Published in final edited form as:
Nat Rev Immunol. 2005 November ; 5(11): 866–879. doi:10.1038/nri1712.

RECONSTRUCTING IMMUNE PHYLOGENY: NEW PERSPECTIVES

Gary W. Litman[*,‡,§], **John P. Cannon**[*,‡], and **Larry J. Dishaw**[‡,§]

[*]Department of Pediatrics, University of South Florida College of Medicine, University of South Florida/All Children's Hospital Children's Research Institute, 830 First Street South, Saint Petersburg, Florida 33701, USA.

[‡]All Children's Hospital, Department of Molecular Genetics, 801 Sixth Street South, Saint Petersburg, Florida 33701, USA.

[§]H. Lee Moffitt Cancer Center & Research Institute, 12902 Magnolia Avenue, Tampa, Florida 33612, USA.

Abstract

Numerous studies of the mammalian immune system have begun to uncover profound interrelationships, as well as fundamental differences, between the adaptive and innate systems of immune recognition. Coincident with these investigations, the increasing experimental accessibility of non-mammalian jawed vertebrates, jawless vertebrates, protochordates and invertebrates has provided intriguing new information regarding the likely patterns of emergence of immune-related molecules during metazoan phylogeny, as well as the evolution of alternative mechanisms for receptor diversification. Such findings blur traditional distinctions between adaptive and innate immunity and emphasize that, throughout evolution, the immune system has used a remarkably extensive variety of solutions to meet fundamentally similar requirements for host protection.

The evolutionary development of the METAZOANS was associated with the diversification of a wide range of specialized cell-surface molecules that mediate key metabolic processes, as well as provide crucial contact interfaces and carry out a broad range of other essential functions. It is not unexpected that some of these molecules also came to function as barriers to pathogenic invasion and, in doing so, began to carry out dedicated innate immune protective functions. Whereas the simplest form of protection, barrier formation, is essentially mechanical in nature, relentless pressure from genetic variation in pathogens probably drove the evolution of such innate immune protective molecules towards diversification and, in parallel, towards integration of signalling pathways to regulate cellular responses to external stimulation. However, despite the sophistication that such innate immune mediators achieved over time, their biological complexity, by definition, would be limited by genome space, so with increasing complexity of body plan and/or increasing pathogen sophistication, they could be overwhelmed.

Correspondence to G.W.L. litmang@allkids.org.

Competing interests statement The authors declare no competing financial interests.

Online links DATABASES The following terms in this article are linked online to: Entrez Gene: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?db=gene BCR | RAG | γδ TCR | TCR α-chain | TCR β-chain
FlyBase: http://flybase.bio.indiana.edu/Dscam

FURTHER INFORMATION Zebrafish: http://www.ensembl.org/Danio_rerio

More than 500 million years ago, a TRANSPOSITION event, probably involving a recombination-activating gene (RAG)-bearing element, might have given rise to the predecessors of the rearranging antigen-binding receptors of the jawed vertebrates, which encompass the vertebrate radiations that extend from the cartilaginous fish through to humans. This is considered the defining point in the emergence of RAG-mediated (conventional) adaptive immunity[1,2], which has evolved to create a mechanism for deriving almost limitless variation from very few genes. Studies in traditional and non-traditional animal models, such as sharks, bony fish and birds, have brought this event and its ramifications for host defence into sharper focus. We can now predict much about how these rearranging antigen-binding receptors probably arose, what alternative pathways of immune-receptor gene evolution have occurred, what relationships exist between B- and T-cell-mediated immunity and natural killer (NK)-cell function, how complex immune recognition might be solved in species that are below the phylogenetic level of the jawed vertebrates, and what general principles link adaptive and innate immune recognition[3,4]. It is now clear that, contrary to traditional views, jawless vertebrates, protochordates and invertebrates have also evolved sophisticated RAG-independent strategies to effect recognition and facilitate elimination of pathogens, to respond to stress, and to distinguish self from non-self [3,5–10]. Some of the molecules and mechanisms that are used to accomplish these processes are related to those used by jawed vertebrates, whereas others seem to be unique solutions to host survival, opening up the possibility that they, or alternative forms, might have broader phylogenetic distributions than are envisioned at present.

The comparative study of immune systems of extant vertebrate and invertebrate species holds important clues about how immune recognition has evolved from generalized to highly (antigen) specific reactions and how integral links between innate and adaptive immunity have been established and maintained. These efforts are being facilitated by the availability of whole-genome information, by the improved methodology for influencing the expression of individual genes (such as transgenesis or RNA interference) and by the development of approaches to examine gene expression and protein interactions in a global context. Studies in key representative vertebrate and invertebrate species — including insects such as *Drosophila melanogaster* (fruitfly), molluscs such as snails, protochordates such as *Botryllus schlosseri* and *Branchiostoma floridae* (amphioxus), and jawless vertebrates such as lampreys and hagfish — are beginning to provide a window on different mechanisms of GERMLINE DIVERSIFICATION and SOMATIC DIVERSIFICATION (FIG. 1). Specialized innovations that influence the generation of immune diversity and the subsequent implementation of powerful selective mechanisms are being recognized[5,8–10].

This Review describes and compares several key immune-related systems that have extensive genomic and/or somatic diversification of innate and adaptive immune-type receptors. Certain other types of receptor that are important in host defence[3] are mentioned, but detailed discussion of these is beyond the scope and intent of the main focus of this Review. In the broadest sense, observations that have been made so far show parallel evolution of various forms of immune defence, and this can be viewed as a continuum in which multiple independent shifts from limited to complex germline variation, and of specialized commitments of somatic cells, have occurred. Even in the context of the limited number of metazoan phyla that have been studied in detail, we can now describe many of the alternative mechanisms of immune recognition that have emerged at varying points in phylogeny and can identify key processes that have bridged and functionally integrated innate and adaptive immunity throughout the evolution of immune competence.

An overview of adaptive immunity in phylogeny

To understand the relationships between innate and adaptive immunity, which are intertwined functionally in modern vertebrates[11–16], it is first instructive to consider some key events in the evolution of adaptive (RAG-mediated) immunity. Two central observations that were made two decades ago showed that markedly disparate mechanisms — differing from each other and from the human and mouse pattern of a single rearranging receptor-gene locus that undergoes marked somatic diversification — give rise to B-cell receptor (BCR) diversity in jawed vertebrates. In sharks, the ancestors of which diverged early in jawed-vertebrate phylogeny, BCRs are encoded by hundreds of independent immunoglobulin-gene loci[17], some of which undergo RAG-dependent recombination and others of which are pre-joined in the germline and therefore do not undergo combinatorial joining or generate JUNCTIONAL VARIATION[18]. By contrast, in birds, which diverged more recently as a group than sharks, the immunoglobulin loci consist of single variable (V) and joining (J) elements (which together encode immunoglobulin light chains), and single V and J elements plus multiple diversity (D) elements (which together encode immunoglobulin heavy chains). GENE CONVERSION, which until this discovery was an underappreciated mechanism of immune-receptor diversification, uses pseudogenes that are encoded upstream of the main gene loci to somatically change V regions in birds[19–21] (as well as in some mammals[22,23]); the 'converted' loci can also be somatically hypermutated[24]. Given these differences in the organization and the mechanisms of somatic diversification between sharks and birds, which strongly indicate co-evolution of immunoglobulin heavy- and light-chain gene organization and diversification mechanisms[25], it comes as little surprise that BCRs in other species also show lineage-dependent divergence of immunoglobulin-gene organization and diversification (FIG. 2). Furthermore, certain characteristics of BCR genomic organization that were previously thought to be restricted to cartilaginous fish, such as cluster-type gene organization, are also present in bony fish[26] (such as catfish) and fleshy-finned fish[27] (such as lungfish). Despite the divergence of heavy-chain types that has now been recognized in cartilaginous[28] and bony[29–31] fish, prototypical heavy-chain CLASS-SWITCH RECOMBINATION is not seen below the phylogenetic level of amphibians[32,33]. Given the DERIVED character of the extant bony fish, the resolution of whole genomes for the pufferfish[34], the zebrafish (see the Online links box) and other bony-fish species could offer considerable insight into the mechanisms that have driven the evolution of BCR diversification.

In marked contrast to the BCRs, T-cell receptors (TCRs) are relatively conserved throughout the jawed vertebrates in terms of the number, organization and multiplicity of rearranging elements[28]. The remarkable conservation of MHC class I (REF. 35) and MHC class II (REFS 36,37) structure and polymorphism throughout jawed-vertebrate phylogeny mirrors the relatively stable history of the TCRs, despite organizational changes that have taken place in the loci encoding these genes[38] (BOX 1). This parallel stability is probably driven by the functional interdependence and possible co-evolution of the genes encoding the TCR α-chain and TCR β-chain and the MHC molecules. The endogenous ligands for the MHC-class-I- and MHC-class-II-independent γδ TCR are less clear[39], and explanations for the organizational stability of the genes encoding the γδ TCR are less apparent. The predicted stability of TCRs, and the nearly continuous variation in complexity, gene number and structure of BCRs, is one of the most consistent themes in the phylogeny of adaptive (RAG-mediated) immunity.

Transitions from innate to adaptive immunity

Understanding the forerunners of BCRs and TCRs, and the mechanisms that promoted their divergence during evolution, might shed light on past and present relationships between innate and conventional adaptive immune recognition. Although the pathways and patterns

of evolution of rearranging antigen-binding receptors will never be known with absolute certainty, one plausible progression could start with the recruitment of an immunoglobulin-gene superfamily (IgSF) member, such as those that are found in species as phylogenetically distant from vertebrates as *Caenorhabditis elegans* and *D. melanogaster*[40]. The most recent common ancestor of the rearranging antigen-binding receptors would have been such an IgSF member; it would have either pre-existed as, or subsequently given rise to, an immunoglobulin V-region-type innate immune cell-surface receptor or a cell-surface adhesion molecule, as is indicated by one interpretation of the crystallographic analysis of shark NAR (new antigen receptor)[41], an immunoglobulin heavy-chain homodimer that binds specific antigens[42]. The J motif (FGXGTXLXV, where X denotes any amino acid), which promotes homo- and heterodimer formation, might have been encoded by a separate exon and become contiguous with the V-encoding region through intron loss, a recognized feature of genome evolution[34]. A VJ structure is the likely target of the key transposition event. Although it is overly simplistic to assume that such a VJ molecule persists in contemporary life forms, several single-V molecules or V-related molecules (or, more precisely, the features of their sequences) are interesting to consider[43]; however, any interpretation of such data also needs to take into account canonical V regions that do not function in immune recognition.

The origins of RAG1 and RAG2 themselves are unclear; however, a convincing homologue of *RAG1* has been found in the sea-urchin genome[3,44] and is contiguous with a structure that is homologous to *RAG2* (REF. 3, and S. Fugmann and J. Rast, personal communication). A single RAG-mediated transposition event into the VJ region could have created a potential junctional interface. Variation at the junctions during RECOMBINATION-SIGNAL-SEQUENCE-mediated recombination would then be introduced by the activation of DNA-binding proteins and repair enzymes that are not specific to the immune system, as well as terminal deoxynucleotidyltransferase, which is an error-prone DNA polymerase of the type X (PolX) family[45] that mediates non-templated base insertion during NON-HOMOLOGOUS END JOINING, the main source of somatic variation in BCRs and TCRs[46]. The capture of seemingly unrelated mechanisms of gene rearrangement and diversification seems to be a generalized phenomenon in the development of immune strategies (discussed later). Class-switch recombination (another form of somatic reorganization that occurs in BCR-gene loci in organisms as phylogenetically distant from mammals as amphibians[32]), as well as somatic hypermutation and gene conversion, is mediated by many of the same DNA-binding proteins that are involved in V, D and J joining, as well as by activation-induced cytidine deaminase (AID)[47]. AID is present in bony fish, and interestingly, zebrafish AID can catalyse class-switch recombination when introduced into mammalian B cells[48]. So, the appearance of a class-switch reaction was dependent on the evolution of switch-region sequences. AID-like and uracil-DNA glycosylase (UNG)-like proteins, which are involved in these processes, had ancient origins, co-evolving with innate immune antiviral defences[49–51].

