CHAPTER 13

Population genetics of insect immune responses

Punita Juneja and Brian P. Lazzaro

13.1 Introduction

The immune system that we can observe and measure today is but a snapshot of a dynamic and evolving process, a moment in an ongoing genetic battle between hosts and their pathogens. Indications of this conflict are etched in the genome as signatures of adaptive evolution in the host immune system. These evolutionary signatures can also be read experimentally to give insight into the nature of host–pathogen interactions. This chapter will examine the evolutionary genetics of insect immune systems over both short and long timescales. In several instances, comparisons and contrasts will be drawn between species with distinct ecologies to elucidate commonalities and idiosyncrasies of insect immune evolution.

Adaptive evolution can manifest in evolutionarily favoured amino acid substitutions within genes as well as in genomic diversification of gene families. Both processes can be measured by comparing homologous genes and gene families across related species. Adaptive amino acid evolution is generally detected as a significantly elevated rate of amino acid substitution relative to an expectation based on the evolutionary rate at genetically silent positions (Box 13.1; Anisimova and Liberles, 2007). Adaptive gene family expansion can be inferred from an increased rate of duplication relative to that of other gene families in the genome (Hahn *et al*., 2005). The recent availability of whole-genome sequences from several insect species allows such comparisons to be made on a wide scale.

Innate immunity, which is shared by homology between vertebrates and insects, is hardwired within the genome and lacks the antibody production that characterizes the adaptive immune response of higher vertebrates. The insect innate immune system is capable of recognition and subsequent eradication of microbes and multicellular parasites through humoral and cellular defence mechanisms (reviewed in Lemaitre and Hoffmann, 2007). Humoral immunity is mediated by production of microbicidal peptides, enzymes, oxidative free radicals, and other compounds that are secreted directly into the insect haemolymph (blood). The humoral defence against microbial infection is genetically well understood in *Drosophila melanogaster*. Invading microbes are detected by recognition molecules performing surveillance, signal is transduced through two primary signalling pathways, and defence is effected in part by abundantly produced antimicrobial peptides (AMPs). The two signalling pathways, termed the Toll and Imd pathways, are conserved between invertebrates and vertebrates. Cellular immunity is defined by encapsulation or engulfment of infective agents by circulating haemocytes. It has been less well characterized at the genetic level, although some genes that mediate cellular recognition and trigger phagocytic engulfment of microbes have been identified. A distinct process, RNA silencing (RNA interference, RNAi), allows specific detection and eradication of RNA viruses (Wang *et al*., 2006). It is expected that functional diversity within the immune response will translate into variation in the selective pressures on different components of the defence response. This chapter will examine the evolutionary genetics of immune defence, interpreting molecular evolutionary patterns in light of protein function to

Box 13.1 Detecting adaptation in the genome

Evidence for natural selection can be revealed through examination of rates of DNA sequence evolution (reviewed in Anisimova and Liberles, 2007). The null model for these studies is that genes evolve by neutral evolutionary processes. Neutral, selectively equivalent mutations arise by chance and, in the absence of natural selection, occasionally become fixed by random genetic drift. The rate with which this happens is the neutral substitution rate. Many mutations that arise within functional genes cause deleterious changes to protein structure or function. These mutations are constrained from rising to high frequency by negative, or purifying, selection and are assumed to rarely fix between species. In contrast, mutations that are advantageous, such as those that confer resistance to disease, may rapidly rise in frequency by positive, or directional, selection. Positive selection leads to a short-term reduction in genetic diversity as the favoured allele replaces existing variation in a population. A sufficiently high number of recurrent adaptive fixations may also increase long-term divergence between species. Alternatively, multiple polymorphisms can be maintained in populations by balancing selection, which increases genetic diversity. In very rare cases, balanced polymorphism can occur when there is a heterozygote advantage, or overdominance, where heterozygote combination of two alleles has a higher fitness than homozygotes of either allele. More frequently, temporal or spatial variation in selection can maintain multiple alleles if each variant is advantageous in a different time or place.

Adaptive evolution can be detected by comparing DNA sequence of homologous genes from closely related species. This is generally achieved by comparing the rate of non-synonymous, amino acid-replacing, substitutions (d_N) to the rate of synonymous substitutions (d_S), which do not affect amino acid sequence. Synonymous substitutions are assumed to be invisible to selection and thus reflect neutral evolution. If all non-synonymous mutations were also selectively neutral, d_N would equal d_S, and the ratio d_N/d_S would equal one. Positive selection on amino acid substitutions would result in an increase in the rate of non-synonymous substitutions, or d_N being greater than d_S. The ubiquity of purifying selection, however, means that the empirically observed rate of non-synonymous substitutions over whole genes is much smaller than the rate of synonymous substitution, and d_N/d_S is almost always much less than one across entire genes. A more sophisticated implementation of this test, phylogenetic analysis by maximum likelihood (Yang *et al.*, 2000), uses gene sequences from multiple species to test the hypothesis that d_N/d_S varies among codons in a gene, allowing localization of the target of selection to particular residues or gene regions. Another test for natural selection, the McDonald–Kreitman test (McDonald and Kreitman, 1991), uses information about polymorphism in species and divergence between species. It tests the null hypothesis that the ratio of non-synonymous and synonymous substitutions segregating within species is the same as the corresponding ratio between species. In this test, positive selection is detected as a proportional excess of non-synonymous fixed differences between species. Selection favouring allelic diversification within species, in contrast, would lead to an excess of non-synonymous polymorphisms. These tests, among others, allow inference of natural selection acting on specific genes and gene regions.

draw insight into how the immune response adapts to pathogen pressures.

13.2 Evolutionary patterns in the antimicrobial immune response

Immune genes tend to evolve more quickly and adaptively than non-immune genes in vertebrates and insects (Murphy, 1991; Schlenke and Begun, 2003; Nielsen *et al.*, 2005; Sackton *et al.*, 2007; Waterhouse *et al.*, 2007). This adaptive evolution is shown by elevated rates of amino acid substitution between species and by elevated rates of duplication within gene families. The availability of whole-genome sequences allows for quantitative contrasts to be made between immune and non-immune genes, as well as for comparisons between functional classes of immune response genes. The recent complete genome sequencing of 12 species of fruit flies in the genus *Drosophila* has allowed particularly fine measurement of rates of substitution and genomic rearrangements between

closely related species. More distant comparative genomic analyses can be achieved by comparing genome sequences of *Drosophila*, the mosquitoes *Anopheles gambiae* and *Aedes aegypti*, the honey bee *Apis mellifera*, and the red flour beetle *Tribolium castaneum* (Figure 13.1).

Genome comparisons between species reveal the distinct selective pressures acting on each species through its unique life history. For example, the honey bee *A. mellifera* has apparently reduced copy number in immune-related gene families, perhaps reflecting decreased emphasis on immunological defence due to hygienic behaviour in the hive (Evans *et al*., 2006). Mosquitoes have expansions in gene families thought to play defensive roles against pathogens borne in vertebrate blood (Christophides *et al*., 2002; Waterhouse *et al*., 2007). Interpretation of these comparisons is often limited, however, because identification of most immune genes in insects stems from functional characterization in only a few species, and primarily in *D. melanogaster*. Novel defence mechanisms in functionally uncharacterized organisms will not be detected through homology searching of genome sequences if they are too divergent to be detected by similarity at the DNA sequence level. Additionally, genes that are evolving extremely rapidly may diverge too quickly to be identified in comparisons between distantly related species. Genomic comparisons will gain power with

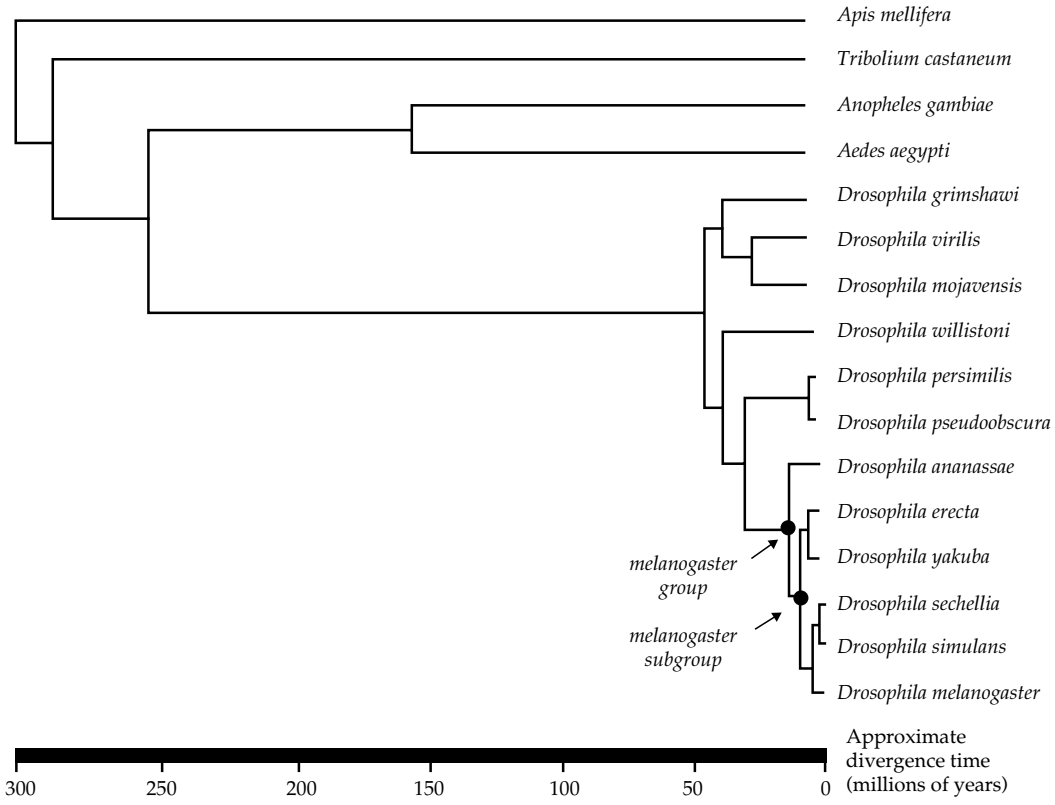

Figure 13.1 Phylogeny of select insect species with sequenced genomes. The *melanogaster* species group and *melanogaster* species subgroup are indicated. Gene-family expansions and contractions were evaluated among *Drosophila* (fruit flies), *Anopheles gambiae* (African malaria mosquito), *Aedes aegypti* (yellow fever mosquito), *Apis mellifera* (honey bee), and *Tribolium castaneum* (red flour beetle) and within the genus *Drosophila*. Adaptive amino acid evolution measurement, which requires shorter phylogenetic distances, was performed primarily in the *melanogaster* species group of *Drosophila* (Tamura *et al*., 2004; Savard *et al*., 2006; *Drosophila* 12 Genomes Consortium, 2007; Waterhouse *et al*., 2007).

