Recombination centres and the orchestration of V(D)J recombination

David G. Schatz and Yanhong Ji‡*

Abstract | The initiation of V(D)J recombination by the recombination activating gene 1 (RAG1) and RAG2 proteins is carefully orchestrated to ensure that antigen receptor gene assembly occurs in the appropriate cell lineage and in the proper developmental order. Here we review recent advances in our understanding of how DNA binding and cleavage by the RAG proteins are regulated by the chromatin structure and architecture of antigen receptor genes. These advances suggest novel mechanisms for both the targeting and the mistargeting of V(D)J recombination, and have implications for how these events contribute to genome instability and lymphoid malignancy.

V(D)J recombination
Somatic rearrangement of variable (V), diversity (D) and joining (J) regions of the genes that encode antigen receptors, leading to repertoire diversity of immunoglobulins and T cell receptors.

Chromosomal translocation
An aberration of chromosome structure in which a portion of one chromosome is broken off and becomes attached to another.

**Department of Immunobiology and Howard Hughes Medical Institute, Yale University School of Medicine, 300 Cedar Street, Box 208011, New Haven, Connecticut 06520-8011, USA.*
‡Department of Immunology & Microbiology, School of Medicine, Xi'an JiaoTong University, Xian Shaanxi, China 710049.
Correspondence to D.G.S.
e-mail:
david.schatz@yale.edu
doi:10.1038/nri2941
Published online 11 March 2011

Vertebrate adaptive immune responses are critically dependent on the expression of a diverse repertoire of antigen receptors — immunoglobulins and T cell receptors (TCRs) — by B and T cells. The genes that encode these antigen receptors are highly unusual in that they exist in a non-functional state in the germline, with the 5′ portion of each gene (which encodes the antigen binding domain) arranged as arrays of variable (V), diversity (D; only present in some loci) and joining (J) gene segments (FIG. 1). Assembly of these genes by V(D)J recombination generates antigen receptor diversity and is the central process around which early lymphocyte development is organized.

Multiple V(D)J recombination events must occur during the genesis of each new lymphocyte, and each event requires the introduction of chromosomal DNA double strand breaks. Such lesions are among the most dangerous that can be inflicted on the genome and, hence, it is not surprising that elaborate mechanisms have evolved to regulate the generation of these DNA breaks and to ensure their efficient repair. Despite this, it has become increasingly clear that the occasional errors that occur during V(D)J recombination are an important source of genome instability (particularly chromosomal translocations) and contribute to the development of lymphomas and leukaemias[1–3].

V(D)J recombination is initiated by the RAG recombinase (referred to hereafter as RAG) — a protein complex consisting primarily of the proteins encoded by recombination activating gene 1 (*RAG1*) and *RAG2*. RAG binds and cleaves the DNA at specific recombination signal sequences (RSSs) that flank each V, D and J gene segment. Thereafter, the DNA ends are processed and repaired by a large group of DNA repair enzymes, many of which are components of the classical non-homologous end joining (NHEJ) repair pathway. The resulting recombination event deletes or inverts a segment of chromosomal DNA that ranges from as small as a few hundred base pairs to as large as several million base pairs.

The initiation of V(D)J recombination is regulated at three distinct levels. First, the RAG proteins are expressed at high levels only during the early stages of lymphocyte development, thereby ensuring that the reaction does not occur in other tissues or cell types. Second, the ability of RAG to initiate V(D)J recombination is dictated by the 'accessibility' of RSSs within chromatin[4]. And third, V(D)J recombination is regulated by the position and three dimensional architecture of antigen receptor loci in the nucleus, with chromosome looping and condensation thought to have a vital role in allowing recombination between widely spaced gene segments.

This Review attempts to integrate prior knowledge in these areas with recent advances in our understanding of how chromosome architecture and covalent histone modifications coordinate the early steps of V(D)J recombination. We discuss the implications of these advances for the physiological and pathophysiological targeting of V(D)J recombination and propose a model for the stepwise engagement of the components of the two RSSs during a V(D)J recombination event.

V(D)J recombination biochemistry

DNA binding and cleavage by the RAG proteins. The RSSs that flank each gene segment consist of conserved heptamer and nonamer elements separated by a less well conserved spacer region of either 12 or 23 base

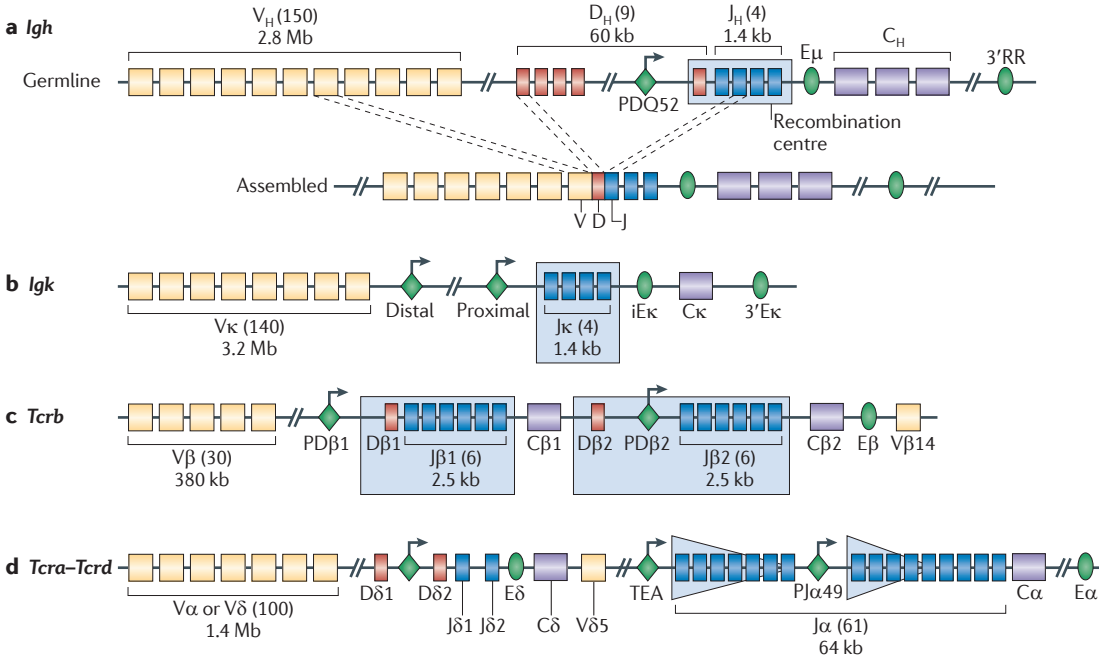


Figure 1 | **The structure of antigen receptor genes.** Schematic diagrams of the four mouse antigen receptor loci for which the binding patterns of the recombination activating gene (RAG) proteins have been reported. Variable (V), diversity (D) and joining (J) gene segments are represented as yellow, red and blue rectangles, respectively, constant (C) regions are shown as grey rectangles, enhancer elements as green ovals and germline promoters associated with recombination centres as green diamonds with arrows. The promoters associated with each V gene segment and the recombination signal sequences (RSSs) are not shown. The approximate sizes of the regions within each locus are indicated, together with the approximate number of gene segments in each region (shown in parentheses). The recombination centres (blue shaded areas) are the regions that are bound by RAG1 and RAG2 when the loci are in their germline configurations[85]. **a** | The immunoglobulin heavy chain (*Igh*) locus in its germline and assembled configurations. **b** | The germline *Igk* locus. **c** | The germline T cell receptor β-chain (*Tcrb*) locus. **d** | The germline *Tcra–Tcrd* locus. The *Tcra* recombination centre is represented by two triangles to depict the 5′ to 3′ gradient of RAG binding observed downstream of the T early α (TEA) element and the increase observed downstream of the Jα49 germline promoter[85].

Non-homologous end joining (NHEJ). A DNA repair process that joins broken DNA ends (double strand breaks) without using homologous DNA as a template. Components of this pathway include the proteins Ku70 (also known as XRCC6), Ku80 (also known as XRCC5), Artemis, X-ray repair cross-complementing protein 4 (XRCC4), DNA ligase IV and the catalytic subunit of DNA-dependent protein kinase (DNA-PKcs).

Chromatin
The combination of DNA, histones and other proteins that comprises eukaryotic chromosomes. The basic repeating unit of chromatin is the nucleosome, which consists of an octamer of histone proteins around which ~146 base pairs of DNA is wound.

Allelic exclusion
In theory, every B cell has the potential to produce two immunoglobulin heavy chains and two immunoglobulin light chains. In practice, however, a B cell produces only one immunoglobulin heavy chain and the majority produce only one immunoglobulin light chain. Similarly, most T cells produce only a single T cell receptor β-chain protein. The process by which the production of two different chains is prevented is known as allelic exclusion. Allelic exclusion is accomplished primarily through regulated V(D)J recombination.

pairs, which defines the 12RSS and 23RSS, respectively (BOX 1). There is a strong preference for recombination between a 12RSS and a 23RSS, a restriction known as the 12–23 rule[5]. RAG (BOX 2) recognizes RSSs and cleaves DNA during V(D)J recombination[6]. RAG1 is the principal DNA-binding component and contains most or all of the active site residues that catalyse DNA cleavage. RAG2 enhances RAG1–heptamer element interactions and is a vital cofactor for DNA cleavage[6,7]. RAG1 and RAG2 contain several important regulatory domains at their amino and carboxyl termini, respectively (BOX 2). DNA cleavage *in vitro* by RAG is strongly enhanced by high-mobility group protein B1 (HMGB1) or HMGB2 (which are ubiquitous, nonspecific DNA binding and bending proteins), but a role for HMGB proteins in V(D)J recombination *in vivo* has yet to be proven[6]. Hereafter, HMGB1 is used to refer to either protein.

RSS recognition is thought to occur via a capture model (FIG. 2a) in which the full complement of RAG1, RAG2 and HMGB1 proteins binds to one RSS to form the signal complex and then captures a second RSS that lacks bound proteins, forming the synaptic or paired complex[8,9]. DNA cleavage is a two-step process (FIG. 2b). RAG first introduces a single strand nick between the heptamer and the gene segment. The 3′ hydroxyl group that is created then attacks the other strand to generate a DNA double strand break, a step referred to as hairpin formation because it generates a covalently sealed hairpin at the end of the gene segment. Hairpin formation only occurs in the paired complex and is thought to take place simultaneously at the two RSSs. Nicking is less tightly regulated and can occur in the signal or paired complex[6,10].

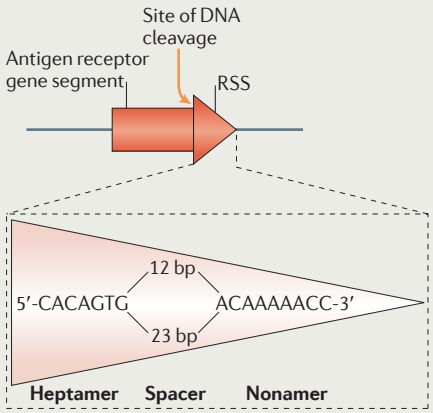
Order during V(D)J recombination. Antigen receptor loci recombine in a well-defined order during lymphocyte development[5,11,12]. The immunoglobulin heavy chain (*Igh*) and TCR β-chain (*Tcrb*) loci are the first to be assembled during the development of B cells and αβ T cells, respectively, with D-to-J recombination invariably preceding V-to-DJ recombination. Following this, recombination of the immunoglobulin light chain locus (either *Igk* or *Igl*) or the TCR α-chain locus (*Tcra*) is initiated. Most mature B cells and αβ T cells express only one functional *Igh* or *Tcrb* allele, respectively, and this phenomenon of allelic exclusion is accomplished through tight control of the recombination process. Immunoglobulin loci recombine fully only in the B cell lineage and TCR loci recombine only in developing

Box 1 | **The recombination signal sequence**

The recombination signal sequence (RSS) lies immediately adjacent to each antigen receptor gene segment and contains two well-conserved DNA elements (the heptamer and the nonamer) separated by a spacer region (see the figure). The consensus sequences of the heptamer and nonamer are the most efficient in mediating V(D)J recombination, but most endogenous RSSs deviate from the consensus at the heptamer, the nonamer, or both. The length of the spacer is tightly conserved, varying at most by 1 base pair from either 12 base pairs or 23 base pairs. Although certain spacer residues show significant conservation, the spacer is much more variable in sequence than the heptamer or the nonamer.