Over time, V-gene duplication and germline substitutions, which were focused in COMPLEMENTARITY-DETERMINING REGION 1 (CDR1) and CDR2, coupled with combinatorial joining, imparted a considerable selective advantage in terms of deriving receptor diversity. Somatic hypermutation of immunoglobulin genes, which is a central feature of BCR maturation, occurs in the earliest extant jawed vertebrates — the cartilaginous fish[52,53] — and might be an accentuation of a genome-wide process[54,55]. ALLELIC EXCLUSION of BCR- and TCR-gene loci, which is an integral component of single-cell immune commitment that is present throughout jawed-vertebrate phylogeny[56], is also a feature of the loci encoding olfactory receptors[57,58] and protocadherins[59], emphasizing again the integration of pre-existing mechanisms in the evolution of immunity. The full potential of these features did not come

to fruition until the co-evolution of mechanisms to select and expand individual clones of immune cells.

Collectively, these observations indicate that the evolution of the V-type antigen-binding receptors from mediators of innate immunity to the fore runners of adaptive (RAG-mediated) immunity probably occurred incrementally through multiple recruitment events rather than as an immediate consequence of an immunological 'big bang'[60], a term that more aptly relates to the apparent absence of RAG in jawless vertebrates. RAG-mediated adaptive immunity is one of the most elaborate examples of the integration of disparate cellular and developmental processes to achieve a highly specialized, lineage-restricted biological function: that is, the use of segmental gene re arrangement to produce a fertile receptor-gene pool for the selection and clonal expansion of individual somatic cell lineages. Many of the elements that we collectively recognize to be common features of jawed-vertebrate adaptive immune receptors, including multigene complexity, somatic hypermutation and gene conversion, also diversify immune-type receptors in jawless vertebrates, protochordates and invertebrates (discussed later).

NK and NK-type receptors in jawed vertebrates

Further insight into relationships between innate and adaptive immunity can be gained from a basic consideration of NK cells, which share a common progenitor with T and B cells but do not somatically rearrange their receptors. NK cells are primary innate immune sentinels for the detection of certain viral infections and malignant transformation. The main classes of NK-cell receptors in mammals are either related to C-type lectins or members of the IgSF, and they are expressed by NK cells and natural killer T (NKT) cells[61]. In humans, the main diversified NK-cell receptors are killer-cell immunoglobulin-like receptors (KIRs), which are activating or inhibitory TYPE I TRANSMEMBRANE PROTEINS with immunoglobulin C2 (constant) ectodomains. In mice, the corresponding receptors are members of the Ly49 gene family and are activating or inhibitory TYPE II TRANSMEMBRANE PROTEINS. The extracellular regions of Ly49 gene products are single C-type-lectin domains. Haplotypic variant and invariant genetic regions of KIRs and Ly49 proteins are interspersed similarly[62]; extensive haplotypic variability in KIRs arises through meiotic recombination[61]. Certain KIRs and Ly49 proteins recognize MHC class I molecules. Other NK-cell receptors recognize a broad range of stress-induced molecules and other non-classical MHC molecules[63,64], as well as viral gene products[65,66] and other NK-cell receptors[67]. NK-cell activity has a broad phylogenetic distribution extending to invertebrates[68]. A lectin-like, NK-cell-type receptor has been characterized in a protochordate[69]; however, its sequence similarity to other NK-cell receptors is remote, and its function is unknown. Putative C-type-lectin NK-cell receptors have been reported in bony fish[70–72]; however, they encode members of the type 2 family of C-type-lectin domains and not the type 5 family, in which mammalian NK-cell receptors are classified.

Determining the evolutionary origins of NK cells and their receptors could be important for understanding the phylogeny of adaptive immunity. However, given the diversity of NK-cell receptors that have already been identified in mammals, this could prove daunting. The NITRs (novel immune-type receptors) — a large multigene family of NK-cell-like activating and inhibitory cell-surface receptors in bony fish — show an intriguing link between conventional adaptive immune receptors and innate immune receptors[73] (TABLE 1). Most NITRs consist of an amino (N)-terminal membrane-distal immunoglobulin V-region-type extracellular domain and a membrane-proximal intermediate (I)-type extracellular domain; other NITRs have only a single V-region-type extracellular domain. Most NITRs have IMMUNORECEPTOR TYROSINE-BASED INHIBITORY MOTIFS (ITIMs) in the cytoplasmic domain and therefore are probably inhibitory. Other NITRs have positively charged residues in their transmembrane regions and probably interact with adaptor molecules that contain

IMMUNORECEPTOR TYROSINE-BASED ACTIVATION MOTIFS (ITAMs)[74]. Some NITRs have amino-acid motifs that are related to the IMMUNORECEPTOR TYROSINE-BASED SWITCH MOTIF (ITSM) that has been described in members of the mammalian CD2, SIRP (signal-regulatory protein), SIGLEC (sialic-acid-binding immunoglobulin-like lectin), CEA (carcinoembryonic antigen) and PIR (paired immunoglobulin-like receptor) families, potentially allowing the modulation of signalling by differential interaction with SRC homology 2 (SH2)-domain-containing adaptor proteins (thereby allowing either activating or inhibitory function)[75]. Other NITRs lack a transmembrane region and resemble the decoy molecules that have been described in other activating and inhibitory receptor systems[76].

High-confidence molecular modelling of NITR V regions shows them to resemble the V domains of immunoglobulin and TCRs[74]; in some cases, they have motifs similar to the J regions of BCRs and TCRs. In pufferfish, 13 main families of NITR V regions are encoded at a single compact locus[73]. In zebrafish, 12 families of NITR V domains are encoded at a single locus on chromosome 7 (REF. 74), and 2 are encoded on chromosome 14 (see zebrafish in the Online links box). Although the NITRs in pufferfish and zebra fish are unequivocal ORTHOLOGUES, the respective gene families do not show absolute SYNTENY. Approximately 150 alleles and 45 structural variants of NITRs have been identified in zebrafish[74]; only one is an unequivocal activating form. Intergenic variation is regionalized within the large NITR1 gene family, and it is reminiscent of the CDR variation that is seen in BCRs and TCRs. NITR genes are transcribed in several different cell lineages in bony fish[77], including NK-type cells and cytotoxic T lymphocytes[78]; at least one NITR protein has a high degree of specificity for an allogeneic target (J.P.C., D. Eason and N. Miller, unpublished observations). The evolution of NITRs is extraordinarily rapid, even within a species, a characteristic that is shared by immunoglobulins, TCRs and KIRs, and this has been attributed to a gene BIRTH-AND-DEATH process[79]. The allelic and haplotypic variation in the NITR genes in a single strain of zebrafish is reminiscent of the variation in KIRs from various human subpopulations[61,62]. KIRs and NITRs share other features, including differential expression by individual cells and expression by NK cells and T cells[80]. Notably, a recent release of the zebrafish genome sequence has uncovered two groups of genes with sequence similarity to human KIRs, on chromosome 2 and chromosome 7 of the zebrafish (see zebrafish in the Online links box).

At least two explanations could account for the evolutionary origins of NITRs: they might be derived from a recombination event between a member of an activating and inhibitory gene family and elements of a locus encoding a rearranging receptor or V-type receptor; or they might have a common origin with the V-region-containing molecules NK-cell protein 30 (NKp30) and NKp44 (REFS 81,82) or, as has been suggested, with a family of activating and inhibitory membrane proteins containing V domains that has been found in a jawless vertebrate[83]. In either case, the NITRs present a compelling case for parallel evolution of BCR- and TCR-type V-region diversity in the context of innate immunity.

Anticipatory immunity in jawless vertebrates

Lymphocytes can be traced to jawless vertebrates. In lampreys, mononuclear cells with morphological features of lymphocytes have been identified in various tissues and in the peripheral blood[84]. Accelerated allograft rejection has been documented in lampreys[84], as have humoral responses to immunization with bacterial and cellular antigens; however, the circulating AGGLUTININS are not related structurally to antibodies[28]. Extensive EXPRESSED SEQUENCE TAG (EST) surveys have failed to identify typical lymphocyte-specific cell-surface determinants in lampreys[85,86], with the exceptions of a predicted gene product that shows a distant relationship to CD4 (among other molecules)[87] and another monomorphic gene product that has cytosolic ITIMs and can be modelled to TCR V regions but shows only remote primary-

structure similarity to jawed-vertebrate TCRs[87]. However, Ikaros and SPI, which are transcription factors that function in higher-vertebrate lymphocyte development and differentiation, have been identified in jawless vertebrates[88–91], as have ETS- (M. Anderson, personal communication), GATA-binding protein 2 (GATA2)- and GATA3-type molecules (J.P.C. and R. Haire, unpublished observations). Using a V-region-directed cloning technique in lamprey lymphocytes, a single gene was identified that resembles *VpreB*, which encodes a surrogate BCR light-chain component that is expressed early in lymphocyte ontogeny in higher vertebrates but does not participate directly in the function of the mature BCR[92]. The lamprey VpreB-like molecule might reflect the character of an early common ancestor that might or might not have mediated an immune function[93]. In addition to the V-region-related sequences in hagfish[83], additional immunoglobulin V-region-related sequences that have so far been identified in lampreys include nectin and Lutheran blood-group-antigen-like genes, neither of which is related closely to BCRs or TCRs[93]. Whole-genome analysis might yet uncover BCR- or TCR-related candidates.

Alternative mechanisms might mediate adaptive immunity in jawless vertebrates. Antigenic or mitogenic stimulation of short-term cultures of lamprey peripheral-blood leukocytes results in considerable proliferation of lymphocytes[9]. Subtracted cDNAs recovered from these cell cultures are highly enriched in sequences encoding a particularly complex set of LEUCINE-RICH REPEAT (LRR)-containing extracellular molecules. Notably, LRRs are a core feature of markedly diverse forms of pattern-recognition molecule that mediate receptor–ligand interactions throughout phylo geny[94–98]. The LRR-encoding cDNAs in lampreys are encoded at a single locus. On the basis of the relatively high degree of hypervariation in the LRR cassette, the lack of variation in the N and carboxyl (C) termini, and expression in lymphocyte-like cells, the molecule has been designated a variable lymphocyte receptor (VLR). VLRs consist of the following regions: a 30–38-residue N-terminal region, an 18-residue N-terminal LRR, followed by up to nine 24-residue LRR cassettes, a 13-residue connecting peptide, a 48–58-residue C-terminal LRR, a threonine- and proline-containing stalk, a glycosylphosphatidylinositol (GPI) anchor and a C-terminal hydrophobic tail (TABLE 2). The N- and C-terminal domains are relatively invariant, but the number and sequence of the internal LRR motifs are variable. The *VLR* cDNAs seem to have arisen by recombination of LRR cassettes, which are found in different orientations in the germline locus and can be inserted into the coding frame of the active *VLR* gene, possibly through gene conversion. PCR amplification of genomic DNA and comparison of product lengths in erythrocytes and lymphocyte-like cells indicates that diversity in the VLR repertoire is derived by somatic recombination but is presumably RAG independent. The structural characteristics of VLRs, as well as cell-surface staining with VLR-specific antibodies, indicate that they are membrane bound. Disruption of the GPI linkage might function as the mechanism to release the membrane-bound VLR, in a manner that is generally reminiscent of lymphocyte-bound immunoglobulin. Single-cell PCR is consistent with haplotypic exclusion, a characteristic of lymphocyte-mediated immunity and a central feature of clonal selection in RAG-mediated adaptive immunity. VLRs that undergo similar genetic rearrangement have also been characterized in hagfish[99].

The postulated role of the VLRs as alternative mediators of adaptive immunity might have an even more ancient history within the broader functional roles of other metazoan LRR-containing proteins, which include the Toll-like receptor (TLR)[94], CATERPILLER (caspase-recruitment domain, transcription enhancer, R (purine)-binding, pyrin, lots of LRRs)[100] and Slit-homologue protein families[101,102]. The TLRs have particularly broad distributions in phylogeny and are closely involved in mammalian immune surveillance[103–105]. TLRs, which comprise one of the better-recognized functional relationships that links innate and adaptive immunity[12,106], recognize molecular patterns that mainly consist of conserved epitopes found on pathogens[6,107]. Pattern recognition by

TLRs does not rely on somatic change; however, there is increasing evidence for functionally advantageous polymorphisms that, for example, improve resistance to bacterial infection[108], as well as increasing evidence for combinatorial pairing[109]. Intracellular LRR-containing proteins that mediate recognition of intracellular infection, modulate signalling and regulate apoptosis have also been described[100,110]. Our preliminary phylogenetic analysis of lamprey and hagfish VLRs indicates that a single LRR protein might have been recruited from a pre-existing family; this protein subsequently would have undergone rapid evolution and divergence (to function as an immune mediator). The VLRs might be a derived TLR family that could supplant and/or complement the function of these receptors in jawless vertebrates[103].

