Analysis of *Schistosoma mansoni* genes using the expressed sequence Tag approach

Amr M. Shabaan[1], Magdy M. Mohamed[1✉], Mohga S. Abdallah[2], Hayat M. Ibrahim[2] and Amr M. Karim[1]

[1]*Department of Biochemistry, Faculty of Science, Ain Shams University, Cairo, Egypt;*
[2]*Department of Chemistry, Faculty of Science, Helwan University, Cairo, Egypt*

Expressed sequence tags (ESTs) are partial cDNA sequences read from both ends of random expressed gene fragments used for discovering new genes. DNA libraries from four different developmental stages of *Schistosoma mansoni* used in this study generated 141 ESTs representing about 2.5% of *S. mansoni* sequences in dbEST. Sequencing was done by the dideoxy chain termination method. The sequences were submitted to GenBank for homology searching in nonredundant databases using Basic Local Alignment Search Tool for DNA (BLASTN) alignment and for protein (BLASTX) alignment at the National Center for Biotechnology Information (NCBI). Among submitted ESTs, 29 were derived from λgt11 sporocyst library, 70 from λZap adult worm library, 31 from λZap cercarial library, and 11 from λZap female B worm library. Homology search revealed that eight (5.6%) ESTs shared homology to previously identified *S. mansoni* genes in dbEST, 15 (10.6%) are homologous to known genes in other organisms, 116 (81.7%) showed no significant sequence homology in the databases, and the remaining sequences (2.1%) showed low homologies to rRNA or mitochondrial DNA sequences. Thus, among the 141 ESTs studied, 116 sequences are derived from noval, uncharacterized *S. mansoni* genes. Those 116 ESTs are important for identification of coding regions in the sequences, helping in mapping of schistosome genome, and identifying genes of immunological and pharmacological significance.

Schistosoma mansoni and *S. japonicum* are two most epidemiologically important species to concentrate research efforts on (Bruce *et al.,* 2000). The main attention has been fo-

cused on creating resources and techniques for development of stage specific libraries, on identifying genes of interest and on drawing low resolution physical maps of the whole parasite genomes. Priority was given to gene discovery using the expressed sequence tag (EST) approach leading to the deposition of 16 815 ESTs in dbEST (Oct./18/02). Despite this significant accomplishment, less than 15% of schistosomal genes have been identified and EST sequencing will therefore continue to be the primary objective of network activities.

Expressed sequence tags can serve the same purpose as random genomic DNA sequence-tagged sites (STSs) and provide the additional feature of pointing directly to an expressed gene. ESTs longer than 150 bp were found to be the most useful for similarity searches and mapping (Adams *et al.*, 1991). ESTs have application in the discovery of new genes, mapping genomes, identification of coding regions in genomic sequences, and the study of the mechanisms of tissue differentiation and ontogeny by developing profiles of sequences that are differentially expressed in particular cell types or at specific developmental stages (Boguski *et al.,* 1993).

Ever since the analysis of complex genomes began in earnest in the 1980s with the assembly of framework physical maps and the advent of methods for large-scale DNA sequencing, genome analysis has been justified in large part by the promise that it would eventually result in discoveries of great biological importance and pave the way for novel approaches in the study of biology. Time has passed and a number of genome projects are either complete or well underway (Hart, 1996; Feng *et al.*, 2002). The ESTs are submitted to the database of expressed sequence tags (dbEST), which is a division of GenBank at the U.S. National Center for Biotechnology Information, and made available to the public through GenBank and its international collaborators, the European Bioinformatics Institute and the DNA database of Japan. The non-

proprietary aspect of EST data and cDNA clone acquisition and use will maximize the enormous value offered to the biomedical community by this initiative (Cohen & Emanuel, 1994). Pharmaceutical companies, biotech companies and academics are now using dbEST to select genes for characterization and functional association (Gerhold & Caskey, 1996). Many commercial enterprises have, however, withheld or delayed submission to GenBank in order to gain a competitive edge in product development from the available information (Baxevanis *et al.*, 1996).

