Historical Comment

Divergent Uses of “Parallel Evolution” during the History of The American Naturalist

Yoel E. Stuart*

Department of Integrative Biology, University of Texas, Austin, Texas 78712
Submitted March 28, 2018; Accepted August 10, 2018; Electronically published November 15, 2018
Online enhancements: appendix.

Abstract: The mechanistic link between natural selection and parallel evolution is well established. Natural selection is the only known deterministic process that can regularly overcome chance and historical contingency to generate the evolution of similar characteristics in independent populations inhabiting similar environments. However, the ready inference of natural selection from parallel evolution has been established only relatively recently. Here, I review the use of “parallel evolution” in the first 125 years of The American Naturalist and show that there were other well-accepted definitions of the term through the history of the field. I discuss the legacy of those alternative ideas and how they helped to shape evolution and ecology as we know them today and finish by discussing a geometric use for “parallel” that may reduce terminological confusion.

Keywords: convergence, development, divergence, mutation, natural selection, parallel.

Introduction

Marquee evidence for the role of natural selection in evolution is provided by the repeated convergence on similar characteristics among independent populations inhabiting similar environments. That is, such parallel evolution (table 1) strongly implies a deterministic mechanism that can overcome the vagaries of chance and historical contingency, and the textbook, mechanistic explanation for this parallel evolution (or, equivalently, convergent evolution; table 1) is natural selection.1 However, in this article, I revisit the first ~125 years of The American Naturalist to show that the link between selection and parallel evolution has not always been so obvious and accepted. Mutational bias and developmental constraint were vibrant, rival, mechanistic explanations for the ubiquity of parallel evolution. Moreover, for much of the history of the use of the term, mechanism was not even considered; rather, parallel evolution was merely a pattern used to describe systematic relationships or development during ontogeny. Thus, use of the term has been quite fluid (table 2; appendix, available online), with the modern understanding (Reznick and Bryga 1996; Reznick et al. 1996; Losos 2011) becoming standardized only in the past 30 years or so, and despite this canalization, there remains controversy (see Arendt and Reznick 2007, who advocate synonymizing “parallelism” with “convergence”). Thus, the definition will likely continue to evolve. In light of the varied meanings of “parallel” reviewed here, I argue that semantic confusion can be reduced by adopting a quantitative, geometric terminology for parallel change. That geometric view has been suggested by recent authors and is being increasingly used, and I discuss the practicality of this new twist on the oft-changing terminology.

Historical Uses of “Parallel” and “Parallelism” in The American Naturalist

In October 2017, I searched “parallel evolution” and “parallelism” (hereafter “parallel” for brevity), using the search function on The American Naturalist’s website (https://www.journals.uchicago.edu/toc/an/current), from the beginning of the journal in 1867 until 2004. I highlight 19 articles from these search results that illustrate the breadth of patterns and processes described by “parallel” through the history of the journal and the field of evolutionary biology (boldfaced in table 2). My purpose here was not to report and discuss every single use of “parallel” but merely to ask whether (and how) the use and meaning of the term changed through time. Then, in late February 2018, I conducted a

