Mini-review:

Defense strategies and immunity-related genes

John Trowsdale1 and Peter Parham2

1 Department of Pathology, and Cambridge Institute for Medical Research, University of Cambridge, Cambridge, GB
2 Department of Structural Biology, Stanford University, Stanford, USA

The immune system is of crucial importance in defense against infection. It has to cope with a large number of different pathogens that relentlessly develop new ways to avoid recognition or elimination. Yet most infections are cleared. Immune-system genes must evolve to keep pace with increasingly sophisticated evasion by pathogens. In this article we examine features of human defense genes that reflect the demands imposed by such intense selection. Key examples are MHC and KIR genes, where features such as polygeny and polymorphism facilitate the comprehensive logistics needed to counteract infection.

Key words: Genetics / Immunity / Defense / Infection

Introduction

Military defense strategists could be advised to study the immune system. A glance at a Ministry of Defence document “The Future Strategic Context for Defence” (http://www.mod.uk/issues/strategic_context/military.htm) reveals a vocabulary that is reminiscent of an immunology textbook. The military and the immune system are, in separate ways, locked into an escalating arms race. Even during periods of peace, defense systems must be in a state of preparedness, which gives them a particular character and sets them apart from civilian activities. Here we consider whether any special characteristics can be recognized in the genes that specify our defense against infection — our immune system.

Differences in immunity genes between different species

Genome-sequencing highlights the importance of immunity, as a considerable fraction of the genome is dedicated to defense. In man and most other mammals up to 5% of the genes are involved in this role. The precise number is difficult to estimate because many sequences related to immunity genes have no known role in defense (Table 1).

A comparison of the genomes of mice and man reveals two classes of gene that differ most in number and in sequence; namely, genes for the immune system and genes for reproduction [1]. It is not difficult to imagine how both sets of loci are crucial to the survival of all species. Intriguingly, there are also hints of intimate connections between the immune system and reproduction. For example, a subset of NK cells appear to play a role in placentation, and the placenta expresses specific MHC molecules such as HLA-G [2].

Features of immunity genes

What features are exploited by immunity genes in order to increase their defense potential? Table 2 lists some of the characteristics (detailed below) that have been noted for genes in either the innate or adaptive immune sys-

[DOI 10.1002/eji.200324693]

Abbreviations: KIR: Killer-cell Ig-like receptor LD: Linkage disequilibrium LILR: Leukocyte Ig-like receptor NKC: Natural killer complex TLR: Toll-like receptor
Table 1. Criteria for defining immune defense genes

<table>
<thead>
<tr>
<th>Category</th>
<th>Comments and examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sequence related to known defense gene</td>
<td>There are sequences related to immune system genes that have no known role in defense, including many Ig superfamily molecules [89]</td>
</tr>
<tr>
<td>Expressed specifically in immune tissues</td>
<td>Not all sequences have such convenient expression patterns</td>
</tr>
<tr>
<td>Expression induced by immuno-modulators such as interferon</td>
<td>Such a list contains many plausible candidates as well as some molecules with more general roles such as cell–cell interaction</td>
</tr>
<tr>
<td>Part of a pathway, such as signal transduction or transcription regulation, that results in expression of defense molecules, such as NFκB</td>
<td>These pathways often encompass molecules with wide applicability</td>
</tr>
<tr>
<td>Interact directly with pathogens or their products</td>
<td>e.g. TLR, lectin superfamily</td>
</tr>
<tr>
<td>Result in immunodeficiency when disrupted</td>
<td>About 80 human genes have been analyzed [74] and many more mouse loci</td>
</tr>
</tbody>
</table>

a) An overriding difficulty defining immune genes is that vertebrate gene products appear to have multiple uses. Incidentally, it has been proposed that the small size of the human genome is a constraint imposed by an efficient adaptive immune system [86].
b) Some mouse class I-related genes may have interesting non-immune functions [87, 88].
c) A table can be found at http://arjournals.annualreviews.org/doi/suppl/10.1146/annurev.immunol.15.1.749

d) Examples of sets of genes that are not necessarily defense related but appear to be evolving rapidly are Morpheus [90] and the mouse urinary proteins (MUPS) [91].

