), PMID: 12691817
- Repasky JAE**, Corbett E, Boboila C, Schatz DG. 2004. Mutational analysis of terminal deoxynucleotidyltransferase-mediated N-nucleotide addition in V(D)J recombination. *Journal of Immunology* **172**:5478–5488. DOI: <https://doi.org/10.4049/jimmunol.172.9.5478>, PMID: 15100289
- Rooney S**, Chaudhuri J, Alt FW. 2004. The role of the non-homologous end-joining pathway in lymphocyte development. *Immunological Reviews* **200**:115–131. DOI: <https://doi.org/10.1111/j.0105-2896.2004.00165.x>, PMID: 15242400
- Sadofsky MJ**, Hesse JE, McBlane JF, Gellert M. 1993. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Research* **21**:5644–5650. DOI: <https://doi.org/10.1093/nar/21.24.5644>, PMID: 8284210
- Saunders C**. 2016. Manta. 1.6.0. GitHub. DOI: <https://doi.org/https://github.com/Illumina/manta>

- Schabla NM**, Perry GA, Palmer VL, Swanson PC. 2018. Vprbp (dcaf1) regulates rag1 expression independently of dicer by mediating rag1 degradation. *Journal of Immunology* **201**:930–939. DOI: <https://doi.org/10.4049/jimmunol.1800054>, PMID: 29925675
- Schatz DG**, Ji Y. 2011a. Recombination centres and the orchestration of V(D)J recombination. *Nature Reviews Immunology* **11**:251–263. DOI: <https://doi.org/10.1038/nri2941>, PMID: 21394103
- Schatz DG**, Swanson PC. 2011b. V(D)J recombination: mechanisms of initiation. *Annual Review of Genetics* **45**:167–202. DOI: <https://doi.org/10.1146/annurev-genet-110410-132552>, PMID: 21854230
- Schjerven H**, Ayongaba EF, Aghajani A, McLaughlin J, Cheng D, Geng H, Boyd JR, Eggesbø LM, Lindeman I, Heath JL, Park E, Witte ON, Smale ST, Frietze S, Müschen M. 2017. Genetic analysis of Ikaros target genes and tumor suppressor function in BCR-ABL1<sup>+</sup> pre-B ALL. *The Journal of Experimental Medicine* **214**:793–814. DOI: <https://doi.org/10.1084/jem.20160049>, PMID: 28190001
- Schlissel MS**, Corcoran LM, Baltimore D. 1991. Virus-transformed pre-B cells show ordered activation but not inactivation of immunoglobulin gene rearrangement and transcription. *The Journal of Experimental Medicine* **173**:711–720. DOI: <https://doi.org/10.1084/jem.173.3.711>, PMID: 1900081
- Scully R**, Panday A, Elango R, Willis NA. 2019. DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nature Reviews Molecular Cell Biology* **20**:698–714. DOI: <https://doi.org/10.1038/s41580-019-0152-0>, PMID: 31263220
- Sekiguchi JA**, Whitlow S, Alt FW. 2001. Increased accumulation of hybrid V(D)J joins in cells expressing truncated versus full-length RAGs. *Molecular Cell* **8**:1383–1390. DOI: [https://doi.org/10.1016/s1097-2765\(01\)00423-3](https://doi.org/10.1016/s1097-2765(01)00423-3), PMID: 11779512
- Shi B**, Dong X, Ma Q, Sun S, Ma L, Yu J, Wang X, Pan J, He X, Su D, Yao X. 2020. The usage of human ighj genes follows a particular non-random selection: the recombination signal sequence may affect the usage of human ighj genes. *Frontiers in Genetics* **11**:524413. DOI: <https://doi.org/10.3389/fgene.2020.524413>, PMID: 33363565
- Silver DP**, Spanopoulou E, Mulligan RC, Baltimore D. 1993. Dispensable sequence motifs in the RAG-1 and RAG-2 genes for plasmid V(D)J recombination. *PNAS* **90**:6100–6104. DOI: <https://doi.org/10.1073/pnas.90.13.6100>, PMID: 8327489
- Talukder SR**, Dudley DD, Alt FW, Takahama Y, Akamatsu Y. 2004. Increased frequency of aberrant V(D)J recombination products in core RAG-expressing mice. *Nucleic Acids Research* **32**:4539–4549. DOI: <https://doi.org/10.1093/nar/gkh778>, PMID: 15328366
- Teng G**, Maman Y, Resch W, Kim M, Yamane A, Qian J, Kieffer-Kwon KR, Mandal M, Ji Y, Meffre E, Clark MR, Cowell LG, Casellas R, Schatz DG. 2015. RAG represents a widespread threat to the lymphocyte genome. *Cell* **162**:751–765. DOI: <https://doi.org/10.1016/j.cell.2015.07.009>, PMID: 26234156
- Thomson DW**, Shahrin NH, Wang PPS, Wadham C, Shanmuganathan N, Scott HS, Dinger ME, Hughes TP, Schreiber AW, Branford S. 2020. Aberrant RAG-mediated recombination contributes to multiple structural rearrangements in lymphoid blast crisis of chronic myeloid leukemia. *Leukemia* **34**:2051–2063. DOI: <https://doi.org/10.1038/s41375-020-0751-y>, PMID: 32076119
- Wiegman AP**, Ward A, Ivanova E, Duijf PHG, Adams MN, Najib IM, Van Oosterhout R, Sadowski MC, Kelly G, Morrical SW, O’Byrne K, Lee JS, Richard DJ. 2021. Genome instability and pressure on non-homologous end joining drives chemotherapy resistance via a DNA repair crisis switch in triple negative breast cancer. *NAR Cancer* **3**:zcab022. DOI: <https://doi.org/10.1093/narcan/zcab022>, PMID: 34316709
- Wong S**, Witte ON. 2004. The BCR-ABL story: bench to bedside and back. *Annual Review of Immunology* **22**:247–306. DOI: <https://doi.org/10.1146/annurev.immunol.22.012703.104753>, PMID: 15032571
- Wong CC**, Martincorena I, Rust AG, Rashid M, Alifrangis C, Alexandrov LB, Tiffen JC, Kober C, Green AR, Massie CE, Nangalia J, Lempidaki S, Döhner H, Döhner K, Bray SJ, McDermott U, Papaemmanuil E, Campbell PJ, Adams DJ. 2014. Inactivating CUX1 mutations promote tumorigenesis. *Nature Genetics* **46**:33–38. DOI: <https://doi.org/10.1038/ng.2846>, PMID: 24316979
- Yuan M**, Wang Y, Qin M, Zhao X, Chen X, Li D, Miao Y, Otieno Odhiambo W, Liu H, Ma Y, Ji Y. 2021. RAG enhances BCR-ABL1-positive leukemic cell growth through its endonuclease activity in vitro and in vivo. *Cancer Science* **112**:2679–2691. DOI: <https://doi.org/10.1111/cas.14939>, PMID: 33949040
- Zhang L**, Reynolds TL, Shan X, Desiderio S. 2011. Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. *Immunity* **34**:163–174. DOI: <https://doi.org/10.1016/j.immuni.2011.02.003>, PMID: 21349429