INTERFACE
FOCUS

rsfs.royalsocietypublishing.org

Review

Cite this article: Stayton CT. 2015 What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution. *Interface Focus* **5**: 20150039.
http://dx.doi.org/10.1098/rsfs.2015.0039

One contribution of 12 to a theme issue 'Are there limits to evolution?'

Subject Areas:
biomathematics

Keywords:
adaptation, constraint, null models, stochastic evolution

Author for correspondence:
C. Tristan Stayton
e-mail: tstayton@bucknell.edu

What does convergent evolution mean? The interpretation of convergence and its implications in the search for limits to evolution

C. Tristan Stayton

Department of Biology, Bucknell University, Lewisburg, PA 17837, USA

Convergent evolution is central to the study of life's evolutionary history. Researchers have documented the ubiquity of convergence and have used this ubiquity to make inferences about the nature of limits on evolution. However, these inferences are compromised by unrecognized inconsistencies in the definitions, measures, significance tests and inferred causes of convergent evolution. I review these inconsistencies and provide recommendations for standardizing studies of convergence. A fundamental dichotomy exists between definitions that describe convergence as a pattern and those that describe it as a pattern caused by a particular process. When this distinction is not acknowledged it becomes easy to assume that a pattern of convergence indicates that a particular process has been active, leading researchers away from alternative explanations. Convergence is not necessarily caused by limits to evolution, either adaptation or constraint; even substantial amounts of convergent evolution can be generated by a purely stochastic process. In the absence of null models, long lists of examples of convergent events do not necessarily indicate that convergence or any evolutionary process is ubiquitous throughout the history of life. Pattern-based definitions of convergence, coupled with quantitative measures and null models, must be applied before drawing inferences regarding large-scale limits to evolution.

1. Introduction

Convergent evolution is a central concept in evolutionary biology, and it is central in two different ways. First, convergent evolution is recognized as a ubiquitous and important feature in the history of life on the Earth. In a book that dealt deeply with convergent evolution, Simon Conway Morris wrote 'A textbook of evolution that fails to mention convergence would be guilty of serious dereliction' [1, p. 13], emphasizing both the omnipresence and the explanatory importance of this phenomenon. Moreover, convergent evolution (or at least homoplasy) is often seen as an inevitable outcome of evolution by natural selection: Colin Patterson suggested that one possible answer to his famous question 'Can you tell me anything you know about evolution, any one thing that is true?' could be 'I know that evolution generates [or explains] homoplasy' [2, pp. 188–189].

But convergence (synonymous with 'convergent evolution' throughout this paper) is also important in the study of evolution for a more utilitarian reason: convergent evolution provides comparative biologists with the replicated events necessary for powerful statistical tests. Although it is possible to provide evidence for adaptation by demonstrating a single functionally relevant 'fit' between organism and environment, the evidence is greatly strengthened if one can show that the fit has evolved independently multiple times [3]. Some researchers have even advised against doing comparative studies if replication is not available [4]. Researchers have taken advantage of such 'convergent adaptation' in many studies [5–11]; as just one example, conclusions regarding the adaptive nature of many traits associated with Caribbean *Anolis* lizard ecomorphs are greatly strengthened by the fact that all ecomorphs have evolved multiple times. Without

THE ROYAL SOCIETY
PUBLISHING

the independent evolution of these correlated suites of traits, the Caribbean *Anolis* system would be just one example among many lizard ecomorphological patterns, rather than the model system in adaptive radiation that it is [6,7,9,12]. Similar arguments apply to studies of phenomena besides adaptation, such as developmental constraint [13,14]. So even if there were no conceptual importance to convergent evolution, it would still ensure that studies can have $n > 1$.

Despite the importance of convergence, there is a surprising and largely unrecognized lack of agreement among researchers concerning its study. This conceptual uncertainty manifests in various ways: by the lack of consistency among definitions of convergence between and sometimes even within textbooks, monographs and papers; by the lack of a single standard set of methods used to quantify or assess the significance of convergence within a dataset; and by the lack of agreement concerning what implications convergence has for an understanding of the history of life. Some definitions frame ‘convergence’ as a simple evolutionary pattern [15–19], while others add adaptive or developmental factors [17,20–24]. Similarly, some studies quantify convergence using a geometric approach [25–27], while others quantify it relative to adaptive peaks [28]. Finally, some researchers see convergence as evidence for large-scale determinism in the universe [13,29–32], while others see it as being compatible with a great deal of stochasticity or contingency in evolution [32–36].

In this paper, I explore these different views of convergence and discuss their implications, with special attention given to implications for understanding limits on the evolutionary process. I do so through two different lenses, which are reflected in the title of this paper. First, I investigate what *the term* convergence means, exploring how a lack of consistency among definitions can lead researchers to make an unwarranted leap from the observation of a *pattern* to the inference of a *mechanism* or *process*. I also discuss how these definitions have informed the way that researchers recognize convergent evolution in a review of measures that have been proposed for quantifying convergence. Second, I discuss what *the phenomenon* of convergence means, or does not mean, for our understanding of evolution and its limits. I focus primarily on convergence in organismal phenotypes, though my arguments will also apply to any biological property for which a measure of difference can be derived, such as nucleotide sequences. Some authors have also discussed convergence at higher levels of biological organization, such as communities or ecosystems. Again, so long as the entities discussed are sufficiently independent and a measure of difference among entities is available, my arguments should apply here as well. I conclude by suggesting definitions, measures, and research approaches that might be most fruitful for investigating the relationships between convergence and evolutionary limits.

2. What does the term convergence mean?

Most students of biology could give a fairly precise definition of convergent evolution or convergence, along with multiple examples of this phenomenon. Conway Morris [1] was right about the importance of the term in biology textbooks—none of the books that I have surveyed, contemporary [15–17,19] or classic [37,38], were derelict regarding convergence. It is also, of course, alluded to in *The Origin of Species* [39]. Convergence is not an obscure concept in biology. However, there is also a pervasive lack of consistency in the definitions given for this phenomenon.

