HHS Public Access

Author manuscript

Cell Rep. Author manuscript; available in PMC 2022 October 12.

Published in final edited form as:

Cell Rep. 2021 October 12; 37(2): 109824. doi:10.1016/j.celrep.2021.109824.

RAG2 abolishes RAG1 aggregation to facilitate V(D)J recombination

Tingting Gan[1], **Yuhong Wang**[1], **Yang Liu**[1], **David G. Schatz**[2], **Jiazhi Hu**[1,3,*]

[1]The MOE Key Laboratory of Cell Proliferation and Differentiation, School of Life Sciences, Peking-Tsinghua Center for Life Sciences, Peking University, Beijing 100871, China.

[2]Department of Immunobiology, Yale University School of Medicine, New Haven, Connecticut, 06519, USA.

[3]Lead contact

Summary

RAG1 and RAG2 form a tetramer nuclease to initiate V(D)J recombination in developing T and B lymphocytes. The RAG1 protein evolves from a transposon ancestor and possesses nuclease activity that requires interaction with RAG2. Here, we show that the human RAG1 aggregates in the nucleus in the absence of RAG2, exhibiting a extremely low V(D)J recombination activity. In contrast, RAG2 does not aggregate by itself, but it interacts with RAG1 to disrupt RAG1 aggregates and thereby to activate robust V(D)J recombination. Moreover, RAG2 from mouse and zebrafish could not disrupt the aggregation of human RAG1 as efficiently as human RAG2 did, indicating a species-specific regulatory mechanism for RAG1 by RAG2. Therefore, we propose that RAG2 coevolves with RAG1 to release inert RAG1 from aggregates and thereby activate V(D)J recombination to generate diverse antigen receptors in lymphocytes.

Introduction

V(D)J recombination shuffles the separate variable (V), diversity (D), and joining (J) gene segments in developing B and T lymphocytes of jawed vertebrates to generate diverse antigen receptors. The tandem V, D, and J segments are flanked by recombination signal sequences (RSSs) containing two relatively conserved motifs: a heptamer (CACAGTG) and a nonamer (ACAAAAACC). The recombination activating genes (*RAGs*) encode two subunits, RAG1 and RAG2, that form the RAG complex to target the RSSs and generate DNA double-stranded breaks (DSBs) between the receptor gene segments and RSSs (Schatz and Ji, 2011). The liberated segments are then joined by the classical nonhomologous end-joining (c-NHEJ) pathway to complete V(D)J recombination (Alt et al., 2013).

*Correspondence to Jiazhi Hu: hujz@pku.edu.cn.

Author Contributions

T.G. and J.H. conceived of the experiments. Experiments were performed by T.G., Y.W., and Y.L.; T.G., Y.W., Y.L., D.G.S, and J.H. analyzed the data; T.G., Y.W., D.G.S, and J.H. wrote the paper.

Declaration of Interests

The authors declare no competing interests.

The human RAG1 contains 1043 residues and includes the catalytic sites responsible for DNA cleavage (Schatz et al., 1989). During DNA cleavage, RAG1 nicks the DNA and subsequently generates DSBs containing a hairpin and a blunt broken end, as is also the case for some DNA transposases and consistent with the hypothesis that RAG1 evolved from a transposase ancestor (Carmona and Schatz, 2017; Lieber, 2019; Teng et al., 2015). Human RAG2 is only 527 residues in length and makes no direct contribution to DNA cleavage, but it enhances RAG1 cleavage activity hundreds of fold (Oettinger et al., 1990; Qiu et al., 2001; Schatz and Swanson, 2011). In this context, V(D)J recombination occurs at levels that are similar to the background in lymphoid cells in the absence of RAG2 (Carmona et al., 2016). The core regions of the RAG complex retain DNA cleavage activity, and recent structural studies of core RAG (cRAG1 and cRAG2) in mouse and zebrafish further reveal the organization of the RAG complex: two core subunits form a tetramer that holds the two paired DNA strands (Chen et al., 2020; Grundy et al., 2009; Kim et al., 2015; Ru et al., 2015; Ru et al., 2018a). Specifically, RAG2 is responsible for stabilizing and conferring specificity to the interaction of RAG1 with substrate DNA (Fugmann et al., 2000; Ji et al., 2010; Swanson, 2004) and also contributes to protein-protein interactions within and between the two halves of the RAG tetramer (Ru et al., 2018b). Both RAG1 and RAG2 are vital for V(D)J recombination in lymphocytes, and many RAG1 or RAG2 loss-of-function mutants have been identified as the cause of human severe combined immunodeficiency (SCID; Kim et al., 2015; Schwarz et al., 1996).

When stimulated by many factors, including cytokines and transcription factors, RAG1 and RAG2 are convergently transcribed in developing lymphocytes (Kuo and Schlissel, 2009; Oettinger et al., 1990). Posttranscriptional modifications are involved in modulating the catalytic activity of RAG complex, including the regulation of RAG degradation and sequestration of RAG activity in a specific phase of cell development (Chaumeil et al., 2013; Kuo and Schlissel, 2009). For instance, RAG2 is expressed in the G1 phase and is modified by cyclin-dependent kinases to initiate rapid degradation when entering the S phase (Li et al., 1996; Lin and Desiderio, 1994). Therefore, V(D)J recombination only occurs in the G1 phase when c-NHEJ dominates DSB repair. In contrast to RAG2, the amount of RAG1 is fairly consistent across the cell cycle (Lee and Desiderio, 1999). In addition, mutation of RAG2 to abolish cell-cycle-control poses threats to genome stability and leads to lymphoid tumorigenesis (Zhang et al., 2011). Very recently, the Schatz lab showed that mouse RAG1 can be stored in nucleoli and escapes in response to G1 arrest (Brecht et al., 2020), providing a novel mechanism for the regulation of RAG1 and V(D)J recombination, and highlighting the importance of understanding the physical nature of the RAG1 and how RAG2 participates in V(D)J recombination.

Protein aggregation or condensation was recently proposed to be a new dimension of protein function regulation (Li et al., 2012; Shin and Brangwynne, 2017; Hnisz et al., 2017; Su et al., 2016). From this conceptual starting point, we found that human RAG1 formed scattered aggregates in the entire nucleus of cells and RAG2 abolished RAG1 aggregation to initiate V(D)J recombination. Our findings provide an additional explanation of how RAG2 restricts RAG1 to function in only the G1 phase of the cell cycle. Upon RAG2 degradation in the S/G2/M phases of the cell cycle, residual RAG1 aggregates, possibly into

some membraneless organelles including nucleolus, to limit its endonuclease activity and potentially reduce the incidence of genome-wide off-target damage.

Results

Human RAG shows robust catalytic activity in nonlymphoid cells

Though RAG1 and RAG2 are mainly expressed in lymphoid cells, the RAG complex also recognizes RSSs and initiates V(D)J recombination-like reactions in nonlymphoid cells. In this context, RAG1 was first identified in NIH 3T3 fibroblasts (Schatz et al., 1989). To provide a detailed comparison of the behaviors of RAG1 and RAG2 in nonlymphoid cells as well as in developing lymphocytes, we integrated a DEL-CJ fragment into HEK293T cells (293T-CJ) via viral delivery, as previously described (Bredemeyer et al., 2006). The DEL-CJ fragment contains a pair of convergent *bona fide* 12 and 23 RSSs flanked by an inverted GFP gene, which allows GFP-deletional recombination initiated by the RAG complex (Fig. 1A). To enhance the sensitivity of detection, we also introduced a quantitative high-throughput sequencing method, termed primer-extension-mediated sequencing (PEM-seq; Yin et al., 2019; Liu et al., 2021; Zhang et al., 2021), to detect the deletional coding joins after ectopic expression of RAG in 293T-CJ cells. PEM-seq relies on a biotinylated primer adjacent to the 23 RSS as bait to capture uncut templates and deletional coding joins and then measures the frequency of deletional coding joins (Fig. 1A).

We fused eGFP to the C-terminus of human RAG1 (RAG1-GFP) and overexpressed it in 293T-CJ cells in the presence or absence of human RAG2. The genomic DNA was harvested four days later to generate PEM-seq libraries. In the absence of RAG2, both RAG1 or RAG1-GFP showed extremely low but detectable levels of recombination between the two RSSs; while in the RAG1 and RAG2 coexpressed cells, the recombination level increased more than 1100 fold (Fig. 1B, S1A, and S1B), though the mRNA level of RAG1-GFP was comparable to that in RAG1-GFP only cells (Fig. S1C). In addition, mCherry-tagged RAG2 also showed a similar ability to activate RAG1-GFP for recombination detected by PEM-seq or PCR (Fig. S1A and S1B). These results indicate dramatically increased activity of the RAG complex as compared to the RAG1 only in a nonlymphoid cell line, similar to the synergetic function of RAG1 and RAG2 observed in lymphocytes. Given the restricted access to human progenitor B or T cells for the study of RAG complex, the consistency of RAG activity in nonlymphoid cells versus lymphocytes allowed us to analyze human RAG1 and RAG2 in available human cell lines.

Human RAG1 but not RAG2 undergoes aggregation in the nucleus

To examine aggregation status, we expressed eGFP-tagged RAG1 or RAG2 individually in 293T-CJ cells. RAG1 formed scattered puncta in the nucleus of approximately 95% of the observed cells, but aggregation was barely observed for RAG2 (Fig. 1C). Similar findings were obtained in the U2OS cells (Fig. 1D, right). Moreover, the RAG1 puncta in the U2OS cells recovered rapidly and could reach a comparable initial intensity at the 20 s time-point after local photobleaching (Fig. 1E), suggesting that these puncta were likely in a liquid phase (Sabari et al., 2018). NALM-6 is a cell line derived from a 19-year-old patient with acute lymphoblastic leukemia (ALL) and is widely accepted to be a precursor B (pre-B) cell

line; these cells exhibit slight endogenous RAG expression (Minowada et al., 1978). Using an ectopic expression strategy similar to that of the 293T-CJ cells, we found that the RAG1 formed puncta in the nucleus of approximately 70% of observed RAG1-deficient NALM-6 cells, while the RAG2 showed puncta in few cells (Fig. 1D, left). The reduced level of RAG1 aggregation in NALM-6 cells could be due to the low expression level of RAG1-GFP (Fig. S1D) or the presence of endogenous RAG2 or both.

