



The ESC: The Dangerous By-Product of V(D)J Recombination

Alastair L. Smith[†], James N. F. Scott and Joan Boyes*

School of Molecular and Cellular Biology, Faculty of Biological Sciences, University of Leeds, Leeds, United Kingdom

OPEN ACCESS

Edited by:

Amy L. Kenter,
University of Illinois at Chicago,
United States

Reviewed by:

Anne Corcoran,
Babraham Institute (BBSRC),
United Kingdom
Richard L. Frock,
Stanford University, United States
Patrick Swanson,
Creighton University, United States

*Correspondence:

Joan Boyes
j.m.boyes@leeds.ac.uk

[†]Present Address:

Alastair L. Smith,
MRC Molecular Haematology Unit,
Weatherall Institute of Molecular
Medicine, John Radcliffe Hospital,
University of Oxford, Oxford,
United Kingdom

Specialty section:

This article was submitted to
B Cell Biology,
a section of the journal
Frontiers in Immunology

Received: 08 April 2019

Accepted: 24 June 2019

Published: 04 July 2019

Citation:

Smith AL, Scott JNF and Boyes J
(2019) The ESC: The Dangerous
By-Product of V(D)J Recombination.
Front. Immunol. 10:1572.
doi: 10.3389/fimmu.2019.01572

V(D)J recombination generates antigen receptor diversity by mixing and matching individual variable (V), diversity (D), and joining (J) gene segments. An obligate by-product of many of these reactions is the excised signal circle (ESC), generated by excision of the DNA from between the gene segments. Initially, the ESC was believed to be inert and formed to protect the genome from reactive broken DNA ends but more recent work suggests that the ESC poses a substantial threat to genome stability. Crucially, the recombinase re-binds to the ESC, which can result in it being re-integrated back into the genome, to cause potentially oncogenic insertion events. In addition, very recently, the ESC/recombinase complex was found to catalyze breaks at recombination signal sequences (RSSs) throughout the genome, via a “cut-and-run” mechanism. Remarkably, the ESC/recombinase complex triggers these breaks at key leukemia driver genes, implying that this reaction could be a significant cause of lymphocyte genome instability. Here, we explore these alternate pathways and discuss their relative dangers to lymphocyte genome stability.

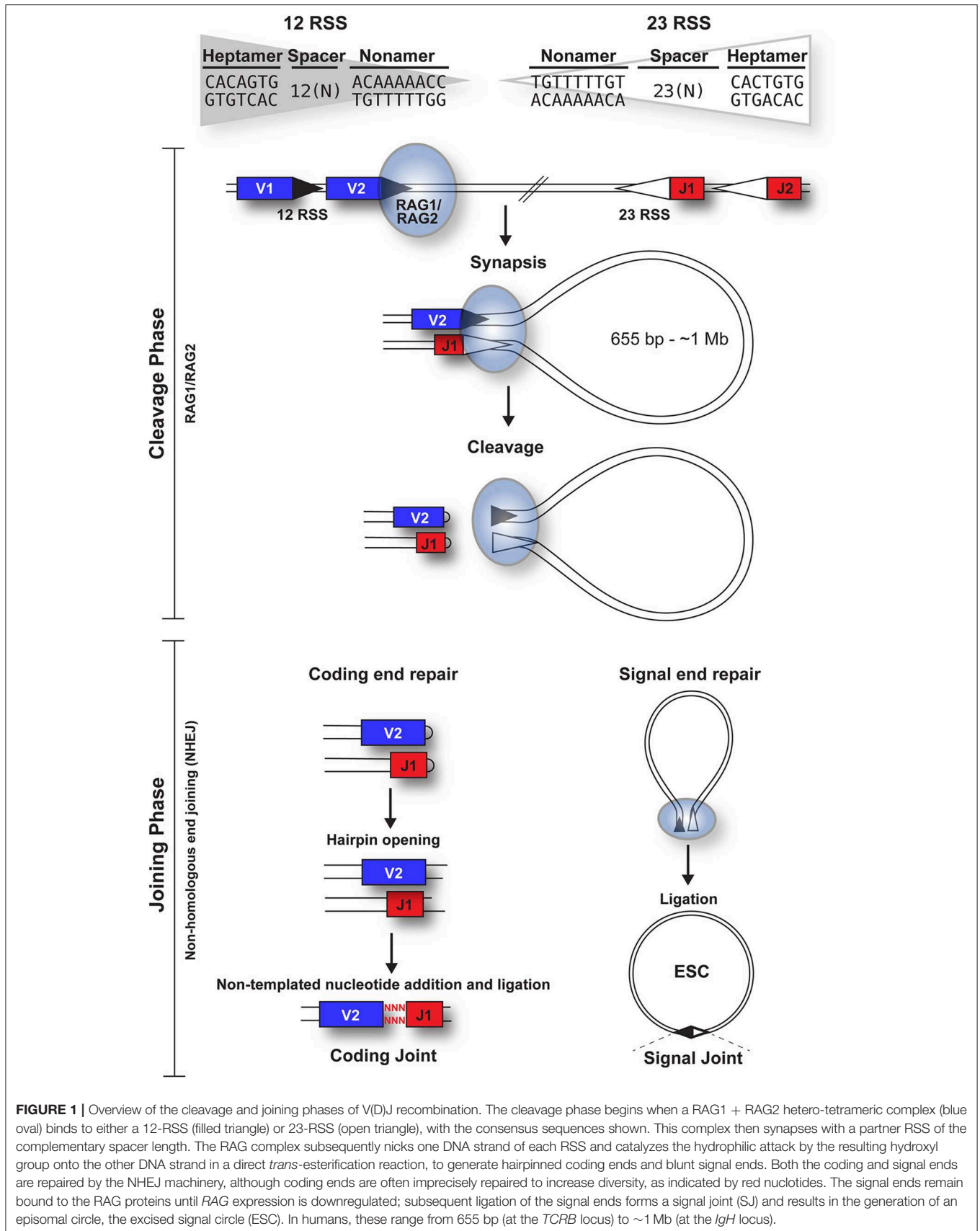
Keywords: V(D)J recombination, RAG proteins, genome instability, leukemia, double strand breaks

