Phylotocol: Promoting Transparency and Overcoming Bias in Phylogenetics

Melissa B. DeBiasse[1,2] and Joseph F. Ryan[1,2]

[1]*Whitney Laboratory for Marine Science, St. Augustine, Florida, USA*

[2]*Department of Biology, University of Florida, Gainesville, Florida, USA*

Corresponding author: Joseph F. Ryan, 9505 Ocean Shore Blvd, St. Augustine, FL, 32080, USA.
Telephone: +1 904-461-4046; E-Mail: joseph.ryan@whitney.ufl.edu

ABSTRACT

The integrity of science requires that the process be based on sound experimental design and objective methodology. Strategies that increase reproducibility and transparency in science protect this integrity by reducing conscious and unconscious biases. Given the large number of analysis options and the constant development of new methodologies in phylogenetics, this field is one that would particularly benefit from more transparent research design. Here, we introduce phylotocol (fī·lō·´tə·kôl), an *a priori* protocol-driven approach in which all analyses are planned and documented at the start of a project. The phylotocol template is simple and the implementation options are flexible to reduce administrative burdens and allow researchers to adapt it to their needs without restricting scientific creativity. While the primary goal of phylotocol is to increase transparency and accountability, it has a number of auxiliary benefits

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26585v4 | CC BY 4.0 Open Access | rec: 23 Dec 2018, publ: 23 Dec 2018

including improving study design and reproducibility, enhancing collaboration and education, and increasing the likelihood of project completion. Our goal with this *Point of View* article is to encourage a dialogue about transparency in phylogenetics and the best strategies to bring transparent research practices to our field.

Keywords: accountability, confirmation bias, open science, phylogenetics, phylotocol, protocol, transparency

The production of reliable scientific results depends upon objective methodology. Reproducibility and transparency are safeguards against conscious and unconscious biases in scientific inquiry. The importance of reproducibility in science has been written about extensively over the past decade (King, R.D., Rowland, J., et al. 2009, McNutt, M. 2014, Markowetz, F. 2015), but its counterpart, transparency, has only recently begun to receive serious consideration (Ihle, M., Winney, I.S., et al. 2017, Nosek, B.A., Ebersole, C.R., et al. 2018). A reproducible study contains methods required to replicate all *reported* results, but it does not necessarily include all decisions that led to the final methodology reported in a study. Therefore, a reproducible study is not necessarily a transparent one. For example, if researchers present only a subset of results along with the methods required to generate those results (reporting bias), the study is technically reproducible, but lacks transparency. This lack of transparency is a problem across scientific disciplines, and is particularly applicable to phylogenetics.

Inferring relationships between genes, genomes, and species is essential for a fundamental understanding of biology. In the nearly 70 years since Hennig formalized phylogenetics (Hennig, W. 1965), the field has matured through the continuous development and improvement of algorithms, models, and data manipulation strategies (Whelan, S., Liò, P., et al. 2001) leading to many advances in phylogenetic methodology. However, the continual nature of methodological improvement and growing number of analysis options impedes standardization of experimental design. While as scientists we strive for objectivity and impartiality, we are all susceptible to conscious and unconscious biases (Kunda, Z. 1990, Christensen-Szalanski, J.J. and Willham, C.F. 1991, Nickerson, R.S. 1998, Pronin, E. and Kugler, M.B. 2007, Nosek, B.A. and Riskind, R.G. 2012), and implementing strategies to reduce the influence of these biases in our experiments is important for the integrity of science.

For most phylogenetic analyses, phylogeneticists are faced with a seemingly infinite combination of algorithms, models, and data manipulation techniques. Some examples related to tree reconstruction include: algorithms (e.g., distance, parsimony, maximum likelihood, and Bayesian inference (Felsenstein, J. 2004)), single-matrix models (e.g., JTT and WAG), criteria to determine model fit (e.g., AIC and BIC (Page, R.D. and Holmes, E.C. 2009)), partitioning and mixture model schemes (Blair, C. and Murphy, R.W. 2010), data filtering (e.g., removing unstable and quickly evolving taxa or genes (Salichos, L. and Rokas, A. 2013)). Other phylogenetic applications (e.g., molecular clock analyses, ancestral state reconstruction, hypothesis testing, and detection of selection) require researchers to make comparable decisions between competing approaches (Baum, D.A. and Smith, S.D. 2013).

