

Phylogenetic Analysis of Diprotodontian Marsupials Based on Complete Mitochondrial Genomes

Maruo Munemasa¹, Masato Nikaido¹, Stephen Donnellan², Christopher C. Austin³,
Norihiko Okada^{1,4*} and Masami Hasegawa^{5,6}

¹Graduate School of Bioscience and Biotechnology, Tokyo Institute of Technology, Kanagawa 226-8501, Japan

²Evolutionary Biology Unit, South Australian Museum, North Terrace, SA 5000, Australia

³Museum of Natural Science, Louisiana State University, LA 70803, USA

⁴Department of Evolutionary Biology and Biodiversity, National Institute for Basic Biology, Okazaki 444-8585, Japan

⁵Institute of Statistical Mathematics, Tokyo 106-8569, Japan

⁶Department of Biosystems Science, Graduate University for Advanced Studies, Kanagawa 240-0193, Japan

(Received 22 October 2005, accepted 25 April 2006)

Australidelphia is the cohort, originally named by Szalay, of all Australian marsupials and the South American *Dromiciops*. A lot of mitochondria and nuclear genome studies support the hypothesis of a monophyly of Australidelphia, but some familial relationships in Australidelphia are still unclear. In particular, the familial relationships among the order Diprotodontia (koala, wombat, kangaroos and possums) are ambiguous. These Diprotodontian families are largely grouped into two suborders, Vombatiformes, which contains Phascolarctidae (koala) and Vombatidae (wombat), and Phalangerida, which contains Macropodidae, Potoroidae, Phalangeridae, Petauridae, Pseudocheiridae, Acrobatidae, Tarsipedidae and Burramyidae. Morphological evidence and some molecular analyses strongly support monophyly of the two families in Vombatiformes. The monophyly of Phalangerida as well as the phylogenetic relationships of families in Phalangerida remains uncertain, however, despite searches for morphological synapomorphy and mitochondrial DNA sequence analyses. Moreover, phylogenetic relationships among possum families (Phalangeridae, Petauridae, Pseudocheiridae, Acrobatidae, Tarsipedidae and Burramyidae) as well as a sister group of Macropodoidea (Macropodidae and Potoroidae) remain unclear. To evaluate familial relationships among *Dromiciops* and Australian marsupials as well as the familial relationships in Diprotodontia, we determined the complete mitochondrial sequence of six Diprotodontian species. We used Maximum Likelihood analyses with concatenated amino acid and codon sequences of 12 mitochondrial protein genomes. Our analysis of mitochondria amino acid sequence supports monophyly of Australian marsupials + *Dromiciops* and monophyly of Phalangerida. The close relatedness between Macropodidae and Phalangeridae is also weakly supported by our analysis.

Key words: complete mitochondrial genome, Marsupial, *Dromiciops*, Phalangerida, phylogeny

INTRODUCTION

Phylogenetic relationships among marsupials have been studied by analyzing complete mitochondrial genomes and concatenated nuclear genes (Amrine-Madsen et al., 2003; Nilsson et al., 2003; Nilsson et al., 2004). Marsupials are separated into two cohorts: the Ameridel-

phia and Australidelphia (Szalay, 1982). Szalay (1982) originally recognized the Australidelphia cohort, which contains the genus *Dromiciops* and Australian marsupials, based on morphological features that both groups share, such as a continuous lower ankle joint pattern. Identification of this cohort was later supported by mitochondrial genome and concatenated nuclear gene sequence analyses (Amrine-Madsen et al., 2003; Nilsson et al., 2003; Nilsson et al. 2004).

Paleontological evidence suggests that all Australian

Edited by Yoko Satta

* Corresponding author. E-mail: nokada@bio.titech.ac.jp

marsupials immigrated to Australia from South America via Antarctica (Woodburne and Case, 1996). Szalay (1982) proposed two hypotheses concerning *Dromiciops*. One hypothesis states that all Australian marsupials are monophyletic and *Dromiciops* is a sister group of all Australian marsupials, and that the single ancestor of all extant Australian marsupials migrated from South America to Australia. The other hypothesis states that Australian marsupials are not monophyletic and that *Dromiciops* is nested in Australian marsupials. In this hypothesis, the ancestor of *Dromiciops* once migrated to Australia then returned to South America, or multiple ancestors of Australian marsupials migrated independently to Australia (Nilsson et al., 2004). An analysis with concatenated 6.4-kb nuclear genes did not resolve this issue (Amrine-Madsen et al., 2003). A mitochondrial genome sequence analysis supports the monophyly of *Dromiciops* and Australian omnivorous marsupials (*Isoodon macrourus*) and, thus, the latter hypothesis (Nilsson et al., 2004). Recent morphological analysis also supports the latter hypothesis (Horovitz and Sanchez-Villagra, 2003).

However, despite the results of these mitochondrial genome and nuclear gene analyses, the familial relationships among Diprotodontia remain unresolved. (Amrine-Madsen et al. 2003; Nilsson et al., 2004). The order Diprotodontia contains Australian marsupials distributed in New Guinea and Australia, including koala, wombat, cuscus and kangaroo. Diprotodontia are morphologically characterized by syndactylus (the second and third digits of the hind limbs are completely fused) and teeth morphology (a pair of incisors in the lower jaw is procumbent) (Marshall et al., 1990). Recent classification recognizes 19 Diprotodontian families and, of these, the extant families are Phascolarctidae (koala), Vombatidae (wombats), Macropodidae (kangaroos and wallabies), Potoroidae (bettongs, potoroos and rat kangaroos), Phalangeridae (brush-tailed possums), Petauridae (gliders and striped possums), Pseudocheiridae (ringtail possums), Acrobatidae (feather-tailed possum and pygmy glider), Tarsipedidae (honey possum) and Burramyidae (pygmy possums) (Table 1; Aplin and Archer, 1987). The suborder Vombatiformes includes two families, namely Phascolarctidae and Vombatidae, and other families are grouped as the suborder Phalangerida. In Phalangerida, there are eight extant families; six of the eight, excluding Macropodidae and Potoroidae, are grouped as possums.

Each family is well recognized by morphological characteristics. Molecular studies with the sequence of the mitochondrial gene, ND2, supports monophyly for each Diprotodontian family (Osborne et al., 2002). The monophyly of suborder Vombatiformes is well supported by several morphological and molecular studies (Kirsch, 1977; Aplin and Archer, 1987; Flannery, 1987; Marshall et al., 1990; Springer et al., 1994; Osborne et al., 2002;

Amrine-Madsen et al., 2003; Asher et al. 2004; Nilsson et al., 2004). Phylogenetic relationships among families in Phalangerida, however, are not well resolved. Flannery (1987) suggested two hypotheses for phylogenetic relationships among these families from morphological evidence:

1. Macropodoidea (Macropodidae + Potoroidae) is a sister group of Phalangeridae, and other possums (Petauridae, Pseudocheiridae, Tarsipedidae, Acrobatidae and Burramyidae) are a sister group of Macropodoidea and Phalangeridae (Tree 1 in Fig. 1).
2. Macropodoidea (Macropodidae + Potoroidae) is a sister group of a monophyletic group composed of all possums (Phalangeridae, Petauridae, Pseudocheiridae, Tarsipedidae, Acrobatidae and Burramyidae) (Tree 2 in Fig. 1).

