1 **Title:** The concept of homoiology and its potential application in
2 palaeontological phylogenetics

3 **Author name:** Roland. B. Sookias
4

5 **Author address:** Museum für Naturkunde, Invalidenstraße 43, Berlin 10115, Germany,
6 sookias.r.b@gmail.com
7

8 **Abstract**
9

10 The neglected concept of homoiology is discussed in the context of palaeontological
11 phylogenetic methods. Homoiology is the case of homoplasy where a parallel or convergent
12 phenotypic development is actually due to shared ancestry, whether a common inherited
13 genetic mechanism or simply a shared initial phenotype. A number of
14 parallelisms/convergences in vertebrate and linguistic evolution are discussed to illustrate
15 the concept of homoiology. It is proposed that parallelisms/convergences, although
16 probably not useful in initial tree inference from morphological data, must not be simply
17 written off as phylogenetically uninformative, but can actually be potentially used to provide
18 additional support values for clades and to help to choose between equally parsimonious or
19 likely trees. An R function is provided to calculate two measures for a given tree and matrix:
20 (a) the potential support for clades based on potential homoiologies; and (b) the fit of the
21 tree to all states given the concept of homoiology. The relationship between homoiology
22 and constraints and to underlying mechanisms is also discussed.
23

24 **Keywords:** homoiology, homology, parallelism, convergence, phylogeny, constraints,
25 evolution
26

27 **Introduction**

28 In this short piece I wish to discuss the somewhat neglected concept of homoiology, and
29 detail a simple method for how it can be potentially useful in morphological, and thus
30 palaeontological, phylogenetic systematics. Homoiology is historically roughly a synonym for
31 parallel evolution (=parallelism), being defined by Plate [1] thus:

32 *“Homoiologie = unabhängig erworbene morphologische und physiologische Ähnlichkeit oder*
33 *Gleichheit bei Tieren, welche sich unabhängig voneinander von derselben Stammform*
34 *ableiten. Sie besitzen eine Anzahl gleicher Erbfaktoren und haben daher die Tendenz, sich in*
35 *ähnlicher Weise weiter zu entwickeln”* (p. 7)
36

37 (homoiology = independently acquired morphological and physiological similarity or
38 sameness in animals, which originate independently from the same parent form. They have

many similar hereditary factors [=genes] and therefore have a tendency to evolve in a similar manner)

Parallel evolution classically differs from convergent evolution in that in parallelism, similar, separately arising features develop from homologous structures, whereas in convergence similar features develop based on structures which are not homologous [2]. However the line between parallelism and convergence is increasingly understood to be blurred - for example, the convergent evolution of eyes in disparate groups in fact involved homologous pigment genes [2, 3] – and a continuum concept is now preferred [2, 4]. However, the fundamental theoretical difference in the concepts is arguably that parallelism considers that common ancestry and environment were causal in the development of a similar phenotype, whereas convergence considers only the environment to have been causal.

Homoiology appears to have been originally coined more or less as a synonym for parallelism, and thence rather neglected (excepting for its very different meaning in anthropology [5]). However I would here like to propose its usage to denote a concept akin to homology, but specifically for features arising separately but due to shared ancestry. Cases of the same genetic information underlying parallel phenotypes are an example of this [6], with the phenotypes thus homoiologues. However, the concept can also be applied to inheritance of a similar ancestral phenotype which then gives rise to parallel phenotypes, but not necessarily with the same genetic mechanisms. In a cladistic, parsimony-based framework these would be treated simply as homoplasies. However, some of these homoplasies are potentially useful in phylogenetic inference, and were explicitly distinguished from other homoplasies by Hennig [7], based on Plate's [8] definition (Fig. 1).

Some examples of homoiology

More or less any example of homoplasy can be framed as a potential homoiology in so far as all life has a common ancestor, but the concept is only of interest where it potentially groups lineages due to a shared ancestral form or shared genotypic/developmental inheritance.

A good example of a case where a parallelism could actually be phylogenetically informative is the loss of limbs in squamates, which has occurred independently some 25 times [9], and which could be considered to support squamate monophyly, despite the clearly homoplastic nature of this development. Similarly, the body plan of grebes and divers appears to be convergent, and leads to them being mistakenly grouped as sister taxa using morphology [10-12]. However, at a deeper level, both clades do appear to be united within Aequiornithes [13, 14], and their similarities may be due to their shared inheritance of a common body plan, which was then subjected to similar environmental pressures (Fig. 2).