Table 2. Some features of defense genes that may reflect their defense role

<table>
<thead>
<tr>
<th>Feature</th>
<th>Comments and examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polygeny</td>
<td>Many examples including MHC, Ig, KIR, FcR</td>
</tr>
<tr>
<td>Polymorphism</td>
<td>MHC, KIR, NOD, TLR</td>
</tr>
<tr>
<td>Clustering</td>
<td>Two types of clusters: 1. duplicated genes in cis 2. genes from different gene families such as MHC</td>
</tr>
<tr>
<td>Linkage disequilibrium</td>
<td>MHC, KIR, FcR</td>
</tr>
<tr>
<td>Sequence exchange</td>
<td>At least two mechanisms: 1. Allele conversion 2. Gene conversion</td>
</tr>
<tr>
<td>Generation of repertoires</td>
<td>Somatic rearrangement in the adaptive immune system (Ig/TCR), Stochastic gene expression (KIR)</td>
</tr>
<tr>
<td>Rapid evolution</td>
<td>The best example is NK receptors, which are very different in arrangement in related species</td>
</tr>
<tr>
<td>Co-evolution</td>
<td>e.g. antigen-processing and -presenting genes in the MHC</td>
</tr>
<tr>
<td>Association with disease</td>
<td>Not always easy to confirm, even in mouse knockouts, because of functional duplication</td>
</tr>
<tr>
<td>Networking</td>
<td>Applies generally, but crucial for defense, in inflammation for example</td>
</tr>
</tbody>
</table>

Rule number one of warfare is to ensure that your army is bigger than theirs. Defense genes have expanded in number many, many times. Duplication of immunity genes underpins many of their other innovative properties. Long before any genome sequencing, Ohno pointed out that modern genes are duplicated progeny of ancestral genes [4]. He realized that duplication provides a way of retaining, through conservation of one duplicate, the currently useful function of the encoded protein, whilst its twin is liberated to mutate and possibly acquire a novel function. This process is prolific in some families of immune defense genes. Examples of such families are MHC class I and class II, Fc receptors, cytokines and killer-cell Ig-like receptor (KIR; for gene designations, see http://www.gene.ucl.ac.uk/cgi-bin/nomenclature/searchgenes.pl). These families appear to be regularly refreshed by duplication and comprise multiple copies with varying degrees of relatedness. These sets of genes are quite often arranged along a chromosome, where they have duplicated in cis and remain next to each other.
Other more extensive duplication mechanisms are invoked, such as en bloc duplication of a complete chromosome, to explain the paralogous clusters of genes that are present on different daughter chromosomes. Complete genome duplication has also been proposed and the 2R hypothesis calls for two such rounds of allopolyploidization. A well-studied example of the outcome of these events is the presence of paralogous clusters of MHC-related genes on four human chromosomes: 1, 6, 9 and 19. Inverted duplication has been associated with the rapid multiplication of gene copies in somatic cells under strong selection. It has been proposed that inverted duplications are generated as a primary event in the amplification process. The ‘hairpin’ arrangement of genes in the leukocyte Ig-like receptor (LILR) cluster (the LILR gene designation replaces ILT and LIR) next to the KIR loci on chromosome 19q13.4 is consistent with this proposal.

The Ig superfamily is the most familiar product of progressive gene duplication. Approaching 1000 human loci encode proteins containing one or more domains that are structurally similar to the Ig domain. As implied by the term superfAMILY, the Ig lineage is ancient, and related protein folds are recognizable in bacterial proteins. However, a proportion of molecules with Ig folds have no apparent role in defense. Most Ig domains are encoded by exons in the same splicing phase so that a domain could be gained or lost with no disturbance of the reading frame of the protein. The lectin-like superfamily is another example of an extensive set of related loci that are present as large clusters on different chromosomes and are the products of a history of repeated duplication.

