

Adapting to a changing world: RAG genomics and evolution

Maristela Martins de Camargo¹ and Laila Alves Nahum^{2*}

¹Department of Immunology, Institute of Biomedical Sciences, University of São Paulo, São Paulo, SP 05508-900, Brazil

²Department of Biological Science, Louisiana State University, Baton Rouge, LA 70803, USA

* Correspondence to: Tel: +11 225 578 8798; Fax: +11 225 578 2597; E-mail: lnahum@lsu.edu

Date received (in revised form): 21th March 2005

Abstract

The origin of the recombination-activating genes (RAGs) is considered to be a foundation hallmark for adaptive immunity, characterised by the presence of antigen receptor genes that provide the ability to recognise and respond to specific peptide antigens. In vertebrates, a diverse repertoire of antigen-specific receptors, T cell receptors and immunoglobulins is generated by V(D)J recombination performed by the RAG-1 and RAG-2 protein complex. RAG homologues were identified in many jawed vertebrates. Despite their crucial importance, no homologues have been found in jawless vertebrates and invertebrates. This paper focuses on the RAG homologues in humans and other vertebrates for which the genome is completely sequenced, and also discusses the main contribution of the use of RAG homologues in phylogenetics and vertebrate evolution. Since mutations in both genes cause a spectrum of severe combined immunodeficiencies, including the Omenn syndrome (OS), these topics are discussed in detail. Finally, the relevance to genomic diversity and implications to immunomics are addressed. The search for homologues could enlighten us about the evolutionary processes that shaped the adaptive immune system. Understanding the diversity of the adaptive immune system is crucially important for the design and development of new therapies to modulate the immune responses in humans and/or animal models.

Keywords: adaptive immune system, recombination-activating genes, human genome, animal models, vertebrate phylogeny, evolutionary genomics

Introduction

When the adaptive immune system arose, approximately 500 million years ago, it provided ancient vertebrates with the ability to recognise sets of amino acids and respond specifically to them. Before the development of the adaptive system, organisms were fated to dealing with environmental molecules in a much simpler and less specific manner; that is, by recognising structural patterns and responding to them in a pre-scripted manner.^{1,2} The inauguration of the adaptive immune system allowed the ancestors of jawed vertebrates to mount sophisticated responses to a varied panel of molecules — from environmental particles, such as allergens, to intracellular parasites and from viruses to intestinal helminthes. Moreover, the adaptive immune system provided these animals with immunological memory. After mounting an effective response against a certain microorganism, the immune system retains the ability rapidly to recognise and neutralise the same microorganism at the onset of a second encounter. This ability to retain memory of previously encountered peptides is the basis for vaccination.

The ability to identify a set of amino acids depends on the presence of antigen-specific receptors. Antigen-specific receptors comprise T cell receptors (TCRs) and

immunoglobulins (Igs or antibodies). An organism's antigen-specific receptors comprise a panel of receptors bearing different specificities being clonally expressed. The higher the variability of the receptors that an organism is capable of expressing, the larger the number of different sets of amino acids that its cells will be able to recognise.

The antigen-specific receptor binding sites are the products of the germline recombination of gene segments named V (variable), D (diversity) and J (joining) within the precursors of T and B cells. In humans, these segments are organised in a translocon configuration — that is, a large number of gene segments are grouped together.³ Several V gene segments lie upstream of grouped D gene segments, which are localised upstream of J gene segments. Recombination of one segment from each group will result in an antigen-specific receptor. The combination of these gene segments is known as V(D)J recombination, and results in a variety of receptors that recognise different peptides, providing the system with a repertoire of possible responses.³ Each gene segment is flanked by elements known as recombination signal sequences, which are conserved heptamer and nonamer sequences, separated by a 12- or 23-base pair spacer.⁴ When two segments are brought together for joining, further combinatorial variation might occur by addition

of nucleotides between the fragments by the terminal deoxynucleotidyl transferase, a template-independent polymerase.⁵

