The Nature and Phylogenomic Impact of Sequence Convergence

Zhengting Zou, Jianzhi Zhang

Chapter 4.6 The Nature and Phylogenomic Impact of Sequence Convergence

Zhengting Zou
Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA
ztzou@umich.edu
https://orcid.org/0000-0003-1716-5090

Jianzhi Zhang
Department of Ecology and Evolutionary Biology, University of Michigan, Ann Arbor, Michigan 48109, USA
jianzhi@umich.edu
https://orcid.org/0000-0001-6141-1290

Abstract

Protein sequence convergence refers to substitutions leading to the same amino acid residue at the same position of a protein in multiple independent evolutionary lineages. Protein sequence convergence is often viewed as adaptive signal so is of great interest to evolutionary biologists. In this article, we review complications in identifying sequence convergences, statistical tests of the null hypothesis that the observed convergence events in a protein are attributable to chance alone, interpretations of genome-wide observations of sequence convergence, and a comparison in the susceptibility of molecular and morphological characters to convergence and its phylogenetic implications. We highlight the substantial progresses made in the last two decades and point out the main challenges at the present.

How to cite: Zhengting Zou and Jianzhi Zhang (2020). The Nature and Phylogenomic Impact of Sequence Convergence. In Scornavacca, C., Delsuc, F., and Galtier, N., editors, *Phylogenetics in the Genomic Era*, chapter No. 4.6, pp. 4.6:1–4.6:17. No commercial publisher | Authors open access book. The book is freely available at https://hal.inria.fr/PGE.

1 Introduction

Convergent evolution, or simply convergence, refers to the independent emergences of the same state of a character in two or more lineages of living organisms. Well-known examples of convergence include the origins of camera-type eyes in cephalopods and vertebrates, and the emergences of wings from forelimbs in birds and bats. Evolutionary biologists are interested in convergence primarily for three reasons. First, because complex characteristics such as camera-type eyes and wings are unlikely to have emerged more than once simply by chance, convergent evolution of complex characteristics is believed to reflect similar adaptations in multiple lineages. Second, convergence indicates that evolution is predictable to some extent, either because there are few viable solutions to a problem or the best solutions are similar in different lineages. Third, convergence confuses phylogenetic analysis, because true phylogenetic signals are based on identity by descent, which, however, is not easy to distinguish from false signals of identity by convergence.

The study of convergence has a long history. Convergence was already discussed in Darwin's Origin of Species as "analogical resemblances"; examples mentioned included body shape and fin-like forelimb of dugongs and whales, morphological resemblance between

Phylogenetics in the genomic era.
Editors: Celine Scornavacca, Frédéric Delsuc and Nicolas Galtier; chapter No. 4.6; pp. 4.6:1–4.6:17

A book completely handled by researchers.
No publisher has been paid.

European and Asian domestic pigs, and electric organs in different fish lineages (Darwin, 1859). Numerous morphological and functional convergences have since been reported, such as similar web architectures of spiders occupying the same habitat type on different Hawaiian islands (Blackledge and Gillespie, 2004), similar bill shape shifts of tidal marsh sparrows in North America (Grenier and Greenberg, 2005), morphological similarities among trunk-ground dwelling anoles on multiple Greater Antillean islands (Langerhans et al., 2006), and intercontinental pairs of desert iguana species with matching habitats (Melville et al., 2006). The list goes on and on with examples across virtually the whole tree of life (Nevo, 1979; Moore and Willmer, 1997; Wittkopp et al., 2003; Fong et al., 2005; Maruyama and Parker, 2017). As our understanding of biology progresses to the molecular level, convergence has also been discovered at the level of molecular phenotypes. For example, the independently evolved pATOM36 protein and MIM complex serve as importers on the mitochondrial outer membrane in trypanosomes and yeast, respectively (Vitali et al., 2018), and rhodopsins of vertebrates and of the visually competent box jellyfish have acquired similar tertiary structures to enable high-fidelity photoreception (Gerrard et al., 2018).

Convergence can occur not only at the phenotypic level but also at the molecular genetic level (Stewart et al., 1987; Doolittle, 1994; Arendt and Reznick, 2008; Manceau et al., 2010). This can include, for example, amino acid sequence changes at different sites of the same protein across multiple lineages (Protas et al., 2006; Rosenblum et al., 2010; Linnen et al., 2013; Zhou et al., 2015; Chikina et al., 2016), or independent formations of a chromosomal cluster of the same set of genes via relocation (Slot and Rokas, 2010). The most studied convergence at the molecular genetic level is, however, sequence convergence. Formally, sequence convergence is defined by independent changes leading to the same nucleotide or amino acid residue at the corresponding sequence positions in multiple lineages. Sequence convergence is often divided into parallel changes and convergent changes, depending on whether the ancestral states prior to the changes are the same or differ among the lineages (Zhang and Kumar, 1997). Hereinafter, we collectively refer to these two types as convergence unless otherwise mentioned. With the rapid accumulation of genome sequences from a variety of organisms, recent years have seen a surge in the report of sequence convergence, prompting a series of questions about the prevalence, adaptiveness, and phylogenetic impacts of sequence convergence. There have also been developments of methods to test whether sequence convergence is attributable to chance. We discuss these aspects of progress in this review.

2 Tests of adaptive sequence convergence in individual genes

Because phenotypic convergences are commonly viewed as strong indications of adaptive evolution, sequence convergences tend to be viewed similarly. However, because there are only four possible states at a nucleotide position and 20 possible states at an amino acid position, and because many of these states are not selectively allowed, the actual number of states permitted per nucleotide or amino acid position is quite small. This makes it possible for sequence convergence to occur simply by chance via neutral evolution instead of by a common selective force. Zhang and Kumar (1997) pioneered the modeling of chance sequence convergence. They proposed a test to examine whether the observed number of parallel or convergent amino acid substitutions is attributable to chance alone. Zou and Zhang (2015a) improved the test by considering different amino acid equilibrium frequencies at different sites, making the neutral model more realistic and the test more reliable. Below we briefly describe Zou and Zhang's test.

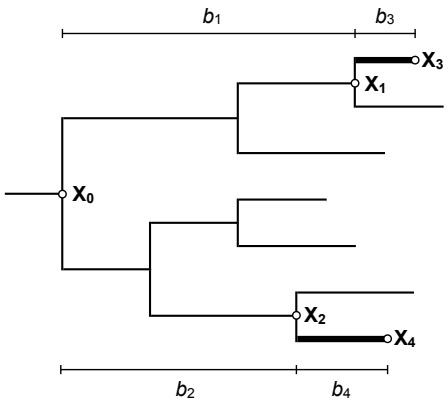


Figure 1 Counting the observed and expected numbers of events of sequence convergence between two branches on a tree. At a given position, the amino acids at nodes X_0 – X_4 are x_0 – x_4, respectively. Thick branches correspond to the converging lineages. The relevant branch lengths are indicated by the b values.

