Punctuated equilibrium comes of age

Stephen Jay Gould & Niles Eldredge

The intense controversies that surrounded the youth of punctuated equilibrium have helped it mature to a useful extension of evolutionary theory. As a complement to phyletic gradualism, its most important implications remain the recognition of stasis as a meaningful and predominant pattern within the history of species, and in the recasting of macroevolution as the differential success of certain species (and their descendants) within clades.

Punctuated equilibrium has finally obtained an unambiguous and incontrovertible majority—that is, our theory is now 21 years old. We also, with parental pride (and, therefore, potential bias), believe that primary controversy has ceded to general comprehension, and that punctuated equilibria have been accepted by most of our colleagues (a more conventional form of majority) as a valuable addition to evolutionary theory.

Kellgren began the best book written to celebrate Darwinism's fiftieth birthday by noting how often (and how continually) various critics had proclaimed the Stebbings (death bed) of natural selection. Punctuated equilibrium has also prospered from announcements of its death or triviality and has been featured in much recent discussion.

As a neophyte in 1972, punctuated equilibrium entered the world in unusual guise. We claimed no new discovery, but only a novel interpretation for the oldest and most robust of palaeontological observations: the geologically instantaneous origination and subsequent stability (often for millions of years) of palaeontological morphospecies. This observation had long been ascribed, by Darwin and others, to the notorious imperfection of the fossil record, and was therefore read in a negative light—as missing information about evolution (defined in standard palaeontological textbooks of the time) as continuous anagenetic transformation of populations, or phyletic gradualism.

In a strictly logical sense, this negative explanation worked and preserved gradualism, then falsely equated with evolution itself, amidst an astonishing lack of evidence for this putative main signal of Darwinism. But think of the practical or heuristic dilemma for working palaeontologists: if evolution meant gradualism, and imperfection precluded the observation of such steady change, then scientists could not access the very phenomenon that both motivated their interest and built life's history. As young, committed and ambitious parents, we therefore proposed punctuated equilibrium, hoping to validate our profession's primary data as signal rather than void. We realized that a standard biological account, Mayr's parietic theory of speciation in small populations peripherally isolated from a parental stock, would yield stasis and punctation when properly scaled into the vastness of geological time—for small populations speciating away from a central mass in tens or hundreds of thousands of years, will translate in almost every geological circumstance as a punctation on a bedding plane, not gradual change up a hill of sediment, whereas stasis tends to characterize the long and recoverable history of successful central populations.

Punctuated equilibrium then grew during its childhood and adolescence, in ways both unruily and orderly. Unruily accidents of history included the misunderstandings of colleagues (who, for example, failed to grasp the key claim about geological scaling, misread geological abruptness as true suddenness, and then interpreted punctuated equilibrium as a saltational theory), and the purposeful misues of creationist foes as this political issue heated up in the United States during the late 1970s (although we took pride in joining with so many colleagues for a successful fight against this philistine scourge, as one of us testified in the Arkansas 'monkey' trial in 1981 and the other wrote a book on creationist distortions).

But orderly extensions, implicit in the undeveloped logic of our original argument, fuelled the useful growth of punctuated equilibrium to fruitful adulthood. (We now realize how poorly we initially grasped the implications of our original argument; we thank our colleagues, especially S. M. Stanley and E. S. Vrba, for development. We originally focused on tempo, but more important theoretical arguments flowed from implications concerning evolution's mode—particularly the causes surrounding our two major claims for equilibrium, or stasis of established species, and the need to reformulate macroevolution, notably the key phenomenon of trends, as an accumulation of discrete speciation events rather than undefinable segments of continuity—a subject encompassed by debate about species selection or species sorting.)

