

Deciphering tsetse's secret partner

Brendan W. Wren

Department of Infectious & Tropical Diseases, London School of Hygiene & Tropical Medicine, London, WC1E 7HT, UK. e-mail: brendan.wren@lshtm.ac.uk

The genome sequence of the bacterial endosymbiont *Wigglesworthia glossinidia* that resides in the gut of the tsetse fly has been determined. Because the tsetse fly relies on this bacterium for fertility and nutrition, this information may be useful in reducing fly populations and halting the spread of the deadly African sleeping disease.

Bacteria are the most populous organisms on the planet, probably followed by insects. Thus it is not surprising that bacteria and arthropods are often intimately linked, depending on each other for their coexistence. One striking example of this dependence is the bloodsucking tsetse fly (Diptera: Glossinidae) that relies on its obligate endosymbiont, *Wigglesworthia glossinidia*, for fertility and nutrition. The tsetse fly transmits the deadly parasite *Trypanosoma brucei*, which causes African sleeping sickness and has a devastating effect on livestock and humans in sub-Saharan Africa (see figure). The World Health Organization estimates that of 60 million people at risk, 300,000 are infected each year. The disease is fatal if untreated and represents the leading cause of mortality in some areas. Current drugs are highly toxic and increasingly ineffective. The potential for disrupting the symbiotic relationship between the bacterium and fly and halting the spread of disease was the driving force behind the *Wigglesworthia glossinidia* genome sequencing project reported by Leyla Akman and colleagues¹ on page 402 of this issue.

that encode the whip-like flagella for motility were identified in the genome. This may be relevant to part of the transmission cycle of the bacterium to new tsetse hosts, such as movement from adult tsetse flies to larvae.



Tsetse fly. *Trypanosoma brucei* are carried in the saliva of the blood-drinking tsetse fly. The protozoa enter a human host through the wound made by the fly when it feeds. *Wigglesworthia glossinidia* is the fly's obligate endosymbiont required for fertility and nutrition.

What's been lost

Shrinking genome size is a common feature of obligate intracellular bacteria. This may arise through a reduction of the genetic repertoire previously required to survive as free-living organisms. One

example of a bacterium associated with an arthropod in which genome downsizing is apparent is the intracellular *Buchnera* bacterium, which lives symbiotically in pea aphids². The genome size of *Buchnera* is 640 kb, compared to 697 kb for *Wigglesworthia*. Although there are similarities between the two genomes, they are actually quite different. For example, *Buchnera* does not have a flagellin system or the dozen cofactor biosynthetic pathways present in *W. glossinidia*. Examples of bacteria whose genomes have been sequenced and who rely on insect carriers and cause severe disease in humans include *Rickettsia prowazekii*³ (typhus), *Borrelia burgdorferi*⁴ (Lyme disease) and *Yersinia pestis*⁵ (plague). These pathogens represent bacteria at various stages of genome decay, with *R. prowazekii* having the most severe mutational meltdown—24% of its genome is non-coding DNA or pseudogenes, compared to 4% for *Y. pestis*. *Y. pestis* has arisen from the enteric pathogen *Y. pseudotuberculosis*

in an eye-blink of evolutionary time (2,000–15,000 years) and thus is at a very early stage of genome decay. Interestingly, many of the genes no longer required by *Y. pestis* are important for enteropathogenesis of *Y. pseudotuberculosis*, evidence of remnants of a redundant enteric lifestyle⁵.

Another arthropod-associated bacterium that manipulates reproduction in millions of insects worldwide is *Wolbachia*. A serendipitous discovery from the genome project of *Onchocerca volvulus*, the filarial nematode parasite that causes river blindness, was the identification of *Wolbachia* genome sequences⁶. It appears that *Wolbachia* have evolved as essential symbionts of filarial nematodes

A free-living history

Wigglesworthia glossinidia was first classified in 1995 and bears the name of the eminent British entomologist Sir Brian Vincent Wigglesworth. The organism is a member of the enterobacteriaceae and is related to *Escherichia coli*. However, the genome size of *W. glossinidia* (697 kb) is one of the smallest sequenced to date, one-seventh the size of *E. coli* K12. *W. glossinidia* has co-evolved with its insect host over millions of years, which has allowed the bacterium to streamline its genome, eliminating genes found in its host. Unexpectedly, the *W. glossinidia* genome still contains remnants of a free-living lifestyle. Although *Wigglesworthia* were thought to be non-motile, genes

The *Wigglesworthia* genome also contains several cofactor biosynthetic pathways, including over 60 genes involved in the synthesis of vitamins, nutrients that the tsetse fly relies on for its fertility. This finding confirms previous studies showing that the fly depends on the bacteria to provide these vitamins, which are not found in its restricted diet of blood. In the absence of the bacteria (and vitamins), the tsetse fly is sterile. Removing the bacteria from tsetse flies could halt the development of offspring, thus reducing fly populations and disease transmission. One remaining question is: does the bacterium also supply essential nutrients to the parasite *T. brucei*?

in addition to arthropods. This finding may explain why some antibiotics are effective against filariasis and may allow the development of new antibiotics for the treatment of filarial diseases.

