

Adaptation, phylogeny, and covariance in milk macronutrient composition

Gregory E. Blomquist

Department of Anthropology, University of Missouri, Columbia, MO, United States of America

ABSTRACT

Background. Milk is a complicated chemical mixture often studied through macronutrient concentrations of fat, protein, and sugar. There is a long-standing natural history tradition describing interspecific diversity in these concentrations. However, recent work has shown little influence of ecological or life history variables on them, aside from maternal diet effects, along with a strong phylogenetic signal.

Methods. I used multivariate phylogenetic comparative methods to revisit the ecological and life history correlates of milk macronutrient composition and elaborate on the nature of the phylogenetic signal using the phylogenetic mixed model. I also identified clades with distinctive milks through nonparametric tests (KSI) and PhylogeneticEM evolutionary modeling.

Results. In addition to the previously reported diet effects, I found increasingly aquatic mammals have milk that is lower in sugar and higher in fat. Phylogenetic heritabilities for each concentration were high and phylogenetic correlations were moderate to strong indicating coevolution among the concentrations. Primates and pinnipeds had the most outstanding milks according to KSI and PhylogeneticEM, with perissodactyls and marsupials as other noteworthy clades with distinct selection regimes.

Discussion. Mammalian milks are diverse but often characteristic of certain higher taxa. This complicates identifying the ecological and life history correlates of milk composition using common phylogenetic comparative methods because those traits are also conservative and clade-specific. Novel methods, careful assessment of data quality and hypotheses, and a “phylogenetic natural history” perspective provide alternatives to these traditional tools.

Subjects Anthropology, Ecology, Evolutionary Studies, Zoology, Nutrition

Keywords Phylogenetic comparative methods, Maternal energetics, Nutrition, Multivariate statistics, Parental care, Mammal, Life history

INTRODUCTION

Patterns of animal parental care reflect the diversity of their life histories and adaptive solutions to ecological challenges (*Clutton-Brock, 1991*). Obligate provisioning of infants by adult females with mammary milk secretions is an ancient and unique aspect of all mammalian life histories. Milk is a complex mixture of chemicals with nutritional, immunological, and hormonal signaling functions (*Power & Schulkin, 2016*), which can change across different phases of maternal care (*Langer, 2008*).

Many hypotheses have been proposed to explain differences in milk composition among mammals (*Ben Shaul, 1963; Oftedal & Iverson, 1995*), but they often have limited

Submitted 3 July 2019
Accepted 22 October 2019
Published 13 November 2019

Corresponding author
Gregory E. Blomquist,
blomquistg@missouri.edu

Academic editor
Patricia Gandini

Additional Information and
Declarations can be found on
page 12

DOI 10.7717/peerj.8085

© Copyright
2019 Blomquist

Distributed under
Creative Commons CC-BY 4.0

OPEN ACCESS

taxonomic scope or explanatory power. Recently, a comprehensive analysis of all available high-quality milk macronutrient data (percentages of fat, protein, and sugar) identified a strong phylogenetic signal in milk composition and limited ecological and life history covariates (*Skibiél et al., 2013*). Additional details of the structure of this phylogenetic signal, such as early burst diversification versus Brownian motion, were not addressed (e.g. *Harmon et al., 2010*). Moreover, as in nearly all previous research on milk composition, each component was treated separately with univariate regression modeling. Both shortcomings are understandable given the rapid, recent development of multivariate phylogenetic comparative methods (*Adams, 2014; Adams & Collyer, 2018*) and techniques for describing phylogenetic signal (*Hardy & Pavoine, 2012; Cornwell et al., 2014; Keck et al., 2016; Bastide et al., 2018*).

Multivariate analysis of milk composition is desirable for several reasons. First, multivariate statistical methods are generally of higher power and can accurately account for correlations among the milk components with or without ecological predictors (*Vargason et al., 2017*). Second, there are strong arguments for a physiological and biochemical basis for coevolution among milk macronutrient concentrations. For example, the main milk sugar lactose draws water from blood into the mammary lumen resulting in higher volume but dilute milks (*Shennan & Peaker, 2000*). The fattiest milks, as seen seals, have little to no sugar and much less water than terrestrial mammals (*Eisert, Oftedal & Barrell, 2013*). Genetic correlations among macronutrient concentrations are also well described in dairy animals. For example, the correlation between protein and fat concentration is strongly positive (≈ 0.8 ; *Analla et al., 1996; Othmane et al., 2002*), mostly likely due to pleiotropic effects of alleles for genes that influence each trait. Third, macronutrient concentrations maybe be related through substitution to accomplish a similar nutritional goal. At least in some taxa, intraspecific and within-individual variation often shows compensatory shifts in fat versus sugar concentration such that the energy content remains stable (*Power et al., 2008; Whittier et al., 2011*).

Finally, many authors have recognized macronutrient compositions covary, such as the low-fat, high-sugar milks of most primates and perissodactyls versus the aforementioned high-fat, low-sugar milks of seals (*Ben Shaul, 1963; Martin, 1984; Oftedal & Iverson, 1995; Hinde & Milligan, 2011*). Moreover, the covariation and phylogenetic clustering were self-evident in visualizations either when simply decorating the tips of a phylogeny or in phylomorphospace plots (*Sidlauskas, 2008*) of the concentrations (*Fig. 1*). The later of these can be particularly illustrative when components were constrained to sum to 100% as in the right-angle mixture model of nutritional geometry (*Raubenheimer, 2011*).

There were two major goals of this paper. First, I further described the phylogenetic signal in milk macronutrient concentrations through univariate and multivariate statistics and visualizations. These were intended to describe the overall pattern of phylogenetic signal (e.g. Brownian motion v. early burst) and identify clades with quantitatively distinctive macronutrient concentrations. Second, I used multivariate phylogenetic regression to revisit the results of *Skibiél et al. (2013)* and tested for ecological predictors of milk composition while estimating the phylogenetic and residual covariance among macronutrients.