Punctuated equilibrium and macroevolution

Stasis and Its meaning. We opened our original paper with a section on what philosopher N. R. Hanson called "the cloven hoofprint of theory," or the structuring of all supposedly objective observation by expectations of prevailing general views. Stasis, as palpable and observable in virtually all cases (whereas rapid punctuations are usually, but not always, elusive), becomes the major empirical ground for studying punctuated equilibrium. Putting together the philosophical insight of ineluctable theoretical bias, with the empirical theme of the tractability of stasis, we devised a motto: "stasis is data." For no bias can be more constraining than invisibility—and stasis, inevitably read as absence of evolution, had always been treated as a non-subject. How odd, though, to define the most common of all palaeontological phenomena as beyond interest or notice! Yet palaeontologists never wrote papers on the absence of change in lineages before punctuated equilibrium granted the subject some theoretical space. And, even more, as palaeontologists didn't discuss stasis, most evolutionary biologists assumed continual change as a norm, and didn't even know that stability dominates the fossil record. Mayr has written: "Of all the claims made in the punctuationalist theory of Eldredge and Gould, the one that encountered the greatest opposition was that of 'pronounced stasis as the usual fate of most species', after having completed the phase of origination...I agree with Gould that the frequency of stasis in fossil species revealed by the recent analysis was unexpected by most evolutionary biologists.

As the most important change in research practice provoked by punctuated equilibrium, stasis has now exited from its closet of non-definition to become a subject of quantitative investigation in all major fossil groups—from microfossils to molluscs, to mammals. Although punctuated equilibrium deals directly only with stability of species through time, the high-conducive analogue of non-trending in larger clades has also graduated from an undefined non-subject to a phenomenon worth documenting. Moreover, because species often maintain stability through such intense climatic change as glacial cycling, stasis must be viewed as an active phenomenon, not a passive response to unaltered environments. Many leading evolutionary...
theorists, while not accepting our preference for viewing stasis in the context of habitat tracking or developmental constraint, have been persuaded by punctuated equilibrium that maintenance of stability within species must be considered as a major evolutionary problem.

**Macровolution as a problem in species sorting.** If punctuated equilibrium has provoked a shift in paradigms for macroevolutionary theory (see ref. 35 for a defense of this view), the main insight for revision holds that all substantial evolutionary change must be reconciled as higher-level sorting based on differential success of certain kinds of stable species, rather than as progressive transformation within lineages (see Eldredge on taxic versus transversational views of evolution; Simpson, however, in the canonical paleontological statement of the generation before punctuated equilibrium, had attributed 90% of macroevolution to the transformational mode, and only 10% to speciation). Figure 1, our original diagram of punctuated equilibrium, shows how a trend may be produced by differential success of certain species without directional change in any species following its origin.

Darwin’s theory of natural selection locates the causality of evolutionary change at one domain on one level: natural selection and taxic by struggle among individual organisms for reproductive success. Given Darwin’s crucial reliance upon lylian uniformity for extrapolating this mode of change to encompass all magnitudes through all times, the interposition of a level for sorting among stable species breaks this causal reduction and truly, in Stanley’s felicitous term, “decouples” macro-evolution from microevolution. Decoupling is not a claim for the falseness or irrelevance of microevolutionary mechanisms, especially natural selection, but a recognition that darwinian extrapolation cannot fully explain large-scale change in the history of life.

The main point may be summarized as follows. Most macroevolution must be rendered by asking what kinds of species within a clade did better than others (specified more frequently, survived longer), or what biases in direction of speciation prevailed among species within a clade. Such questions enjoin a very different programme of research from the traditional “how did natural selection within a lineage build substantial adaptations during long stretches of time?” The new questions require a direct study of species and their differential success; older queries focused downward upon processes within populations and their extrapolation through time. Darwin’s location of causality in organisms must be superseded by a hierarchical model of selection, with simultaneous and important action at genic, organismal and taxic levels, who so stoutly defended classical Darwinism against older, invalid, and very different forms of group selection, now acknowledges the importance of such clade selection in macroevolution. Punctuated equilibrium has been used as a central concept in the development of hierarchy theory in evolutionary biology.

