Molecular phylogenetic studies on Brugia filariae using HHA I repeat sequences

XIE H.*, BAIN O.* and WILLIAMS S. A.*,**

Summary:
This paper is the first molecular phylogenetic study on Brugia parasites (family Onchocercidae) which includes 6 of the 10 species of this genus: B. beaveri Ash et Little, 1964; B. buckleyi Dissanaike et Paramananthan, 1961; B. malayi (Brug, 1927) Buckley, 1960; B. pahangi (Buckley et Edeson, 1956) Buckley, 1960; B. patei (Buckley, Nelson et Heisch, 1958) Buckley, 1960 and B. timori Partono et al., 1977. Hha I repeat sequences are 322 nucleotides long, highly repeated, tandemly arranged and unique to the nuclear genomes of the genus Brugia. Hha I repeat sequence data was collected by PCR, cloning and dideoxy sequencing. The Hha I repeat sequences were aligned and analyzed by maximum parsimony algorithms, distance methods and maximum likelihood methods to construct phylogenetic trees. Bootstrap analysis was used to test the robustness of the different phylogenetic reconstructions. The data indicated that the Hha I repeat sequences are highly conserved within species yet differ significantly between species. The various tree-building methods gave identical results. Bootstrap analyses on the Hha I repeat sequence data set identified at least two clades: the B. pahangi-B. beaveri clade and the B. malayi-B. timori-B. buckleyi clade; the first clade includes parasites of carnivores from Asia and America; the second includes species from primates and lagomorphs from Asiatic region. It was also noted that the Hha I repeat sequences obtained from B. malayi were identical to those obtained from B. timori, indicating very recent speciation.

KEY WORDS: evolution, phylogenetics, PCR, molecular cloning, Hha I repeats, Brugia malayi, B. pahangi, B. beaveri, B. timori, B. patei, B. buckleyi.

Résumé:

MOTS CLES: évolution, phylogénèse, PCR, clonage moléculaire, répétitions HHA I, Brugia malayi, B. pahangi, B. beaveri, B. timori, B. patei, B. buckleyi.

There are ten species in the genus Brugia, six of which were available for study in this paper: B. beaveri Ash and Little, 1964; B. buckleyi Dissanaike and Paramananthan, 1961; B. malayi (Brug, 1927) Buckley, 1960; B. pahangi (Buckley and Edeson, 1956) Buckley, 1960; B. patei (Buckley, Nelson and Heisch, 1958) Buckley, 1960 and B. timori Partono et al., 1977. The morphology and biology of the Brugia species are relatively well known, but the phylogenetic relationships among them are yet to be understood. It is important to understand the evolutionary relationships of the various filarial parasites, since this knowledge can help parasitologists in the design of model systems for species that are difficult to maintain in the laboratory. Furthermore, this study may provide data for morphologists (who base their classification systems entirely on parasite morphology) with a new means to reference their own conclusions. It is very difficult even for morphologists to identify the different species of Brugia at the infective stage (L3 larvae). To resolve fine taxonomical distinctions at the DNA
level, highly repetitive DNA elements and other less conserved sequences can be very useful (Williams et al., 1988). Because most highly repeated DNA sequences do not appear to code for any protein or RNA product, these sequences are able to evolve at a much faster rate than the coding region sequences of genes. McReynolds et al. (1986) did digests of purified B. malayi genomic DNA with a variety of restriction endonucleases. Repeated DNA sequences of 322 bp were observed when the DNA was digested with either Hba I, Alu I or Rsa I. The 322 bp repeat sequence was designated the “Hha I repeat”. By subsequent dot hybridization and DNA renaturation kinetics experiments, it was determined that the Hha I repeats comprise about 12% of the DNA of B. malayi with a copy number of about 30,000 per haploid genome. This repeat is absent in other genera, including the very closely related genus Wuchereria (Williams et al., 1988). Since the ribosomal DNA spacer region sequences were unable to resolve the phylogenetic relationships among the Brugia species in a previous study (Xie et al., 1994), Hha I repeat sequences were used to investigate the phylogenetic relationships among the six species in the genus Brugia.