1. In this essay, I treat convergent and parallel evolution as equivalent: the independent evolution of similar features in different lineages in similar environments; i.e., under similar selection (Arendt and Reznick 2007; Losos 2011). I focus on the term “parallel” in this essay, rather than on “convergence,” because of my own historical contingency. My current research relies on a geometric definition of repeated evolution; i.e., we quantify evolutionary change literally as parallel or not, by calculating angles between vectors of divergence.
<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convergent evolution</td>
<td>Traditionally, the independent evolution of similar features (e.g., phenotype, genotype, function) in similar environments by distantly related lineages with different starting conditions</td>
</tr>
<tr>
<td>Convergent evolution (geometric)</td>
<td>When two derived taxa are closer together in phenotype/genotype/functional space than their ancestral taxa</td>
</tr>
<tr>
<td>Divergent evolution (geometric)</td>
<td>When two derived taxa are farther apart in phenotype/genotype/functional space than their ancestral taxa</td>
</tr>
<tr>
<td>Ontogeny recapitulates phylogeny</td>
<td>Also called the theory of recapitulation, the idea that the embryonic development of an extant organism goes through stages resembling the adults of more primitive species in its lineage</td>
</tr>
<tr>
<td>Parallel evolution</td>
<td>Traditionally, the independent evolution of similar features (e.g., phenotype, genotype, function) in similar environments by closely related lineages with similar starting conditions</td>
</tr>
<tr>
<td>Parallel evolution (geometric)</td>
<td>When evolutionary trajectories through phenotype space have a very low angle between them</td>
</tr>
<tr>
<td>Homology</td>
<td>Shared ancestry of phenotypic traits or genes across different taxa; e.g., the forelimbs of vertebrates, though widely divergent, share the same ancestral tetrapod structure; derived homologous characters (i.e., synapomorphies) are useful for inferring phylogenetic relationships; compare to homoplasy</td>
</tr>
<tr>
<td>Homoplasy</td>
<td>A derived trait present in two taxa, not because of homology but through convergent evolution (e.g., wings in insects and birds); homoplasmatic traits confuse phylogenetic inference</td>
</tr>
<tr>
<td>Serial segmentation</td>
<td>A type of homology, wherein repeated segments in the same organism have a shared developmental origin; some workers considered these repeated segments as a type of parallelism</td>
</tr>
<tr>
<td>Solution space</td>
<td>The complete set of adaptive solutions to a problem set by natural selection</td>
</tr>
</tbody>
</table>

The first use of “parallel” in *The American Naturalist* came in volume 1, in reference to whether human fore- and hind-limbs show homologous (table 1; i.e., parallel) form and function. In the sardonically titled “The Hand as an Unruly Member,” Wilder (1868) ridiculed the idea that in “each bone of the leg we find characters which belong, partly to the ulna, and partly to the radius” (p. 632), as Wilder’s contemporaries believed had to be true for legs and arms that were supposedly parallel in origin and development. Wilder believed that this was forcing data to fit preconceived theory and was not shy in saying so. Hubbs (1944) describes this use of “parallel” as “serial homology” (p. 293); for example, “compare the fore limb and hind limb of a horse [versus] . . . the fore limb of the horse [and] . . . the [teleost] pectoral fin and girdle” (p. 293). There may often be, Hubbs (1944) argued, “vastly greater correspondence in anatomy, and in homology, between serial structures in one individual than between the ‘same’ structures [across taxa]” (p. 293). Unlike our modern use of “parallel” (table 1), this use did not focus on comparisons among populations or species but rather emphasized comparisons among different structures within a single organism. This first use of “parallel” therefore represents an early foray into the field of “evodevo” rather than a description of repeated evolution among replicate populations.

**Variational Series**

“Parallel” was also used to describe what I am calling “variations series.” These are independently evolving lineages that show repeated evolution through the “famous three-fold parallelism between [(1)] the succession of forms in geological time, [(2)] the gradation of [extant] animals from...
low to high and [(3)] the succession of [embryological] stages through which the higher members of a group pass in their individual development” (Wilson 1941, p. 22).

Variational series progress toward higher forms, in parallel, across clades (usually genera): “As a general law . . . the same specific form has existed through a succession of genera, and perhaps in different epochs of geologic time” (Cope 1868, p. 243; see also Cope 1876). For example, Earle (1895) reports “two parallel series” in the “extinct perissodactyle Palaeosyops” that each evolved “increased height of the crowns of the molars, reduction of the intermediate tubercles, increase in size of the skull, and . . . indications of the development of horns” (pp. 622, 625). Similarly, Dobzhansky (1933) argued for parallel variational series in his study of extant lady beetles, noting that “large groups of related species and genera exhibit parallel series of patterns, upholding the rule of homologous series in variation” (p. 97).