**Implications.** Any theory with a claim to novelty in broad perspective must enlighten old problems and suggest extensions. The speciational view of macroevolution, which does not strictly require punctuated equilibrium, but which was nurtured and has thrived in its context, requires a reformulation of nearly all macroevolutionary questions. For example, so-called living fossils, once treated as lineages rendered static by optimal adaptation, unusually stable environment, or lack of genetic variation, should be reconceptualized as members of groups with unusually low speciation rates, and therefore little opportunity to accumulate change. (We have no evidence that the species of ‘living fossil’ groups are particularly old. For example, the western Atlantic horseshoe crab, Limulus polyphemus—the ‘type example’ of the phenomenon—has no fossil record at all, whereas the genus can only be traced to the Miocene.)

Going further, the entire tradition of expressing evolutionary change in darwinian units (where 1 darwin equals character change by a factor of e in 1 million years) makes no sense in a speciational context. (If a lineage goes from species A to D in 10 million years through three episodes of rapid change with intervening stasis, a cited rate of so many millidarwins becomes a meaningless average.) We learn as a received truth of evolution, for example, that human brain size increased at an extraordinary (many say unprecedented) rate during later stages of our lineage. But this entrenched belief may be a chimera born of an error in averaging rates over both punctuations and subsequent periods of stasis. Homo sapiens is a young species, perhaps no more than 200,000 years old. If most of our increment accrued quickly at our origin, but we then express this entirety from our origin to the present time as a darwin rate, we calculate a high value because our subsequent time of stasis has been so short. But if the same speciation event, with the same increment in the same time, had occurred two million years ago (with subsequent stasis), the darwin rate for the identical event would be much lower.

Cope’s rule, the tendency for phyletic increase in body size, had generally been attributed to selective value of large size within anagenetic lineages, but is probably better interpreted as greater propensity for speciation in smaller species, for whom increasing size is the only ‘open’ pathway (see Martin on the negative correlation of generic species richness and body size). Raup and Sepkoski proposed a conventional explanation for decreasing rate of background extinction through geological time: generally better adaptation of later species. But Valentine and Gilinsky (personal communication) offer an interesting speciational alternative: if extinction intensities were constant through time, groups with equally high speciation and extinction

![FIG. 1. Three-dimensional sketch contrasting a pattern of relative stability (A) with a trend (B), where speciation (dashed lines) is occurring in both major lineages. Morphological change is depicted here along the horizontal axes, while the vertical axis is time. Though a retrospective pattern of directional selection might be fitted as a straight line in (B), the actual pattern is stasis within species, and differentiation of more than 200,000 years of change in a particular direction. For further explanation, see ref. 1.](image-url)
rates would fare just as well as groups with equally lower rates. But extinction intensities vary greatly, and the geological record features episodes of high dying, during which extinction-prone groups are more likely to disappear, leaving extinction-resistant groups as life’s legacy. Valentine concludes that these clade-characteristic rates are of course not adaptations per se, but effects flowing from clade properties that were established probably during the early radiations that founded the clades."

The most exciting direct extensions of punctuated equilibrium now involve the study of correlated punctuational events across taxa, and the ecological and environmental sources of such cohesion. Eminent testable are Vrba’s

"turnover-pulse hypothesis" of evolution concentrated in punctuational bursts at times of worldwide climatic pulsing, one of which, about 2.5 million years ago, may have stimulated the origin of the genus Homo; and Brett’s hypothesis of "coordinated stasis" for the structuring of major palaeontological faunas. What might be the ecological source of such striking coherence across disparate taxa through such long times?

The empirics of punctuated equilibrium

Like all major theories in the sciences of natural history, including natural selection itself, punctuated equilibrium is a claim about relative frequency, not exclusivity. Phyletic gradualism has been well documented, again across all taxa from microfossils to mammals. Punctuated equilibrium surely exists in abundance, but validation of the general hypothesis requires a relative frequency sufficiently high to impart the pre-dominant motif and signal to life’s history. The issue remains unsettled, but we consider (in our biased way) that four classes of evidence establish a strong putative case for punctuated equilibrium in this general sense.

