A bacterial artificial chromosome library for *Biomphalaria glabrata*, intermediate snail host of *Schistosoma mansoni*

Coen M Adema/+, Mei-Zhong Luo*, Ben Hanelt, Lynn A Hertel†, Jennifer J Marshall, Si-Ming Zhang, Randall J DeJong/++, Hye-Ran Kim*, David Kudrna*, Rod A Wing*, Cari Soderlund, Matty Knight***, Fred A Lewis***, Roberta Lima Caldeira****, Liana K Jannotti-Passos****, Omar dos Santos Carvalho****, Eric S Loker**

Department of Biology, University of New Mexico, Albuquerque, NM, US *Arizona Genomics Institute, Department of Plant Sciences **Arizona Genomics Computational Laboratory, BIO5 Institute, University of Arizona, Tucson, US ***Biomedical Research Institute, Rockville, MD, US ****Laboratório de Helmintoses Intestinais, Centro de Pesquisas René-Rachou-Fiocruz, Belo Horizonte, MG, Brasil

To provide a novel resource for analysis of the genome of Biomphalaria glabrata*, members of the international* Biomphalaria glabrata *Genome Initiative (biology.unm.edu/biomphalaria-genome.html), working with the Arizona Genomics Institute (AGI) and supported by the National Human Genome Research Institute (NHGRI), produced a high quality bacterial artificial chromosome (BAC) library. The BB02 strain* B. glabrata*, a field isolate (Belo Horizonte, Minas Gerais, Brasil) that is susceptible to several strains of* Schistosoma mansoni*, was selfed for two generations to reduce haplotype diversity in the offspring. High molecular weight DNA was isolated from ovotestes of 40 snails, partially digested with* Hind*III, and ligated into pAGIBAC1 vector. The resulting* B. glabrata *BAC library (BG_BBa) consists of 61824 clones (136.3 kb average insert size) and provides 9.05 × coverage of the 931 Mb genome. Probing with single/low copy number genes from* B. glabrata *and fingerprinting of selected BAC clones indicated that the BAC library sufficiently represents the gene complement. BAC end sequence data (514 reads, 299860 nt) indicated that the genome of* B. glabrata *contains ~ 63% AT, and disclosed several novel genes, transposable elements, and groups of high frequency sequence elements. This BG_BBa BAC library, available from AGI at cost to the research community, gains in relevance because BB02 strain* B. glabrata *is targeted whole genome sequencing by NHGRI.*

Key words: genomics - gene discovery - fingerprinting - schistosomiasis - medical malacology

The application of molecular approaches continues to contribute novel insights into the biology, including genomics of molluscs (Zhang et al. 2004, Mitta et al. 2005). To date, several mitochondrial genomes of molluscs have been sequenced (DeJong et al. 2004, Mizi et al. 2005), but the nuclear genome of a representative of the Phylum Mollusca remains to be fully characterized. In fact, lophotrochozoan protostomes of which mollusca represent the largest group (Rouse 1999), are underrepresented among the animals from which the current assembly of fully sequenced genomes has been obtained. Thus, genomic data from a mollusc will help fill a gap in the information on the evolutionary history of animal life (Collins et al. 2003).

Financial support: the production and distribution of the BG_BBa BAC library at AGI was supported by the funding from the National Human Genome Research Institute under the BAC Library Production program (grant 5U01HG002525; RAW). Parts of this study were supported by NIH grants AI024340 (ESL), AI052363 (CMA), and Fiocruz
+Corresponding author: coenadem@unm.edu.
†Deceased 2 April 2005
++Present address: Laboratory of Malaria and Vector Research, NIAID/NIH, Twinbrook III, Room 2E-20 MSC 8132 Bethesda, MD 20892, US
Received 25 May 2006
Accepted 26 June 2006

Molluscs are a highly diverse group that includes some of the largest, longest living, and most intelligent invertebrates. Genome information will instruct on several remarkable properties of molluscs such as shell formation (biomineralization; Milet et al. 2004), the evolution of body asymmetry (Schilthuizen & Davison 2005), and hermaphrodism (Paraense & Corrêa 1988). Molluscs are also being used to study pharmo-toxicology (Terlau & Olivera 2004); neuroendocrinology (Altelaar et al. 2005); parthenogenesis (Jokela et al. 2003); and the molecular basis of behavior and learning (Williamson & Chrachri 2004, Zhurov et al. 2005). Molluscs serve as bioindicators for monitoring of the environment (Zhao et al. 2005), and (snails especially) are useful for understanding how natural selection operates (Vermeij 2002). Furthermore, molluscs are economically important as a major source of food, can destroy crops, colonize and impact new habitats as invasive species (Pointier et al. 2005), and transmit medically important pathogens.

The latter applies to the freshwater gastropod *Biomphalaria glabrata* (Planorbidae, Basommatophora). This snail serves as one of the most important intermediate hosts for a widespread pathogen of humans, the digenetic trematode *Schistosoma mansoni* (Paraense & Corrêa 1963, Morgan et al. 2001). This parasite causes intestinal schistosomiasis, a debilitating disease that afflicts over 50 million humans (Chitsulo et al. 2004). To a large extent, the geographic distribution of *B. glabrata* defines the

distribution of *S. mansoni* in the Neotropics (Paraense 1986, DeJong et al. 2003). Genetic determinants affect the susceptibility of *B. glabrata* for *S. mansoni* (Lewis et al. 2001), and heterogeneity in genetic composition of *B. glabrata* on smaller scales may further influence the transmission patterns of schistosomiasis (Theron & Coustau 2005). More comprehensive genome sequence data for *B. glabrata* would enable novel investigative approaches to study determinants of transmission, especially in light of an advancing genome project for *S. mansoni* (Loverde et al. 2004).

B. glabrata also hosts a variety of other digenetic trematodes and has been adopted as the most commonly used model host to study the basic biology of digenean-snail interactions (Lie 1982, Adema & Loker 1997, Vergote et al. 2005). As one example, *B. glabrata* has been found to produce after exposure to digeneans a unique family of hemolymph molecules termed FREPs (fibrinogen-related proteins). FREPs consist of a juxtaposition of fibrinogen and immunoglobulin superfamily domains, and have proven to be remarkably diverse in their composition. *B. glabrata* thus serves as a new model system to examine the nature and diversity of non-self recognition molecules produced by invertebrates (Zhang et al. 2004).

Information on the genome of *B. glabrata* will also have relevance for several other *Biomphalaria* species and for yet other species of molluscs which serve as hosts for schistosomes and for a number of other trematode, and some nematode, infectious agents. Besides schistosomiasis, diseases such as fascioliasis, clonorchiasis, and paragonimiasis represent only a few of the snail transmitted diseases with worldwide medical and economic impact (Lockyer et al. 2004a).

In 2001, an international consortium, "the *Biomphalaria glabrata* genome initiative" was founded to develop genome-type projects for this particular pulmonate gastropod species (http://biology.unm.edu/biomphalaria-genome/index.html). Members of this consortium have contributed several gene discovery projects (Jones et al. 2001, Miller et al. 2001, Schneider & Zelck 2001, Raghavan et al. 2003, Lockyer et al. 2004b, Nowak et al. 2004, Jung et al. 2005, Mitta et al. 2005), the full-length sequence of the mitochondrial genome of *B. glabrata* (DeJong et al. 2004), and an estimate of 931 Mb for the size of the nuclear genome of *B. glabrata* (Gregory 2003).

