University of Reading

Phylogenetic non-independence in rates of trait evolution

Article

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Phylogenetic non-independence in rates of

trait evolution

RUNNING HEAD: Phylogenetic non-independence in rates

Manabu Sakamoto[1]* and Chris Venditti[1]*

[1]School of Biological Sciences, University of Reading, Reading, UK, RG6 6BX

*CORRESPONDING AUTHOR CONTACT DETAILS:

m.sakamoto@reading.ac.uk

c.d.venditti@reading.ac.uk

14 **ABSTRACT**

15 Statistical non-independence of species' biological traits is recognized in most traits under

16 selection. Yet, whether or not the evolutionary rates of such biological traits are statistically

17 non-independent remains to be tested. Here we test the hypothesis that phenotypic

18 evolutionary rates are non-independent, i.e. contain phylogenetic signal, using empirical rates

19 of evolution in three separate traits: body mass in mammals; beak shape in birds; and bite force

20 in amniotes. Specifically, we test whether rates are non-independent throughout the

21 evolutionary history of each tree. We find evidence for phylogenetic signal in evolutionary rates

22 in all three case studies. While phylogenetic signal diminishes deeper in time, this is reflective

23 of statistical power owing to small sample and effect sizes. When effect size is large, e.g., owing

24 to the presence of fossil tips, we detect high phylogenetic signals even in deeper time slices.

25 Thus, we recommend that rates be treated as being non-independent throughout the

26 evolutionary history of the group of organisms under study, and any summaries or analyses of

27 rates through time – including associations of rates with traits – need account for the undesired

28 effects of shared ancestry.

29

30 **KEY WORDS:** evolutionary rates; trait evolution; phylogeny; phylogenetic comparative

31 methods; phylogenetic signal

32

33

34   Descent with modification [1] is of fundamental importance to evolution and is recognized in

35   most traits under selection. Through evolutionary time, trait values will be more similar in

36   closely related species compared to distantly related species, since the variance of trait values

37   will be proportional to the divergence in evolutionary time [2]. This equates to shared ancestry,

38   i.e. phylogeny. The degree to which shared ancestry affects biological traits can thus be

39   described by the proportion of variance in trait data across a comparative sample of species

40   that can be explained by phylogenetic relations, or phylogenetic signal – e.g., K [3] or λ [4]. This

41   has statistical implications, i.e., phylogenetic non-independence.

42   While acknowledgement of phylogenetic non-independence in phenotypic trait data has

43   become common in comparative studies [5], it is not so for the rates of evolution (how fast

44   organisms' characteristics evolve). As rates are often used as proxies for adaptations [6, 7], it is

45   of immense importance that we understand their statistical properties, in particular,

46   phylogenetic non-independence. However, we have not been able to identify any study in the

47   literature that explicitly tests for phylogenetic signal in phenotypic evolutionary rates aside

48   from rare instances in which this was implied [8].

49   Here, we test whether evolutionary rates contain phylogenetic signal using three

50   empirical case studies: body mass in mammals [9]; beak shape in birds [10]; and bite force in

51   terrestrial amniotes (ESM). Our basic premise is that if phylogenetic signal is detected in rates,

52   then rates evolve along the branches of a phylogenetic tree in proportion to the passage of

53   time and that closely related species are more similar in rates than distantly related species.

54   Naturally, this necessitates a non-homogenous distribution of rates across the branches of the

55   tree – i.e., variable-rates of phenotypic trait evolution [8, 11-14].

