Available online at www.sciencedirect.com
SCIENCE DIRECT®
Molecular Phylogenetics and Evolution 38 (2006) 363–380
MOLECULAR PHYLOGENETICS AND EVOLUTION
www.elsevier.com/locate/ympev

Different patterns of selection on the nuclear genes IRBP and DMP-1 affect the efficiency but not the outcome of phylogeny estimation for didelphid marsupials

Sharon A. Jansa [a,*], Jessica F. Forsman [a], Robert S. Voss [b]

[a] Bell Museum of Natural History and Department of Ecology, Evolution, and Behavior, University of Minnesota, St. Paul, MN 55108, USA
[b] Department of Mammalogy, American Museum of Natural History, New York, NY 10024, USA

Received 11 April 2005; revised 3 June 2005
Available online 27 July 2005

Abstract

Selection at the protein-level can influence nucleotide substitution patterns for protein-coding genes, which in turn can affect their performance as phylogenetic characters. In this study, we compare two protein-coding nuclear genes that appear to have evolved under markedly different selective constraints and evaluate how selection has shaped their phylogenetic signal. We sequenced 1100+ bp of exon 6 of the gene encoding dentin matrix protein 1 (DMP1) from most of the currently recognized genera of New World opossums (family: Didelphidae) and compared these data to an existing matrix of sequences from the interphotoreceptor retinoid-binding protein gene (IRBP) and morphological characters. In comparison to IRBP, DMP1 has far fewer sites under strong purifying selection and exhibits a number of sites under positive directional selection. Furthermore, selection on the DMP1 protein appears to conserve short, acidic, serine-rich domains rather than primary amino acid sequence; as a result, DMP1 has significantly different nucleotide substitution patterns from IRBP. Using Bayesian methods, we determined that DMP1 evolves almost 30% faster than IRBP, has 2.5 times more variable sites, has less among-site rate heterogeneity, is skewed toward A and away from CT (IRBP has relatively even base frequencies), and has a significantly lower rate of change between adenine and any other nucleotide. Despite these different nucleotide substitution patterns, estimates of didelphid relationships based on separate phylogenetic analyses of these genes are remarkably congruent whether patterns of nucleotide substitution are explicitly modeled or not. Nonetheless, DMP1 contains more phylogenetically informative characters per unit sequence and resolves more nodes with higher support than does IRBP. Thus, for these two genes, relaxed functional constraints and positive selection appear to improve the efficiency of phylogenetic estimation without compromising its accuracy.

Keywords: Mammalian phylogenetics; Phylogenetic utility; Molecular evolution; Positive selection; Bayesian analysis; Marsupials

1. Introduction

During the past decade, molecular systematists have increasingly embraced nuclear gene sequences as a source of phylogenetic information, particularly for resolving relationships at deeper taxonomic levels (Groth and Barrowclough, 1999; Madsen et al., 2001; Murphy et al., 2001). As a result, the identification of new nuclear loci and exploration of their phylogenetic utility have become important aspects of systematics research (e.g., Fang et al., 2000; Griffiths et al., 2004; Roger et al., 1999). Due to the plethora of mitochondrial sequences available from a wide range of taxa, most published gene comparisons have focused on the relative utility of nuclear versus mitochondrial loci (e.g., Baker et al., 2001; Griffiths et al., 2004; Lin and Danforth, 2004; Springer et al., 2001). However, nuclear genes

* Corresponding author. Fax: +1 612 624 6777.
E-mail address: jansa003@umn.edu (S.A. Jansa).

1055-7903/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.ympev.2005.06.007

exhibit a great deal of heterogeneity in various traits including strength of selection, rate of evolution, and base composition, each of which can influence phylogenetic signal. Despite such variation, few studies have focused on the relative efficacy of different nuclear loci for resolving a given phylogenetic problem (but see Lundrigan et al., 2002; van den Bussche et al., 2003). Such explorations should prove fruitful, as they can identify particular aspects of gene evolution that might predict the success of future phylogenetic applications, or reveal undesirable properties that might lead inference astray.

Phylogenetic "utility" can be measured in a number of ways. Among other criteria, efficiency of data collection, levels of homoplasy, strength of resolution, ease of sequence alignment, and agreement with prior studies have all been used in determining the value of particular data sets for phylogenetic inference (Graybeal, 1994; Griffiths et al., 2004; Groth and Barrowclough, 1999). Recently, explicitly model-based approaches have pointed at the relative importance of various nucleotide substitution parameters in determining the phylogenetic utility of sequence data (Lin and Danforth, 2004; Yang, 1996). For example, genes with even base composition and homogeneous patterns of variation across sites tend to outperform those with highly skewed base composition and high levels of rate heterogeneity (Chang and Campbell, 2000; Groth and Barrowclough, 1999; Lin and Danforth, 2004). Although most studies have focused at the nucleotide level, it is important to recognize that patterns of nucleotide substitution for protein-coding loci are determined in part by selection at the protein level. Therefore, examining the role that selection plays in shaping these patterns might yield important insights into the behavior of these loci as phylogenetic markers.

In this paper, we examine how differences in selection on two protein-coding nuclear genes have affected the phylogenetic utility of their DNA sequences. As a measure of the phylogenetic utility of each gene, we consider the number of informative characters per unit sequence, the relative amount of homoplasy exhibited by each gene, each gene's ability to reconstruct relationships among taxa with strong nodal support, and congruence among data sets. Our data consist of sequences from exon 1 of the IRBP gene and exon 6 of the DMP1 locus, and a nonmolecular data set consisting of morphological and karyotypic characters (Voss and Jansa, 2003). Exon 1 of IRBP has been widely used in mammalian molecular systematics to infer relationships across a range of hierarchical levels from the interordinal (Springer et al., 1997; Stanhope et al., 1992) to the interfamilial (deBry and Sagel, 2001) and intergeneric (Jansa and Voss, 2000; Jansa and Weksler, 2004). Exon 6 of DMP1 has been less extensively used in phylogenetic studies; however, its properties have previously been investigated for both deep (Toyosawa et al., 1999; van den Bussche et al., 2003) and shallow (Reeder and Bradley, 2004) hierarchical levels in mammalian systematics.

Didelphid opossums, our focal taxonomic group, are widely distributed from Patagonia to Canada, but they are most diverse in tropical lowland rainforests where numerous sympatric species have been documented at some Amazonian localities (Voss and Emmons, 1996). A total of 89 extant species in 18 genera are currently recognized (Table 1). Of these, five species in three genera are referred to the subfamily Caluromyinae, whereas 84 species in 15 genera belong to the subfamily Didelphinae. Among didelphines, two suprageneric groups are commonly recognized. "Marmosines" include a variety of small (<200 g) mouse opossums formerly classified as or allied with the genus *Marmosa* (sensu Tate, 1933). Another group, the "large $2n = 22$ opossums" includes *Didelphis* and allied genera that are usually distinguished from other didelphids by body size and chromosomal characters. Two genera (*Monodelphis* and *Metachirus*) are didelphines of traditionally uncertain relationships.

Table 1
Current classification of didelphid marsupials[a]

- Caluromyinae
  - *Caluromys* (3 spp.)
  - *Caluromysiops* (1 sp.)
  - *Glironia* (1 sp.)
- Didelphinae
  - "Marmosines"
    - *Chacodelphys* (1 sp.)[b]
    - *Cryptonanus* (5 spp.)[c]
    - *Gracilinanus* (6 spp.)[c]
    - *Lestodelphys* (1 sp.)
    - *Marmosa* (9 spp.)
    - *Micoureus* (6 spp.)
    - *Marmosops* (14 spp.)[d]
    - *Thylamys* (8 spp.)
    - *Tlacuatzin* (1 sp.)[e]
  - "Large $2n = 22$ opossums"
    - *Chironectes* (1 sp.)
    - *Didelphis* (6 spp.)
    - *Lutreolina* (1 sp.)
    - *Philander* (4 spp.)
  - Others
    - *Metachirus* (1 sp.)
    - *Monodelphis* (20 spp.)[f]
- Incertae sedis
  - *Hyladelphys* (1 sp.)[g]

[a] After Gardner (2005), except as noted.
[b] Genus described by Voss et al. (2004a).
[c] Genus described by Voss et al. (2005).
[d] Includes *Marmosops creightoni*, a recently described species, but not *M. dorothea* (a synonym of *M. noctivagus*; see Voss et al., 2004b).
[e] Genus described by Voss and Jansa (2003).
[f] Includes *Monodelphis ronaldi* Solari (2004) and *M. reigi* Lew and Pérez-Hernández (2004).
[g] Although listed as a didelphid by Gardner (2005), the family-level classification of this anomalous taxon has yet to be determined (see Voss et al., 2001).

The first phylogenetic study to include representatives of all then-recognized didelphid genera (Jansa and Voss, 2000) was based on sequence data from exon 1 of IRBP. The resulting phylogeny (Fig. 1) was almost completely resolved and contained many impressively supported nodes. Subsequent phylogenetic studies based on the same sequence data (Voss and Jansa, 2003; Voss et al., 2004a, 2005) were undertaken to extend these results with denser taxon sampling and to assess morphological support for recovered molecular clades. Unexpectedly, denser taxon sampling revealed the existence of several highly divergent lineages represented by species that are clearly not closely related to other members of the genera in which they were traditionally classified. Described as new genera, these include *Chacodelphys* (for "*Thylamys*" *formosus*; Voss et al., 2004a), *Cryptonanus* (for several species formerly classified in *Gracilinanus*; Voss et al., 2005), and *Tlacuatzin* (for "*Marmosa*" *canescens*; Voss and Jansa, 2003).

Unfortunately, another consequence of denser taxon sampling was the erosion of support statistics for certain clades together with some loss of phylogenetic resolution (Voss and Jansa, 2003; Voss et al., 2005). This result was not unexpected insofar as prior studies have emphasized the unpredictable effects that adding taxa can have on phylogenetic inference (Hillis, 1998; Horovitz, 1999; Soltis et al., 1998). Denser taxonomic sampling is always desirable to achieve a better understanding of biological diversification; however, the only solution to the problem of lost resolution and support is to increase the number of characters scored for each taxon. To do so, we have been sequencing additional nuclear exons for didelphid marsupials and examining their phylogenetic properties. As reported herein, sequence data from