Another classic case of convergent evolution is the development of very similar "insectivore-grade" mammals (moles and golden moles, hedgehogs and tenrecs) in the Laurasiatheria

and Afrotheria [15-17]. Although this convergence appears startling, it is not surprising given a probably similar initial form (the ancestor of both groups probably being small and at least partially insectivorous) [15], and the fact that both clades are actually close relatives in the broader scheme of life. Thus, although the convergent spines of tenrecs and hedgehogs cannot group them to the exclusion of other laurasiatheres and afrotheres, they do provide support for eutherian monophyly. From another evolving system, to which phylogenetic methods are widely applied [18, 19] – languages – an example would be the development of presumed historical /i:/ and /u:/ sounds to the diphthongs /ai/ and /aʊ/ in both English and German; the change occurred separately [20], but may provide support to a grouping of the West Germanic languages, and indicate a similar ancestral form or inherited feature which precipitated changes in both languages.

Potential use of homoiologies in morphological phylogenetics

In inference of phylogeny by parsimony, an occurrence of a character state in a part of a tree separated from it by another state is considered simply a homoplasy, and a tree where the occurrences are nearer or further from one another is not more or less parsimonious. However, as can hopefully be seen from the examples discussed above, this approach may end up discarding useful information. Thus it is proposed that the concept of homoiology could be explicitly incorporated into methods of assessing clade support, and choosing between equally parsimonious/likely trees, although not in tree inference itself.

Given the monophyly of life, at some level, all separately developing similar phenotypes are due to common ancestry [2]. Thus incorporating the concept of homoiology in initial tree inference is not advisable, because any feature could be considered indicative of common ancestry at some level, making phylogenetic inference impossible. However, following initial tree inference, homoiologies may provide useful information. One approach to incorporate information from homoiologies using parsimony could be to first infer the tree using a standard approach. Subsequently, the number of character states bracketed by a node could be counted, and those which do not optimize as symplesiomorphies of the clade could be considered as a value of node support (Fig. 3A – black state is potential homoiology of blue node). This “homoiology index” would provide another level of assessment of the support of groupings, even if it did not affect the groupings themselves. Homoiology could also be employed to choose between equally parsimonious trees, with those where the clade containing all instances of a state is smallest being preferred (Fig. 3B as opposed to Fig. 3C).

Both of these approaches are implemented in an R [21] function, based on the ape [22], phytools [23] and phangorn [24] packages, given as supplementary information with this work. The function returns a list of all nodes bracketing character states which are not optimized as the ancestral state for that node, along with the character number and state number/name. Maximum parsimony ancestral state reconstruction is used, and the cutoff of probability can be adjusted, with the default being that states are not considered if the

1 probability of their being the ancestral state of that node is >0 . Also returned is the average
2 number of taxa in the clade bracketed by each state in the dataset; if this is higher, the tree
3 can be considered to fit the data less well.

4
5 Another approach using likelihood or Bayesian probabilistic phylogenetic inference would be
6 to explicitly incorporate clade size bracketing a state or nodal distance between instances of
7 the same state into the likelihood/posterior probability estimate for a tree. What weight in
8 the likelihood calculation the distances should be accorded is unclear however, and a less
9 controversial approach might be to reduce the tree sample *post hoc* to only those trees with
10 the minimum average clade size bracketing all states, and compare this to the original. Nodal
11 probability values could also be recalculated incorporating homoiology, following initial
12 inference.

13 14 **Further theoretical considerations**

15
16 As mentioned above, I consider that a homoiology must not indicate a homologous genetic
17 mechanism (often used to distinguish parallelism from “mere” convergence - [25]). If two
18 taxa share a similar ancestral form, this will tend to mean they will share similar genes, but it
19 is the phenotype – upon which selection acts – which is crucial. Thus, for example, the
20 genetic mechanism or diving adaptations in grebes and divers must not be homologous – the
21 shared, inherited ancestral form (and genes) meant similar adaptations occurred, even if
22 these did not share the same genetic basis. Just as two trajectories starting close to one
23 another are likely to end up close to each other, so may two phenotypes (Fig. 4). Taking a
24 classic mouse example, development of a similar phenotype by different genetic
25 mechanisms [25] can be considered homoiologous, because selection yielded a convergent
26 result because the starting “base” was similar morphologically. Another consideration is
27 whether this shared ancestral body plan/phenotype are equivalent to what are described as
28 evolutionary constraints (e.g. [26, 27]). The answer, I suggest, is that constraints may
29 increase the chance of homoiologies occurring because they will limit potential evolutionary
30 directions, but that homoiologies can also be precipitated by non-constraining similar initial
31 forms, because by chance similar morphology will be likely if starting points are close (Fig. 4).