Per Wilson (1941), such serial progress also manifests itself in the famous idea, championed in America by Louis Agassiz (Packard 1898), that ontogenetic progression through development recapitulates an organism’s phylogenetic history. Cope (1876, p. 219), for example, described “exact parallelism” as a tight resemblance between the adults of lower species and the transitional embryos of higher species. This variational-series use of “parallel” dominated the late 1800s and early 1900s (table 2; fig. 1) alongside the “ontogeny-recapitulates-phylogeny” (table 1) movement. Workers who used this definition often shoehorned their data to fit their preconceived notions of evolution; that is, taking the a priori position that genera and developmental stages should evolve in parallel and then interpreting their data accordingly. Often, data that broke from this pattern were either ignored or explained away as exceptions. For example, Cope (1876, p. 219) created a category for “inexact parallelism,” arguing that ontogeny recapitulating phylogeny is the default state and that deviations are easily explained by derived changes in the ontogenetic timing of different developmental modules across lineages. Similarly, Wilson (1941, p. 27) argued that proponents of this idea never meant to take a “literal interpretation of the statement that ontogeny is a recapitulation of phylogeny. . . . [I]n both generalization and theory the mind is directed to a general resemblance, not to a detailed one. . . . [This is what we mean by] a ’fish-like stage.’” Proponents of this school of

Table 2: Use and meaning of the term “parallel evolution” has changed through time, as revealed through the history of The American Naturalist

<table>
<thead>
<tr>
<th>Use of “parallel”</th>
<th>Citations</th>
</tr>
</thead>
<tbody>
<tr>
<td>Serial segmentation</td>
<td>Wilder 1868; Hubbs 1944</td>
</tr>
<tr>
<td>Variational series</td>
<td>Dawson 1875; Cope 1876; Dall 1877; Ward 1878; Packard 1882, 1898; Hyatt 1884, 1888; Abbott 1887; Osborn 1888, 1891; Beecher 1893; Earle 1895; Cook 1908; Greggs 1909; Dendy 1915; Dobzhansky 1933; Wilson 1941; Copeland 1947; Goodman 1960</td>
</tr>
<tr>
<td>Systematics</td>
<td>Packard 1874; Baur 1891; Ameghino 1893; Osborn 1893, 1902a, 1902b; Hatcher 1894; Bensley 1901; Patten 1902; Lyon 1905; Campbell 1912; Kellogg 1913; Brues 1921; Metcalf 1923; Crampton 1925; Dobzhansky 1933; Gregory 1936; Redfield 1936; Balduf 1938; Wiener 1943; Hubbs 1944; Etkin and Livingston 1947; Bennett 1954; Green 1955; Kuznezov 1956; Nanney 1960; Thorne 1963; Seto et al. 1964; Bergh and Lippert 1965; Kezer et al. 1965; Carson 1969; Hecht and Edwards 1976</td>
</tr>
<tr>
<td>Genetic basis</td>
<td>Bartlett 1916; Gates 1917, 1925, 1936, 1951; Hubbs 1926; Dobzhansky 1933; Miller and Hoy 1939; Huskins 1941; Epling 1947; Goldschmidt 1950; Baker 1959</td>
</tr>
<tr>
<td>Developmental constraint</td>
<td>Ryder 1878; Davenport 1916; Wieland 1925; Wilson 1941; Volpe 1955; Hecht and Edwards 1976</td>
</tr>
<tr>
<td>Selection</td>
<td>Packard 1871; Henn 1912; Clark 1913; Shelford 1914; Gerould 1916; Nichols 1916; Goldschmidt 1918, 1947; Osborn 1927; Dobzhansky 1933; Mozley 1936; Dice 1940; Hubbs 1940; Wilson 1941; Hiesey 1942; Clausen et al. 1947; Moore 1950, 1952; Blair 1951; Kuznezov 1956; Van Valen 1960; Wasserman 1963; Schluter and McPhail 1992; Reznick and Bryga 1996; Reznick et al. 1996; Langerhans and DeWitt 2004</td>
</tr>
</tbody>
</table>