Homology searching is the process of comparing a new sequence against all other known sequences and then attempting to infere the function of the new sequence by assessing the matches and their biological annotation. There are a number of important issues in searching DNA and protein sequence databases, but the most important is the type of database. The schistosome genome project appears to increase our understanding of schistosome genome and its expression as well as provide a new approach to the identification of parasite gene products of immunological or pharmacological inserts (Simpson, 1990).

MATERIALS AND METHODS

Biological materials. The λgt-11 and λZap cDNA libraries from different *S. mansoni* stages were kindly provided by Dr. P.T. LoVerde (SUNY, Buffalo, U.S.A), and Dr. M. Saber (TBRI, Cairo, Egypt).

Parasites. *Schistosoma mansoni* Egyptian strain (Frandsen, 1979) and cercaria were obtained through the Schistosome Biological Supply Program (SBSP) at the Theodor Bilharz Research Institute. Molecular size marker was from Sigma Chemical Co. Bulk dNTPs were from Promega Corp.

Techniques for DNA analysis. *Schistosoma mansoni* Sporocyst λgt11 cDNA library was immunoscreened using affinity purified

antibodies, diluted 1:1000 in 1% BSA/1x PBST20 (Sambrook *et al.,* 1989). Preparation of *in vivo* excised λZap cDNA library was according to Sambrook *et al.* (1989). The λ bacteriophage which may contain insert or not was amplified by the PCR (Saiki *et al.,* 1988) using a pair of oligonucleotides complementary to sequences flanking the unique *Eco*RI site. The enzymatic DNA sequencing (Sanger *et al.,* 1977) system was applied for purified DNA samples. Analysis of the sequencing data was carried out according to the method of Ausubel *et al.* (1995).

Database 'homology' searches were conducted on the sequence data. The searches were conducted using the BLAST algorithm (Altschul *et al.,* 1990). BLASTN searches the homology between the nucleotide query sequence and nucleotide sequence databases, BLASTX searches for homology between the six frame translation of the nucleotide query sequence and amino-acid sequences in protein databases.

RESULTS

EST analysis

The results of a homology search in non-redundant databases for the 141 ESTs produced are summarized in Table 1. A total of 141 ESTs were generated from four different libraries constructed from four distinct stages of the parasite life cycle. The libraries were λgt11 sporocyst, λZap cercaria, λZap adult worm, and λZap female libraries. The sporocyst stage is an important stage for identifying antigens on the surface of 3-h schistosomules. Presumably the RNA messages directing the synthesis of those antigens are present in the sporocyst stage. The generation of expressed sequenced tags in this library depends on the arbitrary selection of indvidual cDNA clones. The efficiency of this process reflects the clonal structure of the libarary used and can be significantly increased using size selected cDNA library. This strategy, however, is not readily applicable when mRNA is limiting as is the case in this study of complete parasite genome. Immunoscreening of this library using affinity purified antibodies derived from rabbits immunized with irradiated cercariae led to the identification of 29 immunoreactive clones.

From the 141 ESTs we found that 5.6% are homologous to previously identified *S. mansoni* genes, 10.5% represent *S. mansoni* genes that are homologous to known genes in other organisms, about 2.1% represent genes of low to moderate homology to ribosomal RNA or mitochondrial DNA, the rest (about 81.6%) represent genes with no known homology in the non-redundant databases.

EST data

Tables 2–5 summarize the positive homology of ESTs submitted for each library.

Table 1. Analysis of homology results as determined by searches in non-redundant databases.

Positive matches are those with BLASTN > 200 or BLASTX > 100.

Libraries	ESTs No.	Genes homologous to *S. mansoni* genes	Genes homologous to known genes in other organisms	Genes with low homology to rRNA, mtDNA	Genes with no database homology
λgt11 sporocyst	29	1 (3.4%)	3 (10.3%)	——	25 (86.2%)
λ-Zap cercaria	31	5 (16.1%)	2 (6.5%)	——	24 (77.4%)
λ-Zap adult worm	70	2 (2.9%)	3 (4.3%)	3 (4.3%)	62 (88.6%)
λ-Zap female (B)	11	0 (0.0%)	7 (63.6%)	——	4 (36.4%)
Total	141	8 (5.7%)	15 (10.6%)	3 (2.1%)	116 (82.3%)