increasing functional characterization of non-model systems and the accumulation of whole-genome sequences for phylogenetically dispersed organisms.

Comparative genomic and molecular evolutionary analyses have revealed that not all genes in the immune system evolve along the same trajectories. Genes in broadly defined functional categories differ in evolutionary mode, suggesting contrasting selective pressures based on gene function. The supporting data and potential selective pressures that drive these evolutionary patterns will be considered in detail.

13.2.1 Toll and Imd signalling pathways

Nearly all core signalling proteins in the Imd and Toll pathways are maintained as strict orthologues among *Drosophila* species (Sackton *et al*., 2007) and between *Drosophila* and mosquitoes (Christophides *et al*., 2002; Waterhouse *et al*., 2007), honey bees (Evans *et al*., 2006), and *Tribolium* (Zou *et al*., 2007). Despite this maintenance of orthology, however, these signalling genes show unexpectedly high levels of amino acid divergence between *D. melanogaster* and mosquitoes and considerable evidence of adaptive evolution within *Drosophila* (see Figure 13.2; see also Figure 6.3 in this volume; Schlenke and Begun, 2003; Jiggins and Kim, 2007; Sackton *et al*., 2007; Waterhouse *et al*., 2007).

The adaptive evolution of innate immune signalling pathways is illustrated dramatically by proteins in the Relish cleavage complex of the Imd signalling pathway (Figure 13.3). Relish is a nuclear factor κB (NF-κB) family transcription factor that is cytoplasmically bound in the absence of infection. Activation of the Imd signalling pathway leads to phosphorylation of Relish, caspase-mediated cleavage of the Relish inhibitory domain, and translocation of the activated transcription factor to the nucleus. Several proteins in the cleavage complex (Dredd, dFADD, IKK_{β}, IKK_{γ}, and Relish itself) appear to be evolving adaptively in *D. melanogaster*, *Drosophila simulans*, and/or the *melanogaster* species group. Adaptive mutations are disproportionately located in protein domains important for the release of activated Relish: the Relish autoinhibitory domain and cleaved linker, the Dredd caspase domain, the dFADD death domain, and the IKK_{β} kinase domain (Figure 13.3; Begun and Whitley, 2000; Schlenke and Begun, 2003; Jiggins and Kim, 2007; Sackton *et al*., 2007). Adaptive evolution of the Relish complex is not universal among *Drosophila*, but is restricted to certain species in the *melanogaster* group (Levine and Begun, 2007; Sackton *et al*., 2007). In an interesting parallel, the *Relish* gene of *Nasutitermes* termites also evolves adaptively, again with positively selected mutations localized in and around the caspase cleavage site and linker (Bulmer and Crozier, 2006), suggesting convergence of selective pressures in these distantly related insects. Nor is adaptive evolution in *Drosophila* restricted to the Relish complex. Many other signal transduction genes in the Imd and Toll pathways (*imd*, *spirit*, *persephone*, *Toll*, *dorsal*, *necrotic*) also show evidence of rapid evolution in *Drosophila* (Schlenke and Begun, 2003; Jiggins and Kim, 2007; Sackton *et al*., 2007).

One hypothesis to explain the preponderance of adaptive mutations in signalling genes is that at least some pathogens may actively interfere with host immune signalling (Begun and Whitley, 2000). Such pathogens could include bacteria that inject immunomodulatory molecules into host cells, immunosuppressive fungi and parasitoid mutualistic polydnaviruses (reviewed in Schmid-Hempel, 2008). In the Relish example, pathogen interference with the assembled cleavage complex could drive co-evolutionary adaptation in several proteins. Alternatively, interference with a single important member of the complex could drive adaptation in that member while promoting compensatory adaptations in the interacting proteins to retain host function. Such compensatory mutations may occur throughout the signalling pathway, amplifying the evidence of natural selection in this gene set (DePristo *et al*., 2005). The convergence of adaptive evolution of genes within the Relish complex in different insect species suggests that some of these genes are common targets of pathogens.

13.2.2 Antimicrobial peptides (AMPs)

The humoral immune response culminates in the production of effector molecules that kill invading microbes. One well-studied class of effector

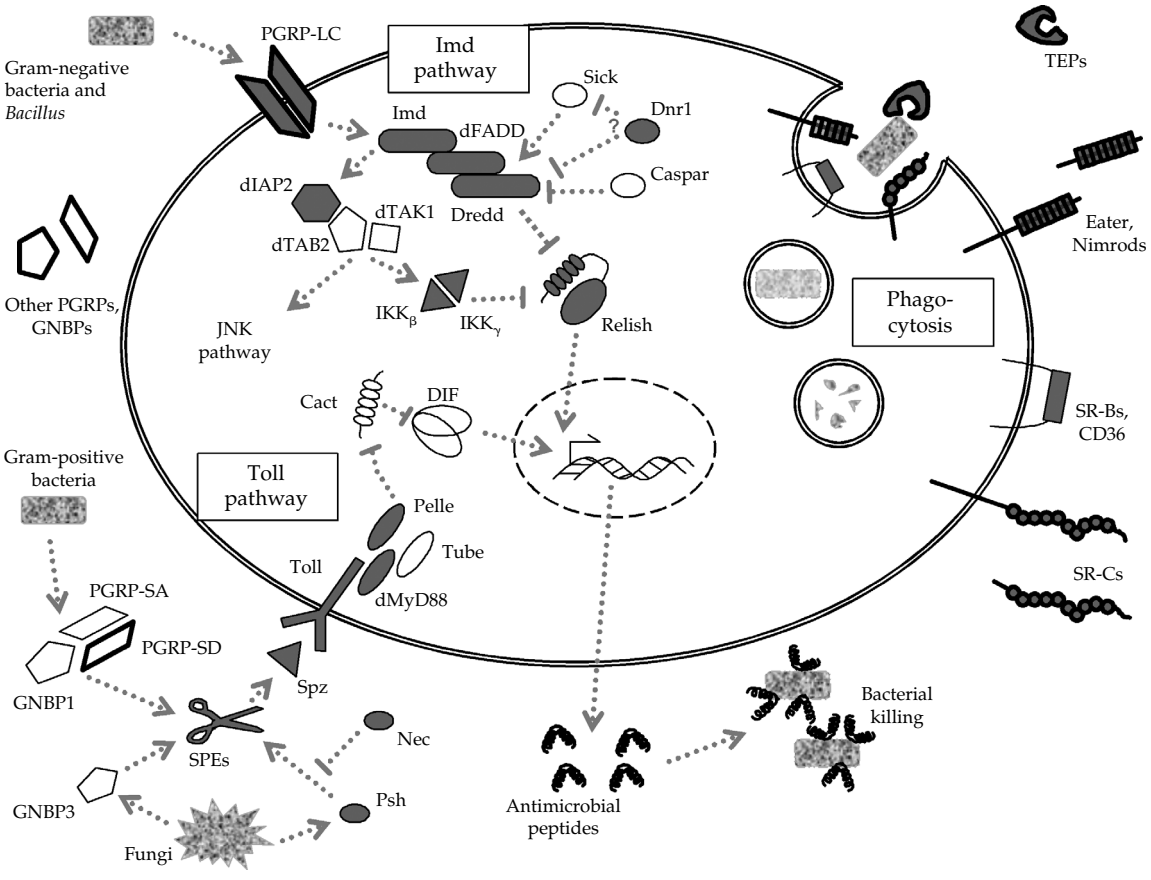

Figure 13.2 A schematic illustration of an idealized *D. melanogaster* immune-responsive cell illustrating prominent proteins required for the activation of a humoral immune response and receptors involved in defensive phagocytosis. Proteins whose gene families have experienced considerable genomic turnover within the genus *Drosophila* and among *Drosophila*, *Anopheles*, *Aedes*, *Apis*, and *Tribolium* are outlined in heavy black. Grey-shaded proteins have been implicated as evolving adaptively at the amino acid sequence level in *D. melanogaster* and/or *D. simulans*. Reproduced with permission from Lazzaro (2008). Cact, cactus; DIF, Dorsal-related immune factor; GNBP, Gram-negative-binding protein; IKK, I-κB kinase; JNK, c-Jun N-terminal kinase; Nec, Necrotic; PGRP, peptidoglycan-recognition protein; SPE, Spätzle-processing enzyme; Spz, Spätzle; TEP, thioester-containing protein.

molecules is AMPs. Most AMPs are short cationic peptides whose microbicidal activity is mediated by direct interaction with the negatively charged lipid membranes of bacteria and fungi (Zasloff, 2002; Lemaitre and Hoffmann, 2007; Yeaman and Yount, 2007). AMPs drew early attention as potential sites of host–pathogen co-evolution (Clark and Wang, 1997; Date *et al*., 1998; Ramos-Onsins and Aguadé, 1998) because of their direct role in the lysis and targeted killing of pathogens. However, systematic study of AMP genes, first in *D. melanogaster* and more recently across six *Drosophila* species, has failed to uncover evidence of adaptive evolution at the amino acid level (e.g. Lazzaro and Clark, 2003; Jiggins and Kim, 2005; Sackton *et al*., 2007). *Drosophila* AMP genes do, however, show extremely high rates of gene family expansion and contraction (Sackton *et al*., 2007). This high rate of genomic turnover extends to other taxa and is characteristic of most AMPs (Figure 13.2). In fact, the majority of *Drosophila* AMPs have no identifiable homologues in the genomes of mosquitoes, honey bees, or *Tribolium* (Christophides *et al*., 2002; Evans *et al*., 2006; Waterhouse *et al*., 2007; Zou *et al*., 2007).