Maintenance of a large force is expensive so it pays to keep just the elite troops on permanent contracts. So too in the immuno-genome. Comparison of MHC class I genes in different species shows how perpetual duplication and turnover of genes is a continuous ‘birth and death’ process. Early work on human and mouse class I sequences led to the assumption that HLA-ABC and H-2kd genes must be orthologues of each other, in view of their analogous positions in the MHC region. This view was not supported by analysis of human and mouse class I sequences. In fact, classical class I sequences within each species are more related to each other than they are between the two species. One explanation is that the sequences originate as direct orthologues but they became homogenized after speciation, by exchange of sequences. A more likely scenario is that diverging species derive their own sets of classical class I genes by duplication of one locus, or a small number of loci, that were present in the common ancestor.

Once they have outlived their usefulness, genes can die and become non-functional pseudogenes. They may remain intact but expressed at negligible levels. Definition of death in the gene world is problematic. There are many MHC class I genes, in different species, in a twilight zone, with unknown levels of animation. Clear roles for these genes, such as HLA-G and -F in humans, remain difficult to define. Not so long ago HLA-C was largely ignored by immunologists because it is expressed at a lower level than HLA-A and -B. Now clear is that HLA-C plays an important role in interaction with NK receptors, possibly one more important than HLA-A and -B. Even unexpressed pseudogenes can remain useful as a source of donor sequences for gene conversion.

Given the promiscuity of MHC class I genes it is understandable that the arrangements of these loci differ markedly between species. An extreme case is the pygmy mouse, where the hundreds of class-I-related sequences, most apparently pseudogenes, are more consistent with a history of ‘boom and bust’ rather than ‘birth and death’. Though not strictly involved in immune defense, but definitely sensing the environment, the olfactory receptor genes are a most interesting example of runaway duplication. Most mammals contain up to 1000 of these genes but in some species, including man, over half are pseudogenes.

NK cell receptors in the Ig superfamily, the KIR gene family in the leukocyte receptor complex on chromosome 19q13.4 in man, and the functional equivalents on mouse chromosome 6, the Ly49 genes, use similar strategies of gene duplication and diversification. The tandem (head–tail) arrangement of the KIR loci facilitates duplication of sets of contiguous loci by unequal crossing-over at meiosis. A recent report is of a haplotype containing two allelic for three contiguous loci, which must have arisen quite recently by this mechanism. The 5' part of KIR loci encodes the extracellular portions of the receptor and the 3' part encodes the cytoplasmic tail. The tails are of two main types, containing either ITIM motifs, which give inhibitory signals, or short tails with a

---

3) An orthologue is one set of homologous genes in different species. A parologue is one set of homologous genes within a single species.

4) The 2R hypothesis states that there were two rounds of duplication of the complete genome early in vertebrate history.
Fig. 1. Equal and unequal crossovers in duplicated KIR loci. (A) Recombination in a family of related genes, such as KIR, arranged head-to-head. A crossover within a gene would result in a novel hybrid molecule. Five genes are shown with the same transcriptional direction (small arrows). The crossover is designated by "X". (B) An example of how misalignment at meiosis could lead to haplotypes with different numbers of genes. The arrangement of many different KIR haplotypes is consistent with other mis-alignments associated with gain or loss of loci in the complex. (C) Proposed historic illegitimate crossover in KIR genes [19].

charged residue in the transmembrane region permitting coupling with activating adapters such as DAP12.

Unequal cross-overs can, in principle, result in switching of a receptor from activating to inhibitory, or vice versa, in one step. There are some obvious examples of such ‘hybrid’ KIR genes. In one case, the KIR3DL/S locus, both inhibitory and activating alleles are present at a single locus. The external ligand-binding parts of the 3DL1 and 3DS1 molecules are almost identical but they are associated with either a C-terminal tail specifying inhibitory motifs (ITIM) or a short tail suitable for recruiting DAP12. Similar features and possibilities apply to mouse Ly49 genes, although they are type II transmembrane lectin-related molecules. The system provides a nice example of the arms race between a virus and host immunity. In susceptible mice, murine cytomegalovirus makes a surrogate class I molecule, m157, that binds to the inhibitory Ly49i receptor and turns off host NK cells. Mice resistant to the virus encode an activating receptor, Ly49h, providing a counter-strategy. The activating and inhibitory receptors are highly homologous, suggesting the possibility that one evolved from the other in response to selective pressure imposed by the pathogen [22].

