

# Chapter 2

## Analysing Convergent Evolution: A Practical Guide to Methods

Kevin Arbuckle and Michael P. Speed

**Abstract** Convergent evolution, or the independent evolution of similar traits, has long been investigated and recognised as an important area of research for evolutionary biology. However, as with many areas of comparative biology, new phylogenetic methods that enhance our ability to study convergence have arisen with greater frequency in recent years. Consequently, we now have a wide range of tools at our disposal and a rapidly developing conceptual framework to guide us in our analyses. This chapter aims to provide a practical guide for those interested in convergent evolution that will enable new entrants to the field to quickly develop a well-rounded research agenda. Although some methods can be performed in other pieces of (stand-alone) software, this guide will focus on the R statistical environment.

### 2.1 Introduction

Convergent evolution is a common phenomenon across the diversity of living organisms. In essence, it refers to the independent evolution of some kind of similarity between two or more organisms, as opposed to any similarity which is a result of inheritance from a common ancestor. Convergent traits may be manifest across a number of levels of biology including both function and form (Losos 2011; Speed and Arbuckle 2016). Convergence can be seen for example in many forms of behaviour, morphology and physiology (McGhee 2011), and in the structure and action of molecules such as toxins or enzymes (Doolittle 1994). For instance, we could consider the phenomenon of mimicry, in which one organism (perhaps a harmless viceroy butterfly, *Limenitis archippus*) evolves to appear like a different

---

K. Arbuckle (✉) · M.P. Speed  
Department of Ecology Evolution and Behaviour Biosciences Building,  
University of Liverpool, Crown Street, Liverpool L69 7ZB, UK  
e-mail: k.arbuckle@liverpool.ac.uk

M.P. Speed  
e-mail: speedm@liverpool.ac.uk

organism (such as a poisonous monarch butterfly, *Danaus plexippus*) in order to deceive another (e.g. a predator) and gain some advantage as a result (Cott 1940; Ruxton et al. 2004). A contrasting example to highlight the wide reach of convergence is the evolution of myoglobin with similar oxygen binding properties in the muscles of several aquatic mammal lineages, which enables prolonged diving ability (Mirceta et al. 2013).

Convergence has long been considered as an important area of research in evolutionary biology. For example, in Chap. 6 of *The Origin of Species*, Darwin (1859) discussed several perceived difficulties with his theory of natural selection and invoked evolutionary convergence to counter some of these potential problems. More recently, convergence has been at the heart of modern comparative biology, albeit in an understated way. In particular, the power and applicability of phylogenetic comparative methods is typically limited by the number of independent replicates, which refers to the number of independent origins in these analyses. Hence, the rate of convergence of traits in a given case study can be a limiting constraint on the power of phylogenetic estimations (Maddison and FitzJohn 2015). Indeed, we often use comparative methods to investigate questions of adaptation or, more broadly, the relationship between different organismal traits across a phylogenetic tree, and our evidence for this is typically gained from several independent gains or losses of a set of traits. Stated in this way, we argue, it becomes clear why convergence should occupy a key position in the quantitative evaluation of macroevolutionary patterns.

As with many areas of phylogenetic comparative biology, methods to investigate convergent evolution have undergone a resurgence, with many new approaches having been devised in recent years. This has coincided with renewed interest in the ways in which we analyse convergence and how we think about it as a concept. While some excellent reviews already exist which focus on particular methods (e.g. Mahler and Ingram 2014), the aim of this chapter is to give a broad guideline of how a researcher might undergo a study of evolutionary convergence.

We will first discuss the types of questions we can ask and what we should set as our analytical goals. We will then give a brief survey of the software which is available to implement the methods discussed herein, followed by guidance on actually running the analyses themselves. For clarity and in the interest of making this chapter a practical guide rather than a review per se, we will not attempt to be entirely comprehensive in the methods we suggest, but instead aim to provide a selection that could be combined to generate a strong piece of research (and see review in Speed and Arbuckle 2016). Note that certain methods will be discussed in different sections, but with a focus on different aspects. In particular, methods that can be used both to detect and quantify convergence will be discussed in both sections covering each of those elements. We will end with a brief summary and some future directions for workers in convergent evolution and method developers.

## 2.2 General Aims in Studies of Convergent Evolution

There are clearly many different possible aims for particular studies of convergent evolution, depending on the system at hand and the interests of particular researchers. Nevertheless, there are some generalities that can be made and it is at this broad level that methodological recommendations can often be directed, which can then be applied to the specific details under investigation at any particular time. In particular, the goal of most studies can be loosely (but not mutually exclusively) divided into the identification or the quantification of evolutionary convergence.