Table 2. Positive homology matches, λgt-11 sporocyst library

No.	EST name	showing homology with
1	MS1MM024.AS6	SM* major female specific polypeptide
2	MS1MM021.AS7	human mitochondrial ATP synthase α-subunit.
3	MS1AO001.AS6	*C. elegans* casein kinase II beta chain
4	MS1MM070.AS8	(Z99759) hypothetical protein

*SM, *Schistosoma mansoni*

λgt-11 sporocyst library

The diluted antibody used to immunoscreen λgt11 library was firstly adsorbed with recombinant λgt11 clones encoding antigens previously identified in our laboratory. This step was necessary in order to decrease the possibility of repeated isolation of the same clones and to facilitate identification of new antigen encoding clones. Twenty nine independent positive plaques were isolated from about 150 000 screened plaques after three rounds of immunoscreening. DNA of those 29 phages was prepared by using PCR with λgt11 forward and backward primers. DNA fragments were sized and separated on 1% agarose gel. The inserts generated from this library were digested by *Eco*R1 and subcloned in pGEM3ZF cloning vector and sequenced by double-stranded sequencing.

The average insert size was 0.62 kbp, while the average EST length was 249 bp. Figure 1 shows the size of randomly selected inserts by PCR. From 29 ESTs generated from this library, one clone is homologous to previously identified *S. mansoni* genes, three *S. mansoni* genes are homologous to known genes in

other organisms, and 26 ESTs have no homology in the databases. Table 2 describes the positive homology matches of the library generated sequences with *S. mansoni* and other organisms' ESTs from the GenBank.

ESTs from λ-Zap cercarial library

Thirty one ESTs were generated from this library and all were sequenced. The average insert size was 0.94 kbp, while the average EST length was 195 bp. Five of them match identified *S. mansoni* genes, two *S. mansoni* genes showed homology to known genes in other organisms, and 24 ESTs have no homology in the databases. Table 3 describes the positive homology matches of library generated sequences with *S. mansoni* and other organisms' ESTs from the GenBank.

ESTs from λ-*Zap* adult worm library

Seventy ESTs were generated from this library. The recombinant pBluescript SK plasmids were sequenced. The average insert size was 1.1 kbp, while the average EST length was 167 bp. Of these ESTs two were homolo-

Table 3. Positive homology matches, λ-Zap cercarial library

No.	EST name	showing homology with
1	MC2NE018.ACS	SM* serpin
2	MC3NE029.ACS	SM histone H1
3	MC3NE030.ACS	SM tegumental associated antigen Sm 20.8
4	MC2AS032.ACS	SM glyceraldhyde-3-P dehydrogenase
5	MC6AS062.ACS	SM aut1 mRNA
6	MC2NE023.ACS	*Drosophila melanogaster* actin 88 F protein
7	MC3NE033.ACS	glycerol-3-P dehydrogenase
8	MC6AS064.ACS	hypothetical protein-1 (*B. subtilis*)

*SM, as in Table 2.

Table 4. Positive homology matches, λ-Zap adult worm library

No.	EST name	showing homology with
1	MA1AS012.AAS	SM* 13 kDa tegumental antigen Sm13
2	MA2AS022.AAS	SM disulfide isomerase homologue precursor
3	MA3AS041.AAS	human RAS related protein RAP-1A
4	MA6ASO67AAS	(Z93379) F21H7.5 (*Caenorhabditis elegans*)
5	MA6ASO69AAS	(Z92834) F39B2.4 (*C. elegans*)

*SM, as in Table 2.

Table 5. Positive homology matches, λ-Zap female (B) library

No.	EST name	showing homology with
1	MFAS053.AFS	(AJ000671) clar protein (streptomyces)
2	MFAS054.AFS	(U28928) coded for by *C. elegans*
3	MFAS055.AFS	(Y09537) β-1,2-*N*-acetylglucose
4	MFAS056.AFS	hypothetical protein 1244-common tobacco
5	MFAS057.AFS	(U42580) a159R (*Paramecium bursar*)
6	MFAS058.AFS	protein transport Protein SEC9
7	MFAS060.AFS	DAZ-like protein (*Xenopus leavis*)

gous to previously identified *S. mansoni* genes, three are *S. mansoni* genes that are homologous to known genes in other organisms, three are low homologous to rRNA or mitochondrial DNA (mtDNA), and 62 ESTs have no significant homology in the databases. Table 4 describes the positive homology matches of the library generated sequences with *S. mansoni* and other organisms' ESTs from GenBank.