The V(D)J recombination is performed by a complex containing the product of the recombination-activating genes (RAGs).^{6,7} *RAG-1* and *RAG-2* were first discovered when these authors attempted to identify the 'recombinase' responsible for V(D)J recombination.⁸ The origin of *RAG* is considered to be a foundation hallmark for adaptive immunity, and *RAG* homologues were identified in many jawed vertebrates. Nevertheless, no *RAG* homologues have been found to date in jawless vertebrates and invertebrates. The sudden acquisition of *RAGs*, which allowed the immune system to create and deal with diversity, has been described by some as the immunological 'big bang'.^{9,10}

In the past few decades, a lot of work has been carried out to elucidate the mechanisms coordinating the adaptive immune system; however, some questions remain unanswered, including the origin of *RAG* proteins, which ultimately gave rise to the adaptive immune system. A series of excellent reviews on several aspects of *RAG* biology and V(D)J recombination was recently published.^{1,11} The present review is not the forum for discussing these aspects; instead, it will focus on recent studies on *RAG* genes and proteins in humans and other vertebrates for which the genome is completely sequenced, and will discuss these in the context of evolutionary genomics, addressing the role of current and future genome projects in understanding the evolution and diversity of the adaptive immune system.

Functional annotation of *RAG* genes and proteins

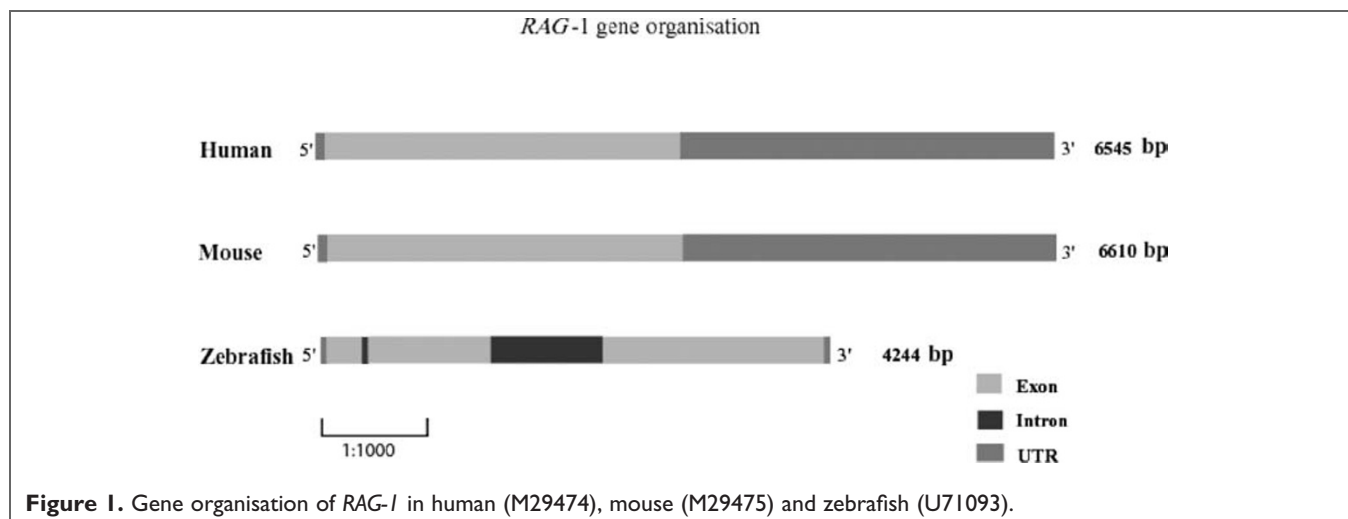
RAGs have been extensively studied over the past few decades.^{1,11,12} The *RAG-1* and *RAG-2* genes are mapped in

the same chromosome, separated by only a few kilobases and placed in opposite orientation to each other.⁷ In all mammals, birds and amphibians characterised so far, they have a compact organisation, with no introns; however, interrupted genes have been described in most fish for which *RAGs* have been identified.^{13,14} Figure 1 illustrates these aspects in *RAG-1* genes in human and two animal models.^{6,15}

RAG proteins are functionally annotated as 'recombination-activating protein', 'recombinase-activating protein', 'recombinase' or even 'V(D)J recombination-activating protein'. They are conserved in size and sequence. *RAG-1* is over 1,000 amino acids in length (1,040 amino acids in the mouse, 1,057 amino acids in zebrafish) and *RAG-2* is over 500 amino acids in length (520 amino acids in frog, 528 amino acids in chicken).

