Rates of Protein Evolution across the Marsupial Phylogeny: Heterogeneity and Link to Life-History Traits

Agusto R. Luzuriaga-Neira* and David Alvarez-Ponce (iD)*

Department of Biology, University of Nevada, Reno, Nevada, USA

*Corresponding authors: E-mails: aluzuriaga@unr.edu; dap@unr.edu

Abstract

Despite the importance of effective population size (N_e) in evolutionary and conservation biology, it remains unclear what factors have an impact on this quantity. The Nearly Neutral Theory of Molecular Evolution predicts a faster accumulation of deleterious mutations (and thus a higher d_N/d_S ratio) in populations with small N_e; thus, measuring d_N/d_S ratios in different groups/species can provide insight into their N_e. Here, we used an exome data set of 1,550 loci from 45 species of marsupials representing 18 of the 22 extant families, to estimate d_N/d_S ratios across the different branches and families of the marsupial phylogeny. We found a considerable heterogeneity in d_N/d_S ratios among families and species, which suggests significant differences in their N_e. Furthermore, our multivariate analyses of several life-history traits showed that d_N/d_S ratios (and thus N_e) are affected by body weight, body length, and weaning age.

Key words: effective population size, d_N/d_S, life-history traits, marsupials, extinction.

Significance

We estimate the rates of protein evolution in each of the branches of the marsupial phylogeny. We found substantial heterogeneity (proteins have evolved faster in certain families and species compared with others), which provides clues about their different effective population sizes (proteins tend to evolve fast in species with small effective population sizes). Rates of protein evolution tend to be higher in lineages with small body size and body length, and with a high weaning age.

Introduction

The effective population size (N_e) is a critical concept in evolutionary and conservation biology because it is directly linked to the effectiveness of natural selection and the amount of neutral variation that a population contains. Species with small N_e tend to be at high risk of extinction because they have low levels of genetic diversity, are susceptible to accumulating deleterious mutations due to genetic drift, and are present potentially low rates of adaptative evolution (Charlesworth 2009). Comparative genomics studies in several groups such as rodents and primates (Wu and Li 1985; Weinreich 2001; Romiguier et al. 2014); mammals, birds, and reptiles (Figuet et al. 2016; Botero-Castro et al. 2017); and large versus small-bodied mammals (Popadin et al. 2007)

have validated the prediction that species with small N_e tend to display a high probability of fixation of deleterious mutations (Ohta 1976; Ohta and Ina 1995). The significance of N_e as an evolutionary and conservation factor is highlighted by several findings showing that N_e values are often lower than the census numbers of breeding individuals in a species (Crow and Morton 1955; Frankham 1995). Understanding what factors affect N_e is essential from the conservation point of view. Whereas a number of factors are known to affect N_e (Charlesworth 2009), it is still unclear which life-history traits impact N_e, and to what extent.

Because of the direct association between N_e and genome evolution, the patterns of genetic variation of a particular species reflect the past variations in its N_e (Nadachowska-Brzyska

et al. 2015). Mutations affecting protein-coding sequences can be classified into synonymous mutations (which largely evolve neutrally) and nonsynonymous mutations (which are often deleterious). The Nearly Neutral Theory of Molecular Evolution predicts a higher nonsynonymous to synonymous divergence ratio ($\omega = d_N/d_S$) in species with a small N_e (Kimura 1983; Ohta 1993). A decrease in N_e increases the fraction of nonsynonymous mutations with selection coefficients below $1/N_e$, which can be fixed by drift (Wright 1931; Kimura 1983). In agreement with this prediction, there is evidence for an accelerated accumulation of nonsynonymous mutations in mammals with low N_e. Ohta (1993) reported a higher d_N/d_S in primates compared with artiodactyls or rodents. This conclusion, based on 17 genes, was confirmed by the analysis of thousands of putative orthologs in primates and rodents (Wu and Li 1985; Weinreich 2001).

Given the general scarcity of direct estimates of N_e in mammals, correlates are usually used as proxies. Body mass is the most frequently used proxy of N_e, based on the general inverse relationship between body mass and N_e (Damuth 1981; Peters and Peters 1986; Damuth 1987). Several studies have shown a positive correlation between body mass and d_N/d_S (Nikolaev et al. 2007; Popadin et al. 2007; Romiguier et al. 2012, 2014). Besides body mass and d_N/d_S, a number of life-history traits such as lifespan and generation time have also been used as proxies of N_e, given their positive correlation with d_N/d_S (Nikolaev et al. 2007; Nabholz et al. 2008; Welch et al. 2008; Lartillot and Poujol 2011). In addition, James and Eyre-Walker (2020) showed that there is a positive correlation between a species' distribution area and its genetic diversity, using mitochondrial DNA data from 639 species of mammals. Because life-history traits and distribution area are related to N_e, they are directly associated with species vulnerability, thus representing intrinsic extinction factors of the species (Purvis et al. 2000; Cardillo et al. 2005; Collen et al. 2011; Rolland et al. 2020).

