Genomes and Genome Projects of Protozoan Parasites

Klaus Ersfeld

School of Biological Sciences, 2.205 Stopford Building, University of Manchester, Oxford Road, Manchester M13 9PT, UK.

Abstract

Protozoan parasites are causing some of the most devastating diseases world-wide. It has now been recognised that a major effort is needed to be able to control or eliminate these diseases. Genome projects for the most important protozoan parasites have been initiated in the hope that the read-out of these projects will help to understand the biology of the parasites and identify new targets for urgently needed drugs. Here, I will review the current status of protozoan parasite genome projects, present findings obtained as a result of the availability of genomic data and discuss the potential impact of genome information on disease control.

Introduction

The resurgence of infectious diseases world-wide has been a major impetus in increasing research activities. The World Health Organisation (http://www.who.int/tdr/) has identified ten major, yet neglected, infectious diseases (African trypanosomiasis, Chagas disease, dengue fever, lymphatic filariasis, leishmaniasis, leprosy, malaria, onchocerciasis, schistosomiasis, and tuberculosis) that are the focus of intense efforts to control or even eradicate the organisms that cause these diseases. Four diseases of this list are caused by protozoan parasites (African trypanosomiasis, Chagas disease, leishmaniasis, and malaria), causing more than 1.3 million deaths annually, possibly as many as 4 million. It is anticipated that the genome information of these parasites will be a major contributing factor in the translation of basic research into applications pertinent to disease control (Hoffman, 2000). In contrast to the human genome project, data of parasite genomes are immediately released into the public domain. It is hoped that the free access and dissemination of the data will de-monopolise research and will lead to a strong involvement of research communities in developing countries where these diseases are most prevalent (Varmus, 2002). Initiatives involving the genome research centres world-wide have been founded around 1996/97 and as a first major landmark in their combined efforts the complete genome sequence of the most important of these parasites, *Plasmodium falciparum*, has recently been published (Gardner *et al.*, 2002a). In this review I illustrate the scientific framework of genome projects using three representative parasites (African trypanosomes, *Leishmania* and *Plasmodium*). Several

other projects (Table 1) are at various stages of progress, but the reasoning and motivations for genomic research is virtually identical between all of them (Tarleton and Kissinger, 2001).

The *Plasmodium falciparum* genome

More than 1 billion people are estimated to carry malaria-causing parasites at any one time. The annual mortality rate is between 0.5-3 million people. A massive increase in population, a deterioration in public health services and infrastructures and the problems associated with now widespread drug resistance has led to a re-emergence of malaria as one of the most serious diseases world-wide (http://www.who.int/tdr/diseases/malaria/default.htm). After a period of optimism, mainly due to the relatively successful reduction of the malaria-transmitting mosquitoes with the pesticide DDT, more people die know of the disease than 40 years ago (Guerin *et al.*, 2002a; Miller and Greenwood, 2002). An important factor contributing to the failure to reduce the burden of this disease is the ever decreasing number of drugs that are still effective in treatment. Chloroquine, once the most important anti-malaria drug, is almost useless due to the spread of resistance parasite populations in virtually all endemic areas. Likewise, resistances are now observed for every anti-malaria drug with the exception of artemisinin derivatives (Wellems and Plowe, 2001; Wellems, 2002). Projects to develop vaccines, a strategy probably more appropriate in areas were continuous drug supplies are difficult to maintain, have so-far failed (Richie and Saul, 2002). Therefore, after a period of four decades of relative stagnation in malaria research, it has been recognised that a massive effort at all levels is required to control this disease. Inspired by the successfully initiated large-scale genome projects (human and yeast) and the rapid completion of the first genomes of pathogenic bacteria (*Haemophilus influenzae* and *Mycoplasma genitalium* in 1995) the *Plasmodium falciparum* genome project consortium was founded in 1996 with the objective to sequence the entire genome of *Plasmodium falciparum*, the most dangerous of the four malaria-causing *Plasmodium* species (Dame *et al.*, 1996; Hoffman *et al.*, 1997). The major motivation and justification for this project was the prospect that the complete genome information would be the basis for a catalogue of new drug and vaccine candidates.

The genome of *P. falciparum* is approximately 23 x 10^6 bp in size, comparable to the size of other protozoan parasites (see Table 1). It consists of 14 chromosomes, ranging from 0.6 - 3.5 x 10^6 bp (Gardner, 2001). Chromosomes, with the exception of 6,7 and 8, can be resolved by pulsed-field gel electrophoresis and therefore, given also the technical and computational limitations at the beginning of the project, a whole chromosome shotgun sequencing strategy was used to sequence the genome. A major technical challenge was the extremely high A/T-content of the genome. On average it is 80.6% but can

*For correspondence. Email k.ersfeld@man.ac.uk.

Table 1. Current genome projects of protozoan parasites.

Parasite	Disease	Genome size (Mb)	No. of genes	Website
Apicomplexa				
Plasmodium falciparum	Malignant tertian malaria	23	5268	http://www.sanger.ac.uk/Projects/P_falciparum/who&what.shtml
Plasmodium vivax	Benign tertian malaria	35-40		http://www.tigr.org/tdb/e2k1/pva1/intro.shtml
Plasmodium yoelii	Rodent malaria	23	5878	http://www.tigr.org/tdb/e2k1/pya1/
Plasmodium chabaudi	Rodent malaria			http://www.sanger.ac.uk/Projects/P_chabaudi/
Plasmodium berghei	Rodent malaria	25-27		http://www.lumc.nl/1040/research/malaria/model.html
Plasmodium knowlesi	Malaria of monkeys (Macaca spp.)			http://www.sanger.ac.uk/Projects/P_knowlesi/
Toxoplasma gondii	Aquired or congenital toxoplasmosis	80		http://www.toxodb.org/links.shtml
Theileria annulata	Theileriosis in domestic ruminants	10		http://www.sanger.ac.uk/Projects/T_annulata/
Theileria parva	East coast fever in ruminants	9		http://www.tigr.org/tdb/e2k1/tpa1/
Eimeria tenella	Coccidiosis in poultry	60		http://www.iah.bbsrc.ac.uk/eimeria/
Babesia bovis	Redwater fever in cattle	9.4		http://www.sanger.ac.uk/Projects/B_bovis/
Cryptosporidium parvum	Cryptosporidiosis	10		http://www.mrc-lmb.cam.ac.uk/happy/CRYPTO/cryptogenome.html
Kinetoplastida				
Trypanosoma brucei	Sleeping sickness	35	8000[1]	http://parsun1.path.cam.ac.uk/
Trypanosoma cruzi	Chagas disease	40		http://www.dbbm.fiocruz.br/genome/tcruzi/tcruzi.html
Leishmania major	Cutaneous leishmaniasis	34	8000[1]	http://www.genedb.org/genedb/leish/index.jsp
Others				
Giardia lamblia	Giardiasis	12		http://jbpc.mbl.edu/Giardia-HTML/index2.html
Entamoeba histolytica	Intestinal amoebiasis	20		http://www.lshtm.ac.uk/itd/units/pmbbu/enta/home.htm
Caenorhabditis elegans		100	19000	http://www.sanger.ac.uk/Projects/C_elegans/
Saccharomyces cerevisiae		12	6275	http://mips.gsf.de/proj/yeast/CYGD/db/index.html

[1] Estimated by extrapolation from the gene number on already finished chromosomes or contigs

rise to more 90% in introns and intergenic regions. In some regions, short 2-3 kb imperfect tandem repeats that are assigned as putative centromeres, the A/T content is higher than 97%. This unusual base composition makes both sequencing and assembly difficult and is one of the reasons why sequencing the whole genome has taken a comparatively long time. New software and improved sequencing tools had do be developed to overcome this major obstacle. Optical mapping techniques were used to support the assembly of sequence data into large contigs (Lai *et al.*, 1999). In 1998 and 1999, respectively, the sequences of chromosomes 3 and 2 were published and in 2002 the entire genome sequence was published (Gardner *et al.*, 1998; Bowman *et al.*, 1999; Gardner *et al.*, 2002a; Gardner *et al.*, 2002b; Hall *et al.*, 2002; Hyman *et al.*, 2002). At the same time both the genome sequence of the model rodent malaria parasite *Plasmodium yoelii yoelii* and the sequence of the genome of the insect vector of *P. falciparum*, the mosquito *Anopheles gambiae* were released (Carlton *et al.*, 2002; Holt *et al.*, 2002). The genome data for a range of different malaria parasites (see Table 1) will be useful in several ways (Carlton *et al.*, 2001; Thompson *et al.*, 2001; Waters, 2002). As mentioned above, the A/T-bias and the presence of introns makes the *in silico* identification of genes difficult. The availability of several genomes with a high degree of synteny, but differences in detail, will aid the completion of gene annotation. More than 60% of genes identified in *P. falciparum* have orthologues in *P. yoelii*. This degree of similarity is, however, surprisingly low given the fact that both parasites cause malaria in mammals. Most strikingly, *P. yoelii* and *P. falciparum* genes that are involved in antigenic variation and virulence share few common features between the sequences (Doolittle, 2002). Many of these genes are, however, located in subtelomeric locations on chromosomes, suggesting that underlying mechanisms that drive antigenic variation are shared

between different species of *Plasmodium* (Scherf *et al.*, 2001). The divergence of genes involved in antigenic variation between species makes comparative genomics/ proteomics with respect to host-parasite interactions less useful than was hoped for (Doolittle, 2002). Nevertheless, laboratory models of malaria will be of great use to functionally characterise and exploit genomic information in the future (Carlton and Carucci, 2002).

The mosquito *Anopheles gambiae* is the most important vector for transmission of malaria parasites (but by no means the only one - there are about 60 other mosquito species capable of malaria transmission) (Spielman and D'Antonio, 2001). The major interest in the *Anopheles* genome will focus on the specific features of the insect enabling co-existence with the parasite (De Gregorio and Lemaitre, 2002). Firstly, comparison with the genome of *Drosophila melongaster*, also of the order Diptera, will reveal potential differences that are the basis the different lifestyles of the two insects. An example is the identification of a 58 of fibrinogen-like molecules in *Anopheles* that probably act as anti-coagulant for an ingested blood meal. In *Drosophila* only 13 of these genes were identified (De Gregorio and Lemaitre, 2002). Secondly, using techniques such as microarray analysis, the impact on the expression of certain genes in infected and non-infected insects can be measured. Such data will be of practical consequence as one of the strategies to combat malaria is to release transgenic mosquitoes into the wild that are incapable of parasite transmission. This will require an extensive knowledge of parasite-vector interactions, but also other issues such as mating behaviour as the modified mosquitoes have to breed efficiently with wild-type insects to drive the transgene into the population. Transcriptional profiling will also help to identify genes involved in insecticide resistance - an increasing problem to malaria control (Hemingway *et al.*, 2002).