In phylogenetics, research plans are generally informal and rarely outlined in detail prior to the start of the project; rather, plans are often constructed gradually, with each next step motivated by the results of the step before, an approach Gelman and Loken (2014) refer to as the “garden of forking paths.” This strategy is problematic because the selection of some paths is more likely than the selection of others, particularly if researchers make downstream methodological decisions consciously or, more often, unconsciously, in response to results that conflict with expected outcomes. Statistically the garden of forking paths is a problem because it makes correcting *p*-values for multiple comparisons impossible, rending them uninterpretable (Tukey, J. 1949, Dunnett, C. 1955).

In clinical trials, where the outcomes of a study can put human lives at risk, biases have been explicitly controlled for, and transparency and reproducibility ensured, through the requirement of *a priori* protocols that outline objective(s), design, methodology, statistical considerations, and study organization (Laine, C., Horton, R., et al. 2007, Zarin, D.A. and Tse, T. 2013, Zarin, D.A., Tse, T., et al. 2017). Protocols must be registered to a governmental regulatory agency, funding agency, and/or an institutional review board prior to the start of a study. Any changes (amendments) to a protocol require explicit justification and an updated version of the protocol (Getz, K.A., Stergiopoulos, S., et al. 2016). Many journals require protocols to be published with clinical trial publications, providing further motivation for their implementation. After the creation of the ClinicalTrials.gov registry (1997) led to the widespread adoption of transparent reporting standards in clinical trials, a dramatic drop in the frequency of positive results was observed, suggesting that bias may have been inflating the number positive results (Kaplan, R.M. and Irvin, V.L. 2015). Protocols greatly reduce, if not eliminate, the potential for researcher bias and in the process ensure the safety of subjects and the integrity of the trial.

Recently, *a priori* analysis plans and the preregistration of research designs have been proposed to promote transparency in the fields of Behavioral Ecology (Ihle, M., Winney, I.S., et al. 2017), Ecology and Evolution (Parker, T.H., Forstmeier, W., et al. 2016), and Psychology

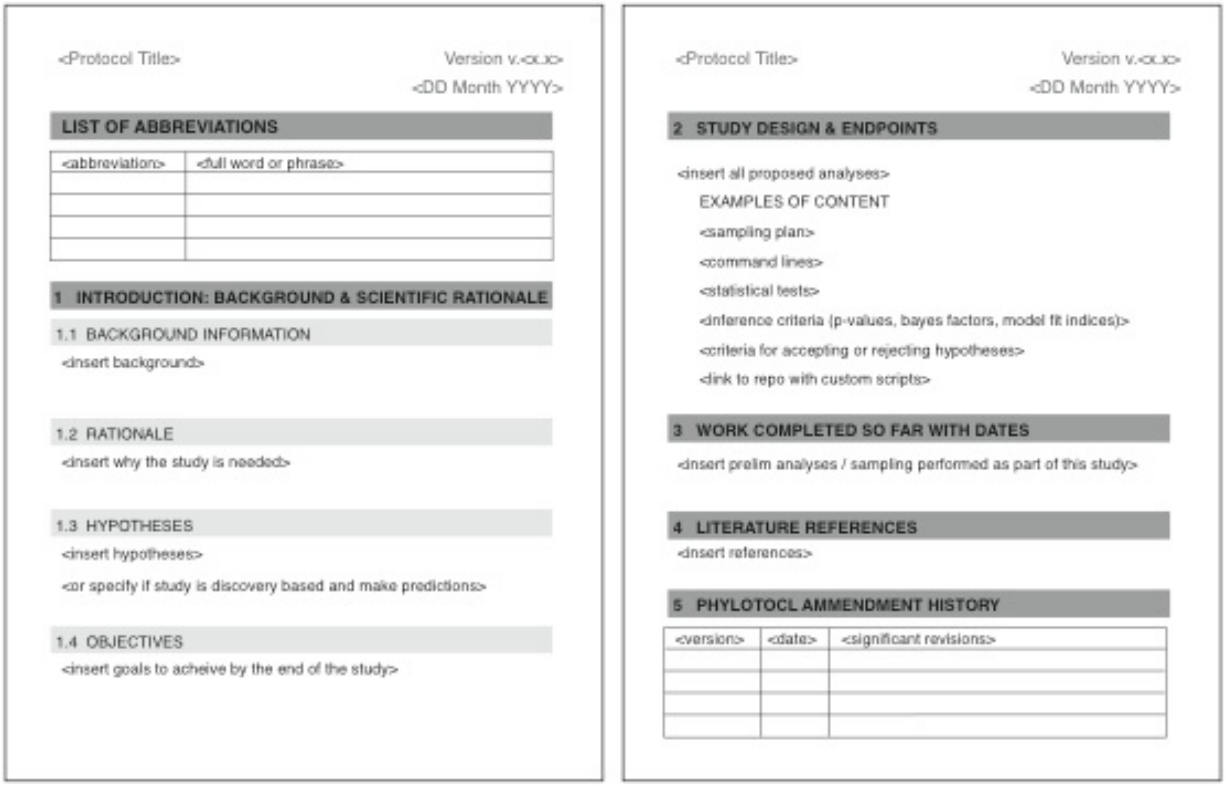
(Hartgerink, C. 2016) and a multidisciplinary working group has established a framework for minimal reporting standards (Aalbersberg, I.J., Appleyard, T., et al. 2018). The proposed measures are comparable to protocol registration in clinical trials and provide effective means to promote transparency in each particular field. Responses to these efforts have been positive (Blumstein, D.T. 2017, Forstmeier, W. 2017, Parker, T.H. and Nakagawa, S. 2017), negative (Koenig, W.D. 2017), and mixed (Cockburn, A. 2017, Hatchwell, B.J. 2017). The biggest barrier to widespread adoption to preregistration is the administrative effort associated with its implementation, perceived restrictions on scientific creativity and exploratory analyses, and concerns that project ideas will be scooped.