The nonamer is a sequence-specific binding site for the nonamer-binding domain of recombination activating gene 1 (RAG1) and functions to anchor the RAG proteins to the DNA. The heptamer has at least two roles. It enhances binding of the RAG recombinase to the RSS and it specifies the site of DNA cleavage, probably owing to its ability to facilitate the formation of an altered DNA structure at the coding flank–heptamer border. The first three nucleotides of the heptamer (5′-CAC) are particularly important in this regard and are essentially invariant in functional RSSs.

The length of the spacer is crucial for efficient V(D)J recombination, suggesting that its primary role is to properly align the heptamer and nonamer. The sequence of the spacer can also influence the efficiency of V(D)J recombination, although usually to only a small extent. Similarly to the heptamer and nonamer, the spacer is also involved in protein–DNA interactions in RAG–high-mobility group protein B (HMGB)–RSS complexes. See REF. 6 for further details.



T cells. Such developmentally regulated and lineage-specific recombination occurs despite the use of fairly well conserved RSSs and the same recombination machinery for all of the loci. For the last 25 years, the dominant and highly successful model used to explain this has been that involving the accessibility of RSSs[4,13]. In the last 10 years, it has become clear that two additional levels of regulation — antigen receptor locus architecture and nuclear location — work together with locus accessibility to ensure order, fidelity and allelic exclusion during V(D)J recombination.

Accessibility and transcription

Chromatin and the inaccessible state. The accessibility model emerged from the observation that *Igh* variable (V_H) gene segments undergo germline transcription (also known as sterile transcription) coincident with their recombination[4]. The model was subsequently reinforced by a wealth of observations demonstrating that V(D)J recombination correlates with numerous markers of open chromatin, including germline transcription, activating histone modifications (such as histone acetylation), nuclease accessibility and DNA hypomethylation[11,12,14–17]. Together, these correlations indicated a tight connection between chromatin structure and the targeting of V(D)J recombination. This notion received strong support from experiments showing that when isolated lymphocyte nuclei were incubated with RAG-containing nuclear extracts, cleavage could be detected at antigen receptor gene RSSs in a pattern that recapitulated the lineage and developmental specificity of V(D)J recombination[18]. These results demonstrated that chromatin can be permissive or repressive for V(D)J recombination and, importantly, that chromatin structure affects the first (DNA cleavage) phase of the reaction.

One simple mechanism to explain the repressive effects of chromatin was revealed when it was discovered that wrapping an RSS around a nucleosome renders it refractory to binding or cleavage by the RAG proteins[19–22]. Proteolytic removal of N-terminal histone tails[20,22] or the action of ATP-dependent chromatin remodelling complexes such as the SWI–SNF ATPase complex[22–25] allows substantial cleavage of nucleosomal RSS substrates by RAG. By contrast, histone acetylation by itself has no[20,21] or only a modest[22] ability to stimulate cleavage of nucleosomal RSS substrates, although it can enhance the effects of chromatin remodelling complexes in such assays[22,24]. Together, these biochemical studies strongly support the idea that the nucleosome is intrinsically inhibitory to DNA cleavage by RAG.

To what extent are RSSs positioned on nucleosomes in the nucleus of developing lymphocytes? An initial study found that the RSS, particularly the nonamer element, serves as a nucleosome-positioning sequence *in vitro* and on plasmid substrates *in vivo*[26]. Therefore, RSSs should exhibit a high frequency of nucleosome occupancy in antigen receptor loci in the absence of mechanisms to reposition nucleosomes. However, a recent analysis of nucleosome occupancy in endogenous antigen receptor loci found no clear correlation between the location of RSSs and the position of nucleosomes[27]. Inaccessible chromatin was characterized by a high nucleosome density but even in this circumstance, some RSSs were at internucleosomal positions[27]. Thus, at least for the portions of the *Tcra* and *Tcrb* loci examined in this study, RSSs do not function as dominant nucleosome-positioning sequences. This is consistent with the emerging view that although the primary DNA sequence directly influences nucleosome positioning, many other factors (such as remodelling complexes, bound proteins, promoter-loaded RNA polymerase II

Germline transcription
Transcription of unrearranged antigen receptor gene loci that begins before or coincident with their activation. It is not thought to produce functional protein, and the promoter and initiation sites are often lost in the subsequent rearrangement events.

Nucleosome
The fundamental structural unit of eukaryotic chromosomes. It consists of pairs of each of the core histones (H2A, H2B, H3 and H4), thereby creating the histone octamer, and a single molecule of the linker histone H1. The nucleosome spans ~146 base pairs of DNA.

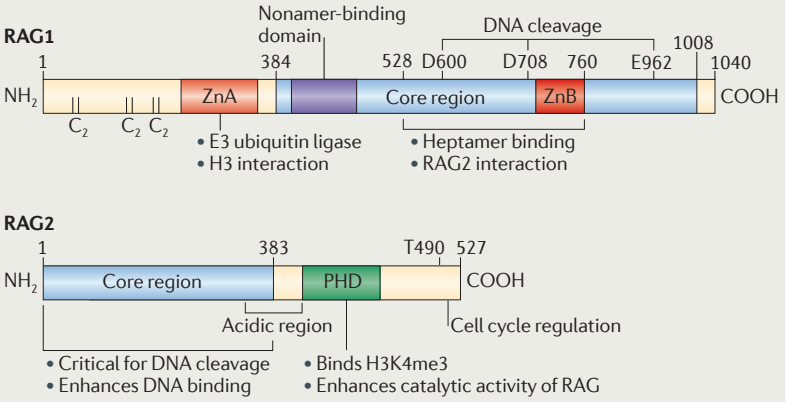
Chromatin remodelling complex
An enzymatic complex that remodels the DNA–nucleosome architecture and thus can determine transcriptional activity. The SWI–SNF ATPase is an example of a complex that remodels chromatin.

Box 2 | **The RAG proteins**

Recombination activating gene 1 (RAG1) and RAG2 are nuclear proteins that are well-conserved in jawed vertebrates. They interact with one another, and thousands of copies are estimated to be present in each developing lymphocyte[104]. 'Core' regions of each RAG protein have been defined as the minimal portion of the protein required for V(D)J recombination activity. The RAG1 core contains a well-defined nonamer-binding domain and a central region (amino acids 528–760) that is responsible for interactions with the heptamer and RAG2 (see the figure). Zinc finger region B (ZnB) is thought to be important for RAG2 interactions. The RAG1 core also contains the active site for DNA cleavage, which includes three acidic amino acids (D600, D708 and E962) that coordinate divalent metal ions and are essential for catalysis. The RAG1 amino-terminal region (amino acids 1–383) enhances V(D)J recombination activity but the mechanism by which it does so is not well understood. This region contains several pairs of conserved cysteine residues (labelled C_2) and ZnA, which homodimerizes, has E3 ubiquitin ligase activity and interacts with and ubiquitylates histone H3 (REF. 88).

The RAG2 core is crucial for DNA cleavage activity, interacts with RAG1 and enhances the DNA-binding affinity and specificity of the RAG complex, although RAG2 has little or no DNA binding activity on its own. The RAG2 carboxy-terminal region (amino acids 384–527) contains a plant homeodomain (PHD) finger that binds specifically to trimethylated histone H3 lysine 4 (H3K4me3), enhances the catalytic activity of the RAG complex and guides RAG2 to regions of active chromatin.

Although RAG1 is expressed throughout the cell cycle, RAG2 is only stable in G0 or G1 phase cells owing to the phosphorylation of T490 in S, G2 and M phase cells[105,106]. RAG2 also contains an acidic region that interacts with histones[79]. Residue numbers refer to the mouse RAG proteins. See REFS 6,107,108 for further details.



and DNA methylation) also have important roles[28–30]. It is also interesting that the sequence AAAAA, which is found in many endogenous nonamers, is poorly incorporated into nucleosomes, which probably reflects intrinsic properties of homopolymeric deoxyadenosine (AAAAA) tracts[28,31]. Taken together, existing data suggest that inaccessibility to the V(D)J recombination machinery is mediated by both direct positioning of RSSs on nucleosomes and steric constraints created by nucleosome packing[21,27].

Transcription and the generation of accessibility. The tight correlation that exists between germline transcription and V(D)J recombination suggests that transcriptional control elements (such as enhancers and promoters), transcription factors and perhaps transcription itself have important roles in rendering RSSs accessible for recombination[11,12,14,32–35]. The deletion of enhancer elements from an antigen receptor locus typically results in a near-total block in recombination of the affected locus, accompanied by a loss of germline transcription and other markers of locus accessibility. By contrast, deletion of promoter elements has a more local effect, reducing recombination specifically in the region downstream of the promoter.

For many years it was not known whether germline transcription was causally related to the generation of accessibility during V(D)J recombination or whether it was merely a side effect of other processes that were instead responsible for accessibility. This issue has now been conclusively resolved, at least for a portion of the *Tcra* locus, through elegant experiments involving the insertion of a transcription terminator into two different locations in the Jα cluster. The result in each case was a block in transcription and recombination of the Jα gene segments downstream of the terminator[36,37]. Hence, transcriptional elongation by RNA polymerase II has a direct and crucial role in creating the accessible chromatin that is required for *Tcra* recombination. Although this is likely to be the case for other antigen receptor loci as well, it is not yet clear if it is a universal rule. Several examples of a dissociation between V(D)J recombination and germline transcription have been reported[12,16,38,39], and recent biochemical and *in vivo* experiments point to a central role for chromatin remodelling complexes in creating accessibility (BOX 3).

The strong connection between transcription and V(D)J recombination extends also to antisense transcription. Developmentally regulated antisense transcription of V_H gene segments[40] and the D_H–J_H region is associated with *Igh* locus assembly[41–43]. Antisense transcription has been suggested to have several different functions in V(D)J recombination. First, it might increase local RSS accessibility by mechanisms similar to that of germline (sense) transcription[40,41]. The strongest support for this comes from a recent study[44] in which the authors created a germline deletion of the interval between the V_H and D_H gene segments. This led to the production of long antisense transcripts running from D_H into the 3′ part of the V_H cluster, and also resulted in V_H-to-D_HJ_H recombination in the thymus, where it is normally not detected. Second, based on the findings that the central D_H gene segments are contained in a repeat unit, are poorly recombined, are transcribed bidirectionally and are associated with repressive histone marks, it has been proposed that antisense transcription contributes to the generation of heterochromatin and the suppression of V(D)J recombination by a double stranded RNA (dsRNA)-dependent repeat-induced gene silencing mechanism[42]. Third, it has been hypothesized that the generation of non-coding RNAs (sense and/or antisense) or the non-coding RNAs themselves influence the three dimensional structure of local chromatin domains or the higher order configuration of the chromatin at antigen receptor loci[34].