Many studies on *RAG* proteins are conducted in the so-called core proteins, mainly because of the difficulties encountered during protein purification. These core proteins correspond to truncated portions of the *RAG* proteins, and have catalytic activity.¹⁶ In humans, the core proteins correspond, approximately, to 63 per cent and 73 per cent of *RAG-1* (residues 384–1040) and *RAG-2* (1–387), respectively. Comparisons between full-length and core proteins have revealed important aspects of their function and regulation.^{17,18} Non-core portions of the *RAG* proteins seem to play important roles in V(D)J recombination that do not influence the specificity of the recombinase site.^{17,18}

Catalytic residues were identified in the *RAG* proteins using different approaches, including sequence analyses, site-directed mutagenesis and functional characterisation of mutant proteins, among others.^{19–21} The *RAG-1* protein contains a DDE motif (D600, D708 and E962), which is critical for both DNA cleavage and hairpin formation.^{19,20,22} Proteins containing the DDE motif are members of the retroviral integrase superfamily, found in bacterial



transposases and retroviral integrases. Therefore, RAG proteins, bacterial transposases and retroviral integrases are functionally related, despite the low sequence similarity shared by them.

Most RAG genes and proteins currently listed in the databases have not been genetically or biochemically characterised, but have been assigned a putative function based on standard sequence similarity methods. This common practice in computational functional genomics seems to have worked reasonably well in the case of the RAG datasets, mainly because their high level of sequence similarity, combined with full coverage of the alignments and the absence of paralogues (originated by gene duplication). Multiple copies of the RAG genes and proteins might exist but data remain unpublished.¹¹

The great majority of RAG genes and proteins found in databases corresponds to partial sequences. In order to find out more about the conserved motifs, critical residues, diversity and potential regulation of RAG-1 and RAG-2, it will be necessary to increase the number of full-length genes and proteins in the databases.

Vertebrate homologues, phylogenetics and evolution

The complete or near-complete sequences of the nuclear genomes of several vertebrates are available in public databases, and many more are underway (<http://www.ensembl.org>, <http://www.genomesonline.org>). This extraordinary amount of data creates a platform for comparative analysis and contributes to the understanding of complex biological systems.

RAG-1 and RAG-2 have been identified in a wide range of jawed vertebrates, including human, chimpanzee, mouse, rat, dog, chicken, zebrafish, fugu and tetraodon (<http://www.ensembl.org>, <http://www.genomesonline.org>). RAG homologues show high levels of sequence similarity across organisms — for example, 90 per cent between human and mouse.⁶ Nevertheless, no RAG homologues have been found to date in jawless vertebrates and invertebrates.

RAG genes have proved to be excellent targets for phylogenetic reconstructions and, in fact, have been extensively applied in studies in a wide range of organisms.^{23–25} Some of the main reasons to choose RAG genes are that: (i) they are composed by long sequences of DNA and protein (providing more sites for statistical analysis); (ii) they show high levels of sequence conservation (helping primer design and use in samples from diverse taxonomic groups); and (iii) there is an absence of introns, as discussed earlier. Moreover, nuclear-encoded genes are under the rules of a different evolutionary context compared with mitochondrial genes and gene products. Like mitochondrial genes, RAG-1 and RAG-2 can be (at least for now) considered as ‘perfect’

orthologues, which adds another advantage of using them as molecular markers in phylogenetic studies.

