

Combined Mitochondrial and Nuclear DNA Sequences Resolve the Interrelations of the Major Australasian Marsupial Radiations

MATTHEW J. PHILLIPS,¹ PATRICIA A. MCLENACHAN,² CHRISTIN DOWN,² GILLIAN C. GIBB,² AND DAVID PENNY²

¹Henry Wellcome Ancient Biomolecules Center, Department of Zoology, Oxford University, South Parks Road, Oxford OX1 3PS, United Kingdom;
E-mail: mattphillips73@hotmail.com

²Allan Wilson Center for Molecular Ecology and Evolution, Massey University, Palmerston North, New Zealand

Abstract.—Australasian marsupials include three major radiations, the insectivorous/carnivorous Dasyuromorphia, the omnivorous bandicoots (Peramelemorphia), and the largely herbivorous diprotodontians. Morphologists have generally considered the bandicoots and diprotodontians to be closely related, most prominently because they are both syndactylous (with the 2nd and 3rd pedal digits being fused). Molecular studies have been unable to confirm or reject this Syndactyla hypothesis. Here we present new mitochondrial (mt) genomes from a spiny bandicoot (*Echymipera rufescens*) and two dasyurids, a fat-tailed dunnart (*Sminthopsis crassicaudata*) and a northern quoll (*Dasyurus hallucatus*). By comparing trees derived from pairwise base-frequency differences between taxa with standard (absolute, uncorrected) distance trees, we infer that composition bias among mt protein-coding and RNA sequences is sufficient to mislead tree reconstruction. This can explain incongruence between trees obtained from mt and nuclear data sets. However, after excluding major sources of compositional heterogeneity, both the “reduced-bias” mt and nuclear data sets clearly favor a bandicoot plus dasyuromorphian association, as well as a grouping of kangaroos and possums (Phalangeriformes) among diprotodontians. Notably, alternatives to these groupings could only be confidently rejected by combining the mt and nuclear data. Elsewhere on the tree, *Dromiciops* appears to be sister to the monophyletic Australasian marsupials, whereas the placement of the marsupial mole (*Notoryctes*) remains problematic. More generally, we contend that it is desirable to combine mt genome and nuclear sequences for inferring vertebrate phylogeny, but as separately modeled process partitions. This strategy depends on detecting and excluding (or accounting for) major sources of nonhistorical signal, such as from compositional nonstationarity. [Base composition; combined data; marsupial; mitochondrial genome; phylogeny.]

Recent Australasian marsupials have evolved into a multitude of niches and provide an extraordinary parallel to the evolution of placental mammals. Aplin and Archer (1987) recognize four orders. Dasyuromorphia (approximately 65 species) includes the insectivorous/carnivorous marsupial mice/cats/wolves/anteater. Peramelemorphia (approximately 20 species) includes the semifossorial and omnivorous bandicoots and bilbies. The most numerous group, the largely herbivorous Diprotodontia (approximately 115 species) includes the possums, kangaroos, koala, and wombats, whereas Notoryctemorphia (marsupial moles) includes just two species in the genus *Notoryctes*. These groups, together with the monotypic Microbiothera (*Dromiciops gliroides*) from South America, make up the cohort Australidelphia, which was first recognized on the basis of tarsal morphology (Szalay, 1982).

The monophyletic status of Australidelphia, to the exclusion of the South American Didelphimorphia (opossums) and Paucituberculata (including caenolestids: shrew opossums), is the only interordinal grouping of marsupials with strong support from DNA sequence data (see Phillips et al., 2001; Amrine-Madsen et al., 2003; Nilsson et al., 2003). Our present focus is on determining the interrelations of the three major australidelphian radiations. The possibilities include:

Bandicoots plus diprotodontians (BDi).—This grouping is consistent with Syndactyla (Jones, 1923–1925), originally diagnosed by the possession of syndactylous hind feet, in which the 2nd and 3rd digits are reduced, equal in length, and fused (or at least held together basally within a common integument) (see Szalay, 1982; Hall, 1987). Retinal morphology (Chase and Graydon, 1990), upper molars with well-defined dilambdodonty (Marshall et al.,

1990), and intracranial extension of the squamosal (see Szalay, 1993) also support Syndactyla. Cladistic analyses of morphological data matrices unanimously favor this relationship (Springer et al., 1997b; Lou et al., 2003; Horovitz and Sánchez-Villagra, 2003). Support for Syndactyla among molecular data, on the other hand, has only been claimed for analyses of the nuclear genes, phosphoglycerate kinase (*Pgk-1*; Colgan, 1999) and recombination activating gene 1 (RAG1, Baker et al., 2004). However, resolution was poor in both studies and in the former case the bandicoots grouped with the dasyuromorphian, *Myrmecobius*, in some analyses.

Dasyuromorphians plus diprotodontians (DaDi).—This grouping is consistent with Eometatheria, in which bandicoots are excluded from the remainder of Australidelphia. Using DNA hybridization data, Kirsch et al. (1997) argued forcefully for this grouping. Although morphology mostly offers contradiction, Asher et al. (2004) consider the continuous facets of the lower ankle joint of eometatherians to be more derived than either the separate facets of didelphids/caenolestids or the intermediate condition in bandicoots. The strongest support for Eometatheria has come from mitochondrial (mt) protein-coding and RNA genes (see Krajewski et al., 1997; Burk et al., 1999). Signal for Eometatheria from concatenated mt genes was strong enough in the combined analyses of Springer et al. (1998) and Asher et al. (2004) to overwhelm largely contradictory signals (see BDi above and BDa below) from the interphotoreceptor retinoid binding protein (IRBP) gene in the former study and from three nuclear genes (IRBP, protamine P1, and *Pgk-1*) plus morphological data in the latter study.

Bandicoots plus dasyuromorphians (BDa).—A “bandicoot-dasyuromorphian” relationship has long

been entertained by mammalogists (e.g., Thomas, 1888; Hughes, 1965). Potentially diagnostic characters for this grouping (see Horovitz and Sánchez-Villagra, 2003), such as the possession of three pairs of lower incisors, a broadened ectotympanic, and a teat count (5–8) intermediate between that of didelphids and diprotodontids, may be plesiomorphic for Australasian marsupials. Accordingly, morphological assessments within cladistic (e.g., Lockett, 1994), and “evolutionary inference” (e.g., Szalay and Sargis, 2001) frameworks have dismissed a bandicoot-dasyuromorphian relationship in favor of Syndactyla. However, with a few exceptions (Retief et al., 1995; Kirsch et al., 1997; Baker et al., 2004), studies of nuclear DNA, or nuclear-encoded proteins, have provided cautious support for a bandicoot-dasyuromorphian grouping. Examples come from serology (Kirsch, 1977), albumin microcomplement fixation (e.g., Baverstock et al., 1987), some DNA hybridization studies (e.g., Westerman, 1991), and IRBP sequences (Springer et al., 1997a; Jansa and Voss, 2000). The most convincing support for grouping dasyuromorphians and bandicoots to the exclusion of diprotodontian comes from the Bayesian inference analysis of Amrine-Madsen et al. (2003) on a concatenation of five nuclear genes: IRBP, von Willebrand factor (vWF), apolipoprotein B (APOB), RAG1, and breast and ovarian cancer susceptibility gene 1 (BRCA1). However, Bayesian posterior probabilities do not adequately account for stochastic error (see Suzuki et al., 2002; Douady et al., 2003; Simmons et al., 2004). Hence, it is notable that neither Eometatheria nor Syndactyla could be rejected even at $P \leq 0.50$ in their maximum likelihood (ML) analysis.

A recent study by Nilsson et al. (2004) in which a number of new marsupial mt genomes are presented deserves attention. Their ML results favored a bandicoot-dasyuromorphian relationship, though the maximum parsimony (MP) and distance analyses instead favored a close relationship between bandicoots and diprotodontians. The statistical significance of the ML results are difficult to gauge because full bootstrapping was not used to explore sampling effects and significance testing compared alternative trees of interest, rather than the alternative trees with the highest likelihoods.

It is possible that relationships among the major groups have been blurred by additional stochastic error and conflicting signals associated with the inclusion of *Dromiciops* and *Notoryctes* as isolated taxa for which long branches cannot be broken. In order to explore this, we run analyses with and without their inclusion. By implication, the primary hypotheses being tested, bandicoots plus diprotodontians (BDi), dasyuromorphians plus diprotodontians (DaDi), and dasyuromorphians plus bandicoots (BDa) define relationships irrespective of the placements of *Dromiciops* and *Notoryctes*. Hence, BDi and DaDi are respectively consistent with, but not equivalent to, Syndactyla (which typically includes *Notoryctes*, e.g., Szalay, 1993) and Eometatheria (which includes both *Dromiciops* and *Notoryctes*).

With the intention of resolving the phylogenetic stalemate on the interrelations of the three major Australasian marsupial groups, we have sequenced three new marsupial mt genomes. These are from a spiny bandicoot (*Echymipera rufescens*) and two dasyuromorphians, a fat-tailed dunnart (*Sminthopsis crassicaudata*) and a northern quoll (*Dasyurus hallucatus*), and are analyzed alongside previously sequenced marsupial mt genomes and a concatenated nuclear protein-coding data set (IRBP, vWF, APOB, RAG1, and BRCA1).

Inferring Marsupial Phylogeny from Mitochondrial Genome Sequences

Phylogenetic analyses of concatenated nuclear gene sequences from placental mammals (e.g., Madsen et al., 2001; Murphy et al., 2001) have, by comparison with mitochondrial efforts (e.g., Lin et al., 2002b; Arnason et al., 2002), cast doubt on the feasibility of reconstructing mammalian interordinal phylogeny from mt genomes. Critical examinations (e.g., Curole and Kocher, 1999; Springer et al., 2001; Corneli, 2002; Schmitz et al., 2002b; Phillips and Penny, 2003) have pointed to saturation-related lack of resolution and nucleotide composition bias as the most serious problems facing mitogenomic phylogenetics. However, an earlier study of marsupial mt genomes (Phillips et al., 2001) showed that phylogenetic signal erosion and composition bias are particularly pronounced at protein 3rd codon positions. This appears to be true of metazoan mt genomes in general and exclusion of 3rd codon positions has increased phylogenetic resolution in numerous studies (e.g., Springer et al., 1999; Lin et al., 2002a; Schwartz et al., 2004). Another approach is to recode purines (A, G) as R and pyrimidines (C, T) as Y. The benefits of RY-coding are twofold. Firstly, it excludes the major sources of mammalian mt genome compositional differences, which relate to biases among purine and particularly among pyrimidine transitions (Gibson et al., 2005). Secondly, RY-coding reduces phylogenetic signal erosion at deep levels by hiding transitions, which saturate rapidly relative to transversions in mt data particularly (see Phillips and Penny, 2003). In light of these observations, we examine whether excluding or recoding putatively biased data is warranted, given the information loss this incurs, and whether the mt and nuclear data should be analyzed together or separately.

A major aim is to assess the potential for compositional nonstationarity to result in an incorrect topology being inferred. Stochastic tests such as the compositional χ^2 test are commonly used for this purpose (e.g., Amrine and Springer, 1999; Negrisolo et al., 2004). As explained in Phillips et al. (2004), these tests only indicate the presence of compositional heterogeneity beyond that expected under a specified level of sampling error and are not truly comparable between data sets because their statistical power depends on factors such as the number of variable sites. Others (e.g., Ho and Jermini, 2004) have articulated the need instead for simple methods for inferring whether any bias is in fact sufficient to

mislead tree reconstruction. Herein we build upon previous work by Lockhart et al. (1994) and Phillips and Penny (2003) to use minimum evolution (ME) trees derived from pairwise base-frequency differences among taxa to assess the topological nature and magnitude of compositional bias. Comparing the resulting ME scores with those derived from standard (absolute, uncorrected) distances allows us to infer the potential for signal related to composition bias to overcome the "historical signal" at a particular node and so result in incorrect tree reconstruction.

We argue that composition bias is sufficient to mislead tree reconstruction for both the mt protein-coding and RNA data sets. However, these data sets pass the base-frequency versus standard (BF/S) distance test after a combination of excluding and recoding (Y and RY-coding) the partitions that are the major sources of compositional heterogeneity. We subsequently argue that combining such "reduced-bias" mitochondrial sequences and nuclear sequences is desirable for inferring marsupial phylogeny. In addition, we briefly discuss the implications of finding a bandicoot-dasyuromorphian relationship on our understanding of the initial diversification of Australasian marsupials.

