

Review

Genomic Catastrophism and the Origin of Vertebrate Immunity

AUSTIN L. HUGHES

Department of Biology, 208 Mueller Laboratory, Pennsylvania State University, University Park, PA 16802, USA

Abstract. Genomic catastrophism is the belief that unique genetic events, unlike those observed in recent evolutionary history, played a key role in the origin of vertebrate adaptations. Catastrophist hypotheses have been particularly popular in accounting for the origin of vertebrate specific immunity. Two major such hypotheses involve genome duplication by polyploidization and horizontal gene transfer. Recent analyses lead to decisive rejection of the widely cited hypothesis that the vertebrate genome underwent two rounds of genome duplication, and theoretical considerations suggest that genome duplication is unlikely to lead to new adaptive advances. Likewise, the evidence that key elements of the vertebrate immune system arose by horizontal transfer from a bacterium or by incorporation of a transposable element into the vertebrate genome remains relatively weak. Thus, at present, a uniformitarian view of the origin of the vertebrate immune system seems more reasonable, especially given the longer time-frame for vertebrate evolution indicated by molecular data.

Key words: genome duplication; horizontal gene transfer; immune system evolution; vertebrate evolution.

The vertebrates are, as far as we know, unique among all organisms in possessing an immune system that is capable of recognizing a wide variety of foreign antigens and of mounting a response to each antigen that is somatically plastic and thus highly specific. The molecules involved in the specific immune system of vertebrates (sometimes misleadingly called “adaptive immunity”, as if the immune responses of vertebrate innate immunity or of the immune systems of other animals were not adaptive in the Darwinian sense) include the immunoglobulins (Ig), T cell receptors (TCR), and class I and class II molecules of the major histocompatibility complex (MHC). All of these are members of the Ig superfamily, whose members are characterized by a distinct folding pattern consisting of two β sheets³. Although Ig superfamily members have been found outside vertebrates, no animals except jawed vertebrates have been shown to possess Ig, TCR, or MHC. The origin of the complex suite of adaptations

associated with vertebrate specific immunity has posed a problem for evolutionary biology. The basal group of jawed vertebrates, the cartilaginous fishes (Chondrichthyes) possess Ig, TCR, and MHC that resemble, in most respects, those of other vertebrates³⁰.

Thus, the vertebrate specific immune system seems to have appeared rather suddenly in evolution. As a consequence, biologists have struggled to devise scenarios explaining this sudden appearance (called a “big bang” by MARCHALONIS and colleagues²⁴). The problem was aggravated by a widespread acceptance of in the evolutionary scenario known as the “Cambrian explosion”. According to this hypothesis, the major animal phyla diverged from each other in the Cambrian period (roughly 590–505 million years ago). Since fossil remains of jawless vertebrates are known from the Upper Cambrian and those of sharks from the Upper Silurian (over 408 million years ago)⁶, it was believed that the deuterostome lineage itself, the chordate body plan, and

the genes of the vertebrate specific immune system must all have appeared in a period of at most 100 million years.

The response of many researchers to this problem has been to seek for solution in what I call “genomic catastrophism”. I use this term in analogy to geology, in which “catastrophism” refers to theories holding that features of the earth’s surface arose as a result of catastrophic events early in the earth’s history which have no counterpart today. Genomic catastrophists hold that the immune system arose through one or more catastrophic events in the evolution of the vertebrate genome; the most commonly invoked of such events are genome duplication by polyploidization^{18, 19, 21} and horizontal gene transfer²¹. The alternative view would be one of genomic uniformitarianism. As in geology, genomic uniformitarianism would hold that the events that occurred in the distant past are likely to have been of the same sort as those we can observe today or in the recent past.

Here I review recent evidence from molecular evolutionary genetics relevant to deciding between the catastrophist or uniformitarian views of the origin of the vertebrate immune system. First, I discuss recent molecular evidence suggesting that vertebrate evolution took place over a much longer time period than was previously supposed. Next I consider the widely held view that vertebrates underwent polyploidization early in their history. Finally, I briefly discuss the hypothesis that horizontal gene transfer may have played a role in the origin of vertebrate specific immunity. My assumption is that genomic uniformitarianism should be the null hypothesis, which we should reject only if the evidence for a unique, “catastrophic” event is strong.

