



HAL
open science

Characterizing and Comparing Phylogenetic Trait Data from Their Normalized Laplacian Spectrum

Eric Lewitus, Leandro Aristide, H el ene Morlon

► **To cite this version:**

Eric Lewitus, Leandro Aristide, H el ene Morlon. Characterizing and Comparing Phylogenetic Trait Data from Their Normalized Laplacian Spectrum. *Systematic Biology*, inPress, 10.1093/sysbio/syz061 . hal-02407709

HAL Id: hal-02407709

<https://hal.science/hal-02407709>

Submitted on 12 Dec 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destin ee au d ep ot et  a la diffusion de documents scientifiques de niveau recherche, publi es ou non,  emanant des  tablissements d'enseignement et de recherche fran ais ou  trangers, des laboratoires publics ou priv es.

Version dated: May 29, 2019

Characterizing and comparing phylogenetic trait data from their normalized Laplacian spectrum

ERIC LEWITUS^{1,2}, LEANDRO ARISTIDE¹, HELENE MORLON¹

¹*Institut de Biologie (IBENS), École Normale Supérieure, Paris, France*

²*Current address: Henry M. Jackson Foundation in support of the US Military HIV Research Program, Walter Reed Army Institute of Research, Silver Spring, MD 20910, USA*

Corresponding author: elewitus@hivresearch.org

Abstract.— The dissection of the mode and tempo of phenotypic evolution is integral to our understanding of global biodiversity. Our ability to infer patterns of phenotypes across phylogenetic clades is essential to how we infer the macroevolutionary processes governing those patterns. Many methods are already available for fitting models of phenotypic evolution to data. However, there is currently no non-parametric comprehensive framework for characterising and comparing patterns of phenotypic evolution. Here we build on a recently introduced approach for using the phylogenetic spectral density profile to compare and characterize patterns of phylogenetic diversification, in order to provide a framework for non-parametric analysis of phylogenetic trait data. We show how to construct the spectral density profile of trait data on a phylogenetic tree from the normalized graph Laplacian. We demonstrate on simulated data the utility of the spectral density profile to successfully cluster phylogenetic trait data into meaningful groups and to characterise the phenotypic patterning within those groups. We furthermore demonstrate how the spectral density profile is a powerful tool for visualising phenotypic space across traits and for assessing whether distinct trait evolution models are distinguishable on a given empirical phylogeny. We illustrate the approach in two empirical datasets: a comprehensive dataset of traits involved in song, plumage and resource-use in tanagers, and a high-dimensional dataset of endocranial landmarks in New World monkeys. Considering the proliferation of morphometric and molecular data collected across the tree of life, we expect this approach will benefit big data analyses requiring a comprehensive and intuitive framework.

Keywords: phylogenetics; macroevolution; traits; primates; tanagers; Laplacian

1 Phylogenetic trait data are essential to understanding the evolution of biodiversity.
2 They have been used to identify adaptive radiations (Harmon et al. 2010), infer stabilizing
3 selection (Hansen 1997; Butler and King 2004), measure the phenotypic effects of species
4 interactions (Drury et al. 2018) and environmental fluctuations (Clavel and Morlon 2017),
5 and generally to estimate the role of the phylogeny in how traits evolve over time
6 (Felsenstein, 1973). They are critical to connecting microevolutionary processes of natural
7 selection to macroevolutionary patterns of phenotypic evolution (Hansen and Martins
8 1996).

9 A wide range of approaches, reflecting the general interest of trait evolution among
10 evolutionary biologists, have been developed to infer the mode and tempo of phenotypic
11 evolution across clades. These include summary statistics that test for the degree of
12 phylogenetic signal in trait data, such as Blomberg's K (Blomberg et al. 2003), and
13 maximum likelihood-based techniques that fit models to phylogenetic trait data and
14 estimate the rate at which traits evolve (see Pennell and Harmon (2013); Manceau et al.
15 (2016); Lewitus (2018) for a review of currently available models). These models rely on
16 the *a priori* formulation of a phenotypic model, which currently can be reduced to whether
17 traits evolve according to a Brownian process along the phylogeny (Felsenstein 1985),
18 towards a trait optimum (Hansen 1997), as an effect of increasing species diversity (Weir
19 and Mursleen 2013) or environmental fluctuations (Clavel and Morlon 2017), or as a result
20 of interspecific interactions (Drury et al. 2016; Manceau et al. 2016). Insofar as they
21 represent a fixed set of biological scenarios, the reliance on parameterized models ultimately
22 limits our ability to characterize the patterns of trait evolution along a phylogeny and
23 compare those patterns between traits independently of pre-defined evolutionary processes.

24 In this paper, we introduce an approach for analysing phylogenetic trait data that
25 requires no assumptions about the underlying generative model. This approach allows for
26 comparisons of the evolutionary histories of traits evolving within a phylogenetic clade and

27 the characterization of trait evolution according to an intuitive graph-theoretical system.
28 Our approach is based on the spectrum of the normalized graph Laplacian, which provides
29 a framework for systematically characterizing and comparing the distribution of trait data
30 across a phylogenetic tree. The normalized graph Laplacian has been successfully utilised
31 in the physical sciences to understand how signal processes are embedded within a graph
32 (Shuman et al. 2013) and has been applied to understanding high-dimensional data
33 produced from, for example, social networks (Rohe et al. 2011), text classification (Apté
34 et al. 1994), and image recognition (Zhang and Hancock 2008). It has also begun to be
35 applied to the biological sciences to aid in big data analysis of metabolic networks (Deyasi
36 et al. 2015) and cancer genomics (Rai et al. 2017). More recently, we introduced an
37 approach for comparing and characterising phylogenies (Lewitus and Morlon 2016a) using
38 the spectral density profile of the graph Laplacian of the distance matrix of a phylogeny,
39 the so-called modified graph Laplacian (MGL), which is able to infer diversification
40 patterns within a phylogeny, as well as directly compare patterns between phylogenies,
41 absent any *a priori* model specification (Lewitus and Morlon 2016b). Together, these
42 applications show the strength of applying the graph Laplacian. However, despite its
43 widespread utility, no such framework has been developed for characterizing and comparing
44 phylogenetic trait data.

45 We first describe how to construct the spectral density profile of the normalized
46 graph Laplacian for phylogenetic trait data and demonstrate how to interpret it in terms of
47 specific properties of phenotypic evolution. We use simulations to show how the profiles
48 relate to conventional metrics of phylogenetic signal and models of trait evolution. We
49 show how to compute the distance between profiles and cluster phylogenetic trait data
50 based on those distances. Finally, we illustrate the utility of this approach for assessing
51 whether distinct trait evolution models are distinguishable using the Cetacean phylogeny.
52 We also illustrate the application of the approach on functional trait data for tanagers

53 (*Thraupidae*) and geometric morphometric data for the endocrania of New World monkeys
54 (*Platyrrhini*). We think that such a non-parametric and comprehensive framework for
55 studying phylogenetic trait diversification will be a valuable complement to existing
56 model-based approaches.

57

MATERIALS AND METHODS

58

Implementation

59 Below, we describe how to use the normalized modified graph Laplacian (nMGL) to
60 construct a spectral density profile for traits (i.e., unidimensional continuous extant tip
61 data) on a phylogeny, how to characterize the profile in terms of evolutionary patterning,
62 and how to compute the distance between profiles. We implemented these functionalities in
63 the R package *RPANDA* freely available on CRAN (Morlon et al. 2016). In the analyses
64 detailed below, phylogenies were simulated using the R package *TESS* (Höhna 2013); trait
65 data for BM, OU and ACDC models were simulated using *mvMORPH* (mvSIM function
66 Clavel et al. (2015)) and for DD and MC models with *RPANDA* (*sim.t.comp* function).
67 Blombergs K was computed using *phytools* (Revell 2012); and MDI was computed using
68 *geiger* (Harmon et al. 2008).

Construction of the Spectral Density Profile for Phylogenetic Trait Data

70 We aim to provide a non-parameteric framework for characterizing and comparing patterns
71 of phylogenetic traits (i.e., tip data) for a given phylogeny. We consider a fully bifurcated
72 tree composed of m terminal branches (Fig. 1A). We note \vec{g} a vector of unidimensional
73 continuous extant trait data associated to this tree. We consider this data as a particular
74 kind of graph, $G = (N, E, w)$, composed of nodes representing extant species, edges
75 delineating the relationships between nodes, and a weight associated to each edge,
76 computed as $w(i, j) = d_{i,j}|g_i - g_j|$ where $d_{i,j}$ is the phylogenetic distance between tips i and
77 j and g_i is the trait value at tip i . Hence, the weight is a combination of phylogenetic and
78 trait distances between two extant species. In Lewitus and Morlon (2016a), the nodes in

79 the graph represent both extant species and internal splitting events in the phylogeny; here
80 we limit the nodes to extant species, as internal splitting events do not have associated
81 trait data. We consider Θ the matrix of weights (Fig. 1B) and D the degree matrix (the
82 diagonal matrix where diagonal element i is computed as $\delta_i = \sum_{k \neq i} w(i, k)$). We construct
83 the normalized modified graph Laplacian (nMGL, see Table 1), defined as
84 $D^{-1/2}(D - \Theta)D^{-1/2}$, which is distinguished from the non-normalized graph Laplacian
85 $(D - \Theta)$ because it is normalized by D . While the normalized version of the graph
86 Laplacian loses some information on the size of the graph compared to the non-normalized
87 version, it is more sensitive to fine-scale features of the graph (Banerjee and Jost 2008).
88 Our approach aims to characterize and compare traits on the same phylogenetic tree
89 (rather than traits between different phylogenetic trees) and so the size of the graph (i.e.,
90 of the tree) is not important. The nMGL is a $m \times m$ positive semi-definite matrix. It
91 therefore has n non-negative eigenvalues, ${}_n\lambda_1 \geq {}_n\lambda_2 \geq \dots \geq {}_n\lambda_m \geq 0$ (throughout, the n
92 subscript preceding symbols highlights that we are considering the normalized graph
93 Laplacian). We convolve them with a Gaussian kernel to ensure a continuous distribution
94 (Banerjee and Jost 2008). The spectral density profile (SDP) of ${}_n\lambda$ from the nMGL,
95 defined as $f(x) = \sum_{i=1} (2\pi\sigma^2)^{-1/2} e^{(-x-{}_n\lambda_i)^2 / 2\sigma^2}$, is plotted as a function of ${}_n\lambda$ as
96 $f^*(x) = \frac{f(x)}{\int f(y)dy}$ (Fig. 1C). Considering the success of previous work showing the capacity
97 of spectral density profiling for differentiating graphs generated by different processes
98 (Banerjee and Jost 2009; Arenas et al. 2006; McGraw and Menzinger 2008; Lewitus and
99 Morlon 2016b), and particularly the framework we recently introduced for characterizing
100 and comparing phylogenies based on their spectral density profiles (Lewitus and Morlon
101 2016a), we hypothesized that the spectral density profile of the nMGL would be a powerful
102 tool for characterizing and comparing trait evolution within phylogenetic clades.

104 The spectrum of ${}_n\lambda$ computed from the nMGL represents primarily global properties of the
105 structure of trait evolution within a phylogenetic clade. Each ${}_n\lambda$ reflects the connectivity
106 (in terms of edge-length) and difference in trait value between one tip and all other tips in
107 a phylogeny. We know from the substantial body of existing work on the normalized graph
108 Laplacian that large ${}_n\lambda$ are characteristic of sparse neighbourhoods typical of highly
109 divergent terminal branches (both in terms of trait value and phylogenetic distance) and
110 small ${}_n\lambda$ are characteristic of denser neighbourhoods typical of barely divergent terminal
111 branches (Chung 1996; Chen et al. 2004). Additionally, for the normalized graph
112 Laplacian, $0 \leq {}_n\lambda \leq 2$ (Bauer and Jost 2009), which in the case of dense matrices (i.e., no
113 zero entries, like Θ) becomes $\sim 1 \leq {}_n\lambda \leq 2$ (Banerjee and Jost 2009). Therefore, as trait
114 differences between closely related tips become smaller, ${}_n\lambda \simeq 1$ accumulate and, as trait
115 differences between closely related tips become larger, ${}_n\lambda > 1$ accumulate. Importantly,
116 trait differences here are relative, so small (or large) trait differences are only small (or
117 large) in regard to the distribution of trait differences across the tree. Also, because the
118 weights used to compute the nMGL are products of phylogenetic and trait distances, it is
119 impossible to separate the relative contribution of each of these distances on the SDP. This
120 is one of the reasons why the nMGL is useful for comparing various trait distributions on a
121 given fixed tree (with fixed phylogenetic distance) and not across different trees.

122 We define three summary statistics computed from the spectrum of ${}_n\lambda$ – the tracer,
123 the fragmenter, and the splitter – that together define the phylogenetic trait space. Traits
124 evolved under different evolutionary scenarios on the same tree occupy different regions of
125 this space (Fig. 2).