The link between life-history traits and N_e remains relatively unexplored. The few studies available are often limited by the number of species and loci studied. Moreover, these studies are largely limited to analyses of a few eutherian mammals or species from a wide variety of orders (Nikolaev et al. 2007; Popadin et al. 2007; Romiguier et al. 2012, 2014). Marsupials are an attractive group to test this link because they are a relatively young order (~160 Myr old; Luo et al. 2011) and are thus relatively homogenous in their biology, yet they exhibit substantial diversity in their life-history traits. In addition, a large exome data set encompassing 45 species of Australasian and American marsupials (representing 18 of the 22 extant families) has recently been generated (Duchêne et al. 2018).

The aim of the current study is two-fold. First, we computed the average rates of protein evolution (d_N/d_S ratios) in each branch of the marsupial phylogeny. This analysis revealed a high amount of heterogeneity in the rates of protein evolution both across families and within each family, which point to differences in N_e. Second, we used multivariate analyses to establish the relationship between the d_N/d_S ratio of each species and several life-history traits. We found that that body length, body mass and weaning age independently impact d_N/d_S, indicating that they are determinants of N_e.

Results

Differences in Rates of Protein Evolution among Marsupial Families and Species

We generated a concatenated alignment that included sequences for 1,472 exons. The concatenome had a total length of 808,941 base pairs for 43 marsupial species. Our analysis included a complete deletion step; thus, from 269,647 possible codons, only 122,308 were retained for analysis. We did not use the remaining codons as they contained missing data and/or ambiguities. We first used the M0 model, which assumes a homogeneous d_N/d_S value for all branches in the phylogenetic tree, to estimate an overall d_N/d_S of 0.1542. The log-likelihood of the alignment under this model was $l_0 = -2,040,465.59$.

We next applied the 19ω model to estimate a separate d_N/d_S for each of the 18 marsupial families included in our analysis, obtaining a log-likelihood value of $l_{19} = -2,039,789.51$. Comparison of the log-likelihoods of both models using a likelihood ratio test indicated that the 19ω model fitted the data significantly better than the M0 model [$2\Delta l = 2 \times (l_{19} - l_0) = 1,352.16$, d.f. $= 18$, $P < 10^{-10}$], this implies significant heterogeneity in the rates of protein evolution among the different families. The d_N/d_S values ranged from 0.1217 (family Dasyuridae) to 0.2397 (family Vombatidae). The d_N/d_S values are detailed in figure 1 and supplementary table S2, Supplementary Material online.

We then used the FR model, which assumes an independent d_N/d_S for each branch in the phylogenetic tree, obtaining a log-likelihood value of $l_{FR} = -2,039,488.51$. The FR model fitted the data significantly better than the M0 model [$2\Delta l = 2 \times (l_{FR} - l_0) = 1,954.16$, d.f. $= 82$, $P < 10^{-10}$] and the 19ω model [$2\Delta l = 2 \times (l_{FR} - l_{19}) = 602$, d.f. $= 64$, $P < 10^{-10}$], implying significant heterogeneity in the rates of protein evolution across the different branches of the phylogeny (even among branches in the same family). Across the studied species, the lowest d_N/d_S is 0.106 (*Pascogale topoatafa*), while the highest d_N/d_S is 0.302 (*Pseudocheirus occidentalis*) (fig. 2, supplementary figs. S1 and S2, Supplementary Material online).

Next, we applied the FR model to each of the 1,472 exon alignments separately. Thus, we obtained a total of 778 (number of exons after removing those with missing data) × 83 (branches) = 64,574 d_N/d_S values. For each pair of branches (a total of 3,403 comparisons), we compared their d_N/d_S values using a paired Wilcoxon test. We obtained
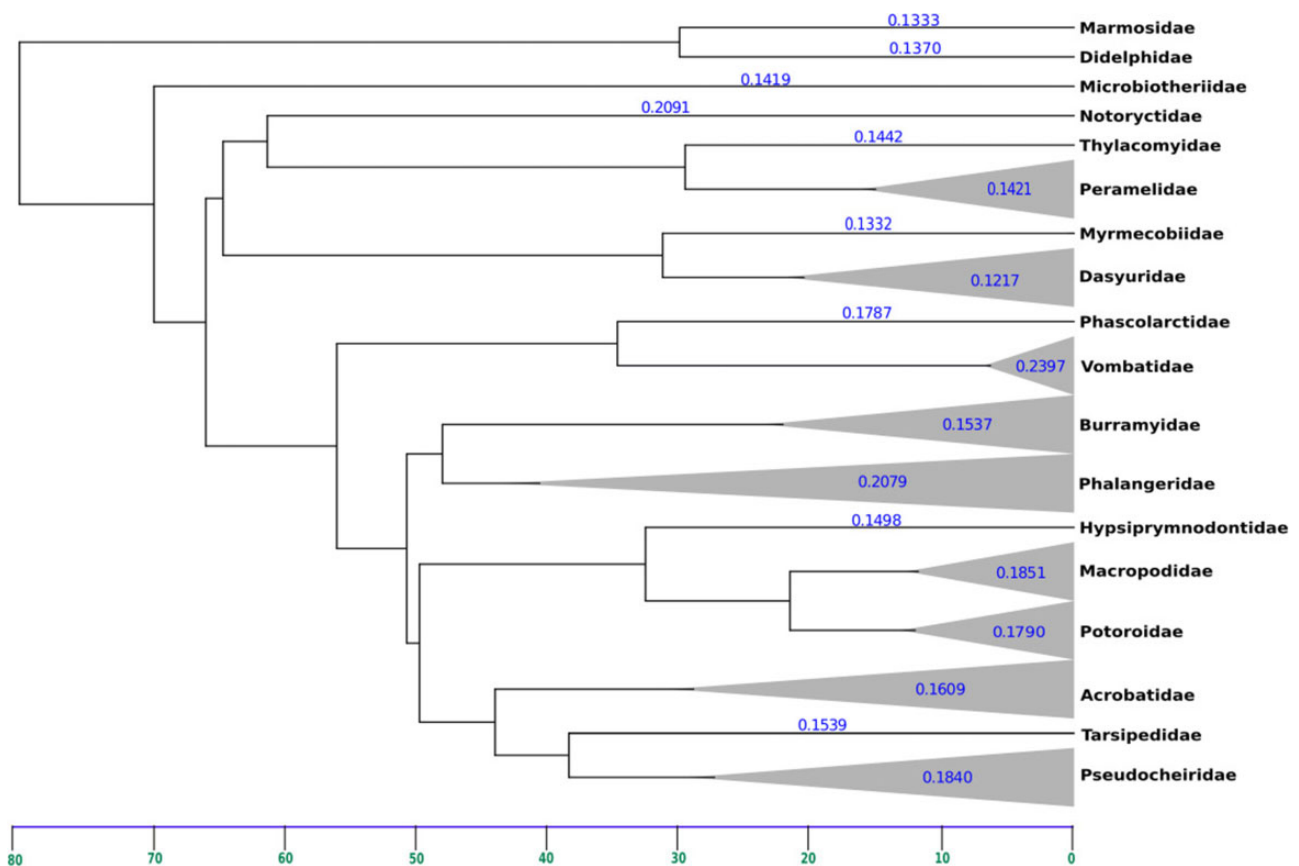
Fig. 1.—Estimated d_N/d_S values for the 18 marsupial families included in our study. Numbers in blue represent the d_N/d_S for each family estimated using the 19ω model. The blue horizontal line represents time in millions of years.

