Evolutionary Dynamics of Immune-Related Genes and Pathways in Disease-Vector Mosquitoes


Mosquitoes are vectors of parasitic and viral diseases of immense importance for public health. The acquisition of the genome sequence of the yellow fever and Dengue vector, *Aedes aegypti* (Aa), has enabled a comparative phylogenomic analysis of the insect immune repertoire: in Aa, the malaria vector *Anopheles gambiense* (Ag), and the fruit fly *Drosophila melanogaster* (Dm). Analysis of immune signaling pathways and response modules reveals both conservative and rapidly evolving features associated with different functional gene categories and particular aspects of immune reactions. These dynamics reflect in part continuous readjustment between accommodation and rejection of pathogens and suggest how innate immunity may have evolved.

Repeatedly during evolution, mosquitoes and other insects have adopted hematophagy to sustain abundant progeny production. In turn, blood feeding provided a new point of entry for pathogens. To counter assaults, innate immunity has evolved to recognize and respond to numerous pathogens, in a dynamic playoff where either host or pathogen may win. Although fundamental concepts mostly derive from Dm, Ag is now an important model for studies of innate immunity. A previous comparative analysis of Ag and Dm immune-related gene families (1) highlighted their diversification and pointed toward an expanded conceptual framework of insect innate immunity. The sequencing of the Aa genome (2) permitted deeper understanding of insect immune systems, as displayed by two quite different mosquito species that diverged ~150 million years ago (Ma) and Dm, which separated from them ~250 Ma. This three-way comparison is considerably more powerful than the previous Dm-Ag study, because it allows measuring true genetic distances rather than unrooted sequence similarities. Taking advantage of the added value from multiple species comparisons, we explore the evolutionary dynamics of innate immunity in insects and how they can address both common and species-specific immune challenges.

Multiple large-scale bioinformatic methods, manual curation, and phylogenetic analyses (3) identified 285 Dm, 338 Ag, and 353 Aa genes from 31 gene families and functional groups implicated in classical innate immunity or defense functions such as apoptosis and response to oxidative stress (table S1). Additional limited analysis of nine sequenced genomes from four holometabolous insect orders, spanning 350 million years of evolution, further defined conserved family features and assisted manual gene model curation by gene family experts. The detailed core analysis (Aa/Ag/Dm) is presented in the supporting online material (SOM) text and in figs. S1 to S22, and the total data set is organized into a web-accessible resource (http://cegg.unige.ch/Insecta/innunodb/), offering a comparative perspective across higher insects. All but 24 previously named Aa genes, as well as 79 previously unnamed Ag genes, were named in accordance with the nomenclature scheme devised for the Ag genome (1) with the use of additional guidelines as described in the SOM; this information will be incorporated in the forthcoming manual annotations of the VectorBase resource (www.vectorbase.org).

Our conservative bioinformatic analysis of the complete genomes identified 4951 orthologous trios (1:1:1 orthologs in the three species) and 886 mosquito-specific orthologous pairs (absent from both Dm and the honeybee, *Apis mellifera*). Combined bioinformatic analysis and manual curation of the immune repertoire identified 91 trios and 57 pairs, plus a combined total of 589 paralogous genes in the three species. Paralogs derive from family expansions and gene losses, or cases of exceptionally high sequence divergence obscuring phylogenetic relationships. Orthologs most likely serve corresponding functions in respective organisms, whereas paralogs may have acquired different functions.

By definition, orthologous trios represent a numerically conserved subset of genes. Nevertheless, a plot of Dm-Aa and Dm-Ag phylogenetic distances, measured in terms of amino acid substitutions, revealed that, on average, immunity trio orthologs are significantly more divergent (~20%) than the totality of trios in the genomes:

1. Division of Cell and Molecular Biology, Faculty of Natural Sciences, Imperial College London, London SW7 2AZ, UK.
2. Department of Structural Biology and Bioinformatics, University of Geneva Medical School, 1211 Geneva, Switzerland.
3. Department of Genetic Medicine and Development, University of Geneva Medical School, 1211 Geneva, Switzerland.
4. Department of Molecular Microbiology and Immunology, Bloomberg School of Public Health, Johns Hopkins University, Baltimore, MD 21205, USA.
5. Department of Entomology and the Institute for Integrative Genome Biology, University of California, Riverside, CA 92521, USA.
6. Department of Entomology, Iowa State University, Ames, IA 50011, USA.
7. Laboratory of Malaria and Vector Research, Twinbrook III Facility, National Institute of Allergy and Infectious Diseases (NIAID), National Institutes of Health (NIH), Bethesda, MD 20892–8132, USA.
8. CNRS Unité Propre de Recherche 9022, Avenir-Inserm, Institut de Biologie Moléculaire et Cellulaire, Strasbourg, France.
9. Department of Animal Health and Biomedical Sciences, University of Wisconsin-Madison, Madison, WI 53706, USA.
10. Department of Entomology and Plant Pathology, Oklahoma State University, Stillwater, OK 74078, USA.
11. Department of Biochemistry, Kansas State University, Manhattan, KS 66506, USA.
13. Russell Labs, Department of Entomology, University of Wisconsin-Madison, Madison, WI 53706, USA.
14. Yale University School of Medicine, Epidemiology, and Public Health, New Haven, CT 06520, USA.
15. Fujian Center for Prevention and Control of Occupational Disease and Chemical Poisoning, Fujian, China.
16. Institute of Plant Physiology and Ecology, Shanghai, China.
17. Department of Biological Sciences, Center for Global Health and Infectious Diseases, University of Notre Dame, Notre Dame, IN 46556, USA.
18. Swiss Institute of Bioinformatics, 1211 Geneva, Switzerland.

*These authors contributed equally to this work.*
The immune recognition phase seems to begin with molecular recognition of microbial patterns, producing immune signals. Some signal transducers are highlighted. Red and blue lines indicate distance means for immunity (red dots) and all trios (blue dots), respectively. (Fig. 2 (4)). The repertoire of putative immune-related gene families. The numbers of 1:1:1 orthologous trios (red), mosquito-specific 1:1 orthologs (orange), and species-specific genes (light brown) are summed to give the total number of genes identified in Dm (first bar), Ag (second bar), and Aa (third bar) for each gene (sub)family. Families are arranged from left to right, according to the decreasing proportion of 1:1:1 orthologous trios within the family. Family acronyms that are not defined in the text include: CASPs, caspases; CATs, catalases; FREP, fibrinogen-related proteins; GALEs, galectins; MLs, MD2-like receptors.
GNBPs, distinct from the two A-type orthologous pairs that resemble fruit fly GNBPs.

Recent studies in Ag identified two types of putative malaria parasite recognition receptors belonging to distinct structural classes: thioester-containing proteins (TEPs) and leucine-rich repeat (LRR) proteins. Members of each class have been associated with the killing and disposal of parasites by lysis or melanization. The TEP family is related to the vertebrate complement factors C3/C4/C5 and pan- protease inhibitors a2-macroglobulins. Ag TEP1 binds to the surface of Plasmodium berghei and mediates parasite killing (9); it also binds to bacteria and promotes phagocytosis (10, 11). TEPs exhibit only one orthologous trio and otherwise form two groups: one with both Dm and mosquito TEPs and another with only mosquito species-specific clades (the latter group includes Ag TEP1) (Fig. 3). The second class of putative receptors include LRR immune gene 1, the pioneer P. berghei LRR antagonist (12); others of similar function are Anopheles Plasmodium-responsive LRR 1 and LRR domain 7, which have been additionally implicated in resistance to P. falciparum, the human malaria parasite (13, 14). Like TEP1, none of the three has identifiable orthologs in Ao or Dm.

Immune modulation is an important process that regulates both the immediate aftermath of recognition and subsequent effector functions and evolves in a “mix and match” mode. Examples are modulation of Toll pathway activation and the melanization reaction, respectively. In both contexts, modulation uses a vast reservoir of serine proteases and their inhibitors [serpins or serine protease inhibitors (SRPNs)] or other regulators, from which particular components are picked to constitute species-specific regulatory modules.

