Estimating body size in early primates: The case of *Archicebus* and *Teilhardina*

Marian Dagosto [a,*], Daniel Gebo [b], Xijun Ni [c], Thierry Smith [d]

[a] *Department of Cell and Molecular Biology, Northwestern University, 303 E. Chicago Ave., Chicago, IL 60611, USA*
[b] *Department of Anthropology, Northern Illinois University, Dekalb, IL 60615, USA*
[c] *Key Laboratory of Vertebrate Evolution and Human Origin, Institute of Vertebrate Paleontology and Paleoanthropology, Chinese Academy of Sciences, China*
[d] *Directorate Earth and History of Life, Royal Belgian Institute of Natural Sciences, 29 rue Vautier, 1000 Brussels, Belgium*

ARTICLE INFO

ABSTRACT

Obtaining accurate estimations of the body mass of fossil primates has always been a subject of interest in paleoanthropology because mass is an important determinant for so many other aspects of biology, ecology, and life history. This paper focuses on the issues involved in attempting to reconstruct the mass of two early Eocene haplorhine primates, *Teilhardina* and *Archicebus*, which pose particular problems due to their small size and temporal and phylogenetic distance from extant primates. In addition to a ranking of variables from more to less useful, the effect of using models of varying taxonomic and size compositions is examined. Phylogenetic correction is also applied to the primate database. Our results indicate that the choice of variable is more critical than the choice of model. The more reliable variables are the mediolateral breadth across the femoral condyles and the area of the calcaneocuboid facet of the calcaneus. These variables suggest a body mass of 39 g (range 33–46 g) for *Archicebus* and 48 g (range 44 –56 g) for *Teilhardina*. The width of the distal femur is found to be the most consistent estimator across models of various composition and techniques. The effect of phylogenetic correction is small but the choice of branch length assumption affects point estimates for the fossils. The majority of variables and models predict the body mass of *Archicebus* and *Teilhardina* to be in the range of the smaller extant mouse lemurs, as expected.

1. Introduction

Archicebus achilles, represented by a single nearly complete skull and skeleton from the early Eocene Yangxi Formation of Hubei, China (Ni et al., 2013), and *Teilhardina belgica*, represented by numerous unassociated dental and skeletal elements representing many parts of the skeleton from the earliest Eocene locality of Dormaal, Belgium (Teilhard de Chardin, 1927; Szalay, 1976; Smith et al., 2006; Gebo et al., 2012, 2015), are among the best known early Eocene primates. Knowledge of their paleobiology is, therefore, particularly vital to our understanding of early primate evolution. Since body mass is often a key to evaluating other important aspects of paleobiology, determination of the likely body mass of these early primate taxa is a significant task for paleoprimatologists. Yet these

particular taxa pose several interesting problems for body mass estimation. First, they are at the far small end of the size distribution of living primates. Secondly, they are not nested within any family of living primates, but lie at or near the base of the tarsiiform radiation. Thirdly, their nearest living relatives (tarsiers) are highly derived dentally, cranially, and postcranially. All of these raise the question of determining an appropriate reference group from which to make the body mass inference.

The body mass of both taxa has previously been compared to that of the small to medium sized extant mouse lemurs (*Microcebus*; species mean body mass of 30–90 g; Rasoloarison et al., 2000; Louis et al., 2006a). The absolute dimensions of most postcranial elements of both taxa fall near those of the mouse lemur species weighing less than 50 g (Table 1). Using conventional regression techniques, the body mass of *Archicebus* has been estimated at about 30 g (95% CI 25–37 g) from its skull length using a primate-wide regression and 32 g (27–39 g) from its body length using a euarchontoglires-wide regression (Ni et al., 2013). A comparison of upper molar size to

---

Table 1
Comparison of dimensions of *Archicebus* and *Teilhardina* with those of *Microcebus* species and small tupaiids.[a]

Species	MASS	FEMBEW	HUMBEW	TROCHAREA	AWID	CWID	CUFACET
Archicebus achilles		3.80	4.48			2.61	2.18
Teilhardina belgica		4.20	4.54	3.49	2.42	2.42	2.52
		$N = 1$	$N = 2$	$N = 3$	$N = 4$	$N = 11$	$N = 7$
Microcebus berthae	30.6	3.79	4.00			2.37	
Microcebus rufus	43.0	3.99	4.43	4.37	2.53	2.37	2.61
Microcebus sambiranensis	44.1	4.10	4.90				
Microcebus myoxinus	47.0	4.50	4.90				
Microcebus griseorufus (1) Amboasary	49.9	4.48	5.06	4.65	2.74	2.55	2.93
Microcebus griseorufus (2) Beza Mahafaly	53.2	4.40	5.00				
Microcebus tavaratra	54.3	4.60	5.20				
Microcebus murinus	61.3	4.60	5.30				
Microcebus ravelobensis	60.0	4.70	5.30				
Dendrogale murina	44.0	4.21	3.63	3.54	1.80	3.22	2.77
Ptilocercus lowii	41.0	4.16		4.60	1.77	3.30	1.80

[a] Body mass (g) from various sources (see SOM Table 1). Data for *Microcebus* HUMBEW and FEMBEW (see Table 2 for definitions) from species other than *M. griseorufus* (1) and *M. rufus* is from Rasoloarison et al. (2000) and tarsal measures are from our own data. Mass is given in g, all linear measurements (FEMBEW, HUMBEW, AWID, CWID) in mm, and the area measures (TROCHAREA, CUFACET) in mm^2. *Archicebus* measurements are from Ni et al. (2013), supplementary information. $N = 1$ for all measures of *Archicebus* with the exception of CUFACET, which is the average of the left and right cuboid width multiplied by the height of the left. For *Teilhardina*, the specimens used and a drawing of the measurements taken can be found in Gebo et al. (2012, 2015).

those of *Teilhardina asiatica* (Ni et al., 2004), however, suggests a lower mass of about 20 g (Ni et al., 2013), while Boyer et al. (2013) estimated a higher mass of 62 g based on calcaneocuboid joint area. The body mass of *T. belgica* has been estimated at 32–58 g from various tarsal measures (Dagosto and Terranova, 1992; Boyer et al., 2013) and 30–135 g based on dental measurements (Gingerich, 1981; Conroy, 1987; Dagosto and Terranova, 1992). Here, we predict body mass of these two taxa based on additional postcranial measurements and compare the results from models with differing sample composition and techniques designed to address some of the consequences of choices in reference model, variables, and phylogenetic correction.

2. Materials and methods

Body mass was regressed against selected postcranial measurements (Table 2) using ordinary least squares regression (OLS). The variables were natural log-transformed prior to the regression. Analyses were run in R (R Development Core Team, 2015). The measurements were chosen based on past success (Conroy, 1987; Dagosto and Terranova, 1992; Ruff, 2003; Squyres and Ruff, 2015; Yapuncich et al., 2015), availability in the taxa in question, and a desire to compare results from different parts of the postcranium. Descriptions and illustrations of the measurements can be found in the references cited in Table 2. The measurements of the extant taxa were taken with calipers by MD, with the exception of the measures FEMBEW and HUMBEW (described in Table 2) for some

Microcebus species that were taken from Rasoloarison et al. (2000). The majority of measurements for *Teilhardina* were made with calipers by MD but a few additional measurements of *Teilhardina* specimens were provided by DLG and TS (Gebo et al., 2012, 2015). It is not possible to take many caliper measurements for *Archicebus*; measurements were taken from the CT reconstructions by XN and calibrated to calcaneal length that could be measured directly on the specimen (Ni et al., 2013). Potential error due to different observers and measurement techniques was not examined in this study. *Teilhardina* is known from multiple specimens; where $N > 1$ the mean is used to make predictions (Table 1). *Archicebus* is represented by a single individual; we use these values as an approximation of the species mean acknowledging all the uncertainty that implies (Smith, 2002).

The value of the predictor variables was judged by R^2, %SEE, and %PE (Smith, 1980, 1984; Van Valkenburgh, 1990; Yapuncich et al., 2015). In addition to the overall %PE, the %PE for taxa similar to the targets in size and/or phylogeny (mouse lemurs, tarsiers) was also calculated (Table 3). Several different analyses were run to examine the effect of taxonomic and body mass composition on the reference sample. We also ran several phylogenetically controlled analyses. The reference group models are as follows:

2.1. Model 1

Euarchonta—98 taxa of strepsirhine, tarsiiform, and platyrrhine primates plus tupaiids and dermopterans. Body masses of living

Table 2
List of measurements used in the study.[a]

Abbreviation	Description	Same measurement	Similar measurement
FEMBEW	Maximum mediolateral breadth across the femoral condyles in mm	Gebo et al. (2012)	FCML in Ruff (2002) and Squyres and Ruff (2015)
HUMBEW	Maximum mediolateral breath across the distal humerus in mm	Measurement BW in Szalay and Dagosto (1980)	
AWID	Maximum width of the talus in mm	A2 in Dagosto and Terranova (1992)	
CWID	Maximum width of the calcaneus in mm	C2 in Dagosto and Terranova (1992)	
CUFACET	Area of the calcaneocuboid facet of the calcaneus; facet height (C5)*facet width (C6); C5 and C6 are defined in Dagosto and Terranova (1992); in mm^2	Index 6 in Dagosto and Terranova (1992)	CCFA in Yapuncich et al. (2015)
TROCHAREA	Area of the lateral talotrochlear facet of the talus; facet length (A3)*facet width (A4); A3 and A4 are defined in Dagosto and Terranova (1992); in mm^2	Index 1 in Dagosto and Terranova (1992)	LTFA in Yapuncich et al. (2015)

[a] Abbreviations of the same or similar measurements from previous studies are given in the third and fourth columns. Illustrations and definitions of the measurements can be found in the cited references. SOM Table 1 gives the sample sizes for the postcranial measurements for the extant taxa.

