J. Protozool. Res., 8, 214-223(1998)

Contribution of Genome Analysis to Understanding of the Biology of, and Diseases Caused by, African Trypanosomes

P.A.O. MAJIWA[1], A. DJIKENG[1], J. E. DONELSON[2] and C. AGUFA[1]

[1]*International Livestock Research Institute (ILRI), P.O. Box 30709, Nairobi, Kenya; and* [2]*Department of Biochemistry, University of Iowa and Howard Hughes Medical Institute, Iowa City, Iowa 52242, USA.*

Received 20 August 1998

Key words: ESTs,comparative genomics, *Trypanosoma*

ABSTRACT

Molecular genetic studies on the African trypanosome have in the past focused on only a few genes or gene products; consequently, a significant number of genes of this organism remain unknown. We are sequencing random directional cDNAs clones of *Trypanosoma brucei rhodesiense* to generate Expressed Sequence Tags (ESTs). Analysis of 2128 clones sequenced so far showed significant similarities with proteins already described in diverse organisms including man, rodents, kinetoplastids, yeasts and plants. A number of the ESTs encode homologues of proteins involved in: signal reception and transduction, cell division, gene regulation, DNA repair and replication, general metabolism, and structural integrity. A large proportion, 562 individual ESTs (26%, representing 299 different transcripts), had significant homology with genes described in organisms other than the African trypanosomes; however, 14% of the ESTs were from genes already described in trypanosomes. We expect that a number of the genes will be present in the other species of the African trypanosomes. This prediction has been tested by hybridization of the random ESTs with total cDNAs of other trypanosome species including *T. evansi*. The data will be discussed in the context of their potential contribution to understanding the biology of these and other parasitic protozoa, and the diseases they cause.

INTRODUCTION

Much useful information can be gained about an organism from accurate and detailed knowledge of its complete gene complement and organization. Genome analyses contribute in different ways to the better understanding of the organisms so analysed. There are many projects in progress that aim at sequencing the complete genomes of various prokaryotic and eukaryotic organisms. Some information from these undertakings is already available for a number of organisms including some bacteria and fungi (Oliver 1996; Mewes et al. 1997). In the context of 'Parasitology', these efforts can be viewed as being directed at either the host or the parasite.

It is expected that within the early part of the next millennium complete sequence information will be available on many organisms. Already there is a wealth of sequence information on many genes archived in public domain databases (Benson et al. 1997). This information is proving to be extremely useful to the biomedical research community. In order to exploit this information, we are generating expressed sequence tags (ESTs) from the cDNA of the human infective *T. b. rhodesiense*. The strategy adopted in EST analysis is to determine the nucleotide sequence composition once from the 5' end of random cDNA clones in a representative library.

The EST approach to gene identification takes advantage of sequence comparisons

across taxa and species, which facilitates the rate at which novel genes are found in an organism under study. Because *T. evansi* is evolutionarily so close to *T. brucei*, it is instructive to consider how molecular genetic studies conducted in the past and those being carried out now, on *T. brucei* can contribute to a better understanding of *T. evansi*, and how the analysis of the genome *T. brucei*, currently in progress, will likely contribute to a better understanding of *T. evansi*.

RFLPs: Cloned sequences encoding different variant surface glycoproteins (VSGs) of *T. brucei* and some anonymous genomic DNA clones were used in Southern blot hybridization analysis of DNA from different populations of *T. brucei*, *T. b. rhodesiense* and *T. b. gambiense*. A stock of *T. evansi* was also included in the analysis. These studies showed that restriction enzyme fragment polymorphisms (RFLPs) analysis is a useful tool for the characterization and finer classification of the trypanosomes within the *Trypanozoon* (Paindavoine et al.1986). In particular, it was demonstrated through such studies that it is possible to distinguish the 'gambiense group' of *T. brucei* from the other sub-species within this subgenus.