MATERIALS AND METHODS

PARASITE MATERIALS

Adult B. pabangi and B. malayi were obtained from Dr. J. McCall (TRS Laboratory, Athens, GA, U.S.A.). Adult B. malayi worms preserved in 70% ethanol were provided by Dr. O. Bain (Paris, France). B. patei adults were kindly provided by Drs. U.R. Rao and A.C. Vickery (University of Southern Florida, Gainesville, U.S.A.). B. beaveri adults preserved in 100mM EDTA were kindly provided by Dr. T.C. Orihel (Tulane University Medical Center, New Orleans, U.S.A.). B. buckleyi adults preserved in 70% ethanol were kindly provided by Dr. A.S. Dissanaike (Colombo, Sri Lanka). B. timori microfilariae preserved in 100mM EDTA were kindly provided by Dr. F. Partono (University of Indonesia, Jakarta, Indonesia). Morphological analyses were conducted to verify the identity of some of the specimens.

EXPERIMENTAL PROCEDURES

Extraction of filarial parasite DNA

Individual adult filarial parasites were picked from 70% ethanol solution and evaporated to dryness. The worms were resuspended in 50 µl of H2O and boiled for 10 minutes to release DNA. These preparations were used as the PCR templates. The same method was used to prepare PCR template from B. timori microfilariae.

Amplification of Hha I repeat DNA sequences from filarial parasite DNA

Two PCR primers matching the consensus sequences of the conserved regions of the Hha I repeat were synthesized by the phosphoramidite method using the Applied Biosystem 391A DNA synthesizer (Foster City, CA, U.S.A.). The sequences of the two primers were:

Primer 1: 5'-GCGCATAATTCATCAGC-3';
Primer 2: 5'-GCAGAAGCTTATCACAAAAGC-3'.

All PCR reagents except primers, templates and double distilled water (ddH2O) were from the GeneAmp PCR kit purchased from Perkin-Elmer Cetus (Norwalk, CT, U.S.A.). The PCR reaction contents included: 5 µl of 10x reaction buffer, 8 µl of 1.25mM dNTP mixture, 20 pmol of each primer, 1-2 µl template DNA solution, 1 unit of Taq polymerase and ddH2O added to a total volume of 50 µl. The PCR cycling programs used were: 95°C, 1 minute; 55°C, 1 minute; 72°C, 1 minute for a total of 30 cycles (Model 480 thermal cycler, Perkin-Elmer Cetus, Norwalk, CT, U.S.A.); and 93°C, 0.5 minute; 55°C, 0.5 minute; 72°C, 0.5 minute for a total of 30 cycles (System 9600 thermal cycler, Perkin-Elmer Cetus, Norwalk, CT, U.S.A.).

Cloning and sequencing of the PCR amplified Hha I repeat sequences and sequence data analysis

The details of the procedures were described in our previous paper (Xie et al., 1994). Please refer to it for more information.

RESULTS

DNA SEQUENCE COLLECTION AND ALIGNMENT

The head-to-tail tandem organization of Hha I repeats is shown in Fig. 1. At least three sequences were obtained for each of the six species. These sequences plus previously unpublished sequence data from our database were used for deriving consensus sequences. Hba I repeat sequences within each species share 99% similarity (data not shown). Consensus sequences for each of the six species were derived and aligned to one another. The portion of the aligned sequence data matrix which is the input for the tree reconstruction algorithms is shown in Table I.
Fig. 1 - *Hha*I repeat structure. PCR primer positions are illustrated by arrows. P1 is *Hha*I primer 1; P2 is *Hha*I primer 2.

