that are the targets of several existing drugs[9].

The complex life cycle of *P. falciparum* means that the parasite has had to adapt to several different environments. So it is also intriguing that, compared with the genome of the free-living budding yeast, the parasite genome[1] encodes a limited number of predicted transporter proteins for the active uptake of nutrients from the environment. In fact, entire classes of transporters seem to be missing. It may be that several genes in this class have been overlooked because they are made up of many small coding regions, which can be missed by gene-prediction algorithms. But, taken at face value, this surprising finding implies that adequate amounts of nutrients recognized by the transporters must be present at all stages of the parasite life cycle, so that there is no selective advantage in having many transporters with differing substrate specificities. Alternatively, the parasite may use previously identified pores or channels to acquire nutrients[10,11].

Regulating protein levels

During its life cycle, *P. falciparum* undergoes several developmental changes. One of the most dramatic is sexual differentiation and the formation of gametes, male and female reproductive cells. The proteomics studies[3,7] of these stages have coincidentally shed light on a fundamental question: how does the parasite regulate the levels of its proteins? The genome[1] encodes relatively few predicted proteins that control the transcription of genes into messenger RNAs (the first step in making a protein). Moreover, there seem to be few transcriptional regulatory elements in the genome — or at least, there are few elements that are known from other organisms. Yet the proteomics analyses and previous studies show that protein abundance is tightly regulated.

The proteomics studies also show that proteins involved in processing mRNAs and in protein synthesis (translation) are expressed at higher levels in gametocytes, particularly female gametocytes, than in other stages. Interestingly, proteins that are present in early zygotes — which are produced from gametocytes — seem to be absent in gametocytes, although the mRNAs encoding these proteins are abundantly present. All of this is consistent with the proposal[12] that the regulation of protein levels is controlled through mRNA processing and translation, rather than by gene transcription. Perhaps this is a general feature of the parasite — another potential drug target.

In addition, one of the proteomics studies[3] reveals groups of genes whose regulation appears to be coordinated. Some simultaneously expressed genes are clustered in the genome; comparison of these genes and their flanking sequences may provide further insight into how they are regulated.

Immune evasion

Arguably the most striking features of the *P. falciparum* genome are the regions near the ends of each chromosome[1]. This is where families of genes that encode surface proteins, such as the *var* genes, are found. These proteins, or antigens, can sometimes be recognized by and thus stimulate the human immune system. But they have a great capacity for change, which occurs partly through the exchange of material between chromosome ends. As the genome sequence shows, the very ends of the chromosomes — the telomeres — have a complex arrangement of sequences that may facilitate such exchange (as described in ref. 13) and thereby lead to immune evasion.

The general structure of the chromosome ends is similar to that in the rodent parasite *P. yoelii yoelii*[2]. But, surprisingly, the genes that encode the variant surface antigens in *P. falciparum* are not found in *P. yoelii yoelii*, which has a different family of variant genes, originally described in a less virulent human parasite, *P. vivax*[14]. This is interesting, because it suggests that *P. yoelii yoelii*, which is often used as a model of *P. falciparum*, is in some respects more similar to *P. vivax*. It is tempting to speculate that, despite their dissimilar sequences, the genes at the ends of the *P. falciparum* and *P. yoelii yoelii* chromosomes have similar functions. But that remains to be seen.

Finally, research on the *P. falciparum var* genes has focused on their role in enabling infected red blood cells to stick to small blood vessels in the brain. This feature is associated with the fatal form of the disease, cerebral malaria. So it is interesting that one of the proteomics analyses[3] reveals that the peptides derived from many of the *var* genes occur in sporozoites, which are produced in mosquitoes and invade the human liver during the initial infection. These results point to possible alternative functions for *var* gene products.

The complete picture

One of the most exciting aspects of this huge undertaking is that it can be related to other work. We now have the genome of the mosquito *A. gambiae*[15], together with draft sequences of the human genome[16,17], and so can get a better handle on the interactions among three species that have long been evolving together. It is well known that certain variations in human genes are associated with a reduced susceptibility to malaria, and analysis of different human populations will no doubt reveal more on this. A close look at the mosquito genome should provide similar insights. Study of the parasite genome will reveal much about how *P. falciparum* interacts with its host and carrier, and more about the genes involved in parasite recognition by the human immune system. Decoding the information in these genomes, and translating it into effective remedies, is both a challenge and an opportunity for the scientific community. ■

Dyann F. Wirth is in the Department of Immunology and Infectious Disease, Harvard School of Public Health, 665 Huntington Avenue, Boston, Massachusetts 02115-6021, USA.
e-mail: dfwirth@hsph.harvard.edu