INTRODUCTION

V(D)J recombination is essential to generate a diverse adaptive immune system that can respond to vast numbers of potential pathogens. This is achieved, in part, from the unique arrangement of antigen receptor loci where multiple copies of variable (V), diversity (D), and joining (J) gene segments lie upstream of constant exon(s). During V(D)J recombination, one of each V, D (if present), and J gene segments are somatically recombined at random to generate the variable exon of the antigen receptor. The stochastic selection of gene segments, along with their imprecise joining enables production of a highly diverse antigen receptor repertoire (1).

The products of the lymphoid specific recombination activating genes 1 and 2 (*RAG1* and *RAG2*) (2), are essential for V(D)J recombination (3, 4). These proteins recognize recombination signal sequences (RSSs) that flank each V, D, and J gene segment and consist of a conserved heptamer (CACAGTG) and nonamer (ACAAAAACC), separated by non-conserved “spacers” of either 12 ± 1 or 23 ± 1 bp. Importantly, efficient recombination only occurs between RSSs with dissimilar spacers, the “12/23 rule” (5).

V(D)J recombination can be divided into cleavage and joining phases [Figure 1; (6)]. Cleavage is initiated when a hetero-tetrameric complex of RAG1 and RAG2 assembles on either a 12- or 23-RSS (7) and subsequently captures a complementary RSS. Upon formation of a stable synaptic complex, the DNA is unwound at the 5' end of the heptamer (8), followed by introduction of a single-strand DNA nick at the heptamer-coding sequence boundary by the DDE catalytic motif of RAG1 (9–11). The exposed free 3' hydroxyl group then attacks the opposite DNA strand in a direct



trans-esterification reaction (2), yielding a pair of covalently sealed hairpins at the coding ends, and blunt signal ends (1).

Repair of the four broken DNA ends is achieved by the non-homologous end joining (NHEJ) machinery (12). Whilst the coding ends undergo extensive processing, resulting in addition or deletion of bases to increase antigen receptor gene diversity, the RSSs at the signal ends are precisely joined in a head-to-head arrangement, generating a signal joint (SJ). Usually, V(D)J recombination deletes the DNA between the gene segments, to generate an excised signal circle [ESC; (1)] that is covalently sealed at the SJ (**Figure 1**). However, recombination, primarily of the *Igκ* locus, can result in inversion of the intervening DNA and retention of the SJ in the genome. Notably, production of every functional antigen receptor gene generates at least one, and up to 10 ESCs, depending on the level of non-productive rearrangement.

The ESC is a non-replicative episome which is likely lost during cell division. Nonetheless, it persists in chicken T cells for approximately 2 weeks (13) and in primates, this appears to be substantially longer (14). Little to no ESC degradation has been observed and both T cell receptor excision circles (TRECs) and KRECs, generated by recombination to the kappa deleting element during allelic and isotypic exclusion, have proved to be excellent markers of recently generated T- and B-cells, respectively, with 70% of newly produced T cells and 50% of transitional and naïve B cells testing positive (15, 16).

Due to their lack of coding capacity and eventual loss from the cell, it is reasonable to ask why cells bother to generate ESCs. One possibility is that SJ formation is required to maintain chromosome integrity during inversional recombination and this mechanism is simply retained during deletional recombination. Generation of a circle during deletional recombination has the further advantage of sequestering potentially reactive DNA ends. Recent studies strongly suggest, however, that far from being inert, ESCs are crucial constituents of reactions that have potentially devastating consequences for lymphocyte genome stability.

RAG MEDIATED SIGNAL END TRANSPOSITION

The first inkling that the ESC poses a potential threat stemmed from the remarkable similarity between the core domain of RAG1, including the catalytic DDE motif, and the *Transib* family of transposons (17). Furthermore, ProtoRAG, a recently discovered transposon in lancelets, is comprised of both RAG1- and RAG2-like genes, flanked by terminal inverted repeats (TIRs), similar to RSSs. This strongly implies that the RAG recombinase evolved from an ancient transposase that was acquired into the jawed vertebrate genome by horizontal gene transfer (18).

Transposases recognize short sequences surrounding the transposon and introduce double strand breaks (DSBs) between these recognition sequences and the flanking chromosomal DNA, in much the same way as the V(D)J recombinase recognizes RSSs (19). The initial breakage of DNA at the antigen

receptor loci and the hairpin end structures formed is also analogous to the process by which hAT family transposons are excised from their host genome (20). Furthermore, RSSs resemble inverted repeats found at either end of a transposon and RAG proteins remain associated with RSSs after DNA cleavage, which is also common in transposition reactions (21). Consequently, it was not surprising that RAG proteins were found to mediate signal end transposition, at least *in vitro*.