significant differences ($P < 0.05$) in 81% of the comparisons (fig. 3 and supplementary fig. S3, Supplementary Material online). The test detected significant differences for most pairs of external branches, except for those involving species from the same genus, such as *Bettongia penicillata* and *Bettongia lesueur*. However, we noted that non-significant differences ($P > 0.05$) could happen between distantly related species, such as *Didelphis virginiana* and *Monodelphis domestica* (fig. 3).

Life-History Traits and Their Association with d_N/d_S in Marsupials

For each of the 43 species included in our analysis, we gleaned information on distribution area and seven life-history trait variables (body mass, body length, sexual maturity age, weaning age, gestation period, litter size, and the number of litters per year; supplementary table S1, Supplementary Material online). Then we evaluated their correlations with d_N/d_S using Spearman's rank correlation coefficients. We detected a significant correlation ($P < 0.05$) for body mass, body length, weaning, and litter size (supplementary fig. S4, Supplementary Material online). Except for litter size, all the correlations were positive.

Then we used a phylogenetic generalized least squares regression (pgls) to avoid the effects of the shared phylogenetic history among the studied species. We found a significant correlation between d_N/d_S and three variables: body mass ($R^2 = 0.37$, $P = 0.008$), body length ($R^2 = 0.37$, $P = 0.009$), and weaning age ($R^2 = 0.35$, $P = 0.02$) (fig. 4, supplementary table S3, Supplementary Material online). Thus, the association between d_N/d_S and litter size was not significant once we controlled for phylogenetic inertia.