Families of pattern-recognition receptors in the innate immune system, such as Toll or NOD1/2, provide other examples of diversification through gene duplication. Both systems of pathogen recognition span organisms as diverse as plants [23], flies and mammals, and there is consistency in structure and function, including signaling pathways, in all of these species [24, 25]. Duplication does not necessarily involve complete genes, and commonly short protein modules, usually encoded on separate exons, are duplicated and combined in various arrangements, such as in the NOD/CARD families [26].

Polymorphism of immune loci

It makes sense to deploy a wide range of defense options and to change them from time to time. Some of the immune system genes are amongst the most variable in the genome. The classic example is of MHC class I and class II, where up to 500 alleles per locus have been reported [27]. Most of the variation is related to binding of peptides from pathogens to the grooves of MHC class I and class II molecules. The variation is most likely driven by episodes of selection for resistance to infection, although other ideas have been suggested. The most intriguing is the notion that MHC polymorphism facilitates out-breeding by providing olfactory cues to guide the female in her choice of mate [28].

Data on person-to-person variation for most genes are lacking but what is clear is that the level of variation within populations differs markedly between different sets of immune defense loci and no other human protein-coding genes are as variable as MHC class I/II. There are some examples of variable genes in other systems, human growth hormone for example [29]. It seems that the loci that exhibit extreme levels of polymorphism — MHC, Toll [30], NOD [25] and NK receptors — may turn out to be those whose products interact directly with pathogens, such as herpes viruses and retroviruses, or their products, such as LPS, although this distinction is blurred since many receptors, including lectin molecules in the natural killer complex (NKC), bind a cellular ligand as well as microbial products [31]. In general the levels of polymorphism of defense molecules may diminish the further away they are from direct interaction with the products of infection. Many cell-surface molecules that interact with other host molecules are practically invariant.

A recent example concerns the mouse Nkrp1 and Clr genes, which are linked at the centromeric end of the NKC [32]. The products of these genes interact as lectin-receptor–lectin-ligand, which is unusual. Unlike other NKC genes, Nkrp1 and Clr are remarkably conserved between mouse haplotypes. Lack of recombination over the region may pose a problem for evolutionary change.
as reciprocal mutations in both receptor and ligand genes are needed to generate new functionally important specificities [33]. At the telomeric end of the NKC, lectin-related NK receptors for non-classical class I molecules are relatively conserved [21].

The ratio of non-synonymous (Ka) to synonymous (Ks) polymorphisms can be used to indicate the nature of the selective pressure on a gene. A high ratio is indicative of diversifying, and a low ratio with purifying, selection. Olfactory, reproductive and immune-related genes generally have higher Ka:Ks values [1].

Many of the KIR loci that are mentioned above also have a large number of alleles [21], as do certain of the related and linked LILR genes, such as LILRB3 and ILT-8. Polymorphism of MHC and KIR loci is not confined to sequence variation. Both gene families are characterized by haplotypes that have or lack particular genes. Thus haplotypes contain a different complement of loci. HLA-DRB is a familiar example, where each haplotype consists of 1–4 DRB loci [27]. Because of these non-uniform arrangements allelism is difficult to define from gene or cDNA sequences and requires more-extensive knowledge of the flanking sequences and loci. In the class III region of the MHC the genes encoding complement component C4 are subject to analogous haplotypic differences in gene numbers although the duplicated copies are highly homogeneous. Presence/absence polymorphism in KIR is proving to be more extensive than for any other gene set, including DRB.

**Clustering of immune loci**

Allies must learn to work with each other in war even if they speak different languages. At the genome level some degree of coordination may be achieved by gene linkage. This is difficult to prove but evidence has been sought by looking for clustering of genes in related pathways, like bacterial operons. Two types of immune gene clusters are evident in mammalian cells: ones comprising homologous genes and ones comprising dissimilar genes. Since genes tend to duplicate in cis, as described above, it is not surprising that loci often lie next to each other in the genome, in linked array. More interesting are the clusters of linked but unrelated immunity loci, such as are present in the MHC. Strong linkage disequilibrium is a well-known feature of the MHC, as well as many other regions of the genome. Advantages may come from linkage of polymorphic genes with each other, where their products interact, either directly as proteins, or indirectly in related pathways.