Key questions regarding VLRs relate to the mechanisms and regulation of somatic-recombination events and to the clonal-selection characteristics of the lympho cyte-like cells that express VLRs. Notwithstanding the differences in terms of the mediators of adaptive immunity, a challenging and broader issue concerns the relationships between the VLR-expressing lympho cytes and those cells that express TCRs or BCRs and are distributed throughout the phylogeny of jawed vertebrates. A corollary issue is whether there are novel mechanisms for the generation of immune-receptor diversity in protochordates and other invertebrates.

VCBPs in protochordates

As indicated previously, the extant jawless vertebrates are highly derived, and despite the likelihood that they have an anticipatory adaptive immune system, their immune mediators might be more similar to other LRR-based systems than to the genes encoding immuno globulin domains. In this regard, inferences drawn from more BASAL species can sometimes provide more insights than those drawn from species that are more related to the jawed vertebrates in terms of phylogenetic position[111]: that is, although a bony fish shares a more recent common ancestor with humans than does a protochordate (FIG. 1), protochordates might still retain certain characteristics that are more closely related to those found in humans, as a consequence of evolutionary derivation. In this regard, the basal morphological and physiological characteristics of the CEPHALOCHORDATE amphioxus might shed light on links between innate immune recognition and immunoglobulin-type receptors that are related to those that mediate RAG-based adaptive immunity.

Early efforts to clone adaptive immune-like genes from amphioxus using methods such as cross hybridization and short-primer PCR were unsuccessful; however, a SECRETION-SIGNAL-PEPTIDE-SELECTION-based approach identified members of a multigene family that encode an immunoglobulin-type molecule (known as V-region-containing chitin-binding protein, VCBP) that consists of a pair of N-terminal V domains and a single C-terminal chitin-binding domain[7] (TABLE 2). Five families of VCBPs are distinguished by sequence differences in their V regions[7]. The N-terminal V regions of VCBP2 and VCBP5 are closely related, although the remainder of each molecule is divergent. The structure of the N-terminal V domain of a VCBP3 molecule[112] has been solved to 1.15 Å (J. Hernandez Prada, R. Haire and D. Ostrov, personal communication).

VCBPs — unlike BCRs, TCRs and some NITRs — lack characteristic J segments and are not membrane anchored. However, the N-terminal V domains of VCBP2 and VCBP5 contain a region of particularly extensive diversity[7,113], which is consistent with immune-type function. Systematic sampling of the local amphioxus population indicates an extra ordinary number of polymorphic alleles across this region of diversity that are inherited in a Mendelian manner (REF. 113, and N. Schnitker and L.J.D., unpublished observations). This hypervariable region probably has a function that is under strong selective pressure, such as

pathogen recognition[113]. VCBPs have also been identified in *Ciona* species[113], and the possibility that somatic variation occurs in VCBPs is being evaluated.

RNA *in situ* hybridization shows that VCBPs are expressed abundantly and (apparently) specifically in the gut of amphioxus, which is an important site of potential pathogen infiltration in this exclusive filter feeder. Amphioxus accounts for up to 70% of the total biomass in certain coastal waters. In such large populations, gene selection can occur at the level of the population rather than on an individual level[113–115], which is in contrast to conventional adaptive immune mediators, for which selection mainly occurs at the level of the somatic cell lineage within an organism. Sea water sustains high levels of bacteria, fungi and viruses and undergoes marked seasonal variation. Although speculative, the function of VCBPs might be to recognize determinants on various pathogens. There is considerable precedent for IgSF-related viral receptors, including those for poliovirus (poliovirus receptor)[116], reovirus (junctional adhesion molecule, JAM)[117], HIV (CD4)[118], mouse hepatitis virus (CEA-related cell-adhesion molecule 1, CEACAM1)[119], measles virus (signalling lymphocytic activation molecule, SLAM)[120] and others. Other potential functions for VCBPs include binding to fungal or bacterial surfaces, direct opsonization, and self- and non-self-recognition[3,7]. It is also possible that the V regions of VCBPs are pathogen specific and that the chitin-binding domains bind microbial *N*-acetylglucosamine, thereby inhibiting biofilm deposition, which is a key step in the colonization of a host by certain bacteria, through the formation of impenetrable barriers[121–123]. Variation in VCBP V regions and BCR and TCR V regions is concentrated in different relative positions and is achieved through different mechanisms, presenting a likely case for evolutionary convergence of the V domain in innate and adaptive immune functions, respectively.

Variation in invertebrate immune receptors

Similar to vertebrates, invertebrate species are under nearly continuous assault from various bacterial, fungal, protozoan, parasitic and viral pathogens. Traditionally, it was thought that several distinct, but relatively invariant, innate immune mediators countered such challenges in invertebrates. However, invertebrates have relatively extensive polymorphism of their immune-gene loci[115], and it is now recognized that they show memory in immune-type defence[124] and can probably mount various anticipatory immune responses[125]. A series of investigations in molluscs has provided insight into a previously unrecognized aspect of innate immune-receptor diversification and, together with a recent report on *D. melanogaster*[10] (discussed later), extends our consideration of ‘adaptive’ immunity beyond the vertebrates and protochordates.

In the snail *Biomphalaria glabrata*, a diverse family of IgSF-domain-containing fibrinogen-related proteins (FREPs) is secreted by circulatory haemocytes. The expression of FREPs is upregulated on infection with a trematode (a type of parasitic flatworm); FREPs then bind to trematode sporocysts and precipitate soluble parasite-derived products from solution[126]. At the molecular level, FREPs consist of one or two tandem IgSF domains at the N terminus, as well as a fibrinogen-like domain at the C terminus, and they have lectin-like activity[126,127] (TABLE 2). Fibrinogen is associated with blood coagulation in vertebrates; proteins that contain fibrinogen-like domains have previously been implicated in innate immune-defence functions in vertebrates and invertebrates[128]. FREPs, similar to the VCBPs in protochordates, are examples of proteins that have recruited more than one type of framework domain during the independent evolution of immune-defence pathways (discussed later).

Sequence variation in FREPs occurs in the IgSF and the fibrinogen-like domains[8]. Each individual snail can express various unique alleles within the different subfamilies. The

FREP alleles seem to be under strong selective pressure, and no single allele has come to dominate the population; instead, a diverse repertoire is maintained, which is potentially an advantage for an immune mediator. Diversity in the sequence encoding the N-terminal immunoglobulin-like domain of FREP3 is three- to fourfold higher than in control, unrelated genes. Taken together, a mathematical estimate and Southern-blot data indicate that a specific number of germline sequences are maintained across each subfamily and could account for all of the known variation, including nucleotide changes and diversification through recombination[8,129]. The observed diversity in *FREP3* cDNAs probably derives from somatic modification of *FREP3* genes[8]. It is unclear whether there is a component of cell-lineage specificity (and possibly selection) in FREPs that amplifies particular sequences from the pool of observed variants.

Several other recent examples show extraordinary mechanisms for generating somatic variation that lie outside the traditional concepts of molecular genetics and evolution. In contrast to the mutation process that is used by FREPs, other invertebrates probably use different methods to diversify immune-related molecules. The penaeidin genes, which encode antimicrobial peptides in crustaceans, occur in sets of relatively few loci, but they express an extensive repertoire[130]. Complex transcriptional processing of a lipopolysaccharide-inducible gene has been shown in sea urchins[131]. The possibilities for alternative spli cing increase markedly when exons have undergone tandem duplication[132], such as for the D. melanogaster Down's syndrome cell-adhesion molecule (*Dscam*) gene, which functions in complex neuronal processing by generating tens of thousands of protein variants from a single gene through differential RNA processing[132,133]. Both the complexity and expression patterns of Dscam isoforms in haemocytes indicate that Dscam might have a role in immune function[133,134]. Recent studies have defined differences in the expression of Dscam isoforms in haemocytes, fat cells and brain tissue[10]. Furthermore, secretion of specific Dscam isoforms has been confirmed at the protein level. *D. melanogaster* larvae carrying transallelic hypomorphic and amorphic mutations had reduced Dscam expression and less efficient phagocytosis of pathogenic bacteria, which was also achieved using Dscam-specific antibody and in RNA-interference studies. Directed binding of specific Dscam isoforms to bacteria was also shown, using soluble Dscam–immunoglobulin fusion proteins. Dscam-isoform variation is seen in multiple insect orders, which is consistent with it having an ancient origin. Collectively, these data underscore a role for complex RNA processing as another mechanism of immunological diversity. Other mechanisms can also introduce high levels of sequence diversity into receptor binding sites[135,136], for example, when accompanied by an adjacent retroelement.

Variation in immune receptors and pathogens

Phylogenetic studies of immune molecules clearly indicate that various mechanisms are used to diversify relatively limited amounts of genetic material to create a diverse set of receptor structures, the complexity of which is immense. The greater the amount of genetic variation that can be derived from a single gene or groups of genes, the more adaptable the host will be in responding to molecular variation shown by pathogens. The events that are likely to have occurred during the evolution of immunoglobulin and the TCR, as well as some of the other diverse immune receptors that are described here, show the successful integration of pathways that share important properties with those that are used to achieve genetic variation in viruses, bacteria, fungi and protozoan parasites. So, there is a co-evolutionary struggle in which selection acts on DNA in the host and in pathogens to mediate sequence variation and diversity for improving detection of infection in one case or for promoting evasion of immunity in the other (BOX 2).

What might be less appreciated is that the functional motifs (for example, LRRs) in the various receptors are related to those found in the pathogens themselves, including in protozoan parasites, bacteria (such as *Listeria monocytogenes* and group B *Streptococcus*) and fungi[137]. Over evolutionary time, a large number of effector molecules, including those that mediate innate and adaptive immune responses, have been created using frameworks of immunoglobulin, lectin and/or LRR domains (FIG. 3). Molecules such as VCBPs, FREPs and Dscam have integrated immunoglobulin and lectin domains that might coordinately function in immune recognition. Molecules that are associated with these structures can recognize and interact with diverse arrays of protein and/or carbohydrate patterns to mediate complex signalling events and effector functions[138]. Some pathogens have recruited these framework structures into some of their effector molecules, in effect creating a case in which selection acts on similar structures in the genetic repertoire of the host and the pathogen (FIG. 3).

Concluding remarks

Initially, adaptive and innate immunity were thought to be temporally separable processes involving different cell types. However, as more details have been learned about both forms of immunity, a growing awareness is emerging of a single, integrated process in which innate and adaptive effectors interact coordinately to counter a wide range of invasive pathological challenges[6]. Whereas innate immunity is generally considered to be the more phylogenetically ancient, this system has diversified and refined its functions over evolutionary time such that it persists in multiple forms and shows lineage-specific diversity. In this light, it is not unexpected that innate immune surveillance and recognition in modern organisms can achieve sophisticated discrimination between the ‘signatures’ of pathogens and those of non-pathogens or self-molecules. Indeed, a combination of specificity and degeneracy in recognition is common to innate and adaptive immune receptors[14]. The evolutionary success of the animal (as well as the plant) phyla outside the jawed vertebrates that do not undergo V(D)J recombination of immunoglobulin and TCR genes is a testament to the ability of non-RAG-mediated systems to defend the host against pathogenic threats.

Studies in various model systems allow an integrated view of adaptive and innate immune function. The dependence of mammalian immunity on T- and B-cell function and, in turn, the dependence of these cell types on RAG-mediated somatic recombination and subsequent diversification of antigen-receptor genes has understandably resulted in a focus on lymphocytes as the central components of mammalian immunity. However, increasing attention has recently been paid to the interactions of T and B cells with innate immune effectors, and this has produced important insights regarding functional bridges between innate and adaptive immune cells during infection[11–16,139,140]. An increasingly rich picture of the diversity of innate immune surveillance and recognition in vertebrates is developing. So, although immunity in modern mammals heavily depends on T and B cells, the activation of these cells and their subsequent roles in elimination of infection depend on their interactions with cells that express innate immune receptors, specifically dendritic cells and other antigen-presenting cells, as well as on the influences of a myriad of cytokines and soluble factors, such as the interleukins, interferons and complement.

The recent discoveries of markedly diversified immune-type genes in species that lack RAG-mediated adaptive immune systems have blurred the distinction between purely innate and purely adaptive immune receptors. The germline-diversified V domains in NITRs are an example of an adaptive immune feature in an innate immune receptor[73,74], whereas germline joining of immunoglobulin-gene clusters in cartilaginous fish is an example of an innate immune feature in an adaptive immune receptor[18]. The Dscam variants in insects[10], the FREPs in snails[8,126], the VLRs in the jawless vertebrates lampreys[9] and hagfish[99], and

the VCBPs in the protochordate amphioxus[7,113] all share certain fundamental characteristics with the adaptive antigen-binding immune receptors of jawed vertebrates.