32 33 **Conclusions**

34
35 With this brief discussion it is hoped to highlight the rather neglected concept of
36 homoiology, to encourage its more general discussion and consideration, and to propose
37 means for its potential use in phylogenetic estimation. Rather than simply throwing away
38 any case of convergence/parallelism as “just” homoplasy, these occurrences – which
39 arguably can be considered the norm rather than the exception in related taxa – may
40 provide useful evidence of relatedness.

Acknowledgments

The author would like to thank E. Ascarrunz (Université de Fribourg) for discussion, E. Amson and F. Bibi (both Museum für Naturkunde), and L. Revell and K. Schliep (both University of Massachusetts Boston) with technical assistance as discussion, and J. Müller (Museum für Naturkunde) for hosting the author. The author was funded by a Humboldt Research Fellowship for postdoctoral researchers from the Alexander von Humboldt Foundation during completion of this work.

References

- [1] Plate, L. 1922 *Allgemeine Zoologie und Abstammungslehre: 1. Teil*. Jena, Gustav Fischer Verlag.
- [2] McGhee, G.R. 2011 *Convergent evolution: limited forms most beautiful*, MIT Press.
- [3] Rey, M., Ohno, S., Pintor-Toro, J.A., Llobell, A. & Benitez, T. 1998 Unexpected homology between inducible cell wall protein QID74 of filamentous fungi and BR3 salivary protein of the insect *Chironomus*. *Proc. Nat. Acad. Sci.* **95**, 6212-6216.
- [4] Arendt, J. & Reznick, D. 2008 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26-32.
- [5] von Cramon-Taubadel, N. 2009 Revisiting the homoiology hypothesis: the impact of phenotypic plasticity on the reconstruction of human population history from craniometric data. *J. Human Evol.* **57**, 179-190.
- [6] Hu, Y., Wu, Q., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L. & Xiu, Y. 2017 Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. *Proc. Nat. Acad. Sci.* **114**, 1081-1086.
- [7] Hennig, W. & Davis, D.D. 1999 *Phylogenetic systematics*, University of Illinois Press.
- [8] Plate, L. 1928 Ueber Vervollkommnung Anpassung und Unterscheidung von niederen und höheren Tieren. *Zool Jahrb* **45**, 745-798.
- [9] Wiens, J.J., Brandley, M.C. & Reeder, T.W. 2006 Why does a trait evolve multiple times within a clade? Repeated evolution of snakelike body form in squamate reptiles. *Evolution* **60**, 123-141.
- [10] Mayr, G. 2004 Morphological evidence for sister group relationship between flamingos (Aves: Phoenicopteridae) and grebes (Podicipedidae). *Zool. J. Linn. Soc.* **140**, 157-169.
- [11] Van Tuinen, M., Butvill, D.B., Kirsch, J.A. & Hedges, S.B. 2001 Convergence and divergence in the evolution of aquatic birds. *Proc. R. Soc. B.* **268**, 1345-1350.
- [12] Manegold, A. 2006 Two additional synapomorphies of grebes Podicipedidae and flamingos Phoenicopteridae. *Acta Ornithologica* **41**, 79-82.
- [13] Jarvis, E.D., Mirarab, S., Aberer, A.J., Li, B., Houde, P., Li, C., Ho, S.Y., Faircloth, B.C., Nabholz, B. & Howard, J.T. 2014 Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science* **346**, 1320-1331.
- [14] Prum, R.O., Berv, J.S., Dornburg, A., Field, D.J., Townsend, J.P., Lemmon, E.M. & Lemmon, A.R. 2015 A comprehensive phylogeny of birds (Aves) using targeted next-generation DNA sequencing. *Nature* **526**, 569.
- [15] Rasnitsyn, A. 2006 Ontology of evolution and methodology of taxonomy. *Paleontol. J.* **40**, S679-S737.
- [16] Stanhope, M.J., Waddell, V.G., Madsen, O., De Jong, W., Hedges, S.B., Cleven, G.C., Kao, D. & Springer, M.S. 1998 Molecular evidence for multiple origins of Insectivora and for a new order of endemic African insectivore mammals. *Proc. Nat. Acad. Sci.* **95**, 9967-9972.
- [17] Madsen, O., Scally, M., Douady, C.J., Kao, D.J., DeBry, R.W., Adkins, R., Amrine, H.M., Stanhope, M.J., de Jong, W.W. & Springer, M.S. 2001 Parallel adaptive radiations in two major clades of placental mammals. *Nature* **409**, 610. (doi:10.1038/35054544).

