Adaptation, phylogeny, and covariance in milk macronutrient composition (#34219)

First submission

Guidance from your Editor

Please submit by **26 Jul 2019** for the benefit of the authors (and your $200 publishing discount).

✓ **Structure and Criteria**
Please read the 'Structure and Criteria' page for general guidance.

▦ **Raw data check**
Review the raw data. Download from the materials page.

🖼 **Image check**
Check that figures and images have not been inappropriately manipulated.

Privacy reminder: If uploading an annotated PDF, remove identifiable information to remain anonymous.

Files	
Download and review all files from the materials page.	4 Figure file(s) 2 Latex file(s) 2 Raw data file(s)

Structure and Criteria

2

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

🗋 You can also annotate this PDF and upload it as part of your review

When ready submit online.

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your guidance page.

BASIC REPORTING

- ✓ Clear, unambiguous, professional English language used throughout.
- ✓ Intro & background to show context. Literature well referenced & relevant.
- ✓ Structure conforms to PeerJ standards, discipline norm, or improved for clarity.
- ✓ Figures are relevant, high quality, well labelled & described.
- ✓ Raw data supplied (see PeerJ policy).

EXPERIMENTAL DESIGN

- ✓ Original primary research within Scope of the journal.
- ✓ Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
- ✓ Rigorous investigation performed to a high technical & ethical standard.
- ✓ Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

- ℹ Impact and novelty not assessed. Negative/inconclusive results accepted. *Meaningful* replication encouraged where rationale & benefit to literature is clearly stated.
- ✓ All underlying data have been provided; they are robust, statistically sound, & controlled.
- ✓ Speculation is welcome, but should be identified as such.
- ✓ Conclusions are well stated, linked to original research question & limited to supporting results.

Standout reviewing tips

The best reviewers use these techniques

Tip	*Example*
Support criticisms with evidence from the text or from other sources	*Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.*
Give specific suggestions on how to improve the manuscript	*Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).*
Comment on language and grammar issues	*The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult.*
Organize by importance of the issues, and number your points	*1. Your most important issue* *2. The next most important item* *3. …* *4. The least important points*
Please provide constructive criticism, and avoid personal opinions	*I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC*
Comment on strengths (as well as weaknesses) of the manuscript	*I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.*

Adaptation, phylogeny, and covariance in milk macronutrient composition

Gregory E Blomquist Corresp. [1]

[1] Department of Anthropology, University of Missouri, Columbia, Missouri, United States

Corresponding Author: Gregory E Blomquist
Email address: blomquistg@missouri.edu

Background. Milk is a complicated chemical mixture often studied through macronutrient concentrations of fat, protein, and sugar. Despite long-standing natural history observations on interspecific diversity in these concentrations, 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 compartive methods to revisit the ecological and life history correlates of milk macronutrient composition and elaborate on the nature of the phylogenetic signal using the phylogentic mixed model, and I 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 less sugary and more fatty milks. Phylogenteic 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 perrisodactyls 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.

Adaptation, Phylogeny and Covariance in Milk Composition

Gregory E. Blomquist[1]

[1]**Department of Anthropology; University of Missouri; Columbia, MO 65211**

Corresponding author:
Gregory E. Blomquist[1]

Email address: blomquistg@missouri.edu

ABSTRACT

Background. Milk is a complicated chemical mixture often studied through macronutrient concentrations of fat, protein, and sugar. Despite long-standing natural history observations on interspecific diversity in these concentrations, 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 compartive methods to revisit the ecological and life history correlates of milk macronutrient composition and elaborate on the nature of the phylogenetic signal using the phylogentic mixed model, ~~and~~ I 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 less sugary and more fatty milks. Phylogenteic 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 perrisodactyls 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.