The most common fundamental aim of studies of convergent evolution is to establish whether convergence is present in a particular trait in a particular group of species. For instance, Ujvari et al. (2015) demonstrated that resistance to cardiac glycoside toxins (and the molecular mechanism of this resistance) has independently evolved several times in animals as distantly related as mammals, reptiles, amphibians and insects. Similarly, Westneat et al. (2005) identified frequent convergence in the biomechanical function of jaws in wrasses (Labridae).

Evolutionary convergence can be quantified in two different ways: first by measuring its frequency and second by measuring its ‘strength’ (e.g. Stayton 2015; Speed and Arbuckle 2016). The aim of both approaches is to provide some way of measuring the influence of evolutionary convergence on the range of traits observed within specified groups, or within specified ‘niches’. Enumerating the cases of convergence of phenotypic categories for example allows us to determine the frequency of convergence (Stayton 2015) within a specified set of organisms. Measuring the strength of convergence, by contrast, requires us to ask ‘how similar are the trait(s) of organisms within a putative convergent subset?’ These two quantitative aspects should be measured separately and considered to be potentially independent. We could, for example, imagine a scenario in which convergence of classes of traits is frequent within a set of organisms, but weak in the sense that the convergent phenotypes are relatively diverse in quantitative terms. Conversely we can imagine rare, but strong convergence.

Examples of both kinds of quantification are becoming increasingly common in the literature. Mahler et al. (2013), for instance, provided several measures of the frequency of morphological convergence (in categorical traits) in Caribbean anoles to provide strong evidence for convergence among different radiations on different islands. More recently, Vidal-García and Keogh (2015) incorporated measures of the strength of convergent evolution into their analysis of (quantitative) morphological convergence in Australian frogs occupying different niches. Both studies benefit from a finer-grained understanding of the questions they were trying to address.

An advantage of the recent move towards quantifying convergence rather than simply documenting it is that it opens up new potential areas of research (Arbuckle et al. 2014; Speed and Arbuckle 2016). For instance, the ability to compare different cases of convergence allows us to explore differences in its frequency or strength when different types of traits or groups of organisms are considered. Asking, for

example, whether some traits are consistently more convergent than others, or whether a particular ecological niche favours stronger convergence than others will provide insights into the creation and limits of biodiversity. Such work will eventually permit a nuanced view of general patterns of convergence, and perhaps the process that generates these patterns, without being limited to a particular model system.

Quantification of convergence will also allow us to assess convergence within a particular system (or multiple systems) at different ‘levels of life’ (Losos 2011; Speed and Arbuckle 2016). By this, we mean levels such as function, form, development and genetics of a particular trait. This will perhaps provide information on precisely how convergence typically evolves, and therefore identify common constraints on the ability of organisms to evolve similarity over a particular period of time. Comparing the attributes of convergence across different levels of life will also point towards traits that are most remarkably convergent: those traits which have independently become similar at many levels (Arbuckle and Speed 2016).

## 2.3 Software Available for Analysing Convergent Evolution

There has been a recent surge of methodological development for the study of convergence, and a related increase in available software for such analyses. This section is not intended to give a comprehensive overview of these, but rather a selection with a focus on R packages. We do this because R has several advantages including (1) many new methods are implemented within it, often the only implementation is in an R package, (2) many biologists will already be familiar with it and (3) data handling and other analyses can be managed within the same environment as convergence methods, which reduces the need for data reformatting. See Table 2.1 for a summary of the software covered below and the following two sections for more details about particular methods and how they can be used.

Although many of the newer quantitative methods specifically designed for analyses of convergence are implemented in R, a simple form of detecting convergence is (of course) via ancestral state reconstruction (ASR). Because ASR is been a long-established technique in comparative biology for more reasons than looking for convergence, there are a number of good stand-alone programs to do this. For instance, BayesTraits (Pagel et al. 2004) reconstructs ancestral states by maximum likelihood (ML) or Bayesian Markov chain Monte Carlo (MCMC) methods. RASP (Yu et al. 2015) was primarily developed for reconstructing geographical ranges of ancestors, but can also be used for other traits and can consider polymorphic ancestral states. RASP uses an MCMC approach for ASR. Finally, Mesquite (Maddison and Maddison 2015) also has functions for reconstructing ancestral states by maximum parsimony (MP) or ML.