In general, two different types of definition of convergent evolution are used by biologists. Both usually require that two or more lineages evolve to be more similar to one another (though see [40] for an exception). A phrase specifying that the lineages are ‘evolutionarily independent’ is often included, either implicitly or explicitly [24,37,41]. However, the two types differ depending on whether a specific evolutionary process or mechanism is required for a pattern of increased similarity to be recognized as convergence. Some definitions do not require any particular mechanism or process; convergence can be defined, for example, as: ‘independent evolution of similar features from different ancestral traits’ [16, p. G7] or ‘the evolution of similar features in independent evolutionary lineages’ [15, p. G8]. These definitions may be called ‘pattern-based’ or ‘process-neutral’, in contrast with ‘process-based’ or ‘causally committed’ definitions (names suggested by Arbuckle *et al*. [25]). Process-based definitions do require a specific mechanism, such as differences in the developmental processes that produce the similar structures [14,18,24,42] or adaptation in response to the same selective pressures. An example of this type is: ‘the independent evolution of similar traits in distantly related organisms due to adaptation to similar environments or a similar way of life’ [17, p. G8] (see also [20–23]). Both types are common in the biological literature, whether in textbooks [15–17,19,21,23,24,42], biological dictionaries and encyclopedia [18,20,22], papers [14,26–28,40,43] or monographs [13,30,31].

Ideas of increased similarity also figure into a type of definition often used in discussions involving discrete character states [44]. Here, convergence refers to situations where two or more lineages with *different* initial character states transition to the same character state. In this context, convergence is distinguished from parallelism, in which two or more lineages with *the same* initial character states both transition to a different shared character state. These situations are usually considered as subcategories of ‘homoplasy’, with the emphasis on ways that taxa could share the same character state without having inherited that state from a common ancestor (thus ‘reversal’ is also often included in these discussions). However, there is a clear analogue with the preceding definitions: convergence occurs when similarity has increased in two lineages, with no necessary assumption of any particular process.

Somewhat confusingly, another definition of convergence also distinguishes it from parallelism, but based on a completely different set of criteria. In this framing, convergence produces similar features ‘by different developmental pathways’ [42, p. G2] or ‘without associated genetic convergence’ [18, p. 174] while ‘parallelism’ produces similar structures by similar developmental or genetic pathways. Clearly, the evolutionary implications of convergence versus parallelism according to such definitions are substantial, even given the difficulties inherent in differentiating between the two [14,45] (though see [36] for an objective distinction). This definition is clearly tied to a given process—increased similarity by itself is not sufficient for convergence. Generally, differences between convergence and parallelism, whatever the framing used, can also be reflected in pattern-based or process-based concepts.

As will be shown below, these differences are not merely superficial: differing definitions of convergent evolution suggest different ways to recognize and quantify convergence, test its significance and study its implications. These different measures, tests and interpretations in turn have profound implications for what the perceived ubiquity of convergence implies regarding limits to evolution.

It is thus very unfortunate that the differences among these types of definition are not often made explicit, and that the definitions themselves are quite often conflated. Authors only rarely present multiple possible definitions for convergent evolution before describing the operational definition used in their study (but see [27,43,46]); perhaps researchers tend to believe that 'their' definition of convergence is shared among all other researchers. Multiple different types of definition can even appear within a single work [16,42], without comment on these differences. This may occur because researchers may believe that convergence can *only* occur in the presence of some kind of limitation on evolution; if it is thought that convergence can only be caused by adaptation in response to a shared set of selective pressures, for example, then it becomes easy for researchers to assume that 'convergence is the independent acquisition of similar characteristics by lineages living in similar environments', and 'convergence is the independent acquisition of similar characteristics' mean essentially the same thing, with different levels of detail.

When pattern-based and process-based definitions become conflated in this way, the differences between definitional *criteria* for convergence and biological *implications* of convergence can become blurred. Convergence documented using a pattern-based definition can serve as evidence for adaptation in response to a shared set of selective pressures, or to the possession of shared developmental processes or to other shared factors—a use to which many biologists would like to put convergent patterns [3,30,31,43,47]. With process-based definitions, by contrast, shared selective pressures, developmental features or other factors must be demonstrated *before* a pattern can be described as convergence.

This illustrates the problems inherent in building 'ultimate'-level explanations into these definitions. With process-based definitions, convergence is no longer a powerful source of evidence that can be used along with other data to document the existence of shared 'limits' among taxa. Instead, it is simply a label used when such limits have already been documented; the important and interesting evidence for the limits is deemphasized. Lack of clarity regarding definitions can produce additional difficulties. Process-based definitions define convergence as a pattern produced by a given shared process, but the complementary phenomenon—the same pattern, not produced by any shared process—is left unnamed. Thus, such definitions do not provide any conceptual space for a potentially important and informative biological phenomenon: a pattern of increased similarity that is *not* produced by any shared process or by any directional process at all. This may in turn bias researchers away from seeking alternative explanations for patterns of increased similarity beyond the preferred or assumed process [40,43], such as evolution in response to *different* selective pressures [43], *different* developmental constraints [14] or stochastic evolution [26,48]. Pattern-based definitions leave room for all such explanations, and can accommodate additional modifiers ('adaptive convergence', for example) or subsets ('developmental parallelism' as a subset of convergence) if desired. For these reasons, pattern-based definitions will be preferable in convergence studies.