Note: The boldface citations are highlighted in the main text and/or were found in October 2017, as described in the text. The citations not in boldface are from a more complete literature review in late February 2018, described in the text, which returned 1,183 articles from the journal’s founding to the present day. I stopped this more complete search at the year 1968, 100 years after the founding of the journal, for a total of 496 articles reviewed. By this point, the varied uses of “parallel” had been well established, and I was traveling along the flat part of a curve of diminishing returns. This citation list is not exhaustive, as I did not include articles where “parallel” was used without any obvious relation to evolution or when it was used (1) as a mere synonym for “similar,” “analogous,” or “correlation” (e.g., between structure and function), (2) by prominent, prolific authors who use “parallel” in passing reference to previous works of their own, or (3) as jargon, whose meaning I could not discern from context. Additional, rarer categories of use for “parallel” that I found during my literature search included (1) a geometric description of traits or characteristics, e.g., parallel deposition of lines on a snail shell, (2) a description of similar behaviors, (3) reference to epigenetics and maternal effects, (4) use as a synonym for coevolution, (5) use as an explanation of biomechanical similarity, (6) use as a synonym for homolog, or (7) the result of chance. Moreover, I do not report “parallel” as a description of communities that have converged in their ecological specialists, e.g., the placental/marsupial radiations, if this pattern was reported without reference to the mechanism that might cause such convergence.
“parallel” were prolific (table 2), and though the mass of data they collected eventually undermined their own notions, their data played a pivotal role in advancing research in development, paleontology, and systematics.

**Systematics**

Starting in the early 1900s, “parallel” became more often used as evidence for or against phylogenetic relationships but without any underlying mechanistic inference about adaptation (table 2; fig. 1). Osborn (1902a, p. 260), for example, discussed “analogy [as] . . . all cases of the similar evolution of organs either of common or [parallel] origin due to similarity of function. . . . [Analogy] may be very misleading as to homology.” Hubbs (1944) spent nearly 20 pages of *The American Naturalist* discussing in part how the “frequency of parallel evolution among related forms . . . often makes it impossible to decide whether or not detailed agreements are due to common origin” (p. 299).

This was part of a broader discussion, summarized by Hubbs (1944) but common throughout *The American Naturalist*, about phylogenetic inference and the search for homologous characters (table 1) in the face of homoplasy (table 1) due to parallel and convergent evolution. This body of work created the foundation for cladistics and eventually modern phylogenetics, which relies on shared (i.e., homologous) but derived characters to infer systematic relationships.

**Genetic Basis of Evolution**

Dobzhansky (1933) observed parallel variation across different lineages of lady beetles and speculated that this “par-

---

3. Notably, Osborn (1902a) helped codify our modern uses of parallel and convergent evolution: “By parallelism is meant the independent acquisition of similar structure in . . . [animals] which are themselves nearly related, and by convergence such acquisition in [animals] which are not closely related” (p. 261).

4. In this framework, many workers used the terms “convergent” and “parallel” almost synonymously (see quotations in the appendix), revealing that the distinction between convergent and parallel evolution (Arendt and Reznick 2007) is not only a modern problem.

5. This use of “parallel,” and the two that follow, differ from the three above in that they directly invoke mechanism.
allelism is, probably, due to the essential similarity of the germ-plasms of the related species” (p. 108). Without naming it directly, Dobzhansky (1933) seems to be arguing for parallelism evolving from standing genetic variation: “differences between the geographical races may be expressed in terms of the relative frequencies of genes responsible for the production of various patterns in the different localities. What we are studying is essentially the geography of the genes . . . not the [color] patterns themselves” (p. 120). In contrast, Gates (1936, p. 514) writes, “the capacity of the germplasm to produce parallel mutations in successive [taxonomic] series . . . appears to be as near to an explanation as we have yet arrived” for parallel evolution. This use of “parallel” argued for a role of repeated de novo mutation in adaptation and anticipated modern findings that regions of a genome with higher mutation rates may be more likely to repeatedly supply adaptive mutations (e.g., Chan et al. 2010). It took nearly 50 years for this view of “parallel” to arise in The American Naturalist (fig. 1), as this use required the modern synthesis of the 1920s–1940s, which reconciled Mendelian inheritance and population genetics into a single evolutionary framework. The modern synthesis also marked the end of the “variational-series” definition of “parallel” (fig. 1) and set the stage for considering the role of selection in parallel evolution (see below). By investigating the genetic basis of evolutionary change, this genetic view of “parallel” inspired whole research programs investigating the relative roles of standing genetic variation and de novo mutation for adaptation.