A novel resource for genomic studies became available when the National Human Genome Research Institute (NHGRI) awarded a white paper application (http://www.genome.gov/Pages/Research/Sequencing/BACLibrary/ BgBACprops.pdf) for funding of the production a high quality bacterial artificial chromosome (BAC) library for *B. glabrata* (http://www.genome.gov/page.cfm ?pageID=10001852). Such a library provides access to large regions of the genome of *B. glabrata*, in an experimentally manageable fashion. Significantly, the NHGRI support guaranteed high quality standards for the finished library, and also made the BAC library publicly available at cost to the research community. The actual development of the BAC library was a collaboration between the Arizona Genomics Institute (AGI; part of the National Institutes of Health BAC Resource Network) and members of *B. glabrata* genome initiative. The genomic DNA from a recent *B. glabrata* field isolate from a schistosomiasis endemic area in Brazil, shown to be susceptible to *S. mansoni*, was used to ensure that the BAC library provides data that are relevant in the context of parasite-snail compatibility. This report describes the *B. glabrata* strain used, and both the production and characterization of the BAC library. Lastly, analysis of sequence data obtained provides first glimpses into the genomic make-up of *B. glabrata*.

MATERIALS AND METHODS

Snails, species identification and susceptibility for schistosome infection - *B. glabrata* snails were collected from a small stream in an endemic site for transmission of *S. mansoni*, in the south east of Brazil, Barreiro, Minas Gerais, (19ºS 59 min/44ºW 02 min). Offspring of these snails are maintained as a laboratory strain designated as BB02 (*Biomphalaria* from Belo Horizonte, Minas Gerais, Brazil 2002).

The species identity of BB02 snails was determined by PCR-RFLP. The ITS1-5.8S-ITS2 sequence region was PCR amplified from DNA of individual snails using primers (all primers are shown 5' -3') ETTS2: TAA CAA GGT TTC CGT A GG TGA A and ETTS1: TGC TTA AGT TCA GCG GGT. The amplicons were digested with *Dde*I and restrictions patterns obtained from BB02 snails were compared to the characteristic banding pattern specific for *B. glabrata* (Vidigal et al. 1998). Also sequences from the 16SrDNA and ND1 genes of one BB02 snail were amplified by PCR, using primer pairs 16Sar: CGC CTG TTT ATC AAA AAC AT - 16Sbr: CCG GTC TGA ACT CAG ATC ACG T (Palumbi et al. 1996) and SNDF1F2: CGR AAA GGA CCT AAY AGT TGG - SND1R4: ART CRA ATG GYG CHC GAT TAG, respectively. (R=A/G Y=C/T H= A/C/T). The sequences from these amplicons were obtained by direct sequencing and analyzed relative to previously generated phylogenies of *Biomphalaria* isolates, all according to DeJong et al. (2003). The sequences of *16S* rDNA and *NADH dehydrogenase 1* were deposited in GenBank under accession numbers AY737280 and AY737281, respectively.

Members of the F1 generation derived from field collected snails were tested for susceptibility to two different *S. mansoni* strains (LE, SJ) at the Section of Molluscs Rearing at the Centro de Pesquisas René-Rachou in Belo Horizonte, Brazil. Groups of 50 juvenile snails (3-6 mm) were exposed individually to 10 miracidia. The parasite-susceptible BB01 strain of *B. glabrata* (maintained over 10 years in the laboratory in Brazil) was used as a control for miracidial infectivity. At 4 weeks post exposure, snails were exposed to artificial light for 30 min and the shedding of cercariae was recorded. Non-shedding snails were dissected to check for developing sporocysts. BB02 *B. glabrata* were similarly tested for susceptibility to the NMRI strain of *S. mansoni* at the Biomedical Research Institute (MD, US).

Preparation of HMW genomic DNA from BB02 B. glabrata - Initial comparisons disclosed that relative to whole body or the digestive gland, the ovotestis of *B. glabrata* was optimal for generation of monocellular sus-

pensions as required to obtain high molecular weight DNA (Luo & Wing 2003). However, the DNA yield from a single snail was insufficient to generate a BAC library. *B. glabrata* is a simultaneous hermaphrodite and offspring were generated by selfing to minimize haplotype diversity. One newly hatched BB02 snail (< 3 mm shell diameter) was kept in isolation to generate F1 progeny by self-fertilization (sF1). A selfed F2 generation (sF2) derived from the sF1 was similarly obtained.

High molecular weight (HMW) genomic DNA was isolated from forty sF2 snails (10-12 mm shell diameter). Following cleaning and removal of shells, live snails were kept briefly in ½199 medium (physiological buffer for snail cells; medium 199 (Sigma) diluted 1:1 [v/v] with distilled water) until all snails were processed. From 4 snails at a time, the ovotestes were dissected and pooled in 800 µl of ½199 in 1.7 ml Eppendorf tubes on ice. All the following manipulations were performed gently to minimize damage to cells and mechanical shearing of DNA. The tissues were disrupted with 3 strokes of a polypropylene pellet pestle (Kontes). The resulting cell suspensions were pooled in a 50 ml Falcon tube on ice. No sediment was evident after 1 h. Cells were pelleted (400 g, 5 min at 4°C) and the cleared supernatant fluid was reduced to 600 µl. The cells were resuspended uniformly by tapping the side of the tube and incubated for 3 min at 42C. Then, 600 µl of 1% Seakem agarose (FMC) in ½199, (pre-warmed to 42°C) was mixed with the cells using minimal agitation. The monocellular cell suspension in agarose was transferred (using a cut-off, wide bore pipette tip) into disposable CHEF plug moulds (Bio-Rad) to obtain plugs with uniform cell numbers embedded in an agarose matrix, and placed on ice for 20 min. The 13 resulting plugs were transferred to 50 ml NDS (0.5 M EDTA, 10 mM Tris, 1% w/v N-lauroyl sarcosine, pH 9.5 (NaOH), supplemented with 1 mg/ml proteinase K (Invitrogen) and incubated overnight at 50°C in a rotary hybridization oven. This treatment lysed the cells while the agarose matrix protected high molecular weight genomic DNA from mechanical shearing. The medium was replaced by NDS and again incubated overnight at 50°C with rotation. DNA quality and susceptibility to *Hind*III digestion were evaluated by contour-clamped homogeneous electric field (CHEF) gel electrophoresis.