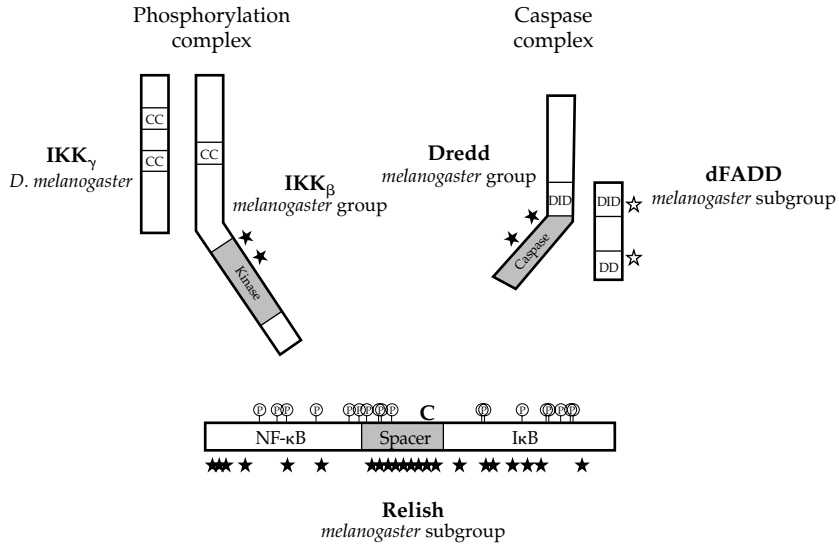

Figure 13.3 Adaptive evolution in the Relish complex. Caspase cleavage of the phosphorylated Relish spacer region allows the nuclear factor κB (NF-κB) domain to be translocated to nucleus, where it drives expression of immune response genes. IKK_γ and IKK_β form a complex through interaction at coiled-coil domains, and IKK_β phosphorylates Relish. The caspase Dredd is activated by dFADD via interaction at death-inducing domains and forms a complex with Relish. Putative Relish activation domains are indicated in grey. Positively selected sites (posterior probability >0.75) are indicated (★=significant at $P<0.01$; ☆=significant at $P<0.02$) and reflect selection along the *D. melanogaster* branch (*Relish*, *IKK*$_\beta$, or *dFADD*) or across the *melanogaster* (*Dredd*) species group (Sackton *et al.*, 2007). Taxonomic lineages where these genes appear to have evolved adaptively are indicated beneath each gene name (Begun and Whitley, 2000; Schlenke and Begun, 2003; Jiggins and Kim, 2007; Sackton *et al.*, 2007). C, caspase cleavage site; CC, coiled-coil domain; DD, death domain; DID, death-inducing domain; IKK, I-κB kinase; IκB, inhibitory κB; P, phosphorylation site.

Instead, these insects each have their own unique peptide families (Bulmer and Crozier, 2004; Evans *et al*., 2006; Waterhouse *et al*., 2007; Zou *et al*., 2007). In some cases, AMP families in different species converge independently on similar tertiary structures and presumably functions (Broekaert *et al*., 1995). Thus, whereas AMPs as a functional class of protein are ubiquitous among higher eukaryotes, there appears to be little homologous retention of peptides over evolutionary time.

The levels of sequence constraint seen in *Drosophila* do not characterize AMP evolution in all taxa. Genomic duplication of AMP genes is occasionally coupled with adaptive diversification at the amino acid level, presumably reflecting functional divergence (Tennessen, 2005; Yeaman and Yount, 2007). Genes encoding a termite-specific class of AMPs, *termicins*, have independently duplicated or triplicated in several termite species, with one duplicate typically sustaining mutations that decrease the polarity of the peptide (Bulmer and Crozier, 2004). These changes, which are driven by positive selection on amino acid sequence, result in peptides with divergent charges. Similarly, the mosquito *A. gambiae* has duplicated members within the *defensin* family (Dassanayake *et al*., 2007). Again, expansion is coupled with elevated rates of amino acid substitutions that change polarity, suggesting adaptive value to having two defensins with slightly different polar affinities. Previous studies in vertebrate AMP families have also found evidence of duplication coupled with positive selection, although in these cases peptide charge is maintained (reviewed in Tennessen, 2005; Yeaman and Yount, 2007). There is compelling evidence from insects and vertebrates that gene-family expansion can sometimes allow adaptive diversification of peptide function (Tennessen, 2005).

AMPs are remarkably efficient at combating infection. Resistance in microbes is seldom observed in nature, and, when it is, it tends to arise in specialized pathogens that are likely to be under

strong selective pressure to resist this form of defence (see Samakovlis *et al*., 1990; Zasloff, 2002). There are several possible explanations for why it may be difficult for most bacteria to evolve resistance. One common AMP mechanism is to disrupt membrane integrity though biochemically simple mechanisms, such as forming open pores (Zasloff, 2002; Yeaman and Yount, 2007). The ability of microbes to evolve resistance to such activities may be limited. However, heritable variation for resistance can be created and selected upon in microbial populations in the laboratory (Perron *et al*., 2006). In natural contexts, hosts simultaneously produce an array of AMPs that differ in charge, hydrophobicity, structure, and activity, probably ensuring that most pathogens are susceptible to at least a subset of them. This is conceptually similar to the application of multiple antibiotics in clinical settings and may serve to delay or eliminate the evolution of resistance (Yeaman and Yount, 2007). If pathogens are slow or fail to evolve resistance to peptides, there may be little selective pressure on insect hosts to adapt their AMPs at the amino acid level over modest evolutionary time. However, divergent bacteria and fungi display a range of susceptibilities to individual peptides (Zasloff, 2002), so diversification in AMP function may be selectively favoured in instances when a host shifts to a new ecological niche and is immediately presented with a novel and distinct set of pathogen pressures.

13.2.3 Recognition molecules in the humoral response

The humoral immune response is activated when circulating recognition factors are stimulated by highly conserved microbial compounds. Gram-negative-binding proteins (GNBPs) and peptidoglycan-recognition proteins (PGRPs) activate the humoral response after recognizing microbial cell-wall peptidoglycans and β-glucans. Some members of the PGRP family downregulate the immune response by degrading free peptidoglycan into non-immunogenic monomers (Lemaitre and Hoffmann, 2007).

PGRP and *GNBP* gene families generally evolve under purifying selection over short evolutionary time, but have undergone substantial genomic turnover on the lineages that separate *Drosophila* from mosquitoes, honey bees, and *Tribolium* (see section 6.4.4 in this volume; Evans *et al*., 2006; Waterhouse *et al*., 2007; Zou *et al*., 2007). Most *GNBPs* and *PGRPs* do not appear to have experienced recent adaptive evolution in *Drosophila* (Jiggins and Hurst, 2003; Schlenke and Begun, 2003; Jiggins and Kim, 2006; Sackton *et al*., 2007), mosquitoes (Little and Cobbe, 2005), or the crustacean *Daphnia* (Little *et al*., 2004). A notable exception, however, is a *Drosophila PGRP* which shows strong indications of adaptive evolution. *PGRP-LC*, an alternatively spliced gene that sits atop the Imd signalling cascade, has sustained a two amino acid insertion in the PGRP-LCa isoform in species of the *melanogaster* subgroup. This insertion is predicted to alter the binding specificity of that isoform, and appears to have been positively selected in conjunction with several additional adaptive substitutions (Sackton *et al*., 2007). Interestingly, the alternatively spliced binding domains of *PGRP-LC* show evidence of either recent independent duplication or concerted evolution in *D. melanogaster* and *A. gambiae* (Christophides *et al*., 2002). These patterns potentially reflect lineage-specific selection for recognition of distinct microbes. In another exception, limited positive selection was also detected in *GNBP* genes of *Nasutitermes* termites (Bulmer and Crozier, 2006). In this case, it was hypothesized that adaptation of recognition capability was driven by a shift in ecology as previously herbivorous termite species adapted to feed on decaying matter, exposing them to a novel community of pathogens.