56

MATERIALS AND METHODS

57

58   We obtained 100 sets of phenotypic evolutionary rates and the associated time-calibrated

59   phylogenetic trees (time-trees) from the authors of the three case studies (ESM). As we would

60   expect rates along all branches of a phylogenetic tree to be affected by shared ancestry, not

61   just the tips, we tested for phylogenetic signal in rates along both terminal and internal

62   branches, by time-slicing the phylogenetic tree. We time-sliced the three time-trees at 1-Myr

63   intervals for the mammals and birds (167 and 109 time slices respectively), and at 5-Myr

64   intervals for amniotes (65 time slices) (see ESM for details). The latter interval was chosen for

65   amniotes owing to their much longer evolutionary history (approx. 350 Myr) compared to

66   mammals and birds. For each time-sliced tree, we matched the terminal branches to the

67   corresponding branches in the complete time-tree (Fig S1). We then assigned the

68   corresponding rates to those terminal branches on the time-sliced tree as tip trait values. We

69   fitted a maximum likelihood (ML) phylogenetic generalized least squares (GLS) model in

70   BayesTraits V3 to estimate phylogenetic signal λ in rates at the tips for each time-slice (GLS_λ).

71   We tested GLS_λ against the null model in which λ is fixed to 0 ($GLS_{\lambda=0}$) as the likelihood ratio

72   (LR_λ) between GLS_λ and $GLS_{\lambda=0}$ and determined significance using the χ^2 distribution (df=1).

73   When λ was significant in >95% of the sample in any given time slice, we determined that

74   phylogenetic signal was present in that time slice. We also compared the significance of an

75   alternative model in which λ is fixed to 1 ($GLS_{\lambda=1}$). The root estimate α of a $GLS_{\lambda=0}$ model is the

76   equivalent of estimating the non-phylogenetic mean rate, while GLS_λ and $GLS_{\lambda=1}$ estimate the

77   phylogenetically corrected mean rates.

80 **RESULTS**

81 Overall, phylogenetic signal at the tips of the complete time trees are high (body mass in

82 mammals, median λ = 0.926; beaks in birds, median λ = 0.729; and bite force in amniotes, λ =

83 1), providing evidence for strong effects of shared ancestry in rates of phenotypic trait

84 evolution along the terminal branches. Phylogenetic signal in rates are generally high and

85 significant in at least 95% of the sample in younger time slices – younger than: 48 Myr ago

86 (mammals); 45 Myr ago (birds); and 30 Myr ago (amniotes) (Fig. 1; ESM). Phylogenetic signal

87 depreciates (drops in strength and significance) rapidly in deeper time slices (Fig. 1; ESM). Fixing

88 λ to 1 (GLS$_{\lambda=1}$) result in qualitatively similar patterns across time slices compared to when λ is

89 estimated (GLS$_\lambda$) (Fig. 1), but depreciation of λ start at younger time slices compared to GLS$_\lambda$

90 (Fig. 1).

91

92 **DISCUSSION**

93 Through our time-sliced GLS models on three datasets, we demonstrate that evolutionary rates

94 of phenotypic traits are indeed phylogenetically non-independent – λ is significant and high,

95 both along the terminal and internal branches (Fig. 1). Crucially, although λ ceases to be

96 significant in deeper time slices in all trees tested (Fig. 1), this reduction in phylogenetic signal

97 most likely depends on two aspects of the rates in the focal time slice: 1) number of tips [15]

98 and 2) rate heterogeneity (ESM). Both reflect issues of statistical power with the former

99 concerning sample size (as determined through simulations; ESM) and the latter effect size (as

100   evident from the effects of fossil tips; Fig 1C; ESM). Incidentally, un-sampled tips of any sort

101   (not just fossils) will likely increase rate heterogeneity should they be sampled. Additionally,

102   information contained at the tips of an ultrametric tree (e.g., trait values) is expected to be lost

103   progressively deeper in the tree (proportional to the phylogenetic variance-covariance

104   structure) as subsequent evolution towards the tips overprints ancestral information – this is an

105   issue plaguing phylogenetic comparative methods in general. Furthermore, since rates are

106   estimated from the phylogeny using models with constant rate evolution as the underlying

107   process of evolution, the resulting rates would inevitably contain phylogenetic signal. Whether

108   this is true or not, this does not alter (rather it reinforces) our argument that inferred rates

109   contain phylogenetic signal (regardless of the reason) and crucially that all downstream

110   summaries and analyses of rates thus must account for phylogenetic non-independence. Thus,

111   we argue that it is safest to assume that phylogenetic signal will be present and strong in

112   deeper time slices [8].