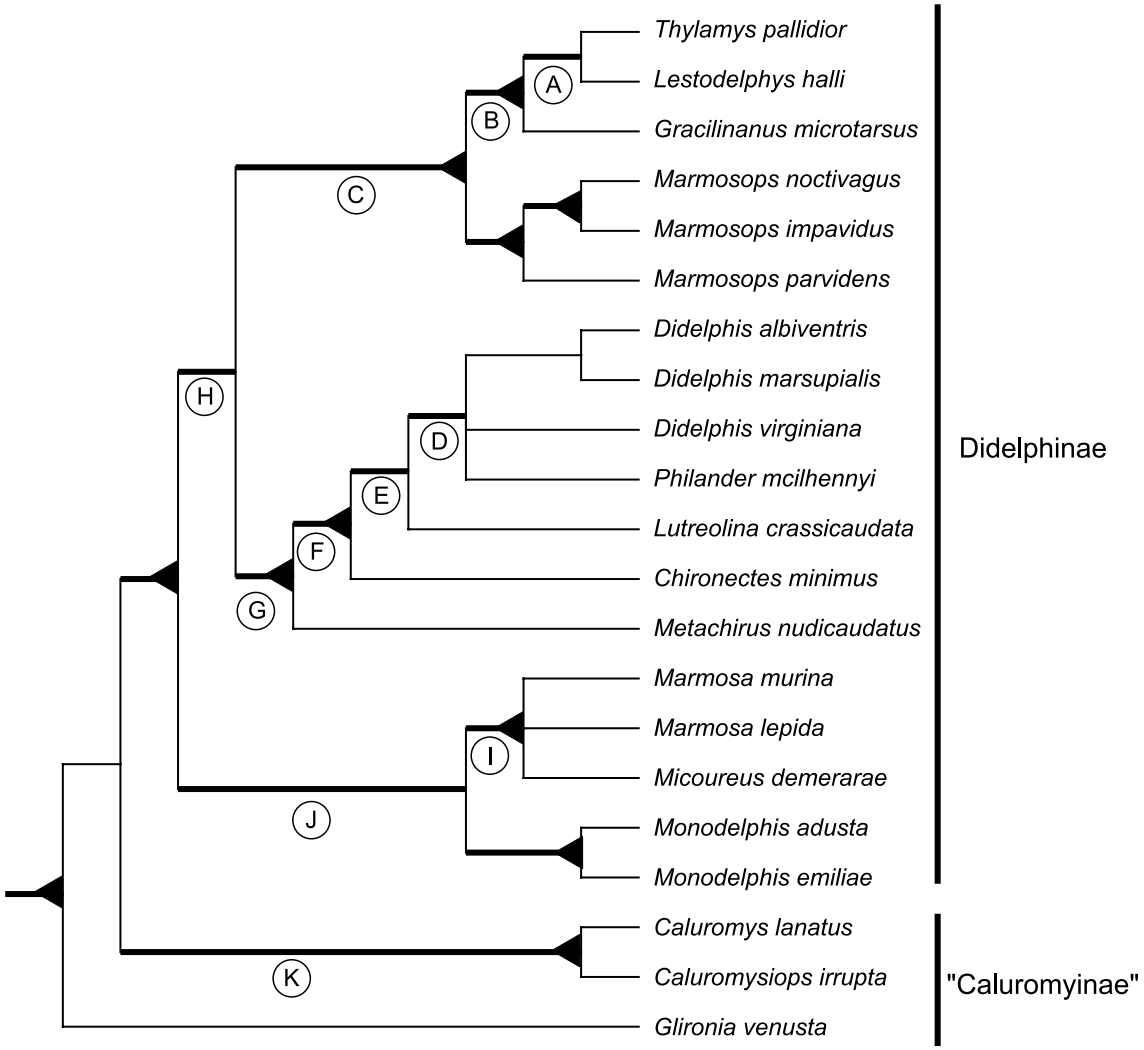


Fig. 1. Tree based on parsimony analysis of IRBP sequences as reported in Jansa and Voss (2000). Letters refer to intergeneric groups as defined in that report and referred to herein. Line thickness and buttressing reflects bootstrap support as follows: a thickened line indicates that the node was recovered in $\geqslant$75% of bootstrap replicates; a thickened line with a terminal buttress indicates that the node was recovered in $\geqslant$95% of bootstrap replicates; a light line represents bootstrap support $\leqslant$50%.

DMP1 exon 6 complement and enhance many of our earlier systematic conclusions and display many desirable qualities as a source of phylogenetic characters.

2. Materials and methods

The IRBP sequences analyzed below are taken from Jansa and Voss (2000), Voss and Jansa (2003), and Voss et al. (2005). Information about voucher material as well as details of our previous analyses appear in these reports as well. All DMP1 sequences were generated for this study as follows.

2.1. DNA amplification and sequencing

We amplified and sequenced part of DMP1 exon 6 using combinations of primers shown in Fig. 2. We first amplified a fragment approximately 1200 bp long using primers DEN-12 and DEN-2 (Toyosawa et al., 1999) from genomic DNA. To generate fragments of a suitable size for sequencing, this product was used as a template in two subsequent PCRs, one using DEN-12 paired with DEN-13, and one using primers DEN-2 and DEN-14; for certain taxa, DEN-2 was replaced with DEN-R1192. Our sample of sonicated DNA from *Lestodelphys* was sufficiently degraded that amplifications were done as a series of 300–400 bp amplifications using the following primer pairs: LesF2/LesR2, LesF3/LesR3, LesF4/LesR4, LesF5/R1192. Due to a 12 bp deletion in the sequence from *Tlacuatzin canescens*, primer DEN-13 was replaced with DEN-R638 for this taxon.

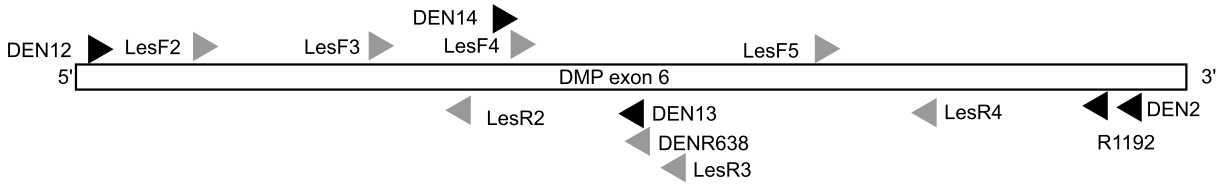
Initial PCR amplifications using genomic DNA as a template were performed in 20 μL reactions using Ampli-Taq Gold polymerase (Applied Biosystems) and reaction conditions as previously described (Jansa and Voss, 2000). Reamplification reactions using this product as a template were performed in 30 μL reactions using Taq DNA Polymerase (Fisher). Resulting PCR products were prepared for sequencing using a QIAquick PCR purification kit (Qiagen) or ExoSAP (Amersham) enzymatic digestion.

PCR products were sequenced in both directions using amplification primers and dye-terminator chemistry (BigDye, Applied Biosystems), and sequencing reactions were run on an ABI 3700 automated sequencer. Sequences were proofed using Sequencher 4.2 (GeneCodes). Base-calling ambiguities were resolved either by choosing the base on the cleanest strand or using the appropriate IUB ambiguity code if both strands showed the same ambiguity. All DMP1 sequences, along with their specimen voucher numbers, have been deposited in GenBank with Accession Nos. DQ083120–DQ083160.

2.2. Sequence analysis

We aligned DMP1 sequences by eye using translated amino acid sequences as a reference. Sequences of DMP1 ranged from 1011 (*Lestodelphys*) to 1176 bp. We were unable to obtain sequence from the first 3 bp of *Gracilinanus microtarsus* and *Tlacuatzin*, from the first 9 bp of *Philander mcillhennyi*, from the first 162 bp of *Lestodelphys*, and from the last 36 bp of *Micoureus paraguayanus* and *M. demerarae*. Twelve alignment-unambiguous internal gaps, ranging in length from 3 to 18 bp, were required to align DMP1 sequences with one another. For phylogenetic analysis, each gap was coded as single insertion–deletion event regardless of its length.

We evaluated IRBP and DMP1 sequences for variation in base composition across taxa and among codon positions of each gene with a χ^2 goodness-of-fit test for each taxon using the average base composition across taxa to calculate relevant expected values. Congruence between the two genes was evaluated using the ILD test (Farris et al., 1995) prior to phylogenetic analysis (but see Barker and Lutzoni, 2002; Dolphin et al., 2000; Hipp



DEN12 5'GATGAAGACGACAGTGGAGATGACACCTT
LesF2 5'AGTCAGAGTGAGAGCAGAAGCCT
LesF3 5'GAAAAAGATTCACAAGACAAGACA
DEN14 5'GAGAACAGCAAAAGCAAATCCCAAG
LesF4 5'AAGAAGATAGCAACAGTCAGTCTCA
LesF5 5'GTCCTCCAAATCAGACAGTGAATC

LesR2 5'TGGAGTCACTCATGGTTTCTTCCA
DEN13 5'ATCRCTGTCCTGAGGTCTAGACTC
DENR638 5'AGACTCACTACTAAGATCCCTCGCAC
LesR3 5'TTCCTTGGATTTACTTTGTTTTTCAT
LesR4 5'CTCACTTGACTCCTTGCTTTCCA
R1192 5'TATCATAACCGACCCAYTGGAA
DEN2 5'ATCTTGGCAATCATTGTCATC

Fig. 2. Names, locations, and sequences of primers used in PCR amplification and sequencing of DMP1 exon 6.

et al., 2004 for interpretation of this test). The ILD test was implemented in PAUP* v4.0b10 (Swofford, 2002) with 500 replicates, using heuristic searches with stepwise addition and 100 random addition sequences per replicate, saving only two trees per replicate. The ILD test appears to exhibit an inflated type I error rate (Barker and Lutzoni, 2002); therefore, we also performed likelihood-based Shimodaira–Hasegawa (SH; Shimodaira and Hasegawa, 1999) tests to determine the relative fit of each data set to alternative topologies. Specifically, we performed three tests: whether the IRBP data could reject either the DMP1 parsimony tree or the combined gene parsimony tree; whether the DMP1 data could reject either the IRBP parsimony tree or the combined gene parsimony tree; and whether the combined gene data could reject either of the individual gene trees. All SH tests were performed using 1000 RELL replicates allowing the parameters for the relevant model of nucleotide substitution (see below) to be estimated.

2.3. Phylogenetic analysis

We performed phylogenetic analyses on the DMP1 and IRBP data sets separately and in combination. For DMP1, we performed parsimony analyses with gaps coded either as phylogenetic characters or as missing data; in the likelihood analyses of DMP1, gaps were treated as missing data. We previously reported our analyses of morphological data alone and in combination with IRBP (Voss and Jansa, 2003; Voss et al., 2005). To assess the effects of adding DMP1 sequences to these data sets, we combined all three sources of character data and analyzed the resulting matrix using parsimony.

Aligned sequences were subjected to phylogenetic analyses using parsimony and maximum-likelihood (ML) as implemented in PAUP* vb10 (Swofford, 2002) and Bayesian phylogenetic analysis as implemented in MrBayes ver. 3.0 or ver. β4 (Huelsenbeck and Ronquist, 2001). For parsimony analysis, all molecular characters were treated as unordered and equally weighted; all parsimony tree searches were heuristic with at least 20 replicates of random taxon addition followed by TBR branch swapping. To determine the best model for use in the ML and Bayesian analyses, we examined the relative fit of various models of nucleotide substitution for the IRBP and DMP1 data separately using both hierarchical likelihood-ratio tests (hLRT) and the Akaike information criterion (AIC) as implemented in ModelTest 3.6 and PAUP* (Posada and Crandall, 1998). Log-likelihood scores and estimated parameter values for the various models were calculated based on a neighbor-joining tree of Jukes–Cantor-corrected distances. Once the best-fit model of nucleotide substitution was chosen, we used additional comparisons (both hLRT and AIC) to evaluate whether a molecular clock could be enforced. The two different approaches to model selection resulted in different best-fit models for our data; we performed full likelihood analyses using the model selected by the AIC for reasons outlined in Posada and Buckley (2004). Subsequent to model evaluation and selection, the maximum-likelihood tree for each gene and for the combined gene data set was determined using a heuristic search in which parameter values were initially fixed according to the best-fit model and a neighbor-joining tree was used as a starting point for TBR branch swapping.

Bootstrap values were calculated for all data sets under the parsimony criterion using 1000 replicates of heuristic searches (five random addition replicates, TBR branch swapping) in PAUP*. For maximum-likelihood analysis, bootstrap values were calculated by generating 500 bootstrap replicates using the seqboot program in PHYLIP 3.6 (Felsenstein, 1993), finding the best-fit tree for each under the appropriate model of nucleotide substitution in PhyML (Guindon and Gascuel, 2003), and calculating the 50% majority rule tree from the resulting trees in PAUP*.