All together, these features have promoted RAG genes as suitable markers for phylogenetic reconstruction. Furthermore, RAGs were selected as targets in the Tree of Life project, a huge initiative that claims to decode the relationships between all living organisms.²⁶ In the search for the Tree of Life, RAG-1 and RAG-2 have been used to study phylogenetic relationships in deep nodes.^{23,24}

Figure 2 shows the proposed phylogeny of the major animal lineages, based on molecular data and fossil records,^{27,28} and indicates the occurrence of the RAG genes and proteins in Mammalia, Reptilia (including birds), Amphibia and Actinopterygii (fish), which correspond to the jawed vertebrates.

No significant sequence similarity has been detected between RAG-1 and RAG-2 that could suggest a possible event of gene duplication of a common ancestral RAG gene.^{3,29}

RAG mutation and immunodeficiencies

Experimental scientists use loss-of-function point mutations widely to assess the function of genes, as well as to evaluate their permissiveness to selective pressures. In humans, RAG-1 and RAG-2 are mapped to chromosome 11p13-p12, and extensively studied spontaneous mutations in both genes cause a spectrum of severe combined immunodeficiencies.

Loss-of-function mutations of RAG-1 and RAG-2 lead to disruption of the V(D)J recombination, disturbing T- and B cell development and causing severe combined immune deficiency (SCID). A patient with SCID does not have mature functional T- and B cells (T-B-SCID), disabling this patient's ability to recognise and respond to specific antigens and to develop immunological memory after infections.^{30,31} Mutations of either one of the RAG genes are associated with another immune deficiency, Omenn Syndrome (OS), which is an autosomal dominant disease with variable expressivity and two causative loci. It is frequently lethal in the homozygous condition. OS patients presenting with partial loss of function of RAG lack circulating B cells, but show a small and highly activated population of circulating T cells. These patients also have lymphadenopathy, hepatomegaly, splenomegaly and erythroderma.^{31,32} It is not clear what other factors are involved in the development of one or other immune deficiencies, since they both require mutations of the RAG-1 and/or RAG-2 genes.³⁰

A study conducted in 45 patients with SCID or OS analysed RAG mutations.³¹ Interestingly, the analysed T-B-SCID patients had nonsense, frameshift or null mutations resulting in biochemical changes that abrogated the recombination ability of RAG-1/RAG-2. In patients with

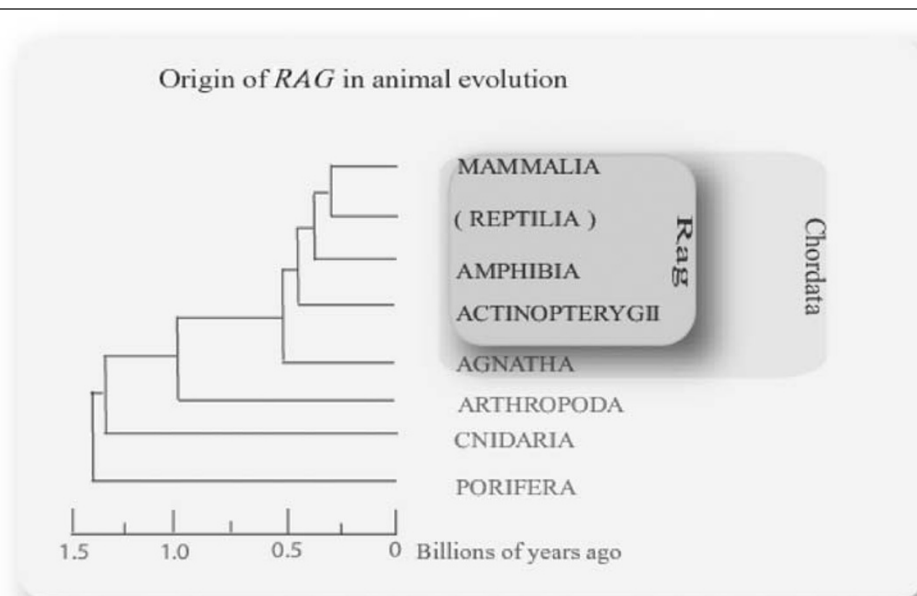


Figure 2. Occurrence of RAG homologues across major animal groups. Reptilia includes birds. Tree topology is based on the phylogenetic analysis²⁷ and time-scale is based on fossil records²⁸ published elsewhere.