Vertebrate Origins

Estimates of the timing of major events of cladogenesis in the vertebrate lineage based on large numbers of genes tested statistically for clock-like evolution have revolutionized our understanding of vertebrate history^{22, 34}. The results of these studies (summarized in Fig. 1) call into question the idea of a Cambrian explosion. Rather, the deuterostome lineage, to which the vertebrates belong, had been evolving independently for nearly a billion years.

If this interpretation is true, it helps explain some recent findings regarding the evolution of vertebrate innate immunity. In addition to the specific immune system, vertebrates possess other less specific immune defenses collectively called the innate immune system. Because vertebrate innate immunity shows some

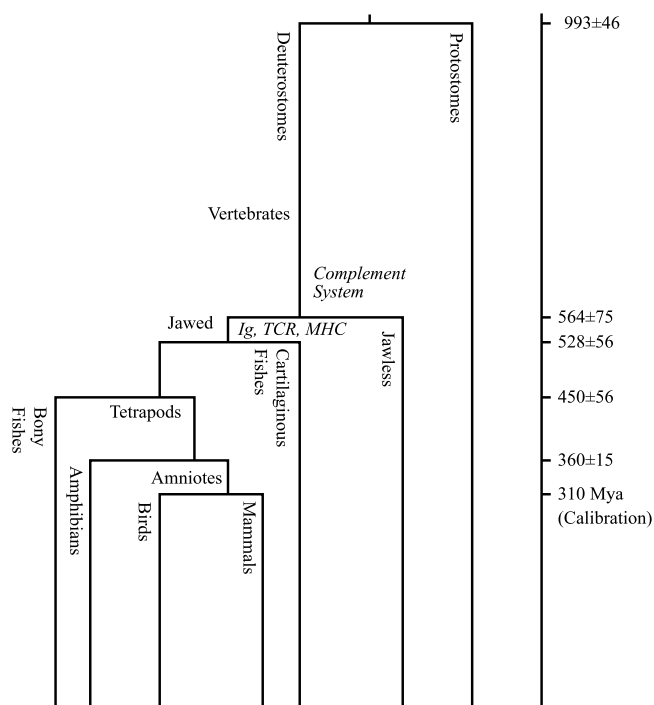


Fig. 1. Diagram of major events of cladogenesis in the history of the vertebrates. Divergence time estimates in millions of years (Mya) ± standard error are based on references^{22, 34}

general resemblance to the immune mechanisms known from invertebrates, it has frequently been suggested that these mechanisms have been conserved since the common ancestors of vertebrates and protostome phyla such as arthropods^{7, 16}. However, phylogenies of gene families having immune system representatives in both vertebrates and invertebrates do not support this hypothesis¹³. Rather, in most cases, vertebrate and invertebrate immune functions seem to have evolved independently¹³. For example, insect hemolin, which is a member of the Ig superfamily and functions in the immune response, is more closely related to Ig family members expressed in insect and vertebrate nervous systems than it is to the Ig family members involved in immunity in vertebrates (Fig. 2). Thus, members of this superfamily independently evolved immune system functions in deuterostomes (including vertebrates) and protostomes (including insects).

Given a longer time frame in which to evolve their unique adaptations for dealing with parasites, it is not really surprising that vertebrates developed a unique immune system, any more than it is surprising that they evolved a unique suite of nervous and sensory adaptations. However, the time between most recent point estimates of the divergence time of jawless vertebrates (about 564 million years ago) and that of cartilaginous fishes (528 million years ago) remains extremely short