3. How is the phenomenon of convergence recognized?

Historically, convergent evolution was probably most commonly recognized intuitively, as a self-evident phenomenon in need of no particularly precise definition. Often convergence has been recognized without any strong phylogenetic hypotheses or quantification of similarity (see [13,31] for collections of hundreds if not thousands of examples). This method can clearly lead to false conclusions, but many classic examples of convergence recognized in this manner still stand: it is difficult to see how the forelimbs of moles and mole crickets could not be more similar to one another now than were the forelimbs of the ancestors of these two groups. Lack of quantification by itself does not invalidate a putative example of convergence, though as will be seen below it may severely limit its utility.

It should also be noted that more definitive evidence for convergence can be derived from very rudimentary phylogenetic and phenotypic data, along with logical principles, if multiple convergent events have occurred between clades. Take for example the classic instances of convergence between marsupials and placentals. Multiple ecological analogues have been identified between the two groups: marsupial and placental 'moles', 'cats', 'wolves', 'anteaters' and 'mice', for example. And it is known that placentals and marsupials constitute mutually exclusive monophyletic groups. Thus, while it is possible that one of the ecological analogue pairs (the 'mice', perhaps) owe their similarity to having retained characteristics from the ancestral therian, it is not possible that all of them should have done so. If members of each pair of analogues are indeed more similar to each other (for certain traits) than to closer relatives, then convergence must have occurred in at least all but one of the pairs.

Multiple avenues are available for more rigorous and quantitative assessments of convergence. In some fortunate cases where direct evidence is available of the phenotypes of individuals in multiple evolving lineages, convergence can be observed more or less directly. Whether in microbial lineages evolving in a laboratory [49–51] or well-preserved fossil sequences [52], it can occasionally be possible to trace the true history of phenotypic similarity between lineages. All that is required in such cases is trait data of individuals and some measure of similarity—if the similarity between the lineages increases over time, then convergence has happened. It should be noted that these are the only situations in which convergence among sister taxa can be *documented*—other sources of data will always mask patterns of increasing similarity among sister lineages, even if such patterns were present.

Such well-documented patterns of change within lineages are rare, however. It is much more common for convergence to be documented with comparative datasets. Researchers have developed many measures and indices for quantifying convergence in comparative data, often on an ad hoc basis for particular studies [53]. Often these methods are not even named, nor propagated beyond their original use. As a side note, this has led to the 'redevelopment' of convergence measures by multiple authors [27,53], a conceptually harmonic but indisputably inefficient phenomenon in convergence studies. At times, separate methods have been developed for discrete [44] and continuous data, but in many cases the same methods can be applied to either type. For example, measures of homoplasy [54] can be modified to quantify the amount of convergence in discrete or continuous data. However, these measures will also track other potentially distinctive phenomena, such as parallelism or reversal.

Most recent methods have focused on measuring convergent evolution only, to the exclusion of other kinds of homoplasy. Some researchers emphasize the increase in similarity that occurs between converging lineages [55,56]. Stayton [26] computed his MCI by comparing the variation in a set of putatively convergent taxa to the variation present in the larger clades to which those taxa belonged. The more similar the putatively convergent taxa, and the more different the close relatives of those taxa, the higher the index. Arbuckle *et al.* [25] further elaborated on this concept in developing the Wheatsheaf index, used to measure the strength of a convergent pattern after convergence itself had been established through other means. The index is calculated as the ratio of the average pairwise distance between all species in a dataset, to the average pairwise distance among all species that are convergent with one another. Larger values of this index indicate that the convergent taxa are much more tightly clustered with each other than are taxa in the dataset in general. This index also increases as the putatively convergent taxa become more distinctive—that is, less similar to all other taxa—which may or may not be desired in a given convergence study.

Stayton [57] provides a more direct measure of increased similarity. For his C_1 measure, ancestral states are reconstructed for two or more putatively convergent lineages, back to their most recent common ancestor. The maximum phenotypic distance between any pair of ancestors (D_{max}) is calculated, and compared with the phenotypic distance between the current putatively convergent taxa (D_{tip}). The greater the difference between D_{max} and D_{tip}, the higher the index. All of these methods are evolutionarily grounded and do not conflate convergence with other concepts; however, they can require inferences regarding ancestral character states, which can be problematic to obtain [58], and they are not appropriate for studies that use a definition of convergence that does *not* require an evolutionary increase in similarity among convergent lineages [40].

Other measures quantify the discrepancy between phylogenetic and phenotypic distance that accompanies convergent evolution. Stayton [27], expanding an idea from [59], suggested that the ratio of patristic distance to phenotypic distance between taxa could be used as a measure of convergence, perhaps with both distances appropriately scaled so that comparisons would be possible between datasets. Muschick *et al.* [10] similarly used a ratio of phylogenetic distance to morphological distance to define different modes of evolution, including a type called ‘convergence/stasis’. These methods are intuitive and conceptually simple, but they do not differentiate between convergent evolution and long-term stasis among the putatively convergent groups (and neither does the Wheatsheaf index, which is why convergence must be established before using this metric)—in both cases, low phenotypic dissimilarity accompanies large phylogenetic distances.

The SURFACE method of Ingram & Mahler [28] is unique among methods of measuring convergence, as it used a model-fitting approach to quantify the similarity in evolutionary patterns among lineages. The method fits a series of increasingly complex Ornstein–Uhlenbeck (OU) models to a set of data, eventually selecting the best one using Akaike information criteria. OU models include both a stochastic component and evolutionary ‘attraction’ of certain lineages towards specific points in phenotypic space [60,61]. The SURFACE method will identify the attractive ‘peaks’ associated with certain clades (that is, it identifies clades for which there is evidence that their members are evolving towards a certain point in phenotypic space), but more crucially, it identifies peaks that are shared among independent lineages. Two or more independent lineages evolving in response to the same peak is considered an instance of convergence, and the total number of shared peaks quantifies the prevalence of convergence in the dataset.