To test whether the presence of RAG2 resulted in the disruption of RAG1 aggregates, we coexpressed RAG1 and RAG2 in the 293T-CJ or U2OS cells. Surprisingly, in the coexpression system, less than 20% of RAG1-GFP/RAG2-mCherry coexpressed cells showed RAG1 puncta, while over 88% of RAG1 only cells containing RAG1 puncta in both types of cells (Fig. S1E). To further test whether RAG2 can disrupt existing RAG1 aggregates, we expressed RAG1 in 293T-CJ cells for 24 hours to allow RAG1 aggregation in advance and then expressed RAG2 in the same cells. RAG2 but not the empty vector could efficiently reduce RAG1 aggregation from 86% to 18% (Fig. S1F), further suggesting that RAG2 can abolish RAG1 aggregates.

Knocked-in human RAG1 undergoes aggregation in mouse pro-B cells

The human tumor-derived NALM-6 pre-B cells show an unphysiological expression level of RAG1 and RAG2, so we next employed the Abelson virus-transformed murine progenitor-B (*vAbl* pro-B) cells (Shinkai et al., 1992; Hu et al., 2015) to validate RAG1 aggregation in lymphocytes. The *vAbl* cell line routinely supports V(D)J recombination in the *IgH* or *Igk* loci in the G1 phase arrested by STI-571 (Hu et al., 2015). We generated a stable *vAbl* pro-B cell line from a RAG2-deficient mouse and then knocked out mouse *Rag1* in the transformed *vAbl* cell lines (Shinkai et al., 1992). Next, we knocked in the human *RAG1* coding sequence fused with *maxGFP* in one allele of the $Rag1^{-/-}Rag2^{-/-}$ *vAbl* cells (Fig. S2A and S2B). To compare the expression level of knocked-in RAG1 in mouse *vAbl* pro-B cells with endogenous RAG1 in human pro-B cells, we isolated the $CD34^{+}CD19^{+}CD10^{+}$ pro-B cells from human cord blood (Reynaud et al., 2003; McWilliams et al., 2013) and then used single-cell qRT-PCR assay (Hagemann-Jensen et al., 2020) to quantify the expression of RAG1 (Fig. S2C and S2D). The single-cell qRT-PCR assay was validated by the consistency of the mRNA levels of *Gapdh* or *GAPDH* in mouse *vAbl* pro-B or human 293T-CJ cells between different genotypes (Fig. S2E). We found the mRNA level of endogenous RAG1 from human pro-B cells was slightly higher than RAG1-maxGFP in G1-arrested mouse *vAbl* pro-B cells, and both were higher than that of the NALM-6 cells and much lower than that of the overexpressed 293T-CJ cells (Fig. 1F and Fig. S2F).

About 71% and 60% of examined cells showed aggregation puncta of RAG1-maxGFP in G1-arrested or cycling *vAbl* cells (Fig. 1G). The G1-arrested cells exhibited a higher level of RAG1 aggregation, possibly due to smaller cell size and higher nuclear RAG1-maxGFP concentration (Fig. S2G and S2H). Of note, most of the RAG1 aggregates in the *vAbl* cells were scattered as showed in human nonlymphoid cells while some of the puncta were bigger (Fig. 1G, middle panel), potentially sequestrated by nucleolus as described previously for mouse RAG1 (Brecht et al., 2020). Consistent with the findings in 293T-CJ and U2OS cells, coexpression of human RAG2 caused a significantly decreased level of RAG1 aggregation

in the *vAbl* pro-B cells (Fig. 1H). Collectively, these findings strongly suggested that RAG2 could regulate the formation of RAG1 aggregates as a potential mechanism for regulating V(D)J recombination.

RAG1 aggregation occurs in the nuclei

To dissect the key residues of RAG1 contributing to its aggregation, we mutated the DDE (D603, D711, and E965) motif of RAG1 at first (Fig. 2A). The DDE motif is responsible for binding magnesium ions and catalyzing DNA cleavage (Fugmann et al., 2000; Kim et al., 1999; Ru et al., 2015). When analyzing the recombination events of the DEL-CJ fragments in 293T-CJ cells by PCR amplification, we found that the D603N, D711N, or E965Q mutations abolished the catalytic activity of RAG1 (Fig. 2B), consistent with previous report (Swanson, 2001). In addition, we found that all three DDE mutants showed a weakened ability to form aggregates though at varying expression levels, with only approximately half of the observed cells forming bright puncta (Fig. 2C and S3A).

To examine the impact of nuclear localization on RAG1 aggregation, we next deleted the N-terminal 288-aa fragment containing the nuclear import signal (NLS) of human RAG1 ($RAG1^{\Delta N}$; Fig. 2A) as well as the IDR (intrinsically disorder regions) predicted by the IUPred2A software (Fig. S3B). The recombination activity of $RAG1^{\Delta N}$ was lower than that of wild-type RAG1, as shown by PCR amplification for recombination at the DEL-CJ fragments (Fig. S3C). The $RAG1^{\Delta N}$ was distributed within the entire cell and formed puncta in only 13% of observed cells (Fig. 2D). Moreover, the puncta were observed in the nucleus but not the cytoplasm of $RAG1^{\Delta N}$-expressing cells, even though the amount of $RAG1^{\Delta N}$ was higher in the cytoplasm (Fig. 2D and 2E). These observations indicate that the nuclear localization of RAG1 is important for its aggregation. To import $RAG1^{\Delta N}$ into nuclei, we fused the aggregation-free NLS fragment (320-328 aa) of human c-MYC to the N-terminus of $RAG1^{\Delta N}$ (Fig. 2A). We found that the amount of nuclear NLS-$RAG1^{\Delta N}$ significantly increased compared to $RAG1^{\Delta N}$, and the NLS-$RAG1^{\Delta N}$ was enriched in the nuclei in over 90% of observed cells (Fig. S3D and S3E). As a consequence, NLS-$RAG1^{\Delta N}$ formed nuclear puncta more frequently than $RAG1^{\Delta N}$ did (Fig. 2D). In the presence of RAG2, NLS-$RAG1^{\Delta N}$ supported V(D)J recombination at levels nearly equal to those of RAG1-GFP and much higher than $RAG1^{\Delta N}$ (Fig. S3C). Similar findings were obtained for cRAG1 with a larger region deleted compared to $RAG1^{\Delta N}$ (Fig. 2A; Fig. S3F and S3G). These data demonstrate that both the integrity of the DDE motif and nuclear location are important for RAG1 aggregation and its catalytic activity.

To further verify the above findings of RAG1 aggregation, we introduced the “OptoDroplet” system (Shin et al., 2017a) which has been used widely to study the intrinsic potential of given protein to form puncta in cells. In this system, the target protein is tagged with Cry2 protein (a photoreceptor protein) and forms small puncta rapidly after blue-light activation if it has the potential to aggregate. When the full-length RAG1(optoRAG1) or the NLS-cRAG1(optoNLS-cRAG1) fused with the Cry2 tag at the C terminus, spherical puncta were observed in nucleus after blue-light activation in the 293T-CJ cells, exhibiting an increasing trend over activation times (Fig. S4A and S4B). Of note, optoRAG1 or optoNLS-cRAG1 still had catalytic activity detected by PCR in the 293T-CJ cells (Fig. S4C). Moreover,

the optoRAG1 or optoNLS-cRAG1 puncta recovered rapidly and reached over 50% of initial intensity at the 60s time-point after photobleaching (Fig. S4D and S4E). Three DDE mutants with undetectable catalytic activity were also showed decreased percentages of cells containing puncta after blue-light activation (Fig. S4F and S4G). The findings in "OptoDroplet" system were highly consistent with overexpressed RAG1 in 293T-CJ cells.

RAG2 prevents RAG1 aggregation through direct interaction

As revealed in *vAbl* and 293T-CJ cells, RAG2 could prevent the RAG1 aggregation in cells (Fig. 1H and S1E). Similar finding was obtained in the "OptoDroplet" system (Fig. S5A). To gain further insight into the impact of RAG2 on RAG1 aggregation, we introduced G35V (104G>T), R39G (115A>G), C41W (123C>G), and R229E (685C>G, 686G>A) mutations at the structural interface of human RAG2 to prevent RAG1 interaction and thereby impair V(D)J recombination (Fig. 3A, 3B and S5B; Kim et al., 2015). The interaction-disrupted RAG2 mutant showed no aggregation in 293T-CJ cells, but was far less effective in blocking RAG1 aggregation than wild-type RAG2, with 13% of RAG2-coexpressed cells showing RAG1 aggregates compared to about 82% with the RAG2 mutant (Fig. 3C). The truncated cRAG2 protein supported V(D)J recombination and was also able to disrupt RAG1 aggregation efficiently (Fig. 3D, 3E, and S5C). Besides, the introduction of the four above mutations into cRAG2 resulted in an elevated level of RAG1 aggregation in the coexpression system, similar to full-length RAG2 (Fig. 3D and 3E).

The plant homeodomain (PHD) of RAG2 is important for the recruitment of RAG complex to H3K4me3 and a W453A mutation within the PHD domain impairs V(D)J recombination but not RAG1 binding (Fig. 3F; Liu et al., 2007; Lu et al., 2015). Either the PHD or the PHD(W453A) domains showed no impact on RAG1 aggregation (Fig. S5D). However, the RAG2(W453A) mutant exhibited a reduced level of RAG1 aggregation from 92% to 37%, slightly higher than 24% of wild-type RAG2 coexpressed cells, which could be due to higher enrichment in the cytoplasm of RAG2(W453A) than RAG2 (Fig. 3G and 3H). The separation of V(D)J recombination and RAG1/RAG2 binding by the W453A mutation implies the key role of direct RAG1/RAG2 interaction in preventing RAG1 aggregation.

RAG1 aggregation impairs V(D)J recombination

To further explore the role of RAG2 in regulating RAG1 aggregation for V(D)J recombination, we fused human *RAG1* or *RAG2* with a fragment of the *FUS* gene to force RAG1 or RAG2 to aggregate in U2OS cells. The FUS fragment showed a robust ability to form puncta (Murray et al., 2017; Fig. 4A and S5E). RAG2-FUS could efficiently form puncta in the entire cell body; the ectopically expressed RAG1-FUS showed tiny puncta in the cytoplasm, comparable in size to FUS puncta, while in the nucleus, RAG1-FUS formed significant larger puncta (Fig. 4A, S5E, and S5F). Moreover, if instead a mutant FUS fragment (FUS^{mut}) with abolished aggregation ability was used (Qamar et al., 2018), we found that RAG1-FUS^{mut} showed no aggregates in the cytoplasm but retained large puncta in the nuclei despite similar expression levels (Fig. 4A, S5F, and S5G). These data indicate that FUS can force RAG1 to aggregate and that the aggregation ability of RAG1 contributes to the formation of nuclear RAG1-FUS aggregates.