A classic example from the MHC class II region is the HLA-DQ molecule (H2A in mouse) which is a heterodimer of α and β chains, both of which are polymorphic [34]. Maintaining linkage of genes for both chains, on the same haplotype, ensures matching sets of products rather than taking pot-luck if the subunits were located on different chromosomes. A similar argument has been proposed for the juxtaposition of antigen-processing genes near to class I loci [35]. In rats, one of the subunits for the TAP molecule that transports peptides for loading onto class I molecules exhibits allelic variability at up to 30 different amino acid sites. Maintaining linkage of the transporter ensures that suitable peptides are supplied to bind the linked, polymorphic class I. This is equivalent to making sure that ammunition supplied with a gun is of the correct caliber. In many species, other relevant genes, such as those encoding tapasin and proteasome components, share the MHC region that contains TAP molecules.

Several explanations have been put forward to explain the persistence of conserved, extended haplotypes in regions like the MHC. One possibility is that recombination is restricted or suppressed in such regions. This has been suggested for the linkage of Nkrp1 receptor and Cir ligand genes referred to above [32]. There is evidence for ‘genetic protection’ in this region of the mouse NKC as there are no documented recombination events in >6400 meioses. A second possibility is that selection favors certain combinations of alleles. A third possibility is that haplotypes are simply the result of rapidly expanding families and insufficient time has elapsed for recombination to randomize the combinations of alleles on founder haplotypes.

The advantages of clustering may be subtle and not necessarily exploited in all species. TAP genes are polymorphic in rats, chickens, hamsters and Xenopus but conserved in most other species including man. Xenopus MHC haplotypes appear to comprise two very ancient allelic lineages, under balancing selection [36]. The chicken MHC appears to contain just the minimal, essential set of immunity genes, arranged as allelic haplotypes, and interestingly includes two C-type lectin genes related to NK receptors [37]. An ancient genetic association between MHC and NK receptor genes is also suggested by the presence of class-I-related genes near the mouse leukocyte receptor complex [38]. In bony fish, which represent about half the number of vertebrate species, class I and class II loci are not linked [39]. This organization is not the ancestral form as more primitive cartilaginous fish have linked MHC class I and class II genes [40].
Other clusters of immunity loci are well known to those laboratories attempting to map loci influencing common, multigenic autoimmune disorders. Susceptibility to complex autoimmune diseases is affected by a variety of genetic, environmental and stochastic factors. Identification of susceptibility alleles is also dogged by extensive genetic heterogeneity and epistatic interactions among the multiple genes involved in disease development. In many cases, the precise etiologic variant is difficult to identify because it lies within a cluster of other potential candidate immunity genes. The syntenic co-localization of susceptibility alleles in both mouse models and human linkage studies indicates either that several susceptibility alleles affect multiple diseases, or alternatively that genomic organization has resulted in the clustering of many immune-system genes. Examples include interleukins, chemokines and receptors, Toll-like receptors (TLR), Fc-receptors and immune co-receptors [41].

**Linkage disequilibrium**

Linkage disequilibrium (LD) is the phenomenon of non-random association between alleles at linked genetic loci. LD varies markedly across the human genome so that there is an irregular relationship between LD and genetic distance. LD has been observed also at distant loci without association of intervening markers. Factors that influence LD include population age, selection and recombination. It has proved extremely difficult to distinguish between the second and third possibilities in the region most studied, namely the MHC. Often alleles of genes in this region have been shown to comprise ‘ancestral haplotypes’ or ‘polymorphic frozen blocks’ [42]. Several careful studies reveal short recombination hotspots within the MHC that may play a major role in maintaining LD in the region [43]. However, this phenomenon is not unique to the MHC and extensive regions of LD feature in many other regions of the genome [44]. Recent data show that recombination hotspots in the MHC are accompanied by biased gene conversion, occurring by gap repair, with eventual extinction of the hotspot [44a]. The KIR region exhibits considerable LD, divided into two halves. There appears to be greater LD within each half than there is between the halves of the complex [45, 46]. Although it is difficult to identify the selective forces driving LD its effect can be to preserve coordinated blocks of alleles.