The closest analogues to SURFACE are a series of methods that count the number of convergence events in a dataset, whatever the similarity or phylogenetic distance that accompanies them. These include an unnamed measure developed by Winemiller [53], the ‘Count’ measure of Stayton [27] and C_5 of Stayton [57]. Along with SURFACE, these methods provide an integer value as a measure of convergence. Winemiller’s measure counts ‘the number of closely related species that [are] actually less similar morphologically to the target species that its morphological nearest neighbor’, scaled by the number of possible comparisons; Stayton’s ‘Count’ sums the number of taxa that are phenotypically more similar to a taxon than to their closest relatives (again, potentially scaled by the number of possible comparisons); and C_5 sums the number of times through the evolution of a clade that lineages evolve into a given region of phenotypic space.

All of the methods described above lend themselves well to null models and hypothesis testing. It is simple to simulate data multiple times according to some null model—Brownian motion, perhaps, or stochastic evolution constrained by certain developmental parameters but otherwise undirected—calculate the measure of convergence for each iteration, and then use the distribution of measures for hypothesis testing (this is even recommended with SURFACE).

Such statistical tests are underlain by an important principle that is often overlooked: some amount of convergence is possible even in datasets where there are no shared selective regimes or shared developmental constraints among lineages, or even no directional influences at all on the evolutionary process. In fact, it has been shown that under some conditions, the frequency and magnitude of convergence can be quite high in evolutionary data generated by a purely stochastic process unconstrained by developmental factors or limits on phenotypes [27]. Far from being evidence against undirected evolution, convergence can be a product of undirected evolution [33]. It may indeed be the case that most instances of convergence in nature are extremely unlikely to have occurred without any non-random influences on the evolutionary process, but this can only be established with null models and empirical data.

All of the methods described here implicitly assume a pattern-based definition. Even SURFACE, which uses an evolutionary model that is often associated with adaptation, leaves open the possibility that a process besides shared selective regimes, even chance, has resulted in lineages showing a substantial tendency to evolve towards the shame regions of phenotypic space [28]. Thus, if a researcher were to prefer a process-based definition of convergence, then presumably a significant value of any of these metrics would be necessary but not sufficient for an assessment of convergence.

Unfortunately, researchers who develop these methods (the present author included) infrequently note that this is the case. As with the discussion of definitions, this makes it easy for researchers to *quantify and test for* convergence using pattern-based concepts, and then *interpret* those patterns using a process-based concept. Rather than being a pattern that may represent convergence if additional facts are demonstrated, or a pattern that can be used as one among many pieces of evidence for shared adaptive regimes [3,9,47,62,63], a pattern of increased similarity becomes *sufficient* evidence for adaptation in response to shared selective pressures even though such patterns can demonstrably occur in the absence of such shared pressures, or even any selective pressures at all [27,48]. While both pattern-based and process-based definitions can be legitimately applied to biological phenomena, it is important to recognize that most convergence measures quantify patterns and do not demonstrate the action of any given process.

4. What are the implications of convergent evolution?

Superficially, it might seem that some of the definitions described above, or methods predicated on those definitions, would also imply certain evolutionary interpretations of convergence. A process-based definition such as 'The independent evolution of similar traits in distantly related organisms due to adaptation to similar environments or a similar way of life' [17], for example, indicates that convergence must always be due to adaptation. However, as the preceding discussions have demonstrated, this implication simply follows from the definition and does not actually say anything about the biological world. Using a definition of convergence that requires adaptation to shared selective regimes does not imply that such adaptation is responsible for all instances where lineages have evolved to be more similar to one another. It simply means that two sources of information are required for an assessment of convergence: first, a significant pattern of increased similarity among lineages, and second, a demonstration of similar selective factors in those lineages. Process-based definitions implicitly leave open the possibility that patterns of increased similarity are *not* due to shared selective regimes or shared developmental constraints. But this possibility is not often mentioned, potentially biasing researchers away from performing significance tests or seeking alternative explanations for the patterns they observe.

These potential biases become particularly acute when it comes to relating convergent evolution to large-scaled conclusions about the history of life on the Earth (or beyond); for example, when attempting to use studies of convergence to draw conclusions regarding the existence and extent of limits to evolution. Clearly, isolated instances of convergence by themselves do not indicate very much about grand patterns of evolution. Thus, multiple authors have attempted to draw larger conclusions by compiling long, impressive lists of examples of convergence [13,30]. The general argument is that the long lists indicate that convergence is surprisingly, unexpectedly ubiquitous, and that this ubiquity indicates the surprising prevalence of some factor—some combination of shared adaptive regimes or shared developmental constraints, perhaps, or an unexpected limit to the number of possible adaptive regimes or developmental programmes—that limits evolution to a greater degree than is currently acknowledged by evolutionary researchers.

The discussion of convergence in this paper suggests three potential problems with these interpretations, all of which can be addressed by greater precision in definitions of convergence and by the greater use of significance tests and null models. First, of course, is the fact that convergence can occur in the absence of shared adaptive regimes, shared developmental constraints, or any other limits or directional influences on evolution [14,27,43,46,64,65]. Thus, long lists of examples of convergent evolution do not by themselves establish the ubiquity of any particular process. However, many of the examples cited in those large studies are accompanied by environmental information (broadly construed), which provides additional evidence for adaptation, and even though repeated correlations between phenotype and environment can be explained by factors other than adaptation [43,48], such cases may be rare, and the data to test for these additional factors can often be readily obtained. Overall, it would be irrational to think that the vast majority of these examples would not stand up to quantitative scrutiny, but certainly there are some that would not, and neither shared adaptation nor constraint should be assumed whenever a convergent pattern is observed. Instead, additional tests for these phenomena need to be conducted. It bears repeating that even without any limits on evolution, any selection at all, or any developmental constraints, fortuitous convergence could still be ubiquitous throughout the tree of life [27].