In the coexpression system of RAG1-FUS and RAG2-mCherry, RAG1-FUS still formed many small puncta in both the nuclei and cytoplasm, even in the presence of RAG2 (Fig. 4B). In contrast, RAG1-FUS^{mut} barely aggregated in the presence of RAG2 (Fig. 4B). Correspondingly, RAG1-FUS^{mut} showed approximately 2.3-fold higher catalytic activity than RAG1-FUS in the presence of RAG2 at the RSS pair of the DEL-CJ segment, as shown by either PEM-seq or PCR (Fig. 4C and 4D). Consistently, the FUS-tagged RAG2 did not stimulate the catalytic activity of RAG1 so efficiently as RAG2-FUS^{mut} did, with a difference of approximately 2.9 folds (Fig. 4C). These findings suggest that aggregated RAG1 possesses low recombination activity and that RAG2 can stimulate RAG activity by overcoming RAG1 aggregation.

Species-specific regulation of RAG1 by RAG2

Since RAG2 is vital for abolishing RAG1 aggregation to initiate V(D)J recombination, we next explored whether RAG2 from different species has the same capacity to stimulate human RAG1. For this purpose, we expressed human RAG1-GFP with human, mouse, or zebrafish RAG2-mCherry in 293T-CJ cells. Mouse RAG2 could also prevent the formation of punctate aggregates of human RAG1, but did so with significantly lower efficiency than human RAG2, while zebrafish RAG2 only modestly reduced human RAG1 aggregation when at a comparable transcription level (Fig. 4E, 4F and S6A and S6B). Similar findings were obtained in U2OS cells (Fig. S6D-F).

Then, we performed PEM-seq and PCR analysis of the RSS recombination events in 293T-CJ cells containing human RAG1 and RAG2 from different species at day 4 after transfection. Compared to human RAG2, mouse RAG2 could support recombination between the DEL-CJ RSS pair but at a much lower level; there was an approximately 6.6-fold decrease (Fig. 4G and S6C). Dramatically, the zebrafish RAG2 failed to stimulate human RAG1 and showed a similar recombination level to that of previously reported 293T-CJ cells containing only human RAG1 (Sun et al., 2020), and V(D)J recombination supported by this RAG2 was negatively correlated with the efficiency of RAG1 aggregation. These data show that human RAG2 is the better form for stimulating human RAG1 for V(D)J recombination, suggesting that coevolution of RAG1 and RAG2 in different species has resulted in species-specific optimized interactions.

Discussion

RAG1 and RAG2 work cooperatively to initiate V(D)J recombination. Biochemical studies have demonstrated that RAG2 enhances the specificity of DNA binding by RAG1 and greatly stimulates RAG1 catalytic activity (Gellert, 2002; Swanson, 2004). We examined the mRNA levels of RAG1 in various cells and found that the V(D)J recombination in nonlymphoid cells is well tolerated by high RAG concentration (Fig. 1F and S2F). We here report that RAG1 aggregates in both B cells and nonlymphoid cells and RAG1 aggregation hinders V(D)J recombination. However, we also noticed that not all the RAG1 aggregates in the nuclei while diffused nuclear RAG1 is still detectable. Why is the remaining diffused RAG1 not sufficient to induce V(D)J recombination? Previous data show that RAG1 and RAG2 bind to H3K4me3 at more than thousands of genome-wide sites, though highly

enriched at the recombination center in antigen receptor loci (Teng et al., 2015). The RAG-enriched antigen receptor loci may show a cooperative effect and are highly sensitive to the loss of RAG. The dose-dependent sensitivity phenomenon has also been found for super-enhancers that BRD4-enriched super-enhancers are much more sensitive to the BRD4 inhibitor than typical enhancers (Loven et al., 2013). Therefore, aggregation-induced reducement of diffused RAG1 may greatly impair V(D)J recombination in antigen receptor loci.

In mice, the N-terminus of RAG1 is very important for enabling RAG1 to localize in the nucleolus and then translocate outside of it in G1 phase (Brecht et al., 2020). Our findings show that human RAG1 can aggregate in nuclei, depending on the N-terminus as well (Fig. 2D). We were able to rescue the formation of human RAG1 puncta by replacing the original N-terminus with the NLS of human c-MYC protein (Fig. 2D), indicating that the nuclear location rather than a specific N-terminal region peptide sequence is important for RAG1 aggregation. The minorly different nuclear phenotypes of human and mouse RAG might be due to the use of nonlymphoid cells or the different ability of each species version of RAG1 to form aggregates. With these regards, we found that full-length mouse RAG1 did not easily form puncta when overexpressed in nonlymphoid cells, suggesting a weak ability of mouse RAG1 to form aggregates (Fig. S6G). In addition, we found that human RAG1 can also be located in the nucleolus of *vAbl* pro-B cells or when coexpressed with the nucleolus marker fibrillarin-mCherry (Fig. 1G and S6H). These data suggest that as it evolved from a DDE transposon, the RAG1 might have adapted different regulation mechanisms in different species (Bischerour et al., 2009; Zhang et al., 2019). The origin of RAG2 is currently unknown (Carmona and Schatz, 2017; Morales Poole et al., 2017). However, RAG2 should occur with RAG1 early in their origins, and they should have evolved simultaneously to enable RAG2 to regulate RAG1. For these reasons, RAG2 from mice or zebrafish cannot fully activate human RAG1 due to decreased ability to bind human RAG1 and abolish RAG1 aggregation.

RAG2 is regulated by cell cycle-dependent degradation to restrict V(D)J recombination in the G1 phase of developing lymphocytes (Li et al., 1996; Lin and Desiderio, 1994). When lymphocytes enter the G1 phase, the newly produced RAG2 directly interacts with inert RAG1 and facilitates the “jailbreak” of RAG1 to subsequentially initiate the V(D)J recombination (Fig. 4H). The nucleolus is a phase separation center *per se* and can sequester some nucleolar proteins to enable protein condensation (Feric et al., 2016), so RAG1 may aggregate with help of nucleolus rather than self-aggregates (Fig. 1G and S6H). In this context, RAG1 only forms puncta in the nuclei, even in situations in which the amount of RAG1 protein in the cytoplasm is slightly higher (Fig. 2D and 2E). Besides the nucleolus as “jail”, other nuclear membraneless organelles or factors may also be involved in holding RAG1 aggregates in the absence of RAG2 (Yang et al., 2020), which needs to be further explored. The jailbreak model ensures that robust V(D)J recombination only occurs in the presence of RAG2 in the G1 phase when c-NHEJ dominates the DSB repair pathway. Mutations of the interacting residues of RAG2 restore RAG1 aggregation and reduce V(D)J recombination, as shown in our experiment and previous report (Fig. 3A-E and Kim et al., 2015). The interaction of RAG1 and RAG2 might change the conformation of RAG1 or block nonspecific multivalent interactions between RAG1 dimers and thereby prevent RAG1

aggregation. Moreover, many SCID-associated RAG2 variants have mutations on the RAG1/RAG2 interface (Kim et al., 2015). Interestingly, mutation of the catalytic DDE carboxylates of RAG1 also reduces the aggregation level of human RAG1, indicating that either the integrity of the DDE motif (which could affect protein conformation) or the catalytic activity of RAG1 is crucial for RAG1 aggregation. Further investigation will be required to resolve this issue.

RAG1 aggregation might be an efficient mechanism for human cells to regulate V(D)J recombination and protect genome integrity. When entering the S phase, degrading only RAG2 is a more economical and efficient form of regulation than degrading both RAG1 and RAG2. Following the timely degradation of RAG2, RAG1 can either undergo slow degradation or can persist in the cells. The RAG complex has the potential ability to target genome-wide CAC motifs (Hu et al., 2015), and RAG1 aggregation could reduce catalytic activity to prevent the remaining RAG1 from damaging the genome. Consistently, mutation of RAG2 to eliminate cell cycle-dependent degradation of RAG2 resulted in increased genome instability, chromosomal translocations, and lymphoid malignancies in mice (Zhang et al., 2011). Whether regulating the noncatalytic partner is a general mechanism for restraining effector proteins, such as nucleases or kinases, remains to be explored. There are some examples that effector proteins have a partner with no distinct functional domain but are crucial for catalytic activity, such as ATRIP for ATR in DNA repair (Cortez et al., 2001) and CDC24 for DNA2 during the maturation of Okazaki fragments (Tanaka et al., 2004).

STAR Methods

RESOURCE AVAILABILITY

Lead contact—Futher information and requests for materials and resources should be addressed to Jiazhi Hu (hujz@pku.edu.cn).

Materials availability—All plasmids and cell lines generated in this study are available from the lead contact with a completed Materials transfer agreement.

Data and code availability—All sequencing data has been deposited at GEO (PRJNA758251). Beside, the processed data and the original western blot images or PCR gels have been deposited at Mendeley Data (DOI: 10.17632/v6hkpgbnpj.1). All the datas are publicly available as of the date of publication. The DOIs are listed in the Key resources table. This paper does not report original code. Any additional information required to reanalyze the data and microscopy data reported in this paper is available from the lead contact upon request.

EXPERIMENTAL MODEL AND SUBJECT DETAILS

Cell lines and cell culture—The NALM-6 cell was cultured in RPMI 1640 (Corning) medium supplemented with 15% FBS (ExCell) (Grawunder et al., 1998). The HEK293T and U2OS cells (ATCC, Washington, USA) were cultured in DMEM (Corning) supplemented with 10% FBS as described. The *vAbl* pro-B cells were cultured in RPMI 1640 (Corning)

medium supplemented with 15% FBS (ExCell). Cells were incubated at 37 °C with 5% CO_2.

METHOD DETAILS

Generation the 293T-CJ cell line by retrovirus—To efficiently and conveniently evaluate the catalytic activity of different RAG complex, we integrated a pair of 12/23 RSSs in the pMX vector and introduced it into HEK293T cells by cotransfected with a helper plasmid-Psi-amphol for two days following the $Ca_3(PO_4)_2$ coprecipitation approach. The integrated sequences in HEK293T cells were detected by PCR and Sanger sequencing. Primers are listed in Key resources table.