Developmental Constraint

Parallel evolutionary change was also used to infer the limited pathways that evolution might take. Wilson (1941) asked whether parallel outcomes arose from “adaptive mechanisms” or “the inertia of heredity” (p. 30). Hecht and Edwards (1976) favored the latter: “the limiting constraints of the developmental processes of large groups of organisms . . . make it all the more likely that parallelisms will occur” (p. 672). That is, parallel evolution might be likely not because of similar selective conditions but rather because of shared biases among lineages in the production of the variation on which selection acts (Losos 2011). Indeed, the (usually unstated) mechanism for recapitulation of phylogeny during ontogeny is constraint; vital developmental pathways are conserved, leading to related lineages showing a “general resemblance” (Wilson 1941, p. 27) between adult and embryonic forms.

Constraint serves as an important alternative hypothesis to the modern textbook explanation of selection as the cause of parallel evolution. Shared patterns of development, genetic correlation, pleiotropy, and epistasis can each keep selection from exploring all solutions to a given environmental problem (Losos 2011). Instead, selection is channeled down a limited set of pathways, leading to shared outcomes. To compare these alternative hypotheses, modern phylogenetic comparative methods have been developed to allow for estimates of “the inertia of heredity” through the quantification of phylogenetic signal—that is, the extent to which phylogenetic history and, presumably, shared constraints might explain patterns of trait evolution (Bloomberg et al. 2002). Another research program investigates the extent to which clades have explored “solution space” (table 1) during the generation of variation (e.g., Raup 1967); restriction to a small part of solution space implies constraint. Finally, evolutionists have experimented tested for the role of constraint in parallel evolution in the lab and in the wild. They have used breeding strategies and other techniques to artificially break genetic correlations (e.g., Beldade et al. 2002; Martin and Wainwright 2013), document epistasis (e.g., Kvitak and Sherlock 2011; Salverda et al. 2011), and generate novel phenotypes (e.g., Beldade et al. 2002) to ask what evolves when constraints on variation are relaxed.

Selection

Inference of the role of selection in parallel evolution arrived early in the history of The American Naturalist (table 2; fig. 1), but the clearest mechanistic reference to selection came from Nichols (1916, p. 567), who distinguished “geographically separated” and perhaps nonadaptive cases of parallel change from “environmental parallels” shaped by selection. From there, selection became increasingly accepted as the mechanistic cause of parallel evolution (fig. 1) and dominates today: “Convergent or parallel evolution is a powerful means for evaluating adaptation because it implies a cause-and-effect relationship between selection and the response to selection” (Reznick and Bryga 1996, p. 339). This canonization of selection with “parallel” in The American Naturalist (and the field) took a long time—only beginning 60 years after The Origin of Species and a half-century after the start of The American Naturalist. This is because it took a long time to discover and develop a mechanism of inheritance that could transmit characteristics favored by natural selection. At first, Darwin favored blending inheritance, but it quickly became apparent that natural selection could not drive evolutionary change under this model (Provine 2001; Bowler 2003). As Mendelians and biometricians argued about the nature of inheritance, natural selection fell out of favor as a mechanistic explanation for evolution.

The resurrection of natural selection and its subsequent link to “parallel” stemmed from four developments. First, as part of the modern synthesis, Mendel was rediscovered and popularized, thereby solving the nature of inheritance and reviving the idea that natural selection could be an im-
Important driver of evolution. Second, the modern synthesis also provided the population-genetic and quantitative-genetic machinery for understanding Darwinian evolution. Third, by inductive reasoning, selection is the only deterministic mechanism that can regularly overcome random mutation and historical contingency to generate repeated outcomes in independent lineages. Fourth, this inductive reasoning has been confirmed by direct measurements of selection coupled with evolutionary experiments in lab and field. Such studies have repeatedly found, across systems and taxa, causal links between shared environments, similar selection, and parallel evolution (e.g., Reznick et al. 1990, 1996; Lenski et al. 1991; Reznick and Bryga 1996; Losos et al. 1997; Woods et al. 2006).