Individual cases. Examples of stasis alone (cited earlier) and simple abrupt replacement, although conforming to expectations of punctuated equilibrium, are not direct evidence for our mechanism: for stasis might just be a null in anagenetic gradualism (though pervasive stasis for long periods in all species of a fauna (a common finding) would require special pleading from gradualists), and replacement might represent rapid transformation without branching, or migration of a distant (phyletic or geographic) relative rather than evolution in situ. A good test of punctuated equilibrium requires (in addition to the obvious need for documented rapidity in an interval known to be sufficiently short) both a phyletic hypothesis to assert sister-group relationship of the taxa involved, and survival of putative ancestors to affirm an event of true branching rather than rapid phyletic transformation.

Given these stringent requirements, and in the light of such an imperfect fossil record, we are delighted that so many cases have been well documented, particularly in the crucial requirement of ancestral survival after punctuated branching. Williamson’s discovery

of multiple molluscan speciation events in isolated African lakes, with return of ancestral lineages upon reconnection with parental water bodies, has been most widely discussed and disputed (although all accept the punctuational pattern). Cheetham’s elegant and meticulously documented story of evolution in the bryozoan Metrarabdotos from Tertiary strata of the Caribbean is particularly gratifying (Fig. 2) in the number of purely punctuational events, the full coverage of the lineage, and the unusual completeness of documentation, especially as Cheetham began his study expecting to reconfirm a gradualistic interpretation (writing to McKinney: "The chronocline I thought was represented . . . is perhaps the most conspicuous casualty of the restudy, which shows that the supposed clade members largely overlap each other in time. Eldredge and Gould were certainly right about the danger of stringing a series of chronologically isolated populations together with a gradualist’s expectations.")

On the subject of punctuational corrections for received gradualistic wisdom, Prothero and Shubin have shown that the most 'firmly' gradualistic part of the horse lineage (the general, and false, exemplar of gradualism in its totality), the Oligocene

![Diagram](https://example.com/diagram.png)

FIG. 2 Reproduced with permission from ref. 65. Phylogenetic (strataphenetic) tree of Caribbean species of Metrarabdotos. Relationships were constructed by calculating euclidean distances between sampled populations using all canonical scores from discriminant analysis, and connecting nearest morphological neighbours in stratigraphic sequence. Morphological distances (horizontal axis) between inferred ancestor and descendant species are to scale; those between species of different pairs are not necessarily so (for example, M. kugleri is morphologically more distant from M. n. sp. 3 than from its ancestor, M. n. sp. 2, even though the diagram makes it appear otherwise). Within-species fluctuation in species-distinquishing morphology is based on pairwise discriminant scores scaled to the distances between species pairs.
transition from *Mesohippus* to *Miophippus*\(^{33}\), conforms to punctuated equilibrium, with stable clades in all species of both lines, transition by rapid branching rather than phylectic transformation, and stratigraphic overlap of both genera (one set of beds in Wyoming has yielded three species of *Mesohippus* and two of *Miophippus*, all contemporaries). Prothero and Shubin conclude: "This is contrary to the widely-held myth about horse species as gradually-evolving, seemingly-younger forms of an unchanging species continuum, with no real distinctions between species. Throughout the history of horses, the species are well-marked and static over millions of years. At high resolution, the gradualistic picture of horse evolution becomes a complex bush of overlapping, closely related species."

**Relative frequencies.** Elegant cases don’t make punctuated equilibrium any more than a swallow makes a summer, but there are a growing number of reports documenting an overwhelming relative frequency (often an exclusivity) for punctuated equilibria in entire groups or faunas. Consider the lifetime testimonies of taxonomic experts on microfossils\(^{34}\), on brachiopods\(^{35-37}\) and on beetles\(^{38}\). Forrey\(^{38}\) has concluded for trilobites and graptolites that the gradualistic mode does occur especially in pelagic or planktonic forms, but accounts for 10% or less of observations of phylectic change, and is relatively slow\(^{39}\).