Steps in local accessibility. These considerations and others led to the following model for how RSSs are freed from the repressive effects of chromatin and the nucleosome to allow RAG binding[11,14–17,32,34,45]. The first step — initial locus opening — might be independent of well-characterized transcriptional control elements.

Antisense transcription
Transcription in the opposite direction and of the opposite strand from that used to generate the normal product of a gene. It is not thought to generate a protein product but instead might alter chromatin structure either directly (via the act of transcription) or indirectly (via the antisense RNA produced).

Heterochromatin
High-density regions in the nucleus that are thought to contain compacted chromatin structures associated with silent genes.

DNase I hypersensitive site
A site of nuclease sensitivity when nuclei from cells are exposed to limiting concentrations of the enzyme DNase I. The digested regions of DNA correspond to sites of open DNA, which might be transcription factor binding sites or areas of altered nucleosome conformation.

Bromodomain
A module of ~110 amino acids that is found in several transcriptional regulators. A bromodomain consists of a four-helix bundle with a single binding pocket for *N*ε-acetyl-lysine on histone tails.

In the *Igh* locus, loss of the repressive histone mark dimethylated histone H3 lysine 9 (H3K9) and an increase in the activating mark dimethylated H3K4 occur in the absence of the core *Igh* intronic enhancer (Eμ) and other known DNase I hypersensitive sites[46]. Although the mechanism for this is unknown, it is attractive to think that such changes allow transcription factors to gain access to Eμ[46]. Methylation of H3K9 is sufficient to suppress the recombination of an otherwise accessible minilocus[47], and the presence of this marker correlates with inactive antigen receptor loci or gene segments[48,49] and can trigger the formation of heterochromatin[50]. Hence, reducing H3K9 methylation is likely to be an important early step in creating accessibility.

In the second step of the model, transcription factors bind to the enhancer and recruit histone acetyltransferases (HATs) that acetylate the N-terminal tails of histones H3 and H4. This provides binding sites for the bromodomains in chromatin remodelling complexes, which in turn free up promoters for binding by RNA polymerase II, thereby allowing the initiation of germline transcription. Many factors (such as HATs, histone methylases, chromatin remodelling complexes and histone chaperones) associate and move with RNA polymerase II through the transcription unit, leading to further nucleosome remodelling and additional activating histone modifications[51,52]. One of the most important of these modifications is H3K4 trimethylation (H3K4me3), which is strongly biased towards the 5′ end of the transcription unit and the region surrounding the promoter[53]. In the fully activated antigen receptor locus, three principal mechanisms probably work together to free RSSs from repressive associations with nucleosomes: one, histone acetylation, which weakens the association of histone tails with DNA, loosens the chromatin fibre[54] and enhances chromatin remodelling[22,24]; two, chromatin remodelling, which could generate nucleosome-free RSSs by repositioning or evicting nucleosomes, or could act more subtly by transiently lifting a loop of DNA off of the nucleosome surface[55]; and three, transcriptional elongation by RNA polymerase II through RSSs, resulting in the disassembly of nucleosomes or their disassociation from DNA[51]. The relative importance of each of these mechanisms probably varies between different RSSs and different loci. In addition to these mechanisms, there are other, RSS-specific, pathways for the delivery of the RAG proteins. For example, the transcription factor FOS (a component of the AP1 dimer) has been implicated in the recruitment of RAG to the 3′ Dβ1 23RSS via an embedded AP1 binding site[56], and the B cell-associated transcription factor PAX5 (paired box protein 5) has been suggested to facilitate RAG recruitment to V_H gene segments[57].

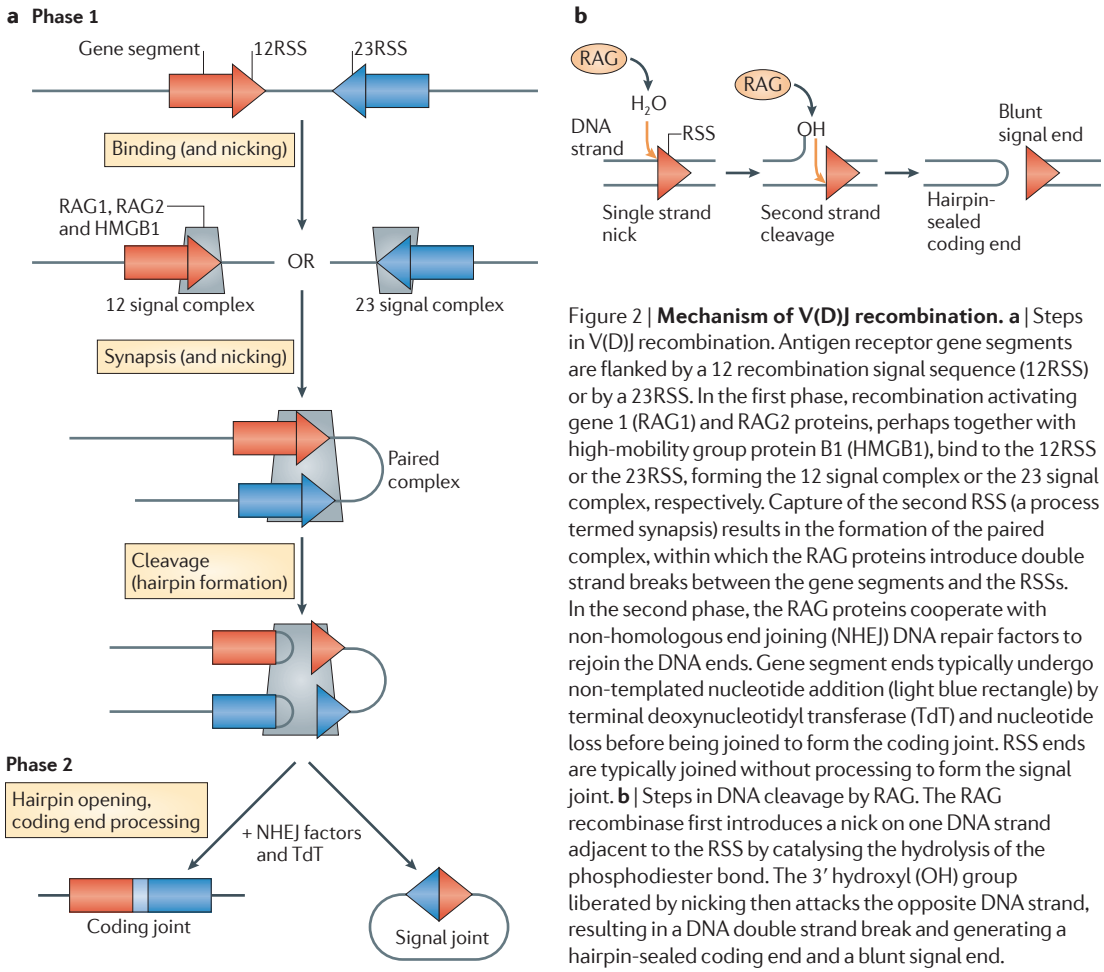


Figure 2 | **Mechanism of V(D)J recombination. a** | Steps in V(D)J recombination. Antigen receptor gene segments are flanked by a 12 recombination signal sequence (12RSS) or by a 23RSS. In the first phase, recombination activating gene 1 (RAG1) and RAG2 proteins, perhaps together with high-mobility group protein B1 (HMGB1), bind to the 12RSS or the 23RSS, forming the 12 signal complex or the 23 signal complex, respectively. Capture of the second RSS (a process termed synapsis) results in the formation of the paired complex, within which the RAG proteins introduce double strand breaks between the gene segments and the RSSs. In the second phase, the RAG proteins cooperate with non-homologous end joining (NHEJ) DNA repair factors to rejoin the DNA ends. Gene segment ends typically undergo non-templated nucleotide addition (light blue rectangle) by terminal deoxynucleotidyl transferase (TdT) and nucleotide loss before being joined to form the coding joint. RSS ends are typically joined without processing to form the signal joint. **b** | Steps in DNA cleavage by RAG. The RAG recombinase first introduces a nick on one DNA strand adjacent to the RSS by catalysing the hydrolysis of the phosphodiester bond. The 3′ hydroxyl (OH) group liberated by nicking then attacks the opposite DNA strand, resulting in a DNA double strand break and generating a hairpin-sealed coding end and a blunt signal end.

Box 3 | **Creating accessibility: transcription versus chromatin remodelling**

In biochemical experiments using chromatinized substrates containing a 12 recombination signal sequence (12RSS) and a 23RSS, cleavage by the recombination activating gene (RAG) recombinase was observed only when a transcriptional activation domain was tethered in close proximity to both RSSs. The activation domain enhanced cleavage by recruiting the SWI–SNF chromatin remodelling complex, and transcription was neither required nor stimulatory for cleavage[25]. Therefore, in this well-defined system, chromatin remodelling was sufficient to confer accessibility in the absence of transcription.

In vivo studies of the T cell receptor β-chain (*Tcrb*) locus also highlight the importance of chromatin remodelling. Recombination of the *Tcrb* locus depends on the action of multiple transcriptional control elements (FIG. 1c), including the *Tcrb* enhancer (Eβ), which is required for recombination and germline transcription of both Dβ–Jβ clusters[38,109–113]; PDβ1, the promoter associated with Dβ1, which is required for recombination and germline transcription of the Dβ1–Jβ1 cluster but not the Dβ2–Jβ2 cluster[114,115]; and PDβ2, a complex promoter that is a strong candidate for controlling recombination of the Dβ2–Jβ2 cluster[116,117]. Recombination of the Dβ1–Jβ1 cluster is associated with DNA looping and the formation of a holocomplex containing Eβ, PDβ1, transcription factors and components of the transcriptional apparatus[112].

The SWI–SNF chromatin remodelling complex is recruited to the Dβ1–Jβ1 region in a PDβ1- and Eβ-dependent manner and is required for efficient Dβ1-to-Jβ1 recombination[118]. Importantly, tethering of the chromatin remodelling factor BRG1 to a site immediately adjacent to Dβ1 compensated for deletion of PDβ1 in a recombination minilocus. However, not only were Dβ1-to-Jβ1 recombination and Dβ1 accessibility restored by BRG1 tethering, but germline transcription of the Dβ1–Jβ1 region was also restored[118]. Hence, although it is clear that recruitment of the SWI–SNF complex can substitute for the PDβ1 promoter, it remains unclear if chromatin remodelling was sufficient to create accessibility in the absence of transcription in this instance. A strong association between recruitment of the SWI–SNF complex, antisense transcription and recombination has also been observed at the immunoglobulin heavy chain locus (*Igh*)[43].

Beyond accessibility

Accessibility, although probably required for the recombination of any particular RSS, is not always sufficient. For example, in PAX5-deficient pro-B cells, 5′ V_H gene segments fail to recombine despite exhibiting wild-type levels of germline transcription and histone acetylation[58]. Instances of accessible *Tcrb* variable (Vβ) gene segments that fail to undergo recombination have also been reported[59]. Current evidence suggests that at least two conditions, in addition to that of accessibility, must often be met before V(D)J recombination can occur. First, the recombining locus needs to move away from repressive chromatin compartments (such as the periphery of the nucleus and pericentric heterochromatin); and second, for long-distance recombination events, the locus must undergo a large-scale structural alteration (for example, compaction or looping) to allow widely separated gene segments to encounter one another with reasonable efficiency[34].

