

# Developmental genomics of the most dangerous animal

Matthew P. Scott\*

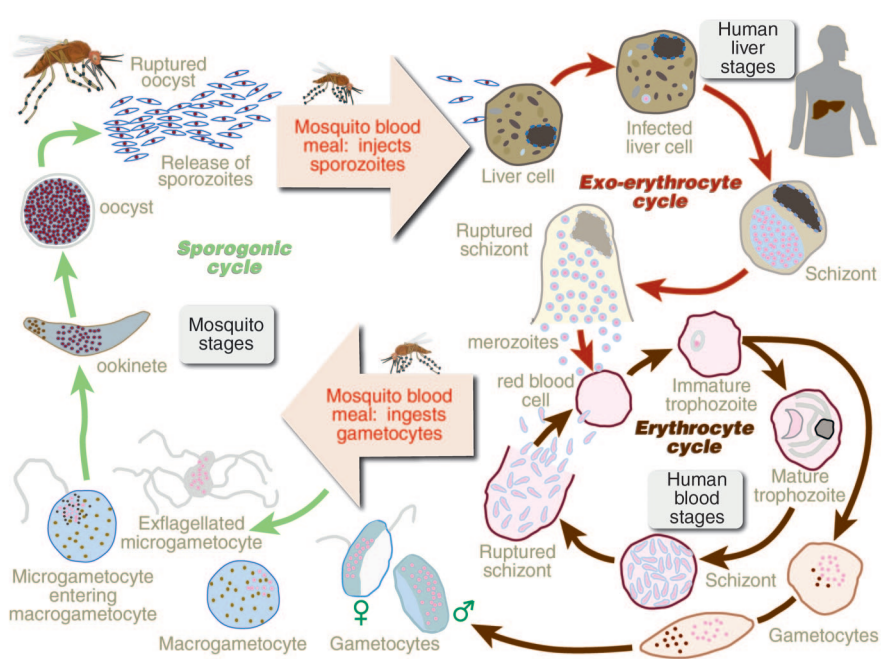
Departments of Developmental Biology, Genetics, and Bioengineering, and Howard Hughes Medical Institute, Clark Center W252, 318 Campus Drive, Stanford University School of Medicine, Stanford, CA 94305-5439

Ronald Ross penned an optimistic poem on August 21, 1897, shortly after discovering malaria parasites in mosquitoes:

This day relenting God  
Hath placed within my hand  
A wondrous thing; and God  
Be praised. At His command,  
Seeking His secret deeds  
With tears and toiling breath,  
I find thy cunning seeds,  
O million-murdering Death.  
I know this little thing  
A myriad men will save.  
O Death, where is thy sting?  
Thy victory, O Grave! (1)

The poem unfortunately overstated the prospects for a victory over malaria. One hundred ten years later, we contemplate a disaster: more than a million deaths annually from malaria, and something on the order of 400 million malaria cases each year. In Africa, more than 100 children die from malaria every hour. The discovery that mosquitoes transmit malaria (1897–1898) earned Ronald Ross the Nobel Prize in 1902, anointing this insect as far and away the most dangerous animal to humans. Mosquitoes also transmit numerous other infections, including yellow fever, dengue, encephalitis, Rift Valley fever, West Nile virus, elephantiasis, . . . a WHO's who of tragedy. The discovery of the malaria parasite, *Plasmodium*, in human blood samples (1880) earned Charles Alphonse Laveran, a French military doctor, the Nobel Prize in 1906. *Plasmodium* is a single-celled protozoan with truly astounding capabilities. Within that single cell lies the information necessary to evade not one but two advanced immune systems, and to go through several dramatic metamorphoses (Fig. 1). The latest contribution to the molecular biology of mosquito-borne diseases is a genomics view of gene activity during the life cycle of the malaria vector *Anopheles gambiae*, reported by Koutsos *et al.* in a recent issue of PNAS (2). The goal is to understand why *A. gambiae* is a malaria vector and how to stop it.

Two mosquito genomes have been completed, *A. gambiae* and *Aedes aegypti* (3, 4), together with the genome of the deadliest malaria organism, *Plasmodium falciparum* (5). More than 90% of the *Anopheles* genes have a clear relative in



**Fig. 1.** The malaria life cycle. When a mosquito carrying sporozoites (“seed animals”) obtains blood, sporozoites enter the human blood stream and travel to the liver. Infected liver cells, called schizonts, rupture and release *Plasmodium* merozoites, which infect erythrocytes and begin a process of asexual multiplication, with reinfection of more blood cells. The lysis of blood cells releases toxins and causes the main symptoms of malaria. During a period of  $\approx 8$  days, some infected blood cells produce gametocytes. These can be ingested by a mosquito to initiate the sporogonic cycle. In the mosquito’s stomach, within an hour, the two types of gametes fuse to form zygotes and develop into an ookinete (“moving egg”). The motile ookinetes penetrate the wall of the midgut and develop into oocysts. Infected midgut cells die by apoptosis and are extruded from the epithelium, but this does not defeat the parasite. Oocysts, which grow  $\approx 1,000$ -fold in volume, are chillingly effective reproductive machines. After  $\approx 12$  days of mitoses, thousands of sporozoites emerge from the oocyst. Sporozoites invade the salivary glands, allowing transmission to a human host during the next blood meal. Altogether  $\approx 2.5$  weeks elapse between ingestion of gametocytes and arrival of the sporozoites at the salivary glands. The drawings are not to scale. The figure is adapted from the Centers for Disease Control and Prevention web site, [www.cdc.gov/dpdx/html/malaria.htm](http://www.cdc.gov/dpdx/html/malaria.htm).