This fact raises a more salient second issue with the use of long lists of examples to infer the presence of large-scale limits on evolution: it has not yet been established for most of these examples that the patterns of increased similarity are strong enough to require explanation from any particular process. As just one example: it has been established that a number of morphotypes have evolved multiple times in fishes. The narrow, elongate 'eel' morphotype is among the most familiar, but there are others; some, such as the 'pike', have even been quantitatively described [53]. However, it has not yet been established that these morphs are unlikely to have evolved multiple times in fishes in the absence of any limits to evolution. Quantitative methods employing null models are required to test the hypothesis that: 'An elongate body form is unlikely to have evolved an observed number of times, in the absence of any directional influences on evolution, given an ancestral teleost body form and known rates of morphological evolution in fishes'. The null models could take many forms, from purely stochastic models that only use one rate of evolution for all measured variables, to those which allow multiple rates of change in different body regions or which incorporate developmental information to reflect correlated evolution among parts; authors may also wish to investigate the influence of different ancestral 'starting points' for simulations. But until such tests are performed, it is premature to claim that the observed phenomena require any explanation at all, much less one that points towards a deep role for any kind of limit on evolution. This issue is more acute for some examples than others: all 'pike' feed in more or less the same way, making chance a less likely explanation for their convergence, but 'eels' come from a variety of clades, occupy a variety of different environments and use a variety of feeding strategies [40], raising the possibility that their convergence may be due at least partly to chance.

Finally, it is not yet clear that very long lists of examples of convergent evolution (again, encompassing hundreds if not thousands of taxa) actually demonstrate that convergence is ubiquitous throughout the history of life. This is true even if every one of the listed examples has been subjected to rigorous statistical testing and demonstrated to have occurred due to some particular evolutionary process. Convergence is widespread, certainly, but some kind of null model is still required whether it is more common than would be expected by chance. This is not just a technical quibble—it has been shown that under certain conditions, the *majority* of evolving taxa end up more similar to some distant relative to their own sister taxa for at least some traits [27]. Because the number of possible comparisons among taxa is astronomical and inevitably much larger than the number of known instances of convergent evolution, and equivalent lists of non-convergent taxa are unlikely to become available (due to lack of attention or publication bias, though see [66,67] for excellent examples of such studies), evolutionary null models for large-scale questions are required to establish that convergence is surprisingly ubiquitous.

Null models for such large-scale questions may seem unreasonably difficult to develop, but a few examples will illustrate possible future directions for research. McGhee [68] has written in support of the development of a ‘periodic table of life’ and Conway Morris [30] cites several studies that might serve as prototypes of such an endeavour. One example is the skeleton space of Thomas & Reif [69]—of the 174 possible pairwise combinations of skeletal element types, all have evolved at least twice, and many evolved early on in the evolution of metazoans [70]. Here, the apparent fullness of this space seems to provide evidence of strong determinism in the history of life.

A null model would give strength to this subjective impression and provide the opportunity to truly understand the significance of this pattern. For example, a researcher might ask: ‘Given the limited number of possible combinations, and known rates of transition among skeleton types, just how unlikely is it that every one of the possible pairwise combinations has evolved more than once?’ or ‘If animals are transitioning between skeletal element types sufficiently rapidly, even at random, is it any surprise that they would land on each region of a 174-square board more than once?’ Or it might be better to put the question as: ‘How rapid would evolution have to be in order to account for the observed pattern?’ It may indeed be the case that the observed pattern is extraordinarily unlikely in the absence of some strong, deeply deterministic influences on evolution, but this needs to be demonstrated in order to assert that the independent ‘discovery’ of one of a predetermined set of parameters can count as evidence for preexisting limits to evolution.

For other datasets, it is the perceived ‘emptiness’ of some phenotypic space, rather than its fullness, that suggests strong limits on the evolutionary process. DNA or protein ‘hyperspaces’ provide an excellent example [71], or even a ‘hyperspace’ of possible human society parameters [72]. The immense size of such spaces (10^{39} possibilities for proteins built by arranging 20 possible amino acids in a chain 100 units long, or 10^8 possible types of human societies) compared to the number of actually observed combinations (a few thousand in both cases) points towards strong limits to the process of evolution, the idea being that the size of the occupied region of this space, relative to the total size of the space, would be much larger, or the amount of convergence observed would be much smaller, if evolution were not limited in some currently poorly understood way. Null models have been developed and used to answer questions regarding proteins, but all systems may be amenable to such modelling. Such a model would allow a researcher to ask: ‘Given the number of human societies that have existed, and some estimate of the rates of transitions between types of societies, is it surprising that humans have only discovered 0.001% of the possible types?’ At least one study has already used null models to demonstrate that the presence of convergence, even within a clade that occupies a very small region of phenotypic space, is not necessarily evidence of limits to evolution: a large amount of convergence will also be observed in a clade in which evolution occurs in random directions, whatever the rate of evolution [27].

Methods that rely on comparing the number of realized or convergent phenotypes to the number of possible phenotypes (often huge but still finite) are only applicable to discrete data. For continuous data, the volume of occupied phenotypic space will often be infinitesimal compared to the volume of unoccupied space, no matter how undirected or unlimited are the evolutionary processes that govern diversification. Instead, arguments that involve continuous characters often cite the unexpected similarities in distantly related organisms (the previously noted convergence between marsupials and placentals is just one set of examples among hundreds of others), or the number of times that certain broadly similar phenotypes have evolved (the multiple independent origins of elongate fishes often called ‘eels’ is again one familiar example among many [30]). As should be obvious from the preceding discussion of measures of convergence, many measures are available for describing such enumerative patterns, and null models for assessing the significance of such patterns should be easy to develop. These null models would provide answers to questions such as: ‘Given known rates of evolution among placental and marsupial mammals (along with known correlations of characters, perhaps) along with an estimate of the ancestral state of each, how likely is it for the observed number of ecomorph pairs to have evolved in the absence of shared adaptive regimes or shared developmental constraints?’ or to the eel questions discussed earlier. Those answers in turn would provide valuable insight on whether these instances of convergence represent evidence for unexpected limitations, constraints or influences on the evolutionary process, or whether they are consistent with evolutionary processes as commonly understood.

