



Deposited via The University of Leeds.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/145528/>

Version: Accepted Version

---

**Article:**

Kirkham, CM and Boyes, J (2019) Genome instability triggered by the V(D)J recombination by-product. *Molecular and Cellular Oncology*, 6 (4). 1610323. ISSN: 2372-3556

<https://doi.org/10.1080/23723556.2019.1610323>

---

© 2019, Taylor & Francis Group, LLC. This is an Accepted Manuscript of an article published by Taylor & Francis Group in *Molecular & Cellular Oncology* on 07/05/2019, available online: <https://www.tandfonline.com/10.1080/23723556.2019.1610323>

**Reuse**

Items deposited in White Rose Research Online are protected by copyright, with all rights reserved unless indicated otherwise. They may be downloaded and/or printed for private study, or other acts as permitted by national copyright laws. The publisher or other rights holders may allow further reproduction and re-use of the full text version. This is indicated by the licence information on the White Rose Research Online record for the item.

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.

# Genome instability triggered by the V(D)J recombination by-product

Christopher M. Kirkham<sup>1</sup> and Joan Boyes<sup>2</sup>

<sup>1</sup>Charles River Laboratories, Inc., Bristol, BS20 7AW, UK.

<sup>2</sup>School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, LS2 9JT, UK.

Corresponding author: Joan Boyes

[j.m.boyes@leeds.ac.uk](mailto:j.m.boyes@leeds.ac.uk)

**Word count:**

Abstract: 62

Main text: 933

## **ABSTRACT**

A newly identified process by which mistargeted V(D)J recombination could cause genome instability in childhood leukemias has been discovered. In this mechanism, called cut-and-run, the excised DNA by-products of V(D)J recombination are re-bound by the recombinase proteins and erroneously trigger double-strand breaks at multiple locations throughout the genome. Many of these breakpoints co-localize with known chromosome alterations in acute lymphoblastic leukemia (ALL).

## **KEYWORDS:**

V(D)J recombination; Double strand breaks; Chromosome translocations; RAG recombinase; Genome instability; Acute lymphoblastic leukemia

## **MAIN TEXT**

The versatility of the adaptive immune system relies on the process of V(D)J recombination, which rearranges different variable (V), diversity (D) and joining (J) gene segments at random to generate a complete coding sequence for the variable region of antibodies and T cell receptors. This poses an inherent threat to genomic stability, however, as it necessitates the creation of double strand breaks (DSBs) in the genomes of developing B and T cells, which must be accurately repaired. Indeed, mistakes in V(D)J recombination have been shown to underpin 30-40% of lymphoid cancers.<sup>1</sup> This is the unavoidably high cost of an immune system that is capable of detecting vast arrays of pathogens.

The proteins responsible for catalyzing V(D)J recombination are the recombination activating gene (RAG) proteins, RAG1 and RAG2. These proteins bind to the recombination signal sequences (RSSs) that flank each V, D and J coding segment (Figure 1A). Cleavage of two RSSs leads to four broken DNA ends, which are then repaired by the proteins of the non-homologous end-joining (NHEJ) pathway.<sup>2</sup>

The DNA that lies between the two RSSs is usually looped out of the genome and re-joined to form a non-replicative episomal piece of DNA called an excised signal circle (ESC), where the RSSs are joined in a head-to-head fashion (Figure 1A). It was previously thought that this is an inert by-product of the recombination reaction. However, two studies published in 2007 independently demonstrated that the ESC can be reinserted back into the genome by a process named reintegration. Here, an ESC and a chromosomal RSS are bound by the RAG proteins which then catalyze recombination via the same chemistry as standard recombination between two RSSs. The end result is that the ESC is reinserted back into the genome at the site of the partner RSS (Figure 1B).<sup>3,4</sup> Reintegration can occur at legitimate RSSs within the antigen receptor loci, or, more dangerously, at one of the many cryptic RSSs (cRSSs) within the genome. These are sequences with enough homology to *bona fide* RSSs that they can be bound by RAG proteins and utilized in a recombination reaction. An estimated 10 million cRSSs are littered throughout the human genome, some of which lie adjacent to proto-oncogenes such as *LIM domain only 2 (LMO2)* and *T-Cell acute*

*lymphoblastic leukemia protein 2 (TAL2)*. If an ESC were to reintegrate into one of these cRSSs, transcription of the oncogene could be upregulated by virtue of one of the promoters that may be present on the ESC (Figure 1B).

