A Standardized Effect Size for Evaluating and Comparing the Strength of Phylogenetic Signal

Dean C. Adams[a,2], **Erica K. Baken**[a,b], **and Michael L. Collyer**[b]

[a]Department of Ecology, Evolution, and Organismal Biology, Iowa State University, Ames, Iowa, 50010. USA.; [b]Department of Science, Chatham University, Pittsburgh, Pennsylvania, 15232. USA.

This manuscript was compiled on July 17, 2020

Macroevolutionary studies frequently characterize the phylogenetic signal in phenotypes, however, analytical tools for comparing the strength of that signal across traits remain largely nderdeveloped. Here we evaluate the efficacy of of Pagel's λ to correctly estimate the strength of phylogenetic signal in phenotypic traits across a range of input values. We find that λ behaves as a Bernoulli random variable, where estimates are increasingly skewed at larger and smaller input levels of phylogenetic signal. Further, the precision of λ varies with input signal. Another measure, Blomberg's κ, is more consistent across a range of tree sizes, and exhibits a positive relationship with input levels of phylogenetic signal. However, that relationship is decidedly nonlinear. Thus, neither λ nor κ are suitable as effect sizes for measuring the strength of phylogenetic signal, and comparing that signal across datasets. As an alternative, we propose a standardized effect size based on κ, (Z_κ), which measures the strength of phylogenetic signal more reliably than does λ, and places that signal on a common scale for statistical comparison. We develop tests based on Z_κ to provide a mechanism for formally comparing the strength of phylogenetic signal across datasets, in much the same manner as effect sizes may be used to summarize patterns in quantitative meta-analysis. Our approach extends the phylogenetic comparative toolkit to address hypotheses that compare the strength of phylogenetic signal between various phenotypic traits, even when those traits are found in different evolutionary lineages or have different units or scales.

phylogenetic signal | macroevolution | lambda | kappa

This PNAS journal template is provided to help you write your work in the correct journal format. Instructions for use are provided below.

Note: please start your introduction without including the word "Introduction" as a section heading (except for math articles in the Physical Sciences section); this heading is implied in the first paragraphs.

Guide to using this template on Overleaf

Please note that whilst this template provides a preview of the typeset manuscript for submission, to help in this preparation, it will not necessarily be the final publication layout. For more detailed information please see the PNAS Information for Authors.

If you have a question while using this template on Overleaf, please use the help menu ("?") on the top bar to search for help and tutorials. You can also contact the Overleaf support team at any time with specific questions about your manuscript or feedback on the template.

Author Affiliations. Include department, institution, and complete address, with the ZIP/postal code, for each author. Use lower case letters to match authors with institutions, as shown in the example. Authors with an ORCID ID may supply this information at submission.

Submitting Manuscripts. All authors must submit their articles at PNAScentral. If you are using Overleaf to write your article, you can use the "Submit to PNAS" option in the top bar of the editor window.

Format. Many authors find it useful to organize their manuscripts with the following order of sections; Title, Author Affiliation, Keywords, Abstract, Significance Statement, Results, Discussion, Materials and methods, Acknowledgments, and References. Other orders and headings are permitted.

Manuscript Length. PNAS generally uses a two-column format averaging 67 characters, including spaces, per line. The maximum length of a Direct Submission research article is six pages and a Direct Submission Plus research article is ten pages including all text, spaces, and the number of characters displaced by figures, tables, and equations. When submitting tables, figures, and/or equations in addition to text, keep the text for your manuscript under 39,000 characters (including spaces) for Direct Submissions and 72,000 characters (including spaces) for Direct Submission Plus.

References. References should be cited in numerical order as they appear in text; this will be done automatically via bibtex,

Significance Statement

Evolutionary biologists wish to quantify and compare the strength of phylogenetic signal across datasets, but analytical tools for these comparisons are generally lacking. Here we develop a standardized effect size, Z_κ, which measures the strength of phylogenetic signal on a common statistical scale. We also provide a test statistic, $\hat{Z}_{12}$, for comparing the strength of phylogenetic signal across datasets. We find that two commonly used parameters (Pagel's λ and Blomberg's κ), not converted to effect sizes, are unsuitable for this purpose. Our effect-size procedure enables biologists to quantitatively address hypotheses that compare the strength of phylogenetic signal between various phenotypic traits, even when those traits are found in different evolutionary lineages or have different units or scales.

D.C.A. designed the research; D.C.A., E.K.B., and M.L.C. performed the research and wrote the paper.