---

**Fig. 2.** Evolution of immune signaling phases in insects. (A) Genes and gene families implicated in two immune signaling pathways, Toll and Imd (green and purple, respectively). The well-recognized phases of signaling, from recognition to effector production, are outlined. Genes known to be part of these pathways in Dm are indicated in blue, with their closest phylogenetic relatives in Ag in red and Aa in yellow (based on the analysis presented in the SOM). Single-copy orthologs (1:1:1) in all three genomes are indicated with solid circles at the branching node and mosquito 1:1 orthologs are indicated with open circles, respectively. Ag genes affecting survival of the malaria parasite P. berghei are marked with stars, and mosquito genes transcriptionally regulated by NF-κB–like mosquito REL factors are marked with diamonds; Aa CECA and Aa DEFA effectors are controlled by both REL1A and REL2 (33, 39); similarly, Ag REL2 controls expression of immune effectors, including CECA1 and GAM (40). Dm LYSs show little response to bacterial infection, but several are up-regulated after infection by microsporidia (41). The mosquito Ag LYSC1/2 and Aa LYSC11 (LysA) genes are up-regulated after bacterial challenge (42, 43), and Ag LYSC2 is controlled by REL1. We constructed radial trees using similarity distances of the conserved sequence cores computed by maximum likelihood. Branch-length scaling is preserved within, but not between, trees. (B) Gene families implicated in the three major immune phases (recognition, signal transduction, and effector production) are clearly different in relative sequence divergence (left panel; sum of branch lengths divided by number of members). Quantitative analysis of evolutionary divergence modes in all six phases defined in (A) is based on gene numbers: trios, mosquito pairs, and genes found in only one species (right panel). All signal transduction genes form trios but are maximally divergent in sequence. In contrast, effector families diversify not by sequence divergence but by gene duplication and creation of new families (e.g., Gambicin in mosquitoes and Diptericin, Drosocin, and others in Dm). This mode results in numerous species-specific effectors but very few trios, contrasting with the pattern seen in signal transduction. The species-specific modulators are selected separately in each species, from very large, divergent families such as SRPNs and CLIPs. Although the Toll and SPZ families are rich in trios, the mosquito genes most closely related to the Dm Toll-1/Spz interaction module are largely species-specific. Finally, the recognition phase shows an intermediate level of diversification, with species-specific genes approximately equal in number to the gene sum of trios and mosquito pairs; in this case, diversification arises by duplication of both genes and domains within genes [see (A)].
Successful triggering of the *Dm* Toll pathway after fungal and Gram-positive recognition engages a dedicated proteolytic activation cascade of serine proteases and SRPNs, of which several have been identified recently (15). None of these proteins exhibit mosquito orthologs, and only Spirit and Grass have recognizable paralogs (Fig. 2). The cascade culminates in cleavage of Spaetzle by the Spaetzle proteolytic enzyme (SPE), releasing a cytokine that binds to Toll. Mosquitoes have several genes encoding Spaetzle-like proteins (SPZs), but their SPE has not been recognized. Suggestively, the short and very specific SPE cleavage site (16) recurs in *Ag* CLIP-domain serine protease B5 (*Ag* CLIPB5) and *Aa* CLIPB38, which are otherwise phylogenetically unrelated.

Similarly, activation of prophenoloxidases (PPOs) to phenoloxidases (POs), the executors of melanization, is induced by a protease cascade (mostly CLIPBs). The cascade is positively and negatively regulated by a network of inactive protease homologs (CLIPAs), CTLs, and SRPNs (Fig. 3). This melanization module is tightly controlled, because it generates toxic byproducts including reactive oxygen species. Reverse genetic analyses have identified a large set of Ag regulators for melanization of *P. berghei* (17–19) or Sephadex beads (20, 21): one SRPN, two CTLs, eight CLIPBs, and three CLIPAs (Fig. 3). Notably, all are members of mosquito-specific expansions, none has a definitive 1:1:1 ortholog, and only SRPN2 has a clear *Aa* ortholog. The reservoir of *Aa* proteases shows an underrepresentation of CLIPAs and massive expansions of CLIPBs as compared with both *Ag* and *Dm*. Finally, the melanization module may encompass additional regulators, because the genetic background determines which components are important in specific *Ag* strains (19).