MATERIAL AND METHODS

Mitochondrial Genome Sequencing

Total genomic DNA was extracted from liver tissue for all three marsupials using the High Pure PCR Template Preparation Kit from Roche Diagnostics. Mitochondrial DNA was amplified in fragments longer than 5 kb, using the Expand Long Template PCR Kit (Roche Diagnostics). Long fragments were nebulized and subcloned using the TOPO Shotgun Subcloning Kit from Invitrogen. Clones with inserts greater than 800 bp were sequenced using universal forward and reverse primers and standard protocols. The reactions were run on an ABI 3730 capillary machine at the AWC Genome Sequencing Centre, Massey University, Palmerston North. We used Sequencher (Gene Codes Corporation) for assembly and editing and for cloned sequences we used consensus sequence from at least three clones. Gaps were covered by direct sequencing of short PCR products generated from the original long templates, using primers from our vertebrate database.

The complete mitochondrial genomes of the new marsupials are available on GenBank: northern quoll (*Dasyurus hallucatus*, AY95973), fat-tailed dunnart (*Sminthopsis crassicaudata*, AY95974), and spiny bandicoot (*Echymipera rufescens*, AY95975). All the new genomes exhibit the gene order and other features typical for marsupials (see Janke et al., 1997; Nilsson et al., 2003); a nonfunctional tRNA^{Lys}, the use of the anticodon GCC for tRNA^{Asp}, and an unusually long sequence of 82 to 87 bp between tRNA^{Trp} and tRNA^{Asn}, which contains the sequence for the origin of replication for the light (L) strand. Two of the genomes contain repeat sequences in the control/noncoding region: the northern quoll sequence has around 170 bp of dimeric repeats (gt interspersed

with at in an apparently random fashion); the dunnart has 29 copies of a 12bp repeat (tgcgtgtgtgta).

Data Matrices

We analyzed the mitochondrial rRNA, tRNA, and 12 H-strand encoded protein-coding genes from 14 marsupials. Only taxa for which both mt and nuclear sequences are available were used, to ensure both the combinability of these mt and nuclear sequences for the BF/S distance comparisons and the comparability of these data sets for signal retention and base composition heterogeneity metrics. Mitochondrial phylogenetic studies typically exclude NADH6 (e.g., Janke et al., 1996; Reyes et al., 2004) because of its L-strand encoding, which for substitution models disrupts the relationship between strand-specific mutational pressures and the selection pressures on the protein sequence, relative to the H-strand proteins.

Alongside the three new mitochondrial sequences, we have included previously published marsupial mt sequences from a wallaroo (*Macropus robustus*), brush-tail possum (*Trichosurus vulpecula*), common wombat (*Vombatus ursinus*), northern brown bandicoot (*Isodon macrourus*), phascogale (*Phascogale tapoatafa*), marsupial mole (*Notoryctes typhlops*), monito del monte (*Dromicops gliroides*), Chilean shrew opossum (*Rhyncholestes raphanurus*), silky shrew opossum (*Caenolestes fuliginosus*), Virginia Opossum (*Didelphis virginiana*), and gray short-tailed opossum (*Monodelphis domestica*). Employing the latter four "Ameridelphian" taxa as the outgroup better serves the aim of inferring relations among the australidelphians than does inclusion of additional placental and monotreme mammals (which diverged from the ingroup twice as long ago; see Nilsson et al., 2003). Inclusion of nonmarsupials substantially decreased the number of alignable sites. Furthermore, including a marsupials-only outgroup appears to be in the interests of modeling substitution among the ingroup. ML estimates of model parameters are similar when obtained for the the ingroup alone or for the ingroup plus marsupial outgroups, but differ considerably when placentals or monotremes are included (particularly for the rate matrix and rate heterogeneity across sites).

As for the original presentation of the nuclear genes (Amrine-Madsen et al., 2003) and in order to be fully compatible with the mt data, some combined sequences are chimeric. For example, the kangaroo subfamily Macropodinae is represented by both *Macropus* (mtDNA, vWF, BRCA1) and *Dendrolagus* (APOB, RAG1, IRBP). GenBank accession details for the mt genomes and nuclear sequences are provided in Table 1.

Sequences were aligned manually within Se-AL v1.0a1 (Rambaut, 1996). With ambiguous sites excluded, the 12 mt protein-coding genes contribute 10,686 sites, the mt RNA genes contribute 3313 sites, and the five nuclear protein-coding genes contribute 5634 sites, for a combined total of 19,633 sites. The data matrix is available online at <http://www.systematicbiology.org>. We designate the DNA sequences on the basis of source: Nuc, Mt-ptn, Mt-RNA, and simply Mt for the combined mitochondrial data.

TABLE 1. GenBank accession numbers. Five combined sequences (Didelphinae, Sminthopsinae, Peramelinae, Phalangeridae, and Macropodinae) are chimeric, using sequences from related genera.

Taxa	Genera	Common name	mt genome	IRBP	vWF	APOB	RAG1	BRCA1
Didelphinae	<i>Didelphis</i>	Virginia opossum	Z29573	Z11814	AF226848	AF548432		AF497261
	<i>Lutreolina</i>	Thick-tailed opossum					AY243390	
Monodelphis	<i>Monodelphis</i>	Gray short-tailed opossum	AJ508398	AF257694	AY243415	AY243431	U51897	AY243453
<i>Caenolestes</i>	<i>Caenolestes</i>	Silky shrew opossum	AJ508400	AF025381	AY243403	AY243418	AY243384	AF355794
<i>Rhyncholestes</i>	<i>Rhyncholestes</i>	Chilean shrew opossum	AJ508399	AY24340	AY243416	AY243432	AY243399	AY243454
<i>Dromiciops</i>	<i>Dromiciops</i>	Monito del monte	AJ508402	AF025384	AY243407	AY243423	AY243389	AY243446
<i>Notoryctes</i>	<i>Notoryctes</i>	Marsupial mole	AJ639874	AF025385	AY243408	AY243424	AY243391	AY243447
Sminthopsinae	<i>Sminthopsis</i>	Dunnart	AY95974		AY243413	AY243428		
	<i>Planigale</i>	Pygmy marsupial mouse		AY243438			AY243396	AY243451
<i>Dasyurus</i>	<i>Dasyurus</i>	Quoll	AY95973	AY243439	AY243414	AY243430	AY243398	AY243452
<i>Phascogale</i>	<i>Phascogale</i>	Phascogale	AJ639869	AF025382	AY243412	AY243427	AY243395	AF355795
<i>Echymipera</i>	<i>Echymipera</i>	Spiny bandicoot	AY95975	AF025383	AY243405	AY243420	AY243386	AF355796
Peramelinae	<i>Perameles</i>	Long-nosed bandicoot		AY243437	AY243411	AY243426	AY243394	AY243450
	<i>Isodon</i>	Short-nosed bandicoot	AF358864					
<i>Vombatus</i>	<i>Vombatus</i>	Common wombat	AJ304826	AF025386	AF497260	AY243429	AY243397	AF284031
Phalangeridae	<i>Phalanger</i>	Cuscus		AY243436	AY243410	AF548431	AY243393	AY243449
	<i>Trichosurus</i>	Brush-tail possum	AF357238					
Macropodinae	<i>Macropus</i>	Kangaroo	Y10524		AJ224670			AF284033
	<i>Dendrolagus</i>	Tree kangaroo		AY243435		AY243422	AY243388	

The analyses were performed without gapped-sites, leaving 17,936 sites for a 12-taxon data set and 17,864 sites for a 14-taxon data set, which respectively exclude and include the two phylogenetically isolated taxa (*Dromiciops* and *Notoryctes*). So, for example, the 12-taxon mitochondrial RNA data set is designated Mt-RNA_(12T), whereas the 14-taxon combined mitochondrial and nuclear data set for all positions is MtNuc_(14T). We focus on the nucleotide sequence primarily because so much of the signal (particularly in the nuclear data) occurs as synonymous changes. Furthermore, available models for amino acid evolution (for use in ML significance testing) are based on empirical studies (e.g., Adachi and Hasegawa, 1996) primarily of nonmarsupials, whereas analysis of nucleotides allows for more realistic models of sequence evolution to be inferred from the marsupial data.

Examination of Phylogenetic Signal

The extent of substitution saturation is usually identified with saturation plots, in which ML (corrected) branch-lengths are plotted against uncorrected branch-lengths (or distances). However, metrics for comparing saturation plots are not well developed and their interpretation is questionable (Yang, 1998; Baker et al., 2001). In order to overcome some of the shortcomings of saturation plots and further explore phylogenetic signal retention/erosion, we use a stemminess metric. In addition, we use relative composition variability (RCV; Phillips and Penny 2003) as a measure of the magnitude (as opposed simply to presence) of compositional nonstationarity and provide a simple test (using base-frequency difference trees) of the potential for this to mislead tree reconstruction. These metrics and tests are described below.

Stemminess.—Fiala and Sokal (1985) defined stemminess as the proportion of overall tree-length that internal branches contribute. As for Phillips et al. (2001), we em-

ploy this definition on minimum evolution (ME) trees derived from uncorrected distances. Stemminess is useful for comparing phylogenetic signal erosion between partitions, because (given the same topology) greater erosion results in shorter internal branches relative to their attendant external branches (hence, lower stemminess).

Relative composition variability (RCV; Phillips and Penny, 2003).—RCV is the average variability in composition between taxa; for nucleotides this is:

$$RCV = \frac{\sum_{i=1}^n (|A_i - A^*| + |T_i - T^*| + |C_i - C^*| + |G_i - G^*|)}{n \cdot t}$$

A_i , T_i , C_i , and G_i are the frequencies of each nucleotide for the i th taxon. A^* , T^* , C^* , and G^* are averages across the n taxa, and t is the number of sites.

Compositional nonstationarity was also examined with χ^2 tests, performed within PAUP* 4.0b10 (Swofford, 2002). Parsimony-uninformative sites cannot directly contribute to phylogenetic biases and effectively dilute compositional heterogeneity and were excluded from both chi-square and RCV calculations.

Base-frequency (BF) distance trees.—ME trees were constructed from a matrix of pairwise base-frequency distances to examine composition bias with respect to specific nodes and assess the potential for this to affect phylogenetic inference. The basic idea is to compare tree-length differences between alternative topologies for trees reconstructed from both standard (absolute, uncorrected) distances and base-frequency differences. In this way, Phillips and Penny (2003) showed that a bias among pyrimidines (C versus T) could account for incorrect rooting of the mammal tree from mt protein-coding genes. Here we extend this methodology to incorporate overall nucleotide composition bias. Base-frequency distances are half the sum of absolute frequency differences

between taxon pairs for each nucleotide category. So the pairwise base-frequency (BF) distance between taxa i and j is:

$$\text{BF distance} = (|A_i - A_j| + |T_i - T_j| + |C_i - C_j| + |G_i - G_j|)/2$$

A_i , T_i , C_i , and G_i , and A_j , T_j , C_j and G_j are the frequencies of each nucleotide for the i th and j th taxa, respectively. Dividing by two is necessary for ME trees on BF distances to be comparable with ME trees on standard distances. This is because a substitution at a site in taxon i that previously had the same base as for taxon j will result in one unit of standard distance, but two units of base-frequency distance. BF distance matrices were calculated within Microsoft Excel, using base frequency data from PAUP*. Parsimony-uninformative characters were excluded from base-frequency distance calculations, as these cannot explain ME differences between standard distance trees. For all of the ME trees, any negative branch-lengths were treated as absolute values for computing tree-length.

Groupings resolved (in agreement with prior molecular and morphological studies) with maximum 100% BPP, ML, and MP support for all data sets and not being tested for composition bias were fixed for the base-frequency versus standard (BF/S) distance comparisons. These constraints include Didelphidae, Caenolestidae, Peramelidae, Dasyuridae, and *Dasyurus* plus *Phascogale* and ensure that the branching patterns adjacent to nodes being tested for composition bias are the same as for our primary phylogenetic analyses. Unless otherwise stated, we focus on the results of our primary (12-taxon) data set, for which artifacts owing to the inclusion of the phylogenetically isolated taxa (*Dromiciops* and *Notoryctes*) can be discounted. However, in Figure 1 we see that the BF distance tree for MtNuc_(14T) shows that *Dromiciops* and *Notoryctes* group closely with diprotodontians and dasyurids with respect to base composition, whereas the bandicoots are similar to the “ameridelphian” didelphids and caenolestids.