Figure 1. PCR amplification of λgt11 sporocyst cDNA showing different insert sizes on 1% agarose gel.

M (1 kbp ladder) molecular size marker.

ESTs from λ-Zap female (B) library

Eleven ESTs were generated from this library. The average insert size was 2.8 kbp, while average EST length was 183 bp. Seven are sequences homologous to identified *S. mansoni* genes, and four ESTs have no homology in the databases. Table 5 describes the positive homology matches of the library generated sequences with *S. mansoni* and other organisms' ESTs from GenBank.
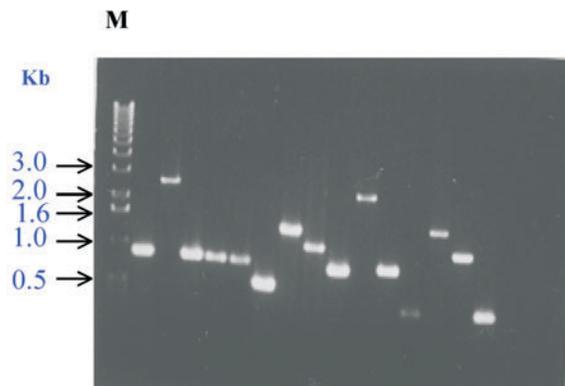
DISCUSSION

Schistosoma are dioecious digenetic trematodes carrying a large (270 Mb) genome. Gaining knowledge about the genome of these parasites is of importance for the understanding of their biology, mechanisms of drug resistance and antigenic variation that determine escape from the host's immune system (Franco *et al.,* 2000). The aim of the present study is a small scalle comparison of *S. mansoni* genes profile in different developmental stages based on the expressed sequence tags (ESTs) approach. Some of those genes may be

of importance for the development of a vaccine against schistosomiasis.

Initially we used λgt-11 sporocyst libraries; this was a nondirectional library so we ended up with a 50% chance for sequencing in the 3′ untranslated region. Most of the results, however, were obtained by sequencing the 5′ end of inserts in clones derived from directional λ-ZAP adult worm, cercariae and female specific libraries.

Most of the ESTs from different libraries were shown to have less than 5% useless clones and more than 80% new genes. The redundancy of each library was also analyzed, showing that one adult worm cDNA library was composed of a small number of highly frequent genes. When comparing ESTs from distinct libraries, we could detect that most genes were present only in a single library, but others were expressed in more than one developmental stage and may represent housekeeping genes in the parasite. When considering only once the genes present in more than one library, a total of 138 informative genes were obtained, corresponding to 116 new *S. mansoni* genes. From the total of unique genes, 10.6% were identified based on homology with genes from other organisms, 5.7% matched *S. mansoni* characterized genes and 82.3% represent unknown genes (Table 1).

We used immunoscreening to identify clones encoding *S. mansoni* vaccine candidates. This approach was successfully used previously in this laboratory to identify and characterize a number of important vaccine candidates (Osman *et al.*, 1995; Keung *et al.*, 1995; Ghazali *et al.*, 1996; Mohamed *et al.*, 1998; Mohamed *et al.*, 2000). This approach is based on the studies which have shown that the early schistosomula stage is an important target for immune elimination (Taylor, 1991).

Random sequencing of cDNA clones was done from excised λZap cDNA libraries for the purpose of generating ESTs; 112 sequences were generated in this study. Homology comparisons with DNA sequences in non-redundant databases revealed that of the 112 EST sequences seven (6.25%) encoded proteins homologous to identified genes in *S. mansoni,* 12 (10.72%) encoded sequences homologous to those of known proteins in organisms other than *S. mansoni.* The availability of those clones will thus facilitate characterization of the schistosome counterparts of proteins identified in other organisms. The remaining 93 (83.04%) EST sequences did not match sequences in the database. These clones are available for studies to identify the function and expression profiles of as yet unidentified proteins.