Our historical considerations of the evolution of adaptive immunity have centred on the origins of the adaptive antigen-binding immune receptors themselves, as well as the MHC class I and class II molecules on which T-cell function depends so closely and the origins in phylogeny of which seem to coincide with the emergence of the antigen-binding receptors. Many efforts have been undertaken to label 'primordial' antigen-binding receptors in species outside the jawed vertebrates, under the presumption that such molecules exist and would appear in a recognizable form today. However, if one takes a more integrated view of the evolution of immunity as a whole and considers the wealth of diversity and discrimination that are available through using innate-type immune receptors, such intense focus on a single arm of a vast system is beginning to seem misplaced. A more realistic consideration of innate immune receptors might be necessary to understand the origins of adaptive immunity, particularly in cases in which modern genes have been assembled from 'parts' of other genes.

A wealth of examples of potential evolutionary vestiges occur in the modern mammalian immune system, including B1 cells and natural IgM, γδ T cells, 'invariant' TCRs, NK cells and their sets of receptors, serum lectins and the various complement pathways. Immune responses in contemporary jawed vertebrates have integrated these diverse systems and cellular interactions to achieve a collective goal of host defence. As these systems are numerous, diverse and under constant evolutionary pressure from the adaptation of pathogens, it is probable that following the lineage of one particular immune effector (even immunoglobulin or the TCR itself) in phylogenetically ancient species will provide only a small part of a tremendously complex picture. It is possible that natural selection has produced (and even eliminated) innovative immune-type genes continuously over evolutionary time and that many such genes will be uncovered as additional non-traditional species are investigated. This view is supported by the diversity of immune mechanisms that are now being discovered in a lineage- and taxon-dependent manner. The successive steps in the evolution of integrated immune complexity might never be known; however, it is probable that unravelling the networks that guide host defence in disparate phyla will continue to provide valuable information that can direct us in approaching and manipulating the complex mechanisms that effect immune function in multicellular organisms.

Acknowledgments

We thank B. Pryor for editorial assistance. We also thank C. Amemiya, M. Cooper, S. Desai, Z. Pancer and J. Yoder for providing valuable comments regarding the subject matter of this Review, and M. Anderson, A. De Tomaso, D. Eason, S. Fugmann, R. Haire, J. Hernandez Prada, N. Miller, D. Ostrov, J. Rast and N. Schnitker for sharing unpublished results. M. Sexton assisted with the preliminary graphic design. G.W.L. is supported by grants from the National Institutes of Health (United States). L.J.D. is supported by a fellowship from the H. Lee Moffitt Cancer Center & Research Institute (United States).

Glossary

METAZOANS	All animals that are above the phylogenetic level of protozoans and sponges.
TRANSPOSITION	Movement of a segment of DNA from one position in the genome to another position (or to a different genome).
GERMLINE DIVERSIFICATION	Changes in DNA sequence that occur in the germline of a species and are heritable.

SOMATIC DIVERSIFICATION	Changes in DNA sequence that occur in individual cells and their progeny. Traditionally, this process has been associated with lymphocytes and is brought about during gene rearrangements, as well as through gene conversion and somatic hypermutation.
JUNCTIONAL VARIATION	The introduction of genetic changes at the joining interfaces of rearranged variable (V), diversity (D) and joining (J) segmental elements in immunoglobulins and T-cell receptors.
GENE CONVERSION	A homologous-recombination event in which the donor gene(s) remains unmodified and an acceptor gene acquires the recombined segment. In chickens, variable (V)-region pseudogenes are donors that modify the functional, rearranged V gene in bursal follicles, and this process generates a diverse pre-immune repertoire.
CLASS-SWITCH RECOMBINATION	The process by which the heavy-chain class that is associated with a specific VDJ rearrangement is changed through recombination during B-cell ontogeny.
DERIVED	Reflecting characteristics that are acquired subsequent to divergence from a common ancestral form.
HISTOCOMPATIBILITY REACTION	The reaction to, and/or rejection of, individual cells or tissues that are introduced from one individual to a genetically different individual.
RECOMBINATION SIGNAL SEQUENCE	A conserved genetic element that constitutes a recognition site for the V(D)J recombinase proteins, which are encoded by the genes recombination-activating gene 1 (RAG1) and RAG2. These sites consist of a palindromic 7-base-pair sequence that is immediately adjacent to the coding gene segments — variable (V), diversity (D) or joining (J) — and is separated by a 12- or 23-base-pair spacer from a relatively conserved 9-base-pair sequence.
NON-HOMOLOGOUS END JOINING	A generic mechanism for repair of double-stranded DNA breaks. It joins hairpin coding ends that form during V(D)J recombination.
COMPLEMENTARITY-DETERMINING REGIONS	(CDRs). The most variable parts of immunoglobulin and the T-cell receptor. These form loops that make contact with specific ligands. There are three such regions (CDR1, CDR2 and CDR3) in each variable domain.
ALLELIC EXCLUSION	A mechanism by which antigen receptors of a single specificity are expressed at the cell surface of a lymphocyte. This is an integral step in the clonal commitment of a cell lineage.

TYPE I TRANSMEMBRANE PROTEIN	An integral membrane protein that is composed of an amino-terminal extracellular region, a transmembrane domain and a carboxy-terminal intracellular region (for example, T-cell receptors).
TYPE II TRANSMEMBRANE PROTEIN	An integral membrane protein that is composed of an amino-terminal intracellular region, a transmembrane domain and a carboxy-terminal extracellular region (for example, Ly49).
IMMUNORECEPTOR TYROSINE-BASED INHIBITORY MOTIF	(ITIM). A structural motif that contains tyrosine residues and is found in the cytoplasmic tails of several inhibitory receptors, such as the low-affinity Fc receptor for IgG FcγRIIB and paired immunoglobulin-like receptor B (PIRB). The prototype six-amino-acid ITIM sequence is (I/V/L/S)XYXX(L/V), where X denotes any amino acid. Ligand-induced clustering of these inhibitory receptors results in tyrosine phosphorylation, often by SRC-family protein tyrosine kinases, and this provides a docking site for the recruitment of cytoplasmic phosphatases that have a SRC homology 2 (SH2) domain.
IMMUNORECEPTOR TYROSINE-BASED ACTIVATION MOTIF	(ITAM). B-cell, T-cell and natural-killer-cell receptors are non-covalently associated with transmembrane proteins that contain one or more ITAMs. The amino-acid sequence of an ITAM is (D/E)XXYXX(L/I)$X_{6–8}$YXX(L/I), where X denotes any amino acid. This is tyrosine phosphorylated after engagement of the ligand-binding subunits, which triggers a cascade of intracellular events that results in cellular activation.
IMMUNORECEPTOR TYROSINE-BASED SWITCH MOTIF	(ITSM). This motif has the amino-acid sequence TXYXX(V/I), where X denotes any amino acid. It recruits many of the same signalling molecules as do immunoreceptor tyrosine-based inhibitory motifs (ITIMs) and immunoreceptor tyrosine-based activation motifs (ITAMs), but it also recruits SAP (signalling lymphocytic activation molecule (SLAM)-associated protein).
ORTHOLOGUES	Similar sequences that are found in different species and can be attributed to descent from a common ancestral sequence.
SYNTENY	The same order of genes occurring on chromosomes that are present in different organisms.
BIRTH AND DEATH	In the context of multigene families encoding immune molecules, this involves the creation and extinction of alleles, and it is co-driven by pathogen load during evolution.
AGGLUTININ	A soluble protein that occurs naturally or in response to immunization and agglutinates a particulate antigen.

	The term can be applied to an immunoglobulin molecule or to proteins that are unrelated to immunoglobulins.
EXPRESSED SEQUENCE TAG	(EST). A genetic sequence that corresponds to an mRNA or a region of an mRNA.
LEUCINE-RICH REPEAT	(LRR). Domains that contain LRRs have a conserved solenoid structure, typically of 20–29 residues and containing an 11 amino-acid consensus sequence, LXXLXLXX(N/C)XL, where X denotes any amino acid. These domains lack considerable identity or similarity in the amino acids surrounding this structure, both between and among families. Sequence substitutions in LRR-containing proteins are associated with changes in specificity and relative affinity towards specific determinants.
BASAL	Reflecting the shared characteristic(s) of a common ancestral form.
CEPHALOCHORDATE	A member of the subphylum Cephalochordata. In this Review, cephalochordates are represented by amphioxus, which has a notochord, dorsal nerve cord and pharyngeal gills that are well developed in the adult stage.
SECRETION SIGNAL-PEPTIDE SELECTION	A cloning method that is based on the rescue of ampicillin resistance in a vector that encodes a defective β-lactamase, by the substitution of an exogenous leader, as well as additional coding sequence, from transmembrane or secreted proteins.

References

1. Agrawal A, Eastman QM, Schatz DG. Transposition mediated by RAG1 and RAG2 and its implications for the evolution of the immune system. Nature. 1998; 394:744–751. [PubMed: 9723614]
2. Hiom K, Melek M, Gellert M. DNA transposition by the RAG1 and RAG2 proteins: a possible source of oncogenic translocations. Cell. 1998; 94:463–470. [PubMed: 9727489]
3. Cannon JP, Haire RN, Rast JP, Litman GW. The phylogenetic origins of the antigen binding receptors and somatic diversification mechanisms. Immunol. Rev. 2004; 200:12–22. [PubMed: 15242392]
4. Chatterji M, Tsai CL, Schatz DG. New concepts in the regulation of an ancient reaction: transposition by RAG1/RAG2. Immunol. Rev. 2004; 200:261–271. [PubMed: 15242411] A comprehensive review of RAG-mediated transposition and its role in the origin of adaptive immunity.
5. Litman GW, Cannon JP, Rast JP. New insights into mechanisms of immune receptor diversification. Adv. Immunol. 2005; 87:209–236. [PubMed: 16102575]
6. Beutler B. Innate immunity: an overview. Mol. Immunol. 2004; 40:845–859. [PubMed: 14698223] A comprehensive review of the wide range of molecules that mediate innate immunity.
7. Cannon JP, Haire RN, Litman GW. Identification of diversified immunoglobulin-like variable region-containing genes in a protochordate. Nature Immunol. 2002; 3:1200–1207. [PubMed: 12415263] The original description of the immune-like VCBPs in the protochordate amphioxus.