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 and 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 and Iverson, 1995), but they often have limited 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 (Skibiel et al., 2013). Additional details of the structure of this phylogenetic signal were not addressed. 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 and Collyer, 2018) and techniques for describing phylogenetic signal (Hardy and 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, sugar draws water from blood into the mammary lumen resulting in higher volume but dilute milks. The fattiest milks, as seen seals, have little to no sugar and much less water than terrestrial mammals (Eisert et al., 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 and Iverson, 1995; Hinde and Milligan, 2011). Moreover, the covariation and phylogenetic clustering are self-evident in visualizations either when simply decorating the tips of a phylogeny or in phylomorphospace plots (Sidlauskas, 2008) of the concentrations (Figure 1). The later of these can be particularly illustrative when components are constrained to sum to 100% as in the right-angle mixture model of nutritional geometry (Raubenheimer, 2011).

There are two major goals of this paper. First, I further describe the phylogenetic signal in milk macronutrient concentrations through univariate and multivariate statistics and visualizations. These are 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 use multivariate phylogenetic regression to revisit the results of Skibiel et al. (2013) and test for ecological predictors of milk composition while estimating the phylogenetic and residual covariance among macronutrients.

MATERIALS & METHODS

All data were initially taken from the supplementary material provided by Skibiel 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 and 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 billardierii*, *Perameles gunnii*, *Notomys cervinus*, and *N. mitchelli*). I omitted the very low fat value for *Diceros bicornis* and very low sugar values for *Mirounga* species because they were originally reported as "trace" amounts or as inequalities below detectable levels. For *Leptonychotes weddellii*, new data including a sensitive assay of sugar were available (Eisert et al., 2013). For the three *Papio* species with differing ecological data but the same milk composition, I used only *Papio anubis* which is one the two species from which the milk data were derived (Roberts et al., 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 Skibiel 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 and Svenning, 2015). matched names by hand between the dataset where nomenclature differed using GenBank's taxonomy for prefered names. The trees were trimmed to the taxa 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 and Hui, 2011; Schmid et al., 2013; Chen et al., 2017). All data manipulation and analysis were carried out in R (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 Skibiel 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).

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. 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). These rank-based tests were complemented with quantitative evolutionary modeling implemented in the PhylogeneticEM R package (Bastide et al., 2018). This identifies clades having different stabilizing selective optima (θs) and quantifies strength of selection (a common α) for the "pull" of selection toward those optima.

RESULTS

Graphical Exploration

Correlations between concentrations of fat, protein, and sugar are clear in bivariate plots (Figure 1). The fat-protein relationship is simplest to display because there is less missing data than sugar. There is also obvious phylogenetic clustering of species when color-coded or when phylogenetic relationships are 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 stand out both displays. Most notable are primates and perrisodactlys 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 some marsupials with modest fat, but high protein and sugar. Additional ecological variables also help interpret the scatter. Nearly all the species classified as aquatic by Skibiel et al. have high fat and protein with low sugar concentrations.