In more recent biochemical experiments that aimed to explore the interactions between RAG proteins and the ESC, we observed, somewhat unexpectedly, that cleavage of an ESC/RSS pair is asymmetric, i.e. cutting at an ESC is around 10-fold less efficient than at an RSS.<sup>5</sup> We were then able to show through bandshift and DNase footprinting experiments, that this asymmetric cutting stems from RAG proteins binding to both RSSs of the ESC simultaneously, thereby blocking efficient cleavage of the ESC in an RSS/ESC complex. This led us to hypothesize that the majority of RSS/ESC pairings would result in cleavage of the RSS whilst the ESC remained intact, instead of reintegrating. We named this reaction “cut-and-run”, as after triggering a “cut” at a genomic RSS, the RAG/ESC complex remains intact and is free to “run” to trigger further genomic DSBs at both legitimate and cRSSs until the RAG proteins are downregulated or the ESC is itself eventually cleaved (Figure 1C). Given the difference in cleavage efficiency between an RSS and ESC, cut-and-run would be around 10-fold more likely to occur than a reintegration reaction, thereby posing a greater threat to genomic stability.

Importantly, the effect is not an *in vitro* anomaly – in a cell line derived from an acute lymphoblastic leukemia (ALL) patient, we also observed low levels of cutting at an ESC relative to a standard RSS.<sup>5</sup> Also, exactly as predicted by the cut-and-run idea, we observed higher rates of  $\gamma$ H2AX foci formation, indicative of DSBs, in cells transfected with the ESC compared to control cells. Furthermore, the breaks caused by the ESC correlate remarkably well with disease. Using linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS),<sup>6</sup> we found that many of the genomic breaks that were formed in the presence of an ESC co-localize with the breaks found in ALL patients that bear the *ETS variant 6/Runx related transcription factor 1 (ETV6/RUNX1)* translocation<sup>7</sup> and furthermore, map to putative cancer driver genes.

The *ETV6/RUNX1* translocation often arises *in utero* but cannot trigger malignancy in the absence of other genetic aberrations. However, the *ETV6/RUNX1* translocation partially stalls B cell progression at a stage where the RAG proteins are still expressed,<sup>8</sup> thus providing an extended window of opportunity for secondary recombination reactions. Involvement of the resulting ESCs in subsequent cut-and-run reactions could generate the secondary mutations that lead to overt leukemia.

Looking beyond ALL, one of the most common forms of translocation caused by aberrant V(D)J recombination is end donation, where a broken RSS formed by RAG cleavage is joined with a DSB formed by an unrelated process, such as ionizing radiation.<sup>1,9</sup> A long-standing question is how the broken RSS becomes available to participate in a chromosome translocation, due to the stringent safeguards in place that shepherd the broken RSSs to the NHEJ pathway.<sup>10</sup> Cut-and-run could be a potent cause of these orphaned breaks, since the non-standard reaction products might

escape efficient repair by the NHEJ proteins. Coupled with the strong correlation between breaks caused by cut-and-run and those found in *ETV6/RUNX1*-positive ALL, it is clear that cut-and-run is likely to be a major source of genomic instability in developing B and T cells.

## FUNDING DETAILS

This work was supported by a Bloodwise research grant (15042) and an Engineering and Physical Sciences Research Council studentship (EP/P505593/1 to C.M.K.).

## DISCLOSURE OF INTEREST

The authors report no conflict of interest.