Let us take amino acid sequence evolution as an example and consider the possibility of sequence convergence between two focal branches (X_1 to X_3 and X_2 to X_4, respectively) of an arbitrary phylogeny shown in Figure 1. Let x_1, x_2, x_3, and x_4 be the amino acids at nodes X_1, X_2, X_3, and X_4, respectively. Convergent changes at a site can be defined by $x_1 \neq x_3$, $x_2 \neq x_4$, $x_1 \neq x_2$, and $x_3 = x_4$, whereas parallel changes can be defined by $x_1 \neq x_3$, $x_2 \neq x_4$, $x_1 = x_2$, and $x_3 = x_4$. Starting from an arbitrary root node X_0 with its state x_0, the probability of observing any conformation $X_1 = x_1, X_2 = x_2, X_3 = x_3, X_4 = x_4$ can be calculated based on a designated Markovian model of protein sequence evolution by $P(x_1, x_2, x_3, x_4 \mid x_0) = P(x_1 \mid x_0)P(x_3 \mid x_1)P(x_2 \mid x_0)P(x_4 \mid x_2)$. Based on the tradition of modelling sequence evolution as a continuous-time Markov process (see Chapter 1.1 [Pupko and Mayrose 2020]), we can calculate each conditional distribution by $P(Y \mid X = x) = I^{(x)}M^b$. Here $I^{(x)}$ is an indicator vector of size 20, with the element corresponding to amino acid x being 1 and all other elements being 0, M is the substitution matrix such as the empirical JTT matrix (Jones et al., 1992) or WAG matrix (Whelan and Goldman, 2001), and b is the branch length between node X and Y. Thus, $P(Y = y \mid X = x)$ is the probability of observing amino acid state y at node Y given state x at node X. The probabilities of parallel and convergent substitutions can thus be calculated by:

$$P_{\text{parallel}} = \sum_{x_1 \neq x_3, x_2 \neq x_4, x_1 = x_2, x_3 = x_4} P(x_1, x_2, x_3, x_4 \mid X_0 = x_0)P(X_0 = x_0)$$

$$P_{\text{convergent}} = \sum_{x_1 \neq x_3, x_2 \neq x_4, x_1 \neq x_2, x_3 = x_4} P(x_1, x_2, x_3, x_4 \mid X_0 = x_0)P(X_0 = x_0)$$

Here $P(X_0 = x_0)$ can be obtained from an indicator vector according to the inferred ancestral state (Zou and Zhang, 2015a). We can further compute $P_{\text{convergence}} = P_{\text{convergent}} + P_{\text{parallel}}$. For a pair of focal branches, such a probability of convergence can be calculated for each amino acid site. Because $P_{\text{convergence}}$ is small, the number of convergences at each site can be modeled by a Poisson distribution with $P_{\text{convergence}}$ as the expectation. So, the number of convergence events for an amino acid sequence is a new Poisson random variable with the expectation equal to the sum of $P_{\text{convergence}}$ across all sites. Given this expectation, one can compute the probability of occurrence of the observed number or more convergence events

from the upper tail probability of the corresponding Poisson distribution (Zou and Zhang, 2015a). A significant test result indicates that the observed convergence events are not fully attributable to chance, implying the involvement of a common selective force in the multiple lineages considered. Similarly, one can also separately test whether the observed parallel changes and convergent changes are more numerous than expected by chance, using P_{parallel} and $P_{\text{convergent}}$, respectively (Zou and Zhang, 2015a).

Applying the above statistical test of convergence or its original version (Zhang and Kumar, 1997) has revealed a number of cases of sequence convergence that are unattributable to chance alone. For example, the prestin protein is a member of the SLC26 anion-transport family providing the electromobility of outer hair cells thought to be responsible for cochlear amplification, an active process that confers sensitivity and frequency selectivity to the mammalian auditory system. Prestin showed significantly more parallel substitutions than expected in the origins of echolocating bats and toothed whales, two mammalian groups that independently evolved echolocation (Li et al., 2010; Liu et al., 2010). Given prestin function[1], this observation suggests prestin contribution, especially by the parallel substitutions, to the evolution of echolocation. Indeed, subsequent cell functional assays showed that the replacement of amino acid N with T at position 7, one of the parallel substitutions observed, converts the prestin of a non-echolocator to that of an echolocator in a biophysical property associated with mammalian high-frequency hearing (Liu et al., 2014). There were also reports of sequence convergence beyond what chance can explain between two lineages of echolocating bats that are thought to have independently evolved echolocation, although the functional role of the sequence convergence has yet to be demonstrated (Liu et al., 2011). More parallel amino acid substitutions than expected by chance were also observed in the independent evolution of a digestive ribonuclease in African and Asian leaf-eating monkeys, and *in vitro* assays showed that the parallel substitutions are responsible for the parallel improvements of the enzyme's activity in the two lineages (Zhang et al., 2002; Zhang, 2006).

Some studies experimentally demonstrated the functional role of sequence convergence without formally testing whether the convergent/parallel substitutions are attributable to chance. This was shown, for instance, for the role of sequence convergence in the elevated oxygen affinity of the hemoglobin in independently adapted hummingbirds in the Andes (Projecto-Garcia et al., 2013). In addition, there are many cases of sequence convergence in lineages that have experienced phenotypic convergence, but neither the causal relation between sequence convergence and phenotypic convergence nor the implausibility of sequence convergence by chance has been established (Christin et al., 2008; Jost et al., 2008; Shen et al., 2010; Feldman et al., 2012; Zhen et al., 2012; Ujvari et al., 2015).

Four criteria have been proposed to establish adaptive parallel/convergent evolution at the protein sequence level (Zhang, 2006). First, similar changes in protein function occur in independent evolutionary lineages. Second, parallel/convergent amino acid substitutions are observed in these proteins. Third, the parallel/convergent substitutions are not attributable to chance alone and therefore must have been driven by a common selective pressure. Fourth, the parallel/convergent substitutions are responsible for the parallel functional changes. These criteria are more stringent than simply showing a significant result from the statistical test mentioned (criterion 3), because one should know the protein functional consequence of the sequence convergence and the selective agent when claiming adaptation. Most claims of adaptive sequence convergence are based on criterion 2 only, some also satisfy criteria 1 and/or 3, but very few satisfy all four criteria.

[1] See Chapter 4.2 (Robinson-Rechavi 2020) for a review of how gene functions are defined.

To detect sequence convergence, one must first know the phylogenetic relationships of the sequences concerned. In theory, the tree used should be the gene tree instead of the species tree, but most people use the species tree instead, probably because the estimation of the gene tree is less reliable than that of species tree, which can be inferred using many genes together. When the species tree is used, some inferred sequence convergences may be false positives due to the discordance between the gene tree and species tree (Mendes et al., 2016). Thus, it is important to ensure that the underlying tree of the sequence evolution is correctly assumed in the study of sequence convergence.