Generation of hRAG1-maxGFP knocked-in *vAbl* pro-B cell.—The Rag2-deficient *vAbl* mouse pro-B cells were generated as previously described (Hu et al., 2015). The mouse *Rag1* genes were knocked out by CRISPR-Cas9 in a pX330 vector (Addgene, 42230) with indicated primers in Key resources table. The guide RNAs were designed at the start of exon 2 and 3' UTR of mouse *Rag1*. Plasmids were transfected into cells at 2 μg plasmids/million cells via nucleofection with an SF Cell Line 4D-Nucleofecto X Kit following the manufacturer's instructions of Lonza 4D (Lonza, GA, USA; DN100 program). Cas9-transfected cells (mCh-positive) were sorted by FACS (MoFlo) into 96-well plates to incubate for 10 days; PCR and RT-qPCR were then used to check the deletions and transcription level of *Rag1*. Cells with homozygous *Rag1* deletion were obtained for the next steps.

We then used guide RNAs listed in Key resources table to knock in the *hRAG1-maxGFP* gene. Donor DNAs containing the *hRAG1* coding gene, a 3x(GGGS) linker, the *maxGFP* sequence and two 700bp homology arms were cloned into a T vector pMK. Then the donor plasmid and pX330-based CRISPR-Cas9 plasmid were co-transfected by nucleofection as described above. Cells with high levels of mCherry (in the Cas9 plasmid) were sorted by MoFlo at 48h after transfection. The sorted cells were then treated by 3 μM STI-571 for 36 h (Hu et al., 2015) prior to being sorted by maxGFP into a 96-well plate. The *hRAG1-maxGFP* knock-in *vAbl* pro-B cells were confirmed with integron-spanning PCR and Sanger sequencing and FACS. PCR primers were listed in Key Resources Table.

We also verified the knock-in of RAG1-maxGFP at Rag1 locus by Southern blot. Briefly, 10 μg genomic DNA was digested overnight with 100U BamHI-HF(NEB) and then subject to 0.8% agarose gel. The 684-bp maxGFP labelled by α-32P-dCTP was used as the detection probe. Standard Southern blot procedures were used.

PEM-seq—Low-frequency coding joins generated by RAG recombinase could be hardly distinguished by PCR. Therefore, we used the PEM-seq method to monitor RAG1or RAG complex-generated breaks at RSS sites and a bait site cleaved by RAG complex at a *bona fide* 23 RSS were used. Translocation junctions between the 23 RSS cutting site and other break sites were retrieved. The plasmids were transfected into 293T-CJ cells with the $Ca_3(PO_4)_2$ coprecipitation approach and the cells were cultured for four days. Cells were harvested, washed with PBS, and subjected to genomic DNA extraction with 1:100 Proteinase K digestion. Next, PEM-Seq libraries were generated as previously described

(Yin et al., 2019). In brief, primer extension was performed with biotinylated primers as listed in Key Resources Table. Enriched biotinylated PCR products were ligated to the bridge adapters, and two subsequent PCR steps were performed to amplify products for sequencing. Sequences were mapped to the hg19 reference genome via the SuperQ pipeline (Yin et al., 2019).

PEM-seq reads were processed as follows. For initial reads preprocessing, both Illumina 2000 adapter sequences and ending low-quality sequences (QC<30) were trimmed by cutadapt package (http://cutadapt.readthedocs.io/en/stable/); remaining reads shorter than 25 bp were discarded. Then reads were de-multiplexed using fastq-multx (https://github.com/brwnj/fastq-multx) to distinguish the index. For reads alignment and clustering, we adapted the corresponding pipeline used in SuperQ to perform reads mapping and break site detection. Note that we used the hg19 genome (modified by DEL-CJ segments insertion at chromosome 1) as reference. Uniquely mapped reads were filtered program as LAM-HTGTS (Hu et al., 2016) did but all the duplicates were kept for the following precession. A molecular barcode (RMB: used to distinguish different reads to remove the replicates from PCR amplification) clustering algorithm38 was adapted to our analysis. The deletional coding joins between the adjacent sequences to 12 and 23 RSSs were counted as V(D)J recombination events. Besides, reads without any detected mutations around break point were identified as germline sequences. Taking uncut primer control and sequencing depth into account, the efficiency of V(D)J recombination for different RAG complex was calculated as bellow: V(D)J recombination frequency (%) = Coding joins/total identified fragments %.

Fluorescence imaging—293T-CJ cells were incubated on 35 mm glass-bottom plates (Cellvix). Cells at approximately 50% confluency were transfected by the $Ca_3(PO_4)_2$ coprecipitation approach with 0.8 μg plasmids for 18 h. Then, the DMEM for culture was replaced with 2 ml of PBS containing 10 μg/mL Hoechst 33342 for 10 min in a CO_2 incubator at 37°C. Cells were washed with prewarmed DMEM twice before imaging. Live-cell images were obtained by a Nikon UltraView VoX microscope (Nikon), equipped with an incubation device (5% CO2, 37C°). The images were taken under the following conditions: 100x oil objective lens, laser sensitivity 111, laser intensity 10.5%, exposure time 50 ms.

U2OS cells at approximately 50% confluency were transfected by X-tremeGENE™ HP DNA transfection reagent following the manufacturer's instructions with 0.8 μg plasmids for 24 h. Living cell images were performed by a Nikon UltraView VoX microscope with an incubation device at 5% CO_2 and 37°C. The images were taken under the following conditions: 100x oil objective lens, laser sensitivity 111, laser intensity 4.5%, exposure time 100 ms.

NALM-6 cells were transfected by nucleofection with an SF Cell Line 4D-Nucleofecto X Kit following the manufacturer's instructions for Lonza 4D (Lonza, GA, USA; DN100 program) with 2 μg plasmids/million cells. Images were taken at 24 h after nucleofection.

vAbl mouse pro-B cells were treated with 3 μM STI-571 at 2 million cells/mL density for 48h. At 24h, the cells were transferred to poly-lysine pre-treated 35 mm glass-bottom plates for another 24h. Living cells images were performed by a Nikon UltraView VoX microscope with an incubation device at 5% CO_2 and 37°C. The images were taken at the following conditions: 100x oil objective lens, laser sensitivity 170, laser intensity 80%, exposure time 300 ms.

All images were acquired using the ImageJ1.48 software package. For counting the numbers and measuring the sizes of puncta, we converted the fluorescence microscopy images to 8 bit and then set the threshold based on the intensity of background before automatic counting performed by ImageJ1.48. The “Median filter of radius 2” and “Huang auto threshold method” of ImageJ1.48 were used for calculating puncta number. The 0.5 to 1.0 of the circularity range for puncta was exclude from calculated. The diameters of puncta were measured by the “line” tool embedded in the ImageJ1.48, source is listed in Key Resources Table.

FRAP assay—U2OS cells at approximately 50% confluency were transfected by X-tremeGENE™ HP DNA transfection reagent following the manufacturer’s instructions with 0.8 μg of plasmids for 24 h, and the FRAP assay was also performed at this time point with a 488 nm laser. The bleaching conditions were performed as previously described: 100x oil objective lens, radius 0.8 μm, laser sensitivity 111, laser intensity 4.5%, and exposure time 1 s. Cells were imaged every 5 s over a 2-min duration. The instruction came from the Brangwynne lab (Shin et al., 2017b).

ImageJ1.48 software package was used to quantify the fluorescence intensity. Briefly, background intensity was first subtracted and the relative intensity of the bleaching area at different time points after bleaching was measured and normalized to pre-bleaching intensity.

Assessment of RAG1 and RAG2 expression levels—293T-CJ cells were harvested at 18 h after transfection. Total RNA was extracted as follows. Approximately 1 million cells were collected in a sterile 1.5 mL tube (Axygen) for each extraction, and then they were pelleted by centrifugation (300 g, 5 min, 4°C). The cell pellet was immediately resuspended in 500 μL of TRIzol reagent (GIBCO-BRL, Gaithersburg, MD) for 5 min at room temperature. Briefly, 200 μL chloroform was added to the sample and vortexed for 30 s. Polysaccharides, membranes, and unlysed cells were eliminated by centrifugation (13,000 g, 10 min, 4°C). The supernatant (approximately 250 μL) was mixed with an equal volume of isopropanol, incubated at −20°C for 1 h, and centrifuged (13,000 g, 10 min, 4°C). Then, the pellet of total RNA was washed with 1 mLof 70% ethanol twice and dried for 30 min at room temperature. The RNA pellet was resuspended in 20 μL of RNA-free sterile water and stored at −80°C for further analysis. The RNA was subjected to cDNA synthesis with a FastQuant RT Kit (Tiangen Biotech). Real-time PCR was performed using a LightCycler 96 PCR system (BioRad) to compare the expression levels of different samples. The expression level of *GAPDH* was used as a control of input cDNA. In addition, the q-PCR primers are listed in Key Resources Table.

Western blotting assay—293T-CJ, NALM-6 (*RAG1*$^{-/-}$), and U2OS cells were harvested at 24 h after transfection by digestion of Trypsin, except the NALM-6 (*RAG1*$^{-/-}$) cells. Then the cells were washed by PBS buffer twice. So, lysed cells by adding SDS buffer (500 μL for a 10 cm diameter plate). Immediately scrape the cells off the plate and transfer the extract to a microcentrifuge tube, which is always located on ice. Next, we sonicated cells for 30 s 10 times to complete cell lysis and sheared genomic DNA. Last, we heated the samples for 5 min at 95°C. After sample preparation, we loaded 20 uL samples onto SDS-PAGE gel for western blot.

Flow cytometry—Human cord blood were diluted by PBS buffer and the light-density mononuclear cells were obtained after centrifugation by treated with the Lymphoprep kit. And then the mononuclear cells were stained by FITC-CD34 (Biolegend), APC-CD10 (Biolegend) and BUV395-CD19 (BD, bioscience) at 4°C for 30min in PBS with 2% FBS. Cells were washed twice after incubation with three antibodies, and then suspended by PBS with 2% FBS at 5 million cells/mL for single cell sorting. The stained mononuclear cells were sorted and analysed on a FACSAria III SORP (Becton Dickinson), which equiped with 355nm, 488nm, and 630nm laser. B cells expressed CD34, CD10 and CD19 simultaneously were sorted as human pro-B cells as previous reported (Reynaud et al., 2003; McWilliams et al., 2013). Antibodies were listed at Key Resources Table. FlowJo_V10 were used for data analysis.

Single-cell qRT-PCR for human pro-B cells—Single-cell suspension from human cord blood were sorted at lysis buffer into a 96-well plate as 1 cell/well on the ice. After sorting, the PCR plate was vortexed and centrifuged (4°C, 1000g, 60s) before performing denaturation at 72°C for 120 seconds. Then the RT mix (Maxima H minus reverse transcriptase mix ,ThermoScientific) and Ribonuclease inhibitor (Invitrogen) and 5% PEG8000 were added to the cell lysate. This mix was vortexed and centrifuged to preparing the reverse transcription at 42°C for 90 min, 85°C for 5 min and 10°C forever.