8. Zhang S-M, Adema CM, Kepler TB, Loker ES. Diversification of Ig superfamily genes in an invertebrate. Science. 2004; 305:251–254. [PubMed: 15247481] The original description of the role of somatic hypermutation in host defence in an invertebrate.
9. Pancer Z, et al. Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey. Nature. 2004; 430:174–180. [PubMed: 15241406] The original description of the VLRs as a somatically variable family of LRR-containing immune-type receptors.
10. Watson FL, et al. Extensive diversity of Ig-superfamily proteins in the immune system of insects. Science. 2005; 309:1874–1878. [PubMed: 16109846] A definitive study showing that complex RNA processing is a mechanism of immune-type-receptor diversification.
11. Schnare M, et al. Toll-like receptors control activation of adaptive immune responses. Nature Immunol. 2001; 2:947–950. [PubMed: 11547333]
12. Hoebe K, Janssen E, Beutler B. The interface between innate and adaptive immunity. Nature Immunol. 2004; 5:971–974. [PubMed: 15454919]
13. Kadowaki N, et al. Subsets of human dendritic cell precursors express different Toll-like receptors and respond to different microbial antigens. J. Exp. Med. 2001; 194:863–869. [PubMed: 11561001]
14. Vivier E, Malissen B. Innate and adaptive immunity: specificities and signaling hierarchies revisited. Nature Immunol. 2005; 6:17–21. [PubMed: 15611777]
15. Molina H, et al. Markedly impaired humoral immune response in mice deficient in complement receptors 1 and 2. Proc. Natl Acad. Sci. USA. 1996; 93:3357–3361. [PubMed: 8622941]
16. Fischer MB, et al. Regulation of the B cell response to T-dependent antigens by classical pathway complement. J. Immunol. 1996; 157:549–556. [PubMed: 8752901]
17. Hinds KR, Litman GW. Major reorganization of immunoglobulin V_H segmental elements during vertebrate evolution. Nature. 1986; 320:546–549. [PubMed: 3083268]
18. Kokubu F, Litman R, Shamblott MJ, Hinds K, Litman GW. Diverse organization of immunoglobulin V_H gene loci in a primitive vertebrate. EMBO J. 1988; 7:3413–3422. [PubMed: 3145194] References 17and 18 are the original descriptions of the multiple cluster-type immunoglobulin heavy-chain genes that are distributed (together with the light-chain genes) throughout the cartilaginous fish. The extensive germline joining of immunoglobulin genes creates an innate-type immune receptor from a rearranging adaptive immune receptor.
19. Thompson CB, Nieman PE. Somatic diversification of the chicken immunoglobulin light chain gene is limited to the rearranged variable gene segment. Cell. 1987; 48:369–378. [PubMed: 3100049]
20. Reynaud C-A, Anquez V, Grimal H, Weill J-C. A hyperconversion mechanism generates the chicken light chain preimmune repertoire. Cell. 1987; 48:379–388. [PubMed: 3100050]
21. Reynaud C-A, Dahan A, Anquez V, Weill J-C. Somatic hyperconversion diversifies the single V_H gene of the chicken with a high incidence in the D region. Cell. 1989; 59:171–183. [PubMed: 2507167] References 19–21 establish another case for extreme differences in BCR organization between species and establish gene conversion as a mechanism for generating complex BCR repertoires.
22. Knight KL. Restricted V_H gene usage and generation of antibody diversity in rabbit. Annu. Rev. Immunol. 1992; 10:593–616. [PubMed: 1590997]
23. Butler JE. Disparate mechanisms drive antibody diversity among mammals: a useful addition to immunology textbooks. Curr. Trends Immunol. 2003; 5:1–18.
24. Kondo T, Arakawa H, Kitao H, Hirota Y, Yamagishi H. Signal joint of immunoglobulin Vλ1–Jλ and novel joints of chimeric V pseudogenes on extrachromosomal circular DNA from chicken bursa. Eur. J. Immunol. 1993; 23:245–249. [PubMed: 8419176]
25. Shamblott MJ, Litman GW. Genomic organization and sequences of immunoglobulin light chain genes in a primitive vertebrate suggest coevolution of immunoglobulin gene organization. EMBO J. 1989; 8:3733–3739. [PubMed: 2511000]
26. Ghaffari SH, Lobb CJ. Structure and genomic organization of immunoglobulin light chain in the channel catfish. J. Immunol. 1993; 151:6900–6912. [PubMed: 8258698]

27. Ota T, Rast JP, Litman GW, Amemiya CT. Lineage-restricted retention of a primitive immunoglobulin heavy chain isotype within the Dipnoi reveals an evolutionary paradox. Proc. Natl Acad. Sci. USA. 2003; 100:2501–2506. [PubMed: 12606718]

28. Litman GW, Anderson MK, Rast JP. Evolution of antigen binding receptors. Annu. Rev. Immunol. 1999; 17:109–147. [PubMed: 10358755]

29. Danilova N, Bussmann J, Jekosch K, Steiner LA. The immunoglobulin heavy-chain locus in zebrafish: identification and expression of a previously unknown isotype, immunoglobulin Z. Nature Immunol. 2005; 6:295–302. [PubMed: 15685175]

30. Bengten E, et al. The IgH locus of the channel catfish, *Ictalurus punctatus*, contains multiple constant region gene sequences: different genes encode heavy chains of membrane and secreted IgD. J. Immunol. 2002; 169:2488–2497. [PubMed: 12193718]

31. Hansen JD, Landis ED, Phillips RB. Discovery of a unique Ig heavy-chain isotype (IgT) in rainbow trout: implications for a distinctive B cell developmental pathway in teleost fish. Proc. Natl Acad. Sci. USA. 2005; 102:6919–6924. [PubMed: 15863615]

32. Zarrin AA, et al. An evolutionarily conserved target motif for immunoglobulin class-switch recombination. Nature Immunol. 2004; 5:1275–1281. [PubMed: 15531884]

33. Stavnezer J, Amemiya CT. Evolution of isotype switching. Semin. Immunol. 2004; 16:257–275. [PubMed: 15522624]

34. Jaillon O, et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. Nature. 2005; 431:946–957. [PubMed: 15496914]

35. Okamura K, Ototake M, Nakanishi T, Kurosawa Y, Hashimoto K. The most primitive vertebrates with jaws possess highly polymorphic MHC class I genes comparable to those of humans. Immunity. 1997; 7:777–790. [PubMed: 9430223]

36. Kasahara M, McKinney EC, Flajnik MF, Ishibashi T. The evolutionary origin of the major histocompatibility complex: polymorphism of class II α-chain genes in the cartilaginous fish. Eur. J. Immunol. 1993; 23:2160–2165. [PubMed: 8370399]

37. Bartl S, Weissman IL. Isolation and characterization of major histocompatibility complex class IIB genes from the nurse shark. Proc. Natl Acad. Sci. USA. 1994; 91:262–266. [PubMed: 8278377]

38. Flajnik MF, Kasahara M. Comparative genomics of the MHC: glimpses into the evolution of the adaptive immune system. Immunity. 2001; 15:351–362. [PubMed: 11567626]

39. Shin S, et al. Antigen recognition determinants of γδ T cell receptors. Science. 2005; 308:252–255. [PubMed: 15821090] References 35–39 describe remarkable similarities in overall structure and complexity of polymorphism of the MHC class I and class II genes between higher vertebrates and cartilaginous fish.

40. Vogel C, Teichmann SA, Chothia C. The immunoglobulin superfamily in *Drosophila melanogaster* and *Caenorhabditis elegans* and the evolution of complexity. Development. 2003; 130:6317–6328. [PubMed: 14623821]

41. Streltsov VA, et al. Structural evidence for evolution of shark Ig new antigen receptor variable domain antibodies from a cell-surface receptor. Proc. Natl Acad. Sci. USA. 2004; 101:12444–12449. [PubMed: 15304650]

42. Stanfield RL, Dooley H, Flajnik MF, Wilson IA. Crystal structure of a shark single-domain antibody V region in complex with lysozyme. Science. 2004; 305:1770–1773. [PubMed: 15319492]

43. Eason DD, et al. Mechanisms of antigen receptor evolution. Semin. Immunol. 2004; 16:215–226. [PubMed: 15522620]

44. Kapitonov VV, Jurka J. RAG1 core and V(D)J recombination signal sequences were derived from *Transib* transposons. PLoS Biol. 2005; 3:e181. [PubMed: 15898832]

45. Ito J, Braithwaite DK. Compilation and alignment of DNA polymerase sequences. Nucleic Acids Res. 1991; 19:4045–4057. [PubMed: 1870963]

46. Jung D, Alt FW. Unraveling V(D)J recombination; insights into gene regulation. Cell. 2004; 116:299–311. [PubMed: 14744439] A comprehensive review of the mechanisms of segmental rearrangement of immunoglobulin genes and the genetic regulation of this process.

47. Muramatsu M, et al. Class switch recombination and hypermutation require activation-induced cytidine deaminase (AID), a potential RNA editing enzyme. Cell. 2000; 102:553–563. [PubMed: 11007474]
48. Barreto VM, et al. AID from bony fish catalyzes class switch recombination. J. Exp. Med. 2005; 202:733–738. [PubMed: 16157688]
49. Bishop KN, et al. Cytidine deamination of retroviral DNA by diverse APOBEC proteins. Curr. Biol. 2004; 14:1392–1396. [PubMed: 15296758]
50. Esnault C, et al. APOBEC3G cytidine deaminase inhibits retrotransposition of endogenous retroviruses. Nature. 2005; 433:430–433. [PubMed: 15674295]
51. Sohail A, Klapacz J, Samaranayake M, Ullah A, Bhagwat AS. Human activation-induced cytidine deaminase causes transcription-dependent, strand-biased C to U deaminations. Nucleic Acids Res. 2003; 31:2990–2994. [PubMed: 12799424]
52. Hinds-Frey KR, Nishikata H, Litman RT, Litman GW. Somatic variation precedes extensive diversification of germline sequences and combinatorial joining in the evolution of immunoglobulin heavy chain diversity. J. Exp. Med. 1993; 178:825–834. [PubMed: 8350056]
53. Lee SS, Tranchina D, Ohta Y, Flajnik MF, Hsu E. Hypermutation in shark immunoglobulin light chain genes results in contiguous substitutions. Immunity. 2002; 16:571–582. [PubMed: 11970880]
54. Wang CL, Harper RA, Wabl M. Genome-wide somatic hypermutation. Proc. Natl Acad. Sci. USA. 2004; 101:7352–7356. [PubMed: 15123833]
55. Wang L, Jackson WC, Steinbach PA, Tsien RY. Evolution of new nonantibody proteins via iterative somatic hypermutation. Proc. Natl Acad. Sci. USA. 2004; 101:16745–16749. [PubMed: 15556995]
56. Eason DD, Litman RT, Luer CA, Kerr W, Litman GW. Expression of individual immunoglobulin genes occurs in an unusual system consisting of multiple independent loci. Eur. J. Immunol. 2004; 34:2551–2558. [PubMed: 15307187]
57. Vassar R, Ngai J, Axel R. Spatial segregation of odorant receptor expression in the mammalian olfactory epithelium. Cell. 1993; 74:309–318. [PubMed: 8343958]
58. Serizawa S, et al. Mutually exclusive expression of odorant receptor transgenes. Nature Neurosci. 2000; 3:687–693. [PubMed: 10862701]
59. Esumi S, et al. Monoallelic yet combinatorial expression of variable exons of the protocadherin-α gene cluster in single neurons. Nature Genet. 2005; 37:171–176. [PubMed: 15640798]
60. Marchalonis JJ, Schluter SF, Bernstein RM, Shen S, Edmundson AB. Phylogenetic emergence and molecular evolution of the immunoglobulin family. Adv. Immunol. 1998; 70:417–506. [PubMed: 9755343]
61. Vilches C, Parham P. KIR: diverse, rapidly evolving receptors of innate and adaptive immunity. Annu. Rev. Immunol. 2002; 20:217–251. [PubMed: 11861603]
62. Parham P. MHC class I molecules and KIRs in human history, health and survival. Nature Rev. Immunol. 2005; 5:201–214. [PubMed: 15719024]
63. Cosman D, et al. ULBPs, novel MHC class I-related molecules, bind to CMV glycoprotein UL16 and stimulate NK cytotoxicity through the NKG2D receptor. Immunity. 2001; 14:123–133. [PubMed: 11239445]
64. Braud VM, et al. HLA-E binds to natural killer cell receptors CD94/NKG2A, B and C. Nature. 1998; 391:795–797. [PubMed: 9486650]
65. Arase H, Mocarski ES, Campbell AE, Hill AB, Lanier LL. Direct recognition of cytomegalovirus by activating and inhibitory NK cell receptors. Science. 2002; 296:1323–1326. [PubMed: 11950999]
66. Smith HRC, et al. Recognition of a virus-encoded ligand by a natural killer cell activation receptor. Proc. Natl Acad. Sci. USA. 2002; 99:8826–8831. [PubMed: 12060703]
67. Iizuka K, Naidenko OV, Plougastel BFM, Fremont DH, Yokoyama WM. Genetically linked C-type lectin-related ligands for the NKRP1 family of natural killer cell receptors. Nature Immunol. 2003; 4:801–805. [PubMed: 12858173]