126 **The tracer** is the peak height of the SDP, denoted ${}_n\eta$, and computed as the
127 ln-transformed maximum value of $f^*(x)$; it represents the iteration of ${}_n\lambda$ around a single
128 value. Higher tracer values mean smaller differences between closely related tips (low
129 within-clade variance) and larger differences between distantly related tips (high

130 among-clade variance). Therefore, we expect the tracer to be a good measure of
131 phylogenetic signal. In order to test this, we compared the tracer to conventional estimates
132 of phylogenetic signal on trait data simulated on a phylogenetic tree. We simulated a single
133 birth-death tree with 100 tips at 20 million years under constant speciation (0.2) and
134 extinction (0.05) rates (throughout, rates of speciation and extinction are expressed in
135 event per lineage per million year) and simulated 500 trait datasets on that tree under a
136 Brownian motion (BM) model of trait evolution with variance $\sigma^2 = 0.01$ (Cavalli-Sforza
137 and Edwards 1967), an exponentially accelerating (AC) model with rate value $\beta = 1.5$ and
138 $\sigma^2 = 0.01$, an exponentially decelerating (DC) model with rate value $\beta = -0.1$ and
139 $\sigma^2 = 0.01$ (Blomberg et al. 2003; Harmon et al. 2010), and a white-noise model by
140 randomly drawing trait values from a normal distribution (with a mean of zero and
141 standard deviation of one). For each of the three first models, we set the root value at 0.
142 For each dataset, we estimated Blomberg's K , which measures the partitioning of variance
143 using a BM model as reference, where $K > 1$ means close relatives resemble one another
144 more than expected under BM, and $K < 1$ means they resemble one another less
145 (Blomberg et al. 2003), and the morphology disparity index (MDI), which is a measure of
146 the difference between the observed diversity through time curve and that expected under
147 a BM model, where a higher MDI indicates that higher subclade disparity than expected
148 under a BM model (Foote 1997; Harmon et al. 2003; Slater et al. 2010). We fit OLS
149 regression models between ${}_n\eta$ and both Blomberg's K and MDI for the 500 trait datasets.

150 **The fragmenter** is the skewness of the SDP, denoted ${}_n\psi$, and computed as the
151 ln-transformed $\frac{\mu_3}{\mu_2^{3/2}}$, where μ_i is the ordinary i th moment of the distribution; it represents
152 the relative abundance of small and large ${}_n\lambda$. Therefore, as trait space becomes more
153 clustered, irrespective of phylogenetic signal, the proportion of small ${}_n\lambda$ increases and so
154 does the fragmenter. Therefore, we expect the fragmenter to be a good measure of the
155 discreteness of trait space. In order to test whether the fragmenter captures discrete

156 clusters of extant trait data, we simulated a single birth-death tree with 200 tips at 20
157 million years under constant speciation (0.2) and extinction (0.05) rates and simulated 200
158 datasets of discrete trait space under low and high phylogenetic signal. For low
159 phylogenetic signal, we simulated trait data on four macroevolutionary landscapes
160 (Boucher et al. 2017), each defined by a different polynomial function: $V(x) = x^2$,
161 $V(x) = x^4 - 0.5x^2$, $V(x) = x^6 - 0.5x^2$, and $V(x) = x^8 - 0.5x^2$, where the landscape is
162 estimated as $e^{-V(x)}$. Here, an increase in the exponent of the first term generates a more
163 discretized trait distribution (i.e., a deeper well in the macroevolutionary landscape). We
164 set $\sigma^2 = 0.5$, the root value equal to 5, and the trait boundaries at $[0, 10]$. We plotted the
165 landscapes as defined by the polynomial functions, histograms of the trait data for each
166 landscape as realized in the simulations, and the spectral density profiles of each dataset.
167 For high phylogenetic signal, we simulated trait data on the same birth-death tree under a
168 DC model with rate value $\beta = -0.6, -0.3, 0$ and $\sigma^2 = 0.1$, where more negative values of β
169 indicate more decelerated rates (Blomberg et al. 2003; Harmon et al. 2010). This generated
170 trait data distributed in discrete monophyletic clusters across the tree. For the low and
171 high phylogenetic signal datasets, we computed the fragmenter and compared values as a
172 function of macroevolutionary landscape and of β .

173 **The splitter** is the principal $n\lambda$, denoted $(n\lambda^*)$; it is diagnostic of the disjointedness
174 of a graph, where larger splitter values imply a more bipartite structure (Banerjee and Jost
175 2008; Bauer and Jost 2009). In macroevolutionary terms, as traits become increasingly
176 bimodally distributed with a strong phylogenetic signal, the splitter increases. As $n\lambda \leq 2$
177 for the nMGL, the splitter $\simeq 2$ when a clade is composed of two phylogenetically distinct
178 subclades with different mean trait values. To assess the relationship between the spectral
179 density profile and differences in mean trait values on a phylogeny, we simulated a single
180 birth-death tree with constant speciation (0.2) and extinction (0.05) rates with 200 tips at
181 20 million years. We then simulated BM models ($\sigma = 0.01$) with q differences in mean trait

182 values for $q = 0 - 4$ by defining different mean trait values for $q + 1$ monophyletic sets of
183 tips, where the mean trait value for q_0 was randomly drawn from a normal distribution
184 with a mean value between $0 - 1$ (and standard deviation of one) and subsequent mean
185 trait values were defined as two-times the previous mean. We then compared $n\lambda^*$ for each
186 set. The value of the splitter is expected to correlate with the disjointedness of the graph,
187 where higher values indicate the nMGL is more bipartite and so can be segregated into two
188 monophyletic groups with distinct mean trait values (Bauer and Jost 2009). To test
189 whether there were, in fact, two monophyletic clusters, we used k-means clustering (for
190 $k=2$) on the nMGL of the phylogenetic trait data. We then calculated the average
191 branch-length distance between tips in cluster 1 and tips in cluster 2. For phylogenetic
192 trait data that can be separated into two monophyletic clusters, the average between-group
193 distance will equal two times the crown age of the tree. We present this as a heuristic test
194 of the monophyly of trait values when the splitter $\simeq 2$. To test the effect of phylogenetic
195 signal on the splitter value, we simulated 10 trait datasets with one difference in mean trait
196 values on a 100-tip constant-rate birth-death tree as above. We then randomized the
197 distribution of tip data within each cluster 100 times and compared the resulting splitter
198 value for the randomized trees against the original splitter value. We compared the splitter
199 values for the two-cluster BM datasets and the randomized two-cluster datasets to 100
200 datasets simulated under a simple BM process (with no clusters and $\sigma = 0.01$).

201 To test the effect of erroneous data on the nMGL, we simulated trait data under a
202 BM process on a 100-tip constant-rate birth-death tree ($\sigma = 0.01$). We tested both the
203 effect of increasing the amount of error and increasing the number of tips with error. For
204 the former, we introduced error on 10% of randomly drawn tips as a sampling variance
205 equal to n times the standard error for $n = 1, 2, 3$. For the latter, we introduced error on
206 20, 30, 40, 100% of tips as a sampling variance equal to the standard error. We simulated
207 100 datasets for each scenario. We compared the resulting splitter, tracer, and fragmenter

208 values to BM datasets ($\sigma = 0.01$) and ACDC datasets ($\beta = -1.1, \sigma = 0.01$) simulated on
209 the same tree and with no introduced error.

210 *Clustering nMGLs from Their Spectral Density Profiles*

211 To demonstrate whether we can distinguish phylogenetic trait data simulated under trait
212 models we know are distinguishable, we clustered nMGLs constructed for trait data on the
213 same phylogeny under different trait models. To cluster nMGLs, we computed the
214 Jensen-Shannon distance between spectral density profiles. The Jensen-Shannon distance is
215 defined as

$$\Delta(\Lambda_1, \Lambda_2) = \sqrt{\frac{1}{2}KL(f_1^*, f^*) + \frac{1}{2}KL(f_2^*, f^*)} \quad (1)$$

216 where f_1^* and f_2^* are spectral densities for profiles 1 and 2, $f^* = \frac{1}{2}(f_1^* + f_2^*)$, and KL is the
217 Kullback-Leibler divergence measure for the probability distribution (Endres and Schindelin
218 2003). We then cluster the matrix of Jensen-Shannon distances for each profile pair using
219 hierarchical clustering with bootstrap resampling and k-medoids clustering using optimal
220 silhouette width, $s(i)$, which is a measure of the between/within-variance of each datapoint
221 i assigned to a cluster; data are typically considered structured at $\bar{s} > 0.51$ (Szekely and
222 Rizzo 2005; Reynolds et al. 2006). In each case, the number of clusters is not set *a priori*.

223 We tested the efficacy of clustering on profiles using trait datasets simulated on
224 birth-death trees. We simulated a total of 1500 trait datasets under a BM model of trait
225 evolution with variance $\sigma^2 = 0.01$ (Cavalli-Sforza and Edwards 1967), an exponentially AC
226 model with rate value $\beta = 1.5$ and $\sigma^2 = 0.01$, and an exponentially DC model with rate
227 value $\beta = -0.1$ and $\sigma^2 = 0.01$ (Blomberg et al. 2003; Harmon et al. 2010). For each model,
228 we set the root value at 0. We visualised this clustering by plotting the profiles in a
229 multidimensional space defined by ${}_n\lambda^*$, ${}_n\psi$, and ${}_n\eta$.

230 We tested the ability of the spectral density profile to find meaningful clusters of
231 trait models on different tree shapes and sizes. To test for the effect of tree shape, we

232 simulated trait datasets on 200-tip birth-death trees (with a max age of 20Ma) with
233 constant speciation (0.2) and extinction (0.02) rates, with decreasing speciation
234 ($0.1 * e^{-0.2t}$) and constant extinction (0.02) rates, and with increasing speciation ($0.1 * e^{0.1t}$)
235 and constant extinction (0.05) rates. We conducted analyses on identical trees without
236 pruning extinct lineages, resulting therefore in non-ultrametric trees, to test whether the
237 profiles of different models were more distinguishable on a non-ultrametric tree compared
238 to an ultrametric tree, as is expected from likelihood-based approaches (Cooper et al.
239 2015). To test for the effect of tree size, we simulated 6000 trait datasets under the same
240 BM, AC, and DC trait model parameters on birth-death trees with constant speciation
241 (0.2) and extinction (0.05) rates with 20, 50, 100, 200, and 500 tips (with a max age of
242 20Ma). As above, phylogenies were simulated using the *R* package *TESS* (Höhna 2013)
243 and trait data were simulated using *mvMORPH* (Clavel et al. 2015).

244

Applications

245 To illustrate our approach, we demonstrate three applications. First, we used the Cetacean
246 phylogeny (87 spp.) (Steeiman et al. 2009) to illustrate how our approach can be used to
247 assess the distinguishability of different trait evolution models in a particular clade. We
248 simulated six trait models under a range of parameter values on the Cetacean phylogeny:
249 BM with $\sigma^2 = 0.1 - 5$; Ornstein-Uhlenbeck (OU) with strength of pull towards an optimum
250 $\alpha = 1 - 20$ and $\sigma^2 = 0.1$; exponential diversity-dependence (DD) with slope parameter
251 $r = -1.1 - -0.1$ and $\sigma^2 = 0.1$; AC with rate value $\beta = 1.1 - 1.5$ and $\sigma^2 = 0.1$; DC with
252 rate value $\beta = -0.5 - -0.1$ and $\sigma^2 = 0.1$; and matching competition (MC) with the
253 strength of competition $S = -1.1 - -0.1$ and $\sigma^2 = 0.1$. For each model, we simulated 500
254 datasets with the root value set to zero. For all datasets, we computed the spectral density
255 profile and clustered them using hierarchical and k-medoid clustering. Second, we used a
256 tanager phylogeny (350 spp.) (*Thraupidae*) with 27 phylogenetically corrected principal

257 component traits (pPC traits) spanning traits related to song, plumage, and resource-use
258 taken from Drury et al. (2018). Ideally, we would have used non-phylogenetically corrected
259 PCs but these were not available. We computed the spectral density profiles for the pPC
260 traits and clustered them using hierarchical and k-medoid clustering and computed their
261 spectral density profile summary statistics. Finally, we used a geometric morphometrics
262 dataset consisting of 399 three-dimensional Procrustes superimposed landmark coordinates
263 describing the external brain shape of 48 species of New World monkeys (*Platyrrhini*)
264 (Aristide et al. 2016). For each landmark, we computed the Euclidean distance between
265 the landmark and the clade mean for that landmark, in order to reduce the dimensionality
266 of the data. We refer to these distances simply as landmarks. We computed the spectral
267 density profile for each of the 399 landmarks and clustered them using hierarchical and
268 k-medoid clustering and plotted their spectral density profile summary statistics in
269 multidimensional space. In order to test how much information was lost in this
270 dimensionality reduction (Monteiro et al. 2000; Uyeda et al. 2015), we also clustered the
271 profiles computed separately for the coordinates along each axis. Even though these axes
272 may not necessarily be aligned with the most biologically meaningful directions of
273 variation, it is a straightforward and convenient way of analyzing the data.

274

RESULTS

275 *Interpreting the Spectral Density Profile for Phylogenetic Trait Data*

276 The shape of the spectral density profile of the nMGL reveals many aspects characteristic
277 of the underlying evolution of a trait within a phylogenetic clade. Specifically, the tracer
278 (peak height, ${}_n\eta$), the fragmenter (skewness, ${}_n\psi$) and the splitter (principal ${}_n\lambda$, ${}_n\lambda^*$), of the
279 profile may be interpreted in terms of the evolutionary history of the trait (Fig. 2).