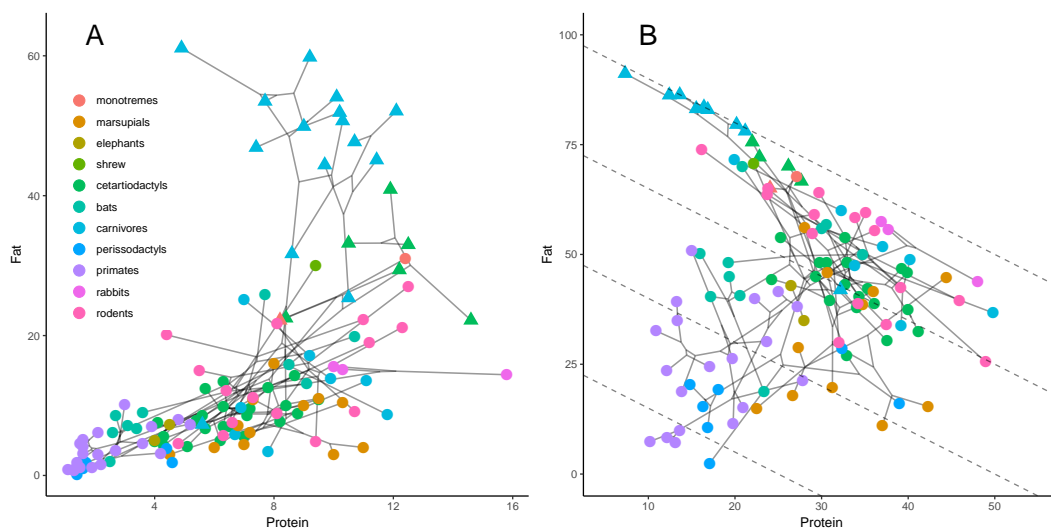


Figure 1 Nutritional geometry of milk as phylomorphospace filling. (A) Raw percentages (g/100 g) of protein and fat are plotted. (B) Percentages out of a sum of fat, protein, and sugar totaling to 100 are shown on the right. Dotted isoclines of sugar concentration are given every 25% in the with the highest percentage from sugar closest to the origin. Triangles are aquatic species in both panels.

Full-size DOI: [10.7717/peerj.8085/fig-1](https://doi.org/10.7717/peerj.8085/fig-1)

MATERIALS & METHODS

All data were initially taken from the supplementary material provided by *Skibiél et al. (2013)*. I made a handful of alterations to the milk concentration database to ensure its quality. A simple check for quality is regression of dry matter concentration against the sum of fat and protein or fat, protein, and sugar (*Oftedal & Iverson, 1995*). Large outliers from this regression were inspected and fixed with values from the original publications (*Myotis velifer* dry matter, *Arctocephalus gazella* all measures) or removed altogether where it was also inconsistent (*Thylogale billardieri*, *Perameles gunnii*, *Notomys cervinus*, and *N. mitchelli*). For *Leptonychotes weddellii*, new data including a sensitive assay of sugar were available (*Eisert, Oftedal & Barrell, 2013*). For the three *Papio* species with differing ecological data but the same milk composition, I used only *Papio anubis* which is one of the two species from which the milk data were derived (*Roberts, Cole & Coward, 1985*) and is very similar ecologically to *P. cynocephalus*. Further augmentation of the database with more recent publications was not necessary to meet the goals identified above. I used an ordinal coding of aquatic adaptation to try and more sensitively capture this feature than the binary coding of Skibiél et al. Three species (*Neovison vison*, *Castor fiber*, *Alces alces*) were categorized as partly aquatic, three families as mostly aquatic (Ornithorhynchidae, Phocidae, Otariidae), and cetaceans were the only group categorized as completely aquatic.

For phylogenetic analyses, I used a set of 1000 mammalian trees (*Faurby & Svenning, 2015*). I matched names by hand between the dataset where nomenclature differed using GenBank's taxonomy for preferred names. The trees were trimmed to the 124 species or subspecies in the dataset with the `drop.tip()` function from the `geiger` package (*Pennell et al., 2014*). After trimming only 209/1000 trees were unique indicating some

phylogenetic uncertainty, but analysis was performed using a single consensus tree from the set computed with TreeAnnotator ([Drummond et al., 2012](#)). Continuous predictors were \log_{10} -transformed to reduce skew and centered by subtracting their means to ease interpretation of intercepts in regression models. Milk macronutrient concentrations were logit-transformed to accurately account for their [0-1] boundaries as proportions. While macronutrient concentrations have traditionally been analyzed on a log scale or untransformed ([Ben Shaul, 1963](#); [Martin, 1984](#)), this approach is usually considered inappropriate when proportions are outside the 0.2–0.8 range ([Warton & Hui, 2011](#); [Schmid et al., 2013](#); [Chen et al., 2017](#)). All data manipulation and analysis were carried out in R ([R Core Team, 2018](#)). Alternatives such as beta regression or logistic generalized linear mixed models (GLMM) have not been adapted for the types of complex multivariate phylogenetic analysis desired here.

I used the MCMCglmm R package to explore a trivariate (fat, protein, sugar concentrations) phylogenetic mixed model ([Hadfield, 2010](#)). The *suite* of milk concentrations is predicted by the ecological variables used by Skibiél et al., while accounting for phylogenetic relatedness. This multivariate approach should be higher power to detect associations and it provides a phylogenetic variance–covariance matrix (VCV) that describes the coevolution of the concentrations under Brownian motion (BM). The variables include the aquatic habitat (noted above) and others as used by Skibiél et al.: arid habitat (binary), maternal body mass, adult diet coded categorically as carnivore, omnivore, herbivore, relative duration of lactation as the ratio of lactation length to the sum of gestation and lactation lengths, reproductive output as the ratio of litter mass and female mass, and an ordinal code for developmental stage at birth (precociality).

To identify clades with quantitatively distinctive milk composition I used the univariate, rank-based test introduced by [Cornwell et al. \(2014\)](#) and implemented in their R package `ksi`. This is an inherently phylogenetic test for clade distinctiveness that works by comparing the frequency distribution of a trait with and without a clade through a sample-sized scaled Kolmogorov–Smirnov test. The iterative algorithm identifies a series of clades of declining distinctiveness and reports ambiguity among neighboring nodes which may be the most recent common ancestor of the distinctive clade. For the milk composition data, I also adapted the package’s R code to allow for bivariate and trivariate versions of the test relying on the `Peacock.test` R package ([Xiao, 2017](#)). This allows testing clade distinctiveness of trait pairs (e.g., Fat-Protein) and the full set of all three concentrations (Fat-Protein-Sugar).