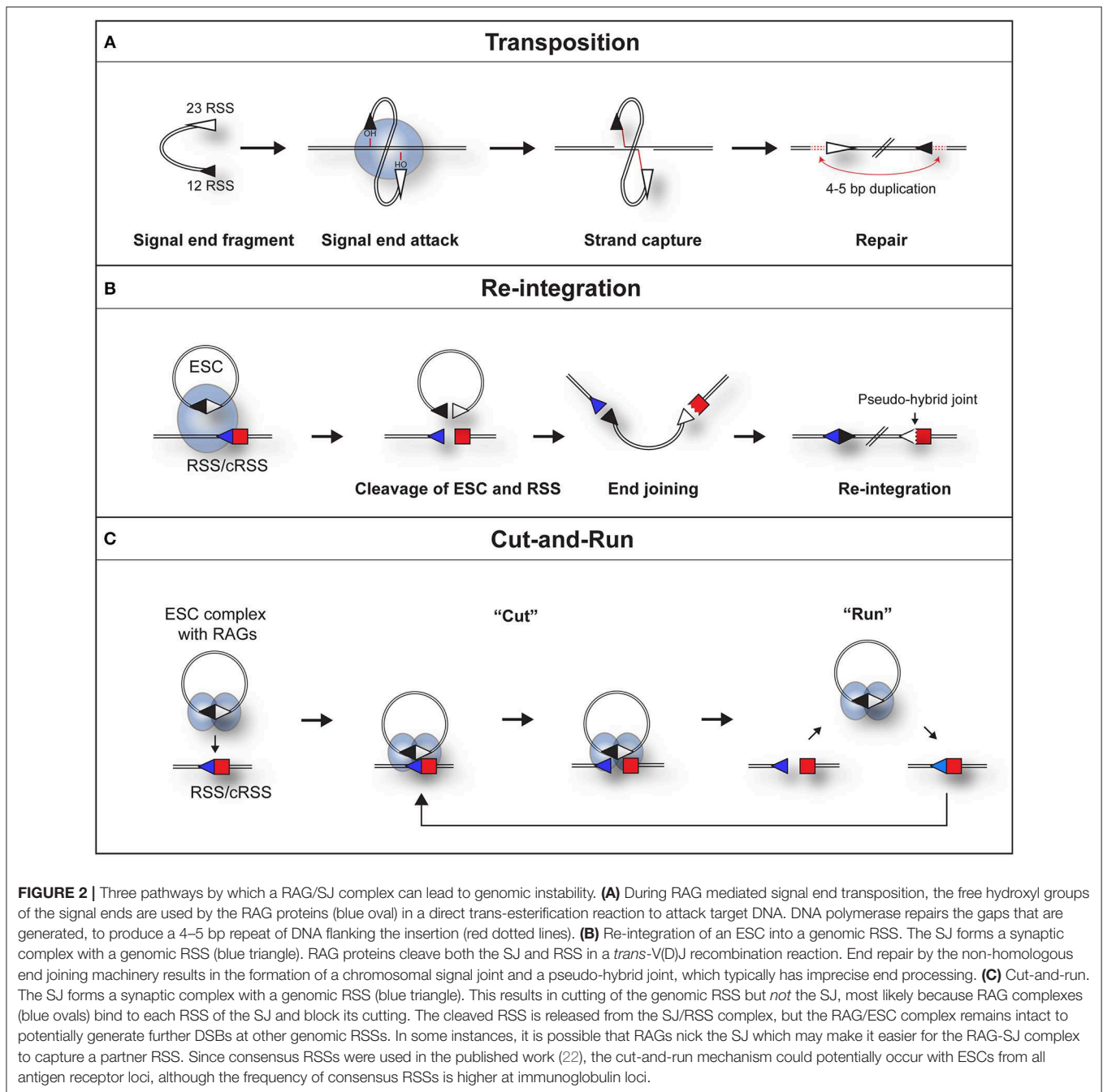
Indeed, both the Gellert and Schatz laboratories demonstrated that in the presence of RAG proteins, DNA substrates flanked by a pair of dissimilar RSSs undergo an intermolecular transposition reaction into DNA targets [**Figure 2A**; (23, 24)]. Furthermore, the SJs were shown to be opened by a “nick-nick” reaction, whereby RAGs sequentially nick at the heptamer-heptamer junction on each strand, to generate the OH⁻ groups required to attack the target DNA during transposition (25).

Despite clear evidence for RAG-mediated transposition events *in vitro*, the impact of these events on genomic stability appears inconsequential. Only two natural integration events have been described (26); in each case, an excised signal end fragment from the *TCRα* locus was integrated into the *HPRT* (hypoxanthine phosphoribosyltransferase) locus in human T cells. Although a handful of transposition events have been observed experimentally, characterized by a 5 bp duplication at the insertion site (27), no single documented case of leukemia or lymphoma can be traced to a transposition event. The most likely explanation is the suppression of transposition by the RAG2 C-terminus (28): Full length RAG2 does not support efficient transposition (29), in contrast to “core” RAG2 (amino acids 1–387), commonly used *in vitro*. This suppression appears to be achieved by the RAG2 C-terminal domain stabilizing the cleaved signal complex (30), to prevent subsequent capture and transposition of signal ends. Consequently, transposition reactions are almost completely suppressed *in vivo*.

RAG MEDIATED ESC RE-INTEGRATION

Although signal ends are not efficiently transposed into the genome, ESCs can be readily cleaved (25) and re-bound by RAG proteins (31) *in vivo*. As a potential consequence, the ESC has been found to undergo relatively efficient re-integration into the genome at RSSs and cryptic RSSs (cRSSs), i.e., RSSs outside the antigen receptor loci that resemble consensus RSSs, via a *trans*-V(D)J recombination reaction. Indeed, in an *ex vivo* assay, Nadel et al. observed that RAGs can integrate genuine SJ sequences into 12-RSS target substrates via synapsis with the 23-RSS of the SJ (32). This generates a new SJ and a pseudo-hybrid joint, where the RSS is joined to a coding end (**Figure 2B**). Consistent with *bona fide* *trans*-V(D)J recombination, the re-integration reaction usually does not violate the 12/23 rule (32) and depends on Artemis to open the hairpin at the coding end. Furthermore, pseudo-hybrid joints, with characteristic processing of the coding end, are observed in mouse thymocytes (32).

