Introduction

The recent publication of the draft human genome sequence (International Human Genome Sequencing Consortium 2001; Venter *et al.* 2001) was hailed by the Presidents of The Royal Society and the National Academy of Sciences of the USA as 'one of the most significant intellectual achievements of all time'. Science at its most triumphalist. But the human genome is only the latest and largest in a rapidly expanding list of sequenced genomes. These include nearly 600 viruses, 30+ Eubacteria, 7 Archaea, *Saccharomyces cerevisiae* ('the fungus'), *Arabidopsis thaliana* ('the plant'), *Caenorhabditis elegans* ('the worm') and *Drosophila melanogaster* ('the fly'). Within a year we expect to add the first of the protozoan genomes to this list (*Plasmodium falciparum*), with other protozoans following soon after. There is also a sizeable list of protozoa for which there is substantial sequence information, although the entire genome of each (down to the last few per cent) will not be sequenced. Clearly the time is right to consider how best to make use of the wonderful new resource that genome databases offer.

This issue of *Philosophical Transactions* is a collection of papers considering different aspects of the theme 'Utilizing the genome sequence of parasitic protozoa', the subject of a meeting of the same name that took place at The Royal Society. The meeting focused on four parasites—*P. falciparum*, *Trypanosoma brucei*, *Leishmania major* and *Toxoplasma gondii*. These species cause important diseases. They now have well-developed genome databases compared with most other parasites and they also afford, to varying extents, the ancillary tools and technologies without which utilization of any genome is impossible. Genome projects in eukaryotes are developed by organism-specific networks of researchers, in collaboration with one or more genome-sequencing centres. This approach to genome project development inevitably leads to organism-specific views on genome projects, which we believe can be unhelpful. The many parallels between genome projects have impressed us. Some of these are quite obvious, such as the methods by which genome sequences are assembled and annotated, or the end objective of improving the biological understanding of disease-causing organisms to further drug and vaccine development. Other parallels seem to be less widely appreciated, such as the importance of database curation and the differences in outputs from transcriptome and proteome analyses. By illustrating some of these parallels, and showing examples of 'best practice' in different parasites, we hope that the meeting and this issue have helped in the progress of functional genomics in parasitology.

Genome projects are curious beasts in two important ways—they are, by and large, *not* hypothesis driven and they require anticompetitive behaviour amongst scientists. Most of us, implicitly or explicitly, take a Popperian view of science, which gives central place to the development of hypotheses, and we spend our scientific lives devising

tests to reject null hypotheses. To then find ourselves devoting time, and substantial resources, to sequencing projects that do not do this is strange, but it highlights the point that genome sequences are resources for the future rather than end products in the Popperian sense. 'Without ideas to give them shape and meaning, these endless successions of base sequences ...are as featureless as a plate of tofu' (Bray 2001). Most of us prefer the security of hypothesis-driven research using the genome databases as new tools. We should not forget, however, that the genome database for a particular parasite is but the central plank of a suite of resources needed to exploit fully its information content. Thus, there is still an important role for non-hypothesis-driven projects in parasitology.

Post-genomic science necessitates an interesting mix of traditional competitive work practices, where ideas are conceived and experiments conducted mainly in a single laboratory, and a continuation of the large-scale collaborative projects that require scientific consortia. This latter style of science was virtually unheard of in parasitology before genome projects. Parasite-specific collaborative networks have developed over several years to make the academic case for genome sequencing. There is still a role for such networks in post-genomic science to tackle projects that are too large for any one laboratory.

Protozoan genomes, when added to the list of sequenced organisms, will prove important in understanding the genetic basis of biological diversity in eukaryotes. The genetic distances between the five higher eukaryotes sequenced to date are very small in the grand scale of eukaryotic phylogeny and most genetic diversity amongst eukaryotes will be exemplified in genome databases compiled for protozoa. One example serves to illustrate this diversity. *Trypanosoma cruzi* and *T. brucei*, the two parasites that cause Chagas' disease in South America and sleeping sickness in Africa, are both members of the same genus and yet the genetic distance between them is approximately the same as that between a human and a starfish.