- [18] Atkinson, Q.D. & Gray, R.D. 2005 Curious parallels and curious connections—phylogenetic thinking in biology and historical linguistics. *Syst. Biol.* **54**, 513-526.
- [19] Gray, R.D. & Atkinson, Q.D. 2003 Language-tree divergence times support the Anatolian theory of Indo-European origin. *Nature* **426**, 435.
- [20] Liberman, A. 1995 The beginning and end of the (great) vowel shifts in Late Germanic In *Insights in Germanic Linguistics: I Methodology in Transition* (ed. R.I. Gerald F. Carr), pp. 203-230. Berlin, Mouton De Gruyter.
- [21] R-Core-Team. 2014 R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. (
- [22] Paradis, E., Claude, J. & Strimmer, K. 2004 APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* **20**, 289-290.
- [23] Revell, L.J. 2012 phytools: an R package for phylogenetic comparative biology (and other things). *Methods in Ecology and Evolution* **3**, 217-223.
- [24] Schliep, K.P. 2011 phangorn: phylogenetic analysis in R. *Bioinformatics* **27**, 592.
- [25] Hoekstra, H.E., Hirschmann, R.J., Bunday, R.A., Insel, P.A. & Crossland, J.P. 2006 A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science* **313**, 101-104.
- [26] Futuyma, D.J. 2010 Evolutionary constraint and ecological consequences. *Evolution* **64**, 1865-1884.
- [27] Sharov, A.A. 2014 Evolutionary constraints or opportunities? *Biosystems* **123**, 9-18.
- [28] Narváez, I., Brochu, C.A., Escaso, F., Pérez-García, A. & Ortega, F. 2015 New crocodyliforms from southwestern Europe and definition of a diverse clade of European Late Cretaceous basal eusuchians. *PLoS ONE* **10**, e0140679.
- [29] Maddison, W.P. & Maddison, D.R. 2017 Mesquite: a modular system for evolutionary analysis. Version 3.2. (<http://mesquiteproject.org>).

Figures

Figure 1. Homoiology as illustrated by Hennig (Hennig and Davis, 1999). The derived form a' develops independently in two closely related taxa from the same original form (a).

Figure 2. Schematic diagram showing how the ancestor of grebes and divers (and flamingos) may have had a similar morphology, despite flamingos and grebes having diverged after either diverged from divers. This scenario makes it unsurprising that a similar morphology could have developed in divers and grebes in response to similar selective pressures, and this final similar morphology can actually be regarded in some way as homologous (i.e. homoiologous), because it is due to a shared ancestral morphology.

Fig.3. Figure showing how homoiology can be used in parsimony analysis to (A) provide additional node support values, and (B,C) choose between trees, as implemented in the function `homoilogies()`. (A) shows a DNA-based phylogeny (with fossil outgroup) for extant crocodilians (given in the supplement as a tree), with character 97 of Narváez et al. [28] (character 66 in supplemental matrix) plotted on it using parsimony ancestral states in Mesquite [29]; state 1 (black; presence of very prominent preorbital ridges) is a potential homoiology of Crocodylinae (blue node). In (B), the smallest clade bracketing all instances of state 0 (pink), indicated by blue node, is smaller than in (C), with the concept of homoiology meaning that tree C should be preferred (only based on this character – an average of all character states should be taken) if both are equally parsimonious for all characters.

Fig. 4. Visualisation of homology and convergence as trajectories in morphospace through time. Each line indicates a taxon, which diverged from a common ancestor (black or grey) at $t=0$. If morphological starting points are the same or close, development of similar morphology is more likely (“parallelism”), and if they are different/distant, less likely (“convergence”).