Table 3
Results of OLS regressions.[a]

Variable	Model	N	R²	Slope	Intercept	SEE	%SEE	%PE	CF	Mouse lemurs within 20%	Mouse lemurs average %PE	Tarsiers within 20%	Tarsiers average %PE
FEMBEW	1. Euarchontan	80	0.986	2.652	0.013	0.175	19.1	12.87	1.015	9/9	7.0	2/2	6.0
	2. Primates	70	0.990	2.677	−0.076	0.143	15.4	10.89	1.010	9/9	7.1	2/2	6.4
	3. Small euarchontan	22	0.951	2.602	0.335	0.129	13.8	10.00	1.008	9/9	6.8	2/2	6.2
	4. Small mammals	54	0.908	2.276	0.603	0.200	22.2	15.54	1.019	8/9	7.6	2/2	6.2
	5a. PIC primate Grafen BL	69	0.925	2.500	0.372	0.171	18.6	14.06	1.014	8/9	16.0	2/2	12.9
	5b. PIC primate divergence time BL	69	0.922	2.552	0.238	0.156	16.9	12.60	1.012	7/9	11.0	2/2	9.4
	5c. Rerooted	69	0.922	2.552	0.232	0.156	16.9	12.60	1.012	8/9	10.4	2/2	8.9
HUMBEW	1. Euarchontan	79	0.964	2.649	−0.203	0.281	32.4	21.44	1.039	7/9	16.1	0	44.2
	2. Primates	70	0.975	2.795	−0.062	0.234	26.4	16.33	1.027	8/9	6.0	0	36.7
	3. Small euarchontan	21	0.688	1.847	1.138	0.326	38.5	27.55	1.049	4/9	21.1	1/2	21.4
	4. Small mammals	55	0.673	2.023	0.436	0.376	45.6	28.58	1.071	4/9	21.8	2/2	8.6
	5a. PIC primate Grafen BL	69	0.873	2.390	0.328	0.316	37.2	27.11	1.048	1/9	26.6	0	44.5
	5b. PIC primate divergence time BL	69	0.871	2.615	−0.293	0.278	32.0	23.55	1.038	9/9	6.6	0	24.4
	5c. Rerooted	69	0.871	2.614	−0.357	0.314	36.9	27.08	1.048	9/9	6.2	0	30.6
AWID	1. Euarchontan	89	0.917	2.341	2.058	0.374	45.4	27.89	1.071	0	42.7	0	50.9
	2. Primates	79	0.938	2.603	1.463	0.299	34.9	22.40	1.045	2/2	17.4	0	37.5
	3. Small euarchontan	15	0.542	1.248	3.110	0.407	50.2	33.58	1.074	0	42.5	1/2	17.0
	5a. PIC primate Grafen BL	78	0.838	2.350	1.943	0.317	37.3	23.70	1.049	0	35.0	0	44.5
	5b. PIC primate divergence time BL	78	0.768	2.430	1.757	0.311	36.5	23.80	1.047	0	27.2	0	40.1
	5c. Rerooted	78	0.768	2.427	1.687	0.332	39.4	27.00	1.054	0	23.4	0	36.1
TROCHAREA	1. Euarchontan	89	0.949	1.399	1.900	0.291	33.8	22.88	1.042	1/2	19.2	0	48.2
	2. Primates	79	0.941	1.432	1.768	0.291	33.8	22.85	1.042	2/2	12.3	0	45.6
	3. Small euarchontan	15	0.727	1.046	2.462	0.314	36.9	28.46	1.044	1/2	21.7	0	28.5
	5a. PIC primate Grafen BL	78	0.748	1.226	2.404	0.340	40.5	27.36	1.056	0	37.5	0	52.3
	5b. PIC primate divergence time BL	78	0.681	1.183	2.493	0.375	45.5	30.89	1.069	0	39.7	0	51.9
	5c. Rerooted	78	0.682	1.183	2.406	0.415	51.4	35.50	1.084	0	35.1	0	48.3
CWID	1. Euarchontan	90	0.871	2.656	1.462	0.477	61.1	36.90	1.112	1/2	16.7	1/2	13.7
	2. Primates	80	0.891	2.503	1.827	0.412	51.0	33.84	1.086	0	32.0	1/2	25.1
	3. Small euarchontan	16	0.655	1.999	1.932	0.377	87.9	27.93	1.069	2/2	11.1	1/2	30.4
	5a. PIC primate Grafen BL	79	0.723	2.170	2.450	0.437	54.8	38.01	1.093	0	51.1	0	38.5
	5b. PIC primate divergence time BL	79	0.685	2.142	2.474	0.442	55.6	38.8	1.095	0	53.5	0	38.5
	5c. Rerooted	79	0.685	2.138	2.458	0.447	56.4	39.66	1.097	0	52.6	0	37.4
CUFACET	1. Euarchontan	77	0.965	1.406	2.519	0.244	27.9	18.75	1.029	2/2	13.3	1/2	14.3
	2. Primates	67	0.965	1.425	2.447	0.221	24.7	16.99	1.024	2/2	8.10	2/2	10.4
	3. Small euarchontan	15	0.861	1.210	2.711	0.224	25.1	16.71	1.022	2/2	12.0	2/2	8.0
	5a. PIC primate Grafen BL	67	0.876	1.277	2.858	0.250	28.4	19.77	1.030	0	29.5	1/2	23.6
	5b. PIC primate divergence time BL	67	0.860	1.298	2.792	0.242	27.4	19.14	1.028	0	26.3	1/2	21.2
	5c. Rerooted	67	0.860	1.298	2.769	0.246	27.9	19.71	1.029	0	24.6	1/2	19.5

[a] SEE = standard error of the estimate; %SEE computes that 68% of the taxa are ± the value; %PE = average prediction error; CF = correction factor. For mouse lemurs and tarsiers, the number of species for which predicted mass was within 20% of actual mass and the average %PE for included taxa is given. The models refer to different sample compositions as described in the Methods section of the text: Model 1 is 98 taxa of euarchontans; Model 2 is 88 taxa of non-catarrhine primates; Model 3 is 22 taxa of euarchontans that weigh less than 200 g; Model 4 is 55 mammals in the size range 10—100 g. Three different PIC models were run, all using the primate database (same as Model 2). Model 5a "Grafen BL" employs branch lengths (BL) following Grafen (1989); Model 5b "Divergence time BL" uses a tree with branch lengths equivalent to divergence times in years estimated from molecular studies (see Fig. 1); Model 5c "Rerooted" follows Garland and Ives (2000) and takes the divergence time BL tree and reroots it as described in the Methods section.

species are derived from numerous sources (e.g., Smith and Jungers, 1997; Taylor and Schwitzer, 2011/12; Butyinski et al., 2013; see Supplementary Online Material [SOM] Table 1 for the full list). We do not have data for every variable for each taxon so the actual N for any regression is <98. The same applies to models 2—5.

2.2. Model 2

Non-catarrhine Primates—88 taxa of strepsirhine, tarsiiform, and platyrrhine primates.

2.3. Model 3

Small euarchontans—22 taxa weighing less than 200 g.

2.4. Model 4

Small mammals—a dataset of 55 tenrecs and shrews from 10 different species in the size range 10—100 g with associated body masses in the collections of the FMNH.

2.5. Model 5

Phylogenetically controlled analysis for the Model 2 database, using phylogenetically independent contrasts (PIC; Felsenstein, 1985) as implemented in the phenotypic diversity analysis program (PDAP) module of Mesquite (Midford et al., 2009; Maddison and Maddison, 2015). We also applied the re-rooting method of Garland and Ives (2000).

2.6. Size control 1

Many papers on body mass estimation in primates are written with the aim of providing an all-purpose estimation and so include data from as many taxa as possible from the smallest primates (*Microcebus*) to the largest (*Gorilla*). In a case, such as this, when the aim is instead to estimate the mass of specific taxa, it would seem that a more biologically relevant context might be provided by restricting the analysis to a range closer to the probable size of the target. Therefore, the results from a broader size range (Models 1 and 2) are compared to a more restricted size range (Model 3).

2.7. Phylogenetic "control" 1

Similarly, it seems reasonable to "control" for phylogeny by including only the most closely related rather than more distantly related taxa since many studies have shown that lower taxonomic levels often give better estimates for members of the taxon than analyses with a broader phylogenetic range (Van Valkenburgh, 1990; Dagosto and Terranova, 1992; Delson et al., 2000; Yapuncich et al., 2015). The estimates from a broad range (Model 1) are compared to those from a more restricted one (Model 2). Given the very basal position of *Archicebus* and *Teilhardina*, in this particular case, the broader phylogenetic bracket that includes the sister taxa of primates might be an equally reasonable approach.

2.8. Size control 2

The most significant challenge in this case is that none of the more phylogenetically relevant samples (Models 1–3, 5) adequately cover the potential size range of the target fossils. In all these analyses, they are at best at the very smallest end and possibly outside of the observed size range, which has the undesirable statistical effect of increased confidence intervals as well as the risk of the model being inappropriate altogether. Model 4 is intended to examine if a more size appropriate, although less taxon appropriate, analysis might be just as useful. Only measurements from the long bones could be examined this way since the tarsals were for the most part not skeletonized.