Hypothesis 2 proposes that Macropodoidea is a group independent of other possums. Flannery showed only one synapomorphy supportive of hypothesis 2. In contrast, he recognized four synapomorphies for hypothesis 1. Kirsch's serological data showed close relatedness between *Macropus* (Macropodidae) and *Trichosurus* (Phalangeridae), supporting hypothesis 1 (Kirsch, 1977). The analysis of the ND2 amino acid sequence data also supports a close relationship between Phalangeridae and Macropodoidea (Macropodidae + Potoroidae) (Osborne et al., 2002), also supporting hypothesis 1.

Another difficulty in determining phylogenetic relationships among families in Phalangerida is that the position of Vombatiformes differs in different analyses. According

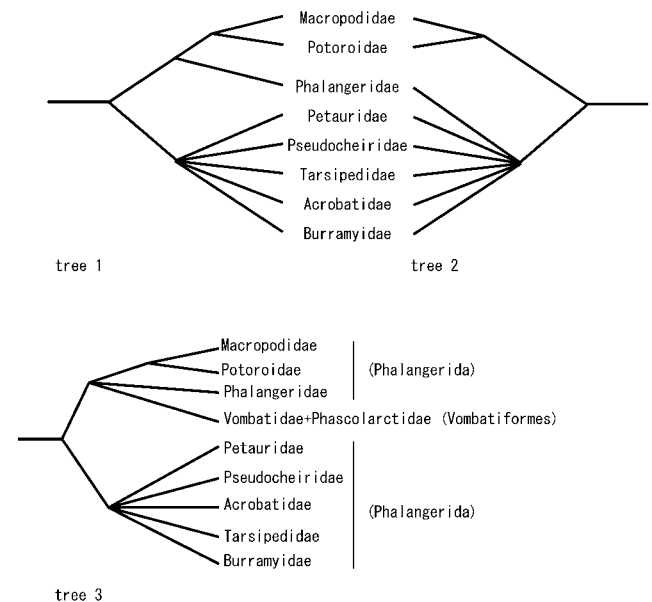


Fig. 1. Two hypotheses by Flannery for the phylogenetic relationships among families in Phalangerida. Tree 1 and 2 indicate Flannery's hypothesis trees predicted by morphological data. Tree3 is indicated by mitochondrial genome ND2 sequence (Osborne et al., 2002)

to the two hypotheses of Flannery, Vombatiformes is considered to be monophyletic as an outgroup taxon. In Osborn's ND2 analysis, however, Vombatiformes is close to Phalangeridae + Macropodoidea (Macropodidae + Potoroidae) (Tree 3 in Fig. 1), implying that suborder Phalangerida is paraphyletic. The same result is seen in the complete mitochondrial genome data analysis (Nilsson et al., 2004). However, these results (Fig. 1, Tree 3) contradict Flannery's monophyletic premise. On the other hand, Maximum Likelihood (ML) analyses of 6.4-kb concatenated nuclear genes suggested the monophyly of Phalangerida with a high bootstrap support (Amrine-Madsen et al., 2003).

In the present study, we sequenced the complete mitochondrial genomes of six Diprotodontian species. Then we performed ML analysis at the amino acid (aa) and nucleotide (nt) levels to elucidate the phylogenetic relationships in Diprotodontia at the family level and to reevaluate the phylogenetic position of *Dromiciops*, currently considered to be a member of Australidelphia.

MATERIALS AND METHODS

Samples of genomic DNA We isolated DNA from ethanol-preserved liver samples from *Phascolarctos cinensis* (koala), *Lagorchestes hirsutus* (western hare wallaby), *Phalanger interpositus* (Stein's cuscus), *Dactylopsila trivirgata* (common striped possum) and *Distoechurus pennatus* (feather-tail possum) provided from South Australian Museum. *Petaurus breviceps* (sugar glider) was purchased from pet shop. Sample information is shown in Table 1. Tissue samples were digested with proteinase K. Total genomic DNA was then isolated by phenol/chloroform extraction and ethanol precipitation (Blin and Stafford, 1976) and stored at 4°C.

Amplification of entire mitochondrial genome sequences We designed universal primers based on the complete mitochondrial sequence of *Macropus robustus* (accession number Y10524) and *Trichosurus vulpecula* (AF357238). The forward and reverse PCR primers are shown in Table 2. We PCR amplified 3–4-kb mitochondrial DNA (mtDNA) fragments of each samples using 30 cycles of denaturation at 94°C for 30 s, annealing at 50°C for 1 min and extension at 72°C for 3 min in thin-walled reaction tubes. The reaction mixture contained 2.5 U Taq polymerase (TaKaRa, Siga, Japan), Taq buffer (Takara), 0.4 mM dNTPs, 2.5 mM MgCl₂, 10 pM primers and 100 ng of genomic DNA. The PCR products completely covered the mitochondrial genomes and were confirmed by electrophoresis in a 2.0% SeaKem GTG agarose gel (TaKaRa, Siga, Japan). The control region of *Phalanger interpositus* and *Petaurus breviceps* were not amplified in this study.

Direct Sequencing and primer walking PCR products of about 2–5 kb fragments were amplified and used as templates for direct sequencing and primer walking. Cycle sequencing using BigDye terminator (ABI, Foster City, CA) was performed with 25 cycles of denaturation at 96°C for 30 s, annealing at 47°C for 15 s, and extension at 60°C for 2 min in a total volume of 10 µl.

Phylogenetic analysis We used the 23 mtDNA sequences. The sequences of 12 mitochondrial proteins (mt-proteins) encoded in the same strand of mtDNA were used for phylogenetic analyses. Alignments of sequences were carefully checked by eye, and all positions with gaps or ambiguous alignments, in addition to overlapping regions between ATP6 and ATP8 and between ND4 and ND4L, were excluded. The total number of remaining

Table 1. Diprotodontian families

| Suborder | Family | Species name | Common English name | Museum ID No. | Accession No. |
|---------------|-------------------------------|--------------------------------|----------------------|---------------|---------------------------------|
| Vombatiformes | Phascolarctidae | <i>Phascolarctos cinereus</i> | Koala | 98.122 | AB241053 (this work) |
| | Vombatidae | <i>Vombatus ursinus</i> | Wombat | | AJ304826 (Janke et al. 2002) |
| Phalangerida | Macropodidae | <i>Macropus robustus</i> | Wallaby euro | | Y10524 (Janke et al. 1997) |
| | | <i>Lagorchestes hirsutus</i> | Western hare wallaby | LH1 | AB241056 (this work) |
| | Potoroidae | <i>Not used in this study</i> | | | |
| | Phalangeridae | <i>Phalanger interpositus</i> | Stein's cuscus | T75 | AB241057 (this work) |
| | | <i>Trichosurus vulpecula</i> | Brush-tailed possum | | AF357238 (Phillips et al. 2001) |
| | Petauridae | <i>Petaurus breviceps</i> | Sugar glider | | AB241055 (this work) |
| | | <i>Dactylopsila trivirgata</i> | Striped possum | T21 | AB241054 (this work) |
| | Pseudocheiridae | <i>Not used in this study</i> | | | |
| | Acrobatidae | <i>Distoechurus pennatus</i> | Feather-tail possum | S65 | AB241052 (this work) |
| | Tarsipedidae | <i>Not used in this study</i> | | | |
| Burramyidae | <i>Not used in this study</i> | | | | |