Next, the single-cell qPCR was performed in sealed 96 well plates with SYBR GreenmasterMix in a Light Cycler (Bio-Rad). Single-cell qPCR reations were prepared into a final volume of 25 μl containing whole 10 μl cDNA mix and 12.5 μl of SYBR Greenmaster mix in the presence of primers at 0.4 μM. After the cDNA amplification, the cycle number of target genes was plotted against the normalized fluorescence intensity to make the PCR amplification efficiencies visualized.

To make sure of the reliability and consistency of the single-cell qRT-PCR, we performed several control as following: 1) We test the expression level of house keeping gene glyceraldehyde 3-phosphate dehydrogenase (GADPH) in different cell lines and we found the expression levels of *GAPDH* or *Gapdh* in human 293T-CJ or the mouse *vAbl* pro-B cells with different genotypes were consistent. 2) The melting curves were detected for each of the samples and checked to ensure only one melting curve was observed from the PCR reaction. 3) Primer controls were performed for each primers to make sure the primer-dimer products has less amplification efficiency in cDNA amplification products (Jeong et al., 2016). All the primers for single-cell qRT-PCR were borrowed from the previous reports (McWilliams et al., 2013) and listed at Key Resources Table.

QUANTIFICATION AND STATISTICAL ANALYSIS

Data are presented as mean ± standard deviation (SD) values. ImageJ1.48 was used for quantitative measurements. And the SuperQ pipeline (Yin et al., 2019) was used for high-throughput sequencing analysis. The statistical details of experiments can be found in the related “METHOD DETAILS” section.

Prism8 was used for data analysis, and the number of replicates and statistical test procedures are indicated in the figure legends. FlowJo_V10 was used for analysis the data of FACS according the instruction of BD company.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Dr. Heng Ru for RAG1 and RAG2 plasmids and his insightful comments. We thank Dr. Dong Li and Dr. Hsiang-Ying Lee for sharing us human cord blood samples and the Flow Cytometry Core at the National Center for Protein Sciences, Peking University, particularly Hongxia Lv and Yinghua Guo, for their technical help. This work was supported by the National Key R&D Program of China (2017YFA0506700 to J.H.), an NSFC grant (31771485 to J.H.) and NIH grant AI32524 to D.G.S. J.H. is an investigator at the PKU-TSU Center for Life Sciences. J.H. is also a Bayer Investigator.

References

Alt FW, Zhang Y, Meng FL, Guo C, and Schwer B (2013). Mechanisms of programmed DNA lesions and genomic instability in the immune system. Cell 152, 417–429. [PubMed: 23374339]

Bischerour J, Lu C, Roth DB, and Chalmers R (2009). Base flipping in V(D)J recombination: insights into the mechanism of hairpin formation, the 12/23 rule, and the coordination of double-strand breaks. Mol. Cell. Biol 29, 5889–5899. [PubMed: 19720743]

Brecht RM, Liu CC, Beilinson HA, Khitun A, Slavoff SA, and Schatz DG (2020). Nucleolar localization of RAG1 modulates V(D)J recombination activity. Proc. Natl. Acad. Sci. USA 117, 4300–4309. [PubMed: 32047031]

Bredemeyer AL, Sharma GG, Huang CY, Helmink BA, Walker LM, Khor KC, Nuskey B, Sullivan KE, Pandita TK, Bassing CH, et al. (2006). ATM stabilizes DNA double-strand-break complexes during V(D)J recombination. Nature 442, 466–470. [PubMed: 16799570]

Carmona LM, Fugmann SD, and Schatz DG (2016). Collaboration of RAG2 with RAG1-like proteins during the evolution of V(D)J recombination. Genes Dev. 30, 909–917. [PubMed: 27056670]

Carmona LM, and Schatz DG (2017). New insights into the evolutionary origins of the recombination-activating gene proteins and V(D)J recombination. FEBS J. 284, 1590–1605. [PubMed: 27973733]

Chaumeil J, Micsinai M, Ntziachristos P, Roth DB, Aifantis I, Kluger Y, Deriano L, and Skok JA (2013). The RAG2 C-terminus and ATM protect genome integrity by controlling antigen receptor gene cleavage. Nat. Commun 4, 2231. [PubMed: 23900513]

Chen X, Cui Y, Wang H, Zhou ZH, Gellert M, and Yang W (2020). How mouse RAG recombinase avoids DNA transposition. Nat. Struct. Mol. Biol 27, 127–133. [PubMed: 32015553]

Cortez D, Guntuku S, Qin J, and Elledge SJ (2001). ATR and ATRIP: partners in checkpoint signaling. Science 294, 1713–1716. [PubMed: 11721054]

Feric M, Vaidya N, Harmon TS, Mitrea DM, Zhu L, Richardson TM, Kriwacki RW, Pappu RV, and Brangwynne CP (2016). Coexisting liquid phases underlie nucleolar subcompartments. Cell 165, 1686–1697. [PubMed: 27212236]

Fugmann SD, Lee AI, Shockett PE, Villey IJ, and Schatz DG (2000). The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu. Rev. Immunol 18, 495–527. [PubMed: 10837067]

Gellert M (2002). V(D)J recombination: RAG proteins, repair factors, and regulation. Annu. Rev. Biochem 71, 101–132. [PubMed: 12045092]

Grawunder U, Zimmer D, Fugmann S, Schwarz K, and Lieber MR (1998). DNA ligase IV is essential for V(D)J recombination and DNA double-strand break repair in human precursor lymphocytes. Mol. Cell 2, 477–484. [PubMed: 9809069]

Grundy GJ, Ramon-Maiques S, Dimitriadis EK, Kotova S, Biertumpfel C, Heymann JB, Steven AC, Gellert M, and Yang W (2009). Initial stages of V(D)J recombination: the organization of RAG1/2 and RSS DNA in the postcleavage complex. Mol. Cell 35, 217–227. [PubMed: 19647518]

Hagemann-Jensen M, Ziegenhain C, Chen P, Ramsköld D, Hendriks GJ, Larsson AJM, Faridani OR, and Sandberg R (2020). Single-cell RNA counting at allele and isoform resolution using Smart-seq3. Nat. Biotechnol 38, 708–714. [PubMed: 32518404]

Hnisz D, Shrinivas K, Young RA, Chakraborty AK, and Sharp PA (2017). A phase separation model for transcriptional control. Cell 169, 13–23. [PubMed: 28340338]

Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW, and Frock RL (2016). Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. Nat. protoc 11, 853–871. [PubMed: 27031497]

Hu J, Zhang Y, Zhao L, Frock RL, Du Z, Meyers RM, Meng FL, Schatz DG, and Alt FW (2015). Chromosomal loop domains direct the recombination of antigen receptor genes. Cell 163, 947–959. [PubMed: 26593423]

Jeong JH, JaeWoo Y, Chua S, and Jo YH (2016). Single-cell gene expression analysis of cholinergic neurons in the arcuate nucleus of the hypothalamus. PLoS One 11, 1–14.

Ji Y, Resch W, Corbett E, Yamane A, Casellas R, and Schatz DG (2010). The in vivo pattern of binding of RAG1 and RAG2 to antigen receptor loci. Cell 141, 419–431. [PubMed: 20398922]

Kim DR, Dai Y, Mundy CL, Yang W, and Oettinger MA (1999). Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. Genes Dev. 13, 3070–3080. [PubMed: 10601033]

Kim MS, Lapkouski M, Yang W, and Gellert M (2015). Crystal structure of the V(D)J recombinase RAG1-RAG2. Nature 518, 507–511. [PubMed: 25707801]

Kuo TC, and Schlissel MS (2009). Mechanisms controlling expression of the RAG locus during lymphocyte development. Curr. Opin. Immunol 21, 173–178. [PubMed: 19359154]

Lee J, and Desiderio S (1999). Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. Immunity 11, 771–781. [PubMed: 10626899]

Li P, Banjade S, Cheng HC, Kim S, Chen B, Guo L, Llaguno M, Hollingsworth JV, King DS, Banani SF, et al. (2012). Phase transitions in the assembly of multivalent signalling proteins. Nature 483, 336–340. [PubMed: 22398450]

Li Z, Dordai DI, Lee J, and Desiderio S (1996). A conserved degradation signal regulates RAG-2 accumulation during cell division and links V(D)J recombination to the cell cycle. Immunity 5, 575–589. [PubMed: 8986717]

Lieber MR (2019). Transposons to V(D)J recombination: evolution of the RAG reaction. Trends Immunol. 40, 668–670. [PubMed: 31307890]

Lin WC, and Desiderio S (1994). Cell cycle regulation of V(D)J recombination-activating protein RAG-2. Proc. Natl. Acad. Sci. USA 91, 2733–2737. [PubMed: 8146183]

Liu M, Zhang W, Xin C, Yin J, Shang Y, Ai C, Li J, Meng F, and Hu J (2021). Global detection of DNA repair outcomes induced by CRISPR-Cas9. Nucleic Acids Res. 49, 8732–8742. [PubMed: 34365511]

Liu Y, Subrahmanyam R, Chakraborty T, Sen R, and 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. [PubMed: 17936034]

Loven J, Hoke HA, Lin CY, Lau A, Orlando DA, Vakoc CR, Bradner JE, Lee TI, and Young RA (2013). Selective inhibition of tumor oncogenes by disruption of super-enhancers. Cell 153, 320–334. [PubMed: 23582323]

Lu C, Ward A, Bettridge J, Liu Y, and Desiderio S (2015). An autoregulatory mechanism imposes allosteric control on the V(D)J recombinase by histone H3 methylation. Cell Rep 10, 29–38. [PubMed: 25543141]

McWilliams L, Su K-Y, Liang X, Liao D, Floyd S, Amos J, Moody MA, Kelsoe G, and Kuraoka M (2013). The human fetal lymphocyte lineage: identification by CD27 and LIN28B expression in B cell progenitors. J. Leukoc. Biol 94, 991–1001. [PubMed: 23901121]

Minowada J, Janossy G, Greaves MF, Tsubota T, Srivastava BI, Morikawa S, and Tatsumi E (1978). Expression of an antigen associated with acute lymphoblastic leukemia in human leukemia-lymphoma cell lines. J. Natl. Cancer Inst 1269–1277.