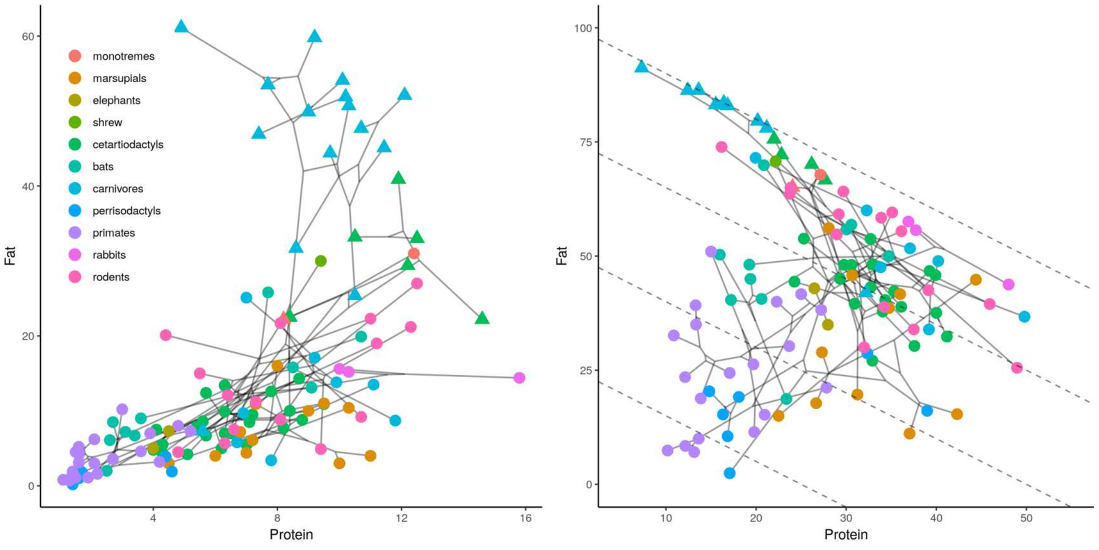

Figure 1. Nutritional geometry of milk as phylomorphospace filling. Raw percentages of protein and fat are plotted on the left while percentages out of a sum of fat, protein, and sugar totalling to 100 are shown on the right. Dotted isoclines of sugar concentration are given every 25% in the right panel. Triangles are aquatic species in both panels.

Distinctive Clades

Non-parametric statistical tests for the distinctiveness of these clades reinforce the graphical patterns (Table 1). The top 5 clades are indicated but typically the distinctiveness drops of steeply from the highest-ranked clade. In the univariate KSI tests, pinnipeds and cetaceans are noted for high fat while perrisodactyls and primates for low fat. The primate genus *Eulemur* is also identified for its further reduction in milk fat. For protein, primates and perrisodactyls stand out for their low values. Primates and pinnipeds are the most distinctive clades for their opposing sugar concentrations, with a marsupial node, perrisodactyls, and *Pteropus* also flagged for high sugar concentrations. Bivariate and trivariate tests generally corroborate these patterns. The only novel clade identified is bovids+cervids which may reflect some ambiguity in the test resolving the nodes around cetaceans.

Evolutionary Modeling

Evolutionary modeling with `PhylogeneticEM` agrees with the distinctiveness of these clades and quantifies the different selective optima for each. The best fitting number of selective regimes was K=6 with K=7 another very good alternative (Figure 2). The regime shifts common to both solutions were pinnipeds, primates, perrisodactyls, 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.

Selective optima (θs) were not dramatically different from the average values seen within each regime (Figure 3, Table 2). The overall strength of selection or "pull" to these optima is weak (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.

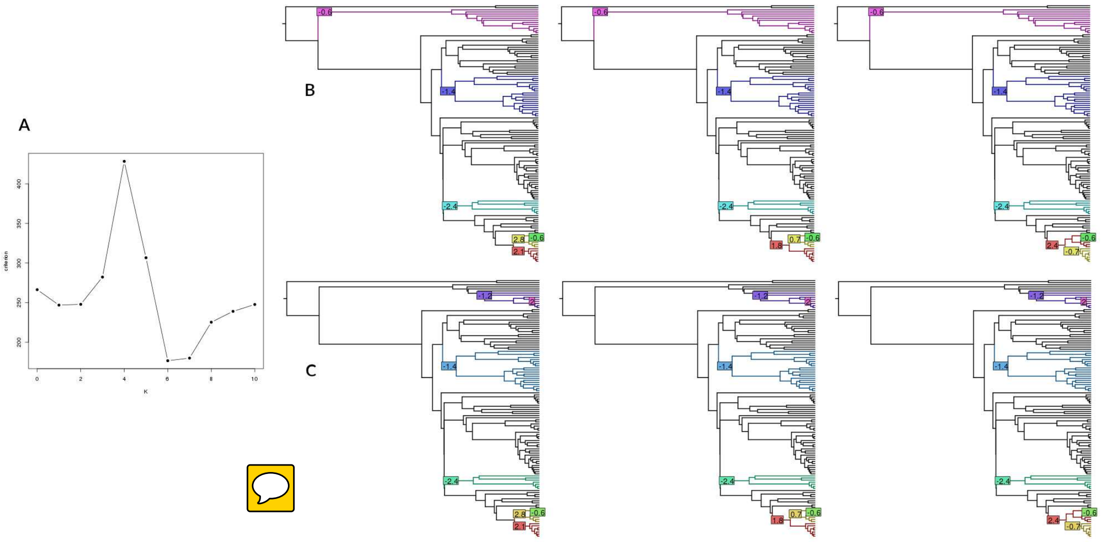

Figure 2. A. PhylogeneticEM K selection criteria. K=6 is the best, though K=7 is a good alternative. B. PhylogeneticEM equivalent clade selection when K=6. Highlighted clades are marsupials, primates, perrisodactyls, and pinniped families or genera. C. PhylogeneticEM equivalent clade selection when K=7. Highlighted clades are the same as B. with the exception of how marsupials are categorized.