We argue that the field of phylogenetics would benefit tremendously from increased transparency. Here we introduce phylotocol, an *a priori* protocol-driven approach in which all analyses are planned and documented at the inception of a project, and optionally are preregistered. Phylotocol can be easily incorporated into phylogenetic studies; we have been using phylotocol since June 2017 and find it improves the rigor and efficiency of our research generally and our experimental design specifically. Here we describe a phylotocol template in detail, propose a set of guidelines for its use, include examples of phylotocols that we have implemented in our own research, and discuss how using a phylotocol can reduce bias and improve transparency and reproducibility in phylogenetics with minimal burdens on researchers' time. Our goal is to start a dialogue about the importance of transparency in phylogenetics and suggest ways to increase transparency and accountability in the field.

ANATOMY OF PHYLOTOCOL

The template phylotocol is based on the clinical trial protocol established by the National Institutes of Health (Hudson, K.L., Lauer, M.S., et al. 2016, NIH 2017a, NIH 2017b) and has seven major sections: 1) Title, 2) Abbreviations, 3) Introduction, 4) Study design, 5) Steps completed, 6) References, and 7) Appendix with version history (Fig. 1). This minimalist format reduces unnecessary burden, lowering the bar for implementation, but is flexible and can be customized to the requirements, preference, and computational expertise of a particular user. As opposed to a detailed template that might stifle creativity, the minimalist strategy is intended to foster the emergence of best practices, which we anticipate will evolve over time. Blank phylotocol templates in Microsoft Word and markdown formats and publicly posted phylotocols for research projects in the Ryan Lab are available at the following link (https://github.com/josephryan/phylotocol) and in the Supplementary Materials (Online Appendices 1-5).

A phylotocol is an outline of all decisions that could affect the final outcome of a study. Some common decisions include: (1) central hypotheses, (2) how taxa and data will be filtered, (3) which methods will be applied, (4) which models will be implemented, and (5) which criteria will be used to validate or reject hypotheses. While not required, we recommend including command lines and parameter settings (e.g., number of starting trees, seeds used for programs with random processes, minimum occupancy of phylogenomic matrices) to maximize clarity. Writing a phylotocol forces researchers to anticipate difficult decisions; for example, when applying different algorithms, models, etc. to the same data matrix, it is important to provide explicit criteria for evaluating conflicting results.