Morales Poole JR, Huang SF, Xu A, Bayet J, and Pontarotti P (2017). The RAG transposon is active through the deuterostome evolution and domesticated in jawed vertebrates. Immunogenetics 69, 391–400. [PubMed: 28451741]

Murray DT, Kato M, Lin Y, Thurber KR, Hung I, McKnight SL, and Tycko R (2017). Structure of FUS protein fibrils and its relevance to self-assembly and phase separation of low-complexity domains. Cell 171, 615–627 e616. [PubMed: 28942918]

Oettinger MA, Schatz DG, Gorka C, and Baltimore D (1990). RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination. Science 248, 1517–1523. [PubMed: 2360047]

Qamar S, Wang G, Randle SJ, Ruggeri FS, Varela JA, Lin JQ, Phillips EC, Miyashita A, Williams D, Strohl F, et al. (2018). FUS phase separation is modulated by a molecular chaperone and methylation of arginine cation-pi interactions. Cell 173, 720–734 e715. [PubMed: 29677515]

Qiu JX, Kale SB, Yarnell Schultz H, and Roth DB (2001). Separation-of-function mutants reveal critical roles for RAG2 in both the cleavage and joining steps of V(D)J recombination. Mol. Cell 7, 77–87. [PubMed: 11172713]

Reynaud D, Lefort N, Manie E, Coulombel L, and Levy Y (2003). In vitro identification of human pro-B cells that give rise to macrophages, natural killer cells, and T cells. Blood 101, 4313–4321. [PubMed: 12560235]

Robinson JT, Thorvaldsdóttir H, Wenger AM, Zehir A, and Mesirov JP (2017). Variant review with the integrative genomics viewer. Cancer Res. 77, e31–e34. [PubMed: 29092934]

Ru H, Chambers MG, Fu TM, Tong AB, Liao M, and Wu H (2015). Molecular mechanism of V(D)J recombination from synaptic RAG1-RAG2 complex structures. Cell 163, 1138–1152. [PubMed: 26548953]

Ru H, Mi W, Zhang P, Alt FW, Schatz DG, Liao M, and Wu H (2018a). DNA melting initiates the RAG catalytic pathway. Nat. Struct. Mol. Biol 25, 732–742. [PubMed: 30061602]

Ru H, Zhang P, and Wu H (2018b). Structural gymnastics of RAG-mediated DNA cleavage in V(D)J recombination. Curr. Opin. Struct. Biol 53, 178–186. [PubMed: 30476719]

Sabari BR, Dall'Agnese A, Boija A, Klein IA, Coffey EL, Shrinivas K, Abraham BJ, Hannett NM, Zamudio AV, Manteiga JC, et al. (2018). Coactivator condensation at super-enhancers links phase separation and gene control. Science 361, eaar3958 [PubMed: 29930091]

Schatz DG, and Ji Y (2011). Recombination centres and the orchestration of V(D)J recombination. Nat. Rev. Immunol. 11, 251–263. [PubMed: 21394103]

Schatz DG, Oettinger MA, and Baltimore D (1989). The V(D)J recombination activating gene, RAG-1. Cell 59, 1035–1048. [PubMed: 2598259]

Schatz DG, and Swanson PC (2011). V(D)J recombination: mechanisms of initiation. Annu. Rev. Genet 45, 167–202. [PubMed: 21854230]

Schneider CA, Rasband WS, and Eliceiri KW (2012). NIH image to ImageJ: 25 years of image analysis. Nat. Methods 9, 671–675. [PubMed: 22930834]

Schwarz K, Gauss GH, Ludwig L, Pannicke U, Li Z, Lindner D, Friedrich W, Seger RA, Hansen-Hagge TE, Desiderio S, et al. (1996). RAG mutations in human B cell-negative SCID. Science 274, 97–99. [PubMed: 8810255]

Shin Y, Berry J, Pannucci N, Haataja MP, Toettcher JE, and Brangwynne CP (2017a). Spatiotemporal control of intracellular phase transitions using light-activated optoDroplets. Cell 168, 159–171.e114. [PubMed: 28041848]

Shin Y, and Brangwynne CP (2017b). Liquid phase condensation in cell physiology and disease. Science 357, eaaf4382. [PubMed: 28935776]

Shinkai Y, Rathbun G, Lam KP, Oltz EM, Stewart V, Mendelsohn M, Charron J, Datta M, Young F, Stall AM, et al. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. Cell 68, 855–867. [PubMed: 1547487]

Su X, Ditlev JA, Hui E, Xing W, Banjade S, Okrut J, King DS, Taunton J, Rosen MK, and Vale RD (2016). Phase separation of signaling molecules promotes T cell receptor signal transduction. Science. 352, 595 – 599. [PubMed: 27056844]

Sun A, Xu K, Liu H, Li H, Shi Y, Zhu X, Liang T, Li X, Cao X, Ji Y, et al. (2020). The evolution of zebrafish RAG2 protein is required for adapting to the elevated body temperature of the higher endothermic vertebrates. Sci. Rep 10, 4126. [PubMed: 32139788]

Swanson PC (2001). The DDE motif in RAG-1 is contributed in trans to a single active site that catalyzes the nicking and transesterification steps of V(D)J recombination. Mol. Cell. Biol 21, 449–458. [PubMed: 11134333]

Swanson PC (2004). The bounty of RAGs: recombination signal complexes and reaction outcomes. Immunol. Rev 200, 90–114. [PubMed: 15242399]

Tanaka H, Ryu GH, Seo YS, and MacNeill SA (2004). Genetics of lagging strand DNA synthesis and maturation in fission yeast: suppression analysis links the Dna2-Cdc24 complex to DNA polymerase delta. Nucleic Acids Res. 32, 6367–6377. [PubMed: 15576681]

Teng G, Maman Y, Resch W, Kim M, Yamane A, Qian J, Kieffer-Kwon KR, Mandal M, Ji Y, Meffre E, et al. (2015). RAG represents a widespread threat to the lymphocyte genome. Cell 162, 751–765. [PubMed: 26234156]

Yang P, Mathieu C, Kolaitis RM, Zhang P, Messing J, Yurtsever U, Yang Z, Wu J, Li Y, Pan Q, et al. (2020). G3BP1 is a tunable switch that triggers phase separation to assemble stress granules. Cell 181, 325–345. [PubMed: 32302571]

Yin J, Liu M, Liu Y, Wu J, Gan T, Zhang W, Li Y, Zhou Y, and Hu J (2019). Optimizing genome editing strategy by primer-extension-mediated sequencing. Cell Discov. 5, 18. [PubMed: 30937179]

Zhang L, Reynolds TL, Shan X, and Desiderio S (2011). Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. Immunity 34, 163–174. [PubMed: 21349429]

Zhang W, Yin J, Zhang-Ding Z, Xin C, Liu M, Wang Y, Ai C, and Hu J (2021). In-depth assessment of the PAM compatibility and editing activities of Cas9 variants. Nucleic Acids Res. 1–11. [PubMed: 33275144]

Zhang Y, Cheng TC, Huang G, Lu Q, Surleac MD, Mandell JD, Pontarotti P, Petrescu AJ, Xu A, Xiong Y, et al. (2019). Transposon molecular domestication and the evolution of the RAG recombinase. Nature 569, 79–84. [PubMed: 30971819]

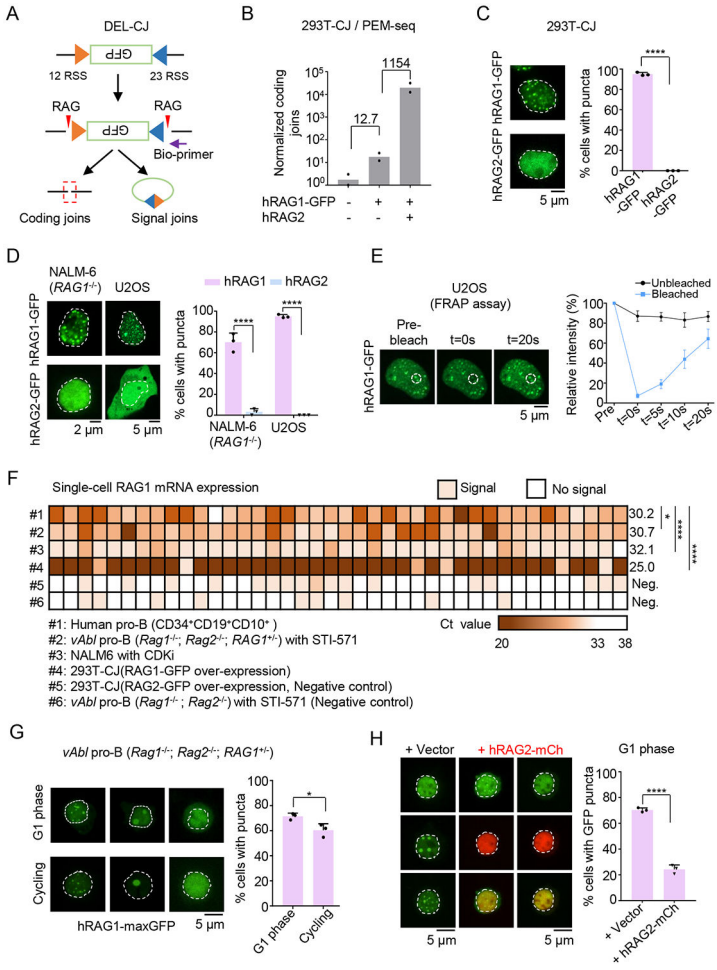

Figure 1. RAG1 undergoes aggregation in lymphocytes and nonlymphoid cells.

(A) Integration of *bona fide* 12 and 23 RSS pair in HEK293T cells (293T-CJ), which can be recognized and cleaved by the RAG complex. Purple arrow indicates the position of the bio-primer for PEM-seq. Red triangles indicate the RAG complex. Green box indicates the inverted GFP segment.

(B) Coding joins detected by the PEM-seq in 293T-CJ cells. Numbers of coding joins were normalized to total mapping reads including uncut and deletional coding joins. Fold changes are indicated. Replicates n=2.

(C) Representative microscopy images of indicated proteins in 293T-CJ cells (left). Bar graph (right) shows the percentage of cells with puncta in the total observed cells; replicates n=3; t-test; ****, p<0.0001. Scale bar: 2 μm (left) and 5 μm (right).

(D) Representative microscopy images of hRAG1-GFP in NALM-6 (*RAG1*$^{-/-}$) and U2OS cells. Percentages on the right; replicates n=3; t-test; ****, p<0.0001. Scale bar: 5 μm.