A genome wide screen of signal end sequences in precursor T-cells, similarly confirmed the presence of ESC insertions *in vivo*. Approximately half were due to re-integration of an ESC,



with a distinct absence of transposition events (33), suggesting that re-integration is the principal mechanism of ESC insertion. A high rate of ESC re-integration was also observed using episomes carrying *LMO2* and *TAL2* cRSSs (32), implying that re-integration at proto-oncogenes is feasible. Given that ESCs are highly likely to carry promoters, such as those adjacent to V gene segments, such re-integration could upregulate oncogenes, contributing to malignant transformation. Notably, the RAG2 C-terminus also suppresses ESC re-integration by about 7-fold (33). This may relate to the degradation of RAG2 outside of G1 by phosphorylation of threonine 490 (34), which has been

shown to suppress genome instability (35). Nevertheless, based on the experiments described above, ESC re-integration has been estimated to occur in 1 in 1,000,000 to 1 in 10,000 thymocytes (32, 33). Given that millions of lymphocytes are generated each day, this is equivalent to ~5,000 re-integration events per genome per day (32).

Despite the high estimated number of re-integration events, none has been unequivocally linked to malignant transformation. Moreover, only one natural re-integration has been reported, where a *TCR α* -derived ESC was inserted into a cRSS in the *HPRT* locus (36). The large discrepancy between the estimated

re-integration frequency and actual carcinogenic events begs the question of why more disease-causing re-integrations are not observed. One possibility is that re-integration results in B- or T- cell death, either by insertional mutagenesis into critical genes or because re-integration of an ESC, which can be up to 1 Mb in humans, is error-prone and generates DSBs that trigger apoptosis via the p53 surveillance pathway. A second possibility is that re-insertion occurs, but is not detected (32). Since recombination is stochastic, it is difficult to predict which ESCs will be generated, and moreover, the ESC could be potentially inserted at any of 10 million cRSSs (37), which makes screening for reintegration difficult. Indeed, over a decade after re-integration was first described, its contribution to lymphoid malignancies remains unknown.

Although not strictly involving ESCs, the Robbiani laboratory recently described a related mechanism of RAG-mediated genome instability (38). Using bespoke translocation capture (TC-seq) and insertion capture sequencing (IC-seq), they found that RAGs can release DNA fragments (with signal ends, coding ends or hybrid ends) from antigen receptor loci, independent of normal recombination, and these fragments reintegrate into a RAG-independent DSB elsewhere in the genome. By developing a novel pipeline to analyse whole genome sequencing data, they found 5 out of 34 acute lymphoblastic leukemia (ALL) and follicular lymphoma patients displayed genomic insertions from the antigen receptor loci. However, the authors believe this is an underestimate due to limitations in sequencing depth and read lengths. Notably, with such improvements in whole-genome sequencing, this analysis pipeline could be used additionally to detect RAG-mediated ESC re-integration in lymphoid malignancies.

CUT-AND-RUN

Whilst re-integration events clearly occur, it was shown recently that synaptic complex formation between an ESC and RSS *in vitro* results in efficient RSS cleavage, whereas the ESC itself is barely cut (22). This asymmetric cleavage of an ESC/RSS complex does not appear to be an *in vitro* artifact as assays using a cell line derived from an ALL patient, showed that episomes carrying a 12- or 23-RSS are also cleaved readily whereas substrates bearing a SJ (ESC) exhibit negligible cutting (22).

This phenomenon can be best explained by RAGs binding to both RSSs of the ESC (22), thereby occluding the heptamer-heptamer boundary of the SJ, preventing its cleavage. What makes this reaction particularly dangerous is following cleavage, the broken RSS is released, generating a DSB (22) whereas the ESC/RAG complex remains intact. Although not yet observed directly, hypothetically, the RAG/ESC complex could catalyze further breaks at new cRSSs in a reaction termed “cut-and-run” (Figure 2C). This reaction could potentially continue until the RAG proteins are downregulated and/or the ESC is eventually cleaved—with disastrous consequences for the lymphocyte genome. Indeed, analysis of chromosome breakpoints caused by the RAG/ESC complex, using linear amplification-mediated, high-throughput genome-wide sequencing [LAM-HTGTs; (39)]

showed a significant overlap between cut-and-run-mediated breakpoints and those observed in *ETV6/RUNX1*-positive ALL patients (22, 40). Moreover, breakpoints in eight out of the eleven most commonly mutated genes in B-cell ALL were observed in the presence of the ESC (22). This strongly implies that cut-and-run has a role in the development of *ETV6/RUNX1*-positive ALL, and potentially other lymphoid cancers.

THE DANGERS OF CUT-AND-RUN COMPARED TO RE-INTEGRATION

The discovery of two distinct mechanisms by which the ESC triggers genome instability raises the fundamental question of which poses the greater danger. This will be influenced by both the reaction frequency and damage caused by each reaction.