Discussion

Our analyses show that rates of protein evolution, as measured from the d_N/d_S ratios, widely vary among the different marsupial families (fig. 1). Families with the lowest d_N/d_S ratios are expected to be the ones with largest N_e. Our results revealed that the Vombatidae family is the one with the highest d_N/d_S (0.2397), meaning that its species have accumulated more nonsynonymous divergence per unit of synonymous divergence compared with the other 17 families analyzed. *Vombatus ursinus* and *Lasiorinhus latifrons* (both represented in the tree) are species with a large body size and mass (26,650 g and 876.5 mm on average). Furthermore, both species present a very low reproduction rate (0.75 litters per year),
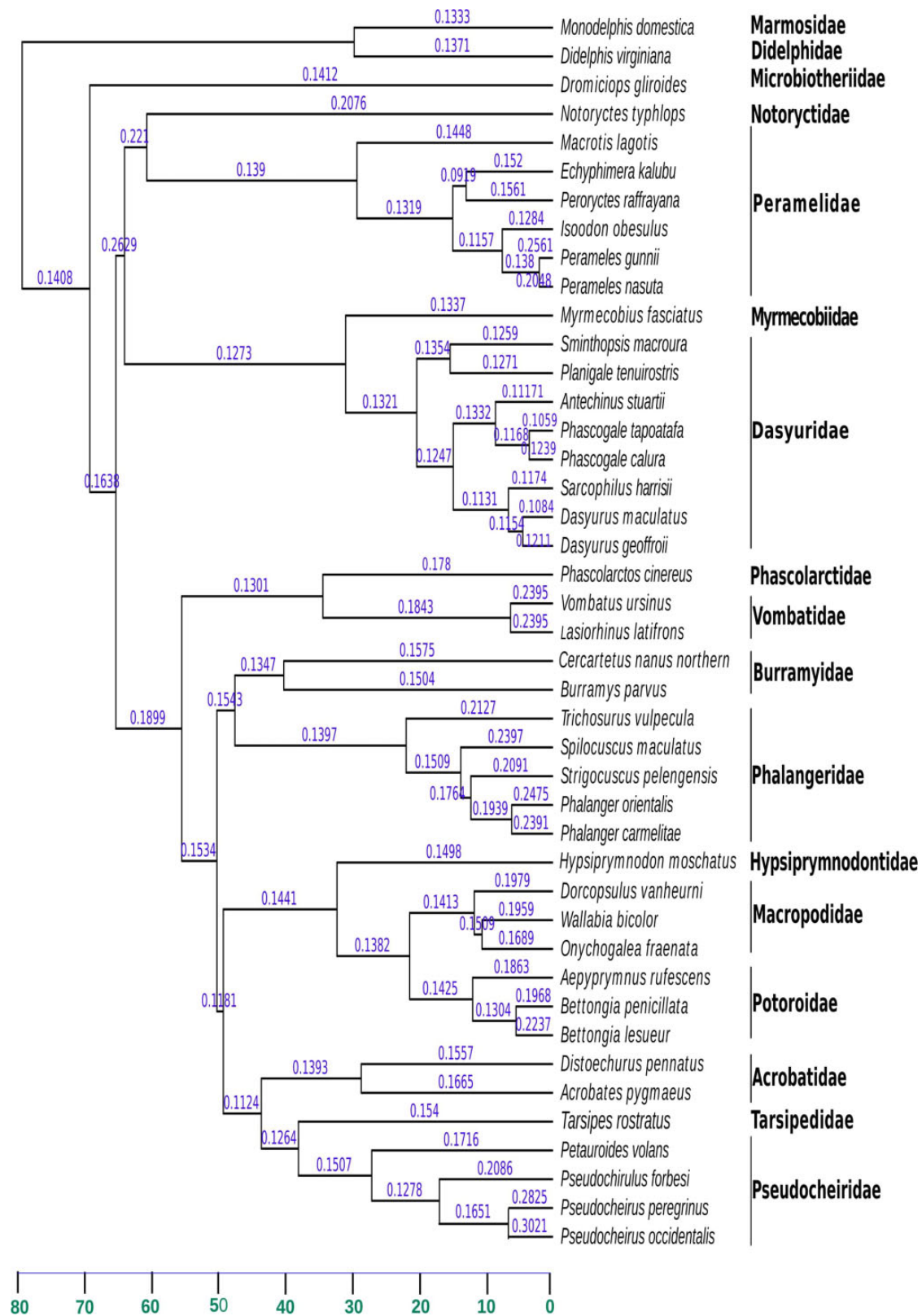
Fig. 2.—Estimated d_N/d_S at each branch of the marsupial phylogenetic tree. The numbers in blue on each branch represent the estimated d_N/d_S (ω) values. The blue horizontal line represents time in millions of years.
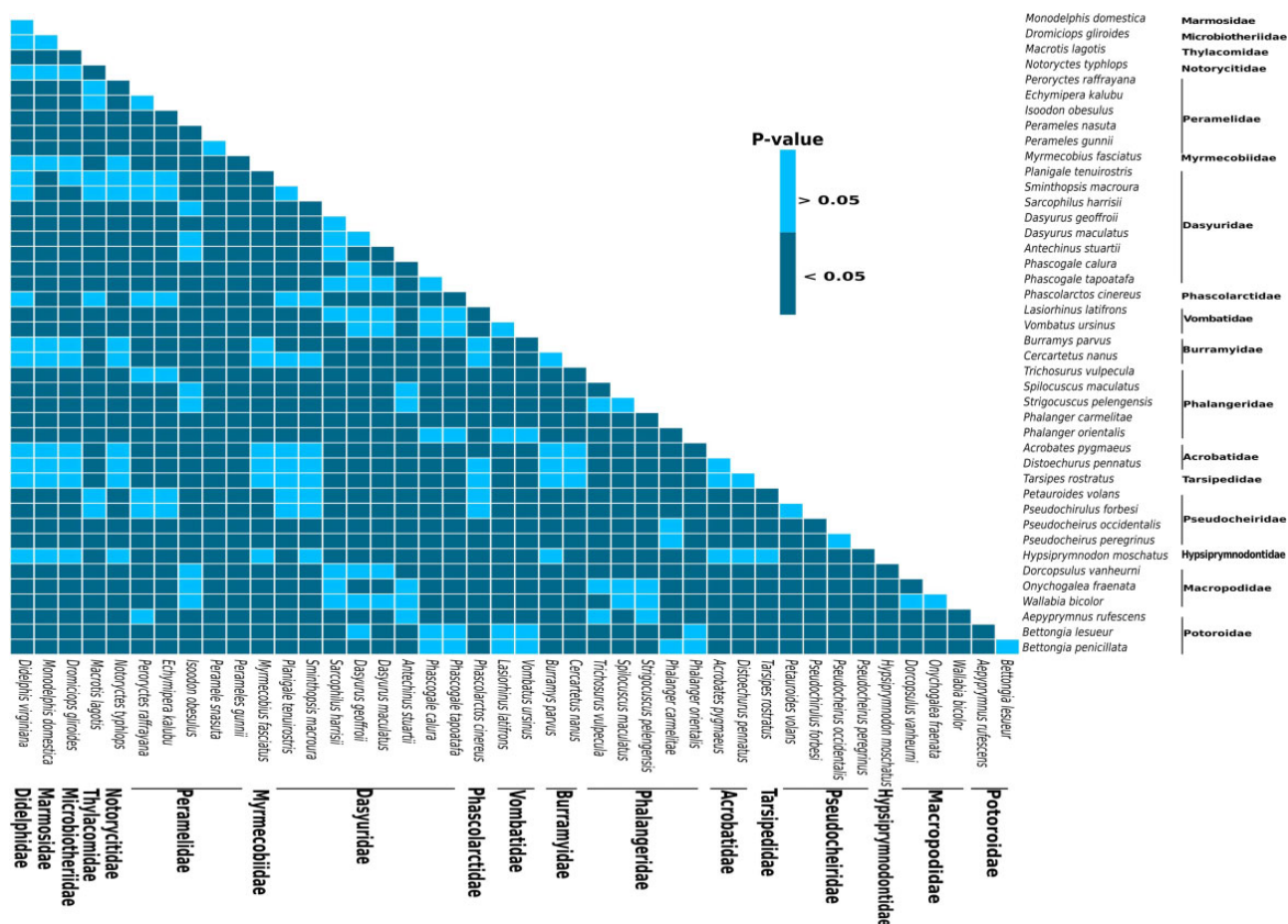
FIG. 3.—Heatmap representing the paired Wilcoxon tests using the d_N/d_S values for 778 exon sequences from 43 marsupial species. Each cell represents a P-value calculated under the alternative hypothesis of difference between each pair of species. The dark color for most of the cells represents P-values lower than 0.05, indicating significant differences between species. The light color represents P-values greater than 0.05, showing no significant differences between species. An expanded version of this figure (including internal branches) is presented in supplementary figure S3, Supplementary Material online.