**PHYLOGENETIC TREE RECONSTRUCTION**

*Analysis by maximum parsimony method (Swofford, 1993)*

Search settings:

Exhaustive search option was used. Gapmode was set to newstate. Branches having maximum length zero was collapsed to yield polytomies; topological constraints were not enforced; trees were unrooted; multi-state taxa was interpreted as polymorphism; character-state optimization was by accelerated transformation.

Search results:

One shortest tree of length 87 was found in this search (Fig. 2), which is one step shorter than the next two trees. The next two trees are both 88 steps long. They differ from the shortest tree in their relative positions of *B. buckleyi* and *B. timori*. The statistics of the shortest tree are the following (Kluge and Farris, 1969; Farris, 1989): Consistency index (CI) =

Table 1. – *Hha*I repeat DNA sequence data matrix used for phylogenetic tree reconstruction. There are 257 nucleotide positions in the data matrix. Twenty seven of them (about 10%) are informative. ("-" : gap; Constant: invariable site; Uninf.: uninformative site, as calculated by PAUP 3.1)
0.977; CI excluding uninformative characters = 0.708; Retention index (RI) = 0.931; Rescaled consistency index (RC) = 0.910.

**Analysis by Neighbor-joining method (Saitou and Nei, 1987) and by Maximum likelihood method (Felsenstein, 1981 and 1993)**

Both neighbor-joining and maximum likelihood methods yield trees identical to the tree generated by the parsimony analysis (Fig. 2). They are not shown here.

**Analysis by bootstrap methods (Felsenstein, 1985)**

Bootstrap analyses were done with parsimony algorithm, neighbor-joining algorithm and with maximum likelihood algorithm (Table II and Fig. 2). All the three methods give very similar results. It is clear from these analyses that the B. malayi-B. timori-B. buckleyi clade, B. malayi-B. timori-B. buckleyi-B. patei clade and the B. pabangi-B. beaveri clade are strongly supported (Fig. 2 and Table II).

<table>
<thead>
<tr>
<th>Clade</th>
<th>a</th>
<th>b</th>
<th>c</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parsimony</td>
<td>78%</td>
<td>98%</td>
<td>100%</td>
</tr>
<tr>
<td>Neighbor-joining</td>
<td>83%</td>
<td>94%</td>
<td>100%</td>
</tr>
<tr>
<td>Maximum likelihood</td>
<td>90%</td>
<td>95%</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table II. – Summary of the bootstrap values. a, b and c are the clades noted in Figure 2.

**DISCUSSION**

**DNA sequence data collection**

The Hba I repeat sequences were chosen to examine the phylogenetic relationships of the Brugia species because it was shown that Hba I repeats possess species-specific regions within the repeat sequences. Most of the parasite materials were available only in very limited quantities for this study, so it was not practical to use traditional methods to extract genomic DNA and to perform endonuclease digestion and cloning. A very effective method was developed using PCR technology. This method proved very efficient and reliable. For Hba I repeat sequences, no difference was seen in sequence data collected by the PCR cloning method and the traditional genomic DNA cloning method (data not shown). To further test the error rate of Taq polymerase used in PCR reactions, one already sequenced Hba I repeat clone was subjected to PCR and direct cloning procedures. Four clones were randomly selected and sequenced using the CircumVent Thermal Cycle Sequencing Kit. No nucleotide misincorporation was detected from the 1200 bases sequenced.

**Tree reconstruction algorithms**

In this study, three main categories of tree reconstruction methods were used: parsimony, distance and maximum likelihood analyses. The Hba I repeat sequence data set gave very similar tree topologies no matter which phylogenetic reconstruction algorithms were used.