The authors declare no conflict of interest.

[2]To whom correspondence should be addressed. E-mail: dcadamsiastate.edu

Fig. 1. Placeholder image of a frog with a long example caption to show justification setting.

e.g. (**?**) and (**? ?**). All references should be included in the main manuscript file.

Data Archival. PNAS must be able to archive the data essential to a published article. Where such archiving is not possible, deposition of data in public databases, such as GenBank, ArrayExpress, Protein Data Bank, Unidata, and others outlined in the Information for Authors, is acceptable.

Language-Editing Services. Prior to submission, authors who believe their manuscripts would benefit from professional editing are encouraged to use a language-editing service (see list at www.pnas.org/site/authors/language-editing.xhtml). PNAS does not take responsibility for or endorse these services, and their use has no bearing on acceptance of a manuscript for publication.

Digital Figures. Only TIFF, EPS, and high-resolution PDF for Mac or PC are allowed for figures that will appear in the main text, and images must be final size. Authors may submit U3D or PRC files for 3D images; these must be accompanied by 2D representations in TIFF, EPS, or high-resolution PDF format. Color images must be in RGB (red, green, blue) mode. Include the font files for any text.

Figures and Tables should be labelled and referenced in the standard way using the `\label{}` and `\ref{}` commands.

Figure **??** shows an example of how to insert a column-wide figure. To insert a figure wider than one column, please use the `\begin{figure*}...\end{figure*}` environment. Figures wider than one column should be sized to 11.4 cm or 17.8 cm wide. Use `\begin{SCfigure*}...\end{SCfigure*}` for a wide figure with side captions.

Tables. In addition to including your tables within this manuscript file, PNAS requires that each table be uploaded to the submission separately as a "Table" file. Please ensure that each table .tex file contains a preamble, the `\begin{document}` command, and the `\end{document}` command. This is necessary so that the submission system can convert each file to PDF.

Table 1. Comparison of the fitted potential energy surfaces and ab initio benchmark electronic energy calculations

Species	CBS	CV	G3
1. Acetaldehyde	0.0	0.0	0.0
2. Vinyl alcohol	9.1	9.6	13.5
3. Hydroxyethylidene	50.8	51.2	54.0

nomenclature for the TSs refers to the numbered species in the table.

Single column equations. Authors may use 1- or 2-column equations in their article, according to their preference.

To allow an equation to span both columns, use the `\begin{figure*}...\end{figure*}` environment mentioned above for figures.

Note that the use of the `widetext` environment for equations is not recommended, and should not be used.

Supporting Information (SI). Authors should submit SI as a single separate PDF file, combining all text, figures, tables, movie legends, and SI references. PNAS will publish SI uncomposed, as the authors have provided it. Additional details can be found here: policy on SI. For SI formatting instructions click here. The PNAS Overleaf SI template can be found here. Refer to the SI Appendix in the manuscript at an appropriate point in the text. Number supporting figures and tables starting with S1, S2, etc.

Authors who place detailed materials and methods in an SI Appendix must provide sufficient detail in the main text methods to enable a reader to follow the logic of the procedures and results and also must reference the SI methods. If a paper is fundamentally a study of a new method or technique, then the methods must be described completely in the main text.

SI Datasets. Supply Excel (.xls), RTF, or PDF files. This file type will be published in raw format and will not be edited or composed.

SI Movies. Supply Audio Video Interleave (avi), Quicktime (mov), Windows Media (wmv), animated GIF (gif), or MPEG files and submit a brief legend for each movie in a Word or RTF file. All movies should be submitted at the desired reproduction size and length. Movies should be no more than 10 MB in size.

3D Figures. Supply a composable U3D or PRC file so that it may be edited and composed. Authors may submit a PDF file but please note it will be published in raw format and will not be edited or composed.

Materials and Methods

Please describe your materials and methods here. This can be more than one paragraph, and may contain subsections and equations as required. Authors should include a statement in the methods section describing how readers will be able to access the data in the paper.

Subsection for Method. Example text for subsection.

ACKNOWLEDGMENTS. Please include your acknowledgments here, set in a single paragraph. Please do not include any acknowledgments in the Supporting Information, or anywhere else in the manuscript.

Fig. 2. This caption would be placed at the side of the figure, rather than below it.

$$\begin{aligned}(x+y)^3 &= (x+y)(x+y)^2 \\ &= (x+y)(x^2+2xy+y^2) \\ &= x^3+3x^2y+3xy^3+x^3.\end{aligned} \quad [1]$$