Generation of the BG_BBa BAC library - The methods of Luo and Wing (2003) were used to produce the BAC library. Briefly, following testing to determine optimal conditions, HMW DNA embedded in plugs was partially digested with *Hind*III. Following separation on CHEF gels twice, DNA fragments in the 150-300 kilobase (kb) range were eluted and ligated into pAGIBAC1. This BAC vector carries a resistance marker for chloramphenicol and incorporates a high signal for blue/white screening of non-insert transformants. The resulting constructs were introduced into DH10B-T1 phage resistant *Escherichia coli* cells by electroporation and plated on LB containing 12.5 µg/ml chloramphenicol and 80 µg/ml X-gal, 100 µg/ml IPTG for blue/white screening. Guided by video recognition of successful transformants, clones were picked and gridded into 384 well plates by a Q-bot (Genetix). Clones were stored as glycerol cultures at –80°C. Also, the clones from the BAC library were inoculated on four 22.5 × 22.5 cm Hybond N+ filters (Amersham) in high density, double spots and 4 × 4 patterns with a Q-bot (Genetix). The resulting filters a, b, c each contained 18432 clones in duplicate in six fields, the last filter (d) held 6528 clones in the same layout. The membranes were placed on LB agar plates containing 12.5 µg/ml chloramphenicol and incubated overnight to obtain colonies of 1 to 2 mm diameter. The membranes were placed (colony side up) on absorbent filter paper (Whatman Cat. No. 3030 700) soaked in the following solutions: (1) solution 1 (0.5 N NaOH, 1.5 M NaCl) for 7 min; (2) solution 2 (1.5 M NaCl, 0.5 M Tris-HCl, pH 8.0), 7 min; (3) air dry for more than 1 h; (4) solution 3 (0.4 N NaOH), 20 min; (4) solution 4 (5X SSPE), 7 min, and air dried overnight. The complete library (as frozen stocks), high density filters, and individual clones are available at cost from AGI. Protocols for screening of high density BAC library filters and address determination of positive signals are publicly available from AGI (www.genome.arizona.edu).

Isolation and sequencing of BAC DNA - At AGI, BAC DNA was isolated from 1.2 ml 2 × YT (Fisher) overnight cultures using alkaline lysis (96-well format) with a Quadra 96 Model 320 (Tomtec). Both ends of BAC inserts were sequenced using T7: TAA TAC GAC TCA CTA TAG GG as ''forward'' primer and BES_HR: CACT CAT TAG GCA CCC CA as the ''reverse'' primer. Cycle sequencing (BigDye Terminator v 3.1, Applied Biosystems) was performed using PTC-200 thermal cyclers (MJ Research) in 384-well format applying 150 cycles of 10 s at 95°C, 5 s at 55°C, and 2.5 min at 60°C. Extension products were purified by CleanSeq magnetic beads (Agencourt). Samples were eluted into 20 µl of ddH_20 and separated on ABI 3730xl capillary sequencers with default conditions. Sequence data were collected by data collection software (Applied Biosystems), and transferred to a UNIX workstation. Sequences were base-called using the program Phred (Ewing & Green 1998, Ewing et al. 1998); vector and low-quality (Phred value <16) sequences were removed using the program Lucy (Chou & Holmes 2001). The methods applied at UNM included Montage BAC96 (Millipore) and Perfectprep BAC 96 (Eppendorf) for isolation of BACs. BAC ends were sequenced (Big Dye v. 3.1, ABI), also with T7 and BES_HR primers, using Biometra T-gradient thermal cyclers in 96 well format. The temperature profile was 1 min at 94°C, 100 cycles of 30 s at 94°C, 1 min at 55°C, 1 min at 72°C, and 7 min at 72°C. Following cleanup (Montage SEQ96; Millipore), extension products were read on an ABI 3700. Sequencher (GC codes) was used to remove vector sequences and edit chromatograms by eye.

Quality control of the BAC library - To estimate the average insert size of the BAC library, BACs were extracted from 361 randomly selected clones at AGI. The DNA was digested to completion with *Not*I (3 h/37°C) and separated on 1% CHEF gels to determine the size of the insert DNA. These data were applied to calculate the estimated coverage of the genome of *B. glabrata* by the BAC library. Absence of insert DNA was monitored to determine the proportion of empty vector in the BAC library.

The non-redundancy of BAC inserts was tested by sequencing (AGI) both termini of a random set of 192 clones. The clones were arbitrarily selected from wells A01, A02, A03 from plates 1-32, and well B23 from plates 1-96 in which the library is stored.

The representation of the genome of *B. glabrata* in the BAC library was investigated by screening the BAC library for sequences representing low- or single copy genes of *B. glabrata* (UNM). The probe sequences were selected from the literature, or chosen arbitrarily (see Table II). The probes were amplified by PCR from genomic or cDNA templates, labeled with ^{32}P α dCTP (Perkin Elmer) by random priming (Prime-it RT, Stratagene), and used as hybridization probes to screen filters that contained spotted BAC clones. The initial screening of high density filters representing the whole library (as available from AGI) was performed with two sets of five pooled probes (see http://www.genome.arizona.edu/information/protocols/index.html). The filters were prehybridized at 65°C for at least 4 h with hybridization buffer (0.5 M sodium phosphate pH 7.2, 7% SDS, 1 mM EDTA, 10 µg/ml sheared salmon DNA). After an exchange with fresh buffer, prehybridization was continued for 2 h. The probes were added and hybridized (>18 h, 65°C). The filters were washed sequentially with 2X SSC, 1X, and 0.1X SSC (all containing 0.1% SDS), 2 times each (20 min, 65°C), then autoradiography was performed. Positive clones were identified and obtained from AGI as bacterial stab cultures. These clones were used to manually prepare macroarrays (96 well format) applying similar methods as described above for the high density filters. The macroarrays were screened with individual probes to determine which clones contained specific target sequence. The BAC clones were also end-sequenced.

Contig alignment of BACs by fingerprinting - The BACs from clones that strongly hybridized the low- or single copy probes were subjected to the fingerprinting methods described by Luo et al. (2003). The resulting digestion patterns were compared for similarities to identify and contiguously align (partially) overlapping BACs using FPC software for the contig assembly (Soderlund et al. 2000). Also see http://www.genome.arizona.edu/BAC_special_projects/

Computational analysis and annotation of BAC end sequences - A contig analysis of the BAC end sequences was performed using Sequencher (GC codes). The clustering criteria were arbitrarily set at 98% identity over 100 nucleotides. The AT content was calculated for all non-redundant (sequence contigs were used instead of individual cluster mates) BAC end sequence data combined. BLAST searches were used to investigate the likelihood that BAC inserts were of snail origin, as well as to uncover similarities between BAC end sequences and the protein and nucleotide databases of GenBank, with special consideration of sequence entries from *B. glabrata*. E-values ≤ 10^{-4} were considered significant. Discrepancies in sequence similarities between genomic and cDNA sequences were analyzed for the presence of non-coding sequences, including introns. Repetitive sequences were identified by direct inspection of sequence data and by analysis of results from BLAST searches. The BAC end sequence data were submitted as genome survey sequences (GSS) to GenBank under accession numbers CZ547921-CZ548269; DX360039-DX360203.

RESULTS

Characteristics of the BB02 strain of B. glabrata - Snails of the field isolate collected in September 2002, morphologically consistent with being *B. glabrata*, were identified as the species *B. glabrata* by PCR-RFLP (Fig. 1). Additionally, the *16S* rDNA (GenBank accession AY737280) and *NADH dehydrogenase 1* (ND1; accession AY737281) sequences from one BB02 snail were each 99% identical to previously characterized sequences from other *B. glabrata* isolates. Phylogenetic analysis based on these sequences placed the BB02 strain within the "B1" Brazilian clade of *B. glabrata* that was designated by DeJong et al. (2003). Bootstrap support was 85-98%, depending on the use of distance, maximum parsimony, or maximum likelihood methods. Phylogenetic trees are not shown; they were essentially identical to those presented in DeJong et al. (2003).

