OUTLOOK

Evolving adaptive immunity

Geoffrey A. Lovely and Ranjan Sen

Laboratory of Molecular Biology and Immunology, National Institute on Aging, Baltimore, Maryland 21224, USA

Generation of a diverse repertoire of antigen receptor specificities via DNA recombination underpins adaptive immunity. In this issue of *Genes & Development*, Carmona and colleagues (pp. 909–917) provide novel insights into the origin and function of recombination-activating gene 1 (RAG1) and RAG2, the lymphocyte-specific components of the recombinase involved in the process.

> "*It appears therefore that immunologically competent cells have evolved a pattern of somatic genetic behavior which is radically different from anything normally found in modern molecular genetics.*"
>
> —Dreyer and Bennett (1965).

We now know that the "radically different" behavior of immune cells involves assembling antigen receptor genes of B and T lymphocytes via a DNA cut-and-paste mechanism known as V(D)J recombination (Jung et al. 2006). For immunoglobulin (*Ig*) heavy chain and T-cell receptor (TCR) β-chain genes, this involves two somatic recombination events that put together variable (V), diversity (D), and joining (J) gene segments that are widely separated in the germline. *Ig* light chains and TCRα chain genes are assembled by one recombination event that juxtaposes V and J gene segments. A random assortment of recombining gene segments and the imprecise nature of V(D)J recombination ensure that millions of antigenic specificities are generated prior to antigen exposure.

Central to this process are the recombination-activating gene 1 (*RAG1*) and *RAG2*. RAG1 and RAG2 proteins initiate V(D)J recombination by introducing DNA double-strand breaks at specialized recombination signal sequences (RSSs) that flank antigen receptor gene segments. These breaks are processed and religated by ubiquitously expressed enzymes of the nonhomologous end-joining (NHEJ) pathway. RSSs have a very well-defined organization, consisting of a highly conserved heptanucleotide sequence and a somewhat less conserved nonamer sequence that are separated by either 12 or 23 nucleotides (referred to as 12RSS and 23RSS, respectively). RAG1/2-initiated DNA breaks require recognition and synapsis between a 12RSS and 23RSS but not between 12/12RSSs or 23/23RSSs. This feature of asymmetry is referred to as the 12/23 rule. An interesting consequence is that gene segments with the same RSS configuration can undergo evolutionary amplification without adversely affecting recombination. Thus, the hundreds of V gene segments present in the genome (each with either a 12RSS or 23RSS) will only participate in functional V-to-D (or V-to-J) rearrangements but not in nonproductive V-to-V rearrangements. The mechanistic basis of the 12/23 rule and its origin are therefore of fundamental importance to adaptive immunity. Another enigmatic aspect of V(D)J recombination reaction is that RSS recognition and catalytic residues that cleave DNA are both located in RAG1. However, both RAG1 and RAG2 are required to assemble antigen receptor genes. The molecular mechanisms by which RAG2 "reveals" RAG1 functionality are not known.

A critical insight as to the origins of the V(D)J recombinase came from the hypothesis that the RAG1/RAG2 combination is a "disassembled transposon" (Thompson 1995). Soon thereafter, it was shown that vertebrate RAG1/RAG2 could carry out transposition reactions in vitro (Agrawal et al. 1998; Hiom et al. 1998). This idea gained further traction with the identification of a RAG1-like domain in a transposase used by the *Transib* transposon (Kapitonov and Jurka 2005). Remarkably, the DNA sequence recognized by the *Transib* transposase is very similar to the highly conserved heptamer sequence (CACAGTC) of vertebrate RSSs, suggesting that the *Transib* transposase and RAG1/RAG2 share target specificity. In other words, target specificity of these proteins has not changed substantially over 1000 million years of evolution (Fig. 1). Once a functional transposon that worked with a RAG1-like transposase had been identified, the next question was whether RAG2 was ever a part of a functional transposon or whether RAG2 is a relatively new non-transposon-associated component coopted by the immune system.