**Figure 2:** The extent and magnitude of parallel evolution can be described geometrically. A, Hypothetical trait means are plotted along two phenotypic axes for four different parapatric population pairs (a–d) that each span a habitat boundary (open and filled circles represent different habitats). Pairs are connected by a solid line, which represents that pair’s divergence vector across habitats. The angle (θ) between divergence vectors can be calculated, as can each vector’s length (L) and the difference between vectors in length (ΔL). B, Vectors can evolve along a continuum from parallel (θ ≈ 0°) to antiparallel (i.e., θ ≈ 180°), and parallel change can lead to convergence or divergence. That is, under this framework, “parallel” and “convergent” describe two different phenomena, depending on the starting and ending positions of each population and the vector directions. Axes in B are hypothetical trait axes and could represent axes of phenotypic (genetic, transcriptomic, etc.) variation. C, D, There is a (non)parallel continuum for lake-stream divergence for 16 parapatric threespine stickleback population pairs (see Stuart et al. 2017). Plotted in C are population means (± standard errors) for gill raker number and first dorsal spine length. Lines connect paired lake and stream populations. D presents the histograms of the all the observed pairwise angles and vector length differences calculated for the empirical vectors in C. Most vector comparisons reveal orthogonal trajectories through trait space, but with a wide distribution from parallel to antiparallel. Similarly, most vectors have similar lengths, but with large variance.
The multiple uses of “parallel” throughout the history of *The American Naturalist* confirm that the link between “parallel” and selection is by no means the only possible usage (table 2) and suggest that the definition will continue to evolve. Indeed, for example, the distinction between parallel and convergent evolution has become even more difficult to pinpoint as genetic and molecular approaches have improved over the past decades, and it has been suggested that all instances of repeated evolutionary outcomes in replicated populations might be encompassed by a single term: “convergence” (Arendt and Reznick 2007).

Were parallel evolution and convergence to be collapsed in this way, the term “parallel” would be freed up for a more specific, and perhaps more final, use. That is, parallel evolution could be used as a geometric (table 1), quantitative description of the trajectory of evolutionary change. Imagine a vector of divergence describing multtrait evolutionary change between an ancestor population and its descendant population (fig. 2A). Now imagine multiple such vectors for replicate ancestor-descendant pairs, and then imagine calculating the pairwise angles separating those replicate vectors as well as the differences in vector lengths (fig. 2A; e.g., Adams and Collyer 2009; Stuart et al. 2017). Parallel evolution might then be defined as those cases where there are small angles between vectors; that is, divergence proceeds, literally, in parallel across replicates (fig. 2B). Moreover, this approach would allow the quantification of the extent of parallel change across pairs. That is, rather than binning cases into binary, yes/no parallel states, one can consider a geometric continuum of parallel evolution, from convergent (table 1), to parallel, to divergent (table 1), to antiparallel (fig. 2B). For example, figure 2C plots lake-stream divergence in stickleback gill raker number and first dorsal spine length (data from Stuart et al. 2017). The lake-stream divergence vectors for some of these pairs have similar magnitudes and small angles between them (i.e., highly parallel), whereas others have different magnitudes and large angles between them (i.e., highly nonparallel; fig. 2D). This precise, geometric, continuum-friendly approach to “parallel” would reduce terminological confusion by separating parallel/nonparallel change from convergent/divergent change (fig. 2B; table 1), all while enabling a more comprehensive program of study. That is, as Kusnezov (1956) argued in *The American Naturalist*, “lacks of parallelisms are perhaps just as interesting for our understanding of the evolutionary mechanisms as are the instances of parallel developments” (p. 349). Or, from a more recent issue of the journal, “Assessing the nature and relative importance of shared and unique responses should aid in elucidating the relative generality or peculiarity in evolutionary divergence” (Langersmans and Dewitt 2004, p. 335). This mathematical definition gives us a way to quantify those shared and unique responses, and it may be a useful new way to consider “parallel.”

**Acknowledgments**

I thank Daniel Bolnick, David Reznick, Alice Winn, and an anonymous reviewer for thoughtful discussion of these ideas that improved the manuscript. I thank the American Society of Naturalists for inviting me to speak in the “150 Years of *The American Naturalist*” symposium at the biannual meeting of the society in Asilomar. I was supported by National Science Foundation grant DEB-1456462.

**Literature Cited**


Editor: Alice A. Winn