Kinetoplast DNA: The analysis of kinetoplast DNA (kDNA), first carried out in *T. brucei* (Donelson et al. 1979; Chen and Donelson 1980) and later extended to *T. evansi* (Borst et al. 1987; Lun et al. 1992) indicated that *T. evansi* lacks the compact kDNA network. These studies showed also that *T. evansi* have no maxicircle DNA molecules (Borst et al. 1987; Ou et al. 1991), which are the equivalent of mitochondrial DNA. This provided an explanation in part for failure of these trypanosomes to undergo cyclical development in the tsetse vector. Whereas it was evident that the maxicircles were the equivalents of mitochondrial DNA, the function of the minicircles remained obscure for many years. It is now known that the minicircle molecules are transcribed (Pollard et al. 1990) into guide RNA (gRNA) which play an essential role in RNA editing (Thiemann et al. 1994), a process which occurs primarily in the procyclic forms of the parasite. Implicitly, RNA editing does not occur in *T. evansi*.

The minicircles of *T. evansi*, in contrast to those of *T. brucei*, are homogeneous in sequence, comprising two main types designated A and B. Thus, in any one clonal population of *T. evansi*, the minicircles are normally of one or another of the two sequence types (Borst et al. 1987). *Ttypanosoma evansi* with the type A minicircles are found among isolates of the parasite from East Africa, West Africa, and South America, whereas those with the type B minicircles have been found so far only among the East African isolates. Curiously, the type B minicircle appears to share the same degree of sequence homology with the type A minicircles, and some *T. brucei* minicircles. In *T. brucei*, there are at least three different sequence classes of minicircles whose proportions vary among different cells (Chen and Donelson 1980). Based on this information, it has been possible to design oligonucleotide primers for specific detection of *T. evansi* by the polymerase chain reaction (PCR) and to distinguish it from *T. brucei*. This observation provided one of the first indications that these two closely related parasites can be distinguished at the molecular level.

Repetitive elements: A search for repetitive elements in the genome of *T. brucei* revealed two molecules designated RIME (Hasan et al. 1984) and *ingi* (Kimmel et al. 1987), both of which occur at several loci, sometimes together, in the genome of the *Trypanozoon*. Although the functions of these elements remain unknown, they most likely contribute to parasite genome mobility. Because of its ubiquity among the *Trypanozoon*, the sequences of both elements form the basis for the PCR detection of trypanosomes within this subgenus.

An additional property of *ingi* that makes it attractive as a target for sensitive detection of the *Trypanozoon* is the fact that it is transcribed into RNA (Kimmel et al. 1987). Assays directed at detecting this RNA could provide a means of revealing the presence of live parasites in an infection. The other repetitive element which is the target for detection of the *Trypanozoon* is the 177bp repeat initially described by Sloof et al (1983).

Since both of these elements are found in the genome of *T. evansi*, they can been exploited as targets for the DNA-based systems for detection of this parasite. Indeed in our experience, we found no difference in the specificity and sensitivity with which *T. evansi* or *T. brucei* can be detected by PCR when either *ingi* or the 177bp repeat is the target of the amplification. This observation supports the notion that these repetitive elements are represented at approximately the same copy numbers in the genomes of both parasites.

Chromosome polymorphisms: Although trypanosomes do not condense their chromosomes at any stage sufficiently to be visualised cytologically, chromosome-sized DNA molecules can be prepared from these organisms for different analyses. One such analysis is the separation of the chromosomes by pulsed field gel electrophoresis (Schwartz and Canto 1984) to give molecular karyotypes. Among the *Trypanozoon*, molecular karyotypes appear to be quite variable, in both size and number. In general, however, the overall pattern of the chromosomes appear to be similar for members of a single antigenic repertoire. Therefore, the chromosome pattern or molecular karyotype of trypanosomes can be used as an index of the identity of populations of the parasite in a particular locality. This approach has been used by Waitumbi and Young (1994) to characterise *T. evansi* isolates from camel herds in Kenya. In this way, it was possible to distinguish different molecular karyotypic groups among the parasite isolates in a manner that could not have been possible by the analysis of kDNA minicircles. Thus, although the molecular karyotypes do not correlate with kDNA minicircle patterns, the information obtained from the analysis can be useful in epidemiological surveys. The separation of trypanosome chromosomes was previously performed under conditions which could separate only the medium-sized chromosomes, less than 1Mb. Presently, all the 11 chromosomes of *T. brucei* can be resolved by PFGE.