**Sequence exchanges and immune loci**

A defensive force must keep up to date. Some immune system genes refresh their sequences by incorporating fragments from related genes rather than relying on new mutations. This was discovered by developing histo-incompatible (bm) mutants of inbred strains of mice, which were found to have complex alterations in MHC class I molecules, particularly H-2Kb [47]. The clustered pattern of alterations was consistent with genetic recombination rather than point mutation. Potential donor genes were identified in several other regions, including K, D and Q. The extent of the DNA region transferred ranged from 5–95 nucleotides. Genealogical analysis of the mutants was consistent with the sequences being exchanged by gene-conversion events occurring during mitotic amplification of germ cells.

These data indicate that genetic exchange plays a major role in diversification of evolution of MHC class I genes. Comparison of the nucleotide sequences of human MHC class I indicates that exchange of short sequences, usually ~10–15 nucleotides, is the source of many alleles, whereas alleles produced by exchange between loci are rarer [48]. Similar exchanges may account for the ‘patchwork’ of sequences evident in MHC class II although, like class I, shuffling may take place by allele conversion (double-crossovers between alleles) as well as gene conversion [27]. Analysis of single sperm was used to detect new HLA-DPB1 alleles generated by interallelic gene conversion [49].

How much diversification of other sequences, in immune defense, or the genome in general, is by gene conversion? This is extremely difficult to measure but there are documented examples suggesting that it is extensive [50, 51], and it is possible that most variation is generated by this mechanism. For example, comparison of orthologues and paralogues of olfactory-receptor genes in primates revealed a multiplicity of gene-conversion events, which have led to segment shuffling in the odorant-binding site [52]. Conversely, it has been suggested that frequent intra-chromosomal gene conversion is responsible for maintaining a high degree of sequence conservation of massive palindromic arrays associated with testis-specific genes on the Y chromosome [50].

**Generation of repertoires**

The problem of responding to every eventuality has been neatly solved by the adaptive immune system. Variation in certain immune-system genes, through somatic gene rearrangements dependent on RAG (recombination-activating gene) functions, is well documented elsewhere [53]. These genes include those encoding immunoglobulins, TCR and cadherins [54], with cadherins exploiting the duplication of gene segments. This mechanism for providing a large repertoire of receptors is confined to the adaptive immune system, and species can...
To remain effective any defense strategy needs continu-
than selection [66]. Together, these findings suggest that the high allele num-
but the HLA-B alleles differ between communities. The total number of alleles in each community remains small
have been selected in response to novel parasite loads. The inter-allelic conversion. These new alleles may have
new HLA-B alleles appear to have arisen, mostly by
North America are identical to those in founding popula-
ations but in some South American communities many
new HLA-B alleles appear to have arisen, mostly by
inter-allelic conversion. These new alleles may have
been selected in response to novel parasite loads. The total number of alleles in each community remains small
but the HLA-B alleles differ between communities. Together, these findings suggest that the high allele num-
bers that characterize urban populations are more a reflection of migration and population admixture rather than selection [66].

A different mechanism is used by KIR and Ly49 genes, to
provide a repertoire of human and mouse NK cells, respectively, expressing a range of receptors. In these cases, stochastic expression of different combinations of receptors by NK cells, followed by maintenance of expression on daughter cells, results in a repertoire of NK clones with a variety of specificities [57, 58]. Not known is how stochastic expression of receptor genes is regulated. The promoters of differentially expressed KIR genes are extremely similar, consistent with a ‘random’ control mechanism based on limiting transcription fact-
ors or promoter methylation [59, 60]. Stochastic expres-
sion of Ly49 genes may be controlled in a similar way. Other mechanisms of variegation have arisen indepen-
dently several times in evolution [61]. There may be par-
allels in the control of the vast array of olfactory receptor genes, whereby each neuron apparently expresses only one product out of the thousand available [62]. These stochastic control systems may exploit general, rather than gene-specific, mechanisms because mice trans-
genic for a large DNA fragment containing the human KIR genes exhibit elements of stochastic expression on mouse NK cells even though mice may not have func-
tional KIR genes [63]. There is also evidence that the rep-
ertoire of KIR-expressing NK clones is conditioned by engagement with self class I molecules [64, 65].