We performed two versions of Bayesian phylogenetic analysis, one in which parameters were allowed to vary independently between the two genes (mixed model analysis) and one using a single model of nucleotide substitution applied across both genes simultaneously (uniform model analysis). For the mixed model analysis, we conducted one run of Metropolis-coupled Markov chain Monte Carlo (with four incrementally heated chains each), in which we specified a model with six categories of base substitution, a Γ-distributed rate parameter, and a proportion of invariant sites (GTR + I + Γ). Uniform interval priors were assumed for all parameters except base composition, which assumed a Dirichlet prior. We decoupled parameter estimation (including branch lengths) across the data set, thus allowing substitution parameters to be estimated for the DMP1 and IRBP data sets independently. Runs were allowed to proceed for 2 million generations each, and trees were sampled every 100 generations. To ensure that each run converged on the same average log-likelihood, we plotted the log-likelihood values against generation time for each. We discarded the first 500,000 generations (5000 trees) from the run as burn-in (as determined by plotting $-\ln L$ scores for each generation and discarding those prior to the point of stationarity), and calculated estimated parameter values and posterior probabilities for each node based on the remaining trees.

To evaluate whether the mixed model Bayesian analysis returned different posterior probability estimates than an analysis that assumed a single model across both genes, we conducted a second Bayesian analysis using a single model applied across the combined gene data set. Prior probability values were assigned as above, and a single run of four incrementally heated

chains was allowed to proceed for 2 million generations, sampling trees every 100 generations. The initial 500,000 generations (5000 trees) were discarded as burn-in, and the remaining trees were used to calculate estimated parameter values and posterior probabilities for each node.

To quantify nucleotide substitution patterns for the two genes, we obtained parameter estimates using likelihood analysis of each gene separately, Bayesian analysis with parameter estimation decoupled between the genes, and model averaging using Akaike weights (Posada and Buckley, 2004). The decoupled Bayesian analysis also allowed us to determine whether given parameter estimates were significantly different between the two genes. To do so, we calculated the 95% credibility interval for the difference between parameter estimates for the two genes. If this interval contained zero, then there was no significant difference between the genes for the given parameter. The model averaging approach also allowed us to evaluate the relative importance of various nucleotide substitution parameters for each gene (Posada and Buckley, 2004).

2.4. Tests for selection

To evaluate the types and strength of selection acting on DMP1 and IRBP protein sequences, we compared different models of protein evolution using the codeml program in the computer package PAML 3.14 (Yang, 1997). These tests allow evaluation of whether particular sites in a protein have evolved under positive selection or purifying selection (Nielsen and Yang, 1998; Yang et al., 2000). A ratio of replacement to silent substitutions greater than one (dN/dS > 1) constitutes evidence of positive selection, whereas a dN/dS ratio less than one suggests purifying selection. Specifically, we performed two tests for positive selection on the combined gene data set. We performed these tests using one of the minimum-length trees resulting from parsimony analysis of all data combined. The first test compares the relative fit of two models, one of which (M0) allows for only one dN/dS ratio (hereafter referred to by its parameter designation ω) across the data set versus one (M3 with four site categories) that allows sites to be distributed across four classes of ω. The second test is similar to the M3–M0 comparison, but uses a β-distribution of site categories rather than four discrete categories to model the distribution of ω across sites. This test compares the likelihood of a neutral model (M7) with one that allows for a proportion of positively selected sites (M8). Comparison of these models constitutes a test for positive selection: if M3 or M8 provides a better fit to the data than their comparable neutral models (M0 or M7, respectively), and if a proportion of sites have an estimated $\omega > 1$, then these sites are inferred to have evolved under positive selection. In addition to this explicit test for positive selection, we used the Bayes empirical Bayes approach implemented in PAML 3.14 to obtain posterior mean estimates of ω for each site (Yang et al., 2005). We then evaluated the distribution of ω values between the two genes and among sites within each gene to test explicit hypotheses of molecular evolution (see below).

3. Results

3.1. Base, codon, and amino acid composition

Base, codon, and amino acid compositional characteristics for didelphid IRBP sequences were discussed by Jansa and Voss (2000). Although the taxon sampling differs between that study and the present one, the base-compositional characteristics for the IRBP gene remain similar for our current taxon set. Briefly, didelphid IRBP sequences exhibit a slight overall bias toward GC (average GC content = 56%), with most of this bias contained in third codon positions (average $GC_3 = 64\%$). In contrast, didelphid DMP1 sequences exhibit a bias toward AG (average AG content = 65%), which is strongly reflected in first and second codon positions (average $AG_1 = 75\%$, average $AG_2 = 73\%$). No taxon in our study exhibits significant departures from base-compositional stationarity for either gene when assessed with a χ^2 goodness-of-fit test.

Dentin matrix protein 1 in humans has been characterized as a serine-rich, acidic protein with a high concentration of aspartic (Asp) and glutamic acid (Glu) residues (Srinivasan et al., 1999). These characteristics apply for DMP1 amino acid sequences across didelphid taxa as well, and are consistent with the observed base-compositional characteristics of the gene. For the didelphids we sequenced, serine (coded by TCN or AGY) is the single most prevalent amino acid, comprising an average of 25% of the protein (range across taxa = 23.8–26.1%). Aspartic (GAY) and glutamic (GAR) acid are the next most common residues and together comprise 26.9% (range 25.3–28.4%) of the protein. For these three amino acids, there is a strong bias toward AGT and AGC among the possible serine codons, toward GAT as an Asp codon, and toward GAA as a Glu codon.

3.2. Substitution characteristics and model evaluation

For IRBP, among-site rate variation (α) and the proportion of invariant sites (P_{inv}) were deemed the most important properties of the substitution process; allowing the different transversion types to have different rates was more important than allowing different rates for the transition parameters; while allowing different frequencies of the four bases had relatively low importance

Table 2
Maximum-likelihood parameter estimates for each gene under its best-fit model (IRBP: TVMef+I+Γ+clock; DMP1: GTR+Γ) and the combined gene data set under its best-fit model (TIM+I+Γ), as well as model averaged estimates and the relative importance of each parameter as determined using a weighted AIC approach (Posada and Buckley, 2004)

Parameter	IRBP			DMP1			Combined		
	Likelihood	Model averaged	Importance	Likelihood	Model averaged	Importance	Likelihood	Model averaged	Importance
rCT	7.29	7.07	0.42	7.14	7.08	1.0	5.85	6.38	1.0
rCG	1.36	1.28	0.72	0.86	0.86	0.81	0.70	0.92	0.47
rAT	1.57	1.62	0.72	0.42	0.42	0.81	0.70	0.74	0.47
rAG	7.29	5.90	0.42	2.78	2.76	1.0	3.33	3.6	1.0
rAC	2.19	2.13	0.72	1.08	1.08	0.81	1.00	1.3	0.47
πA	0.25	0.25	0.59	0.38	0.38	1.0	0.31	0.31	1.0
πC	0.25	0.26	0.59	0.17	0.17	1.0	0.22	0.22	1.0
πG	0.25	0.27	0.59	0.26	0.26	1.0	0.27	0.27	1.0
πT	0.25	0.22	0.59	0.19	0.19	1.0	0.20	0.20	1.0
α	0.89	0.87	0.99	0.67	0.83	1.0	1.00	1.0	1.0
P_{inv}	0.50	0.50	0.97	0	0.18	0.40	0.35	0.35	0.93
Tree length	0.39	NA	NA	0.62	NA	NA	0.51	NA	NA
$-\ln L$	4369.38	NA	NA	5645.18	NA	NA	10134.24	NA	NA

(Table 2). Therefore, a model describing the evolution of this gene should include a proportion of invariant sites, a γ-distributed rate parameter, different transversion rate parameters, and equal base frequencies. Hierarchical likelihood-ratio tests (hLRTs) resulted in a best-fit model with even base frequencies and separate rates for transitions and transversions (K80); the AIC settled on a slightly more complex model that allowed for one transition rate but different transversion rates (TVMef). Both methods of model choice included a γ-distributed rate parameter (α) and a proportion of invariant sites (P_{inv}) as part of the best-fit model. In addition, we could not reject a molecular clock for IRBP data using the hLRT criterion ($-2\Delta\ln L = 45.98$, df = 39, $p = 0.12$); a molecular clock also provided a better fit to the IRBP data according to the AIC.

For DMP1, base frequency, among-site rate variation, and transition parameters all received high relative importance; transversion parameters were less important; whereas including a proportion of invariant sites was relatively unimportant (Table 2). A model that allows for different proportions of bases, two different transversion rates and different rates for transition types (TIM) was chosen through likelihood-ratio testing. The AIC, however, suggests that a more complex model that includes different rates for each substitution type (GTR) is more appropriate. Both approaches included a γ-distributed rate parameter but no invariant sites in the best-fit model. A molecular clock was rejected for the DMP1 data set using either criterion for model choice (hLRT: $-2\Delta\ln L = 112.88$, df = 39, $p \ll 0.001$).

For the combined gene data set analyzed under a uniform model, base frequencies, transition parameters, site-specific rates, and a proportion of invariant sites were all considered important parameters (Table 2). Both the hLRT and AIC suggest a TIM+I+Γ model as the best fit for both genes analyzed simultaneously, and a molecular clock was rejected for this data set (hLRT: $-2\Delta\ln L = 91.90$, df = 39, $p \ll 0.001$).

Comparisons of mean parameter estimates from the partitioned Bayesian analysis (Table 3, Fig. 3) suggest that the two genes have similar CT and CG substitution

Table 3
Mean parameter estimates and 95% credibility intervals (in parentheses) from a Bayesian analysis assuming a single model across both genes (uniform model) and one that allowed parameter estimation to be decoupled between genes (mixed model)

Parameter	Uniform model	Mixed model	
	IRBP + DMP1	IRBP	DMP1
rCT	7.98 (6.11–10.32)	8.58 (5.09–14.29)	7.47 (5.39–10.36)
rCG	1.06 (0.77–1.43)	1.14 (0.58–2.05)	0.88 (0.52–1.37)
rAT	0.86 (0.60–1.19)	**1.93** (0.91–3.73)	**0.43** (0.25–0.67)
rAG	4.54 (3.54–5.84)	**6.62** (3.88–11.24)	**2.87** (2.05–3.95)
rAC	1.56 (1.15–2.08)	**2.12** (1.12–3.66)	**1.12** (0.73–1.64)
πA	0.31 (0.30–0.32)	**0.24** (0.22–0.27)	**0.37** (0.35–0.40)
πC	0.22 (0.20–0.23)	**0.27** (0.24–0.29)	**0.18** (0.16–0.20)
πG	0.27 (0.25–0.28)	0.27 (0.25–0.30)	0.26 (0.24–0.28)
πT	0.20 (0.19–0.22)	**0.22** (0.20–0.24)	**0.19** (0.17–0.21)
α	1.11 (0.58–1.89)	**1.14** (0.36–2.87)	**1.31** (0.66–2.83)
P_{inv}	0.34 (0.18–0.46)	**0.48** (0.23–0.62)	**0.19** (0.02–0.36)
TL	0.57 (0.53–0.61)	**0.57** (0.50–0.66)	**0.73** (0.66–0.81)
$-\ln L$	10183.02 (10171.49–10195.71)	10113.7 (10096.32–10132.21)	

Values in bold indicate that these parameter estimates differed significantly between the two genes.