Returning to the *Plasmodium* genome - what are the

main findings? A total of 5268 genes is predicted, 60% of these genes have no assigned function (Gardner *et al.*, 2002a). The overall structure of the chromosomes confirms in much detail previous observations. The subtelomeric regions of all 14 chromosomes are enriched in genes involved in immune evasion and pathogenesis (Crabb and Cowman, 2002). The three known gene families (59 *var*, 149 *rif* and 28 *stevor* genes in the genome reference strain 3D7) are preferentially clustered in telomeric regions: a search of the *P. falciparum* genome database PlasmoDB (http://plasmodb.org) reveals that 29 *var*, 44 *rif* and 5 *stevor* genes are within 20 kB distance of telomeres (Kissinger *et al.*, 2002). The proteins encoded by *var* and *rif*-genes (PfEMP1 and rifins) are expressed on the surface of infected erythrocytes (Kyes *et al.*, 1999; Kyes *et al.*, 2001) whereas *stevor* gene products appear to localise to internal membrane structures of erythrocytes (Kaviratne *et al.*, 2002). The function of PfEMP1 proteins in sequestration of infected red blood cells is well established but the involvement of rifins and stevor proteins in antigenic variation is hypothetical. The host is able to mount an efficient immune response against PfEMP1 which the parasite undermines by transcriptional switching between different *var* genes (Wahlgren and Bejarano, 1999; Cooke *et al.*, 2000; Kyes *et al.*, 2001). As we will see when discussing the genome data of another parasite capable of antigenic variation, *Trypanosoma brucei*, the telomeric localisation of genes involved in immune evasion is a common feature of such parasites (Barry *et al.*, 2003). In addition to the presence of these protein coding genes a complex array of repetitive sequence elements has been identified in telomeric locations. Although some of these elements were known the complexity and extent of their arrangement become only clear after assembling large telomeric and subtelomeric chromosomal regions (Gardner *et al.*, 2002a). It has been hypothesised that the presence of large blocks of near-identical interchromosomal sequences leads to the observed clustering of chromosome ends, thus facilitating ectopic recombinational activities (Freitas-Junior *et al.*, 2000; Scherf *et al.*, 2001; O'Donnell *et al.*, 2002). The high degree of conservation observed between subtelomeric regions in *P. falciparum* could be the result of frequent inter-chromosomal exchange leading to sequence homogenisation. Genes involved cytoadherence and antigenic variation, associated non-coding regions (promoters, intergenic regions and sub-telomeric repeats) occupy around 10% of the total genome of *P. falciparum*.

A surprising result of the gene analysis is the parasite's lack of some enzymes (such as the F0 *a* and *b* subunits of the ATP synthase complex) that are required for a functional mitochondrion in terms of energy (ATP) production using an electrochemical membrane gradient (Wirth, 2002). Enzymes of other metabolic pathways (TCA-cycle and oxidative phosphorylation) have been identified. Overall, the proportion of enzymes within the entire genome is much smaller than that calculated for other organisms. Whereas in *P. falciparum* ~8% of all predicted genes were assigned Enzyme Commission (EC) numbers, the proportion in *S. cerevisiae* is 17%. This could point to a reduced complexity of metabolic pathways, but could also be a consequence of the technical difficulties to identify orthologous genes *in silico*. A similar picture emerges for the analysis of transporter proteins. In comparison to other eukaryotic genomes, *P. falciparum* possesses a very limited number of membrane transporters that are required for nutrient uptake. This is particularly valid for transporters of organic nutrients and less obvious for inorganic ion transporters. A potential explanation for this observation is the lifecycle of the parasite. It is likely that a stable supply of nutrients is present both during the insect stage and the mammalian (intracellular) stage of the development. These relatively stable environments obviate the need for maximum adaptability to both qualitatively and quantitatively fluctuations in nutrients encountered by free-living organisms. It will be interesting to see whether genome statistics of other eukaryotic parasites mirror the situation in *Plasmodium*.

One of the most significant findings of the genome analysis so far are related to the function of an organelle unique to Apicomplexa, the apicoplast (Roos *et al.*, 2002). The apicoplast is a spherical structure of approximately 1 μm in diameter, surrounded by four (in *Toxoplasma*) or three (in *Plasmodium*) membranes (Köhler *et al.*, 1997; Hopkins *et al.*, 1999). This plastid harbours a 35 kb circular genome (Wilson and Williamson, 1997). The presence of more than two membranes and the phylogenetic analysis of the 35 kb DNA suggests that this organelle was acquired by secondary endosymbiosis of a red alga that contained the evolutionary remnants of a cyanobacterium . A number of drugs (fluoroquinolones, clindamycin), known to inhibit certain functions in prokaryotes, have been shown to be effective against apicomplexan parasites by interfering with apicoplast function (Fichera and Roos, 1997). This not only supports the endosymbiosis hypothesis but also shows that a functional apicoplast is essential for parasite survival. As only few apicoplast proteins are actually encoded by the 35 kb circle, most organelle-specific proteins are encoded by the nuclear genome and have to be targeted to and imported into the apicoplast (Waller *et al.*, 1998; Roy *et al.*, 1999). Targeting is achieved by the presence of a bipartite signal sequence at the N-terminus of apicoplast proteins consisting of a 15-40 amino acid signal sequence of hydrophobic nature responsible for ER translocation and a subsequent 100-150 amino acids transit peptide sequence that targets proteins into the apicoplast, presumably by interacting with an as yet unidentified import machinery. Mining the *Plasmodium* database by applying algorithms that identify typical biochemical features (mainly amino acid features that affect charge distribution) of the transit peptide sequence it was possible to identify more than 450 putative apicoplast proteins (Zuegge *et al.*, 2001; Foth *et al.*, 2003). Given a total of about 5300 genes this represents more than 8% of the total genome. Using these data it was possible to piece together a number of metabolic pathways that are either partially or entirely localised inside the apicoplast. Some of these pathways are not present in the mammalian host and enzymes involved in such specific functions are potential drug targets. One enzyme (DOXP reductoisomerase) involved in the 1-deoxy-D-xylulose 5 phosphate (DOXP) pathway of isoprenoid synthesis is efficiently inhibited by fosmidomycin, a substance originally

identified to inhibit the same enzyme in bacteria and plants (Lell *et al.*, 2003). By inhibiting this essential and parasite-specific metabolic pathway parasitemia can be cleared in a mouse malaria model (*P. vinckei*), in red blood cell cultures of *P. falciparum* and, more recently, in initial clinical trials on malaria-infected humans (Jomaa *et al.*, 1999; Missinou *et al.*, 2002). Due to the high rate of recrudescence observed after a short treatment regime with fosmidomycin, combination therapies with clindamycin (see above) are now explored to increase the clinical suitability. Likewise, other components of the DOXP-pathway have been identified and are now evaluated as anti-malaria drug targets (Altincicek *et al.*, 2002; Kollas *et al.*, 2002). In the context of identifying new drug targets it is remarkable that the biocide triclosane, previously thought to be a fairly unspecific bactericide, has now been found to specifically inhibit the enzyme enoylreductase of the type II fatty acid synthesis pathway in bacteria (Beeson *et al.*, 2001; Kapoor *et al.*, 2001; Suguna *et al.*, 2001). Surprisingly, this pathway is also present in Apicomplexa and localised in the apicoplast (Waller *et al.*, 2003). The currently widespread use of triclosane in many consumer products will most likely lead to the emergence of resistance against this drug and render it useless both as a bactericide and a potential anti-malaria drug. A ban of the indiscriminate use of triclosane is therefore urgently needed (Schweizer, 2001; Tan *et al.*, 2002).

Although some of the metabolic pathways are not found in the host, they are not specific to *P. falciparum* but are present in other Apicomplexa as well. This offers the possibility that model systems can be used to investigate the molecular components and characteristics of such pathways in much more detail than it is possible in *P. falciparum* (Mota *et al.*, 2001; Carlton and Carucci, 2002; Waters, 2002). For example, the related parasite *Toxoplasma gondii* is a much more tractable organism where the whole gamut of molecular techniques (e.g. mutagenesis, gene replacement, ectopic expression) can be employed with ease, something currently still difficult to do in *P. falciparum*. This demonstrates one the advantages of conducting parallel genome projects on related organisms, such as *Toxoplasma gondii*, *Theileria ssp*. or *Eimeria tenella*. The latter two parasites are also of importance as they cause diseases in domestic animals (Table 1).

Analysing the *P. falciparum* proteome

Both the development of drugs and vaccines against the parasite requires knowledge about the potential targets and the expression profile of targets. Most parasites undergo complex lifecycles involving developmental changes in the host and vector. On a molecular level these developmental cycles are mirrored by differential expression patterns of a number of genes that encode proteins relevant to a particular stage of the life cycle. Whereas the genome data set gives a blueprint of the entire gene complement only a detailed transcriptional or, preferably, proteomic analysis will provide information about the temporal and developmental implementation of this genetic information (Carucci *et al.*, 2002). Therefore, concomitant with the

genome project, a proteome analysis of different stages of the life cycle of *P. falciparum* was undertaken (Florens *et al.*, 2002; Lasonder *et al.*, 2002). Protein lysates prepared from various life cycle stages were analysed by tryptic peptide fingerprinting using HPLC and tandem mass spectrometry. Three main general findings emerged from this analysis. First, the number of differentially expressed proteins was surprisingly large. Almost half of the detected proteins of the sporozoite stage (the mosquito stage that is infective) were specific to that stage. Even different stages within the mammalian host (trophozoites, merozoites, gametocytes (Bannister *et al.*, 2000; Smith *et al.*, 2002)) had between 20 and 30% unique proteins. Only 6% of proteins detected in this study were shared between all four stages (Florens *et al.*, 2002). Also, in sexual stages of the parasite (gametocytes and gametes) 575 out of a total of 1289 proteins were not found in asexual stages (Lasonder *et al.*, 2002). It remains to be seen whether these findings reflect in all cases true differences in expression patterns between stages or can partially be explained by technical issues. Between 75% (Lasonder *et al.*, 2002) and 54% (Florens *et al.*, 2002) of all predicted gene products were not detected by the two published studies. The reasons can be of technical (solubility, sample preparation etc.) or biological nature (low abundance, expression in stages not analysed etc.). Secondly, the expression pattern of one of the major virulence factors, PfEMP1 and rifins (encoded by *var* and *rif* genes) was unexpected because they were both found to be already expressed by the sporozoite (insect) stage. This, however, is reminiscent of the situation in *Trypanosoma brucei* where a particular variant of the major surface glycoprotein that has its function in immune evasion in the bloodstream of the mammalian host, is already expressed in the insect stage of the parasite that is infective to mammals (Ginger *et al.*, 2002). It will be interesting to find out whether in *Plasmodium* a particular subset of *var* and *rif* genes is expressed in sporozoites. Thirdly, the proteome analysis identified several groups of coexpressed proteins. Some of these co-expressed groups are physically linked as small gene clusters on chromosomes. This observed co-expression indicates co-regulation of expression and possible functional associations between co-expressed proteins. It should be noted that a proteome analysis is not only helpful to analyse expression patterns of predicted genes but is also valuable to achieve a complete annotation of the genome. Currently, the cut-off length of automatically annotated putative genes is arbitrarily set at around 300 bp. Proteome analysis of yeast has, however, shown that even smaller open reading frames code for proteins (Oshiro *et al.*, 2002). Also, missed exons in the genome of *Plasmodium* could lead to the identification of incomplete genes. Proteome analysis and comparative genomics will both be valuable tools to improve the quality of gene identification.