These rank-based tests were complemented with quantitative evolutionary modeling implemented in the PhylogeneticEM R package ([Bastide et al., 2018](#)). This is a multivariate phylogenetic method designed to detect instantaneous evolutionary shifts in sets of correlated traits, which may correspond to the invasion of new habitats or evolution of novel traits that characterize a clade. The method identifies clades that have different stabilizing selection optima (θ s) and quantifies strength of selection (a common α) for the “pull” of selection toward those optima. It is similar in scope to other packages such as `l1ou` and `mvSLOUCH`, but it uses a more complex model allowing for trait correlations ([Bartoszek et al., 2012](#); [Khabbazian et al., 2016](#)). All R code for the analysis is provided as an [Supplemental Information 1](#).

RESULTS

Graphical exploration

Correlations between concentrations of fat, protein, and sugar were clear in bivariate plots (Fig. 1). The fat-protein relationship was simplest to display because there is less missing data than sugar. There was also obvious phylogenetic clustering of species when color-coded or when phylogenetic relationships were overlaid with inferred ancestral states in a phylomorphospace plot. Constraining the three percentages to sum to 100 produces a right-angle mixture model of nutritional geometry, where sugar concentrations are diagonal isoclines in the bivariate fat-protein plot. Unusual clades stood out in both displays. Most notable were primates, perissodactyls, elephants, and some bats with low fat, low protein, high sugar milks; pinnipeds with extraordinarily high fat, modest protein, and low sugar content; cetaceans with high fat and protein and low sugar; and most marsupials with modest fat, but high protein and sugar. Additional ecological variables also helped interpret the scatter. Nearly all the species classified as aquatic by Skibieli et al. had high fat and protein with low sugar concentrations.

Distinctive clades

Non-parametric statistical tests for the distinctiveness of these clades reinforced the graphical patterns. The top 5 clades were tabulated but typically the importance statistic drops off steeply from the highest-ranked clade (Table 1). In the univariate KSI tests, pinnipeds and cetaceans were noted for high fat while perissodactyls and primates were for low fat. The primate genus *Eulemur* was also identified for its further reduction in milk fat. For protein, primates, perissodactyls and *Pteropus* stood out for their low values. Ruminants or bovids + cervids were also flagged for protein which may reflect some ambiguity in testing the cetacean node. Primates and pinnipeds were the most distinctive clades for their opposing sugar concentrations, with a marsupial node, perissodactyls, and *Pteropus* also flagged for high sugar concentrations. Bivariate and trivariate tests generally corroborated these patterns.

Evolutionary modeling

Evolutionary modeling with PhylogeneticEM agreed with the distinctiveness of these clades and quantified the different selective optima for each. The best fitting number of selective regimes was $K = 6$ with $K = 7$ another very good alternative by the package's penalized likelihood selection criterion (BGHml, Fig. 2). The regime shifts common to both solutions were pinnipeds, primates, perissodactyls, otariids, and the phocid genus *Mirounga*. The $K = 6$ and $K = 7$ solutions only differed in how they described selective regimes within marsupials: with $K = 6$ marsupials were placed within a common regime, while with $K = 7$ diprotodonts and the diprotodont species *Setonix* were placed in separate regimes. For either $K = 6$ or $K = 7$ multiple equivalent solutions were identified (degeneracy), but these only differed in the order of shifts within pinnipeds (see Supplemental Information 1).

Selective optima (θ s) were close to the average values seen within each clade (Fig. 3, Table 2). The clade optima for primates, marsupials, perissodactyls and pinnipeds were all deviations from the inferred root (12.01% fat, 7.62% protein, 3.56% sugar). Optima for

Table 1 Distinctive clades from the Kolmogorov–Smirnov Importance (KSI) tests of *Cornwell et al. (2014)*. Well-known clades names are given with others left as number. See Fig. S1 for a phylogeny with nodes labeled.

Node	Rank	KSI	KSI/max	Nodesets
Fat				
Pinnipeds	1	3.268	1.000	Pinnipeds; nd12; arctoids; nd10; carnivores; phocids
Cetaceans	2	2.221	0.680	Cetaceans; whales
Perissodactyls	3	2.160	0.661	Perissodactyls
Primates	4	2.272	0.695	Primates; anthropoids; catarrhines
<i>Eulemur</i>	5	1.677	0.513	<i>Eulemur</i>
Protein				
Primates	1	3.150	1.000	Primates; anthropoids; catarrhines; cercopithecoids
Perissodactyls	2	2.186	0.694	Perissodactyls; <i>Equus</i> ; nd33
<i>Pteropus</i>	3	2.178	0.691	<i>Pteropus</i> ; bats; nd71
Ruminants	4	1.906	0.605	Ruminants; bovids+cervids; bovids
<i>Eulemur</i>	5	1.565	0.497	<i>Eulemur</i>
Sugar				
Primates	1	3.301	1.000	Primates; anthropoids
Pinnipeds	2	2.700	0.818	Pinnipeds; nd12; phocids
nd116	3	2.169	0.657	nd116; diprotodonts; nd114; nd113; marsupials; nd117
Perissodactyls	4	1.926	0.584	Perissodactyls; <i>Equus</i>
<i>Pteropus</i>	5	1.947	0.590	<i>Pteropus</i> ; bats; nd71
Fat - Protein				
Pinnipeds	1	3.396	1.000	Pinnipeds; nd12; arctoids; nd10
Primates	2	3.056	0.900	Primates; anthropoids; catarrhines; cercopithecoids
Perissodactyls	3	2.480	0.730	Perissodactyls
Cetaceans	4	2.237	0.659	Cetaceans
<i>Pteropus</i>	5	2.162	0.637	<i>Pteropus</i> ; bats
Fat - Sugar				
Primates	1	3.433	1.000	Primates; anthropoids
Pinnipeds	2	2.700	0.786	Pinnipeds; nd12
Perissodactyls	3	2.497	0.727	Perissodactyls
Bovids+cervids	4	2.390	0.696	Bovids+cervids
nd114	5	2.141	0.624	nd114; nd113; marsupials; diprotodonts; nd116
Protein - Sugar				
Primates	1	3.653	1.000	Primates; anthropoids
Pinnipeds	2	2.700	0.739	Pinnipeds; nd12
Marsupials	3	2.395	0.656	Marsupials; nd113; nd114
Perissodactyls	4	2.276	0.623	Perissodactyls
<i>Pteropus</i>	5	2.147	0.588	<i>Pteropus</i> ; nd71
Fat - Protein - Sugar				
Primates	1	3.697	1.000	Primates; anthropoids
Bovids+cervids	2	2.846	0.770	Bovids+cervids
nd12	3	2.698	0.730	nd12; pinnipeds
Perissodactyls	4	2.447	0.662	Perissodactyls
Marsupials	5	2.292	0.620	Marsupials; nd113; nd114