2.9. Phylogenetic control 2

Taxa cannot be considered truly independent data points since there is a natural hierarchical structure that violates some of the assumptions of ordinary regression (Felsenstein, 1985). The method of PIC provides one way to account for such structure within the dataset. This was implemented using the PDAP module (Midford et al., 2009) within Mesquite (Maddison and Maddison, 2015). Of course the first question to be asked is if such correction is needed (Revell, 2010), and indeed virtually all of the regressions constructed for this paper exhibited phylogenetic signal in the residuals as measured by Pagel's λ or Blomberg's K (Pagel, 1999; Blomberg et al., 2003), as implemented in the R routine phytools (Revell, 2012).

The tree used for the phylogenetic analyses is illustrated in Figure 1. As the tree was constructed from multiple sources, it was difficult to obtain a single standard for branch lengths. Branch lengths were determined from divergence dates derived from molecular studies (referred to as model 5b; the PIC Divergence branch length [BL] Model in Tables 3 and 4). The majority of nodes on the tree use the divergence dates of Perelman et al. (2011), but were supplemented and sometimes supplanted using information from other studies (Fausser et al., 2002; Yoder and Yang, 2004; Andriaholinirina et al., 2006; Andriantompohavana et al., 2007; Finstermeier et al., 2013; Markolf and Kappeler, 2013; Masters et al., 2013; Thiele et al., 2013; Lei et al., 2014; Buckner et al., 2015; Driller et al., 2015; Lynch Alfaro et al., 2015; Pozzi et al., 2015; Ruiz-García et al., 2015). In some cases, dates were estimated from genetic distances scaled to another close node on the tree.

Given that many of these dates have wide confidence intervals and derive from studies using different genes and techniques, they must be considered only an estimate. Therefore, other options for branch length were explored. One commonly used option is to set all branch lengths to one, but as this did not result in a non-significant correlation between the standardized phylogenetically independent contrasts and their standard deviations as required by PIC (Garland et al., 1992), it could not be employed. The transformation suggested by Grafen (1989) was used instead (referred to as Model 5a; the PIC Grafen BL in Tables 3 and 4).

Given a tree with branch lengths PIC can be used to obtain a phylogenetically corrected regression equation. With this "generic" equation, the prediction intervals for a taxon with unknown "y" are calculated as if the phylogenetic position of the target species is unknown and are therefore generally larger than those resulting from the conventional OLS regression (Garland and Ives, 2000). More precise intervals can be obtained if the phylogenetic position of the taxon can be specified and the tree rerooted to put the target and its sister at the basal node. This proved to be difficult to apply in this case since forming a trichotomy among the fossil taxon, the tarsier clade, and the rest of the primates forced platyrrhines and strepsirhines to be a clade, an arrangement that differs from the preferred phylogeny. Since the fossils considered here are very taxonomically and temporally distant from the model taxa, this procedure is essentially the same as placing this trichotomy very conservatively at the base of the primate radiation. Nevertheless, given the exploratory spirit of this study, the results of this rerooting analysis are reported (referred to as Model 5c Rerooted in Tables 3 and 4; this tree uses the divergence dates for branch length).

Although an ideal analysis would have matched values for an individual's morphological measurement with its body mass, such a dataset does not exist for lower primates. In this database, for example, less than five primate specimens have an associated body mass. All analyses except #4 were therefore run with taxon means as is typical in such studies. This seemingly straightforward statement does, however, gloss over some fundamental issues of sample composition that have been raised in many previous critiques (e.g., Jungers, 1990; Martin, 1990; Smith, 2002), but are always worth reiterating:

1. The problem of "species." Due to an explosion of fieldwork and new genetic approaches the past decades have seen a revolution in primate systematics. There is increasing recognition of more taxonomic diversity than was accepted even a few years ago. Researchers may disagree about the taxonomic level at which many of these newly proposed "species" should be recognized (e.g., Tattersall, 2007); nevertheless, many of these taxa have been demonstrated to vary not only geographically and genetically, but also in physical characteristics including body mass. The result is that there are now many more taxonomic units available to be included in an analysis that may improve its
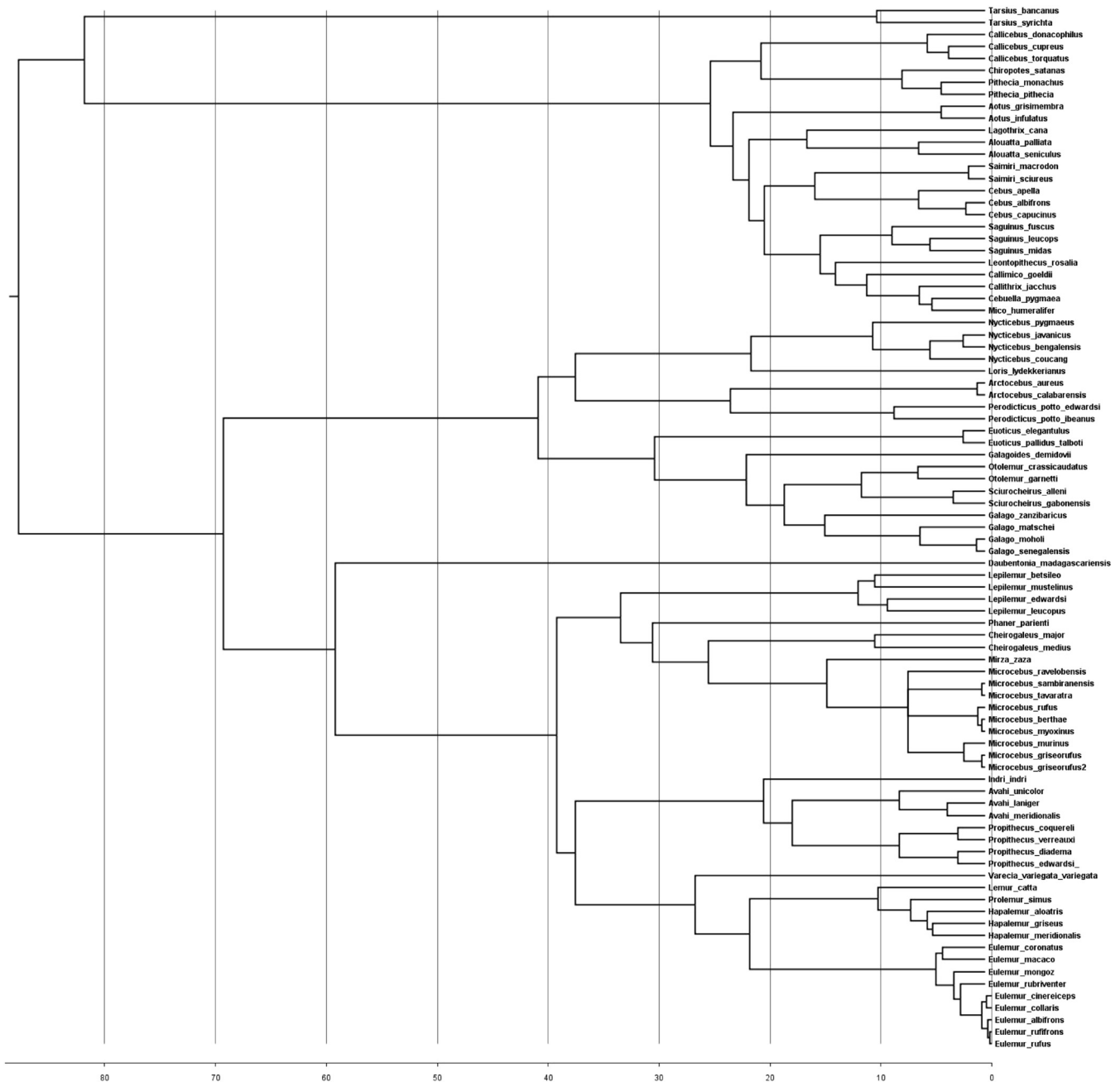
Figure 1. Extant primate phylogeny used in analyses 5a–c. The general primate backbone is from Perelman et al. (2011). For Malagasy lemurs, the tree of Yoder (2013) is the base, adding some modifications for *Eulemur* species (Markolf and Kappeler, 2013), *Lepilemur* species (Louis et al., 2006b), *Avahi* (Andriantompohavana et al., 2007), and cheirogaleids (Weisrock et al., 2012). The Lorisiform tree follows Pozzi et al. (2014, 2015). For Platyrrhini, the general backbone is from Schneider and Sampaio (2015), with callithrichid species following Buckner et al. (2015). The X-axis divergence date in MYA is estimated from molecular studies (analyses 5b and 5c). Analysis 5a uses different branch lengths following the method of Grafen (1989).

statistical qualities (Smith, 2002; but see Harmon and Losos [2005]) and provide a better match between the populations from which the morphological data and body masses are derived. On the other hand, correlating the morphological data from museum specimens labeled under potentially outmoded taxonomic schemes with the appropriate new designation is a major challenge. Some specimens have only the most vague locality data (e.g., "Madagascar," "east coast of Africa") and even specific locality data may not be enough when "new" taxa are sympatric or nearly so with others (e.g., *Avahi*, *Microcebus*, and *Saguinus* species and subspecies). The taxa we recognized are listed in SOM Table 1. We acknowledge that the "tips" or operational taxonomic units (OTUs) used here may not all be equivalent phylogenetic units; many may represent more than one species (e.g., the *Cheirogaleus major*, *Varecia variegata*, and *Galagoides demidovii* samples) and some are likely to be geographic variants that may be lumped in the future. The results of the PIC analyses are of course sensitive to the structure of the tree, something that is likely to be in constant flux especially for lower level taxa (e.g., *Microcebus*; Weisrock et al., 2012), but also for higher level relationships (e.g., Lepilemuridae + Cheirogaleidae, position of *Aotus*).