The genomes of Kinetoplastida

The order of Kinetoplastida includes three parasite complexes of medical importance: *Leishmania* (several species)*, Trypanosoma cruzi* and *Trypanosoma brucei* (several species). The species *Leishmania major* and *L.*

donovani are widespread in South- and Central America, the Mediterranean region, Africa, the Middle East and Indian subcontinent (Herwaldt, 1999). *L. major* is the cause of cutaneous leishmaniasis, usually a mild and often self-healing disease, and the *L. donovani* species complex causes visceral leishmaniasis (Kala-Azar), a much more severe and life-threatening disease that affects internal organs such as spleen or liver. The parasites is transmitted by about 20 different species of sandflies. The second group of kinetoplastid parasites comprises *Trypanosoma cruzi*, the causative agent of Chagas disease (Rodriques Coura and de Castro, 2002). Its geographical distribution is restricted to South- and Central America. The vector is mainly transmitted by triatome bugs via faecal contamination of bite sites. The third group includes the *Trypanosoma brucei* complex (Hide, 1999). This parasite causes sleeping sickness and is restricted to sub-Saharan Africa. It is transmitted via the bite of tsetse flies. Two different species, *T. brucei rhodesiense and T. b. gambiense*, are responsible for an acute and chronic form of the disease, respectively. For disease control it is important to note that, in contrast to malaria, transmission of *Leishmania ssp.* and *Trypanosoma ssp.* is zoonotic with considerable host reservoirs amongst wild and domestic animals. Although kinetoplastid parasites share many aspects of their basic biology, their parasitic lifestyles and the associated biology related to pathogenicity are fundamentally different. The most obvious difference is that *Leishmania* and *T. cruzi* are intracellular parasites, whereas *T. brucei* is extracellular. Therefore, I will deal with their respective genome biology separately and point out common features where relevant. Due to the status of progression of the respective genome projects, only the *Leishmania* and *T. brucei* projects will be covered in this review.

The *Leishmania major* genome

Leishmania major (strain 'Friedlin') was chosen as the genome reference strain for technical and biological reasons: although a laboratory strain, it can still be passaged through the sandfly vector, it has served as a model for immunological aspects of infection and some stages of its lifecycle can be reproduced *in vitro* (Ravel e*t al.*, 1998). Throughout its lifecycle it is diploid and sexual stages and meiosis/genetic exchange have not been observed. The nuclear (haploid) genome size is approximately 34 Mb and the genome is composed of 36 chromosomes, ranging from 2.5 Mb to 0.3 Mb (Wincker *et al.*, 1996) . To sequence the entire genome, large insert genomic cosmid libraries were constructed, mapped to individual chromosomes separated by pulsed field gel electrophoresis, and finally sheared into shorter fragments and sequenced. The complete sequence of the smallest chromosome (chromosome I,) was published in 1999 and several more (chr 3, 4, 5, 24) have been completed since (Myler *et al.*, 1999). The estimated total number of genes is in the region of 8000. In January 2003 about 7500 genes were deposited in the central *Leishmania major* database GeneDB (http://www.genedb.org/genedb/leish/index.jsp). 53% of these genes are currently unclassified in terms of

potential functions, a proportion similar to that found in *Plasmodium falciparum*.

The overall sequence and organisation of the published chromosome I data revealed interesting features congruent with our understanding of some molecular mechanisms of *Leishmania* and other kinetoplastids. One of the unique features of the genome of trypanosomatids is the way genes are transcribed (Johnson *et al.*, 1987; Lee and Van der Ploeg, 1997). Unlike other eukaryotes, genes are transcribed into mRNA as polycistronic pre-mRNAs that are co-transcriptionally spliced into gene-sized units. In contrast to bacterial operons co-transcribed genes are rarely functionally related and no promoters have as yet been identified that drive transcription of polycistronic transcription units. Polycistronic transcription is reflected in the regional organisation of genes on chromosomes as genes that are part of a particular transcription unit are located on the same coding strand. It was, however, surprising to see that chromosome I has a strictly bi-directional structure: from an inversion point located about 70 kb distal from the left telomere, 29 genes are encoded on the same strand facing towards the left telomere and the remaining 50 genes are encoded by the opposite strand towards the right telomere (Donelson *et al.*, 1999). Transcription is directed away from the point of inversion (divergent). It is tempting to speculate that we are looking at two massive polycistronic transcription units. Nuclear run-on analysis suggests that this is indeed the case (Myler *et al.*, 2001). Analysis of the sequences of other chromosomes has revealed a similar organisation, although transcription can also be directed away from the telomeres (convergent). On the largest of these chromosomes, chromosome 35, several inversion points appear to exist, but the sequence of this chromosome is incomplete and the final product may change the current picture. Interest has focused on the role of the inversion point. It was hypothesised that they might function as origins of replication, promoters for RNA-polymerase II based transcription or act as centromeres (Donelson *et al.*, 1999; McDonagh *et al.*, 2000). In a recent study deletion mutants were created that lacked the inversion point on chromosome I (Dubessay *et al.*, 2002a). Analysis of the cell lines revealed that mitotic stability of the truncated chromosomes was not affected and neither was the ability to express marker genes introduced in close proximity of the inversion point. This argues against a function as centromere, origin of replication or initiation point of transcription. It was, however, remarkable that it was impossible to create a null mutant of this region. The attempt to delete this region on all alleles of chromosome I was countered by the cell by producing an extra copy of the entire chromosomes that contains the inversion point. The conclusion of this result was that the deleted region is essential for parasite survival. The same group also created cell clones with truncations of chromosome 1 that delete all coding DNA sequences from this chromosome (Dubessay *et al.*, 2002b). This truncated chromosome was mitotically stable and it was speculated that telomeric or subtelomeric repeat elements may act as centromeres. Unusually, this chromosome also contained a core region made up of tandemly repeated copies of ectopic plasmid

vector sequence, used for creating the deletion and selecting for transformants, stretching over 105 kb. Whether this additional repetitive DNA sequence has an influence on stability is not yet established.

One of the driving forces for this, as for any other, genome project is the prospect that the analysis of the genetic information will identify new drug targets. Experience from *Leishmania*, however, demonstrates that this route can be very tricky indeed. The surface of *Leishmania* cells is covered with a dense coat of glycoconjugates, amongst them membrane-bound lipophosphoglycan (LPG). Their extracellular location and abundance has led to the suggestion that these molecules are major virulence factors and may be important in establishing an infection and in parasite survival inside the host (Turco *et al.*, 2001). Using biochemical approaches and aided by available genome sequence information the biosynthetic pathways of LPG assembly have been established. Knowing the key enzymes it was possible to experimentally test the virulence hypothesis by knocking-out an essential enzyme of the biosynthetic pathway and thereby preventing the synthesis of the final LPG. When this experiment was done in *L. major*, the genome strain, the hypothesis was confirmed as the parasite lost its virulence (Späth *et al.*, 2000). However, when the equivalent LPG was knocked-out in *L. mexicana*, a closely related species also causing cutaneous leishmaniasis, no effect on virulence was observed (Ilg, 2000). These data demonstrate clearly the need for very detailed comparative studies even between closely related parasites to identify common biological traits of pharmacological interest. New drugs against the various clinical manifestations of leishmaniasis are urgently needed. Only three classes of drugs, pentavalent antimony compounds and, as a second line of treatment, amphotericin B and pentamidine, are available. All these compounds have disadvantages, such as difficult administration or toxicity (Guerin *et al.*, 2002b). In recent years only one compound, miltefosine, has been added to the list of promising anti-leishmaniasis drugs (licensed for use only in India so far) (Sundar *et al.*, 2002). The history of miltefosine, an alkylphosphocholine probably interfering with lipid metabolism (Lux *et al.*, 2000), demonstrates another paradigm in the development of anti-parasitic drugs. Miltefosine was first developed as an antitumor agent but turned out to be clinically ineffective. Only much later its effectiveness against leishmaniasis was discovered (Croft *et al.*, 1996; Croft *et al.*, 2003). As the development of new substances from synthesis to clinical evaluation is long and extremely expensive, it is difficult to see a commitment from the private sector to initiate such programmes in order to develop drugs specifically directed against parasitic diseases prevalent mainly in poor countries. Therefore, the 'hi-jacking' of already existing compounds for the treatment of parasitic diseases is a perfectly reasonable approach to reduce time of development and costs. As we have seen in the case of drugs effective against malaria (e.g. fosmidomycin), this 'piggy-bag' approach to increase the number of potential drugs appears to be the most promising option to date (Gelb and Hol, 2002). Genomic information will be of immense use to identify cellular components of the parasite (individual molecules, biochemical pathways, or compartments such as membranes) that might interact with already existing pharmacological substances.

The *Trypanosoma brucei* genome

Parasites belonging to the *Trypanosoma brucei* species complex are the causative agents of trypanosomiasis or African sleeping sickness (*T. b. rhodesiense* and *T. b. gambiense*) and nagana, a wasting disease of domestic cattle of major economical impact (*T. b. brucei*). The geographical distribution of these disease-causing trypanosome species is restricted to Sub-Saharan West- and East-Africa (Hide, 1999). 60 million people are currently at risk from the disease and according to WHO estimates between 300,000 - 500,000 are infected at any one time (http://www.who.int/emc/diseases/tryp/). The parasite is transmitted through the bite of the tsetse fly. The life cycle of the parasite, probably with the exception of *T. b. gambiense*, is zoonotic with large wild animal reservoirs, making control measures difficult. Recent years have seen a resurgence of the disease, even in areas that were previously unaffected (Fevre *et al.*, 2001; Hutchinson *et al.*, 2003). Clinically, the disease caused by the two human-infective forms are classified into a chronic form, that can last symptom-free for several years (caused by *T.b.gambiense)* and an acute form (*T. b. rhodesiense*). During stage I of the disease, the parasite circulates in the blood and other body fluids or tissues. Only in stage II, after the parasite has penetrated the blood-brain barrier the disease has an inevitable fatal outcome if not treated. There is a limited choice of drugs available for treatment of trypanosomiasis (Legros *et al.*, 2002). Pentamidin and suramin are the first choice drugs against stage I of the disease and have been in use for more than 60 years; melarsoprol, eflorithine and nifurtimox are used against stage II. Melarsoprol, since its first use 1949 the first choice for stage II treatment, is an organo-arsenical compound and highly toxic. Up to 5% of the fatality rate of treated trypanosomiasis patients can be attributed to toxic side effects of this drug. Moreover, recent therapeutic failure indicate the emergence of melarsoprolresistant parasites (Matovu *et al.*, 2001a; Matovu *et al.*, 2001b). Eflornithine, first used in 1981, is the only "new" compound registered for trypanosomiasis treatment. The drug is difficult to administer and is less effective against *T.b. rhodesiense*. Nifurtimox is still under assessment and, although licensed for treating Chagas disease caused by *Trypanosoma cruzi* in South America, large scale controlled trials have not yet been undertaken to assess its usefulness against African sleeping sickness (Burchmore *et al.*, 2002; Legros *et al.*, 2002). Furthermore, the key drugs to treat the disease (pentamidine, melarsoprol, eflornithine) are currently produced and supplied free of charge as gesture of good will by a pharmaceutical company (Aventis) and the long-term supply is not guaranteed.