5. Summary

Convergent evolution is undoubtedly an important phenomenon, but its study has been hindered by inconsistencies and unclarity in its definitions, by the lack of a standard set of techniques for quantifying its magnitude and testing its significance, and by infrequent application of null models for assessing its implications in biology. Although multiple definitions of convergent evolution may be appropriate, given the large number of questions that researchers ask regarding this phenomenon, pattern-based or process-neutral definitions—ones that merely describe convergence as an increase in similarity—are preferred, as they avoid many potential problems that process-based or causally committed definitions can produce. Fortunately, most measures of convergence are already built within a pattern-based framework. Even given this, a variety of different measures are available, each quantifying a different aspect of convergent evolution. Researchers should choose the definitions and methods that best fit their particular questions, although in general measures that only quantify convergence, and do not conflate it with some other process (such as long-term evolutionary stasis) should be preferred.

Caution should be used when it comes to interpreting convergent patterns. The presence of convergence by itself does not indicate that any particular influence, such as adaptation in response to shared selective pressures, or shared developmental constraint, has been acting on the evolutionary process. In fact, it does not even mean that any influence has been acting at all, although even when adaptation is driving the evolution of individual lineages, convergence does not necessarily imply that the lineages share a common selective regime. Additional evidence, often readily available, is needed to support a non-stochastic mechanism for convergent evolution [48]. The fact that a large number of examples of convergence can be compiled does not demonstrate the ubiquity of convergence. For this, studies are needed that compare the observed incidence of convergence to that expected under null models. Given that these are not yet widely available, conclusions about the meaning of convergence for deep limits to evolution seem premature at present.

More generally, the field of convergence studies is sorely in need of greater application of null models. Although many instances of convergence appear striking, it is known that large amounts of convergence can be generated even through a purely stochastic evolutionary process. Viewing convergence as a self-evident phenomenon, documenting its ubiquity with long lists of individual convergent events, and noting researchers' intuitive reactions as evidence of convergence's deep implications—these methods have little persuasive power for researchers who are sceptical of the profound conclusions reached in certain works [1,29–31] (sceptics include [33–36]).

Thus, it is important to subject 'self-evident' instances of convergence to significance testing, to determine whether they are indeed unlikely under a null model of undirected, unlimited evolution. Fortunately, many strong methods are already available for quantifying convergent evolution [10,25–28,53,57] and all lend themselves well to hypothesis testing. All that is needed is greater effort towards quantifying and assessing the significance of convergent patterns. Quantification can be difficult, especially if comparisons are desired between very distantly related taxa. However, given that convergence among such taxa (in different phyla, for example) constitutes some of the most compelling evidence for deep limits on evolution [30,36] and represents some of the most fruitful ground for future studies (why has the 'mantis' body form evolved multiple times among arthropods, but never among amniotes?), efforts should be directed towards such quantification.

Such quantification and significance testing need not be exclusively destructive of researchers' intuitions regarding the 'astounding' nature of some convergent events—it may be that such instances of convergence are even more surprising than our intuition implies! In any case, a more rigorous approach would not only help to provide support for currently subjective assessments of convergence, but it would allow researchers to determine just how unlikely certain instances of convergence actually are. Only then can convergent evolution provide strong substantial evidence concerning the nature, if any, of limits to evolution.

Competing interests. I declare I have no competing interests.

Funding. This work is based on research conducted at the National Evolutionary Synthesis Center (NESCent), which provided a short-term fellowship to the author. NESCent was supported by NSF EF-0905606.

Acknowledgements. This research benefited greatly from conversations with M. Burd. The author is grateful to S. Conway Morris, V. Ling and all other organizers of the 'Are There Limits to Evolution?' symposium held at Cambridge University. R. Powell and one anonymous reviewer provided valuable feedback on an earlier draft.

References

1. Conway Morris S. 1998 *The crucible of creation: the Burgess shale and the rise of animals.* Oxford, UK: Oxford University Press.
2. Patterson C. 1994 Null or minimal models. In *Models in phylogeny reconstruction* (eds RW Scotland, DJ Siebert, DM Williams), pp. 173–192. Oxford, UK: Clarendon Press.
3. Currie A. 2012 Convergence as evidence. *Brit. J. Phil. Sci*. **64**, 763–786. (doi:10.1093/bjps/axs027)
4. Garland T, Adolph SC. 1994 Why not to do two-species comparative studies: limitation on inferring adaptation. *Physiol. Zool.* **67**, 797–828.
5. Kocher TD, Conroy JA, McKaye KR, Stauffer JR. 1993 Similar morphologies of cichlid fishes in Lakes Tanganyika and Malawi are due to convergence. *Mol. Phyl. Evol.* **2**, 158–165. (doi:10.1006/mpev.1993.1016)
6. Losos JB, Jackman TR, Larson A, de Quiroz K, Rodriguez-Schettino L. 1998 Contingency and determinism in replicated adaptive radiations of island lizards. *Science* **279**, 2115–2118. (doi:10.1126/science.279.5359.2115)
7. Schluter D. 2000 *The ecology of adaptive radiation.* Oxford, UK: Oxford University Press.
8. Blackledge TA, Gillespie RG. 2004 Convergent evolution of behavior in an adaptive radiation of Hawaiian web-building spiders. *Proc. Natl Acad. Sci. USA* **101**, 16 228–16 233. (doi:10.1073/pnas.0407395101)
9. Losos JB. 2009 *Lizards in an evolutionary tree: ecology and adaptive radiation of anoles.* Berkeley, CA: University of California Press.
10. Muschick M, Indermauer A, Salzburger W. 2012 Convergent evolution within an adaptive radiation of cichlid fish. *Curr. Biol.* **22**, 1–7. (doi:10.1016/j.cub.2012.10.048)
11. Martin CH, Wainwright PC. 2013 Multiple fitness peaks on the adaptive landscape drive adaptive radiation in the wild. *Science* **399**, 208–211. (doi:10.1126/science.1227710)
12. Roughgarden J. 1995 *Anolis lizards of the Caribbean: ecology, evolution and plate tectonics*. Oxford, UK: Oxford University Press.
13. McGhee G. 2011 *Convergent evolution: limited forms most beautiful.* Cambridge, MA: The MIT Press.
14. Wake DB, Wake MH, Specht CD. 2011 Homoplasy: from detecting pattern to determining process and mechanism of evolution. *Science* **331**, 1032–1035. (doi:10.1126/science.1188545)
15. Reece JB, Urry LA, Cain ML, Wasserman SA, Minorsky PV, Jackson RB. 2013 *Biology.* San Francisco, CA: Benjamin Cummings.
16. Sadava D, Hillis DM, Heller HC, Berenbaum MR. 2011 *Life: the science of biology,* 9th edn. Sunderland, MA: Sinauer Associates.
17. Freeman S, Allison L, Black M, Podgorski G, Quillin K, Moroe J, Taylor E. 2014 *Biological science*, 5th edn. New York, NY: Pearson.
18. Thain M, Hickman M. 2004 *The Penguin dictionary of biology*, 11th edn. New York, NY: Penguin Books.
19. Starr C, Evers CA, Starr L. 2015 *Biology: concepts and applications*, 9th edn. Stamford, CT: Cengage Learning.
20. Pagel M. 2002 *Encyclopedia of evolution,* vol. 1. Oxford, UK: Oxford University Press.