BB02 snails proved highly susceptible to three different strains of *S. mansoni*. At 4 weeks following experimental exposure, 89.6% or more of the snails harbored viable parasite infections (Table I).

Generation of the BG_BBa BAC library - The genomic DNA sample form 40 twice-selfed BB02 snails yielded sufficient quantity of HMW DNA (Fig. 2). The cloning of fragments ranging from 150-300 kb (partial *Hind*III digest) allowed isolation of 61,824 transformants, which were distributed over 161 plates with 384 wells. The BAC library was designated BG_BBa ("BG" is the first letter of genus and species, the "B" is for BB02 strain, the second "B" designates BAC library, "a" is the first library made).

TABLE I

BB02 *Biomphalaria glabrata*: susceptibility for different strains of *Schistosoma mansoni*

Snail	Parasite strain	Positive/negative/died	% susceptible
F1	LE	43/5/2	89.6
F1	SJ	40/4/6	90.9
Selfed F2	NMRI	47/1/2	97.9

Fifty BB02 snails, either F1 offspring of field collected snails or snails resulting from two generations of selfing were each exposed to 10 miracidia from different *S. mansoni* strains and checked for infection four weeks later. Snails that had died were not included in the calculation of % susceptible snails.

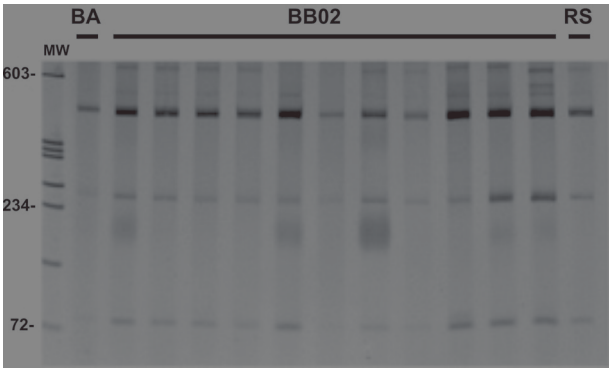

Fig. 1: molecular identification of BB02 strain snails as *Biomphalaria glabrata* through PCR-RFLP on a polyacrylamide gel. The *Dde*I digestion pattern of the ITS region (amplified by PCR) is the same for reference *B. glabrata* from Jacobina, state of Bahia, Brazil (BA), from Esteio, state of Rio Grande do Sul, Brazil (RS) and for BB02 *B. glabrata*. Markers indicated in bp.

Fig. 2: high molecular weight DNA from BB02 *Biomphalaria glabrata,* separated on a CHEF gel. MW markers (Lambda ladder from New England Biolabs) indicated in kb.

Properties of the BG_BBa BAC library - The average insert size observed from clones of the BG_BBa library was 136.3 kb (n=361). The distribution of different insert size categories is shown in Fig. 3. Over 90% (or 328) of the BACs had an insert size greater than 100 kb. No empty clones (vector without insert) were recorded. The BG_BBa library consists of 61824 clones with an average insert of 136.3 kb. This provides a 9.05-fold coverage of the genome of *B. glabrata* based on a size estimate of 931 Mb (Gregory 2003). The sequencing of BAC ends of 192 clones yielded 349 reads totaling 242270 nucleotides (nt; designated the AGI set). Contig analysis indicated that all of the sequences obtained from this random sample were unique. Some BAC inserts (1.4% of the total) shared highly similar sequences at one terminal end, but differed from each other on the other side of the insert. These BACs were BG_BBa0012A03 and BG_BBa0064N23 (GenBank accessions of the sequence reads from the termini are CZ548214 and CZ548008, respectively) displaying 12 differences over 770 nt; and BG_BBa0023A03 (CZ548090), BG_BBa0095B23 (CZ548268), BG_BBa0024B23 (CZ548158) that shared a 464 nt sequence (6 differences) that was highly similar to a transposable element (GenBank XP_791680).

All 10 low- or single copy probes hybridized with clones on the high density filters representing the complete BAC library. Verification of putative positives by colony hybridization (using macroarrays) identified some false positives but also confirmed the representation of each target sequence in the BAC library (see Table II).

Contig alignment of BACs by fingerprinting - Analysis of the multiple restriction patterns of 55 BAC clones provided data that were sorted into 13 contigs. Two of these contigs (numbers 2 and 8) combined BACs that had hybridized different probe sequences. In total, 11 different contigs provided alignment of BACs that had each hybridized with the same probe sequence. These assemblies revealed the relative position of several BAC clones within the genome of *B. glabrata* (Fig. 4, for all contigs see http://www.genome. arizona.edu/cgi-bin/ BAC_ special_proj).

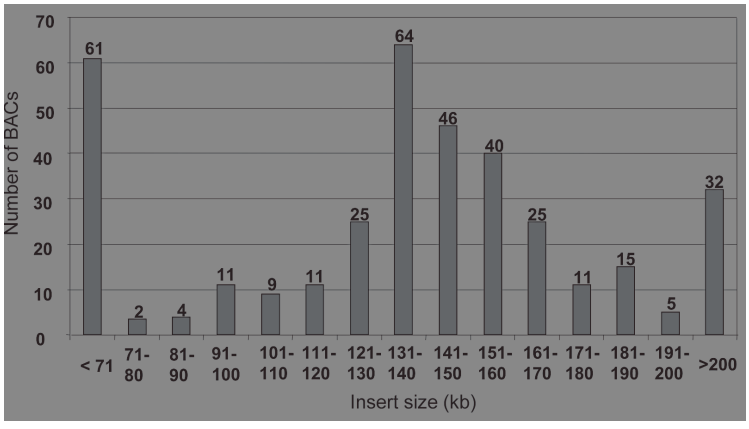

Fig. 3: bar graph showing the size distribution of insert size determined from 361 BAC clones. The number of clones with a particular insert size range is indicated on top of the bar. These data were used to calculate the average BAC insert size at 136.3 kb.

TABLE II
Representation of single/low copy genes in the BAC library

Gene	Probe	Rationale to consider as single/low copy	Positive clones
β *actin*	1037-1607 of AF329436	3-5 bands on Southern Adema 2002	45
FREP4	626 nt of AY012701 [a]	2-3 bands on Southern Zhang & Loker 2004	4
FREP13	182-661 of AF515468	1-2 bands on Southern Zhang & Loker 2004	23
HSP70	295-1014 of L44127	2-4 bands on Southern Laursen et al. 1997	13
MnSOD	317 - 640 nt AY500813	1-5 bands on Southern Jung et al. 2005	10
myoglobin	AY314976 (379 nt)	low sequence diversity Zhang et al. 2004	80
ferritin	AW739853 (492 nt)	arbitrary	2
L32 ribosomal protein	330 nt [b]	arbitrary	15
sialic acid binding lectin	CN445843 (318 nt)	arbitrary	3
ubiquitin	AF268491 (450 nt)	arbitrary	3

These sequences were selected as likely single or low copy genes of *Biomphalaria glabrata*. These were used to probe the BG_BBa BAC library to check for representation of less common gene sequences. Indicated are name of the source gene, the region used as probe (partial interval or full length of GenBank entries), rationale for assumption as likely single or low copy number gene (or arbitrary decision), the number of clones in the BAC library that hybridized the probe. *a*:only the coding sequences from the gene positions 921-6074 were amplified from cDNA. *b*: sequence amplified PCR using primers (5' -3') QCL32F AAA GAA ACA AGA AGT TTA TTC GC and QCL32R GGG ATT TGT CAC TTT AAT GGC.