<Protocol Title> Version v.<x.x>
<DD Month YYYY>

LIST OF ABBREVIATIONS

<abbreviation>	<full word or phrase>

1 INTRODUCTION: BACKGROUND & SCIENTIFIC RATIONALE

1.1 BACKGROUND INFORMATION

<insert background>

1.2 RATIONALE

<insert why the study is needed>

1.3 HYPOTHESES

<insert hypotheses>

<or specify if study is discovery based and make predictions>

1.4 OBJECTIVES

<insert goals to acheive by the end of the study>

<Protocol Title> Version v.<x.x>
<DD Month YYYY>

2 STUDY DESIGN & ENDPOINTS

<insert all proposed analyses>

EXAMPLES OF CONTENT

<sampling plan>

<command lines>

<statistical tests>

<inference criteria (p-values, bayes factors, model fit indices)>

<criteria for accepting or rejecting hypotheses>

<link to repo with custom scripts>

3 WORK COMPLETED SO FAR WITH DATES

<insert prelim analyses / sampling performed as part of this study>

4 LITERATURE REFERENCES

<insert references>

5 PHYLOTOCL AMMENDMENT HISTORY

<version>	<date>	<significant revisions>

Figure 1 Phylotocol template. Based on the NIH clinical trial protocol, the phylotocol layout has been tailored to match the needs of phylogenetic research. The format and the information required are flexible. The figure displays the Microsoft Word version of the template, but there is also a markdown version. A phylotocol can be used as the basis for preregistration, uploaded to any online repository, or kept as a personal document (see Implementation section).

Ideally, a researcher would plan all steps in an analysis pipeline before testing is started, but in many cases, adjustments to the plan are needed once experiments are underway. The appendix section of the phylotocol is designed to accommodate changes to the analysis pipeline, for example, including an improved method that has recently been released, adding newly available data to a study, adjusting parameter settings, or correcting obvious mistakes. Each change should be accompanied by a justification and documentation of work completed so far, the latter making it possible to determine at which stage of a project a change was made.

PRIMARY OBJECTIVES OF PHYLOTOCOL

The primary objective of phylotocol is to increase transparency and accountability in phylogenetics. By outlining analyses *a priori*, phylotocol promotes transparency and reduces biases on the part of researchers. While many decisions made during the course of a study are obviously free of bias, and others clearly driven by bias are avoided by the majority of researchers, most decisions fall somewhere along this spectrum. By integrating transparency into a study, researchers provide readers with the ability to evaluate the validity of these decisions.

A transparent study reports all steps in the pipeline, even those that were replaced by other methods, or those that motivated downstream analyses but were not explicitly addressed in the final manuscript. In this way, phylotocol differs from traditional methods or supplementary methods sections, which typically only describe methodology for results that are reported in a manuscript. As is the case with supplementary materials, it is likely that a casual reader of the study will not be interested in the technical details supplied in a phylotocol; however, these details will be extremely important to researchers who are replicating or building upon the results of the study.

Accountability is a natural by-product of transparency (Mellor, D., Vazire, S., et al. 2018). In phylogenetics, as in other fields, it can be tempting to modify analyses when results conflict with our expectations. By implementing a phylotocol, researchers acknowledge that they are accountable for changes made during the period of a study and will be more motivated to deeply consider the implications of post-hoc decisions on the outcome of the analyses and the interpretation of the results.

AUXILIARY BENEFITS OF PHYLOTOCOL

While the primary goal of phylotocol is to increase transparency and accountability, the process offers a number of auxiliary benefits, which we describe below:

Designing a Better Study

Outlining each step of a study in a phylotocol before analyses are started can bring about a more robust plan. The process of transcribing procedures and guidelines for the interpretation of results can identify important steps and logical flaws that might otherwise be overlooked in a more patchwork experimental design. Catching these obstacles early in the process can lead to huge savings in time and/or money.

Documentation

Unlike in wet-lab based experimental biology, keeping a formal notebook to record the details of an analysis is less commonplace in phylogenetics. Creating a phylotocol that is updated throughout duration of a project helps serve many of the same purposes of a lab notebook. In this manner, a phylotocol serves as a key reference document for constructing the methods section of a manuscript.

Collaboration

Creating and executing a phylotocol can facilitate seamless collaborations among research groups. Getting input early from collaborators can strengthen a study while also ensuring that effort between collaborators does not overlap. Listing all steps also allows computational, personnel, budgetary, and/or other resource needs to be assessed. When collaborators agree on the analyses before a project is initiated, it helps prevent misunderstandings and/or conflicts down the line.