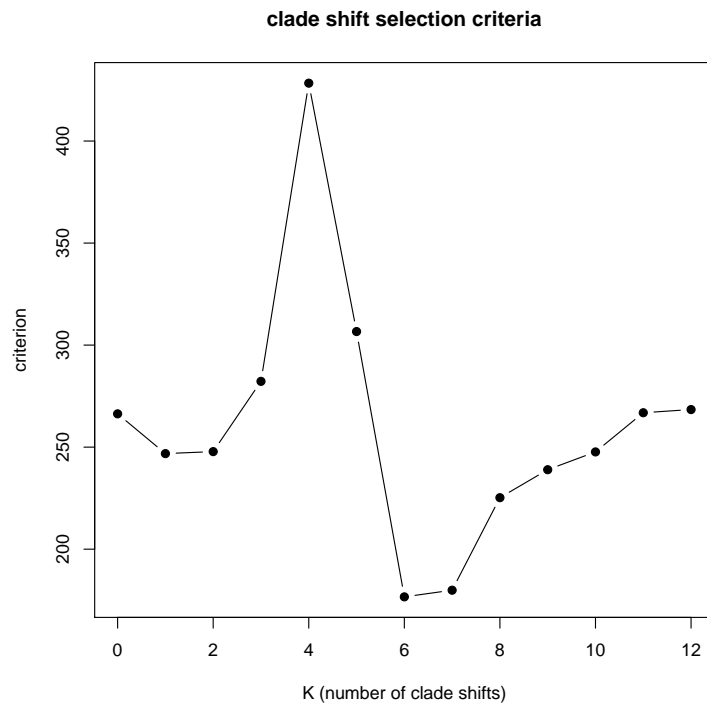


Figure 2 PhylogeneticEM number of selection regime (K) penalized likelihood criteria. $K = 6$ is the best, though $K = 7$ is a good alternative.

Full-size DOI: [10.7717/peerj.8085/fig-2](https://doi.org/10.7717/peerj.8085/fig-2)

Table 2 PhylogeneticEM selection regime optima for $K = 6$ clades. Values are back-transformed to raw percentages (g/100 g).

	Primates	Marsupials	Perissodactyls	Pinnipeds	<i>Mirounga</i>	otariids
Fat	3.18224	6.83607	1.20587	59.78315	44.72456	43.04597
Protein	2.36287	7.95330	2.37018	7.97800	8.36046	10.83272
Sugar	7.18186	7.71032	6.06746	0.76505	0.00007	0.08824

otariids and *Mirounga* were additional deviations from the inferred pinniped optimum. The overall strength of selection or “pull” to these optima is weak ($\alpha = 0.07$) which translates into a phylogenetic half-life of about 9.5 times the total height of the mammalian phylogeny. Thus, the regime shifts can be thought of as instantaneous jumps to novel values embedded in a process that is well approximated by Brownian motion.

Phylogenetic mixed model

Multivariate phylogenetic mixed model prediction of milk composition from ecological and life history traits was largely consistent with results of Skibiél et al., despite removal of some taxa from their dataset, recoding of aquatic habitat, and use of a multivariate technique to incorporate correlations among the milk variables (Table 3). Increased carnivory resulted in a large significant increase in fat concentration, a small non-significant increase in protein concentration, and modest non-significant reduction in sugar concentration. Increased lactation length also caused a large reduction in fat concentration. However,