Table 4
A comparison of body mass predictions (estimate), 95% confidence intervals (lci-uci), and 95% prediction intervals (lpi-upi) for *Archicebus* and *Teilhardina* from conventional OLS regressions and from phylogenetically informed analyses using PIC; all based on the primate database (Model 2).[a]

	Model	Archicebus					Teilhardina				
		Estimate	lci	uci	lpi	upi	Estimate	lci	uci	lpi	upi
FEMBEW	Model 2 Conventional OLS	33.4	30.8	36.1	24.8	44.8	43.6	40.6	46.9	32.5	58.5
	Model 5a PIC Grafen BL	41.3	18.5	92.4	8.3	206.0	53.1	23.9	118.2	10.7	264.0
	Model 5b PIC Divergence time BL	38.8	23.8	63.1	13.1	114.3	50.0	30.9	80.9	17.0	147.2
	Model 5c Rerooted tree	38.5	22.2	66.8	17.2	86.1	49.7	28.8	85.8	22.4	110.5
CUFACET	Model 2 Conventional OLS	35.8	30.6	41.9	22.4	57.2	44.0	37.9	51.1	27.6	70.1
	Model 5a PIC Grafen BL	48.4	17.3	135.6	6.2	378.7	58.2	20.9	162.5	7.5	454.6
	Model 5b PIC Divergence time BL	46.0	24.1	87.8	11.0	193.2	55.5	29.3	105.2	13.2	232.4
	Model 5c Rerooted tree	45.0	21.7	93.5	15.5	130.9	54.3	26.3	112.2	18.8	156.8
TROCHAREA	Model 2 Conventional OLS						36.6	30.2	44.3	19.9	67.4
	Model 5a PIC Grafen BL						54.2	12.7	231.2	3.0	985.8
	Model 5b PIC Divergence time BL						56.7	21.9	146.9	6.9	470.2
	Model 5c Rerooted tree						52.8	18.0	154.9	11.0	253.1
HUMBEW	Model 2 Conventional OLS	36.4	32.1	41.3	22.5	59.1	37.5	33.1	42.4	23.1	60.7
	Model 5a PIC Grafen BL	52.7	19.0	146.1	6.9	404.2	54.4	19.6	150.7	7.1	417.2
	Model 5b PIC Divergence time BL	39.1	20.8	73.4	9.7	157.0	40.5	21.6	75.9	10.1	162.5
	Model 5c Rerooted tree	37.0	18.1	75.4	13.1	104.2	38.3	18.8	78.0	13.6	107.5
CWID	Model 2 Conventional OLS	74.5	60.2	92.3	31.9	173.9	61.6	49.1	77.3	26.3	144.3
	Model 5a PIC Grafen BL	100.7	22.2	456.1	4.8	2115.5	85.4	18.8	388.6	4.1	1798.4
	Model 5b PIC Divergence time BL	101.5	40.2	256.2	12.2	843.9	86.3	33.9	219.4	10.3	719.7
	Model 5c Rerooted tree	99.9	34.6	287.8	20.8	479.3	84.8	29.2	245.9	17.7	406.8
AWID	Model 2 Conventional OLS						45.1	37.5	54.3	24.1	84.2
	Model 5a PIC Grafen BL						58.5	18.2	188.2	5.6	609.5
	Model 5b PIC Divergence time BL						51.9	23.2	116.4	8.6	314.9
	Model 5c Rerooted tree						48.7	19.5	121.8	12.8	185.3

[a] Results are shown for the generic PIC equations using two different methods to determine branch lengths: Grafen's (1989) method (Model 5a) and a model based on divergence dates (Model 5b). The results from the rerooting technique of Garland and Ives (2000) using the divergence date tree are also shown (Model 5c). The values are detransformed using the ML correction factor.

2. Small intrataxon sample size. Another issue for this study is that the intrataxon sample size for postcranial measurements, which was small enough to begin with (see Dagosto and Terranova, 1992), was often further reduced to accommodate new taxa and to remove unclassifiable specimens. In the 1992 database, 100 percent of taxa had $N \geq 5$ and 57% had $N \geq 10$; in the current database (Model 1) only 44% of taxa had $N \geq 5$ and only 14% had $N \geq 10$. One concern is that small sample size can greatly increase the chance of error in matching literature body masses with skeletal measurements (Smith, 2002). Another is that small sample sizes may not adequately take into account intrataxon variation and that results (parameter estimation, significance levels, hypothesis testing) may be skewed (Garamszegi, 2014). This is known to be particularly problematic in PIC, in which poorly estimated tips (whether from sampling or measurement error) in closely related species can have a great effect on the analysis (Ricklefs and Starck, 1996; Purvis and Webster, 1999; Felsenstein, 2008). Harmon and Losos (2005) note that in a nonphylogenetic correlation context, small intraspecific sample size is more likely to lead to Type II rather than Type I error, while in a phylogenetic context the possibility of Type I error is increased. Their simulations, however, show that the effect of low intraspecific sample size is less where correlations are high (as with the relationships analyzed here) and where interspecific variation is large compared to intraspecific variation (e.g., Models 1 and 2), but has more potential for impact at a restricted size range (e.g., Model 3 here) depending on the distribution of the variance. Still, increasing intraspecific sample size can generally lead to improvement in statistical power.

3. Heterogeneity of sample sizes. Sample sizes are unequal among taxa, which means that some points are more reliably estimated than others, an issue that is admittedly not controlled for in these analyses. The generally high repeatability of morphological measurements should in theory ameliorate such effects (Garamszegi, 2014).

4. Determining "mean" body mass. Whereas the morphological sample may be unsatisfactorily small, at least the process of determining the arithmetic mean is straightforward (assuming of course that arithmetic mean is the best measure of central tendency to use). Not so for body mass. The recent explosion of fieldwork on primates has much improved our knowledge since the compendium of Smith and Jungers (1997). However, it is again a two-edged sword. The recent update for Malagasy species by Taylor and Schwitzer (2011/12), for example, highlights the presence of populational, geographic, sexual, and seasonal variation in mass. With a few exceptions, only means and ranges (sometimes separated by sex) are reported, not individual values. It is not such a simple matter to decide on a reasonable value for central tendency when there are multiple such values reported from different studies. This is what we have done:

   a. Since most prosimian primates are not strongly dimorphic and the morphological specimens were for the most part unsexed, we lumped male and female values for both morphological variables and mass; if male and female values were reported separately, we "averaged" them (of course, not a true average if N is different). We did the same for the platyrrhines in our sample, even though in some cases they exhibit a greater degree of dimorphism than the prosimians.

   b. We did the same when a taxon was represented by more than one population with different weights, but we cannot match the morphological sample to either one (e.g., *Propithecus verreauxi* from Kirindy and Beza Mahafaly), or if we cannot determine which of the sympatric (or nearly so) taxa the morphological specimens belong to (e.g., *Avahi laniger* or *Avahi peyrierasi*)

   c. When there were several means (or midranges) from different studies presented for a species, we computed 1) the arithmetic mean of all the values, 2) the geometric mean, 3) the median, and 4) a sort of weighted average based on the sample size to give higher influence to the better sampled means (again, not a true weighted average since only the mean value of a sample is reported). Luckily, in most cases these all yield reasonably similar estimates for taxon central tendency (SOM Table 1). We used the arithmetic mean in our analyses.

5. Correction factors. Estimates derived from regressions calculated in log form require correction to be expressed in the arithmetic scale (see reviews in Smith, 1993; Clifford et al., 2013). One solution is to simply leave the results in logarithmic form as is done here in Figures 2–4. When it is desirable to state results in the original scale (e.g., Tables 4 and 5), a correction is employed. Some methods yield a single correction factor (CF) that is applied to any X (e.g., REML [=QMLE], Smear factor, ML, Finney's; see Smith 1993; Clifford et al., 2013 for definitions), others yield CFs that vary with X such as the uniformly minimum variance unbiased (UMVU), which is the only unbiased method. Simulations have shown that the differences among the methods can be small (Clifford et al., 2013) and the same is true for our data (Table 6). Applying different CFs, or even no CF, would not significantly alter any of the major conclusions of this paper. CFs vary from about 1% for the better predictors to 11% for the poorer ones (Table 3).

Since the regressions for the PIC analyses are computed using standardized independent contrasts forced through the origin, it is more difficult to calculate the preferred UMVU correction. For the conventional OLS regressions considered here, the results of the ML method [$(SEE^2/2)*(N-2/N)$] differed least from UMVU, therefore, it was used to approximate a correction for both the PIC and conventional regressions to facilitate comparisons between them. The estimates in Tables 4 and 5 use this correction.