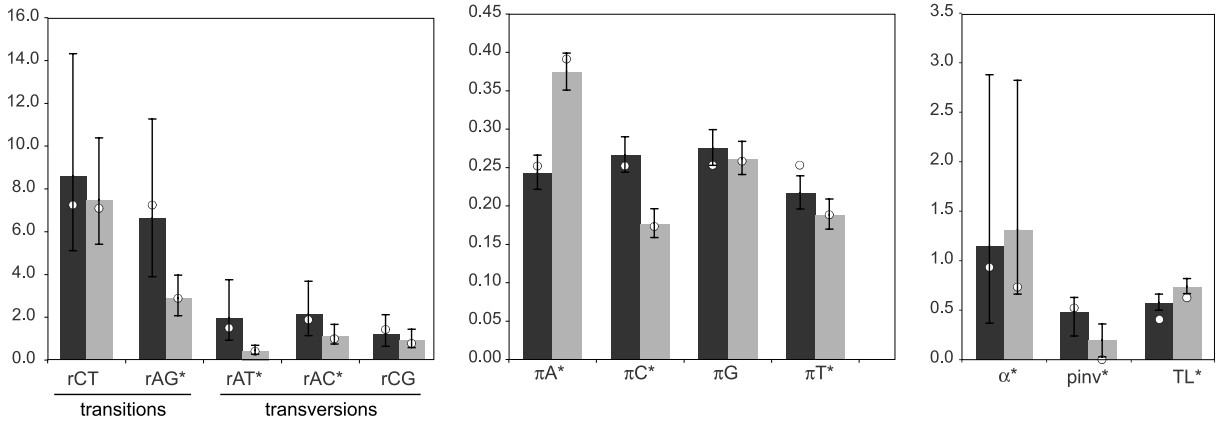


Fig. 3. Average values and 95% credibility intervals of parameter estimates for IRBP (dark gray) and DMP1 (light gray) as determined through a Bayesian analysis that decoupled parameter estimation between the two genes. An asterisk indicates that parameter estimates differ significantly between the two loci. Open circles represent point estimates of parameter values from a likelihood analysis of each gene under its best-fit model.

rates and a similar proportion of G. However, DMP1 evolves almost 30% faster than IRBP, has 2.5 times more variable sites, has less among-site rate heterogeneity, is skewed toward A and away from CT (IRBP has relatively even base frequencies), and has a significantly lower rate of change between adenine and any other nucleotide (Fig. 3). Likelihood point estimates for certain parameters fall outside the 95% credibility interval of the Bayesian estimates. However, model-averaged parameter estimates are all within the 95% credibility interval for values obtained using Bayesian methods (compare Tables 2 and 3).

3.3. Phylogenetic results

Parsimony analysis of the IRBP data (containing 22% variable sites and 14% parsimony-informative sites; Table 4) resulted in 21 minimum-length trees; the strict consensus of these trees resolved 31 of a possible 39 nodes (Fig. 4A). Separate analysis of the DMP1 data (containing 36% variable sites and 20% parsimony-informative sites; Table 4) resulted in 30 minimum-length trees, the strict consensus of which resolves 33 nodes (Fig. 4B), two more than the IRBP parsimony consensus. Moreover, a higher proportion of nodes in the DMP1 consensus topology received bootstrap support $\geqslant$75% than those in the IRBP consensus tree (Table 4). Simultaneous analysis of IRBP and DMP1 returned eight equally parsimonious trees, the strict consensus of which resolved 33 nodes (not shown); 88% of these had bootstrap support $\geqslant$75%. Analyses of the matrix that included both genes plus morphology resulted in more equally parsimonious trees and the loss of one node in the strict consensus; however, the number of well-supported nodes (bootstrap $\geqslant$75%) increased slightly over that recovered from parsimony analysis of just the combined gene data (Fig. 5, Table 4).

Table 4
Tree statistics from parsimony analysis of the various partitioned and combined data sets

	IRBP	DMP1[a]	DMP1[b]	DMP1[b] + IRBP	All data combined[c]
No. of characters	1158	1176	1188	2346	2417
No. of variable	255	418	430	685	755
No. of informative	159	239	247	406	472
No. of MPTs	21	30	30	8	20
Tree length	427	673	691	1124	1301.5
$CI_{(excl.\ uninf.)}$	0.63	0.61	0.61	0.61	0.59
$CI_{(incl.\ uninf.)}$	0.71	0.73	0.72	0.72	0.69
RI	0.84	0.82	0.82	0.83	0.82
No. of nodes resolved[d]					
bs $\geqslant$ 95%	14	16	18	22	22
75 $\geqslant$ bs < 95	9	12	10	7	8
50 $\geqslant$ bs < 75	9	7	8	5	3
Total	31	34	33	33	32

[a] With gaps coded as missing data.
[b] With gaps coded as phylogenetic characters.
[c] IRBP + DMP1 (with gaps as characters) + morphology.
[d] Nodes resolved with given bootstrap support (bs); total refers to resolved nodes in strict consensus tree.

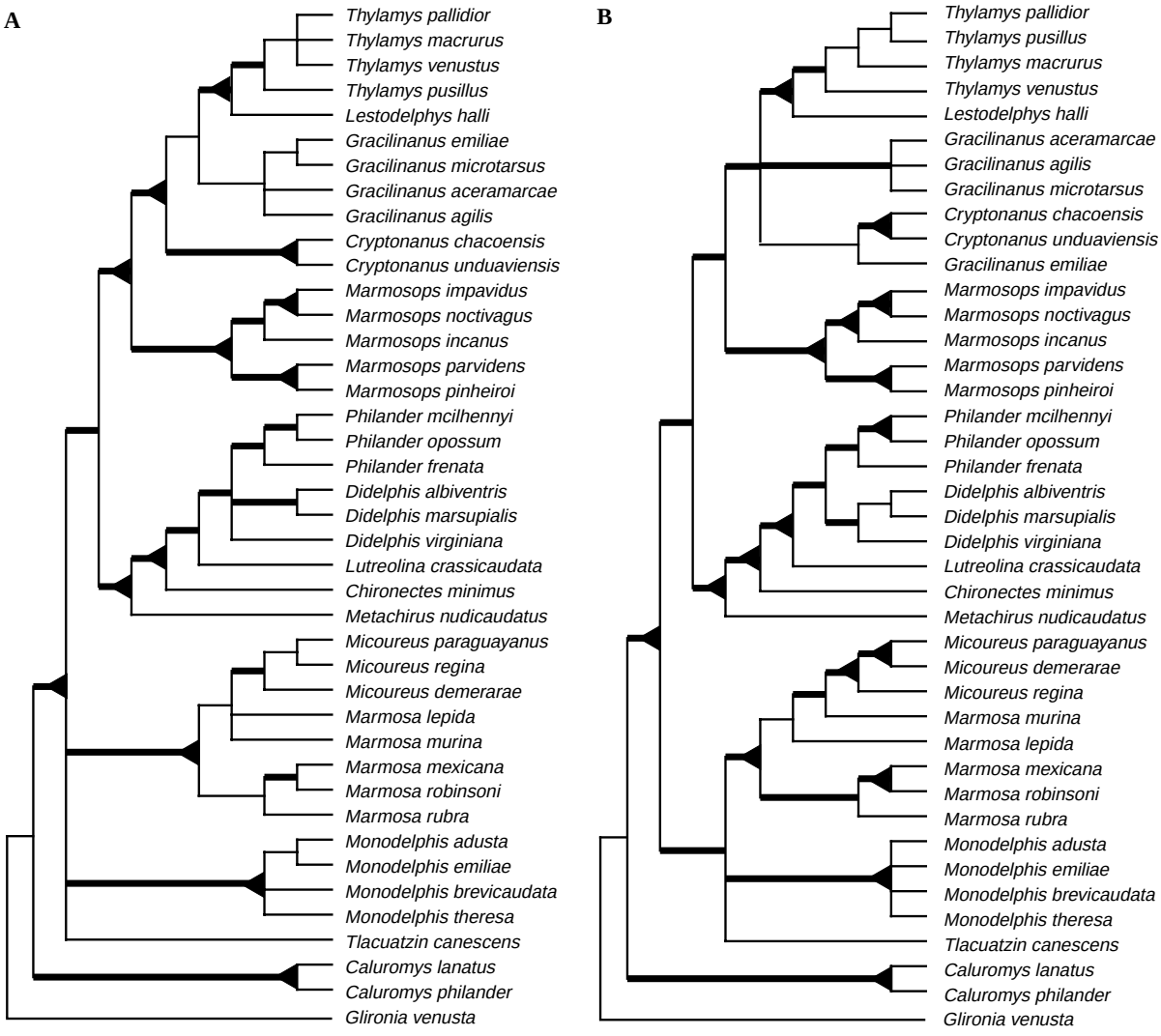


Fig. 4. The strict consensus of equally parsimonious trees resulting from separate analysis of (A) the IRBP data, (B) the DMP1 data with gaps coded as phylogenetic characters. Line thickness and buttressing are as described in Fig. 1.

A few minor differences characterize the trees based on DMP1 sequences with indels coded as phylogenetic characters versus those in which indels were treated as missing data. In the former analysis, *Marmosa lepida* joins at the base of a clade containing species of *Micoureus* and *Marmosa murina* with bootstrap support of 68%. When gaps are coded as missing data, this taxon forms an unresolved trichotomy with that clade and the remaining species of *Marmosa*. Moreover, two poorly supported clades—one uniting *G. aceramarcae* and *G. agilis* (bootstrap = 59%) and one uniting all species of *Gracilinanus* and *Cryptonanus* in a single clade (bootstrap = 55%)—are only recovered when indels are treated as missing data. Although most of the remaining nodes have similar bootstrap values between the two analyses, support for the two species of *Cryptonanus* + *G. emiliae* rises from 66 to 82% when indels are coded as missing, whereas support for *M. demerarae* + *M. paraguayanus* decreases from 95 to 87%.

Four clades in the combined data parsimony tree are characterized by unique and unreversed deletions in DMP1 (indicated by stars in Fig. 5). All species of *Marmosops* share a 3 bp deletion; *Didelphis*, *Philander*, and *Lutreolina* share an independent 3 bp deletion; species of *Micoureus*, *M. murina*, and *M. lepida* share an 18 bp deletion; and the two species of *Caluromys* share two different 3 bp deletions. The remaining seven deletions (ranging from 3 to 15 bp in length) are either autapomorphic ($n = 4$) or homoplastic ($n = 3$).

Likelihood analysis of the IRBP data under its best-fit model (TVMef+I+Γ+clock) resulted in rooting the topology on *Metachirus*. We noted a similar phenomenon in previous likelihood analyses of IRBP for a slightly different set of didelphid taxa (Voss and Jansa, 2003); however, this topology is not significantly different from those that place the root so that didelphines are monophyletic as assessed by a SH test. Once this topology is re-rooted with *Glironia*, it is nearly identical to and