One potential explanation for the observation that *PGRPs* and *GNBPs* tend to exhibit little indication of adaptive amino acid evolution is that these proteins recognize highly conserved pathogen sugar moieties. The cell-wall components recognized by these proteins are indispensable for most microbes, and, generally speaking, may not be easily modifiable. There thus may be little pressure on these genes to adapt over short time periods. Additionally, these recognition proteins are active against molecules that are conserved across a wide range of microbial taxa. There are, however, a limited number of examples of positive selection on *PGRPs* and *GNBPs*. Coupled with the observations of gene family duplication and divergence among

species, instances of positive selection may reflect bursts of diversification as recognition function is fine-tuned to species-specific selective pressures.

13.2.4 Recognition molecules in the cellular response

Recognition is also a necessary prerequisite for pathogen clearance via cellular immunity, and several gene families have been identified that encode membrane-bound phagocytic receptors. Phagocytosis is also promoted by 'tagging' of microbes with extracellularly secreted opsonins. Several genes encoding both phagocytic receptors and opsonins show evidence of adaptive amino acid evolution within the genus *Drosophila* (Sackton *et al*., 2007) and frequent genomic turnover within *Drosophila* and between *Drosophila* and other insects (Figure 13.2; Evans *et al*., 2006; Sackton *et al*., 2007; Waterhouse *et al*., 2007; Zou *et al*., 2007). In *Drosophila*, recognition genes are significantly more likely to show evidence of positive selection than genes with signalling or microbicidal functions (Sackton *et al*., 2007). This difference is largely driven by recognition genes that trigger the cellular response, with nine of 10 recognition genes that yield significant evidence of positive selection having been either experimentally confirmed to be involved in phagocytosis or homologous to known phagocytosis genes. Specifically, these are genes encoding thioester-containing proteins (TEPs) (Jiggins and Kim, 2006; Sackton *et al*., 2007), the Eater and Nimrod families (Sackton *et al*., 2007), the class C scavenger receptors (Lazzaro, 2005), and the CD36 homologue epithelial membrane protein (emp) (Sackton *et al*., 2007).

TEPs have been directly implicated as opsonins mediating the cellular clearance of microbes including bacteria and malaria-causing *Plasmodium* in *Drosophila* and *Anopheles* (Levashina *et al.,* 2001; Blandin *et al*., 2004; Stroschein-Stevenson *et al*., 2006). Proteolytic cleavage of a hypervariable spacer, or 'bait', domain exposes the thioester motif, which then covalently binds microbes and labels them for phagocytosis. TEPs appear to be hotspots of adaptation in several species. In *D. melanogaster*, there are six *Tep* genes, four of which have intact thioester domains and thus are likely to function as opsonizing agents (Blandin and Levashina, 2004). One of four of the intact *Teps* show evidence of adaptive divergence between *D. melanogaster* and *D. simulans* and three show evidence for directional selection in the *melanogaster* species group (Table 13.1; Jiggins and Kim, 2006; Sackton *et al*., 2007). Interestingly, one of the adaptively evolving *Tep* genes is constitutively expressed at higher levels in European than African populations of *D. melanogaster*, suggesting that expression of this *Tep* may be locally adapted (Hutter *et al*., 2008). *Tep* genes in mosquitoes and the more distantly related crustacean *Daphnia* also show evidence of adaptive amino acid evolution (Little *et al*., 2004; Little and Cobbe, 2005). In all cases, positively selected amino acid mutations are overrepresented in the bait domain that is cleaved to expose the thioester motif. It is unknown whether the proteases that cleave TEPs are produced by host or pathogen, so it is not yet possible to say whether adaptation in this domain is due to co-evolution with pathogen proteases or with pathogen molecules that interfere with host proteolysis.

Tep gene families are expanded in mosquitoes, with 13 *Tep* genes found in the *Anopheles gambiae* genome and eight in the *Aedes aegypti* genome (Christophides *et al*., 2002; Waterhouse *et al*., 2007). The expansions in size of the *Tep* gene family appear to have been independent in each of these two taxa and potentially reflect elevated pressure on cellular immunity. The *A. gambiae Tep1* gene is segregating for two sharply divergent alleles, one of which, when homozygous, confers absolute resistance to experimental infection with the rodent malaria *Plasmodium berghei* (Blandin *et al*., 2004; Baxter *et al*., 2007). Individuals homozygous for the susceptible allele sustain robust *P. berghei* infections. These two alleles differ by multiple amino acid substitutions, including several that are clustered around the thioester domain. It is currently unclear which substitutions cause the phenotypic differences in susceptibility, or whether it is an epistatic phenotype involving substitution in multiple domains of the protein. Both alleles are found at high frequencies in natural populations (Obbard *et al*., 2008), suggesting selective forces maintain these two alleles in the wild. This system provides a tantalizing opportunity to understand the mechanisms

Table 13.1 Evolutionary genetics of the *Tep* gene family of phagocytic recognition molecules in *Drosophila*.

	Tep 1	*Tep 2*	*Tep 3*	*Tep 4*	*Tep 5*	*Tep 6* (*Mcr*)	Reference
Functional data							
Overview	Upregulated in response to infection	Upregulated in response to infection		Upregulated in response to infection	Not expressed; likely to be a pseudogene	Lacks a thioester domain	Reviewed in Blandin and Levashina (2004)
Phagocytic activity		Required for efficient phagocytosis of the bacterium *Escherichia coli*	Required for efficient phagocytosis of the bacterium *Staphylococcus aureus*			Required for efficient phagocytosis of the fungus *Candida albicans*	Stroschein-Stevenson *et al.* (2006)
Species divergence:							
d_N/d_S (Box 13.1)	Exceptionally elevated d_N/d_S between *D. melanogaster* and *D. simulans* clustered around the bait domain; elevated d_N/d_S in the *melanogaster* species subgroup	Elevated d_N/d_S in the *melanogaster* species subgroup	Not significant[a]	Not significant			Jiggins and Kim (2006)
	Elevated d_N/d_S in the *melanogaster* group with trend towards an excess of positively selected sites at the bait domain	Elevated d_N/d_S across the entire gene in the *melanogaster* species group	Not significant	Elevated d_N/d_S in the *melanogaster* species group with an excess of positively selected sites at the bait domain		Not significant	Sackton *et al.* (2007)
McDonald–Kreitman test (Box 13.1)	Elevated amino acid replacements across entire gene in *D. melanogaster*	Not significant		Not significant			Jiggins and Kim (2006)
Population divergence							
Differential expression		Expression levels significantly higher in European than in African *D. melanogaster* populations	Not significant	Not significant			Hutter *et al.* (2008)

[a]Not significant indicates genes that were included in the referenced studies but not found to depart from the null expectation. Empty cells indicate that no information has been obtained.

that lead to the maintenance of immune response polymorphisms in a natural context.

Whole-genome comparisons within the genus *Drosophila* indicate that, in striking contrast to recognition molecules that trigger the humoral response, recognition molecules that initiate the cellular response show abundant evidence of adaptive evolution. Deeper investigation of the *Tep* gene family reveals that adaptive evolution extends beyond *Drosophila* to include mosquitoes and *Daphnia*, and demonstrates extant functional variation in a mosquito *Tep* gene. The signals of adaptive evolution suggest that these recognition molecules interact with evolutionarily labile pathogen motifs or that, like signalling molecules in humoral defence, they are potentially subject to interference by pathogen-produced proteins.

13.2.5 Summary

The diverse evolutionary trajectories of various genes in the insect immune response (Figure 13.2) can be interpreted in light of their molecular functions and interactions with pathogens. Pathogen-recognition molecules that stimulate the humoral response interact with highly conserved microbial cell-wall components. Although obligate pathogens are sometimes able to reduce their cell walls to escape detection, most microbes are evolutionarily constrained because they must also be able to persist in non-infectious environments. Similarly, there may be few ways in which microbes can evolve resistance to AMPs, especially when host insects simultaneously employ multiple peptides with distinct activities. If there is little adaptation in pathogens to escape host humoral recognition and antibiotic killing, then it may be expected that there would be little indication of adaptive amino acid evolution in the host genes over short evolutionary time. Both humoral recognition factors and AMPs exhibit rapid rates of genomic duplication and deletion, and in some taxa duplication is coupled with a burst of amino acid diversification that presumably increases breadth of function.

In contrast, signal transduction proteins in the humoral immune response are largely maintained in strict orthology across insect species, but frequently show indications of adaptive amino acid evolution within species. A hypothesized explanation is that the strong maintenance of orthology in these pathways makes them attractive targets for immune suppression by generalist pathogens. This may be a particularly successful strategy for microbes that are unable to evade or resist the recognition and microbicidal stages of humoral immunity. Gene duplication and diversification are not commonly observed here, perhaps because this is not a successful strategy for escaping pathogen interference. Genomic retention of a duplicated gene that can be manipulated by pathogens would be detrimental because host signalling function would be impaired. Instead, rapid fixation of amino acid 'escape' variants in signalling genes seems to be the most effective host strategy, and coordinate compensatory mutation in physically interacting proteins may amplify the signal of adaptive evolution in this functional category.

Recognition factors and opsonins in the cellular immune response evolve by adaptive amino acid evolution and frequent genomic turnover. In general, little is known about the specific activities and recognition profiles of these genes, making it difficult to interpret the evolutionary patterns in a functional context. The evolutionary genetics, however, do lead to functional predictions, including that the cellular recognition factors bind evolutionarily labile pathogen epitopes or are subject to pathogen interference, both of which could drive rapid amino acid evolution. At the moment, virtually nothing is known about the molecular evolution or population genetics of host genes that drive phagocytosis after pathogen recognition. Microbes are capable of manipulating host cells both to promote and inhibit phagocytic uptake (Schmid-Hempel, 2008), leading to the prediction that genes encoding the machinery of phagocytosis will, like genes in humoral signalling pathways, show abundant evidence of adaptive evolution.

