

TECHNICAL NOTE

Phylogenetic utility of dynamin and triose phosphate isomerase

NATE B. HARDY

Department of Entomology, University of California, One Shields Avenue, Davis, CA 95616, U.S.A.

Abstract. Dynamin and triose phosphate isomerase (Tpi), two markers suggested as potentially useful for insect phylogenetics, have been sequenced for 12 scale insect taxa (Hemiptera: Coccoidea). Protocols are given for their amplification using conventional polymerase chain reaction, and their phylogenetic utility has been evaluated using several qualitative criteria and a modified version of the partition addition bootstrap alteration approach. Dynamin and Tpi fragments are easy to amplify and are evolving at a rate comparable with widely used nuclear ribosomal markers. The dynamin fragment has a single short intron. The Tpi fragment has three introns. One possible drawback to the use of the dynamin fragment as a phylogenetic marker is that its short length limits accurate modelling of complex substitution processes.

Introduction

If there were no conflict between datasets, we would not need sequences from multiple loci to reconstruct phylogenies. Incongruence between gene trees can result from a number of analytical (choice of optimality criterion, taxon sampling, nucleotide substitution model) and biological (variable rates of nucleotide evolution, hybridization, horizontal transfer, lineage sorting of ancestral polymorphisms) factors (Rokas *et al.*, 2003a). From an analysis of 106 genes from seven *Saccharomyces* species, Rokas *et al.* (2003b) concluded that between eight and 20 genes were required to reconstruct the genome tree. The extent to which this result can be generalized remains to be seen, especially in terms of the ability of multiple loci to overwhelm confounding analytical factors (Kolaczkowski & Thornton, 2004; Soltis *et al.*, 2004). Insofar as incongruence between gene trees is prevalent, multiple loci will be required. Between January 2000 and September 2006, of 92 papers using molecular data to reconstruct insect phylogenies published in the journals *Systematic Entomology* and *Molecular Phylogenetics and Evolution*, the median number of loci used was two (interquartile range = 2) with only seven studies using five or more loci. This survey ignored four studies that used entire mitochondrial genomes, but I treated tightly linked mitochondrial genes as separate, effectively overestimating the

number of data partitions evolving independently. Seemingly, insect systematists have been using too few loci to reconstruct organismal phylogenies with confidence.

Regier (2005) provided a list of primers that are reported to amplify 59 nuclear protein-coding markers across arthropods, with 20 markers highlighted as being sufficiently variable to merit testing below the ordinal level. Regier's group used reverse transcriptase polymerase chain reaction (RT-PCR), a PCR technique that uses mRNA transcripts as a starting material and deploys an RT to make single-stranded DNA used as the template in the PCR. The main advantage of RT-PCR is that introns are not present in mRNA. Unfortunately, because of the relative instability of RNA, RT-PCR requires fresh material, which is often unavailable to molecular phylogeneticists.

Published molecular phylogenetic studies on scale insects (Hemiptera: Coccoidea) have relied upon the small ribosomal subunit (18S), the large ribosomal subunit (28S), and to a more limited extent elongation factor 1 α , and the cytochrome oxidase subunits I and II (e.g. Gullan *et al.*, 2003; Cook & Gullan, 2004; Morse & Normark, 2006). Cytochrome oxidase subunit II has been difficult to amplify, as has elongation factor 1 α , the utility of which is further compromised by the presence of paralogous copies in at least some taxa (Downie & Gullan, 2004). Thus, there is an acute need for new markers for scale insect phylogenetics.

In the course of studying a monophyletic group of Australian felt scale species (Coccoidea: Eriococcidae) that induce galls on *Eucalyptus*, I attempted to amplify eight of the 'Regier markers' using conventional PCR. I was unable

Correspondence: N. B. Hardy, Department of Entomology, University of California, One Shields Avenue, Davis, CA 95616, U.S.A. E-mail: nbhardy@ucdavis.edu

to amplify six markers (in the following list each marker is denoted first by the name of the encoded protein and second by the code appearing in the Regier laboratory protocol, an integer with the suffix '*fin*'): syntaxin (36*fin*); glucosamine phosphate isomerase (44*fin*); gelsolin (109*fin*), hypothetical protein (270*fin*); tetrahydrofolate synthase (3017*fin*); and phosphatidylinositol kinase (8053*fin*). I could amplify two markers, dynamin (3006*fin*) and triose phosphate isomerase (Tpi; 197*fin*) and have assessed their utility for insect phylogenetics, using criteria extracted from Danforth *et al.*'s (2005) comparison of nuclear protein coding versus ribosomal genes: (1) ease of amplification and sequencing; (2) absence of long introns; (3) straightforward alignment; and (4) appropriate substitution rate. The importance of the absence of paralogues was also noted by Danforth *et al.* (2005), but this is difficult to demonstrate and I note only the number of bands visible on preparative gels. To this list I have added a fifth criterion: sequence length adequate for modelling the substitution process. If the sequence is too short (or if the data are partitioned into segments, such as codon positions, that are too short), parameter estimates may be subject to an unacceptable amount of random error. Phylogenetic utility will always be a function of a particular phylogenetic question. One summary method of evaluating the phylogenetic utility of a marker is simply to take a test sample of taxa and see how well the marker reconstructs a well-supported phylogeny.