3. Results

3.1. Performance of variables

Based on R^2, %SEE, and %PE (Table 3; Figs. 2 and 3), the most reliable variable of those tested was FEMBEW. In a study of catarrhine primates, Ruff (2003) found his very similar measurement ("femoral total condyle mediolateral breadth [FCML]") to be "locomotor blind" and thus very useful for body mass prediction. Neither the phylogenetic composition nor the body size range of the sample had a very strong effect on the outcome. The %PE for the target-similar species (mouse lemurs and tarsiers) was quite reasonable (7−8%) regardless of model. This is especially noteworthy since the body mass of tarsiers has proven difficult to predict (Gingerich, 1981; Gingerich et al., 1982; Dagosto and Terranova, 1992; Yapuncich et al., 2015). Of all the regressions run here, the residuals from this model were the only ones not to exhibit phylogenetic signal as measured by Pagel's λ, although they were significant using Blomberg's K. Therefore, it is doubtful that a phylogenetic correction is actually necessary here, but the results are presented for exploratory purposes. The PIC regressions yielded lower R^2 and slightly higher %SEEs and %PEs, but the results were still very good and the best of any phylogenetically informed variable.

As expected from previous work (Dagosto and Terranova, 1992; Yapuncich et al., 2015), CUFACET also performed fairly well across all conventional models but less well for the target-similar species after phylogenetic correction. TROCHAREA also predicted the mass of mouse lemurs fairly well in most conventional models. The overall %PE (>20%) and the %PE for tarsiers (>40%) were, however, much higher than with FEMBEW or CUFACET. After phylogenetic correction, the %PEs for mouse lemurs were also unacceptably high. HUMBEW was similar in performing reasonably well for mouse lemurs but poorly for tarsiers.

The other variables (AWID and CWID) are judged to be less reliable since they have higher overall %PE and/or do a poorer job predicting the mass of the target-similar species. They also were more affected by alterations in the sample composition and had much larger confidence intervals and prediction intervals (Figs. 2 and 3) than the better predictors.

3.2. Performance of different models

In general, all the models performed similarly; the better the variable, the less difference among the models. Expanding the taxonomic range of model 2 (noncatarrhine primates) to include tree shrews and dermopterans (Model 1) never gave better results (in terms of R^2, %SEE, overall %PE, or target-similar %PE) than the primate-only model. The only exception is the prediction error for mouse lemurs and tarsiers from CWID that decreased using the euarchontan model. Similarly, limiting the reference group to small euarchontans (Model 3) or small mammals (Model 4) did not give better (or particularly worse) results than those using Model 2.

Analyses using phylogenetic correction had lower R^2s and higher %PEs both for overall and for target species, indicating that phylogeny had some, but not an overwhelming influence on the strong relationship between body mass and these variables. A
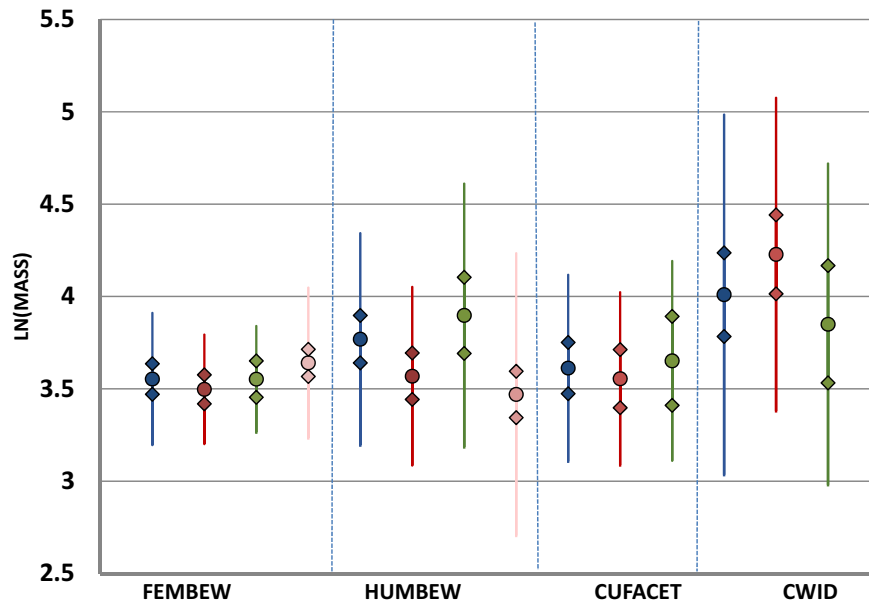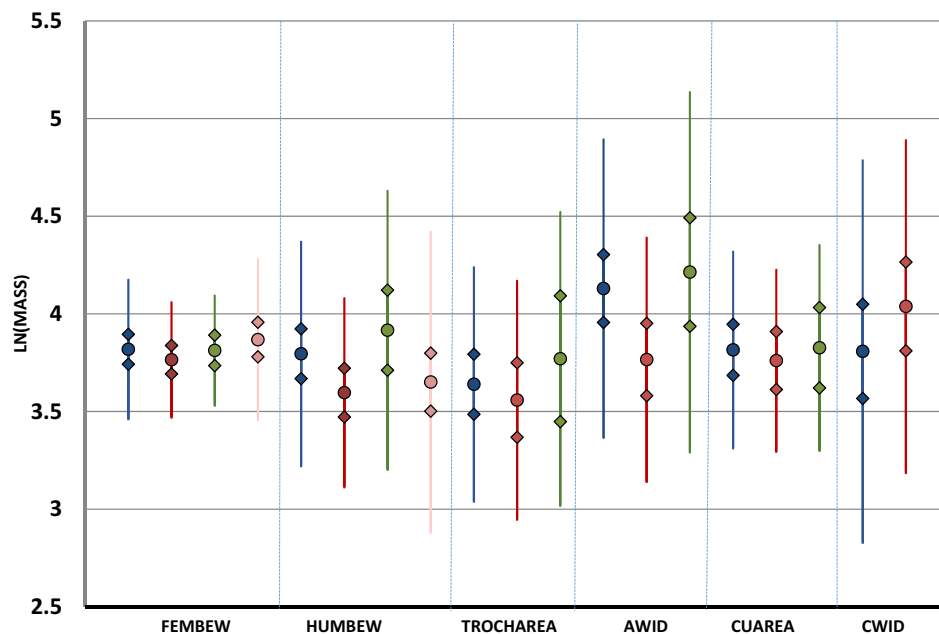
Figure 2. LN (body mass) predictions made for *Archicebus achilles* based on the conventional OLS models of the measurements on the X-axis. The point estimate (circle), the 95% confidence interval (diamonds), and the 95% prediction intervals (line) are shown. For each predictor from left to right, blue indicates the euarchontan model (1), red the primate model (2), green the small euarchontan model (3), and pink the small mammal model (4; only calculated for FEMBEW and HUMBEW). Note the narrower confidence and prediction intervals for the better (FEMBEW, CUFACET) versus poorer models (HUMBEW, CWID) in each pair. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Figure 3. LN (body mass) predictions made for *Teilhardina belgica* based on the conventional OLS models of the measurements on the X-axis. The point estimate (circle), the 95% confidence interval (diamonds), and the 95% prediction intervals (line) are shown. For each predictor from left to right, blue indicates the euarchontan model (1), red the primate model (2), green the small euarchontan model (3), and pink the small mammal model (4; only calculated for FEMBEW and HUMBEW). Note the narrower confidence and prediction intervals for the better (FEMBEW, TROCHAREA, CUFACET) versus poorer models (HUMBEW, AW, CW) in each pair. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

similar result was reported in a study of phocids (Churchill et al., 2014).

The ranking of variables from better (FEMBEW, CUFACET) to worse (AWID, CWID) was the same as with conventional regression. The difference between PIC and non-PIC results was smallest with FEMBEW (the only one of the variables that may not require phylogenetic correction) and CUFACET, but generally increased more in the other variables. There were only minor differences in the regression and prediction results (R^2, %SEE, overall %PE) between the trees using different branch length assumptions, but the divergence date tree generally gave better results for the target-similar species than the tree using Grafen method lengths.
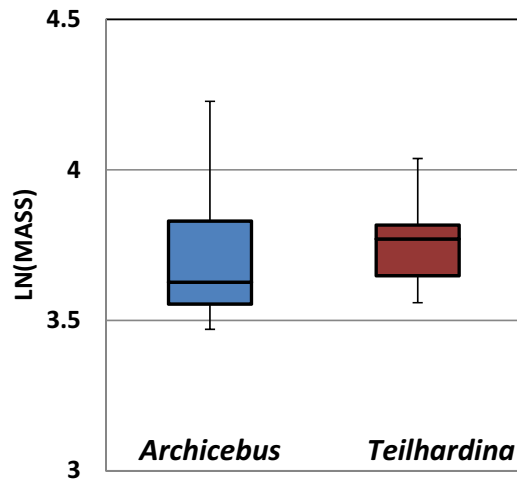
Figure 4. A boxplot of point estimates of body mass (in natural log) from conventional OLS models. Note that *Teilhardina* has a narrower interquartile range than does *Archicebus*. See Table 5 for values.

3.3. Predictions for Archicebus and Teilhardina from conventional OLS models

Table 5 and Figures 2 and 3 illustrate the predictions made for *A. achilles* and *T. belgica* based on the conventional OLS models. The more reliable variables (FEMBEW, CUFACET) yield estimates for *Archicebus* that average 36 g (range 33–39 g across different models), values generally lower than those predicted from the less reliable estimators (HUMBEW, CWID; mean = 51 g; range 34–75 g). For *Teilhardina*, the more reliable estimators suggest a body mass of 46 g (range 44–49 g), while the range from the less reliable estimators is 35–73 g (mean = 49 g). These results are similar to those generated previously from skull length, body length, and tarsal variables. As could be concluded from the raw measurements (Table 1), *Archicebus* was likely smaller than *T. belgica*. The point estimates from the better estimates do not overlap, although their confidence and prediction intervals do, as do the predictions from the poorer estimators.