Rapid evolution

The colonization of the Americas by southward migration of small Asian populations from the Behring land bridge ~30,000 years ago provides an invaluable experiment of nature. The HLA alleles of descendant populations in North America are identical to those in founding populations but in some South American communities many new HLA-B alleles appear to have arisen, mostly by
inter-allelic conversion. These new alleles may have
been selected in response to novel parasite loads. The total number of alleles in each community remains small
but the HLA-B alleles differ between communities. Together, these findings suggest that the high allele num-
bers that characterize urban populations are more a reflection of migration and population admixture rather than selection [66].

To remain effective any defense strategy needs continu-
ous innovation. Examples are emerging of genetic mech-
anisms to underpin such demands on the immune sys-
tem. Natural killer receptors differ dramatically between different mammalian species. The two most widely studied species, man and mouse, have exploited genes from different superfamilies — Ig superfamily and lectin super-
family, respectively — to carry out analogous tasks. As touched on above, some key properties of the receptors are that they are expressed clonally on NK cells, that they engage MHC class I and that they give either posi-
tive or negative signals through ITAM or ITIM motifs, respectively. These properties are shared by the Ly49 family of lectin-like receptors mice and the KIR family of Ig-superfamily receptors in man. Comparisons of differ-
ent species for their complement of NK receptor loci reveal a rapid evolution for some receptors, like KIR and Ly49, but not for others like CD94 and NKG2D. Some species appear to use both KIR and Ly49 [66a]. Most pri-
mates, with the possible exception of orangutan but
including man, rely on KIR, and the vestigial Ly49 copy is a pseudogene. Different primates carry widely different sets of KIR genes, many of which are species specific. These findings are consistent with evolution of NK receptors at a rate even faster than MHC class I, which
between chimpanzee and man is well conserved, although KIR differ between these two species [67].

Some of the most interesting defense strategies seem to have evolved by harnessing non-immune, civilian func-
tions, an example being the novel use of DNA deaminase to destroy viral sequences, a proposed innate defense system [68]. It may turn out that retroviruses such as HIV exploit this to generate a high level of variation in viral sequence to throw up class I escape variants [69]. Defense demands creativity and competition, the innovation or stagnation proposed in the Red Queen hypothesis [70] because many pathogens have sophisticated and rapidly evolving evasion mechanisms, some of which employ mechanisms not too dissimilar to those of the host. The human malaria parasite Plasmodium falciparum for example varies the antigenic character of infected erythrocytes by selective expression of var genes from an extremely large, rearranging gene family [71].

Co-evolution

A successful general knows the importance of communi-
cation and in modern warfare the TV and radio transmit-
ters are usually amongst the first targets of an aggressive campaign. How is communication achieved between

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA, Weinheim www.eji.de
highly networked gene systems that probably use many more layers of sophisticated communication and cooperation than any military operation? It is generally believed that the modest increase in gene number between worms and humans is not sufficient to account for their different phenotypes and that complex, subtle interactions between genes, and between their products, are at the heart of the differences [72]. At present we have few tools to dissect such interactions, which may include any system we can imagine, such as proteins in large complexes [73], signaling pathways, transcription factors, epigenetics and differential splicing.

Innovation is key to evolution and the immune system is full of examples of novel use of a pre-existing protein. Some of the most interesting are those genes that span a variety of species. The NOD/CARD gene family, for example, performs defensive roles in plants as animals [25]. In the former it is associated with wailing off infected portions of the plant and in vertebrates with apoptosis.

Association with disease

To survive it is important to learn from retreats and defeats as well as famous victories. Immune-defense genes by definition are important for resistance to infection and this has been confirmed in over 80 primary human immunodeficiencies, where loss of diverse properties that include transduction, cell surface phenotype, transcription or phagocytosis results in vulnerability to infection [74]. Often, these defects lead to susceptibility to a precise range of organisms, such as mycobacteria in individuals with mutations in the interferon pathway [75]. For reasons not understood, individuals compromised for loading peptides onto class I molecules are vulnerable to Pseudomonas colonization of the lung [76]. For many of the defense genes, including some of the large, polymorphic sets of genes discussed here, it has been difficult to demonstrate a direct association with disease resistance. Many MHC class I and class II molecules are associated with different autoimmune diseases but it has been difficult to determine which infectious diseases are responsible for driving the variation. One previously advanced explanation is that, for humans, prevalent diseases that drove MHC variation are a thing of the past [77]. The recent AIDS, BSE and SARS epidemics have shifted such views on our invulnerability to infection in modern cities.