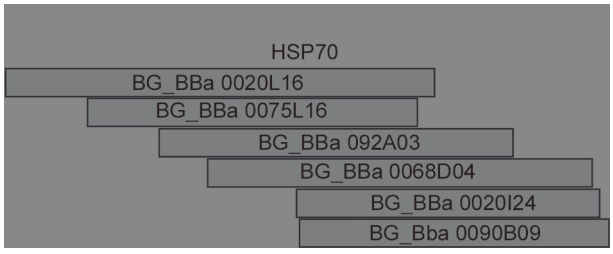

Fig. 4: fingerprinting by restriction enzyme digestion patterns allowed the contig alignment of BACs. This graphical depiction shows the relative position of six DNA inserts of BAC clones that each reacted positive with an HSP70-specific probe (see Table II). The aligned BACs represent a 234 kb (estimated) region of genomic DNA of *Biomphalaria glabrata.* The BAC inserts are identified by the BB_BGa clone number.

Analysis of sequence data from BAC ends - The end sequencing of the BAC clones used for macroarrays yielded another 165 sequence reads totaling 62402 nt (designated as the UNM set). These BACs were not chosen randomly; rather they were selected based on the first screening of high density filters with 2 different pools of 5 probes representing single- or low copy sequences as described above. Nevertheless, the majority of BAC end sequences recovered from this group was unique. In some cases, multiple clones had identical insert sequence from at least one of their termini; these were DX360154-DX360158 and DX360097- DX360114 (both BACs of the latter combination reacted with the FREP13 probe). Two sequence reads from BACs that bound the myoglobin probe (DX360049, DX360199) combined into a 639 nt contig, consistent with the grouping of these BAC inserts in "contig 5" resulting from the fingerprinting approach.

All the BAC end sequences are considered to derive from the nuclear genome of the snail because a direct sequence comparison showed no similarities with the mitochondrial genome of *B. glabrata* (NC_005439; DeJong et al. 2004). Thus, the base usage determined from these BAC end sequences (Table III) indicated that the nuclear genome of *B. glabrata* has an AT content of just over 63%.

BLAST analysis showed that the majority (430 or 83.7%) of the 514 BAC end sequences were novel; they did not display significant similarity with previously known sequences. However, several putative gene sequences were identified that had not been recorded previously from *B. glabrata*. Sequences from five BAC ends most resembled previous entries derived from *B. glabrata* (summarized in Table IV). Two segments of the contig of BAC end sequences DX360049 and DX360199 showed BLAST similarities to *glyceraldehyde-3-phosphate dehydrogenase* in a way that also revealed intron-exon structure, with splice sites that display the general GT-AG consensus (Fig. 5).

Several types of sequence repeats were observed directly from the BAC-derived sequences. These included simple dinucleotide microsatellites (e.g a string of TA repeats in DX360102), but also more complex sequence repeats such as CZ548151 that showed 18 almost exact repeats of "ACCCTGGTATGCCTTAGTGCTTGTATTGG". Furthermore, BLAST results indicated the presence of various transposable elements in 50 BAC end reads (9.7% of the total).

Remarkably, the genomic sequence collected here also contained an enigmatic type of high frequency sequence elements (HFSE). Stretches of between about 60 and 250 nucleotides embedded within almost 12% of the BAC end sequences showed significant similarity to (non protein-encoding) intron sequences of two separate kinds of *B. glabrata* genes for which full-length sequences have been provided previously; *myoglobin* and several *fibrinogen-related protein* genes (*FREP2.1*; *FREP3.1*; *FREP4*; *FREP6*, *FREP7.1*, *FREP13.1*). Thus similarities were detected only at nucleotide level (BLASTN), not at the de-

TABLE III

Base content of BAC end sequence data sets

Base	AGI 349 reads, 238576 nt [a] number (% of total)	UNM 165 reads, 60219 nt [a] number (% of total)	Combined 514 reads, 298795 nt [a] number (% of total)
A	76474 (32.05)	18605 (30.89)	95079 (31.82)
T	75823 (31.78)	18769 (31.16)	94592 (31.65)
C	43270 (18.13)	10586 (17.57)	53856 (18.02)
G	42973 (18.01)	10986 (18.24)	53959 (18.05)
N	36 (0.15)	1273 (2.08)	1309 (0.44)
	AT = 63.86%	AT = 62.05%	AT = 63.47%

Nucleotide content calculated for the BAC end sequence data from *Biomphalaria glabrata* (GenBank accession numbers CZ547921-CZ548269; DX360039-DX360203). Because of the use of different methods, results are also shown separately for sequence data from the Arizona Genomics Institute (AGI), the University of New Mexico (UNM). The number of A, T, C, G nucleotides is indicated, and the percent of the total number of nucleotides; N: unresolved nucleotides; *a*: number of non-redundant nucleotides, this excludes duplicate sequences.

TABLE IV

Sequence similarities within BAC end sequence reads

(Accession#)	BLAST similarity (e-values ≤ 10^{-4})
CZ547944	Similar to mitotic checkpoint protein
CZ547961	P1/eIF-2A protein kinase
CZ547969	*Biomphalaria glabrata* repetitive polymorphic DNA marker[a]
CZ547982	Similar to Ubiquitin-conjugating enzyme E2 D2
CZ547988	Multiple ankyrin repeats
CZ547993	tRNA synthetase
CZ548003	IAP-3
CZ548026	*Biomphalaria glabrata* repetitive polymorphic DNA marker[a]
CZ548030	Ring finger protein
CZ548065	Dynein, axonemal
CZ548081	Structure specific recognition protein 1
CZ548093	GTF2IRD2
CZ548107	Helicase SKI2W
CZ548120	High conductance calcium-activated potassium channel
CZ548136	Ubiquitin specific protease 20
CZ548153	GTF2IRD2
CZ548185	Delta-like 4 protein
CZ548186	General transcription factor II i repeat domain 2
CZ548197	CG10806-PB
CZ548200	Fatty acid synthase
CZ548212	GA19464-PA
CZ548217	Zinc finger protein
CZ548220	CG10244-PA
CZ548223	Glutaminyl-tRNA synthetase
CZ548231	*Biomphalaria glabrata* repetitive polymorphic DNA marker[a]
CZ548246	Nitric oxide synthase
DX360039	NAD(P) dependent steroid dehydrogenase-like
DX360045	Tubulin tyrosine ligase-like protein
DX360047	Thioredoxin
DX360081	Dermatopontin 3 (*Biomphalaria glabrata*)
DX360104	Glyceraldehyde-3-phosphate dehydrogenase
DX360123	MEGF10
DX360157	HSP70 (*Biomphalaria glabrata*)
DX360198	Xanthine dehydrogenase
DX360199	Glyceraldehyde-3-phosphate dehydrogenase

Accession numbers starting with CZ designate randomly chosen BACs, DX identifies BACs selected by the probes listed in Table II. Most results were derived from BLASTX; *a*: identifies BLASTN results.