The biology of kinetoplastids is distinguished by the presence of a number of biological processes that are either unique to these parasites or have been discovered first in these organisms. The posttranscriptional extensive editing of mitochondrial mRNA by insertion or deletion of uridylate

residues and the compartmentalisation of glycolysis in a specific organelle, the glycosome, are examples of unique features (Clayton and Michels, 1996; Guerra-Giraldez *et al.*, 2002; Madison-Antenucci *et al.*, 2002). *Trypanosoma brucei* has, however, made its entry into general molecular biology textbooks by the phenomenon of antigenic variation (Vickerman, 1985; 1989; Cross, 1996). In contrast to other eukaryotic parasites, African trypanosomes are extracellular parasites and do, at no stage during their life cycle, hide inside host cells like *Leishmania* or *Plasmodium*. In order to evade the host immune system the parasite has evolved a strategy that is now considered a general paradigm for monoallelic expression of a multigene family. In the bloodstream of the mammalian host the surface of the parasite is covered by a dense coat of about 10^7 identical molecules of a glycoprotein (variant surface glycoprotein, VSG). This molecule is highly immunogenic and within about seven days post-infection the host has mounted a strong immune response and is able to eliminate the parasite. However, within the population of parasites circulating in the blood a few have changed their surface coat to a different isotype of VSG, a process called VSG switching (Vanhamme *et al.*, 2001). This population is unrecognisable to the immune system until a new immune response is developed. By then, another subpopulation with yet another VSG coat will have emerged. The molecular mechanisms of antigenic variation are complex and not fully understood. The overwhelming importance of antigenic variation for parasite survival is, however, reflected in the genome organisation of trypanosomes and appears to be one of the important forces that shapes the overall organisation of the genome. In order to maintain the ability to switch and use new VSG variants the parasite maintains a reservoir of approximately one thousand VSG genes. The distribution of these genes is non-random. A relatively small set of genes is located within so-called expression sites (ES). About 20 ES are located - one per site - in subtelomeric location on most (but not all) of 11 pairs of diploid chromosomes and are, as the name suggests, the sites from where VSG genes are transcribed. However, only one ES is active at a given time within a single cell and all other ES are silenced by an hitherto unknown mechanism. In addition to the few VSG genes harboured in ES, the vast majority of VSG genes are located on internal sites of chromosomes, often in small clusters. The chromosomes referred to so far are also labelled, due to their size range of 1-6 Mbp, megabase chromosomes (Melville *et al.*, 1998). They contain, in addition to VSG-ES and internal VSG genes, all other "housekeeping" genes of the cell. In order to expand its VSG reservoir even further, trypanosomes maintain two extra classes of chromosomes that are, as far as is known, entirely devoted to VSG maintenance. Approximately 5 intermediate chromosomes (ca. 200 kilobase pairs in size) also harbour subtelomeric expression sites and about one hundred minichromosomes (30-150 kilobase pairs) also contain subtelomeric VSG genes, although not in the context of an ES (Ersfeld *et al.*, 1999; El-Sayed *et al.*, 2000). Minichromosomal VSG genes have been shown to be a preferred source for a genetic recombination mechanism that leads to the transposition of silent VSG genes into ES

on megabase chromosomes from where they can be potentially expressed (Robinson *et al.*, 1999). Therefore, minichromosomes appear to function exclusively as a VSG gene reservoir and are stably maintained during the cell division (Van der Ploeg *et al.*, 1984; Ersfeld and Gull, 1997). To make things even more complex, two different types of ES sites have been described: bloodstream ES, that are potentially active during the time the parasites resides in the mammalian host, and metacyclic ES, that are active in the tsetse fly during the life cycle stage that is infective to the mammalian host and seem to have the purpose to prime the parasite for life inside the bloodstream after transmission (Barry *et al.*, 1998; Bringaud *et al.*, 2001). Although both ES-types are subtelomeric, their internal architecture is very different (Ginger *et al.*, 2002). Taken together, trypanosomes devote more than 10 % of their genome, which is about 35 Mbp in size, to antigenic variation. In a recent publication several different bloodstream ES sites were characterised at DNA sequence level (Berriman *et al.*, 2002). The theme that emerged from this and previous studies is one of variation and conservation. ES are heterogeneous with respect to their gene content. In addition to a single VSG gene that is located close to the telomeric end of ES, ES contain numerous other genes collectively called ES associated genes. A function for most of these genes has, with one exception, not yet been established. Up to eight different ES-associated genes (ESAGs) can be harboured in one ES, flanked by a promoter and the telomere-proximal VSG gene. These genes, including the VSG gene, are under the control of a RNA-polymerase I-type promoter, an unusual situation as normally all protein coding genes in eukaryotes are transcribed by RNA-pol II (Zomerdijk *et al.*, 1990). A possible hypothesis for the presence of pol I VSG transcription in trypanosomes is the fact that the VSG constitutes about 10% of total cellular protein and that only a highly active polymerase, such as polI, permits transcription at such high levels. Transcription of all ES described so far runs towards the telomere. Analysis of different ES for their content of ESAGs shows that the number of these genes can vary considerably. None of the ES harbours all genes known as ESAGs, some harbour up to seven, some as few as three. The only ESAGs that are always present are ESAGs 6 and 7. These are the only ESAGs with an assigned function and they code for the dimeric transferrin receptor of trypanosomes (Salmon *et al.*, 1994). Except for these two essential genes, all other ESAGs seem to be non-essential. It cannot be excluded that they confer some sort of advantage in the natural environment of the parasite, not obvious under *in vitro* culture conditions. It is also remarkable that, in the ES analysed, ESAGs 6 and 7 are always located immediately downstream of the pol I promoter. Given a certain 'leakiness' of the pol I promoter, transcription of ESAGs 6 and 7 might also occurs on silent ES, where transcription of the VSG gene, located up to 60 kbp downstream of the promoter, is completely shut down. The role of differential expression of ferritin receptor variants from different ES in host adaptation has been discussed elsewhere (Bitter *et al.*, 1998). This diversity of ES is counterbalanced by a significant conservation in other sequence elements

(Berriman *et al.*, 2002). A characteristic 50 bp repeat is located immediately upstream of the ES promoter. The VSG genes are always preceded by a 70 bp repeat element. These elements are probably required for efficient recombination by duplicative transposition between silent VSG genes and expression site. It has, however, also been shown that the absence of the 70 bp repeat does not prevent recombination-based switching in cultured trypanosomes (McCulloch *et al.*, 1997). Towards the telomeres, VSG genes are associated with other repetitive element (Aline and Stuart, 1989). In addition to conservation of repetitive, noncoding DNA elements it was also shown that some genes and pseudogenes show a remarkable conservation between different ES within a single cell. It has been hypothesised that this conservation could indicate that expression sites have only recently been arisen by duplication events from a single ES (Berriman *et al.*, 2002). However, especially when considering a similar high degree of conservation between subtelomeric regions in the *Plasmodium* genome (see above) it seems equally likely that this conservation is the result of sequence homogenisation due to the recombinational activities that take place in subtelomeric regions of chromosomes. Such homogenisation of sequences will especially affect genes or pseudogenes with no inherent requirement for diversity and will be less obvious in genes where selective pressure drives genes towards sequence diversity (e.g. VSG genes, transferrin receptors).

A unique feature of subtelomeric regions in the trypanosome genome is the presence of a recently discovered family of genes, termed retrotransposon hot spot (RHS) genes (Bringaud *et al.*, 2002). Based on sequence similarities the RHS genes have been divided into six subfamilies (RHS 1-6). In addition to intact genes, many RHS pseudogenes coexist within the RHS clusters. These pseudogenes arise either by frameshifts and stop codons within the open reading frame, or by the insertion of retrotransposon sequences. These insertions occur at exactly the same relative position in all members of this family. The total copy number of RHS genes/pseudogenes has been estimated to be about 130. The RHS sequences show no significant similarities to other proteins and their function is unknown. Antibodies raised against recombinant RHS proteins of each subfamily have shown that, with the possible exception of RHS2, the proteins are located in the nucleus. RHS-related sequence have also been identified in *Trypanosoma cruzi*. RHS genes/pseudogenes are clustered in the genome and frequently occur as tandem repeats. The clusters analysed so far are found upstream of bloodstream expression sites, but also in subtelomeric location on chromosomes not carrying a bloodstream ES. On the publicly accessible sequence (acc.no. AL929603) of chromosome I of *T. brucei* stock TREU927, the genome reference strain, the RHS cluster stretches over a length of about 190 kbp from immediately upstream of the VSG promoter towards the centre of the left hand side of this chromosome.

The overall architecture of chromosomes of *T. brucei* is very similar to that of *Leishmania*: many genes are located on single continuous coding strand with few inversion points on one chromosome. No centromeres have been identified yet. In contrast to *T. brucei*, *Leishmania* lacks the presence of retroelements within the genome. As they are also found in the kinetoplastids *T. cruzi* and *Crithidia fasciculata* their absence is remarkable and could hint at some fundamental differences in genome maintenance and dynamics (Bhattacharya *et al.*, 2002).

What will be the potential impact of the *T. brucei* genome project on combating the disease? As the project is not yet completed only few studies employing a genome wide analysis to identify novel features of trypanosome biology have been published. When searching the databases for proteins that are located in the glycosome, a kinetoplastid specific organelle related to peroxisomes that contains most glycolytic enzymes, Hannaert et al identified several genes encoding for homologs of plant enzymes (Guerra-Giraldez *et al.*, 2002; Hannaert *et al.*, 2003). A total of 16 different genes were identified coding for proteins involved in a variety of cellular functions ranging from carbohydrate metabolism to fatty acid synthesis. Interestingly, most of these proteins contain a targeting signal typical for glycosomal proteins. In contrast to the presence of the apicoplast in *Plasmodium* and related organisms, trypanosomes do not contain the evolutionary remnants of a plant organelle. The authors therefore speculate that an extensive gene transfer occurred from an endosymbiontic organelle genome to the nuclear genome before this endosymbiont of algal origin was lost from the precursor of today's trypanosomes. Whatever the mechanism of acquisition of these plant-like genes was, it now offers, analogous to malaria parasites, the possibility that herbicide-type substances could be used to target the enzymes encoded by these genes. Orthologous genes have also been identified in *Leishmania* (Hannaert *et al.*, 2003). In addition to such a genome wide analysis to search for novel drug targets, the genome information will also greatly aid gene discovery within other projects, particularly in combination with mass spectrometry technologies (Oliver, 2002; Beverley, 2003). Examples are the elucidation of the pathway GPI-anchor biosynthesis and fatty acid biosynthesis (Ferguson, 2000; Morita *et al.*, 2000a; Morita *et al.*, 2000b; Chang *et al.*, 2002) or the identification of components of the RNA-editing machinery (Schnaufer *et al.*, 2001; Stuart *et al.*, 2002).