The precise characterization of the diversity of ESTs and physiological activity in different stage libraries can be ascertained by sorting the identified genes into functional categories. The detection of genes involved in tegument and membranes, transcriptional and translational activities or with regulatory and signaling functions is considered significant in cercariae followed by sporocyst then adult worm. Cytoskeletal and structural genes were also detected in ESTs that are essential for muscle contraction, movement through water or blood stream or laying out eggs. Genes that might be related to enzymes involved in energy metabolism formed the largest category in the case of the cercarial library which is not surprising considering that cercariae need much energy for rapid swimming and for penetrating the skin of the host. This finding is in accordance with previous expression studies of energy metabolism genes (Skelly & Shoemaker, 1995). The cercariae need to be fully prepared for the rapid morphological and physiological changes that will be necessary for their transformation into schistosomula to adult worm.

Some interesting matches to genes of *S. mansoni* or other organisms (<100% homology at amino-acid level) were found, suggesting that these may be new members of gene families that include several genes with similar sequences and probably similar functions. These new genes may provide important in-

sights into the physiology of the sporocyst, cercariae and adult worms. Amongst the informative genes we identified major female specific polypeptide, Sm20.8 tegumental protein, Sm disulfide isomerase, Sm glyceraldhyde-3-phosphate dehydrogenase, α-subunit of ATP synthase, *D. melanogaster* actin 88 protein, Sm13 tegumental protein. This is interesting because previous analyses had identified some of those genes from sporocysts to adult worms. This suggested that the protein is synthesized at the sporocyst stage and retained throughout the life of the worm (Grossman *et al.,* 1990), leading to the proposal that many genes for S. *mansoni* proteins were subject to stage-specific regulation.

EST MS1MM024.AS6 with insert size 0.6 kbp was shown to share 51% identity at the amino-acid level with the database *S. mansoni* major female specific polypeptide transcript of 0.36 kbp which encodes a protein of 104 amino acids. Homology occurs at the region from bp 5–103 of the EST, corresponding to amino acids 66–98 of the SM major female specific polypeptide. The sexual maturation of adult schistosomes, pairing of male and female worms and the subsequent massive and continuous production of eggs are the major developmental events of the parasite that occur within the mammalian host. The end product of maturation is the production of viable eggs and thus studies of the maturation process have concentrated on the development of the female vitelline gland within which the major egg proteins are synthesized. Indeed, in females which pair with males, the appearance of differentiating vitelline cells is the earliest sign of reproductive maturation. Apparently, a single type of mRNA, which is absent from male and immature female worms dominates the message population of egg producing female schistosomes. For simplicity we will refer to the product of this gene as the major female specific polypeptide (FSP) and the gene as the *FSP* gene. It has been demonstrated that the presence of the male worm is required to induce female specific polypeptide

gene expression although it is not yet known whether an agent is transferred from one sex to the other, which acts directly on the nuclear DNA of the other sex or if an internal signal is induced. Although the gene is very highly expressed, it is present in a very low copy number and is found in the DNA of both sexes. It is not rearranged or amplified during expression. The gene is first expressed five weeks after infection of the mammalian host (Simpson *et al.,* 1987).

EST MC3NE030.ACS with insert size 0.8 kbp was shown to share 58% identity at the amino-acid level with the database *S. mansoni* tegumental antigen Sm20.8 transcript of 0.7 kbp which encodes a protein of 181 amino acids. Homology occurs at the region from bp 34–207 of the EST, corresponding to amino acids 1–58 of Sm20.8. Sm20.8 shares homology with Sm21.7, Sm22.6 and Sj22.6, schistosome vaccine candidates which are members of a family of tegument associated antigens. It is not clear what role these proteins play in the tegument of schistosomes, they may help to maintain this organ. Members of this family contain a sequence motif characteristic of Ca-binding proteins known as the EF hand motif, Sm20.8 possesses a similar motif at amino acids 38–56. Sm20.8 has an isoelectric point of 7.27 and its expression is developmentally regulated, with the highest concentration found in cercariae, Sm20.8 shows immunoreactivity with sera from infected humans and rabbit vaccinated with irradiated cercariae. Confocal microscopy demonstrates that Sm20.8 localizes to the tegument of adult worms and 3 h schistosomula. The IgG fraction specific to Sm20.8 mediated complement killing of schistosomula *in vitro* by 35%. Vaccination of mice with naked DNA containing the Sm20.8 gene and subsequently challenged with cercariae showed 30% reduction in worm burden compared to controls (Mohamed *et al.,* 1998).