1. Gardner, M. J. *et al. Nature* **419,** 498–511 (2002).
2. Carlton, J. M. *et al. Nature* **419,** 512–519 (2002).
3. Florens, L. *et al. Nature* **419,** 520–526 (2002).
4. Hall, N. *et al. Nature* **419,** 527–531 (2002).
5. Gardner, M. J. *et al. Nature* **419,** 531–534 (2002).
6. Hyman, R. W. *et al. Nature* **419,** 534–537 (2002).
7. Lasonder, E. *et al. Nature* **419,** 537–542 (2002).
8. Jomaa, H. *et al. Science* **285,** 1573–1576 (1999).
9. Waller, R. F. *et al. Proc. Natl Acad. Sci. USA* **95,** 12352–12357 (1998).
10. Desai, S. A., Bezrukov, S. M. & Zimmerberg, J. *Nature* **406,** 1001–1005 (2000).
11. Kirk, K. *Nature* **406,** 949–951 (2000).
12. Dechering, K. J. *et al. Mol. Cell. Biol.* **19,** 967–978 (1999).
13. Freitas-Junior, L. H. *et al. Nature* **407,** 1018–1022 (2000).
14. del Portillo, H. A. *et al. Nature* **410,** 839–842 (2001).
15. Holt, R. A. *et al. Science* **298,** 129–149 (2002).
16. International Human Genome Sequencing Consortium *Nature* **409,** 860–921 (2001).
17. Venter, J. C. *et al. Science* **291,** 1304–1351 (2001).

The mosquito genome

The post-genomic era opens

Ennio De Gregorio and Bruno Lemaitre

The mosquito *Anopheles gambiae* is the main agent in the transmission of human malaria. Its genome sequence will in time help to devise control strategies, but will be a more immediate boon for insect biologists.

The papers that appear in this issue, describing the genome of the human malaria parasite *Plasmodium falciparum*, are published simultaneously with others in *Science* tackling the genome of the mosquito *Anopheles gambiae*. The connection is obvious: the parasite requires a mosquito to complete its complex life cycle and for transmission from one host to another. These two species are respectively the major parasite causing malaria and the major vector.

Plasmodium is taken up by mosquitoes in blood meals drawn from infected humans (see the life-cycle diagram on page 495). The parasite then undergoes several developmental stages, and crosses two mosquito cell layers that enclose the insect's midgut and salivary glands. Ultimately, *Plasmodium* is passed on when the mosquito bites a new human host, about two weeks after ingesting the first infected blood meal. For more than a century, an objective of malaria control programmes

INSTITUT PASTEUR/P. BREY

E. DE GREGORIO

Figure 1 **The mosquito and the fruitfly in typical pose — *Anopheles* (top) on human skin, *Drosophila* on a banana.**

has been to block parasite transmission by mosquitoes. These approaches will clearly benefit from the improved understanding of mosquito biology and mosquito interactions with *P. falciparum* that the genome sequences will make possible.

The *A. gambiae* genome[1] was sequenced by a collaboration between Celera Genomics, the French National Sequencing Centre (Genoscope) and The Institute for Genomics Research (TIGR), in association with several university laboratories. These groups used the same 'shotgun' strategy as that applied for sequencing the human, mouse and fruitfly (*Drosophila melanogaster*) genomes. Random fragments of genomic DNA were first cloned in bacteria, and sequenced, and the overlapping clones were then assembled into contiguous sequences. Unexpectedly, the high levels of genetic variation (polymorphisms) in the reference strain of *A. gambiae* used for sequencing — the PEST strain — made the genomic assembly step difficult. The genetic variation might be explained by the fact that two distinct populations of *A. gambiae* have contributed to the PEST strain, thereby creating a mosaic genome structure. This unprecedented situation required the development of new sequence-assembly strategies, and these will be a considerable asset for future genome projects — as with mosquitoes, not all organisms are available as inbred laboratory strains.

Comparison with the fruitfly

Much of the interest in the *A. gambiae* genome will centre on comparisons with that of *D. melanogaster*, which was published two years ago[2]. These two insects belong to the same taxonomic order, the Diptera, but inhabit distinct environments and have different lifestyles (Fig. 1). *Drosophila melanogaster* feeds on decaying organic matter, such as damaged or rotting fruit, where it also completes its life cycle, whereas *A. gambiae* feeds on sugar nectar and on the blood of vertebrate hosts. Blood meals are required for female mosquitoes to produce eggs; these are laid in water, where larvae develop and hatch. Blood feeding exposes the insect to viruses and parasites — like *Plasmodium*, these other pathogens exploit *Anopheles* as a vector for transmission.