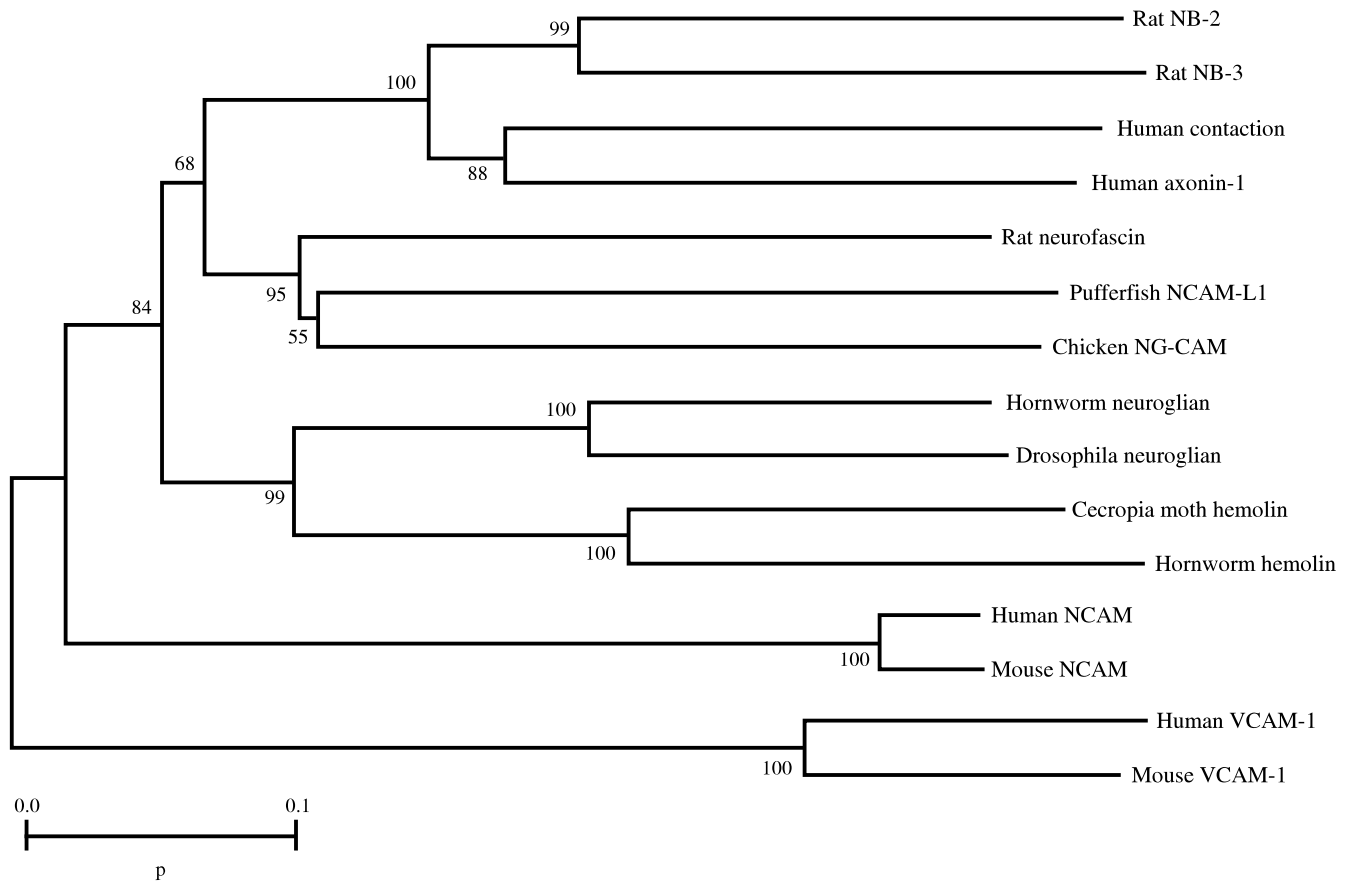


Fig. 2. Phylogenetic tree of insect hemolin and related proteins of insects and vertebrates, illustrating the close relationship between hemolin and insect neuroglians. The tree was constructed by the neighbor-joining method³¹ on the basis of the proportion of amino acid difference (p)

(less than 30 million years). It is important to recognize that the standard errors of these estimates are quite large, bounding a range of 167 million years. Even so, this may seem a short time for all of the mechanisms of specific immunity to appear and diversify.

One possibility is that the ancestors of modern jawless vertebrates had at least the rudiments of a specific immune system. If so, such a system may remain in modern jawless vertebrates (lampreys and hagfish), but may have been undetected as yet. Alternatively, although present in ancestral jawless vertebrates, the specific immune system may have subsequently been lost in their modern descendants. Lampreys and hagfish are highly specialized organisms, and they may not bear much resemblance physiologically to ancient jawless vertebrates such as ostracoderms.

Genome Duplication

One of the most widely cited hypotheses in evolutionary biology is Ohno's hypothesis that two rounds of duplication of the entire genome by polyploidization

(the 2R hypothesis) occurred early in vertebrate history^{23, 25, 32, 33}. A number of authors have asserted that these alleged events of genome duplication played a major role in the evolution of vertebrate specific immunity^{18, 19, 21}, however, this literature is very incoherent, and none of the authors explains clearly how genome duplication is supposed to have done this. Of course, genome duplication would provide a mechanism for duplicating individual genes or gene clusters, but the problem in this case is one of explaining the origin of the genes involved in vertebrate specific immunity, not merely their duplication. For example, genome duplication might seem a plausible way of explaining why there are 4 distinct types of TCR (α , β , δ and γ), but it cannot explain the origin of TCR themselves nor how the interaction of TCR and MHC molecules evolved.