Using *Plasmodium falciparum, Leishmania* and *Trypanosoma brucei* I have tried to review the status of current protozoan parasite genome projects and given examples of read-outs of the sequence information at various levels of analysis. Many parallels can be drawn when comparing parasite genomes. Metabolic pathways have been identified that are unique to the parasites and offer potential drug targets. The organisation of the genetic information on individual chromosomes strongly reflects the necessities imposed on the parasite due to host/parasite interactions. Genes involved in antigenic variation/immune evasion are located near the telomeric ends of chromosomes where they are exposed to the desired effects of an active recombination machinery to maintain as much diversity as possible (Cano, 2001; Barry *et al.*, 2003). Without doubt the availability of genome information will have a major impact on basic research in parasitology, particular in molecular science and biochemistry. However,

justification to fund parasite genome projects was mainly the prospect to develop new drugs or vaccines against infectious diseases. How likely is it that this aim will be realised? They are certainly voices that doubt the impact of genomics in parasite control (Curtis, 2000) and drug discovery in general (Coghlan, 2002). There are indeed data that indicate an overall decrease in productivity in pharmaceutical research (Taylor, 2003). The number of drugs containing new active ingredients has steadily dropped over the previous two decades and it is hard to see that the development of drugs against parasitic diseases largely confined to poor countries would run contrary to this trend. It has been estimated that 20 to 30 new drugs will be needed for a sustainable control of the most important protozoan diseases (Gelb and Hol, 2002). It appears very unlikely that private initiative by companies will raise the necessary funds to develop new drugs against tropical diseases. The costs for developing a new drug is put at around $ 500 million, and, unless different safety standards and regulatory procedures are applied, these costs will be of a similar magnitude for drugs against tropical diseases. If only for ethical reasons, a scenario of "cutting corners" is not acceptable. On the other hand, the objective of drug development against non-infectious diseases is much more complex in comparison to develop a drug against a parasite. In the former case modulation (activation/inhibition/substitution) of endogenous gene products (proteins or metabolic products) is usually required whereas drugs against infectious diseases are designed with only one (simple) aim: to kill the parasite with minimal side-effects on the patient. Screening of compound libraries is therefore more straightforward. Another feasible way of reducing costs has been already mentioned: the exploitation of already existing substances. Drugs that show a good efficiency against malaria (dapson, fosmidomycin) and visceral leishmaniasis (miltefosine), have originally been developed for other applications but are now 'hijacked' and being evaluated to treat infectious diseases (Guerin et al., 2002a; Sundar et al., 2002). The identification of a relatively large number of gene products closely related to plant enzymes both in malaria and kinetoplastids opens possibilities to test a relatively defined set of pharmacological substances (e.g. herbicides) on their effect on these target molecules. In addition it will be necessary to redefine the role of academia in applied research. More development leading towards the clinical application of basic science will need to be done within academic institutions. Closer links between public and private institutions will be necessary to translate research into applications. Finally, a note of caution concerning the informative value of genomic information. As extensively discussed for the human genome project, a genome project is not completed by sequencing a single genome. This is particularly true for parasite genomes strains that are adapted to in vitro cultures. Single genomes can only serve as a reference for detailed comparative studies (Kissinger et al., 2002). Although conservation of syntenic groups of genes has been demonstrated for various related parasites (Bringaud et al., 1998; Gull, 2001; Carlton et al., 2002), recent publications have revealed a surprising genetic variety between various Plasmodium falciparum isolates (Volkman et al., 2002; Wootton et al., 2002; Duraisingh et al., 2003). In one study it was shown that receptor proteins involved in red blood cell invasion are expressed variantly in different lines. It was hypothezised that this expression patterns are an adaptive response to heterogeneity of erythrocyte surface proteins to maximise the parasites' opportunities to survive inside the host (Duraisingh et al., 2003). In a second study it was demonstrated that mutations in a gene associated with chloroquine resistance spread through the parasite population in relatively short period of time (Wootton et al., 2002). Such genetic micro-heterogeneity has direct implications for drug and vaccine development (McConnell, 2002; Wongsrichanalai et al., 2002). Population analysis will be necessary on a genome wide scale to avoid focusing on targets that will turn out to be too heterogeneous in wild-type populations to be of use as drug targets.

Acknowledgements

References

Aline, R.F., and Stuart, K. 1989 Trypanosoma brucei: conserved sequence organization 3' to telomeric variant surface glycoprotein genes. Experimental Parasitology 68: 57-66.

Altincicek, B., Duin, E.C., Reichenberg, A., Hedderich, R., Kollas, A.K., Hintz, M., Wagner, S., Wiesner, J., Beck, E., and Jomaa, H. 2002 LytB protein catalyzes the terminal step of the 2-C-methyl-D-erythritol-4-phosphate pathway of isoprenoid biosynthesis. FEBS Lett 532: 437-440.

Bannister, L.H., Hopkins, J.M., Fowler, R.E., Krishna, S., and Mitchell, G.H. 2000 A brief illustrated guide to the ultrastructure of Plasmodium falciparum asexual blood stages. Parasitol Today 16: 427-433.

Barry, J.D., Ginger, M.L., Burton, P., and McCulloch, R. 2003 Why are parasite contingency genes often associated with telomeres? Int J Parasitol 33: 29-45.

Barry, J.D., Graham, S.V., Fotheringham, M., Graham, V.S., Kobryn, K., and Wymer, B. 1998 VSG gene control and infectivity strategy of metacyclic stage Trypanosoma brucei. Mol Biochem Parasitol 91: 93-105.

Beeson, J.G., Winstanley, P.A., McFadden, G.I., and Brown, G.V. 2001 New agents to combat malaria. Nat Med 7: 149-150.

Berriman, M., Hall, N., Sheader, K., Bringaud, F., Tiwari, B., Isobe, T., Bowman, S., Corton, C., Clark, L., Cross, G.A., Hoek, M., Zanders, T., Berberof, M., Borst, P., and Rudenko, G. 2002 The architecture of variant surface glycoprotein gene expression sites in Trypanosoma brucei. Mol Biochem Parasitol 122: 131-140.

Beverley, S.M. 2003 Protozomics: trypanosomatid parasite genetics comes of age. Nat Rev Genet 4: 11-19.

Bhattacharya, S., Bakre, A., and Bhattacharya, A. 2002 Mobile genetic elements in protozoan parasites. J Genet 81: 73-86.

Bitter, W., Gerrits, H., Kieft, R., and Borst, P. 1998 The role of transferrin-receptor variation in the host range of

Trypanosoma brucei. Nature 391**:** 499-502.

Bowman, S., Lawson, D., Basham, D., Brown, D., Chillingworth, T., Churcher, C.M., Craig, A., Davies, R.M., Devlin, K., Feltwell, T., Gentles, S., Gwilliam, R., Hamlin, N., Harris, D., Holroyd, S., Hornsby, T., Horrocks, P., Jagels, K., Jassal, B., Kyes, S., McLean, J., Moule, S., Mungall, K., Murphy, L., Barrell, B.G., and et al. 1999 The complete nucleotide sequence of chromosome 3 of *Plasmodium falciparum*. Nature 400**:** 532-538.

Bringaud, F., Biteau, N., Donelson, J.E., and Baltz, T. 2001 Conservation of metacyclic variant surface glycoprotein expression sites among different trypanosome isolates. Mol Biochem Parasitol 113**:** 67-78.

Bringaud, F., Biteau, N., Melville, S.E., Hez, S., El-Sayed, N.M., Leech, V., Berriman, M., Hall, N., Donelson, J.E., and Baltz, T. 2002 A new, expressed multigene family containing a hot spot for insertion of retroelements is associated with polymorphic subtelomeric regions of *Trypanosoma brucei*. Eukaryot Cell 1**:** 137-151.

Bringaud, F., Vedrenne, C., Cuvillier, A., Parzy, D., Baltz, D., Tetaud, E., Pays, E., Venegas, J., Merlin, G., and Baltz, T. 1998 Conserved organization of genes in trypanosomatids. Mol Biochem Parasitol 94**:** 249-264.

Burchmore, R.J., Ogbunude, P.O., Enanga, B., and Barrett, M.P. 2002 Chemotherapy of human African trypanosomiasis. Curr Pharm Des 8**:** 256-267

Cano, M.I. 2001 Telomere biology of trypanosomatids: more questions than answers. Trends Parasitol 17**:** 425-429.

Carlton, J.M., Angiuoli, S.V., Suh, B.B., Kooij, T.W., Pertea, M., Silva, J.C., Ermolaeva, M.D., Allen, J.E., Selengut, J.D., Koo, H.L., Peterson, J.D., Pop, M., Kosack, D.S., Shumway, M.F., Bidwell, S.L., Shallom, S.J., van Aken, S.E., Riedmuller, S.B., Feldblyum, T.V., Cho, J.K., Quackenbush, J., Sedegah, M., Shoaibi, A., Cummings, L.M., Florens, L., Yates, J.R., Raine, J.D., Sinden, R.E., Harris, M.A., Cunningham, D.A., Preiser, P.R., Bergman, L.W., Vaidya, A.B., van Lin, L.H., Janse, C.J., Waters, A.P., Smith, H.O., White, O.R., Salzberg, S.L., Venter, J.C., Fraser, C.M., Hoffman, S.L., Gardner, M.J., and Carucci, D.J. 2002 Genome sequence and comparative analysis of the model rodent malaria parasite *Plasmodium yoelii yoelii*. Nature 419**:** 512-519.

Carlton, J.M., and Carucci, D.J. 2002 Rodent models of malaria in the genomics era. Trends Parasitol 18**:** 100-102.

Carlton, J.M., Muller, R., Yowell, C.A., Fluegge, M.R., Sturrock, K.A., Pritt, J.R., Vargas-Serrato, E., Galinski, M.R., Barnwell, J.W., Mulder, N., Kanapin, A., Cawley, S.E., Hide, W.A., and Dame, J.B. 2001 Profiling the malaria genome: a gene survey of three species of malaria parasite with comparison to other apicomplexan species. Mol Biochem Parasitol 118**:** 201-210.

Carucci, D.J., Yates, J.R., and Florens, L. 2002 Exploring the proteome of *Plasmodium*. Int J Parasitol 32**:** 1539-1542.

Chang, T., Milne, K.G., Guther, M.L., Smith, T.K., and Ferguson, M.A. 2002 Cloning of *Trypanosoma brucei* and *Leishmania major* genes encoding the GlcNAc-PI de-Nacetylase of glycosylphosphatidylinositol biosynthesis that is essential to the african sleeping sickness parasite.

J Biol Chem 2**:** 2.

Clayton, C.E., and Michels, P. 1996 Metabolic compartmentation in African trypanosomes. Parasitology Today 12**:** 465-471.

Coghlan, A. 2002 Gene data slows drug discovery. New Scientists 175**:** 11.

Cooke, B., Coppel, R., and Wahlgren, M. 2000 Falciparum malaria: sticking up, standing out and out-standing. Parasitol Today 16**:** 416-420.

Crabb, B.S., and Cowman, A.F. 2002 *Plasmodium falciparum* virulence determinants unveiled. Genome Biol 3**:** 1-4.

Croft, S.L., Seifert, K., and Duchene, M. 2003 Antiprotozoal activities of phospholipid analogues. Molecular and Biochemical Parasitology 126**:** 165-172.

Croft, S.L., Snowdon, D., and Yardley, V. 1996 The activities of four anticancer alkyllysophospholipids against *Leishmania donovani*, *Trypanosoma cruzi* and *Trypanosoma brucei*. J Antimicrob Chemother 38**:** 1041-1047.

Cross, G.A. 1996 Antigenic variation in trypanosomes: secrets surface slowly. Bioessays 18**:** 283-291.

Curtis, C.F. 2000 The case for deemphasizing genomics in malaria contro. Science 290**:** 1508.

Dame, J.B., Arnot, D.E., Bourke, P.F., Chakrabarti, D., Christodoulou, Z., Coppel, R.L., Cowman, A.F., Craig, A.G., Fischer, K., Foster, J., Goodman, N., Hinterberg, K., Holder, A.A., Holt, D.C., Kemp, D.J., Lanzer, M., Lim, A., Newbold, C.I., Ravetch, J.V., Reddy, G.R., Rubio, J., Schuster, S.M., Su, X.Z., Thompson, J.K., Werner, E.B., and et al. 1996 Current status of the *Plasmodium falciparum* genome project. Mol Biochem Parasitol 79**:** 1-12.

De Gregorio, E., and Lemaitre, B. 2002 The post-genomic era opens. Nature 419**:** 496-497.

Donelson, J.E., Gardner, M.J., and El-Sayed, N.M. 1999 More surprises from Kinetoplastida. Proc. Natl. Acad. Sci .USA. 96**:** 2579-2581.