The higher predictions from CWID are statistical outliers for both species. This phenomenon has been noted previously for other

Table 5
Point estimates and prediction intervals of body mass (in g) for *Archicebus* and *Teilhardina* from the conventional OLS models detransformed with the correction factor.[a]

Taxon	Measurement	Euarchontan Model 1	Non-catarrhine Primates Model 2	Small euarchontans Model 3	Small mammals Model 4
Archicebus	FEMBEW	35.5 (24–50) 3.553334	33.4 (25–45) 3.497424	35.2 (26–47) 3.552638	38.9 (26–58) 3.641462
	CUFACET	38.1 (23–63) 3.612459	35.8 (22–57) 3.554956	39.4 (23–68) 3.651537	
	HUMBEW	45.0 (25–80) 3.769022	36.4 (22–59) 3.568436	51.7 (25–105) 3.897825	34.4 (27–61) 3.469737
	CWID	61.6 (23–163) 4.009866	74.5 (32–174) 4.228153	50.2 (21–120) 3.849789	
Teilhardina	FEMBEW	46.2 (32–66) 3.818751	43.6 (32–59) 3.765341	45.6 (34–60) 3.813035	48.8 (32–74) 3.869252
	CUFACET	46.8 (28–77) 3.815926	44.0 (28–70) 3.761063	46.9 (28–79) 3.82667	
	TROCHAREA	39.7 (22–72) 3.639797	36.6 (20–67) 3.558822	45.3 (21–96) 3.770345	
	HUMBEW	46.3 (26–82) 3.795907	37.5 (23–61) 3.596807	52.7 (26–108) 3.916506	35.4 (19–88) 3.496651
	AWID	66.5 (31–143) 4.129881	45.1 (24–84) 3.766217	72.6 (29–182) 4.214121	
	CWID	50.4 (19–134) 3.808113	61.6 (26–144) 4.038054	43.1 (18–105) 3.697928	

[a] The untransformed log (e) value of the estimate is also given. The models refer to different sample compositions as described in the Methods section of the text: Model 1 is 98 taxa of euarchontans, Model 2 is 88 taxa of non-catarrhine primates, Model 3 is 22 taxa of euarchontans that weigh less than 200 g, Model 4 is 55 mammals in the size range 10–100 g. Precise sample sizes and other regression statistics for each model are given in Table 3.

Table 6
A comparison of the estimates for *Microcebus* species derived from the regression of mass on FEMBEW in primates (Model 2) after application of various correction factors (CF).[a]

	Naive	UMVU	REML	ML	Finney's	Ratio	Smear
Microcebus berthae	32.80	33.11	33.14	33.13	33.13	33.44	33.12
M. griseorufus 1	51.17	51.66	51.69	51.68	51.69	52.18	51.68
M. griseorufus 2	48.90	49.38	49.41	49.39	49.40	49.87	49.39
M. murinus	55.08	55.62	55.65	55.63	55.64	56.17	55.63
M. myoxinus	51.94	52.44	52.47	52.46	52.46	52.96	52.45
M. ravelobensis	58.35	58.92	58.95	58.93	58.94	59.50	58.93
M. rufus	37.64	38.00	38.03	38.01	38.02	38.38	38.01
M. sambiranensis	40.48	40.87	40.90	40.89	40.89	41.28	40.88
M. tavaratra	55.08	55.62	55.65	55.63	55.64	56.17	55.63
CF		1.009977*	1.010274	1.009979	1.010125	1.019706	1.009954

*UMVU yields a different correction factor for each X; the average value is given here.
[a] The "naïve" estimate is the uncorrected antilog. Other methods are described in Smith (1993) and Clifford et al. (2013). The R code of Clifford et al. (2013) was used to calculate CFs for the OLS regressions.

Eocene primates (Dagosto and Terranova, 1992) and is due to the more distal placement of the peroneal tubercle in fossil taxa compared to living primates. *Archicebus* is also noted to have a particularly wide calcaneus (Ni et al., 2013).

A boxplot (Fig. 4) shows that *Teilhardina* has a narrower interquartile range than does *Archicebus*. This may be because for most variables, *Teilhardina* is represented by more than one individual giving a more reliable estimate of the species mean, while *Archicebus* is known from only one individual and thus the predictions are accompanied by much greater uncertainty (Smith, 2002). Also, the measurements of *Archicebus* were taken from the CT reconstructions rather than from direct caliper measurements as in the majority of other specimens in the reference sample, although these were calibrated to calcaneal length that could be measured directly on the specimen (Ni et al., 2013). If and how this difference in measurement technique might affect the results is unknown.

3.4. Predictions for Archicebus and Teilhardina from phylogenetically informed models

Estimates for *Archicebus* and *Teilhardina* from phylogenetically corrected models using the all-primate database (Model 2) were generated following procedures outlined in Garland and Ives (2000). Despite the fact that the overall regression statistics of the Grafen and divergence date trees are similar, indicating that the analyses are reasonably robust to different assumptions about branch lengths (Diaz-Uriarte and Garland, 1998), the point estimates for the fossils are higher using the Grafen tree than the divergence tree (Table 4). The predictions from the divergence date tree are closer to those of the conventional regression.

Point estimates for *Archicebus* and *Teilhardina* using the generic PIC equation are 2% (HUMBEW) to 44% (TROCHAREA) higher than those from the conventional OLS regressions (Table 4) and the range of point estimates expands for both taxa. The values derived from CWID are again strong outliers. For *Archicebus*, the mean of point estimates from the Grafen tree from the more reliable estimators (FEMBEW and CUFACET) was 45 g and 56 g for *Teilhardina*, both 20–25% higher than the results from the conventional regressions. Estimates from the divergence date tree best estimators yield lower averages of 42 g for *Archicebus* and 52 g for *Teilhardina*. Generic prediction intervals are, as expected, much higher than in conventional regression (Garland and Ives, 2000). The rerooting procedure was successful at reducing these (Table 4).

4. Conclusions

The purpose of this study was to compare the results of applying different reference models and techniques to the prediction of body mass in the early Eocene fossil primate taxa *Archicebus* and *Teilhardina*. In general, the choice of a variable was more crucial than the choice of reference population. Model composition had a minimal effect on the better performing variables but a greater effect on the more poorly performing ones. For example, restricting the analysis to primates (Model 2) generally yielded better results than the more inclusive euarchontan model (Model 1)—but this could be due not just to the expanded taxonomic range but to the smaller morphological sample sizes and less well known body masses for the nonprimate euarchontans. Restricting the analysis to a smaller size range (either small euarchontans [Model 3] or small mammals [Model 4]) had relatively little effect on the results for FEMBEW or CUFACET, but was greater for the other variables.

PIC yielded slightly different results than conventional OLS, although the ranking of variables from good (FEMBEW, CUFACET) to poor (AWID, CWID) was very similar. Estimates derived from PIC models will of course vary according to the phylogenetic hypothesis and branch lengths employed. Given the particular set of assumptions used here, PIC analyses had lower R^2, higher %PEs, and larger prediction intervals than conventional regression of the same variables. This does not mean that one should not choose to apply phylogenetic correction if it seems necessary (i.e., if there is phylogenetic signal in the model residuals). On the contrary, these approaches may provide a more realistic picture of the relationship between body mass and morphological variables and avoid the potentially false sense of statistical strength produced by conventional OLS (Felsenstein, 1985; Garland et al., 1992; Pyenson and Sponberg, 2011).

Predictions for the fossil taxa were 2–44% higher and had larger confidence and prediction intervals than the conventional OLS regressions. The effect of rerooting the tree was minimal on the point estimates from the better estimators but had a greater effect on the poorer estimators. It was helpful in reducing the wider prediction intervals of the generic PIC regression. In this particular case, the basal position of the fossil taxa and their phylogenetic and temporal remoteness from the extant model taxa posed particular problems for the rerooting analysis. It is likely that the same issue will apply to any Eocene omomyiform or adapiform and we may have to be resigned to generic predictions and their larger prediction intervals for such taxa. This approach, however, should prove to be very useful for fossils that can be nested within or near a speciose extant taxon (Smith, 2002).

This study shows that the body mass of living euarchontans and primates can be reasonably well estimated from various postcranial measures (particularly FEMBEW and CUFACET). FEMBEW in particular seems to be not only "locomotor blind" (Ruff, 2002, 2003), but also nearly "phylogenetically blind," at least for the non-catarrhine primate reference population considered here. The "poorer" variables (AWID, CWID) have a weaker relationship with body mass in the taxa examined, suggesting that they may also contain functional and/or phylogenetic signal.

We used these equations to estimate the body mass of two Eocene primates, *A. achilles* and *T. belgica*. Most variables and methods addressed here did predict values in the range of the small and medium sized mouse lemurs as expected from the raw values of the measures and the results of previous predictions. The values generated from our best estimators for *Archicebus* (33–46 g) are in the range of the 30–33 g estimated from skull length and body length by Ni et al. (2013), but higher than their 20 g estimate derived from upper molar size. Our best estimates are lower than Boyer et al.'s (2013) 62 g estimate based on their very similar measure of cuboid facet area. Whether this is due to different sample composition (their regression is based on a euarchontan sample that includes catarrhine primates) or a different measurement technique we cannot say. The best estimators for *T. belgica* suggest a body mass of 44–56 g, similar to what has been predicted previously from smaller samples of various tarsal measures (Dagosto and Terranova, 1992; Boyer et al., 2013) but lower than most estimates made from dental variables (Gingerich, 1981; Conroy, 1987; Dagosto and Terranova, 1992).