Although *T. evansi* cannot be distinguished from other species or subspecies within the *Trypanozoon*, strictly on the basis of molecular karyotypes, the molecular karyotype can be very useful in determining that the trypanosome is indeed a member of this subgenus. The relative stability of molecular karyotypes enabled Waitumbi et al (1994) to show that drug resistant *T. evansi* can persist in the same locale for extended periods of time.

In order to maximize the benefits that can be derived from the analyses of gene organization in chromosomes, it has become necessary to develop chromosome-specific markers, such that individual chromosomes can be identified irrespective of size fluctuations. The markers that have proven to be most useful for this are genes discovered in the course of studies on the basic biology of *T. brucei*. The number of markers that can be used for tagging the chromosomes has been increased many times fold by the availability of *T. brucei* expressed sequence tags (ESTs) (el-Sayed et al. 1995).

The ESTs are generated by single-pass sequencing of random cDNAs, directionally inserted into a plasmid. Normally, the sequencing is performed from the 5'end of the cloned insert since this is the one nearest to the coding region. Over the past three years or so, we have generated ESTs from *T. brucei* (Djikeng et al. 1998). It is evident from the EST analysis that a number of the novel genes found in this trypanosome will be very useful in understanding the biology of this trypanosome and other parasites that cause diseases in

man and livestock.

Comparative genomics and hybridization in silico: Because of functional constraints, genes that perform similar functions in different organisms tend to be structurally related. Any differences that may be found when such genes are compared are often more pronounced at the nucleotide sequence level than at the amino acid sequence level. When a short sequence derived from an EST is compared with the other sequences in a database such as GenBank, hybridization *in silico* occurs with all the sequences that have been deposited in the databases.

The nucleotide sequences obtained in EST analyses can at best be only 98% accurate. However, the compensatory algorithms of the computer programs such as BLAST (Altschul et al. 1990), used in the comparisons, require only a short stretch of nucleotide sequence to identify a gene contained in a piece of DNA. Thus, by allowing inter-genome sequence comparisons across taxonomic boundaries, EST analysis is proving to be an effective alternative to the traditional approach to gene identification, in which only one gene is cloned at a time. The majority of genes discovered through the analysis of *T. brucei* EST are likely to be present also in *T. evansi*.

EST analysis of T. brucei: The cDNA library we used was constructed in lZAP II, from mRNA of *T. b. rhodesiense* WRaTat1.24 (Alarcon et al. 1994; el-Sayed et al. 1995). Plasmids from this library were rescued *en masse* by co-infection of bacterial cells with helper phage, which were subsequently grown on plates supplemented with ampicillin. The sequencing reactions (Sanger et al. 1977) were performed on individual plasmid cDNA clones, to obtain approximately 400 nt of reliable sequence data. The sequence data were edited and then used as queries to search the GenBank using the BLAST program (Altschul et al. 1990). An EST was considered to have a significant match if P(n) <1x10-5. In assigning putative functional roles to ESTs which had significant homologues in the databases, we adopted the general broad categories of cellular functional roles described by Adams et al. (1995).

RESULTS

A total of 2128 independent randomly selected directional cDNA clones were analysed by sequence similarity searches against the GenBank databases, most recently in March 1998, using the BLAST program (Altschul et al. 1990). As a check regarding the quality of the cDNA library, we determined the frequency with which sequences encoding a few known trypanosomal genes were encountered among the ESTs. This indicated that the most redundant homologues are the ribosomal protein ESTs (7.9%), followed by a and b tubulins (3%). Although the trypanosome from which the library was constructed expressed predominantly mVAT4, different VSGs were also found among the ESTs. In total 30 VSGs were found, nearly half of which were mVAT4. The others were homologues of VSGs in different VAT repertoires including MITaR1, ILTaR1, BoTaR and AnTaR.