21. Freeman S, Herron JC. 2007 *Evolutionary analysis,* 4th edn. Upper Saddle River, NJ: Pearson Prentice Hall.
22. Hine R. 2008 *Oxford dictionary of biology.* Oxford, UK: Oxford University Press.
23. Russel PJ, Wolfe SL, Hertz PE, Starr C, McMillan B. 2008 *Biology: the dynamic science.* Stamford, CT: Cengage Learning.
24. Travis J, Reznick DN. 2009 Adaptation. In *Evolution: the first four billion years* (eds M Ruse, J Travis), pp. 105–131. Cambridge, MA: Belknap Press of Harvard University Press.
25. Arbuckle KK, Bennett CM, Speed MP. 2014 A simple measure of the strength of convergent evolution. *Meth. Ecol. Evol.* **5**, 685–693. (doi:10.1111/2041-210X.12195)
26. Stayton CT. 2006 Testing hypotheses of convergence with multivariate data: morphological and functional convergence among herbivorous lizards. *Evolution* **60**, 824–841. (doi:10.1111/j.0014-3820.2006.tb01160.x)
27. Stayton CT. 2008 Is convergence surprising? An examination of the frequency of convergence in simulated data sets. *J. Theor. Biol.* **252**, 1–14. (doi:10.1016/j.jtbi.2008.01.008)
28. Ingram T, Mahler DL. 2013 SURFACE: detecting convergent evolution from comparative data by fitting Ornstein–Uhlenbeck models with stepwise AIC. *Meth. Ecol. Evol.* **4**, 416–425. (doi:10.1111/2041-210X.12034)
29. Dennett DC. 1995 *Darwin's dangerous idea.* New York, NY: Simon & Schuster.
30. Conway Morris S. 2004 *Life's solution: inevitable humans in a lonely universe*. Cambridge, UK: Cambridge University Press.
31. Conway Morris S. 2008 *The deep structure of biology: is convergence sufficiently ubiquitous to give a directional signal*? West Conshohocken, PA: Templeton Press.
32. Vermeij GJ. 2006 Historical contingency and the purported uniqueness of evolutionary innovations. *Proc. Natl Acad. Sci. USA* **103**, 1804–1809. (doi:10.1073/pnas.0508724103)
33. Gould SJ. 1989 *Wonderful life: the Burgess shale and the nature of history.* New York, NY: Norton.
34. Gould SJ. 2002 *The structure of evolutionary theory.* Cambridge, MA: Harvard University Press.
35. Powell R. 2009 Contingency and convergence in macroevolution: a reply to John Beatty. *J. Phil.* **106**, 390–403.
36. Powell R. 2012 Convergent evolution and the limits of natural selection. *Eur. J. Phil. Sci.* **2**, 355–373. (doi:10.1007/s13194-012-0047-9)
37. Simpson GG. 1949 *The meaning of evolution.* New Haven, CT: Yale University Press.
38. Keeton WT. 1967 *Biological science.* New York, NY: W.W. Norton & Co.
39. Darwin CR. 1959 *On the origin of species by means of natural selection, or, the preservation of favoured races in the struggle for life*. London, UK: John Murray.
40. Collar DC, Reece JS, Alfaro ME, Wainwright PC, Mehta RS. 2014 Imperfect morphological convergence: variable changes in cranial structures underlie transitions to durophagy in moray eels. *Am. Nat.* **183**, E168–E184. (doi:10.1086/675810)
41. Mayr E. 1970 *Populations, species, and evolution: an abridgment of 'animal species and evolution'.* Cambridge, MA: Harvard University Press.
42. Futuyma DJ. 1998 *Evolutionary biology*, 4th edn. Sunderland, MA: Sinauer Associates, Inc.
43. Losos JB. 2011 Convergence, adaptation, and constraint. *Evolution* **65**, 1827–1840. (doi:10.1111/j.1558-5646.2011.01289.x)
44. Sanderson MJ, Hufford L. 1996 *Homoplasy: the recurrence of similarity in evolution.* London, UK: Academic Press.
45. Arendt J, Reznik D. 2007 Convergence and parallelism reconsidered: what have we learned about the genetics of adaptation? *Trends Ecol. Evol.* **23**, 26–32. (doi:10.1016/j.tree.2007.09.011)
46. Wake DB. 1991 Homoplasy—the result of natural selection, or evidence of design limitations? *Am. Nat.* **138**, 543–567. (doi:10.1086/285234)
47. Harmon LJ, Kolbe JJ, Cheverud JM, Losos JB. 2005 Convergence and the multidimensional niche. *Evolution* **59**, 409–421. (doi:10.1111/j.0014-3820.2005.tb00999.x)
48. Sober E. 2005 Is drift a serious alternative to natural selection as an explanation of complex adaptive traits? In *Philosophy, biology, and life* (ed. A O'Hear), pp. 125–154. Cambridge, UK: Cambridge University Press.