13.3 Evolutionary patterns in the antiviral immune response

Early characterization of the immune response focused primarily on antimicrobial defence. Antiviral defence is at least partially distinct from

that against microbes, and currently is only poorly understood. Both the Toll and Imd pathways are activated during the course of some viral infections; however, only the Toll pathway seems to confer protection (Lemaitre and Hoffmann, 2007). RNAi provides an independent mechanism of defence that is specific against RNA viruses (Wang *et al*., 2006). Viruses are formidable opponents for the immune system. They are capable of rapid evolution owing to their fast generation times, large population sizes, high mutation rates and obligate pathogen lifestyles. These factors hint that the evolutionary patterns of antiviral defence genes will be different from those described previously for the antimicrobial defence.

Short-term evolution of an antiviral defence gene has been studied at the *D. melanogaster* locus *ref(2)P*, which is proposed to function in the Toll pathway (Avila *et al*., 2002). This locus is polymorphic for alleles that explain a large component of the variation in susceptibility or resistance to the rhabdovirus Sigma (Contamine *et al*., 1989; Bangham *et al*., 2007, 2008). A single domain, termed PB1, of *ref(2)P* is required for viral replication (Carré-Mlouka *et al*., 2007). Sigma is infective if a permissive allelic variant of this domain is present, but not with a restrictive allele or genetic knock-out of the domain. This domain has an excess of amino acid polymorphisms (Wayne *et al*., 1996), consistent with natural selection acting to maintain allelic diversity. A random sample of 10 phenotypically random alleles identified six amino acid polymorphisms in the PB1 domain (Wayne *et al*., 1996). A single complex mutation, with a single glycine residue substituted for glutamine and asparagine residues, was found on restrictive but not permissive alleles. The remaining polymorphisms are shared by both restrictive and permissive alleles. The frequency of the complex mutation varies between populations, ranging from absent in some African and European populations to 23% in some North American populations (Bangham *et al*., 2007). There is greatly reduced variation in the restrictive haplotype in a North American population, suggesting that it has recently risen to high frequency by directional selection (Box 13.1). This indicates that selection is acting on localized spatial scales, likely in concert with Sigma virus, which also varies in frequency and genotype between populations (Carpenter *et al*., 2007).

The fact that there is an excess of non-synonymous polymorphism in *ref(2)P* PB1 domain but that only a single complex mutation separates restrictive and permissive alleles suggests that current Sigma virus populations have become adapted to some of the remaining polymorphisms. Indeed, analysis of all combinations of polymorphisms on the restrictive allele in artificially generated constructs indicates that no fewer than two of the three mutations are required to create a restrictive allele (Carré-Mlouka *et al*., 2007). These data suggest a model wherein novel mutations have been driven to high frequency by directional selection, but that the sweeps are incomplete because the virus quickly adapts to the increasingly common allele before it fixes in the population. Host resistance then requires the repeated reintroduction of novel restrictive mutations. The most escalated rates of evolution are expected when host and pathogen are co-evolving, such that host adaptations to escape infection are met by a gene-for-gene pathogen adaptation to maintain virulence (Dawkins and Krebs, 1979). Over the evolutionary long term, there is evidence for elevated amino acid substitution at this domain, with more adaptive mutations becoming fixed in *D. melanogaster* when compared with *D. simulans*, a species in which Sigma infection is rare or absent (Wayne *et al*., 1996). Restrictive polymorphisms that are driven to high frequencies during partial selective sweeps will fix by genetic drift more often than mutations that are selectively neutral over their entire evolutionary history, which may lead to elevated amino acid divergence between species.

A distinct pathway using RNAi presents an important defence against RNA viruses. In *D. melanogaster*, double-stranded viral RNA (dsRNA) is recognized and cleaved into small interfering RNA (siRNA) by Dicer-2 (Wang *et al*., 2006). These siRNAs then guide cleavage of matching RNA via formation of an RNA-induced silencing complex (RISC). Some viruses produce proteins that suppress RNA silencing. For example, *Drosophila* picornavirus C produces a dsRNA-binding protein that interferes with Dicer-2 activity and promotes viral establishment and proliferation (van Rij *et al*., 2006). *Dicer-2*, along with RISC genes *R2D2* and *Argonaute-2*,

are among the most rapidly evolving genes in the *D. melanogaster* genome. These antiviral genes, but not their paralogues with housekeeping regulatory function, show indications of adaptive evolution by recurrent fixation of novel amino acid mutations (Obbard *et al*., 2006).

The unique patterns of evolution of antiviral defence yield a useful system for integrating measures of short- and long-term evolution. In the case of *ref(2)P* in *D. melanogaster*, rapid evolution is driven by a gene-for-gene interaction between host and virus, and is evidenced by reduced genetic variation within the selectively favoured allele in the short term and increased amino acid divergence in the long term. Rates of long-term evolution in RNAi antiviral genes in *D. melanogaster* are dramatically higher than the genome average. Evidence suggests that the selective pressures are different from those that act on antimicrobial defence, leading to elevated rates of evolution. This may reflect either rapid viral evolution or high host specificity in viruses, either of which would facilitate co-evolution. Like humoral signalling pathways in the antimicrobial defence, RNAi pathways are also subject to pathogen interference to overcome host defences, indicating that they too are a potential site of direct conflict. Thus, evidence from both types of defence suggests that sites of pathogen interference display elevated evolutionary rates. As antiviral defence becomes better characterized at the molecular level, this system will yield further insights into genetic adaptation to pathogen pressures and serve as a comparison for evolutionary patterns observed in antimicrobial defence.

13.4 From genotype to phenotype

All the patterns discussed thus far have pertained to the long-term evolution of the immune system. It is important to remember, however, that all adaptive evolution is based on phenotypic polymorphism that segregates in populations at some point in time. Indeed, extant natural populations harbour considerable genetic variation for immunocompetence. This segregating phenotypic variation is the substrate for short-term evolution. Understanding its genetic basis and the forces governing its persistence is essential for predicting the evolutionary response to natural or artificial perturbations in infectious pressure in natural populations.

In organisms with well-characterized genomes, it is possible to directly test the phenotypic effects of allelic variation in pre-chosen 'candidate' genes though genotype–phenotype association mapping. These studies have been employed most effectively in *D. melanogaster*. For instance, natural allelic variation in the *ref(2)P* gene clearly determines resistance to the vertically transmitted Sigma virus in *D. melanogaster* females in an almost purely Mendelian fashion (Contamine *et al*., 1989; Bangham *et al*., 2008). Genetic variation in Sigma viral transmission through males, however, does not map to *ref(2)P* (Bangham *et al*., 2008). Variation in the ability of *D. melanogaster* to suppress bacterial infection has been mapped to polymorphisms in pathogen-recognition factors and signalling genes within the Toll and Imd pathways (Lazzaro *et al*., 2004, 2006). Expression levels, but not polymorphisms, of AMPs are also associated with resistance to infection (T.B. Sackton, B.P. Lazzaro, and A.G. Clark, unpublished data). These observations, coupled with evaluation of transcriptional activity of the immune system, indicate that signalling flux through the Toll and Imd pathways is a tremendously important determinant of resistance to bacterial infection. In contrast to the antiviral resistance determined by *ref(2)P*, polymorphisms mapped in the antibacterial association studies each make relatively small contributions to variance in the resistance phenotype, suggesting that resistance to bacterial infection is a combinatorial function of multiple genes of individually small effect. Even in sum, the mapped antibacterial factors do not explain the entirety of the genetic variance, indicating that other unstudied genes also contribute to variation in resistance.

If pathogen infection can be so detrimental to the condition of the host, and host alleles that confer high resistance to infection exist in natural populations, why then does resistance not spread to all individuals? Genetic trade-offs, whereby immunocompetence comes at a cost to another phenotype within an organism, can constrain natural selection from fixing resistant genotypes (Roff and Fairbairn, 2007). Potential costs of resistance include direct damage to host tissues due to immune activity and correlated reduction in investment in

other physiological traits, including alternative immune functions, metabolism, and reproduction. Which investment strategy is most favourable will depend on the strength of pathogen pressures and on selection acting on other fitness traits of the organism.

An experimental approach that has been used to study genetic trade-offs is artificial selection for increased resistance to infection and subsequent measurement of correlated changes in other fitness traits. This method identifies costs of resistance, defined as changes in traits that reduce fitness in selected lines compared with unselected lines. Artificially selecting the Indian meal moth, *Plodia interpunctella*, for increased resistance to granulosis virus infection led to correlated increases in larval development time and pupal weight and a decrease in egg viability in selected lines (Boots and Begon, 1993). Selection in *D. melanogaster* for resistance to parasitoid or fungal infection led to a correlated decreases in larval competitive ability and adult fecundity, respectively, in the absence of infection (Kraaijeveld and Godfray, 1997, 2008). Costs that are measured in artificial selection lines should be interpreted with caution, however, as selection experiments can sometimes result in the fixation of rare alleles with large phenotypic effects that are not representative of functional genetic variation in natural contexts. For example, *A. gambiae* mosquitoes selected for refractoriness to *Plasmodium* infection achieve this through an increased melanization response (Collins *et al*., 1986) and high levels of cellular oxidative free radicals that are extremely damaging to host cells (Kumar *et al*., 2003). Natural resistance in wild populations of *A. gambiae*, however, is generally accomplished with a melanization-independent mechanism (Riehle *et al*., 2006), and is likely to be less costly or damaging than mechanisms seen in laboratory-selected lines.