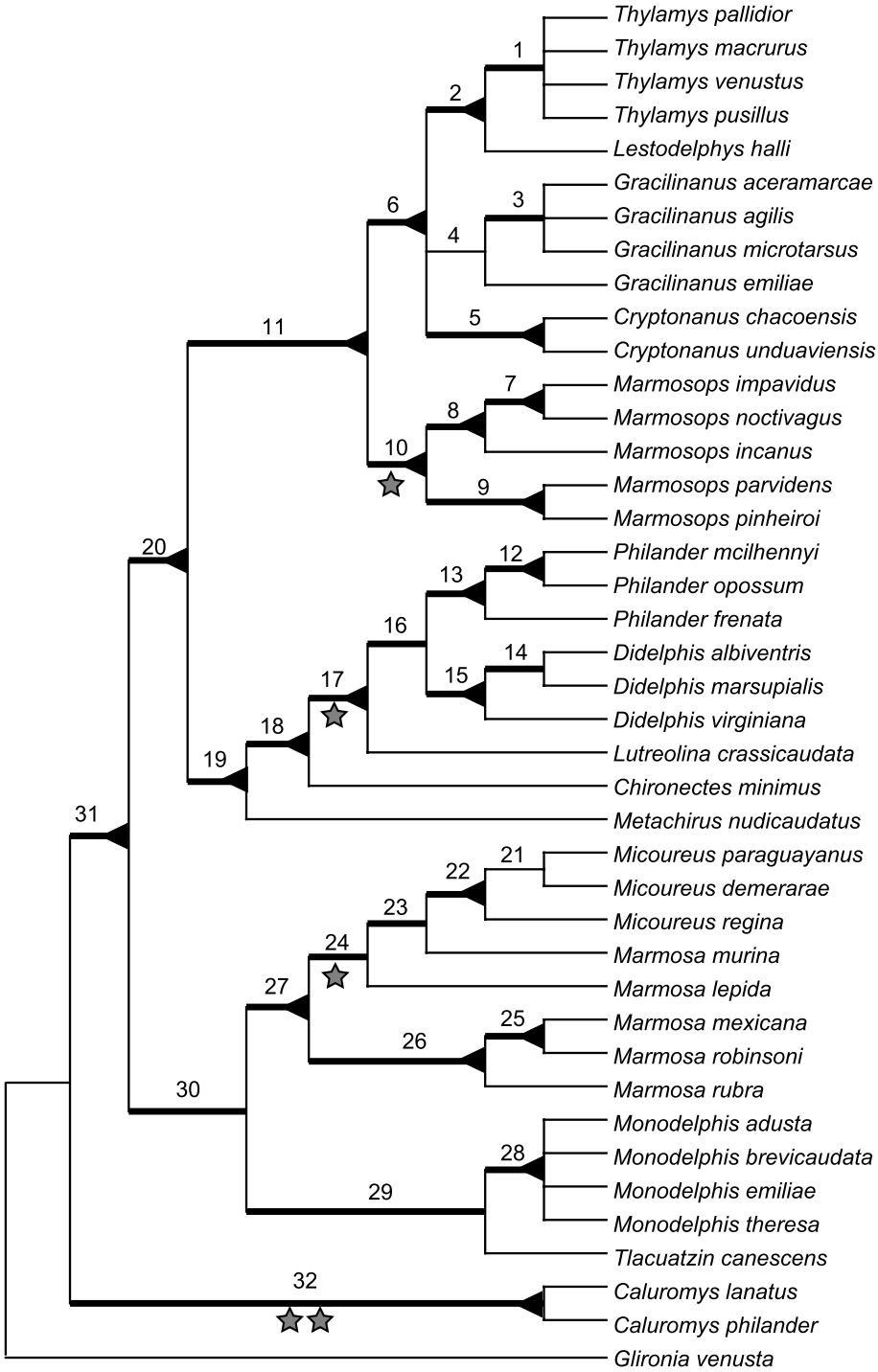


Fig. 5. The strict consensus of 20 equally parsimonious trees resulting from analysis of the combined molecular (DMP1 + IRBP) and nonmolecular data. Numbers identify nodes and refer to Table 5 for support values. Gray stars indicate unique and unreversed deletions in the DMP1 sequences. Line thickness and buttressing are as described in Fig. 1.

entirely congruent with the strict consensus of trees from the unweighted parsimony analysis (Fig. 4A). Likelihood analyses of the DMP1 data under its best-fit model (GTR + Γ) resulted in a tree that is fully consistent with the parsimony DMP1 tree (Fig. 4B). Moreover, both likelihood and parsimony analysis of the separate gene data sets resulted in similar nodal support values Table 5).

Despite the inferred differences in nucleotide substitution dynamics between the two genes described above, we found no significant phylogenetic incongruence between the IRBP and DMP1 data sets when assessed with

Table 5
Bootstrap support (parsimony and likelihood) and nodal posterior probability estimates (Bayesian uniform model analysis) from each dataset for nodes that appear in the combined data parsimony (Fig. 5) or likelihood trees (Fig. 6)

Node[a]	Parsimony				Likelihood			Bayesian
	IRBP	DMP1[b]	DMP1[b] + IRBP	All data[c]	IRBP	DMP1[d]	DMP1[d] + IRBP	DMP1[d] + IRBP
1	87	76	95	86	95	80	97	1.0
2 (A)	98	100	100	100	99	100	100	1.0
3	—	91	89	83	—	91	89	1.0
4	61	32	55	70	75	17	48	—
5	100	100	100	100	100	100	100	1.0
6 (B)	100	97	100	100	100	99	100	1.0
7	100	100	100	100	100	100	100	1.0
8	87	95	97	96	74	95	87	1.0
9	100	100	100	100	100	100	100	1.0
10	99	100	100	100	99	100	100	1.0
11 (C)	97	85	97	97	95	81	97	1.0
12	78	99	100	100	95	72	100	1.0
13	80	92	98	98	88	91	99	1.0
14	75	62	85	76	78	98	87	1.0
15	32	88	91	99	82	97	98	1.0
16 (D)	77	79	96	92	89	94	98	1.0
17 (E)	81	98	99	100	94	99	100	1.0
18 (F)	100	100	100	100	100	100	100	1.0
19 (G)	98	100	100	100	97	98	100	1.0
20 (H)	77	78	96	97	77	73	97	1.0
21	—	95	61	62	—	89	31	—
22	89	99	99	100	94	100	100	1.0
23	—	80	82	77	13	90	92	1.0
24 (I)	74	68	97	86	79	77	92	1.0
25	100	100	100	100	100	100	100	1.0
26	54	96	98	99	59	95	98	1.0
27	98	98	100	100	99	96	100	1.0
28	100	100	100	100	100	100	100	1.0
29	39	58	57	81	42	44	47	—
30 (J)	57	81	89	89	67	89	95	1.0
31	100	100	100	100	98	89	97	1.0
32 (K)	100	100	100	100	100	100	100	1.0
33	—	66	44	30	—	82	51	—
34	—	30	18	—	—	48	34	—
35	68	—	31	30	68	5	48	—
36	28	39	35	9	35	55	50	—

Dashes indicate bootstrap support <5% or Bayesian posterior probabilities <0.95.
[a] Node numbers refer to Figs. 5 or 6. Letters refer to intergeneric groupings as shown in Fig. 1.
[b] Gaps treated as phylogenetic characters.
[c] DMP1 (gaps as characters) + IRBP + morphology.
[d] Gaps treated as missing data.

the parsimony-based ILD test ($p = 0.378$). In contrast, likelihood-based SH tests examining the fit of the DMP1 data to the IRBP tree and vice versa suggested that two genes contain conflicting phylogenetic signal (DMP1 on IRBP tree: $-2\Delta\ln L = 49.72$, $p = 0.002$; IRBP on DMP1 tree: $-2\Delta\ln L = 23.68$, $p = 0.034$). However, none of the nodes that differ between the two trees are supported by bootstrap values >50%. Moreover, neither gene alone was able to reject the combined data tree (DMP1 on combined data tree: $-2\Delta\ln L = 9.87$, $p = 0.305$; IRBP on combined data tree: $-2\Delta\ln L = 5.87$, $p = 0.387$).

Likelihood analysis of the combined gene data set under its best-fit model (TIM + I + Γ) also resulted in a tree (Fig. 6) that was largely congruent with the combined gene parsimony tree. The principal difference between the two topologies concerns the placement of *Tlacuatzin*. Whereas parsimony analysis of the combined gene data set places *Tlacuatzin* sister to *Monodelphis* with weak bootstrap support (57%, tree not shown), likelihood analysis of the combined gene data set under a single model places *Tlacuatzin* sister to the clade including *Monodelphis*, *Micoureus*, and *Marmosa* with similarly weak support (50%, Fig. 6). However, when the nonmolecular data are added to the combined gene data for parsimony analysis, bootstrap support for the relationship between *Tlacuatzin* and *Monodelphis* increases to 81%, despite the fact that nonmolecular data alone cannot resolve the position of *Tlacuatzin* with any certainty (Voss and Jansa, 2003).

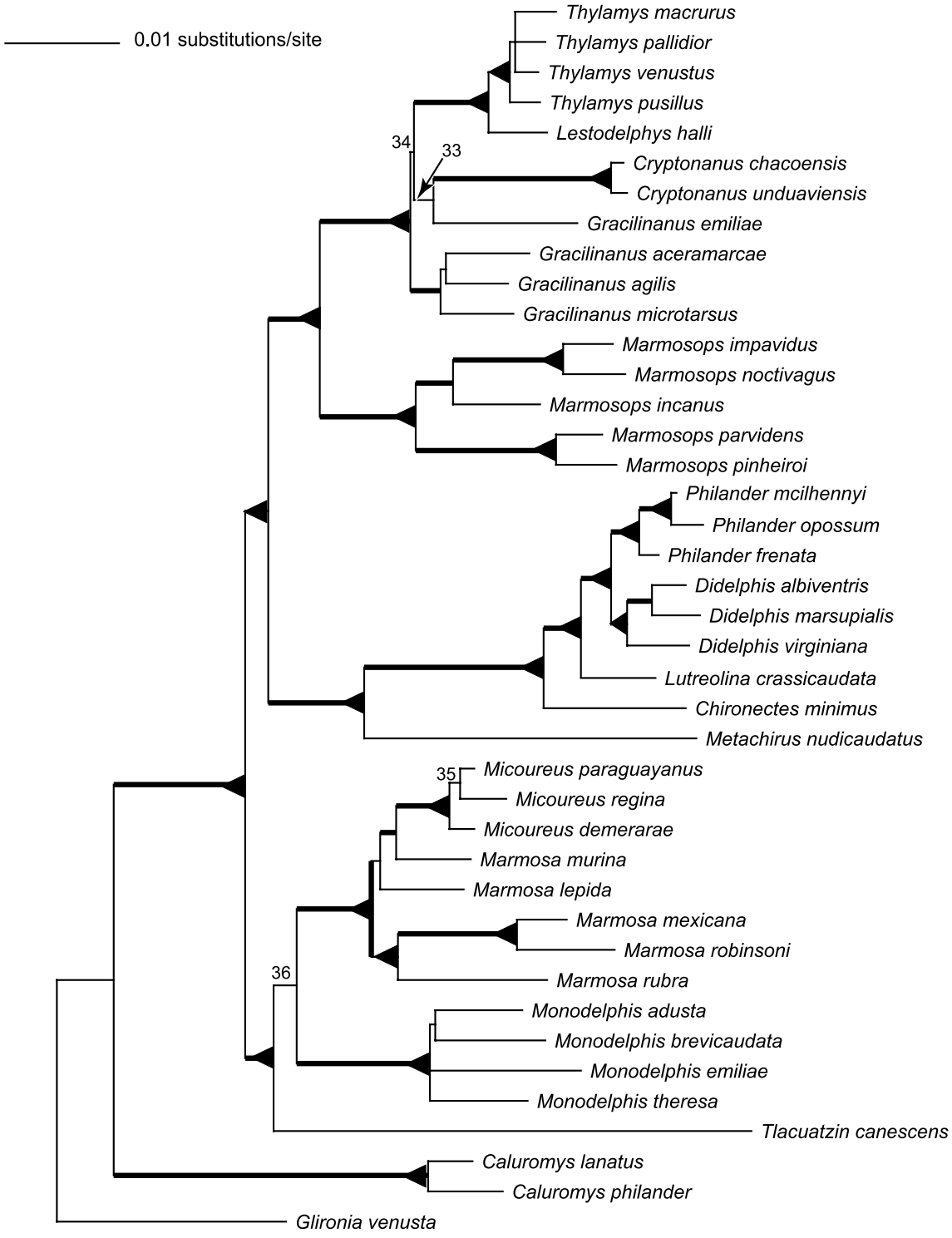


Fig. 6. Tree resulting from likelihood analysis of the combined gene data set under its best-fit model (TIM+I+Γ). Line thickness and buttressing are as described in Fig. 1. Node numbers identify additional nodes that were not recovered in the parsimony analysis (Fig. 5) and refer to Table 5 for support values.

For the Bayesian analyses, there were no nodes with markedly different posterior probability estimates between the uniform and mixed model analyses; all nodes with posterior probabilities >0.95 in the partitioned analysis also received values >0.95 in the uniform model analysis. Conversely, any node with a posterior probability <0.95 in the partitioned analysis was similarly poorly supported in the uniform model analysis. Therefore, although the two genes have significantly different nucleotide substitution patterns, assuming a single mod-

el of evolution across both genes does not markedly affect our phylogenetic conclusions. Bayesian nodal posterior probability values do not appear to overestimate support for nodes as inferred from parsimony or likelihood bootstrapping (contra Suzuki et al., 2002). Only three nodes receive posterior probability values <95%; each of these also receives bootstrap support <70% in either likelihood or parsimony analyses. Moreover, all of the nodes that are supported by posterior probabilities >0.95 also receive bootstrap support >75% in both the parsimony and uniform model likelihood analyses.