**Table 2.1** A selection of software that are useful for analyses of convergent evolution

Platform	Software	Type of measure	Method(s) implemented
Stand-alone	BayesTraits	Identification	ASR (ML and MCMC)
	Mesquite	Identification	ASR (MP and ML)
	RASP	Identification	ASR (MCMC)
R packages	ape	Identification	ASR (MP and ML)
	ouch	Identification	OU models on custom regimes
	phytools	Identification	ASR (ML, MCMC and stochastic character mapping)
	corHMM	Identification	ASR (ML; including hidden rates models)
	surface	Identification and quantification	OU models over the phylogeny with tests for convergence
	convevol	Identification and quantification	Phylomorphospace and distance-based measures ('C-metrics')
	windex	Quantification	Wheatsheaf index

Additional details of software and methods can be found in the main text. Abbreviations are as follows: *ASR* = Ancestral state reconstruction, *ML* = Maximum likelihood, *MCMC* = Markov chain Monte Carlo, *MP* = Maximum parsimony, *OU* = Ornstein–Uhlenbeck

Within R, several options are available for identifying cases of convergence. This is once again mostly using ASR methods. For instance, *ape* (Paradis et al. 2004) can reconstruct ancestral states using MP or ML, while *corHMM* (Beaulieu et al. 2013) uses ML to reconstruct categorical traits and can do so using either standard transition rate models or hidden rates models. The *phytools* package (Revell 2012) has numerous functions for estimating ancestral states using ML, MCMC and stochastic character mapping and has a particularly wide range of models from which to perform the ASR, including the incorporation of trends and threshold models. We note that most R functions for ASR, including those listed above, make use of a numerical optimisation algorithm rather than calculating an exact solution (for cases where this is possible). Nevertheless, the latter can be implemented using the ‘reconstruct’ function in *ape*, which calculates ancestral states using exact calculus. Finally, the package *surface* (Ingram and Mahler 2013) implements the method of the same name by acting as a wrapper for functions from *ouch* (Butler and King 2004).

The most simple way of quantifying convergent evolution in a trait is simply to identify it via a method such as ASR or SURFACE (note that the method is written in block capitals whereas the package is written in lower case) and then count the number of independent origins of the trait (e.g. Foote et al. 2015). This gives one measure of the frequency of convergence in the data set. However, a few software packages have been designed to provide additional options for measuring convergence. We have already mentioned the *surface* package in the context of its identification of convergent regimes, but it also provides additional information which can be used as a form of quantification. Specifically, in addition to the number of

convergent regimes found in the SURFACE analysis, other values such as the proportion of total regime shifts that are convergent and the reduction in the complexity of phenotypic evolution once convergence is accounted for can also be extracted (Ingram and Mahler 2013). Therefore, surface provides several measures of the frequency of convergence.

Several methods recently developed by Stayton (2015), herein referred to as ‘C-metrics’, are implemented in the *convevol* package (Stayton 2015) and provide measures of both frequency and strength of convergent evolution in a particular data set. This package also enables significance testing of whether the amount of convergence is ‘surprising’ (*sensu* Stayton 2008) using simulations under a Brownian motion model. Finally, another recently developed method, the Wheatsheaf index (Arbuckle et al. 2014), aims to quantify the relative strength of convergence in a subset of species (defined *a priori*) within a larger data set and is implemented in the *windex* package (Arbuckle and Minter 2015).

## 2.4 Detecting Convergent Evolution

The most basic and fundamental question we can ask about convergence is whether or not it occurs in a data set, or at least whether or not we have evidence for it in a data set. Not only is this an important question in itself, but it is of course vital to know prior to (or at least simultaneously with) quantification whether there is anything to quantify in the first place. It should also be noted that in the case of categorical traits, quantification is limited to frequency of convergence, and so its detection becomes proportionately more important in these cases.

*Ancestral state reconstruction*—The most basic way of testing for the presence of convergence (and again the only way for categorical traits) is to perform an ASR. As alluded to in the previous section, there are a multitude of methods of estimating ancestral states and even more options of software for their implementation (Table 2.1). Once you have completed your ASR there are many options for visualisation (e.g. Revell 2014) and the interpretation is very straightforward for categorical characters; the trait is convergent if it has arrived at the same state more than once across the phylogeny. For continuous traits, the interpretation can often be a little more subjective; how similar do such traits have to be before there is good evidence of convergence? In most cases, this will not be a problem (particularly where convergence is strong) as there will often be a clear shift in trait values that is shared between convergent lineages but is notably different from intermediate estimates. However, where convergence is not particularly strong, detection from ASR will be more difficult and inference will typically be informed by knowledge of what level of variation in the trait is biologically important in the system under study.