Location in the nucleus. Both the periphery of the nucleus and regions of pericentric heterochromatin have been implicated in gene silencing[60] — perhaps most directly by the finding that forced recruitment of an expression cassette to the inner nuclear membrane results in transcriptional repression[61]. Considerable evidence now indicates that antigen receptor loci move away from these nuclear compartments when they undergo recombination and associate with them following the termination of recombination[16,34,62]. Regulated[63–65] or stochastic[66] associations of antigen receptor loci with repressive nuclear compartments have been implicated in the initiation and maintenance of allelic exclusion. After the successful recombination of one allele, recombination of the other must often be suppressed. In the case of the *Igh* locus, this has been associated with interaction of the second allele with an *Igk* allele[67] and relocation to pericentric heterochromatin[68].

The *Tcra* and *Tcrd* loci, which are not subject to allelic exclusion, do not undergo a developmentally regulated association with heterochromatin[63]. However, the connection between loss of accessibility, suppression of V(D)J recombination and association with a repressive chromatin compartment is neither direct nor simple. Germline transcription of the Vβ region has been demonstrated to occur on both alleles in pro-T cells[69], even though many or most of the alleles are associated with a repressive compartment[63,66]. And *Igk* undergoes biallelic germline transcription in pre-B cells[70,71] despite the fact that one allele is associated with heterochromatin. Hence, the mechanism by which locus association with pericentric heterochromatin might inhibit V(D)J recombination remains unknown and the inhibitory nature of the interaction remains to be proven.

Antigen receptor locus architecture. Fluorescence *in situ* hybridization (FISH) studies have demonstrated a striking correlation between V(D)J recombination and changes in the large scale chromatin structure of antigen receptor loci. The *Igh*, *Igk*, *Tcrb* and *Tcra* loci all adopt a contracted configuration coincident with their recombination and revert to a decondensed state after their recombination is complete[62,63,68,72]. Contraction is at least in part the result of chromatin looping that brings distal V gene segments into close proximity with the (D)J segment[63,72,73], and this should facilitate long range recombination events involving distal V gene segments. The strongest evidence for this comes from the analysis of PAX5-deficient pro-B cells, in which *Igh* locus contraction and distal V_H gene segment recombination both fail to occur; expression of PAX5 in these cells overcomes both defects[74].

Pro-B cell
A cell in the earliest stage of B cell development in the bone marrow. Pro-B cells are characterized by incomplete immunoglobulin heavy chain rearrangements and are defined as $CD19^+$ and cytoplasmic IgM^- or, sometimes, as $B220^+CD43^+$ (by the Hardy classification scheme).

Pericentric heterochromatin
Regions of very densely packed chromatin fibres located near the centromere of each chromosome. These regions are typically inactive and often cluster to form discrete clumps in the nucleus.

Fluorescence *in situ* hybridization
(FISH). The use of fluorescent probes to visually label specific DNA sequences in the nuclei of cells that are in the interphase or metaphase stages of mitosis.

A major advance in our understanding of the structure and dynamics of antigen receptor loci comes from a recent high resolution (<50 nm) FISH analysis of the *Igh* locus[75]. Using data derived from 12 markers spanning the 3 Mb *Igh* locus, the authors could calculate an average trajectory for the *Igh* locus in haematopoietic cells before commitment to the B cell lineage (pre-pro-B cells) and in pro-B cells (in which *Igh* recombination occurs). In pre-pro-B cells, the *Igh* locus is relatively decondensed, with distal V_H gene segments in a distinct domain that is well separated from the D_H–J_H region. The V_H portion of the locus could best be modelled as a series of 1 Mb DNA rosettes made up of multiple loops of ~120 kb, with the rosettes separated by linkers of ~60 kb. It seems likely that the rosettes are dynamic structures in which loops of variable sizes undergo frequent folding and unfolding[34,75]. In pro-B cells, the locus contracts into a single condensed domain, in which the more distal V_H gene segments have moved substantially so that all the V_H gene segments are approximately equidistant from the D_H–J_H region. Interestingly, despite its greater overall condensation, the *Igh* locus in pro-B cells exhibits greater conformational flexibility than in pre-pro-B cells[75]. Recombinationally active loci are therefore proposed to adopt a condensed but dynamic chromatin architecture that ensures that both proximal and distal V gene segments have a roughly equal chance of encountering and recombining with the (D)J segment[34,75].

The idea that regulation of locus architecture has a key role in controlling V gene segment usage received further support from a recent study of the *Tcra–Tcrd* locus[76]. In pro-T cells, in which all of the V gene segments must have the chance to undergo recombination during *Tcrd* locus assembly, both the 3′ and 5′ halves of the Vδ region are contracted and presumably in close proximity to the Dδ–Jδ region. However, in pre-T cells, in which initial *Tcra* rearrangements preferentially involve 3′ Vα gene segments, the 3′ portion of the locus is contracted but the 5′ portion is not[76]. Recombination of 5′ Vα gene segments would therefore be unlikely to occur until after an initial 3′ Vα-to-Jα recombination event has deleted a portion of the locus and brought the 5′ Vα segments closer to the Jα region.

In summary, large scale movement and structural reorganization of antigen receptor loci provide important layers of regulation, in addition to that of RSS accessibility, for the proper orchestration of V(D)J recombination.

Targeting of RAG proteins

How do the RAG proteins find their way to accessible RSSs? Is it a passive process that relies on their sequence-specific DNA binding properties, or is it facilitated by other interactions? And where in the large antigen receptor loci do the RAG proteins bind to initiate V(D)J recombination? Do they bind uniformly to the many available RSSs, or is binding restricted to a particular portion of each locus? Recent advances have begun to answer these fundamental questions.

RAG2: reader of the histone code. The first major advance in this area was triggered by the observations that the C-terminal domain of RAG2 contains a non-canonical plant homeodomain (PHD) finger[77,78] and that this region of RAG2 can interact with histones[79]. Two groups then made the seminal discovery that the RAG2 PHD finger specifically recognizes H3K4me3 and that this interaction is important for efficient V(D)J recombination[80,81]. RAG2 mutations that abrogate this interaction substantially reduce the efficiency of V(D)J recombination[80,81], as does a reduction in H3K4me3 levels[80]. Furthermore, mutation of a single critical residue within the RAG2 PHD finger (tryptophan 453) is associated with Omenn syndrome, which is a rare condition characterized by severe immunodeficiency that is strongly linked to inefficient V(D)J recombination[82]. Because H3K4me3 is highly enriched near the 5′ end of transcription units, the PHD finger of RAG2 should facilitate recruitment of RAG to RSSs located close to active promoters. Interestingly, recognition of H3K4me3 by RAG2 does more than simply enhance RAG binding; recent biochemical experiments indicate that it also enhances the catalytic activity of RAG[83,84].

Focal RAG binding in recombination centres. It is surprising that V(D)J recombination has been studied intensively for many years without an understanding of where the RAG proteins bind within antigen receptor loci. The recent development of a RAG chromatin immunoprecipitation (ChIP) assay has allowed for substantial insights into this problem[85]. In the *Igh*, *Igk*, *Tcrb* and *Tcra* loci, both RAG1 and RAG2 were found to bind specifically to a small region containing J gene segments (and the D gene segments that lie very close to the J region in *Igh* and *Tcrb*) (FIG. 1). In no case was binding detected at V or more distal D gene segments. These focal regions of RAG binding, referred to as recombination centres, display the features one would expect for accessible RSSs — namely, substantial germline transcription and high levels of histone H3 acetylation, H3K4me3 and RNA polymerase II occupancy. The dynamics of RAG binding closely mirror the developmental stage- and lineage-specific pattern of V(D)J recombination events, with two interesting exceptions. The binding of RAG to *Igh* or *Tcrb* is robust in pre-B and pre-T cells, respectively, cells in which V-to-DJ recombination is suppressed and one allele is associated with pericentric heterochromatin[63,68]. Hence, allelic exclusion of *Igh* in pre-B cells or *Tcrb* in pre-T cells does not involve the exclusion of RAG from these loci, and association with pericentric heterochromatin might be compatible with RAG binding. It remains unclear how recombination of V_H or Vβ gene segments that are proximal to a recombination centre is prevented in pre-B and pre-T cells, respectively.

Interestingly, each RAG protein can recapitulate its highly specific pattern of binding in the absence of the other, with one exception noted below[85]. How might this occur? For RAG1, specific recruitment

Chromatin immunoprecipitation
An experimental technique that analyses direct binding of an endogenous transcription factor to chromatin by fixation with formaldehyde followed by immunoprecipitation with a transcription factor-specific antibody. Gene-specific enrichment is then assessed by polymerase chain reaction analysis of the immunoprecipitated DNA.

Recombination centre
A region of an antigen receptor locus that is characterized by strong binding of recombination activating gene 1 (RAG1) and RAG2 and high levels of germline transcription, RNA polymerase II, histone acetylation and trimethylated histone H3 lysine 4 (H3K4me3).

into recombination centres in the absence of RAG2 is probably mediated, at least in part, by direct RSS recognition. This suggestion is supported by experiments demonstrating that both the RSS and the nonamer-binding domain of RAG1 are required for the binding of RAG1 to a chromosomal recombination substrate[85]. Furthermore, RAG1 by itself has been shown to bind relatively tightly and with a slow dissociation kinetics to the 12RSS *in vitro*[86]. Hence, it appears that RAG1 has substantial RSS-specific binding activity *in vivo* in the absence of RAG2.

For RAG2, the specific pattern of binding observed in the absence of RAG1 is probably due to the recognition of H3K4me3 by the RAG2 PHD finger. In support of this, a genome wide ChIP analysis revealed that RAG2 is recruited to thousands of active regions of the genome, in a pattern essentially identical to that of H3K4me3 (REF. 85). Intriguingly, several promoter regions in genes that do not encode antigen receptors showed substantial RAG2 binding but were found to lack detectable RAG1 binding. This indicates that developing lymphocytes contain a pool of RAG2 that is not associated with RAG1, although how widespread this phenomenon is remains to be determined.

The tight link between accessibility and transcription leads to the strong prediction that binding of RAG depends on transcription and on transcriptional control elements. Indeed, the deletion of enhancer elements results in a global loss of RAG1 binding at the *Tcrb* and *Tcra* loci, and promoter deletion causes a local defect in RAG1 binding in the region downstream of the promoter[87]. Importantly, blocking transcriptional elongation in the Jα cluster results in defective RAG1 binding immediately downstream of the blockade. Hence, promoters, enhancers and transcription elongation have a crucial role in creating the environment needed for RAG binding[87].

The recombination centre model. These findings have led to the proposal that the complex series of events that occur during V(D)J recombination are orchestrated within recombination centres[85]. The recombination centre model (FIG. 3) begins with the creation of a focal region of highly accessible RSSs characterized by elevated levels of germline transcription, histone acetylation and H3K4me3. These conditions are ideal for the recruitment of RAG1 and RAG2, which probably engage both H3K4me3 and the RSSs within the recombination centre. This is followed by capture of a partner RSS to form the paired complex, and this in turn leads to DNA cleavage and recombination. When the partner is a distal V gene segment, architectural reorganization (such as looping or contraction) of the locus is probably required to bring the V gene segment into the proximity of the recombination centre. Our current simplistic conception is that V gene segments form a 'cloud' around the recombination centre and that the dynamic formation, movement and breakdown of chromatin loops containing the V gene segments gives each gene segment a reasonable chance of encountering the recombination centre.