3.4. Tests for selection

In the majority of pairwise comparisons, DMP1 exhibits higher divergence values across the whole gene than does IRBP (Fig. 7A), but this divergence is proportioned differently among codon positions for the two genes. DMP1 has a lower proportion of third position transitions than does IRBP (Figs. 7B and C) but a much higher proportion of change at first and second codon positions (not shown). Given that changes in first and second codon positions are predominantly replacement substitutions, we should expect to see more sites in DMP1 with an elevated ratio of replacement to silent substitutions (ω) than in IRBP.

As expected given these patterns of divergence, models that allow for a proportion of sites to assume $\omega > 1$ (M3 and M8) fit the data much better than corresponding models that do not (M0 and M7, respectively). M3 provides a much better fit to the data than M0 with log-likelihood values $\ell_{M3} = -9904.42$ and $\ell_{M0} = -10050.15$, respectively ($2\Delta\ell = 291.46$, $df = 6$, $p \ll 0.001$). Similarly, M8 provides a better fit than M7 with $\ell_{M8} = -9906.78$ and $\ell_{M7} = -9923.58$ ($2\Delta\ell = 33.59$, $df = 2$, $p \ll 0.001$). Parameter estimates from the M3 model suggest that about 6% of sites across the two genes are under positive selection, with 5.6% of sites having $\hat{\omega} = 1.87$ and 0.2% have $\hat{\omega} = 7.54$; M8 estimates far fewer sites under positive selection with 3% of sites having $\hat{\omega} = 2.63$. In comparisons discussed below, we report estimates under the more conservative M8 model.

Although models that allow for positively selected sites fit the combined gene data set better than those that

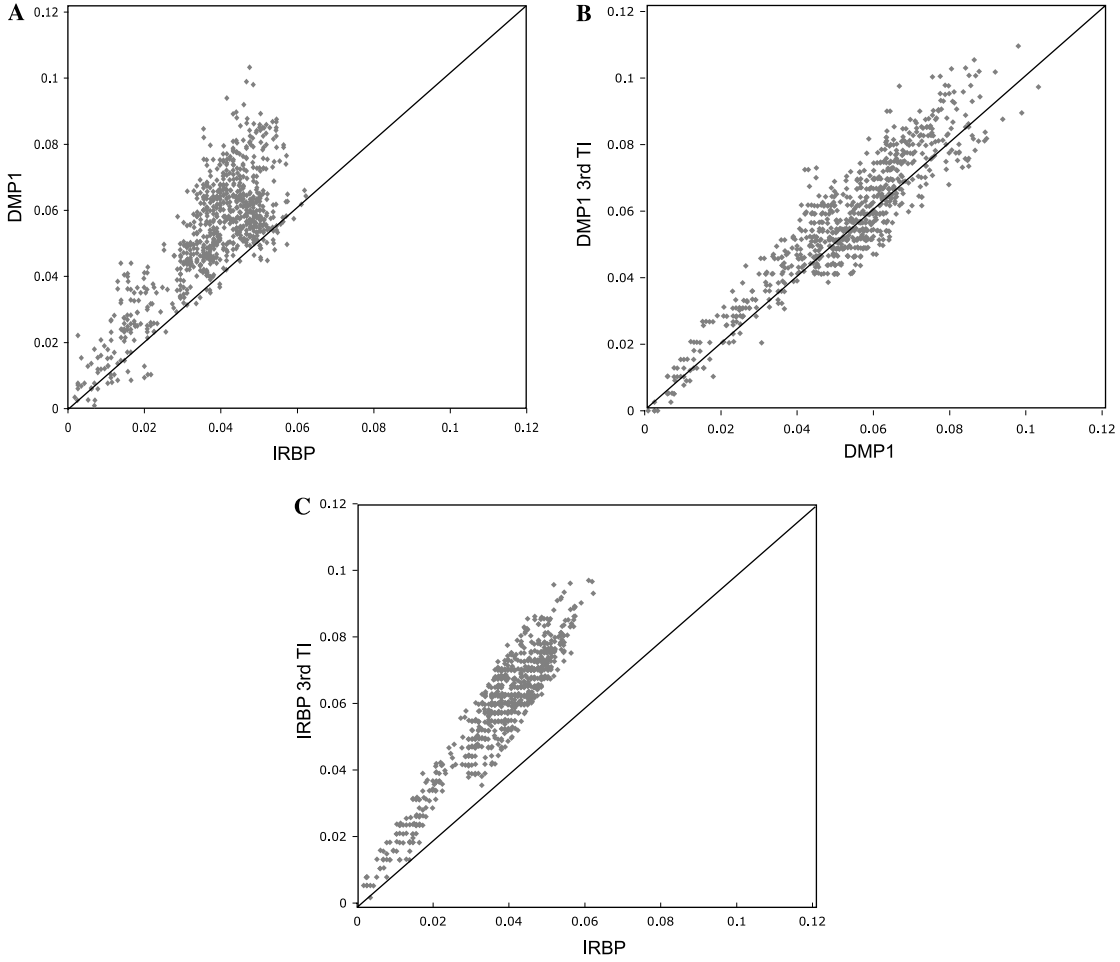


Fig. 7. Scatterplots of distances (*p*-distance) for pairwise comparisons of (A) overall sequence divergence in DMP1 vs. IRBP, (B) third position transitions in DMP1 vs. overall DMP1 divergence, and (C) third position transitions in IRBP vs. overall IRBP divergence.

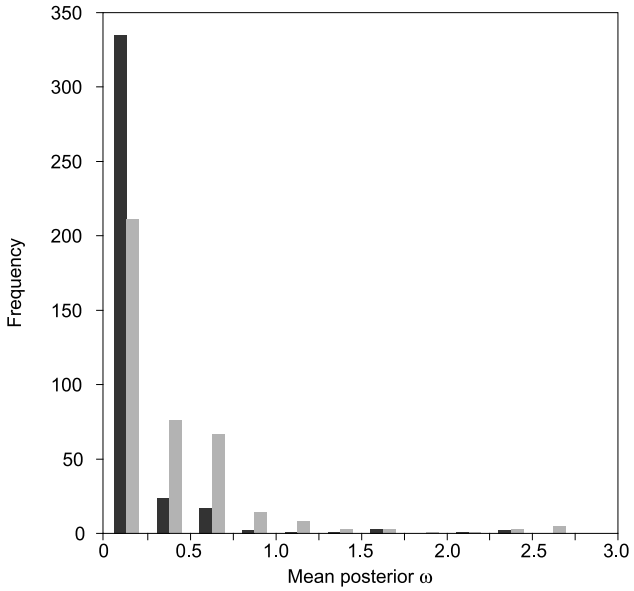


Fig. 8. Distribution of mean posterior ω values for sites in IRBP (dark gray) and DMP1 (light gray) as estimated with the M8 model as implemented in the codeml program of PAML (Yang, 1997).

do not, the number of positively selected sites differs markedly between the two genes (Fig. 8). IRBP has only seven sites with an estimated $\omega > 1$, none of which have posterior probability values >95% of being under positive selection. In contrast, DMP1 has 17 sites with $\omega > 1$, five of which have posterior probabilities >95% of being under positive selection. The proportion of sites exhibiting an estimated $\omega > 1$ differs significantly between the two genes when assessed with a χ^2 test ($\chi^2 = 4.14$, $\mathrm{df} = 1$, $p = 0.04$). Not only does DMP1 exhibit a higher proportion of sites under positive selection, but it also has far fewer sites that are under strong purifying selection than does IRBP. Because there is no objective value of $\omega < 1$ that indicates strong purifying selection, we arbitrarily define sites under strong purifying selection as those with $\omega \leqslant 0.25$, noting that our results are robust to more and less stringent criteria. For IRBP, 87% of the codon positions are under strong purifying selection as compared with 54% of codon positions in DMP1. This proportion is highly significant when assessed with a χ^2 test ($\chi^2 = 101.0$, $\mathrm{df} = 1$, $p \ll 0.001$). Therefore, it appears that the DMP1 protein is under considerably less selective constraint than the IRBP protein and has a few sites that may be evolving under positive selection.

4. Discussion

4.1. Phylogenetic relationships

Although IRBP and DMP1 clearly differ in function, selective pressures, and nucleotide substitution patterns, estimates of didelphid relationships based on separate phylogenetic analyses of these genes are remarkably congruent whether patterns of nucleotide substitution are explicitly modeled or not. Moreover, the phylogenetic information provided by IRBP and DMP1 tends to be complementary rather than redundant in simultaneous analyses of our taxon-dense data sets. A brief review of recent progress in didelphid phylogeny estimation based on IRBP and morphology provides a necessary context for these assessments and for discussion of the comparative evolution of these two genes.

The strength of our initial phylogenetic conclusions based on IRBP sequences alone (Jansa and Voss, 2000) was not compromised by adding new nonmolecular characters, so much as it was by adding new taxa in subsequent analyses (Voss and Jansa, 2003; Voss et al., 2005). In effect, nonmolecular (morphological + karyotypic) data provide strong support for just a few clades, all of which were also strongly supported by IRBP, and no examples of "hard" incongruence were discovered. Instead, all observed conflicts between nonmolecular and IRBP topologies resulting from separate analyses involved nodes with weak support in one or both data sets. Although certain nodes that were strongly supported by IRBP lost some support when the nonmolecular data were added, there was a net increase in nodal support when these data sets were combined in a simultaneous analysis (Table 6 in Voss and Jansa, 2003).

In contrast, adding new taxa substantially eroded support for some clades. Thus, certain relationships that were completely resolved in our initial study of 21 didelphid species collapsed entirely when new taxa were added, and nodal support was generally lower for the more taxon-dense analysis. For example, clade J (Fig. 1) was strongly supported (92% parsimony bootstrap) in our 21-taxon tree but was not recovered in a subsequent analysis with 31 taxa. The resulting polytomy is attributable to including additional species of *Marmosa*, *Monodelphis*, and *Micoureus*, as well as the enigmatic taxon *Tlacuatzin canescens* (a species previously assigned to *Marmosa*). Apparently, these new taxa destabilized a node that may have been artefactually supported due to sparse taxon sampling in our initial study.

Our expectation that DMP1 sequences would help recover some of this lost resolution proved well-founded. As documented herein, these new characters restored a group containing *Micoureus*, *Marmosa*, and *Monodelphis*, to which *Tlacuatzin* also appears to belong. This revised concept of clade J (differing from our original concept by containing *Tlacuatzin*) receives most of its support from DMP1 (Table 5). Although clade J does not appear in the strict consensus of minimum-length IRBP trees (Fig. 4A), there is in fact some support for this relationship in the IRBP data set (Table 5). Our combined gene analysis recovers clade J with strong

support whether explicit models of nucleotide substitution are utilized (1.0 Bayesian posteriors in uniform and mixed model analyses; 95% likelihood bootstrap) or not (89% parsimony bootstrap), nor does support for this grouping change when nonmolecular characters are included in a simultaneous analysis of all three data sets. However, the exact position of *Tlacuatzin* within clade J remains uncertain. Only the combined data parsimony analysis recovers *Tlacuatzin* as sister to *Monodelphis* with any substantial support (bootstrap = 81%), whereas the likelihood analysis recovered a different arrangement that placed *Tlacuatzin* sister to the clade including *Marmosa*, *Micoureus*, and *Monodelphis* (bootstrap = 50%).