OS, the authors found a predominance of missense mutations that allowed partial activity of *RAG-1/RAG-2*.^{33,34} One of these studies has also identified three homozygous missense mutations in *RAG-1* in one patient and in *RAG-2* in three patients, located outside the *RAG* core regions.³¹ Taken together with recent results from group, another which showed that removal of regions outside the core impaired the processing of recombination substrates,¹⁷ these results point to the possibility that the core region alone accounts for the whole recombination activity.

Up until now, there have been no spontaneous mutations of *RAG-1/RAG-2* described in animals other than humans.³⁵ The closest phylogenetic model available for studies on T-B-SCID is the *RAG-1/RAG-2* knockout mouse.^{36,37} These mice have their *RAG-1/RAG-2* genes disrupted and do not have circulating T- and B cells. More recently, mutations of *RAG-1* in zebrafish have become a second model of choice.³⁸

Nevertheless, other forms of SCID have been described in other animals, but all of these have been due to defects in proteins downstream of *RAG-1/RAG-2* in the V(D)J recombination cascade.

SCID in foals is limited to the Arabian breed and is an autosomal recessive disorder due to an absence of DNA-dependent protein kinases (DNA-PKcs) in the cells of these animals.^{39,40} In addition, spontaneous SCID due to DNA-PKcs defects was described in C.B-17 mice of BALB/c origin⁴¹ and in Jack Russell terrier dogs.⁴²

A third class of spontaneous SCID is found in dogs from the Basset and Cardigan Welsh Corgi breeds and in humans.^{43–45} This form of SCID is linked to the X chromosome and is thus called XSCID or SCID-X1, and involves a mutation on the gene that encodes the common chain of interleukin-2 (IL-2) receptor (also named γ chain or γc). As this chain is common to receptors of several interleukins involved in T- and B cell maturation (IL-2, -4, -7, -9, -15 and -21), a defect on this chain disrupts the signalling pathway downstream of these receptors, preventing cell maturation.⁴⁴

Regardless of whether their defect affects *RAG* genes or downstream components of the V(D)J recombination cascade, these animal models teach us how crucial the adaptive immune system — or, in other words, the acquisition of functional *RAG* genes — is to the survival of complex organisms. The world surrounding us poses constant challenges that must be identified and neutralised, not to mention our own internal challenges, such as the development of cancers.

Relevance and perspectives

In order to have a better understanding of the function, interaction, regulation and evolution of *RAG* genes and proteins, it is very important to access their full-length sequences and protein structures. Primary sequence analysis may fail to detect distant relationships among proteins, while secondary and tertiary structures might add to the analyses. Most sequences available are partial sequences, and so far no

crystal structures have been published for whole RAG proteins.

Dense sampling across diverse taxonomic groups, defining genomic diversity (or biodiversity), can improve phylogenetic reconstruction considerably and give us insights into evolutionary patterns and novelties in the molecular world.⁴⁶ As mentioned previously, there is a strong bias in databases towards partial sequences of RAG genes and proteins and also concerning mammals and birds. Evolutionary approaches would benefit tremendously from a rational choice of targets for future genome sequencing projects, especially if the aspects discussed here are taken into consideration.

The study of RAG proteins and other signature proteins of the immune system is relevant for immunomics, an interdisciplinary field that integrates immunology, genomics, proteomics, bioinformatics and related areas. In the 'omics' era, this important approach aims to unravel aspects of the origin, function, regulation and generation of diversity of the immune system.

Understanding the acquisition and evolution of the V(D)J recombination machinery will enlighten us on how the challenges posed by a changing ancient world selected the earliest immune systems. Moreover, these studies will disentangle the present pressures shaping our fellow animals and our own immune systems.