68. Cossarizza A, et al. Earthworm leukocytes that are not phagocytic and cross-react with several human epitopes can kill human tumor cell lines. Exp. Cell Res. 1996; 224:174–182. [PubMed: 8612683]
69. Khalturin K, Becker M, Rinkevich B, Bosch TCG. Urochordates and the origin of natural killer cells: identification of a CD94/NKR-P1-related receptor in blood cells of *Botryllus*. Proc. Natl Acad. Sci. USA. 2003; 100:622–627. [PubMed: 12518047]
70. Sato A, Mayer WE, Overath P, Klein J. Genes encoding putative natural killer cell C-type lectin receptors in teleostean fishes. Proc. Natl Acad. Sci. USA. 2003; 100:7779–7784. [PubMed: 12802013]
71. Prouty MG, Correa NE, Barker LP, Jagadeeswaran P, Klose KE. Zebrafish-*Mycobacterium marinum* model for mycobacterial pathogenesis. FEMS Microbiol. Lett. 2003; 225:177–182. [PubMed: 12951238]
72. Zelensky AN, Gready JE. C-type lectin-like domains in *Fugu rubripes*. BMC Genomics [online]. 2004; 5:51.
73. Strong SJ, et al. A novel multigene family encodes diversified variable regions. Proc. Natl Acad. Sci. USA. 1999; 96:15080–15085. [PubMed: 10611341]
74. Yoder JA, et al. Resolution of the NITR gene cluster in zebrafish. Proc. Natl Acad. Sci. USA. 2004; 101:15706–15711. [PubMed: 15496470] . References 73 and 74 describe NITR-gene clusters and the multiple structural forms of NITRs.
75. Sidorenko SP, Clark EA. The dual-function CD150 receptor subfamily: the viral attraction. Nature Immunol. 2003; 4:19–24. [PubMed: 12496974]
76. Dietrich J, Nakajima H, Colonna M. Human inhibitory and activating Ig-like receptor which modulates the function of myeloid cells. Microbes Infect. 2000; 2:323–329. [PubMed: 10758410]
77. Hawke NA, et al. Extraordinary variation in a diversified family of immune-type receptor genes. Proc. Natl Acad. Sci. USA. 2001; 98:13832–13837. [PubMed: 11698645]
78. Shen L, et al. Identification and characterization of clonal NK-like cells from channel catfish (*Ictalurus punctatus*). Dev. Comp. Immunol. 2004; 28:139–152. [PubMed: 12969799]
79. Ota T, Nei M. Divergent evolution and evolution by the birth-and-death process in the immunoglobulin V_H gene family. Mol. Biol. Evol. 1994; 11:469–482. [PubMed: 8015440]
80. Litman GW, Hawke NA, Yoder JA. Novel immune-type receptor genes. Immunol. Rev. 2001; 181:250–259. [PubMed: 11513146]
81. Pende D, et al. Identification and molecular characterization of NKp30, a novel triggering receptor involved in natural cytotoxicity mediated by human natural killer cells. J. Exp. Med. 1999; 190:1505–1516. [PubMed: 10562324]
82. Cantoni C, et al. NKp44, a triggering receptor involved in tumor cell lysis by activated human natural killer cells, is a novel member of the immunoglobulin superfamily. J. Exp. Med. 1999; 189:787–796. [PubMed: 10049942]
83. Suzuki T, Shin I, Fujiyama A, Kohara Y, Kasahara M. Hagfish leukocytes express a paired receptor family with a variable domain resembling those of antigen receptors. J. Immunol. 2005; 174:2885–2891. [PubMed: 15728499]
84. Good, RA.; Finstad, J.; Litman, GW. The Biology of Lampreys II: Immunology. Hardisty, MV.; Potter, IC., editors. Academic; London: 1972. p. 405-432.
85. Uinuk-Ool T, et al. Lamprey lymphocyte-like cells express homologs of genes involved in immunologically relevant activities of mammalian lymphocytes. Proc. Natl Acad. Sci. USA. 2002; 99:14356–14361. [PubMed: 12391333]
86. Mayer WE, et al. Isolation and characterization of lymphocyte-like cells from a lamprey. Proc. Natl Acad. Sci. USA. 2002; 99:14350–14355. [PubMed: 12388781] A definitive description of lymphocytes in a jawless vertebrate.
87. Pancer Z, Mayer WE, Klein J, Cooper MD. Prototypic T cell receptor and CD4-like coreceptor are expressed by lymphocytes in the agnathan sea lamprey. Proc. Natl Acad. Sci. USA. 2004; 101:13273–13278. [PubMed: 15328402]
88. Haire RN, Miracle AL, Rast JP, Litman GW. Members of the Ikaros gene family are present in early representative vertebrates. J. Immunol. 2000; 165:306–312. [PubMed: 10861066]

89. Shintani S, et al. Do lampreys have lymphocytes? The Spi evidence. Proc. Natl Acad. Sci. USA. 2000; 97:7417–7422. [PubMed: 10840049]

90. Anderson MK, Sun X, Miracle AL, Litman GW, Rothenberg EV. Evolution of hematopoiesis: three members of the PU.1 transcription factor family in a cartilaginous fish, *Raja eglanteria*. Proc. Natl Acad. Sci. USA. 2001; 98:553–558. [PubMed: 11149949]

91. Cupit PM, et al. Ikaros family members from the agnathan *Myxine glutinosa* and the urochordate *Oikopleura dioica*: emergence of an essential transcription factor for adaptive immunity. J. Immunol. 2003; 171:6006–6013. [PubMed: 14634112]

92. Melchers F, et al. Repertoire selection by pre-B-cell receptors and B-cell receptors, and genetic control of B-cell development from immature to mature B cells. Immunol. Rev. 2000; 175:33–46. [PubMed: 10933589]

93. Cannon JP, et al. Variable domains and a VpreB-like molecule are present in a jawless vertebrate. Immunogenetics. 2005; 56:924–929. [PubMed: 15650874]

94. Bell JK, et al. Leucine-rich repeats and pathogen recognition in Toll-like receptors. Trends Immunol. 2003; 24:528–533. [PubMed: 14552836]

95. Girardin SE, Sansonetti PJ, Philpott DJ. Intracellular vs extracellular recognition of pathogens — common concepts in mammals and flies. Trends Microbiol. 2002; 10:193–199. [PubMed: 11912027]

96. Kobe B, Kajava AV. The leucine-rich repeat as a protein recognition motif. Curr. Opin. Struct. Biol. 2001; 11:725–732. [PubMed: 11751054]

97. Hayes AJ, et al. Recombination within a nucleotide-binding-site/leucine-rich-repeat gene cluster produces new variants conditioning resistance to soybean mosaic virus in soybeans. Genetics. 2004; 166:493–503. [PubMed: 15020438]

98. Mondragon-Palomino M, Meyers BC, Michelmore RW, Gaut BS. Patterns of positive selection in the complete NBS-LRR gene family of *Arabidopsis thaliana*. Genome Res. 2002; 12:1305–1315. [PubMed: 12213767]

99. Pancer Z, et al. Variable lymphocyte receptors in hagfish. Proc. Natl Acad. Sci. USA. 2005; 102:9224–9229. [PubMed: 15964979]

100. Ting JP, Davis BK. CATERPILLER: a novel gene family important in immunity, cell death, and diseases. Annu. Rev. Immunol. 2005; 23:387–414. [PubMed: 15771576]

101. Taguchi A, et al. Molecular cloning of novel leucine-rich repeat proteins and their expression in the developing mouse nervous system. Brain Res. Mol. Brain Res. 1996; 35:31–40. [PubMed: 8717337]

102. Wong K, Park HT, Wu JY, Rao Y. Slit proteins: molecular guidance cues for cells ranging from neurons to leukocytes. Curr. Opin. Genet. Dev. 2002; 12:583–591. [PubMed: 12200164]

103. Roach JC, et al. The evolution of vertebrate Toll-like receptors. Proc. Natl Acad. Sci. USA. 2005; 102:9577–9582. [PubMed: 15976025]

104. Cook DN, Pisetsky DS, Schwartz DA. Toll-like receptors in the pathogenesis of human disease. Nature Immunol. 2004; 5:975–979. [PubMed: 15454920]

105. Philpott DJ, Girardin SE. The role of Toll-like receptors and Nod proteins in bacterial infection. Mol. Immunol. 2004; 41:1099–1108. [PubMed: 15476921]

106. Iwasaki A, Medzhitov R. Toll-like receptor control of the adaptive immune responses. Nature Immunol. 2004; 5:987–995. [PubMed: 15454922]

107. Kopp E, Medzhitov R. Recognition of microbial infection by Toll-like receptors. Curr. Opin. Immunol. 2003; 15:396–401. [PubMed: 12900270]

108. Hawn TR, et al. Toll-like receptor 4 polymorphisms are associated with resistance to Legionnaires' disease. Proc. Natl Acad. Sci. USA. 2005; 102:2487–2489. [PubMed: 15699327]

109. Liew FY, Xu D, Brint EK, O'Neill AJ. Negative regulation of Toll-like receptor-mediated immune responses. Nature Rev. Immunol. 2005; 5:446–458. [PubMed: 15928677]

110. Inohara N, Chamaillard M, McDonald C, Nunez G. NOD-LRR proteins: role in host–microbial interactions and inflammatory disease. Annu. Rev. Biochem. 2005; 74:355–383. [PubMed: 15952891]

111. Kusserow A, et al. Unexpected complexity of the *Wnt* gene family in a sea anemone. Nature. 2005; 433:156–160. [PubMed: 15650739]
112. Hernandez Prada JA, Haire RN, Cannon JP, Litman GW, Ostrov DA. Crystallization and preliminary X-ray analysis of VCBP3 from *Branchiostoma floridae*. Acta Crystallogr. D Biol. Crystallogr. 2004; 60:2022–2024. [PubMed: 15502315]
113. Cannon JP, Haire RN, Schnitker N, Mueller MG, Litman GW. Individual protochordates possess unique immune-type receptor repertoires. Curr. Biol. 2004; 14:R465–R466. [PubMed: 15203016] A description of a high degree of interindividual sequence polymorphism in VCBPs in the protochordate amphioxus.
114. Kwiatkowski DP. The complexity of genetic variation in a simple immune system. Trends Genet. 2005; 21:197–199. [PubMed: 15797611]
115. Lazzaro BP, Sceurman BK, Clark AG. Genetic basis of natural variation in *D. melanogaster* antibacterial immunity. Science. 2004; 303:1873–1876. [PubMed: 15031506]
116. Mendelsohn CL, Wimmer E, Racaniello VR. Cellular receptor for poliovirus: molecular cloning, nucleotide sequence, and expression of a new member of the immunoglobulin superfamily. Cell. 1989; 56:855–865. [PubMed: 2538245]
117. Barton ES, et al. Junction adhesion molecule is a receptor for reovirus. Cell. 2001; 104:441–451. [PubMed: 11239401]
118. Dalgleish AG, et al. The CD4 (T4) antigen is an essential component of the receptor for the AIDS retrovirus. Nature. 1984; 312:763–767. [PubMed: 6096719]
119. Williams RK, Jiang GS, Holmes KV. Receptor for mouse hepatitis virus is a member of the carcinoembryonic antigen family of glycoproteins. Proc. Natl Acad. Sci. USA. 1991; 88:5533–5536. [PubMed: 1648219]
120. Tatsuo H, Ono N, Tanaka K, Yanagi Y. SLAM (CDw150) is a cellular receptor for measles virus. Nature. 2000; 406:893–897. [PubMed: 10972291]
121. Fedtke I, Gotz F, Peschel A. Bacterial evasion of innate host defenses — the *Staphylococcus aureus* lesson. Int. J. Med. Microbiol. 2004; 294:189–194. [PubMed: 15493829]
122. Roberts ME, Stewart PS. Modelling protection from antimicrobial agents in biofilms through the formation of persister cells. Microbiology. 2005; 151:75–80. [PubMed: 15632427]
123. Sutherland IW. The biofilm matrix — an immobilized but dynamic microbial environment. Trends Microbiol. 2001; 9:222–227. [PubMed: 11336839]
124. Kurtz J, Franz K. Evidence for memory in invertebrate immunity. Nature. 2003; 425:37–38. [PubMed: 12955131] One example of a recent line of new evidence that shows that specificity and memory are a general hallmark of invertebrate immune defence.
125. Little TJ, Hultmark D, Read AF. Invertebrate immunity and the limits of mechanistic immunology. Nature Immunol. 2005; 6:651–654. [PubMed: 15970937]
126. Adema CM, Hertel LA, Miller RD, Loker ES. A family of fibrinogen-related proteins that precipitates parasite-derived molecules is produced by an invertebrate after infection. Proc. Natl Acad. Sci. USA. 1997; 94:8691–8696. [PubMed: 9238039] The original description of FREPs as a family of immune-type molecules.
127. Zhang S-M, Loker ES. The *FREP* gene family in the snail *Biomphalaria glabrata*: additional members, and evidence consistent with alternative splicing and *FREP* retrosequences. Dev. Comp. Immunol. 2003; 27:175–187. [PubMed: 12590969]
128. Loker ES, Adema CM, Zhang S-M, Kepler TB. Invertebrate immune systems — not homogeneous, not simple, not well understood. Immunol. Rev. 2004; 198:10–24. [PubMed: 15199951]
129. Zhang S-M, Loker ES. Representation of an immune responsive gene family encoding fibrinogen-related proteins in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*. Gene. 2004; 341:255–266. [PubMed: 15474308]
130. Cuthbertson BJ, Shepard EF, Chapman RW, Gross PS. Diversity of the penaeidin antimicrobial peptides in two shrimp species. Immunogenetics. 2002; 54:442–445. [PubMed: 12242595]
131. Nair SV, Del Valle H, Gross PS, Terwilliger DP, Smith LC. Macroarray analysis of coelomocyte gene expression in response to LPS in the sea urchin: identification of unexpected immune diversity in an invertebrate. Physiol. Genomics. 2005; 22:33–47. [PubMed: 15827237]