EST MA2AS022.AAS with insert size 0.7 kbp was shown to share 88% identity at the amino-acid level with the database *S. mansoni*

protein disulfide-isomerase (PDI) homologue precursor transcript of 5.3 kbp which encodes a protein of 482 amino acids. Homology occurs at the region from bp 6–110 of the EST, corresponding to amino acids 445–479 of the protein disulfide-isomerase homologue precursor. Also, another EST MA2AS022.AAS with insert size 0.4 kb was shown to share 95% identity with the same protein but the homology occurs at the region from bp 1–60 of the EST, corresponding to amino acids 463–482 of the protein disulfide-isomerase homologue precursor. PDI is a ubiquitous enzyme that is localized in the lumen of the endoplasmic reticulum. The deduced protein sequence contains two thioredoxin like domains with a particularly high sequence identity of 68% between *S. mansoni* and man. Thioredoxin domains characterize all PDI genes hitherto sequenced but they are also found in some other genes with very different functions. The schistosome PDI gene and the human PDI gene, however, share also significant similarity (34%) in other parts of the coding region besides the thioredoxin domains. In vertebrates PDI is mainly accumulated in secretory cells where its abundance correlates with the level of secretory protein synthesis. Its major function is to catalyze isomerization of intramolecular disulfide/sulfhydryl bonds, a reaction that is required for folding of nascent polypeptides in the endoplasmic reticulum. The PDI genes of *S. mansoni* and man are also very similar in their genomic structure, the sequence positions of the introns in the schistosome gene are exactly the same as in the human PDI gene (Finken *et al.,* 1994*).*

EST MC2AS032.ACS with insert size 1.6 kbp was shown to share 90% identity at the amino-acid level with the database *S. mansoni* glyceraldhyde-3-phosphate dehydrogenase transcript of 1.5 kbp which encodes a protein of 338 amino acids molecular mass 37 kDa. Homology occurs at the regions from bp 3–110, 112–138, 159–200 of the EST, corresponding to amino acids 14–49, 50–58, 65–78 of the protein. Schistosomes ingest large

quantities of glucose, as much as 26% of its dry body mass per hour, since they depend on glycolysis for their energy production, so glyceraldhyde-3-phosphate dehydrogenase is certainly critical to parasite survival. The enzyme was identified as a surface associated vaccine candidate. The function of the surface-located enzyme is, however, not clear. So far there is no evidence indicating that it plays a metabolic role at the surface since it would have to be associated with other glycolytic enzymes to perform its catalytic activity (Goudot-Crozel *et al.,* 1989).

EST MS1MM021.AS7 with insert size 0.6 kbp was shown to share 94% identity at the amino-acid level with the database human mitochondrial α subunit of ATP synthase transcript of 1.8 kbp which encodes a protein of 553 amino acids. Homology occurs at the region from bp 2–265 of the EST, corresponding to amino acids 364–451 of the mitochondrial α subunit of the ATP synthase. The ATP synthase complexes of mitochondria, chloroplasts, and bacterial membranes are structurally and functionally similar. Functionally, ATP synthase complexes synthesize ATP from ADP and P_i utilizing energy generated from the electron transport chain (Breen, 1988).

EST MC2NE023.ACS with insert size 0.6 kbp was shown to share 98% identity with *S. mansoni* actin mRNA of 1.5 kbp. Homology occurs at the region from bp 1–167 of the EST, corresponding to bases 687–853 of the actin mRNA. Actins constitute a highly conserved family of proteins found in all eukaryotes. These proteins are found predominantly in the cytoplasm of cells as monomers polymerize to form microfilaments. Microfilaments of actin participate in various cell functions such as muscle contraction, cell cytoskeleton and motility. Actin microfilaments also function in bundles where they form microvilli, stereocilia, and other cellular structures. The tegument of the human blood fluke *S. mansoni* is limited by two lipid membrane bilayers that contain numerous crystal-

line spines. Optical diffraction patterns suggested that these spines are composed predominantly of actin bundles. Actin has since been shown by fluorescence microscopy to be present on the surface of schistosomula, and by immunofluorescence to be present in the muscle, tegumental tubercles and spines of male and female adult worms. Two-dimensional gel electrophoresis and immunoblotting have detected several isoforms of actin. Actin, either associated with tegumental spines or free in the cytoplasm, has been implicated in the repair and maintenance of the integrity of the adult worm's tegument. *S. mansoni* actin genes are more homologous to invertebrate actin than vertebrate cytoplasmic actin genes, as expected due to their evolutionary relationship. Although highly conserved, the differences observed in the primary sequence of the parasite actin may account for the immunogenicity of the protein (Oliveira & Kemp, 1995).