A more relevant, but much subtler, measurement of genetic trade-offs is obtained by measuring genetic correlations between traits in naturally occurring, unselected genotypes. This is commonly done by measuring phenotypes in genetic clones or in individuals' with known genetic relatedness and estimating the genetic contribution to the phenotype. In *D. melanogaster*, genotypes with high resistance to bacterial infection had low fecundity in the absence of infection in a food-limited environment (McKean *et al*., 2008). In the pea aphid *Acyrthosiphon pisum*, clonal lines with high resistance to attack by the parasitoid wasp *Aphidius ervi* had reduced fecundity (Gwynn *et al*., 2005). However, in this case, resistance to parasitoids can be conferred by bacterial endosymbionts, so the genetic basis for this trade-off may be mediated by factors outside the host genome. In both examples, the cost of resistance is a decrease in reproductive fitness.

The ultimate goal is to identify the genetic architecture underlying trade-offs. Quantitative trait locus (QTL) mapping has been used to locate these genetic regions. This approach relies on contrived crosses between chosen parents to establish phenotypically variable recombinant progeny. Genetic markers are then genotyped at periodic intervals across the genome, allowing the localization of genomic regions encoding the phenotypic variation without relying on *a priori* candidate genes. QTL mapping, however, lacks the resolution to identify specific genes or alleles. In the red flour beetle *T. casteneum* and in the bumble bee *Bombus terrestis*, simultaneous mapping of immune and fitness traits found that loci associated with immune phenotypes occasionally co-localized with QTLs involved in fecundity, viability, and body size (Zhong *et al*., 2005; Wilfert *et al*., 2007a). There are two potential genetic mechanisms that could cause genetic correlations between immune and fitness traits. Genetic correlations can be caused by pleiotropy, where a single gene influences multiple traits. Trade-offs are due to antagonistic pleiotropy, where a single allelic variant of a gene has a positive effect on one trait but a negative effect on the other. Alternatively, allelic variants of distinct genes affecting the two traits may be in linkage disequilibrium due to physical proximity on a chromosome, and thus these variants are coordinately passed to the offspring. Selection acts simultaneously on traits that are correlated by either pleiotropy or linkage disequilibrium. However, only antagonistic pleiotropy places a long-term constraint on selection because recombination can degrade correlations based on linkage disequilibrium. QTL mapping relies on experimentally generated linkage disequilibrium that spans much

greater physical distances than are observed in natural populations, so it is relevant to follow QTL-based studies of genetic correlations with field-based studies to determine whether the traits co-segregate in nature.

Trade-offs have been also identified within the immune response. For example, in *B. terrestis*, lines selected for increased resistance to trypanosome infection also had a higher investment in a phenoloxidase response coupled with a lower investment in AMP response (Wilfert *et al*., 2007b). The Egyptian cotton leafworm, *Spodoptera littoralis*, demonstrated positive genetic correlations among haemocyte density, cuticular melanization, and phenoloxidase activity, but a negative genetic correlation between haemocyte density and lysozyme-like antibacterial activity (Cotter *et al*., 2004). A different result is obtained from females of the mealworm beetle *Tenebrio molitor*, where cuticular melanization shows a negative genetic correlation with haemocytes and phenoloxidase, suggesting that the genetic architecture of these correlations can vary between species (Rolff *et al*., 2005). These results demonstrate that increased investment in one component of the immune response can come at a cost to other immune functions, and indicate the potential for trade-offs within the immune response to place constraints on the evolution of global resistance.

Thus far, all resistance measures have been considered only in a single environment; however, the optimal immune strategy can be expected to vary based on environmental conditions (Lazzaro and Little, 2009). Selective pressures are heterogeneously distributed in the environment. Abiotic factors such as day length, temperature, and moisture vary between populations, affecting development time, metabolic flux, and other traits, and also altering the composition of pathogen communities and nutrient availability. Allelic variants in some genes respond differently to changes in the environment, termed genotype-by-environment interactions. If a genotype is particularly favoured in certain conditions, local adaptation to the proximate environment can occur. Temperate and tropical populations of *D. melanogaster* varied significantly in their resistance to the generalist fungal pathogen *Beauveria bassiana* (Tinsley *et al*., 2006) and bacterial pathogen *Providencia rettgeri* (Lazzaro *et al*., 2008). Considerable genotype-by-environment interaction was observed in resistance of *D. melanogaster* to *P. rettgeri* infection across multiple temperatures. Despite that observation, temperature populations were on average more resistant to *P. rettgeri* than the tropical one at lower temperatures, which potentially reflects adaptation to the local temperature. Spatial heterogeneity in the environment can lead to the maintenance of multiple resistance alleles if local adaptation is sufficiently strong to withstand erosion by migration and gene flow.

The magnitude, or even the existence, of genetic trade-offs can also vary between environments. In natural and laboratory settings, infestation by the mite *Macrocheles subbadius* negatively affects the fertility and body size of its host, *Drosophila nigrospiracula* (Luong and Polak, 2007). There is genetic variation for resistance to mites, which in this case is mediated by an avoidance behaviour. It has been demonstrated that, similar to *D. melanogaster* selected for parasitoid resistance, lines selected for mite resistance also suffer a cost in terms of decreased larval competitive ability. Manipulating the environment with high temperatures and increased larval density to create stressful conditions tends to increase costs of resistance. For instance, in previously considered examples from *D. melanogaster*, resistance to bacterial infection was correlated with low fecundity only in a food-limited environment (McKean *et al*., 2008), and larval success of parasitoid-resistant larvae was compromised only under crowded conditions (Kraaijeveld and Godfray, 1997). In all of these cases, selection can act independently on the traits in a non-stressful environment but the traits are constrained to each other under resource-limited conditions. Genetic variation for different allocations of resources between resistance and fitness traits can be maintained by environmental heterogeneity since the optimal investment strategy will be context-dependent (Roff and Fairbairn, 2007). Selection on these variants will be inefficient because trade-offs will only be apparent in certain conditions.

The host immune response faces a special obstacle in evolving immunity: the immune system must respond to living organisms that are them-

selves free to evolve. Its pathogen 'environment' is capable of rapid evolution, often much more quickly than the host. Analogous to genotype-by-environment interactions, a genotype-by-genotype interaction occurs when the efficacy of a host resistance genotype is dependent on the genotype of the pathogen. Antagonistic pleiotropy can occur in this context if resistance to one pathogen genotype comes with susceptibility to another. The specificity of these interactions can allow for temporal fluctuations in host and parasite genotypes in a frequency-dependent manner. Such fluctuations are generally difficult to measure experimentally, but have been observed natural populations of the snail host *Potamopyrgus antipodarum* and trematode parasite *Microphallus* sp. as well as in the crustacean host *Daphnia magna* and bacterial parasite *Pasteuria ramosa* (Dybdahl and Lively, 1998; Decaestecker *et al*., 2007). In both cases, resistant host genotypes are at an advantage when they are rare because their infective parasite genotypes are also rare, allowing resistant host genotypes to then to rise in frequency. This leads to a time-lagged increase in the infective parasite genotype, causing the host advantage to decline, subsequently reducing the frequency first of the host genotype and then the parasite genotype. This type of co-evolution is probably rare, occurring only when a parasite infects a narrow species range of hosts, allowing for specific, reciprocal adaptation, and when the parasite greatly reduces the fitness of the host such that selective pressure on resistance is high. In reality, many parasites are likely adapting to multiple host and impose only small reductions of fitness, placing more diffuse selective pressures on their hosts.

Environmental heterogeneity in pathogens and pathogen genotypes can lead to spatial adaptation to local pathogen pressures (Woolhouse *et al*., 2002). Genotype-by-genotype interactions between hosts and pathogens allow for adaptation to proximate pathogen pressures. Experimental evolution has been used to demonstrate the potential for local adaptation. In an experiment where *P. ramosa* was serially passaged for several generations on *D. magna*, it evolved high levels of infectivity on the host used for passage and in some cases lost virulence on non-passaged hosts (Little *et al*., 2006). This indicates that parasites can adapt to current hosts, perhaps at a cost of infecting alternate hosts, in only a few generations. Spatial variation in resistance can be detected by comparing the success of infection between host–parasite combinations that are either sympatric (local) or allopatric (foreign). Although most theoretical models predict that the parasite should be most successful in sympatric infections, in practice both parasite local adaptation and maladaption are observed (Woolhouse *et al*., 2002). In *A. gambiae*, a locus that was found to control encapsulation response to the malaria parasite *Plasmodium falciparum* was strongest against allopatric infections (Niaré *et al*., 2002). Another locus restricting infection intensity was strongest against sympatric infections. Despite the opposite directions of these responses, both findings demonstrate population variation in resistance. In some cases, host resistance and parasite virulence have been observed to covary. The parasitoid *Asobara tabida* has been reported to have the highest virulence in the Mediterranean and lower virulence in northern Europe (Kraaijeveld and Godfray, 1999). *D. melanogaster*, the host, was observed to have the highest resistance in the Mediterranean and southern Europe, and low resistance in northern Europe, evidence of adaptation to local parasitoid pressures.

Tremendous variation in immunocompetence exists in extant natural populations. Trade-offs within the immune response and between immunocompetence and other fitness components constrain the ability of natural selection to drive resistant genotypes to fixation. Variation in trade-offs is maintained in part by environmental variation, whereby the costs associated with a particular genotype are context-dependent. Genotype-by-environment interactions and local adaptation can potentially lead to the maintenance of multiple polymorphisms in heterogeneous environments. Furthermore, the pathogen 'environment' is itself evolving. These forces in combination oftentimes limit the evolution of a single globally resistant genotype.