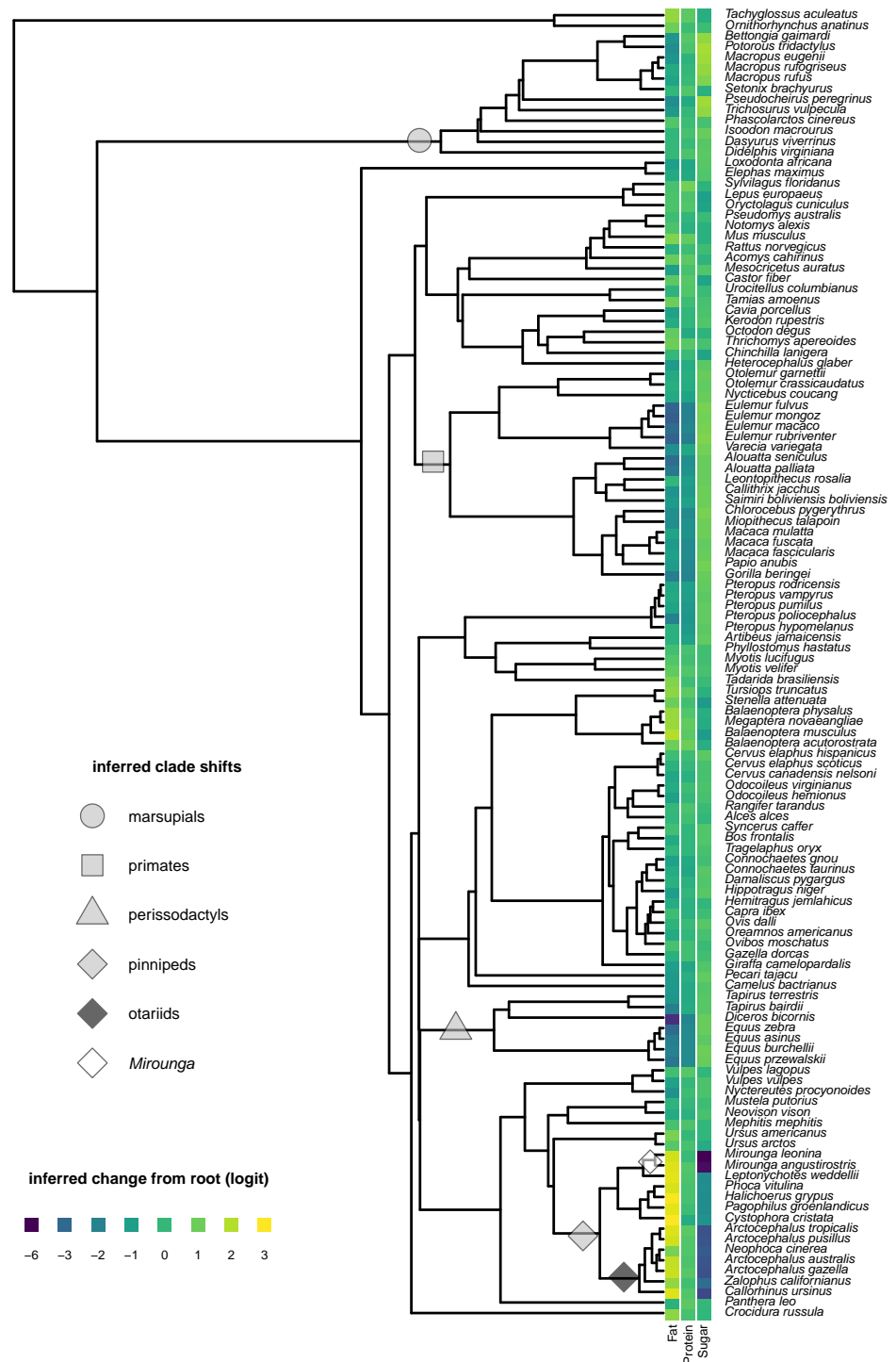


Figure 3 Clade shifts identified by PhylogeneticEM with $K = 6$. Clades with different selection regimes are marked with icons. Colored blocks for each concentration show departures from the root value for each species on the logit scale. Missing data for sugar concentration were imputed from the PhylogeneticEM model for 15 species. See Table 2 for clade names and selection optima (θ).

Full-size DOI: 10.7717/peerj.8085/fig-3

Table 3 Multivariate phylogenetic mixed model regression coefficients, credible interval and MCMC *p*-values. Regression coefficients are for milk concentrations as logit-transformed proportions. Non-intercept *P* < 0.10 are in italics and *P* < 0.05 in bold.

	Fat			Protein			Sugar		
	β	CI	<i>P</i>	β	CI	<i>P</i>	β	CI	<i>P</i>
Intercept	-2.161	(-3.945, -0.317)	0.024	-2.347	(-3.448, -1.252)	0.000	-3.204	(-4.599, -1.805)	0.000
Arid (0/1)	-0.190	(-0.507, 0.141)	0.255	-0.166	(-0.355, 0.022)	<i>0.083</i>	-0.019	(-0.345, 0.310)	0.925
Aquatic (ord.)	0.220	(-0.146, 0.621)	0.260	0.044	(-0.174, 0.268)	0.680	-0.376	(-0.734, -0.014)	0.049
Diet, omnivore	0.194	(-0.197, 0.570)	0.324	-0.029	(-0.248, 0.191)	0.797	0.089	(-0.277, 0.484)	0.648
Diet, carnivore	0.799	(0.226, 1.396)	0.007	0.162	(-0.173, 0.490)	0.326	-0.376	(-0.978, 0.289)	0.245
Female mass	-0.049	(-0.262, 0.175)	0.652	-0.034	(-0.152, 0.090)	0.562	-0.103	(-0.298, 0.103)	0.313
Rep. output	0.150	(-0.178, 0.519)	0.372	0.167	(0.001, 0.344)	<i>0.053</i>	-0.246	(-0.624, 0.136)	0.199
Lactation length	-0.901	(-1.621, -0.220)	0.013	0.105	(-0.269, 0.479)	0.595	0.398	(-0.443, 1.227)	0.338
Precociality (ord.)	0.066	(-0.127, 0.264)	0.500	-0.031	(-0.139, 0.088)	0.598	-0.039	(-0.234, 0.164)	0.703

Table 4 Multivariate phylogenetic mixed model phylogenetic (upper triangle) and residual (lower triangle) correlations with 95% credible intervals. Phylogenetic heritabilities are on the diagonal. Correlations and heritabilities excluding zero from their credible interval are in bold.

	Fat	CI	Protein	CI	Sugar	CI
Fat	0.976	(0.934, 0.994)	0.675	(0.447, 0.815)	-0.750	(-0.920, -0.464)
Protein	0.156	(-0.775, 0.688)	0.997	(0.980, 1.000)	-0.473	(-0.70, -0.119)
Sugar	0.380	(-0.069, 0.994)	0.251	(-0.536, 0.898)	0.872	(0.680, 0.979)

novel patterns also emerged from the reanalysis. There was a nearly significant reduction of protein concentration in arid-adapted mammals, and a nearly significant increase with increasing reproductive output. Finally, increasing aquatic-adapted mammals had significantly reduced milk sugar concentration. There was a non-significant trend for increasing fat concentration with aquatic adaptation.

Phylogenetic heritabilities and correlations reaffirmed a strong phylogenetic signal in milk composition (Table 4). All of the phylogenetic heritabilities were very high (0.872, 0.976, 0.997). The phylogenetic correlations were all moderate to strong indicating coevolution of concentrations. The fat-protein correlation was moderate and positive (0.675), while fat-sugar was strongly negative (-0.750) and protein-sugar was moderately negative (-0.473). Residual correlations were weaker and all included zero within their credible intervals.

DISCUSSION

I used current comparative methods to describe the phylogenetic signal and ecological correlates of milk macronutrient concentrations. The multivariate phylogenetic mixed model results recover those from the previous analysis by *Skibił et al. (2013)*, especially on the influence of diet and relative lactation length on milk fat and protein. However, I also found statistical support for increasingly aquatic mammals having milk that is lower in sugar and higher in fat (cf. *Oftedal & Iverson, 1995*). Other non-significant trends for reduction of protein concentration in arid-adapted mammals and increased protein with higher reproductive output may be biologically meaningful. Finally, phylogenetic

correlations indicate that all three concentrations have coevolved during mammalian evolution history. Differences from the previous report could be due to different coding of predictors (aquatic), more stringent data filtering, and the Bayesian multivariate framework adopted here.

In general, there were few ecological and life history correlates of milk macronutrient composition with detectable statistical associations by these phylogenetic comparative methods. This contrasts with long-standing characterizations of milks as finely attuned to the reproductive ecology and maternal energetics of different mammalian clades (*Oftedal & Iverson, 1995; Ben Shaul, 1963*). The other tools implemented here were more consistent with these characterizations. Clades with distinctive milks (especially primates and pinnipeds) were routinely identified with rank-based tests. This was corroborated with statistical models that identified shifting selection regimes for each clade.