280 The tracer summary statistic represents the peak height of the spectral density
281 profile. In macroevolutionary terms, this is indicative of the phylogenetic signal of a trait,
282 where larger ${}_n\eta$ indicate more phylogenetic signal (Fig. 3A-C). We show that the tracer is
283 strongly correlated with conventional summary statistics of phylogenetic signal, with ${}_n\eta$
284 increasing with Blomberg's K ($y = 3.44 - 4.13x + 1.36x^2$, $R^2 = 0.96$, $P < 0.01$) and
285 decreasing with MDI ($y = -2.65 + 2.23x - 0.34x^2$, $R^2 = 0.93$, $P < 0.01$) (Fig. 3D).
286 White-noise models fall at the lowest end of tracer values, converging with AC models
287 simulated with $\beta = 1.5$ in terms of tracer values (Fig. 3).

288 The fragmenter summary statistic represents the relative abundance of small *versus*
289 large ${}_n\lambda$. In macroevolutionary terms, larger ${}_n\psi$ indicate a more discrete distribution of
290 trait means in trait space. We show that for trait data simulated on increasingly
291 discretized macroevolutionary landscapes, spectral density profiles have correspondingly
292 higher fragmenter values (Fig. 4A-C). We also show for trait data simulated with DC
293 models with an increasingly negative rate parameter, β , which produce increasingly
294 discretized trait space, that spectral density profiles have correspondingly higher
295 fragmenter values (Fig. 4D-F). Notably, the discrete clusters of mean trait values generated
296 by macroevolutionary landscapes are generally not monophyletic, whereas those generated
297 by DC models are monophyletic.

298 The splitter summary statistic, which is the principal $n\lambda$ computed from the nMGL,
299 is diagnostic of the bipartiteness of the nMGL. Specifically, it is indicative of how easily the
300 graph can be disjointed into two components. We show that splitter values increase
301 (approaches 2) as the number of monophyletic groups with different trait means
302 approaches two (Fig. 5A-C). When groups are defined using k-means clustering (with
303 $k = 2$) on the nMGL, the average phylogenetic distance between groups approaches two
304 times the crown age of the phylogeny when there are two monophyletic groups,
305 demonstrating that clustering on the nMGL allows recovering these two groups (Fig. 5D).
306 Splitter values obtained from the randomized datasets are similar to those obtained from
307 the original datasets, suggesting that phylogenetic signal has little effect on splitter values
308 (Supplemental Fig. 1).

309 Importantly, the fragmenter and tracer values are sensitive to the introduction of
310 erroneous data, although not dramatically (Supplemental Fig. 2). When a considerable
311 amount of sampling variance (equal to three times the standard error) is introduced on
312 10% of tips, fragmenter and tracer values decrease only slightly. The impact of erroneous
313 data only becomes appreciable when it is introduced on a large proportion of tips ($\geq 30\%$).
314 However, it is only when 100% of tips are affected by erroneous data that the inference of
315 fragmenter and tracer values begins to approach that of AC models ($\beta = 1.5$), which shows
316 that the nMGL is in large robust to error-prone data.

317 *Clustering Models of Phylogenetic Trait Data*

318 For the traits simulated on the constant-rate birth-death trees under the three different
319 trait evolution models, we found that the spectral density profiles were optimally clustered
320 into three groups (bootstrap probability > 0.95) (Fig. 6A). Separate clusters could be
321 overwhelmingly ($> 95\%$) assigned to AC, BM, and DC models with an average silhouette
322 width= 0.6. The DC cluster is considerably farther from the AC and BM clusters than the

323 AC and BM clusters are from each other, based on Euclidean distance. Trait models
324 simulated on increasing-rate and decreasing-rate trees show slightly different abilities to
325 cluster trait models using spectral density profiles. They also show different configurations
326 of profiles in multidimensional space, although this is expected because the nMGL is
327 computed using the phylogenetic distance matrix, which is sensitive to tree shape. For the
328 increasing-rate tree, the profiles were optimally clustered into three groups (bootstrap
329 probability > 0.95), each of which could be exclusively assigned to either AC, BM, or DC
330 models with an average silhouette width= 0.79 (Fig. 6B). Similarly to the constant-rate
331 tree, the DC cluster is considerably farther from the AC and BM than the AC and BM are
332 from each other. For the decreasing-rate tree, we found two significant clusters (bootstrap
333 probability > 0.95), one of which can be exclusively assigned to DC models and another
334 that is a hodgepodge of AC and BM trait models with an average silhouette width= 0.55
335 (Fig. 6C). When plotted in multidimensional space, the AC and BM models simulated on
336 the decreasing-rate tree occupy the same region and are therefore indistinguishable based
337 on their spectral density profile summary statistics for this tree.

338 For the constant-rate non-ultrametric tree, trait models are also distinguishable
339 based on hierarchical and k-medoids clustering (Supplemental Fig. 3A). The average
340 silhouette width for clusters of traits on the non-ultrametric tree is 0.82, compared to only
341 0.6 on the ultrametric tree (Supplemental Fig. 3B), which demonstrates that the trait
342 models are more distinguishable on the non-ultrametric tree. We similarly found trait
343 models to be distinguishable on increasing-rate (Supplemental Fig. 3C) and decreasing-rate
344 (Supplemental Fig. 3D) non-ultrametric trees.

345 We estimated the effects of tree size on spectral density profile summary statistics.
346 Fragmenter and tracer values increase with tree size, while splitter values decrease with
347 tree size (Supplemental Fig. 4A). At 20 tips, the profiles of AC, BM, and DC models
348 occupy the same phylogenetic trait space, but at 50 tips the models are distinguishable

349 (Supplemental Fig. 4B). While the nMGL loses some information on the size of the graph
350 compared to the non-normalized version, clearly there is still some effect of size. This is
351 likely because size and shape are integrated in phylogenies.

352 *Applications*

353 Traditionally, likelihood-based models are fit to phylogenetic trait data and the model
354 showing the best support is inferred as the generative one. Oftentimes the difference in
355 support between models is small and therefore finding traits with similar evolutionary
356 histories or comparing those evolutionary histories can be difficult. The ability of our
357 approach to directly compare the spectral density profiles of the nMGLs of different traits
358 on the same tree allows us to find clusters of traits with similar evolutionary histories and
359 then compare those histories in a multidimensional space defined by interpretable
360 parameters without needing to qualify differences based on estimated likelihoods.

361 By simulating datasets under different trait models on the Cetacean phylogeny, we
362 are able to visualize how distinguishable these models are from one another under different
363 parameter values (Fig. 7). When all parameter values are taken together, we are not able
364 to clearly distinguish between all models using hierarchical clustering (Fig. 7). While
365 under certain parameter values each model occupies its own space, there is nonetheless
366 overlap for parameter values, suggesting that, for the Cetacean phylogeny, trait evolution
367 under different phenotypic models are quite similar. Particularly similar models are DC
368 and DD, although these diverge in phylogenetic trait space for large parameter values; and
369 OU and AC, but unsurprisingly, because these two models are algebraically identical on
370 ultrametric trees (Uyeda et al. 2015).

371 We clustered the spectral density profiles for 27 pPC traits in the tanager
372 phylogeny. We identified two clusters using hierarchical clustering (bootstrap
373 probability= 0.96) (Fig. 8A); and found the same two clusters using k-medoid clustering,

374 where the inferred axes explained 69% of variance among the spectral density profiles (Fig.
375 8B). Cluster 1 was comprised of 10 plumage traits, 6 resource-use traits, and 1 song trait,
376 whereas Cluster 2 was comprised of 9 song traits and 1 resource-use trait, suggesting
377 different evolutionary histories for different types of traits (Fig. 8B). Cluster 1 showed
378 significantly higher ($T > 2.8$, $P < 0.01$) splitter, fragmenter, and tracer values compared to
379 Cluster 2 (Fig. 8C). This suggests that that plumage and resource-use traits have a
380 stronger phylogenetic signal and evolve into more discrete trait space, indicative of
381 monophyletic clusters of traits. While the plumage and resource-use cluster have a
382 significantly higher splitter value than the song cluster, both have low splitter values (i.e.,
383 $\ll 2$) and therefore little evidence of bipartiteness.

384 The landmark data for New World monkeys clustered into three groups according to
385 k-medoid clustering, with a minimum average silhouette width of 0.51, and according to
386 hierarchical clustering (bootstrap probability > 0.9) (Fig. 8D). Cluster 1 showed
387 significantly higher ($T > 1.96$, $P \leq 0.05$) fragmenter and tracer values, suggesting a
388 stronger phylogenetic signal and evolution into more discrete trait space compared to the
389 other clusters of landmarks (Fig. 8E). Cluster 2 showed significantly lower fragmenter and
390 tracer values than the other clusters, suggesting it evolves with little phylogenetic signal
391 into a more uniform trait space. Cluster 3 showed intermediary fragmenter and tracer
392 values, but significantly higher splitter values, indicative of more bipartiteness. The
393 relationship between fragmenter and tracer, which is indicative of the amount of
394 convergence in trait space, shows that tracer values increase as a function of fragmenter
395 values faster based on a one-sided t-test ($P < 0.05$) in cluster 1 compared to clusters 2 and
396 3, suggesting lower levels of convergence in cluster 1. Interestingly, the three clusters
397 broadly correspond to well-defined brain regions (Fig. 8F). Specifically, cluster 1 comprises
398 landmarks mostly located on the parietal, cerebellar, and the anterior portion of the frontal
399 region, cluster 2 landmarks mainly correspond to the temporal, occipital, and stem regions,

400 and cluster 3 comprises landmarks on the posterior and ventral areas of the frontal region
401 and parts of the temporal. These results suggest that different brain regions evolved with
402 different evolutionary histories. When we clustered the landmark data along each axis
403 separately, treating the coordinates as tip data, we identified the same three clusters along
404 each axis according to k-medoid clustering, with a minimum average silhouette width for
405 each cluster of 0.46, and according to hierarchical clustering (bootstrap probability > 0.9).

406

DISCUSSION

407 We recently introduced an approach for characterizing and comparing phylogenies using
408 the spectrum of the graph Laplacian (Lewitus and Morlon 2016a). Here, we have extended
409 this approach to analyse the evolution of traits within phylogenetic clades. We have shown
410 how to compute the spectral density profile of the nMGL for phylogenies with associated
411 trait data and demonstrated how to use these profiles to characterize and compare trait
412 data within a phylogenetic clade. This provides a broad, scalable framework for
413 characterizing the distribution of traits within a phylogenetic clade without classifying
414 those distributions according to pre-defined models of phenotypic evolution. This
415 non-parametric approach therefore provides a complement to existing model-based
416 approaches to studying phenotypic evolution.

417 Because the spectral density profile of the nMGL is computed directly from the
418 phylogeny and trait data, it provides a comprehensive rendering of the structure of trait
419 evolution across a phylogenetic clade. Consequently, the spectral density profiles of
420 different traits on a phylogenetic tree, unlike likelihood values or summaries of phylogenetic
421 signal, can be clustered absent any *a priori* model specification. We show that this is
422 successful in distinguishing between phylogenetic trait data generated under different
423 macroevolutionary processes and sensitive to the parameter values under which those
424 processes are generated. Hence, in the same way that spectral density profiles have been
425 used for identifying principal patterns of diversification in vertebrates (Lewitus and Morlon
426 2016b), they can be used for identifying principal patterns of phenotypic evolution across
427 multiple traits within clades, as we have illustrated here with two empirical datasets.
428 Spectral density profiles can also be used to quickly evaluate how distinguishable different
429 trait evolutionary processes are, as we have illustrated here on the Cetacean phylogeny.
430 This can be very useful when developing new models, to make sure they will be

431 distinguishable before putting all the effort into develop likelihood-based inferences for
432 these models. Similarly, although it is impossible to separate the relative contribution of
433 phylogenetic and trait distances on the SDP, it is possible to compare SDPs for the same
434 trait data across multiple versions of a phylogeny (e.g., a posterior distribution of trees
435 generated by Bayesian inference) and thus estimate the effect of tree construction on
436 inferences of trait evolution. We can also anticipate that spectral density profiles will be
437 useful to compute the distance between simulated and real data in Approximate Bayesian
438 Computation approaches (Beaumont 2010) for fitting models of phenotypic evolution that
439 are not amenable to likelihood computation (e.g., Clarke et al. (2017)). Although currently
440 limited to the analysis of continuous traits, an extension of the nMGL to incorporate
441 discrete binary traits would be straightforward: the trait distance between species would
442 be 0 or 1 if pairs have the same or a different trait, respectively. Existing work on signed
443 graph Laplacians (Kunegis et al. 2010), which attach a positive, negative, or neutral sign to
444 each edge, already show the potential for using graph Laplacians to explore graphs with
445 associated discrete values. As there are a wide range of discrete traits that are the focus of
446 many macroevolutionary questions (e.g., Beaulieu et al. (2013)). we think development of
447 the nMGL for the analysis of discrete trait evolution is an important direction for future
448 work to move in.