other species (6). How are these genes deployed? Insect transcription was first observed as chromosome “puffs,” stage-specific swellings of the chromosomes that are often indicators of transcription activity. A map of puffs induced by ecdysone, the first indication in any animal of steroid-induced gene activity, set the stage for global views of transcription later (7). A comprehensive view began with the first insect genome completed, that of *Drosophila melanogaster* (8), and continues with the mosquito genomes (3, 4) and the honey bee (9). *Drosophila* and *Anopheles* ancestors diverged  $\approx 250$  million years ago. The *Drosophila* genome allowed analyses of the flux of RNA species during developmental time (10, 11), which can now be compared with *Anopheles*.

Insect genomes, and knowledge of their activation patterns, will be useful in several ways, such as understanding evolution over the past half-billion years. Insects have evolved sophisticated chemical processing systems for constructing materials as amazing as spider silk and as pliable and impervious as cuticle, and for digesting and altering a vast range of natural substances. Genomics will reveal the enzymes that produce these materials and offer

Author contributions: M.P.S. wrote the paper.

The author declares no conflict of interest.

See companion article on page 11304 in issue 27 of volume 104.

\*E-mail: [msscott@stanford.edu](mailto:msscott@stanford.edu).

© 2007 by The National Academy of Sciences of the USA

chances to make or alter the enzymes and the materials ourselves. A third boon will be a deep understanding of sensory systems and signal processing, building upon current knowledge of vision, olfaction, and pheromones. One example is the recent identification of the *Drosophila* and *Anopheles* receptor proteins that detect CO<sub>2</sub> (12). With this information, it will be possible to look for new molecules that make it hard for mosquitoes to find us. We will also have the chance to discover the genetic basis of inherited, programmed behaviors (13), which are so dramatically evident in social insect castes, migratory insect navigation, and the deep attraction *Anopheles* mosquitoes feel for the human body.

A more complete view of mosquito molecular biology may allow the discovery or invention of selective anti-mosquito agents that do not destroy harmless and beneficial insects and may help us keep effective agents from being neutralized by genetic variation. One consequence of mosquitoes' prodigious reproduction, many generations per year in tropical climes, is that drug resistance can spread rapidly through the population. DDT (dichlorodiphenyltrichloroethane) was an effective agent that saved many lives at the cost of substantial environmental damage, but then mosquitoes resistant to DDT arose. Genomics approaches are leading us to the genes that change to confer resistance to DDT (14). The most common insecticide used to coat protective bed nets is a set of compounds called pyrethroids, similar in chemical structure to pyrethrin, which is obtained from chrysanthemum flowers. A genomics approach was used to determine that genes for two P450 enzymes are hyper-

active in *Anopheles* strains that are resistant to pyrethrins (15).

To learn how the mosquito genome is put to work, Koutsos *et al.* (2) prepared RNA from eight stages and several dissected tissues. Experimental samples were hybridized in combination with a reference standard to microarrays of  $\approx 19,680$  expressed sequence tags (ESTs). The arrays contain  $\approx 8,800$  genes, described for now as "contigs." Koutsos *et al.* found that 1,571 contigs changed RNA levels more than 2-fold during development, and they were grouped into 30 clusters of genes with similar behaviors. For example, two midgut-specific clusters emerged, an important set of genes given the events of malaria parasite maturation (Fig. 1). Previous analyses have directly examined the RNA repertoire in the midgut of infected mosquitoes (16). They found *Anopheles* and *Plasmodium* genes that are induced during ookinete differentiation and midgut invasion.

The vast amount of information about *Drosophila* development and physiology can be usefully applied to the biology of important vector species such as *Anopheles*, if it can be linked accurately. Koutsos *et al.* (2) did an extensive comparison of the gene expression patterns of *Drosophila* (10, 11) with those of *A. gambiae*. There was a gratifying level of agreement, with  $>1,000$  orthologous gene pairs behaving similarly. The greatest similarity was seen in early development programs, with increasing discrepancies in later stages that presumably reflect the different larval and adult lifestyles of mosquito and fruit fly. Despite this trend, adult males of the two species have similar gene expression programs. Interestingly, the global similarity between gene expression patterns seems to be partially independent of the similarity

of coding sequences, so they may evolve separately or at least are capable of doing so.