Education

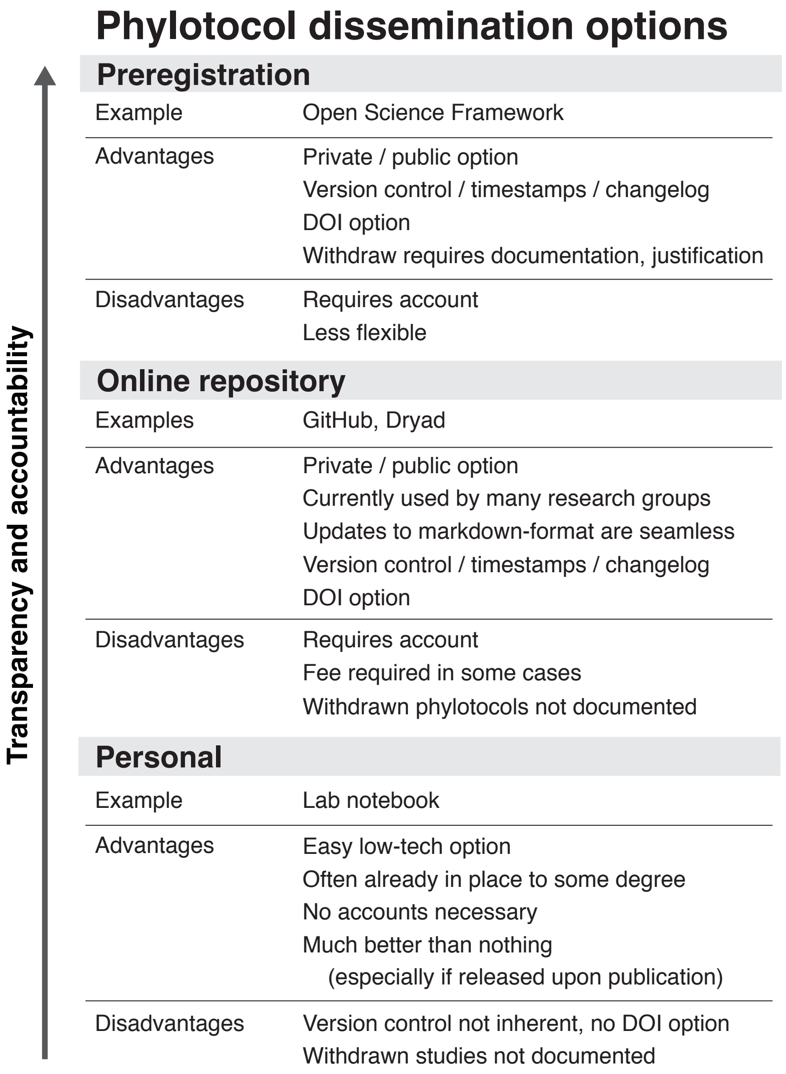
Phylotocol provides an excellent framework from which to train early career scientists. During the process of constructing a phylotocol, students gain a deeper understanding of the components of the study. Later, they have a roadmap from which to work throughout the project and mentors can be sure that effort is focused appropriately. Furthermore, previous phylotocols are useful references for new lab members who want to quickly get up to speed on how the lab performs particular analyses and can act as a template from which to start new analyses. Phylotocols can easily be incorporated into undergraduate and graduate courses as a tool to teach methodology, the importance of robust experimental design, and to reinforce the concepts of transparency and reproducibility in science.

Project Completion

The inherent open-endedness of science can often be intimidating and create a barrier to project completion. Implementing a phylotocol can remove this barrier by providing explicit starting and stopping points for a project and the motivation to complete the study as planned. The phylotocol quantifies the number of objectives a project requires and helps researchers prioritize each step. Beginning and completing a manuscript for the project will also be less daunting because the background information, study justification, methods, and references will already be compiled in the phylotocol. Starting new projects hinders the ability to complete existing projects; a phylotocol serves as a gentle impediment to spontaneously starting tangential projects and therefore increases productivity.