with weaning periods ranging from 225 to 545 days. Contrarily, the Dasyuridae family, which has the lowest d_N/d_S (0.1217), is represented by species with a much smaller body mass (ranging from 7 to 8,000 g) and size (63–763 mm), and a much higher reproductive rate (3–12 litters per year and weaning periods between 95 and 22 days). The contrast between these families with the highest and lowest d_N/d_S ratios in the phylogenetic tree suggests that the above-mentioned life-history traits can be used as a proxies of N_e in marsupials, similar to previous studies in other species.

Our analyses based on the free-ratios model also revealed significant differences in the rates of protein evolution of most of the species studied. Species from the genus *Pseudocheirus* exhibited the highest d_N/d_S values, whereas species from the genus *Phascogale* presented the lowest d_N/d_S. In our study, the critically endangered species *Pseudocheirus occidentalis* has the highest d_N/d_S ratio (0.3021). Currently, *P. occidentalis* is confined to a small area in southwestern Australia (Bader et al. 2019). Our

estimations suggest a potential distribution area of 18,010 km². This species is characterized by a low reproduction rate (1 litter per year) and a considerable long weaning time (210 days). In contrast, the species with the lowest d_N/d_S, (0.1059), *Phascogale tapoatafa*, considered in the near-threatened category according to the UICN criteria, has a much larger potential distribution area (1,068,122 km²), a higher reproduction rate (6 litters per year), and a lower weaning age (122 days). The body mass and length are the most used N_e proxies. Comparing the species with the lowest and highest d_N/d_S values, *Pascogale tapoatafa* has a considerably lower body mass and length (211 g and 195 mm) than *Pseudocheirus occidentalis* (700 g and 360 mm), supporting the inverse relationship between these two variables and N_e.