Where inference from ASR is not clear but higher certainty is wanted, then quantitative methods such as cluster analysis or dendrograms of trait values (also known as ‘phenograms’) may be useful. In these scenarios, evidence of

convergence would be gained from the clustering of phylogenetically disparate lineages occupying distinct clusters, or ‘clades’ in a phenogram (see examples in Speed and Arbuckle 2016). Although evolutionary stasis could produce similar patterns to convergence here, the distribution of other species within clusters may provide some idea of whether this is likely to be a good explanation or not.

An additional consideration when using ASR is which type of method to use. This can be considered first on two levels: maximum parsimony (MP) and model-based methods. Although contrasting views still exist in the literature, we recommend against using MP for this purpose. General criticisms of MP for comparative analyses apply here (e.g. Cunningham et al. 1998; Currie and Meade 2014). For instance, MP ignores information contained in branch lengths and therefore essentially assumes that all branch lengths are equal; an assumption that is unlikely to apply to most phylogenies. It also assumes that evolutionary rates have been slow, which may be the case but this can be explicitly estimated in model-based approaches and accounted for as part of model fitting. MP is also poor at reflecting the level of uncertainty in the reconstruction compared to model-based approaches which can provide a probability for each state at each node in the tree. Finally, specifically in the context of convergence, MP explicitly tries to minimise changes in the trait and therefore is likely to be biased against recovering convergent evolution since this necessitates additional changes.

Having decided to use a model-based method for ASR, the next choice is which model you want to fit. There are a range of options here depending firstly on whether the trait in question is categorical or continuous. Categorical trait models are described in terms of a transition rate matrix, and this could involve each rate parameter in the matrix being different or include various constraints such as certain rates forced to be equal or forced to be zero (such that a particular change can’t happen). There are also recent but increasingly used categorical trait models such as threshold or hidden rates models that might be appropriate for some systems. In terms of continuous trait models, there are again a range of options including Brownian motion, Ornstein–Uhlenbeck (OU), and trend models. Whatever range of models you feel is plausible, it is usually worth comparing the fit of each model in some way to give an empirical justification for a particular model choice (e.g. using AIC, likelihood ratio tests for nested models or Bayes factors).

Finally, there is a choice of how to fit the model. For example, you can fit many models in either a frequentist or a Bayesian framework, although the implementation of some models may only be currently available in one or the other. However, where there is a choice, this decision essentially comes down to informed personal preference since there are pros and cons to both approaches and strong and convincing advocates of each.

*Ornstein–Uhlenbeck (OU) models*—OU models essentially represent evolution of a continuous trait which is changing around an ‘optimum’ value (Felsenstein 1988; Hansen 1997). As the trait changes away from the optimum value, it experiences a ‘pull’ back to the optimum, the strength of which increases as the trait evolves further from the optimum. This is arguably a better representation of adaptive convergent evolution than other models because the optimum in the model

can be considered to represent an adaptive peak. Ingram and Mahler (2013) used OU modelling combined with stepwise model selection via AIC to test for convergence in one or (ideally) multiple continuous traits; a method they called SURFACE.

Briefly, SURFACE fits OU models over a phylogeny and uses AIC to identify where a change in the parameters of the model (or a ‘regime shift’) has occurred. This stage identifies where on the tree, for example, a trait has started to evolve around a new optimum/adaptive peak. In the next stage, SURFACE combines the different regimes estimated in the first stage and tests whether different combinations improve the model fit (again with AIC); in other words, it tests whether any of the regimes have evolved independently in different lineages. Interpretation of whether convergence has been detected is made simple by plotting the estimated regimes on the phylogeny in different colours, also implemented in the R package *surface* (Ingram and Mahler 2013), and looking to see whether the same regime appears multiple times.

It should be noted that a method very similar to SURFACE has recently been developed by Khabbazzian et al. (2016) and implemented in the R package *l1ou*. The latter method uses a faster algorithm, making it more suitable for very large phylogenies, and can use a newly devised and more conservative information criteria for model selection. For practical purposes, this method can be substituted for SURFACE in the discussions herein, but as it is only very recently been released it has not yet been used as often as SURFACE.