Doolittle, R.F. 2002 The grand assault. Nature 419**:** 493-494.

Dubessay, P., Ravel, C., Bastien, P., Crobu, L., Dedet, J.P., Pages, M., and Blaineau, C. 2002a The switch region on *Leishmania major* chromosome 1 is not required for mitotic stability or gene expression, but appears to be essential.PG. Nucleic Acids Res 30**:** 3692-3697.

Dubessay, P., Ravel, C., Bastien, P., Stuart, K., Dedet, J.P., Blaineau, C., and Pages, M. 2002b Mitotic stability of a coding DNA sequence-free version of *Leishmania major* chromosome 1 generated by targeted chromosome fragmentation. Gene 289**:** 151-159.

Duraisingh, M.T., Triglia, T., Ralph, S.A., Rayner, J.C., Barnwell, J.W., McFadden, G.I., and Cowman, A.F. 2003 Phenotypic variation of *Plasmodium falciparum* merozoite proteins directs receptor targeting for invasion of human erythrocytes. EMBO J. 22**:** 1047-1057.

El-Sayed, N.M., Hegde, P., Quackenbush, J., Melville, S.E., and Donelson, J.E. 2000 The African trypanosome genome. Int J Parasitol 30**:** 329-345.

Ersfeld, K., and Gull, K. 1997 Partitioning of large and minichromosomes in *Trypanosoma brucei*. Science 276**:** 611-614.

Ersfeld, K., Melville, S.E., and Gull, K. 1999 Nuclear and genome organization of *Trypanosoma brucei*. Parasitol Today 15**:** 58-63.

Ferguson, M.A.J. 2000 Glycosylphosphatidylinositol biosynthesis validated as a drug target for African sleeping sickness. PNAS 97**:** 10673-10675.

Fevre, E.M., Coleman, P.G., Odiit, M., Magona, J.W., Welburn, S.C., and Woolhouse, M.E. 2001 The origins of a new *Trypanosoma brucei rhodesiense* sleeping sickness outbreak in eastern Uganda. Lancet 358**:** 625-628.

Fichera, M.E., and Roos, D.S. 1997 A plastid organelle as a drug target in apicomplexan parasites. Nature 390**:** 407-409.

Florens, L., Washburn, M.P., Raine, J.D., Anthony, R.M., Grainger, M., Haynes, J.D., Moch, J.K., Muster, N., Sacci, J.B., Tabb, D.L., Witney, A.A., Wolters, D., Wu, Y., Gardner, M.J., Holder, A.A., Sinden, R.E., Yates, J.R., and Carucci, D.J. 2002 A proteomic view of the *Plasmodium falciparum* life cycle. Nature 419**:** 520-526.

Foth, B.J., Ralph, S.A., Tonkin, C.J., Struck, N.S., Fraunholz, M., Roos, D.S., Cowman, A.F., and McFadden, G.I. 2003 Dissecting apicoplast targeting in the malaria parasite *Plasmodium falciparum*. Science 299**:** 705-708.

Freitas-Junior, L.H., Bottius, E., Pirrit, L.A., Deitsch, K.W., Scheidig, C., Guinet, F., Nehrbass, U., Wellems, T.E., and Scherf, A. 2000 Frequent ectopic recombination of virulence factor genes in telomeric chromosome clusters of *P. falciparum*. Nature 407**:** 1018-1022.

Gardner, M.J. 2001 A status report on the sequencing and annotation of the *P. falciparum* genome. Mol Biochem Parasitol 118**:** 133-138.

Gardner, M.J., Hall, N., Fung, E., White, O., Berriman, M., Hyman, R.W., Carlton, J.M., Pain, A., Nelson, K.E., Bowman, S., Paulsen, I.T., James, K., Eisen, J.A., Rutherford, K., Salzberg, S.L., Craig, A., Kyes, S., Chan, M.S., Nene, V., Shallom, S.J., Suh, B., Peterson, J., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., Haft, D., Mather, M.W., Vaidya, A.B., Martin, D.M., Fairlamb, A.H., Fraunholz, M.J., Roos, D.S., Ralph, S.A., McFadden, G.I., Cummings, L.M., Subramanian, G.M., Mungall, C., Venter, J.C., Carucci, D.J., Hoffman, S.L., Newbold, C., Davis, R.W., Fraser, C.M., and Barrell, B. 2002a Genome sequence of the human malaria parasite *Plasmodium falciparum*. Nature 419**:** 498-511.

Gardner, M.J., Shallom, S.J., Carlton, J.M., Salzberg, S.L., Nene, V., Shoaibi, A., Ciecko, A., Lynn, J., Rizzo, M., Weaver, B., Jarrahi, B., Brenner, M., Parvizi, B., Tallon, L., Moazzez, A., Granger, D., Fujii, C., Hansen, C., Pederson, J., Feldblyum, T., Peterson, J., Suh, B., Angiuoli, S., Pertea, M., Allen, J., Selengut, J., White, O., Cummings, L.M., Smith, H.O., Adams, M.D., Venter, J.C., Carucci, D.J., Hoffman, S.L., and Fraser, C.M. 2002b Sequence of Plasmodium falciparum chromosomes 2, 10, 11 and 14. Nature 419**:** 531-534.

Gardner, M.J., Tettelin, H., Carucci, D.J., Cummings, L.M., Aravind, L., Koonin, E.V., Shallom, S., Mason, T., Yu, K., Fujii, C., Pederson, J., Shen, K., Jing, J., Aston, C., Lai, Z., Schwartz, D.C., Pertea, M., Salzberg, S., Zhou, L., Sutton, G.G., Clayton, R., White, O., Smith, H.O., Fraser, C.M., Hoffman, S.L., and et al. 1998 Chromosome 2

sequence of the human malaria parasite *Plasmodium falciparum*. Science 282**:** 1126-1132.

Gelb, M.H., and Hol, W.G. 2002 Drugs to combat tropical protozoan diseases. Science 297**:** 343-344.

Ginger, M.L., Blundell, P.A., Lewis, A.M., Browitt, A., Gunzl, A., and Barry, J.D. 2002 Ex Vivo and In Vitro Identification of a Consensus Promoter for VSG Genes Expressed by Metacyclic-Stage Trypanosomes in the Tsetse Fly. Eukaryot Cell 1**:** 1000-1009.

Guerin, P.J., Olliaro, P., Nosten, F., Druilhe, P., Laxminarayan, R., Binka, F., Kilama, W.L., Ford, N., and White, N.J. 2002a Malaria: current status of control, diagnosis, treatment, and a proposed agenda for research and development. Lancet Infect Dis 2**:** 564-573.

Guerin, P.J., Olliaro, P., Sundar, S., Boelaert, M., Croft, S.L., Desjeux, P., Wasunna, M.K., and Bryceson, A.D. 2002b Visceral leishmaniasis: current status of control, diagnosis, and treatment, and a proposed research and development agenda. Lancet Infect Dis 2**:** 494-501.

Guerra-Giraldez, C., Quijada, L., and Clayton, C.E. 2002 Compartmentation of enzymes in a microbody, the glycosome, is essential in *Trypanosoma brucei*. J Cell Sci 115**:** 2651-2658.

Gull, K. 2001 The biology of kinetoplastid parasites: insights and challenges from genomics and post-genomics. Int J Parasitol 31**:** 443-452.

Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D., Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K., Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C., Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C., Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J., Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N., Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S., Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D., Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N., Lawson, D., Lennard, N., Line, A., Maddison, M., McLean, J., Mooney, P., Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A., Rabbinowitsch, E., Rajandream, M.A., Rutter, S., Rutherford, K.M., Sanders, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R., Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L., Whitehead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C., and Barrell, B.G. 2002 Sequence of *Plasmodium falciparum* chromosomes 1, 3-9 and 13. Nature 419**:** 527-531.

Hannaert, V., Saavedra, E., Duffieux, F., Szikora, J.-P., Rigden, D.J., Michels, P.A.M., and Opperdoes, F.R. 2003 Plant-like traits associated with metabolism of *Trypanosoma* parasites. PNAS**:** 0335769100.

Hemingway, J., Field, L., and Vontas, J. 2002 An overview of insecticide resistance. Science 298**:** 96-97.

Herwaldt, B.L. 1999 Leishmaniasis. Lancet 354**:** 1191-1199.

Hide, G. 1999 History of sleeping sickness in East Africa. Clin Microbiol Rev 12**:** 112-125.

Hoffman, S.L. 2000 Research (genomics) is crucial to attacking malaria. Science 290**:** 1509.

Hoffman, S.L., Bancroft, W.H., Gottlieb, M., James, S.L., Burroughs, E.C., Stephenson, J.R., and Morgan, M.J.

1997 Funding for malaria genome sequencing. Nature 387**:** 647.

Holt, R.A., Subramanian, G.M., Halpern, A., Sutton, G.G., Charlab, R., Nusskern, D.R., Wincker, P., Clark, A.G., Ribeiro, J.M., Wides, R., Salzberg, S.L., Loftus, B., Yandell, M., Majoros, W.H., Rusch, D.B., Lai, Z., Kraft, C.L., Abril, J.F., Anthouard, V., Arensburger, P., Atkinson, P.W., Baden, H., de Berardinis, V., Baldwin, D., Benes, V., Biedler, J., Blass, C., Bolanos, R., Boscus, D., Barnstead, M., Cai, S., Center, A., Chatuverdi, K., Christophides, G.K., Chrystal, M.A., Clamp, M., Cravchik, A., Curwen, V., Dana, A., Delcher, A., Dew, I., Evans, C.A., Flanigan, M., Grundschober-Freimoser, A., Friedli, L., Gu, Z., Guan, P., Guigo, R., Hillenmeyer, M.E., Hladun, S.L., Hogan, J.R., Hong, Y.S., Hoover, J., Jaillon, O., Ke, Z., Kodira, C., Kokoza, E., Koutsos, A., Letunic, I., Levitsky, A., Liang, Y., Lin, J.J., Lobo, N.F., Lopez, J.R., Malek, J.A., McIntosh, T.C., Meister, S., Miller, J., Mobarry, C., Mongin, E., Murphy, S.D., O'Brochta, D.A., Pfannkoch, C., Qi, R., Regier, M.A., Remington, K., Shao, H., Sharakhova, M.V., Sitter, C.D., Shetty, J., Smith, T.J., Strong, R., Sun, J., Thomasova, D., Ton, L.Q., Topalis, P., Tu, Z., Unger, M.F., Walenz, B., Wang, A., Wang, J., Wang, M., Wang, X., Woodford, K.J., Wortman, J.R., Wu, M., Yao, A., Zdobnov, E.M., Zhang, H., Zhao, Q., et al. 2002 The genome sequence of the malaria mosquito *Anopheles gambiae*. Science 298**:** 129-149.

Hopkins, J., Fowler, R., Krishna, S., Wilson, I., Mitchell, G., and Bannister, L. 1999 The plastid in *Plasmodium falciparum* asexual blood stages: a three-dimensional ultrastructural analysis. Protist 150**:** 283-295.

Hutchinson, O.C., Fevre, E.M., Carrington, M., and Welburn, S.C. 2003 Lessons learned from the emergence of a new *Trypanosoma brucei rhodesiense* sleeping sickness focus in Uganda. Lancet Infect Dis 3**:** 42-45.