With sequences of 18S, dynamin and Tpi for 12 scale insect taxa, I used a modification of the partition addition bootstrap alteration (PABA) approach (Struck *et al.*, 2006) to assess the relative contribution of each gene to the nodal support values resulting from an analysis of the combined data. The PABA approach addresses a very general question: how are bootstrap (BS) values affected by different sequences of partition addition? For the present study, I was concerned most with a specific case of partition addition; I wanted to estimate the contribution of each marker to the net BS values estimated from all of the data (our best estimate of the organismal phylogeny). To this end, I performed a partition subtraction bootstrap alteration (PSBA) analysis, explained below. The results of this diagnostic were compared with calculations of partitioned decay index (DI) values.

Dynamin is a large (600 amino acid) force-producing GTPase involved in exocytosis and endocytosis, causing the scission of nascent membrane vesicles (Sweitzer & Hinshaw, 1998). It is a member of the dynamin superfamily of proteins, which all share three primary structural features: a GTPase domain at the C-terminus, a central region of unknown function, and a GTPase effector domain at the N-terminus effecting oligomerization and regulation of GTPase activity. Regier and Cunningham (Regier, 2005) developed completely degenerate primers that yielded a 222 bp RT-PCR fragment. To my knowledge, this is the first study to test dynamin as a candidate for phylogenetics.

Tpi (or TIM) is an approximately 250 amino acid glycolytic enzyme catalysing the interconversion of glyceraldehyde-3-phosphate with dihydroxyacetone phosphate, and is essential for efficient energy production (Lolis *et al.*,

1990). The gene has played a central role in the debate regarding the origin of introns (Gilbert *et al.*, 1986; Tittiger *et al.*, 1992). Using primers flanking intron 3 of Tpi, Beltrán *et al.* (2002) and Blum *et al.* (2003) estimated relationships among closely related butterfly species, following the hypothesis that introns in nuclear protein-coding genes may be evolving at a rate comparable with mitochondrial DNA, and be useful for phylogenetics at the finest scale. In both studies, alignment of the intron sequences proved to be impossible for all but sister species. Peterson *et al.* (2004) used multiple genes, including Tpi, to estimate the age of the last common ancestor of the Bilateria, with four insect species representing the arthropods. In humans, a single functional Tpi gene occurs along with three intronless pseudogenes (Brown *et al.*, 1985). Avoiding introns by using RT-PCR, Regier (2005) developed primers yielding a 441 bp fragment. Sequences of Tpi from three cockroach species appear in GenBank, associated with an unpublished paper (Knobloch *et al.*, unpubl.).

Materials and methods

Taxon sampling

Nucleotide sequences were obtained for 11 felt scale (Eriococcidae) species and one undescribed mealybug (Pseudococcidae) outgroup (Table 1). All but two of the included eriococcids are part of a monophyletic group of gall-inducing Australian species (Gondwanan Clade; Cook & Gullan, 2004). The age of this lineage is unknown. *Cylindrococcus casuarinae* Maskell and a species of *Olliffia* Fuller were included as more proximal outgroups (compared with the mealybug). The undescribed mealybug DNA voucher will be deposited at the Bohart Museum of Entomology, University of California, Davis (UCD). All other DNA vouchers will be deposited in the Australian National Insect Collection (ANIC).

Primers and PCR

Fragments of four genes were amplified: 18S (*c.* 600 bp) (von Dohlen & Moran, 1995); a fragment overlapping the D2 and D3 expansion regions of 28S (*c.* 750 bp) [primers of Morse & Normark (2006) modified from Downton & Austin (1998) and Whiting *et al.* (1997)]; dynamin (*c.* 300 bp); and Tpi (*c.* 950 bp). Fragments of dynamin and Tpi were amplified with the Regier primers (Table 2).

Regier & Shi (2005) found that the addition of M13 tails to the 5' ends of degenerate primers markedly increased amplicon yield. In the Regier laboratory protocol it is specified that all primers tested had M13 tails, and that they were polyacrylamide gel electrophoresis (PAGE) purified. The authors acknowledged that they were uncertain of the necessity of PAGE purification. Initially, I tested primers with M13 tails that had not been PAGE purified, but these failed to produce observable amplicons. Subsequently, PCRs

Table 1. Taxa of Coccoidea sampled.