EST MA1AS012.AAS (GenBank accession # AA269243) with insert size about 0.65 kbp was shown to share 100% identity at the amino-acid level with the database *S. mansoni* protein Sm13 transcript of 0.96 kbp which encodes a protein of 320 amino acids. Sequence analysis revealed a 104 amino acid open reading frame (ORF) identical to that described by Abath *et al.* (2000). Our sequence also shows 100% identity with GenBank genomic Sm13 sequence AF07886 of 960 bp and with Sm13 mRNA sequence U67153 of 396 bp. However, EST database searches revealed three almost identical sequences (GenBank accession numbers AAS17940, N20684 and AA269243). The independent isolation of homologous cDNA confirmed the sequence of Sm13.

This study will provide additional data to the Schistosoma Gene Discovery Program, which is part of the Schistosoma Genome Project created in 1992. One of the main objectives of this program is the discovery and characterisation of new genes of *S. mansoni* and *S. japonicum* in an attempt to search for new targets for drugs and vaccine development.

The success of the Schistosoma Gene Discovery Program is demonstrated by the number of the catalogued genes that now reaches 15 to 20% of the full gene complement of its genome (Franco *et al.,* 2000).

REFERENCES

Abath GC, Fredrerico EM, Xavier RA, Yara MG, Norma L, Marcilio B, Simpson JG. (2000) Characterization of Sm13, a tegumental antigen of *Schistosoma mansoni. Exp Parasitol.*; **19**: 12–9.

Adams MD, Kelley JM, Gocayne JD, Dubnick M, Polymeropoulos MH, Xiao H, Merril CR, Wu A, Olide B, Moreno RF, Kerlavage AR, McCobie WR, Venter C. (1991) Complementary DNA sequencing: expressed sequence tags and human genome project. *Science.*; **252**: 1651–6.

Ausubel FM, Brent R, Kingston RE, Moore DD, Sadman JG, Smith JA, Struhl K, Albright LM, Coen DM, Vaki A. (1995) Informatics for molecular biologist. In *Current protocols in molecular biology.* Ausubel FM, Brent R, Kingston RE, *et al*. eds, Chap 19, John Wiley.

Baxevanis AD, Boguski MS, Ouellette BF. (1996) Computational analysis and annotation of sequence data. In *Genome analysis: a laboratory manual.* Birren B, Green E, Hieter P, Klapholz S, Myers R. (eds).

Bergquist R. (1993) Schistosomiasis. In *Tropical disease research progress 1991–1992*, 11th programme report of the UNDP/World Bank/WHO, pp. 29–36.

Boguski MS, Lowe TM, Tolstoshev CM. (1993) dbEST — database for "expressed sequence tags". *Nat Genet.*; **4**: 332–3.

Breen GA. (1988) The major parasite surface antigen associated with human resistance to schistosomiasis is a 37-kD glyceraldehyde-3P-dehydrogenase. *J Exp Med.*; **170**: 2065–80.

Bruce JL, Sandra Z, John MW. (2000) Ecological factor in schistosome transmission, and an environmentally benign method for con-

trolling snails in recreational lake with a re-cord of schistosome. *Dermatitis*.; **49**: 9–17.

Feng XG, Lin JJ, Zhu JG, Zhang L, Yuan CX, Cai YM. (2002) Protective effect induced by vaccination with a partial cDNA expression library of *Schistosoma japonicum* in mice. *Sheng Wu Gong Cheng Xue Bao*.; **18**: 231–4.

Finken M, Sobek A, Symmons P, Kunz W. (1994) Characterization of the complete pro-tein disulfide isomerase gene of *Schistosoma mansoni* and identification of the tissues of its expression. *Mol Biochem Parasitol*.; **64**, 135–44.