Other studies access all available lineages in entire faunas and assert the dominance of punctuated equilibrium. Stanley and Yang\(^{30}\) found no gradualism at all in the classic Tertiary molluscan sequences of the Gulf and Atlantic Coasts. With the exception of *Gryphaea*, Hallam\(^{40}\) detected no phylectic change in shape (but only for body size) in any Jurassic bi valve in Europe. Kelley\(^{41-43}\) documented the prevalence of punctuation for mol luscs in the famous Maryland Miocene sequence, and Verbeek\(^{44}\) has done the same for African bivalves. Even compilations from the literature are strongly biased by previous traditions for ignoring stasis as non-data and only documenting putative gradualism, grant a majority to punctuated equilibrium, as in Barnovskey’s\(^{45}\) compendium for Quaternary mammals, with punctuated equilibrium “supported twice as often as phylectic gradualism . . . .” Even many of the species considered exhibit most of their morphological change near a speciation event, and most species seem to be discrete entities”. When controlled studies are done by one team in the field, punctuated equilibrium almost always seems to predominate. Prothero\(^{46}\) “examined all the mammals with a reasonably complete record from the Eocene Oligocene boundary in South Dakota and related areas in Wyoming and Nebraska . . . . With one exception (gradually increasing in the orendocton *Miniochoerus*), we found that all of the Badlands mammals were static through millions of years, or speciated abruptly (if they changed at all)."

**Inductive patterns.** Even more the general inductive patterns should be explored as criteria. Stanley\(^{38-40}\) has proposed a series of tests, all carried out to punctuated equilibrium’s advantage. Others suggest that certain environments and ecologies should be conducive to one preferred mode along the continuum of possibilities. Johnson\(^{47-49}\) suggests that punctuated equilibrium should dominate in the benthic environments that yield most of our fossil record, while gradualism might prevail in pelagic realms. Sheldon\(^{49}\) proposes the counter-intuitive but not unreasonable idea that punctuated equilibrium may prevail in unstable environments, gradualism in stable regimes.

**Tests from living organisms.** Distinct evolutionary modes yield disparate patterns as results; punctuated equilibrium might therefore be tested by studying the morphological and taxonomic distributions of organisms, including living faunas. (Sev eral of Stanley’s tests\(^{38}\) use modern organisms, and other criteria from fossils should be explored—especially the biometric discordance or orthogonality, favourable to punctuated equilibrium and actually found where investigated\(^{50-53}\), of within and between species trends.)

Cladistic patterns should provide a good proving ground. Avise\(^{54}\) performed an interesting and much quoted test, favourable to gradualism, by comparing genetic and morphological differences in two fish clades of apparently equal age and markedly different speciation frequencies. But as Mayden\(^{55}\) argued, this test was wrong in its particular case, and non-optimal as a general procedure; a better method would compare cladistic sister groups, guaranteed by this status to be equal in age. Mindell et al.\(^{56-58}\) have now performed such a test on allozyme data in general, and have validated punctuated equilibrium’s key claim for positive correlation of evolutionary distance and speciation frequency. Lemen and Freeman’s\(^{59}\) interesting proposal for additional cladistic tests cannot be sustained because they must assume that unbranched arms of their cladograms truly feature no speciation events or their roots, whereas numerous transient and extinct species must populate most of these pathways. Wagner\(^{60}\) has developed a way of estimating rapidly branching speciation versus gradual speciation or transformation from cladograms, and his initial results favour predominant rapid branching in Palaeozoic gastropods.

**Difficulties and prospects.** Many semantic and terminological muddles that once impeded resolution of this debate have been clarified. Opponents now accept that punctuated equilibrium was never meant as a saltational theory, and that stasis does not signify rock-hard immobility, but fluctuation of little or no accumulated consequence, and temporal spread within the range of geographic variability among contemporary populations—by Stanley’s proper criterion, so strikingly validated in his classic study\(^{61}\). We trust that everyone now grasps the centrality of relative frequency as a key criterion (and will allow, we may hope, that enough evidence has now accumulated to make a case, if not fully prove the point).