**Parsimony analysis**: Because the sequence region studied was non-coding, all of the characters were equally weighted. The same tree was obtained even when transversions were weighted twice as heavy as transversions (data not shown). Different sequence addition methods were tried and all gave the same topology. Multi-state taxa were generally treated as polymorphism according to the sequence data. However, the same topologies were still obtained even if multi-state taxa were treated as uncertainty, in which case the consistency indices were improved dramatically (data not shown). This observation is due to the presence of a large amount of nucleotide polymorphism in the data set. The tree length distribution is heavily skewed (data not shown) indicating that a significant degree of phylogenetic information is contained in the data set (Hillis, 1991). The exhaustive search yielded one shortest tree (Fig. 2) which unambiguously resolved the branching order of the six species of Brugia. However, this tree was not rooted because there is no outgroup available (Hba I repeats were only found in Brugia species).

**Distance analysis and maximum likelihood analysis**: Distance methods including UPGMA (Sneath and Sokal, 1973), neighbor-joining, Fitch-Margoliash (Fitch-Margoliash, 1967) methods and the maximum likelihood analysis were used in this analysis. Distance matrices were calculated with the maximum
likelihood method to transform the observed pairwise distances. All of the methods gave identical tree topologies (data not shown). These results are in good agreement with the exhaustive search tree obtained by the parsimony method (Fig. 2).

CONFIDENCE LIMITS ON PHYLOGENIES AND THE HYPOTHESIS

In this study, algorithms from all three major categories were applied to the bootstrapped replicates. All three bootstrap analyses (Table II) support the grouping of the B. malayi-B. timori-B. buckleyi clade, B. malayi-B. timori-B. buckleyi-B. patei and the B. beaveri-B. pabangi clade which clearly identified the branching pattern of the six Brugia species (Fig. 2).

The Hba I repeat sequence data in this study showed distinctive patterns with respect to individual species (Except B. malayi and B. timori, Table I). This is a good example of the congruence among morphological data. B. timori and B. malayi were shown to be very closely related from both the Hba I repeat data and the 5S rDNA spacer region sequence data (Xie et al., 1994). Both the 5S spacer region sequences and the Hba I repeat sequences from the two species are virtually identical. Morphologically (Partono et al., 1994), these two species distinguish themselves from the other Brugia species by having a left to right spicular ratio of 3.0 (in the males) which is the biggest in the genus. The microfilariae of B. timori differ from the microfilariae of B. malayi in having a greater total body length and a long cephalic space. Overall, these two species are the most closely related among the Brugia species morphologically and molecularly. These two species may have been separated too recently to show any DNA sequence variation between them.

It should be noted that B. malayi has a relatively wide distribution in Southeast Asia and has marked intraspecific variations both morphologically (Bain et al., 1988) and biologically (e.g. periodicity). Hba I repeat sequences from B. malayi were also collected with regard to different geographical origin with different morphology (strains from China and India), to their periodicity and to their host specificity by Freedman (1991) and Xie (1993). The molecular data failed to reveal any variations among the different strains of B. malayi.

It should also be noted that of the three species from carnivores (B. pabangi, B. patei and B. beaveri), only B. pabangi (localized in the oriental region) and B. beaveri (localized in the nearctic region) are closely connected; Oriental species from humans and other mammals (B. malayi and B. timori) are closely connected to the species from Lagomorphs (B. buckleyi). More studies are necessary if one is to elucidate the more detailed evolution pathways of the Brugia species.

ACKNOWLEDGMENTS

The authors wish to thank J. McCall, U. R. Rao, A. C. Vickery, T. C. Orihel, L. A. McReynolds and F. Partono for providing the parasite materials used in this study; J. Felsenstein and D. L. Swofford for technical assistance with their computer programs; Dr. Pierre Darlu for his constructive suggestions in the preparation of this manuscript. Support for this study was provided by the Blakeslee fund for genetics research at Smith College.

REFERENCES


BUCKLEY J.J.C. and EDeson J.F.B. : On the adult morphology of Wuchereria sp. (malayi?) from a monkey (Macaca irus) and from cats in Malaya, and on Wuchereria pabangi n.sp. from a dog and a cat. Journal of Helminthology, 1956, 30, 1-20.


Accepté le 13 juillet 1994