Using Base-Frequency Distance Trees to Infer the Potential Influence of Composition Bias

Composition χ^2 testing (Table 2) reveals highly significant ($P < 0.0001$) base-frequency heterogeneity among each of the Mt-ptn, Mt-RNA, and Nuc data sets

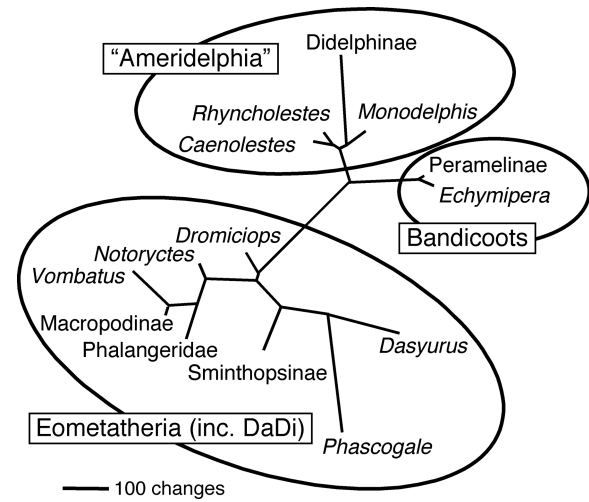


FIGURE 1. MtNuc_(14T) base frequency (BF) distance tree. Compositional bias favors Eometatheria (including a dasyuromorphian plus diprotodontian association), which is separated from bandicoots, caenolestids, and didelphids by the longest branch on the tree.

under standard (nt) coding. This only tells of the presence of compositional heterogeneity. BF/S distance comparisons allow us to infer the topological nature and the magnitude of composition bias affecting specific hypotheses—in the present case, BDa versus BDi versus DaDi.

In order for composition bias to result in an incorrect topology being reconstructed on standard distances (or other methods), we generally expect that the BF distance tree also favors that putatively incorrect topology (condition 1). In addition, support for that topology on BF distance would have to be of greater magnitude than support on standard distance (condition 2). This second condition arises because at least a small proportion of the BF signal will be attributable to sites that are uninformative with respect to the node of interest, even though sites that are uninformative for the tree as a whole are excluded. In Table 3, condition 1 is met where both the standard tree and BF tree favor the same topology and condition 2 is met where the BF/S ratio is >1.00 . The BF/S ratio is the ME score difference between the favored (f) and alternative (a) trees on BF distance, divided by the ME score difference between these trees on standard (absolute, uncorrected) distances:

$$\text{BF/S ratio} = \text{ME}(\text{BF})_{f-a} / \text{ME}(\text{Std.})_{f-a}$$

TABLE 2. Statistics indicating phylogenetic signal retention (stemminess) and compositional heterogeneity (composition χ^2 test P -values and relative composition variability [RCV]) for the 12-taxon data sets. These are shown in a for the original nuclear (Nuc) and mitochondrial (Mt-ptn, Mt-RNA) data sets, which have standard nucleotide coding. In b, sources of potentially misleading composition bias (indicated by BF/S distance comparisons, see Table 3) are excluded, such that RB-Mt-RNA is RY-coded and for RB-Mt-ptn, 3rd codon positions are excluded, and 1st codon positions are Y-coded.

	Data set	Informative sites	Stemminess	Composition χ^2 test	RCV	RCV/stemminess
a.	Nuc	1085	0.4403	<0.0001	0.1002	0.2276
	Mt-ptn	4170	0.0861	<0.0001	0.1244	1.4448
	Mt-RNA	574	0.2171	<0.0001	0.1173	0.5403
b.	RB-Mt-ptn	915	0.2254	<0.0001	0.0556	0.2467
	RB-Mt-RNA	235	0.2866	0.0218	0.0723	0.2523

TABLE 3. Potential influence of composition bias among the 12-taxon nuclear (a), mt protein (b–d), and mt RNA (e–f) data sets on tree selection between either bandicoot-diprotodontian (BDi), dasyuromorphian-diprotodontian (DaDi), or bandicoots-dasyuromorphian (BDa) associations. Which of these trees' minimum evolution (ME) favors is shown for both standard, uncorrected distances (Std. tree) and base-frequency distances (BF tree)? When both favor the same topology, if the ME difference between that topology (f) and an alternative (a) is greater on BF distances than on Std. distances (i.e., the BF/S ratio = $ME(BF)_{f-a} / ME(Std.)_{f-a} > 1.00$), then compositional heterogeneity alone may be sufficient to mislead tree selection. As an example, for data set b (nt-coded Mt-ptn), comparisons of the BF tree (tree-f) with both alternatives (BDa and BDi) produce BF/S ratios > 1.00 , so we recommend further bias reduction. Inclusion (coding) indicates the data included, for which numbers refer to codon positions, and whether the data is standard (nt), Y-coded (Y), or RY-coded (RY).

	Data set	Inclusion (coding)	Std. tree	BF tree	BF/S ratio	Suggestion
a.	Nuc	All (nt)	BDa	DaDi	—	Use
b.	Mt-ptn	All (nt)	DaDi	DaDi	1.84 over BDa, 3.90 over BDi	Reduce bias
c.	Mt-ptn	1, 2 (nt)	DaDi	DaDi	4,171.53 over BDa, 3.16 over BDi	Reduce bias
d.	Mt-ptn	1 (Y), 2 (nt)	BDa	DaDi	—	Use
e.	Mt-RNA	All (nt)	DaDi	DaDi	0.19 over BDa, 1.95 over BDi	Reduce bias
f.	Mt-RNA	All (Y)	DaDi	DaDi	0.58 over BDa, 163.73 over BDi	Reduce bias
g.	Mt-RNA	All (RY)	BDi	BDi	0.73 over BDa, 0.07 over DaDi	Use

In the present context, the BF/S ratio is relevant only when condition 1 is met. In Table 3 we see that base-frequency trees favor DaDi for each of the nt-coded data sets. Standard distance supports a conflicting topology for the nuclear data set. Hence, we can be confident that the topology favored (BDa) on standard distance is not an artifact of the base-frequency bias (which favors DaDi). In contrast, we need to be aware that for both of the full nt-coded mitochondrial data sets, Mt-ptn (b) and Mt-RNA (e), the inferred interrelations of the major Australasian orders may be an artifact of composition bias. In each case DaDi is favored by both BF and standard distances (meeting condition 1) and the BF/S ratio is well above 1.00 in comparing the ME score for this topology with that of at least one of the alternatives (meeting condition 2).

Among marsupial (Phillips et al. 2001) and other mammalian (e.g., Gibson et al., 2005) mt genomes, base-frequency biases are most pronounced among pyrimidines, then among purines, and decrease from 3rd to 1st to 2nd positions for protein-coding sequences. Hence, as a strategy for reducing bias among Mt-ptn and Mt-RNA, we consider in successive steps, Y-coding (C, T lumped as Y) as advocated by Gibson et al. (2005), RY-coding (C, T lumped as Y and A, G lumped as R), and data exclusion. The procedure continues until either condition 1 or 2 is no longer met.

In Table 3c we see that even with 3rd codon positions excluded from Mt-ptn, the base-frequency bias may be strong enough to result in the incorrect tree being inferred on standard distances. Here, $ME(BF)_{DaDi-BDa}$ and $ME(BF)_{DaDi-BDi}$ are similar, 82.93 and 101.09, respectively. However, it is differences on standard distances that account for the disparity in BF/S ratios between DaDi and the two alternatives. $ME(Std.)$ for DaDi is only lower than for BDa by 0.01988 (giving a BF/S ratio of 4171.53), whereas $ME(Std.)_{DaDi-BDi}$ is 31.97 (giving a BF/S ratio of 3.16).

In addition to excluding 3rd codon positions, to be confident that phylogenetic analysis of the Mt-ptn data do not simply reflect base-frequency biases, it was also necessary to Y-code the 1st codon positions (Table 3d).

Furthermore, Table 3 (f, g) shows that RY-coding is warranted for Mt-RNA as purine transitions provide a strong enough composition bias signal to result in an incorrect topology being inferred. Although both BF and standard distances favor BDi when the RNA data are RY-coded, condition 2 is not met because the BF/S ratios are < 1.00 . These reduced-bias data sets are henceforth referred to as RB-Mt-ptn and RB-Mt-RNA. The coding of the nine data sets used for phylogenetic analysis is described in Table 4.

Phylogenetic Analysis

Following the examination of saturation and base composition, we inferred the marsupial phylogeny from the original and reduced-bias data sets. Maximum parsimony (MP), minimum evolution (ME), and maximum likelihood (ML) analyses were performed within PAUP*4.0b10. Bootstrapping (1000 replicates) applied heuristic searches from random starting trees. Substitution was modeled for the ML bootstrap analyses using GTR+I+ Γ_4 , the model category favored by AIC scores derived for each of the concatenated (nonpartitioned) data sets using ModelTest 3.06 (Posada and Crandall, 1998). ModelTest parameter estimates (which

TABLE 4. Partition inclusion and coding regimes for the data sets used in phylogenetic analyses. Bayesian inference and ML significance testing models are partitioned, ML bootstrap analyses are not. Standard (AGCT) coding is assumed unless otherwise stated. Y-coding lumps pyrimidines (C,T) and RY-coding lumps purines (A, G) and pyrimidines (C, T).

Data set	No. of partitions	Partitions (and coding)
Nuc	3	Nuclear protein codons: 1, 2, 3
Mt-ptn	3	mt protein codons: 1, 2, 3
Mt-RNA	2	mt RNA: stems, loops
Mt	5	Mt-ptn and Mt-RNA combined
MtNuc	8	Mt-ptn, Mt-RNA and Nuc combined
RB- Mt-ptn	2	mt protein codons: 1(Y), 2
RB- Mt-RNA	2	mt RNA: stems (RY), loops (RY)
RB-Mt	4	RB-Mt-ptn and RB-Mt-RNA combined
RB- MtNuc	7	RB-Mt-ptn, RB-Mt-RNA and Nuc combined

are based on an NJ tree) were used in heuristic searches to find an ML tree on which parameters were reestimated for ML bootstrapping. Bayesian inference (MrBayes 3.0b4; Huelsenbeck and Ronquist 2001) also employed GTR+I+ Γ_4 models, except for the RY-coded partitions (F81+I+ Γ_4), which only have one substitution type. Four chains were run for 3,000,000 (6,000,000 for 14-taxa) generations with trees being sampled every 200 generations. The burn-in for each MrBayes run was set at 500,000 (2,000,000 for 14-taxa) generations, after having checked for convergence and examined estimated sample sizes for the parameter estimates (using Tracer v1.0 Rambaut and Drummond, 2003).

Additional MP analyses were performed within PAUP* for examining congruence between the Mt and Nuc data sets by partition homogeneity testing (ILD test of Farris et al., 1994). Each partition homogeneity test employed 1000 heuristic replicates.

MP, ML, and Bayesian inference on each of the 12- and 14-taxon MtNuc and RB-MtNuc data sets resolve for all but between one and five nodes with 100% bootstrap (and Bayesian posterior probability, BPP) support. Support among alternative topologies was further examined with KH (Kishino and Hasegawa, 1989) and approximately unbiased (AU; Shimodaira, 2002) tests, using the RELL method (100,000 replications) within CONSEL (Shimodaira and Hasegawa, 2001). The AU test is related to the SH test (Shimodaira and Hasegawa, 1999) and has been developed in order to overcome tree selection biases that affect the latter test, particularly when many topologies are being simultaneously compared. The ML significance tests were applied to the separate and combined RB-Mt and Nuc data sets, with all parameters ML-optimized for each tree hypothesis.