Individual ESTs were placed into putative functional groups according to the known function of the most significant homologue of each one. In this preliminary classification of trypanosome ESTs, each gene or gene product has been assigned to only one functional group, although one gene or gene product can have more than one role. The proportion of ESTs in each of the six different functional groups is summarised in the pie diagram shown in Figure 1.

Potentially interesting genes among ESTs: There were within each functional category, ESTs which are significantly homologous to genes or gene products which can be

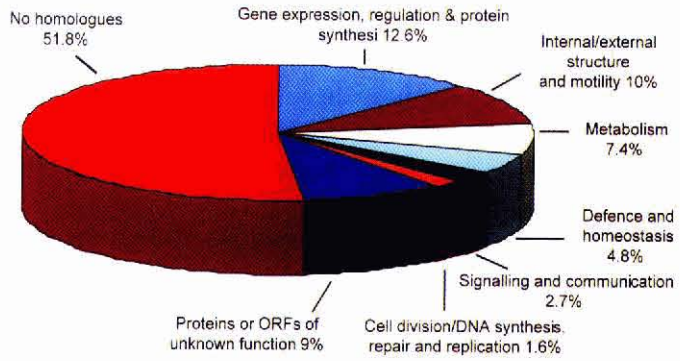

Figure 1. Pie diagram showing the proportion of ESTs assigned to each functional category, based upon the best homologue in the public domain database.

Table 1. Partial list of potentially interesting *T. b. rhodesiense* ESTs

Functional category	**EST**	**Homologue**	**P(n)**	**Description of function**
Signalling & communication	T1260	Extracellular signal-regulated kinase	$6.1x10^{-26}$	ACTIVATED BY TYROSINE AND THREONINE PHOSPHORYLATION; BELONGS TO THE DC2/CDC28 SUBFAMILY OF SER/THR PROTEIN KINASES
	T2248	LDL receptor protein	$8.5x10^{-10}$	LIPOPROTEIN RECEPTOR
	T2290	Protein kinase C delta type	$3.2x10^{-19}$	CAN DESENSITIZE MATING PHEROMONE SIGNAL TRANSDUCTION IN YEAST
Cell Division & DNA Synthesis	T1864	Cell division control protein 40	$1.0x10^{-13}$	INVOLVED IN REPLICATION AND MITOTIC SPINDLE FORMATION AND/OR MAINTENANCE; BELONGS TO THE BETA TRANSDUCIN FAMILY
	T4019	RAD3	$9.5x10^{-16}$	DNA HELICASE COMPONENT OF RNA POLYMERASE TRANSCRIPTION INITIATION FACTOR TFIIH (FACTOR B) AND THE NUCLEOTIDE EXCISION REPAIROSOME; COMPONENT OF NUCLEOTIDE EXCISION REPAIR FACTOR 3 (NEF3) COMPLEX.
Gene Expression & Protein Synthesis	T1195	*ede*B protein	$2.4x10^{-38}$	ONE OF THE GENES UP-REGULATED IN DEDIFFERENTIATING NICOTANIA GLAUCA (TOBACCO) PITH TISSUE
	T2097	Ubiquitin-conjugating enzyme E2	$3.0x10^{-36}$	CATALYSES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER PROTEINS. MEDIATE THE SELECTIVE DEGRADATION OF SHORT-LIVED AND ABNORMAL PROTEINS
Defence & homeostasis	T2159	STI (stress inducible protein)	$3.4x10^{-36}$	A STRESS-INDUCIBLE GENE FROM SOYBEAN CODING FOR A PROTEIN BELONGING TO THE TPR (TETRATRICOPEPTIDE REPEATS) FAMILY
Metabolism	T3176	Proteasome	$1.3x10^{-15}$	A MULTICATALYTIC PROTEINASE COMPLEX, CHARACTERIZED BY ITS ABILITY TO CLEAVE PEPTIDES WITH ARG, PHE, TYR, LEU, AND GLU ADJACENT TO THE LEAVING GROUP AT NEUTRAL OR SLIGHTLY BASIC pH; IS INVOLVED IN AN ATP/UBIQUITIN-DEPENDENT NON-LYSOSOMAL PROTEOLYTIC PATHWAY
Structure & Motility	T2506	Centromere / microtubule binding protein	$7.1x10^{-7}$	BINDS TO CENTROMERES AND MICROTUBULES; IS INVOLVED IN CHROMOSOME SEGREGATION AND IS ESSENTIAL FOR CELL GROWTH
Undetermined	T3347	Vegetative specific protein	$6.2x10^{-16}$	THE PROTEIN IS EXPRESSED IN DEACTIVATED UPON THE H7GROWING CELLS AND INITIATION OF DEVELOPMENT.