The observed diversity of modulation components suggests that related but distinct regulatory modules may evolve in different species and even in subspecific taxa. Recruitment of individual members from very large multigenic families may be followed by modulatory fine-tuning through selection imposed by particular microbes. For example, several of the genes that negatively control *P. berghei* melanization in *Ag* [CTL4, CTL mannose-binding 2 (CTLM2), and SRPN2] do not affect *P. falciparum* (22, 23). Because *Ag* is a natural vector of *P. falciparum* but not of *P. berghei*, it is appealing to speculate that the sets of regulators of the melanization module evolve with and are manipulated by parasites. This modular mix and match evolution hinders detailed knowledge transfer between vector species but reinforces its importance in shaping the immune response. Future experimental studies of the melanization module in *Aa*, which can melanize bacteria and filarial worms, as well as sporozoites of the avian parasite *P. gallinaceum* (24, 25), will be fruitful in further exploring this fascinating mode of immune evolution.

Although Toll-like receptors (TLRs) are found throughout the animal kingdom, phylogenetic and functional studies have suggested that insect Tolls and mammalian TLRs evolved independently (26). Most *Dm* Tolls serve development functions, and the recruitment of the Toll (Toll-1) receptor to immune signaling has been ascribed to convergent evolution. Even within insects, our analysis detects diversity: species-specific Toll expansions and only three trios. *Dm* Toll-1 has no clear orthologs; replications have created a clade of four *Ag* and four *Aa* genes, all related to both *Dm* Toll-1 and *Dm* Toll-5 (Fig. 2). In addition to its role in antifungal and antibacterial responses, Toll-1 has been implicated in cellular antiviral responses (27). Thus, the possibility that the expanded Toll-1/ Toll-5 clade in mosquitoes is related to their interactions with viruses merits detailed functional investigation. An unexpected evolutionary pattern was also observed for Spaetzle, the cytokine partner of *Dm* Toll-1, which shows three *Aa* paralogs and no identifiable *Ag* ortholog. *Aa* SPZ1C acts together with *Aa* TOLLS to activate antifungal responses (28); however, the absence of an *Ag* Spaetzle ortholog raises questions about the evolution of this pair of molecules as an immune module, especially because the cytokine-Toll interaction is not required for mammalian TLR signaling. The only insect Tolls that cluster with TLRs are *Dm* Toll-9, *Ag* TOLL9, and *Aa* TOLL9A/9B. Because *Dm* Toll-9 is the only other Toll linked to *Drosophila* immunity (29), it is possible that this clade represents the most ancient immune-related insect Tolls. Whether these receptors can directly recognize microbial or viral immune inducers remains to be seen; it is worth noting that they are
more similar to lipid-binding TLRs rather than to nucleic acid–binding TLRs.

Signal transduction components exhibit an unexpected mode of evolution. Rather than duplicating to create novel cascades responding to distinct challenges, or picking up members of multiprotein families to promote adaptive interactions, these components show robustness, maintaining their distinctive identity and functionality in the face of sequence evolution. The cytoplasmic signal transduction of the Toll pathway includes a chain of interacting partners, almost invariably encoded by orthologous trios: myeloid differentiation factor 88 (MYD88), TUBE, PELLE, tumor necrosis factor receptor–associated factor 6 (TRAF6), and CACT (Fig. 2).

The same is true for the components of the IMD pathway: IMD, Fas-associated death domain protein (FADD), Dredd (CASPL1), IAP2, transforming growth factor–β–activated kinase (TAK1), and inhibitor of nuclear factor kβ kinase subunits γ and β (IKKγ and IKKβ). Despite persistent orthology, these components show marked divergence in sequence (Fig. 1A). A similar pattern is observed in the signal transducers Dome and Hop of the immune signaling Janus kinase–signal transducers and activators of transcription (JAK–STAT) pathway, which is activated in Dm by virus infections (30). We hypothesize that the requirement for these factors to interact productively with others in the same chain causes escalating sequence divergence: A mutation in one may enhance the acceptability of certain mutations in its interacting partner, maintaining pathway function through coherent evolution rather than stasis. Consistent with this interpretation, evidence has been reported for an association between natural sequence variation of core signaling pathway components and immune competence in Drosophila (31). Similar evolutionary patterns are detected among members of the RNA interference antiviral pathway, Dicer-2 and Ago-2 (32), which also form highly divergent trios.