Unanswered questions and future directions

Pathways of RAG binding. The finding that each RAG protein can independently find its way into recombination centres suggests that there are at least three different pathways of RAG recruitment: one, binding of RAG1 to the RSS followed by recruitment of RAG2 to the bound RAG1; two, binding of RAG2 to H3K4me3 followed by recruitment of RAG1 to the bound RAG2; and three, recruitment of a preformed RAG1–RAG2 complex via interactions with the RSS and H3K4me3.

The first pathway might be important in cycling cells. In the S, G2 and M phases of the cell cycle, RAG2 levels are very low and RAG1 would be predicted to bind accessible RSSs within recombination centres. Then, as cells re-enter the G1 phase and RAG2 levels rise, RAG2 could be recruited to the pre-bound RAG1. The second pathway might occur at the large number of sites outside of antigen receptor loci that lack a strong binding site for RAG1 but contain substantial levels of H3K4me3. The third pathway might dominate in cells expressing substantial levels of both RAG1 and RAG2, as the two proteins interact with one another and the resulting complex binds to the RSS with greater specificity and affinity than RAG1 alone[6]. It is appealing to think that the RAG1–RAG2 complex, with its ability to interact with both the RSS and H3K4me3, would be recruited more rapidly and efficiently into recombination centres than either RAG protein alone. The *Igh* locus appears to provide an example of this, as binding of RAG1 to the *Igh* recombination centre is markedly reduced in the absence of RAG2 (REF. 85). It is not clear why this was observed at *Igh* but not at the other antigen receptor loci examined[85].

It will be important to determine the extent to which the two RAG proteins exist separately from one another *in vivo* and how dynamic their association is. It will also be important to determine the extent to which the association of RAG1 with chromatin is influenced by its recently reported ability to interact with and ubiquitylate histone H3 (REF. 88). This information, together with additional genome-wide binding studies, should shed considerable light on the relative importance of the three pathways of RAG recruitment.

Properties of recombination centres. Many of the basic features of recombination centres remain to be characterized. All of the recombination centres defined thus far contain multiple J gene segments (and associated RSSs) within a 2–10 kb region, but it is not known whether the relatively close spacing of the RSSs or even the presence of more than one RSS in the recombination centre contributes to RAG recruitment. It also remains unclear how many RAG molecules are present within a recombination centre, how many RSSs within a recombination centre are engaged by RAG at any one time in any particular cell and how dynamic the RAG–RSS interaction is.

Regulatory potential of recombination centres. A particularly attractive feature of recombination centres is their potential to regulate both the DNA cleavage and the repair phases of V(D)J recombination. Because initial binding of RAG appears to be restricted to the J and proximal D

Cryptic RSS
A region of DNA that resembles a true recombination signal sequence (RSS) in some of its functionally important sequence features but does not lie adjacent to an antigen receptor gene segment.

gene segments of antigen receptor loci, V and distal D gene segments have a reduced ability to engage in undesirable V(D)J recombination events with one another, with RSSs in another antigen receptor locus, or with cryptic RSSs or alternative DNA structures[89] in genes that do not encode antigen receptors. This initial binding preference should also limit the incidence of single strand nicking by RAG at V and distal D RSSs. RAG-generated nicks have been demonstrated to stimulate homologous recombination and thus could contribute to genome instability[90].

Simultaneous engagement of multiple RSSs in a single recombination centre, if it occurs, would raise the possibility that more than one recombination event occurs on an allele at one time. This is neither necessary for gene assembly nor desirable from the standpoint of genome stability. So, do deterministic mechanisms exist to prevent this, or is it made unlikely by virtue of inefficient capture of a partner RSS? Similarly, how do developing lymphocytes avoid initiating V(D)J recombination on both alleles of the *Igh*, *Igk* or *Tcrb* loci simultaneously and thus prevent the risk of a violation of allelic exclusion? Interestingly, it was recently observed that both the *Igh* and *Igk* loci exhibit a phenomenon of developmentally regulated pairing of homologous chromosomes, and that this is reduced in the absence of RAG1 (REF. 64). Other data has suggested that homologue pairing and

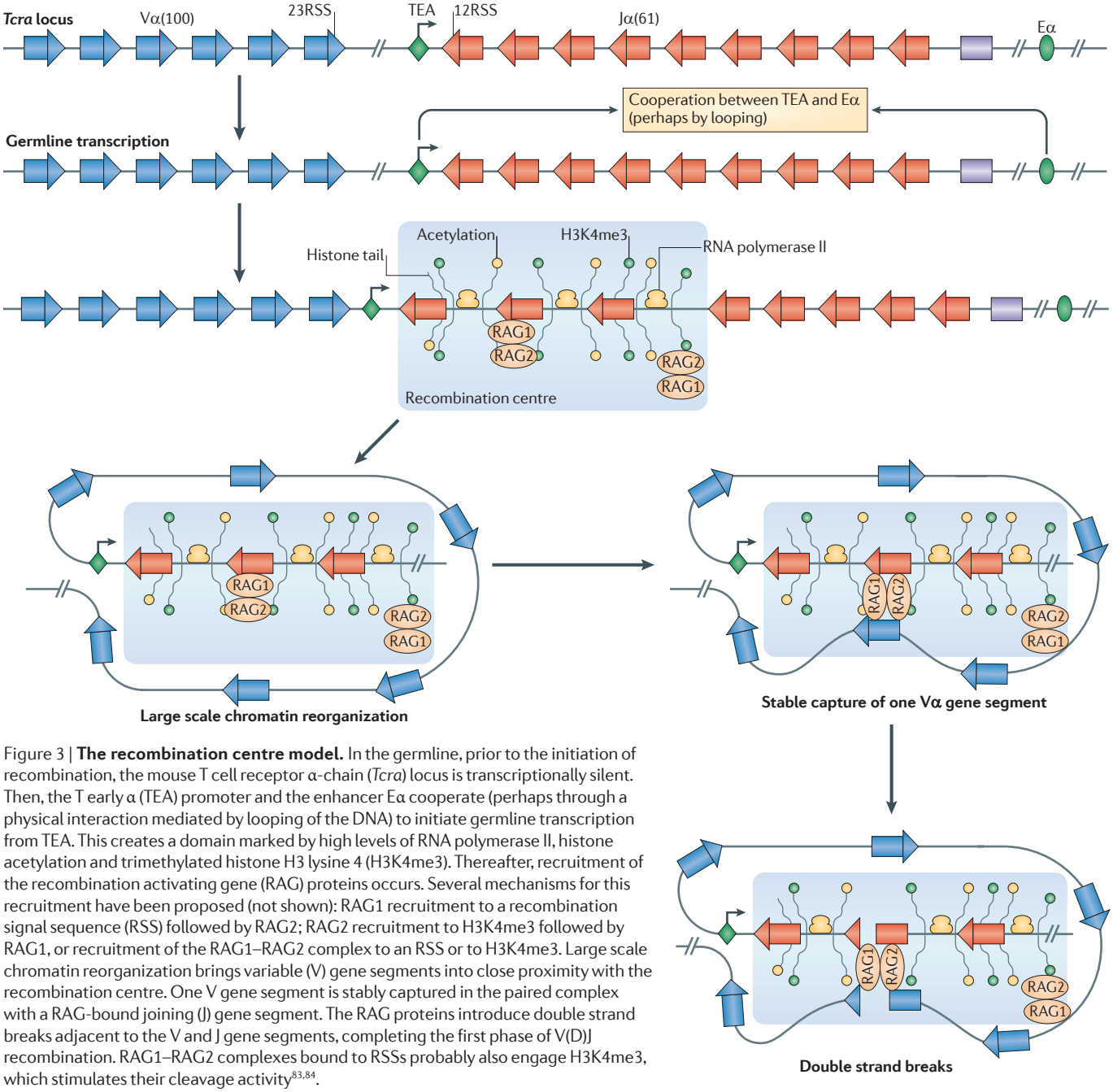


Figure 3 | **The recombination centre model.** In the germline, prior to the initiation of recombination, the mouse T cell receptor α-chain (*Tcra*) locus is transcriptionally silent. Then, the T early α (TEA) promoter and the enhancer Eα cooperate (perhaps through a physical interaction mediated by looping of the DNA) to initiate germline transcription from TEA. This creates a domain marked by high levels of RNA polymerase II, histone acetylation and trimethylated histone H3 lysine 4 (H3K4me3). Thereafter, recruitment of the recombination activating gene (RAG) proteins occurs. Several mechanisms for this recruitment have been proposed (not shown): RAG1 recruitment to a recombination signal sequence (RSS) followed by RAG2; RAG2 recruitment to H3K4me3 followed by RAG1, or recruitment of the RAG1–RAG2 complex to an RSS or to H3K4me3. Large scale chromatin reorganization brings variable (V) gene segments into close proximity with the recombination centre. One V gene segment is stably captured in the paired complex with a RAG-bound joining (J) gene segment. The RAG proteins introduce double strand breaks adjacent to the V and J gene segments, completing the first phase of V(D)J recombination. RAG1–RAG2 complexes bound to RSSs probably also engage H3K4me3, which stimulates their cleavage activity[83,84].

the DNA damage response protein ataxia telangiectasia mutated (ATM) are involved in ensuring RAG-mediated cleavage of only one allele at a time[64]. One speculative possibility is that homologue pairing is mediated by a physical interaction between the recombination centres on the two alleles; this interaction could then facilitate communication between the two recombination centres so as to prevent initiation of V(D)J recombination simultaneously on the two alleles.

It is attractive to think that the recombination centre serves as a hub, or 'V(D)J recombination factory'[91], within which both the DNA cleavage and DNA repair phases of the reaction are orchestrated. The other factors (proteins or RNA) found within recombination centres are not known, although it is likely that RNA polymerase II, other components of the transcription apparatus, histone modifying enzymes and chromatin remodelling complexes are present before DNA cleavage takes place. After cleavage, factors involved in the DNA damage response (such as ATM) and DNA repair (such as NHEJ proteins) are probably recruited into the recombination centre, and RAG might have an active role in this recruitment. DNA repair during V(D)J recombination is channelled into the NHEJ pathway by the RAG proteins[90,92,93], and RAG1 is able to interact with the NHEJ factors Ku70 (also known as XRCC6) and Ku80 (also known as XRCC5)[94] and with the DNA damage response factor MDC1 (mediator of DNA damage checkpoint protein 1) (G. Coster, D. Chen, D.G. Schatz and M. Goldberg, unpublished observations). An important open question is whether any DNA damage response or repair factors are recruited into recombination centres before DNA cleavage.

Steps leading to the formation of the paired complex. The locations of the recombination centres identified by ChIP analysis show no correlation with the presence of a particular type of RSS[85], as J_H and Jκ gene segments are flanked by 23RSSs, whereas Jβ and Jα gene segments are flanked by 12RSSs. It is therefore puzzling that an earlier study detected RAG-mediated nicks at 12RSSs but not 23RSSs in all antigen receptor loci examined, including *Igk*, leading to a '12RSS-first' model for paired complex formation during V(D)J recombination[95]. Why were nicks observed at Vκ but not Jκ gene segments when RAG binding was detected at Jκ but not Vκ gene segments? One potential explanation is that the ChIP analysis failed to detect binding to Vκ, perhaps because the RAG proteins are bound to only a small subset of the many (~140) Vκ gene segments in any one cell. This possibility cannot be ruled out, although it does not readily explain the absence of nicks at the highly accessible Jκ gene segments. In addition, as discussed previously[85], the 12RSS-first model cannot easily accommodate the RAG binding patterns observed at the *Igh* locus.