Even though our correlation and pgls regression plots provide insights into the relationship between d_N/d_S and several life-history traits frequently used as proxies of N_e, correlation tests are significant for only four factors: body mass, body length, litter size, and weaning age, with low-d_N/d_S species
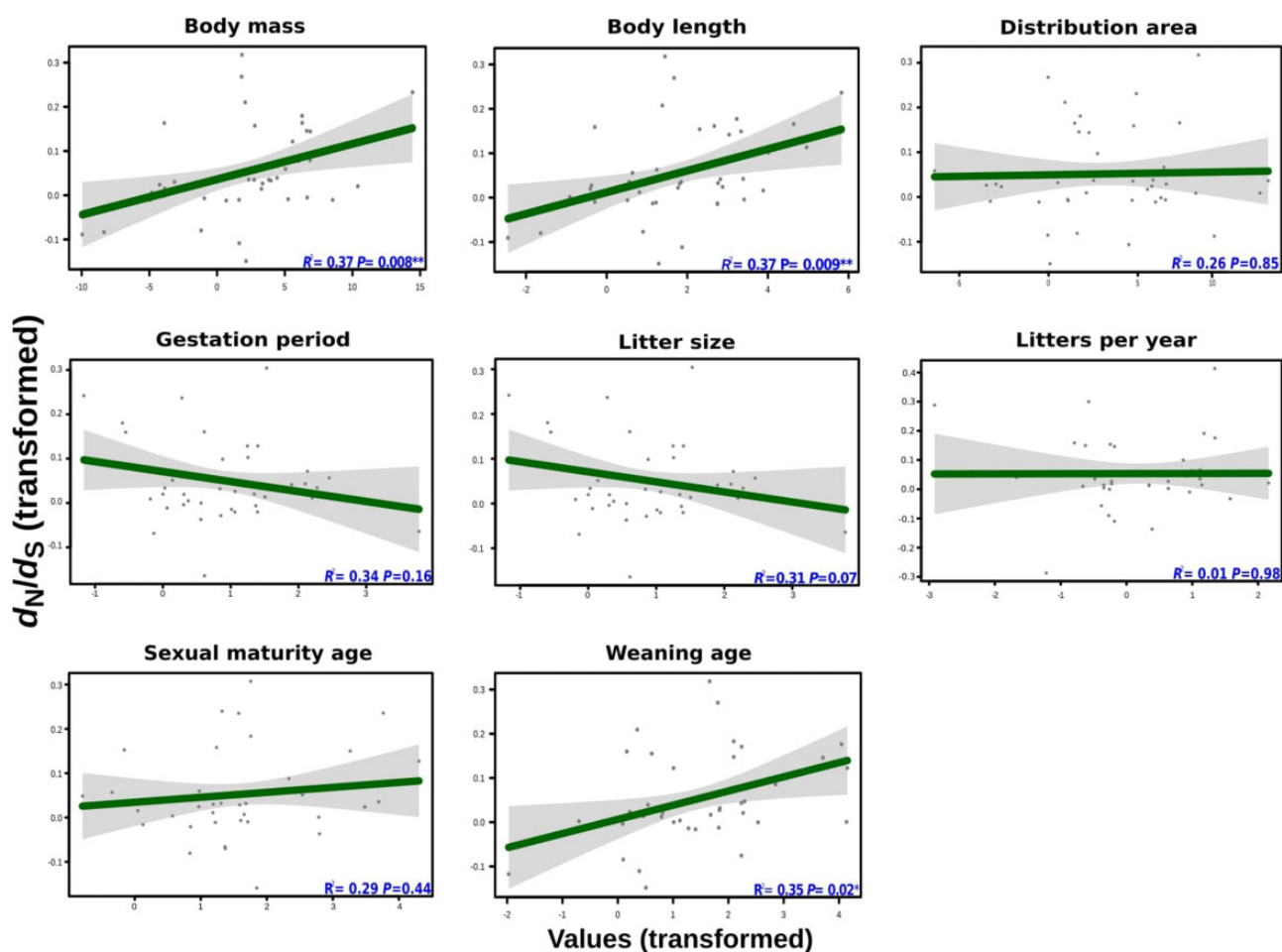
Fig. 4.—Phylogenetic generalized least squares regressions between d_N/d_S and distribution area and seven life-history traits in marsupials. Each panel displays the regression lines, R^2 coefficient, and P-values. Shaded gray areas represent 95% confidence intervals. *$P < 0.05$, **$P < 0.01$, ***$P < 0.001$.

tending to exhibit low body mass, low body length, high litter size, and low weaning age (supplementary fig. S4, Supplementary Material online). These results differ from previous findings in placental mammals, where sexual maturity age and generation time were more correlated with d_N/d_S, indicating that they are better predictors of N_e than body mass (Nikolaev et al. 2007; Lartillot 2013; Romiguier et al. 2013). In addition, our pgls regression analyses indicate that only body mass, body length, and weaning age have an impact on d_N/d_S once phylogenetic effects are accounted for (fig. 4).

Fisher et al. (2003) studied the influence of four life-history traits (body size, reproductive rate, habitat specialization, and diet) and former distribution area on marsupial susceptibility to decline and extinction. They found that even though extrinsic factors such as geographical range overlap with domestic animals played a major role, intrinsic factors such as body size and other life-history traits could also contribute to the extinction of the Australian marsupials. In this context, our results also imply that besides well-known proxies of N_e such as body mass and size, other variables like weaning

age can be used as proxies of N_e in marsupials. These variables are commonly considered as intrinsic factors in the decline and extinction in marsupials.

In summary, using exome data from 43 marsupial species, we report considerable variation in rates of protein evolution among families and species, indicative of important differences in N_e. We found that this variation is associated with a number of life-history traits. Some of these traits have been described before as intrinsic extinction factors in marsupials and other vertebrates.

Materials and Methods

Molecular Data and Phylogeny

We performed our analyses using a data set of 1,550 aligned exons and the phylogenetic tree obtained by Duchêne et al. (2018). The length of the alignments ranged from 141 to 3,660 bp (47–1,220 codons). Each alignment contained sequences for 45 species of marsupials representing 18 of the 22 extant families. To generate each alignment, the

authors identified the orthologs of their targeted exons using as reference the genome of *Monodelphis domestica* (Mikkelsen et al. 2007), using the condition of a single BlastN hit with a bit score >380. The phylogenetic tree was estimated with RAxML v8.1.1 (Stamatakis 2014), using a concatenated alignment of 867,000 bp, including all the sequenced loci. For the phylogenetic tree calculation, they used 12 fossil-based age constraints on internal nodes in the tree for calibration (see further details in Duchêne et al. 2018). All the alignments and the phylogenetic tree generated by Duchêne et al. (2018) are available online (github.com/duchene/marsupial_family_phylogenomics; last accessed January 21, 2021).