Signal transduction culminates in the next phase: nuclear translocation of transcription factors. The cytoplasmic nuclear factor κB (NF-κB) transcription factors remain inactive until a processed immune signal frees them from inhibitors, permitting their entry into the nucleus and transcription of effector genes. The evolutionary pattern in this phase combines aspects observed in other phases. The NF-κBIs of the Imd pathway [Relish in Dm and Rel-like NF-κB protein 2 (REL2) in mosquitoes] form an orthologous trio that displays high sequence divergence, as in signal transducer trios (Figs. 1A and 2). A recent duplication in Aa has resulted in an orthologous quartet (Ag REL1, Dm Dorsal, Aa REL1A, and Aa REL1B). In contrast, Dif is absent from both mosquito species, although the intronless Aa REL1B gene may have originated by retrotransposition. Transgenic analysis has shown that REL1A controls Aedes antifungal responses, as does Dif in Dm (33); this represents an interesting case of functional transfer between paralogs. STAT, the transcription factor of the JAK–STAT pathway, shows high sequence divergence like REL2 and has been duplicated in Ag.

Immune effectors are required to target and neutralize the microbial source of the immune signal. We observed varied evolutionary dynamics for different categories of effectors, reflecting their modes of action. Those acting directly on microbes diversify rapidly or are species-specific, whereas effector enzymes that produce chemical cues to attack invaders remain conserved but independently expand in each species.

The production of AMPs, which act on bacterial membranes causing lysis, is a classic immune-inducible effector response (Fig. 2). Seven AMP families exist in Dm, but only three of them were detected in mosquitoes: Defensins (DEFs), cecropins (CECs), and attacins (ATIs) are highly diverse, together displaying no orthologous trio and only one confident 1:1 orthologous pair. Conversely, gambicins are only encountered in mosquitoes. The apparent paucity of mosquito AMPs in contrast to Dm may be attributable to different prevalence of bacteria in the two mosquito species, which together serve to provide a flexible system capable of adapting to new challenges. The repertoire of recognition receptors of microbial groups such as bacteria and fungi, which are encountered by all species, is achieved through expansion and fine-tuning of model genes. New functions (e.g., recognition of malaria parasites) are acquired from genes bearing powerful and ancient recognition domains such as LRRs. Protein networks modulating immune signals are assembled independently in each species, in the mix and match mode of evolution described as “bricolage” by François Jacob; they therefore coevolve with pathogens and may be subject to evasion. Pathways of signal transduction, on the other hand, remain highly conserved, and their constituent genes seem to evolve always in concert. Finally, effector mechanisms follow evolutionary patterns that depend on their mode of action; most are highly divergent or even species-specific, in contrast to the ancient, conserved oxidative defense mechanisms.

Recognition of the role of Toll in Drosophila immunity led directly to the identification of TLRs as a fundamental aspect of mammalian innate immunity. Similarly, the diverse evolutionary modes of insect immunity that we detected in the present study can guide future studies on the evolution of innate immune mechanisms in vertebrates and other animals. They can also facilitate targeted studies of immunity in the two mosquito species, which together transmit some of the most devastating infectious diseases of humankind.

References and Notes
3. Materials and methods are available as supporting material on Science Online.

www.sciencemag.org

Culling Prey Promotes Predator Recovery—Alternative States in a Whole-Lake Experiment

Lennart Persson,1* Per-Anne Amundsen,2 André M. De Roos,3 Anders Klemetsen,2 Rune Knudsen,2 Raul Primicerio2

Many top-predator fish stocks in both freshwater and marine systems have collapsed as a result of overharvesting. Consequently, some of these communities have shifted into seemingly irreversible new states. We showed, for predators feeding on prey that exhibit food-dependent growth, that overharvesting. Consequently, some of these communities have shifted into seemingly irreversible new states. We showed, for predators feeding on prey that exhibit food-dependent growth, that

