# Functional Genomics in Trypanosomes: Using RNA Interference to Assign Gene Function and Validate Drug Targets in Tropical Disease Agents

By Dr Mark C Field, Department of Biological Sciences at Imperial College, London, and Dr Mark Carrington, Department of Biochemistry at the University of Cambridge, UK

**Dr Mark C Field is at the Department of Biological Sciences, Imperial College. He was educated at Oxford Univeristy and completed his doctoral work there at the Glycobiology Institute on glyoprotein structure and function. His postdoctoral career began in the US where he started to work on trypanosomatids at the Rockefeller and Stanford Universities, as well as a short period at Genentech Inc. In 1994, he established his own research group at Imperial College, London, which focuses on the molecular cell biology of *Trypanosoma brucei*.**

**Dr Mark Carrington is at the Department of Biochemistry, University of Cambridge. He was educated at Cambridge, obtaining his PhD in Plant Molecular Biology. He has worked on trypanosomes since the start of his postdoctoral career and has extensive experience in the study of protein structure and RNA processing and is interested in the development of therapeutics based on surface proteins and on host-resistance mechanisms.**

The World Health Organization has designated many tropical protozoan pathogens as high priority re-emergent disease agents. Asymmetric funding – many of these pathogens affect citizens of countries with low or no economic wealth – has meant that the development of new therapeutics is slow or non-existent. For researchers interested in *Trypanosoma brucei*, the causative agent of human African sleeping sickness, a combination of renewed interest, completion of the genome project and RNA interference technology may herald a new period of advance in both therapeutic development and understanding of basic biology. Can the research community deliver?

## THE DISEASE: THE NEED FOR NEW APPROACHES AND NEW DRUGS

Trypanosomes are unicellular protozoa and members of the order kinetoplastida, a highly divergent branch of the eukaryota which probably separated from the mammalian lineage ~3 x $10^9$ years ago. Pathogenic trypanosomatids include the causative agents for Chagas' disease or South American trypanosomiasis and Kala Azar or Leishmaniasis. They are equally divergent from the Apicomplexan protozoa that cause Malaria, Toxoplasmosis and Tularemia.

Human African sleeping sickness, or trypanosomiasis, is caused by infection with *Trypanosoma brucei rhodesiense* or *T. b. gambiensie*. A related wasting disease of bovines, Nagana, is caused by the nonhuman infective *T. brucei brucei*, *T. congolense* and *T. vivax*. Domestic animals act as a reservoir for *T. b. rhodesiense* and possibly also *T. b. gambiense*. In the case of *T. b. rhodesiense*, the ability to infect humans is due to the presence of a single gene for the serum resistance-associated protein, and transfer of this gene to *T. b. brucei* confers human infectivity (1). Transmission is caused by the biting Tsetse fly of the *Glossina* genus (see Figure 1). Over 30 countries in sub-Saharan Africa are affected, including Uganda, and the Democratic Republic of Congo, with 11 countries classified as having the disease at 'endemic' status and with 60 million people at risk (2). The combined impact on human health, as well as on agriculture due to both the real and perceived inability to raise cattle for food and use horses as draught animals in endemic areas has led to huge economic hardship and has in all likelihood depressed the rate of development of many African states.

Fear of the disease alone has caused abandonment of much fertile land.

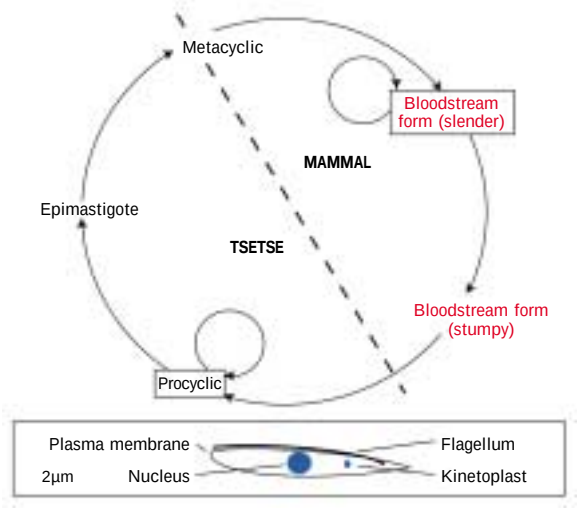
Current WHO/TDR numbers for fatalities are 50,000 per annum with >1.5 million DALYs, but these numbers are likely to be a major underestimate.