For ML significance testing (as for Bayesian inference) we partitioned on the basis of source (mt, nuclear) as well as codon position (1, 2, 3) for proteins and secondary structure (stems, loops) for RNA. The number of (fully independent) partitions and their coding is shown in Table 4 for each data set. Because all of the sites are not evolving under the same processes, partitioning the data for ML analysis allows for more accurate modeling (e.g., Yang, 1996; Caterino et al., 2001). In these analyses, standard nt-coded partitions were modeled under GTR+I+ Γ_4 , whereas Y-coded and RY-coded partitions were modeled under 3 (A-G, A-Y, G-Y) and 1 (R-Y) substitution-type subsets of GTR+I+ Γ_4 , respectively.

The benefits of partitioning into more homogeneously evolving subsets can be offset by estimating parameters from fewer sites, reducing the signal to variance ratio (DeBry, 1999; Pupko et al., 2002). For RB-MtNuc_(14T), the additional 192 branch-length and substitution process parameters that partitioning incurs is easily justified under the Akaike Information Criterion (AIC; $-2\ln L + 2 \times \text{no. of free parameters}$) and improves $-\ln L$ by 2096.02. This translates into a likelihood "partitioning advantage" (the difference between partitioned and unpartitioned $-\ln L$, divided by the unpartitioned $-\ln L$) of 4.1%. This increases to 7.3% for MtNuc_(14T).

RESULTS

Preliminary Data Examination

The potential for base frequency heterogeneity to influence tree selection depends on both the magnitude of the composition bias and the degree to which substitution saturation has eroded "historical" phylogenetic signal. We have employed relative composition variability (RCV) as a measure of the magnitude of composition bias and stemminess as a measure of phylogenetic signal retention. These measures (see Table 2a) help to explain why Nuc appears (from the BF/S distance comparisons) to be less susceptible to composition artifacts than the nt-coded mt data sets. Nuc has a lower RCV and more than double the stemminess of either Mt-ptn or Mt-RNA.

In Table 2b we see that excluding 3rd codon positions, Y-coding 1st positions, and RY-coding RNA results in substantial increases in stemminess and decreases in RCV for both bias-reduced mt data sets. Indeed, for RCV/stemminess, RB-Mt-ptn (0.2467) and RB-Mt-RNA (0.2523) are almost as low as for Nuc (0.2276). This combined factor proxy for the potential influence of composition bias on tree selection is considerably higher for original Mt-ptn (1.448) and Mt-RNA (0.5403).

Incongruence testing provides further evidence for the bias reduction strategy enhancing the fidelity of the mt data. ILD testing with the 12-taxon data sets find the original Mt data to be incongruent with Nuc ($P = 0.007$). Given the dominance of the Mt data in terms of informative sites in this analysis, the significance levels may even be conservative (see Dowton and Austin, 2002). In contrast, the ILD test indicates that the RB-Mt and Nuc data are fully congruent ($P = 1.000$).

The situation is less clear for the 14-taxon data sets. The significance level at which congruence between RB-Mt and Nuc remains high ($P = 0.609$), but is similar to that for Mt and Nuc ($P = 0.728$). The latter result is curious because the inclusion of *Dromiciops* and *Notoryctes* in fact tends to increase support both for DaDi from Mt (not shown) and for BDa from Nuc (as will be seen later). It is likely that the 14-taxon partition homogeneity results are explained by the additional instability or "noise" these taxa lend the trees, which reduces the power of the ILD test to reject congruence.

Phylogenetic Reconstruction

The primary (12-taxon combined mt and nuclear) data sets for examining the interrelations of the Australasian marsupial orders resolve seven of the nine groupings with maximum BPP and bootstrap support (Fig. 2). Each of these, Didelphidae, Caenolestidae, Australidelphia, Diprotodontia, Peramelidae, and Phascogalinae plus Dasyurinae, is well corroborated in previous studies (e.g., Amrine-Madsen et al., 2003; Cardillo et al., 2004). The influence of the efforts to reduce composition bias is most prominent in resolving groupings that have been more elusive. ML and MP bootstrap support for a macropodid plus phalangerid grouping (Phalangeriformes) increases from 84% and 85%, respectively, for MtNuc to 97% and 100% for RB-MtNuc. The most

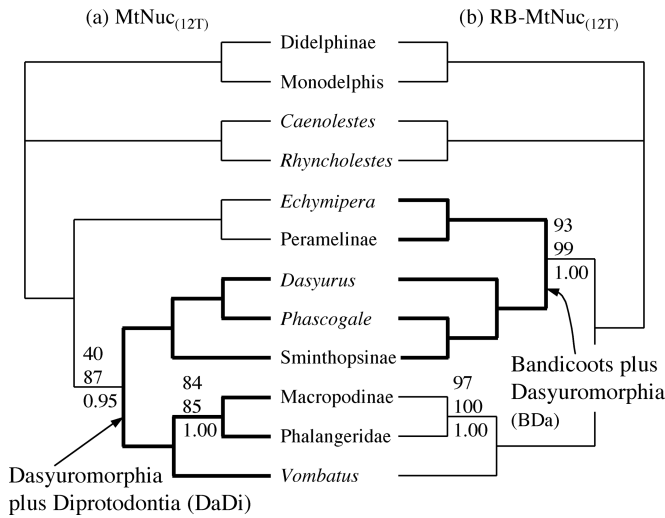


FIGURE 2. Marsupial phylogenies inferred from the MtNuc_(12T) original (a) and reduced-bias (b) data sets. Bootstrap supports are shown for ML (top) and MP (middle), with BPP values (bottom). Values are not shown for nodes that receive maximum support in all analyses.

profound impact of the bias reduction is a shift in support for DaDi (MP = 87%, BPP = 0.95; Fig. 2a) to support for BDa (MP = 99%, BPP = 1.00; Fig. 2b). Notably, ML bootstrap support favors BDa even for MtNuc at 60%, though this rises to 93% for RB-MtNuc.

The results for the individual data sets are teased apart in Table 5. DaDi is favored in analyses of both of the original Mt-ptn and Mt-RNA data sets. However, as a consequence of the bias reduction, ML and BPP support for BDa increase from 15% and 0.00 to 79% and 1.00, respectively, for the mt proteins and from 14% and 0.00 to 52% and 0.55 for the RNA. MP analyses tell essentially the same story (Appendix 1). The upshot of excluding and recoding data on the basis of potential for tree selection artifacts resulting from composition bias is that each of the 12-taxon Nuc, RB-Mt-ptn, and RB-Mt-RNA data sets favor the same tree.

The 14-taxon RB-MtNuc (Fig. 3) and MtNuc (not shown) data sets favor the same relationships as for their respective 12-taxon counterparts (Fig. 2). More specifically, the bias-reduced and original data sets respectively favor BDa and DaDi among the major Australasian groups. However, ML and MP bootstrap support is much reduced in the 14-taxon cases. This may be misleading

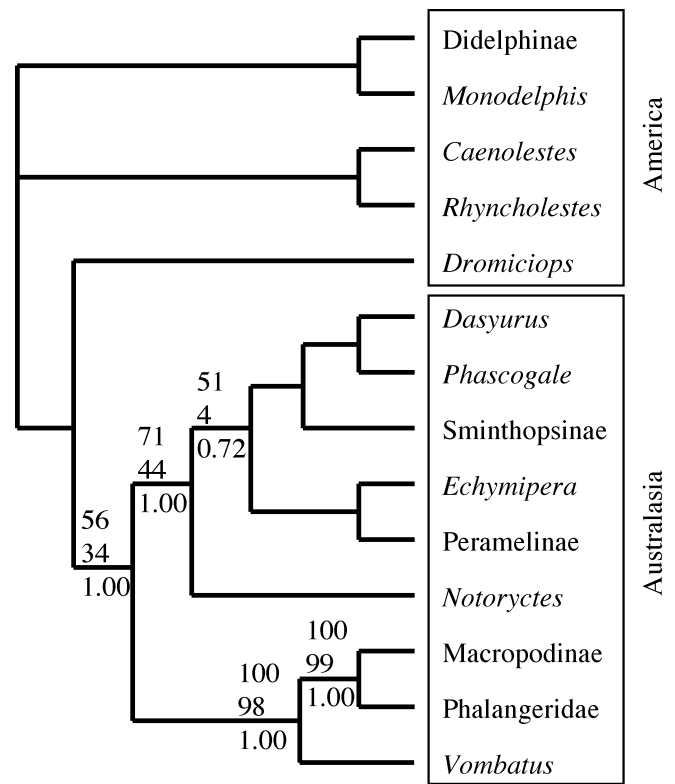


FIGURE 3. Marsupial phylogeny inferred from the reduced-bias 14-taxon combined data set, RB-MtNuc_(14T). Bootstrap supports are shown for ML (top) and MP (middle), with BPP values (bottom). These are not shown for nodes that receive maximum support in all analyses. All resolve for an monophyletic Australasian grouping.

and was predicted earlier. That the low resolution results from the instability of *Dromiciops* and *Notoryctes* on the trees can be shown by constraining their placements in accordance with the tree in Figure 3. With *Dromiciops* and *Notoryctes* constrained respectively to fall outside and within the Australasian taxa, BDa (including or excluding *Notoryctes*) receives 98% and 99% ML and MP bootstrap support for RB-MtNuc_(14T)—higher even than for the 12-taxon trees.

Phylogenetic Hypothesis Testing

Our ML hypothesis testing analyses have two key advantages. Firstly, in contrast to the bootstrap analyses,

TABLE 5. Maximum-likelihood bootstrap support (and Bayesian posterior probabilities) from the 12-taxon mitochondrial protein, mitochondrial RNA, and nuclear data sets for marsupial relationships, including for bandicoot-dasyuromorphian (BDa), bandicoot-diprotodontian (BDi), and dasyuromorphian-diprotodontian (DaDi) associations.

	Alternative associations among the major Australasian orders				
	BDa	BDi	DaDi	Australidelphia	Phalangeriformes
Mt-ptn	15 (0.00)	4 (0.00)	78 (1.00)	67 (0.99)	49 (1.00)
RB-Mt-ptn	79 (1.00)	4 (0.00)	16 (0.00)	96 (1.00)	96 (1.00)
Mt-RNA	14 (0.00)	1 (0.00)	79 (0.99)	96 (1.00)	52 (0.62)
RB-Mt-RNA	52 (0.55)	3 (0.17)	19 (0.27)	95 (1.00)	38 (0.73)
Nuc	89 (1.00)	1 (0.00)	10 (0.00)	100 (1.00)	97 (1.00)

TABLE 6. Log-likelihood differences between trees and their statistical significance under KH and AU tests for the reduced-bias mt (RB-Mt), nuclear (Nuc), and combined (RB-MtNuc) data sets. ML models are partitioned for proteins (by codon) and RNA (stems, loops). Twelve-taxon analyses (a: trees 1–9) compare affinities among the major Australasian orders and within Diprotodontia. Fourteen-taxon analyses (b: trees 10–15) compare australidelphian interordinal affinities. Each data set favors BDa (\pm *Notoryctes*). Among the 105 possibilities, the three highest $-\ln L$ trees are shown, as well as the best BDi and DaDi trees. P -values > 0.05 are bold. Taxon abbreviations: B, bandicoots (Peramelemorphia); Da, Dasyuromorphia; Di, Diprotodontia; Dr, *Dromiciops*; M, Macropodinae; P, Phalangeridae; N, *Notoryctes*; V, *Vombatus*.