Evolutionary biologists have also raised a number of theoretical issues from their domain of microevolution. Some, like the frequency of sibling speciation, seem to us either irrelevant or unanswerable, rather than far, our view (as we then underestimate the amount of true speciation from palaeontologically defined morphospecies, and such an underestimate works against punctuated equilibrium). Others, like the potential lack of correspondence between bi species and palaeontological morphospecies, might be worrisome, but available studies, done to assess the problem in the light of punctuated equilibrium, affirm the identity of palaeontological taxa with true bi species (see Jackson and Cheetham\(^{62}\) on bryozoan species, and Michaux\(^{63}\) on palaeontological stasis in gastropod morphospecies that persist as good genetic bio species). But, as we stress unhappily, that issue does not disturb our position, instead, focuses upon claims that speciation causes significant morphological change, for no validation of such a position has emerged (while the frequency and efficacy of our original supporting notion, Mayr’s “genetic revolution” in peripheral isolates, has been questioned). Moreover, reasonable arguments for potential change throughout the history of lineages have been advanced\(^{64}\), although the empirics of stasis throws the efficacy of such processes into doubt. The pattern of punctuated equilibrium exists (at predominant relative frequency, we would argue) and is robust. *Eprr non si muove*; but why then? For the association of morphological change with speciation remains as a major pattern in the fossil record.

We believe that the solution to this dilemma may be provided in a brilliant but neglected suggestion of Futuyma\(^{65}\). He holds that morphological change may accumulate anywhere along the geological trajectory of a species. But unless that change be “locked up” by acquisition of reproductive isolation (that is, speciation), it cannot persist or accumulate and must be washed out during the complexity of interdigitation through time among varying populations of a species. Thus, species are not special because their origin permits a unique moment for instigating change, but because they provide the only mechanism for protecting change. Futuyma writes: “In the absence of reproductive
isolation, differentiation is broken down by recombination. Given reproductive isolation, however, a species can retain its distinctive complex of characters as its spatial distribution changes along with that of its habitat or niche... Although speciation does not accelerate evolution within populations, it provides for ecological changes to the extent of being able to be registered in the fossil record. Thus, it is plausible to expect many evolutionary changes in the fossil record to be associated with speciation... By an extension of the same argument, sequences of speciation are then required for trends: "Each step has a more ephemeral existence only because reproductive isolation prevented the slippage consequent on interbreeding with other populations... Speciation may facilitate agamogenesis by retaining, stepwise, the advances made in one direction... Futuyma's simple yet profound insight may help to heal the remaining rifts and integrate punctuated equilibrium into an evolutionary theory hierarchically enriched in its light..."

In summarizing the impact of recent theories upon human concepts of nature's order, we cannot yet know whether we have witnessed a mighty gain in insight about the natural world (against anthropocentric hopes and biases that always hold us down), or just another transient blip in the history of correspondence between misperceptions of nature and prevailing social realities of war and uncertainty. Nonetheless, contemporary science has massively substituted notions of indeterminacy, historical contingency, chaos and punctuation for previous convictions about gradual, progressive, predictable determinism. These transitions have occurred in field after field: Kuhn's celebrated notion of scientific revolutions is, for example, a punctuational theory for the history of scientific ideas. Punctuated equilibrium, in this light, is only palaeontology's contribution to a Zeitgeist, and Zeitgeists, as (literally) transient ghosts of time, should never be trusted. Thus, in developing punctuated equilibrium, we have either been loadoes and pandurers to fashion, and therefore destined for history's ashheap, or we had a spark of insight about nature's constitution. Only the punctuational and unpredictable future can tell.

Stephen Jay Gould is at the Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138, USA; and Niles Eldredge is at the Department of Invertebrates, American Museum of Natural History, Central Park West at 73th Street, New York, New York 10024, USA.