132. Kondrashov FA, Koonin EV. Origin of alternative splicing by tandem exon duplication. Hum. Mol. Genet. 2001; 10:2661–2669. [PubMed: 11726553]

133. Schmucker D, et al. *Drosophila* Dscam is an axon guidance receptor exhibiting extraordinary molecular diversity. Cell. 2000; 101:671–684. [PubMed: 10892653]

134. Neves G, Zucker J, Daly M, Chess A. Stochastic yet biased expression of multiple Dscam splice variants by individual cells. Nature Genet. 2004; 36:240–246. [PubMed: 14758360]

135. Liu M, et al. Reverse transcriptase-mediated tropism switching in *Bordetella* bacteriophage. Science. 2002; 295:2091–2094. [PubMed: 11896279]

136. Doulatov S, et al. Tropism switching in *Bordetella* bacteriophage defines a family of diversity-generating retroelements. Nature. 2004; 431:476–481. [PubMed: 15386016]

137. Kedzierski L, Montgomery J, Curtis J, Handman E. Leucine-rich repeats in host–pathogen interactions. Arch. Immunol. Ther. Exp. (Warsz.). 2004; 52:104–112. [PubMed: 15179324]

138. Vasta GR, Ahmed H, Odom EW. Structural and functional diversity of lectin repertoires in invertebrates, protochordates and ectothermic vertebrates. Curr. Opin. Struct. Biol. 2004; 14:617–630. [PubMed: 15465324]

139. Ito T, et al. Interferon-α and interleukin-12 are induced differentially by Toll-like receptor 7 ligands in human blood dendritic cell subsets. J. Exp. Med. 2002; 195:1507–1512. [PubMed: 12045249]

140. Honda K, et al. Selective contribution of IFN-α/β signaling to the maturation of dendritic cells induced by double-stranded RNA or viral infection. Proc. Natl Acad. Sci. USA. 2003; 100:10872–10877. [PubMed: 12960379]

141. Dehal P, et al. The draft genome of *Ciona intestinalis*: insights into chordate and vertebrate origins. Science. 2002; 298:2157–2160. [PubMed: 12481130]

142. Scofield VL, Schlumpberger JM, West LA, Weissman IL. Protochordate allorecognition is controlled by a MHC-like gene system. Nature. 1982; 295:499–502. [PubMed: 7057909] An important genetic description of the fusibility phenomenon in protochordates.

143. Buss LW, Moore JL, Green DR. Autoreactivity and self-tolerance in an invertebrate. Nature. 1985; 313:400–402.

144. Cadavid LF, Powell AE, Nicotra ML, Moreno M, Buss LW. An invertebrate histocompatibility complex. Genetics. 2004; 167:357–365. [PubMed: 15166160]

145. De Tomaso AW, Weissman IL. Evolution of a protochordate allorecognition locus. Science. 2004; 303:977. [PubMed: 14963321]

146. De Tomaso A, Weissman IL. Initial characterization of a protochordate histocompatibility locus. Immunogenetics. 2003; 55:480–490. [PubMed: 14520503]

147. De Tomaso AW, et al. Isolation and characterization of a protochordate histocompatibility locus. Nature. (in the press).

148. Barbet AF, et al. Expression of multiple outer membrane protein sequence variants from a single genomic locus of *Anaplasma phagocytophilum*. Infect. Immun. 2003; 71:1706–1718. [PubMed: 12654783]

149. Brayton KA, Palmer GH, Lundgren A, Yi J, Barbet AF. Antigenic variation of *Anaplasma marginale msp2* occurs by combinatorial gene conversion. Mol. Microbiol. 2002; 43:1151–1159. [PubMed: 11918803]

150. Hamrick TS, Dempsey JA, Cohen MS, Cannon JG. Antigenic variation of gonococcal pilin expression *in vivo*: analysis of the strain FA1090 pilin repertoire and identification of the *pilS* gene copies recombining with *pilE* during experimental human infection. Microbiology. 2001; 147:839–849. [PubMed: 11283280]

151. Wang D, Botkin DJ, Norris SJ. Characterization of the *vls* antigenic variation loci of the Lyme disease spirochaetes *Borrelia garinii* Ip90 and *Borrelia afzelii* ACAI. Mol. Microbiol. 2003; 47:1407–1417. [PubMed: 12603744]

152. Coyne MJ, Weinacht KG, Krinos CM, Comstock LE. Mpi recombinase globally modulates the surface architecture of a human commensal bacterium. Proc. Natl Acad. Sci. USA. 2003; 100:10446–10451. [PubMed: 12915735]

153. Krinos CM, et al. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. Nature. 2001; 414:555–558. [PubMed: 11734857]

154. Weinacht KG, et al. Tyrosine site-specific recombinases mediate DNA inversions affecting the expression of outer surface proteins of *Bacteroides fragilis*. Mol. Microbiol. 2004; 53:1319–1330. [PubMed: 15387812]

155. Donelson JE. Antigenic variation and the African trypanosome genome. Acta Trop. 2003; 85:391–404. [PubMed: 12659976]

156. McCulloch R. Antigenic variation in African trypanosomes: monitoring progress. Trends Parasitol. 2004; 20:117–121. [PubMed: 15036032]

157. Duraisingh MT, et al. Heterochromatin silencing and locus repositioning linked to regulation of virulence genes in *Plasmodium falciparum*. Cell. 2005; 121:13–24. [PubMed: 15820675]

158. Freitas-Junior LH, et al. Telomeric heterochromatin propagation and histone acetylation control mutually exclusive expression of antigenic variation genes in malaria parasites. Cell. 2005; 121:25–36. [PubMed: 15820676]

159. Haire RN, Rast JP, Litman RT, Litman GW. Characterization of three isotypes of immunoglobulin light chains and T-cell antigen receptor α in zebrafish. Immunogenetics. 2000; 51:915–923. [PubMed: 11003385]

160. Zapata A, Amemiya CT. Phylogeny of lower vertebrates and their immunological structures. Curr. Top. Microbiol. Immunol. 2000; 248:67–107. [PubMed: 10793475]

Box 1

Alternative mediators of histocompatibility

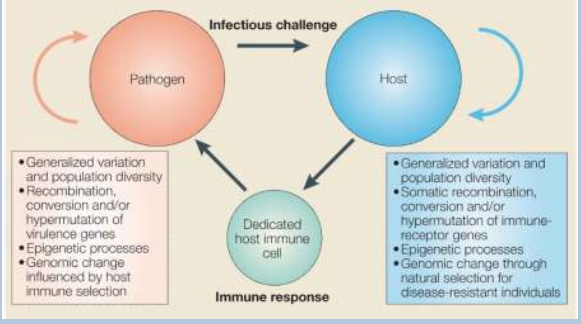
Similar to the B-cell receptors (BCRs) and T-cell receptors (TCRs), MHC class I and class II molecules seem to be restricted to the jawed vertebrates. A rigorous genome-based search for homologues of the genes encoding these molecules in *Ciona intestinalis* (a solitary protochordate) was unsuccessful[141]. It is not yet possible to search whole genomes in the two orders of extant jawless vertebrates, the lampreys and hagfish; however, several molecular genetic searches have failed to detect the presence of MHC class I or class II molecules in these species. The apparent absence of MHC class I or class II molecules in species that occupy positions below the phylogenetic level of jawed vertebrates[38] confounds efforts to assign the ancestral form of the conventional adaptive immune receptors as being either BCR- or TCR-like and might present an intractable problem. Notably, despite the apparent absence of MHC molecules, *Botryllus schlosseri* (a colonial protochordate)[142] and *Hydractinia* species (cnidarians; which are invertebrates)[143,144] show HISTOCOMPATIBILITY REACTIONS (that is, allotransplantation reactions) through a process known as fusibility, which protects the germline by inhibiting chimerism and/or parasitism[145]. Fusibility is determined by a single Mendelian trait and, in *Botryllus schlosseri*[146], is mediated by highly polymorphic cell-surface receptors that contain immunoglobulin-superfamily and epidermal-growth-factor domains that are encoded at a single locus[147]. Notably, again, a chimeric structure is involved in an immune-type function.

Box 2

Mechanistic similarities between pathogenicity and protection

In a co-evolutionary struggle, the host and the pathogen use similar mechanisms to modify selection and outcome. In various pathogens, including bacteria and protozoan parasites, several fundamental mechanisms of DNA change are used during invasion of the host. DNA recombination[148–151], gene conversion[152–154] and gene hypermutation can lead to the evasion of host detection and to the subsequent clonal selection of the pathogenic variant. For example, in trypanosomes, DNA recombination and gene conversion rapidly modify variant surface glycoproteins and allow evasion of host detection[155,156]. In the parasite that causes malaria, these genetic events, which include gene activation and regulation, are modulated by various complex epigenetic mechanisms[157,158].

In the host, analogous processes, including DNA recombination and hypermutation of recognition receptors, counter these pathogenic challenges. For example, genome diversification occurs during meiotic recombination and can be further increased by the creation and maintenance (through sexual outbreeding) of allelic diversity (which is generated by selection and drift) and by the epigenetic control of gene expression. These events help to establish generalized variation and population diversity. The host immune cell, through its receptors, directly influences selection of the host or the pathogen. In vertebrates, somatic modification of the immune-cell receptor repertoire is common and can further improve recognition potential in clonal lineages. The direct interaction between immune cell and pathogen affects the selection of virulence genes and the diversity of pathogen populations. Since the origin of metazoans, a continuous exchange of selective pressures through a pathogen–host–immune-cell loop has driven the simultaneous diversification of cell-surface receptors, using parallel mechanisms of genetic modification, which has helped to drive the lineage-dependent divergences that are seen in the innate and adaptive immune repertoires.



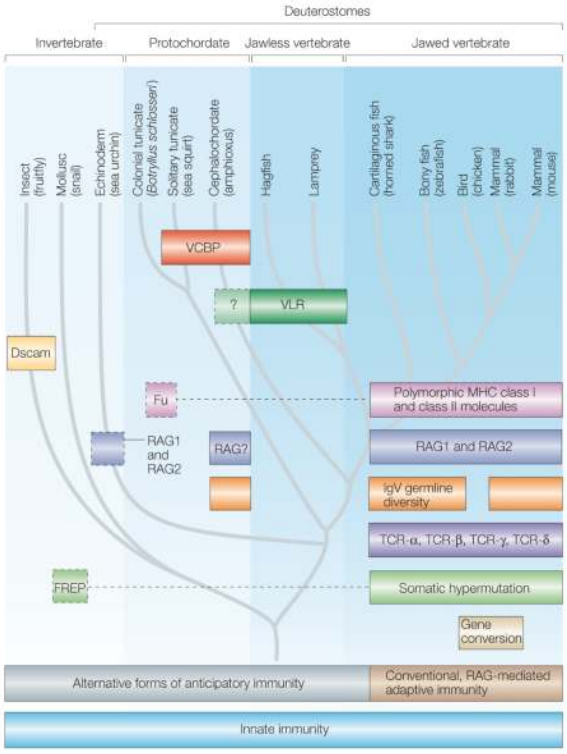


Figure 1. Phylogenetic occurrence of various immune functions, genetic organizations and mechanisms of receptor variation in selected representative species

The selection of individual invertebrate, protochordate, jawless-vertebrate and jawed-vertebrate species is based on emphasizing certain effects and is not intended to be comprehensive. Innate immunity uses various mediators that are not necessarily shared among the main groups that are represented; conventional, recombination-activating gene (RAG)-mediated adaptive immunity is thought to be confined to the jawed vertebrates, and alternative forms of anticipatory immunity are found in species that are below this level of phylogenetic development. Dashed lines are used to emphasize certain general or parallel relationships. Fusibility (*Fu*) genes are unrelated to the genes encoding MHC class I and class II molecules, but they are an alternative histocompatibility locus[146]. A *Fu* locus has also been identified in *Hydractinia* species (cnidarians; which are invertebrates), and there might be homologous forms of *Fu* gene products in other species. *RAG1*-like gene sequences are found in amphioxus, which are cephalochordates; *RAG2* sequences have not yet been characterized. Differential RNA splicing creates diversity in insect Down's syndrome cell-adhesion molecule (Dscam), which is associated with bacterial binding and phagocytic function. Hypermutation occurs in the genes encoding fibrinogen-related proteins (FREPs) in a mollusc. Various leucine-rich repeat (LRR)-encoding genes, which are distantly related to the genes encoding the variable lymphocyte receptors (VLRs) of lampreys and hagfish and the Toll-like receptors of vertebrates, can be identified in the amphioxus genome (indicated by light-green shading); however, their relatedness to the genes encoding VLRs and to other LRR-encoding genes has yet to be determined, and the possibility that there might be VLRs in jawed vertebrates awaits further investigation. Throughout the metazoans, a huge number of different receptors and effector molecules are used to effect innate immunity, and these typically vary between the major phylogenetic groups. The background shows a simplified phylogeny of many of the species that feature in current considerations of the evolution of innate and adaptive immunity. IgV,

immunoglobulin variable region; TCR, T-cell receptor; VCBP, variable-region-containing chitin-binding protein.