*T. brucei* is a remarkable organism in that, unlike many pathogenic protozoa, the parasite persists as a free living cell within the circulatory system and tissue spaces of its host. In this location, it is under constant challenge by the humoral arm of the host immune system and the parasite has evolved highly efficient mechanisms for immune evasion, mainly based on antigenic variation of the surface coat. This process makes vaccine development very unlikely and, given the poor record of development of vaccines against protozoan parasites in general, this route has essentially been abandoned. Following a period in the circulatory system, the parasite eventually enters and proliferates within the central nervous system (CNS) where disrupted sleep patterns, behavioural changes, coma and eventual death follow.

Without therapeutic intervention, African trypanosomiasis is 100 per cent fatal, and current drug options are limited: Suramin, the original treatment for trypanosomiasis, was introduced by Bayer over 70 years ago and remains the drug of choice for early stage disease, which is defined as being before the infection enters the CNS. Unfortunately, due to poor public health and communications in many endemic areas, patients frequently fail to present at the clinic sufficiently early and drugs for later stage disease must then be used. These are invariably toxic and expensive with numerous side effects, and resistance is emerging to all late stage drugs; their utility is unlikely to last more than a decade. Currently, DFMO (or eflornithine) and melarsoprol are the drugs of choice, although only the latter is of real value against *T. b. rhodesiense*. Melarsoprol is an arsenical-based drug with horrendous side effects, including rapid fatality following administration in approximately five per cent of patients, this is most likely due to acute encephalitic shock due to lysis of parasites in the CNS. DFMO, the 'resurrection drug', was introduced in the early 1990s but was discontinued in 1995. The drug has recently been reintroduced following a five year hiatus due to the inability of agencies to secure a manufacturer; a fortuitous discovery of a new use for DFMO as a cosmetic depilatory by Bristol Myers Squibb for facial hair, mainly of marketability to women in the developed world, means that the compound is now available again (3).

**Figure 1: The Life Cycle of the African Trypanosome**

Human infective stages are in red, and stages that are experimentally accessible are boxed. The life cycle is digenetic, with the ultimate host being man or a number of bovines and other mammals. The mammalian forms are differentiated by morphology into the rapidly dividing and self-renewing population (slenders) and stumpy forms, which have been proposed to be terminally differentiated for transmission to the insect vector. In the absence of authenticated markers the true nature of the stumpy form remains to be fully elucidated. In the insect vector (tsetse fly) the bloodstream forms rapidly differentiate into a self-renewing procyclic stage. Following a period in the insect gut these forms migrate into the haemoceole and differentiate into epimastigotes. A further differentiation accompanies migration to the tsetse mouthparts to produce the metacyclic form, which is adapted to initiate infection in the mammalian stage. Many details of the life cycle remain controversial. Box: the overall morphology of a trypanosome cell; the spindle-shaped cell has a single nucleus, a flagellum that is attached to the plasma membrane, and also a structure, the kinetoplast. Originally thought to be part of the locomotory apparatus, this organelle is in fact the mitochondrial genome.

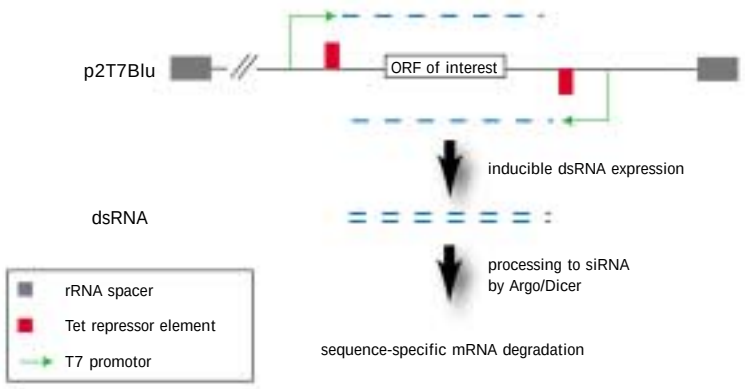



Asymmetric investment in drug development, which means a heavy focus on certain conditions primarily of interest to the developed world that are also frequently of low morbidity, is a characteristic of the current pharmaceutical scene. In a survey conducted by MSF/DND in 2001, it was observed that 1,393 new drugs entered the clinic in the 25 year period from 1975 to 1999, but of these only 13 were cited as having utility against

**Without therapeutic intervention, African trypanosomiasis is 100 per cent fatal, and current drug options are limited: Suramin, the original treatment for trypanosomiasis, was introduced by Bayer over 70 years ago and remains the drug of choice for early stage disease, which is defined as being before the infection enters the CNS. Unfortunately, due to poor public health and communications in many endemic areas, patients frequently fail to present at the clinic sufficiently early and drugs for later stage disease must then be used. These are invariably toxic and expensive with numerous side effects, and resistance is emerging to all late stage drugs.**