Life-History Traits and Distribution Area Data

For each of the marsupial species included in our analysis, we collected data of distribution area and seven life-history traits from the PantTHERIA database (Jones et al. 2009) and the Animal Diversity Web (available at http://animaldiversity.ummz.umich.edu; last accessed May 20, 2021). The life-history traits used in this study include body mass, body length, litter size, litters per year, sexual maturity age, weaning age, and gestation time. We estimated the distribution area for each species with the QGIS program (QGIS Development Team 2016), using the data retrieved from The UICN Red List of Threatened Species (IUCN 2020). The retrieved life-history traits information and estimated distribution areas are shown in supplementary table S1, Supplementary Material online.

Estimation of Rates of Protein Evolution

Out of the 45 species in the original data set, we discarded two species (*Potorous tridactylus* and *Pseudochirops corinnae*) because they exhibited more than 10% missing data. We also removed 78 loci because their sequence files contained more than 30% missing or ambiguous data. We then created a concatenated alignment comprising 1,472 exons (808,941 bp; 269,947 codons) for 43 marsupial species. We removed the branches for the two discarded species from the phylogenetic tree using the R-project package ape v5.4 (Paradis et al. 2004).

Using the concatenated alignment, we computed rates of evolution using model 0 (M0), the free-ratios model (FR), and a 19-ratios model (which we called 19ω), as implemented in the CODEML module of the PAML package, version 4.8d (Yang 1997). Each model computes a number of nonsynonymous to synonymous divergence ratios ($\omega = d_N/d_S$). The M0 model assumes a constant d_N/d_S for all branches in the tree, while the FR model assumes an independent d_N/d_S for each branch. The other model, which is intermediate between M0 and FR, estimates a separate d_N/d_S for any user-defined set of branches. In our case, we used the model to calculate a separate d_N/d_S for each of the 18 marsupial families, plus an additional ω for internal branches not belonging to any of these families. Significant differences between pairs of nested

models (M0 vs FR, M0 vs 19ω, and 19ω vs FR) imply heterogeneity in the d_N/d_S values among the different species and families represented in the tree. To compare the fit of the models, we compared the results of the likelihood ratio test with a χ^2 distribution (using as degrees of freedom the difference between the number of parameters of each pair of nested models).

In addition to the analysis of the concatenated alignment, we used the FR model to estimate the d_N/d_S for all branches for each exon separately. We tested for significant differences for each pair of branches using a paired Wilcoxon test (as implemented in the R-project v3.5.2 software; R Core Team 2018); for each pair of branches, the d_N/d_S values of all pairs of orthologous exons were compared.

The Relationship between Rates of Protein Evolution and Morpho-Ecological Variables

We evaluated the correlation between each of the morpho-ecological factors (life-history traits and distribution area) and the d_N/d_S of each species using Spearman's rank correlation. However, simple correlations do not take into account the effect of phylogenetic inertia. Consequently, we complemented our analyses by performing a phylogenetic generalized least squares regression (pgls). To visualize de pgls regression and correct for nonindependence caused by the phylogeny, we performed a transformation of the d_N/d_S and life-history variables, assuming a normal distribution and equal variance (as in Rolland et al. 2020). For the pgls regression analysis, we used the method and the R-project code implemented by Rolland et al. (2020).

Supplementary Material

Supplementary data are available at *Genome Biology and Evolution* online.

Acknowledgments

Data Availability

All data used in this work are publicly available, as described in the Materials and Methods.

Literature Cited

Bader JC, Van Helden BE, Close PG, Speldewinde PC, Comer SJ. 2019. Sheoak woodlands: a newly identified habitat for western ringtail possums. J Wild Mgmt. 83(5):1254–1260.

Botero-Castro F, Figuet E, Tilak MK, Nabholz B, Galtier N. 2017. Avian genomes revisited: hidden genes uncovered and the rates versus traits paradox in birds. Mol Biol Evol. 34(12):3123–3131.

Cardillo M, et al. 2005. Multiple causes of high extinction risk in large mammal species. Science 309(5738):1239–1241.

Charlesworth B. 2009. Fundamental concepts in genetics: effective population size and patterns of molecular evolution and variation. Nat Rev Genet. 10(3):195–205.

Collen B, et al. 2011. Predicting how populations decline to extinction. Philos Trans R Soc Lond B Biol Sci. 366(1577):2577–2586.

Crow JF, Morton NE. 1955. Measurement of gene frequency drift in small populations. Evolution 9(2):202–214.

Damuth J. 1987. Interspecific allometry of population density in mammals and other animals: the independence of body mass and population energy-use. Biol J Linn Soc. 31(3):193–246.

Damuth J. 1981. Population density and body size in mammals. Nature 290(5808):699–700.

Duchêne DA, et al. 2018. Analysis of phylogenomic tree space resolves relationships among marsupial families. Syst Biol. 67(3):400–412.

Figuet E, et al. 2016. Life history traits, protein evolution, and the nearly neutral theory in amniotes. Mol Biol Evol. 33(6):1517–1527.