113   An important implication here is that as rates will be statistically non-independent at

114   various time intervals throughout the history of the clade of interest, patterns gleaned from

115   simple summaries (e.g., interval means) of rates-through-time (RTT) can potentially be

116   misleading. Simple RTT plots are prevalent in recent literature [e.g., 10, 16, 17-21], the profiles

117   of which routinely interpreted at face value, with peaks and troughs representing periods of

118   bursts and declines in rates [16, 17, 19]. However, accounting for phylogenetic non-

119   independence by assuming strong phylogenetic signal uniformly across all time slices [8] – i.e.

120   phylogenetic mean α from our GLS$_{\lambda=1}$ models across time slices – results in phylo-RTT profiles

121   that are often different from those of non-phylogenetic RTT (Fig. 2). Thus, non-phylogenetic

122 RTT profiles cannot be taken at face value without knowledge of phylogenetic signal through

123 time. More crucially, this implies that statistical analyses of rates need also account for

124 phylogenetic non-independence. Testing hypotheses of external influences (ecological or

125 environmental) on rates of evolution would require the application of appropriate phylogenetic

126 statistical methods – e.g. phylogenetic regression models [14, 22]. Not doing so will run the risk

127 of resulting in misleading statistical results [2, 5].

128   As phenotypic evolutionary rates have been interpreted as reflecting the intensity of

129 natural selection [6, 14], that they contain phylogenetic signal implies that ancestors and

130 descendants as well as closely related species either: 1) share intrinsic mechanisms for selection

131 responsiveness (e.g., genetic predisposition); 2) share similar levels of extrinsic selection

132 pressures (e.g., similar ecological niches, environments, etc); or 3) both. Two (or more) species

133 descended from a parent species would be expected to start their respective independent

134 evolution with the same level of intrinsic responsiveness as well as extrinsic selection pressures,

135 and thus at the phenotypic evolutionary rate, of the parent species. The daughter species then

136 would be subject to independent genetic mutations and selection pressures depending on their

137 respective environments.

138   However, this is not to say that descendent rates are rigidly constrained by ancestry;

139 exceptional rate shifts along individual branches are widely observed in many traits across

140 various groups of organisms [8, 9, 11, 14, 23]. Such exceptional rate shifts can often be orders

141 of magnitude greater than the background rate and occur instantaneously (with respect to

142 geological time) such that the effects of ancestry may be marginal.

143   In conclusion, our analyses demonstrate that rates of phenotypic evolution estimated

144   from phylogenetic trees using models of trait evolution are statistically non-independent (most

145   likely owing to shared ancestry), across the tips and through time – we posit that our results are

146   conservative with phylogenetic signal actually being more prevalent. Thus, we recommend that

147   phylogenetic non-independence be accounted for in summaries and analyses of evolutionary

148   rates through time, using appropriate phylogenetic comparative methods.