It is well known that polyploidization has occurred more recently in certain lineages of bony fishes and amphibians. For example, it is well known that repeated polyploidization events have occurred in the frog genus *Xenopus*⁵. Observing that there are 7 *hox* clusters in zebrafish but only 4 in tetrapods, AMORES et al.² recent-

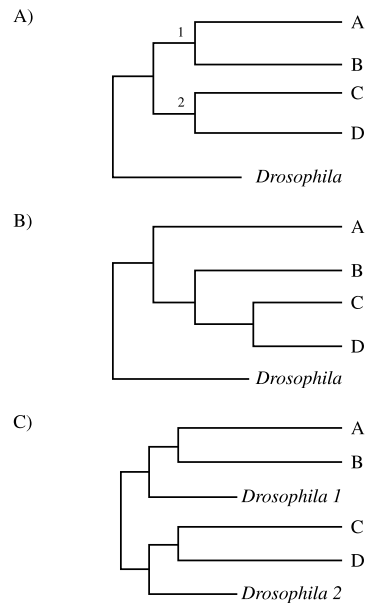


Fig. 3. Possible phylogenies of gene families having 4 members in vertebrates (A-D). A – phylogeny of the form (AB) (CD), consistent with the hypothesis of two rounds of genome duplication; B – phylogeny of the form (A) (BCD); C – phylogeny indicating that one duplication of vertebrate genes preceded the divergence of deuterostomes (including vertebrates) and protostomes (including arthropods)

ly proposed that the ancestors of bony fish underwent a round of genome duplication. However, the conclusion that this duplication occurred in the ancestors of all bony fish is unwarranted. Pufferfish have only 4 *hox* clusters, suggesting that the duplication may have occurred independently in the zebrafish lineage. In any event, there is no compelling evidence of polyploidization in the ancestors of all vertebrates^{12, 33}.

Supporters of the 2R hypothesis point to the fact that there are certain gene families having one member in *Drosophila* and 4 in vertebrates³². If this situation in fact results from 2 rounds of polyploidization, then the phylogeny of the vertebrate genes is expected to show a specific topology, showing 2 clusters of 2 genes each (Fig. 3A). I call this a topology of the form (AB) (CD). On the other hand, a topology of the form (A) (BCD), in which one vertebrate gene duplicated before the others (Fig. 3B), does not support the 2R hypothesis. Likewise, if the vertebrate genes duplicated before the origin of vertebrates, as indicated by a phylogeny like that of Fig. 3C, the 2R hypothesis is not supported.

HUGHES¹² tested the 2R hypothesis by examining phylogenies of 13 developmentally important gene families having one member in *Drosophila* and 4 in vertebrates; these include the very families listed by SIDOW³² as supporting the 2R hypothesis. In fact, only one of these families showed a topology of the form

(AB) (CD), and that received weak statistical support. In 6 families, there was statistically significant support for a topology of the form (A) (BCD)¹². In 2 families, there was significant support for duplication of the vertebrate genes before the divergence of deuterostomes and protostomes, and in one family before the divergence of vertebrates and urochordates¹². Thus, there was essentially no support for the 2R hypothesis.

Theoretical considerations also lead us to question how duplication of the entire genome could realistically lead genes encoding proteins with new functions. There are 2 models of how new protein function evolves: 1) OHNO²⁶ proposed that, after gene duplication, one gene copy is redundant and thus free to accumulate mutations at random. Most such redundant copies will eventually become pseudogenes, but a few will by chance hit upon some new beneficial function; 2) various authors^{11, 14, 17, 27} have proposed that, when duplicate genes adapt to new functions, gene duplication is ordinarily preceded by a period of “gene sharing”²⁸, when a single gene encodes a bifunctional protein product. Under this latter model, positive Darwinian selection after gene duplication is expected to play a role in adaptation of daughter genes to their specific functions. Molecular evidence strongly argues against Ohno’s model as a general explanation for the origin of new protein function and generally supports the latter model^{11, 14}.