In the early 1900s, the top-predator brown trout (Salmo trutta) was the only species in the low-productive Lake Takvatn, in northern Norway (14–16). Overharvesting reduced trout to low levels, and Arctic char (Salvelinus alpinus)—prey for, but also a potential competitor of, small brown trout for invertebrates (14, 16–19)—was introduced in about 1930. The char soon dominated the fish community, and by 1980 trout were almost absent (Figs. 1 and 2, A and D). To improve lake fisheries, 666,000 charr (31.3 metric tons) were removed during 1984 to 1989 (14, 15). By 1991 char density had decreased by 80%, subsequently rebounded to less than half its 1984 density, and ultimately exhibited a decelerating decrease toward a new steady state (1992 to 2006, regression $F_{1,14} = 14.7, P = 0.002$) (Fig. 1). Trout density increased from 1989 to 1992, remaining steady afterward (1992 to 2006, regression $F_{1,14} = 0, P = 0.99$). On average (1992 to 2006), the trout density was 12% of the char density, similar to the value of 15% observed in a control lake (16). Although char may compete with small trout for food (14, 17, 18), we found no evidence for a negative density–dependent effect of either total char density or density of char <150 mm in size on the body condition of 100-mm trout (regressions, $F_{1,16} = 0.10$ to 2.27, $P > 0.1$).

We generated seven testable expectations of this prey-culling management strategy with an existing resource-prey-predator, food-chain model (7, 13), in which prey exhibit food-dependent growth and reproduction and predators forage on small prey only. At low densities, invading predators cannot increase in density in the stable prey-resource state (Fig. 3). Culling prey induces oscillations in prey density and strong pulses in prey recruitment, which allow predators to increase in numbers and reach high densities. Predators can subsequently control the prey, driving the system toward an alternative, stable resource-prey-predator equilibrium (Fig. 3) characterized by (1) lower prey density (smaller-size

RAPID CHANGES OBSERVED IN MANY ECOLOGICAL SYSTEMS, SUCH AS THE COLLAPSE OF MAJOR FISH STOCKS (1–4), HAVE PROMPTED AN INCREASED INTEREST IN ALTERNATIVE STABLE STATES DURING RECENT YEARS (5, 6). THEOREETICAL STUDIES (7, 8) SUGGEST THAT SIZE-SELECTIVE PREDATION (9–11) MAY BE A MAJOR MECHANISM BEHIND SHIFTS BETWEEN ALTERNATIVE STABLE STATES IF REDUCED COMPETITION FOR RESOURCES AMONG REMAINING PREDATORS (6, 12) ACCELERATES PREY GROWTH. PREDATION ON SMALL INDIVIDUALS IN THIS CASE LEADS TO AN OVERCOMPENSATING RESPONSE BECAUSE SURVIVING PREY MATURE MORE RAPIDLY AND ACHIEVE HIGHER POPULATION PRODUCTIVE OUTPUTS. COUNTERINTUITIVELY, DENSITIES OF SMALL PREY HENCE INCREASE AND NOT DECREASE WHEN PREDATORS FORAGE ON SUCH SMALL PREY INDIVIDUALS (7, 13). THUS, SIZE-SELECTIVE PREDATORS SHAPE THE BIOTIC ENVIRONMENT TO THEIR OWN ADVANTAGE. THESE PREDATOR-PREY SYSTEMS ARE, HOWEVER, PRONE TO IRREVERSIBLE COLLAPSE OF THE PREDATOR IF OVERHARVESTED. A DROP IN PREDATOR DENSITY CAUSES PREY TO GROW AND REPRODUCE MORE SLOWLY AND CONSEQUENTLY PRODUCE LOWER ABUNDANCES OF VULNERABLE, SMALL-SIZED PREY. THIS CHANGE IN PREY-SIZE DISTRIBUTION SUBSEQUENTLY PROHIBITS RECOVERY OF THE PREDATOR, MAKING THE COLLAPSE SEEMINGLY IRREVERSIBLE (7, 8). NEXT TO THE PREDATOR-PREY STATE THE COMMUNITY TAKES UP AN ALTERNATIVE STABLE STATE WITH ONLY PREY (7). SIZE-SELECTIVE HARVESTING OF PREY MAY OFFER A ROUTE TO PREDATOR RECOVERY, BECAUSE IT SHOULD STIMULATE RATES OF PREY GROWTH AND REPRODUCTION AND THEREBY SHIFT PREY-SIZE DISTRIBUTION TOWARD SMALLER INDIVIDUALS. ONCE THE PREY-SIZE CONSTRAINT FOR RECOVERY IS LIFTED, THE RECOVERED PREDATOR POPULATION SHOULD ITSELF BE ABLE TO SUSTAIN THE SYSTEM IN THE NEW STATE.

SOM Text

SOM Text

Supporting Online Material

Table S1 and S2

References

12 January 2007; accepted 23 May 2007

10.1126/science.1139862