**Figure 2: RNA Interference Technology in Trypanosomes**

The basic vector, p2T7Blu, is a third generation version of a system originally developed by John Donelson and Doug LaCount at the University of Iowa. The present version, developed by David Horn from the London School of Hygiene and Tropical Medicine, for the TrypanoFAN project, includes a number of improvements for high throughput cloning. A sequence corresponding to the ORF of interest is inserted into the vector, usually by PCR from genomic DNA. The recombinant plasmid is then electroporated into the trypanosome and stable genomic integration achieved by recombination between the rRNA spacer elements incorporated into p2T7Blu and the rRNA locus (gray box). RNAi is controlled by the presence of the Tetracycline repressor protein and its DNA-binding site either side of the sequence to be targeted (red box). RNAi is achieved by addition of Tetracycline, which displaces the repressor protein, facilitating head to head transcription by T7 bacterophage DNA polymerase from dual promotors (green). The basic T7 expression system used for the project was developed by George Cross and Elisabeth Wirtz at the Rockefeller University. dsRNA is produced and processed by mechanisms presumed to be analogous to those present in higher eukaryotes – the precise details of the RNAi system in trypanosomes have not been fully elucidated. The // indicates omission of several sequence features, including those required for selection in trypanosomes and for propagation of the plasmid in *E. coli*.




tropical diseases (4). Even more telling is that in 2000 a total of 137 drug agents were in the development pipeline, but only one was cited as being of potential utility against African trypanosomes or malaria. By contrast, seven were cited for obesity and eight for erectile dysfunction in the same period. One of the authors has listened to one executive explaining to an African audience that whereas his company was investing heavily in the development of antiobesity treatments for companion animals (in other words. pets), it saw no future in investing in treatments for livestock diseases endemic to Africa. The dismal state of drug availability for African trypanosomiasis, a major disease, is unlikely to improve much in the next few years, although efforts by WHO and associated agencies do provide some hope.

## THE GENOME, RNA INTERFERENCE AND THE CHALLENGE!

Can new approaches provide some means to break an apparent deadlock in drug development? Much hope lies in the availability of genome data, informatics strategies and associated technologies in this area, as anticipated in so many others. The era of post-genomics has now arrived for *T. brucei*; the genome is complete and due for formal publication in late 2004 (5). The genome contains approximately 9,000 open reading frames (ORFs), distributed across 11 chromosomes. The organisation of the genes in a chromosome is unusual in trypanosomatids and most genes are transcribed by RNA polymerase II in a polycistronic fashion, with the nascent RNA resolved by means of a *trans*-splicing reaction. Conventional *cis*-splicing is rare.

Early genome annotation highlights two major challenges for the field. Firstly, approximately 50 per cent of ORFs have no clear homology beyond occasional basic structural motifs to ORFs outside of other kinetoplastida genomes, suggesting that a large proportion of the biology of these organisms remains uncharted. Secondly, for those genes where some homology does exist, accurate annotation of function is frequently not possible because of weak sequence similarity. This level of divergence at the genome level was predicted by multiple examples of divergent biochemistry and evolutionary distance and has the consequence that potential investigation of gene function by homology to more tractable systems is, for many ORFs, a non-starter in *T. brucei*.

By contrast, the RNA interference techniques that have emerged in the last few years often feel like a piece of science fiction – the ability to precisely, rapidly and cheaply ablate expression of particular genes has become a reality. This technology is based on the introduction of a double-stranded RNA complementary to the mRNA of the gene of interest. The precise mechanistic details differ between species and have yet to be fully elucidated, but in essence the dsRNA is cleaved by endonucleases to 21 to 23 base pair fragments that bind to the cognate mRNA and initiate further endonuclease cleavage, this time of the mRNA. The process is rapid, efficient and specific and is most likely an ancient mechanism for genome defence from retroviral elements (6). Trypanosomes were one of the first systems in which RNAi was described (7) and since the earliest reports, highly reliable methods have been developed for efficient RNAi based ablation of expression of any target gene.

The challenge that faces the trypanosome research community then is clear. There is a divergent genome with a vast potential for drug development and possibly hundreds of ORFs are available with sufficient divergence to present a viable target. There are now the means to identify many of these. How can this promise be translated into therapeutics in the face of poor funding?