	RB-Mt			Nuc			RB-MtNuc		
	$-\ln L$	KH	AU	$-\ln L$	KH	AU	$-\ln L$	KH	AU
a. 12-taxon trees									
1. BDa: (B, Da), (V, (M, P))	<26801.6>	—	—	<17658.4>	—	—	<44460.0>	—	—
2. BDi: (Da, (B, (V, (M, P))))	7.7	0.112	0.192	11.0	0.035	0.009	18.7	0.016	0.014
3. DaDi: (B, (Da, (V, (M, P))))	6.6	0.162	0.277	6.2	0.193	0.275	12.8	0.097	0.127
4. BDa: (B, Da), (P, (M, V))	12.5	0.080	0.147	14.3	0.032	0.026	26.9	0.009	0.023
5. BDi: (Da, (B, (P, (M, V))))	19.7	0.038	0.056	25.5	0.006	0.003	45.1	0.001	0.001
6. DaDi: (B, (Da, (P, (M, V))))	18.1	0.053	0.074	20.3	0.027	0.014	38.5	0.006	0.006
7. BDa: (B, Da), (M, (V, P))	15.9	0.021	0.034	12.2	0.059	0.079	28.1	0.007	0.012
8. BDi: (Da, (B, (M, (V, P))))	22.9	0.014	0.019	23.5	0.010	0.002	46.4	<0.001	<0.001
9. DaDi: (B, (Da, (M, (V, P))))	21.5	0.021	0.035	18.4	0.044	0.038	39.9	0.003	0.001
b. 14-taxon trees									
10. BDa: (Dr, (Di, (N, (B, Da))))	0.6	0.429	0.548	1.2	0.400	0.517	<49451.2>	—	—
11. BDa: (Dr, (Di, (B, (N, Da))))	<29796.3>	—	—	2.7	0.342	0.413	0.9	0.437	0.549
12. BDa: (Dr, (Di, (Da, (B, N))))	0.2	0.491	0.641	4.9	0.189	0.199	3.3	0.279	0.368
13. BDa: ((Dr, Di), (N, (B, Da)))	6.6	0.129	0.079	<19653.0>	—	—	4.7	0.235	0.294
14. BDi: (Dr, (Da, (Di, (B, N))))	5.3	0.257	0.263	17.4	0.037	0.013	20.9	0.034	0.033
15. DaDi: (Dr, ((Da, Di), (B, N)))	8.4	0.115	0.033	14.7	0.062	0.051	21.2	0.029	0.023

ML models are partitioned on the basis of codon positions and RNA secondary structure, for which substitution processes differ markedly. Secondly, type 2 error rates associated with Kishino-Hasegawa (KH) and approximately unbiased (AU) hypothesis tests can provide a closer reflection of sampling error than do bootstrap values (Shimodaira and Hasegawa, 1999), which in turn are far more faithful than typically overconfident BPP values (Suzuki et al., 2002; Gontcharov et al., 2004).

Nine 12-taxon trees attracted $\geq 1\%$ bootstrap (or BPP) support in at least one of the Nuc, Mt, or RB-Mt analyses (either MP, ML, or Bayesian). These include the nine combinations of the three possible sister groupings among the diprotodontians (Phalangeridae, Macropodinae, and *Vombatus*) with each of BDi, DaDi, and BDa. The results in Table 6a reveal that the same topology (as for Fig. 2b) is favored for each of the data sets RB-Mt_(12T), Nuc_(12T), and the combined RB-MtNuc_(12T). Log likelihood differences from that best tree and significance (P) values are also given in Table 6. Although the KH and AU tests provide similar results, the AU P -values are expected to be more reliable (Shimodaira, 2002) and were usually more conservative when significance was borderline. Hence, unless otherwise stated, we refer to the AU test P -values.

Although each of the 12-taxon analyses favors both a bandicoot-dasyuromorphian (BDa) relationship as well as Phalangeriformes, alternatives to both of these cannot be rejected at $P \leq 0.05$ for either the mitochondrial or nuclear data alone. Nonrejected alternatives are shown (in bold) in Table 6a for RB-Mt_(12T) and Nuc_(12T). DaDi is the strongest alternative to BDa for both data sets. Combining the mt and nuclear data provides clear statistical advantages. Individually, both data sets allow alternatives to Phalangeriformes and BDa to be rejected at $P \approx 0.075$ and $P \approx 0.275$, respectively. On the combined 12-

taxon data, Phalangeriformes and BDa alternatives are rejected at $P = 0.023$ and $P = 0.127$, respectively.

Table 6b shows the three trees (10, 11, 12) with the highest likelihoods for the 14-taxon mt, nuclear, and combined analyses. Each is consistent with a bandicoot-dasyuromorphian association among the major groups and places *Dromiciops* outside of the Australasian radiation. Trees with the highest likelihoods that are consistent with BDi (tree 14) and DaDi (tree 15) are also shown. The best BDi tree, which is 20.9 $-\ln L$ units inferior to the favored BDa tree, is in fact a variant of *Syndactyla*, in this case with bandicoots and *Notoryctes* grouping together. Note that the best DaDi tree (21.2 $-\ln L$ units inferior to the favored BDa tree) is not a variant of Eometatheria (Dasyuromorphia, Diprotodontia, *Dromiciops*, and *Notoryctes*), which fares worse still, 22.0 $-\ln L$ units inferior to BDa.

The most striking effect of including *Dromiciops* and *Notoryctes* (Table 6b) is that $-\ln L$ support for BDa increases further, such that the BDi and DaDi alternatives can be strongly rejected with the combined mt and nuclear data, at $P = 0.033$ (tree 14) and $P = 0.023$ (tree 15). Alternatives to Phalangeriformes among the 14-taxon trees are also rejected at $P < 0.05$ (not shown). Of the monogeneric orders, *Notoryctes* is sister to bandicoots plus dasyuromorphians and *Dromiciops* is sister to the Australasian marsupials. However, as foreshadowed by the bootstrap results (Fig. 3), these placements are weakly supported. Indeed, among the separate data sets, RB-Mt_(14T) instead places *Notoryctes* as sister to the dasyuromorphians, whereas *Dromiciops* groups with Diprotodontia in the Nuc_(14T) analysis.

Clearly, the exact placements of *Dromiciops* and *Notoryctes* within Australidelphia must still be considered as uncertain. However, finding that bandicoots and

dasyuromorphians associate together among the major orders is a significant result and has important implications for understanding the morphological evolution and initial diversification of Australasian marsupials.

DISCUSSION

We have presented three new mt genomes, one for a bandicoot (*Echymipera rufescens*) as well as two for Dasyuromorphia (*Dasyurus hallucatus* and *Sminthopsis crassicaudata*). These data have helped us resolve for both a bandicoot-dasyuromorphian association among the major Australian marsupial orders and a grouping of kangaroos and possums (Phalangeriformes) among the diprotodontians. We have achieved this by differentiating between sources of compositional nonstationarity that provide either sufficient or insufficient signal to mislead tree reconstruction and by combining only the putatively unbiased mt and nuclear data.

Is Composition Bias Misleading Tree Reconstruction?

The qualitative similarity between MP, ML, and Bayesian inference for the shift in support from DaDi on the original data set (Fig. 2a) to BDa on the reduced-bias data set (Fig. 2b, see also Table 5) suggests a mechanism is at work that can similarly compromise MP and highly parametric (ML and Bayesian) approaches. Compositional nonstationarity fits this description (Galtier and Gouy, 1998; Foster, 2004) and is a usual suspect for phylogenetic bias among animal mtDNA studies. The nonhistorical signals this induces have been shown to mislead tree reconstruction for invertebrates (Jermini and Crozier, 1994; Delsuc et al., 2003) and basal vertebrates (Naylor and Brown, 1998), through to birds (Paton et al., 2002) and mammals (Schmitz et al., 2002a; Phillips and Penny, 2003). Hence, our first step in assessing the reliability of the different DNA sequence data sets was to examine the potential for nucleotide composition bias to mislead tree reconstruction.

In large data sets, the reduced importance of stochastic error leads to compositional homogeneity being rejected by the composition χ^2 test even when the signal for a bias is very small relative to historical signals (Phillips et al., 2004). In the present case, excluding data on the basis of χ^2 testing rejecting homogeneity ($P \leq 0.05$) would leave us without either the mitochondrial or nuclear data (see Table 2). Importantly, though, the base-frequency versus standard (BF/S) distance comparisons reveal that composition bias is only of sufficient magnitude to mislead phylogeny reconstruction if Mt-RNA transitions, Mt-ptn 3rd positions, and 1st position pyrimidine transitions are included. That is, where the BF distances favor the same tree as standard distances, but with substantially greater ME support on the BF distances than on standard distances (BF/S ratios >1.00 ; see Table 3). The latter condition is necessitated by some of the BF signal being attributable to sites that are uninformative for the node of interest. Excluding these sources of compositional heterogeneity leaves reduced-bias mt data sets (RB-Mt-ptn and RB-Mt-RNA)

from which to infer phylogeny alongside the nuclear sequences.

Presently the BF/S distance comparisons are not an explicit test of compositional stationarity. However, data sets could be bootstrapped to incorporate stochastic error. The presence/absence of highly supported nodes on BF distance trees run on the pseudoreplicates would provide a more powerful test of compositional homogeneity than does the χ^2 . This is because the former test is tree-based and factors out compositional correlation due to shared phylogenetic history. Moreover, BF/S distance comparisons gain considerable advantage in explanatory power over χ^2 testing by indicating whether or not the magnitude of composition bias at specific nodes may be sufficient to account for the apparent phylogenetic signal.

If compositional heterogeneity is the major source of nonhistorical signal in the data and enough sites are sampled to overcome stochastic error, then we would expect BF/S distance comparisons and BF distance trees to predict patterns of incongruence and clade support between partitions. This expectation is borne out remarkably well. ILLD test congruence between Nuc_(12T) and RB-Mt_(12T), but incongruence ($P = 0.007$) between Nuc_(12T) and Mt_(12T) is consistent with the BF/S distance comparisons, indicating that composition bias is sufficient to potentially mislead tree reconstruction only for the original Mt-ptn and Mt-RNA data. More specifically, the two differences in the MtNuc tree (Fig. 2a) from the RB-MtNuc tree (Fig. 2b) are the DaDi grouping and reduced support for Phalangeriformes (concurrent with increased support for a kangaroo plus wombat grouping). The BF distance tree in Figure 1 shows that both results are consistent with support from composition bias.

It is particularly interesting that for RB-MtNuc, the inclusion of *Dromiciops* and *Notoryctes* increases the $-\ln L$ advantage of BDa over DaDi from 12.8 to 21.2 (Table 6). This is unlikely to be due to the phylogenetically isolated taxa effectively reducing stochastic error for inferring the interrelations of the major orders. Despite the $-\ln L$ advantage of BDa over BDi also increasing slightly (from 18.7 to 20.9) for the 14-taxon analysis, stochastic effects are in fact greater, such that the AU-test P -value for rejecting BDi actually increases (from 0.014 to 0.037). Again, however, examination of BF distances can explain reduced support for DaDi relative to BDa.

Although the BF/S distance comparisons (Table 3) suggest that composition bias alone is not sufficient to result in incorrect tree selection for the reduced-bias data sets, levels of tree support may still be considerably affected by remnant compositional heterogeneity. ME (BF distance) analysis finds that this remnant heterogeneity favors DaDi in both RB-MtNuc_(12T) and RB-MtNuc_(14T). With 12 taxa, BDa and BDi require 154.0 and 153.0 changes more than DaDi. With the addition of *Dromiciops* and *Notoryctes*, this reduces to 17.3 and 9.3 changes, respectively. Confidence in BDa increases with this decrease in conflict from compositional bias and the significance level at which AU testing rejects DaDi falls from $P = 0.127$ to $P = 0.028$.

As for other studies of composition bias among mammalian mt genomes (e.g., Reyes et al., 1998; Gibson et al., 2005), it is among-pyrimidine bias that is the major source of the compositional signal for DaDi (see Fig. 1). High average thymine:cytosine ratios at informative sites among the nt-coded data that were excluded or recoded are found among the bandicoots (1.340) and outgroup “ameridelphians” (1.102), so these tend to attract relative to the lower thymine dasyuromorphians (0.939) and diprotodontians (0.601). The strong compositional bias on these mt data also provides a likely explanation for earlier studies of one or a few mt genes (e.g., Krajewski et al., 1997; Burk et al., 1999) favoring Eometatheria (including DaDi). Furthermore, these mt data will have provided the majority of signal to the total evidence analysis of Asher et al. (2004) and the supertree of Cardillo et al. (2004), which again both favored Eometatheria.

The conclusion that the BF/S distance comparisons warrant the alterations made to the original mt data is further supported by the consequent increase in phylogenetic signal retention (high stemminess; Table 2) and congruence with the nuclear data. We also looked at the possibility of branch-length effects (e.g., long-branch attraction; see Felsenstein, 1978; Hendy and Penny, 1989) misleading tree reconstruction before considering phylogenetic analysis of the remaining RB-MtNuc to be reliable.