Species	Code	Collection site	Tpi Genbank accession no.	Dynammin Genbank accession no.
Pseudococcidae sp.	NH121	Fiji, Viti Levu	EF125993	EF125972
<i>Olliffia</i> sp.	NH128	Vic., Bunyip	EF125995	EF125983
<i>Cylindrococcus casuarinae</i> Maskell	NH068	Vic., Bellarine Peninsula	EF125991	EF125980
Gondwanan clade				
<i>Lachnodioides hirtus</i> Maskell	NH054	Vic., Tyabb	EF125990	EF125979
Genus X, sp. 1	NH034	Vic., Cardinia Reservoir Park	EF125985	EF125974
Genus X, sp. 2	NH126	Vic., Cranbourne	EF125994	EF125982
<i>Opisthoscelis conica</i> Fuller	NH048	Vic., Hattah	EF125987	EF125976
<i>Opisthoscelis mammularis</i> Froggatt	NH052	Vic., Cranbourne	EF125989	EF125978
<i>Opisthoscelis subrotunda</i> Schrader	NH075	Vic., Mildura	EF125992	EF125981
<i>Opisthoscelis</i> sp. 1	NH031	Vic., Grampians	EF125984	EF125973
<i>Opisthoscelis</i> sp. 2	NH050	Vic., Bunyip	EF125988	EF125977
<i>Sphaerococcopsis</i> sp.	NH045	Vic., Bundoora	EF125986	EF125975

Tpi, triose phosphate isomerase.

were performed with primers without the M13 tails. I have not tested PAGE-purified primers with M13 tails. Once several sequences of dynamin were obtained, COUNTCODON (<http://www.kazusa.or.jp/codon/countcodon.html>) was used to describe codon-usage bias and inform the design of less degenerate primers for felt scale insects (3006F1.1, CCG GAY ATG GCG TTC GAA GCT A; 3006R2.1, TCT TCG TGG TTG GTG TTC ATG TAC GC). It was then possible to run touchdown PCR programs with higher annealing temperatures and amplify fragments that had been elusive with Regier's primers at lower annealing temperatures.

PCR was performed with 18 µL reaction volumes: 4 µL of ddH₂O, 10 µL of Takara Ex *Taq* DNA Polymerase Premix, 1 µL of each primer (10 pmol/µL), and 2 µL of DNA template. Dynammin and Tpi fragments were amplified with similar touchdown (Don *et al.*, 1991) programs. Dynammin: initial denaturation of 95 °C for 3 min; six cycles of 95 °C for 30 s, 52 °C for 30 s, 72 °C for 1 min; six cycles of 95 °C for 30 s, 49 °C for 30 s, 72 °C for 1 min; six cycles of 95 °C for 30 s, 47 °C for 30 s, 72 °C for 1 min; six cycles of 95 °C for 30 s, 45 °C for 30 s, 72 °C for 1 min; 31 cycles of 95 °C for

30 s, 42 °C for 30 s, 72 °C for 1 min; a final extension of 72 °C for 10 min. The program used to amplify Tpi differed in the following respects: a 94 °C denaturation temperature was used to preserve enzyme activity, and longer extension times were used to accommodate the larger amplicon; 1 min 30 s for the cycles with annealing temperatures from 52 to 45 °C and 2 min for the final 31 cycles with a 42 °C annealing temperature.

DNA sequencing, alignment, variability and analysis

PCR products were purified by exonuclease I and shrimp alkaline phosphatase digestion of single-stranded DNA (primers) and dNTPs (ExoSAP-IT, USB Corp., Cleveland, OH, U.S.A.). With the same primers used for the PCR, amplicons were sequenced in both directions using the ABI Big Dye V3 terminator sequencing reaction kit (Perkin-Elmer/ABI, Weiterstadt, Germany) on an ABI Prism 3100 automatic sequencer on 5% acryl/bisacryl Long Ranger gels at the College of Biological Sciences DNA Sequencing Facility at UCD.

Sequences were compiled and edited using SEQUENCHER 4.2 (Gene Codes Corporation, Ann Arbor, MI, U.S.A.) and then aligned using MUSCLE (Edgar, 2004). Alignments of protein-coding genes were subsequently imported into MACCLADE (Maddison & Maddison, 2002), and introns delimited using the GT-AG rule (Rogers & Wall, 1980). Introns were excluded prior to all phylogenetic analyses.

PAUP* 4.0b10 (Swofford, 2003) was used to: (1) compute base frequencies for each nucleotide position and introns for both dynammin and Tpi; (2) perform χ^2 tests for homogeneity of base frequencies across taxa; (3) recover the most-parsimonious tree(s) for each locus and the combined dataset using the branch-and-bound algorithm; and (4) estimate nodal support for each optimal topology with nonparametric bootstrapping (Felsenstein, 1985). One thousand BS replicates were conducted for each dataset, with each search using the branch-and-bound algorithm.

Table 2. Primers from Regier (2005) for dynammin and triose phosphate isomerase (Tpi; 5'-3'). Primer names ending in F bind to the antisense strand, and those ending in R to the sense strand. The numbers in brackets following the sequences designate where the primers bind to the in silico translated protein sequences of the two loci in *Drosophila melanogaster* (GenBank accession numbers: Tpi: NM 176587; dynammin: NM 206745).