considered interesting in the context of trypanosome biology. Some of these are listed in Table 1 and are noteworthy because they: (a) should have been found earlier, given that the processes in which they are thought to be involved have been investigated in the African trypanosomes; and (b) could provide new and unexpected information about the trypanosomes. A growing but incomplete listing of *T. brucei* ESTs can be viewed at http://parsun1.path.cam.ac.uk/est.htm.

We performed both northern and Southern blot hybridizations, using the ESTs as probes. These blot hybridizations confirmed that the ESTs were indeed of trypanosome origin, and could be detected in the trypanosome RNA or DNA under standard hybridizations and post-hybridization wash conditions. Additionally, many of the ESTs identified genes which are transcribed in a developmental stage-specific manner, being present in some but absent from other life-cycle stages of the trypanosome.

We sought to determine the number of random ESTs containing sequences common to both *T. evansi* and *T. brucei*. Inserts of fifty ESTs homologous to known genes were hybridized to radio-labelled total cDNA of either *T. brucei* or *T. evansi*, shown in Figure 2.

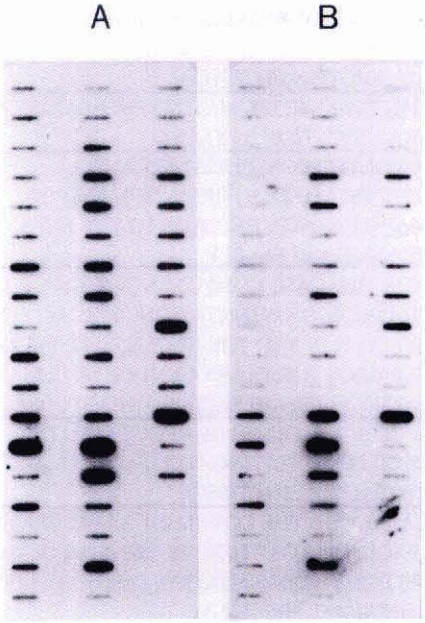

Figure 2. Inserts of fifty random ESTs of *T. brucei* were slot blotted and hybridized with radio-labelled cDNA from either *T. brucei* (panel A) or *T. evansi* (panel B). The filters were washed with 0.1xSSC, 0.1%SDS at 650C, then exposed to film for autoradiography for 6 hrs to obtain the autoradiogram shown.

Twenty of these hybridized equally strongly to the *T. evansi* cDNA, as they did to the homologous *T. brucei* cDNA. A majority of the remaining hybridized very weakly to the *T. evansi* cDNAs probe, indicating that they represent rare transcripts. Only a minority of the ESTs did not hybridize at all to the *T. evansi* cDNA probe. In the another experiment, we used oligonucleotide primers specific to molecules extensively characterized in *T. brucei* to determine if the same molecules can be amplified from *T. evansi* cDNA. The results we obtained are shown in Figure 3. The VAT primer, which forms the core of a sequence conserved in the 3' end of all *T.brucei* VSG-encoding sequences, produced a fragment of a similar size in the cDNA of both *T. evansi* and *T. brucei*. The product from the *T. brucei* cDNA was more abundant than that obtained from the *T. evansi* cDNA, but both were of similar size. This primer pair is specific for the *Trypanozoon* VSGs, and is expected to amplify all VSG transcripts within a pool of cDNAs. The VSGs so amplified can be cloned and sequenced in mass, allowing one to determine the entire repertoire of VSGs expressed at

the time the RNA was purified from the parasite.