3 Genomic patterns of sequence convergence: adaptive or neutral?

The abundance of genome sequence data now allows researchers to discover sequence convergence at the genomic scale. Without finding the specific reason of convergence at individual sites, one can ask whether the total number of convergence events observed in a genome exceeds the neutral expectation. Four approaches have been used to address this question. The first is the statistic $\Delta SSLS$ (difference in site-specific likelihood support), which is the difference in log likelihood value of a site under two alternative tree topologies. For example, Liu et al. (2011) evaluated each nucleotide site of the mitochondrial genome sequence alignment containing squamate reptile species for its likelihood support of a widely accepted nuclear tree topology versus a radically different mitochondrial topology splitting the supposedly monophyletic Iguania. It was found that most sites support the nuclear tree, while a small number of sites strongly favor the convergent mitochondrial topology. The log likelihood nature of $\Delta SSLS$ allows summation over sites to derive a gene-specific statistic (Parker et al., 2013). However, there is no explicit neutral expectation of $\Delta SSLS$ for a site or gene; consequently, the $\Delta SSLS$ distribution for different sites or genes is empirical. One can identify sites or genes in each tail of the $\Delta SSLS$ distribution, but cannot prove based on this information whether the convergence signals of these sites or genes are due to positive selection, because any distribution has tails (Zou and Zhang, 2015b). Therefore, while this method allows identifying sites/genes with the strongest convergence signals, it does not allow testing whether the observed convergence signals result from positive selection.

The second approach is to use observed rates of sequence divergence as a control when studying sequence convergence. A divergence event at a site between two branches refers to independent substitutions at the site in the two branches resulting in different nucleotide or amino acid states. The numbers of convergence (Cv) and divergence (Dv) events for a pair of branches are strongly correlated such that the $\frac{Cv}{Dv}$ ratio typically does not vary greatly among different pairs of branches (Castoe et al., 2009; Thomas and Hahn, 2015). A significantly higher $\frac{Cv}{Dv}$ ratio for a focal branch pair relative to other branch pairs would suggest a deviation in the focal branches. However, because the $\frac{Cv}{Dv}$ ratio under neutral evolution is unknown, one cannot prove that the deviation results from adaptive convergence in the focal branches. Furthermore, recent studies showed that, even under neutral evolution, the $\frac{Cv}{Dv}$ ratio decreases with the divergence between the branches concerned (Goldstein et al., 2015; Zou and Zhang, 2017), violating the assumption that the $\frac{Cv}{Dv}$ ratio is expected to be constant among all pairs of branches. However, one can classify both convergence and divergence events into two types: substitutions starting from the same state and those starting from different states. The two types of convergence events are precisely parallel and convergent substitutions, respectively. The convergent $\frac{Cv}{Dv}$ ratio and parallel $\frac{Cv}{Dv}$ ratio are each expected to be constant irrespective of the divergence between the branches (Zou and Zhang, 2017).

The third approach to testing adaptive convergence between a pair of branches is to use comparable control branch pairs (Projecto-Garcia et al., 2013; Foote et al., 2015; Zou and Zhang, 2015b; Natarajan et al., 2016; Xu et al., 2017). For instance, to test whether there is an excess in sequence convergence between echolocating bats and dolphins (focal branch pair), one could compare the focal branch pair with the control branch pair of echolocating bats and the cow, which represents a (non-echolocating) sister lineage of dolphins (Zou and Zhang, 2015b). Interestingly, in this case, the focal branch pair has fewer sequence convergences than the control branch pair (Zou and Zhang, 2015b). This comparison can be further controlled by the number of divergence events in the two branch pairs, accounting for potential differences in branch lengths. That is, one can construct a contingency table with Cv and Dv values of the two branch pairs, which can be statistically compared by a G-test. Notably, because sister taxa can have different branch lengths, the expected $\frac{Cv}{Dv}$ ratio is only equal when the convergent $\frac{Cv}{Dv}$ and divergent $\frac{Cv}{Dv}$ are separately considered, as mentioned above. Recently, Xu et al. (2017) applied a more stringent criterion in counting sequence convergence events in order to increase the probability of identifying adaptive sequence convergences. They compared three mangrove species with their respective non-mangrove sister species as well as a species that is the outgroup of all six species, and counted a convergence event at a site only when all mangrove species share the same amino acid state that differs from the amino acid state conserved among all four non-mangrove species. Their simulation showed that applying this criterion of convergence at conserved sites (CCS) substantially reduces chance convergence or convergence due to incorrect inference of ancestral states (Xu et al., 2017). Nevertheless, not all CCS events are necessarily adaptive, and Xu et al. (2017) selected candidate genes for adaptive convergence according to the number of CCS events per gene, based on an arbitrary cutoff. Thus, the CCS method provides candidates for adaptive convergence rather than proving adaptive convergence.

None of the above three approaches estimate the number of convergence events expected under neutral evolution. Consequently, they can compare the amount of sequence convergence among genes or among branch pairs, but cannot tell whether the amount of convergence observed exceeds the neutral expectation. The fourth and final approach differs from the above approaches in that it compares the observed amount of convergence with the neutral expectation. The neutral expectation is estimated by conducting computer simulations of sequence evolution or is probabilistically calculated. For instance, Rokas and Carroll (2008) used simulations to generate sequence alignments of the same size as the real data, using relatively simple models whose parameters are estimated from the actual data. They then regarded the number of convergence events observed from the simulated alignments as the neutral expectation. Zou and Zhang (2015a) directly calculated the expected number of convergence events between focal branch pairs as described in the previous section. Regardless of the method used in deriving the neutral expectation, the key is the substitution model and its parameters, because using different models or parameters results in drastically different neutral expectations (Zou and Zhang, 2015a). Rokas and Carroll (2008) reported that the number of convergence events observed at the genomic scale is much greater than the neutral expectation and suggested that this excess may be due to positive selection. However, in estimating the neutral expectation, they assumed equal amino acid compositions across sites, which is unrealistic and may lead to underestimation of the neutral expectation. In a subsequent study, Zou and Zhang (2015a) showed that the amount of sequence convergence observed at the genomic scale is compatible with neutral expectations derived under realistic substitution models. Below we summarize their analyses and results.

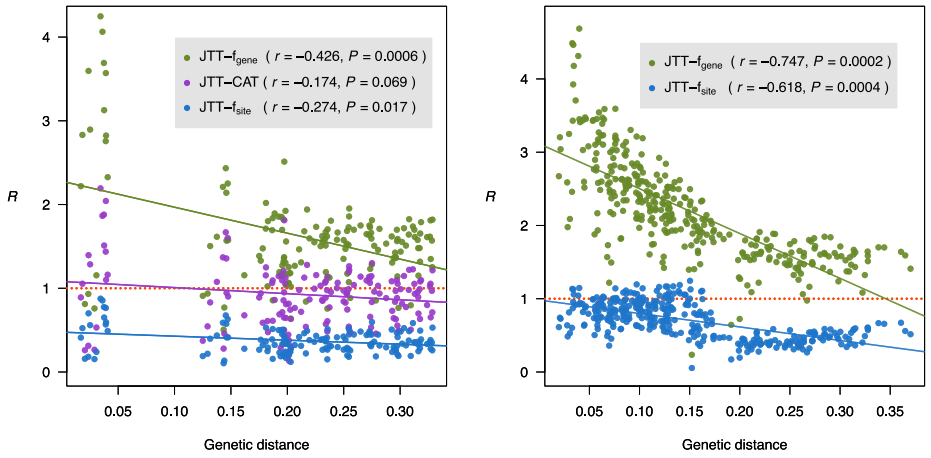
In 5,935 orthologous protein alignments of 12 *Drosophila* species, totaling 2,028,428