Table 2. Primers to amplify complete mitochondrial DNA

| Primer | Sequence(5'-3') |
|------------|-------------------------|
| MarMitF-2 | AAAGCAAAGCACTGAAA |
| MarMitR1 | TTAGCYAGTCAAACCGG |
| MarMitF-1 | GAGCTTAATTGAAATAGGCAA |
| MarMitR2-1 | AGCCTGAATAGGMCTAGAAAT |
| MarMitF2a | GATAACAGCGCAATCCTATT |
| MarMitF2b | GGTTCAAATCCTCTCCTTAA |
| MarMitR2-2 | AATCCAGGTCGGTTTCTATC |
| PRIM1 | AAGCTATCGGGCCCATA |
| PRIM3a | GAATTTACAGTTCAATGCTTA |
| Mar4R | CCAAAGGCCTTCAAAGCCTT |
| PRIM4R | GGTGATGTGGCGTCTTGAA |
| PRIM3a | GAATTTACAGTTCAATGCTTA |
| PRIM6a | GGTCAATGYTCAGAAATYGT |
| PRIM7a | ATGAACGAAAATTTATTTGCC |
| PRIM8a | TAATGACYACCAAACACA |
| MarMitF8 | TTTCCAATCATTAAGTTCTGG |
| MarMitR8-1 | AGATAWTTGATTTCCGACTCAA |
| MarMitF8 | TTTCCAATCATTAAGTTCTGG |
| MarMitR11b | TCCACCRATTCTTCAARTAC |
| MarMitF11 | TGAGACCTHCAACAYATCTTCAT |
| MarMitR14b | TGGACTCTAACATAACCTA |
| MarMitF14 | GTAGCYATAGCAGMAGTGTAAC |
| PRIM15r | CATCAACWCCCAAAGCTGAC |

Forward and reverse primer sets for amplify each samples are follows; *Phascolarctos cinereus* [MarMitF-2, MarMitR2-2; MarMitF2, MarMitR2-1; ND1F8, MarMitR5; PRIM3a, MarMitR8; PRIM6a, MarMitR9-2; MarMitF9-2, PRIM15r; MarMitF14, MarMitR1], *Lagorhynchus hirsutus* [PRIM15, MarMitR2-2, MarMitF2, MarMitR2-1; PRIM1, MarMitR5; PRIM3a, PRIM7R; PRIM4, MarMitR9-2; MarMitF8-1, MarMitR14b; MarMitF14, PRIM15r], *Phalanger interpositus* [MarMitF14, MarMitR1; MarMitF-1, MarMitR2-2; MarMitF2, MarMitR2-1; PhaInM2F2-1, MarMitR5; PRIM3a, PRIM4R; PRIM4, PRIM7R; PRIM6a, MarMitR8-1; MarMitF8, MarMitR11b; MarMitF11a, MarMitR14b], *Petaurus breviceps* [PRIM15, MarMitR2-1; MarMitF2, MarMitR2-1; ND1F8, MarMitR5; PRIM3a, PRIM7R; PRIM6a, MarMitR9-2, MarMitF9-2, RPIM1r] *Dactylopsila trivirgata* [PRIM15, MarMitR2-2; MarMitF2, MarMitR2-1; PRIM1, Mar4R; PRIM3a, MarMitR8; MarMitF8, MarMitR11b; PRIM9a, MarMitR14b; MarMitF11a, PRIM15r], *Distoechurus pennatus* [PRIM15, MarMitR2-2; MarMitF2, MarMitR2-1; ND1F8, MarMitR5; PRIM3a, PRIM7R; PRIM6a, MarMitR9-2; MarMitF8-1, MarMitR14b; MarMitF14c, PRIM15r].

codons was 3,453.

The phylogenetic analyses were carried out at both the aa and nt levels with the ML method (Felsenstein, 1981; Kishino et al., 1990). We used the ProtML program in the MOLPHY package (ver. 2.3) (Adachi and Hasegawa, 1996a), the TREE-PUZZLE program for quartet-puzzling (QP) analysis (Strimmer and von Haeseler, 1996), and the

CodeML program in the PAML package (ver. 3.14) (Yang, 1997) for analysis of the aa sequences with the mtREV-F model (Adachi and Hasegawa, 1996b) and analysis of the nucleotide sequences of the protein-encoding genes with the codon-substitution (CS) model (Goldman and Yang, 1994; Yang et al., 1998). For the CS model, we used Miyata et al.'s (1979) distance with geometric formulae (Yang et al., 1998). We also used the BaseML program in the PAML package with the GTR + Γ model (Rodriguez et al., 1990) to analyze nt sequences, in which only the first and second codon positions were used.

In using the CodeML, BaseML, and TREE-PUZZLE programs, the discrete Γ distribution model (with eight categories, except for the CS model for which four categories were used because of the computational burden) for the site-heterogeneity (Yang, 1996) was adopted, and the shape parameter (α) of the Γ distribution was optimized. Bootstrap probabilities (BPs) were estimated by the RELL (resampling of estimated log-likelihoods) method (Kishino et al., 1990) with 10,000 bootstrap resamplings. The RELL method has been shown to be efficient in estimating BPs without performing ML estimation for each resampled data (Hasegawa and Kishino, 1994).

From preliminary analyses of the concatenated aa sequences of the 12 mt-proteins using the NJ, MP and TREE-PUZZLE methods, we were able to fix some clades that were supported with high BPs or QP supports and had no biological controversy surrounding them. This reduced the number of candidate trees provided to more sophisticated analyses. Even so, the number of candidate trees was still too large to allow exhaustive analysis with a computationally intensive method, and therefore we carried out an approximate likelihood analysis with the ProtML program for all the candidate trees. The most serious problem of the ML method when applied to data from many species is the explosively increasing number of possible trees, but most of these trees are very poor and unpromising. By quickly eliminating these trees by an approximate method, the ML method can be applied to a large phylogeny. In estimating the branch lengths for each tree topology by ML, we usually use the time-consuming Newton-Raphson method. The approximate likelihood option implemented in ProtML avoids this process and estimates an "approximate likelihood" from the initial values for the Newton-Raphson method given by the ordinary least squares. We could examine all the possible trees with the approximate likelihood method, and, by excluding unpromising trees by this approximate criterion, we could select the best 20,000 trees for the full likelihood analysis. There is a strong correlation between the approximate likelihood and the maximum likelihood, and, in a practical sense, this is a good method to reduce the computational burden (Adachi and Hasegawa, 1996a). However, even 20,000 trees may be too many for most sophisticated models. Therefore,

we further reduced the number of candidate trees by selecting the best trees among the 20,000 that had log-likelihood scores differing by less than 2 SEs from that of the highest likelihood tree derived from a simpler model.

RESULTS

Genome features of the six new mitochondrial genomes

The mitochondrial genomes of six marsupials determined in our analysis contain 13 proteins, 2 ribosomal RNA and 22 transfer RNA genes. The gene orders are same as typical marsupials (Janke et al., 1994, 1997). The tRNA-Lys is highly mutated, and the anticodons of tRNA-Asp are GCC. These phenomena are also typical to marsupial mitochondria where it is reported that tRNA-Lys from nuclear DNA is imported to mitochondria and help translation (Döner et al., 2001) and anticodon of tRNA-Asp GCC which usually code for Gly is modified to GUC after transcription (Börner et al., 1996).