(E) Representative images of hRAG1-GFP in U2OS cells for FRAP experiment. The dashed line indicates the circled bleached dots. Bleaching occurred at t=0 s, and recovered intensity was detected at t=20 s (left). And the intensity recovery lines for hRAG1-GFP puncta (right). The background-subtracted fluorescence intensities are normalized by pre-bleach values. Cell number n=15. Scale bar: 5 μm.

(F) Heatmap showing *RAG1* gene expression patterns across individual human pro-B cells or other types of cells in Fig. S2D. The mRNA of *RAG1* in single-cell level were quantified using the single-cell qPCR and the Ct values were presented as color gradation. Average of Ct was marked for each sample. Cell number n=40; t-test; *, $p<0.05$; ****, $p<0.0001$; Neg., negative control.

(G) Representative microscopy images of hRAG1-maxGFP in *vAbl* (*Rag1*$^{-/-}$; *Rag2*$^{-/-}$; *RAG1*$^{+/-}$) cells at cycling or G1 phase (left). Bar graph shows the percentage of cells with puncta (right); replicates n=3; t-test; *, $p<0.05$. Scale bar: 5 μm.

(H) Representative microscopy images of hRAG1-maxGFP with over-expressed hRAG2-mCherry in *vAbl* pro-B (*Rag1*$^{-/-}$; *Rag2*$^{-/-}$; *RAG1*$^{+/-}$) cells at G1 phase (left). Bar graph (right) shows the percentage of cells with puncta; replicates n=3; t-test; ****, $p<0.0001$. Scale bar: 5 μm.

All error bars represent mean ± SD.

See also in Fig. S1 and Fig. S2.

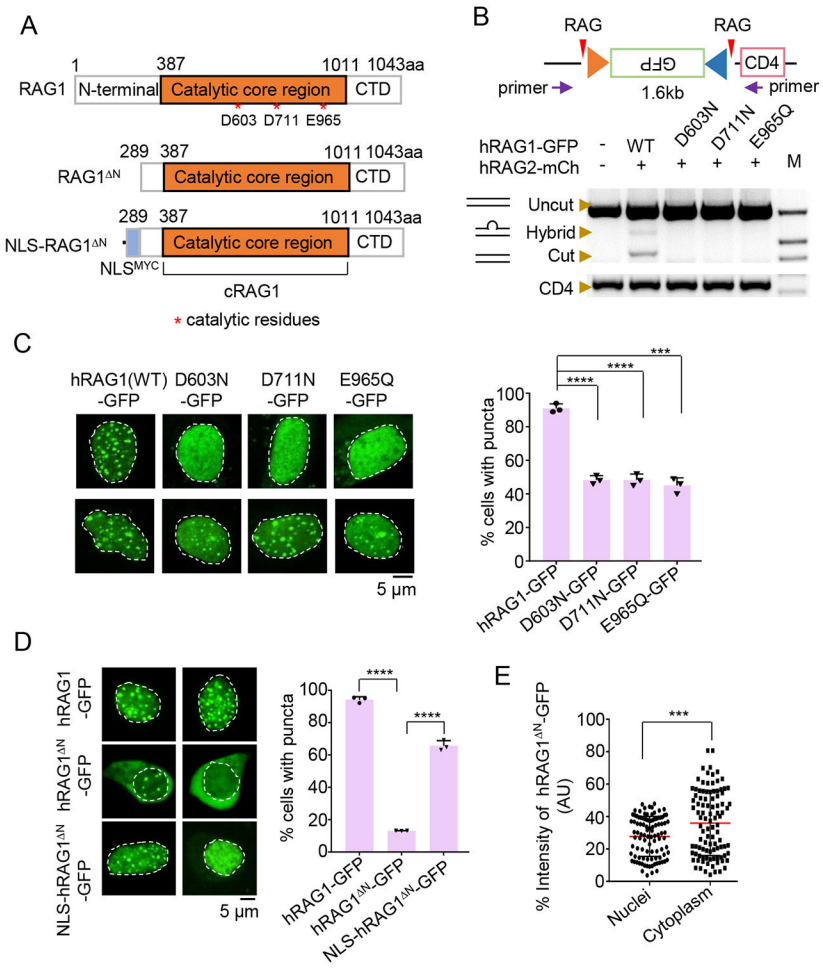

Figure 2. RAG1 aggregates rely on its nuclease activity and nuclear localization.

(A) Schematic diagram of RAG1, RAG1$^{\Delta N}$ and NLS-RAG1$^{\Delta N}$ fusion proteins.

(B) Point mutation at the DDE motif of RAG1 hardly supports the recombination of RSSs. PCR was used to quantify the deletional events between 12 RSS and 23 RSS. The purple arrows indicate the primers used for PCR amplification. The CD4 segment was used to quantify the input 293T-CJ genomic DNA. “M” indicates the DNA marker. mCh, mCherry.

(C) Representative microscopy images of DDE mutants in 293T-CJ cells (left). Bar graph (right) shows the percentage of cells with puncta; replicates n=3; t-test; ****, p<0.0001; ***, p<0.001. Scale bar: 5 μm.

(D) Representative microscopy images of RAG1, RAG1$^{\Delta N}$ and NLS-RAG1$^{\Delta N}$ in 293T-CJ cells (left), and the bar graph (right) shows the percentage of cells with puncta. Replicates n=3; t-test; ****, p<0.0001. Scale bar: 5 μm.

(E) Distribution of GFP intensity for RAG1$^{\Delta N}$ in 293T-CJ cells. T-test; ***, p<0.001.

All error bars represent mean ± SD.

See also in Fig. S3 and Fig. S4.

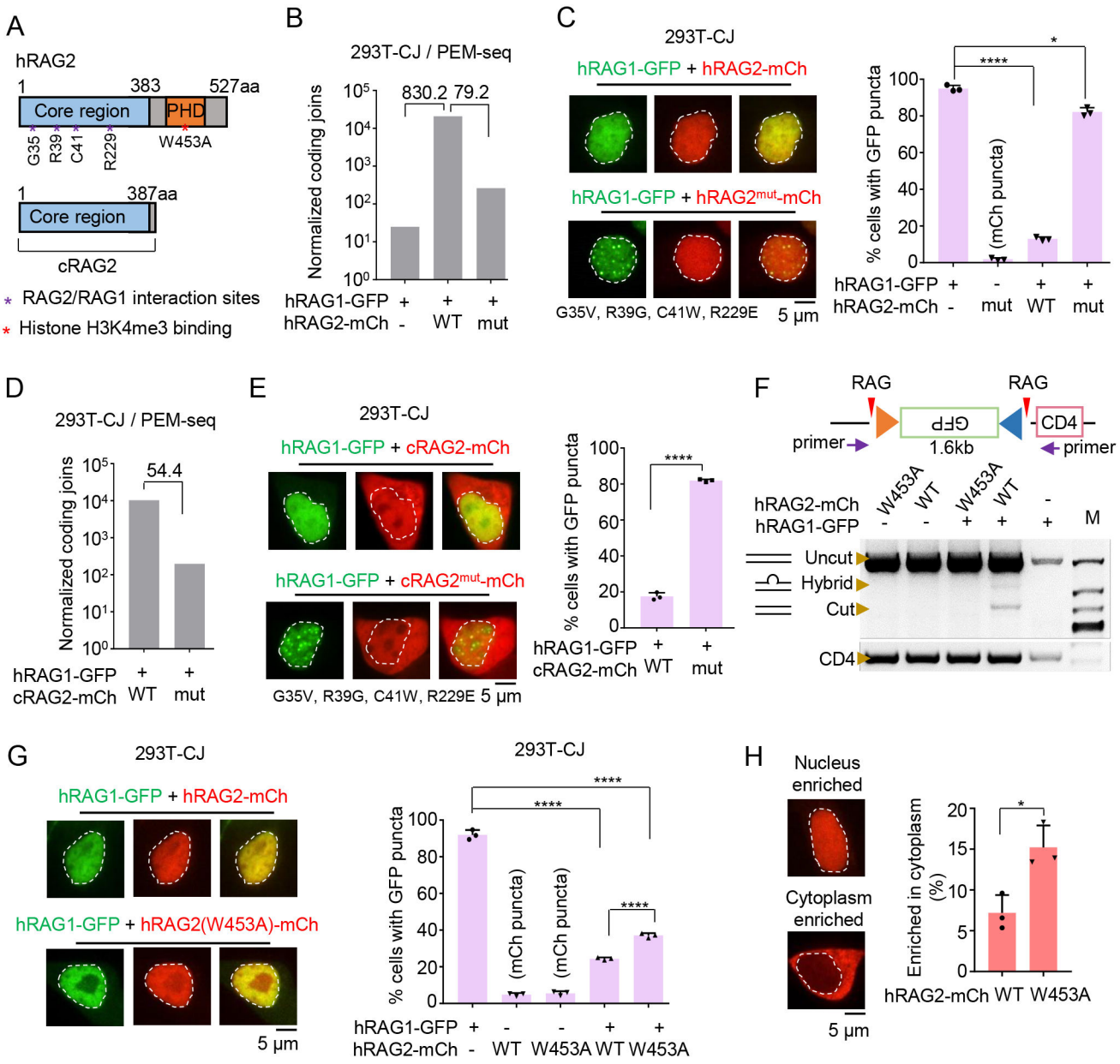

Figure 3. RAG2 abolishes aggregation through direct interaction with RAG1.

(A) Schematic diagram of RAG2 and cRAG2.

(B) Normalized coding joins detected by PEM-seq in 293T-CJ cells with indicated proteins. Fold changes are indicated. Replicate n=1.

(C) Representative microscopy images of RAG1-GFP coexpressed with mutated RAG2-mCh in 293T-CJ cells. Mutation sites at the interface of RAG2 are indicated. Bar graph shows the percentage of cells with puncta; replicates n=3; t-test; *, $p<0.05$; ****, $p<0.0001$. Scale bar: 5 µm.

(D) Normalized coding joins detected by the PEM-seq of hRAG1 with cRAG2 or cRAG2^{mut} in 293T-CJ cells. Fold changes are indicated. Replicate n=1.

(E) Representative microscopy images of RAG1-GFP coexpressed with mutated cRAG2-mCh in 293T-CJ cells. Percentages on the right. Replicates n=3; t-test; ****, $p<0.0001$. Scale bar: 5 μm.
(F) Quantitation of the catalytic activity of RAG1-GFP coexpressed with indicated forms of RAG2-mCh in 293T-CJ cells by PCR.
(G) Representative microscopy images of indicated proteins in 293T-CJ cells. Percentages on the right. Replicates n=3; t-test; ****, $p<0.0001$. Scale bar: 5 μm.
(H) Distribution of wild-type or W453A of RAG2-mCh in 293T-CJ cells. Percentages on the right. Replicates n=3; t-test; *, $p<0.05$. Scale bar: 5 μm.
All error bars represent mean ± SD.
See also in Fig. S5.