The disconnect between traditional phylogenetic comparative methods (PCMs) as regressions that “control for phylogeny” (*Freckleton, Harvey & Pagel, 2002*) with these distinctive clade and selection regime identification methods is striking. Because mammalian life histories and ecology are fairly conservative, there are few independent cases of mammals evolving similar ecological or life history traits that PCMs rely on to identify associations. For example, true aquatic adaptation is only found in cetaceans, pinnipeds, and the platypus. PCMs are unlikely to identify statistical associations in this case. In contrast, diet categories are more diverse within mammalian clades such that a robust diet-milk fat association can be found by PCMs. While the success of traditional PCMs capturing this association may be due to a strong causal biological mechanism, failure to capture other hypothesized associations with precociality, or aquatic or arid habitats should not be used as evidence against these ecological factors influencing milk composition. Indeed, their authors advocate these newer methods as a “natural history tool” that effectively complement graphical and other descriptive methods (*Uyeda, Zenil-Ferguson & Pennell, 2018*). In particular, they are sensitive to clade-wide adaptations and novel lineage-specific traits that traditional PCMs fail to capture.

The most dramatic of these contrasts in clades was between primates and pinnipeds. These groups are well-known to have strikingly different strategies of infant provisioning and growth rates (*Power & Schullkin, 2016; Skibieli et al., 2013; Hinde & Milligan, 2011; Langer, 2008; Oftedal & Iverson, 1995; Martin, 1984; Ben Shaul, 1963*). Anthropoid primates in particular have the lowest average growth rates among placental mammals and feed infants “on demand” over a prolonged period of lactation where mother and infant are in close proximity allowing for frequent nursing (*Charnov & Berrigan, 1993; Case, 1978*). Pinnipeds grow much more rapidly over shorter lactation periods. Rapid energy transfer to infants is likely facilitated by maternal nutrient stores in blubber and, in some habitats, strongly selected for by the thermoregulatory demands of cold ocean water. Phocid mothers fast during very brief lactation periods while otariid mothers forage and nurse infants during rare visits over longer lactation periods that are short for their body sizes (*Sapriza, 2019; Schulz & Bowen, 2005*). In both pinniped families, milk sugars other than lactose predominate. This allows for their milk to have very little water because lactose draws water into the mammary lumen (*Shennan & Peaker, 2000*).

Perissodactyl milks are similar to those of anthropoid primates. While they share “on demand” feeding with primates, the high lactose and thus water content of equid and rhino milks are argued to be necessary for evaporative cooling of infants in hot environments ([Hinde & Milligan, 2011](#); [Oftedal & Iverson, 1995](#)). Traditional PCMs are ill-suited for handling these convergent milk phenotypes for the different ecological and life history correlates in these clades.

Marsupial milks are distinctive for high sugar concentrations coupled with moderate protein and fat concentrations. Such high sugar concentrations are accomplished by having very little lactose which allows marsupial milks to avoid becoming extremely dilute. In general, marsupials grow at a slow average rate over a long period of lactation (g/day, [Case, 1978](#)), which may allow brain growth to be dissociated from metabolic rate in marsupials ([Weisbecker & Goswami, 2010](#)). The clade shift identified here for marsupials is consistent with many recent studies arguing marsupial development is derived compared to the last common ancestor of marsupials and placentals (e.g., [Werneburg et al., 2016](#)). Comparative milk databases have used composition at pouch emergence (or teat detachment). The comparability of this stage to mid-lactation in placentals is uncertain as another distinctive feature of marsupial milks is their pronounced change in composition over the lactation period ([Oftedal & Iverson, 1995](#)).

The comparative database of milk macronutrient composition, while the best resource available, was inevitably limited. Standardization of data collected across taxa is not always clear and intraspecific variation is not always documented. This will appear as “measurement error” in comparative analysis and will reduce both phylogenetic signal and the strength of regression coefficients in PCMs or other analyses ([Silvestro et al., 2015](#); [Hardy & Pavoine, 2012](#)). For example, prior to eliminating some of the concerning data points in the original database, these outlier species were often assigned their own selection regime. This is likely for elephant seals (*Mirounga*) assignment to their own selection regime. The milk sugar concentrations reported for these two species are very low and were treated as inequalities below detectable levels in original publications. Moreover, while the database is also adequate for many analyses its size will limit the power of recent statistical models designed to measure phylogenetic signal or discriminate among different patterns of selection (e.g., stabilizing, early burst) versus Brownian motion ([Housworth, Martins & Lynch, 2004](#); [Boettiger, Coop & Ralph, 2012](#); [Silvestro et al., 2015](#); [Uyeda & Harmon, 2014](#)).

CONCLUSIONS

Mammalian milks are diverse but often characteristic of certain higher taxa. This makes the ecological and life history correlates of milk composition difficult to identify using traditional phylogenetic comparative methods because those traits are often conservative and clade-specific, too. Primates and pinnipeds have the most outstanding milks according to multiple newly devised tests, with perissodactyls and marsupials as other interesting clades.

ACKNOWLEDGEMENTS

Thanks to Steve Leigh for an introduction to phylogenetic comparative methods and to Katie Hinde and Lauren Newmark for valuable discussions of lactation biology. Comments from two reviewers and the associate editor improved this manuscript.

ADDITIONAL INFORMATION AND DECLARATIONS

Funding

The author received no funding for this work.

Competing Interests

The author declares there are no competing interests.

Author Contributions

- Gregory E. Blomquist conceived and designed the experiments, performed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Raw data and code are available in the [Supplementary Files](#).

Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.8085#supplemental-information>.