IMPLEMENTATION OF PHYLOTOCOL

There are several ways to implement phylotocol (Fig. 2). The option with the highest returns on transparency and accountability is preregistration with an organization such as the Open Science Framework (OSF, https://osf.io/) (Nosek, B.A., Alter, G., et al. 2015). If using OSF to post a phylotocol, we recommend choosing the “Open-Ended Registration” option and pasting a text version of phylotocol into the box. The OSF registry has an embargo system which keeps a registration private for up to four years, but ensures that a preregistered study is eventually released, whether published or not. A preregistration can be withdrawn but the title is still released and a justification is required. OSF also allows users to connect registrations to workflow management tools (e.g., Dataverse, Dropbox, figshare, Github, and others, see: http://help.osf.io/m/addons), so that contributions from different members of a research team can be connected, persistently stored, and cited in one location. When researchers are ready to disseminate early findings, any file on the OSF can be given a digital object identifier (DOI) and shared as a preprint (https://osf.io/preprints) prior to publication in a journal. One drawback to posting a phylotocol on OSF is that the original document cannot be edited. If changes to the phylotocol are needed, a new version must be uploaded. Another small drawback is that posting to OSF requires registering for an account and keeping track of credentials.



Phylotocol dissemination options

Preregistration	
Example	Open Science Framework
Advantages	Private / public option Version control / timestamps / changelog DOI option Withdraw requires documentation, justification
Disadvantages	Requires account Less flexible
Online repository	
Examples	GitHub, Dryad
Advantages	Private / public option Currently used by many research groups Updates to markdown-format are seamless Version control / timestamps / changelog DOI option
Disadvantages	Requires account Fee required in some cases Withdrawn phylotocols not documented
Personal	
Example	Lab notebook
Advantages	Easy low-tech option Often already in place to some degree No accounts necessary Much better than nothing (especially if released upon publication)
Disadvantages	Version control not inherent, no DOI option Withdrawn studies not documented

Figure 2 Implementation options. There are three frameworks for implementing a phylotocol, each with increasing returns on transparency and accountability, as indicated by the arrow. For each framework, an example strategy is listed with its associated advantages and disadvantages. The preregistration framework provides a superior level of transparency, but the repository and personal frameworks still provide benefits and are especially useful for getting started with phylotocol.

A second way to implement phylotocol is to post the document to an online software or data repository, such as GitHub or Dryad. Many users have experience with one or more of these repositories, so the learning curve with this option is minimal. The specific features of different online repositories vary, but most have a timestamp feature to provide transparency as to when a phylotocol is posted and edited, version control, which allows for seamless updating (especially when implementing a markdown version of phylotocol), and DOI assignment. Most repositories allow documents to remain private, but a drawback in terms of promoting transparency is that if a study is discontinued or substantially changed, there is no requirement to release the phylotocol and/or justify the retraction. This could present a transparency problem if a future work relies on data generated as part of an unfinished study. Like preregistration, online repositories also have the minor inconvenience of requiring users to create an account and keep track of credentials.

The third way to implement a phylotocol is to create a private document on a personal computer or in a lab notebook. This is the most simple, low tech, and flexible option and does not require making an account or remembering a password. This strategy lacks the built-in version control and timestamp features of the above options, which is a disadvantage (although version control software can be implemented secondarily). In addition, like a phylotocol privately posted to an online repository, there is no requirement that a document kept in a lab notebook be made public, limiting the transparency of the process. However, this strategy can greatly increase the transparency of a project and researchers who choose this option will greatly benefit from implementing a personal phylotocol.

The multiple flexible options for implementing phylotocol, each with various levels of commitment, make it easy to try out the process. Researchers interested in incorporating more transparent practices in their research could ease into phylotocol by first making private documents for their own use. Once familiar with the process, they can transition to posting the phylotocol to an online repository, and then move towards preregistration, which is the gold standard for transparency and accountability. Each step along this progression requires a higher level of commitment, but we predict that the structure will serve many researchers well.

DISCUSSION

The production of reliable and bias-free results is an indisputable goal of all phylogenetic studies. By planning analyses before a study begins, and making methodological choices transparent, a phylotocol reduces the likelihood of confirming a false hypothesis. A phylotocol, therefore, makes considerable contribution towards reaching the goal of strong, bias-free research results.