ML models that incorporate among-site rate heterogeneity are usually expected to perform far better than MP when rates vary considerably across the tree (Sullivan and Swofford, 2001). Hence, similar trends in bootstrap support under MP and ML implies that branch-length biases are not the source of signal for BDa. In fact, inspection of the unrooted RB-MtNuc ML tree (Fig. 4) shows that the “ameridelphian” and dasyurid branches are long, relative to the diprotodontian and bandicoot branches, such that long branch-length attraction would tend to favor a bandicoot-diprotodontian (BDi) association among the major groups. We provisionally consider RB-MtNuc to be unbiased for inferring the present marsupial tree. This is not to say that these data do not contain nonhistorical signals, but rather that finding a bandicoot-dasyuromorphian association may reflect historical signal but is not an artifact of either compositional nonstationarity or the distribution of branch-lengths.

Combining Nuclear and Mitochondrial Data

The mt protein, mt RNA, and nuclear protein data sets are clearly different process partitions in the sense of Bull et al. (1993). Before even considering the empirical data on differences in substitution processes (e.g., Lin and Danforth, 2004), we know they differ in fundamental aspects of replication, exposure to and buffering of oxidative damage, error correction, and mode of inheritance (see Ballard and Whitlock, 2004). Given the differences in evolutionary processes, some proponents of taxonomic congruence (e.g., Miyamoto and Fitch, 1995) would encourage separate analysis of these partitions. However,

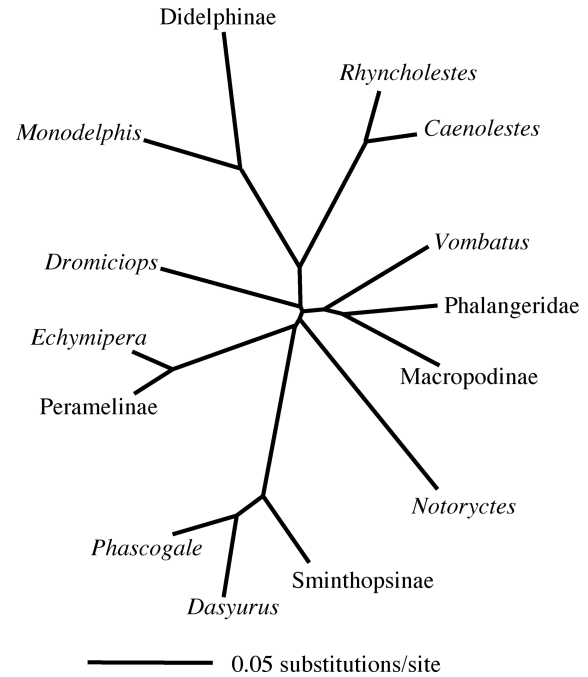


FIGURE 4. Maximum likelihood (GTR+I+ Γ_4) tree for the 14-taxon reduced-bias mt and nuclear combined data set, RB-MtNuc_(14T).

the primary arguments for separate analysis are to aid in identifying incongruence (Chen et al., 2003) and to account for differences in substitution models (Buckley et al., 2002). Neither argument applies in the present situation. Firstly, the nuclear and reduced-bias mt data sets display high ILD-test congruence and, secondly, in our combined analysis we have allowed ML and Bayesian models to be optimized for each partition. The important question here is whether combining RB-Mt with Nuc reduces the overall influence of sampling error and increases our confidence in the tree or, alternatively, whether including the “noisier” (lower stemminess; Table 2) mt data effectively pollutes the signal from the nuclear data.

AU significance testing reveals that combining the mt and nuclear data does increase the prominence of apparent phylogenetic signal relative to the influence of sampling error. Alternatives to Phalangeriformes (Table 6a) and to BDa (even for the less composition-biased 14-taxon analysis; Table 6b) can only be rejected at $P = 0.079$ and $P = 0.051$, respectively, with the nuclear sequences alone. These P -values fall to 0.023 and 0.033 with the inclusion of the mt data. Such increases in resolution upon combining data that are not significantly incongruent are essentially what has been sought by advocates of conditional data concatenation (e.g., Bull et al., 1993; Huelsenbeck et al., 1996).

The present study represents a partial unification of the ideals of both total evidence (Kluge, 1989) and taxonomic congruence (Mickevich, 1978), with a number of major concerns of both approaches being met. Even among advocates of total evidence there is room for excluding possible heritable characters from phylogenetic

inference. Kluge (1989), for example, wrote that “including all relevant evidence can be seen as a harmless activity, unless one is prepared to argue a priori that certain evidence will confound the analysis and must therefore be eliminated.” To this end we have inferred the marsupial phylogeny from all the putatively unbiased evidence from among the mt and nuclear data. With respect to taxonomic congruence, incongruence among partitions has been explained and its source excluded, whereas partitioned ML allowed differing patterns of substitution between the Nuc and Mt-ptn codon positions and Mt-RNA loops and stems to be modeled separately. Combined, partitioned analysis also avoids many of the vagaries of weighting (or not) separately analyzed partitions (Hillis, 1987; and see Levasseur and Lapointe, 2001).

Our method only differs from conditional data concatenation in our emphasis on partitioning the ML analysis. A further amendment worth consideration is whether or not to include data that are putatively unbiased and congruent with other partitions, but retain so little signal relative to “noise” that they only serve to increase variance and so reduce the power to reject hypotheses. A number of possible solutions to such data quality versus quantity problems have been proposed (e.g., Mindell and Thacker, 1996; Cunningham, 1997). In practice, though, and certainly in the present case, such a decision will often not be required with our strategy of only including putatively unbiased data. This is because BF/S distance comparisons will often identify noisy data as being potentially misleading for tree reconstruction, because composition bias tends to be accentuated when saturation has eroded much of the ancestral signal. Hence, the partition would be excluded before concerns are raised over the affect of saturation on statistical power.

Inference on the Evolution of the Major Radiations of Australasian Marsupials

We can now be confident that among the three major lineages of Australasian marsupials, it is the primarily herbivorous diprotodontians that diverged first, with the semifossorial/omnivorous bandicoots and insectivorous/carnivorous dasyuromorphians diverging from each other shortly after. This phylogeny can be used to guide future efforts for reconstructing the ecological niche and morphology of the last common ancestor (LCA) of these groups and for determining which traits were retained in the earliest stem diprotodontians, dasyuromorphians, and bandicoots.

Presently, though, we warn of the possibility of over-interpretation. Absence of, or incorrect resolution in rooting the Australasian marsupials in morphological cladistic analyses (e.g., Springer et al., 1997a; Horovitz and Sánchez-Villagra 2003), implies either the presence of major sources of nonhistorical signal, or that few synapomorphies of the BDa stem lineage have been retained. Hence, it may be difficult to accurately track transformations in character-mapping studies. Indeed, given the relative shortness of those internal branch-

lengths on the RB-MtNuc_(14T) phylogeny (Fig. 4), the morphology of the LCA of bandicoots and dasyuromorphians may have changed little from that of the LCA of all Australasian marsupials.

The near-absence of an Australian terrestrial vertebrate fossil record from the latest Oligocene (≈ 25 Mya) back to the Early Cretaceous (≈ 105 Mya) is the biggest stumbling block to inferring the nature of the initial diversification of Australasian marsupials. Thus it is critical to determine the relationships of Paleocene and Eocene fossils with alleged close affinities to the Australasian crown group, such as *Andinodelphys* (Marshall et al., 1990) from Bolivia and the fragmentary specimens from Murgon, Australia (Godthelp et al., 1999; also see Woodburne and Case, 1996). Hopefully, using the molecular tree as a scaffold will remove some of the uncertainty involved in placing fossil taxa on the tree.

Beyond the interrelations of the major Australasian orders, we can dismiss two key hypotheses, Syndactyla, a variant of BDi that typically includes *Notoryctes*, and Eometatheria, a variant of DaDi that also includes *Dromiciops* and *Notoryctes*. The latter hypothesis influenced Kirsch et al. (1997) to flag the possibility that bandicoots and other australidelphians were independently derived from among South American marsupials. Our well-corroborated finding of a bandicoot-dasyuromorphian association instead implies that bandicoot origins lie firmly within the Australasian radiation. Hence, the osseous patella, “mortise and tenon” upper ankle joint, chorioallantoic placenta, and extended corpus luteum life of bandicoots must be considered to have evolved in convergence with placental mammals. The alternative, retention from early therian ancestry, requires at least five independent losses among marsupials for each of these characters if *Dromiciops* is basal among australidelphians and “Ameridelphia” is paraphyletic (e.g., Amrine-Madsen et al., 2003). It would be interesting to explore whether all four of these traits, which appear to be adaptive for either a cursorial (running) lifestyle (Muizon, 1998; Szalay and Sargis, 2001) or rapid reproduction (Tyndale-Biscoe and Renfree, 1987) are evolutionarily linked.

A further conclusion is that syndactyly of the 2nd and 3rd pedal digits is not a synapomorphy of bandicoots and diprotodontians, as has often been advocated (see Szalay, 1993; Hall, 1987; Luckett, 1994). It has already been pointed out that basic aspects of syndactyly are under simple genetic control (Muragaki et al., 1996) and that syndactyly has evolved independently in placentals, including in gibbons and otter shrews (Nowak, 1991). Indeed, syndactyly appears to be incipient in some didelphimorph marsupials (Bensley, 1903; Kirsch, 1977), whereas Sargeant and Thulborn (1986) even attribute syndactylous footprints from the Cretaceous to an early marsupial.

The upshot of syndactyly being homoplastic among marsupials is that it can be added to the growing list of complex characters that were considered as robust support for mammal clades that have since been shown to be incorrect. Other examples include tribosphenic

molars (which evolved independently in Gondwanan and Laurasian mammals; see Lou et al., 2001) and the double-trochlea astragali (ankle bones) of artiodactyls, which are now known to be paraphyletic (Gatesy and O'Leary, 2001). We need to be very careful about giving any character special phylogenetic status, no matter how complex it appears to be, especially if the facets of that complexity may be functionally, genetically, or developmentally linked.

CONCLUSIONS AND FURTHER WORK

We have employed a simple strategy for resolving incongruence between mitochondrial and nuclear DNA process partitions for marsupial interrelations. The key was to examine compositional nonstationarity using base-frequency versus standard (BF/S) distance comparisons. Unlike stochastic tests (e.g., composition χ^2), BF/S distance comparisons are not affected by phylogenetic correlation, they allow biases to be explored across specific nodes and, most importantly, indicate whether sources of bias are of sufficient magnitude to potentially mislead tree reconstruction.

The BF/S distance comparisons show impressive explanatory power, allowing us to predict patterns of incongruence and clade support among partitions. One drawback, however, is that only standard, uncorrected distances (for inferring tree support) are directly comparable with the BF distances. The difficulty is in appropriately weighting the BF distances such that they are comparable with more realistic corrected distance models. This problem deserves further attention, although ML/Bayesian approaches for exploring the contribution of composition bias to relative support among competing hypotheses would be more desirable and may even be more tractable.

In the first part of our strategy, the BF/S distance comparisons show signal among Mt-ptn and Mt-RNA to be potentially misleading. In the second, composition bias was reduced by excluding Mt-ptn 3rd positions, Y-coding Mt-ptn 1st positions, and RY-coding Mt-RNA, with the resulting reduced bias data sets combined with the nuclear sequences. Consequently, we have for the first time been able to confidently determine affinities among the three major radiations of Australasian marsupials ((Diprotodontia, (Peramelemorphia, Dasyuromorphia)).

Further resolving the interordinal relations of marsupials requires the monogeneric *Dromiciops* and *Notoryctes* to be placed with respect to the short internal branches of the current tree. It appears likely that *Dromiciops* is basal among Australidelphia (leaving a monophyletic Australasian clade; Fig. 3), with AU-testing on the reduced-bias mt data set rejecting alternatives at $P < 0.1$ (e.g., tree 13 in Table 6). Although the present analysis of the nuclear sequences marginally favors *Dromiciops* as sister to Diprotodontia, with increased taxon sampling, Bayesian inference on these sequences (Amrine-Madsen et al., 2003) also favored a basal placement for *Dromiciops* among Australidelphia. *Notoryctes* is more prob-

lematic still, with little separating alternative sister relations (with either Dasyuromorphia, bandicoots, or a bandicoot-dasyuromorphian clade) for either the mt or nuclear analyses.