As an alternative, we propose a 'nonamer-first' RSS recognition model that invokes asymmetry in the events that occur during RAG recognition of 12RSSs and 23RSSs *in vivo*. In this speculative model (FIG. 4), the initial step in RSS recognition is the freeing of a loop of RSS DNA from the surface of a nucleosome by a chromatin remodelling complex[55], followed by binding of RAG1 (and RAG2 if present) and perhaps HMGB1 to the nonamer (FIG. 4b,c). The nonamer is the most likely site of initial DNA binding by RAG for two reasons. First, the nonamer serves as the stable 'anchor' for RAG on the RSS[6], and second, the synapsis of two RSSs in solution requires RAG1, HMGB1 and the nonamer but not RAG2 or the heptamer[96]. In the case of a 12RSS, nonamer binding is followed by heptamer engagement and nicking. However, in the case of a 23RSS, in which the heptamer and nonamer are farther apart on the face of the nucleosome, RAG anchored on the nonamer might have greater difficulty engaging the heptamer, so nicking might not occur as efficiently. The model therefore predicts that in a recombination centre containing 23RSSs (for example, at *Igk*), the predominant presynaptic species is a complex in which only the 23RSS nonamer is bound by RAG (FIG. 4c). We propose that synapsis of this 23RSS nonamer-only complex with a 12RSS allows for rapid 12RSS nicking (FIG. 4d,e) and slower nicking of the 23RSS (FIG. 4f). Once both RSSs are nicked, hairpin formation occurs rapidly at both RSSs[97] (FIG. 4g), with the result that few nicked 23RSSs would accumulate within the cell.

The nonamer-first model predicts that if a 23RSS were completely freed from its inhibitory interactions with a nucleosome, then nicking should occur before synapsis and perhaps be detectable. Consistent with this, a recent study observed nicking at the Dβ1 23RSS and not at Jβ 12RSSs[98]. Given that germline transcription from PDβ1 initiates within the Dβ1 23RSS[99] and that Dβ1 resides in a long internucleosomal interval[27], it is plausible that the Dβ1 23RSS is fully accessible for RAG binding and is readily nicked. Overall, current evidence suggests that V(D)J recombination need not initiate with binding of RAG to a 12RSS.

Ectopic RAG binding and genome instability. The mouse and human genomes contain millions of cryptic RSS sequences, as well as many sequences capable of forming alternative (non-B form) DNA structures[5,100,101]. Current evidence indicates that meaningful levels of RAG-mediated cleavage occur at some of these DNA sequences and that this contributes to lymphoma-associated genome alterations, such as deletion of the *SCL–SIL* locus in T cell acute lymphoblastic leukaemia and translocations of B cell lymphoma 2 (*BCL2*) to the *IGH* locus in follicular lymphoma[1–3]. The finding that RAG2 binds to thousands of transcriptionally active sites outside of antigen receptor genes raises several questions concerning the specificity and fidelity of V(D)J recombination. To what extent is RAG1 also recruited to these sites? If such recruitment occurs, is it mediated primarily through interactions with RAG2 or with cryptic RSSs, or by other mechanisms (such as recognition of alternative DNA structures[86])? Do the RAG proteins interact with components of the transcriptional apparatus, and might the RAG2–H3K4me3 interaction influence histone modifications, chromatin structure or transcription? Most importantly, how often do the RAG proteins create nicks and double strand breaks at ectopic sites in the genome,

Homologous recombination
Genetic recombination that occurs between regions of DNA with long stretches of homology. This occurs with a low frequency in somatic cells and at a much higher frequency in germ cells.

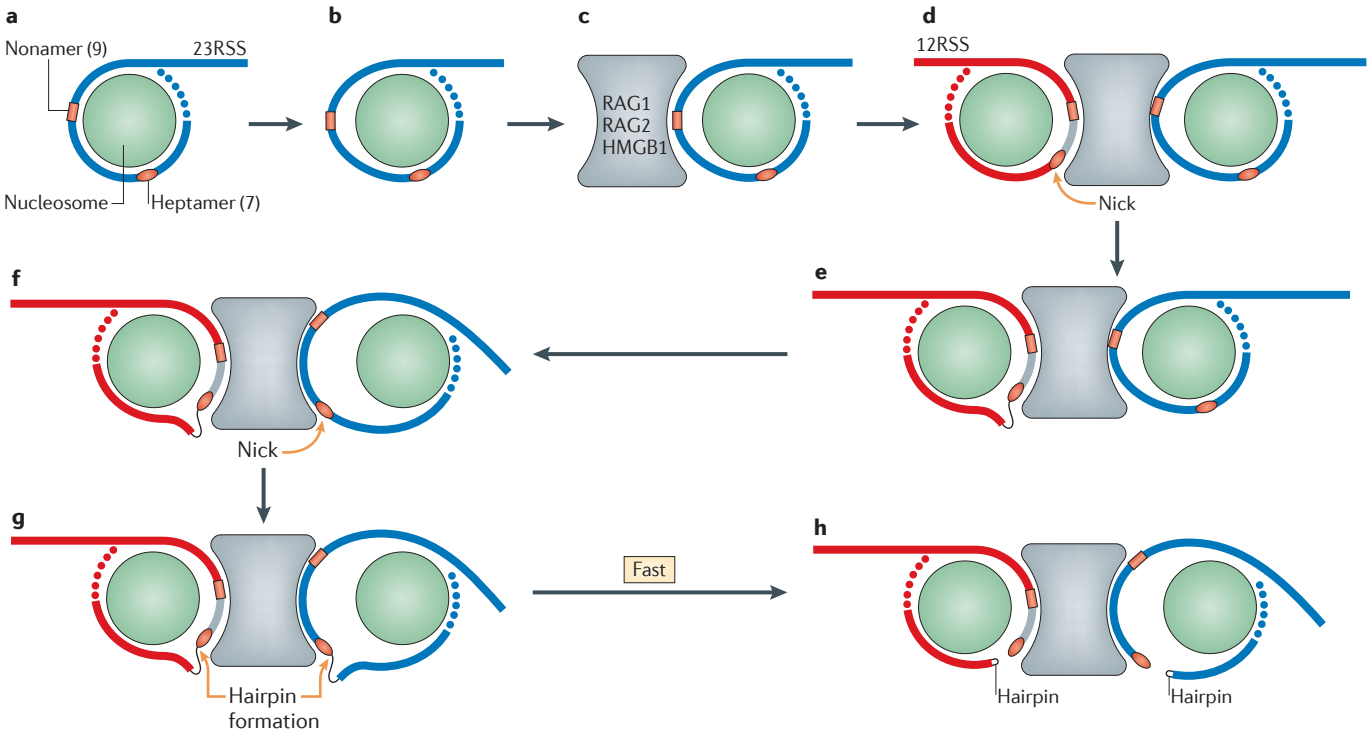


Figure 4 | **The nonamer-first model for RSS recognition. a** | A schematic diagram of a 23 recombination signal sequence (23RSS) wrapped around a nucleosome, with the nonamer and heptamer indicated. **b** | Chromatin remodelling creates a loop of DNA, which might contain the nonamer, free of the nucleosome surface. **c** | A RAG1–RAG2–HMGB1 complex binds the 23RSS nonamer but fails to engage the heptamer. **d** | The nonamer-only 23RSS complex synapses with the 12RSS, and engagement of the 12RSS heptamer allows for nicking. **e** | The nicked 12RSS resides in a relatively stable synaptic complex, but the heptamer of the 23RSS is not engaged. **f** | Eventually, the heptamer of the 23RSS is freed from the nucleosome surface and can be engaged by the RAG complex. Nicking of the 23RSS can now occur. **g** | The paired complex in which both RSSs are nicked rapidly undergoes hairpin formation at both RSSs. **h** | Hairpin formation yields the doubly cleaved product. This pathway is proposed to be important in situations where the recombination centre contains 23RSSs (for example, in the *Igk* locus). HMGB1, high-mobility group protein B1; RAG, recombination activating gene.

what mechanisms exist to suppress such ectopic DNA damage and under what circumstances do the RAG proteins circumvent these controls? A comprehensive cataloguing of sites of RAG binding and of RAG-mediated instability will begin to answer some of these questions. Interestingly, both RAG and activation-induced cytidine deaminase (AID), which is crucial for immunoglobulin gene somatic hypermutation and class-switch recombination, are recruited to active regions of chromatin[85,102] and to altered DNA structures[2,89,103]. Furthermore, it has been suggested that the combined action of RAG and AID can lead to DNA nicks and breaks and to genome instability[3]. Hence, a convergence of the mechanisms that recruit these two dangerous enzyme systems to chromatin might increase the risk of genome instability in B cells and help explain why the majority of human lymphoid malignancies originate from B cells[3].

Conclusions

V(D)J recombination is essential for the development of the adaptive immune system in humans and most other vertebrate species. The reaction is capable of generating an extraordinarily diverse array of antigen receptors, but this diversity comes at the cost of occasional genome instability and lymphoid malignancies. Recent advances have revealed important new aspects of how the RAG proteins are guided to appropriate target sites and have suggested new pathways that might explain mistargeting. These discoveries have revealed that complex processes related to nuclear dynamics and higher order chromatin architecture are layered on top of mechanisms that create gene segment accessibility so as to achieve lineage-specific and developmentally appropriate patterns of recombination. The RAG proteins are intimately involved in the regulation of V(D)J recombination through their ability to read a component of the histone code and to localize in focal recombination centres in antigen receptor loci. But some of these same properties of the RAG proteins might well contribute to their mistargeting to loci that do not encode antigen receptors and, ultimately, to the creation of ectopic DNA breaks and genome instability. Future experiments will probably reveal additional layers in the regulation of V(D)J recombination, as well as new regulatory functions inherent in RAG1 and RAG2.