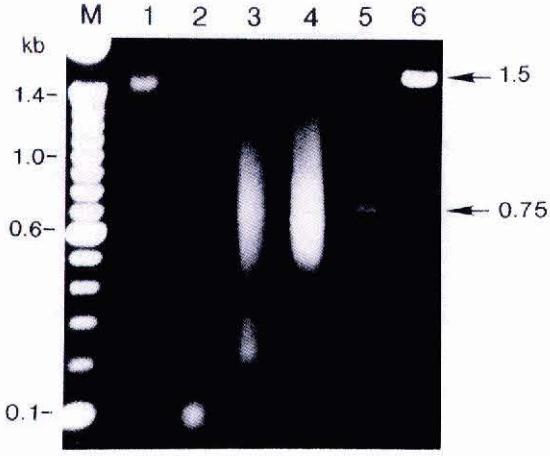

Figure 3. Mass amplification of sequences from either *T. evansi* or *T. brucei* total cDNA. Oligonucleotide primers specific for different genes sequences initially identified in *T. brucei* were used in the amplification of the sequences from the bloodstream forms of either *T. brucei* or, *T. evansi*. Lanes contain samples as follows: 1, *T. evansi* cDNA, amplified with VAT primer (TGA TAT ATT TTA ACA CCT) and the 3' segment of the miniexon(ACA GTT TCT GTA CTA TAT TG); 2, *T. brucei* cDNA amplified with VAT primer and 3' segment of the miniexon; 3, *T.brucei* cDNA amplified with Not 1 primer-adaptor (CAA TTC GCG GCC GCT TTT TTT TTT TTT TT) and the 3' segment of the miniexon; 4, *T. evansi* cDNA amplified with Not 1 primer adaptor and the 3' segment of the miniexon; 5, *T. evansi* cDNA amplified with primers specific for the *T. brucei* homologue of the receptor for activated protein kinase C, designated *Tb*RACK; 6, *T. brucei* cDNA amplified with the VAT primer and the 3' segment of the miniexon. In each case, the cDNAs used had been synthesized by priming the mRNA with Not1 primer adaptor during the reverse transcription to single-strand cDNA.

DISCUSSION

Trypanosoma evansi and *T. brucei* are evolutionarily very close to each other, both being in the same subgenus. Since *T. brucei* is certainly the best studied of the African trypanosomes, reagents derived from it should be of direct benefit to understanding of the biology of the other trypanosomes, specifically *T. evansi*. The examples cited in the foregoing paragraphs illustrate the ways in which such benefits can be realised.

One of the reasons for undertaking the EST analysis of *T. brucei* was to find new genes of this parasite and to make such genes generally available for other studies. New trypanosome genes fall in one of two broad categories: (a) the genes which have not been identified before in trypanosomes, but which have been studied in another organism; or (b) genes which have not been described in any other organism and, therefore, have no homologues among the sequences in the public domain databases. Among the ESTs we have generated so far are >1000 distinct new trypanosome genes. A number of these have already attracted the attention of investigators working either on trypanosomes or on other organisms.

Studies on the biology of *T. evansi* will benefit directly from these ESTs. We know, for example that the major antigen detected in the sera of animals infected by *T. evansi* is the hsp70, a protein encoded by one of the ESTs we have cloned.

We recognise that it would be impractical for a small group of investigators or a single institute to study all the genes described here. Therefore, we encourage investigators interested in specific ESTs to request for them in writing or by sending an e-mail to either P.M. (for investigators resident in Europe and the rest of the world) or J.D. (e-mail: john-donelson@uiowa.edu for investigators in the Americas). Alternatively, investigators can download the sequence of ESTs relevant to their work and use the information to synthesize oligonucleotides for cloning the corresponding gene by PCR or hybridization. The information deposited at GenBank can also be used to obtain the nearest homologue of a gene of interest; such a gene can then be used as a probe in heterologous screening of a recombinant library. The putative functional groups into which the ESTs fall represent the

major processes of the cell; from other experiments we determined that some of the gene transcripts are detectable at some life-cycle stages and undetectable in others. Such pieces of information indicate the possible involvement of the genes in processes that are modulated during the developmental cycle of the trypanosome.