149 **LITERATURE CITED**

150 [1] Darwin, C. 1859 *On the Origin of Species by Means of Natural Selection, or the Preservation*
151 *of Favoured Races in the Struggle for Life*. First Edition ed. London, UK.
152 [2] Felsenstein, J. 1985 Phylogenies and the Comparative Method. *American Naturalist* **125**, 1-
153 15. (doi:Doi 10.1086/284325).
154 [3] Blomberg, S.P., Garland, T. & Ives, A.R. 2003 Testing for phylogenetic signal in comparative
155 data: Behavioral traits are more labile. *Evolution* **57**, 717--745.
156 [4] Pagel, M. 1997 Inferring evolutionary processes from phylogenies. *Zoologica Scripta* **26**, 331-
157 348. (doi:10.1111/j.1463-6409.1997.tb00423.x).
158 [5] Harvey, P.H. & Pagel, M.D. 1991 *The comparative method in evolutionary biology*, Oxford
159 University Press.
160 [6] Simpson, G.G. 1944 *Tempo and mode in evolution*. New York, Columbia University Press; 237
161 p.
162 [7] Pagel, M. 1994 The adaptationist wager. In *Phylogenetics and Ecology* (eds. P. Eggleton &
163 R.I. Vane-Wright), pp. 29-51. London, Academic Press.
164 [8] Venditti, C., Meade, A. & Pagel, M. 2011 Multiple routes to mammalian diversity. *Nature*
165 **479**, 393-396. (doi:10.1038/nature10516).
166 [9] Baker, J., Meade, A., Pagel, M. & Venditti, C. 2015 Adaptive evolution toward larger size in
167 mammals. *Proceedings of the National Academy of Sciences* **112**, 5093-5098.
168 (doi:10.1073/pnas.1419823112).
169 [10] Cooney, C.R., Bright, J.A., Capp, E.J.R., Chira, A.M., Hughes, E.C., Moody, C.J.A., Nouri, L.O.,
170 Varley, Z.K. & Thomas, G.H. 2017 Mega-evolutionary dynamics of the adaptive radiation of
171 birds. *Nature* **542**, 344-347. (doi:10.1038/nature21074
172 http://www.nature.com/nature/journal/v542/n7641/abs/nature21074.html#supplementary-
173 information).
174 [11] Eastman, J.M., Alfaro, M.E., Joyce, P., Hipp, A.L. & Harmon, L.J. 2011 A Novel Comparative
175 Method for Identifying Shifts in the Rate of Character Evolution on Trees. *Evolution* **65**, 3578-
176 3589. (doi:10.1111/j.1558-5646.2011.01401.x).
177 [12] Rabosky, D.L. 2014 Automatic Detection of Key Innovations, Rate Shifts, and Diversity-
178 Dependence on Phylogenetic Trees. *Plos One* **9**, e89543. (doi:10.1371/journal.pone.0089543).
179 [13] Elliot, M.G. & Mooers, A.Ø. 2014 Inferring ancestral states without assuming neutrality or
180 gradualism using a stable model of continuous character evolution. *Bmc Evol Biol* **14**, 226.
181 (doi:10.1186/s12862-014-0226-8).
182 [14] Baker, J., Meade, A., Pagel, M. & Venditti, C. 2016 Positive phenotypic selection inferred
183 from phylogenies. *Biological Journal of the Linnean Society* **118**, 95-115.
184 (doi:10.1111/bij.12649).
185 [15] Freckleton, R.P., Harvey, P.H. & Pagel, M. 2002 Phylogenetic analysis and comparative
186 data: A test and review of evidence. *American Naturalist* **160**, 712--726.
187 [16] Close, Roger A., Friedman, M., Lloyd, Graeme T. & Benson, Roger B.J. 2015 Evidence for a
188 Mid-Jurassic Adaptive Radiation in Mammals. *Current Biology*. (doi:10.1016/j.cub.2015.06.047).
189 [17] Felice, R.N. & Goswami, A. 2018 Developmental origins of mosaic evolution in the avian
190 cranium. *Proc Natl Acad Sci U S A* **115**, 555-560. (doi:10.1073/pnas.1716437115).