We are not claiming that these particular characters, models, or methods are the best possible choices for determination of body mass for these taxa. Our intent was simply to compare their effects. There are certainly other variables and procedures not explored here that have been demonstrated to improve the accuracy of predictions (e.g., multivariate regression, partial least squares regression, see Smith, 2002; Pyenson and Sponberg, 2011; Churchill et al., 2014; Yapuncich et al., 2015). There are alternative models such as classical calibration and reduced major axis regression that are recommended when it is necessary to extrapolate outside the range of living taxa, as might be the case here (e.g., Jungers, 1988; Konigsberg et al., 1998; Hens et al., 2000; but see Smith, 2002).

Many methods for determining, transforming, and simulating the effects of alternative trees, branch lengths, and evolutionary models can be employed (e.g., Felsenstein, 2004), as well as techniques for taking into account intraspecific variation (e.g., Garamszegi, 2014).

Even though the estimates for the fossil taxa generated here seem reasonable, we would still urge some caution in using them to make further inferences. These analyses are based on small samples where the source of the morphological data and the body masses are unmatched. The best variable (FEMBEW) as measured by overall performance still yields %PEs for taxa used to construct the equation ranging from almost zero to nearly 100% and is only known from one individual of each fossil species.

Even the best performing variables, those that are most unaffected by sample composition and phylogenetic correction, yield confidence and prediction intervals wide enough that it would be wise to take them into account when using the estimates to make inferences about paleobiology. For example, to determine if middle phalanx length of *Teilhardina* was long compared to that of a typical primate, Gebo et al. (2012) used a range of body mass (30—50 g) to make the comparison. Even at 80 g, *Teilhardina* would still have a relatively long middle phalanx; thus, this conclusion seems fairly robust. On the other hand, answering the question "Are *Teilhardina* and *Archicebus* the same size?" is more difficult to say with high confidence when prediction intervals are taken into account.

Acknowledgments

Supplementary Online Material

Supplementary online material to this article can be found at http://dx.doi.org/10.1016/j.jhevol.2017.02.005.

References

Andriaholinirina, N., Fausser, J.-L., Roos, C., Zinner, D., Thalmann, U., Rabarivola, C., Ravoarimanana, I., Ganzhorn, J.U., Meier, B., Hilgartner, R., Walter, L., Zaramody, A., Langer, C., Hahn, T., Zimmermann, E., Radespiel, U., Craul, M., Tomiuk, J., Tattersall, I., Rumpler, Y., 2006. Molecular phylogeny and taxonomic revision of the sportive lemurs (*Lepilemur*, Primates). BMC Evol. Biol. 6, 17.

Andriantompohavana, R., Lei, R., Zaonarivelo, J., Engberg, S., Nalanirina, G., McGuire, S., Shore, G., Andrianasolo, J., Herrington, K., Brenneman, R., Louis, E.E., 2007. Molecular Phylogeny and Taxonomic Revision of the Woolly Lemurs, Genus *Avahi* (Primates: Lemuriformes). Special Publications of the Museum of Texas Tech Univeristy, Lubbock, TX, pp. 1—59.

Blomberg, S.P., Garland, T., Ives, A.R., 2003. Testing for phylogenetic signal in comparative data: behavioral traits are more labile. Evolution 57, 717—745.

Boyer, D.M., Seiffert, E.R., Gladman, J.T., Bloch, J.I., 2013. Evolution and allometry of calcaneal elongation in living and extinct primates. PLoS One 8, e67792.

Buckner, J.C., Lynch Alfaro, J.W., Rylands, A.B., Alfaro, M.E., 2015. Biogeography of the marmosets and tamarins (Callitrichidae). Mol. Phylo. Evol. 82, Part B, 413—425.

Butyinski, T., Kingdon, J., Kalina, J., 2013. Mammals of Africa. In: Volume II. Primates. Bloomsbury, London.

Churchill, M., Clementz, M.T., Kohno, N., 2014. Predictive equations for the estimation of body size in seals and sea lions (Carnivora: Pinnipedia). J. Anat. 225, 232—245.

Clifford, D., Cressie, N., England, J.R., Roxburgh, S.H., Paul, K.I., 2013. Correction factors for unbiased, efficient estimation and prediction of biomass from log—log allometric models. Forest Ecol. Manag. 310, 375—381.

Conroy, G.C., 1987. Problems of body-weight estimation in fossil primates. Intl. J. Primatol. 8, 115—137.

Dagosto, M., Terranova, C.J., 1992. Estimating the body size of Eocene primates: a comparison of results from dental and postcranial variables. Intl. J. Primatol. 13, 307—344.

Delson, E., Terranova, C.J., Jungers, W.L., Sargis, E.J., Jablonski, N.G., Dechow, P.C., 2000. Body mass in Cercopithecidae (Primates, Mammalia): estimation and scaling in extinct and extant taxa. Anthropol. Pap. Am. Mus. Natl. Hist. 83, 1—159.

Diaz-Uriarte, R., Garland, T., 1998. Effects of branch length errors on the performance of phylogenetically independent contrasts. Syst. Biol. 47, 654—672.

Driller, C., Merker, S., Perwitasari-Farajallah, D., Sinaga, W., Anggraeni, N., Zischler, H., 2015. Stop and go—waves of tarsier dispersal mirror the genesis of Sulawesi Island. PLoS One 10, e0141212.

Fausser, J.-L., Prosper, P., Donati, G., Ramanamanjato, J.-B., Rumpler, Y., 2002. Phylogenetic relationships between *Hapalemur* species and subspecies based on mitochondrial DNA sequences. BMC Evol. Biol. 2, 1—9.

Felsenstein, J., 1985. Phylogenies and the comparative method. Am. Nat. 125, 1—15.

Felsenstein, J., 2004. Inferrring Phylogenies. Sinauer, Sunderland.

Felsenstein, J., 2008. Comparative methods with sampling error and within-species variation: contrasts revisited and revised. Am. Nat. 171, 713—725.

Finstermeier, K., Zinner, D., Brameier, M., Meyer, M., Kreuz, E., Hofreiter, M., Roos, C., 2013. A mitogenomic phylogeny of living primates. PLoS One 8, e69504.

Garamszegi, L., 2014. Uncertainties due to within-species variation in comparative studies: measurement errors and statistical weights. In: Garamszegi, L. (Ed.), Modern Phylogenetic Comparative Methods and Their Application in Evolutionary Biology. Springer, New York, pp. 157—199.

Garland, T., Harvey, P.H., Ives, A.R., 1992. Procedures for the analysis of comparative data using phylogenetically independent contrasts. Syst. Biol. 41, 18—32.

Garland, T., Ives, A.R., 2000. Using the past to predict the present: confidence intervals for regression equations in phylogenetic comparative methods. Am. Nat. 155, 346—364.

Gebo, D.L., Smith, T., Dagosto, M., 2012. New postcranial elements for the earliest Eocene fossil primate *Teilhardina belgica*. J. Hum. Evol. 63, 205—218.

Gebo, D.L., Smith, R., Dagosto, M., Smith, T., 2015. Additional postcranial elements of *Teilhardina belgica*: The oldest European primate. Am. J. Phys. Anthropol. 156, 388—406.

Gingerich, P.D., 1981. Early Cenozoic Omomyidae and the evolutionary history of tarsiiform primates. J. Hum. Evol. 10, 345—374.

Gingerich, P.D., Smith, H., Rosenberg, K., 1982. Allometric scaling in the dentition of Primates and prediction of body weight from tooth size in fossils. Am. J. Phys. Anthropol. 58, 81—100.

Grafen, A., 1989. The phylogenetic regression. Phil. Trans. Roy. Soc. Lond. 326, 119—157.

Harmon, L.J., Losos, J.B., 2005. The effect of intraspecific sample size on type I and type II error rates in comparative studies. Evolution 59, 2705—2710.

Hens, S.M., Konigsberg, L.W., Jungers, W.L., 2000. Estimating stature in fossil hominids: which regression model and reference sample to use? J. Hum. Evol. 38, 767—784.

Jungers, W.L., 1988. Lucy's length: Stature reconstruction in *Australopithecus afarensis* (A.L.288—1) with implications for other small-bodied hominids. Am. J. Phys. Anthropol. 76, 227—231.

Jungers, W.L., 1990. Problems and methods in reconstructing body size in fossil primates. In: Damuth, J., MacFadden, B. (Eds.), Body SIze in Mammalian Paleobiology. Cambridge University Press, Cambridge, pp. 103—118.

Konigsberg, L.W., Hens, S.M., Jantz, L.M., Jungers, W.L., 1998. Stature estimation and calibration: Bayesian and maximum likelihood perspectives in physical anthropology. Am. J. Phys. Anthropol. 107, 65—92.

Lei, R., Frasier, C.L., McLain, A.T., Taylor, J.M., Bailey, C.A., Engberg, S.E., Ginter, A.L., Randriamampionona, R., Groves, C.P., Mittermeier, R.A., Louis, Edward E., 2014. Revision of Madagascar's dwarf lemurs (Cheirogaleidae: *Cheirogaleus*): Designation of species, candidate species status and geographic boundaries based on molecular and morphological data. Prim. Cons. 28, 9—35.