13.5 Conclusion

Genes involved in the immune response show signals of rapid evolution, with the precise evolutionary mode varying among components of

the immune system. Extant populations harbour tremendous genetic and phenotypic variation in resistance, providing the substrate upon which selection acts. Examination of both evolutionarily ancient and current patterns has only rarely been performed. The most complete example is from the *ref(2)P* locus in *D. melanogaster*, which is polymorphic for the ability to permit or restrict Sigma virus infection. In natural populations, this locus shows evidence for elevated polymorphism, partial selective sweeps, and spatial heterogeneity in allele frequencies, all of which reflect an on-going battle between host and pathogen. These polymorphisms also often become fixed, driving long-term adaptive amino acid evolution. Other parts of the immune system could be equivalently studied, such as a polymorphic locus in the mosquito *A. gambiae* that confers resistance to malaria. In general, characterization of forces that facilitate or inhibit the spread of host resistance through populations, combined with genome-scale comparisons between species, will allow the linkage of short-term and long-term patterns to fully define the lability and constraint on adaptive evolution across the immune system.

Understanding the factors that influence the evolution of the immune response has important ramifications for diverse fields of study. Evaluation of the feasibility of applications such as the proposed engineering of transgenic disease-vector insects to control transmission and the use of pathogens to implement biological control of pest populations benefits from the most complete understanding possible of how resistance arises and propagates through natural populations. These are inherently evolutionary biological questions. The evolutionary dynamics of insect–pathogen interactions also has clinical importance in so far as insects can serve as model hosts for humans. Evolutionary inferences about how pathogens interact and interfere with different components of the immune system inform studies in molecular immunology. Advances in immunology, in turn, will test these predictions and identify new sets of genes and pathways in a wider range of organisms, further broadening the field of evolutionary genetics.

13.6 Acknowledgements

We thank Madeline Galac for helpful advice during the preparation of this manuscript and Jacob Crawford, Sarah Short, and Gerardo Marquez for feedback on the manuscript. We thank Tim Sackton for providing portions of the data illustrated in Figure 13.3. Work in the Lazzaro lab is funded by grants from the US National Science Foundation and National Institutes of Health.

13.7 References

Anisimova, M. and Liberles, D.A. (2007) The quest for natural selection in the age of comparative genomics. *Heredity* 99, 567–579.

Avila, A., Silverman, N., Diaz-Meco, M.T., and Moscat, J. (2002) The *Drosophila* atypical protein kinase C-ref(2)P complex constitutes a conserved module for signaling in the toll pathway. *Molecular and Cellular Biology* 22, 8787–8795.

Bangham, J., Obbard, D.J., Kim, K.W. *et al*. (2007) The age and evolution of an antiviral resistance mutation in *Drosophila melanogaster*. *Proceedings of the Royal Society of London Series B Biological Sciences* 274, 2027–2034.

Bangham, J., Kim, K.W., Webster, C.L., and Jiggins, F.M. (2008) Genetic variation affecting host-parasite interactions: different genes affect different aspects of sigma virus replication and transmission in *Drosophila melanogaster*. *Genetics* 178, 2191–2199.

Baxter, R.H., Chang, C.I., Chelliah, Y. *et al*. (2007) Structural basis for conserved complement factor-like function in the antimalarial protein TEP1. *Proceedings of the National Academies of Sciences USA* 104, 11615–11620.

Begun, D.J. and Whitley, P. (2000) Adaptive evolution of Relish, a *Drosophila* NF-kappaB/IkappaB protein. *Genetics* 154, 1231–1238.

Blandin, S. and Levashina, E.A. (2004) Thioester-containing proteins and insect immunity. *Molecular Immunology* 40, 903–908.

Blandin, S., Shiao, S.H., Moita, L.F. *et al*. (2004) Complement-like protein TEP1 is a determinant of vectorial capacity in the malaria vector *Anopheles gambiae*. *Cell* 116, 661–670.

Boots, M. and Begon, M. (1993) Trade-offs with resistance to a granulovirus in the Indan meal moth, examined by a laboratory evolution experiment. *Functional Ecology* 7, 528–534.

Broekaert, W.F., Terras, F.R., Cammue, B.P., and Osborn, R.W. (1995) Plant defensins: novel antimicrobial

peptides as components of the host defense system. *Plant Physiology* 108, 1353–1358.

Bulmer, M.S. and Crozier, R.H. (2004) Duplication and diversifying selection among termite antifungal peptides. *Molecular Biology and Evolution* 21, 2256–2264.

Bulmer, M.S. and Crozier, R.H. (2006) Variation in positive selection in termite GNBPs and Relish. *Molecular Biology and Evolution* 23, 317–326.

Carpenter, J.A., Obbard, D.J., Maside, X., and Jiggins, F.M. (2007) The recent spread of a vertically transmitted virus through populations of *Drosophila melanogaster*. *Molecular Ecology* 16, 3947–3954.

Carré-Mlouka, A., Gaumer, S., Gay, P. *et al*. (2007) Control of sigma virus multiplication by the *ref(2)p* gene of *Drosophila melanogaster*: an *in vivo* study of the PB1 domain of Ref(2)P. *Genetics* 176, 409–419.

Christophides, G.K., Zdobnov, E., Barillas-Mury, C. *et al*. (2002) Immunity-related genes and gene families in *Anopheles gambiae*. *Science* 298, 159–165.

Clark, A.G. and Wang, L. (1997) Molecular population genetics of *Drosophila* immune system genes. *Genetics* 147, 713–724.

Collins, F.H., Sakai, R.K., Vernick, K.D. *et al*. (1986) Genetic selection of a *Plasmodium*-refractory strain of the malaria vector *Anopheles gambiae*. *Science* 234, 607–610.

Contamine, D., Petitjean, A.M., and Ashburner, M. (1989) Genetic resistance to viral infection: the molecular cloning of a *Drosophila* gene that restricts infection by the rhabdovirus sigma. *Genetics* 123, 525–533.

Cotter, S.C., Kruuk, L.E., and Wilson, K. (2004) Costs of resistance: genetic correlations and potential trade-offs in an insect immune system. *Journal of Evolutionary Biology* 17, 421–429.

Dassanayake, R.S., Silva Gunawaradene, Y.I., and Tobe, S.S. (2007) Evolutionary selective trends of insect/mosquito antimicrobial defensin peptides containing cysteine-stabilized alpha/beta motifs. *Peptides* 28, 62–75.

Date, A., Satta, Y., Takahata, N., and Chiqusa, S.I. (1998) Evolutionary history and mechanism of the *Drosophila cecropin* gene family. *Immunogenetics* 47, 417–429.

Dawkins, R. and Krebs, J.R. (1979) Arms races between and within species. *Proceedings of the Royal Society of London Series B Biological Sciences* 205, 489–511.

Decaestecker, E., Gaba, S., Raeymaekers, J.A. *et al*. (2007) Host-parasite 'Red Queen' dynamics archived in pond sediment. *Nature* 450, 870–873.

DePristo, M.A., Weinreich, D.M., and Hartl, D.L. (2005) Missense meanderings in sequence space: a biophysical view of protein evolution. *Nature Reviews Genetics* 6, 678–687.

Drosophila 12 Genomes Consortium (2007) Evolution of genes and genome on the *Drosophila* phylogeny. *Nature* 450, 203–218.

Dybdahl, M.F. and Lively, C.M. (1998) Host-parasite coevolution: evidence for rare advantage and time-lagged selection in a natural population. *Evolution* 52, 1057–1066.

Evans, J.D., Aronsteins, K., Chen, Y.P. *et al*. (2006) Immune pathways and defence mechanisms in honey bees *Apis mellifera*. *Insect Molecular Biology* 15, 645–656.

Gwynn, D.M., Callaghan, A., Gorham, J. *et al*. (2005) Resistance is costly: trade-offs between immunity, fecundity and survival in the pea aphid. *Proceedings of the Royal Society of London Series B Biological Sciences* 272(1574), 1803–1808.

Hahn, M.W., De Bie, T., Stajich, J.E. *et al*. (2005) Estimating the tempo and mode of gene family evolution from comparative genomic data. *Genome Research* 15, 1153–1160.

Hutter, S., Saminadin-Peter, S.S., Stephan, W., and Parsch, J. (2008) Gene expression variation in African and European populations of *Drosophila melanogaster*. *Genome Biology* 21, R12.

Jiggins, F.M. and Hurst, G.D. (2003) The evolution of parasite recognition genes in the innate immune system: purifying selection on *Drosophila melanogaster* peptidoglycan recognition proteins. *Journal of Molecular Evolution* 57, 598–605.

Jiggins, F.M. and Kim, K.W. (2005) The evolution of antifungal peptides in *Drosophila*. *Genetics* 171, 1847–1859.

Jiggins, F.M. and Kim, K.W. (2006) Contrasting evolutionary patterns in *Drosophila* immune receptors. *Journal of Molecular Evolution* 63, 769–780.

Jiggins, F.M. and Kim, K.W. (2007) A screen for genes evolving under positive selection in *Drosophila*. *Journal of Evolutionary Biology* 20, 965–970.

Kraaijeveld, A.R. and Godfray, H.C. (1997) Trade-off between parasitoid resistance and larval competitive ability in *Drosophila melanogaster*. *Nature* 389, 278–280.

Kraaijeveld, A.R. and Godfray, H.C. (1999) Geographic patterns in the evolution of resistance and virulence in *Drosophila* and its parasitoids. *American Naturalist* 153, S61–S74.