In addition to increased resolution, several clades receive improved support with the addition of DMP1 sequences to previously analyzed (IRBP + nonmolecular) data sets. Support for all relationships among species of *Didelphis* and *Philander* improves, as does support for clades C, E, and G (Fig. 5; Table 5). However, the most dramatic increase in support occurs for clade H. This node is particularly significant as it effectively falsifies the hypothesized monophyly of the small "marmosine" opossums (Creighton, 1984; Reig et al., 1987), an obviously artificial group of genera containing numerous species formerly classified as or allied with *Marmosa* (sensu lato; Tate, 1933). The lack of any strong evidence for "marmosine" monophyly was reviewed by Jansa and Voss (2000), who remarked that clade H had previously appeared in analyses of DNA–DNA hybridization data reported by Kirsch et al. (1995) and Kirsch and Palma (1995). Although clade H was not supported by Palma and Spotorno's (1999) analysis of mitochondrial 12S DNA sequences, it received reasonably high nodal support in a subsequent study of the nuclear transthyretin intron (Steiner et al., 2005). Neither IRBP nor DMP1 alone provide particularly strong support for clade H, but when combined, they provide compelling evidence for "marmosine" paraphyly.

Our present analysis, while not conclusive, also supports the emerging picture of *Micoureus* as a monophyletic group nested within a paraphyletic *Marmosa* (Fig. 5). Among the species of *Marmosa* that we have sampled, three (*mexicana, robinsoni,* and *rubra*) now form a well-supported cluster, while the remaining two (*murina* and *lepida*) form a clade with *Micoureus*. Although this latter clade receives only moderate bootstrap support (86%) in our combined data analysis, we note that all members of this group share an 18 bp deletion in DMP1. Conclusive evidence of relationships among species of *Marmosa* and *Micoureus* will require far denser species-level sampling, as well as improved nodal support from additional characters. However, it should be noted that no phylogenetic study to date has provided compelling evidence for the monophyly of *Marmosa*, even in the strictest sense.

Although DMP1 generally improves resolution and support across didelphine phylogeny, an obvious example of conflict between this gene and IRBP concerns relationship among species of *Cryptonanus* and *Gracilinanus* (compare Figs. 4A and B). Members of *Cryptonanus* are morphologically distinguishable from species of *Gracilinanus*, and previous analyses of IRBP sequences have supported the reciprocal monophyly of these genera (Voss et al., 2005). Surprisingly, we recovered a sister-group relationship between *G. emiliae* and *Cryptonanus* in our separate analyses of DMP1 sequences (Fig. 4B) and in our likelihood analysis of the combined gene data set (Fig. 6). However, we note that this relationship receives only weak support and does not appear in our combined data parsimony analysis (Fig. 5). At most, therefore, this is an example of decidedly "soft" conflict between two otherwise impressively congruent data sets.

4.2. Molecular evolution

Several aspects of molecular evolution can affect the phylogenetic utility of a given gene sequence. Among these, rate heterogeneity, skewed base frequencies, and departures from base-compositional stationarity across taxa are thought to adversely affect phylogenetic accuracy (Galtier and Gouy, 1995; Lin and Danforth, 2004; Yang, 1996). In particular, a high evolutionary rate coupled with severe rate heterogeneity is most frequently associated with the apparently poor performance of particular gene sequences (Lin and Danforth, 2004; Yang, 1996, 1998). If rapid evolutionary change is restricted to only a few sites (as indicated by a low value for the α parameter describing the shape of the gamma distribution), these sites tend to become saturated such that phylogenetic signal is overwhelmed by noise. Such problems do not, however, appear to affect either IRBP or DMP1. Neither gene departs from base-compositional stationarity, and both are relatively slowly evolving, with a maximum uncorrected distance among our taxa of 6.2% for IRBP and 10.3% for DMP1. Although IRBP exhibits significantly more rate heterogeneity in our mixed model Bayesian analysis than does DMP1 (Table 3), neither gene shows evidence of sequence saturation (Figs. 7B and C).

Based on these qualities, IRBP and DMP1 show similar promise as phylogenetic markers. Indeed, they are closely comparable in several measures of phylogenetic utility, including topological congruence and degree of homoplasy (CI and RI values, Table 4). However, DMP1 outperforms IRBP in terms of the number of informative sites per sequenced base pair, number of resolved nodes in strict consensus topologies, and the relative support for these nodes (Table 5). Moreover, the DMP1 alignment exhibits a number of parsimony-informative insertion–deletion events whereas the IRBP

alignment does not. This improved resolution and support is due in part to the fact that DMP1 evolves about 30% more quickly than IRBP but also exhibits less among-site rate heterogeneity (α values, Table 3). This reduced rate heterogeneity can in turn be explained by the fact that DMP1 has considerably more sites free to vary than does IRBP (P_{inv} values, Table 3). Because changes in DMP1 are not concentrated at a particular site (e.g., third position) but are distributed across sites in the gene, DMP1 and IRBP have similarly low levels of homoplasy despite the accelerated substitution rate for DMP1.

Our analyses of selection suggest that IRBP and DMP1 have experienced very different evolutionary constraints at the protein level (Fig. 8), and that these constraints determine, at least in part, the phylogenetic behavior of the two genes. In particular, DMP1 has far fewer sites under strong purifying selection than does IRBP, and it exhibits a number of sites that may have experienced positive directional selection. Previous studies suggest that this pattern of selection on DMP1 may be widespread across mammalian taxa. For example, van den Bussche et al. (2003) noted that pairwise estimates of dN/dS ratios were higher for DMP1 than the nuclear RAG2 locus for several species of bats. Thus, these different selective pressures might explain why DMP1 has more sites free to vary and less among-site rate heterogeneity than IRBP. However, why DMP1 should show reduced selective constraints at the protein level compared to other nuclear genes remains to be fully understood.

One possible explanation concerns the functional properties of the proteins in question. The IRBP gene, for example, is expressed in the vertebrate retina, where it plays a role in transporting retinoids during the visual cycle (Gonzalez-Fernandez, 2003). Studies suggest that the IRBP protein folds so as to create a specific, hydrophobic ligand-binding site that requires tryptophan residues to successfully bind retinol (Gonzalez-Fernandez, 2003; Loew and Gonzalez-Fernandez, 2002). Plausibly, proper folding and biological activity of the IRBP protein requires conservation of a linear amino acid sequence and maintenance of specific residues at particular sites. Perhaps for this reason, the IRBP gene may be unable to tolerate insertion–deletion events (at least for the first 1000 bp of exon 1) and is highly conserved at the amino acid level.

In contrast, the DMP1 gene is expressed primarily in vertebrate teeth and bones, where the protein plays a role in the calcification of mineralized tissues (Butler and Ritchie, 1995; Feng et al., 2003). The full-length DMP1 protein has been difficult to isolate from living tissues, and recent work suggests that a functional, full-length product may simply not exist in vivo (Qin et al., 2003). Rather, DMP1 appears to be posttranslationally cleaved into two shorter molecules that are biologically active. Additional studies suggest that the proper functioning of DMP1 does not depend on *intra*molecular folding, so much as on formation of a *inter*molecular matrix composed of short protein fragments that complex calcium ions (He et al., 2003). To successfully bind calcium, these short fragments need to contain phosphorylated serine residues in close proximity to the acidic amino acids aspartic and glutamic acid. Therefore, selection on this gene presumably acts primarily to conserve peptide cleavage sites (typically at Asp residues; Qin et al., 2003) along with Ser, Asp, and Glu (SDE) residues (Kawasaki et al., 2004).

If this functional model is correct, then we should expect SDE residues in DMP1 to be under strong purifying selection while other residues should have higher ω values consistent with relaxed selection. To test this prediction, we examined the posterior mean ω estimates for sites that were SDE in an arbitrarily chosen taxon (*Caluromys lanatus*) versus ω estimates for all other sites. The distribution of these ω estimates differs markedly between the two categories. Consistent with our predictions, 97% of the sites containing Ser, Asp, or Glu were under strong purifying selection ($\omega \leqslant 0.25$ by our arbitrary threshold criterion), while sites containing any other residue had a much lower proportion (85%) of sites with $\omega \leqslant 0.25$, a difference that is significant when assessed with a χ^2 test ($\chi^2 = 16.6$; df $= 1$; $p \ll 0.001$). Moreover, a similar analysis conducted for IRBP resulted in no significant difference in selective constraints between SDE sites and other amino acids ($\chi^2 = 0.56$; df $= 1$; $p = 0.45$). Therefore, the unique functional characteristics of DMP1 appear to constrain only a few type of amino acids while allowing others to change more freely.

We hypothesize that such functional constraints at the protein level affect the nucleotide substitution parameters for the DMP1 gene and influence its phylogenetic utility. The functional need to maintain Ser, Asp, and Glu, all of which have either A or G in the first and second codon position, plausibly explains why DMP1 sequences are rich in these nucleotides and exhibit relatively few changes between adenine and any other nucleotide (Fig. 3). The apparent lack of functional constraints on other residues allows DMP1 to have several sites free to vary, a high rate of evolution, and relatively little among-site rate heterogeneity. On the other hand, because the biological activity of DMP1 may not require conservation of long series of amino acids (Kawasaki et al., 2004), the gene can tolerate insertion–deletion events and becomes very difficult to align at higher taxonomic levels (van den Bussche et al., 2003; personal observation). Thus, patterns of selection may restrict the phylogenetic utility of DMP1 to relatively well-circumscribed groups that do not exhibit extensive regions of alignment ambiguity, but within these groups, its relaxed selective constraints make it a valuable phylogenetic tool.

Acknowledgments

We are grateful to the curators and collections support personnel of the many museums that provided tissue samples for this work. In particular, we thank Phil Myers and Steve Hinshaw (UMMZ), Bruce Patterson and Bill Stanley (FMNH), Mark Engstrom and Burton Lim (ROM), Jim Patton (MVZ), and Robert Baker (TTU). We thank Keith Barker and Susan Weller for numerous discussions and for their careful reading of earlier versions of this paper. The University of Minnesota Supercomputing Institute provided valuable computing time and resources. This work was supported by NSF grant DEB-0211952 and by funds provided by the University of Minnesota to S.A.J.

References

Baker, R.H., Wilkinson, G.S., DeSalle, R., 2001. Phylogenetic utility of different types of molecular data used to infer evolutionary relationships among stalk-eyed flies (Diopsidae). Syst. Biol. 50, 87–105.

Barker, F.K., Lutzoni, F.M., 2002. The utility of the Incongruence Length Difference test. Syst. Biol. 51, 625–637.

Chang, B.S.W., Campbell, D.L., 2000. Bias in phylogenetic reconstruction of vertebrate rhodopsin sequences. Mol. Biol. Evol. 17, 1220–1231.

Creighton, G.K., 1984. Systematic studies on opossums (Didelphidae) and rodents (Cricetidae). Ph.D. Dissertation, University of Michigan, Ann Arbor.

deBry, R.W., Sagel, R.M., 2001. Phylogeny of Rodentia (Mammalia) inferred from the nuclear-encoded gene IRBP. Mol. Phylogenet. Evol. 19, 290–301.

Dolphin, K., Belshaw, R., Orme, C.D.L., Quicke, D.L.J., 2000. Noise and incongruence: interpreting results of the incongruence length difference test. Mol. Phylogenet. Evol. 17, 401–406.

Fang, Q.Q., Mitchell, A., Regier, J.C., Mitter, C., Friedlander, T.P., Poole, R.W., 2000. Phylogenetic utility of the nuclear gene dopa decarboxylase in noctuoid moths (Insecta: Lepidoptera: Noctuoidae). Mol. Phylogenet. Evol. 15, 472–486.

Farris, J.S., Källersjö, M., Kluge, A.G., Bult, C., 1995. Constructing a significance test for incongruence. Syst. Biol. 44, 570–572.

Felsenstein, J., 1993. PHYLIP (Phylogeny Inference Package), version 3.5c. Department of Genetics, University of Washington.