## REFERENCES

1. Kupperts R. Mechanisms of B Cell Lymphoma Pathogenesis. *Nat. Rev. Cancer.* 2005; 5: 251-262. Doi: 10.1038/nrc1589. PMID: 15803153.
2. Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Ann. Rev. Biochem.* 2002; 71: 101–32. doi: 10.1146/annurev.biochem.71.090501.150203. PMID: 12045092.
3. Curry JD, Schulz D, Guidos CJ, Danska JS, Nutter L, Nussenzweig A, Schlissel MS. Chromosomal reinsertion of broken RSS ends during T cell development. *J. Exp. Med.* 2007; 204: 2293–2303. doi: 10.1084/jem.20070583. PMID: 17785508.
4. Vanura K, Montpellier B, Le T, Spicuglia S, Navarro JM, Cabaud O, Roulland S, Vachez E, Prinz I, Ferrier P, et al. In vivo reinsertion of excised episomes by the V(D)J recombinase: a potential threat to genomic stability. *PLoS Biol.* 2007; 5: e43. doi: 10.1371/journal.pbio.0050043. PMID: 17298184.
5. Kirkham CM, Scott JNF, Wang X, Smith AL, Kupinski AP, Ford AM, Westhead DR, Stockley PG, Tuma R, Boyes J. Cut-and-Run: A distinct mechanism by which V(D)J recombination triggers genome instability. *Mol. Cell.* Forthcoming. <https://doi.org/10.1016/j.molcel.2019.02.025>.
6. Hu J, Zhang Y, Zhao L, Frock RL, Du Z, Meyers RM, Meng FL, Schatz DG, Alt FW. Chromosomal Loop Domains Direct the Recombination of Antigen Receptor Genes. *Cell.* 2015; 163: 947-59. doi: 10.1016/j.cell.2015.10.016. PMID: 26593423.
7. Papaemmanuil E, Rapado I, Li Y, Potter NE, Wedge DC, Tubio J, Alexandrov LB, Van Loo P, Cooke SL, Marshall J, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in *ETV6-RUNX1* acute

lymphoblastic leukemia. *Nat Genet.* 2014; 46: 116–25. doi: 10.1038/ng.2874. PMID: 24413735.

8. Tsuzuki S, Seto M, Greaves M, Enver T. Modeling first-hit functions of the t(12;21) TEL-AML1 translocation in mice. *Proc. Natl. Acad. Sci. USA.* 2004; 101: 8443-8. doi: 10.1073/pnas.0402063101. PMID: 15155899.
9. Roth DB. Restraining the V(D)J recombinase. *Nat Rev Immunol.* 2003; 3: 656-66. doi: 10.1038/nri1152. PMID: 12974480.
10. Helmink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks. *Ann. Rev. Immunol.* 2012; 30: 175-202. doi: 10.1146/annurev-immunol-030409-101320. PMID: 22224778.

### FIGURE CAPTION

**Figure 1. Overview of V(D)J recombination and the mechanisms by which its by-product triggers genome instability.** **A.** V(D)J recombination occurs between 12- and 23-recombination signal sequences (RSSs) that flank all variable (V), diversity (D) and joining (J) gene segments. The recombination activating gene (RAG) proteins first bind to a 12 and 23-RSS and bring the coding segments together in a synaptic complex. The RAG proteins then catalyze formation of double strand breaks precisely at the RSS-coding segment border, resulting in a post-cleavage complex containing two coding ends and two signal ends, which are then repaired by the non-homologous end-joining (NHEJ) proteins to give a chromosomal coding joint and an extra-chromosomal signal joint on an excised signal circle (ESC). **B.** Reintegration involves synapsis between an excised signal circle (ESC) and a chromosomal RSS, which can either be a legitimate RSS or a cryptic RSS (cRSS). A standard recombination reaction occurs, resulting in the reintegration of the ESC back into the genome. **C.** The more frequent outcome of an ESC/RSS synaptic complex is the cleavage of the RSS, but not the ESC. This leaves the ESC/RAG complex free to stimulate cleavage at additional RSSs in a “cut-and-run” reaction. Light blue ovals = RAG1/2 complex; deep blue square = V coding segment; red square = D coding segment; white triangle = 12-RSS; black triangle = 23-RSS; striped triangle = cRSS; P = promoter; ESC = excised signal circle.