REFERENCES

- Adams DC. 2014.** A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution* **68(9)**:2675–2688 DOI [10.1111/evo.12463](https://doi.org/10.1111/evo.12463).
- Adams DC, Collyer ML. 2018.** Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. *Systematic Biology* **67(1)**:14–31 DOI [10.1093/sysbio/syx055](https://doi.org/10.1093/sysbio/syx055).
- Analla M, Jiménez-Gamero I, Muñoz-Serrano A, Serradilla JM, Falagán A. 1996.** Estimation of genetic parameters for milk yield and fat and protein contents of milk from murciano-granadina goats. *Journal of Dairy Science* **79(10)**:1895–1898 DOI [10.3168/jds.S0022-0302\(96\)76558-X](https://doi.org/10.3168/jds.S0022-0302(96)76558-X).
- Bartoszek K, Pienaar J, Mostad P, Andersson S, Hansen TF. 2012.** A phylogenetic comparative method for studying multivariate adaptation. *Journal of Theoretical Biology* **314**:204–215 DOI [10.1016/j.jtbi.2012.08.005](https://doi.org/10.1016/j.jtbi.2012.08.005).
- Bastide P, Ané C, Robin S, Mariadassou M. 2018.** Inference of adaptive shifts for multivariate correlated traits. *Systematic Biology* DOI [10.1093/sysbio/syy005](https://doi.org/10.1093/sysbio/syy005).

- Ben Shaul DM. 1963.** The composition of the milk of wild animals. *The International Zoo Yearbook* **4**(1):333–342 DOI [10.1111/j.1748-1090.1963.tb03691.x](https://doi.org/10.1111/j.1748-1090.1963.tb03691.x).
- Boettiger C, Coop G, Ralph P. 2012.** Is your phylogeny informative? Measuring the power of comparative methods. *Evolution* **66**(7):2240–2251 DOI [10.1111/j.1558-5646.2011.01574.x](https://doi.org/10.1111/j.1558-5646.2011.01574.x).
- Case TJ. 1978.** On the evolution and adaptive significance of postnatal growth rates in the terrestrial vertebrates. *The Quarterly Review of Biology* **53**(3):243–282 DOI [10.1086/410622](https://doi.org/10.1086/410622).
- Charnov EL, Berrigan D. 1993.** Why do female primates have such long lifespans and so few babies? or life in the slow lane. *Evolutionary Anthropology* **1**(6):191–194.
- Chen K, Cheng Y, Berkout O, Lindhiem O. 2017.** Analyzing proportion scores as outcomes for prevention trials: a statistical primer. *Prevention Science* **18**(3):312–321 DOI [10.1007/s11121-016-0643-6](https://doi.org/10.1007/s11121-016-0643-6).
- Clutton-Brock TH. 1991.** *The evolution of parental care*. Princeton: Princeton University Press.
- Cornwell WK, Westoby M, Falster DS, FitzJohn RG, O’Meara BC, Pennell MW, McGlenn DJ, Eastman JM, Moles AT, Reich PB, Tank DC, Wright IJ, Aarssen L, Beaulieu JM, Kooyman RM, Leishman MR, Miller ET, Niinemets U, Oleksyn J, Ordonez A, Royer DL, Smith SA, Stevens PF, Warman L, Wilf P, Zanne AE. 2014.** Functional distinctiveness of major plant lineages. *Journal of Ecology* **102**(2):345–356 DOI [10.1111/1365-2745.12208](https://doi.org/10.1111/1365-2745.12208).
- Drummond AJ, Suchard MA, Xie D, Rambaut A. 2012.** Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Molecular Biology and Evolution* **29**(8):1969–1973 DOI [10.1093/molbev/mss075](https://doi.org/10.1093/molbev/mss075).
- Eisert R, Oftedal OT, Barrell GK. 2013.** Milk composition in the Weddell seal *Leptonychotes weddellii*: evidence for a functional role of milk carbohydrates in pinnipeds. *Physiological and Biochemical Zoology* **86**(2):159–175 DOI [10.1086/669036](https://doi.org/10.1086/669036).
- Faurby S, Svenning J-C. 2015.** A species-level phylogeny of all extant and late Quaternary extinct mammals using a novel heuristic-hierarchical Bayesian approach. *Molecular Phylogenetics and Evolution* **84**:14–26 DOI [10.1016/j.ympev.2014.11.001](https://doi.org/10.1016/j.ympev.2014.11.001).
- Freckleton R, Harvey P, Pagel M. 2002.** Phylogenetic analysis and comparative data: a test and review of evidence. *The American Naturalist* **160**(6):712–726 DOI [10.1086/343873](https://doi.org/10.1086/343873).
- Hadfield JD. 2010.** MCMC methods for multi-response generalized linear mixed models: the MCMCglmm R package. *Journal of Statistical Software* **33**(2):1–22.
- Hardy OJ, Pavoine S. 2012.** Assessing phylogenetic signal with measurement error: a comparison of Mantel tests, Blomberg et al.’s K, and phylogenetic distograms. *Evolution* **66**(8):2614–2621 DOI [10.1111/j.1558-5646.2012.01623.x](https://doi.org/10.1111/j.1558-5646.2012.01623.x).
- Harmon LJ, Losos JB, Jonathan Davies T, Gillespie RG, Gittleman JL, Bryan Jennings W, Kozak KH, McPeck MA, Moreno-Roark F, Near TJ, Purvis A, Ricklefs RE, Schluter D, Schulte II JA, Seehausen O, Sidlauskas BL, Torres-Carvajal O, Weir JT, Mooers AØ. 2010.** Early bursts of body size and shape evolution are rare in comparative data. *Evolution* **64**(8):2385–2396 DOI [10.1111/j.1558-5646.2010.01025.x](https://doi.org/10.1111/j.1558-5646.2010.01025.x).