## FUNCTIONAL GENOMICS – WHAT CAN REALISTICALLY BE ACHIEVED WITH RNAI

In trypanosomes, RNAi is currently performed by creation of a stably transfected cell line using a vector that contains head to head promoters between which is placed a DNA fragment homologous to the gene of interest, usually produced by PCR (8,9, see Figure 2). This construct is introduced into the trypanosome genome via electroporation, and transfected cells are selected by drug treatment. This rather laborious process is required due to very low transfection efficiency for trypanosomatids, which essentially negates any possibility of

**The ability to ascertain if a gene product is essential is a vital aspect for drug target discovery as well as basic research. The potential to target a large cohort of genes that are predicted as having a specific function in, for example mitosis, could provide a rapid mechanism by which those genes that are of potential interest as targets may be rapidly identified and taken forward for more detailed analysis, in other words incorporation into drug screens, biochemical analysis and structural studies.**

using the technology as a reliable forward screen for a biological process or biochemical activity.

### What can this Technology be Used For?

RNAi has been successfully used to investigate a wide range of biological processes in trypanosomes from RNA metabolism and recombination to cytoskeletal function and intracellular transport. All of these studies have provided new insights into the workings of the parasite. The major advantage of this method is that the RNAi is inducible and consequently a gene may be suppressed at will. This is vital as a great many of the gene products that have been selected by workers in the field turn out to be essential and result in rapid cell death. The traditional approach of gene knockout would have yielded no information as null mutants for these genes are, by definition, inviable.

The ability to ascertain if a gene product is essential is a vital aspect for drug target discovery as well as basic research. The potential to target a large cohort of genes that are predicted as having a specific function in, for example mitosis, could provide a rapid mechanism by which those genes that are of potential interest as targets may be rapidly identified and taken forward for more detailed analysis, in other words incorporation into drug screens, biochemical analysis and structural studies. Results from such studies are not available as yet, but the promise is very real and several projects are ongoing – again, the combination of RNAi and the complete genome sequence are prerequisites for such efforts.

In a second application, RNAi now allows a potential drug target that is suggested by biochemical or other data to be rapidly evaluated. For example, trypanothione reductase has

been suggested as a potent therapeutic target as this enzyme is unique to trypanosomatids. Analysis of this pathway by gene knockout was reported about two decades after initial suggestions of the enzyme as a candidate, together with data indicating the levels of the enzyme required for parasite viability (10). With RNAi, such information can now be acquired within months.

A third example of the application of RNAi in the trypanosome system is the potential for assignment of gene function based on systematic analysis of the genome. The large number of unknowns in the trypanosome gene annotation databases (5) suggests a rather daunting task in terms of assignment of function, one that is clearly well beyond the ability of the research community or the funding agencies to complete by traditional approaches in a realistic timeframe. The goal of a pilot project to undertake systematic targeted RNAi of all genes of chromosome 1 (which represents about four per cent of the genome) has been undertaken with support from the UK's Wellcome Trust and is a collaborative effort between several UK laboratories. The project has sought to devise improved methods for RNAi, to assess the utility of simple assays to reveal phenotypes, to build robust standard operating procedures for high throughput analysis of trypanosome RNAi mutants and also to use bioinformatics approaches to assess the ability to predict phenotype based on several criteria (including expression profile, predicted function based on sequence, position on the chromosome and similarity of phenotype to

**The Trouble with Trypanosomes**

Trypanosomes present several technical challenges to researchers interested in basic biology or development of therapeutics beyond asymmetric funding. As a counterbalance several major breakthroughs in the last few years, which exploit some fortuitous aspects of trypanosome biology, makes the immediate future for trypanosome research look better than ever. It remains to be seen if the optimism within the research community can be translated into new knowledge, and ultimately therapeutics within a reasonable timeframe.

**Challenges**

- Diploid and a large (26Mb) complex genome: difficult, slow and inelegant knockout system
- Complex media and growth requirements: expensive for large-scale studies
- Low transfection efficiency: difficult to use 'library' strategies for knockout or forward genetics. Must electroporate and select for transformants, requiring 2-4 weeks
- A lack of biologically relevant genetic screens
- Growth on solid media difficult: cloning is complex, expensive and labour intensive
- Genomic instability: some loss of selectable markers and RNAi machinery in minority of cases
- Non-conventional transcription: all molecular tools must be built *de novo* by the research community
- High degree of divergence from model organisms: much of the biology is potentially novel

**Advantages**

- Very reliable RNA interference technology is now available
- A completed genome project
- Robust homologous recombination: targeted insertions and replacements very reliable
- Mature set of molecular tools
- *In vitro* culture provides access to major life cycle stages, including the human and cattle infective forms

other RNAi mutants). The project makes both data and biological materials freely available via the web (11).