The first *RAG1*-like (*RAG1L*) and *RAG2*-like (*RAG2L*) genes outside the immune system were identified in the purple sea urchin *Strongylocentrotus purpuratus* (Fugmann et al. 2006). While the genomic organization of *spRAG1L* and *spRAG2L* (*RAG1L* genes in *S. purpuratus*)

[*Keywords*: RAG; Transib; V(D)J recombination; evolution; transposition]
Corresponding author: senra@mail.nih.gov
Article is online at http://www.genesdev.org/cgi/doi/10.1101/gad.281014.116.

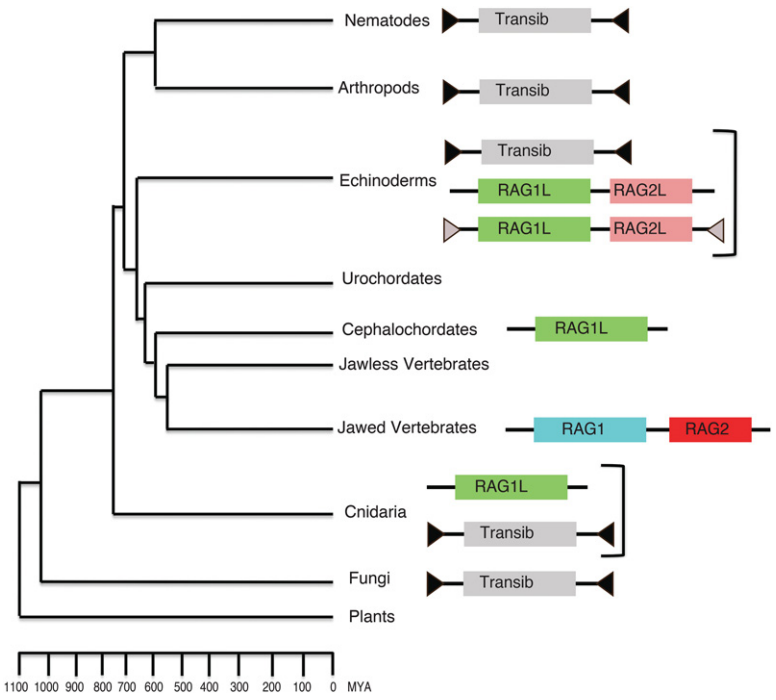

Figure 1. A phylogenic tree of the *Transib* superfamily adapted from Kapitonov and Jurka (2005). The presence of terminal inverted repeats (TIRs; black triangles) indicates a functional transposon that is free to move, while gray triangles indicate apparent TIRs. *Transib* (gray rectangle), *RAG1*-like (RAG1L; green rectangle), *RAG2*-like (RAG2L; pink rectangle), *RAG1* (blue rectangle), and *RAG2* (red rectangle) genes or transposons are shown next to the phylum in which they were discovered on the phylogenic tree. A scale in millions of years (MYA) is shown *below* the tree. The diagram depicts the presence of paired *RAG1L* and *RAG2L* genes in green and purple sea urchins (*middle* line in Echinoderms), paired *RAG1L* and *RAG2L* genes with apparent TIRs in *Patiria minata* (*bottom* line in Echinoderms), and *RAG1L* alone in several other lineages.

was very similar to the organization of vertebrate *RAG1* and *RAG2*, there was no evidence that the two genes were part of a functional transposon (Fig. 1). Similarly organized *RAG1L* and *RAG2L* genes were later found in the green sea urchin *Lytechinus variegates*, again with no indication of a transposon origin (Kapitonov and Koonin 2015). However, recent analyses of the sea star (*Patiria minata*) genome revealed *RAG1L* and *RAG2L* genes that appear to be flanked by terminal inverted repeats (TIRs) (Kapitonov and Koonin 2015) that are considered to be the "smoking gun" for a functional transposon (Fig. 1). These observations hinted that a *Transib* transposon containing only a *RAG1L* gene had incorporated a *RAG2L* gene while still a functional transposon. Thereafter, in some lineages, the transposon-derived gene pair became a part of the genome, while, in other lineages, the gene pair was lost. The former likely served as the precursor to the vertebrate V(D)J recombinase.