Fisher DO, Blomberg SP, Owens IP. 2003. Extrinsic versus intrinsic factors in the decline and extinction of Australian marsupials. Proc Biol Sci. 270(1526):1801–1808.

Frankham R. 1995. Effective population size/adult population size ratios in wildlife: a review. Genet Res. 66(2):95–107.

IUCN. 2020. The IUCN Red List of Threatened Species [Internet]. Available from: http://www.iucnredlist.org. Accessed December 23, 2020.

James J, Eyre-Walker A. 2020. Mitochondrial DNA sequence diversity in mammals: a correlation between the effective and census population sizes. Genome Biol Evol. 12(12):2441–2449.

Jones KE, et al. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant and recently extinct mammals: ecological Archives E090-184. Ecology 90(9):2648.

Kimura M. 1983. Rare variant alleles in the light of the neutral theory. Mol Biol Evol. 1(1):84–93.

Lartillot N. 2013. Interaction between selection and biased gene conversion in mammalian protein-coding sequence evolution revealed by a phylogenetic covariance analysis. Mol Biol Evol. 30(2):356–368.

Lartillot N, Poujol R. 2011. A phylogenetic model for investigating correlated evolution of substitution rates and continuous phenotypic characters. Mol Biol Evol. 28(1):729–744.

Luo ZX, Yuan CX, Meng QJ, Ji Q. 2011. A Jurassic eutherian mammal and divergence of marsupials and placentals. Nature 476(7361):442–445.

Mikkelsen TS, et al.; Broad Institute Whole Genome Assembly Team. 2007. Genome of the marsupial *Monodelphis domestica* reveals innovation in non-coding sequences. Nature 447(7141):167–177.

Nabholz B, Glémin S, Galtier N. 2008. Strong variations of mitochondrial mutation rate across mammals—the longevity hypothesis. Mol Biol Evol. 25(1):120–130.

Nadachowska-Brzyska K, Li C, Smeds L, Zhang G, Ellegren H. 2015. Temporal dynamics of avian populations during Pleistocene revealed by whole-genome sequences. Curr Biol. 25(10):1375–1380.

Nikolaev SI, et al.; Program NIoHISCCS. 2007. Life-history traits drive the evolutionary rates of mammalian coding and noncoding genomic elements. Proc Natl Acad Sci USA. 104(51):20443–20448.

Ohta T. 1993. An examination of the generation-time effect on molecular evolution. Proc Natl Acad Sci USA. 90(22):10676–10680.

Ohta T. 1976. Role of very slightly deleterious mutations in molecular evolution and polymorphism. Theor Popul Biol. 10(3):254–275.

Ohta T, Ina Y. 1995. Variation in synonymous substitution rates among mammalian genes and the correlation between synonymous and non-synonymous divergences. J Mol Evol. 41(6):717–720.

Paradis E, Claude J, Strimmer K. 2004. APE: analyses of phylogenetics and evolution in R language. Bioinformatics 20(2):289–290.

Peters RH, Peters RH. 1986. The ecological implications of body size. Cambridge: Cambridge University Press.

Popadin K, Polishchuk LV, Mamirova L, Knorre D, Gunbin K. 2007. Accumulation of slightly deleterious mutations in mitochondrial protein-coding genes of large versus small mammals. Proc Natl Acad Sci USA. 104(33):13390–13395.

Purvis A, Gittleman JL, Cowlishaw G, Mace GM. 2000. Predicting extinction risk in declining species. Proc Biol Sci. 267(1456):1947–1952.

QGIS Development Team. 2016. QGIS geographic information system. Beaverton, Oregon: Open Source Geospatial Foundation Project. Available from: http://qgis.osgeo.org. Accessed December 15, 2020.

R Core Team. 2018. R: a language and environment for statistical computing. Vienna (Austria): R Foundation for Statistical Computing. Available from: https://www.R-project.org. Accessed May 10, 2021.

Rolland J, Schluter D, Romiguier J. 2020. Vulnerability to fishing and life history traits correlate with the load of deleterious mutations in teleosts. Mol Biol Evol. 37(8):2192–2196.

Romiguier J, et al. 2012. Fast and robust characterization of time-heterogeneous sequence evolutionary processes using substitution mapping. PLoS One. 7(3):e33852.

Romiguier J, et al. 2014. Comparative population genomics in animals uncovers the determinants of genetic diversity. Nature 515(7526):261–263.

Romiguier J, Ranwez V, Douzery EJ, Galtier N. 2013. Genomic evidence for large, long-lived ancestors to placental mammals. Mol Biol Evol. 30(1):5–13.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30(9):1312–1313.

Weinreich DM. 2001. The rates of molecular evolution in rodent and primate mitochondrial DNA. J Mol Evol. 52(1):40–50.

Welch JJ, Bininda-Emonds OR, Bromham L. 2008. Correlates of substitution rate variation in mammalian protein-coding sequences. BMC Evol Biol. 8:53.

Wright S. 1931. Evolution in Mendelian Populations. Genetics 16(2):97–159.

Wu CI, Li WH. 1985. Evidence for higher rates of nucleotide substitution in rodents than in man. Proc Natl Acad Sci USA. 82(6):1741–1745.

Yang Z. 1997. PAML: a program package for phylogenetic analysis by maximum likelihood. Comput Appl Biosci. 13(5):555–556.

Associate editor: Jay Storz