One of the main differences between the two species is that, at 278 million base pairs, the *A. gambiae* genome is much bigger than that of *D. melanogaster* (estimated to be 180 million base pairs). But this difference is not reflected in the total number of genes, which, with 13,000–14,000 genes so far identified in both insects, is surprisingly similar. It seems that, in the course of evolution, *Drosophila* has experienced a progressive reduction both in the regions between genes and in the introns, the non-protein-coding stretches of DNA within genes.

Comparison of the coding sequences reveals that the genomes of *Anopheles* and *Drosophila* are less similar than would be expected for two species that diverged 'only' 250 million years ago. Only half of the genes in the two genomes can be interpreted as orthologues — genes in different species that have common ancestry, although their functions may differ. *Anopheles* and *Drosophila* orthologues show an average of about 56% identity in DNA sequence. As Zdobnov *et al.* point out in another of the papers in *Science*[3], from the sequence standpoint, the two species differ more than do humans and pufferfish — species that diverged 450 million years ago. Some of the protein families present in both mosquito and fruitfly appear to have evolved from a common ancestral gene through independent gene-duplication in each species. The *Anopheles* genome shows several cases of such expansion which might reflect adaptation to its lifestyle. An example is the family of fibrinogen-like proteins (of which there are 58 in *Anopheles* and 13 in *Drosophila*), which in the mosquito are probably used as anticoagulant for the ingested blood meals.

Defence mechanisms

Insects have efficient immune systems for combating the various pathogens they encounter, and most of our knowledge in this area comes from genetic and molecular studies in *Drosophila*. Finding out how *Anopheles* responds to *Plasmodium* infection is essential for obtaining clues to controlling malaria. Christophides *et al.*[4] analysed the gene families in *A. gambiae* that are linked to insect immunity, and show that they diverge widely from those in *Drosophila*. Good examples are the prophenoloxidase enzymes (nine in the mosquito, three in the fruitfly); these enzymes catalyse the synthesis of melanin, which is associated with several defence reactions in insects.

The study by Christophides *et al.* suggests that *Anopheles* employs the same general defence mechanisms as *Drosophila*, and uses similar pathogen-activated signal-transduction pathways, but that it has adapted recognition and effector immune genes to different types of aggressors. The best characterized effector system in insects consists of antimicrobial peptides, which display a wide spectrum of antibiotic activities. Interestingly, out of seven families of these peptides found in *Drosophila*, only two are also evident in *Anopheles*: five, then, are specific to *Drosophila*. Conversely, at least one mosquito-specific antimicrobial peptide has already been identified and others might be discovered by functional studies in the future. The expression profiles of some *A. gambiae* immune genes also suggest that, like the fruitfly, the mosquito mounts specific immune responses adapted to different types of pathogen[4,5].

The availability of the entire DNA sequence, together with tools such as DNA microarrays and targeted gene disruption[6–8], will make *Anopheles* a powerful model system for studying insect biology. The genomic data will also help in developing strategies to combat malaria and other mosquito-borne human diseases, for example yellow fever, dengue, filariasis and encephalitis. Such strategies will include reducing the number and lifespan of infectious mosquitoes, analysing what attracts them to their human targets, and limiting the capacity of parasites to develop within the insect vector. Malaria is characterized by a highly complex set of interactions between the parasite, the vector and the host. Now that the genomes of all three players have been fully sequenced, the post-genomic era in combating this dreadful disease can really begin. ■

Ennio De Gregorio and Bruno Lemaitre are at the Centre de Génétique Moléculaire, CNRS, 91198 Gif-sur-Yvette, France.
e-mails: gregorio@cgm.cnrs-gif.fr
lemaitre@cgm.cnrs-gif.fr

1. Holt, R. A. *et al. Science* **298,** 129–149 (2002).
2. Special section on the *Drosophila* genome *Science* **287,** 2181–2224 (2000).
3. Zdobnov, A. M. *et al. Science* **298,** 149–159 (2002).
4. Christophides, G. K. *et al. Science* **298,** 159–165 (2002).
5. Dimopoulos, G. *et al. Proc. Natl Acad. Sci. USA* **99,** 8814–8819 (2002).
6. Catteruccia, F. *et al. Nature* **405,** 959 (2000).
7. Grossman, G. L. *et al. Insect Mol. Biol.* **10,** 597–604 (2001).
8. Blandin, S. *et al. EMBO Rep.* **3,** 852–856 (2002).