Confidently resolving the affinities of *Dromiciops* and *Notoryctes* may require data from transposable elements or genomic rearrangements. Nevertheless, the present results could aid these efforts and morphological investigations by providing a phylogenetic scaffold for the three major radiations. Furthermore, as shown in Table 2, there is considerable compositional heterogeneity even among nuclear data. As such, sequence analyses focusing on family-level or deeper marsupial phylogeny should routinely employ objective means for identifying specific branches that may result from composition bias. We further advise that although sequencing effort should concentrate on the generally higher quality nuclear data, the mitochondrial data can offer valuable independent confirmation and should not be ignored.

ACKNOWLEDGEMENTS

We thank Don Colgan and the Australian Museum (Sydney) for the dasyurid and bandicoot tissue samples. Olga Kardailsky and Charmaine Carlisle made valuable contributions to the laboratory work. We also thank Mike Westerman, Cary Krajewski, Mike Lee, and Peter Foster for helpful comments on the manuscript. This project is supported by the New Zealand Marsden Fund.

REFERENCES

- Adachi, J., and M. Hasegawa. 1996. Model of amino acid substitution in proteins encoded by mitochondrial DNA. *J. Mol. Evol.* 42:459–468.
- Amrine, H. M., and M. S. Springer. 1999. Maximum likelihood analysis of the tethythere hypothesis based on a multigene data set and a comparison of different models of sequence evolution. *J. Mamm. Evol.* 6:161–176.
- Amrine-Madsen, H., M. Scally, M. Westerman, M. J. Stanhope, C. Krajewski, and M. S. Springer. 2003. Nuclear gene sequences provide evidence for the monophyly of australidelphian marsupials. *Mol. Phyl. Evol.* 28:186–196.
- Aplin, K. P., and M. Archer. 1987. Recent advances in marsupial systematics with a new syncretic classification. Pages xv–lxxii in *Possums and opossums: Studies in evolution* (M. Archer, ed.). Surrey Beatty and Sons, Chipping Norton, New South Wales, Australia.
- Arnason, U., J. A. Adegoke, K. Bodin, E. W. Born, Y. B. Esa, A. Gullberg, M. Nilsson, R. V. Short, X. Xu, and A. Janke. 2002. Mammalian mitogenomic relationships and the root of the eutherian tree. *Proc. Natl. Acad. Sci. USA* 99:8151–8156.
- Asher, R. J., I. Horovitz, and M. R. Sánchez-Villagra. 2004. First combined cladistic analysis of marsupial mammal interrelationships. *Mol. Phyl. Evol.* 33:240–250.
- Baker, M. L., J. P. Wares, G. A. Harrison, and R. D. Miller. 2004. Relationships among the families and orders of marsupials and the major mammalian lineages based on recombination activating gene-1. *J. Mamm. Evol.* 11:1–16.
- Baker, R. H., G. S. Wilkinson, and R. DeSalle. 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diptera: Diopsidae). *Syst. Biol.* 50:87–105.
- Ballard, J. W. O., and M. C. Whitlock. 2004. The incomplete natural history of mitochondria. *Mol. Ecol.* 13:729–744.
- Baverstock, P. R., J. Birrell, and M. Krieg. 1987. Albumin immunologic relationships among Australian possums: a progress report. Pages 229–234 in *Possums and opossums: Studies in evolution* (M. Archer, ed.). Surrey Beatty and Sons, Chipping Norton, New South Wales, Australia.

- Bensley, B. A. 1903. On the evolution of the Australian Marsupialia: With remarks on the relationships of the marsupials in general. *Trans. Linn. Soc. Lond.* 9:83–217.
- Buckley, T. R., P. Arensburger, and C. Simon. 2002. Combined data, bayesian phylogenetics, and the origin of the New Zealand cicada genera. *Syst. Biol.* 51:4–18.
- Bull, J. J., J. P. Huelsenbeck, C. W. Cunningham, D. L. Swofford, and P. J. Waddell. 1993. Partitioning and combining data in phylogenetic analysis. *Syst. Biol.* 42:384–397.
- Burk, A., M. Westerman, D. J. Kao, J. R. Kavanagh, and M. S. Springer. 1999. An analysis of marsupial interordinal relationships based on 12S rRNA, tRNA valine, 16S rRNA, and cytochrome *b* sequences. *J. Mamm. Evol.* 6:317–334.
- Cardillo, M., O. R., P. Bininda-Emonds, E. Boakes, and A. Purvis. 2004. A species-level phylogenetic supertree of marsupials. *J. Zool.* 264:11–31.
- Caterino, M. S., R. D. Reed, M. M. Kuo, and F. A. H. Sperling. 2001. A partitioned likelihood analysis of swallowtail butterfly phylogeny (Lepidoptera: Papilionidae). *Syst. Biol.* 50:106–127.
- Chase, J., and M. L. Graydon. 1990. The eye of the northern brown bandicoot, *Isodon macrourus*. Pages 117–122 in *Bandicoots and Bilbies* (J. H. Seebeck, R. R. Brown, R. L. Wallis, and C. M. Kemper, eds.). Surrey Beatty and Sons, Chipping Norton, New South Wales, Australia.
- Chen, W.-J., C. Bonillo, and G. Lecointre. 2003. Repeatability of clades as a criterion of reliability: A case study for molecular phylogeny of Acanthomorpha (Teleostei) with larger number of taxa. *Mol. Phyl. Evol.* 26:262–288.
- Colgan, D. J. 1999. Phylogenetic studies of marsupials based on phosphoglycerate kinase DNA sequences. *Mol. Phyl. Evol.* 11:13–26.
- Corneli, P. S. 2002. Complete mitochondrial genomes and eutherian evolution. *J. Mamm. Evol.* 9:281–305.
- Cunningham, C. W. 1997. Is congruence between data partitions a reliable predictor of phylogenetic accuracy? Empirically testing an iterative procedure for choosing among phylogenetic methods. *Syst. Biol.* 46:464–478.
- Curole, J. P., and T. D. Kocher. 1999. Mitogenomics: Digging deeper with complete mitochondrial genomes. *Trends Ecol. Evol.* 14:394–398.
- DeBry, R. W. 1999. Maximum likelihood analysis of gene-based and structure-based process partitions, using mammalian mitochondrial genomes. *Syst. Biol.* 48:286–299.
- Delsuc, F., M. J. Phillips, and D. Penny. 2003. Comment on “Hexapod origins: monophyletic or paraphyletic?” *Science* 301:1887–1889.
- Douady, C. J., F. Delsuc, Y. Boucher, W. F. Doolittle, and E. J. Douzery. 2003. Comparison of Bayesian and maximum likelihood bootstrap measures of phylogenetic reliability. *Mol. Biol. Evol.* 20:248–254.
- Dowton, M., and A. D. Austin. 2002. Increased congruence does not necessarily indicate increased phylogenetic accuracy—the behavior of the incongruence length difference test in mixed-model analyses. *Syst. Biol.* 51:19–31.
- Farris, J. S., M. Källersjö, A. G. Kluge, and C. Bult. 1994. Testing significance of incongruence. *Cladistics* 10:315–319.
- Felsenstein, J. 1978. Cases in which parsimony or compatibility methods will be positively misleading. *Syst. Zool.* 27:401–410.
- Fiala, K. L., and R. R. Sokal. 1985. Factors determining the accuracy of cladogram estimation: evaluation using computer simulation. *Evolution* 39:609–622.
- Foster, P. G. 2004. Modeling compositional heterogeneity. *Syst. Biol.* 53:485–495.
- Galtier, N., and M. Gouy. 1998. Inferring pattern and process: Maximum likelihood implementation of a nonhomogeneous model of DNA sequence evolution for phylogenetic analysis. *Mol. Biol. Evol.* 15:871–879.
- Gatesy, J., and M. A. O’Leary. 2001. Deciphering whale origins with molecules and fossils. *Trends Ecol. Evol.* 16:562–570.
- Gibson, A., V. Gowri-Shankar, P. G. Higgs, and M. Rattray. 2005. A comprehensive analysis of mammalian mitochondrial genome base composition and improved phylogenetic methods. *Mol. Biol. Evol.* 22:251–264.
- Godthelp, H., S. Wroe, and M. Archer. 1999. A new marsupial from the Early Eocene Tingamarra local fauna of Murgon, southeastern Queensland: A prototypical Australian marsupial? *J. Mamm. Evol.* 6:289–313.
- Gontcharov, A. A., B. Marin, and M. Melkonian. 2004. Are combined analyses better than single gene phylogenies? A case study using SSU rDNA and rbcL sequence comparisons in the Zygnematophyceae (Streptophyta). *Mol. Biol. Evol.* 21:612–624.
- Hall, L. S. 1987. Syndactyly in marsupials—problems and prophecies. Pages 245–255 in *Possums and opossums: Studies in evolution* (M. Archer, ed.). Surrey Beatty and Sons, Chipping Norton, New South Wales, Australia.
- Hendy, M., and D. Penny. 1989. A framework for the quantitative study of phylogenetic data. *Syst. Zool.* 38:297–309.
- Hillis, D. M. 1987. Molecular versus morphological approaches to systematics. *Annu. Rev. Ecol. Syst.* 18:23–42.
- Ho, S. Y. W., and L. S. Jermini. 2004. Tracing the decay of the historical signal in biological sequence data. *Syst. Biol.* 53:623–637.
- Horovitz, L., and M. R. Sánchez-Villagra. 2003. A morphological analysis of marsupial mammal higher-level phylogenetic relationships. *Cladistics* 19:181–212.
- Huelsenbeck, J., J. J. Bull, and C. W. Cunningham. 1996. Combining data in phylogenetic analysis. *Trends Ecol. Evol.* 11:152–158.
- Huelsenbeck, J., and F. Ronquist. 2001. MrBayes: Bayesian inference on phylogenetic trees. *Bioinformatics* 17:754–755.
- Hughes, R. L. 1965. Comparative morphology of spermatozoa from five families. *Aust. J. Zool.* 13:533–543.
- Janke, A., N. J. Gemmel, G. Feldmaier-Fuchs, A. von Haeseler, and S. Pääbo. 1996. The mitochondrial genome of a monotreme—the platypus (*Ornithorhynchus anatinus*). *J. Mol. Evol.* 42:153–159.
- Janke, A., X. Xu, and U. Arnason. 1997. The complete mitochondrial genome of the wallaroo (*Macropus robustus*) and the phylogenetic relationship of among Monotremata, Marsupialia and Eutheria. *Proc. Nat. Acad. Sci. USA* 94:1276–1281.
- Jansa, S. A., and R. S. Voss. 2000. Phylogenetic studies of didelphid marsupials I. Introduction and preliminary results from nuclear IRBP gene sequences. *J. Mamm. Evol.* 7:43–77.
- Jermini, L. S., and R. H. Crozier. 1994. The cytochrome *b* region in the mitochondrial DNA of the ant *Tetraponera rufoniger*: Sequence divergence in Hymenoptera may be associated with nucleotide content. *J. Mol. Evol.* 38:282–294.
- Jones, F. W. 1923–1925. *The Mammals of South Australia*. Government Printers, Adelaide.
- Kirsch, J. A. W. 1977. The comparative serology of Marsupialia, and the classification of marsupials. *Aust. J. Zool. (Suppl.)* 38:1–152.
- Kirsch, J. A. W., F.-J. Lapointe, and M. S. Springer. 1997. DNA-hybridisation studies of marsupials and their implications for metatherian classification. *Aust. J. Zool.* 45:211–280.
- Kishino, H., and M. Hasegawa. 1989. Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data and the branching order Hominidae. *J. Mol. Evol.* 29:170–179.
- Kluge, A. G. 1989. A concern for evidence and a phylogenetic hypothesis of relationships among *Epicrates* (Boidae, Serpentes). *Syst. Zool.* 38:7–25.
- Krajewski, C., L. Buckley, and M. Westerman. 1997. DNA phylogeny of the marsupial wolf resolved. *Proc. Roy. Soc. Lond. B* 264:911–917.
- Levasseur, C., and F.-J. Lapointe. 2001. War and peace in phylogenetics: A rejoinder on total evidence and consensus. *Syst. Biol.* 50: 881–891.
- Lin, C.-P., and B. N. Danforth. 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined data sets. *Mol. Phyl. Evol.* 30:686–702.
- Lin, Y.-H., P. A. McLenachan, A. R. Gore, M. J. Phillips, R. Ota, M. Hendy, and D. Penny. 2002a. Four new mitochondrial genomes, and the increased stability of evolutionary trees of mammals from improved taxon sampling. *Mol. Biol. Evol.* 19:2060–2070.
- Lin, Y.-H., P. J. Waddell, and D. Penny. 2002b. Pika and vole mitochondrial genomes increase support for both rodent monophyly and glires. *Gene* 294:119–124.
- Lockhart, P. J., M. Steel, M. Hendy, and D. Penny. 1994. Recovering evolutionary trees under a more realistic model of sequence evolution. *Mol. Biol. Evol.* 11:605–612.
- Lou, Z.-X., R. Cifelli, and Z. Kielan-Jaworowska. 2001. Dual origin of tribosphenic mammals. *Nature* 409:53–57.
- Lou, Z.-X., Q. Ji, J. R. Wible, and C.-X. Yuan. 2003. An Early Cretaceous tribosphenic mammal and metatherian evolution. *Science* 302:1934–1940.