In placental mitochondrial control region, the presence of three conserved sequence block (CSB I, CSB II and CSB III) were reported (Walberg and Clayton, 1981). And two of CSBs (CSB II and CSB III) were found in *Didelphis* and *Macropus* (Janke et al., 1994, 1997). In our analysis, all CSBs are found in *Isodon*, *Distoechirus*, *Lagorchestes*, *Phascolarctos*, *Macropus* and *Vombatus*.

Estimation of the phylogenetic tree using mt-genome data

Figure 2 shows a QP tree of the concatenated aa sequences of 12 mt-proteins. This tree indicates five stable monophyletic clades. These are *Macropus* + *Lagorchestes* clade (Macropodidae) (1), the *Trichosurus* + *Phalanger* clade (Phalangeridae) (2), the *Vombatus* + *Phascolarctos* clade (Vombatiformes) (6), the *Caenolestes* + *Rhyncholestes* clade (9), and the *Didelphis* + *Thylamys* clade (10), and were confirmed with 95, 85, 89, 98, and 100% QP support, respectively. These clades were also supported with high BPs from the NJ and MP analyses (data not shown); therefore, we fixed these clades in the subsequent ML analyses. We used some eutherians and monotremes as outgroups (the relationships among the outgroup species were also fixed, as shown in Fig. 2). The number of possible trees among (1) *Macropus* + *Lagorchestes*, (2) *Trichosurus* + *Phalanger*, (3) *Distoechurus*, (4) *Petaurus*, (5) *Dactylopsila*, (6) *Vombatus* + *Phascolarctos*, (7) *Isodon*, (8) *Dromiciops*, (9) *Caenolestes* + *Rhyncholestes*, (10) *Didelphis* + *Thylamys* is 34,459,425. These trees were examined using the approximate likelihood option of ProtML for the concatenated mt-proteins, and the best 20,000 trees were selected for full likelihood analyses by the ProtML program with the mtREV-F model. The best 1,155 trees with log-likelihood scores within 2 SEs compared to the

ML tree by the ProtML analysis were selected for more sophisticated analyses using the CodeML program with the mtREV-F + Γ model and by the BaseML with the GTR + Γ model applied to the nucleotides at the first and second positions in each codon.

Figure 3 shows the ML tree given by the CodeML analysis. Strong support was given to the *Petaurus* + *Dactylopsila* clade (Petauridae), which had 99% BP (100% BP by the BaseML analysis of the first and second codon positions), and to the Diprotodontia clade that consists of groups (or a taxa) 1–6 with 98% (94%) BP. The grouping of the Australian (Au) marsupials with the South American (Am) *Dromiciops* (the Australidelphia hypothesis by Szalay) was supported with 92% (87%) BP. The ML tree suggests that *Dromiciops* is nested within the Australian marsupials. The BP for the monophyly of Australian marsupials, excluding *Dromiciops* as an outgroup, was only 2% (4%), and thus this relationship was not supported (Tree 5 in Table 3).

In the ML tree, the *Didelphis* + *Thylamys* clade (Didelphimorpha) represents the most basal lineage in the marsupial tree, and the *Caenolestes* + *Rhyncholestes* clade (Paucituberculata) is the sister group of the Australian marsupials + *Dromiciops*. The BP for the monophyly of *Didelphis* + *Thylamys* + *Caenolestes* + *Rhyncholestes* was 2.5% (8.9%); therefore, monophyly of the two major South American families was unlikely by our analysis. The monophyly of South American marsupials, including *Dromiciops*, was 0.07% (0.00%). Therefore, from our analyses, these groupings seem unlikely, consistent with previous analyses of mitochondrial genomes and nuclear genes (Springer et al., 1998; Amrine-Madsen et al., 2003, Nilsson et al. 2004). However, the South American clade including *Dromiciops* cannot be rejected by the conservative test of the Weighted Shimodaira-Hasegawa (WSH) test (Tree 6 in Table 3).

The monophyly of Phalangerida, including (1) Macropodidae, (2) Phalangeridae, (3) *Distoechurus*, (4) *Petaurus* and (5) *Dactylopsila*, was supported by 83% BP in the aa analysis and by 84% BP in the nt analysis of first and second positions within codons. The *Petaurus* + *Dactylopsila* + *Distoechurus* clade received 80% BP in the aa analysis but only 32% in the nt analysis of first and second positions within codons. We carried out further ML analysis of nt sequences using the CS model for a limited number of species (Fig. 4). This is the ML tree of 105 unrooted trees of the following six groups: (1) *Macropus* + *Lagorchestes*, (2) *Trichosurus* + *Phalanger*, (3) *Distoechurus*, (4) *Petaurus*, (5) *Dactylopsila* and (6) *Vombatus* + *Phascolarctos*. This tree differs from that in Fig. 3 in that *Distoechurus* is not a sister group of the *Petaurus* + *Dactylopsila* clade, and the grouping of these three species received only 46% BP. All the BPs were low except for the *Petaurus* + *Dactylopsila* grouping; the relationships among *Macropus* + *Lagorchestes*, *Distoechurus*,