*Distance-contrast plots*—Muschick et al. (2012) further developed a method to identify convergence that was proposed by Winemiller (1991): distance-contrast plots. The method makes use of the idea that convergent evolution will lead to short phenotypic distances (more similar phenotypes) relative to the phylogenetic distance between a convergent pair of species. Therefore, if pairwise distances (scaled between 0 and 1) are plotted with phylogenetic distance on the x-axis and phenotypic distances on the y-axis, convergence is indicated by species pairs which fall below the line of unity (where  $x = y$ ) and especially species pairs which fall in the lower right-hand side. Note that the same pattern would also arise from evolutionary stasis, not only convergence, but this also applies to many methods and can be assessed using a method such as ASR to ensure independent origins.

The biggest problem with distance-contrast plots as a means of detecting convergence in a data set is which cut-off to use. In other words, when is a data point far enough below the line of unity to be considered evidence for convergence? Muschick et al. (2012) addressed this by conducting simulations of a trait under Brownian motion (to represent a ‘neutral’ trait). These simulated traits were used to generate a null expectation, and the distribution of pairwise distances from the putatively convergent trait is compared to this null expectation. Interpretation of convergence (or stasis) is therefore possible when many more species pairs fall into parts of the plot below the line of unity than can be explained by the neutral trait simulations. This method is not fully implemented within R as a separate function, but each part (deriving the required distances, simulations and plotting functions) is available either using standard R functions or those from *ape* (Paradis et al. 2004).



*Position and movement in phylomorphospace*—The use of the so-called phylomorphospace plots provides an intuitive way to identify convergent evolution, albeit one which is difficult to decide a threshold for considering a given pattern as evidence for convergence. Briefly, a phylomorphospace plot is a standard plot of at least two continuous traits measured in different species, but with a phylogenetic tree of the species superimposed over the plot and linking the data points. Furthermore, using an ASR method allows the plotting of estimated ancestral states on the same plot via the position of internal nodes of the plotted tree. This therefore allows tracing of the trait's evolution in phenotypic space by following the phylogeny. Phylomorphospace plots can be generated in *phytools* (Revell 2012), and an interpretation of convergent evolution is obtained from instances of multiple branches independently arriving in the same area of the plot. A specified area of the phylomorphospace plot can be highlighted using *convevol* (Stayton 2015) which helps to bring some level of objectivity to whether two species are similar enough to be considered convergent (although the initial specification of the area is still arbitrary).

## 2.5 Quantifying Convergent Evolution

Although the identification of convergence is a laudable and interesting goal in its own right, in most cases we can derive a better understanding of the system in question if we can quantify the convergence in some way (Arbuckle et al. 2014; Speed and Arbuckle 2016). As mentioned earlier, this quantification can relate to one of two aspects of convergence: its frequency and its strength. It is worth highlighting that for categorical traits, only the frequency can be measured and this can be done simply by counting the number of independent gains of such a trait from an ASR used to identify convergence. In fact, this also applies to all methods of identifying convergence above.

More broadly, Stayton (2008) highlighted that convergence can occur simply by chance when even a randomly evolving trait may become similar in multiple unrelated species. Depending on the research question at hand, we may be interested in convergence as a broad pattern (however it is generated) or we may be interested only in convergence that is more frequent (or stronger) than expected by chance. Of course, both of these measures could be interesting but the emphasis put on each one (total versus 'unexpected' convergence) will likely vary with the aim of the study.

For most methods of quantifying convergent evolution, either frequency or strength, simulations can be used to estimate how much convergence we would expect at random and use this for comparison with the observed amount of convergence in the data set to evaluate how much of the total convergence is 'unexpected'. Traits can be simulated in a wide range of R packages (e.g. *ape*) under many different models, but Brownian motion is commonly used as a model to simulate suitable 'neutral' traits for generating null expectations. In practice, many

(perhaps 1000 or more) traits could be simulated under your null model of choice, the method for measuring convergence applied to each one and the desired value (e.g. number of convergent events) extracted and used to plot a histogram. The actual values from the data can then be denoted on the histogram and used to judge how unexpected the amount of convergence is. Similarly, a P-value could be generated by taking the proportion of the simulated traits that showed the same amount or more convergence than observed in the data.