In the case of duplication of the entire genome, if Ohno’s model were true, we would predict that the vast majority of duplicated genes would become pseudogenes. Thus, contrary to the view of the genomic catastrophists, ancient genome duplication would be largely irrelevant to modern functional genomics. On the alternative model, because of the role of positive selection, simultaneous adaptation of large numbers of duplicate genes to new functions would impose a substitutional load^{8, 15} that no population could bear. Either model, then, yields the prediction that polyploidization in itself will not lead to major adaptive innovations.

This prediction is consistent with what is observed in recent polyploids. For example, the numerous polyploidization events in the frogs of the genus *Xenopus*⁵ have had no detectable phenotypic effects. They have certainly led to no changes in body plan or other major adaptive advances. Contrary to Ohno’s hypothesis, duplicate genes in *Xenopus laevis* are subject to purifying selection as long as they are expressed²⁰ and thus are not free to accumulate mutations at random, but none is known to have achieved an important new function. On the uniformitarian view, this is to be expected, and if recent polyploidization has had no major phenotypic

	Zinc Finger Domain	
	293	334
Human RAG1	CQICEHILADP	VETNCKHVFCRVCILRCLKVMGSYCP
Mouse RAG1	CQICEHILADP	VETSCKHLFCRICILRCLKVMGSYCP
Chicken RAG1	CQICEHILADP	VETTCRHLFCRTCILSCIRVMGCYCP
Clawed frog RAG1	CLVCEHILSDP	VQTSCKHLFCRICILKYIKLMGCYCP
Rainbow trout RAG1	CQVCDHLLSEP	VQSPCRHLFCRSCIAYIISLGPHCP
Bull shark RAG1	CQVCEHLLADP	VQSPGMHLFCRTCILKCVTVLGKYCP
<i>E. nidulans</i> usvH	CQVCKDFFDN	PVITSCSHTFCSLCIRRCLST-EGKCPT
<i>N. crassa</i> USV2	CHVCKDFYD	SPMLTSCNHTFCSLCIRRCLSV-DSKCPL
Yeast RAD18	CHICKDFLK	VPVLTFCGHTFCSLCIRTHLNN-QPNCPL
	* . * . .	* . * * * * . * * *
	396	408
Human RAG1	RQHLLSLTRRAQK	
Mouse RAG1	RQHLLSLTRRAQK	
Chicken RAG1	RQHLLSLTRRAQK	
Clawed frog RAG1	RQHLLTLTRRAQK	
Rainbow trout RAG1	RQHLLSLTRRAQK	
Bull shark RAG1	RQHLLSLTRRAQK	
<i>E. nidulans</i> usvH	RPSILQLARMA-Q	
<i>N. crassa</i> USV2	RKVLEFARTP-P	
Yeast RAD18	RSSLDDALRIPKP	
	* . * * .	
	701	727
Human RAG1	FKF-IFRGTGYDEKLVREVEGLEASGSV	
Mouse RAG1	FKF-IFRGTGYDEKLVREVEGLEASGSV	
Chicken RAG1	FKF-IFRGTGYDEKLVREVEGLEASGST	
Clawed frog RAG1	FKF-SFRGTGYDEKLVREVEGLEASGSL	
Rainbow trout RAG1	FRF-HFRGTGYDEKLVREVEGLEASGST	
Bull shark RAG1	FQF-IFRGTGYDEKLVREVEGLEASGSI	
<i>E. nidulans</i> usvH	INYSLLKDTVLRKKL-KDL-GIPNWGPR	
<i>N. crassa</i> USV2	LAYSIVRDTALRKKL-SEL-GLSTHGSR	
Yeast RAD18	INFTSMTQSQIKQKL-SSL-GLSTNGTR	
	*	* *

Fig. 4. Portions of an alignment (constructed with the CLUSTAL V program⁹) of vertebrate RAG1 with RAD18 of yeast (*Saccharomyces cerevisiae*) and related fungal DNA-binding proteins (from *Emericella nidulans* and *Neurospora crassa*). The numbers of residues in human RAG1 are indicated. “*” indicates a residue conserved in all sequences; “.” indicates a position having chemically similar residues in all sequences

effects, there is no reason to believe that more ancient polyploidizations did either.