The idea of including additional steps to an already time-consuming research process will almost certainly be met with hesitation, if not objection, but we contend that the time spent on phylotocol is easily recovered both in the short and long term. In practice, we have found that time invested in phylotocol pays dividends downstream, particularly when training junior researchers, writing manuscripts, and keeping projects on track towards completion. In the long run, wide adoption of phylotocol will lead to less confirmation bias in the scientific record and therefore huge savings in time that would otherwise be spent building upon or rebutting questionable results.

A major concern is that implementing a phylotocol will stifle scientific creativity and data exploration (Koenig, W.D. 2017). We contend that phylotocol and creative data exploration are not mutually exclusive, and that in some ways, phylotocol enhances the creative process. Writing a phylotocol explicitly requires that researchers dedicate time to planning a study start to finish, which can be an inherently creative process, potentially more so than planning the analyses haphazardly or informally. Furthermore, phylotocol does include built-in support for unplanned exploratory analyses through appendix updates. Decisions to add, change, or disregard planned analyses require only that changes be documented and justified.

As scientists, our ultimate goal is to make discoveries and formulate theories that stand up to rigorous testing, and eventually become widely accepted as truth. The possibility that bias can inadvertently influence our research results should not be minimized or neglected. By implementing phylotocol phylogeneticists will show dedication to scientific integrity, which will lead to confidence in the reliability of their work. In this way, transparent research practices like phylotocol help maximize research impact.

CONCLUSION

Phylotocol is a powerful tool to increase transparency and accountability in phylogenetics. It has great potential to improve how phylogenetic research is conducted, interpreted, communicated, and perceived. The implementation is straightforward and offers a range of auxiliary benefits, including making contributions to study design, reproducibility, collaboration, and education. Phylotocol can bolster scientific productivity both at the level of the individual researcher as well as in the broader context of the scientific record. While phylotocol is a simple idea, its repercussions could be far reaching if widely implemented.

ACKNOWLEDGEMENTS

We thank Malika Ihle, Jessica Whelpley, Daniel Sasson, David Mellor, Tim Parker, Lyndon Coghill, Daniel Cooke, Jacob Esselstyn, Teisha King, Laura Lagomarsino, Rafael Marcondes, Genevieve Mount, Zachary Rodriguez, and Mark Swanson, for comments on earlier versions of the manuscript. We thank Bryan Carstens for including phylotocol in a forum on Data Standards and Community Resources Discussion at the 2018 meeting of the Society of Systematic Biologists, and for feedback from attendees at that meeting. The views expressed in this paper do not necessarily reflect the views of those acknowledged. This material is based in part upon work supported by the National Science Foundation under Grant Number 1542597. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

LITERATURE CITED

Aalbersberg IJ, Appleyard T, Brookhart S, Carpenter T, Clarke M, Curry S, Dahl J, DeHaven A, Eich E, Franko M. 2018. Making Science Transparent By Default; Introducing the TOP Statement.

Baum DA, Smith SD. 2013. Tree thinking: an introduction to phylogenetic biology. Roberts Greenwood Village, CO.

Blair C, Murphy RW. 2010. Recent trends in molecular phylogenetic analysis: where to next? Journal of Heredity, 102:130-138.

Blumstein DT. 2017. Research credibility: the devil is in the details: a comment on Ihle et al. Behavioral Ecology, 28:355-355.

Christensen-Szalanski JJ, Willham CF. 1991. The hindsight bias: A meta-analysis. Organizational behavior and human decision processes, 48:147-168.

Cockburn A. 2017. Long-term data as infrastructure: a comment on Ihle et al. Behavioral Ecology, 28:357-357.

Dunnett C. 1955. A multiple comparison procedure for comparing several treatments with a control. Journal of the American Statistical Association, 50:1096-1121.

Felsenstein J. 2004. Inferring phylogenies. Sinauer associates Sunderland, MA.

Forstmeier W. 2017. Preregister now for an upgrade to Behavioral Ecology 2.0: a comment on Ihle et al. Behavioral Ecology, 28:358-359.

Gelman A, Loken E. 2014. The statistical crisis in science. American scientist, 102:460.

Getz KA, Stergiopoulos S, Short M, Surgeon L, Krauss R, Pretorius S, Desmond J, Dunn D. 2016. The impact of protocol amendments on clinical trial performance and cost. Therapeutic Innovation & Regulatory Science, 50:436-441.

Hartgerink C. 2016. Research practices and assessment of research misconduct. ScienceOpen Research.