In this issue of *Genes & Development*, Carmona et al. (2016) provide insights on several outstanding questions regarding the origin and function of RAG1 and RAG2. First, they demonstrate that the vertebrate RAG1 by itself can carry out the complete V(D)J recombination reaction, albeit at much lower efficiency compared with the combination of RAG1 and RAG2. These observations provide direct experimental evidence that RAG1 can function alone, as would have been predicted from the occurrence of a RAG1L protein in the *Transib* transposon. Interestingly, recombination mediated by RAG1 alone does not follow the 12/23 rule, as reflected in comparable levels of recombination between substrates carrying two 12RSSs or a 12RSS and 23RSS. Substrates with only two 23RSSs did not recombine detectably, perhaps reflecting lower affinity of RAG1 for 23RSS. Thus, RAG2 not only increased the overall reaction efficiency of RAG1 for both 12/12 and 12/23 substrates but also enforced the 12/23 rule by introducing an ~30-fold preference for asymmetric RSSs. Recent structural studies identify conformational changes in RAG1/2 bound to 12RSS or 23RSS (Ru et al. 2015). It remains to be determined whether and how RAG2 contributes to these changes.

Having detected functional activity of vertebrate RAG1 alone, Carmona et al. (2016) were prompted to test the distantly related RAG1L proteins from *Transib* (Hztransib) and *S. purpuratus* (spRAG1L). Neither of these proteins alone activated 12/23-dependent recombination of substrates integrated into fibroblasts. However, cotransfection of vertebrate *RAG2* with either *RAG1L* gene produced detectable levels of 12/23-dependent recombination. Moreover, this activity was abolished by mutating residues in the *RAG1L* genes predicted to catalyze DNA strand scission. These observations are interesting because they reveal (1) enzymatic activity of RAG1L proteins from these organisms, (2) the ability of RAG1L proteins to recognize and work with vertebrate RSSs, and (3) that these proteins can be functionally assisted by vertebrate RAG2. Carmona et al. (2016) took these studies a step further by assaying in vitro transposition potential of Hztransib in the presence or absence of vertebrate RAG2. They found that Hztransib alone transposed 12RSS-containing donors but not 23RSS-containing donors. Transposition using 12/23RSS or 23RSS substrates required coincubation with vertebrate RAG2, leading the investigators to propose that acquisition of a *RAG2L* gene may have permitted an ancestral *Transib* element to accommodate 23RSS recognition. Since asymmetry of the V(D)J recombination reaction is an essential ingredient of adaptive immunity (as discussed above), these

observations illuminate the origin and function of each component of the V(D)J recombinase.

References

Agrawal A, Eastman QM, Schatz DG. 1998. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature* **394:** 744–751.

Carmona LM, Fugmann SD, Schatz DG. 2016. Collaboration of RAG2 with RAG1-like proteins during the evolution of V(D)J recombination. *Genes Dev* (this issue). doi: 10.1101/gad.278432.116.

Dreyer WJ, Bennett JC. 1965. The molecular basis of antibody formation: a paradox. *Proc Natl Acad Sci* **54:** 864–869.

Fugmann SD, Messier C, Novack LA, Cameron RA, Rast JP. 2006. An ancient evolutionary origin of the Rag1/2 gene locus. *Proc Natl Acad Sci* **103:** 3728–3733.

Hiom K, Melek M, Gellert M. 1998. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell* **94:** 463–470.

Jung D, Giallourakis C, Mostoslavsky R, Alt FW. 2006. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol* **24:** 541–570.

Kapitonov VV, Jurka J. 2005. RAG1 core and V(D)J recombination signals sequences were derived from *Transib* transposons. *PLoS Biol* **3:** e181.

Kapitonov VV, Koonin EV. 2015. Evolution of the RAG1–RAG2 locus: both proteins came from the same transposon. *Biol Direct* **10:** 20.

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

Thompson CB. 1995. New insights into V(D)J recombination and its role in the evolution of the immune system. *Immunity* **3:** 531–539.