Kraaijeveld, A.R. and Godfray, H.C. (2008) Selection for resistance to a fungal pathogen in *Drosophila melanogaster*. *Heredity* 100, 400–406.

Kumar, S., Christophides, G.K., Cantera, R. *et al*. (2003) The role of reactive oxygen species on *Plasmodium* melanotic encapsulation in *Anopheles gambiae*. *Proceedings of the National Academies of Sciences USA* 100, 14139–14144.

Lazzaro, B.P. (2005) Elevated polymorphism and divergence in the class C scavenger receptors of

Drosophila melanogaster and *D. simulans*. *Genetics* 169, 2023–2034.

Lazzaro, B.P. (2008) Natural selection on the *Drosophila* antimicrobial immune system. *Current Opinion in Microbiology* 11, 284–289.

Lazzaro, B.P. and Clark, A.G. (2003) Molecular population genetics of inducible antibacterial peptide genes in *Drosophila melanogaster*. *Molecular Biology and Evolution* 20, 914–923.

Lazzaro, B.P. and Little, T.J. (2009) Immunity in a variable world. *Philosophical Transactions of the Royal Society of London Series B Biological Sciences* 364, 15–26.

Lazzaro, B.P., Sceurman, B.K., and Clark, A.G. (2004) Genetic basis of natural variation in *D. melanogaster* antibacterial immunity. *Science* 303, 1873–1876.

Lazzaro, B.P., Sackton, T.B., and Clark, A.G. (2006) Genetic variation in *Drosophila melanogaster* resistance to infection: a comparison across bacteria. *Genetics* 174, 1539–1554.

Lazzaro, B.P., Flores, H.A., Lorigan, J.G., and Yourth, C.P. (2008) Genotype-by-environment interactions and adaptation to local temperature affect immunity and fecundity in *Drosophila melanogaster*. *PLoS Pathogens* 14, e1000025.

Lemaitre, B. and Hoffmann, J. (2007) The host defense of *Drosophila melanogaster*. *Annual Review of Immunology* 25, 697–743.

Levashina, E.A., Moita, L.F., Blandin, S. *et al*. (2001) Conserved role of a complement-like protein in phagocytosis revealed by dsRNA knockout in cultured cells of the mosquito, *Anopheles gambiae*. *Cell* 104(5), 709–718.

Levine, M.T. and Begun, D.J. (2007) Comparative population genetics of the immunity gene, Relish: is adaptive evolution idiosyncratic? *PLoS ONE* 2, e442.

Little, T.J. and Cobbe, N. (2005) The evolution of immune-related genes from disease carrying mosquitoes: diversity in a peptidoglycan- and a thioester-recognizing protein. *Insect Molecular Biology* 14, 599–605.

Little, T.J., Colbourne, J.K., and Crease, T.J. (2004) Molecular evolution of *Daphnia* immunity genes: polymorphism in a gram-negative binding protein gene and alpha-2-macroglobulin gene. *Journal of Molecular Evolution* 59, 498–506.

Little, T.J., Watt, K., and Ebert, D. (2006) Parasite-host specificity: experimental studies on the basis of parasite adaptation. *Evolution* 60, 31–38.

Luong, L.T. and Polak, M. (2007) Environment-dependent trade-offs between ectoparasite resistance and larval competitive ability in the *Drosophila-Macrocheles* system. *Heredity* 99, 632–640.

McDonald, J.H. and Kreitman, M. (1991) Adaptive protein evolution at the *Adh* locus in *Drosophila*. *Nature* 351, 652–654.

McKean, K.A., Yourth, C.P., Lazzaro, B.P., and Clark, A.G. (2008) The evolutionary costs of immunological maintenance and deployment. *BMC Evolutionary Biology* 8, 76.

Murphy, P.M. (1991) Molecular mimicry and the generation of host defense protein diversity. *Cell* 72, 823–826.

Niaré, O., Markianos, K., Volz, J. *et al*. (2002) Genetic loci affecting resistance to human malaria parasites in West African mosquito vector population. *Science* 298, 213–216.

Nielsen, R., Bustamante, C., Clark, A.G. *et al*. (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology* 3, e170.

Obbard, D.J., Jiggins, F.M., Halligan, D.L., and Little, T.J. (2006) Natural selection drives extremely rapid evolution in antiviral RNAi genes. *Current Biology* 16, 580–585.

Obbard, D.J., Callister, D.M., Jiggins, F.M. *et al*. (2008) The evolution of *TEP1*, an exceptionally polymorphic immunity gene in *Anopheles gambiae*. *BMC Evolutionary Biology* 8, 274.

Perron, G.G., Zasloff, M., and Bell, G. (2006) Experimental evolution of resistance to an antimicrobial peptide. *Proceedings of the Royal Society of London Series B Biological Sciences* 273, 251–256.

Ramos-Onsins, S. and Aguadé, M. (1998) Molecular evolution of the *Cecropin* multigene family in *Drosophila*: functional genes vs. pseudogenes. *Genetics* 150, 157–171.

Riehle, M.M., Markianos, K., Niaré, O. *et al*. (2006) Natural malaria infection in *Anopheles gambiae* is regulated by a single genomic control region. *Science* 312, 577–575.

Roff, D.A. and Fairbairn, D.J. (2007) The evolution of trade-offs: where are we? *Journal of Evolutionary Biology* 20, 433–447.

Rolff, J., Armitage, S.O., and Coltman, D.W. (2005) Genetic constraints and sexual dimorphism in immune defense. *Evolution* 59, 1844–1850.

Sackton, T.B., Lazzaro, B.P., Schlenke, T.A. *et al*. (2007) Dynamic evolution of the innate immune system in *Drosophila*. *Nature Genetics* 39, 1461–1468.

Samakovlis, C., Kimbrell, D.A., Kylsten, P. *et al*. (1990) The immune response of *Drosophila*: pattern of cecropin expression and biological activity. *EMBO Journal* 9, 2969–2976.

Savard, J., Tautz, D., Richards, S. *et al*. (2006) Phylogenomic analysis reveals bees and wasps (Hymenoptera) at

the base of the radiation of Holometabolous insects. *Genome Research* 16, 1334–1338.

Schlenke, T.A. and Begun, D.J. (2003) Natural selection drives *Drosophila* immune system evolution. *Genetics* 164, 1471–1480.

Schmid-Hempel, P. (2008) Parasite immune evasion: a momentous molecular war. *Trends in Ecology and Evolution* 23, 318–326.

Stroschein-Stevenson, S.L., Foley, E., O'Farrell, P.H., and Johnson, A.D. (2006) Identification of *Drosophila* gene products required for phagocytosis of *Candida albicans*. *PLoS Biology* 4, e4.

Tamura, K., Subramanian, S., and Kumar, S. (2004) Temporal patterns of fruit fly (*Drosophila*) evolution revealed by mutation clocks. *Molecular Biology and Evolution* 21, 36–44.

Tennessen, J.A. (2005) Molecular evolution of animal antimicrobial peptides: widespread positive selection. *Journal of Evolutionary Biology* 18, 1387–1394.

Tinsley, M.C., Blanford, S., and Jiggins, F.M. (2006) Genetic variation in *Drosophila melanogaster* pathogen susceptibility. *Parasitology* 132, 767–773.

van Rij, R.P., Saleh, M.C., Berry, B. *et al*. (2006) The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in *Drosophila melanogaster*. *Genes and Development* 20, 2985–2995.

Wang, X.H., Aliyari, R., Li, W.X. *et al*. (2006) RNA interference directs innate immunity against viruses in adult *Drosophila*. *Science* 312, 452–454.

Waterhouse, R.M., Kriventseva, E.V., Meister, S. *et al*. (2007) Evolutionary dynamics of immune-related genes and pathways in disease-vector mosquitoes. *Science* 316, 1738–1743.

Wayne, M.L., Contamine, D., and Kreitman, M. (1996) Molecular population genetics of *ref(2)p*, a locus which confers viral resistance in *Drosophila*. *Molecular Biology and Evolution* 13, 191–199.

Wilfert, L., Gadau, J., Baer, B., and Schmid-Hempel, P. (2007a) Natural variation in the genetic architecture of a host-parasite interaction in the bumblebee *Bombus terresis*. *Molecular Ecology* 16, 1327–1339.

Wilfert, L., Gadau, J. and Schmid-Hempel, P. (2007b) The genetic architecture of immune defense and reproduction in male *Bombus terrestis* bumblebees. *Evolution* 61, 804–815.

Woolhouse, M.E., Webster, J.P., Domingo, E. *et al*. (2002) Biological and biomedical implications of the co-evolution of pathogens and their hosts. *Nature Genetics* 32, 569–577.

Yang, Z., Nielsen, R., Goldman, N., and Pedersen, A.M. (2000) Codon-substitution models for heterogeneous selection pressure at amino acid sites. *Genetics* 155, 431–449.

Yeaman, M.R. and Yount, N.Y. (2007) Unifying themes in host defense effector polypeptides. *Nature Reviews Microbiology* 5, 727–740.

Zasloff, M. (2002) Antimicrobial peptides of multicellular organisms. *Nature* 415, 389–395.

Zhong, D., Pai, A., and Yan, G. (2005) Costly resistance to parasitism: evidence from simultaneous quantitative trait loci mapping for resistance and fitness in *Tribolium castaneum*. *Genetics* 169, 2127–2135.

Zou, Z., Evans, J.D., and Lu, Z. (2007) Comparative genomic analysis of the *Tribolium* immune system. *Genome Biology* 8, R177.