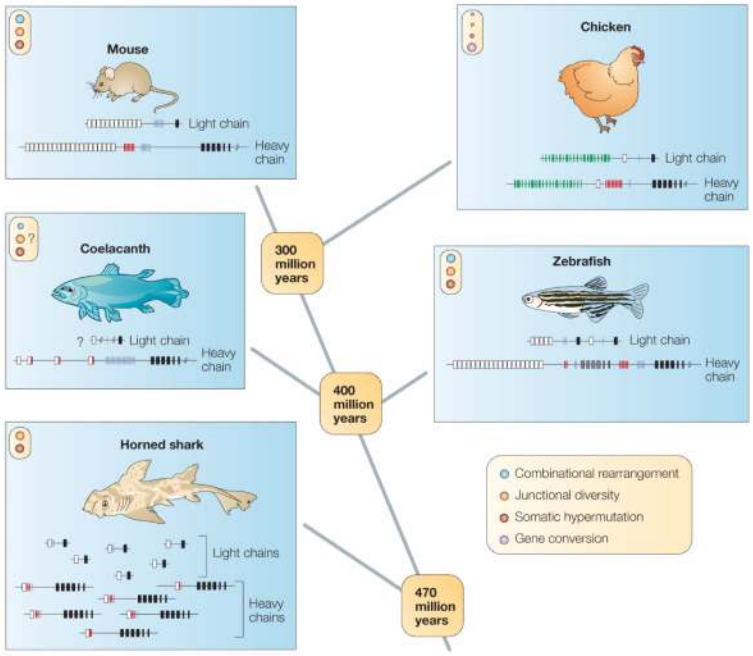


Figure 2. Main differences in the organization of rearranging segmental elements in immunoglobulin heavy and light chains in selected jawed-vertebrate models

Each jawed vertebrate is indicated by its common name: horned shark, zebrafish, coelacanth, chicken and mouse. The selection of the horned shark, chicken and mouse emphasizes the co-evolution of the organization of genes and the main mechanisms of diversification. In the case of mice, only an Igκ light-chain locus is shown; the Igλ light-chain locus has a cluster-type organization. Neither the number nor the relative position that is shown for variable (white), diversity (red), joining (blue) and pseudogene (green; used in gene conversion) elements in heavy and light chains is absolute. For the heavy chain, only the most proximal exons encoding the constant region (black) are indicated, except in the case of zebrafish, for which the exons encoding an additional heavy-chain class (grey) are shown. Additional pseudogenes are present in the continuous variable-element-encoding region in mice and zebrafish and probably also in coelacanths. The organization of genes in the type II light chain[159] in the zebrafish has been inferred by annotation of the zebrafish genome and is not complete as represented; note the presence of two constant regions in the cluster. Light-chain gene organization has not been reported in coelacanths. The relative contributions of various diversifying mechanisms have been approximated in circles: combinatorial rearrangement (blue), junctional diversity (orange); somatic hypermutation (red) and gene conversion (pink). The differences in circle area roughly approximate the relative differences in overall effect and are not quantitative. The relative time of divergence from the lineage that gave rise to mammals is indicated in million-year increments; a ~50-million-year error is assumed[160]. This figure is produced from data taken from REFS 29,33.

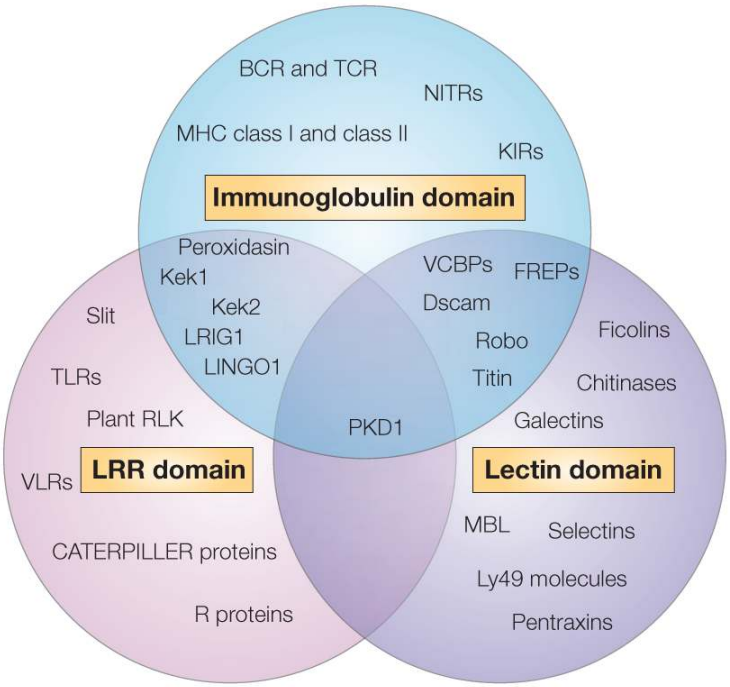
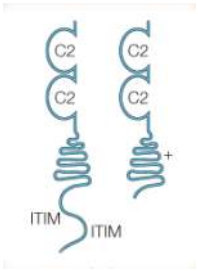
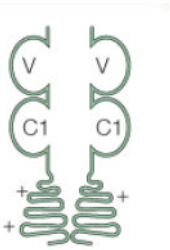
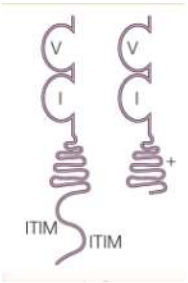


Figure 3. Three main framework domains are incorporated in various proteins, including those that mediate adaptive and innate immunity

The immunoglobulin-domain framework includes a superfamily of homologous types: variable (V), intermediate (I) and constant (C1 and C2). Single or multiple copies of one or more types of immunoglobulin domain are present in different receptors: for example, novel immune-type receptors (NITRs) can have V and I domains. Other proteins have more than one type of framework domain: for example, immunoglobulin V domains and a chitin-binding lectin domain in V-region-containing chitin-binding proteins (VCBPs). Fibrinogen-related proteins (FREPs) and Down's syndrome cell-adhesion molecule (Dscam) also contain immunoglobulin and lectin domains. Proteins that have both lectin and leucine-rich repeat (LRR) domains are rare; only polycystic kidney disease 1 (PKD1) is known to have all three domains. Proteins that share a specific domain type might not be related by common ancestry, so only the domains themselves have a homologous past (as products of recruitment). LRR-containing proteins are a special case. The LRR framework is a small motif (~11 amino-acid residues) that is generally encoded in contiguous copies. The functional site can be a single motif or a combination of motifs. LRR motifs are so divergent that even though the 11-residue-motif structure is highly conserved, together with the resulting solenoid structure of the LRR-containing protein, it is not possible to predict whether any two LRR motifs evolved from a common ancestor or whether they arose through evolutionary convergence. BCR, B-cell receptor; CATERPILLER, caspase-recruitment domain, transcription enhancer, R (purine)-binding, pyrin, lots of LRRs family; ficolin, collagen- and fibrinogen-domain-containing lectin; galectin, galactoside-binding lectin; Kek, kekkon; KIR, killer-cell immunoglobulin-like receptor; LINGO1, LRR- and immunoglobulin-domain-containing, NOGO (neurite-outgrowth inhibitor)-receptor-interacting protein; LRIG1, LRR and immunoglobulin-like domains 1; MBL, mannose-binding lectin; RLK, receptor-like kinase; Robo, roundabout; R proteins, plant disease-resistance proteins; TCR, T-cell receptor; TLR, Toll-like receptor; VLR, variable lymphocyte receptor.

Table 1

Natural-killer-cell, T-cell and related receptors

Feature*	KIR‡	TCR‡	NITR‡
			
Number of IgSF domains	2–3	2	1–2
Vdomains	No	Yes	Yes
Somatic change	No	Yes	No
Activating and inhibitory forms	Yes	No	Yes
Haplotypic variation	Yes	–	Yes
Maximum haplotype number §	14	~70 Vα, 52 Vβ	36

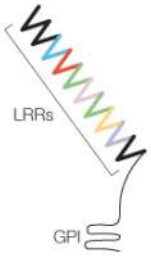
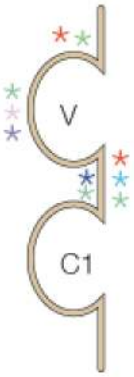
*The various characteristics that are indicated allow comparisons between the three receptor classes and are neither inclusive nor necessarily indicative of functional analogy, but they show that novel immune-type receptors (NITRs) share various characteristics with killer-cell immunoglobulin-like receptors (KIRs) and T-cell receptors (TCRs).

‡The KIRs and αβ-TCR shown are of human origin; the NITRs that are described are from zebrafish. The structures have been drawn to emphasize potential similarities between these receptor types; the domain types — variable (V), intermediate (I) and constant (C1 and C2) — are indicated. Inhibitory motifs (immunoreceptor tyrosine-based inhibitory motifs, ITIMs) and positively charged residues in the transmembrane regions of KIRs and NITRs are shown; the pairings that are depicted are not intended to imply that KIRs and NITRs function as heterodimers. Positively charged residues in the TCR, which does function as a heterodimer, are also indicated. The structures are not drawn to scale.

§Haplotype comparisons reflect the highest allele content for various human populations and a laboratory line of zebrafish. IgSF, immunoglobulin superfamily; Vα, variable region of α-chain; Vβ, variable region of β-chain.

Table 2

Organization of some alternative immune-type receptors

Feature*	Dscam‡	VLR‡	VCBP‡	FREP‡	IgL‡
					
IgSF domains	Yes	No	Yes	Yes	Yes
V domains	No	No	Yes	No	Yes
Leucine-rich repeats	No	Yes	No	No	No
Somatic reorganization	No	Yes	No	No	Yes
Junctional diversity	No	No	No	No	Yes
Complex RNA processing	Yes	No	No	No	No
Somatic hypermutation	No	No	No	Yes	Yes
Extensive population polymorphism	No	No	Yes	Yes	No
Regionalized hypervariation	Yes	No	Yes	No	Yes

*The various properties and characteristics that are indicated focus on those that best reflect the similarities and differences between the various receptors shown.

‡Schematic organization is shown for immune-type receptors that have various characteristics of innate and adaptive immune receptors and are found in traditional model systems: from jawless vertebrates, variable lymphocyte receptors (VLRs); from invertebrates, fibrinogen-related proteins (FREPs) and Down's syndrome cell-adhesion molecule (Dscam); from protochordates, variable (V)-region-containing chitin-binding proteins (VCBPs); and from jawed vertebrates, immunoglobulin light chains (IgLs). Coloured segments in the VLR indicate the regions in which extensive somatic variation occurs in leucine-rich repeats (LRRs). The relative sites of sequence variation in the various receptors are indicated by coloured symbols, irrespective of whether the variation occurs in a single gene (as is the case for IgL). Variation arises through several mechanisms: complex RNA processing, for Dscam; recombination of LRR cassettes, for VLRs; allelic complexity and population polymorphism, for VCBPs; and somatic mutations in multiple germline genes, for FREPs. Variation in Dscam is localized to only three immunoglobulin domains. Gene conversion (not shown) is a diversification mechanism for avian IgL and possibly for VLRs, although the latter has not been established. C1 domain, type 1 immunoglobulin constant-region domain; CBD, chitin-binding domain; FBG domain, fibrinogen-like domain; FnIII domain, fibronectin type III domain; GPI, glycosylphosphatidylinositol; IgSF, immunoglobulin superfamily.