Reaction Frequency

The overall reaction frequency will depend on its actual frequency as well as the availability of reaction components. Both re-integration and cut-and-run require a complex between RAGs and the ESC and thus will be restricted to cells where both are present, such as pro- and pre-B cells as well as immature B cells, where RAGs are upregulated for receptor editing and receptor revision (41, 42). Moreover, whilst re-integration can theoretically occur with either a covalently closed ESC or open SJ, cut-and-run requires the ESC to be covalently closed. Notably, SJs remain unligated following recombination until RAGs are down-regulated, either as a result of cell replication (34, 43) or following productive antigen receptor recombination (44). This will therefore further restrict cut-and-run; nonetheless, it could occur in normal lymphocytes, for example, by using ESCs in pre-B cells that were generated by IgH recombination in pro-B cells. Substrates for both reactions are likely to be substantially increased, however, in cancer cells. Indeed, a number of pre-leukaemic (45) and leukaemic cells (46) continually express RAGs, triggering ongoing recombination and increased production of ESCs (47). Because these cells continually divide, the generation of covalently closed ESCs is likely to be particularly high, thereby enhancing the risks of further genome instability.

Nonetheless, the window in which cut-and-run or re-integration can occur will be restricted by the short half-life of RAG1 [~15–30 min; (48, 49)] and by the cell-cycle dependent degradation of RAG2 outside of G1 (34, 35). Not only this, but it appears that the lymphocyte genome has tried to protect itself against off-target RAG cleavage: Only ~3500 of the millions of cRSSs are occupied by RAG1 (50), substantially limiting where breaks could occur. In addition, genomic regions outside of the antigen receptor loci that are enriched for RAG binding were found to be depleted of RSSs, a mechanism suggested to protect active transcriptional start sites from off-target RAG cleavage (50).

Yet further restrictions on RSS cleavage are imposed by local chromatin modifications. Indeed, RSSs need to be accessible to RAGs and to have proximal nucleosomes marked by acetylation of lysine 27 of H3 (H3K27Ac) for RAG1 binding (51) and

trimethylation of lysine 4 of H3 (H3K4me3) for RAG2 (52, 53). Whilst these factors undoubtedly provide some protection, cleavage at cRSSs clearly still occurs. Indeed, breaks at cRSSs in cancer driver genes were found to be a predominant cause for cancer progression in *ETV6/RUNX1* ALL, and our LAM-HTGTS experiments suggest the ESC could play a role in causing some of these breaks (22, 40).

Interaction of RAG2 with H3K4me3 via its PHD finger has a further regulatory role, namely to overcome the auto-inhibition of RAG1 cleavage, imposed by RAG2 (54–56). Since the chromatin modifications present on ESCs are currently unknown, it is difficult to determine if the ESC/RAG complex is affected by altered H3K4me3 levels. Nonetheless, formation of the RAG/ESC complex generates half of the synaptic complex required for either cut-and-run or re-integration, meaning that these reactions require just one accessible cRSS in the genome (22). This, together with the intrinsic mobility of the RAG/ESC complex, is expected to increase the chances of the RAG/ESC complex “finding” genomic (c)RSSs with the correct chromatin modifications. Furthermore, RAG proteins have been shown to nick the ESC *in vitro* and *in vivo* (25), which could also increase the chances of the RAG/ESC complex capturing a partner RSS, although the level of ESC nicking that we observed was noticeably lower than at an RSS (22).

Considering the reaction rate, cut-and-run is far more likely to occur. In fact, *in vitro* reactions suggest there is a 10-fold greater likelihood of cutting just at the RSS (for cut-and-run) compared to cutting at both the RSS and ESC (22), to enable re-integration. Furthermore, the fact that the ESC/recombinase complex remains intact following cutting at one site (22) means that cut-and-run could potentially trigger a series of DSBs, resulting in significant damage. By contrast, re-integration if it occurs, is very likely to occur only once as the event typically mutates one of the RSSs (32), preventing further reactivity. Whilst it is difficult to know the actual frequency of cut-and-run, using γ H2AX as a marker for DSB, physiological levels of RAGs (57) and close to physiological levels of ESCs (3–6 ESCs per cell), an average of two additional DSBs were observed in ~20% of cells in the presence of the ESC, compared to control cells (22). This is considerably higher than the estimates for re-integration of 1 in 100,000 cells (33), suggesting that cut-and-run could be both frequent and dangerous.

Damage Caused by Cut-and-Run or Re-integration

Both re-integration of the ESC and the cut-and-run reaction pose significant threats to genomic stability but theoretically, a single ESC re-integration event has the potential to cause a greater degree of genomic disruption due to the insertion of up to 1 Mb of DNA. Furthermore, due to the presence of strong promoters within the antigen receptor loci, re-integration of an ESC could upregulate genes next to the insertion site, including proto-oncogenes such as *LMO2*. Not only this, but Nadel et al. observed two cases of translocations

which appeared to be associated with insertion events (32), suggesting secondary re-arrangement events may occur after integration.

A single cut-and-run event has, in theory, a lower probability of genomic disruption compared to ESC re-integration. However, the damage caused will depend entirely on the processing of the break. Asymmetric ESC/RSS cleavage generates one hairpinned and one blunt end (22); crucially, these broken ends are released and are not chaperoned by RAGs to the NHEJ machinery (58). Although it is possible that the NHEJ proteins are recruited independently, resulting in hairpin opening and repaired DNA that closely resembles the sequence prior to cleavage [akin to open and shut joints; (59)], it is also possible that the released ends are processed by the error-prone alternative NHEJ pathway or that they may be used as substrates for chromosome translocations. Indeed, DSBs generated by cut-and-run were found to readily undergo translocation, as determined by the LAM-HTGTS assay (22, 39). Moreover, one of the most common V(D)J recombination errors is the end donation reaction, where a broken RSS becomes joined with an independently broken DNA end (19). It is possible that some of the breaks generated by cut-and-run are substrates for such reactions.