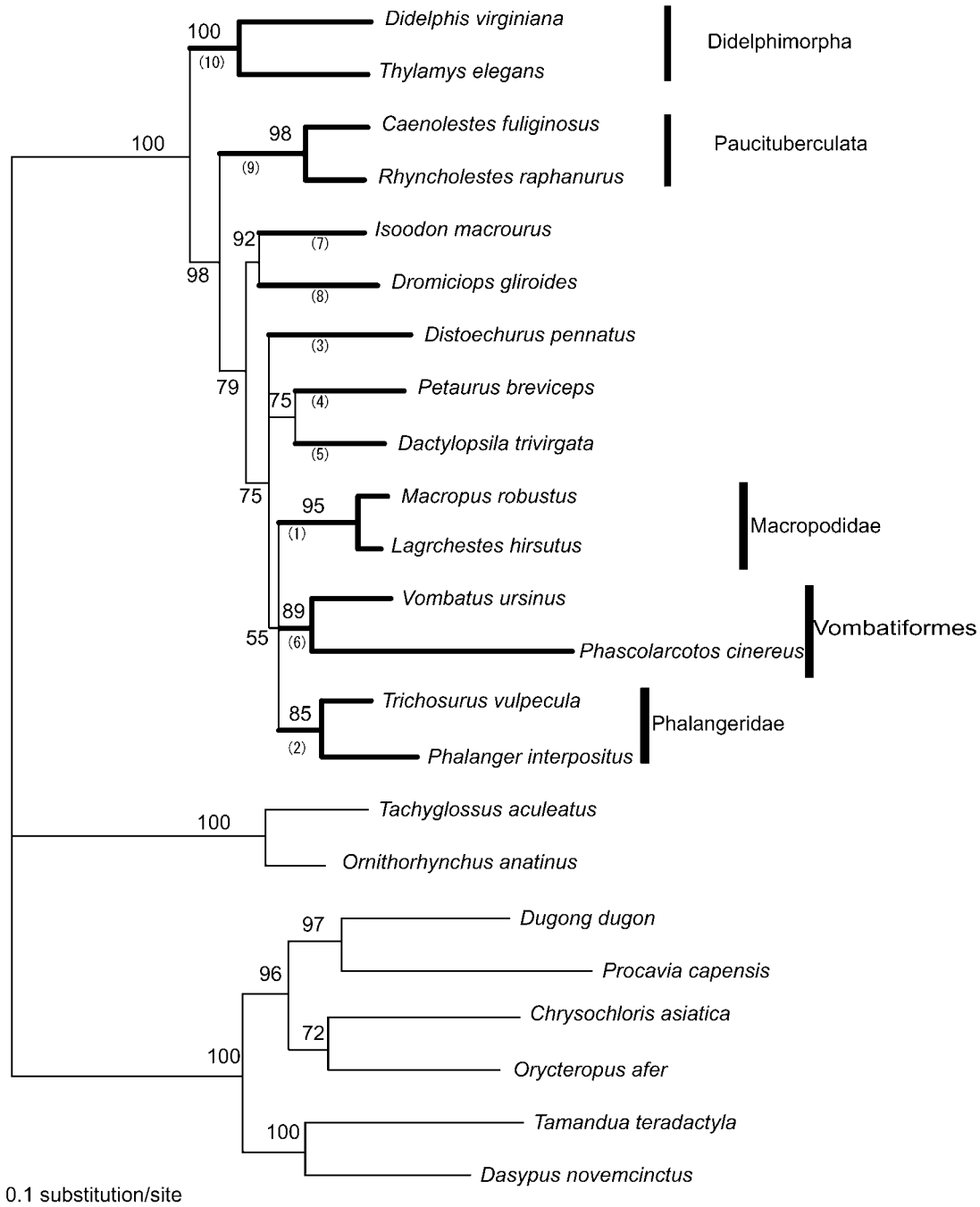


Fig. 2. A quartet-puzzling tree of concatenated amino acid sequences of 12 mt-proteins using the mtREV-F + Γ model. This tree was used for making the conclusive tree shown in Fig. 3. The horizontal length of each branch is proportional to the estimated number of amino acid substitutions. Numbers indicate QP support values. Eutherians and monotremes were used as outgroups. The numbers (1)–(10) under the branch indicate monophyletic groups and single taxa. The five stable monophyletic groups are shown in bold. *Dasypus novemcinctus*, nine-banded armadillo (accession number Y11832) *Tamandua tetradactyla*, southern tamandua (AJ421450), *Chrysochloris asiatica*, Cape golden mole (AB096866), *Dugong dugon*, dugong (AY075116), *Procavia capensis*, cape hyrax (AB096865), *Orycteropus afer*, aardvark (Y18475), *Didelphis virginiana*, Virginia opossum (Z29573), *Trichosurus vulpecula*, common brush-tailed possum (AF357238), *Isoodon macrourus*, brindled bandicoot (AF358864), *Macropus robustus*, wallaroo euro (Y10524), *Tachyglossus aculeatus*, short-beaked echidna (AJ303116), *Ornithorhynchus anatinus*, platypus (X83427), *Caenolestes fuliginosus*, Ecuadorian caenolestid (AJ508400), *Dromiciops gliroides*, colocolos (AJ508402), *Rhyncholestes raphanurus*, Chilean caenolestid (AJ50839), *Thylamys elegans*, elegant fat-tailed opossum (AJ508401), *Vombatus ursinus*, wombat (AJ304826), *Phascolarctos cinensis* koala (this work ; AB241053), *Lagrchestes hirsutus* western hare wallaby (this work; AB241056), *Phalanger interpositus* Stein's cuscus (this work; AB241057), *Petaurus breviceps* sugar glider (this work; AB241055), *Dactylopsila trivirgata* common striped possum (this work; AB241054) and *Distoechrus pennatus* feathertail possum (this work; AB241052).

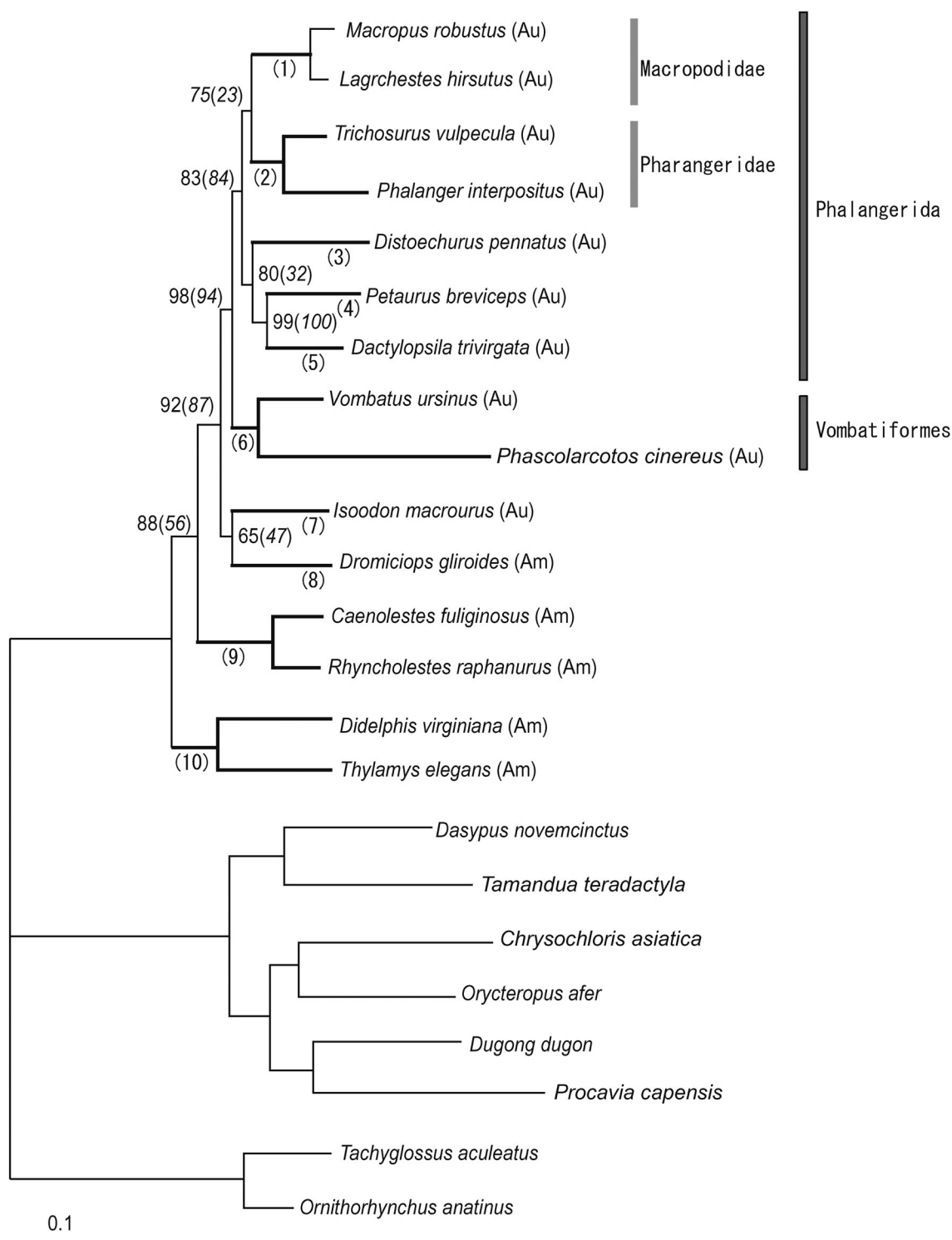


Fig. 3. An ML tree of concatenated amino acid sequences of 12 mt-proteins using the mtREV-F + Γ model. The monophyly of *Isoodon* and *Dromiciops*, the monophyly of Phalangerida and the sister relationship between Macropodidae and Phalangeridae were supported in this tree. The horizontal length of each branch is proportional to the estimated number of amino acid substitutions. Numbers indicate percent BPs estimated by the RELL method for amino acid analysis, and numbers in parentheses indicate percent BPs estimated by the RELL method for nt analysis of the first and second positions in codons. Numbers (1)–(10) under the branch indicate the monophyletic groups shown in Fig. 2.