Primer	Sequence
197fin (Tpi) 1F	AA Y TGG AAR ATG AAY GG [138]
197fin (Tpi) 2R	GCC CAN ACN GGY TCR TA [291]
3006fin	CCN GAY ATG GCN TTY GA [458]
(dynammin) 1F	
3006fin	TCY TCR TGR TTN GTR TTC ATR
(dynammin) 2R	TA [540]

In their PABA approach, Struck *et al.* (2006) determined BS values for each partition and every combination of partitions. They then determined, for each node, δ , the change in the BS value resulting from the addition of one partition to an existing dataset for all possible combinations and orders of partition addition. Mean values for δ were calculated for each node. The results were tabulated and inspected for trends. For the PSBA diagnostic used here, I was concerned only with two types of partition: (1) that containing all of the data and (2) all possible partitions missing exactly one of the markers. BS values were determined for each of these and then δ , the change in the BS value resulting from the subtraction of one of the markers from the analysis, was determined for each subtraction. As a way of summarizing the contribution of one partition in reconstructing the combined data tree, positive and negative values for δ can be separately summed across nodes. Larger negative sums indicate that much of the support for the combined tree is derived from that marker. Conversely, large positive values for a partition indicate conflict between that partition and the combined data. DI for the unpartitioned combined dataset and the combined dataset partitioned by gene were computed with TREEROT.V2 (Sorenson, 1999), in conjunction with PAUP*.

Phylogenies were also reconstructed using Bayesian inference methods using MRBAYES 3.1.1 (Ronquist & Huelsenbeck, 2003). Substitution models and priors on model parameters were selected for: (1) the combined dataset; (2) each gene; and (3) each codon position for each protein-coding gene using the Akaike information criterion as implemented in MODELTEST 3.7 (Posada & Crandall, 1998). Each gene was analysed separately under the selected model(s) and parameter priors. Bayes factor comparison (as in Nylander *et al.*, 2004; Brandley *et al.*, 2005) was used to determine if the protein-coding genes were better fit by one global model or separate models for each codon position. Bayes factor estimates (twice the difference between harmonic means of $-\ln$ likelihoods) were interpreted with the suggestions of Kass & Raftery (1995). The combined data were analysed with the model and partitioning strategy selected for each gene. Each analysis consisted of four Markov chains, three hot and one cold (the program defaults), run simultaneously for one million generations, with trees sampled every 100 generations, discarding the first 1000 trees from the burn-in period. Each analysis was repeated three times, and recovered similar parameter estimates, suggesting that stationarity had been achieved.

Results

PCRs using the dynamin primers invariably yielded a single amplicon, with one intron ranging from 65–85 bp in length, along with nucleotides encoding 63 amino acids from the central region of the gene. There was a slight A/T bias (60%) and near significant heterogeneity of base frequencies at the third codon positions across taxa (Table 3).

Table 3. Nucleotide base frequencies for dynamin and triose phosphate isomerase (Tpi). Codon positions are designated nt1, nt2, and nt3.

Dynamin	All nt	nt1	nt2	nt3	Introns
A	0.3286	0.3458	0.2968	0.3432	0.3074
C	0.1729	0.2158	0.1283	0.1747	0.1365
G	0.2286	0.2694	0.2206	0.1957	0.2195
T	0.2699	0.1689	0.3543	0.2865	0.3366
$\chi^2 P$	0.8191	0.9999	1	0.0785	0.0129
Tpi	All nt	nt1	nt2	nt3	Introns
A	0.3089	0.2707	0.3112	0.3445	0.2905
C	0.1726	0.1275	0.2213	0.1687	0.1444
G	0.2447	0.3962	0.1714	0.1679	0.1555
T	0.2737	0.2056	0.2961	0.3188	0.4096
$\chi^2 P$	0.9167	0.9999	1	0.0135 ^a	0.1335

^aWhen the outgroups (Pseudococcidae sp.; *Olliffia* sp.; *Cylindrococcus casuarinae*) are excluded this *P*-value becomes 0.7606.

PCRs using the Tpi primers also yielded single bands. The reads from the sequencing machine were reliable only for 600–800 bp. The first of three introns varied from 50–135 bp across species, the second ranged from 120–260 bp, and the third from 75–225 bp. After discarding introns and areas of poor sequence quality at the beginning and end of the chromatographs, nucleotides encoding 104–108 amino acids remained. The amino acid indels occurred in the first and third exons. The A/T bias was similar to that of dynamin (59%). Base frequencies were significantly heterogeneous at the third codon positions, but this could be rectified by the exclusion of the outgroups (*P* = 0.7606).

Tree searches under maximum parsimony (MP) are summarized in Table 4. The MP tree reconstructed from the combined data (Fig. 1) had strong support (BS values > 70%) at every node except F, *Opisthoscelis conica* + *O. mammularis*. No two MP gene trees (not shown) were completely congruent, and none matched exactly the combined data topology, but no conflicting relationship was well supported. All MP trees supported the monophyly of the Gondwanan clade (Fig. 1, node B). Partitioned DI are given in Table 5. Five of six negative DI values came from 18S and dynamin, but no negative values were greater than four, and

Table 4. Summary of parsimony searches.