Relationship to Development of Cancer

The relative dangers of these reactions can be further estimated by considering their links to cancer. To date, no documented examples exist of cancers triggered by re-integration. This may be because re-integration is hard to detect and it was argued that since approximately one third of T-ALL cases have oncogene activation without abnormal karyotypes (32), oncogenesis triggered by pathways such as re-integration may be involved. However, many whole genome sequences from patients have since become available, which show relatively small chromosome changes that would not be detected karyotypically but nonetheless lead to oncogene activation. Crucially, many of these small changes in *ETV6/RUNX1*-positive ALL occur at RSSs (40) and there is a strong correlation between the breaks in patients and those caused by cut-and-run (22, 40). Furthermore, about half of the chromosome abnormalities mapped via whole genome sequencing show breaks at just one RSS (40). It is difficult to account for such breaks during normal V(D)J recombination given the stringency for coupled RSS cleavage and end joining. Cut-and-run could account for such breaks.

CONCLUDING STATEMENT

It is clear that the ESC is far from inert but instead poses a significant threat to genome stability. Although re-integration has the potential for significantly greater damage, its low frequency compared to cut-and-run and the low probability of integrating at a proto-oncogene significantly reduces its overall danger. Cut-and-run, on the other hand, appears to be frequent and could be a source of breaks for the major chromosome alterations associated with errors in V(D)J recombination. However, to fully understand the impact of cut-and-run, further experiments are required to determine (a) the outcomes of the

released broken ends, (b) if the ESC “runs” to trigger subsequent genomic breaks, and (c) if cut-and-run contributes to other B and T cell cancers.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

REFERENCES

- Gellert M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu Rev Biochem.* (2002) 71:101–32. doi: 10.1146/annurev.biochem.71.090501.150203
- McBlane JF, van Gent DC, Ramsden DA, Romeo C, Cuomo CA, Gellert M, et al. Cleavage at a V(D)J recombination signal requires only RAG1 and RAG2 proteins and occurs in two steps. *Cell.* (1995) 83:387–95. doi: 10.1016/0092-8674(95)90116-7
- Mombaerts P, Iacomini J, Johnson RS, Herrup K, Tonegawa S, Papaioannou VE. RAG-1-deficient mice have no mature B and T lymphocytes. *Cell.* (1992) 68:869–77. doi: 10.1016/0092-8674(92)90030-G
- Shinkai Y, Rathbun G, Lam KP, Oltz EM, Stewart V, Mendelsohn M, et al. RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement. *Cell.* (1992) 68:855–67. doi: 10.1016/0092-8674(92)90029-C
- Tonegawa S. Somatic generation of antibody diversity. *Nature.* (1983) 302:575–81. doi: 10.1038/302575a0
- Schatz DG, Swanson PC. V(D)J recombination: mechanisms of initiation. *Annu Rev Genet.* (2011) 45:167–202. doi: 10.1146/annurev-genet-110410-132552
- Swanson PC. Fine structure and activity of discrete RAG-HMG complexes on V(D)J recombination signals. *Mol Cell Biol.* (2002) 22:1340–51. doi: 10.1128/MCB.22.5.1340-1351.2002
- Ru H, Mi W, Zhang P, Alt FW, Schatz DG, Liao M, et al. DNA melting initiates the RAG catalytic pathway. *Nat Struct Mol Biol.* (2018) 25:732–42. doi: 10.1038/s41594-018-0098-5
- Kim DR, Dai Y, Mundy CL, Yang W, Oettinger MA. Mutations of acidic residues in RAG1 define the active site of the V(D)J recombinase. *Genes Dev.* (1999) 13:3070–80. doi: 10.1101/gad.13.23.3070
- Landree MA, Wibbenmeyer JA, Roth DB. Mutational analysis of RAG1 and RAG2 identifies three catalytic amino acids in RAG1 critical for both cleavage steps of V(D)J recombination. *Genes Dev.* (1999) 13:3059–69. doi: 10.1101/gad.13.23.3059
- Fugmann SD, Viley IJ, Ptaszek LM, Schatz DG. Identification of two catalytic residues in RAG1 that define a single active site within the RAG1/RAG2 protein complex. *Mol Cell.* (2000) 5:97–107. doi: 10.1016/S1097-2765(00)80406-2
- Helmkink BA, Sleckman BP. The response to and repair of RAG-mediated DNA double-strand breaks. *Annu Rev Immunol.* (2012) 30:175–202. doi: 10.1146/annurev-immunol-030409-101320
- Kong F-K, Chen CH, Six A, Hockett RD, Cooper MD. T cell receptor gene deletion circles identify recent thymic emigrants in the peripheral T cell pool. *Proc Natl Acad Sci USA.* (1999) 96:1536–40. doi: 10.1073/pnas.96.4.1536
- Sodora DL, Douek DC, Silvestri G, Montgomery L, Rosenzweig M, Igarashi T, et al. Quantification of thymic function by measuring T cell receptor excision circles within peripheral blood and lymphoid tissues in monkeys. *Eur J Immunol.* (2000) 30:1145–53. doi: 10.1002/(SICI)1521-4141(200004)30:4<1145::AID-IMMU1145>3.0.CO;2-7
- Mensen A, Ochs C, Stroux A, Wittenbecher F, Szyska M, Imberti L, et al. Utilization of TREC and KREC quantification for the monitoring of early T- and B-cell neogenesis in adult patients after allogeneic hematopoietic stem cell transplantation. *J Transl Med.* (2013) 11:188. doi: 10.1186/1479-5876-11-188

FUNDING

We gratefully acknowledge support from Bloodwise (research grant: 15042) and the National Center for the Replacement, Refinement, and Reduction of Animals in Research (NC3Rs; studentship: NC/K001639/1). Bloodwise contribute to a fund for Open Access publication fees that is administered at the University of Leeds by the library.