amino acid sites after the removal of gaps and ambiguous sites, 650 and 292 sites respectively experienced parallel and convergent substitutions in the two exterior branches leading to *D. yakuba* and *D. mojavensis*. Are these observed numbers of sites with parallel and convergent substitutions significantly greater than the corresponding neutral expectations? Zou and Zhang (2015a) examined three different neutral models. The first is the gene-specific JTT-f_{gene} model, which is based on the average substitution patterns of many proteins (Jones et al., 1992) with the equilibrium frequencies of the 20 amino acids in the model replaced with the observed amino acid frequencies of the protein concerned. The second neutral model considered is the site-specific JTT-f_{site} model, in which the equilibrium amino acid frequencies are replaced with the observed amino acid frequencies at the site concerned across all sequences in the alignment. One caveat in applying the JTT-f_{site} model is that, because the number of taxa used is smaller than 20 and because the total branch length of the *Drosophila* tree is also much smaller than 20, the observation of a limited number of different amino acids at a site may not mean that only those observed amino acids are acceptable but could be due to insufficient evolutionary time and taxon sampling for all acceptable amino acids to appear. Zou and Zhang (2015a) thus tried a third neutral model, JTT-CAT (Lartillot and Philippe, 2004) to estimate the expected numbers of convergent and parallel sites. Instead of having one set of equilibrium amino acid frequencies for all sites of a protein (JTT-f_{gene}) or one set per site (JTT-f_{site}), CAT uses a Bayesian mixture model for among-site heterogeneities in amino acid frequencies (see Chapter 1.4 [Lartillot 2020]). It estimates the total number of classes of sites and their respective amino acid frequencies, as well as the affiliation of each site to a given class. Due to the computational intensity of parameter estimation under JTT-CAT, Zou and Zhang (2015a) analyzed 1,081 relatively long proteins from the entire set of 5,935 proteins in an attempt to acquire the most information with the least amount of computer time.

The expected numbers of convergent and parallel sites, as well as the ratios (R) of the observed to expected numbers, are presented in Table 1 under each of the three neutral models. One can see that R varies from significantly above 1 to significantly below 1 among different neutral models (Table 1).

Type of sites	Number of sites examined	Observed number of sites	Expected number of sites		R	P-value
			Substitution model	Number of sites		
Convergent sites	2,028,428	292	JTT-f_{gene}	194.2	1.50	3.8E-11
	2,028,428	292	JTT-f_{site}	475.2	0.61	9.4E-20
	780,615	93	JTT-CAT	118.0	0.79	1.0E-3
Parallel sites	2,028,428	650	JTT-f_{gene}	388.6	1.67	3.2E-34
	2,028,428	650	JTT-f_{site}	2125.7	0.31	8.8E-309
	780,615	218	JTT-CAT	184.8	1.18	9.4E-3

Table 1 Observed numbers of sites experiencing convergent and parallel substitutions and the corresponding numbers expected under various neutral models of amino acid substitution. Reprinted with permission from Zou and Zhang (2015a). Results presented are for the two exterior branches leading to *D. yakuba* and *D. mojavensis*, respectively. R is defined as the ratio between the observed number and expected number. For the computation of the P-value, a statistical test is conducted under the assumption that the number of convergent (or parallel) sites follows a Poisson distribution with the mean equal to the expected number. When the observed number is smaller than the expected, the lower tail probability is given; when the observed number is larger than the expected, the upper tail probability is given.



(a) (b)

Figure 2 Ratios (R) of observed numbers of molecular convergences to the expected numbers in the protein evolution of *Drosophila* and mammals. (a) Scatter plot showing R against the genetic distance between the two branches concerned in the phylogeny of 12 *Drosophila* species. The R values under JTT-f_{gene} and JTT-f_{site} are based on all 5,935 proteins, whereas those under JTT-CAT are based on a subset of 1,081 proteins. (b) Scatter plot showing R against the genetic distance between the two branches considered for 2,759 proteins in the phylogeny of 17 mammals. In both panels, each dot represents one branch pair, and different colors show the results under different neutral models. Genetic distance is the number of amino acid substitutions per site between the two younger ends of the two branches considered. Solid lines show linear regressions. The r values are Pearson's correlation coefficients. P values are from Mantel tests. The horizontal red dotted line shows $R = 1$. The figure was redrawn using data from Zou and Zhang (2015a).

Thus, the answer to the question of whether there are more sequence convergences than the chance expectation depends on the neutral model assumed. Similar patterns were found when other branch pairs in the *Drosophila* tree were examined (Figure 2(a)). Zou and Zhang (2015a) further repeated this analysis in a set of 17 mammals. The data comprised 2,759 one-to-one orthologous proteins, with a total length of 1,079,696 amino acid sites. While the large data size prohibited them from using JTT-CAT, the analysis showed that R tends to exceed 1 under JTT-f_{gene} but becomes close to or even smaller than 1 under JTT-f_{site} (Figure 2(b)). Because models considering among-site heterogeneity in equilibrium amino acid frequencies almost always fit actual protein sequences better than comparable models assuming among-site homogeneity (Lartillot and Philippe, 2004, 2006), the findings from using the fourth approach suggest that the observed sequence convergence at the genomic scale is generally explainable by chance.

Figure 2 also shows an interesting pattern that R decreases with the genetic distance (number of amino acid substitutions per site) between the two younger ends of the two branches considered. A similar trend was independently reported for vertebrate mitochondrial proteins (Goldstein et al., 2015). Two explanations have been proposed. First, incomplete lineage sorting could make a gene tree different from the species tree, causing false inferences of convergences under the species tree. When the genetic distance between the two lineages considered increases, such false positive errors are expected to reduce, resulting in a negative correlation between R and the genetic distance (Mendes et al., 2016). Second, due to interactions among amino acid residues, the amino acids acceptable at a

site may change in evolution as a result of substitutions at other sites, such that an amino acid allowed at a site in one part of a tree becomes prohibited in another part of the tree, reducing the probability of sequence convergence with the genetic distance. Because the neutral models considered here do not include this factor, the neutral expectation of convergence is presumably overestimated, and R underestimated, when the genetic distance is large (Zou and Zhang, 2015a). While empirical evidence for the first reason exists, further analysis after excluding this factor still shows a negative correlation between R and the genetic distance, suggesting that the second reason may also exist in the data analyzed (Zou and Zhang, 2017). Importantly, evolutionary shifts in amino acid compositions at a site were observed when large alignments of hundreds to thousands of orthologous proteins were examined (Zou and Zhang, 2015a), and case studies showed that the same amino acid substitution can sometimes cause different or even opposite functional effects in homologous proteins (Zhang, 2003; Natarajan et al., 2016).

Note that, in all of the above analyses, amino acid substitutions are assumed to follow the JTT matrix with a set of equilibrium amino acid frequencies that could vary among sites or proteins. Recent studies found that the JTT substitution matrix does not apply universally and that different species show species-specific, genome-wide substitution patterns (Zou and Zhang, 2019). This means that amino acid substitution patterns are more diverse across species than generally thought. Consequently, one should be cautious in interpreting results from using the JTT matrix when studying sequence convergence.