A wake-up call

Recent weeks have heralded the most significant step forward in vector-borne parasitology with the publication of the genome sequences of the malaria parasite⁷ (*Plasmodium falciparum*) and its carrier the mosquito⁸ (*Anopheles gambiae*). This is in stark

contrast to the handful of known tsetse fly genes and the apparent stalling of the *T. brucei* genome project. There is a danger that parasitologists will be seduced into researching exclusively malaria. This would be detrimental, as comparative genetic information from arthropod bloodsucking vectors and eukaryotic parasites, which have distinct life cycles, would yield significant added value. The sequencing of *W. glossinidia* marks the start of an effort to compile the genetic information of the players in the African sleeping disease cycle. The genome

sequences of the tsetse fly and *T. brucei* are now required to develop a comprehensive understanding of the *T. brucei* life cycle that will serve as the basis to design an effective intervention strategy against the deadly African sleeping disease. □

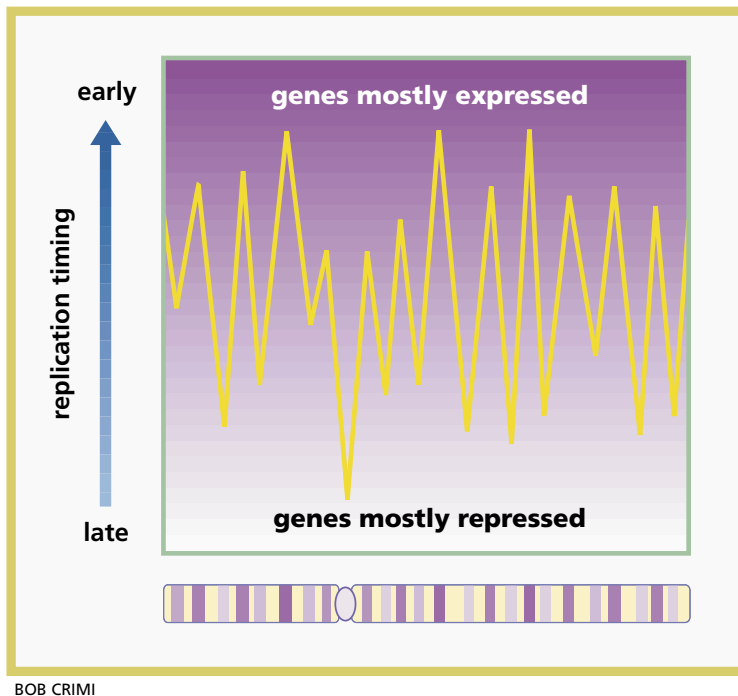
1. Akman, L. et al. *Nature Genet.* **32**, 402–407 (2002).
2. Shigenobu, S., Watanabe, H., Hattori, M., Sakaki, Y. & Ishikawa, H. *Nature* **407**, 81–86 (2000).
3. Andersson, S.G. et al. *Nature* **396**, 133–140 (1998).
4. Fraser, C.M. et al. *Nature* **390**, 580–586 (1997).
5. Parkhill, J. et al. *Nature* **413**, 523–527 (2001).
6. Unnasch, T.R. & Williams, S.A. *Int. J. Parasitol.* **30**, 543–552 (2000).
7. Gardner, M.J. et al. *Nature* **419**, 498–511 (2002).
8. Holt, R.A. et al. *Science* **298**, 129–149 (2002).

Replication timing and metazoan evolution

David M. Gilbert

Department of Biochemistry and Molecular Biology, S.U.N.Y. Upstate Medical University, Syracuse, New York 13210, USA. e-mail: gilbertd@upstate.edu

The long-presumed relationship between transcriptional activity of genes and their replication early in S-phase was challenged when a whole-genome analysis of replication timing in budding yeast found no such relationship. A new study reports the first genome-wide comparison of replication timing and gene expression in a multicellular organism, revealing a strong correlation between the two. This difference may reflect levels of nuclear organization that are important in the context of tissue-specific gene regulation.



A link between replication and expression. Replication timing analysis along the length of a chromosome reveals domains that replicate at characteristic times. Early-replicating genes have a higher probability of being expressed, whereas late-replicating genes are less likely to be expressed. The molecular basis of this probabilistic relationship and the extent to which it has a role in the developmental regulation of gene expression are important questions for future research.

For several decades, a correlation between gene activity and replication in the first half of S-phase has been appreciated, but its molecular basis remains a mystery¹. What has been desperately needed is a more complete sampling of genes to determine whether the apparent relationship between replication timing and gene expression could withstand a test of statistical significance. The advent of DNA array technology made this possible. Although the evaluation of replication timing for individual genes is a formidable task, the effort in these experiments is expended in synchronizing cells and isolating labeled replication intermediates. Once these replication intermediates are isolated, a profile of replication timing of the entire genome with a DNA array does not require much more effort than evaluating individual genes one at a time by traditional methods. The first whole-genome analysis of replication timing, performed in budding yeast², could find no correlation between gene expression and replication time, underscoring the need to perform similar analyses in a multicellular organism. On page 438 of this issue, Dirk Schübeler and colleagues³ report such an analysis, using an established *Drosophila melanogaster* cell line.