449 When reduced to their constituent properties (i.e., splitter, fragmenter, and tracer
450 values), spectral density profiles are useful in summarizing the structure of phylogenetic
451 trait data and in visualizing differences between them. The tracer is a measure of
452 phylogenetic signal and correlates well with conventional summary statistics. Blomberg's
453 K , as a measure of the partitioning of within- *versus* among-clade variance, resembles what
454 the tracer is measuring, which is the iteration of $n\lambda$ around a single value. When within-clade
455 variance is low and among-clade variance is high, then the majority of $n\lambda$ will have a
456 similar value, the tracer will be high and so will Blomberg's K . The fragmenter measures

457 the discreteness of phenotypic space. Higher fragmenter values indicate that trait values
458 are distributed in more discrete groups in phenotypic space, as would occur under an early
459 burst model of trait diversification or high levels of convergence to multiple optima. The
460 relationship between the tracer and fragmenter gives some indication as to whether
461 convergence has likely occurred: the ratio of tracer to fragmenter will be higher if the
462 discretization of trait values in phenotypic space shows a strong phylogenetic signal (i.e., in
463 the absence of convergence). We show, for example, that a two-peak macroevolutionary
464 landscape results in high fragmenter values, but relatively lower tracer values than occur
465 under a DC model, indicative of the high level of phenotypic convergence in the
466 macroevolutionary landscape model and low level of phenotypic convergence in the DC
467 model. Of course, we cannot assign a threshold value for convergence, above which the
468 tracer to fragmenter ratio conclusively evinces phenotypic convergence. However, for a
469 given analysis of different trait data on a tree, we recommend comparing tracer to
470 fragmenter ratios between analyses, in order to deduce the comparative levels of
471 convergence between datasets. Finally, the splitter of the nMGL is diagnostic of the
472 bipartiteness of the graph and therefore, in terms of phylogenetic trait data, higher splitter
473 values indicate a bimodal distribution of trait values with high phylogenetic signal.

474 We analyse a previously published dataset on pPCs for tanagers (Drury et al. 2018)
475 to show the usefulness of clustering phylogenetic trait data to identify and characterize
476 traits with similar evolutionary histories among a set. Our result, that the evolution of
477 song-related traits is distinct from that of plumage- and resource-use-related traits, is
478 consistent with those found in Drury et al. (2018) for species that are year-round territorial
479 and/or found in dense habitats. The high tracer and fragmenter values in plumage and
480 resource-use traits suggests the discretized trait space of these traits possesses a high
481 phylogenetic signal, while the low tracer and fragmenter values in song traits suggests low
482 phylogenetic signal and non-discretized trait space.

483 We analyse a dataset of 399 landmarks on the endocrania of 48 species of New
484 World monkeys. We show that these landmarks cluster into three groups. Landmarks
485 within each cluster delineate meaningful regions of the external brain morphology, which
486 suggests that each of these regions evolved differently. Cluster 1, which mostly represents
487 the anterior frontal, parietal, and cerebellar regions, shows these regions have evolved into
488 a discretized trait space with high phylogenetic signal, whereas cluster 2, which defines the
489 temporal, occipital, and stem regions, shows these regions have evolved in a more uniform
490 space with low phylogenetic signal. Cluster 3, which defines the posterior and ventral areas
491 of the frontal region and part of the temporal region, also shows evidence of these regions
492 evolving into a discretized trait space, but with higher levels of convergence than the
493 regions of cluster 1. Despite the differences in approach, these results align well with a
494 previous analysis of the same dataset conducted using PCA (Aristide et al. 2016), which
495 suggests that, during the adaptive radiation of New World monkeys, brain shape evolved
496 first into discrete regions of morphospace, with subsequent bursts of evolution generating
497 convergence among clades. Moreover, according to Aristide et al. (2016), the different
498 stages of this diversification can be associated to the evolution of particular regions of the
499 brain. For example, coincident with our results for cluster 1, the anterior frontal region
500 would have diversified early into discrete trait optima, while convergent changes would be
501 mostly associated with other areas of the frontal region, in agreement with our cluster 3.
502 Overall, our results support the idea that there has been differential selection on different
503 brain regions in New World monkeys, due both to an early adaptive radiation and
504 convergence on ecologically relevant traits (Rosenberger 1992; Gavrilets and Losos 2009;
505 Aristide et al. 2015, 2016).

506 A major focus of work on phenotypic evolution relates to the study and
507 identification of co-evolving traits using multivariate models (Clavel et al. 2015).
508 Specifically, the correlated evolution of multiple traits resulting in evolutionary integration

509 expects such sets of traits to have shared evolutionary histories (Goswami 2007). We would
510 therefore also expect that these traits, whether they are biologically integrated or
511 co-evolving with some shared variable, will have similar spectral density profiles; and so
512 clustering profiles may be a way to identify different sets of integrated traits from
513 multivariate data. This can become particularly useful when there are many traits, as is
514 more often becoming the case with the proliferation of trait data (e.g., Jones et al. (2009);
515 Hamish et al. (2014)).

516 We have developed an approach, implemented in user-friendly software, which is a
517 valuable addition to existing PCMs and provides a new way to analyse and conceive
518 phenotypic evolution.

519 **Acknowledgments** We would like to thank Alex Bjarnason, Julien Clavel,
520 Jonathan Drury, Carmelo Fruciano, Odile Maliet, Marc Manceau, Olivier Missa, and
521 Guilhem Sommeria Klein for helpful comments on the manuscript. EL would like to thank
522 Evan Charles for helpful discussion. Funding was provided by the CNRS and a grant
523 (PANDA) from the European Research Council (ERC) attributed to HM. The views
524 expressed are those of the authors and should not be construed to represent the positions
525 of the U.S. Army, the Department of Defense, or the Department of Health and Human
526 Services.

*

527

528 References

- 529 Apté, C., F. Damerau, and S. M. Weiss. 1994. Automated learning of decision rules for text categorization.
530 ACM Trans. Inf. Syst. 12:233–251.
- 531 Arenas, A., A. Diaz-Guilera, and C. J. Perez-Vicente. 2006. Synchronization reveals topological scales in
532 complex networks. Phys. Rev. Lett. 96:114102.
- 533 Aristide, L., S. F. dos Reis, A. C. Machado, I. Lima, R. T. Lopes, and S. I. Perez. 2016. Brain shape
534 convergence in the adaptive radiation of new world monkeys. Proceedings of the National Academy of
535 Sciences 113:2158–2163.
- 536 Aristide, L., A. L. Rosenberger, M. F. Tejedor, and S. I. Perez. 2015. Modeling lineage and phenotypic
537 diversification in the new world monkey (platyrrhini, primates) radiation. Molecular Phylogenetics and
538 Evolution 82:375–385.
- 539 Banerjee, A. and J. Jost. 2008. On the spectrum of the normalized graph laplacian. Linear Algebra and its
540 Applications 428:3015–3022.
- 541 Banerjee, A. and J. Jost. 2009. Graph spectra as a systematic tool in computational biology. Networks in
542 Computational Biology 157:2425–2431.
- 543 Bauer, F. and J. Jost. 2009. Bipartite and neighborhood graphs and the spectrum of the normalized graph
544 Laplacian. ArXiv e-prints .
- 545 Beaulieu, J. M., B. C. O'Meara, and M. J. Donoghue. 2013. Identifying Hidden Rate Changes in the
546 Evolution of a Binary Morphological Character: The Evolution of Plant Habit in Campanulid
547 Angiosperms. Systematic Biology 62:725–737.
- 548 Beaumont, M. A. 2010. Approximate Bayesian Computation in Evolution and Ecology. Annual Review of
549 Ecology, Evolution, and Systematics 41:379–406.
- 550 Blomberg, S. P., T. Garland Jr, and A. R. Ives. 2003. Testing for phylogenetic signal in comparative data:
551 behavioral traits are more labile. Evolution 57:717–745.

- 552 Boucher, F. C., V. Démetry, E. Conti, L. J. Harmon, and J. Uyeda. 2017. A general model for estimating
553 macroevolutionary landscapes. *Systematic biology* .
- 554 Butler, M. A. and A. A. King. 2004. Phylogenetic comparative analysis: a modeling approach for adaptive
555 evolution. *The American Naturalist* 164:683–695.
- 556 Cavalli-Sforza, L. L. and A. W. Edwards. 1967. Phylogenetic analysis: models and estimation procedures.
557 *Evolution* 21:550–570.
- 558 Chen, G., G. Davis, F. Hall, Z. Li, K. Patel, and M. Stewart. 2004. An interlacing result on normalized
559 laplacians. *SIAM Journal on Discrete Mathematics* 18:353–361.
- 560 Chung, F. 1996. *Spectral Graph Theory* 92. American Mathematical Society, Fresno, CA.
- 561 Clarke, M., G. H. Thomas, and R. P. Freckleton. 2017. Trait evolution in adaptive radiations: modeling
562 and measuring interspecific competition on phylogenies. *The American Naturalist* 189:121–137.
- 563 Clavel, J., G. Escarguel, and G. Merceron. 2015. mvmorph: an r package for fitting multivariate
564 evolutionary models to morphometric data. *Methods in Ecology and Evolution* 6:1311–1319.
- 565 Clavel, J. and H. Morlon. 2017. Accelerated body size evolution during cold climatic periods in the
566 Cenozoic. *Proceedings of the National Academy of Sciences of the United States of America*
567 114:4183–4188.
- 568 Cooper, N., G. H. Thomas, C. Venditti, A. Meade, and R. P. Freckleton. 2015. A cautionary note on the
569 use of ornstein uhlenbeck models in macroevolutionary studies. *Biological Journal of the Linnean*
570 *Society* 118:64–77.
- 571 Deyasi, K., A. Banerjee, and B. Deb. 2015. Phylogeny of metabolic networks: A spectral graph theoretical
572 approach. *Journal of Biosciences* 40:799–808.
- 573 Drury, J., J. Clavel, M. Manceau, and H. Morlon. 2016. Estimating the Effect of Competition on Trait
574 Evolution Using Maximum Likelihood Inference. *Systematic Biology* 65:700–710.
- 575 Drury, J. P., J. A. Tobias, K. J. Burns, N. A. Mason, A. J. Shultz, and H. Morlon. 2018. Contrasting
576 impacts of competition on ecological and social trait evolution in songbirds. *PLOS Biology* 16:1–23.

- 577 Endres, D. and J. Schindelin. 2003. A new metric for probability distributions. *IEEE Transactions on*
578 *Information Theory* 49:1858–60.
- 579 Felsenstein, J. 1985. Phylogenies and the Comparative Method. *The American Naturalist* 125:1–15.
- 580 Foote, M. 1997. The evolution of morphological diversity. *Annual Review of Ecology and Systematics*
581 28:129–152.
- 582 Gavrillets, S. and J. B. Losos. 2009. Adaptive radiation: Contrasting theory with data. *Science*
583 323:732–737.
- 584 Goswami, A. 2007. Cranial modularity and sequence heterochrony in mammals. *Evolution & Development*
585 9:290–298.
- 586 Hamish, W., B. Jonathan, S. Jennifer, de la Rosa Carolina, R. M. M., and J. Walter. 2014. Eltonraits 1.0:
587 Specieslevel foraging attributes of the world’s birds and mammals. *Ecology* 95:2027–2027.
- 588 Hansen, T. F. 1997. Stabilizing selection and the comparative analysis of adaptation. *Evolution*
589 51:1341–1351.
- 590 Hansen, T. F. and E. P. Martins. 1996. Translating between microevolutionary process and
591 macroevolutionary patterns: the correlation structure of interspecific data. *Evolution* 50:1404–1417.
- 592 Harmon, L. J., J. B. Losos, T. Jonathan Davies, R. G. Gillespie, J. L. Gittleman, W. Bryan Jennings,
593 K. H. Kozak, M. A. McPeck, F. Moreno-Roark, T. J. Near, A. Purvis, R. E. Ricklefs, D. Schluter, J. A.
594 Schulte II, O. Seehausen, B. L. Sidlauskas, O. Torres-Carvajal, J. T. Weir, and A. Ø. Mooers. 2010.
595 Early bursts of body size and shape evolution are rare in comparative data. *Evolution; International*
596 *Journal of Organic Evolution* 64:2385–2396.
- 597 Harmon, L. J., J. A. Schulte, A. Larson, and J. B. Losos. 2003. Tempo and mode of evolutionary radiation
598 in iguanian lizards. *Science (New York, N.Y.)* 301:961–964.
- 599 Harmon, L. J., J. T. Weir, C. D. Brock, R. E. Glor, and W. Challenger. 2008. GEIGER: investigating
600 evolutionary radiations. *Bioinformatics* 24:129–131 PMID: 18006550.
- 601 Höhna, S. 2013. Fast simulation of reconstructed phylogenies under global time-dependent birth-death
602 processes. *Bioinformatics (Oxford, England)* 29:1367–1374.