4 Convergence as noise in phylogenetics

When discussing "analogical resemblances", Darwin pointed out that such resemblances "will not reveal—will rather tend to conceal their blood-relationship", so are "almost valueless to the systematist" (Darwin, 1859). Convergence is actually worse than being valueless, because it confuses phylogenetic inference and should be removed in phylogenetics if at all possible. Traditionally, phylogenetic trees of different organisms are inferred using morphological, physiological, or behavioral characters, collectively referred to as morphological characters hereinafter. The advent of molecular biology, especially the accumulation of sequenced genomes, supplied numerous molecular characters in the form of DNA and protein sequences, which are often considered more suitable than morphological characters for phylogenetic inference (Jousselin et al., 2003; Perelman et al., 2011; Wake et al., 2011; Legg et al., 2013; Springer et al., 2013; Jarvis et al., 2014). A major reason for this consideration concerns convergence. Compared with morphological characters, molecular characters are believed by many to be less susceptible to convergence (Givnish and Sytsma, 1997; Page and Holmes, 1998; Jousselin et al., 2003; Gaubert et al., 2005; Wiens et al., 2010; Wake et al., 2011; Davalos et al., 2012; Legg et al., 2013; Springer et al., 2013; Davalos et al., 2014). Nevertheless, this belief appears to have arisen in the early days of molecular systematics when morphological convergence had long been known while molecular convergence had not. As mentioned above, recent genetic and genomic studies revealed a large number of convergence events in protein sequence evolution. Zou and Zhang (2016) therefore compared the two character types, focusing on a large dataset containing both morphological and molecular characters that was previously used for jointly inferring the mammalian species tree. The data consist of 3,414 parsimony informative morphological characters and 5,722 parsimony informative amino acid sites for 46 extant and 40 fossil species (O'Leary et al., 2013). Below we summarize the analyses and findings from Zou and Zhang (2016).

Identifying character convergence requires the correct phylogeny, but because the mam-

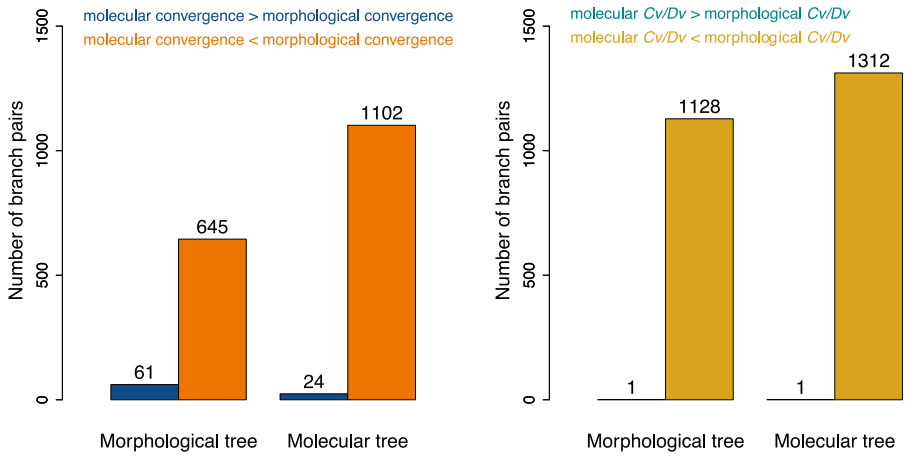
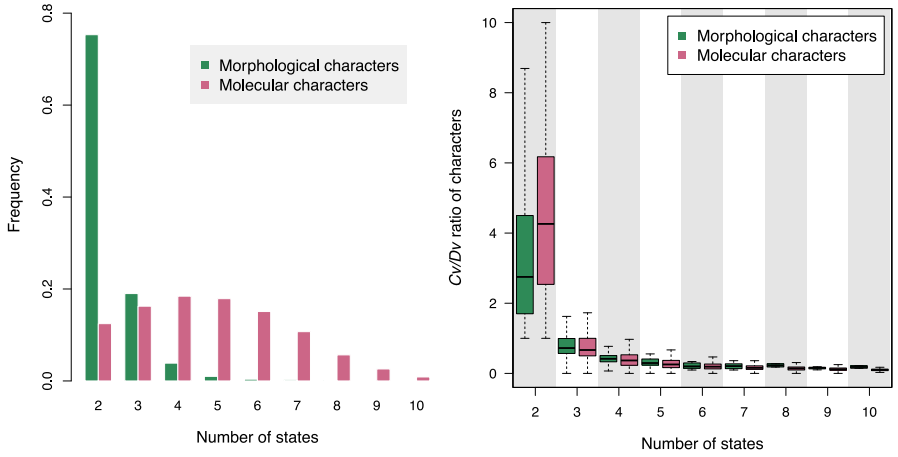


Figure 3 Comparison between morphological and molecular (sequence) convergences in mammalian evolution. (a) Comparison between the number of branch pairs for which the mean number of convergences per morphological character significantly exceeds that per molecular character (orange) and the number of branch pairs for which the number of convergences per molecular character significantly exceeds that per morphological character (blue) under each of two trees considered. (b) Comparison between the number of branch pairs for which the convergence/divergence ($\frac{Cv}{Dv}$) ratio is significantly greater for morphological characters than molecular characters (yellow) and the number of branch pairs for which $\frac{Cv}{Dv}$ is significantly lower for morphological characters than molecular characters (green) under each of two trees considered. In both panels, significance is defined by Q-value < 0.05. Number of branch pairs for a bar is indicated above the bar. The figure was redrawn using data from Zou and Zhang (2016).

malian tree is not completely resolved, Zou and Zhang (2016) considered three trees, respectively reconstructed using the morphological characters only, molecular characters only, and both types of characters in the data. Under each tree, they inferred the ancestral states at all interior nodes for each character by parsimony. For each pair of independent branches that can be investigated for convergence, they identified characters that showed convergence and compared the mean number of convergences per character between morphological and molecular characters. Among 3,396 investigated pairs of branches in the morphological tree, the number of branch pairs with a significantly higher number of convergences per morphological character than that per molecular charter substantially exceeds the number of branch pairs with a significantly lower number of convergences per morphological character than that per molecular character (Figure 3(a)). The mean number of convergence per morphological character is 1.7 times that per molecular character. When comparing the $\frac{Cv}{Dv}$ ratio introduced early, they also found morphological characters to exhibit overwhelmingly larger $\frac{Cv}{Dv}$, compared with molecular characters (Figure 3(b)). The mean $\frac{Cv}{Dv}$ ratio of morphological characters is 4.0 times that of molecular characters. When the above analyses were repeated under the molecular tree, even more convergences and higher $\frac{Cv}{Dv}$ ratios were found for morphological characters relative to those for molecular characters (Figure 3). Similar results were obtained under the total evidence tree.