*SURFACE*—Although SURFACE is primarily a tool for detecting convergence (see previous section), the summary function for the method in surface allows more quantitative information to be gleaned on the frequency of convergence than simply counting the number of convergent regimes (Ingram and Mahler 2013). This was well illustrated in a study by the same authors that developed the method (and others) which investigated convergence in anole lizards (Mahler et al. 2013). The output from SURFACE can be summarised to generate additional metrics such as the number of shifts to convergent regimes (as above), but also the number of different convergent regimes, and the proportion of all regimes which are convergent. Therefore, SURFACE provides several measures of the frequency of convergence over and above simply counting the number of convergent events.

*C-metrics*—Stayton (2015) developed five measures of convergent evolution, termed  $C_1$ – $C_5$  and collectively referred to here as ‘C-metrics’. All five C-metrics are related to the phylomorphospace concept outlined above, which allows the tracing of changes in similarity over time and across the tree.  $C_5$  was considered separately from the other four metrics by Stayton (2015) as it deals with measuring the frequency of convergence rather than its strength. In essence,  $C_5$  counts the number of lineages that cross into a prespecified area of phylomorphospace (the area denotes a minimum level of similarity to be considered convergent). The area acts as a boundary, and since only crossing the boundary is considered in the  $C_5$  metric, it only counts independent origins of similarity. Furthermore, because estimates of ancestral states are considered in phylomorphospace, the value of  $C_5$  can potentially include instances of convergent ancestors that have since diverged again.

The remaining four C-metrics ( $C_1$ – $C_4$ ) are also related to the phylomorphospace concept and so also consider the change in similarity through time.  $C_1$  measures the change in similarity of two species as a proportion of the maximum distance the lineages have experienced.  $C_2$  is conceptually similar to  $C_1$  but is measured on an absolute scale in contrast to being relative to the maximum phenotypic distance.  $C_3$  and  $C_4$  are based on standardising  $C_2$  for the total amount of evolutionary change in the clade of interest (defined in different ways), which allows comparison between data sets. Consequently,  $C_1$ – $C_4$  are quantitative measures of the strength of convergent evolution rather than simply its frequency, and can therefore separate convergent changes of different magnitudes, even when the actual number of times convergence has arisen remains the same. The interpretation of all of the C-metrics is based on higher values indicating stronger convergence, or a greater frequency of convergence in the case of  $C_5$ .

*Wheatsheaf index*—The Wheatsheaf index was developed as a way to quantify convergent evolution in a particular subset of species within a clade (Arbuckle et al.

2014). This subset is termed the ‘focal group’ and may be one of two types. The focal group can be species occupying a particular ‘niche’ (used loosely), or exploiting the environment in similar ways, for which convergence is expected—for instance burrowing animals when considering body shape, which will give a measure of how strongly the convergence is in a particular trait (or several traits) for the ‘niche’ of interest. Alternatively, the focal group can be a set of species previously identified as convergent using one of the methods for detecting convergence covered above. In this case, the Wheatsheaf index quantifies how strong the convergence is within the known convergent species without reference to potential drivers of that convergence. In either case, the Wheatsheaf index is designed to be implemented after convergence has been identified in the data set—strictly a means of quantifying convergence rather than detecting it. The interpretation is simple in that stronger convergence will produce larger values of the index.

The Wheatsheaf index combines two aspects of trait evolution. It considers convergence to be stronger when species are more phenotypically similar to one another, but also as the focal group become more different from other species (as this implies a greater distance traversed across an adaptive landscape to achieve the similarity). The windex package (Arbuckle and Minter 2015) also provides a test of whether the Wheatsheaf index is higher (i.e. convergence is stronger) than expected by chance given the topology of the underlying phylogeny. Note that because it assumes convergence is present in the data set, it is not designed to detect convergence *per se*, but simply whether it is substantially stronger than expected.

As a final point related to the Wheatsheaf index, we can consider it as also providing a measure of divergent selection—in that very low values of the index would tend to indicate more diversification in the traits in the focal subgroup compared to the overall set of species in the tree in question. As with its use to study convergence, we can use the significance tests implemented in windex to determine whether the level of phenotypic disparity is greater than expected (using  $P > 0.95$  in that case). Hence, with this one measure, we can ask whether evolution has followed either a convergent or a divergent path.

## 2.6 Summary and Future Directions

In this chapter, we have attempted to provide an introduction to some of the practical aspects of studies of convergent evolution. Our hope is that readers with an interest in convergence can use this chapter as a starting point and as a set of guidelines to begin their research programmes. At the very least, we hope to have encouraged serious thought about how such a project might be structured and what types of methods are most desirable.