Phylogenetic Mixed Model

Multivariate phylogenetic mixed model prediction of milk composition from ecological and life history traits was largely consistent with results of Skibiel 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. Increased carnivory resulted 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, novel patterns also emerge 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,

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 supplementary figure for a phylogeny with nodes labeled.

node	rank	KSI	KSI/max	nodesets
Fat				
pinnipeds	1	3.268	1	pinnipeds; nd27; arctoids; nd21; carnivores; phocids
cetaceans	2	2.221	0.68	cetaceans; whales
perrisodactyls	3	2.16	0.661	perrisodactyls
primates	4	2.272	0.695	primates; anthropoids; catarrhines
Eulemur	5	1.677	0.513	Eulemur
Protein				
primates	1	3.15	1	primates; anthropoids; catarrhines; cercopithecoids
perrisodactyls	2	2.186	0.694	perrisodactyls; Equus; nd46
Pteropus	3	2.178	0.691	Pteropus; bats; nd82
ruminants	4	1.906	0.605	ruminants; bovids+cervids; bovids
Eulemur	5	1.565	0.497	Eulemur
Sugar				
primates	1	3.301	1	primates; anthropoids
pinnipeds	2	2.7	0.818	pinnipeds; nd27; phocids
nd7	3	2.169	0.657	nd7; diprotodonts; nd5; nd4; marsupials; nd8
perrisodactyls	4	1.926	0.584	perrisodactyls; Equus
Pteropus	5	1.947	0.59	Pteropus; bats; nd82
Fat - Protein				
pinnipeds	1	3.396	1	pinnipeds; nd27; arctoids; nd21
primates	2	3.056	0.9	primates; anthropoids; catarrhines; cercopithecoids
perrisodactyls	3	2.48	0.73	perrisodactyls
cetaceans	4	2.237	0.659	cetaceans
Pteropus	5	2.162	0.637	Pteropus; bats
Fat - Sugar				
primates	1	3.433	1	primates; anthropoids
pinnipeds	2	2.7	0.786	pinnipeds; nd27
perrisodactyls	3	2.497	0.727	perrisodactyls
bovids+cervids	4	2.39	0.696	bovids+cervids
nd5	5	2.141	0.624	nd5; nd4; marsupials; diprotodonts; nd7
Protein - Sugar				
primates	1	3.653	1	primates; anthropoids
pinnipeds	2	2.7	0.739	pinnipeds; nd27
marsupials	3	2.395	0.656	marsupials; nd4; nd5
perrisodactyls	4	2.276	0.623	perrisodactyls
Pteropus	5	2.147	0.588	Pteropus; nd82
Fat - Protein - Sugar				
primates	1	3.697	1	primates; anthropoids
bovids+cervids	2	2.846	0.77	bovids+cervids
nd27	3	2.698	0.73	nd27; pinnipeds
perrisodactyls	4	2.447	0.662	perrisodactyls
marsupials	5	2.292	0.62	marsupials; nd4; nd5

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

	primates	marsupials	perrisodactyls	pinnipeds	*Mirounga*	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