Feng, J.Q., Huang, H., Lu, Y., Ye, L., Xie, Y., Tsutsui, T.W., Kunieda, T., Castranio, T., Scott, G., Bomewald, L.B., Mishina, Y., 2003. The Dentin Matrix Protein 1 (DMP1) is specifically expressed in mineralized, but not soft, tissues during development. J. Dent. Res. 82, 776–780.

Galtier, N., Gouy, M., 1995. Inferring phylogenies from DNA sequences of unequal base compositions. Proc. Natl. Acad. Sci. USA 92, 11317–11321.

Gardner, A.L., 2005. Didelphimorphia. In: Wilson, D.E., Reeder, D.M. (Eds.), Mammal Species of the World, third ed. Johns Hopkins University Press, Baltimore, in press.

Gonzalez-Fernandez, F., 2003. Interphotoreceptor retinoid-binding protein – an old gene for new eyes. Vis. Res. 43.

Graybeal, A., 1994. Evaluating the phylogenetic utility of genes: a search for genes informative about deep divergences among vertebrates. Syst. Biol. 43, 174–193.

Griffiths, C.S., Barrowclough, G.F., Groth, J.G., Mertz, L., 2004. Phylogeny of the Falconidae (Aves): a comparison of the efficacy of morphological, mitochondrial, and nuclear data. Mol. Phylogenet. Evol. 32, 101–109.

Groth, J.G., Barrowclough, G.F., 1999. Basal divergences in birds and the phylogenetic utility of the nuclear RAG-1 gene. Mol. Phylogenet. Evol. 12, 115–123.

Guindon, S., Gascuel, O., 2003. PHYML: A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52, 696–704.

He, G., Dahl, T., Veis, A., George, A., 2003. Nucleation of apatite crystals in vitro by self-assembled dentin matrix protein 1. Nat. Mater. 2, 552–558.

Hillis, D.M., 1998. Taxonomic sampling, phylogenetic accuracy, and investigator bias. Syst. Biol. 47, 3–8.

Hipp, A.L., Hall, J.C., Sytsma, K.J., 2004. Congruence versus phylogenetic accuracy: revisiting the incongruence length difference test. Syst. Biol. 53, 81–89.

Horovitz, I., 1999. A phylogenetic study of living and fossil platyrrhines. Am. Mus. Novit. 3269, 1–40.

Huelsenbeck, J.P., Ronquist, F., 2001. MrBayes: Bayesian inference of phylogeny. Bioinformatics 17.

Jansa, S.A., Voss, R.S., 2000. Phylogenetic studies on didelphid marsupials I. Introduction and preliminary results from nuclear IRBP gene sequences. J. Mammal. Evol. 7, 43–77.

Jansa, S.A., Weksler, M., 2004. Phylogeny of muroid rodents: relationships within and among major lineages as determined by IRBP gene sequences. Mol. Phylogenet. Evol. 31, 256–276.

Kawasaki, K., Suzuki, T., Weiss, K.M., 2004. Genetic basis for the evolution of vertebrate mineralized tissue. Proc. Natl. Acad. Sci. USA 101, 11356–11361.

Kirsch, J.A.W., Dickerman, A.W., Reig, O.A., 1995. DNA/DNA hybridization studies of carnivorous marsupials IV. Intergeneric relationships of the opossums (Didelphidae). Marmosiana 1, 57–78.

Kirsch, J.A.W., Palma, R.E., 1995. DNA/DNA hybridization studies of carnivorous marsupials. V.A further estimate of relationships among opossums (Marsupialia: Didelphidae). Mammalia 59, 403–425.

Lew, D., Pérez-Hernández, R., 2004. Una nueva especie del género *Monodelphis* (Didelphimorphia: Didelphidae) de la sierra de Lema, Venezuela. Memoria de la Fundación La Salle de Ciencias Naturales 2004 (2003), pp. 159–160.

Lin, C.P., Danforth, B.N., 2004. How do insect nuclear and mitochondrial gene substitution patterns differ? Insights from Bayesian analyses of combined datasets. Mol. Phylogenet. Evol. 30, 686–702.

Loew, A., Gonzalez-Fernandez, F., 2002. Crystal structure of the functional unit of interphotoreceptor binding protein (IRBP). Structure 10, 43–49.

Lundrigan, B.L., Jansa, S.A., Tucker, P.K., 2002. Phylogenetic relationships in the genus *Mus*, based on paternally, maternally, and biparentally inherited characters. Syst. Biol. 51, 410–431.

Madsen, O., Scally, M., Douady, C.J., Kao, D.J., DeBry, R.W., Adkins, R., Amrine, H.M., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Parallel adaptive radiations in two major clades of placental mammals. Nature 409, 610–614.

Murphy, W.J., Eizirik, E., O'Brien, S.J., Madsen, O., Scally, M., Douady, C.J., Teeling, E., Ryder, O.A., Stanhope, M.J., de Jong, W.W., Springer, M.S., 2001. Resolution of the early placental mammal radiation using Bayesian phylogenetics. Science 294, 2348–2351.

Nielsen, R., Yang, Z., 1998. Likelihood models for detecting positively selected amino acid sites and applications to the HIV-1 envelope gene. Genetics 148, 929–936.

Palma, R.E., Spotorno, A.E., 1999. Molecular systematics of marsupials based on the rRNA 12S mitochondrial gene: The phylogeny of didelphimorphia and of the living fossil microbiotheriid *Dromiciops gliroides* Thomas. Mol. Phylogenet. Evol. 13, 525–535.

Posada, D., Buckley, T.R., 2004. Model selection and model averaging in phylogenetics: advantages of Akaike Information Criterion and Bayesian approaches over likelihood ratio tests. Syst. Biol. 53, 793–808.

Posada, D., Crandall, K.A., 1998. Modeltest: testing the model of DNA substitution. Bioinformatics 14, 817–818.

Qin, C., Brunn, J.C., Cook, R.G., Orkiszewski, R.S., Malone, J.P., Veis, A., Butler, W.T., 2003. Evidence for the proteolytic processing of Dentin Matrix Protein 1. J. Biol. Chem. 278, 34700–34708.

Reeder, S.A., Bradley, R.D., 2004. Molecular systematics of neotomine-peromyscine rodents based on the dentin matrix protein 1 gene. J. Mammal. 85, 1194–1200.

Reig, O.A., Kirsch, J.A.W., Marshall, L.G., 1987. Systematic relationships of the living and Neocenozoic American ''opossum-like'' marsupials (Suborder Didelpimorphia), with comments on the classification of these and of the Cretaceous and Paleogene New World and European Metatherians. In: Archer, M. (Ed.), Possums and Opossums: Studies in Evolution. Surrey Beatty and Sons and the Royal Zoological Society of New South Wales, Sydney, pp. 1–89.

Roger, A., Sandbloom, O., Doolittle, W., Philipppe, H., 1999. An evaluation of elongation-factor-1-alpha as a phylogenetic marker for eukaryotes. Mol. Biol. Evol. 16, 218–233.

Shimodaira, H., Hasegawa, M., 1999. Multiple comparisons of log-likelihoods with applications to phylogenetic inference. Mol. Biol. Evol. 16, 1114–1116.

Solari, S., 2004. A new species of *Monodelphis* (Didelphimorphia: Didelphidae) from southeastern Peru. Mamm. Biol. 69, 145–152.

Soltis, D.E., Soltis, P.S., Mort, M.E., Chase, M.W., Savolainen, V., Hoot, S.B., Morton, C.M., 1998. Inferring complex phylogenies using parsimony: an empirical approach using three large DNA data sets for angiosperms. Syst. Biol. 47, 32–42.

Springer, M.S., Burk, A., Kavanagh, J.R., Waddell, V.G., Stanhope, M.J., 1997. 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., DeBry, R.W., Douady, C., Amrine, H.M., Madsen, O., de Jong, W.W., Stanhope, M.J., 2001. Mitochondrial versus nuclear gene sequences in deep-level mammalian phylogeny reconstruction. Mol. Biol. Evol. 18, 132–143.

Srinivasan, R., Chen, B., Gorski, J., George, A., 1999. Recombinant expression and characterization of Dentin Matrix Protein 1. Connect. Tissue Res. 40, 251–258.

Stanhope, M.J., Czelusniak, J., Si, J.-S., Nickerson, J., Goodman, M., 1992. A molecular perspective on mammalian evolution from the gene encoding interphotoreceptor retinoid binding protein, with convincing evidence for bat monophyly. Mol. Phylogenet. Evol. 1, 148–160.

Steiner, C., Tilak, M., Douzery, E.J.P., Catzeflis, F.M., 2005. New DNA data from a transthyretin nuclear intron suggest an Oligocene to Miocene diversification of living South America opossums (Marsupialia: Didelphidae). Mol. Phylogenet. Evol. 35, 363–379.

Suzuki, Y., Glazko, G.V., Nei, M., 2002. Overcredibility of molecular phylogenesis 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. Sinauer Associates.

Tate, G.H.H., 1933. A systematic revision of the marsupial genus *Marmosa*, with a discussion of the adaptive radiation of the murine opossums (*Marmosa*). Bull. Am. Mus. Nat. Hist. 66, 1–251.

Toyosawa, S., O'hUigin, C., Klein, J., 1999. The dentin matrix protein 1 gene of prototherian and metatherian mammals. J. Mol. Evol. 48, 160–167.

van den Bussche, R.A., Reeder, S.A., Hanse, E.W., Hoofer, S.R., 2003. Utility of the dentin matrix protein 1 (DMP1) gene for resolving mammalian intraordinal phylogenetic relationships. Mol. Phylogenet. Evol. 26, 89–101.

Voss, R.S., Emmons, L.H., 1996. Mammalian diversity in Neotropical lowland rainforests: a preliminary assessment. Bull. Am. Mus. Nat. Hist. 230, 115.

Voss, R.S., Lunde, D.P., Simmons, N.P., 2001. The mammals of Paracou, French Guiana: a Neotropcial rainforest fauna. Part 2. Nonvolant species. Bull. Am. Mus. Nat. Hist. 263, 1–236.

Voss, R.S., Jansa, S.A., 2003. Phylogenetic studies on didelphid marsupials II. Nonmolecular data and new IRBP sequences: Separate and combined analyses of didelphine relationships with denser taxon sampling. Bull. Am. Mus. Nat. Hist., 1–82.

Voss, R.S., Gardner, A.L., Jansa, S.A., 2004a. On the relationships of *Marmosa formosa* Shamel, 1930 (Marsupialia: Didelphidae), a phylogenetic puzzle from the Chaco of northern Argentina. Am. Mus. Novit. 3442, 1–18.

Voss, R.S., Yensen, E., Tarifa, T., 2004b. An introduction to *Marmosops* (Marsupialia: Didelphidae), with the description of a new species from Bolivia and notes on the taxonomy and distribution of other Bolivian congeners. Am. Mus. Novit. 3466, 1–40.

Voss, R.S., Lunde, D.P., Jansa, S.A., 2005. On the contents of *Graciliananus* Gardner and Creighton, 1989, with the description of a previously unrecognized clade of small didelphid marsupials. Am. Mus. Novit. (in press).

Yang, Z., 1996. Among-site rate variation and its impact on phylogenetic analysis. Trends Ecol. Evol. 11, 367–372.

Yang, Z., 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput. Appl. BioSci. 13, 555–556.

Yang, Z., 1998. On the best evolutionary rate for phylogenetic analysis. Syst. Biol. 47, 125–133.

Yang, Z., Nielsen, R., Goldman, N., Krabbe Pedersen, A.-M., 2000. Codon-substitution models for heterogeneous selection pressure at amino acid sites. Genetics 155, 431–449.

Yang, Z., Wong, W.S.W., Nielsen, R., 2005. Bayes empirical Bayes inference of amino acid sites under positive selection. Mol. Biol. Evol. 22, 1107–1118.