Louis, E., Coles, M., Andriantompohavana, R., Sommer, J., Engberg, S., Zaonarivelo, J., Mayor, M., Brenneman, R., 2006a. Revision of the mouse lemurs (*Microcebus*) of eastern Madagascar. Intl. J. Primatol. 27, 347—389.

Louis, E.E., Engberg, S., Lei, R., Geng, H., Sommer, J., Randriamampionona, R., Randrianmanana, J., Zaonarivelo, J., Andriantompohavana, R., Randria, G., Prosper, Ramaromilanto B., Rakotoarisoa, G., Rooney, A., Brenneman, R., 2006b. Molecular and Morphological Analyses of the Sportive Lemurs (Family Megaladpidae: Genus *Lepilemur*) Reveals 11 Previously Unrecognized Species. Special Publications, Museum of Texas Tech Univeristy, pp. 1—47.

Lynch Alfaro, J.W., Boubli, J.P., Paim, F.P., Ribas, C.C., da Silva, M.N.F., Messias, M.R., Röhe, F., Mercês, M.P., Silva, J.S., Silva, C.R., Pinho, G.M., Koshkarian, G., Nguyen, M.T.T., Harada, M.L., Rabelo, R.M., Queiroz, H.L., Alfaro, M.E., Farias, I.P., 2015. Biogeography of squirrel monkeys (genus *Saimiri*): South-central Amazon origin and rapid pan-Amazonian diversification of a lowland primate. Mol. Phylo. Evol. 82, Part B, 436—454.

Maddison, W.P., Maddison, D.R., 2015. Mesquite: a modular system for evolutionary analysis. Version 3.04. http://mesquiteproject.org.

Markolf, M., Kappeler, P., 2013. Phylogeographic analysis of the true lemurs (genus *Eulemur*) underlines the role of river catchments for the evolution of microendemism in Madagascar. Front. Zool. 10, 70.

Martin, R.D., 1990. Estimating body mass and correlated variables in extinct mammals: travels in the fourth dimension. In: Damuth, J., McFadden, B.J. (Eds.), Body Size in Mammalian Paleontology. Cambridge University Press, New York, pp. 49—68.

Masters, J.C., Silvestro, D., Génin, F., DelPero, M., 2013. Seeing the wood through the trees: the current state of higher systematics in the Strepsirhini. Folia Primatol. 84, 201—219.

Midford, P.E., Garland, T., Maddison, W.P., 2009. PDAP: PDTREE package for Mesquite. version 1.15. http://mesquiteproject.org/pdap_mesquite/.

Ni, X., Gebo, D.L., Dagosto, M., Meng, J., Tafforeau, P., Flynn, J.J., Beard, K.C., 2013. The oldest known primate skeleton and early haplorhine evolution. Nature 498, 60—64.

Ni, X.J., Wang, Y.Q., Hu, Y.M., Li, C.K., 2004. A euprimate skull from the early Eocene of China. Nature 427, 65—68.

Pagel, M.D., 1999. Inferring the historical patterns of biological evolution. Nature 401, 877—884.

Perelman, P., Johnson, W.E., Roos, C., Seuánez, H.N., Horvath, J.E., Moreira, M.A.M., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M.P.C., Silva, A., O'Brien, S.J., Pecon-Slattery, J., 2011. A molecular phylogeny of living primates. PLoS Genet. 7, e1001342.

Pozzi, L., Disotell, T., Masters, J., 2014. A multilocus phylogeny reveals deep lineages within African galagids (Primates: Galagidae). BMC Evol. Biol. 14, 72.

Pozzi, L., Nekaris, K.A.-I., Perkin, A., Bearder, S.K., Pimley, E.R., Schulze, H., Streicher, U., Nadler, T., Kitchener, A., Zischler, H., Zinner, D., Roos, C., 2015. Remarkable ancient divergences amongst neglected lorisiform primates. Zool. J. Linn. Soc. 175, 661—674.

Purvis, A., Webster, A.J., 1999. Phylogenetically independent comparisons and primate phylogeny. In: Lee, P.C. (Ed.), Comparative Primate Socioecology. Cambridge University Press, Cambridge, pp. 44—68.

Pyenson, N.D., Sponberg, S.N., 2011. Reconstructing body size in extinct crown cetacea (Neoceti) using allometry, phylogenetic methods and tests from the fossil record. J. Mamm. Evol. 18, 269—288.

R Development Core Team, X., 2015. R: A Language and Environment for Statistical Computing, https://www.R-project.org ed. R Foundation for Statistical Computing, Vienna, Austria.

Rasoloarison, R.M., Goodman, S.M., Ganzhorn, J.U., 2000. Taxonomic revision of mouse lemurs (*Microcebus*) in the western portions of Madagascar. Intl. J. Primatol. 21, 963—1020.

Revell, L.J., 2010. Phylogenetic signal and linear regression on species data. Methods Ecol. Evol. 1, 319—329.

Revell, L.J., 2012. phytools: an R package for phylogenetic comparative biology (and other things). Methods Ecol. Evol. 3, 217—223.

Ricklefs, R.E., Starck, J.M., 1996. Applications of phylogenetically independent contrasts: a mixed progress report. Oikos 77, 167—172.

Ruff, C., 2002. Long bone articular and diaphyseal structure in Old World monkeys and apes. I: locomotor effects. Am. J. Phys. Anthropol. 119, 305—342.

Ruff, C.B., 2003. Long bone articular and diaphyseal structure in Old World monkeys and apes. II: estimation of body mass. Am. J. Phys. Anthropol. 120, 16—37.

Ruiz-García, M., Luengas-Villamil, K., Leguizamon, N., de Thoisy, B., Gálvez, H., 2015. Molecular phylogenetics and phylogeography of all the *Saimiri* taxa (Cebidae, Primates) inferred from mt COI and COII gene sequences. Primates 56, 145—161.

Schneider, H., Sampaio, I., 2015. The systematics and evolution of New World primates—a review. Mol. Phylo. Evol. 82, Part B, 348—357.

Smith, R.J., 1980. Rethinking allometry. J. Theor. Biol. 87, 97—111.

Smith, R.J., 1984. Allometric scaling in comparative biology: problems of concept and method. Am. J. Physiol. 246, R152—R160.

Smith, R.J., 1993. Logarithmic transformation bias in allometry. Am. J. Phys. Anthropol. 90, 215—228.

Smith, R.J., 2002. Estimation of body mass in paleontology. J. Hum. Evol. 43, 271—287.

Smith, R.J., Jungers, W.L., 1997. Body mass in comparative primatology. J. Hum. Evol. 32, 523—559.

Smith, T., Rose, K.D., Gingerich, P.D., 2006. Rapid Asia-Europe-North America geographic dispersal of earliest Eocene primate *Teilhardina* during the Paleocene-Eocene Thermal Maximum. Proc. Natl. Acad. Sci. USA 103, 11223—11227.

Squyres, N., Ruff, C.B., 2015. Body mass estimation from knee breadth, with application to early hominins. Am. J. Phys. Anthropol. 158, 198—208.

Szalay, F.S., 1976. Systematics of the Omomyidae (Tarsiiformes, Primates): taxonomy, phylogeny, and adaptations. Bull. Am. Mus. Nat. Hist. 156, 157—450.

Szalay, F.S., Dagosto, M., 1980. Locomotor adaptations as reflected on the humerus of Paleogene primates. Folia Primatol. 34, 1—45.

Tattersall, I., 2007. Madagascar's lemurs: cryptic diversity or taxonomic inflation? Evol. Anthropol. 16, 12—23.

Taylor, L.A., Schwitzer, C., 2011/12. Body masses of wild lemurs. Lemur News 16, 34—40.

Teilhard de Chardin, P., 1927. Les mammifères de l'Eocène inférieur de la Belgique. Mém. Mus. R. Hist. Natl. Belgique 36, 1—33.

Thiele, D., Razafimahatratra, E., Hapke, A., 2013. Discrepant partitioning of genetic diversity in mouse lemurs and dwarf lemurs—Biological reality or taxonomic bias? Mol. Phylo. Evol. 69, 593—609.

Van Valkenburgh, B., 1990. Skeletal and dental predictors of body mass in carnivores. In: Damuth, J., MacFadden, B.J. (Eds.), Body Size in Mammalian Paleontology. Cambridge Univeristy Press, New York, pp. 181—205.

Weisrock, D.W., Smith, S.D., Chan, L.M., Biebouw, K., Kappeler, P.M., Yoder, A.D., 2012. Concatenation and concordance in the reconstruction of mouse lemur phylogeny: an empirical demonstration of the effect of allele sampling in phylogenetics. Mol. Biol. Evol. 29, 1615—1630.

Yapuncich, G.S., Gladman, J.T., Boyer, D.M., 2015. Predicting euarchontan body mass: a comparison of tarsal and dental variables. Am. J. Phys. Anthropol. 157, 472—506.

Yoder, A.D., 2013. The lemur revolution starts now: The genomic coming of age for a non-model organism. Mol. Phylogenet. Evol. 66, 442—452.

Yoder, A.D., Yang, Z.H., 2004. Divergence dates for Malagasy lemurs estimated from multiple gene loci: geological and evolutionary context. Mol. Ecol. 13, 757—773.