



ITS2 variability of *Biomphalaria* (Mollusca, Planorbidae) species from the Paranapanema Valley (São Paulo State, Brazil): Diversity patterns, population structure, and phylogenetic relationships

Roseli Tuan and Priscila dos Santos

Laboratório de Bioquímica e Biologia Molecular, Superintendência de Controle de Endemias, Secretaria da Saúde do Estado de São Paulo, São Paulo, SP, Brazil.

Abstract

The ribosomal DNA internal transcribed spacer 2 (ITS2) has been shown to be a useful genetic marker for species identification and phylogenetic reconstruction in the genus of freshwater snails *Biomphalaria* (Preston 1910). Additionally, ITS2 studies in *Biomphalaria* have uncovered significant intra-specific genetic variability suggesting the presence of cryptic species complexes. We obtained ITS2 sequences for the *Biomphalaria* species *B. glabrata*, *B. tenagophila*, *B. occidentalis* and *B. peregrina* from the Paranapanema Valley (São Paulo State, Brazil) and compared them with a comprehensive set of published *Biomphalaria* ITS2 sequences using Bayesian inference of phylogeny. Analysis of the resulting trees showed that the newly obtained *B. glabrata* sequences did not cluster with those from other Brazilian localities and that sub-structuring occurred among Brazilian *B. tenagophila* populations. Moreover, although ITS2 sequences seem to indicate clear genetic differentiation within both *B. glabrata* and *B. tenagophila*, evidence in support of the occurrence of cryptic species is more compelling for the latter. We discuss the significance and implications of the detected patterns of ITS2 variability for taxonomic studies in *Biomphalaria*.

Key words: Bayesian inference, *Biomphalaria*, DNA sequences, internal transcribed spacer 2 (ITS2), ribosomal DNA.

Received: December 13, 2005; Accepted: August 9, 2006.

Species of planorbid snails belonging to the genus *Biomphalaria*, which includes the vectors of the blood fluke *Schistosoma mansoni*, have recently received increasingly attention in molecular systematic and phylogeography studies (Campbell *et al.*, 2000; Vidigal *et al.*, 2000, 2004; DeJong *et al.*, 2001; Mavárez *et al.*, 2002; Pointier *et al.*, 2005). These studies have largely supported the pivotal body of work on the taxonomy of Neotropical *Biomphalaria* carried out by malacologist Lobato Paraense (Paraense, 2001) and, additionally, uncovered interesting patterns of intraspecific variability (Mavárez *et al.*, 2002; Vidigal *et al.*, 2004).

Among the genetic markers used so far in studies on *Biomphalaria*, the nuclear ribosomal DNA (rDNA) internal transcribed spacers (ITS) have been shown to be particularly useful for species identification and phylogenetic reconstruction (Campbell *et al.*, 2000; Vidigal *et al.*, 2000; Pointier *et al.*, 2005). Furthermore, ITS sequence studies have revealed and/or confirmed sub-structuring within widely distributed species such as *B. glabrata* and *B.*

tenagophila (Mavárez *et al.*, 2002; Vidigal *et al.*, 2004), adding to the body of evidence indicating that these nominal taxa may constitute complexes of cryptic species (Woodruff and Mulvey, 1997; Spatz *et al.*, 1999). In these circumstances, the investigation of genetic variability patterns across the range of distribution of these species takes a prominent role in helping to elucidate how many independently evolving entities there are and which geographical areas they occupy.

In this study, we obtained nuclear ribosomal DNA (rDNA) internal transcribed spacer 2 (ITS2) sequences from *B. glabrata*, *B. tenagophila*, *B. occidentalis* and *B. peregrina* specimens collected in the Paranapanema Valley region (São Paulo State, Brazil). The newly obtained sequences were compared to all other ITS2 sequences published to date for the same species. In doing so, we addressed the following issues: polymorphism and divergence patterns, phylogenetic relationships, and the possible occurrence of species complexes in *Biomphalaria*.

We collected 16 *Biomphalaria* snails (Table 1) along the banks of different streams of the Paranapanema River in Ourinhos (lat -22.98, lon -49.87) and Ipaucu (lat -23.05, lon -49.62) and extracted total genomic DNA from the foot tissue of individual specimens using DNeasy Tissue Kits

Table 1 - *Biomphalaria* rDNA ITS2 sequences used in the study.