- 603 Jones, K. E., J. Bielby, M. Cardillo, S. A. Fritz, J. O'Dell, C. D. L. Orme, K. Safi, W. Sechrest, E. H.
604 Boakes, C. Carbone, C. Connolly, M. J. Cutts, J. K. Foster, R. Grenyer, M. Habib, C. A. Plaster, S. A.
605 Price, E. A. Rigby, J. Rist, A. Teacher, O. R. P. Bininda-Emonds, J. L. Gittleman, G. M. Mace, and
606 A. Purvis. 2009. PanTHERIA: a species-level database of life history, ecology, and geography of extant
607 and recently extinct mammals. *Ecology* 90:2648–2648.
- 608 Kunegis, J., S. Schmidt, A. Lommatzsch, J. Lerner, E. W. De Luca, and S. Albayrak. 2010. Spectral
609 analysis of signed graphs for clustering, prediction and visualization. Pages 559–570 *in* Proceedings of
610 the 2010 SIAM International Conference on Data Mining SIAM.
- 611 Lewitus, E. 2018. Inferring evolutionary process from neuroanatomical data. *Frontiers in Neuroanatomy*
612 12:54.
- 613 Lewitus, E. and H. Morlon. 2016a. Characterizing and Comparing Phylogenies from their Laplacian
614 Spectrum. *Systematic Biology* 65:495–507.
- 615 Lewitus, E. and H. Morlon. 2016b. Natural Constraints to Species Diversification. *PLOS Biology*
616 14:e1002532.
- 617 Manceau, M., A. Lambert, and H. Morlon. 2016. A unifying comparative phylogenetic framework including
618 traits coevolving across interacting lineages. *Systematic biology* 66:551–568.
- 619 McGraw, P. N. and M. Menzinger. 2008. Laplacian spectra as a diagnostic tool for network structure and
620 dynamics. *Phys. Rev. E* 77:031102.
- 621 Monteiro, L. R., B. Bordin, and S. F. dos Reis. 2000. Shape distances, shape spaces and the comparison of
622 morphometric methods. *Trends in ecology & evolution* 15:217–220.
- 623 Morlon, H., E. Lewitus, F. L. Condamine, M. Manceau, J. Clavel, and J. Drury. 2016. RPANDA: an R
624 package for macroevolutionary analyses on phylogenetic trees. *Methods in Ecology and Evolution*
625 7:589–597.
- 626 Pennell, M. W. and L. J. Harmon. 2013. An integrative view of phylogenetic comparative methods:
627 connections to population genetics, community ecology, and paleobiology. *Annals of the New York*
628 *Academy of Sciences* 1289:90–105.

- 629 Rai, A., P. Pradhan, J. Nagraj, K. Lohitesh, R. Chowdhury, and S. Jalan. 2017. Understanding cancer
630 complexome using networks, spectral graph theory and multilayer framework. *Scientific Reports* 7:41676
631 EP –.
- 632 Revell, L. J. 2012. phytools: An r package for phylogenetic comparative biology (and other things).
633 *Methods in Ecology and Evolution* 3:217–223.
- 634 Reynolds, A., G. Richards, B. de la Iglesia, and V. Rayward-Smith. 2006. Clustering rules: A comparison
635 of partitioning and hierarchical clustering algorithms. *Journal of Mathematical Modelling and*
636 *Algorithms* 5:475–504.
- 637 Rohe, K., S. Chatterjee, and B. Yu. 2011. Spectral clustering and the high-dimensional stochastic
638 blockmodel. *Ann. Statist.* 39:1878–1915.
- 639 Rosenberger, A. L. 1992. Evolution of feeding niches in new world monkeys. *American Journal of Physical*
640 *Anthropology* 88:525–562.
- 641 Shuman, D., S. Narang, P. Frossard, A. Ortega, and P. Vanderghelynst. 2013. The Emerging Field of Signal
642 Processing on Graphs: Extending High-Dimensional Data Analysis to Networks and Other Irregular
643 Domains. *IEEE Signal Processing Magazine* 30:83–98.
- 644 Slater, G. J., S. A. Price, F. Santini, and M. E. Alfaro. 2010. Diversity versus disparity and the radiation of
645 modern cetaceans. *Proceedings of the Royal Society of London B: Biological Sciences* 277:3097–3104.
- 646 Steeman, M. E., M. B. Hebsgaard, R. E. Fordyce, S. Y. Ho, D. L. Rabosky, R. Nielsen, C. Rahbek,
647 H. Glenner, M. V. Sørensen, and E. Willerslev. 2009. Radiation of extant cetaceans driven by
648 restructuring of the oceans. *Systematic Biology* 58:573–585.
- 649 Szekeley, G. and M. Rizzo. 2005. Hierarchical Clustering via Joint Between-Within Distances: Extending
650 Ward’s Minimum Variance Method. *Journal of Classification* 22:151–183.
- 651 Uyeda, J. C., D. S. Caetano, and M. W. Pennell. 2015. Comparative analysis of principal components can
652 be misleading. *Systematic Biology* 64:677–689.
- 653 Weir, J. T. and S. Mursleen. 2013. Diversity-dependent cladogenesis and trait evolution in the adaptive
654 radiation of the auks (aves: alcidae). *Evolution; International Journal of Organic Evolution* 67:403–416.

655 Zhang, F. and E. R. Hancock. 2008. Graph spectral image smoothing using the heat kernel. Pattern
656 Recognition 41:3328 – 3342.

<i>symbol</i>	<i>descriptor</i>	<i>significance</i>	<i>value</i>
nMGL	normalized modified graph Laplacian	the distance matrix of the phylogeny weighted by the differences in trait values between tips and normalized by the degree matrix	positive semi-definite symmetric matrix
${}_n\lambda$	eigenvalues	eigenvalue calculated from the nMGL	$\sim 1 < {}_n\lambda \leq 2$
SDP	spectral density profile of the nMGL	the density profile of eigenvalues calculated from the nMGL	kernel density estimate of ${}_n\lambda$
${}_n\lambda^*$	splitter	the maximum eigenvalue; reflects bipartiteness (i.e., monophyletic clustering of trait data)	$1 < {}_n\lambda^* \leq 2$
${}_n\psi$	fragmenter	the skewness of the SDP; reflects discreteness	$0 < {}_n\psi < \infty$
${}_n\eta$	tracer	the maximum height of the SDP; reflects phylogenetic signal	$0 < {}_n\eta < \infty$

Table1: Glossary of graphical and statistical terms.

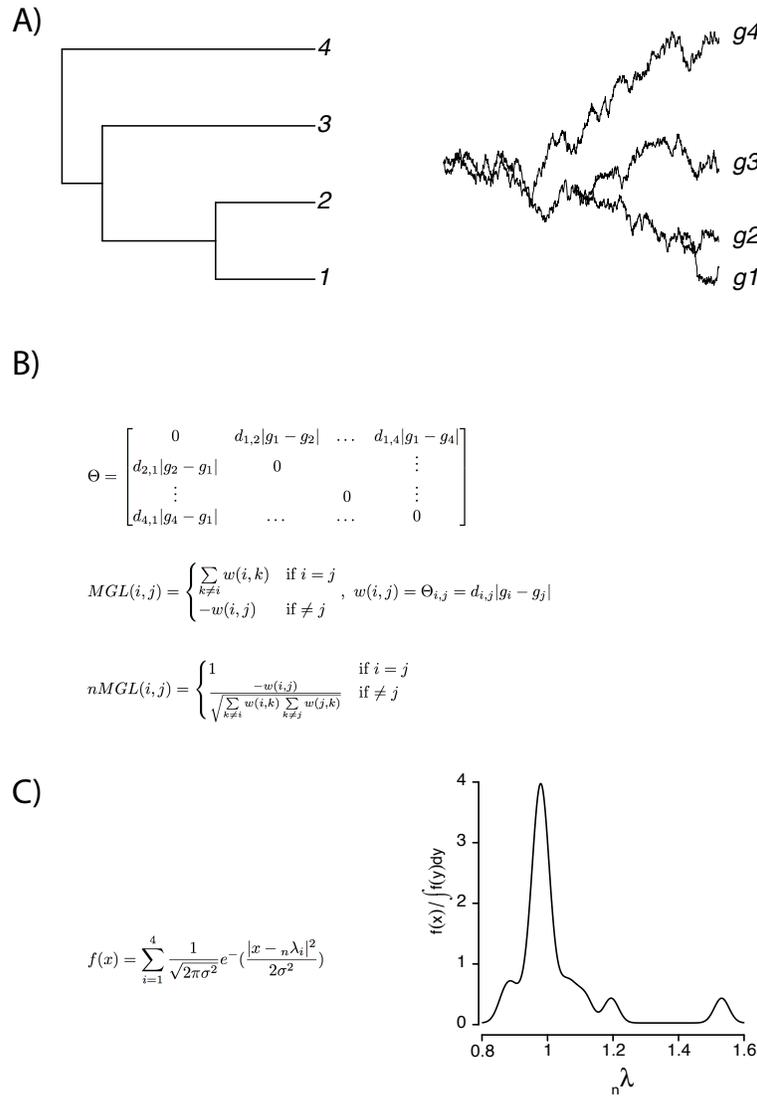


Figure 1: Pipeline for constructing the spectral density profile for the nMGL of phylogenetic trait data. (A) Given a phylogenetic tree with m terminal branches and unidimensional, continuous, extant trait data for m tips, (B) take the Hadamard product of the difference matrix of the trait data ($|g_i - g_j|$) and the matrix of phylogenetic branch-lengths between tips ($d_{i,j}$), such that $\Theta = d_{i,j}|g_i - g_j|$ at $i \neq j$ and zero along the diagonal. The weighted MGL, $D - \Theta$, where D is the degree matrix of Θ , is computed as the weighted value of (i, j) , $-\Theta(i, j) = -w(i, j)$, at $i \neq j$ and as $\sum w_{i,k}$ for $i = j$. The normalized MGL (nMGL) is normalized by D , so that $nMGL = D^{-1/2}(D - \Theta)D^{-1/2}$, resulting in unity along the diagonal and negative the weighted value of (i, j) divided by the square-root of the product of δ_i and δ_j for $i \neq j$. (C) The spectral density is obtained by convolving the eigenvalues, $n\lambda$, computed from the nMGL with a Gaussian kernel and then plotting the density of $n\lambda$.

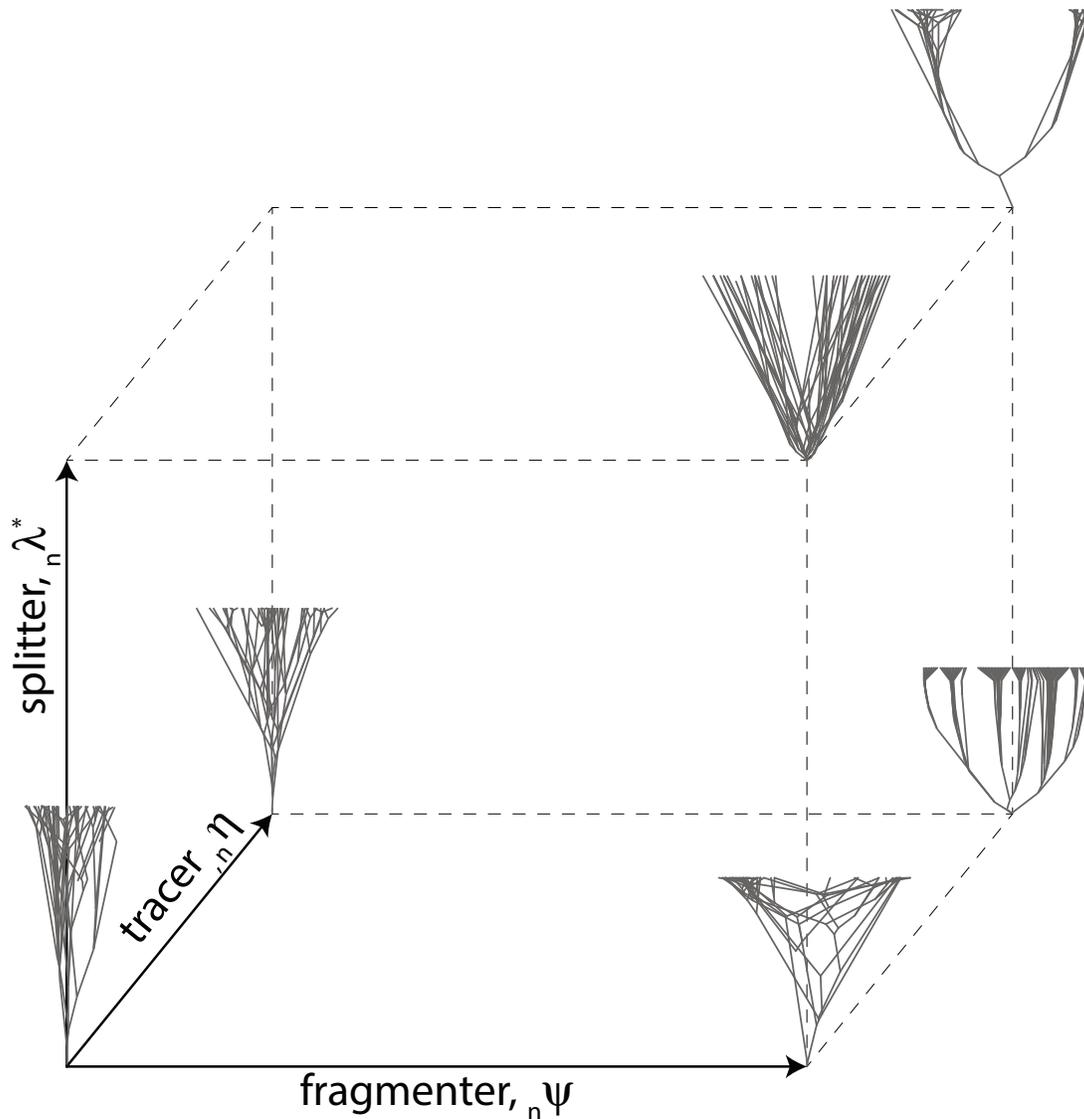


Figure 2: **Defining phylogenetic trait space** Any phylogenetic trait data can be placed in a three-dimensional space defined by the splitter (${}_n\lambda^*$), the fragmenter (${}_n\psi$), and the tracer (${}_n\eta$), which broadly represent the bipartiteness, discreteness, and phylogenetic signal, respectively, of the phylogenetic trait data. Hypothetical traitgrams are placed in the corners of the space, illustrating the type of patterns expected in those corners. Traitgrams are generated on the same phylogenetic tree under different trait evolution parameters.