Data	Total number of characters	Parsimony informative characters	Tree length	Number of trees
18S	621	52	162	2
28S (D2, D3)	808	129	574	1
Dynamin	189	49	162	2
Tpi	326	108	348	2
Combined	1944	338	1257	1

Tpi, triose phosphate isomerase.

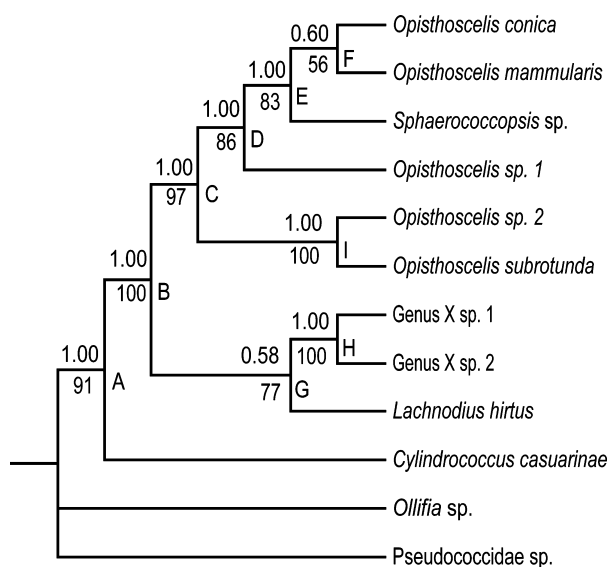


Fig. 1. Strict consensus and majority rule cladogram resulting from maximum parsimony and Bayesian analyses, respectively. Posterior probabilities are shown above each node and bootstrap values are shown below each node.

thus did not provide strong support for conflict among datasets (DeBry, 2001). For every node except node I, either Tpi or 28S had the largest positive DI. The results of the PSBA diagnostic (Table 6) reflected the partitioned DI values. The removal of Tpi from the dataset had the most negative impact on BS values across the tree. Excluding dynamin or 18S from the dataset appeared to have little impact on total tree support, yet each was important for individual nodes.

The models selected for each partition are given in Table 7. The most complex model, GTR + I + G, was not selected for any partition. Bayes factor comparison strongly favoured separate models for each codon position of Tpi ($B_{10} = 98$). This was not the case for dynamin, where Bayes factor comparison supported the application of a single model to the partition ($B_{10} = -60$). The results

Table 5. Decay index (DI) values. Nodes are marked on Fig. 1.

Node	Unpartitioned DI	18S DI	Dynamin DI	28S DI	Tpi DI
A	12	4	4	5	-1
B	37	10	4	13	10
C	10	-4	-1	7	8
D	6	-1.5	0.5	5	2
E	5	1	0	1	3
F	1	-1	1	1	0
G	4	1	-2	1	4
H	15	1	3	6	5
I	17	6	3	5	3

Tpi, triose phosphate isomerase.

Table 6. Partition subtraction bootstrap alteration bootstrap (BS) values. Nodes are marked in Fig. 1. These estimates fail to take into account the variance in BS values.

Node	Combined BS		Without 18S BS		Without dynamin BS		Without 28S BS		Without Tpi BS	
	BS	δ	BS	δ	BS	δ	BS	δ	BS	δ
A	91	-3	88, 0	-10	80, 0	-11	95, 0	+4	100, 0	0
B	100	0	99, +2	0	82, -15	0	76, -21	0	55, -31	-21
C	97	0	93, +7	-1	51, -35	0	43, -40	0	39, -17	-31
D	86	0	84, +1	+2	79, -4	0	66, -11	0	97, -3	-40
E	83	0	68, +12	+4	43, -13	0	79, +2	0	100, 0	-17
F	56	0	66, -11	-10	96, -4	0	100, 0	0	+2	-11
G	77	0	100, 0	+8	100, 0	0	100, 0	0	+4	-11
H	100	0	100, 0	0	+14	0	+2	0	+4	-123
I	100	0	+22	-11	-11	0	-82	0	-82	-123
Total + δ										
Total - δ										

Tpi, triose phosphate isomerase.

Table 7. Akaike information criterion selected substitution models. Codon positions are designated nt1, nt2 and nt3.

Partition	18S	28S	Dynamin	nt1	nt2	nt3	Tpi	nt1	nt2	nt3
Model	SYM + I + G	GTR + G	GTR + G	GTR + I	HKY + I	HKY	TrN + I + G	TrN + G	TIMeF + G	TVM + I

from the Bayesian analysis of the combined data under separate models for each partition [18S; D2D3; dynamin; Tpi codon position (nt1); Tpi nt2; Tpi nt3] were very similar to results recovered under MP, with all nodes except F (*Opisthoscelis conica* + *O. mammularis*) and G (*Lachnodius*

hirtus + genus X) with >95% posterior probabilities (Fig. 1). Phylograms resulting from the analysis of each marker separately, under the selected model(s) (Fig. 2) showed that dynamin and Tpi are evolving at rates intermediate to those of 18S and 28S.