Grasping the diversity of the adaptive immune system, including the function and regulation of its key elements — such as RAG proteins — is crucial for the design and development of new therapies to modulate immune responses in humans. As researchers continue to study key components of the adaptive immune system, their interaction and regulation, they will be contributing to a better understanding of this important system and for the development of alternative new therapies.

Acknowledgments

This work was supported by FAPESP — Fundação de Amparo à Pesquisa do Estado de São Paulo, Brazil (Grant 01/02584-2), MBL Associates, Frederick B. Bang and Catherine F. Shouse Fellowships and was partly supported by grants to David D. Pollock from the National Institutes of Health (GM065612-01 and GM065580-01), the National Science Foundation (EPS-0346411) and the State of Louisiana Board of Regents (Support Fund and the Millennium Research Program's Biological Computation and Visualization Center). M.M.D.C. is an assistant professor at the University of São Paulo (USP), Brazil, and is a recipient of a Young Investigators Grant from FAPESP. L.A.N. is currently a postdoctoral associate with David D. Pollock at Louisiana State University, USA. Special thanks to Mitchell L. Sogin (Marine Biological Laboratory) for hosting M.M.D.C. in 2002; to Auro Nomizo (USP), José Gustavo Amarante-Mendes (USP) and Sérgio Luiz Pereira (Royal Ontario Museum) for valuable suggestions to this manuscript; and José Adriano de Sousa for providing artwork.

Note in proof

While this manuscript was being reviewed, another work was published, describing the evolution of duplicated RAG-1 genes in the polyploid clawed frogs, *Xenopus* and *Silurana*.⁴⁷