1. Mills, K. D., Ferguson, D. O. & Alt, F. W. The role of DNA breaks in genomic instability and tumorigenesis. *Immunol. Rev.* **194**, 77–95 (2003).
2. Lieber, M. R., Yu, K. F. & Raghavan, S. C. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. *DNA Repair* **5**, 1234–1245 (2006).
3. Tsai, A. G. *et al.* Human chromosomal translocations at CpG sites and a theoretical basis for their lineage and stage specificity. *Cell* **135**, 1130–1142 (2008). **This study proposed, and provided support for, the fascinating idea that RAG and AID collaborate to create DNA nicks and breaks at methylated CpG sequences.**
4. Yancopoulos, G. D. & Alt, F. W. Developmentally controlled and tissue-specific expression of unrearranged V_H gene segments. *Cell* **40**, 271–281 (1985). **This classic study provided the first evidence for, and proposed, the accessibility model for the control of V(D)J recombination.**
5. Lewis, S. M. The mechanism of V(D)J joining: lessons from molecular, immunological, and comparative analyses. *Adv. Immunol.* **56**, 27–150 (1994).
6. Swanson, P. C. The bounty of RAGs: recombination signal complexes and reaction outcomes. *Immunol. Rev.* **200**, 90–114 (2004).
7. Fugmann, S. D. & Schatz, D. G. Identification of basic residues in RAG2 critical for DNA binding by the RAG1–RAG2 complex. *Mol. Cell* **8**, 899–910 (2001).
8. Mundy, C. L., Patenge, N., Matthews, A. G. W. & Oettinger, M. A. Assembly of the RAG1/RAG2 synaptic complex. *Mol. Cell. Biol.* **22**, 69–77 (2002).
9. Jones, J. M. & Gellert, M. Ordered assembly of the V(D)J synaptic complex ensures accurate recombination. *EMBO J.* **21**, 4162–4171 (2002).
10. Gellert, M. V(D)J recombination: RAG proteins, repair factors, and regulation. *Annu. Rev. Biochem.* **71**, 101–132 (2002).
11. Cobb, R. M., Oestreich, K. J., Osipovich, O. A. & Oltz, E. M. Accessibility control of V(D)J recombination. *Adv. Immunol.* **91**, 45–109 (2006).
12. Hesslein, D. G. & Schatz, D. G. Factors and forces controlling V(D)J recombination. *Adv. Immunol.* **78**, 169–232 (2001).
13. Yancopoulos, G. D. & Alt, F. W. Regulation of the assembly and expression of variable-region genes. *Annu. Rev. Immunol.* **4**, 339–368 (1986).
14. Jung, D., Giallourakis, C., Mostoslavsky, R. & Alt, F. W. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu. Rev. Immunol.* **24**, 541–570 (2006).
15. Goldmit, M. & Bergman, Y. Monoallelic gene expression: a repertoire of recurrent themes. *Immunol. Rev.* **200**, 197–214 (2004).
16. Krangel, M. S. T cell development: better living through chromatin. *Nature Immunol.* **8**, 687–694 (2007).
17. Schlissel, M. S. Regulating antigen-receptor gene assembly. *Nature Rev. Immunol.* **3**, 890–899 (2003).
18. Stanhope-Baker, P., Hudson, K. M., Shaffer, A. L., Constantinescu, A. & Schlissel, M. S. Cell type-specific chromatin structure determines the targeting of V(D)J recombinase activity *in vitro*. *Cell* **85**, 887–897 (1996). **This classic study demonstrated that chromatin structure is a critical determinant of DNA cutting by RAG.**
19. Kwon, J., Imbalzano, A. N., Matthews, A. & Oettinger, M. A. Accessibility of nucleosomal DNA to V(D)J cleavage is modulated by RSS positioning and HMG1. *Mol. Cell* **2**, 829–839 (1998).
20. Golding, A., Chandler, S., Ballestar, E., Wolffe, A. P. & Schlissel, M. S. Nucleosome structure completely inhibits *in vitro* cleavage by the V(D)J recombinase. *EMBO J.* **18**, 3712–3723 (1999).
21. McBlane, F. & Boyes, J. Stimulation of V(D)J recombination by histone acetylation. *Curr. Biol.* **10**, 483–486 (2000).
22. Kwon, J., Morshead, K. B., Guyon, J. R., Kingston, R. E. & Oettinger, M. A. Histone acetylation and hSWI/SNF remodeling act in concert to stimulate V(D)J cleavage of nucleosomal DNA. *Mol. Cell* **6**, 1037–1048 (2000).
23. Patenge, N., Elkin, S. K. & Oettinger, M. A. ATP-dependent remodeling by SWI/SNF and ISWI proteins stimulates V(D)J cleavage of 5 S arrays. *J. Biol. Chem.* **279**, 35360–35367 (2004).
24. Nightingale, K. P. *et al.* Acetylation increases access of remodelling complexes to their nucleosome targets to enhance initiation of V(D)J recombination. *Nucl. Acids Res.* **35**, 6311–6321 (2007).
25. Du, H., Ishii, H., Pazin, M. J. & Sen, R. Activation of 12/23-RSS-dependent RAG cleavage by hSWI/SNF complex in the absence of transcription. *Mol. Cell* **31**, 641–649 (2008).
26. Baumann, M., Mamais, A., McBlane, F., Xiao, H. & Boyes, J. Regulation of V(D)J recombination by nucleosome positioning at recombination signal sequences. *EMBO J.* **22**, 5197–5207 (2003).
27. Kondilis-Mangum, H. D. *et al.* Transcription-dependent mobilization of nucleosomes at accessible TCR gene segments *in vivo*. *J. Immunol.* **184**, 6970–6977 (2010).
28. Kaplan, N. *et al.* The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* **458**, 362–366 (2009).
29. Segal, E. & Widom, J. What controls nucleosome positions? *Trends Genet.* **25**, 335–343 (2009).
30. Zhang, Y. *et al.* Intrinsic histone–DNA interactions are not the major determinant of nucleosome positions *in vivo*. *Nature Struct. Mol. Biol.* **16**, 847–852 (2009).
31. Segal, E. & Widom, J. Poly(dA:dT) tracts: major determinants of nucleosome organization. *Curr. Opin. Struct. Biol.* **19**, 65–71 (2009).
32. Chowdhury, D. & Sen, R. Regulation of immunoglobulin heavy-chain gene rearrangements. *Immunol. Rev.* **200**, 182–196 (2004).
33. Krangel, M. S. Mechanics of T cell receptor gene rearrangement. *Curr. Opin. Immunol.* **21**, 133–139 (2009).
34. Jhunjhunwala, S., van Zelm, M. C., Peak, M. M. & Murre, C. Chromatin architecture and the generation of antigen receptor diversity. *Cell* **138**, 435–448 (2009).
35. Schlissel, M. S. Regulation of activation and recombination of the murine Igκ locus. *Immunol. Rev.* **200**, 215–223 (2004).
36. Abarrategui, I. & Krangel, M. S. Regulation of T cell receptor-α gene recombination by transcription. *Nature Immunol.* **7**, 1109–1115 (2006). **This study provided the first direct support for the long held idea that transcriptional elongation is important for creating accessibility for V(D)J recombination.**
37. Abarrategui, I. & Krangel, M. S. Noncoding transcription controls downstream promoters to regulate T-cell receptor α recombination. *EMBO J.* **26**, 4380–4390 (2007).
38. Sikes, M. L., Meade, A., Tripathi, R., Krangel, M. S. & Oltz, E. M. Regulation of V(D)J recombination: a dominant role for promoter positioning in gene segment accessibility. *Proc. Natl Acad. Sci. USA* **99**, 12309–12314 (2002).
39. Fernex, C., Capone, M. & Ferrier, P. The V(D)J recombinational and transcriptional activities of the immunoglobulin heavy-chain intronic enhancer can be mediated through distinct protein-binding sites in a transgenic substrate. *Mol. Cell. Biol.* **15**, 3217–3226 (1995).
40. Bolland, D. J. *et al.* Antisense intergenic transcription in V(D)J recombination. *Nature Immunol.* **5**, 630–637 (2004).
41. Bolland, D. J. *et al.* Antisense intergenic transcription precedes *Igh* D-to-J recombination and is controlled by the intronic enhancer Eμ. *Mol. Cell. Biol.* **27**, 5523–5533 (2007).
42. Chakraborty, T. *et al.* Repeat organization and epigenetic regulation of the D_H-Cμ domain of the immunoglobulin heavy-chain gene locus. *Mol. Cell* **27**, 842–850 (2007).
43. Osipovich, O. A., Subrahmanyam, R., Pierce, S., Sen, R. & Oltz, E. M. Cutting edge: SWI/SNF mediates antisense *Igh* transcription and locus-wide accessibility in B cell precursors. *J. Immunol.* **183**, 1509–1513 (2009).
44. Giallourakis, C. C. *et al.* Elements between the IgH variable (V) and diversity (D) clusters influence antisense transcription and lineage-specific V(D)J recombination. *Proc. Natl Acad. Sci. USA* **107**, 22207–22212 (2010).
45. Oettinger, M. A. How to keep V(D)J recombination under control. *Immunol. Rev.* **200**, 165–181 (2004).
46. Chakraborty, T. *et al.* A 220-nucleotide deletion of the intronic enhancer reveals an epigenetic hierarchy in immunoglobulin heavy chain locus activation. *J. Exp. Med.* **206**, 1019–1027 (2009).
47. Osipovich, O. *et al.* Targeted inhibition of V(D)J recombination by a histone methyltransferase. *Nature Immunol.* **5**, 309–316 (2004).
48. Morshead, K. B., Ciccone, D. N., Taverna, S. D., Allis, C. D. & Oettinger, M. A. Antigen receptor loci poised for V(D)J rearrangement are broadly associated with BRG1 and flanked by peaks of histone H3 dimethylated at lysine 4. *Proc. Natl Acad. Sci. USA* **100**, 11577–11582 (2003).
49. Johnson, K. *et al.* B cell-specific loss of histone 3 lysine 9 methylation in the V_H locus depends on Pax5. *Nature Immunol.* **5**, 853–861 (2004).
50. Smale, S. T. The establishment and maintenance of lymphocyte identity through gene silencing. *Nature Immunol.* **4**, 607–615 (2003).
51. Workman, J. L. Nucleosome displacement in transcription. *Genes Dev.* **20**, 2009–2017 (2006).
52. Li, B., Carey, M. & Workman, J. L. The role of chromatin during transcription. *Cell* **128**, 707–719 (2007).
53. Barski, A. *et al.* High-resolution profiling of histone methylations in the human genome. *Cell* **129**, 823–837 (2007).
54. Shogren-Knaak, M. *et al.* Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* **311**, 844–847 (2006).
55. Zhang, Y. *et al.* DNA translocation and loop formation mechanism of chromatin remodeling by SWI/SNF and RSC. *Mol. Cell* **24**, 559–568 (2006).
56. Wang, X. *et al.* Regulation of *Tcrb* recombination ordering by c-Fos-dependent RAG deposition. *Nature Immunol.* **9**, 794–801 (2008).
57. Zhang, Z. X. *et al.* Transcription factor Pax5 (BSAP) transactivates the RAG-mediated V_H-to-DJ_H rearrangement of immunoglobulin genes. *Nature Immunol.* **7**, 616–624 (2006).
58. Hesslein, D. G. *et al.* Pax5 is required for recombination of transcribed, acetylated, 5′ IgH V gene segments. *Genes Dev.* **17**, 37–42 (2003).
59. Jackson, A., Kondilis, H. D., Khor, B., Sleckman, B. P. & Krangel, M. S. Regulation of T cell receptor β allelic exclusion at a level beyond accessibility. *Nature Immunol.* **6**, 189–197 (2005).
60. Kosak, S. T. & Groudine, M. Form follows function: the genomic organization of cellular differentiation. *Genes Dev.* **18**, 1371–1384 (2004).
61. Reddy, K. L., Zullo, J. M., Bertolino, E. & Singh, H. Transcriptional repression mediated by repositioning of genes to the nuclear lamina. *Nature* **452**, 243–247 (2008).
62. Kosak, S. T. *et al.* Subnuclear compartmentalization of immunoglobulin loci during lymphocyte development. *Science* **296**, 158–162 (2002). **This paper provided the first evidence that V(D)J recombination is associated with contraction and movement away from the nuclear periphery of the recombining locus.**
63. Skok, J. A. *et al.* Reversible contraction by looping of the *Tcra* and *Tcrb* loci in rearranging thymocytes. *Nature Immunol.* **8**, 378–387 (2007).
64. Hewitt, S. L. *et al.* RAG-1 and ATM coordinate monoallelic recombination and nuclear positioning of immunoglobulin loci. *Nature Immunol.* **10**, 655–664 (2009).
65. Goldmit, M. *et al.* Epigenetic ontogeny of the *Igk* locus during B cell development. *Nature Immunol.* **6**, 198–203 (2005).
66. Schlimgen, R. J., Reddy, K. L., Singh, H. & Krangel, M. S. Initiation of allelic exclusion by stochastic interaction of *Tcrb* alleles with repressive nuclear compartments. *Nature Immunol.* **9**, 802–809 (2008).
67. Hewitt, S. L. *et al.* Association between the *Igk* and *Igh* immunoglobulin loci mediated by the 3′ *Igk* enhancer induces 'decontraction' of the *Igh* locus in pre-B cells. *Nature Immunol.* **9**, 396–404 (2008).
68. Roldan, E. *et al.* Locus 'decontraction' and centromeric recruitment contribute to allelic exclusion of the immunoglobulin heavy-chain gene. *Nature Immunol.* **6**, 31–41 (2005).
69. Jia, J., Kondo, M. & Zhuang, Y. Germline transcription from T-cell receptor Vβ gene is uncoupled from allelic exclusion. *EMBO J.* **26**, 2387–2399 (2007).
70. Singh, N., Bergman, Y., Cedar, H. & Chess, A. Biallelic germline transcription at the κ immunoglobulin locus. *J. Exp. Med.* **197**, 743–750 (2003).
71. Amin, R. H. *et al.* Biallelic, ubiquitous transcription from the distal germline Igκ locus promoter during B cell development. *Proc. Natl Acad. Sci. USA* **106**, 522–527 (2009).