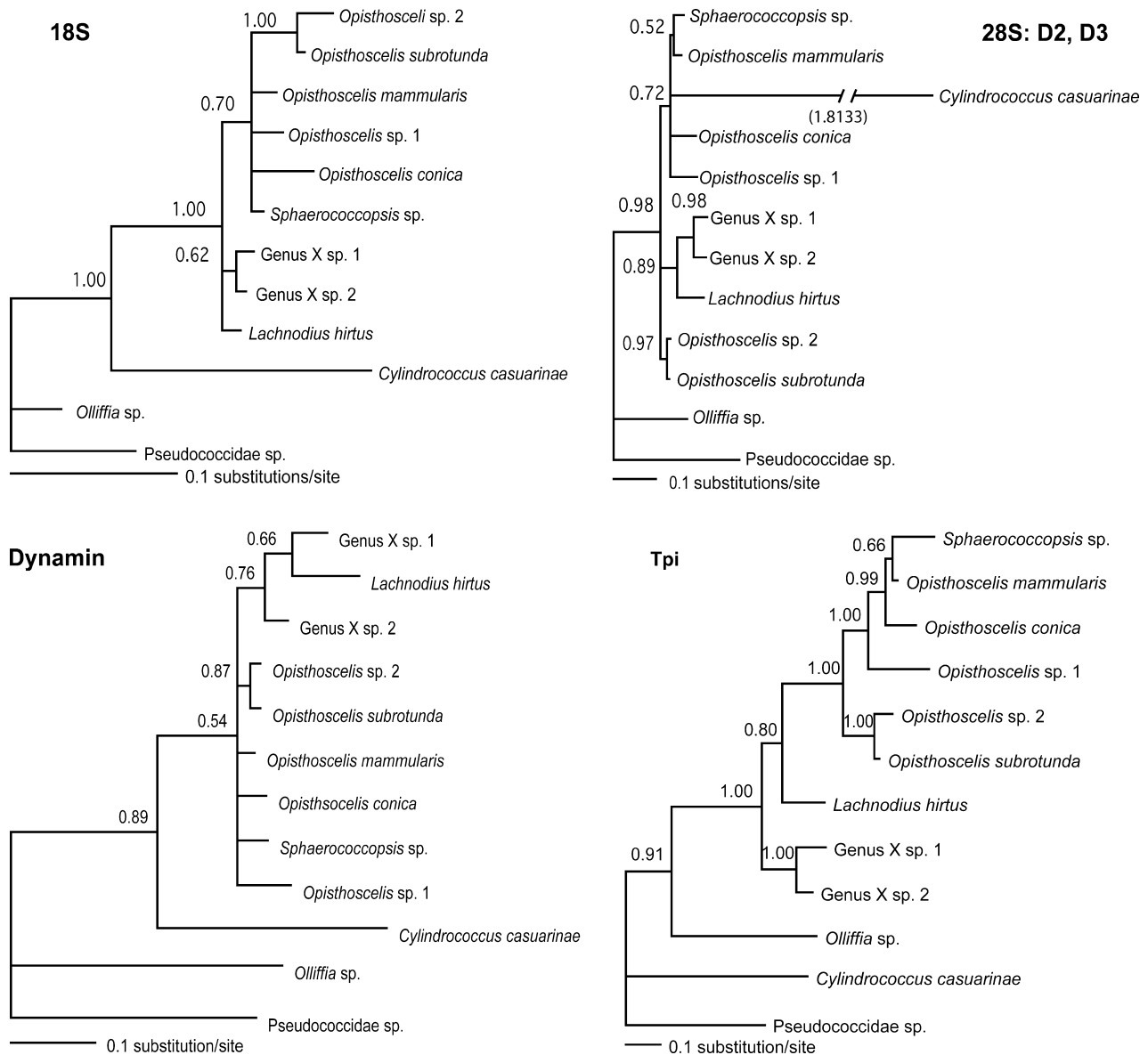


Fig. 2. Phylograms resulting from Bayesian analysis of each locus under the Akaike information criterion selected model(s). On the 28S tree, the branch subtending *Cylindrococcus casuarinae* is broken, with the total length shown in parentheses beneath the break.