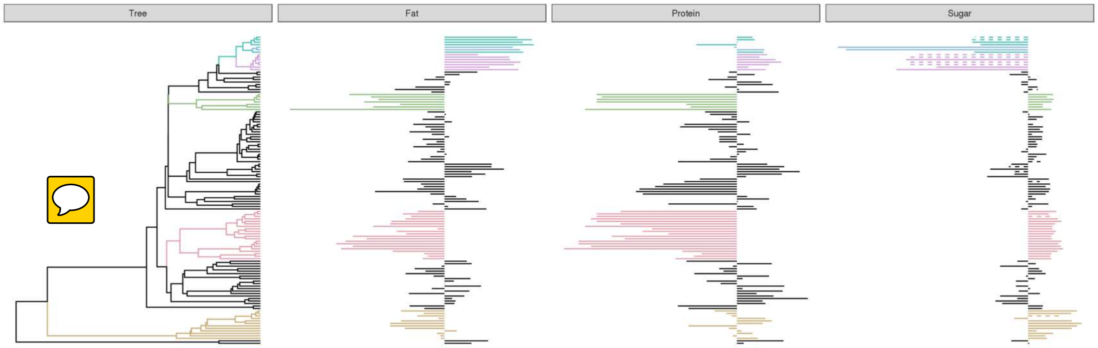

Figure 3. Clade shifts identified by PhylogeneticEM with K=6. Clades with different selection regimes have different colors. Bar plots for each concentration show departures from the root value for each species. Dashed bars were missing data imputed from the PhylogeneticEM model. See Table 2 for clade names and selection optima (θ).

increasing aquatic-adapted mammals have significantly reduced milk sugar concentration. There was a non-significant trend for increasing fat concentration with aquatic adaptation.

Phylogenetic heritabilities and correlations reaffirm a strong phylogenetic signal in milk composition. All of the phylogenetic heritabilities were very high (0.893, 0.976 ,0.997). The phylogenetic correlations were all moderate to strong indicating coevolution of concentrations. The fat-protein correlation was moderate and positive (0.644), while fat-sugar was strongly negative (-0.766) and protein-sugar was moderately negative (-0.425). Residual correlations were weaker and all included zero within their credible intervals.

Table 3. Multivariate phylogenetic mixed model regression coefficients, credible interval and MCMC p-values. Regression coefficients are for milk concentrations as logit-transformed proportions.

	Fat	Protein	Sugar
Intercept	-2.228 (-3.991,-0.488) 0.020	-2.365 (-3.402,-1.286) 0.000	-3.159 (-4.588,-1.792) 0.000
Arid (0/1)	-0.191 (-0.501,0.181) 0.262	-0.162 (-0.338,0.025) 0.084	-0.006 (-0.370,0.309) 0.972
Aquatic (ord.)	0.226 (-0.162,0.581) 0.256	0.048 (-0.159,0.262) 0.696	-0.383 (-0.739,-0.020) 0.042
Diet, omnivore	0.195 (-0.201,0.571) 0.312	-0.025 (-0.250,0.193) 0.820	0.080 (-0.292,0.434) 0.674
Diet, carnivore	0.812 (0.242,1.426) 0.014	0.170 (-0.153,0.491) 0.336	-0.376 (-0.954,0.267) 0.214
Female mass	-0.051 (-0.262,0.158) 0.644	-0.033 (-0.153,0.086) 0.578	-0.102 (-0.299,0.090) 0.306
Rep. output	0.147 (-0.222,0.476) 0.398	0.172 (-0.005,0.343) 0.056	-0.231 (-0.636,0.149) 0.234
Lactation length	-0.894 (-1.595,-0.217) 0.020	0.111 (-0.295,0.445) 0.544	0.409 (-0.456,1.152) 0.338
Precociality (ord.)	0.070 (-0.144,0.271) 0.482	-0.033 (-0.143,0.085) 0.570	-0.042 (-0.253,0.137) 0.698

Table 4. Multivariate phylogenetic mixed model phylogenetic (upper triangle) and residual (lower triangle) correlations with 95% credible intervals. Phylogenetic heritabilities are on the diagonal.

	Fat		Protein		Sugar	
Fat	0.976	(0.930,0.995)	0.644	(0.468,0.825)	-0.776	(-0.903,-0.425)
Protein	0.092	(-0.774,0.681)	0.997	(0.979,1.000)	-0.425	(-0.709,-0.122)
Sugar	0.366	(-0.060,0.994)	0.073	(-0.553,0.864)	0.893	(0.660,0.963)

DISCUSSION

I used current comparative methods to describe the phylogenetic signal and ecological correlates of milk macronutrient concentrations. The multivariate phylogenetic mixed model results were largely consistent with previous analysis by Skibiel et al. (2013), especially on the influence of diet and relative lactation length on milk fat and protein. I also found statistical support for increasingly aquatic mammals having less sugary and more fatty milks (cf. Oftedal and 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 are 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 and 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 et al., 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 adaption 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. Indeed, their authors advocate these newer methods as a "natural history tool" that effectively complement graphical and other descriptive methods (Uyeda et al., 2018). In particular, they are sensitive to clade-wide adaptations and novel lineage-specific traits that traditional PCMs fail to capture.