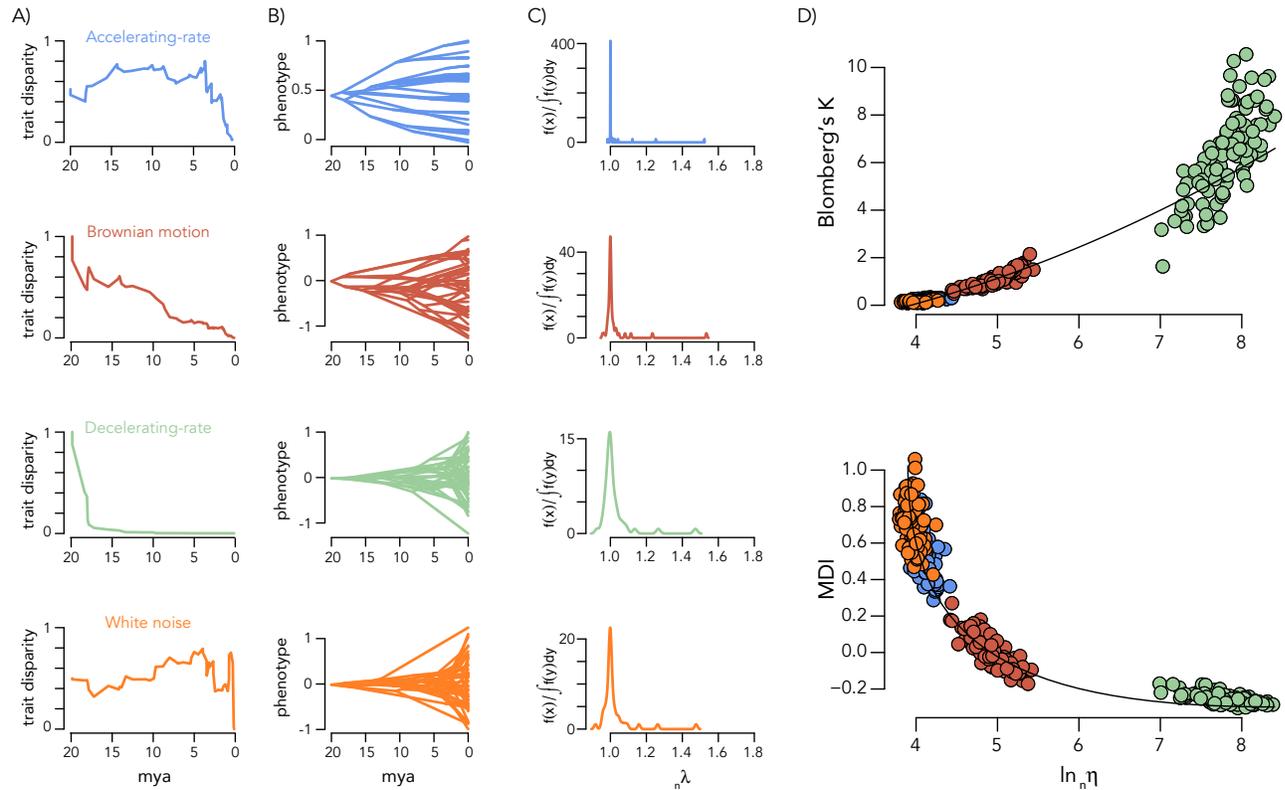


Figure 3: Interpreting the tracer of spectral density profiles. (A) Disparity-through-time plots for traits evolved under an AC, BM, DC, and white-noise model on the same 100-tip constant-rate birth-death phylogeny. (B) Traitgrams and (C) spectral density profiles for the phylogenetic traits in (A). Note the difference y-axis range in (C). Pairwise plots of Blomberg's K and MDI as a function of the tracer for phylogenetic trait data simulated under AC, BM, DC, and white-noise models. The best-fit regression slopes are shown for each plot.

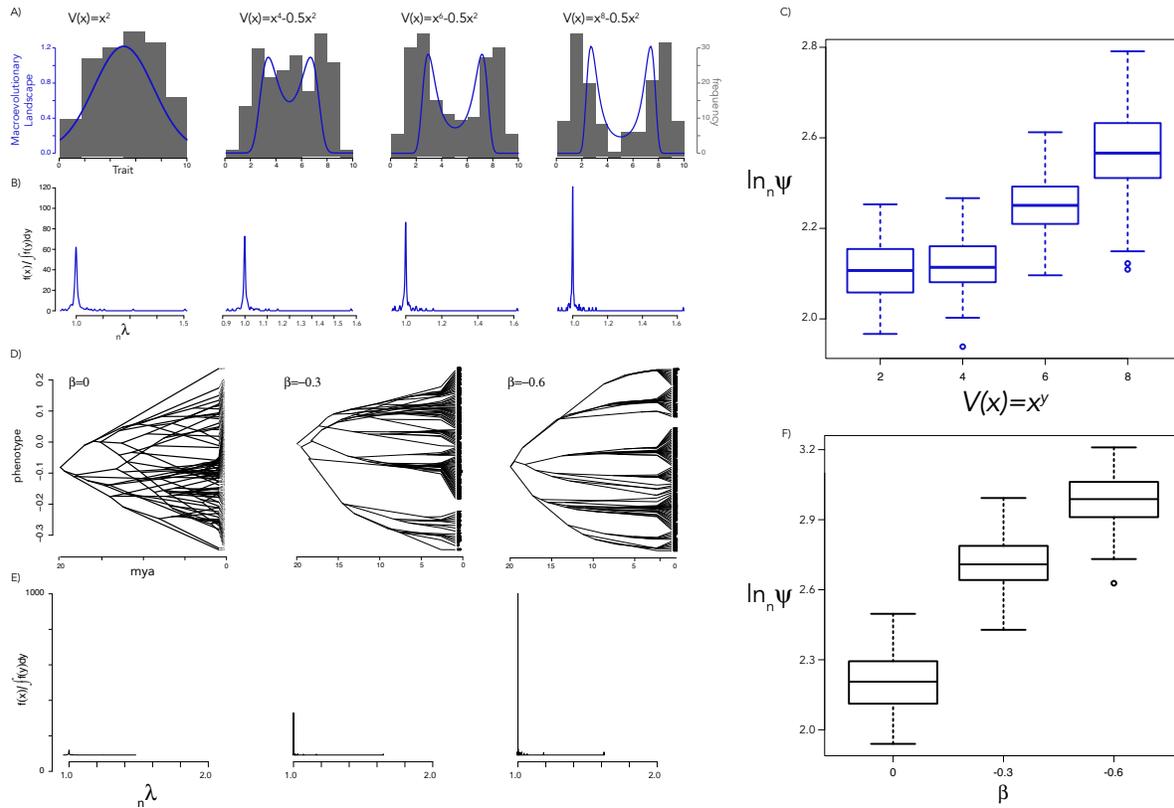


Figure 4: Interpreting the fragmenter of spectral density profiles. (A) Histograms of simulated trait values (grey) under four macroevolutionary landscapes (blue). (B) Spectral density profiles for phylogenetic trait data simulated under each landscape in (A). (C) Boxplot of fragmenter values for spectral density profiles generated under each macroevolutionary landscape in (A). (D) Traitgrams of phylogenetic trait data simulated under ACDC models with different rate parameter values, β . (E) Spectral density profiles for the phylogenetic trait data in (D). (F) Boxplot of fragmenter values for phylogenetic trait data simulated under each DC model in (C).

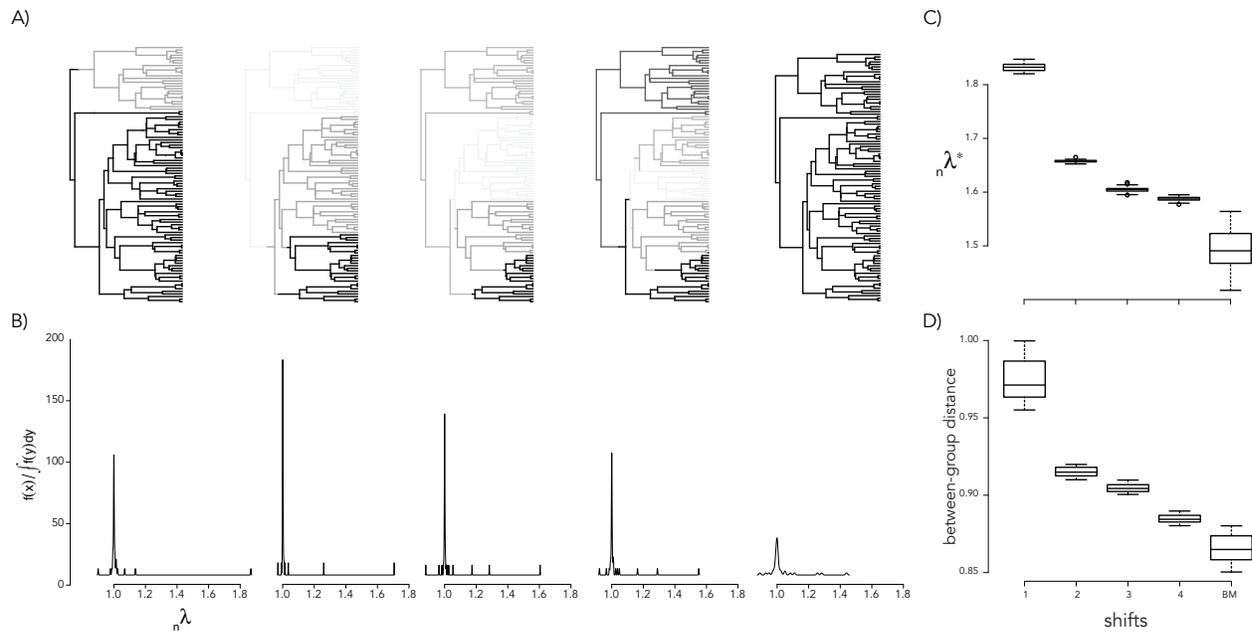


Figure 5: **Interpreting the splitter of spectral density profiles.** (A) Phylogenies simulated with 1 – 4 monophyletic shifts in mean trait values and no shifts in trait value. Different mean trait values are represented in grey scale. (B) Spectral density plots for the phylogenetic trait data in (A). (C) Boxplot of splitter values for phylogenetic trait data simulated under different numbers of monophyletic shifts in mean trait value. (D) Boxplot of the between-cluster branch-length distances (as a ratio over two times the crown age of the tree) for phylogenetic trait data simulated under different shifts in mean trait value, where clusters are defined by k-means clustering on nMGL (with $k=2$). Both splitter and between-cluster branch-length distance increase as the nMGL approaches bipartiteness (splitter= 2).