Horizontal Gene Transfer

The process of segmental joining by which vertebrate Ig and TCR are assembled involves a number of unique proteins, including the recombination activators genes 1 (RAG1) and RAG2. BERNSTEIN et al.⁴ observed some amino acid sequence similarity between RAG1 and RAG2 and certain bacterial integrases, involved in the bacterial site-specific recombination system, and proposed that the ancestors of RAG1 and RAG2 were horizontally transferred from bacteria to vertebrates early in vertebrate history. However, RAG1 and RAG2

also show similarity to eukaryotic DNA-binding proteins. For example, RAG1 shows similarity throughout its length to yeast RAD18 and related fungal DNA-binding proteins; the regions of greatest similarity are illustrated in Fig. 4. These are the zinc-finger domain and two other regions of unknown function, one of which overlaps the region of similarity with bacterial Fim B pointed out by BERNSTEIN et al.⁴ RAG2 shows similarity to another yeast DNA-binding protein, SAS. These resemblances suggest that the ancestors of RAG1 and RAG2 may have been present in the eukaryotic ancestors. Thus, sequence comparisons alone provide no compelling reason for accepting the extraordinary mechanism of horizontal gene transfer.

Recently, experimental evidence that RAG1 and

RAG2 together can act as a transposase *in vitro* has been taken as supporting the hypothesis that these genes originated in an ancient transposable element that was somehow “tamed” and adapted to its immune system function by ancient vertebrates^{1, 10, 32}. While consistent with this hypothesis, the transposase-like features of RAG1 and RAG2 do not in themselves prove it. We know at present essentially nothing about the origin of transposable elements. For example, it is possible that transposable elements have themselves originated from recombination-promoting genes that have “escaped” from genomes of cellular organisms, rather than such genes being “tamed” transposable elements. Furthermore, it is possible that the transposase-like features of RAG1 and RAG2 have simply arisen as a fortuitous by-product of these molecules’ function in segmental rearrangement. Certainly, these results raise an interesting possibility regarding the origin of RAG1 and RAG2, but at present it is no more than a possibility.

Conclusions

Two major types of “catastrophic” events have been alleged to play a role in the origin of vertebrate specific immunity: 1) genome duplication; and 2) horizontal gene transfer. There is no good evidence for the former, and indeed there are theoretical reasons for doubting that genome duplication could give rise to important adaptive changes. As regards the exogenous origin of RAG1 and RAG2, it remains only one of several viable hypotheses. From the point of view of scientific method, it seems important to be skeptical of claims of unique genetic events early in vertebrate history until more substantial evidence becomes available. Molecular data are now giving us indications that the independent evolution of the vertebrate lineage has occupied a much longer time than was previously supposed. This longer time frame makes it quite conceivable that the unique adaptations of vertebrates, including those of the specific immune system, arose by the ordinary processes of gene duplication, recombination, drift, and natural selection that we can observe in more recent populations.