Species and localities where collected	Sequence code	ITS2 GenBank accession numbers
<i>B. glabrata</i>		
Ourinhos (SP), Brazil	<i>BglOur 1 BglOur 2 BglOur 3</i>	DQ143952 DQ143953 DQ143954
Esteio (RS), Brazil	<i>BglEst 1 BglEst 2</i>	AF198660 ¹ AF198661 ¹
Belo Horizonte (mg), Brazil	<i>BglBH</i>	AF449582 ³
Salvador (BA), Brazil	<i>BglSal 1 BglSal 2</i>	AY030376 ² AF449581 ³
Belém (PA), Brazil	<i>BglBel</i>	AF198659 ¹
Aragua, Venezuela	<i>BglAra</i>	AY030377 ²
Calicanto, Venezuela	<i>BglCal</i>	AF449589 ³
Caripe, Venezuela	<i>BglCar</i>	AF449583 ³
Chuao, Venezuela	<i>BglChu</i>	AF449588 ³
Los Naranjos, Venezuela	<i>BglLN 1 BglLN 2</i>	AF449584 ³ AF449585 ³
Portuguesa, Venezuela	<i>BglPor</i>	AF198662 ¹
Punta Cabito, Venezuela	<i>BglPC</i>	AF449591 ³
Tres Lances, Venezuela	<i>BglTL 1 BglTL 2</i>	AF449586 ³ AF449587 ³
Zuata, Venezuela	<i>BglZua</i>	AF449590 ³
Pierrot, Saint-Lucia	<i>BglPie</i>	AF449596 ³
Maupéou, Martinique	<i>BglMau</i>	AF449593 ³
Grand Camp, Guadeloupe	<i>BglGC</i>	AF449595 ³
Vallet, Guadeloupe	<i>BglVal</i>	AF449592 ³
Vieux-Fort, Guadeloupe	<i>BglVF</i>	AF449594 ³
Rio Grande Town, Puerto Rico	<i>BglRG</i>	AY030375 ²
Jarabacoa, Dominican Republic	<i>BglJar</i>	AY030374 ²
<i>B. tenagophila</i>		
Ourinhos (SP), Brazil	<i>BttOur 1-2 BttOur 3 BttOur 4 BttOur 5</i>	DQ143955 DQ143956 DQ143957 DQ143958
Ipauçu (SP), Brazil	<i>BttIpa 1-4</i>	DQ143959
Imbé (RS), Brazil	<i>BttImb</i>	AF198655 ¹
Taim (RS), Brazil	<i>BttTai</i>	AY631860 ³
Contagem (MG), Brazil	<i>BttCon</i>	AF449614 ³
Vespasiano (MG), Brazil	<i>BttVes</i>	AF198654 ¹
Formosa (GO), Brazil	<i>BttFor 1 BttFor 2</i>	AF198656 ¹ AY030388 ²
Itamaraju (BA), Brazil	<i>BttIta</i>	AY631859 ³
Asunción, Paraguay	<i>BttAsu</i>	AY030387 ²
Araza, Argentina	<i>BttAra</i>	AY425746 ⁴
Bonpland, Argentina	<i>BttBon</i>	AY425745 ⁴
Kinshasa, Democratic Republic of Congo	<i>BttKin</i>	AY631861 ³
<i>B. t. guaibensis</i>		
Esteio (RS), Brazil	<i>BtgEst</i>	AY425749 ⁴
Guaíba (RS), Brazil	<i>BtgGua</i>	AY425750 ⁴
Termas, Uruguay	<i>BtgTer</i>	AY425751 ⁴
<i>B. occidentalis</i>		
Ourinhos (SP), Brazil	<i>BocOur 1-2</i>	DQ143960
Caceres (MG), Brazil	<i>BocCac</i>	AY030389 ²
Capetinga (MG), Brazil	<i>BocCap</i>	AF198658 ¹
Campo Grande (MS), Brazil	<i>BocCG</i>	AF198657 ¹
Triângulo, Argentina	<i>BocTri</i>	AY425748 ⁴
Villa Chica, Argentina	<i>BocVC</i>	AY425747 ⁴
<i>B. peregrina</i>		
Ourinhos (SP), Brazil	<i>BpeOur 1 BpeOur 2</i>	DQ143961 DQ143962
Taim (RS), Brazil	<i>BpeTai</i>	AF198677 ¹
Alfenas (MG), Brazil	<i>BpeAlf</i>	AF198676 ¹
Bom Jesus da Penha (MG), Brazil	<i>BpeBJ</i>	AF198678 ¹
Nova Lima (MG), Brazil	<i>BpeNL</i>	AY030401 ²
San Antonio, Uruguay	<i>BpeSA</i>	AY030400 ²