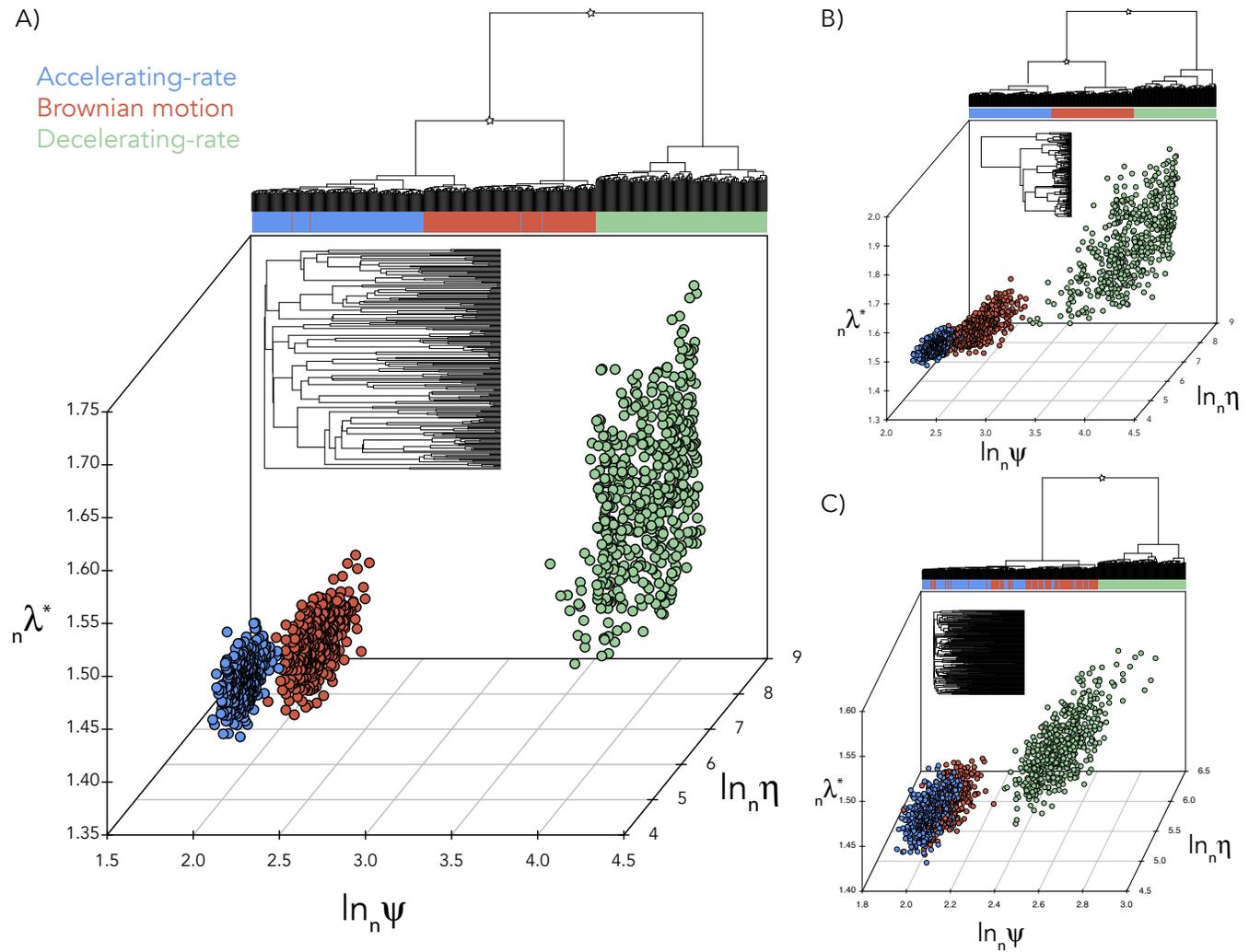


Figure 6: **Clustering phylogenetic trait data using the spectral density profile of the nMGL.** Hierarchical clustering of spectral density profiles and three-dimensional plotting of spectral density profile summary statistics for phylogenetic trait data simulated under AC, BM, and DC models of trait evolution on (A) a constant-rate birth-death tree, (B) an increasing-rate birth-death tree, and (C) a decreasing-rate birth-death tree. The trees are shown as insets. Asterisks denote bootstrap probabilities > 0.95 at the split.

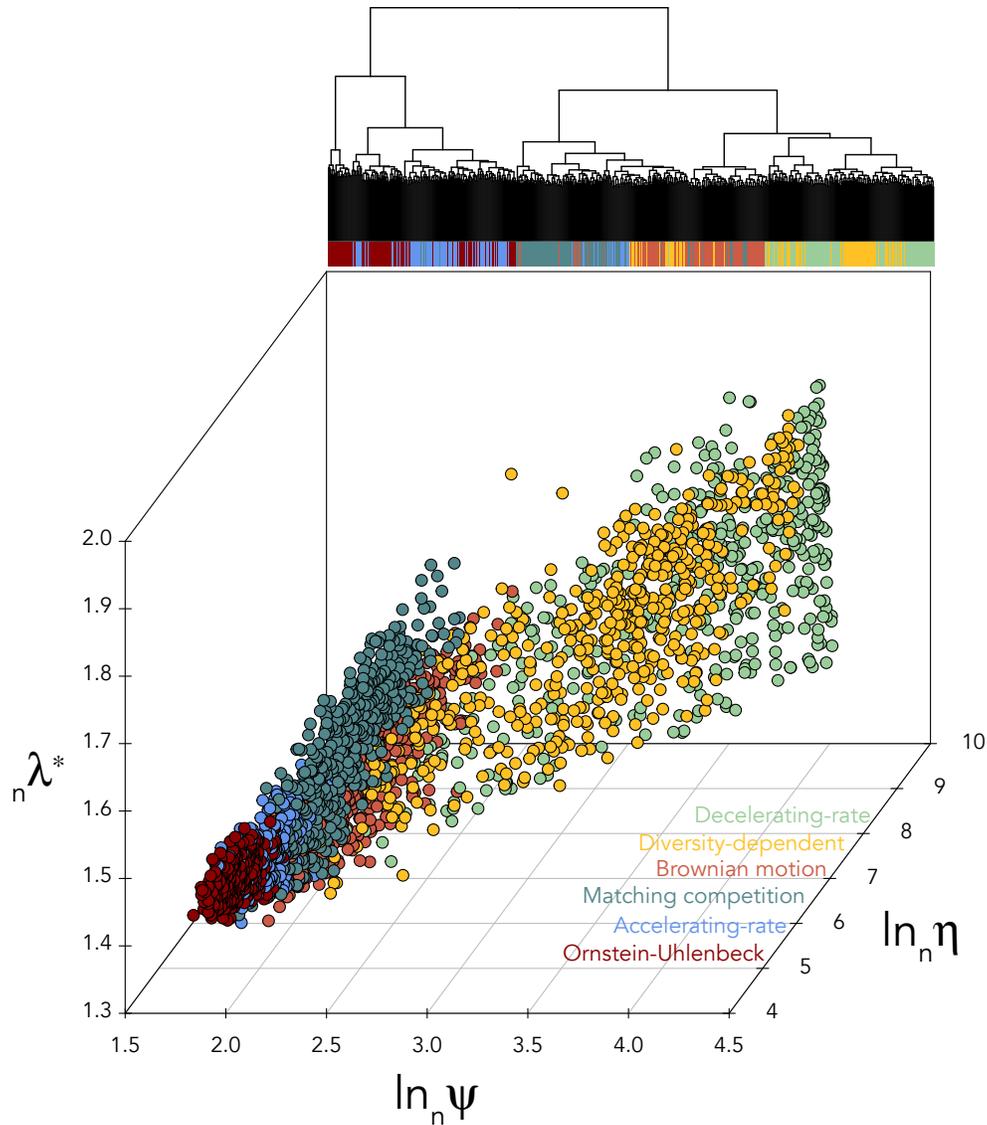


Figure 7: **Spectral density profiles for simulated trait models on the Cetacean phylogeny.** Hierarchical clustering and multidimensional plot of spectral density profile summary statistics for trait data simulated under AC, BM, DC, DD, MC, and OU models under varying parameter values on the Cetacean phylogeny.

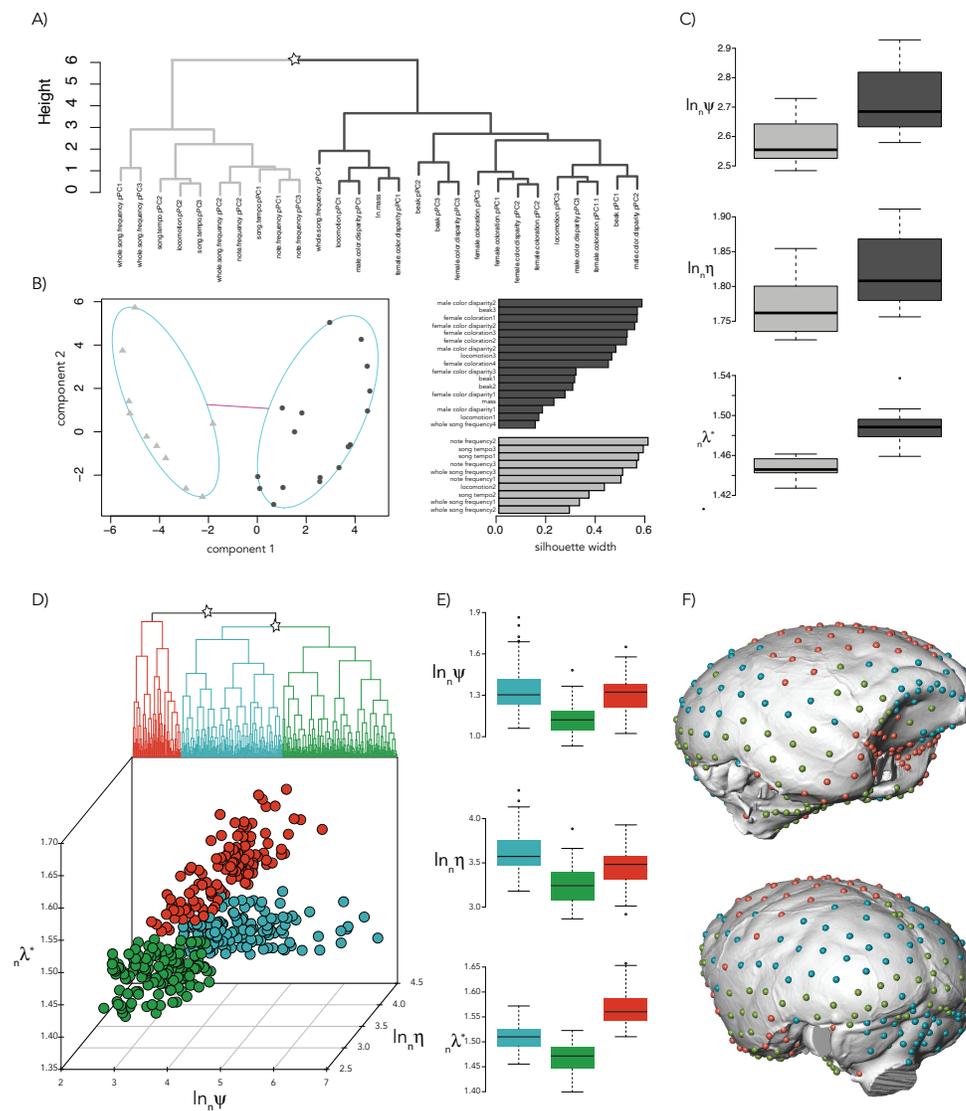


Figure 8: Spectral density profiling of traits in tanagers and New World monkeys. (A) Hierarchical and (B) k-medoids clustering on the spectral density profiles of the nMGLs constructed from 27 pPC traits on the tanager phylogeny. Silhouette widths are shown for each pPC trait in the k-medoid clustering. (C) Spectral density profile summary statistics for pPC traits within each cluster identified in (A,B). (D) Hierarchical clustering of spectral density profiles and multidimensional plot of spectral density profile summary statistics for 399 landmarks on New World monkey endocrania: cluster 1 (blue), cluster 2 (green), cluster 3 (red). (E) Boxplot of summary statistics for each cluster identified in (D). (F) Three-dimensional representation of the New World monkey endocranium with placement of the clusters of landmarks corresponding to (D).

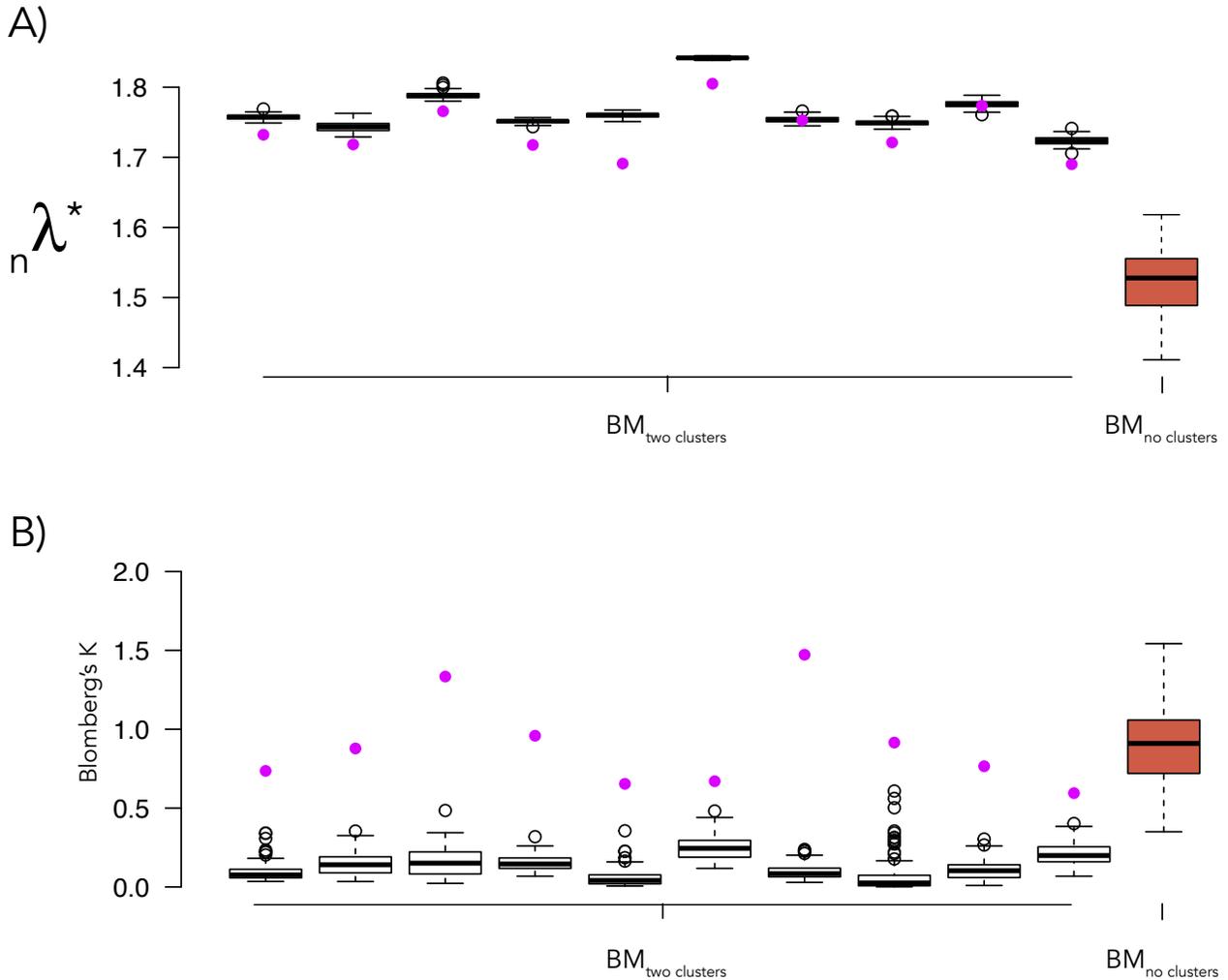


Figure S1: **Measuring the effect of phylogenetic signal on splitter values.** (A) Boxplot of the splitter values for 100 randomized datasets (white) obtained for each of the ten datasets with two monophyletic clusters. Splitter values for the initial BM datasets with two clusters are shown in purple. Boxplot of 100 datasets simulated under a simple BM process with no clusters on a single tree (coral) is shown for comparison. (B) Boxplot of Blomberg's K for each randomized dataset (white); values for the initial BM datasets with two clusters are shown in purple. Boxplot of 100 datasets simulated under a simple BM process with no clusters on a single tree (coral) is shown for comparison.