72. Fitzsimmons, S. P., Bernstein, R. M., Max, E. E., Skok, J. A. & Shapiro, M. A. Dynamic changes in accessibility, nuclear positioning, recombination, and transcription at the Igκ locus. *J. Immunol.* **179**, 5264–5273 (2007).
73. Sayegh, C., Jhunjhunwala, S., Riblet, R. & Murre, C. Visualization of looping involving the immunoglobulin heavy-chain locus in developing B cells. *Genes Dev.* **19**, 322–327 (2005).
74. Fuxa, M. *et al.* Pax5 induces V-to-DJ rearrangements and locus contraction of the immunoglobulin heavy-chain gene. *Genes Dev.* **18**, 411–422 (2004).
75. Jhunjhunwala, S. *et al.* The 3D structure of the immunoglobulin heavy-chain locus: implications for long-range genomic interactions. *Cell* **133**, 265–279 (2008).
This landmark study provided the highest resolution picture to date of the higher order chromatin architecture and dynamics of an antigen receptor locus.
76. Shih, H. Y. & Krangel, M. S. Distinct contracted conformations of the *Tcra*/*Tcrd* locus during *Tcra* and *Tcrd* recombination. *J. Exp. Med.* **207**, 1835–1841 (2010).
77. Callebaut, I. & Mornon, J. P. The V(D)J recombination activating protein RAG2 consists of a six-bladed propeller and a PHD fingerlike domain, as revealed by sequence analysis. *Cell. Mol. Life Sci.* **54**, 880–891 (1998).
78. Elkin, S. K. *et al.* A PHD finger motif in the C terminus of RAG2 modulates recombination activity. *J. Biol. Chem.* **280**, 28701–28710 (2005).
79. West, K. L. *et al.* A direct interaction between the RAG2 C terminus and the core histones is required for efficient V(D)J recombination. *Immunity* **23**, 203–212 (2005).
80. Matthews, A. G. *et al.* RAG2 PHD finger couples histone H3 lysine 4 trimethylation with V(D)J recombination. *Nature* **450**, 1106–1110 (2007).
This study provided a high resolution structure for the PHD finger of RAG2 bound to H3K4me3, as well as evidence for the biological importance of this interaction.
81. 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* **27**, 561–571 (2007).
This study demonstrated that the RAG2 PHD finger binds H3K4me3 and provided evidence for the biological importance of this interaction.
82. Gomez, C. A. *et al.* Mutations in conserved regions of the predicted RAG2 kelch repeats block initiation of V(D)J recombination and result in primary immunodeficiencies. *Mol. Cell. Biol.* **20**, 5653–5664 (2000).
83. Shimazaki, N., Tsai, A. G. & Lieber, M. R. H3K4me3 stimulates the V(D)J RAG complex for both nicking and hairpinning in *trans* in addition to tethering in *cis*: implications for translocations. *Mol. Cell* **34**, 535–544 (2009).
84. Grundy, G. J., 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* **107**, 22487–22492 (2010).
85. Ji, Y. *et al.* The *in vivo* pattern of binding of RAG1 and RAG2 to antigen receptor loci. *Cell* **141**, 419–431 (2010).
This study demonstrated that the binding of RAG1 and RAG2 to antigen receptor loci is developmentally regulated, lineage restricted and focused on small, highly active regions of chromatin referred to as recombination centres.
86. Zhao, S., Gwyn, L. M., De, P. & Rodgers, K. K. A non-sequence-specific DNA binding mode of RAG1 is inhibited by RAG2. *J. Mol. Biol.* **387**, 744–758 (2009).
87. Ji, Y. *et al.* Promoters, enhancers, and transcription target RAG1 binding during V(D)J recombination. *J. Exp. Med.* **207**, 2809–2816 (2010).
88. Grazini, U. *et al.* The RING domain of RAG1 ubiquitylates histone H3: a novel activity in chromatin-mediated regulation of V(D)J joining. *Mol. Cell* **37**, 282–293 (2010).
89. Raghavan, S. C., Swanson, P. C., Wu, X., Hsieh, C. L. & Lieber, M. R. A non-B-DNA structure at the *Bcl-2* major breakpoint region is cleaved by the RAG complex. *Nature* **428**, 88–93 (2004).
90. Lee, G. S., Neiditch, M. B., Salus, S. S. & Roth, D. B. RAG proteins shepherd double-strand breaks to a specific pathway, suppressing error-prone repair, but RAG nicking initiates homologous recombination. *Cell* **117**, 171–184 (2004).
91. Matthews, A. G. W. & Oettinger, M. A. RAG: a recombinase diversified. *Nature Immunol.* **10**, 817–821 (2009).
92. Corneo, B. *et al.* Rag mutations reveal robust alternative end joining. *Nature* **449**, 483–486 (2007).
93. 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* **104**, 17046–17051 (2007).
94. Raval, P., Kriatchko, A. N., Kumar, S. & Swanson, P. C. Evidence for Ku70/Ku80 association with full-length RAG1. *Nucl. Acids Res.* **36**, 2060–2072 (2008).
95. Curry, J. D., Geier, J. K. & Schlissel, M. S. Single-strand recombination signal sequence nicks *in vivo*: evidence for a capture model of synapsis. *Nature Immunol.* **6**, 1272–1279 (2005).
96. Yin, F. F. *et al.* Structure of the RAG1 nonamer binding domain with DNA reveals a dimer that mediates DNA synapsis. *Nature Struct. Mol. Biol.* **16**, 499–508 (2009).
97. Yu, K. F. & Lieber, M. R. Mechanistic basis for coding end sequence effects in the initiation of V(D)J recombination. *Mol. Cell. Biol.* **19**, 8094–8102 (1999).
98. Franchini, D. M., Benoukraf, T., Jaeger, S., Ferrier, P. & Payet-Bornet, D. Initiation of V(D)J recombination by Dβ-associated recombination signal sequences: a critical control point in TCRβ gene assembly. *PLoS ONE* **4**, e4575 (2009).
99. Sikes, M. L., Gomez, R. J., Song, J. & Oltz, E. M. A developmental stage-specific promoter directs germline transcription of DβJβ gene segments in precursor T lymphocytes. *J. Immunol.* **161**, 1399–1405 (1998).
100. Merelli, I. *et al.* RSSsite: a reference database and prediction tool for the identification of cryptic recombination signal sequences in human and murine genomes. *Nucl. Acids Res.* **38**, W262–W267 (2010).
101. Zhao, J., Bacolla, A., Wang, G. & Vasquez, K. M. Non-B DNA structure-induced genetic instability and evolution. *Cell. Mol. Life Sci.* **67**, 43–62 (2010).
102. Yamane, A. *et al.* Deep-sequencing identification of the genomic targets of the cytidine deaminase AID and its cofactor RPA in B lymphocytes. *Nature Immunol.* **12**, 62–69 (2010).
103. Staszewski, O. *et al.* Activation-induced cytidine deaminase induces reproducible DNA breaks at many non-Ig loci in activated B cells. *Mol. Cell* **41**, 232–242 (2011).
104. Leu, T. M. & Schatz, D. G. rag-1 and rag-2 are components of a high-molecular-weight complex, and association of rag-2 with this complex is rag-1 dependent. *Mol. Cell. Biol.* **15**, 5657–5670 (1995).
105. Lin, W.-C. & Desiderio, S. Cell cycle regulation of V(D)J recombination-activating protein RAG-2. *Proc. Natl Acad. Sci. USA* **91**, 2733–2737 (1994).
106. Lee, J. & Desiderio, S. Cyclin A/CDK2 regulates V(D)J recombination by coordinating RAG-2 accumulation and DNA repair. *Immunity* **11**, 771–781 (1999).
107. Fugmann, S. D., Lee, A. I., Shockett, P. E., Villey, I. J. & Schatz, D. G. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. *Annu. Rev. Immunol.* **18**, 495–527 (2000).
108. De, P. & Rodgers, K. K. Putting the pieces together: identification and characterization of structural domains in the V(D)J recombination protein RAG1. *Immunol. Rev.* **200**, 70–82 (2004).
109. Mathieu, N., Hempel, W. M., Spicuglia, S., Verthuy, C. & Ferrier, P. Chromatin remodeling by the T cell receptor (TCR)-β gene enhancer during early T cell development: implications for the control of TCR-β locus recombination. *J. Exp. Med.* **192**, 625–636 (2000).
110. Bouvier, G. *et al.* Deletion of the mouse T-cell receptor β gene enhancer blocks α-β T-cell development. *Proc. Natl Acad. Sci. USA* **93**, 7877–7881 (1996).
111. Bories, J. C., Demengeot, J., Davidson, L. & Alt, F. W. Gene-targeted deletion and replacement mutations of the T-cell receptor β-chain enhancer: the role of enhancer elements in controlling V(D)J recombination accessibility. *Proc. Natl Acad. Sci. USA* **93**, 7871–7876 (1996).
112. Oestreich, K. J. *et al.* Regulation of TCR β gene assembly by a promoter/enhancer holocomplex. *Immunity* **24**, 381–391 (2006).
113. Spicuglia, S. *et al.* Promoter activation by enhancer-dependent and -independent loading of activator and coactivator complexes. *Mol. Cell* **10**, 1479–1487 (2002).
114. Whitehurst, C. E., Chattopadhyay, S. & Chen, J. Control of V(D)J recombinational accessibility of the Dβ1 gene segment at the TCRβ locus by a germline promoter. *Immunity* **10**, 313–322 (1999).
115. Whitehurst, C. E., Schlissel, M. S. & Chen, J. Deletion of germline promoter PDβ1 from the TCRβ locus causes hypermethylation that impairs Dβ1 recombination by multiple mechanisms. *Immunity* **13**, 703–714 (2000).
116. McMillan, R. E. & Sikes, M. L. Differential activation of dual promoters alters Dβ2 germline transcription during thymocyte development. *J. Immunol.* **180**, 3218–3228 (2008).
117. McMillan, R. E. & Sikes, M. L. Promoter activity 5' of Dβ2 is coordinated by E47, Runx1, and GATA-3. *Mol. Immunol.* **46**, 3009–3017 (2009).
118. Osipovich, O. *et al.* Essential function for SWI-SNF chromatin-remodeling complexes in the promoter-directed assembly of *Tcrb* genes. *Nature Immunol.* **8**, 809–816 (2007).

Acknowledgements
The authors wish to thank E. Oltz, G. Teng, K. Shetty and J. Banerjee for comments on the manuscript. We apologize that not all of the relevant literature could be cited owing to space constraints.

Competing interests statement
The authors declare no competing financial interests.

FURTHER INFORMATION
David G. Schatz's homepage: http://medicine.yale.edu/immuno/people/david_schatz.profile

ALL LINKS ARE ACTIVE IN THE ONLINE PDF