Fig. 5: alignment of two BAC end sequences shows intron-exon structure of a partial *Biomphalaria glabrata* gene encoding *glyceraldehyde-3-phosphate dehydrogenase* (e-value of $1e^{-22}$ for BLASTX). Note that the boundaries of the introns display the GT-AG consensus for slice sites. Solid underline: DX360199 (nucleotides 1-271), dotted underline: DX3601049 (131-639), lower case: intron sequences, bold lower case: termini of intron sequences, upper case: exons, bold upper case: predicted amino acid sequence, hyphens indicate partial codons.

duced aminoacid level (BLASTX). A comparison disclosed that parts of the introns of *myoglobin* and *FREPs* share regions with considerable sequence similarities. These particular sequence regions generated a surprisingly high number of significant BLAST hits when compared against genomic and cDNA sequences of *B. glabrata* (Fig. 6). BLAST analysis showed that the different HFSE do not share significant similarities with known sequences from other organisms, either at nucleotide or deduced amino acid levels.

DISCUSSION

The development of genomic projects for *B. glabrata* will inform on the biology of molluscs in general. But of course, such an effort must also be considered in light of schistosomiasis. The medical importance of a better understanding of snail-parasite interactions and the possibility of gaining novel insights into the transmission of schistosomiasis by *B. glabrata* was part of the motivation for NHGRI to support the production of a high quality BAC library. Importantly, the NHGRI support also makes the library available at cost to the research community as a resource for the research community to enable advanced level study as a complement to ongoing research into the biology of *B. glabrata*, both in the laboratory and in the field.

Before generating the BAC library, it was of paramount importance to confirm that snails collected from the field were indeed *B. glabrata*. Two independent methods, PCR-RFLP analysis of nuclear sequences (Vidigal et al. 1998) and phylogenetic analysis of sequence segments from mitochondrial genes (DeJong et al. 2003) each confirmed that BB02 snails were *B. glabrata.* Likewise, different research groups tested and confirmed that BB02 *B. glabrata* is susceptible to infection by various strains of *S. mansoni*. Several research groups in the US, the UK and in Brazil now maintain the BB02 snails. This also minimizes the risk of accidental loss of the strain.

The isolation of sufficient amounts of good quality, HMW DNA from *B. glabrata* proved to be a challenge. Routine methods for DNA extraction from whole snail bodies yielded DNA fragments ≤ 50 kb, too low for generation of large inserts in a BAC library. The methods of Luo and Wing (2003), originally developed for plant tissues, did produce HWM DNA > 800 kb, but only from the ovotestis of *B. glabrata* and not in adequate amounts from a single snail. The additional snails needed were generated by selfing. This is possible because *B. glabrata* is a simultaneous hermaphrodite that can self fertilize (producing both male and female gametes) to produce offspring. Although this does not equate to cloning snails, the haplotype diversity of selfed progeny is thus lowered compared to that of outcrossing snails.

Once produced, the BG_BBa BAC library was found to exceed the quality standards set by NHGRI. At an average size of 136.3 kb, inserts of snail-derived sequence are large enough to accommodate several genes in genomic sequence context, even when considering the 16 kb size of *FREP*7.1 (AY028462) the largest sequence containing a full-length gene from *B. glabrata*, characterized to date. Non-transformed clones were not observed and the proportion of empty vectors (no insert) is considered negligible. Furthermore, the quality control showed that the inserts are diverse in size and sequence content. At the same time, each of 10 low copy number gene sequences of *B. glabrata* was present. The recovery of HSP70-encoding sequence (DX360157) confirmed the integrity of the library and showed that genes can be screened for and recovered for additional analysis. The contig assembly based on fingerprinting further demonstrated the con-

```
Myoglobin    tggatggctgcctggtcgtgc-ggtttgagcgctggactgtcgtttggat
FREP2.1      ....................c......c..............c....
FREP4        -----........a..ga...-a.....c..t.............c.a.c
FREP6        ----.ta..............-a.....c..............c....
FREP7.1      -..t.................-......c..............c....
CZ548223     ----.................-......ca.............ca...
CZ547960     ----................t-......c.....................
AW740133     -.........................c..............c....
DW474487     ---------------...........c.c..............c....

Myoglobin    ttatcgatggtgccgggttcgaacc-tgcccgctcccatccccc-gtcgt
FREP2.1      ..t....c...cga......a....c.................a-....
FREP4        ...........c.aa.....a..t.t.tt..a...........t-...c.
FREP6        ..t....c...cga......a....c.................a--....
FREP7.1      .......c...cga......a....c....a...t.........c.....
CZ548223     ...........c........a....c.................-.....
CZ547960     ........a...........a....c....a.............-.....
AW740133     .......c...cga......a....c.................-.....
DW474487     .......c...caa......a....c.................-.....

Myoglobin    cctgcggg-agatttagactaggaagtatattatttcctactctgaaggg
FREP2.1      ........a..g...g............a.c...c.t.a..........a
FREP4        ........-..gg..g............a.....c.t-------------
FREP6        .-.......-..g...g............a.c...c.t.a.........ra
FREP7.1      .....a..-..g...g............a.....c.t.a..........a
CZ548223     ........-..g...g............a-....c.t.a..........a
CZ547960     .......a-..g...g............a.c...c.t.a..........a
AW740133     ........-..g...g.....t......a.c...c.t.a.........aa
DW474487     ...........g...g............a.c...c.t.a..........a

Myoglobin    acatccgaaacatgtaaaaca
FREP2.1      ...................t.
FREP4        ---------------------
FREP6        -....................
FREP7.1      ..............c.....
CZ548223     .....................
CZ547960     ...................---
AW740133     .....................
DW474487     ...................--
```

Fig. 6: the genome of *Biomphalaria glabrata* contains high frequency sequence elements (HFSE) that occur in BAC end sequences, introns of previously characterized genes encoding fibrinogen-related genes (*FREPs*) and *myoglobin*, as wells as other genomic and cDNA sequence entries for this snail. The alignment shows an example (more exist) of one group of HFSE present in introns of *myoglobin* (GenBank accession BGU89283; nucleotide residues 150-324) and several *FREP* genes: FREP2.1 (AY012700, 4863-5030), FREP4.1 (AY012701; 3644-3769), *FREP6* (AY012702; 690-847), *FREP7.1* (AY028462; 1648-1812), and in BAC end sequences CZ548223 (67-226), CZ547960 (492-652) and ESTs AW740133 (110-281) DW474487 (215-366) collected from *B. glabrata.* Identity is indicated by dots, hyphens represent gaps in the alignment.

sistency of retrieving relevant, related sequences in the genome context from the BAC library. The apparent diversity of BACs (failure to form a single contig) recovered after screening with the *myoglobin* sequence must be considered against a rapid increase of (hemo)globin-like EST sequences in GenBank (85 entries in February 2006). *Myoglobin* may be a low copy sequence, but the abundance of now available information suggests a high likelihood for cross-reactivity with sequences from other genes.