Although we have, to some degree, dealt with different facets of convergence separately here, a combined strategy is usually the best option. By this, we mean that unless a research question is deliberately narrow and specific, a program aimed at understanding convergent evolution in a particular system should incorporate

both the detection *and* quantification of convergence (Speed and Arbuckle 2016). Moreover, because different measures of convergence address different features of convergence, we recommend that a thorough investigation at least quantifies both the frequency and strength of convergence, and preferably uses multiple methods to do each of these.

In relation to using a multitude of methods, we wish to highlight a prime example of the type of approach we advocate. A recent paper by Friedman et al. (2016) investigated convergent evolution in surgeonfishes with respect to feeding niche. They employed a range of methods to detect convergence in morphology as well as to quantify both its frequency and strength, and the pluralistic approach taken provided them with a particularly rich source of information. This in turn is reflected in their interpretation of the results, which displays remarkable nuance and detail that would have been impossible to give with a more limited methodological approach.

Although we have come a long way in recent years in the development of a methodological toolbox for convergent evolution, we are likely to see additions continue to appear. This is partly due to an increasing emphasis on the conceptual framework within which we consider convergence (e.g. Arbuckle et al. 2014; Stayton 2015; Speed and Arbuckle 2016), as new ways of thinking about convergence are likely to generate new ways to study it. However, all methods are fraught with caveats and limitations, and attempts to address some of these will almost certainly result in the modification and extension of current methods as well as the design of entirely new methods. This has been a frequent pattern in the development of phylogenetic comparative methods more generally (e.g. Rabosky and Huang 2015), and there is no reason to think that the subset of those which deal specifically with convergent evolution will not show the same trend. Each method will have its own particular set of attributes which might be developed, but many are likely to benefit from modifications to allow fossil data to be incorporated where available as this typically improves inference from comparative methods (Slater et al. 2012). Nevertheless, we now work at a time when the potential for understanding fundamental evolutionary phenomena such as convergence is greater than ever before, and we just have to make good use of the tools available.

## References

- Arbuckle K, Minter A (2015) Windex: analyzing convergent evolution using the WheatSheaf index in R. *Evol Bioinform* 11:11–14
- Arbuckle K, Bennett CM, Speed MP (2014) A simple measure of the strength of convergent evolution. *Methods Ecol Evol* 5:685–693
- Beaulieu JM, O'Meara BC, Donoghue MJ (2013) Identifying hidden rate changes in the evolution of a binary morphological character: the evolution of plant habitat in campanulid angiosperms. *Syst Biol* 62:725–737
- Butler MA, King AA (2004) Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *American Naturalist* 164:683–695