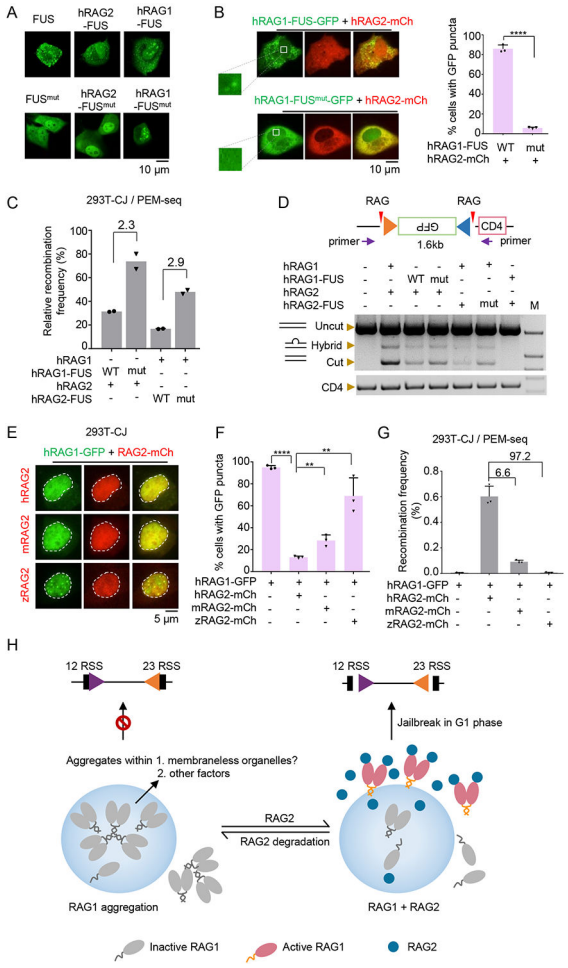

Figure 4. Human RAG2 abolishes the aggregation of human RAG1 to initiate V(D)J recombination.

(A) Representative microscopy images of indicated proteins or corresponding mutants (bottom) in U2OS cells. Scale bar: 10 μm.

(B) Representative microscopy images of indicated proteins coexpressed with RAG2-mCh in U2OS cells. Magnified view of “Boxed regions” shows the RAG1-FUS-GFP puncta in the nucleus. Percentages on the right. Replicates n=3; t-test; ****, p<0.0001. Scale bar: 10 μm.

(C) Coding joins detected by PEM-seq in 293T-CJ cells with indicated proteins. Fold changes are showed. Replicates n=2.

(D) Quantitation of the catalytic activity of indicated coexpression of RAG1 and RAG2 fusion proteins in 293T-CJ cells by PCR.

(E-F) Representative microscopy images of RAG1-GFP coexpressed with indicated RAG2-mCh in 293T-CJ cells (E); Percentages on the panel (F). Replicates n=3; t-test; **, p<0.01;****, p<0.0001. Scale bar: 5 μm.

(G) Recombination frequency mediated by RAG1-GFP with indicated RAG2-mCh detected by PEM-seq. Fold changes are indicated. Replicate n=3; t-test was used for analysis.

(H) “Jailbreak” working model of RAG1. Left: when the RAG2 is degraded, the remaining RAG1 proteins aggregated in the nucleus may with the help of other factors or membraneless organelles. Right: when cells enter the G1 phase, new RAG2 travels to

interact with aggregated RAG1, and releases the RAG1 from aggregation jail to initiate the V(D)J recombination.

All error bars represent mean ± SD.

See also in Fig. S6.

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
Rabbit anti-GFP	Beyotime	AF1483
Rabbit anti-actin	CST	#4967
Goat anti-rabbit IgG-HRP	Abcam	ab6721
APC anti-human CD10	Biolegend	312209
FITC anti-human CD34	Biolegend	343503
BUV395 anti-human CD19	BD bioscience	563549
Chemicals, peptides, and recombinant proteins		
RNase A	Thermo Fisher	EN0531
Proteinase K	Invitrogen	25530015
Streptavidin C1 beads	Invitrogen	65002
Hoechst 33342	Invitrogen	H3570
SF Cell Line 4D-Nucleofecto X Kit	Lonza	V4XC-2024
Calcium chloride	Lab stock	N/A
X-tremeGENE™ HP	Roche	Cat. 06 366 236 001
Deposited data		
Mendeley Data	This paper	10.17632/v6hkpgbnpj.1
GEO database	This paper	PRJNA758251
Experimental models: Cell lines		
HEK293T	Lab stock	N/A
U2OS	ATCC	HTB-96™
NALM-6	From Frederick W. Alt	N/A
vAbl (*Rag2*$^{-/-}$)	From Frederick W. Alt	N/A
Oligonucleotides		
Integration check: CATGACAAGAGTTACTAACAG	This paper	N/A
Integration check: TCTGCTTGGCGCCTTCAGTGC	This paper	N/A
Recombination detect: GACCTTACACAGTCCTGCTG	This paper	N/A
Recombination detect: CTTGCCGAGCATGGTTGTG	This paper	N/A
Recombination detect: GCACTTGCTTCTGGTGCTGC	This paper	N/A
Recombination detect: GCTTCTTGCCCATCTGGAGC	This paper	N/A
PEM-seq: TTGCCAAAAGACGGCAATATGGTGG	This paper	N/A
PEM-seq: CATATAGACAAACGCACACCG	This paper	N/A
RT-PCR: ACAACTTTGGTATCGTGGAAGG	This paper	N/A
RT-PCR: GCCATCACGCCACAGTTTC	This paper	N/A
RT-PCR: ACCCTCAGGCAAGCTTAGGG	This paper	N/A
RT-PCR: TCGACCAGGATGGGCACCAC	This paper	N/A
RT-PCR: GGATCTGTGAACGGTCACGAG	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
RT-PCR: ACGTAAGCCTTAGAGCCGTAC	This paper	N/A
Rag1 knock-out: CCATGTTGGCTAAGCTACCTGGG	This paper	N/A
Rag1 knock-out: TTTGGGCATTGAGGACTCTCTGG	This paper	N/A
RAG1 knock-in: GCTTTCCAGCTCAGGGTAGACGG	This paper	N/A
Rag2 knock-out: GGTGTCTTATAATTCTAAGACAAGC	This paper	N/A
Rag2 knock-out: GTAAGAGTTCTATACTTTATCACTG	This paper	N/A
Rag1 knock-out check: CCTAATAGGTACCAGGGACG	This paper	N/A
Rag1 knock-out check: CTAAACGGCTCAGGCAATCTC	This paper	N/A
RAG1 knock-in check: GAGGGAGCTGATGGATCTTTAC	This paper	N/A
RAG1 knock-in check: CACGAAGCTGTAGTAGCCGC	This paper	N/A
RAG1 knock-in check: GTTTTCATCTTAGGAGTCTG	This paper	N/A
RAG1 knock-in check: CTGTAGCTAACAGGTGCATTC	This paper	N/A
Southern blot probe: GAGGTTCAGGGGGTGGCAGC	This paper	N/A
Southern blot probe: CGAAGGCGATGGGGGTCTTG	This paper	N/A
Single-cell qRT-PCR: TGGAGTGGCACCCCCACACA	This paper	N/A
Single-cell qRT-PCR: GTGCTGACGGGCTTGTCTTGCT	This paper	N/A
Single-cell qRT-PCR: AGGTCGGTGTGAACGGATTTG	This paper	N/A
Single-cell qRT-PCR: TGTAGACCATGTAGTTGAGGTCA	This paper	N/A
Recombinant DNA		
pX332-GFP	This paper	N/A
pX332-hRAG1	This paper	N/A
pX332-hRAG2	This paper	N/A
pX332-hRAG1-GFP	This paper	N/A
pX332-hRAG2-GFP	This paper	N/A
pMX-DEL-CJ	This paper	N/A
pMX-GFP	This paper	N/A
pMX-hRAG1-GFP	This paper	N/A
pMX-hRAG2-mCherry	This paper	N/A
pX332-D603N-GFP	This paper	N/A
pX332-D711N-GFP	This paper	N/A
pX332-E965Q-GFP	This paper	N/A
pX332-hRAG1$^{\Delta N}$-GFP	This paper	N/A
pX332-NLS- hRAG1$^{\Delta N}$-GFP	This paper	N/A
pX332-hRAG2-mCherry	This paper	N/A
pX332-hRAG2^{mut}-mCherry	This paper	N/A
pX332-FUS-GFP	This paper	N/A
pX332-FUS^{mut}-GFP	This paper	N/A
pX332-hRAG1-FUS-GFP	This paper	N/A
pX332-hRAG1-FUS^{mut}-GFP	This paper	N/A

REAGENT or RESOURCE	SOURCE	IDENTIFIER
pX332-hRAG2-FUS-GFP	This paper	N/A
pX332-hRAG2-FUS^{mut}-GFP	This paper	N/A
pX332-mRAG1-GFP	This paper	N/A
pX332-mCherry	This paper	N/A
pX332-FBL-mCherry	This paper	N/A
pX332-mRAG2-mCherry	This paper	N/A
pX332-zRAG2-mCherry	This paper	N/A
pX332-cRAG1-GFP	This paper	N/A
pX332-NLS-cRAG1-GFP	This paper	N/A
pX332-cRAG2-mCherry	This paper	N/A
pX332-cRAG2^{mut}-mCherry	This paper	N/A
pX332-hRAG2(W453A)-mCherry	This paper	N/A
pX332-PHD-mCherry	This paper	N/A
pX332-PHD(W453A)-mCherry	This paper	N/A
pX330	Addgene	42230
pX330-Cas9-CMV-mCherry	This paper	N/A
pMK vector	Addgene	72835
Software and algorithms		
ImageJ1.48	Schneider et al., 2012	https://imagej.nih.gov/
IGV2.3.97	Robinson et al., 2017	https://software.broadinstitute.org/software/igv
Xshell 6	Google website	https://www.netsarang.com/zh/xshell/
FlowJo_V10	BD company website	https://www.flowjo.com/solutions/flowjo/
SuperQ	Yin et al., 2019	Doi:10.1038/s41421-019-0088-8