Hyman, R.W., Fung, E., Conway, A., Kurdi, O., Mao, J., Miranda, M., Nakao, B., Rowley, D., Tamaki, T., Wang, F., and Davis, R.W. 2002 Sequence of *Plasmodium falciparum* chromosome 12. Nature 419**:** 534-537.

Ilg, T. 2000 Lipophosphoglycan is not required for infection of macrophages or mice by *Leishmania mexicana*. Embo J 19**:** 1953-1962.

Johnson, P.J., Kooter, J.M., and Borst, P. 1987 Inactivation of transcription by UV irradiation of *T. brucei* provides evidence for a multicistronic transcription unit including a VSG gene. Cell 51**:** 273-281.

Jomaa, H., Wiesner, J., Sanderbrand, S., Altincicek, B., Weidemeyer, C., Hintz, M., Turbachova, I., Eberl, M., Zeidler, J., Lichtenthaler, H.K., Soldati, D., and Beck, E. 1999 Inhibitors of the nonmevalonate pathway of isoprenoid biosynthesis as antimalarial drugs. Science 285**:** 1573-1576.

Kapoor, M., Dar, M.J., Surolia, A., and Surolia, N. 2001 Kinetic determinants of the interaction of enoyl-ACP reductase from *Plasmodium falciparum* with its substrates and inhibitors. Biochem Biophys Res Commun 289**:** 832-837.

Kaviratne, M., Khan, S.M., Jarra, W., and Preiser, P.R. 2002 Small Variant STEVOR Antigen Is Uniquely Located within Maurer's Clefts in *Plasmodium falciparum*-Infected Red Blood Cells. Eukaryot Cell 1**:** 926-935.

Kissinger, J.C., Brunk, B.P., Crabtree, J., Fraunholz, M.J., Gajria, B., Milgram, A.J., Pearson, D.S., Schug, J., Bahl, A., Diskin, S.J., Ginsburg, H., Grant, G.R., Gupta, D., Labo, P., Li, L., Mailman, M.D., McWeeney, S.K., Whetzel, P., Stoeckert, C.J., and Roos, D.S. 2002 The *Plasmodium* genome database. Nature 419**:** 490-492.

Köhler, S., Delwiche, C.F., Denny, P.W., Tilney, L.G., Webster, P., Wilson, R.J., Palmer, J.D., and Roos, D.S. 1997 A plastid of probable green algal origin in Apicomplexan parasites. Science 275**:** 1485-1489.

Kollas, A.K., Duin, E.C., Eberl, M., Altincicek, B., Hintz, M., Reichenberg, A., Henschker, D., Henne, A., Steinbrecher, I., Ostrovsky, D.N., Hedderich, R., Beck, E., Jomaa, H., and Wiesner, J. 2002 Functional characterization of GcpE, an essential enzyme of the non-mevalonate pathway of isoprenoid biosynthesis.
FEBS Lett 532**:** 432-436.

Kyes, S., Horrocks, P., and Newbold, C. 2001 Antigenic variation at the infected red cell surface in malaria. Annu Rev Microbiol 55**:** 673-707.

Kyes, S.A., Rowe, J.A., Kriek, N., and Newbold, C.I. 1999 Rifins: a second family of clonally variant proteins expressed on the surface of red cells infected with *Plasmodium falciparum*. Proc. Natl. Acad. Sci .USA. 96**:** 9333-9338.

Lai, Z., Jing, J., Aston, C., Clarke, V., Apodaca, J., Dimalanta, E.T., Carucci, D.J., Gardner, M.J., Mishra, B., Anantharaman, T.S., Paxia, S., Hoffman, S.L., Craig Venter, J., Huff, E.J., and Schwartz, D.C. 1999 A shotgun optical map of the entire *Plasmodium falciparum* genome. Nat Genet 23**:** 309-313.

Lasonder, E., Ishihama, Y., Andersen, J.S., Vermunt, A.M., Pain, A., Sauerwein, R.W., Eling, W.M., Hall, N., Waters, A.P., Stunnenberg, H.G., and Mann, M. 2002 Analysis of the *Plasmodium falciparum* proteome by high-accuracy mass spectrometry. Nature 419**:** 537-542.

Lee, M.G., and Van der Ploeg, L.H. 1997 Transcription of protein-coding genes in trypanosomes by RNA polymerase I. Annu Rev Microbiol 51**:** 463-489.

Legros, D., Ollivier, G., Gastellu-Etchegorry, M., Paquet, C., Burri, C., Jannin, J., and Buscher, P. 2002 Treatment of human African trypanosomiasis—present situation and needs for research and development. Lancet Infect Dis 2**:** 437-440.

Lell, B., Ruangweerayut, R., Wiesner, J., Missinou, M.A., Schindler, A., Baranek, T., Hintz, M., Hutchinson, D., Jomaa, H., and Kremsner, P.G. 2003 Fosmidomycin, a novel chemotherapeutic agent for malaria. Antimicrob Agents Chemother 47**:** 735-738.

Lux, H., Heise, N., Klenner, T., Hart, D., and Opperdoes, F.R. 2000 Ether—lipid (alkylphospholipid) metabolism and the mechanism of action of ether—lipid analogues in *Leishmania*. Mol Biochem Parasitol 111**:** 1-14.

Madison-Antenucci, S., Grams, J., and Hajduk, S.L. 2002 Editing machines: the complexities of trypanosome RNA editing. Cell 108**:** 435-438.

Matovu, E., Enyaru, J.C., Legros, D., Schmid, C., Seebeck, T., and Kaminsky, R. 2001a Melarsoprol refractory *T. b. gambiense* from Omugo, north-western Uganda. Trop Med Int Health 6**:** 407-411.

Matovu, E., Seebeck, T., Enyaru, J.C., and Kaminsky, R.

2001b Drug resistance in *Trypanosoma brucei* spp., the causative agents of sleeping sickness in man and nagana in cattle. Microbes Infect 3**:** 763-770.

McConnell 2002 Malaria number game. The Lancet Infectious Diseases 2**:** 263.

McCulloch, R., Rudenko, G., and Borst, P. 1997 Gene conversions mediating antigenic variation in *Trypanosoma brucei* can occur in variant surface glycoprotein expression sites lacking 70-base-pair repeat sequences. Molecular and Cellular Biology 17**:** 833-843.

McDonagh, P.D., Myler, P.J., and Stuart, K. 2000 The unusual gene organization of *Leishmania major* chromosome 1 may reflect novel transcription processes. Nucleic Acids Res 28**:** 2800-2803.

Melville, S.E., Leech, V., Gerrard, C.S., Tait, A., and Blackwell, J.M. 1998 The molecular karyotype of the megabase chromosomes of *Trypanosoma brucei* and the assignment of chromosome markers. Mol Biochem Parasitol 94**:** 155-173.

Miller, L.H., and Greenwood, B. 2002 Malaria—a shadow over Africa. Science 298**:** 121-122.

Missinou, M.A., Borrmann, S., Schindler, A., Issifou, S., Adegnika, A.A., Matsiegui, P.B., Binder, R., Lell, B., Wiesner, J., Baranek, T., Jomaa, H., and Kremsner, P.G. 2002 Fosmidomycin for malaria. Lancet 360**:** 1941-1942.

Morita, Y.S., Acosta-Serrano, A., Buxbaum, L.U., and Englund, P.T. 2000a Glycosyl phosphatidylinositol myristoylation in African trypanosomes. New intermediates in the pathway for fatty acid remodeling. J Biol Chem 275**:** 14147-14154.

Morita, Y.S., Paul, K.S., and Englund, P.T. 2000b Specialized fatty acid synthesis in African trypanosomes: myristate for GPI anchors. Science 288**:** 140-143.

Mota, M.M., Thathy, V., Nussenzweig, R.S., and Nussenzweig, V. 2001 Gene targeting in the rodent malaria parasite *Plasmodium yoelii*. Mol Biochem Parasitol 113**:** 271-278.

Myler, P.J., Audleman, L., deVos, T., Hixson, G., Kiser, P., Lemley, C., Magness, C., Rickel, E., Sisk, E., Sunkin, S., Swartzell, S., Westlake, T., Bastien, P., Fu, G., Ivens, A., and Stuart, K. 1999 *Leishmania major* Friedlin chromosome 1 has an unusual distribution of protein-coding genes. Proc. Natl. Acad. Sci .USA. 96**:** 2902-2906.

Myler, P.J., Beverley, S.M., Cruz, A.K., Dobson, D.E., Ivens, A.C., McDonagh, P.D., Madhubala, R., Martinez-Calvillo, S., Ruiz, J.C., Saxena, A., Sisk, E., Sunkin, S.M., Worthey, E., Yan, S., and Stuart, K.D. 2001 The *Leishmania* genome project: new insights into gene organization and function. Med Microbiol Immunol (Berl) 190**:** 9-12.

O'Donnell, R.A., Freitas-Junior, L.H., Preiser, P.R., Williamson, D.H., Duraisingh, M., McElwain, T.F., Scherf, A., Cowman, A.F., and Crabb, B.S. 2002 A genetic screen for improved plasmid segregation reveals a role for Rep20 in the interaction of *Plasmodium falciparum* chromosomes. Embo J 21**:** 1231-1239.

Oliver, S.G. 2002 Functional genomics: lessons from yeast. Philos Trans R Soc Lond B Biol Sci 357**:** 17-23.

Oshiro, G., Wodicka, L.M., Washburn, M.P., Yates, J.R., Lockhart, D.J., and Winzeler, E.A. 2002 Parallel identification of new genes in *Saccharomyces cerevisiae*. Genome Res 12**:** 1210-1220.

Ravel, C., and al., e. 1998 The complete chromosomal organization of the reference strain of the *Leishmania* genome project, *L. major* 'Friedlin'. Parasitol Today 14**:** 301-303.

Richie, T.L., and Saul, A. 2002 Progress and challenges for malaria vaccines. Nature 415**:** 694-701.

Robinson, N.P., Burman, N., Melville, S.E., and Barry, J.D. 1999 Predominance of duplicative VSG gene conversion in antigenic variation in African trypanosomes. Molecular and Cellular Biology 19**:** 5839-5846.

Rodriques Coura, J., and de Castro, S.L. 2002 A critical review on Chagas disease chemotherapy. Mem Inst Oswaldo Cruz 97**:** 3-24.

Roos, D.S., Crawford, M.J., Donald, R.G., Fraunholz, M., Harb, O.S., He, C.Y., Kissinger, J.C., Shaw, M.K., and Striepen, B. 2002 Mining the *Plasmodium* genome database to define organellar function: what does the apicoplast do? Philos Trans R Soc Lond B Biol Sci 357**:** 35-46.

Roy, A., Cox, R.A., Williamson, D.H., and Wilson, R.J. 1999 Protein synthesis in the plastid of *Plasmodium falciparum*. Protist 150**:** 183-188.

Salmon, D., Geuskens, M., Hanocq, F., HanocqQuertier, J., Nolan, D., Ruben, L., and Pays, E. 1994 A novel heterodimeric transferrin receptor encoded by a pair of VSG expression site-associated genes in *T. brucei*. Cell 78**:** 75-86.

Scherf, A., Figueiredo, L.M., and Freitas-Junior, L.H. 2001 *Plasmodium* telomeres: a pathogen's perspective. Curr Opin Microbiol 4**:** 409-414.

Schnaufer, A., Panigrahi, A.K., Panicucci, B., Igo, R.P., Jr., Wirtz, E., Salavati, R., and Stuart, K. 2001 An RNA ligase essential for RNA editing and survival of the bloodstream form of *Trypanosoma brucei*. Science 291**:** 2159-2162.