¹Vidigal *et al.* (2000); ²DeJong *et al.* (2001); ³Mavárez *et al.* (2002); ⁴Vidigal *et al.* (2004); ⁵Pointier *et al.* (2005).

Acronyms for Brazilian states shown between parentheses: BA = Bahia; GO = Goiás; mg = Minas Gerais; PA = Pará; RS = Rio Grande do Sul; and SP = São Paulo.

Bayesian majority rule consensus tree obtained for this data set, using the parameters form of the GTR + G model, as recommended by MrModeltest 2.0. The ML analysis produced two trees of $-\ln L = 1610.16$, differing minimally from one another: tree # 1 had *Bg*/BH sister to *Bg*/Est 1/2 and tree # 2 displayed the same topology for the *B. glabrata* clade as the Bayesian one. Both ML trees also differed from the Bayesian one by a minor variation in the *B. peregrina* clade (*Bpe*BJ sister to *Bpe*Tai/*Bpe*SA). Evaluation of clade support for the ML trees via bootstrapping was precluded by computing limitations. The MP analysis yielded an incalculable number of trees of length 160 (CI = 0.863, RI = 0.977), but the strict consensus of the first 100,000 retained trees produced the same topology as the Bayesian tree, except that *Bg*/Est 1 and *Bg*/Est 2 did not cluster together. MP bootstrap values (maximum number of trees retained per iteration set to 200) are indicated in Figure 2.

All *B. glabrata* sequences were recovered in a monophyletic clade, containing a set of defined branches denominated V1, V2 and B in Figure 2. These branches were also characterized by Mavárez *et al.* (2002), using concatenated ITS2 and mitochondrial 16S rDNA sequences. Venezuela harbors two distinct ITS2 lineages (V1 and V2) and, interestingly, the Venezuelan sample reported in Vidigal *et al.* (2000) clearly belongs to subgroup V2, showing that this clade is more widespread than previously thought. The Brazilian clade (B), also reported by Mavárez *et al.* (2002) for two sequences (*Bg*/BH and *Bg*/Sal 2), contained all other Brazilian sequences published to date (*Bg*/Sal 1, DeJong *et al.*, 2001; *Bg*/Est 1-2, *Bg*/Bel, Vidigal *et al.*, 2000) except the ones from Ourinhos (*Bg*/Our 1-3), these latter sequences appearing in an unstructured, more basal, region of the *B. glabrata* clade, along with the sequences from Central America (*Bg*/Pie, *Bg*/Mau, *Bg*/GC, *Bg*/Val, *Bg*/VF, *Bg*/RG, *Bg*/Jar).

In contrast to *B. glabrata*, the clade containing *B. tenagophila*, *B. t. guaibensis*, and *B. occidentalis* sequences displayed a polyphyletic pattern (the *B. tenagophila* complex, Spatz *et al.*, 1999; Vidigal *et al.*, 2004). The *B. tenagophila* sequences were observed in two different branches (T1 and T2). The two *B. occidentalis* samples from Argentina (*Boc*Tri and *Boc*VC) also appeared on separate branches, one clustered with *B. t. guaibensis* sequences (G/O), the other with the remaining *B. occidentalis* (O). However, unlike in Vidigal *et al.* (2004), support for joint monophyly of branches G/O and O was not obtained.