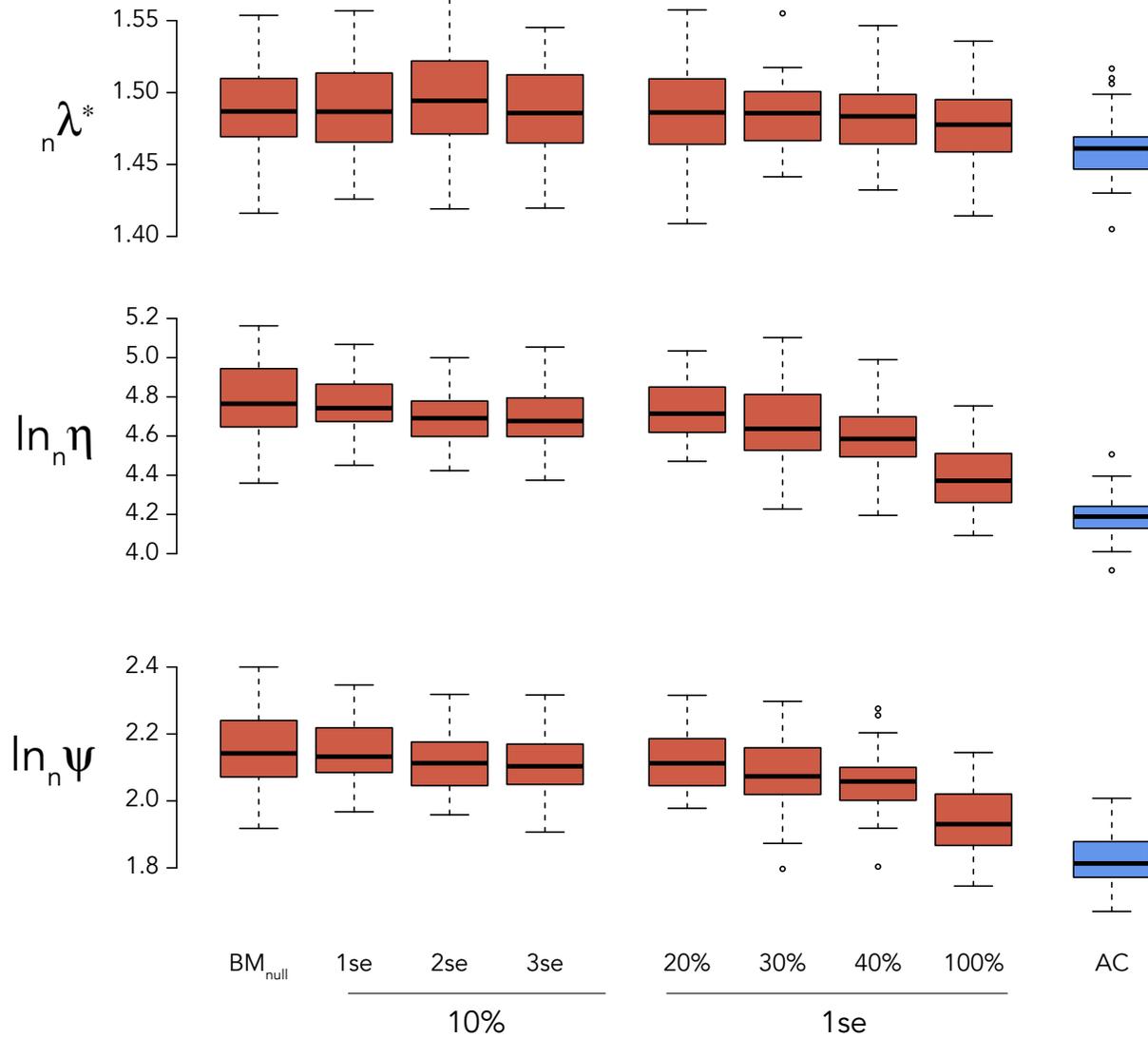


Figure S2: Measuring the effect of erroneous trait data on spectral density profile summary statistics. Spectral density profile summary statistics for data simulated under a BM process (coral) with introduced error for 10% of tips with a sampling variance equal to one, two, and three times the standard error of the simulated BM data; and with a sampling variance equal to one times the standard error for 10, 20, 30, 100% of tips. Spectral density profile summary statistics for data simulated on the same tree under an ACDC process ($\beta = 1.5$) is also shown (cornflowerblue).

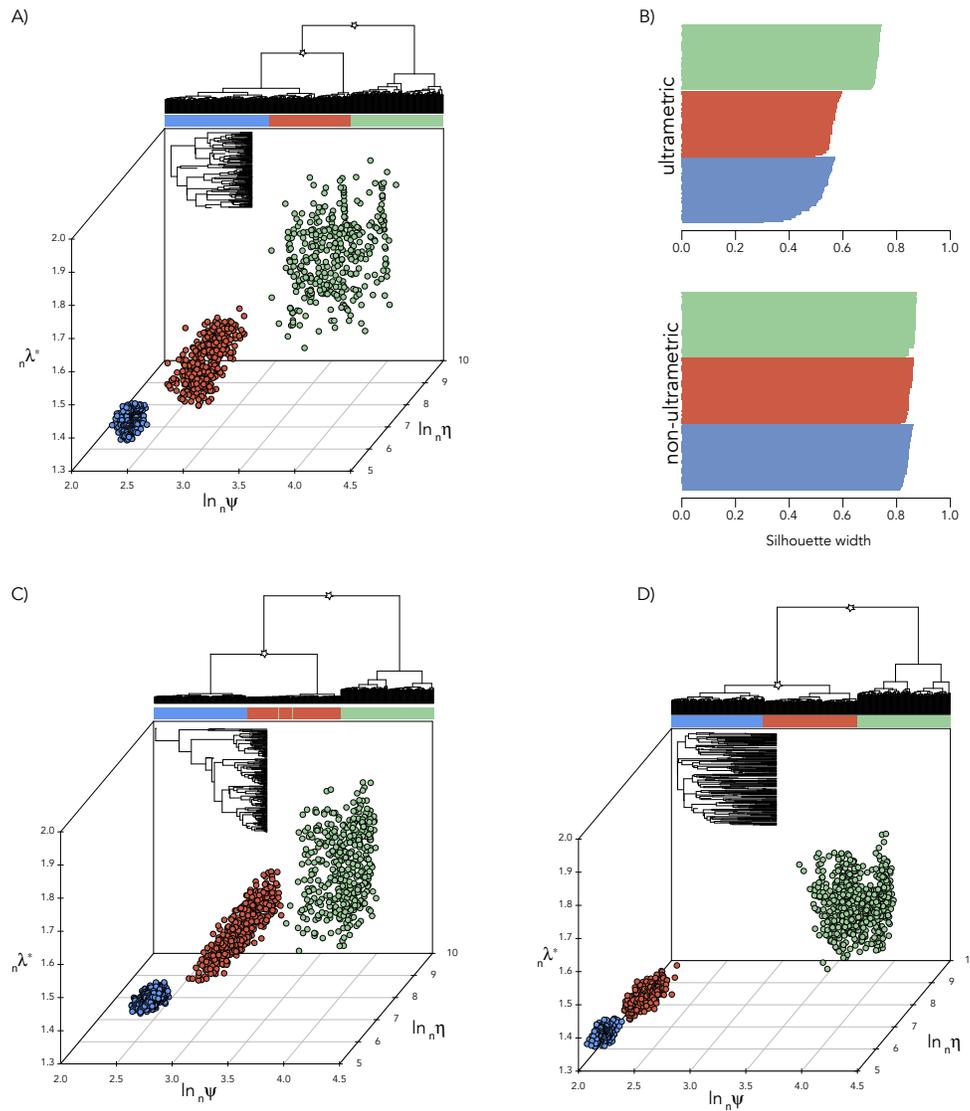


Figure S3: Clustering phylogenetic trait data using the spectral density profile of the nMGL on a non-ultrametric tree. Hierarchical clustering of spectral density profiles and three-dimensional plotting of spectral density profile summary statistics for phylogenetic trait data simulated under AC (cornflower blue), BM (coral), and DC (sea green) models of trait evolution on a single (A) constant-rate, (C) increasing-rate, and (D) decreasing-rate birth-death tree without pruning extinct lineages. Tree is shown in inset. Asterisks denote bootstrap probabilities > 0.95 at the split. (B) Silhouette widths for profiles comprising each trait model cluster simulated on the ultrametric or non-ultrametric tree (see Fig. 5A).

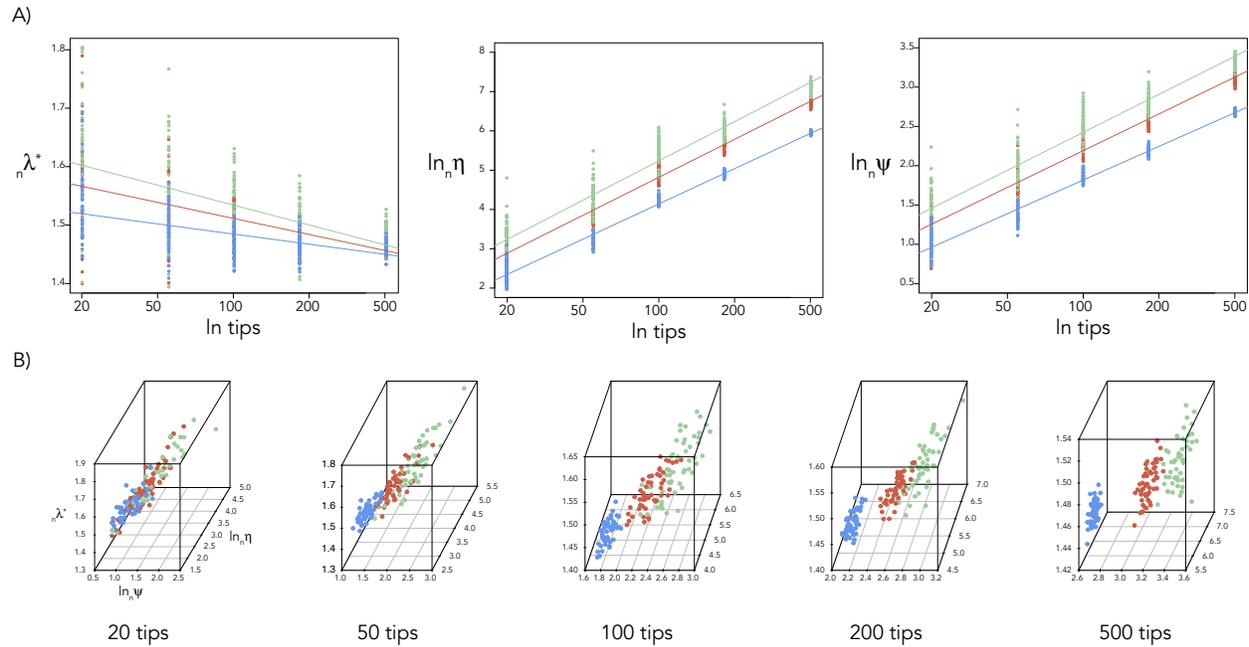


Figure S4: **Effect of tree size on the nMGL.** (A) Scatterplots and OLS regression slopes for spectral density profile summary statistics for trait data simulated under DC (sea green), BM (coral), and AC (cornflower blue) models on constant-rate birth-death trees with different numbers of tips. (B) Phylogenetic trait space for trait models simulated under AC, BM, and DC models on trees with different numbers of tips.