- Serana F, Airò P, Chiarini M, Zanotti C, Scarsi M, Frassi M, et al. Thymic and bone marrow output in patients with common variable immunodeficiency. *J Clin Immunol.* (2011) 31:540–549. doi: 10.1007/s10875-011-9526-6
- Kapitonov VV, Jurka J. RAG1 core and V(D)J recombination signal sequences were derived from transib transposons. *PLoS Biol.* (2005) 3:e181. doi: 10.1371/journal.pbio.0030181
- Zhang Y, Cheng TC, Huang G, Lu Q, Surleac MD, Mandell JD, et al. Transposon molecular domestication and the evolution of the RAG recombinase. *Nature.* (2019) 569:79–84. doi: 10.1038/s41586-019-1093-7
- Roth DB. Restraining the V(D)J recombinase. *Nat Rev Immunol.* (2003) 3:656–66. doi: 10.1038/nri1152
- Zhou L, Mitra R, Atkinson PW, Hickman AB, Dyda F, Craig NL. Transposition of hAT elements links transposable elements and V(D)J recombination. *Nature.* (2004) 432:995–1001. doi: 10.1038/nature03157
- Roth DB, Craig NL. VDJ recombination : a transposase goes to work. *Cell.* (1998) 94:411–4. doi: 10.1016/S0092-8674(00)81580-9
- Kirkham CM, Scott JNF, Wang X, Smith AL, Kupinski AP, Ford AM, et al. Cut-and-run: a distinct mechanism by which V(D)J recombination causes genome instability. *Mol Cell.* (2019) 74:584–97.e9. doi: 10.1016/j.molcel.2019.02.025
- Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell.* (1998) 94:463–70. doi: 10.1016/S0092-8674(00)81587-1
- Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature.* (1998) 394:744–51. doi: 10.1038/29457
- Neiditch MB, Lee GS, Huye LE, Brandt VL, Roth DB. The V(D)J recombinase efficiently cleaves and transposes signal joints. *Mol Cell.* (2002) 9:871–8. doi: 10.1016/S1097-2765(02)00494-X
- Messier TL, O'Neill JP, Hou S-M, Nicklas JA, Finette BA. *In vivo* transposition mediated by V(D)J recombinase in human T lymphocytes. *EMBO J.* (2003) 22:1381–8. doi: 10.1093/emboj/cdg137
- Reddy YV, Perkins EJ, Ramsden DA. Genomic instability due to V(D)J recombination-associated transposition. *Genes Dev.* (2006) 20:1575–82. doi: 10.1101/gad.1432706
- Elkin SK, Matthews AG, Oettinger MA. The C-terminal portion of RAG2 protects against transposition *in vitro*. *EMBO J.* (2003) 22:1931–8. doi: 10.1093/emboj/cdg184
- Swanson PC, Volkmer D, Wang L. Full-length RAG-2, and not full-length RAG-1, specifically suppresses RAG-mediated transposition but not hybrid joint formation or disintegration. *J Biol Chem.* (2004) 279:4034–44. doi: 10.1074/jbc.M311100200
- Matthews AG, Oettinger MA. Regulation of RAG transposition. *Adv Exp Med Biol.* (2009) 650:16–31. doi: 10.1007/978-1-4419-0296-2_2
- Canela A, Sridharan S, Sciascia N, Tubbs A, Meltzer P, Sleckman BP, et al. DNA breaks and end resection measured genome-wide by end sequencing. *Mol Cell.* (2016) 63:898–911. doi: 10.1016/j.molcel.2016.06.034
- Vanura K, Montpellier B, Le T, Spicuglia S, Navarro J-M, Cabaud O, 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