Interestingly, sub-structuring within the *B. tenagophila* T1 branch was observed, extending the differentiation characterized for the Ourinhos/Ipauçu data set to other samples (T1 groups a and b). Additionally, some un-derived *B. tenagophila* sequences were recovered, in the same fashion of the *B. glabrata* clade. In fact, sub-structuring within *B. tenagophila* T1 makes the nucleotide diversity in this branch only ($\pi = 0.893\% \pm 0.232$) comparable to that found in the entire *B. glabrata* data set

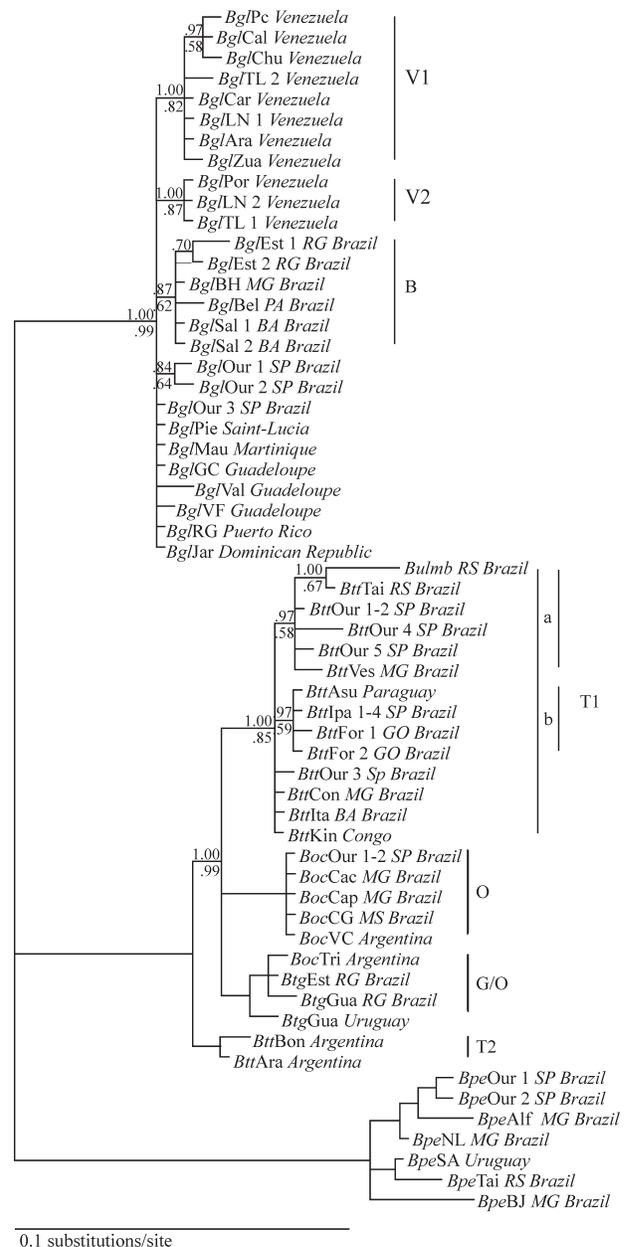


Figure 2 - Bayesian phylogenetic tree for *Biomphalaria* ITS2 sequences. Values above and below nodes indicate clade posterior probabilities and maximum parsimony (MP) bootstrap proportions, respectively.

($\pi = 0.729\% \pm 0.084$). Divergence between branches within *B. glabrata* and *B. tenagophila* T1 is also similar (e.g. $D_{V1-V2} = 1.277\% \pm 0.338$, $D_{T1a-b} = 1.215\% \pm 0.434$). However, when the entire *B. tenagophila* complex is considered it becomes apparent that not only is it comprised of highly diverse sequences ($\pi = 1.604\% \pm 0.176$) but also that divergence within the nominal *B. tenagophila* is far larger ($D_{T1-T2} = 2.029\% \pm 0.532$) than within the nominal *B. glabrata* (for which D_{V1-V2} gives the highest pairwise value). Finally, the *B. peregrina* clade also encompasses a set of divergent sequences, which is reflected in a remark-

ably high level of nucleotide diversity ($\pi = 1.823\% \pm 0.412$).