Zou and Zhang (2016) noted that 75.2% of parsimony-informative morphological characters are binary in the data of O'Leary et al. (2013) (Figure 4(a)). Because binary characters can only have one kind of change given an ancestral state, it is obvious that they are susceptible to convergence once multiple changes occur. By contrast, only a small fraction (12.4%)



(a) (b)

Figure 4 Morphological characters tend to have fewer states than molecular characters. (a) Frequency distribution of the number of states per character. (b) $\frac{Cv}{Dv}$ ratio of a character decreases as the number of states increases. $\frac{Cv}{Dv}$ ratio of a character is the sum of convergences across all branch pairs divided by that of divergences. The top and bottom edges of a box represent the first and third quartiles of the distribution, respectively, while the thick line inside the box represents the median. The two whiskers show the maximum value not greater than the first quartile plus 1.5 times the box height and the minimum value not smaller than the third quartile minus 1.5 times the box height, respectively. $\frac{Cv}{Dv}$ ratios are calculated under the morphological tree. The same pattern is observed when $\frac{Cv}{Dv}$ ratios are calculated under the molecular tree. The figure was redrawn using data from Zou and Zhang (2016).

of molecular characters are binary (Figure 4(a)). The median number of states is five for molecular characters, significantly higher than that (two) for morphological characters ($P < 10^{-300}$). The probability of convergence relative to that of divergence for a character is expected to decrease with the number of states. Let the $\frac{Cv}{Dv}$ ratio of a character be the sum of Cv values across all branch pairs divided by the sum of Dv values across all branch pairs for the character. Indeed, the $\frac{Cv}{Dv}$ ratio decreases with the number of states for both types of characters (Figure 4(b)) and this trend remains after the control of evolutionary rate (represented by number of steps inferred on the tree). It was estimated that the $\frac{Cv}{Dv}$ ratio of an average morphological character is 0.89 times that of a molecular character with the same number of states. These results indicate that, compared with molecular characters, the higher convergence of morphological characters is caused by having fewer states rather than intrinsically higher susceptibilities to adaptive convergent evolution, because morphological characters are no more prone to convergence than molecular characters once the number of states is controlled for.

Because the vast majority of molecular convergences are explainable by chance (Foote et al., 2015; Thomas and Hahn, 2015; Zou and Zhang, 2015a,b), the fact that average morphological characters have even smaller $\frac{Cv}{Dv}$ ratios than those of molecular characters of the same numbers of states suggests that most morphological convergences observed in the data analyzed are probably also attributable to chance. If convergence is owing to chance rather than lineage-specific selection, it is possible to identify and remove convergence-prone characters using species with reliable phylogenetic relationships and then infer the tree for species of uncertain relationships using the remaining characters. This approach would be

especially beneficial to phylogenetic inference that includes morphological data because of the relatively frequent convergence in such data. Zou and Zhang proposed a method to identify convergence-susceptible (morphological or molecular) characters and demonstrated that removing such characters improves phylogenetic accuracy (Zou and Zhang, 2016). Interestingly, applying this method to O'Leary et al.'s data alters the phylogenetic relationships among echolocating bats (Zou and Zhang, 2016).

5 Conclusions

Sequence convergence in any given gene is generally rare. However, when the entire genome is analyzed, hundreds of sites may show convergence. But because some neutral models predict even more convergence events than what has been observed, the vast majority of convergences observed in genome-wide analysis are attributable to chance. Nevertheless, this conclusion about sequence convergence at the genomic scale does not exclude the possibility of some adaptive events of sequence convergence. In fact, adaptive sequence convergence has been clearly demonstrated by statistical and experimental tests in a few genes. Experience suggests that genome-wide identification of sequence convergence, coupled with considerations of gene functions and relevant phenotypic effects, can provide candidates for adaptive convergence that should be followed up with experimental validation.

Appropriately modelling sequence evolution in the absence of positive selection is critical for a proper detection of adaptive convergence. This is a major methodological issue in current, and presumably future, literature on the subject. The processes of incomplete lineage sorting (Mendes et al., 2016) and introgression (Witt and Huerta-Sanchez, 2019) complicate the identification of genuine convergence events between closely related species (Lee and Coop, 2019).

Apart from potential indications of adaptation, convergence is a major source of phylogenetic noise. Comparative analyses of a large dataset of morphological and molecular characters used by systematists for inferring the mammalian phylogeny showed that morphological characters experienced more convergent evolution than molecular characters. Hence, molecular trees are expected to be more reliable than morphological trees with comparable data sizes. Interestingly, however, the reason behind the higher convergence of morphological than molecular characters is not that morphological characters are intrinsically more prone to convergence as a result of frequent positive selection. Instead, at least for the O'Leary et al. (2013) data, the reason is that morphological characters used by systematists tend to have fewer states than molecular characters, and the propensity for convergence is not higher for morphological than molecular characters once the number of states is controlled for. It has been shown than convergence-prone characters can be identified and removed to improve the accuracy of phylogenetic inference. This practice would be especially important for phylogenetic analysis involving morphological characters due to their higher probability of convergence. While the rapid accumulation of genome sequences will eventually dwarf the morphological data of any extant species, morphological data will remain useful in phylogenetic analysis that needs to contain fossils (see Chapter 5.1 [Pett and Heath 2020]), whose value to understanding evolution is indispensable. In this sense, better modeling of morphological convergence and development of methods for detecting convergence-prone traits will potentially improve the accuracy of phylogenetic reconstruction.

References

Arendt, J. and Reznick, D. (2008). Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol Evol*, 23(1):26–32.

Blackledge, T. A. and Gillespie, R. G. (2004). Convergent evolution of behavior in an adaptive radiation of hawaiian web-building spiders. *Proc Natl Acad Sci U S A*, 101(46):16228–33.

Castoe, T. A., de Koning, A. P. J., Kim, H. M., Gu, W. J., Noonan, B. P., Naylor, G., Jiang, Z. J., Parkinson, C. L., and Pollock, D. D. (2009). Evidence for an ancient adaptive episode of convergent molecular evolution. *Proc Natl Acad Sci U S A*, 106(22):8986–91.

Chikina, M., Robinson, J. D., and Clark, N. L. (2016). Hundreds of genes experienced convergent shifts in selective pressure in marine mammals. *Mol Biol Evol*, 33(9):2182–92.

Christin, P. A., Salamin, N., Muasya, A. M., Roalson, E. H., Russier, F., and Besnard, G. (2008). Evolutionary switch and genetic convergence on rbcl following the evolution of c4 photosynthesis. *Mol Biol Evol*, 25(11):2361–8.

Darwin, C. (1859). *On the Origin of Species by Means of Natural Selection*. J. Murray, London,.

Davalos, L. M., Cirranello, A. L., Geisler, J. H., and Simmons, N. B. (2012). Understanding phylogenetic incongruence: lessons from phyllostomid bats. *Biol. Rev. Camb. Philos. Soc.*, 87(4):991–1024.

Davalos, L. M., Velazco, P. M., Warsi, O. M., Smits, P. D., and Simmons, N. B. (2014). Integrating incomplete fossils by isolating conflicting signal in saturated and non-independent morphological characters. *Syst. Biol.*, 63(4):582–600.

Doolittle, R. F. (1994). Convergent evolution: the need to be explicit. *Trends Biochem Sci*, 19(1):15–8.

Feldman, C. R., Brodie, E. D., Brodie, E. D., and Pfrender, M. E. (2012). Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc Natl Acad Sci U S A*, 109(12):4556–61.

Fong, S. S., Joyce, A. R., and Palsson, B. O. (2005). Parallel adaptive evolution cultures of escherichia coli lead to convergent growth phenotypes with different gene expression states. *Genome Res*, 15(10):1365–72.