The group of ESTs that holds even greater potential is represented by the 52% of genes which have no homologue in the public domain databases. As additional information becomes available from the whole genome sequencing of other organisms, homologues of some ESTs in this category will be found, thus decreasing with time the number trypanosome ESTs without homologues. That small number is likely to comprise the transcripts which encode parasite-specific proteins. The challenging phase of this investigation will be elucidation of the functions of these trypanosome-specific genes. Such an undertaking becomes feasible in the face of the rapid progress in the emerging field of functional genomics whose goal is the elucidation of the functions of the new genes (Hieter and Boguski 1997; McKusick 1997; Oliver 1996).

Functional genomics is normally conducted in a model organism such as the yeast *S. cerevisiae*, whose complete sequence information is now available (Mewes et al. 1997). Indeed, yeast is increasingly being employed in the elucidation of a number of biological phenomena which could not be easily approached in some organisms. These include (a) assays on drug resistance in *Plasmodium falciparum* (Wooden et al. 1997), (b) isolation of genes encoding secreted proteins (Jacobs et al. 1997; Klein et al. 1996), and (c) the identification of interacting protein partners (Gietz et al. 1997). With the trypanosome genes at hand, it should be possible to incorporate the parasite in this new field of functional genomics.

ACKNOWLEDGEMENTS

We thank Loren Donelson, Samwel Oyola, Lorraine Nyaoke and the staff of the ILRI sequencing unit for technical assistance. This investigation received financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR).

REFERENCES

Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D. & White, O. et al. 1995. Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence. *Nature 377: 3-174.*

Alarcon, C.M., Son, H.J., Hall, T. & Donelson, J.E. .1994. A monocistronic transcript for a trypanosome variant surface glycoprotein. *Mol. Cell. Biol. 14: 5579-5591.*

Altschul, S.F., Gish, W., Miller, W., Myers, E.W. & Lipman, D.J. 1990. Basic local alignment search tool. *J. Mol. Biol. 215: 403-410.*

Benson, D.A., Boguski, M.S., Lipman, D.J. & Ostell, J. 1997). GenBank. *Nucleic Acids Res. 25: 1-6.*

Borst, P., Fase-Fowler, F. & Gibson, W.C. 1987. Kinetoplast DNA of *Trypanosoma evansi. Mol. Biochem. Parasitol. 23: 31-38.*

Chen, K.K. & Donelson, J.E. 1980. Sequences of two kinetoplast DNA minicircles of *Trypanosoma brucei. Proc. Natl. Acad. Sci. U. S. A. 77: 2445-2449.*

Djikeng, A., Agufa, C., Donelson, J. E . & Majiwa, P.A.O. 1998. Generation of expressed sequence tags as physical landmarks in the genome of *Trypanosoma brucei. Gene* 221:

93-106.

Donelson, J.E., Majiwa, P.A. & Williams, R.O. 1979. Kinetoplast DNA minicircles of *Trypanosoma brucei* share regions of sequence homology. *Plasmid 2: 572-588.*

el-Sayed, N.M.A., Alarcon, C.M., Beck, J.C., Sheffield, V.C.& Donelson, J.E. 1995. cDNA expressed sequence tags of *Trypanosoma brucei rhodesiense* provide new insights into the biology of the parasite. *Mol. Bioch. Parasitol.* 73: 75-90.