Table 3. Comparison of trees among marsupials

| Tree | $\Delta l \pm SE$ | BP | AU | PP | KH | WSH |
|------|-----------------------------|--------------------|-------|---------------------|-------|-------|
| 1 | $\langle -44,819.8 \rangle$ | 0.376 | 0.970 | 0.995 | 0.718 | 1.000 |
| 2 | -21.5 ± 17.8 | 0.011 | 0.376 | 5×10^{-10} | 0.115 | 0.917 |
| 3 | -5.4 ± 9.7 | 0.109 | 0.795 | 0.004 | 0.282 | 0.995 |
| 4 | -12.2 ± 7.7 | 0.014 | 0.201 | 5×10^{-6} | 0.059 | 0.660 |
| 5 | -12.3 ± 8.2 | 0.007 | 0.235 | 5×10^{-6} | 0.067 | 0.820 |
| 6 | -48.2 ± 17.8 | 2×10^{-5} | 0.046 | 1×10^{-21} | 0.004 | 0.136 |

a) Amino acid sequences

b) Nucleotide sequences (first and second positions in codons)

| Tree | $\Delta l \pm SE$ | BP | AU | PP | KH | WSH |
|------|-----------------------------|-------|-------|---------------------|-------|-------|
| 1 | -3.9 ± 15.9 | 0.031 | 0.616 | 0.005 | 0.405 | 0.999 |
| 2 | $\langle -51,096.1 \rangle$ | 0.051 | 0.78 | 0.225 | 0.511 | 1.000 |
| 3 | -4.2 ± 12.4 | 0.020 | 0.652 | 0.003 | 0.369 | 0.999 |
| 4 | -7.8 ± 17.4 | 0.018 | 0.354 | 9×10^{-5} | 0.328 | 0.969 |
| 5 | -11.8 ± 14.7 | 0.001 | 0.168 | 2×10^{-6} | 0.211 | 0.938 |
| 6 | -39.4 ± 19.7 | 0.000 | 0.005 | 2×10^{-18} | 0.023 | 0.247 |

The differences in the log-likelihoods of alternative trees from that of the ML tree (Δl) are given (estimated with Kishino and Hasegawa's (1989) formulae) followed by ± 1 SE. Log-likelihood of the ML tree is given in the parentheses. The p -values of the following tests by the CONSEL program (Shimodaira and Hasegawa, 2001) are also shown: Bootstrap Probability (BP; Felsenstein, 1985), Approximately Unbiased test (AU; Shimodaira, 2002); Bayesian Posterior Probability (PP; Huelsenbeck et al., 2001), Kishino-Hasegawa test (KH; Kishino and Hasegawa, 1989), and Weighted Shimodaira-Hasegawa test (WSH; Shimodaira and Hasegawa, 1999). Among the 1,155 trees analyzed with the sophisticated models, only two trees were represented by ML trees in either analysis. Also shown are the second best tree by the amino acid analysis, the best Ameridelphia monophyly tree, the best tree in which Australian marsupials are monophyletic, and the best tree in which South American marsupials including *Dromiciops* are monophyletic.

Tree 1: (((((Mac,Pha),(Dis,(Pet,Dac))),VomPha),(Iso,Dro)),Pau),Did);

Tree 2: ((((((Mac,Dis),(Pet,Dac)),Pha),VomPha),(Dro,Iso),Pau),Did);

Tree 3: ((((((Mac,Pha),(Dis,(Pet,Dac))),VomPha),(Dro,Iso),Pau),Did) (the second best tree by amino acid analysis);

Tree 4: (((((Mac,Pha),(Dis,(Pet,Dac))),VomPha),(Iso,Dro)),(Pau,Did)) (the best among Ameridelphia monophyly trees);

Tree 5: ((((((Mac,Pha),(Dis,(Pet,Dac))),VomPha),(Iso,Dro),Pau),Did) (the best tree in which Australian marsupials are monophyletic);

Tree 6: ((((((Mac,Pha),(Dis,(Pet,Dac))),VomPha),(Iso),(Dro,Pau),Did)) (the best tree in which American marsupials including *Dromiciops* are monophyletic).

The abbreviations are as follows; Mac; *Macropus robustus*, Pha; *Phalanger interpositus*, Dis; *Distoechurus pennatus*, Pet; *Petaurus breviceps*, Dac; *Dactylopsila trivirgata*, VomPha; *Vombatus ursinus* and *Phascolarctos cinereus*, Iso; *Isodon macrourus*, Dro; *Dromiciops gliroides*, Pau; *Caenolestes fuliginosus* and *Rhyncholestes raphanurus*, Did; *Didelphis virginiana*.

Petaurus + *Dactylopsila* and *Trichosuru* + *Phalanger* remain ambiguous by our ML analyses.

DISCUSSION

Phylogenetic position of South American Microbiotheria Although the hypothesis of the South American order Microbiotheria (*Dromiciops*) being nested among Australian marsupials (Fig. 3 and Table3, Trees-1, 2, 3 and 4) was the most likely by our analysis consistently with the previous mitochondrial genome analysis (Nilsson et al., 2004), the support was not conclusively strong. Tree-5, where *Dromiciops* is sister to all Australian marsupials (Marshall et al., 1990), could not be rejected by

the AU, KH and WSH tests (Table 3). Indeed, as also was obtained in the mitochondria genome analysis of Nilsson et al. (2004), the most likely tree was that of *Dromiciops* nested within Australian marsupials. However, Tree 4 of Australidelphia monophyly had a log-likelihood score lower than that of the ML tree by only 15.0 ± 10.3 (± 1 SE) and could not be rejected with the SH test ($p = 0.63$).

Szalay (1982) proposed the cohorts Ameridelphia and Australidelphia (including *Dromiciops* in South America), and the monophyly of Australidelphia has been supported by previous molecular studies (e.g., Springer et al., 1998; Amrine-Madsen et al., 2003; Nilsson et al., 2004). Although the conservative WSH test could not reject an