Foote, A. D., Liu, Y., Thomas, G. W. C., Vinar, T., Alfoldi, J., Deng, J. X., Dugan, S., van Elk, C. E., Hunter, M. E., Joshi, V., Khan, Z., Kovar, C., Lee, S. L., Lindblad-Toh, K., Mancia, A., Nielsen, R., Qin, X., Qu, J. X., Raney, B. J., Vijay, N., Wolf, J. B. W., Hahn, M. W., Muzny, D. M., Worley, K. C., Gilbert, M. T. P., and Gibbs, R. A. (2015). Convergent evolution of the genomes of marine mammals. *Nat. Genet.*, 47(3):272–5.

Gaubert, P., Wozencraft, W. C., Cordeiro-Estrela, P., and Veron, G. (2005). Mosaics of convergences and noise in morphological phylogenies: what's in a viverrid-like carnivoran? *Syst. Biol.*, 54(6):865–94.

Gerrard, E., Mutt, E., Nagata, T., Koyanagi, M., Flock, T., Lesca, E., Schertler, G. F. X., Terakita, A., Deupi, X., and Lucas, R. J. (2018). Convergent evolution of tertiary structure in rhodopsin visual proteins from vertebrates and box jellyfish. *Proc Natl Acad Sci U S A*, 115(24):6201–6.

Givnish, T. J. and Sytsma, K. J. (1997). Consistency, characters, and the likelihood of correct phylogenetic inference. *Mol. Phylogenet. Evol.*, 7(3):320–30.

Goldstein, R. A., Pollard, S. T., Shah, S. D., and Pollock, D. D. (2015). Non-adaptive amino acid convergence rates decrease over time. *Mol Biol Evol*, 32(6):1373–81.

Grenier, J. L. and Greenberg, R. (2005). A biogeographic pattern in sparrow bill morphology: parallel adaptation to tidal marshes. *Evolution*, 59(7):1588–95.

Jarvis, E. D., Mirarab, S., [...], Gilbert, M. T. P., and Zhang, G. (2014). Whole-genome analyses resolve early branches in the tree of life of modern birds. *Science*, 346(6215):1320–31.

Jones, D. T., Taylor, W. R., and Thornton, J. M. (1992). The rapid generation of mutation data matrices from protein sequences. *Computer Applications in the Biosciences*, 8(3):275–82.

Jost, M. C., Hillis, D. M., Lu, Y., Kyle, J. W., Fozzard, H. A., and Zakon, H. H. (2008). Toxin-resistant sodium channels: parallel adaptive evolution across a complete gene family. *Mol Biol Evol*, 25(6):1016–24.

Jousselin, E., Rasplus, J. Y., and Kjellberg, F. (2003). Convergence and coevolution in a mutualism: evidence from a molecular phylogeny of ficus. *Evolution*, 57(6):1255–69.

Langerhans, R. B., Knouft, J. H., and Losos, J. B. (2006). Shared and unique features of diversification in greater antillean anolis ecomorphs. *Evolution*, 60(2):362–9.

Lartillot, N. (2020). The bayesian approach to molecular phylogeny. In Scornavacca, C., Delsuc, F., and Galtier, N., editors, *Phylogenetics in the Genomic Era*, chapter 1.4, pages 1.4:1–1.4:17. No commercial publisher | Authors open access book.

Lartillot, N. and Philippe, H. (2004). A bayesian mixture model for across-site heterogeneities in the amino-acid replacement process. *Mol Biol Evol*, 21(6):1095–109.

Lartillot, N. and Philippe, H. (2006). Computing bayes factors using thermodynamic integration. *Syst Biol*, 55(2):195–207.

Lee, K. M. and Coop, G. (2019). Population genomics perspectives on convergent adaptation. *Philos Trans R Soc Lond B Biol Sci*, 374(1777):20180236.

Legg, D. A., Sutton, M. D., and Edgecombe, G. D. (2013). Arthropod fossil data increase congruence of morphological and molecular phylogenies. *Nat. Commun.*, 4:2485.

Li, Y., Liu, Z., Shi, P., and Zhang, J. (2010). The hearing gene prestin unites echolocating bats and whales. *Curr Biol*, 20(2):R55–6.

Linnen, C. R., Poh, Y. P., Peterson, B. K., Barrett, R. D. H., Larson, J. G., Jensen, J. D., and Hoekstra, H. E. (2013). Adaptive evolution of multiple traits through multiple mutations at a single gene. *Science*, 339(6125):1312–6.

Liu, Y., Cotton, J. A., Shen, B., Han, X., Rossiter, S. J., and Zhang, S. (2010). Convergent sequence evolution between echolocating bats and dolphins. *Curr Biol*, 20(2):R53–4.

Liu, Z., Li, S., Wang, W., Xu, D., Murphy, R. W., and Shi, P. (2011). Parallel evolution of kcnq4 in echolocating bats. *PLoS One*, 6(10):e26618.

Liu, Z., Qi, F. Y., Zhou, X., Ren, H. Q., and Shi, P. (2014). Parallel sites implicate functional convergence of the hearing gene prestin among echolocating mammals. *Mol Biol Evol*, 31(9):2415–24.

Manceau, M., Domingues, V. S., Linnen, C. R., Rosenblum, E. B., and Hoekstra, H. E. (2010). Convergence in pigmentation at multiple levels: mutations, genes and function. *Philos Trans R Soc Lond B Biol Sci*, 365(1552):2439–50.

Maruyama, M. and Parker, J. (2017). Deep-time convergence in rove beetle symbionts of army ants. *Curr Biol*, 27(6):920–6.

Melville, J., Harmon, L. J., and Losos, J. B. (2006). Intercontinental community convergence of ecology and morphology in desert lizards. *Proc Biol Sci*, 273(1586):557–63.

Mendes, F. K., Hahn, Y., and Hahn, M. W. (2016). Gene tree discordance can generate patterns of diminishing convergence over time. *Mol Biol Evol*, 33(12):3299–307.

Moore, J. and Willmer, P. (1997). Convergent evolution in invertebrates. *Biol Rev*, 72(1):1–60.

Natarajan, C., Hoffmann, F. G., Weber, R. E., Fago, A., Witt, C. C., and Storz, J. F. (2016). Predictable convergence in hemoglobin function has unpredictable molecular underpinnings. *Science*, 354(6310):336–9.

Nevo, E. (1979). Adaptive convergence and divergence of subterranean mammals. *Annu. Rev. Ecol. Evol. Syst.*, 10:269–308.

O'Leary, M. A., Bloch, J. I., Flynn, J. J., Gaudin, T. J., Giallombardo, A., Giannini, N. P., Goldberg, S. L., Kraatz, B. P., Luo, Z. X., Meng, J., Ni, X. J., Novacek, M. J., Perini, F. A., Randall, Z. S., Rougier, G. W., Sargis, E. J., Silcox, M. T., Simmons, N. B., Spaulding, M., Velazco, P. M., Weksler, M., Wible, J. R., and Cirranello, A. L. (2013). The placental mammal ancestor and the post-k-pg radiation of placentals. *Science*, 339(6120):662–7.

Page, R. D. M. and Holmes, E. C. (1998). *Molecular evolution: a phylogenetic approach*. Blackwell Science.

Parker, J., Tsagkogeorga, G., Cotton, J. A., Liu, Y., Provero, P., Stupka, E., and Rossiter, S. J. (2013). Genome-wide signatures of convergent evolution in echolocating mammals. *Nature*, 502(7470):228–31.