Hatchwell BJ. 2017. Replication in behavioural ecology: a comment on Ihle et al. Behavioral Ecology, 28:360-360.

Hennig W. 1965. Phylogenetic systematics. Annual review of entomology, 10:97-116.

Hudson KL, Lauer MS, Collins FS. 2016. Toward a new era of trust and transparency in clinical trials. Jama, 316:1353-1354.

Ihle M, Winney IS, Krystalli A, Croucher M. 2017. Striving for transparent and credible research: practical guidelines for behavioral ecologists. Behavioral Ecology, 28:348-354.

Kaplan RM, Irvin VL. 2015. Likelihood of null effects of large NHLBI clinical trials has increased over time. PLoS One, 10:e0132382.

King RD, Rowland J, Oliver SG, Young M, Aubrey W, Byrne E, Liakata M, Markham M, Pir P, Soldatova LN. 2009. The automation of science. Science, 324:85-89.

Koenig WD. 2017. Striving for science that is transparent, credible—and enjoyable: a comment on Ihle et al. Behavioral Ecology, 28:358-358.

Kunda Z. 1990. The case for motivated reasoning. Psychological bulletin, 108:480.

Laine C, Horton R, DeAngelis CD, Drazen JM, Frizelle FA, Godlee F, Haug C, Hébert PC, Kotzin S, Marusic A. 2007. Clinical trial registration—looking back and moving ahead. Mass Medical Soc.

Markowetz F. 2015. Five selfish reasons to work reproducibly. Genome biology, 16:274.

McNutt M. 2014. Reproducibility. American Association for the Advancement of Science.

Mellor D, Vazire S, StephenLindsay D. 2018. Transparent science: A more credible, reproducible, and publishable way to do science.

Nickerson RS. 1998. Confirmation bias: A ubiquitous phenomenon in many guises. Review of general psychology, 2:175.

NIH. 2017a. National Institutes of Health Word version of final clinical trials protocol.

NIH. 2017b. NIH and FDA Request for Public Comment on Draft Clinical Trial Protocol Template for Phase 2 and 3 IND/IDE Studies.

Nosek BA, Alter G, Banks GC, Borsboom D, Bowman SD, Breckler SJ, Buck S, Chambers CD, Chin G, Christensen G. 2015. Promoting an open research culture. Science, 348:1422-1425.

Nosek BA, Ebersole CR, DeHaven AC, Mellor DT. 2018. The preregistration revolution. Proceedings of the National Academy of Sciences:201708274.

Nosek BA, Riskind RG. 2012. Policy implications of implicit social cognition. Social Issues and Policy Review, 6:113-147.

Page RD, Holmes EC. 2009. Molecular evolution: a phylogenetic approach. John Wiley & Sons.

Parker TH, Forstmeier W, Koricheva J, Fidler F, Hadfield JD, Chee YE, Kelly CD, Gurevitch J, Nakagawa S. 2016. Transparency in ecology and evolution: real problems, real solutions. Trends in Ecology & Evolution, 31:711-719.

Parker TH, Nakagawa S. 2017. Practical models for publishing replications in behavioral ecology: a comment on Ihle et al. Behavioral Ecology, 28:355-357.

Pronin E, Kugler MB. 2007. Valuing thoughts, ignoring behavior: The introspection illusion as a source of the bias blind spot. Journal of Experimental Social Psychology, 43:565-578.

Salichos L, Rokas A. 2013. Inferring ancient divergences requires genes with strong phylogenetic signals. Nature, 497:327.

Tukey J. 1949. Comparing individual means in the analysis of variance. Biometrics, 5.

USFDA. 1997. Food and Drug Administration Modernization Act of 1997. Public Law, 105:111.

Whelan S, Liò P, Goldman N. 2001. Molecular phylogenetics: state-of-the-art methods for looking into the past. TRENDS in Genetics, 17:262-272.

Zarin DA, Tse T. 2013. Trust but verify: trial registration and determining fidelity to the protocol. Annals of internal medicine, 159:65-67.

Zarin DA, Tse T, Williams RJ, Rajakannan T. 2017. Update on trial registration 11 years after the ICMJE policy was established. New England Journal of Medicine, 376:383-391.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.26585v4 | CC BY 4.0 Open Access | rec: 23 Dec 2018, publ: 23 Dec 2018