Franco GR, Valadao AF, Azevedo V, Rabelo EM. (2000) The Schistosoma gene discovery pro-gram: state of the art. *Int J Parasitol*.; **10**: 53–63.

Frandson F. (1979) Studies of the relationship between *Schistosoma mansoni* and their inter-mediate hosts III the genus *Biomphalaira* and *Schistosoma mansoni* from Egypt, Kenya, Sudan, Uganda, West Indies (St. Lucia) and Zaire (two different strains: Katanga and Kinshasa). *J Helmin*.; **53**: 321–48.

Cohen MM, Emanuel BS. (1994) Expressed Se-quence Tags. *Science*.; **266**: 1790–1.

Gerhold D, Caskey CT. (1996) It's the genes! EST access to human genome content. *BioEssays*.; **18**: 973–81.

Ghazalie NE, Kiang D, Medhat AM, Abdel-Fatah M, Karim AM, LoVerde PT. (1996) Identifica-tion and characterization of *Schistosoma mansoni* p177, a cyclophilin. *Biochem Parasitol*.; **76**: 73–82.

Grossman Z, Ram D, Markovies,A, *et al.* (1990) Sm: stage specific expression of muscle-spe-cific genes. *Exp Parasitol*.; **70**: 62–71.

Hart DL. (1996) EST! EST!! EST!!! *BioEssays*.; **18**: 1021–3.

Jeffs SA, Hagan P, Allen R, Oliveira CR, Smithers SR, Simpson AJG. (1991) Molecu-lar cloning and characterization of the 22 kDa adult *Schistosoma mansoni* recognized by antibodies from mice protectively vaccinated with isolated tegumental surface membranes. *Mol Biochem Parasitol*.; **46**: 159–68.

Keung WL, Shalaby KF, Amina MM, Hong SY, Karim AM, LoVerde PT. (1995) *Schistosoma mansoni:* characterization of gene encoding Sm23 an integral membrane protein. *Exp Parasitol*.; **80**: 155–8.

Mohamed MM, Shalaby KF, LoVerde PT, Karim AM. (1998) Characterization of Sm208, a member of a family schistosome tegumental antigens. *Mol Biochem Parasitol*.; **96**: 15–25.

Mohamed MR, Shalaby KF, Abdellah NM, LoVerde PT, Karim AM. (2000) *Schistosoma mansoni* Myosin heavy chain: characteriza-tion of an immunoreactive fragment with vaccine potential *N Egypt J Med*.; **22**: 297–303.

Oliveira GC, Kemp WM. (1995) Cloning of two actin genes from *Schistosoma mansoni. Mol Biochem Parasitol*.; **75**: 119–22.

Osman A, Kian D, LoVerde PT, Karim AM. (1995) *Schistosoma mansoni*: characterization of p50, an immunophilin. *Exp Parasitol*.; **80**: 550–9.

Saiki RK, Gelfand DH, Stoffel S, Scharf SJ, Higuchi R, Horn GT, Mullis KB, Erlich HA. (1988) Primer-directed enzymatic amplifica-tion of DNA with thermostable DNA polymer-ase. *Science*.; **239**: 487–91.

Sanger F, Nichlen S, Coulson AR. (1977) DNA sequencing with chain terminating inhibi-tors. *Proc Natl Acad Sci U S A*.; **74**: 5463–7.

Skelly PJ, Shoemaker CB. (1995) A molecular genetic study of the variations in metabolic function during schistosome development. *Mem Inst Oswaldo Cruz*.; **90**: 281–4.

Simpson AJ, Knight M, Kelly C, Hackett F, Omer-Ali P, Smithers SR. (1987) The cloning of schistosome antigens. *Mol Biochem Parasitol*.; **64**: 135–44.

Simpson AJG. (1990) Schistosoma surface anti-gens: developmental expression and immuno-logical function. *Parasitol Today*.; **6**: 40–5.

Taylor DW. (1991) Schistosoma vaccine. *Experientia*.; **47**: 126–38.

Young RA, Davis RW. (1983) Yeast RNA poly-merase II genes: isolation with antibody probes. *Science*.; **222**: 778–81.