Perelman, P., Johnson, W. E., Roos, C., Seuanez, H. N., Horvath, J. E., Moreira, M. A., Kessing, B., Pontius, J., Roelke, M., Rumpler, Y., Schneider, M. P., Silva, A., O'Brien, S. J., and Pecon-Slattery, J. (2011). A molecular phylogeny of living primates. *PLoS Genet.*, 7(3):e1001342.

Pett, W. and Heath, T. A. (2020). Inferring the timescale of phylogenetic trees from fossil data. In Scornavacca, C., Delsuc, F., and Galtier, N., editors, *Phylogenetics in the Genomic Era*, chapter 5.1, pages 5.1:1–5.1:18. No commercial publisher | Authors open access book.

Projecto-Garcia, J., Natarajan, C., Moriyama, H., Weber, R. E., Fago, A., Cheviron, Z. A., Dudley, R., McGuire, J. A., Witt, C. C., and Storz, J. F. (2013). Repeated elevational transitions in hemoglobin function during the evolution of andean hummingbirds. *Proc Natl Acad Sci U S A*, 110(51):20669–74.

Protas, M. E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W. R., Zon, L. I., Borowsky, R., and Tabin, C. J. (2006). Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat Genet*, 38(1):107–11.

Pupko, T. and Mayrose, I. (2020). A gentle introduction to probabilistic evolutionary models. In Scornavacca, C., Delsuc, F., and Galtier, N., editors, *Phylogenetics in the Genomic Era*, chapter 1.1, pages 1.1:1–1.1:21. No commercial publisher | Authors open access book.

Robinson-Rechavi, M. (2020). Molecular evolution and gene function. In Scornavacca, C., Delsuc, F., and Galtier, N., editors, *Phylogenetics in the Genomic Era*, chapter 4.2, pages 4.2:1–4.2:20. No commercial publisher | Authors open access book.

Rokas, A. and Carroll, S. B. (2008). Frequent and widespread parallel evolution of protein sequences. *Mol Biol Evol*, 25(9):1943–53.

Rosenblum, E. B., Rompler, H., Schoneberg, T., and Hoekstra, H. E. (2010). Molecular and functional basis of phenotypic convergence in white lizards at white sands. *Proc Natl Acad Sci U S A*, 107(5):2113–7.

Shen, Y. Y., Liu, J., Irwin, D. M., and Zhang, Y. P. (2010). Parallel and convergent evolution of the dim-light vision gene rh1 in bats (order: Chiroptera). *PLoS One*, 5(1):e8838.

Slot, J. C. and Rokas, A. (2010). Multiple gal pathway gene clusters evolved independently and by different mechanisms in fungi. *Proc Natl Acad Sci U S A*, 107(22):10136–41.

Springer, M. S., Meredith, R. W., Teeling, E. C., and Murphy, W. J. (2013). Technical comment on "the placental mammal ancestor and the post-k-pg radiation of placentals". *Science*, 341(6146):613.

Stewart, C. B., Schilling, J. W., and Wilson, A. C. (1987). Adaptive evolution in the stomach lysozymes of foregut fermenters. *Nature*, 330(6146):401–4.

Thomas, G. W. and Hahn, M. W. (2015). Determining the null model for detecting adaptive convergence from genomic data: A case study using echolocating mammals. *Mol Biol Evol*, 32(5):1232–6.

Ujvari, B., Casewell, N. R., Sunagar, K., Arbuckle, K., Wuster, W., Lo, N., O'Meally, D., Beckmann, C., King, G. F., Deplazes, E., and Madsen, T. (2015). Widespread convergence in toxin resistance by predictable molecular evolution. *Proc Natl Acad Sci U S A*, 112(38):11911–6.

Vitali, D. G., Kaser, S., Kolb, A., Dimmer, K. S., Schneider, A., and Rapaport, D. (2018). Independent evolution of functionally exchangeable mitochondrial outer membrane import complexes. *Elife*, 7:e34488.

Wake, D. B., Wake, M. H., and Specht, C. D. (2011). Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science*, 331(6020):1032–5.

Whelan, S. and Goldman, N. (2001). A general empirical model of protein evolution derived from multiple protein families using a maximum-likelihood approach. *Mol Biol Evol*, 18(5):691–9.

Wiens, J. J., Kuczynski, C. A., Townsend, T., Reeder, T. W., Mulcahy, D. G., and Sites, J. W., J. (2010). Combining phylogenomics and fossils in higher-level squamate reptile phylogeny: Molecular data change the placement of fossil taxa. *Syst. Biol.*, 59(6):674–88.

Witt, K. E. and Huerta-Sanchez, E. (2019). Convergent evolution in human and domesticate adaptation to high-altitude environments. *Philos Trans R Soc Lond B Biol Sci*, 374(1777):20180235.

Wittkopp, P. J., Williams, B. L., Selegue, J. E., and Carroll, S. B. (2003). Drosophila pigmentation evolution: Divergent genotypes underlying convergent phenotypes. *Proc Natl Acad Sci U S A*, 100(4):1808–13.

Xu, S. H., He, Z. W., Guo, Z. X., Zhang, Z., Wyckoff, G. J., Greenberg, A., Wu, C. I., and Shi, S. H. (2017). Genome-wide convergence during evolution of mangroves from woody plants. *Mol Biol Evol*, 34(4):1008–15.

Zhang, J. (2003). Parallel functional changes in the digestive rnases of ruminants and colobines by divergent amino acid substitutions. *Mol Biol Evol*, 20(8):1310–7.

Zhang, J. (2006). Parallel adaptive origins of digestive rnases in asian and african leaf monkeys. *Nat Genet*, 38(7):819–23.

Zhang, J. and Kumar, S. (1997). Detection of convergent and parallel evolution at the amino acid sequence level. *Mol Biol Evol*, 14(5):527–36.

Zhang, J., Zhang, Y. P., and Rosenberg, H. F. (2002). Adaptive evolution of a duplicated pancreatic ribonuclease gene in a leaf-eating monkey. *Nat Genet*, 30(4):411–5.

Zhen, Y., Aardema, M. L., Medina, E. M., Schumer, M., and Andolfatto, P. (2012). Parallel molecular evolution in an herbivore community. *Science*, 337(6102):1634–7.

Zhou, X. M., Seim, I., and Gladyshev, V. N. (2015). Convergent evolution of marine mammals is associated with distinct substitutions in common genes. *Sci Rep*, 5:16550.

Zou, Z. and Zhang, J. (2015a). Are convergent and parallel amino acid substitutions in protein evolution more prevalent than neutral expectations? *Mol Biol Evol*, 32(8):2085–96.

Zou, Z. and Zhang, J. (2015b). No genome-wide protein sequence convergence for echolocation. *Mol Biol Evol*, 32(5):1237–41.

Zou, Z. and Zhang, J. (2016). Morphological and molecular convergences in mammalian phylogenetics. *Nat Commun*, 7:12758.

Zou, Z. and Zhang, J. (2017). Gene tree discordance does not explain away the temporal decline of convergence in mammalian protein sequence evolution. *Mol Biol Evol*, 34(7):1682–8.

Zou, Z. and Zhang, J. (2019). Amino acid exchangeabilities vary across the tree of life. *Sci. Adv.*, 5(12):eaax3124.