References

- Du Pasquier, L. (2004), 'Speculations on the origin of the vertebrate immune system', *Immunol. Lett.* Vol. 92, pp. 3–9.
- Du Pasquier, L., Zucchetti, I. and De Santis, R. (2004), 'Immunoglobulin superfamily receptors in protochordates: Before RAG time', *Immunol. Rev.* Vol. 198, pp. 233–248.
- Schatz, D.G. (2004), 'Antigen receptor genes and the evolution of a recombinase', *Semin. Immunol.* Vol. 16, pp. 245–256.
- Tonegawa, S. (1983), 'Somatic generation of antibody diversity', *Nature* Vol. 302, pp. 575–581.
- Nadel, B. and Feeney, A.J. (1997), 'Nucleotide deletion and P addition in V(D)J recombination — A determinant role of the coding-end sequence', *Mol. Cell. Biol.* Vol. 17, pp. 3768–3778.
- Schatz, D.G., Oettinger, M.A. and Baltimore, D. (1989), 'The V(D)J recombination activating gene, RAG-1', *Cell* Vol. 59, pp. 1035–1048.
- Oettinger, M.A., Schatz, D.G., Gorka, C. et al. (1990), 'RAG-1 and RAG-2, adjacent genes that synergistically activate V(D)J recombination', *Science* Vol. 248, pp. 1517–1523.
- Schatz, D.G. and Baltimore, D. (2004), 'Uncovering the V(D)J recombinase', *Cell* Vol. 116(2 Suppl.), pp. S103–S106.
- Marchalonis, J.J., Schluter, S.F., Bernstein, R.M. et al. (1998), 'Antibodies of sharks: Revolution and evolution', *Immunol. Rev.* Vol. 166, pp. 103–122.
- Schluter, S.E., Bernstein, R.M., Bernstein, H. et al. (1999), 'Big Bang' emergence of the combinatorial immune system', *Dev. Comp. Immunol.* Vol. 23, pp. 107–111.
- Schatz, D.G. (2004), 'V(D)J recombination', *Immunol. Rev.* Vol. 200, pp. 5–11.
- Eason, D.D., Cannon, J.P., Haire, R.N. et al. (2004), 'Mechanisms of antigen receptor evolution', *Semin. Immunol.* Vol. 16, pp. 215–226.
- Hansen, J.D. and Kaattari, S.L. (1995), 'The recombination activation gene 1 (RAG-1) of rainbow trout (*Oncorhynchus mykiss*): Cloning, expression, and phylogenetic analysis', *Immunogenetics* Vol. 42, pp. 188–195.
- Schluter, S.F. and Marchalonis, J.J. (2003), 'Cloning of shark RAG-2 and characterization of the RAG-1/RAG-2 gene locus', *FASEB J.* Vol. 17, pp. 470–472.
- Willett, C.E., Cherry, J.J. and Steiner, L.A. (1997), 'Characterization and expression of the recombination activating genes (*rag1* and *rag2*) of zebrafish', *Immunogenetics* Vol. 45, pp. 394–404.
- Gellert, M. (2002), 'V(D)J recombination: RAG proteins, repair factors, and regulation', *Annu. Rev. Biochem.* Vol. 71, pp. 101–132.
- Steen, S.B., Han, J.O., Mundy, C. et al. (1999), 'Roles of the 'dispensable' portions of RAG-1 and RAG-2 in V(D)J recombination', *Mol. Cell. Biol.* Vol. 19, pp. 3010–3017.
- Swanson, P.C., Volkmer, D. and Wang, L. (2004), 'Full-length RAG-2, and not full-length RAG-1, specifically suppresses RAG-mediated transposition but not hybrid joint formation or disintegration', *J. Biol. Chem.* Vol. 279, pp. 4034–4044.
- Kim, D.R., Dai, Y., Mundy, C.L. et al. (1999), 'Mutations of acidic residues in RAG-1 define the active site of the V(D)J recombinase', *Genes Dev.* Vol. 13, pp. 3070–3080.
- Landree, M.A., Wibbenmeyer, J.A. and Roth, D.B. (1999), 'Mutational analysis of RAG-1 and RAG-2 identifies three catalytic amino acids in RAG-1 critical for both cleavage steps of V(D)J recombination', *Genes Dev.* Vol. 13, pp. 3059–3069.
- De, P. and Rodgers, K.K. (2004), 'Putting the pieces together: Identification and characterization of structural domains in the V(D)J recombination protein RAG-1', *Immunol. Rev.* Vol. 200, pp. 70–82.
- Fugmann, S.D., Villey, I.J., Ptaszek, L.M. et al. (2000), 'Identification of two catalytic residues in RAG-1 that define a single active site within the RAG-1/RAG-2 protein complex', *Mol. Cell* Vol. 5, pp. 97–107.
- Groth, J.G. and Barrowclough, G.F. (1999), 'Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene', *Mol. Phylogenet. Evol.* Vol. 12, pp. 115–123.
- Murphy, W.J., Eizirik, E., Johnson, W.E. et al. (2001), 'Molecular phylogenetics and the origins of placental mammals', *Nature* Vol. 409, pp. 614–618.