- Luckett, W. P. 1994. Suprafamilial relationships within Marsupialia: Resolution and discordance from multidisciplinary data. *J. Mammal. Evol.* 2:255–288.
- Madsen, O., M. Scally, C. J. Douady, D. Kao, R. W. DeBry, R. Adkins, H. M. Amrine, M. J. Stanhope, W. W. de Jong, and M. S. Springer. 2001. Parallel adaptive radiations in two major clades of placental mammals. *Nature* 409:610–614.
- Marshall, L. G., J. A. Case, and M. O. Woodburne. 1990. Phylogenetic relationships of the families of marsupials. Pages 433–505 in *Current mammalogy, Volume 2* (H. H. Genoways, ed.). Plenum Press, New York.
- Mickevich, M. F. 1978. Taxonomic congruence. *Syst. Zool.* 27:143–158.
- Mindell, D. P., and C. E. Thacker. 1996. Rates of molecular evolution: Phylogenetic issues and applications. *Annu. Rev. Ecol. Syst.* 27:279–303.
- Miyamoto, M. M., and W. M. Fitch. 1995. Testing species phylogenies and phylogenetic methods with congruences. *Syst. Biol.* 44:64–76.
- Muizón, C., de. 1998. *Mayulestes ferrox*, a borhyaenoid (Metatheria, Mammalia) from the early Palaeocene of Bolivia. *Phylogenetic and palaeobiologic implications. Geodiversitas* 20:19–142.
- Muragaki, Y., S. Mundios, J. Upton, and B. R. Olsen. 1996. Altered growth and branching patterns in synpolydactyly caused by mutations in HOXD13. *Science* 272:548–551.
- Murphy, W. J., E. Eizirik, W. E. Johnson, Y. P. Zhang, O. A. Ryder, and S. J. O'Brien. 2001. Molecular phylogenetics and the origins of placental mammals. *Nature* 409:614–618.
- Naylor, G. J. P., and W. M. Brown. 1998. Amphioxus mitochondrial DNA, Chordate phylogeny, and the limits of inference based on comparisons of sequences. *Syst. Biol.* 47:61–76.
- Negriso, E., A. Minelli, and G. Valle. 2004. The mitochondrial genome of the house centipede *Scutigera* and the monophyly versus paraphyly of myriapods. *Mol. Biol. Evol.* 21:770–780.
- Nilsson, M., U. Arnason, P. B. S. Spencer, and A. Janke. 2004. Marsupial relationships and a timeline for marsupial radiation in South Gondwana. *Gene* 340:189–196.
- Nilsson, M., A. Gullberg, A. E. Spotorno, U. Arnason, and A. Janke. 2003. Radiation of extant marsupials after the K/T boundary: Evidence from complete mitochondrial genomes. *J. Mol. Evol.* 57 (Suppl.):S3–S12.
- Nowak, R. M. 1991. *Walker's Mammals of the World*, 5th edition. Johns Hopkins University Press, Baltimore.
- Paton, T., O. Haddrath, and A. J. Baker. 2002. Complete mitochondrial DNA genome sequences show that modern birds are not descended from transitional shorebirds. *Proc. Roy. Soc. Lond. B* 269:839–846.
- Phillips, M. J., F. Delsuc, and D. Penny. 2004. Genome-scale phylogeny and the detection of systematic biases. *Mol. Biol. Evol.* 21:1455–1458.
- Phillips, M. J., Y.-H. Lin, G. L. Harrison, and D. Penny. 2001. Mitochondrial genomes of a bandicoot and a brushtail possum confirm the monophyly of Australiadelphian marsupials. *Proc. Roy. Soc. Lond. B* 268:1533–1538.
- Phillips, M. J., and D. Penny. 2003. The root of the mammalian tree inferred from whole mitochondrial genomes. *Mol. Phyl. Evol.* 28:171–185.
- Posada, D., and K. A. Crandall. 1998. Model Test: Testing the model of DNA substitution. *Bioinformatics* 14:817–818.
- Pupko, T., D. Huchon, Y. Cao, N. Okada, and M. Hasegawa. 2002. Combining multiple data sets in a likelihood analysis: which models are the best? *Mol. Biol. Evol.* 19:2294–2307.
- Rambaut, A. 1996. Sequence Alignment Editor, version v1.0 alpha.1.
- Rambaut, A., and A. J. Drummond. 2003. Tracer v1.0. Available at <http://evolve.zoo.ox.ac.uk>.
- Retief, J. D., C. Krajewski, M. Westerman, R. J. Winkfein, and G. H. Dixon. 1995. Molecular phylogeny and evolution of marsupial protamine P1 genes. *Proc. Roy. Soc. Lond. B* 259:7–14.
- Reyes, A., C. Gissi, F. Catzeflis, E. Nevo, G. Pesole, and C. Saccone. 2004. Congruent mammalian trees from mitochondrial and nuclear genes using Bayesian methods. *Mol. Biol. Evol.* 21:397–403.
- Reyes, A., C. Gissi, G. Pesole, and C. Saccone. 1998. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. *Mol. Biol. Evol.* 15:957–966.
- Sarjeant, W. A. S., and R. A. Thulborn. 1986. Probable marsupial footprints from the Cretaceous sediments of British Columbia. *Can. J. Earth Sci.* 23:1223–1227.
- Schmitz, J., M. Ohme, B. Suryobroto, and H. Zischler. 2002a. The colugo (*Cynocephalus variegatus*, Dermoptera): the primates' gliding sister? *Mol. Biol. Evol.* 19:2308–2312.
- Schmitz, J., M. Ohme, and H. Zischler. 2002b. The complete mitochondrial sequence of *Tarsius bancanus*: evidence for extensive nucleotide compositional plasticity of primate mitochondrial DNA. *Mol. Biol. Evol.* 19:544–553.
- Schwarz, M. P., S. M. Tierney, S. J. B. Cooper, and N. J. Bull. 2004. Molecular phylogenetics of the allodapine bee genus *Braunsapis*: A-T bias and heterogeneous substitution parameters. *Mol. Phyl. Evol.* 32:110–122.
- Shimodaira, H. 2002. An approximately unbiased test of phylogenetic tree selection. *Syst. Biol.* 51:492–508.
- Shimodaira, H., and M. Hasegawa. 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. *Mol. Biol. Evol.* 16:1114–1116.
- Shimodaira, H., and M. Hasegawa. 2001. CONSEL: For assessing the confidence of phylogenetic tree selection. *Bioinformatics* 17:1246–1247.
- Simmons, M. P., K. M. Pickett, and M. Miya. 2004. How meaningful are Bayesian support values? *Mol. Biol. Evol.* 21:188–199.
- Springer, M. S., H. M. Amrine, A. Burk, and M. J. Stanhope. 1999. Additional support for Afrotheria and Paenungulata, the performance of mitochondrial versus nuclear genes, and the impact of data partitions with heterogeneous base composition. *Syst. Biol.* 48:65–75.
- Springer, M. S., A. Burk, J. R. Kavanagh, V. G. Waddell, and M. J. Stanhope. 1997a. The interphotoreceptor retinoid binding protein gene in therian mammals: implications for higher level relationships and evidence for loss of function in the marsupial mole. *Proc. Natl. Acad. Sci. USA* 94:13754–13759.
- Springer, M. S., R. W. DeBry, C. J. Douady, H. M. Amrine, O. Madsen, W. W. de Jong, and M. J. Stanhope. 2001. Mitochondrial versus nuclear gene sequences in deep-level phylogeny reconstruction. *Mol. Biol. Evol.* 18:132–143.
- Springer, M. S., J. A. W. Kirsch, and J. A. Case. 1997b. The chronicle of marsupial evolution. Pages 129–161 in *Molecular evolution and adaptive radiation* (T. J. Givnish, and K. J. Sytsma, eds.). Cambridge University Press, New York.
- Springer, M. S., M. Westerman, J. R. Kavanagh, A. Burk, M. Woodburne, D. J. Kao, and C. Krajewski. 1998. The origin of the Australasian marsupial fauna and the phylogenetic affinities of the enigmatic monito del monte and the marsupial mole. *Proc. Roy. Soc. Lond. B* 265:2381–2386.
- Sullivan, J., and D. L. Swofford. 2001. Should we use model-based methods for phylogenetic inference when we know that assumptions about among-site rate variation and nucleotide substitution pattern are violated. *Syst. Biol.* 50:723–729.
- Suzuki, Y., G. V. Glazko, and M. Nei. 2002. Overcredibility of molecular phylogenies obtained by Bayesian phylogenetics. *Proc. Natl. Acad. Sci. USA* 99:16138–16143.
- Swofford, D. L. 2002. PAUP*. Phylogenetic analysis using parsimony (*and other methods), version 4.0b10 (PPC). Sinaur Associates, Sunderland, Massachusetts.
- Szalay, F. S. 1982. A new appraisal of marsupial phylogeny and classification. Pages 621–640 in *Carnivorous Marsupials* (M. Archer, ed.). Royal Zoological Society of New South Wales, Sydney.
- Szalay, F. S. 1993. Metatherian taxon phylogeny: Evidence and interpretation from the cranioskeletal system. Pages 216–241 in *Mammal phylogeny: Mesozoic differentiation, multituberculates, monotremes, early eutherians, and marsupials* (F. S. Szalay, M. J. Novacek, and M. C. McKenna, eds.). Springer-Verlag, New York.
- Szalay, F. S., and E. J. Sargis. 2001. Model-based analysis of postcranial osteology of marsupials from the Palaeocene of Itaboraí (Brazil) and the phylogenetics and biogeography of Metatheria. *Geodiversitas* 23:139–302.
- Thomas, O. 1888. *Catalogue of the Marsupialia and Monotremata in the Collection of the British Museum (Natural History)*. British Museum, London.
- Tyndale-Biscoe, C. H., and M. B. Renfree. 1987. *Reproductive physiology of marsupials*. Cambridge University Press, Cambridge.
- Westerman, M. 1991. Phylogenetic relations of the marsupial mole, *Notoryctes typhlops* (Marsupialia: Notoryctidae). *Aust. J. Zool.* 39:529–537.

- Woodburne, M. O., and J. A. Case. 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. Maximum-likelihood models for combined analyses of multiple sequence data. *J. Mol. Evol.* 42:587–596.
- Yang, Z. 1998. On the best evolutionary rate for phylogenetic analysis. *Syst. Biol.* 47:125–133.
- First submitted 5 November 2004; reviews returned 31 January 2005; final acceptance 25 August 2005*
Associate Editor: Mike Lee

APPENDIX 1. Maximum-parsimony bootstrap support from the 12-taxon mitochondrial protein, mitochondrial RNA and nuclear data sets for marsupial relationships, including for bandicoot-dasyuromorphian (BDa), bandicoot-diprotodontian (BDi), and dasyuromorphian-diprotodontian (DaDi) associations.

	Alternative associations among the major Australasian orders			Australidelphia	Phalangeriformes
	BDa	BDi	DaDi		
Mt-ptn	7	—	84	32 ^a	4 ^b
RB-Mt-ptn	93	1	4	96	90
Mt-RNA	10	6	78	96	25 ^c
RB-Mt-RNA	38	12	43	91	19 ^d
Nuc	99	—	1	100	96

^aBanditoos plus caenolestids favored (49%).

^bVombatus plus Phalangeridae favored (63%).

^cVombatus plus Phalangeridae favored (71%).

^dVombatus plus Macropodinae favored (47%).

Copyright of *Systematic Biology* is the property of Taylor & Francis Ltd and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.