References

1. AGRAWAL A., EASTMANN Q. E. and SCHATZ D. G. (1998): Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. *Nature*, **394**, 744–751.
2. AMORES A., FORCE A., YAN Y. -L., JOLY L., AMEMIYA C., FRITZ A., HO R. K., LANGELAND J., PRINCE V., WANG Y. -L., WESTERFIELD M., EKKER M. and POSTLETHWAIT J. H. (1998): Zebrafish *hox* clusters and vertebrate genome evolution. *Science*, **282**, 1711–1714.
3. BARCLAY A. N., BEYERS A. D., BIRKELAND M. L., BROWN M. H., DAVIS S. J., SOMOZA C. and WILLIAMS A. F. (1993): *The leucocyte antigens factsbook*. Academic Press, London.
4. BERNSTEIN R. M., SCHLUTER S. F., BERNSTEIN H. and MARCHALONIS J. J. (1996): Primordial emergence of the recombination activating gene 1 (RAG1): sequence of the complete shark gene indicates homology to microbial integrases. *Proc. Natl. Acad. Sci. USA*, **93**, 9454–9459.
5. BISBEE C. A., BAKER M. A., WILSON A. C., IRANDOKHT H. -A. and FISCHBERG M. (1977): Albumin phylogeny for clawed frogs (*Xenopus*). *Science*, **195**, 785–787.
6. CARROLL R. L. (1988): *Vertebrate paleontology and evolution*. W. H. Freeman, New York.
7. HABICHT G. S. (1994): Primordial immunity: foundations for the vertebrate immune system. *Ann. NY Acad. Sci.*, **712**, ix–xi.
8. HALDANE J. B. S. (1957): The cost of natural selection. *J. Genet.*, **55**, 511–524.
9. HIGGINS D. G., BLEASBY A. J. and FUCHS R. (1992): Clustal V: improved software for multiple sequence alignment. *Comput. Appl. Biosci.*, **8**, 189–191.
10. HIOM K., MELEK M. and GELLERT M. (1998): DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. *Cell*, **94**, 463–470.
11. HUGHES A. L. (1994): The evolution of functionally novel proteins after gene duplication. *Proc. R. Soc. Lond. Biol. Sci.*, **256**, 119–124.
12. HUGHES A. L. (1999): Phylogenies of developmentally important proteins do not support the hypothesis of two rounds of genome duplication early in vertebrate history. *J. Mol. Evol.*, **48**, 565–576.
13. HUGHES A. L. (1998): Protein phylogenies provide evidence of a radical discontinuity between arthropod and vertebrate immune systems. *Immunogenetics*, **47**, 283–296.
14. HUGHES A. L. (1999): *Adaptive evolution of genes and genomes*. Oxford University Press, New York.
15. HUGHES M. K. and HUGHES A. L. (1993): Evolution of duplicate genes in a tetraploid animal, *Xenopus laevis*. *Mol. Biol. Evol.*, **10**, 1360–1369.
16. HULTMARK D. (1993): Immune reactions in *Drosophila* and other insects: a model for innate immunity. *Trends Genet.*, **9**, 178–183.
17. JENSEN R. A. (1976): Enzyme recruitment in the evolution of new function. *Ann. Rev. Microbiol.*, **30**, 409–425.
18. KASAHARA M. (1998): What to the paralogous regions in the genome tell us about the origin of the adaptive immune system? *Immunol. Rev.*, **166**, 159–175.
19. KASAHARA M., NAKAYA J., SATTA Y. and TAKAHATA N. (1997): Chromosomal duplication and the emergence of the adaptive immune system. *Trends Genet.*, **13**, 90–92.
20. KIMURA M. (1968): Evolutionary rate at the molecular level. *Nature*, **217**, 624–626.
21. KLEIN J., SATO A. and O’HUGIN C. (1998): Evolution by gene duplication in the major histocompatibility complex. *Cytogenet. Cell Genet.*, **80**, 123–127.
22. KUMAR S. and HEDGES S. B. (1998): A molecular timescale for vertebrate evolution. *Nature*, **392**, 917–920.
23. LUNDIN L. G. (1993): Evolution of the vertebrate genome as reflected in paralogous chromosomal regions in man and house mouse. *Genomics*, **16**, 1–19.

24. MARCHALONIS J. J., SCHLUSTER S. F., BERNSTEIN R. M. and HOHMAN V. S. (1998): Antibodies of sharks: revolution and evolution. *Immunol. Rev.*, **166**, 103–122.
25. OHNO S. (1970): *Evolution by gene duplication*. Springer, New York.
26. OHNO S. (1973): Ancient linkage groups and frozen accidents. *Nature*, **244**, 259–262.
27. ORGEL L. E. (1977): Gene-duplication and the origin of proteins with novel functions. *J. Theor. Biol.*, **67**, 773.
28. PIATIGORSKY J. and WISTOW G. (1991): The recruitment of crystallins: new functions precede gene duplication. *Science*, **252**, 1078–1079.
29. PLASTERK R. (1998): Ragtime jumping. *Nature*, **394**, 718–719.
30. RAST J. P. and LITMAN G. W. (1998): Towards understanding the evolutionary origins and early diversification of rearranging antigen receptors. *Immunol. Rev.*, **166**, 79–86.
31. SAITOU N. and NEI M. (1987): The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.*, **4**, 406–425.
32. SIDOW A. (1996): Gen(om)e duplications in the evolution of early vertebrates. *Curr. Opin. Genet. Dev.*, **6**, 715–722.
33. SKRABANEK L. and WOLFE K. H. (1998): Eukaryote genome duplication: where is the evidence? *Curr. Opin. Genet. Dev.*, **8**, 694–700.
34. WANG D. Y., KUMAR S. and HEDGES S. B. (1999): Divergence time estimates for the early history of animal phyla and the origin of plants, animals, and fungi. *Proc. R. Soc. Lond. Biol. Sci.*, **266**, 163–171.

Received in April 1999

Accepted in June 1999