The comparative database of milk macronutrient composition, while the best resource available, is 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 and Pavoine, 2012). For example, prior to eliminating some of the concerning data points in the original database, these outlier species were often assigned there own selection regime. Moreoever, 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 et al., 2004; Boettiger et al., 2012; Silvestro et al., 2015; Uyeda and 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 perrisodactyls 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.

REFERENCES

Adams, D. C. (2014). A method for assessing phylogenetic least squares models for shape and other high-dimensional multivariate data. *Evolution*, 68(9):2675–2688.

Adams, D. C. and Collyer, M. L. (2018). Multivariate phylogenetic comparative methods: evaluations, comparisons, and recommendations. *Syst Biol*, 67(1):14–31.

Analla, M., Jiménez-Gamero, I., Muñoz-Serrano, A., Serradilla, J. M., and Falagán, A. (1996). Estimation of genetic parameters for milk yield and fat and protein contents of milk from murciano-granadina goats. *J Dairy Sci*, 79(10):1895–1898.

Bastide, P., Ané, C., Robin, S., and Mariadassou, M. (2018). Inference of adaptive shifts for multivariate correlated traits. *Syst Biol*.

Ben Shaul, D. M. (1963). The composition of the milk of wild animals. *Int. Zoo Yb.*, 4(1):333–342.

Boettiger, C., Coop, G., and Ralph, P. (2012). Is your phylogeny informative? Measuring the power of comparative methods. *Evolution*, 66(7):2240–2251.

Chen, K., Cheng, Y., Berkout, O., and Lindhiem, O. (2017). Analyzing proportion scores as outcomes for prevention trials: a statistical primer. *Prev Sci*, 18(3):312–321.

Clutton-Brock, T. H. (1991). *The Evolution of Parental Care*. Princeton University Press, Princeton, N.J.

Cornwell, W. K., Westoby, M., Falster, D. S., FitzJohn, R. G., O'Meara, B. C., Pennell, M. W., McGlinn, D. J., Eastman, J. M., Moles, A. T., Reich, P. B., Tank, D. C., Wright, I. J., Aarssen, L., Beaulieu, J. M., Kooyman, R. M., Leishman, M. R., Miller, E. T., Niinemets, U., Oleksyn, J., Ordonez, A., Royer, D. L., Smith, S. A., Stevens, P. F., Warman, L., Wilf, P., and Zanne, A. E. (2014). Functional distinctiveness of major plant lineages. *J Ecol*, 102(2):345–356.

Drummond, A. J., Suchard, M. A., Xie, D., and Rambaut, A. (2012). Bayesian phylogenetics with BEAUti and the BEAST 1.7. *Mol Biol Evol*, 29(8):1969–1973.

Eisert, R., Oftedal, O. T., and Barrell, G. K. (2013). Milk composition in the weddell seal leptonychotes weddellii: evidence for a functional role of milk carbohydrates in pinnipeds. *Physiol Biochem Zool*, 86(2):159–175.

Faurby, S. and Svenning, J.-C. (2015). A species-level phylogeny of all extant and late quaternary extinct mammals using a novel heuristic-hierarchical bayesian approach. *Mol Phylogenet Evol*, 84:14–26.

Freckleton, R., Harvey, P., and Pagel, M. (2002). Phylogenetic analysis and comparative data: A test and review of evidence. *The American Naturalist*, 160(6):712–726.

Hadfield, J. D. (2010). MCMC methods for multi-response generalized linear mixed models: the MCMCglmm r package. *J Stat Softw*, 33(2):1–22.

Hardy, O. J. and 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.

Hinde, K. and Milligan, L. A. (2011). Primate milk: proximate mechanisms and ultimate perspectives. *Evol Anthropol*, 20(1):9–23.