Because some species and variants of mosquitoes are resistant to *Plasmodium* infection, studies of immunity have focused on how the mosquito's immune system can affect the probability of parasite reproduction. Even certain *A. gambiae* variants do not transmit *Plasmodium*. Insects depend on the innate immunity system, and multiple lines of defense are used to resist *Plasmodium* (17). Major loci controlling susceptibility to *P. falciparum* are clustered at a single chromosome site (18). Koutsos *et al.* (2) found that in general, the immune genes are in the "developmentally increasing" clusters, the genes more highly expressed in pupae and adults. Perhaps during metamorphosis there is a greater chance for pathogens introduced during larval periods to proliferate and cause disease. Substantial enrichment for immune system transcripts was observed in adult females relative to other stages, suggesting that blood meals raise the danger of infection (2, 19). Males sip only nectar.

The potential for using genetic alteration of *Anopheles* to make them more resistant to *Plasmodium* infection has already been demonstrated with a transgenic mosquito that expresses a peptide that blocks the parasites from binding midgut cells (20). There are also successes in employing RNA interference strategies to alter mosquito gene expression and reduce infection by parasites (21). Even modest improvements in prevention or treatment of malaria could ease suffering for tens of millions of people. Koutsos *et al.* (2), in classifying genes by their times and places of action, have taken another step toward this goal.

- Baton LA, Ranford-Cartwright LC (2005) *Trends Parasitol* 21:573–580.
- Koutsos AC, Blass C, Meister S, Schmidt S, MacCallum RM, Soares MB, Collins FH, Benes V, Zdobnov E, Kafatos FC, Christophides GK (2007) *Proc Natl Acad Sci USA* 104:11304–11309.
- Holt RA, Subramanian GM, Halpern A, Sutton GG, Charlab R, Nusskern DR, Wincker P, Clark AG, Ribeiro JM, Wides R, *et al.* (2002) *Science* 298:129–149.
- Nene V, Wortman JR, Lawson D, Haas B, Kodira C, Tu ZJ, Loftus B, Xi Z, Megy K, Grabherr M, *et al.* (May 17, 2007) *Science*, 10.1126/science.1138878.
- Gardner MJ, Hall N, Fung E, White O, Berriman M, Hyman RW, Carlton JM, Pain A, Nelson KE, Bowman S, *et al.* (2002) *Nature* 419:498–511.
- Zdobnov EM, von Mering C, Letunic I, Torrents D, Suyama M, Copley RR, Christophides GK, Thomasova D, Holt RA, Subramanian GM, *et al.* (2002) *Science* 298:149–159.
- Ashburner M, Chihara C, Meltzer P, Richards G (1974) *Cold Spring Harb Symp Quant Biol* 38:655–662.
- Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, Scherer SE, Li PW, Hoskins RA, Galle RF, *et al.* (2000) *Science* 287:2185–2195.
- Honeybee Genome Sequencing Consortium (2006) *Nature* 443:931–949.
- Arbeitman MN, Furlong EE, Imam F, Johnson E, Null BH, Baker BS, Krasnow MA, Scott MP, Davis RW, White KP (2002) *Science* 297:2270–2275.
- Stol V, Gauhar Z, Mason C, Halasz G, van Batenburg MF, Rifkin SA, Hua S, Herreman T, Tongprasit W, Barbano PE, *et al.* (2004) *Science* 306:655–660.
- Jones WD, Cayirlioglu P, Kadow IG, Vossall LB (2007) *Nature* 445:86–90.
- Sinha S, Ling X, Whitfield CW, Zhai C, Robinson GE (2006) *Proc Natl Acad Sci USA* 103:16352–16357.
- David JP, Strode C, Vontas J, Nikou D, Vaughan A, Pignatelli PM, Louis C, Hemingway J, Ranson H (2005) *Proc Natl Acad Sci USA* 102:4080–4084.
- Muller P, Donnelly MJ, Ranson H (2007) *BMC Genomics* 8:36.
- Abraham EG, Islam S, Srinivasan P, Ghosh AK, Valenzuela JG, Ribeiro JM, Kafatos FC, Dimopoulos G, Jacobs-Lorena M (2004) *J Biol Chem* 279:5573–5580.
- Christophides GK, Vlachou D, Kafatos FC (2004) *Immunol Rev* 198:127–148.
- Riehle MM, Markianos K, Niare O, Xu J, Li J, Toure AM, Podiougou B, Oduol F, Diawara S, Diallo M, *et al.* (2006) *Science* 312:577–579.
- Warr E, Aguilar R, Dong Y, Mahairaki V, Dimopoulos G (2007) *BMC Genomics* 8:37.
- Ito J, Ghosh A, Moreira LA, Wimmer EA, Jacobs-Lorena M (2002) *Nature* 417:452–455.
- Brown AE, Catteruccia F (2006) *BioTechniques Suppl*:38–44.