Already, the BAC library has provided new information on the *B. glabrata* genome. The BAC end sequence data not only identified several genes, but as shown in Fig. 5, also provided information regarding intron-exon structure of genes. The latter is relevant since most *B. glabrata* genes are only characterized from cDNA. The large amount of BAC end sequence data indicated that the genome of *B. glabrata* has a 63% AT content. This is higher than previously estimated, based on the reported 54% AT content for the related *B. alexandrina* (Nabih & El Ansary 1980). However, it is not as high as the 74.6% AT content of the mitochondrial genome of *B. glabrata* (DeJong et al.2003).

The sequence data provided an indication of abundance of mobile elements: these occurred in 9.7% of the BAC ends characterized. This sampling is incomplete, but 9.7% is low compared to the genome of *S. mansoni* that may consist of 50% mobile elements (Laha et al. 2005). Another surprise were the groups of HFSEs, embedded within BAC end sequences, introns of *myoglobin* and *FREP* genes and shared with cDNA sequences (Fig. 6). Perhaps HFSEs are a peculiarity of the genome of *B.*

glabrata, possibly these are transcriptionally active regions that function in regulation of gene expression (e.g. Claverie 2005, Mattick 2005). Likely, elucidation of the origin and function of these will benefit from continued genome level investigations of *B. glabrata*.

In conclusion, a high quality BAC library for *B. glabrata* is now available. This helps to keep pace with similar research developments for *S. mansoni* such that this parasite-snail interaction can be analyzed at the genomics level. Because the current full genome sequencing project also employs the BB02 strain *B. glabrata* (http://genome.wustl.edu/genome.cgi?GENOME=Biomphalaria%20glabrata), the usefulness of the BG_BBa BAC library as a research resource will continue to increase by making the *B. glabrata* genome accessible to study snail biology in novel ways.

ACKNOWLEDGMENTS

Ulrike Zelck (Tûbingen University, Germany) provided helpful discussion for selection of low copy genes. At the University of Arizona, Chris Mueller, Kristi Collura, Nick Sisneros, Marina Wissotski, Dan Smart, David Campos and Kiran Rao provided excellent technical assistance. At the University of New Mexico, George Rosenberg assisted with sequencing of BACs, technical support was provided by the Molecular Biology Facility, which is supported by NIH grant number 1P20RR18754 from the Institute Development Development Award (IDeA) Program of the National Center for Research Resources.

REFERENCES

Adema CM 2002. Comparative study of cytoplasmic actin DNA sequences from six species of Planorbidae (Gastropoda: Basommatophora). *J Molluscan Stud 68*: 17-23.

Adema CM, Loker ES 1997. Specificity and immunobiology of larval digenean-snail associations. In B Fried, TK Graczyk (eds), *Advances in Trematode Biology*, CRC Press, Boca Raton, FL, p. 230-263.

Altelaar AF, van Minnen J, Jimenez CR, Heeren RM, Piersma SR 2005. Direct molecular imaging of *Lymnaea stagnalis* nervous tissue at subcellular spatial resolution by mass spectrometry. *Anal Chem 77*: 735-741.

Chou HH, Holmes MH 2001. DNA sequence quality trimming and vector removal. *Bioinformatics 17*: 1093-1104.

Claverie J-M 2005 Fewer Genes, More Noncoding RNA. *Science 309*: 1529-1530.

Collins FS, Green ED, Guttmacher AE, Guyer MS 2003. A vision for the future of genomics research. *Nature 422*: 835-847.

Chitsulo L, Loverde R, Engels D, Barakat R, Colley D, Cioli D, Engels D, Feldmeier H, Loverde P, Olds GR, Ourna J, Rabello A, Savioli L, Traore M, Vennerwald B 2005. Schistosomiasis. *Nat Rev Microbiol 2*: 12-13.

DeJong RJ, Emery AM, Adema CM 2004. The mitochondrial genome of *Biomphalaria glabrata* (Gastropoda: Basommatophora), intermediate host of *Schistosoma mansoni*. *J Parasitol 90*: 991-997.

DeJong RJ, Morgan JAT, Wilson WD, Al-Jaser MH, Appleton CC, Coulibaly G, D'andrea PS, Doenhoff MJ, Haas W, Idris MA, Magalhães LA, Moné H, Mouahid G, Mubila L, Pointier J-P, Webster JP, Zanotti-Magalhães EM, Paraense WL, Mkoji GM, Loker ES 2003. Phylogeography of *Biomphalaria glabrata* and *B. pfeifferi*, important intermediate hosts of *Schistosoma mansoni* in the New and Old World tropics. *Mol Ecol 12:* 3041-3056.

Ewing B, Green P 1998.Base-calling of automated sequencer traces using Phred: II. Error probabilities. *Genome Res 8*: 186-194.

Ewing B, Hillier L, Wendl MC, Green P 1998. Base-calling of automated sequencer traces using Phred: I. Accuracy assessment. *Genome Res 8*: 175-185.

Gregory TR 2003. Genome size estimates for two important freshwater molluscs, the zebra mussel (*Dreissena polymorpha*) and the schistosomiasis vector snail (*Biomphalaria glabrata*). *Genomics 46*: 841-844.

Jokela J, Lively CM, Dybdahl MF, Fox JA 2003. Genetic variation in sexual and clonal lineages of a freshwater snail. *Biol J Linn Soc Lond 79*: 165-181.

Jones CS, Lockyer AE, Rollinson D, Noble LR 2001. Molecular approaches in the study of *Biomphalaria glabrata* - *Schistosoma mansoni* interactions: linkage analysis and gene expression profiling. *Parasitology 123* (Suppl. S): S181-S196.

Jung Y, Nowak TS, Zhang S-M, Hertel LA, Loker ES, Adema CM 2005 Manganese superoxide dismutase from *Biomphalaria glabrata*. *J Invertebr Pathol 90*: 59-63.

Laha T, Kewgrai N, Loukas A, Brindley PJ 2005 Characterization of SR3 reveals abundance of non-LTR retrotransposons of the RTE clade in the genome of the human blood fluke, *Schistosoma mansoni*. *BMC Genomics 6*: 154.

Laursen JR, di Liu H, Wu XJ, Yoshino TP 1997 Heat-shock response in a molluscan cell line: characterization of the response and cloning of an inducible HSP70 cDNA. *J Invertebr Pathol 70*: 226-233.

Lewis FA, Patterson CN, Knight M, Richards CS 2001. The relationship between *Schistosoma mansoni* and *Biomphalaria glabrata*: genetic and molecular approaches. *Parasitology 123* (Suppl. S): S169-S179.

Lie KJ 1982. Survival of *Schistosoma mansoni* and other trematode larvae in the snail *Biomphalaria glabrata* - a discussion of the interference theory. *Trop Geogr Med 34*: 111-122.