In fact, there are now some good examples of associations of single MHC alleles with infectious disease resistance or susceptibility, such as HLA-B*27 and *57 and slow progression to AIDS [78]. But, it may be difficult to show a direct association in most cases because an allele that initially provides resistance eventually predominates. A common MHC allele is more likely to become a susceptibility locus for the next wave of pathogen that is continuously coming up with escape variants, as in the ‘frequency-dependence’ model. This has been proposed for the association of malaria with different HLA-B loci in different parts of Africa [79]. The HLA-B*5301 allele in West Africa, derived from *3501 by a gene conversion event, provides some protection against malaria. In East Africa the malaria parasite seems to have adapted such that *5301 is no longer protective. The benefits of MHC polymorphism seem to be subtle but in the long-term the combination of gene conversion and weak selection are effective.

The complexity of association of variation in immune defense genes with disease is also illustrated by Fc receptor variation in mice. Two alleles of FcγRIIb (CD32), an inhibitory IgG receptor, are prevalent in wild mice. The allele expressing a low level results in enhanced antibody responses and macrophage activation. This promotes resistance to bacteria but at the cost of increase in septic shock after infection. It may also increase autoimmune conditions such as systemic lupus erythematosus. Conversely, the animal carrying the high-level allele is resistant to septic shock but susceptible to infection, and is less likely to suffer from autoimmunity (Clatworthy and Smith, personal communication).

This trade-off between resistance to disease yet susceptibility to autoimmunity, the friendly fire of immune defense, may predominate where polymorphic immune receptors are involved. Immunity to infection could be balanced to provide an adequate response for defense whilst avoiding the consequences of excessive inflammation. An explanation proposed for the insidious increase of allergies and autoimmune diseases in western populations is that reduced exposure to infection during childhood can make the immune system more likely to overreact to infection or innocuous environmental antigens later in life. There is a trade-off between effective defense and self-damage. According to the hygiene hypothesis [80], lack of infections in this age of cleanliness, vaccination and antibiotics, means that key regulatory networks between cells and genes which prevent allergy and autoimmunity are poorly adjusted, producing an army that, having been trained at a time of peace, is ill equipped at distinguishing friend from foe.

How can evidence for epistatic interactions, such as those proposed between MHC class I and KIR, be identified and studied? As discussed above, infectious disease is one of the most difficult areas to probe epistatic interactions because the interactions are all finely tuned by evolution. Nevertheless, data have been obtained
which show such interactions in models of both infection [81] and autoimmunity [82]. Transplantation provides a more demanding test situation as unusual combinations of polymorphic products are juxtaposed unnaturally. Evidence from transplants is indeed starting to demonstrate ways in which combinations of MHC class I and KIR are detrimental or can be exploited to combat leukemia [83].

Networking

The immune system is capable of recognizing every potential threat. As explained earlier, the adaptive immune system achieves this by creating vast repertoires of receptors at random and keeping the useful ones. Traditionally, it was believed that innate immunity comprised a much simpler system of receptors, such as TLR, each for a separate pathogen product. This appears to be an oversimplification. Subtle collaboration and networking between groups of receptors, signaling pathways and protein complexes appear to be needed to control and define the net activation state of a cell and an immune response, when encountering a potential pathogen or a related attenuated strain [84, 85].

Responding to might with might can result in escalation of a problem that spirals out of control. Some immune responses are measured; examples including tuberculosis and leprosy and worm infections, where a degree of infection is contained in a kind of stand-off, since excessive inflammation can be life-threatening. The peaceful coexistence of host and commensal is worth striving for.

References


50 Hawley, R. S., The Human Y chromosome: rumours of its death have been greatly exaggerated. Cell 2003. 113: 825–828.


Tanimachi, D. M., Hanke, T., Takizawa, H., Jamieson, A. M., Belkin, D., Torkar, M., Chang, C., Barten, R., Tolaini, M., Zuckerkandl, E.,...