33. Curry JD, Schulz D, Guidos CJ, Danska JS, Nutter L, Nussenzweig A, et al. Chromosomal reinsertion of broken RSS ends during T cell development. *J Exp Med.* (2007) 204:2293–303. doi: 10.1084/jem.20070583
34. Lin WC, Desiderio S. Regulation of V(D)J recombination activator protein RAG-2 by phosphorylation. *Science.* (1993) 260:953–9. doi: 10.1126/science.8493533
35. Zhang L, Reynolds TL, Shan X, Desiderio S. Coupling of V(D)J recombination to the cell cycle suppresses genomic instability and lymphoid tumorigenesis. *Immunity.* (2011) 34:163–74. doi: 10.1016/j.immuni.2011.02.003
36. Messier TL, O'Neill JP, Finette BA. V(D)J recombinase mediated inter-chromosomal HPRT alterations at cryptic recombination signal sequences in peripheral human T cells. *Hum Mutat.* (2006) 27:829. doi: 10.1002/humu.9440
37. Lewis SM, Agard E, Suh S, Czyzyk L. Cryptic signals and the fidelity of V(D)J joining. *Mol Cell Biol.* (1997) 17:3125–36. doi: 10.1128/MCB.17.6.3125
38. Rommel PC, Oliveira TY, Nussenzweig MC, Robbiani DF. RAG1/2 induces genomic insertions by mobilizing DNA into RAG1/2-independent breaks. *J Exp Med.* (2017) 214:815–31. doi: 10.1084/jem.20161638
39. Hu J, Meyers RM, Dong J, Panchakshari RA, Alt FW, Frock RL. Detecting DNA double-stranded breaks in mammalian genomes by linear amplification-mediated high-throughput genome-wide translocation sequencing. *Nat Protoc.* (2016) 11:853–71. doi: 10.1038/nprot.2016.043
40. Papaemmanuil E, Rapado I, Li Y, Potter NE, Wedge DC, Tubio J, et al. RAG-mediated recombination is the predominant driver of oncogenic rearrangement in ETV6-RUNX1 acute lymphoblastic leukemia. *Nat Genet.* (2014) 46:116–25. doi: 10.1038/ng.2874
41. Jankovic M, Casellas R, Yannoutsos N, Wardemann H, Nussenzweig MC. RAGs and regulation of autoantibodies. *Annu Rev Immunol.* (2004) 22:485–501. doi: 10.1146/annurev.immunol.22.012703.104707
42. Han S, Dillon SR, Zheng B, Shimoda M, Schlissel MS, Kelsoe G. V(D)J recombinase activity in a subset of germinal center B lymphocytes. *Science.* (1997) 278:301–5. doi: 10.1126/science.278.5336.301
43. Jiang H, Chang F-C, Ross AE, Lee J, Nakayama K, Nakayama K, et al. Ubiquitylation of RAG-2 by Skp2-SCF links destruction of the V(D)J recombinase to the cell cycle. *Mol Cell.* (2005) 18:699–709. doi: 10.1016/j.molcel.2005.05.011
44. Ramsden DA, Gellert M. Formation and resolution of double-strand break intermediates in V(D)J rearrangement. *Genes Dev.* (1995) 9:2409–20. doi: 10.1101/gad.9.19.2409
45. Dong Y, Liu F, Wu C, Li S, Zhao X, Zhang P, et al. Illegitimate RAG-mediated recombination events are involved in *IKZF1* Δ 3-6 deletion in *BCR-ABL1* lymphoblastic leukaemia. *Clin Exp Immunol.* (2016) 185:320–31. doi: 10.1111/cei.12812
46. Heinänenmi M, Vuorenmaa T, Teppo S, Kaikkonen MU, Bouvy-Liivrand M, Mehtonen J, et al. Transcription-coupled genetic instability marks acute lymphoblastic leukemia structural variation hotspots. *Elife.* (2016) 5:e13087. doi: 10.7554/eLife.13087
47. Bashford-Rogers RJM, Nicolaou KA, Bartram J, Goulden NJ, Loizou L, Koumas L, et al. Eye on the B-ALL: B-cell receptor repertoires reveal persistence of numerous B-lymphoblastic leukemia subclones from diagnosis to relapse. *Leukemia.* (2016) 30:2312–21. doi: 10.1038/leu.2016.142
48. Sadofsky MJ, Hesse JE, McBlane JE, Gellert M. Expression and V(D)J recombination activity of mutated RAG-1 proteins. *Nucleic Acids Res.* (1993) 21:5644–50. doi: 10.1093/nar/21.24.5644
49. Schabla NM, Perry GA, Palmer VL, Swanson PC. VprBP (DCAF1) regulates RAG1 expression independently of dicer by mediating RAG1 degradation. *J Immunol.* (2018) 201:930–9. doi: 10.4049/jimmunol.1800054
50. Teng G, Maman Y, Resch W, Kim M, Yamane A, Qian J, et al. RAG represents a widespread threat to the lymphocyte genome. *Cell.* (2015) 162:751–65. doi: 10.1016/j.cell.2015.07.009
51. Maman Y, Teng G, Seth R, Kleinstein SH, Schatz DG. RAG1 targeting in the genome is dominated by chromatin interactions mediated by the non-core regions of RAG1 and RAG2. *Nucleic Acids Res.* (2016) 44:9624–37. doi: 10.1093/nar/gkw633
52. Matthews AG, Kuo AJ, Ramón-Maiques S, Han S, Champagne KS, Ivanov D, et al. RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature.* (2007) 450:1106–10. doi: 10.1038/nature06431
53. Liu Y, Subrahmanyam R, Chakraborty T, Sen R, Desiderio S. A plant homeodomain in RAG-2 that binds Hypermethylated lysine 4 of histone H3 is necessary for efficient antigen-receptor-gene rearrangement. *Immunity.* (2007) 27:561–71. doi: 10.1016/j.immuni.2007.09.005
54. Grundy GJ, Yang W, Gellert M. Autoinhibition of DNA cleavage mediated by RAG1 and RAG2 is overcome by an epigenetic signal in V(D)J recombination. *Proc Natl Acad Sci USA.* (2010) 107:22487–92. doi: 10.1073/pnas.1014958107
55. Lu C, Ward A, Bettridge J, Liu Y, Desiderio S. An autoregulatory mechanism imposes allosteric control on the V(D)J recombinase by histone H3 methylation. *Cell Rep.* (2015) 10:29–38. doi: 10.1016/j.celrep.2014.12.001
56. Bettridge J, Na CH, Pandey A, Desiderio S. H3K4me3 induces allosteric conformational changes in the DNA-binding and catalytic regions of the V(D)J recombinase. *Proc Natl Acad Sci USA.* (2017) 114:1904–9. doi: 10.1073/pnas.1615727114
57. Bories JC, Cayuela JM, Loiseau P, Sigaux F. Expression of human recombination activating genes (RAG1 and RAG2) in neoplastic lymphoid cells: correlation with cell differentiation and antigen receptor expression. *Blood.* (1991) 78:2053–61.
58. Cui X, Meek K. Linking double-stranded DNA breaks to the recombination activating gene complex directs repair to the nonhomologous end-joining pathway. *Proc Natl Acad Sci USA.* (2007) 104:17046–51. doi: 10.1073/pnas.0610928104
59. Lewis SM, Hesse JE, Mizuuchi K, Gellert M. Novel strand exchanges in V(D)J recombination. *Cell.* (1998) 55:1099–107. doi: 10.1016/0092-8674(88)90254-1

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Smith, Scott and Boyes. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.