- Hinde K, Milligan LA. 2011.** Primate milk: proximate mechanisms and ultimate perspectives. *Evolutionary Anthropology* **20**(1):9–23 DOI [10.1002/evan.20289](https://doi.org/10.1002/evan.20289).
- Housworth EA, Martins EP, Lynch M. 2004.** The phylogenetic mixed model. *The American Naturalist* **163**(1):84–96 DOI [10.1086/380570](https://doi.org/10.1086/380570).
- Keck F, Rimet F, Bouchez A, Franc A. 2016.** phyloSignal: an R package to measure, test, and explore the phylogenetic signal. *Ecology and Evolution* **6**(9):2774–2780 DOI [10.1002/ece3.2051](https://doi.org/10.1002/ece3.2051).
- Khabbazian M, Kriebel R, Rohe K, Ané C. 2016.** Fast and accurate detection of evolutionary shifts in Ornstein–Uhlenbeck models. *Methods in Ecology and Evolution* **7**(7):811–824 DOI [10.1111/2041-210X.12534](https://doi.org/10.1111/2041-210X.12534).
- Langer P. 2008.** The phases of maternal investment in eutherian mammals. *Zoology (Jena)* **111**(2):148–162 DOI [10.1016/j.zool.2007.06.007](https://doi.org/10.1016/j.zool.2007.06.007).
- Martin RD. 1984.** Scaling effects and adaptive strategies in mammalian lactation. *Symposia of the Zoological Society of London* **51**:87–117.
- Oftedal OT, Iverson SJ. 1995.** Phylogenetic variation in the gross composition of milks. In: Jensen RG, ed. *Handbook of milk composition*. San Diego: Academic Press, 749–789.
- Othmane MH, De La Fuente LF, Carriedo JA, San Primitivo F. 2002.** Heritability and genetic correlations of test day milk yield and composition, individual laboratory cheese yield, and somatic cell count for dairy ewes. *Journal of Dairy Science* **85**(10):2692–2698 DOI [10.3168/jds.S0022-0302\(02\)74355-5](https://doi.org/10.3168/jds.S0022-0302(02)74355-5).
- Pennell MW, Eastman JM, Slater GJ, Brown JW, Uyeda JC, FitzJohn RG, Alfaro ME, Harmon LJ. 2014.** geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics* **30**(15):2216–2218 DOI [10.1093/bioinformatics/btu181](https://doi.org/10.1093/bioinformatics/btu181).
- Power ML, Schulkin J. 2016.** *Milk: the biology of lactation*. Baltimore: Johns Hopkins University Press.
- Power ML, Verona CE, Ruiz-Miranda C, Oftedal OT. 2008.** The composition of milk from free-living common marmosets (*Callithrix jacchus*) in Brazil. *American Journal of Primatology* **70**(1):78–83 DOI [10.1002/ajp.20459](https://doi.org/10.1002/ajp.20459).
- R Core Team. 2018.** R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing. Available at <http://www.R-project.org/>.
- Raubenheimer D. 2011.** Toward a quantitative nutritional ecology: the right-angled mixture triangle. *Ecological Monographs* **81**(3):407–427 DOI [10.1890/10-1707.1](https://doi.org/10.1890/10-1707.1).
- Roberts SB, Cole TJ, Coward WA. 1985.** Lactational performance in relation to energy intake in the baboon. *American Journal of Clinical Nutrition* **41**(6):1270–1276 DOI [10.1093/ajcn/41.6.1270](https://doi.org/10.1093/ajcn/41.6.1270).
- Sapriza FGR. 2019.** Lactation strategies and milk composition in pinnipeds. In: *Lactation*. London: IntechOpen DOI [10.5772/intechopen.85386](https://doi.org/10.5772/intechopen.85386).
- Schmid M, Wickler F, Maloney KO, Mitchell R, Fenske N, Mayr A. 2013.** Boosted beta regression. *PLOS ONE* **8**(4):e61623 DOI [10.1371/journal.pone.0061623](https://doi.org/10.1371/journal.pone.0061623).
- Schulz TM, Bowen WD. 2005.** The evolution of lactation strategies in pinnipeds: a phylogenetic analysis. *Ecological Monographs* **75**(2):159–177 DOI [10.1890/04-0319](https://doi.org/10.1890/04-0319).

- Shennan DB, Peaker M. 2000.** Transport of milk constituents by the Mammary Gland. *Physiological Reviews* **80**(3):925–951 DOI [10.1152/physrev.2000.80.3.925](https://doi.org/10.1152/physrev.2000.80.3.925).
- Sidlauskas B. 2008.** Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution* **62**(12):3135–3156 DOI [10.1111/j.1558-5646.2008.00519.x](https://doi.org/10.1111/j.1558-5646.2008.00519.x).
- Silvestro D, Kostikova A, Litsios G, Pearman PB, Salamin N. 2015.** Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution* **6**(3):340–346 DOI [10.1111/2041-210X.12337](https://doi.org/10.1111/2041-210X.12337).
- Skibił AL, Downing LM, Orr TJ, Hood WR. 2013.** The evolution of the nutrient composition of mammalian milks. *Journal of Animal Ecology* **82**(6):1254–1264 DOI [10.1111/1365-2656.12095](https://doi.org/10.1111/1365-2656.12095).
- Uyeda JC, Harmon LJ. 2014.** A novel bayesian method for inferring and interpreting the dynamics of adaptive landscapes from phylogenetic comparative data. *Systematic Biology* **63**(6):902–918 DOI [10.1093/sysbio/syu057](https://doi.org/10.1093/sysbio/syu057).
- Uyeda JC, Zenil-Ferguson R, Pennell MW. 2018.** Rethinking phylogenetic comparative methods. *Systematic Biology* **67**(6):1091–1109 DOI [10.1093/sysbio/syy031](https://doi.org/10.1093/sysbio/syy031).
- Vargason T, Howsmon D, McGuinness D, Hahn J. 2017.** On the use of multivariate methods for analysis of data from biological networks. *Processes* **5**(3):36 DOI [10.3390/pr5030036](https://doi.org/10.3390/pr5030036).
- Warton DI, Hui FKC. 2011.** The arcsine is asinine: the analysis of proportions in ecology. *Ecology* **92**(1):3–10 DOI [10.1890/10-0340.1](https://doi.org/10.1890/10-0340.1).
- Weisbecker V, Goswami A. 2010.** Brain size, life history, and metabolism at the marsupial/placental dichotomy. *Proceedings of the National Academy of Sciences of the United States of America* **107**(37):16216–16221 DOI [10.1073/pnas.0906486107](https://doi.org/10.1073/pnas.0906486107).
- Werneburg I, Laurin M, Koyabu D, Sánchez-Villagra MR. 2016.** Evolution of organogenesis and the origin of altriciality in mammals. *Evolution & Development* **18**(4):229–244 DOI [10.1111/ede.12194](https://doi.org/10.1111/ede.12194).
- Whittier CA, Milligan LA, Nutter FB, Cranfield MR, Power ML. 2011.** Proximate composition of milk from free-ranging mountain gorillas (*Gorilla beringei beringei*). *Zoo Biology* **30**(3):308–317 DOI [10.1002/zoo.20363](https://doi.org/10.1002/zoo.20363).
- Xiao Y. 2017.** A fast algorithm for two-dimensional Kolmogorov–Smirnov two sample tests. *Computational Statistics & Data Analysis* **105**:53–58 DOI [10.1016/j.csda.2016.07.014](https://doi.org/10.1016/j.csda.2016.07.014).