25. Pereira, S.L., Baker, A.J. and Wajntal, A. (2002), 'Combined nuclear and mitochondrial DNA sequences resolve generic relationships within the Cracidae (Galliformes Aves)', *Syst. Biol.* Vol. 51, pp. 946–958.
26. Maddison, D.R., Schulz, K.-S., (2004), The Tree of Life Web Project. URL: <http://tolweb.org> (last accessed in January 2005).
27. Hedges, S.B., Blair, J.E., Venturi, M.L. *et al.* (2004), 'A molecular time-scale of eukaryote evolution and the rise of complex multicellular life', *BMC Evol. Biol.* Vol. 4, p. 2.
28. Benton, M.J. (1993), *The fossil record*, 1st editions Chapman&Hall, London, p. 845.
29. Hughes, A.L. (1999), 'Genomic catastrophism and the origin of vertebrate immunity', *Arch. Immunol. Ther. Exp.* Vol. 47, pp. 347–353.
30. Corneo, B., Moshous, D., Gungor, T. *et al.* (2001), 'Identical mutations in *RAG-1* and *RAG-2* genes leading to defective V(D)J recombinase activity can cause either T-B-severe combined immune deficiency or Omenn syndrome', *Blood* Vol. 97, pp. 2772–2776.
31. Villa, A., Sobacchi, C., Notarangelo, L.D. *et al.* (2001), 'V(D)J recombination defects in lymphocytes due to *RAG* mutations: Severe immunodeficiency with a spectrum of clinical presentations', *Blood* Vol. 97, pp. 81–88.
32. Omenn, G.S. (1965), 'Familial reticuloendotheliosis with eosinophilia', *N. Engl. J. Med.* Vol. 273, pp. 427–432.
33. Schwarz, K., Gauss, G.H., Ludwig, L. *et al.* (1996), 'Rag mutations in human B cell-negative SCID', *Science* Vol. 274, pp. 97–99.
34. Villa, A., Santagata, S., Bozzi, F. *et al.* (1998), 'Partial V(D)J recombination activity leads to Omenn syndrome', *Cell* Vol. 93, pp. 885–896.
35. Perryman, L.E. (2004), 'Molecular pathology of severe combined immunodeficiency in mice, horses, and dogs', *Vet. Pathol.* Vol. 41, pp. 95–100.
36. Mombaerts, P., Iacomini, J., Johnson, R.S. *et al.* (1992), '*RAG-1* deficient mice have no mature B and T lymphocytes', *Cell* Vol. 68, pp. 869–877.
37. Shinkai, Y., Rathbun, G., Lam, K.P. *et al.* (1992), '*RAG-2* deficient mice lack mature lymphocytes owing to inability to initiate V(D)J rearrangement', *Cell* Vol. 68, pp. 855–867.
38. Wienholds, E., Schulte-Merker, S., Walderich, B. *et al.* (2002), 'Target-selected inactivation of the zebrafish *rag1* gene', *Science* Vol. 297, pp. 99–102.
39. McGuire, T.C., Poppie, M.J. and Banks, K.L. (1974), 'Combined (B-and T-lymphocyte) immunodeficiency: A fatal genetic disease in Arabian foals', *J. Am. Vet. Med. Assoc.* Vol. 164, pp. 70–76.
40. Wiler, R., Leber, R., Moore, B.B. *et al.* (1995), 'Equine severe combined immunodeficiency: A defect in V(D)J recombination and DNA-dependent protein kinase activity', *Proc. Natl. Acad. Sci. USA* Vol. 92, pp. 11485–11489.
41. Bosma, G.C., Custer, R.P. and Bosma, M.J. (1983), 'A severe combined immunodeficiency mutation in the mouse', *Nature* Vol. 301, pp. 527–530.
42. Meek, K., Kienker, L., Dallas, C. *et al.* (2001), 'SCID in Jack Russell terriers: A new animal model of DNA-PKcs deficiency', *J. Immunol.* Vol. 167, pp. 2142–2150.
43. Jezyk, P.F., Felsburg, P.J., Haskins, M.E. *et al.* (1989), 'X-linked severe combined immunodeficiency in the dog', *Clin. Immunol. Immunopathol.* Vol. 52, pp. 173–189.
44. Noguchi, M., Yi, H., Rosenblatt, H.M. *et al.* (1993), 'Interleukin-2 receptor gamma chain mutation results in X-linked severe combined immunodeficiency in humans', *Cell* Vol. 73, pp. 147–157.
45. Henthorn, P.S., Somberg, R.L., Fimiani, V.M. *et al.* (1994), 'IL-2Rgamma gene microdeletion demonstrates that canine X-linked severe combined immunodeficiency is a homologue of the human disease', *Genomics* Vol. 23, pp. 69–74.
46. Pollock, D.D. (2002), 'Genomic biodiversity, phylogenetics and coevolution in proteins', *Appl. Bioinformatics* Vol. 1, pp. 81–92.
47. Evans, B.J., Kelley, D.B., Melnick, D.J. *et al.* (2005), 'Evolution of *RAG-1* in polyploid clawed frogs', *Mol. Biol. Evol.* Vol. 22, pp. 1193–1207.