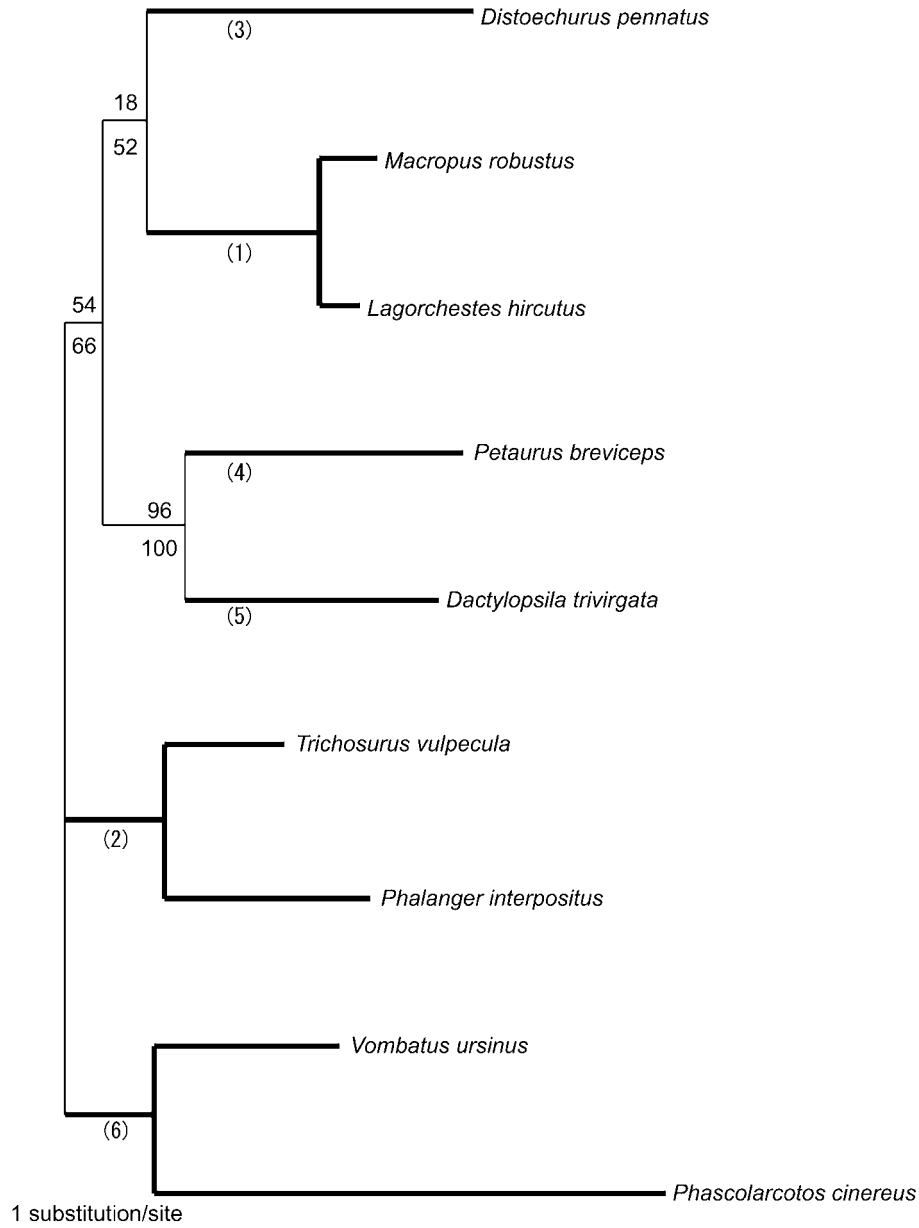


Fig. 4. An ML tree of concatenated codon sequences of 12 mt-protein genes using the CS model. Numbers indicate percent BPs estimated by the RELL method for amino acid analysis (above the branch) and CS analysis (below the branch). Numbers (1)–(10) under the branch indicate the monophyletic groups shown in Fig. 2.

alternative possibility (Tree 6 in Table 3), the Australidelphia clade including *Dromiciops* was also very strongly supported by our analyses.

Is Australidelphia monophyletic? The best tree differed between the aa analysis and the nt analysis of the first and second position in codons. The former analysis preferred Tree 1, in which Australidelphia is paraphyletic, whereas the latter preferred Tree 2, in which Australidelphia is monophyletic. However, the difference was not statistically significant. The log-likelihood differences between the two trees were only 21.5 ± 17.8 and

3.9 ± 15.9 for the aa and nt analyses, respectively, and the AU, KH and WSH tests could not reject the alternative tree. However, the Bayesian posterior probability (PP; Huelsenbeck et al., 2001) rejected Tree 2 with a p -value of 5×10^{-6} , but this might reflect the tendency of Bayesian PPs to predict with overconfidence when the model assumed in the analysis differs from reality (which is always the case to some degree) (Waddell et al., 2001). Therefore, this molecular analysis should be regarded as having only limited resolution concerning Australidelphia monophyly.

Phylogenetic relationships among Phalangerida, Vombatiformes and Macropodoidea Diprotodontia is separated into suborders, Vombatiformes and Phalangerida (Aplin and Archer, 1987). The monophyly of Phalangerida is ambiguous because the synapomorphy of Phalangerida is poorer than that of Vombatiformes (Aplin and Archer, 1987). While monophyly of Phalangerida is supported by serological analysis, not enough samples were analyzed in that study to conclusively determine monophyly (Kirsch, 1977). Only the concatenated nuclear gene sequence analysis strongly supports the monophyly of Phalangerida (Amrine-Madsen, 2003); all other analyses using mitochondrial genome sequences do not strongly support this assertion (Springer et al., 1994; Osborne et al., 2002; Nilsson et al., 2004). In our study, the QP tree (Fig. 2) indicated that Phalangeridae (*Trichosurus* + *Phalanger*), Macropodidae (*Macropus* + *Lagorchestes*) and Vombatiformes (*Vombatus* + *Phascogale*) form a monophyletic group. But the more sophisticated ML analysis supports the monophyly of Phalangerida with 83% BP by the aa analysis, and 84% by the nt analysis of first and second positions in codons (Fig. 3). Thus, our analysis of the complete mitochondrial genome supports the monophyly of Phalangerida.

In addition, our ML analysis with mitochondrial aa sequence supports a close relationship between Phalangeridae and Macropodoidea. Osborne's mitochondrial ND2 analysis indicated that Phalangeridae is a sister group of Macropodoidea (Osborne et al., 2002). Our analysis of mitochondrial genome aa sequence partially supports a close relationship between Phalangeridae and Macropodidae, (and, at the same time, that Phalangerida is monophyletic). Therefore our analysis is consistent with Flannery's hypothesis 1.

Familial relationships in Phalangerida remain ambiguous, although some molecular analyses have attempted to clarify relationships among kangaroos (Macropodidae) and possums (Phalangeridae, Petauridae, Pseudocheiridae, Acrobatidae, Tarsipedidae and Burramyidae) (Springer et al., 1994; Osborne et al., 2002). Our analysis of aa sequences supports the monophyly of Acrobatidae (*Distoechurus*) and Petauridae (*Petaurus* + *Dactylopsila*), but the codon sequence analysis does not. The elucidation of relationships among possum families is difficult because a large amount of molecular data are not yet available. More possum DNA sequences, in addition to other diagnostic information (for example, SINE (Short Interspersed Repetitive Element) insertions), are still needed to resolve phylogenetic relationships among the Diprotodontian families.