The project has improved methods for RNAi allowing larger scale projects to be attempted with confidence. A constant concern is with false negative results, in other words no observable phenotype when in fact RNAi of the gene of interest should produce an effect but for technical reasons the experiment fails; the methodology has been validated to the level that essentially any gene may now be targeted with the expectation that data are reliable. So far, most of the genes of chromosome 1 have been targeted with a high level of growth defects. Interestingly, this suggests that a large number of genes are non-redundant and required for robust replication in a non-selective environment, that is in *in vitro* culture. This cohort may represent a fundamental set of genes essential to viability, and as such represent a promising collection of candidate drug targets. Equally important, the identification of a large number of genes that have no apparent phenotype in culture allows rapid rejection of those genes as targets. Clearly, much more work is required, both to increase the size of the current dataset and also to follow through on a number of potential therapeutic targets that have been identified by this work.

Finally, in both of the most closely related systems, *Trypanosoma cruzi* (causative agent of Chagas' disease) and *Leishmania* (Kala Azar and Leishmaniasis), RNAi has not been successful, despite some fairly exhaustive efforts (12,13), suggesting a fundamental biological problem, rather than a technical issue. This, of course, throws the burden even more heavily onto *T. brucei*, as this organism must now serve as a model for these other parasites as well.

## CONCLUSION

Systematic analysis of trypanosome gene function by RNAi is a realistic possibility and significant inroads have been built already. New advances in the technology, together with completion of the genome, mean that comparatively large numbers of genes may be screened for potential therapeutic value and at least validated at a preliminary level. This, however, represents only a beginning, and design of candidate drugs, validation of specificity and subsequent production represent, continued hurdles that must be tackled if African trypanosomiasis is to be brought under control. ◆

*The authors can be contacted at mfield@ic.ac.uk and mc115@mole.bio.cam.ac.uk*

*Acknowledgements*

*Work on African trypanosomes in the authors' laboratories is provided by the Wellcome Trust and the Leverhulme Trust, and both authors are supported by HEFC. TrypanoFAN is supported by the Wellcome Trust. We also thank the people directly involved in the experimental aspects of trypanofan; Chandra Subramanium, Seth Redmond and Paul Veazey.*

**References**

1. **Vanhamme L, Paturiaux-Hanocq F, Poelvoorde P, Nolan DP, Lins L, Van Den Abbeele J, Pays A, Tebabi P, Van Xong H, Jacquet A, Moguilevsky N, Dieu M, Kane JP, De Baetselier P, Brasseur R and Pays E, Apolipoprotein L-I is the trypanosome lytic factor of human serum, *Nature* Mar 6;422(6927): pp83-7, 2003**

2. **http://www.who.int/tdr/diseases/tryp/diseaseinfo.htm**

3. **http://www.vaniqa.com**

4. ***Fatal Imbalance: The Crisis in Research and Development for Drugs for Neglected Diseases*, A report by the MSF Access to Essential Medicines Campaign and the Drugs for Neglected Diseases Working Group**

5. **http://www.genedb.org/genedb/tryp/index.jsp**

6. **Tijsterman M Ketting RF and Plasterk RH, The genetics of RNA silencing, *Annu Rev Genet* 36: pp489-519, Epub, 2002**

7. **Ngo H, Tschudi C, Gull K and Ullu E, Double-stranded RNA induces mRNA degradation in Trypanosoma brucei, *Proc Natl Acad Sci USA*, Dec 8;95(25): pp14,687-92, 1998**

8. **LaCount DJ, Bruse S, Hill KL and Donelson JE, Double-stranded RNA interference in Trypanosoma brucei using head-to-head promoters, *Mol Biochem Parasitol*. Nov;111(1): pp67-76, 2000**

9. **http://www.trypanofan.org, and Redmond S, Vadivelu J and Field MC, RNAit: an automated web-based tool for the selection of RNAi targets in *Trypanosoma brucei, Mol Biochem parasitol* Apr 25; 128 (1): pp115-8, 2003**

10. **Krieger S, Schwarz W, Ariyanayagam MR, Fairlamb AH, Krauth-Siegel RL and Clayton C, Trypanosomes lacking trypanothione reductase are avirulent and show increased sensitivity to oxidative stress, *Mol Microbiol*. Feb;35(3): pp542-52, 2000**

11. **www.trypanofan.org**

12. **DaRocha WD, Otsu K, Teixeira SM and Donelson JE, Tests of cytoplasmic RNA interference (RNAi) and construction of a tetracycline-inducible T7 promoter system in Trypanosoma cruzi, *Mol Biochem Parasito*. Feb;133(2): pp175-86, 2004**

13. **Robinson KA and Beverley SM, Improvements in transfection efficiency and tests of RNA interference (RNA1) approaches in the protozoan parasite Leishmania, *Mol Biochem Parasitol* May; 128 (2): pp217-28, 2003**