Schweizer, H.P. 2001 Triclosan: a widely used biocide and its link to antibiotics. FEMS Microbiol Lett 202**:** 1-7.

Smith, T.G., Walliker, D., and Ranford-Cartwright, L.C. 2002 Sexual differentiation and sex determination in the Apicomplexa. Trends Parasitol 18**:** 315-323.

Späth, G.F., Epstein, L., Leader, B., Singer, S.M., Avila, H.A., Turco, S.J., and Beverley, S.M. 2000 Lipophosphoglycan is a virulence factor distinct from related glycoconjugates in the protozoan parasite *Leishmania major*. Proc. Natl. Acad. Sci. USA. 97**:** 9258-9263.

Spielman, A., and D'Antonio, M. 2001 Mosquito: Faber and Faber. Stuart, K., Panigrahi, A.K., Schnaufer, A., Drozdz, M., Clayton, C., and Salavati, R. 2002 Composition of the editing complex of *Trypanosoma brucei*. Philos Trans R Soc Lond B Biol Sci 357**:** 71-79.

Suguna, K., Surolia, A., and Surolia, N. 2001 Structural basis for triclosan and NAD binding to enoyl-ACP reductase of *Plasmodium falciparum*. Biochem Biophys Res Commun 283**:** 224-228.

Sundar, S., Jha, T.K., Thakur, C.P., Engel, J., Sindermann, H., Fischer, C., Junge, K., Bryceson, A., and Berman, J. 2002 Oral miltefosine for Indian visceral leishmaniasis. N Engl J Med 347**:** 1739-1746.

Tan, L., Nielsen, N.H., Young, D.C., and Trizna, Z. 2002 Use of antimicrobial agents in consumer products. Arch Dermatol 138**:** 1082-1086.

Tarleton, R.L., and Kissinger, J. 2001 Parasite genomics: current status and future prospects. Curr Opin Immunol 13**:** 395-402.

Taylor, D. 2003 Fewer new drugs from the pharmaceutical industry. British Medical Journal 326**:** 408-409.

Thompson, J., Janse, C.J., and Waters, A.P. 2001 Comparative genomics in *Plasmodium*: a tool for the identification of genes and functional analysis. Mol Biochem Parasitol 118**:** 147-154.

Turco, S.J., Spath, G.F., and Beverley, S.M. 2001 Is lipophosphoglycan a virulence factor? A surprising diversity between *Leishmania* species. Trends Parasitol 17**:** 223-226.

Van der Ploeg, L.H.T., Cornelissen, A.W.C.A., Barry, J.D., and Borst, P. 1984 Chromosomes of Kinetoplastida. EMBO Journal 3**:** 3109-3115.

Vanhamme, L., Pays, E., McCulloch, R., and Barry, J.D. 2001 An update on antigenic variation in African trypanosomes. Trends Parasitol 17**:** 338-343.

Varmus, H. 2002 Building a global culture of science. Lancet 360**:** S1-4.

Vickerman, K. 1985 Developmental cycles and biology of pathogenic trypanosomes. British Medical Bulletin 41**:** 105-114.

Vickerman, K. 1989 Trypanosome sociology and antigenic variation. Parasitology 99**:** S37-S47.

Volkman, S.K., Hartl, D.L., Wirth, D.F., Nielsen, K.M., Choi, M., Batalov, S., Zhou, Y., Plouffe, D., Le Roch, K.G., Abagyan, R., and Winzeler, E.A. 2002 Excess polymorphisms in genes for membrane proteins in *Plasmodium falciparum*. Science 298**:** 216-218.

Wahlgren, M., and Bejarano, M.T. 1999 A blueprint of 'bad air'. Nature 400**:** 506-507.

Waller, R.F., Keeling, P.J., Donald, R.G., Striepen, B., Handman, E., Lang-Unnasch, N., Cowman, A.F., Besra, G.S., Roos, D.S., and McFadden, G.I. 1998 Nuclearencoded proteins target to the plastid in *Toxoplasma gondii* and *Plasmodium falciparum*. Proc. Natl. Acad. Sci .USA. 95**:** 12352-12357.

Waller, R.F., Ralph, S.A., Reed, M.B., Su, V., Douglas, J.D.,

Minnikin, D.E., Cowman, A.F., Besra, G.S., and McFadden, G.I. 2003 A type II pathway for fatty acid biosynthesis presents drug targets in *Plasmodium falciparum*. Antimicrob Agents Chemother 47**:** 297-301.

Waters, A.P. 2002 Orthology between the genomes of *Plasmodium falciparum* and rodent malaria parasites: possible practical applications. Philos Trans R Soc Lond B Biol Sci 357**:** 55-63.

Wellems, T.E. 2002 *Plasmodium* chloroquine resistance and the search for a replacement antimalarial drug. Science 298**:** 124-126.

Wellems, T.E., and Plowe, C.V. 2001 Chloroquine-resistant malaria. J Infect Dis 184**:** 770-776.

Wilson, R.J., and Williamson, D.H. 1997 Extrachromosomal DNA in the Apicomplexa. Microbiol Mol Biol Rev 61**:** 1-16.

Wincker, P., Ravel, C., Blaineau, C., Pages, M., Jauffret, Y., Dedet, J.P., and Bastien, P. 1996 The *Leishmania* genome comprises 36 chromosomes conserved across widely divergent human pathogenic species. Nucleic Acids Res 24**:** 1688-1694.

Wirth, D.F. 2002 Biological revelations. Nature 419**:** 495-496.

Wongsrichanalai, C., Pickard, A.L., Wernsdorfer, W.H., and Meshnick, S.R. 2002 Epidemiology of drug-resistant malaria. Lancet Infect Dis 2**:** 209-218.

Wootton, J.C., Feng, X., Ferdig, M.T., Cooper, R.A., Mu, J., Baruch, D.I., Magill, A.J., and Su, X.Z. 2002 Genetic diversity and chloroquine selective sweeps in *Plasmodium falciparum*. Nature 418**:** 320-323.

Zomerdijk, J., Ouellette, M., Ten Asbroek, A., Kieft, R., Bommer, A.M.M., Clayton, C.E., and Borst, P. 1990 The promoter for a variant surface glycoprotein gene expression site in *Trypanosoma brucei*. EMBO Journal 9**:** 2971-2801.

Zuegge, J., Ralph, S., Schmuker, M., McFadden, G.I., and Schneider, G. 2001 Deciphering apicoplast targeting signals—feature extraction from nuclear-encoded precursors of *Plasmodium falciparum* apicoplast proteins. Gene 280**:** 19-26.

Further Reading

- **MALDI-TOF Mass Spectrometry in Microbiology**
Edited by: M Kostrzewa, S Schubert (2016)
www.caister.com/malditof

- *Aspergillus* and *Penicillium* in the Post-genomic Era
Edited by: RP Vries, IB Gelber, MR Andersen (2016)
www.caister.com/aspergillus2

- **The Bacteriocins: Current Knowledge and Future Prospects**
Edited by: RL Dorit, SM Roy, MA Riley (2016)
www.caister.com/bacteriocins

- **Omics in Plant Disease Resistance**
Edited by: V Bhadauria (2016)
www.caister.com/opdr

- **Acidophiles: Life in Extremely Acidic Environments**
Edited by: R Quatrini, DB Johnson (2016)
www.caister.com/acidophiles

- **Climate Change and Microbial Ecology: Current Research and Future Trends**
Edited by: J Marxsen (2016)
www.caister.com/climate

- **Biofilms in Bioremediation: Current Research and Emerging Technologies**
Edited by: G Lear (2016)
www.caister.com/biorem

- **Microalgae: Current Research and Applications**
Edited by: MN Tsaloglou (2016)
www.caister.com/microalgae

- **Gas Plasma Sterilization in Microbiology: Theory, Applications, Pitfalls and New Perspectives**
Edited by: H Shintani, A Sakudo (2016)
www.caister.com/gasplasma

- **Virus Evolution: Current Research and Future Directions**
Edited by: SC Weaver, M Denison, M Roossinck, et al. (2016)
www.caister.com/virusevol

- **Arboviruses: Molecular Biology, Evolution and Control**
Edited by: N Vasilakis, DJ Gubler (2016)
www.caister.com/arbo

- *Shigella*: **Molecular and Cellular Biology**
Edited by: WD Picking, WL Picking (2016)
www.caister.com/shigella

- **Aquatic Biofilms: Ecology, Water Quality and Wastewater Treatment**
Edited by: AM Romaní, H Guasch, MD Balaguer (2016)
www.caister.com/aquaticbiofilms

- **Alphaviruses: Current Biology**
Edited by: S Mahalingam, L Herrero, B Herring (2016)
www.caister.com/alpha

- **Thermophilic Microorganisms**
Edited by: F Li (2015)
www.caister.com/thermophile
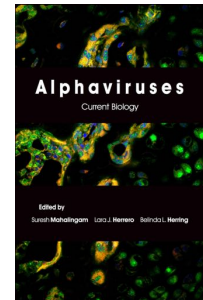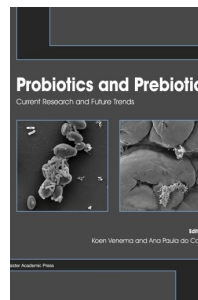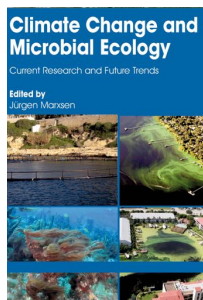
- **Flow Cytometry in Microbiology: Technology and Applications**
Edited by: MG Wilkinson (2015)
www.caister.com/flow

- **Probiotics and Prebiotics: Current Research and Future Trends**
Edited by: K Venema, AP Carmo (2015)
www.caister.com/probiotics

- **Epigenetics: Current Research and Emerging Trends**
Edited by: BP Chadwick (2015)
www.caister.com/epigenetics2015

- *Corynebacterium glutamicum*: **From Systems Biology to Biotechnological Applications**
Edited by: A Burkovski (2015)
www.caister.com/cory2

- **Advanced Vaccine Research Methods for the Decade of Vaccines**
Edited by: F Bagnoli, R Rappuoli (2015)
www.caister.com/vaccines

- **Antifungals: From Genomics to Resistance and the Development of Novel Agents**
Edited by: AT Coste, P Vandeputte (2015)
www.caister.com/antifungals

- **Bacteria-Plant Interactions: Advanced Research and Future Trends**
Edited by: J Murillo, BA Vinatzer, RW Jackson, et al. (2015)
www.caister.com/bacteria-plant

- *Aeromonas*
Edited by: J Graf (2015)
www.caister.com/aeromonas

- **Antibiotics: Current Innovations and Future Trends**
Edited by: S Sánchez, AL Demain (2015)
www.caister.com/antibiotics

- *Leishmania*: **Current Biology and Control**
Edited by: S Adak, R Datta (2015)
www.caister.com/leish2

- *Acanthamoeba*: **Biology and Pathogenesis (2nd edition)**
Author: NA Khan (2015)
www.caister.com/acanthamoeba2

- **Microarrays: Current Technology, Innovations and Applications**
Edited by: Z He (2014)
www.caister.com/microarrays2

- **Metagenomics of the Microbial Nitrogen Cycle: Theory, Methods and Applications**
Edited by: D Marco (2014)
www.caister.com/n2

Order from caister.com/order