Whereas cross validation with independent biological markers remains a critical step in assessing the nature and significance of ITS2 variability in *Biomphalaria*, the patterns of genetic differentiation revealed in ITS2 studies nonetheless provide significant information about barriers to gene flow. Thus, it is remarkable to observe fixed differences between *B. tenagophila* samples from Ourinhos and Ipaucu, localities no more than 30 km apart, and this finding may indicate low connectivity between these sites and deserves further investigation. It is also noteworthy that *B. glabrata* sequences from Ourinhos did not cluster with the remaining samples from Brazil, and in this case additional sampling within São Paulo and other Brazilian states will be needed to elucidate the significance of this finding on a broader phylogeographical scale. Nonetheless, we note that in both cases the low magnitude of change ($\pi < 1\%$) seems to indicate these are likely to be examples of intraspecific polymorphism. Actually, a cohesive status for *B. tenagophila* T1 and *B. glabrata* samples is suggested not only by the magnitude, but also by the pattern of distribution of ITS2 variability. The latter is apparent in the structure of both clades, which exhibit a few distinct but short branches and a set of undifferentiated sequences displayed in a “rake” fashion – the typical pattern of continuous sequence variability characteristic of single species (Crandall and Templeton, 1993). We investigated this pattern further by examining statistical parsimony networks obtained with the program TCS (Clement *et al.*, 2000) separately for *B. glabrata* and *B. tenagophila* T1 sequences. In both cases, the most undifferentiated (zero length branches) sequences (*Bgl*Our 3, *Bgl*Pie, *Bgl*Mau, *Bgl*GC, *Bgl*RG and *Bgl*Jar for *B. glabrata*, *Btt*Con, *Btt*Ita and *Btt*Kin for *B. tenagophila* T1) occupied central (ancestral) positions in the networks, holding most of the mutational connections with other sequences (data not shown). One direct implication of this finding is that, considering the presumably recent introduction of *B. tenagophila* into Africa (Pointier *et al.*, 2005), ITS2 genotyping would offer limited power for source estimation of the Kinshasa population – the use of multiple and/or more sensitive genetic markers would be required in this case. However, since the reliability of network inferences (in particular the assignment of ancestral sequences) depends on dense populational sampling (Crandall and Templeton, 1993), and is, in general, an inadequate tool for use on data sets obtained mainly for taxonomic studies, these results must be regarded with caution, solely as preliminary conjectures to be tested with additional data.

When the *B. glabrata* clade was compared with the entire *B. tenagophila* complex clade, however, a more intricate picture emerged: *vis-à-vis*, the *B. glabrata* clade was shallower with less differentiated branches than the *B.*

tenagophila complex clade (Figure 2). Here, both the degree and the pattern of differentiation of *B. tenagophila* T1 and T2 branches supply compelling evidence that these samples may belong to distinct, yet cryptic, species.

Finally, in our reappraisal of the ITS2 data it was interesting to identify that the highest level of sequence diversity among all the species analyzed was for *B. peregrina*. Historically, species of greater epidemiological importance such as *B. glabrata* have received more attention in population genetic studies, and their respective levels of genetic diversity, therefore, emphasized (*B. glabrata* was once called “the most differentiated mollusc species yet described”, Woodruff and Mulvey, 1997). However, as data from multi-species studies accumulates it becomes evident that several Neotropical *Biomphalaria* species (such as *B. tenagophila* and *B. peregrina*) show remarkably high levels of genetic variability. This is a relevant finding from both epidemiological and evolutionary viewpoints and supports, for instance, the hypothesis of a Neotropical origin for the genus (Campbell *et al.*, 2000; DeJong *et al.*, 2001). Moreover, it highlights the need for further studies aimed at investigating patterns of population structure and species limits for these taxa.

Acknowledgments

Funding for this work was provided by the Brazilian agency FAPESP (Fundação de Amparo à Pesquisa do Estado de São Paulo), grant 03/00039-2.