REFERENCES

Adachi, J., and Hasegawa, M. (1996a) MOLPHY: Programs for Molecular Phylogenetics ver. 2.3, Computer Science Mono-

- graphs, 28. Institute of Statistical Mathematics, Tokyo.
- Adachi, J., and Hasegawa, M. (1996b) Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* **42**, 459–468.
- Amrine-Madsen, H., Scally, M., Westerman, M., Stanhope, M.J., Krajewski, C., and Springer, M. S. (2003) Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. *Mol. Phylogenet. Evol.* **28**, 186–196.
- Aplin, K. P., and Archer, M. (1987) Recent advances in marsupial systematics with a new systematic classification. In: *Possums and Opossums: Studies in Evolution*. (ed.: Archer, M.) pp. xv–lxxii. Surrey-Geatty and Sons and the Zoological Society of New South Wales, Sydney.
- Asher, R. J., Horovitz, I., and Sanchez-Villagra, M. R. (2004) First combined cladistic analysis of marsupial mammal interrelationships. *Mol. Phylogenet. Evol.* **33**, 240–250.
- Blin, N., and Stafford, D. W. (1976) A general method for isolation of high molecular weight DNA from eukaryotes. *Nucleic Acids Res.* **3**, 2303–2308.
- Börner, G. V., Mörl, M., Janke, A., and Pääbo, S. (1996) RNA editing changes the identity of a mitochondrial tRNA in marsupials. *EMBO J.* **15**, 5949–5957.
- Dörner, M., Altmann, M., Pääbo, S., and Mörl, M. (2001). Evidence for import of a Lysyl-tRNA into marsupial mitochondria. *Mol. Biol. Cell.* **12**, 2688–2698.
- Felsenstein, J. (1981) Evolutionary trees from DNA sequences: a maximum likelihood approach. *J. Mol. Evol.* **17**, 368–376.
- Felsenstein, J. (1985) Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* **39**, 783–791.
- Flannery, T. (1987) The relationships of the macropodoids (Marsupialia) and the polarity of some morphological features within the phalangeriformes. in "Possums and Opossums: studies in evolution" ed by M. Archer. Surrey Beatty & Sons and the Royal Zoological Society of New South Wales Sydney, 741–747.
- Goldman, N., and Yang, Z. (1994) A codon-based model of nucleotide substitution for protein-coding DNA sequences. *Mol. Biol. Evol.* **11**, 725–736.
- Hasegawa, M., and Kishino, H. (1994) Accuracies of the simple methods for estimating the bootstrap probability of a maximum likelihood tree. *Mol. Biol. Evol.* **11**, 142–145.
- Horovitz, I., and Sánchez-Villagra, M. R. (2003) A morphological analysis of marsupial mammal higher-level phylogenetic relationships. *Cladistics* **19**, 181–212.
- Huelsenbeck, J. P., Ronquist, F., Nielsen, R., and Bollback, J. P. (2001) Bayesian inference of phylogeny and its impact on evolutionary biology. *Science* **294**, 2310–2314.
- Janke, A., Feldmaier-Fuchs, G., Thomas, W. K., Haeseler, A., and Pääbo, S., (1994) The marsupial mitochondrial genome and the evolution of placental mammals. *Genetics* **137**, 243–256.
- Janke, A., Xu, X., and Arnason, U. (1997) The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship among Monotremata, Marsupialia, and Eutheria. *Proc. Natl. Acad. Sci. USA* **94**, 1276–1281.
- Janke, A., Magnell, O., Wiczeorek, G., Westerman, M., and Arnason, U. (2002) Phylogenetic analysis of 18S rRNA and mitochondrial genomes of the wombat, *Vombatus ursinus*, and the spiny anteater, *Tachyglossus aculeatus* : increased support for the Marsupionta hypothesis. *J. Mol. Evol.* **54**, 71–80.
- Kirsh, J. A. W. (1977) The comparative serology of marsupialia, and a classification of marsupials. *Aust. J. Zool. Suppl. Ser.* **52**, 1–152.

- Kishino, H., and Hasegawa, M., (1989) Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in hominoidea. *J. Mol. Evol.* **29** (2) 170–179
- Kishino, H., Miyata, T., and Hasegawa, M. (1990). Maximum likelihood inference of protein phylogeny, and the origin of chloroplasts. *J. Mol. Evol.* **31**, 151–160.
- Marshall, L. G., Case, J. A., and Woodburne, M. O. (1990) Phylogenetic relationships of the families of marsupials. In *Current Mammalogy*, vol. 2 (ed.: H. Genoways), pp.433–505. New York, Plenum Press.
- Miyata, T., Miyazawa, S., and Yasunaga, T. (1979) Two types of amino acid substitutions in protein evolution. *J. Mol. Evol.* **12**, 219–236.
- Nilsson, M. A., Gullberg, A., Spotorno, A. E., Arnason, U., and Janke, A. (2003) Radiation of extant marsupials after the K/T boundary: Evidence from complete mitochondrial genomes. *J. Mol. Evol.* **57**, S3–S12.
- Nilsson, M. A., Arnason, U., Spencer, P. B. S., and Janke, A. (2004) Marsupial relationships and a timeline for marsupial radiation in South Gondwana. *Gene* **340**, 189–196.
- Osborne M. J., Christidis, L., and Norman, J. A. (2002) Molecular phylogenetics of the Diprotodontia (kangaroos, wombats, koala, possums, and allies). *Mol. Phylogenet. Evol.* **25** (2), 219–228.
- Phillips, M. J., Lin, Y., Harrison, G. L., and Penny, D. (2001) Mitochondrial genomes of a bandicoot and a brushtail possum confirm the monophyly of australidelphian marsupials. *Procl. R. Soc. Lond. B* **268**, 1533–1538.
- Rodriguez, F., Oliver, J. L., Marin A., and Medina, J. R. (1990) The general stochastic model of nucleotide substitution. *J. Theoret. Biol.* **142**, 485–501.
- Shimodaira, H. (2002) An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* **51**, 492–508.
- Shimodaira, H., and Hasegawa, M. (1999) Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* **16**, 1114–1116.
- Shimodaira, H., and Hasegawa, M. (2001) CONSEL: a program for assessing the confidence of phylogenetic tree selection. *Bioinformatics* **17**, 1246–1247.
- Springer, M. S., Westerman, M., and Kirsch, J. A. W., (1994) Relationships among orders and families of marsupials based on 12S ribosomal DNA sequences and the timing of the marsupial radiation. *J. Mammal. Evol.* **2**(2), 85–115.
- Springer, M. S., Westerman M., Kavanagh, J. R., Burk, A., Woodburne, M. O., Kao, D. J., and Krajewski, C. (1998) The origin of the Australian marsupial fauna and the phylogenetic affinities of the enigmatic monito del monte and marsupial mole. *Proc. R. Soc. Lond. B* **265**, 2381–2386.
- Strimmer, K., and von Haeseler, A. (1996) Quartet puzzling: a quartet maximum-likelihood method for reconstructing tree topologies. *Mol. Biol. Evol.* **13**, 964–969.
- Szalay, F.S. (1982) A new appraisal of marsupial phylogeny and classification. In *Carnivorous marsupials* (ed.: M. Archer), pp. 621–640. Sydney: Royal Zoological Society of New South Wales.
- Walberg, M. W., and Clayton, D. A. (1981) Sequence and properties of the human KB cell and mouse L cell D-loop regions of mitochondrial DNA. *Nucleic Acids Res.* **9** (20), 5411–5421.
- Waddell, P. J., Kishino, H., and Ota, R. (2001) A phylogenetic foundation for comparative mammalian genomics. *Genome Informatics* **12**, 141–154.
- Woodburne, M. O., and Case, J. A. (1996). Dispersal, vicariance, and the late Cretaceous to early Tertiary land mammal biogeography from South America to Australia. *J. Mammal. Evol.* **3**, 121–161.
- Yang, Z. (1996) Among-site rate variation and its impact on phylogenetic analyses. *TREE* **11**, 367–372.
- Yang, Z. (1997) PAML: a program package for phylogenetic analysis by maximum likelihood. *CABIOS* **13**, 555–556.
- Yang, Z., Nielsen, R., and Hasegawa, M. (1998) Models of amino acid substitution and applications to mitochondrial protein evolution. *Mol. Biol. Evol.* **15**, 1600–1611.