191 [18] Lloyd, G.T., Wang, S.C. & Brusatte, S.L. 2012 Identifying heterogeneity in rates of
192 morphological evolution: Discrete character change in the evolution of lungfish (Sarcopterygii;
193 Dipnoi). *Evolution* **66**, 330-348. (doi:10.1111/j.1558-5646.2011.01460.x).
194 [19] Hopkins, M.J. & Smith, A.B. 2015 Dynamic evolutionary change in post-Paleozoic echinoids
195 and the importance of scale when interpreting changes in rates of evolution. *Proceedings of the*
196 *National Academy of Sciences of the United States of America* **112**, 3758-3763.
197 (doi:10.1073/pnas.1418153112).
198 [20] Rabosky, D.L., Donnellan, S.C., Grundler, M. & Lovette, I.J. 2014 Analysis and Visualization
199 of Complex Macroevolutionary Dynamics: An Example from Australian Scincid Lizards.
200 *Systematic Biology* **63**, 610-627. (doi:10.1093/sysbio/syu025).
201 [21] Evans, A.R., Jones, D., Boyer, A.G., Brown, J.H., Costa, D.P., Ernest, S.K.M., Fitzgerald,
202 E.M.G., Fortelius, M., Gittleman, J.L., Hamilton, M.J., et al. 2012 The maximum rate of mammal
203 evolution. *Proceedings of the National Academy of Sciences* **109**, 4187-4190.
204 [22] Orme, D., Freckleton, R., Thomas, G., Petzoldt, T., Fritz, S., Isaac, N. & Pearse, W. 2013
205 caper: Comparative Analyses of Phylogenetics and Evolution in R. R package version 0.5.2. (R
206 package version 0.5.2 ed.
207 [23] Chira, A.M. & Thomas, G.H. 2016 The impact of rate heterogeneity on inference of
208 phylogenetic models of trait evolution. *Journal of Evolutionary Biology* **29**, 2502-2518.
209 (doi:10.1111/jeb.12979).
210 [24] Sakamoto, M. & Venditti, C. 2018 Phylogenetic non-independence in rates of trait
211 evolution. (osf.io/pn4ma.
212

213

221

222   **AUTHOR CONTRIBUTIONS:**

223   All authors made substantial contributions to conception and design, acquisition of data,

224   analysis and interpretation of data; drafted and revised the article critically for important

225   intellectual content; made final approval of the version to be published; and made agreements

226   to be accountable for all aspects of the work in ensuring that questions related to the accuracy

227   or integrity of any part of the work are appropriately investigated and resolved.

228

229   **DATA ACCESSIBILITY:**

230   Supporting data are made available through the Open Science Framework

231   (https://osf.io/pn4ma/?view_only=2de49c7ad61944ed97c373783a7d1956)[24], and described

232   in ESM.

233

234   **COMPETING INTERESTS:**

235   The authors declare no competing interests.

FIGURE LEGENDS

237 **Figure 1. Phylogenetic signal in rates of trait evolution through time.**

238 Phylogenetic signal (λ) was estimated across time sliced phylogenetic trees (top row) in three

239 independent data sets: A, mammalian body mass; B, avian beak shape; and C, aminote bite

240 force. Faint lines represent each of the 100 samples with the bold line representing the median

241 λ. The percentage of the sample in which LR_λ (likelihood ratio between GLS_λ and $GLS_{\lambda=0}$) was

242 significant is shown for each time slice (second row). Further, the fit of $GLS_{\lambda=1}$ is shown as the

243 percentage of the sample in which $LR_{\lambda=1}$ (likelihood ratio between $GLS_{\lambda=1}$ and $GLS_{\lambda=0}$) was

244 significant for each time slice (third row). Red dashed line represents the 95% threshold. Blue

245 dashed line (top) represents the time slice for the 95% threshold as determined through

246 simulations (Fig. S3). The relationship between the percentage of significant λ and N_{Tips} (bottom

247 row) shows a clear drop off in the percentage from 95% of the sample (red box).

248

249 **Figure 2. Mean evolutionary rates through time compared to phylogenetically corrected**

250 **mean rates.** Simple mean values of evolutionary rates at each time slice across the three

251 datasets (A, mammalian body mass; B, avian beak shape; C, amniote bite force) show distinctive

252 patterns of rates through time. However, these patterns are far less prominent in

253 phylogenetically corrected mean rates (α) through time. α are the phylogenetic root estimates

254 of the $GLS_{\lambda=1}$ model. Faint lines represent each MCMC run while the bold line shows the median

255 value for each time slice.

256