- Cott HB (1940) Adaptive coloration in animals. Methuen and Company, London, UK
- Cunningham CW, Omland KE, Oakley TH (1998) Reconstructing ancestral character states: a critical reappraisal. *Trends Ecol Evol* 13:361–366
- Currie TE, Meade A (2014) Keeping yourself updated: bayesian approaches in phylogenetic comparative methods with a focus on Markov chain models of discrete character evolution. In Garamszegi LZ (ed) *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*, Springer-Verlag, Berlin, Heidelberg
- Darwin CR (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray Publishers, London, UK
- Doolittle RF (1994) Convergent evolution: the need to be explicit. *Trends Biochem Sci* 19:15–18
- Felsenstein J (1988) Phylogenies and quantitative characters. *Annu Rev Ecol Syst* 19:445–471
- Footo AD, Liu Y, Thomas GWC, Vinař T, Alföldi J, Deng J, Dugan S, van Elk CE, Hunter ME, Joshi V, Khan Z, Kovar C, Lee SL, Lindblad-Toh K, Mancina A et al (2015) Convergent evolution of the genomes of marine mammals. *Nat Genet* 47:272–275
- Friedman ST, Price SA, Hoey AS, Wainwright, PC (2016); Ecomorphological convergence in planktivorous surgeonfishes. *J Evol Biology* 29:965–978
- Hansen TF (1997) Stabilizing selection and the comparative analysis of adaptation. *Evolution* 51:1341–1351
- Ingram T, Mahler DL (2013) SURFACE: detecting convergent evolution from comparative data by fitting Ornstein-Uhlenbeck models with stepwise AIC. *Methods Ecol Evol* 4:416–425
- Khabbazian M, Kriebel R, Rohe K, Ané C (2016) Fast and accurate detection of evolutionary shifts in Ornstein-Uhlenbeck models. *Methods in Ecol Evolution* (Early View)
- Losos JB (2011) Convergence, adaptation, and constraint. *Evolution* 65:1827–1840
- Maddison WP, FitzJohn RG (2015) The unsolved challenge to phylogenetic correlation tests for categorical characters. *Syst Biol* 64:127–136
- Maddison WP, Maddison DR (2015) Mesquite: a modular system for evolutionary analysis. Version 3.04. <http://mesquiteproject.org>
- Mahler DL, Ingram T (2014) Phylogenetic comparative methods for studying clade-wide convergence. In Garamszegi LZ (ed.) *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*, Springer-Verlag, Berlin, Heidelberg
- Mahler DL, Ingram T, Revell LJ, Losos JB (2013) Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* 341:292–295
- McGhee G (2011) Convergent evolution: limited forms most beautiful. Massachusetts Institute of Technology Press, Cambridge, Massachusetts
- Mirceta S, Signore AV, Burns JM, Cossins AR, Campbell KL, Berenbrink M (2013) Evolution of mammalian diving capacity traced by myoglobin net surface charge. *Science* 340:1234192
- Muschick M, Indermaur A, Salzburger W (2012) Convergent evolution within an adaptive radiation of cichlid fishes. *Curr Biol* 22:2362–2368
- Pagel M, Meade A, Barker D (2004) Bayesian estimation of ancestral character states on phylogenies. *Syst Biol* 53:673–684
- Paradis E, Claude J, Strimmer K (2004) APE: analyses of phylogenetics and evolution in R language. *Bioinformatics* 20:289–290
- Rabosky DL, Huang H (2015) A robust semi-parametric test for detecting trait-dependent diversification. *Syst Biol* 2015:syv066
- Revell LJ (2012) phytools: an R package for phylogenetic comparative biology (and other things). *Methods Ecol Evol* 3:217–223
- Revell LJ (2014) Graphical methods for visualizing comparative data on phylogenies. In Garamszegi LZ (ed.) *Modern phylogenetic comparative methods and their application in evolutionary biology: concepts and practice*, Springer-Verlag, Berlin, Heidelberg
- Ruxton GD, Sherratt TN, Speed MP (2004) Avoiding attack: the evolutionary ecology of crypsis. *Warning Signals and Mimicry*. Oxford University Press, Oxford, UK
- Slater GJ, Harmon LJ, Alfaro ME (2012) Integrating fossils with molecular phylogenies improves inference of trait evolution. *Evolution* 66:3931–3944

- Speed MP Arbuckle K (2016) Quantification provides a conceptual basis for convergent evolution. *Biol Rev* (Early View)
- Stayton CT (2008) Is convergence surprising? An examination of the frequency of convergence in simulated datasets. *J Theor Biol* 252:1–14
- Stayton CT (2015) The definition, recognition, and interpretation of convergent evolution, and two new measures for quantifying and assessing the significance of convergence. *Evolution* 69:2140–2153
- Ujvari B, Casewell NR, Sunagar K, Arbuckle K, Wüster W, Lo N, O’Meally D, Beckmann C, King GF, Deplazes E, Madsen T (2015) Widespread convergence in toxin resistance by predictable molecular evolution. *Proc Natl Acad Sci USA* 112:11911–11916
- Vidal-García M, Keogh JS (2015) Convergent evolution across the Australian continent: ecotype diversification drives morphological convergence in two distantly related clades of Australian frogs. *J Evol Biol* 2015:2136–2151
- Westneat MW, Alfaro ME, Wainwright PC, Bellwood DR, Grubich JR, Fessler JL, Clements KD, Smith LL (2005) Local phylogenetic divergence and global evolutionary convergence of skull function in reef fishes of the family Labridae. *Proc R Soc B* 272:993–1000
- Winemiller KO (1991) Ecomorphological diversification in lowland freshwater fish assemblages from five biotic regions. *Ecol Monogr* 61:343–365
- Yu Y, Harris AJ, Blair C, He X (2015) RASP (Reconstruct Ancestral State in Phylogenies): a tool for historical biogeography. *Mol Phylogenet Evol* 87:46–49

Evolutionary Biology

Convergent Evolution, Evolution of Complex Traits,  
Concepts and Methods

Pontarotti, P. (Ed.)

2016, XI, 422 p. 50 illus., 36 illus. in color., Hardcover

ISBN: 978-3-319-41323-5