49. Blount ZD, Borland CZ, Lenski RE. 2008 Historical contingency and the evolution of a key innovation in an experimental population of *Escherichia coli. Proc. Natl Acad. Sci. USA* **105**, 7899–7906. (doi:10.1073/pnas.0803151105)
50. Jasmin J-N, Zeyl C. 2012 Life-history evolution and density-dependent growth in experimental populations of yeast. *Evolution* **66**, 3789–3802. (doi:10.1111/j.1558-5646.2012.01711.x)
51. Wiser MJ, Ribeck N, Lenski RE. 2013 Long-tern dynamics of adaptation in asexual populations. *Science* **342**, 1364–1367. (doi:10.1126/science.1243357)
52. Purnell MA, Bell MA, Baines DC, Hart PJB, Travis MP. 2007 Correlated evolution and dietary change in fossil stickleback. *Science* **317**, 1887. (doi:10.1126/science.1147337)
53. Winemiller KO. 1991 Ecomorphological diversification in lowland freshwater fish assemblages from five biotic regions. *Ecol. Monogr.* **61**, 343–365. (doi:10.2307/2937046)
54. Ackerly DD, Donoghue MJ. 1998 Leaf size, sampling allometry, and Corner's rules: phylogeny and correlated evolution in maples (*Acer*). *Am. Nat.* **152**, 767–791. (doi:10.1086/286208)
55. Osborn HF. 1905 The ideas and terms of modern philosophical anatomy. *Science* **21**, 959–961. (doi:10.1126/science.21.547.959-a)
56. Haas O, Simpson GG. 1946 Analysis of some phylogenetic terms, with attempts at redefinition. *Proc. Am. Phil. Soc.* **90**, 319–349.
57. 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. (doi:10.1111/evo.12729)
58. Omland KE. 1999 The assumptions and challenges of ancestral state reconstructions. *Syst. Biol.* **48**, 604–611. (doi:10.1080/106351599260175)
59. Raup DM, Gould SJ. 1974 Stochastic simulation and evolution of morphology—towards a nomothetic paleontology. *Syst. Biol.* **23**, 305–322. (doi:10.1093/sysbio/23.3.305)
60. Butler MA, King AA. 2004 Phylogenetic comparative analysis: a modeling approach for adaptive evolution. *Am. Nat.* **164**, 683–695. (doi:10.1086/426002)
61. Beaulieu JM, Jhwueng D-C, Boettiger C, O'Meara BC. 2012 Modeling stabilizing selection: expanding the Ornstein-Uhlenbeck model of adaptive evolution. *Evolution* **66**, 2369–2383. (doi:10.1111/j.1558-5646.2012.01619.x)
62. Mahler DL, Revell LJ, Glor RE, Losos JB. 2010 Ecological opportunity and the rate of morphological evolution in the diversification of Greater Antillean anoles. *Evolution* **64**, 2731–2745. (doi:10.1111/j.1558-5646.2010.01026.x)
63. Mahler DL, Ingram T, Revell LJ, Losos JB. 2013 Exceptional convergence on the macroevolutionary landscape in island lizard radiations. *Science* **341**, 292–295. (doi:10.1126/science.1232392)
64. Donoghue MJ. 2005 Key innovations, convergence, and success: macroevolutionary lessons from plant phylogeny. *Paleobiology* **31**, 77–93. (doi:10.1666/0094-8373(2005)031[0077:KICASM]2.0.CO;2)
65. Jaekel M, Wake DB. 2007 Developmental processes underlying the evolution of a derived foot morphology in salamanders. *Proc. Natl Acad. Sci. USA* **104**, 20 437–20 442. (doi:10.1073/pnas.0710216105)
66. Stein BE. 2000 Morphology of subterranean rodents. In *Life underground: the biology of subterranean rodents* (eds EA Lacey, JL Patton, GN Cameron), pp. 19–61. Chicago, IL: The University of Chicago Press.
67. Leal M, Knox AK, Losos JB. 2002 Lack of convergence in aquatic *Anolis* lizards. *Evolution* **56**, 785–791. (doi:10.1111/j.0014-3820.2002.tb01389.x)
68. McGhee G. 2014 Convergent evolution: a periodic table of life? In *The deep structure of biology: is convergence sufficiently ubiquitous to give a directional signal*? (ed. S Conway Morris), pp. 17–31. West Conshohocken, PA: Templeton Press.
69. Thomas RDK, Reif W-E. 1993 The skeleton space: a finite set of organic designs. *Evolution* **47**, 341–356. (doi:10.2307/2410056)
70. Thomas RDK, Shearman RM, Stewart GW. 2000 Evolutionary exploitation of design options by the first animals with hard skeletons. *Science* **288**, 1239–1242. (doi:10.1126/science.288.5469.1239)
71. Smith T, Morowitz H. 1982 Between physics and history. *J. Mol. Evol.* **18**, 265–282. (doi:10.1007/BF01734104)
72. Cronk L. 1999 *That complex whole: culture and the evolution of human behavior*. Boulder, CO: Westview.