The papers in this issue can be considered in four groups. The first group comprises case studies in utilizing functional genomics in parasites (Almeida *et al.* 2002; Beverley *et al.* 2002; Carucci 2002; Sibley *et al.* 2002) using both microarray technology and proteomics. Three papers provide an illustration of the power of mining genome databases, which requires an understanding of how these databases are compiled, annotated and curated to exploit fully their potential, and of combining informatics with experimentation to speed progress (Fairlamb 2002; Roos *et al.* 2002; Stuart *et al.* 2002). Utilizing a genome database successfully requires continued development of experimental technologies to manipulate genomes using both forward and reverse genetics (Cowman *et al.* 2002; Tait *et al.* 2002; Ullu *et al.* 2002; Waters 2002). Finally, two papers indicate possible ways forward for parasite

functional genomics, in the widest sense, by illustrating what is being done with yeast (Oliver 2002) and pathogenic bacteria (Moxon *et al.* 2002), systems that are genetically more tractable than those of protozoa and with smaller, simpler genomes.

We thank the speakers for the quality of their presentations, which contributed to a very successful meeting, and the quality of the papers produced for this volume, which we believe provide excellent illustrations of the current exciting developments in parasitology.

Mike Turner¹

Jennie Blackwell²

Chris Newbold³

Keith Vickerman⁴

August 2001

¹*Division of Infection and Immunity, Joseph Black Building, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK*

²*Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Wellcome Trust/MRC Building, Addenbrooke's Hospital, Hills Road, Cambridge CB10 1SA, UK*

³*Institute of Molecular Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK*

⁴*Division of Environmental and Evolutionary Biology, Graham Kerr Building, Institute of Biomedical and Life Sciences, University of Glasgow, Glasgow G12 8QQ, UK*

REFERENCES

- Almeida, R. (and 10 others) 2002 From genomes to vaccines: *Leishmania* as a model. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Beverley, S. M. (and 14 others) 2002 Putting the *Leishmania* genome to work: functional genomics by transposon trapping and expression profiling. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Bray, D. 2001 Reasoning for results. *Nature* **412**, 863.
- Carucci, D. J. 2002 Technologies for the study of gene and protein expression in *Plasmodium*. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Cowman, A. F., Baldi, D. L., Duraisingh, M., Healer, J., Mills, K. E., O'Donnell, R. A., Thompson, J., Triglia, T., Wickham, M. E. & Crabb, B. S. 2002 Functional analysis of *Plasmodium falciparum* merozoite antigens: implications for erythrocyte invasion and vaccine development. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Fairlamb, A. H. 2002 Metabolic pathway analysis in trypanosomes and malaria parasites. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- International Human Genome Sequencing Consortium (IHGSC) 2001 Initial sequencing and analysis of the human genome. *Nature* **409**, 860–821.
- Moxon, E. R., Hood, D. W., Saunders, N. J., Schweda, E. K. H. & Richards, J. C. 2002 Functional genomics of pathogenic bacteria. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Oliver, S. G. 2002 Functional genomics: lessons from yeast. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Roos, D. S., Crawford, M. J., Donald, R. G. K., Fraunholz, M., Harb, O. S., He, C. Y., Kissinger, J. C., Shaw, M. K. & Striepen, B. 2002 Mining the *Plasmodium* genome database to define organellar function: what does the apicoplast do? *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Sibley, L. D., Mordue, D. G., Su, C., Robben, P. M. & Howe, D. K. 2002 Genetic approaches to studying virulence and pathogenesis in *Toxoplasma gondii*. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Stuart, K., Panigrahi, A. K., Schnauffer, A., Drozd, M., Clayton, C. & Salavati, R. 2002 Composition of the editing complex of *Trypanosoma brucei*. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Tait, A., Masiga, D., Ouma, J., MacLeod, A., Sasse, J., Melville, S., Lindegard, G., McIntosh, A. & Turner, M. 2002 Genetic analysis of phenotype in *Trypanosoma brucei*: a classical approach to potentially complex traits. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Ullu, E., Djikeng, A., Shi, H. & Tschudi, C. 2002 RNA interference: advances and questions. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).
- Venter, J. C. (and 273 others) 2001 The sequence of the human genome. *Science* **291**, 1304–1351.
- Waters, A. P. 2002 Orthology between the genomes of *Plasmodium falciparum* and rodent malaria parasites: possible practical applications. *Phil. Trans. R. Soc. Lond. B* **357** (this issue).