Lockyer AE, Jones CS, Noble LR, Rollinson D 2004a Trematodes and snails: an intimate association. *Can J Zool 82*: 251-269.

Lockyer AE, Noble LR, Rollinson D, Jones CS 2004b. *Schistosoma mansoni*: resistant specific infection-induced gene expression in *Biomphalaria glabrata* identified by fluorescent-based differential display. *Exp Parasitol 107*: 97-104.

LoVerde PT, Hirai H, Merrick JM, Lee NH, El-Sayed N 2004. *Schistosoma mansoni* genome project: an update. *Parasitol Int 53*: 183-192.

Luo M, Wing RA 2003. An improved method for plant BAC library construction. In E Grotewold, *Plant Functional Genomics: Methods and Protocols*, Humana Press, Totowa, NJ, p. 3-19.

Luo M, Thomas C, You FM, Hsiao J, Ouyang S, Buell CR, Malandro M, McGuire PE, Anderson OD, Dvorak J 2003.

High-throughput fingerprinting of bacterial artificial chromosomes using the snapshot labeling kit and sizing of restriction fragments by capillary electrophoresis. *Genomics 82*: 378-389.

Mattick JS 2005. The Functional Genomics of Noncoding RNA. *Science 309*: 1527-1528.

Milet C, Berland S, Lamghari M, Mouries L, Jolly C, Borzeix S, Doumenc D, Lopez E, 2004. Conservation of signal molecules involved in biomineralisation control in calcifying matrices of bone and shell Source. *C R Palevol 3*: 493-501.

Miller AN, Raghavan N, FitzGerald PC, Lewis FA, Knight M 2001. Differential gene expression in haemocytes of the snail *Biomphalaria glabrata*: effects of *Schistosoma mansoni* infection. *Int J Parasitol 31*: 687-696.

Mitta G, Galinier R, Tisseyre P, Allienne JF, Girerd-Chambaz Y, Guillou F, Bouchut A, Coustau C 2005. Gene discovery and expression analysis of immune-relevant genes from *Biomphalaria glabrata* hemocytes. *Dev Comp Immunol 29*: 393-407.

Mizi A, Zouros E, Moschonas N, Rodakis GC 2005. The complete maternal and paternal mitochondrial genomes of the Mediterranean mussel *Mytilus galloprovincialis*: implications for the doubly uniparental inheritance mode of mtDNA. *Mol Biol Evol 22*: 952-967.

Morgan JAT, DeJong RJ, Snyder SD, Mkoji GM, Loker ES 2001. *Schistosoma mansoni* and *Biomphalaria*: past history and future trends. *Parasitology 123* (Suppl. S): S211-S228.

Nabih I, El Ansary A 1980. Genetic-studies on fresh-water snails specific intermediate hosts for schistosomiasis .2. Isolation and base composition determination of deoxyribonucleic-acid. *Cell Mol Biol 26*: 455-458.

Nowak TS, Woodards AC, Jung Y, Adema CM, Loker ES 2004. Identification of transcripts generated during the response of resistant *Biomphalaria glabrata* to *Schistosoma mansoni* infection using suppression subtractive hybridization. *J Parasitol 90*: 1034-1040.

Paraense WL 1986. Distribuição dos caramujos no Brasil. In FA Reis, I Faria, N Katz (eds), *Modernos Conhecimentos sobre Esquistossomose Mansônica*, Suplemento dos Anais 1983/84, vol. 14, Academia Mineira de Medicina, Belo Horizonte, p. 117-128.

Paraense WL, Corrêa L 1963. Variation in susceptibility of populations of Australorbis glabratus to a strain of *Schistosoma mansoni*. *Rev Inst Med Trop Sã Paulo 5*: 15-22.

Paraense WL, Corrêa LR 1988. Self-fertilization in the freshwater snails *Helisoma duryi* and *Helisoma trivolvis*. *Mem Inst Oswaldo Cruz 83*: 405-410.

Pointier JP, David P, Jarne P 2005. Biological invasions: the case of planorbid snails. *J Helminthol 79*: 249-256.

Raghavan N, Miller AN, Gardner M, FitzGerald PC, Kerlavage AR, Johnston DA, Lewis FA, Knight M 2003. Comparative gene analysis of *Biomphalaria glabrata* hemocytes pre- and post-exposure to miracidia of *Schistosoma mansoni*. *Mol Biochem Parasitol 126*: 181-191.

Rouse GW 1999. Trochophore concepts: ciliary bands and the evolution of larvae in spiralian metazoa. *Biol J Linn Soc Lond 66*: 411-464.

Schilthuizen M, Davison A 2005. The convoluted evolution of snail chirality. *Naturwissenschaften 92*: 504-515.

Schneider O, Zelck UE 2001. Differential display analysis of hemocytes from schistosome-resistant and schistosome-susceptible intermediate hosts. *Parasitol Res 87*: 489-491.

Soderlund C, Humphrey S, Dunhum A, French L 2000. Contigs built with fingerprints, markers and FPC V4.7. *Genome Res 10*: 1772-1787.

Terlau H, Olivera BM 2004. Conus venoms: A rich source of novel ion channel-targeted peptides. *Physiol Rev 84*:41-68.

Theron A, Coustau C 2005. Are *Biomphalaria* snails resistant to *Schistosoma mansoni*? *J Helminthol 79*: 187-191.

Vergote D, Bouchut A, Sautiere PE, Roger E, Galinier R, Rognon A, Coustau C, Salzet M, Mitta G 2005. Characterisation of proteins differentially present in the plasma of *Biomphalaria glabrata* susceptible or resistant to *Echinostoma caproni*. *Int J Parasitol 35*: 215-224.

Vermeij GJ 2002. Characters in context: molluscan shells and the forces that mold them. *Paleobiology 28*: 41-54.

Vidigal TH, Spatz L, Nunes DN, Simpson AJ, Carvalho OS, Dias Neto E 1998. *Biomphalaria* spp: identification of the intermediate snail hosts of *Schistosoma mansoni* by polymerase chain reaction amplification and restriction enzyme digestion of the ribosomal RNA gene intergenic spacer. *Exp Parasitol 89*: 180-187.

Williamson R, Chrachri A 2004. Cephalopod neural networks. *Neurosignals 13*: 87-98.

Zhao X, Zheng M, Liang L, Zhang Q, Wang Y, Jiang G 2005. Assessment of PCBs and PCDD/Fs along the Chinese Bohai Sea coastline using mollusks as bioindicators. *Arch Environ Contam Toxicol 49*: 178-185.

Zhang S-M, Loker ES 2004. Representation of an immune responsive gene family encoding fibrinogen-related proteins in the freshwater mollusc *Biomphalaria glabrata*, an intermediate host for *Schistosoma mansoni*. *Gene 341*: 255-266.

Zhang S-M, Adema CM, Kepler TB, Loker ES 2004. Diversification of Ig superfamily genes in an invertebrate. *Science 305*: 251-254

Zhurov Y, Proekt A, Weiss KR, Brezina V 2005. Changes of internal state are expressed in coherent shifts of neuromuscular activity in *Aplysia* feeding behavior. *J Neurosci 25*: 1268-1280.