Discussion

Partitioned DI values and the PSBA diagnostic revealed similar features of the dataset. Both indicated the dominance of Tpi and 28S in the reconstructions and the relative equivocacy of 18S and dynamin. However, raw DI values have been shown to be a function of both tree shape and the number of parsimony informative characters at the node of interest (DeBry, 2001), making the DI and any method employing the DI difficult to interpret. As noted by Struck *et al.* (2006), the usefulness of posterior probabilities for assessing the effect of a partition on nodal support is limited by the tendency of posterior probability values to be very close to either unity or zero. For these reasons, the use of methods based on the BS, such as the PSBA, may be the most appropriate. Dynamin and Tpi each have their shortcomings. Tpi amplicons vary in length, due to the presence of multiple introns and a number of amino acid indels in the first and third exons, resulting in more time spent on alignment. Also, the number of introns increases the chance that in some taxa at least one of the introns will be sufficiently large to preclude PCR using these primers. Although dynamin is characterized by less length polymorphism, the fragment is small, approaching the lower limit of what justifies the effort of sequencing. This problem is illustrated by our inability to justify the selection of separate models for each codon position. Of the numerous partitioning strategies evaluated by Brandley *et al.* (2005), partitioning by codon position was the most important in terms of increasing the likelihood of the data. The contribution of dynamin to the BS values of the concatenated dataset was not as substantial as that of Tpi, but dynamin was useful for certain nodes. Dynamin and especially Tpi appear to be evolving at rates appropriate for reconstructing relationships among species and genera in the Eriococcidae. The long branch leading to *Cylindrococcus casuarinae* in all of the genealogies is striking, but is consistent with previous estimates of Australian eriococcid phylogeny based on 18S (L.G. Cook, unpubl. data) and with its unusual morphology and ecology (Gullan *et al.*, 2004).

Amplification of Tpi is not as easy as for dynamin, but neither is difficult. It was not necessary to use more expensive Hot-Start *Taq* packages, nor was it necessary to use PAGE-purified primers with M13 tails added to the 5' ends. Further testing of Regier's markers seems warranted.

Acknowledgements

Brian O'Meara assisted with technical aspects of the analysis, in addition to suggesting certain methods. Peter Cranston, Penny Gullan, Geoff Morse and three anonymous reviewers provided valuable comments on earlier drafts of the manuscript. Specimens were collected under permit no. FA 05128 from the Department of Primary Industries and Water, Tasmania, and permit no. 10002997 from the Department of Sustainability and Environment, Victoria. This research was supported by a Henry A. Jastro and Peter

J. Shields Graduate Research Scholarship from the UCD College of Agricultural and Environmental Sciences to N.B. Hardy, and by grant DEB-0118718 from the U.S. National Science Foundation (Partnerships for Enhancing Expertise in Taxonomy program) to P.J. Gullan.

References

- Beltrán, M., Jiggins, C.D., Bull, V., Linares, M., Mallet, J., McMillan, W.O. & Bermingham, E. (2002) Phylogenetic discordance at the species boundary: comparative gene genealogies among rapidly radiating *Heliconius* butterflies. *Molecular Biology and Evolution*, **19**, 2176–2190.
- Blum, M.J., Bermingham, E. & Dasmahapatra, K. (2003) A molecular phylogeny of the neotropical butterfly genus *Anartia* (Lepidoptera: Nymphalidae). *Molecular Phylogenetics and Evolution*, **26**, 46–55.
- Brandley, M.C., Schmitz, A. & Reeder, T.W. (2005) Partitioned Bayesian analyses, partition choice, and the phylogenetic relationships of scinid lizards. *Systematic Biology*, **54**, 373–390.
- Brown, J.R., Daar, I.O., Krug, J.R. & Maquat, L.E. (1985) Characterization of the functional gene and several processed pseudogenes in the human triosephosphate isomerase gene family. *Molecular and Cellular Biology*, **5**, 1694–1706.
- Cook, L.G. & Gullan, P.J. (2004) The gall-inducing habit has evolved multiple times among the eriococcid scale insects (Sternorrhyncha: Coccoidea: Eriococcidae). *Biological Journal of the Linnean Society*, **83**, 441–452.
- Danforth, B.N., Lin, C. & Fang, J. (2005) How do insect nuclear ribosomal genes compare to protein-coding genes in phylogenetic utility and nucleotide substitution patterns? *Systematic Entomology*, **30**, 549–562.
- DeBry, R.W. (2001) Improving interpretation of the decay index for DNA sequence data. *Systematic Biology*, **50**, 742–752.
- von Dohlen, C.D. & Moran, N.A. (1995) Molecular phylogeny of the Homoptera: a paraphyletic taxon. *Journal of Molecular Evolution*, **41**, 211–223.
- Don, R.H., Cox, P.T., Wainwright, B.J., Baker, K. & Mattick, J.S. (1991) 'Touchdown' PCR to circumvent spurious priming during gene amplification. *Nucleic Acids Research*, **19**, 4008.
- Downie, D.A. & Gullan, P.J. (2004) Phylogenetic analysis of mealybugs (Hemiptera: Coccoidea: Pseudococcidae) based on DNA sequences from three nuclear genes, and a review of the higher classification. *Systematic Entomology*, **29**, 238–259.
- Downton, M. & Austin, A.D. (1998) Phylogenetic relationships among the microgastroid wasps (Hymenoptera: Braconidae): combined analysis of 16S and 28S rDNA genes and morphological data. *Molecular Phylogenetics and Evolution*, **10**, 354–366.
- Edgar, R.C. (2004) MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research*, **32**, 1792–1797.
- Felsenstein, J. (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evolution*, **39**, 783–791.
- Gilbert, W., Marchionni, M. & McKnight, G. (1986) On the antiquity of introns. *Cell*, **46**, 151–154.
- Gullan, P.J., Downie, D.A. & Steffan, S.A. (2003) A new pest species of the mealybug genus *Ferrisia* Fullaway (Hemiptera: Pseudococcidae) from the United States. *Annals of the Entomological Society of America*, **96**, 723–737.
- Gullan, P.J., Miller, D.R. & Cook, L.G. (2004) Gall-inducing scale insects (Hemiptera: Sternorrhyncha: Coccoidea). *Biology, Ecology and Evolution of Gall-inducing Arthropods* (ed. by A. Raman,