References

- Campbell G, Jones CS, Lockyer AE, Hughes S, Brown D, Noble LR and Rollinson D (2000) Molecular evidence supports an african affinity of the neotropical freshwater gastropod, *Biomphalaria glabrata*, say 1818, an intermediate host for *Schistosoma mansoni*. *Proc Biol Sci* 267:2351-8.
- Clement M, Posada D and Crandall KA (2000) TCS: A computer program to estimate gene genealogies. *Mol Ecol* 9:1657-1659.
- Crandall KA and Templeton AR (1993) Empirical tests of some predictions from coalescent theory with applications to intraspecific phylogeny reconstruction. *Genetics* 134:959-969.
- DeJong RJ, Morgan JA, Paraense WL, Pointier JP, Amarista M, Ayeh-Kumi PF, Babiker A, Barbosa CS, Bremond P, Pedro Canese A, de Souza CP, Dominguez C, File S, Gutierrez A, Incani RN, Kawano T, Kazibwe F, Kpikpi J, Lwambo NJ, Mimpfoundi R, Njiokou F, Noel Poda J, Sene M, Velasquez LE, Yong M, Adema CM, Hofkin BV, Mkoji GM and Loker ES (2001) Evolutionary relationships and biogeography of *Biomphalaria* (Gastropoda, Planorbidae) with implications regarding its role as host of the human bloodfluke, *Schistosoma mansoni*. *Mol Biol Evol* 18:2225-2239.
- Hall TA (1999) BioEdit: A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* 41:95-98.
- Huelsenbeck JP and Ronquist F (2001) MrBayes: Bayesian inference of phylogenetic trees. *Bioinformatics* 17:754-755.

- Mavárez J, Steiner C, Pointier J-P and Jarne P (2002) Evolutionary history and phylogeography of the schistosome-vector freshwater snail *Biomphalaria glabrata* based on nuclear and mitochondrial DNA sequences. *Heredity* 89:266-272.
- Nei M (1987) *Molecular Evolutionary Genetics*. Columbia University Press, New York, 512 pp.
- Nylander JAA (2004) MrModeltest 2.0. Program distributed by the author. Evolutionary Biology Centre, Uppsala University.
- Paraense WL (2001) The schistosome vectors in the Americas. *Mem Inst Oswaldo Cruz* 96(Suppl.):7-16.
- Pointier J-P, DeJong RJ, Tchuem Tchuenté LA, Kristensen TK and Loker ES (2005) A neotropical snail host of *Schistosoma mansoni* introduced into Africa and consequences for the schistosomiasis transmission: *Biomphalaria tenagophila* in Kinshasa (Democratic Republic of Congo). *Acta Tropica* 93:191-199.
- Rozas J, Sanchez-Del Barrio JC, Messeguer X and Rozas R (2003) DnaSP, DNA polymorphism analyses by the coalescent and other methods. *Bioinformatics* 19:2496-2497.
- Spatz L, Vidigal THDA, Caldeira RL, Dias Neto E, Cappa SMG and Carvalho OS (1999) Study of *Biomphalaria tenagophila tenagophila*, *B. t. guaiensis* and *B. occidentalis* by polymerase chain reaction amplification and restriction enzyme digestion of the ribosomal RNA intergenic spacer regions. *J Molluscan Stud* 65:143-149.
- Swofford DL (2002) PAUP*. *Phylogenetic Analysis Using Parsimony (*And Other Methods)*. Version 4. Sinauer Associates, Sunderland.
- Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F and Higgins DG (1997) The ClustalX windows interface: Flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* 25:4876-4882.
- Vidigal THDA, Kissinger JC, Caldeira RL, Pires ECR, Monteiro E, Simpson AJG and Carvalho OS (2000) Phylogenetic relationships among Brazilian *Biomphalaria* species (Mollusca, Planorbidae) based upon analysis of ribosomal ITS2 sequences. *Parasitology* 121:611-620.
- Vidigal TH, Spatz L, Kissinger JC, Redondo RA, Pires EC, Simpson AJ and Carvalho OS (2004) Analysis of the first and second internal transcribed spacer sequences of the ribosomal DNA in *Biomphalaria tenagophila* complex (Mollusca, Planorbidae). *Mem Inst Oswaldo Cruz* 99:153-158.
- Woodruff DS and Mulvey M (1997) Neotropical schistosomiasis: African affinities of the host snail *Biomphalaria glabrata* (Gastropoda: Planorbidae). *Biol J Linn Soc* 60:505-516.

Associate Editor: Fábio de Melo Sene