Gietz, R.D., Triggs-Raine, B., Robbins, A., Graham, K.C. & Woods, R.A. 1997. Identification of proteins that interact with a protein of interest: applications of the yeast two-hybrid system. *Mol. Cell. Biochem. 172: 67-79.*

Hasan, G., Turner, M.J. & Cordingley, J.S. 1984. Complete nucleotide sequence of an unusual mobile element from *Trypanosoma brucei. Cell 37: 333-341.*

Hieter, P. & Boguski, M. 1997. Functional genomics: it's all how you read it. *Science 278: 601-602.*

Jacobs, K.A., Collins-Racie, L.A., Colbert, M., Duckett, M., Golden-Fleet, M., Kelleher, K., Kriz, R., LaVallie, E.R., Merberg, D., Spaulding, V., Stover, J., Williamson, M.J. &
McCoy, J.M. 1997. A genetic selection for isolating cDNAs encoding secreted proteins. *Gene 198: 289-296.*

Kimmel, B.E., ole-MoiYoi, O.K. & Young, J.R. 1987. *Ingi*, a 5.2-kb dispersed sequence element from *Trypanosoma brucei* that carries half of a smaller mobile element at either end and has homology with mammalian LINEs. *Mol. Cell. Biol. 7: 1465-1475.*

Klein, R.D., Gu, Q., Goddard, A. & Rosenthal, A. 1996. Selection for genes encoding secreted proteins and receptors. *Proc. Natl. Acad. Sci. U. S. A. 93: 7108-7113.*

Lun, Z.R., Brun, R. & Gibson, W. 1992. Kinetoplast DNA and molecular karyotypes of *Trypanosoma evansi* and *Trypanosoma equiperdum* from China. *Mol. Biochem. Parasitol. 50: 189-196.*

McKusick, V.A. 1997. Genomics: structural and functional studies of genomes. *Genomics 45: 244-249.*

Mewes, H.W., Albermann, K., Bahr, M., Frishman, D., Gleissner, A., Hani, J., Heumann, K., Kleine, K., Maierl, A., Oliver, S.G., Pfeiffer, F. & Zollner, A. 1997. Overview of the yeast genome. *Nature 387: 7-65.*

Oliver, S.G. 1996. From DNA sequence to biological function. *Nature* 379: 597-600.
Ou, Y.C., Giroud, C. and Baltz, T. 1991. Kinetoplast DNA analysis of four *Trypanosoma evansi* strains. *Mol. Biochem. Parasitol. 46: 97-102.*

Paindavoine, P., Pays, E., Laurent, M., Geltmeyer, Y., Le Ray, D., Mehlitz, D. & Steinert, M. 1986. The use of DNA hybridization and numerical taxonomy in determining relationships between *Trypanosoma brucei* stocks and subspecies. *Parasitology 92: 31-50.*

Pollard, V.W., Rohrer, S.P., Michelotti, E.F., Hancock, K. & Hajduk, S.L. 1990. Organization of minicircle genes for guide RNAs in *Trypanosoma brucei. Cell 63: 783-790.*

Sanger, F., Nicklen, S. & Coulson, A.R. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. U. S. A. 74: 5463-5467.*

Schwartz, D.C. & Cantor, C.R. 1984. Separation of yeast chromosome-sized DNAs by pulsed field gradient gel electrophoresis. *Cell 37: 67-75.*

Sloof, P., Menke, H.H., Caspers, M.P. & Borst, P. 1983. Size fractionation of *Trypanosoma brucei* DNA: localization of the 177-bp repeat satellite DNA and a variant surface glycoprotein gene in a mini-chromosomal DNA fraction. *Nucleic Acids Res. 11:*

3889-3901.

Thiemann, O.H., Maslov, D.A. & Simpson, L. 1994. Disruption of RNA editing in *Leishmania tarentolae* by the loss of minicircle-encoded guide RNA genes. *EMBO J. 13: 5689-5700.*

Waitumbi, J.N. & Young, J.R. 1994. Electrophoretic karyotyping is a sensitive epidemiological tool for studying *Trypanosoma evansi* infections. *Vet. Parasito.l 52: 47-56.*

Wooden, J.M., Hartwell, L.H., Vasquez, B. & Sibley, C.H. 1997. Analysis in yeast of antimalaria drugs that target the dihydrofolate reductase of *Plasmodium falciparum. Mol. Biochem. Parasitol. 85: 25-40.*