- C.W. Schaefer and T.M. Withers), pp. 195–196. Science Publishers, Enfield.
- Kass, R.E. & Raftery, A.E. (1995) Bayes factors. *Journal of the American Statistical Association*, **90**, 773–795.
- Kolaczowski, B. & Thornton, J.W. (2004) Performance of maximum parsimony and likelihood phylogenetics when evolution is heterogeneous. *Nature*, **431**, 980–984.
- Lolis, E., Alber, T., Davenport, R.C., Rose, D., Hartman, F.C. & Petsko, G.A. (1990) Structure of yeast triosephosphate isomerase at 1.9-Å resolution. *Biochemistry*, **29**, 6609–6618.
- Maddison, W.P. & Maddison, D.R. (2002) *Macclade*, version 4.05. Sinauer Associates, Sunderland, MA.
- Morse, G.E. & Normark, B.B. (2006) A molecular phylogenetic study of armoured scale insects (Hemiptera: Diaspididae). *Systematic Entomology*, **31**, 338–349.
- Nylander, J.A.A., Ronquist, F., Huelsenbeck, J.P. & Nieves-Aldrey, J.L. (2004) Bayesian phylogenetic analysis of combined data. *Systematic Biology*, **53**, 47–67.
- Peterson, K.J., Lyons, J.B., Nowak, K.S., Takacs, C.M., Wargo, M.J. & McPeck, M.A. (2004) Estimating metazoan divergence times with a molecular clock. *Proceedings of the National Academy of Sciences*, **101**, 6536–6541.
- Posada, D. & Crandall, K.A. (1998) Modeltest: testing the model of DNA substitution. *Bioinformatics*, **14**(9), 817–818.
- Regier, J.C. (2005) Protocols, concepts, and reagents for preparing DNA sequencing templates, version 4/01/05 (www.umbi.umd.edu/jcrlab/PCR_primers.pdf).
- Regier, J.C. & Shi, D. (2005) Increased yield of PCR product from degenerate primers with nondegenerate, nonhomologous 5'-tails. *Biotechniques*, **38**, 34, 36, 38.
- Rogers, J. & Wall, R. (1980) A mechanism for RNA splicing. *Proceedings of the National Academy of Sciences*, **77**, 1877–1879.
- Rokas, A., King, N., Finnerty, J. & Carroll, S.B. (2003a) Conflicting phylogenetic signals at the base of the metazoan tree. *Evolution and Development*, **5**, 346–359.
- Rokas, A., Williams, B.L., King, N. & Carroll, S.B. (2003b) Genome-scale approaches to resolving incongruence in molecular phylogenies. *Nature*, **425**, 798–804.
- Ronquist, F. & Huelsenbeck, J.P. (2003) MRBAYES 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics*, **19**, 1572–1574.
- Soltis, D.E., Albert, V.A., Savolainen, V. et al. (2004) Genome-scale data, angiosperm relationships, and 'ending incongruence': a cautionary tale in phylogenetics. *Trends in Plant Science*, **9**, 477–483.
- Sorenson, M.D. (1999) *Treerot*, version 2. Boston University, Boston.
- Struck, T.H., Purschke, G. & Halanych, K.M. (2006) Phylogeny of Eunicida (Annelida) and exploring data congruence using a partition addition bootstrap alteration (PABA) approach. *Systematic Biology*, **55**, 1–20.
- Sweitzer, S.M. & Hinshaw, J.E. (1998) Dynamin undergoes a GTP-dependent conformational change causing vesiculation. *Cell*, **93**, 1021–1029.
- Swofford, D.L. (2003) *PAUP**. *Phylogenetic Analysis Using Parsimony (*and Other Methods)*, version 4. Sinauer Associates, Sunderland, MA.
- Tittiger, C., Whyard, S. & Walker, V.K. (1992) A novel intron site in the triosephosphate isomerase gene from the mosquito *Culex tarsalis*. *Nature*, **361**, 470–472.
- Whiting, M.F., Carpenter, J.C., Wheeler, Q.D. & Wheeler, W.C. (1997) The Strepsiptera problem: phylogeny of the holometabolous insect orders inferred from 18s and 28s ribosomal DNA sequences and morphology. *Systematic Biology*, **46**, 1–68.

Accepted 22 November 2006