Housworth, E. A., Martins, E. P., and Lynch, M. (2004). The phylogenetic mixed model. *Am Nat*, 163(1):84–96.

Keck, F., Rimet, F., Bouchez, A., and Franc, A. (2016). phylosignal: an r package to measure, test, and explore the phylogenetic signal. *Ecol Evol*, 6(9):2774–2780.

Langer, P. (2008). The phases of maternal investment in eutherian mammals. *Zoology (Jena)*, 111(2):148–162.

Martin, R. D. (1984). Scaling effects and adaptive strategies in mammalian lactation. *Symposia of the Zoological Society of London*, 51:87–117.

Oftedal, O. T. and Iverson, S. J. (1995). Phylogenetic variation in the gross composition of milks. In Jensen, R. G., editor, *Handbook of Milk Composition*, pages 749–789. Academic Press, San Diego.

Othmane, M. H., De La Fuente, L. F., Carriedo, J. A., and 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. *J Dairy Sci*, 85(10):2692–2698.

Pennell, M. W., Eastman, J. M., Slater, G. J., Brown, J. W., Uyeda, J. C., FitzJohn, R. G., Alfaro, M. E., and Harmon, L. J. (2014). geiger v2.0: an expanded suite of methods for fitting macroevolutionary models to phylogenetic trees. *Bioinformatics*, 30(15):2216–2218.

Power, M. L. and Schulkin, J. (2016). *Milk: The Biology of Lactation*. Johns Hopkins University Press, Baltimore, Maryland.

Power, M. L., Verona, C. E., Ruiz-Miranda, C., and Oftedal, O. T. (2008). The composition of milk from free-living common marmosets (callithrix jacchus) in brazil. *Am J Primatol*, 70(1):78–83.

Raubenheimer, D. (2011). Toward a quantitative nutritional ecology: the right-angled mixture triangle. *Ecological Monographs*, 81(3):407–427.

Roberts, S. B., Cole, T. J., and Coward, W. A. (1985). Lactational performance in relation to energy intake in the baboon. *Am J Clin Nutr*, 41(6):1270–1276.

Schmid, M., Wickler, F., Maloney, K. O., Mitchell, R., Fenske, N., and Mayr, A. (2013). Boosted beta regression. *PLoS ONE*, 8(4):e61623.

Sidlauskas, B. (2008). Continuous and arrested morphological diversification in sister clades of characiform fishes: a phylomorphospace approach. *Evolution*, 62(12):3135–3156.

Silvestro, D., Kostikova, A., Litsios, G., Pearman, P. B., and Salamin, N. (2015). Measurement errors should always be incorporated in phylogenetic comparative analysis. *Methods in Ecology and Evolution*, 6(3):340–346.

Skibiel, A. L., Downing, L. M., Orr, T. J., and Hood, W. R. (2013). The evolution of the nutrient composition of mammalian milks. *J Anim Ecol*, 82(6):1254–1264.

Team, R. C. (2018). R: A language and environment for statistical computing.

Uyeda, J. C. and Harmon, L. J. (2014). A Novel Bayesian Method for Inferring and Interpreting the Dynamics of Adaptive Landscapes from Phylogenetic Comparative Data. *Systematic Biology*, 63(6):902–918.

Uyeda, J. C., Zenil-Ferguson, R., and Pennell, M. W. (2018). Rethinking phylogenetic comparative methods. *Systematic Biology*, 67(6):1091–1109.

Vargason, T., Howsmon, D., McGuinness, D., and Hahn, J. (2017). On the use of multivariate methods for analysis of data from biological networks. *Processes*, 5(3):36.

Warton, D. I. and Hui, F. K. C. (2011). The arcsine is asinine: the analysis of proportions in ecology. *Ecology*, 92(1):3–10.

Whittier, C. A., Milligan, L. A., Nutter, F. B., Cranfield, M. R., and Power, M. L. (2011). Proximate composition of milk from free-ranging mountain gorillas (gorilla beringei beringei). *Zoo Biol*, 30(3):308–317.

Xiao, Y. (2017). A fast algorithm for two-dimensional kolmogorov–smirnov two sample tests. *Comput Stat Data Anal*, 105:53–58.