



**The Disparity of the Burgess Shale Arthropod Fauna and the Limits of Cladistic Analysis: Why We Must Strive to Quantify Morphospace**

Stephen Jay Gould

*Paleobiology*, Vol. 17, No. 4. (Autumn, 1991), pp. 411-423.

Stable URL:

<http://links.jstor.org/sici?sici=0094-8373%28199123%2917%3A4%3C411%3ATDOTBS%3E2.0.CO%3B2-H>

*Paleobiology* is currently published by Paleontological Society.

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/paleo.html>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

---

The JSTOR Archive is a trusted digital repository providing for long-term preservation and access to leading academic journals and scholarly literature from around the world. The Archive is supported by libraries, scholarly societies, publishers, and foundations. It is an initiative of JSTOR, a not-for-profit organization with a mission to help the scholarly community take advantage of advances in technology. For more information regarding JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

## The disparity of the Burgess Shale arthropod fauna and the limits of cladistic analysis: why we must strive to quantify morphospace

Stephen Jay Gould

*Abstract.*—Three major arguments have been raised against the crucial claim, documented by Whittington and colleagues for the Burgess Shale fauna, and so contrary to traditional views, that disparity of anatomical design reached an early maximum in the history of multicellular life: (1) the presence of many early taxa with low membership and high rank is an artifact of naming; (2) cladistic analysis of Burgess arthropods negates the claim for greater early disparity; and (3) Whittington's argument is a retrospective fallacy based on assigning high rank to differentia only by virtue of their later capacity to define major branches. I show that all these arguments are either false or illogical, and that the claim for increased early disparity is justified: (1) Taxonomic rank is an artifact, but no one has ever based a claim for greater disparity on this false criterion. (2) Cladistics can only deal with branching order, whereas disparity is a phenetic issue. These two legitimate aspects of evolutionary "relationship" are logically distinct. The rooting of a cladogram only illustrates monophyletic ancestry (which no one doubts, as we are not creationists), and cannot measure disparity. (3) The active stabilization of the differentia of *Baupläne* (for genetic and developmental reasons only dimly understood) provides a powerful rationale for weighting these characters in considerations of disparity; nothing had so stabilized in the Burgess fauna. If these differentia were steadily changing contingencies, rather than actively stabilized features with "deep" architectural status, then the retrospective argument would be justified. Although the three arguments are wrong, the claim for greater early disparity cannot be confidently established until we develop quantitative techniques for the characterization of morphospace and its differential filling through time. This is a dauntingly difficult problem, much harder than cladistic ordering, but not intractable.

Stephen Jay Gould. *Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138*

Accepted: May 7, 1991

### I. The New Iconography of the Burgess Shale

C. D. Walcott, committed to viewing the history of life as progressing upward and expanding outward (Gould 1989), interpreted the arthropods of the Burgess Shale as primitive members of five ancient lines, extending deep into the Precambrian and, with the exception of trilobites, persisting to the present day (Walcott 1912: pp. 160–161). Paleontological opinion followed this basic model of later Phanerozoic expansion from a few, primitive Cambrian precursors ("the cone of increasing diversity" of Gould 1989: p. 38) until the radical revision set in motion by Whittington's (1971) monograph on *Marrella* and still continuing today (Whittington 1985; Whittington and Briggs 1985; Whittington and Conway Morris 1985). For example, the later and canonical opinion of Størmer (1959), enshrined in the *Treatise on Invertebrate Paleontology*, lumped almost all peculiar Burgess ar-

thropods into the "wastebucket" class Trilobitoidea, invalidly defined (as we now recognize) by the symplesiomorphic retention of biramous limbs behind the head.

Whittington and colleagues, in a series of monographs extending over 20 years, have literally overturned this received wisdom and proposed a reverse geometry of maximal early disparity in anatomical design followed by extinction of most body plans. The few surviving *Baupläne* may diversify greatly (as in the unparalleled example of insects), but seem unable ever again to generate markedly different developmental architectures that would, at least in conventional taxonomies, be construed as taxa of highest rank. I have referred to this inverted cone, "Christmas tree" or "bamboo field with most ramets lopped off" as a model of "decimation and diversification" (Gould 1989: p. 46).

For the most important arthropod component of the Burgess Shale fauna, this inver-

sion of traditional views recognizes members of all four conventional divisions: Trilobita (19 species; Whittington 1975, 1977), Crustacea (*Canadaspis*; Briggs 1978), Chelicerata (*Sanctacaris*; Briggs and Collins 1988), and Uniramia (more tentatively, if onychophores are part of the uniramian line, and if *Aysheaia* is an onychophore; Whittington 1978; Robison 1985). But none of the other Burgess arthropods can be placed into these great groups of later success, and all present unique features, primarily of tagmosis, and arrangement and form of appendages, outside the subsequent range of arthropod body plans. Thirteen unique arthropod *Baupläne* have been described by Whittington and his students (see table in Gould 1989: pp. 210–211), while Whittington (1985) estimated that the addition of species not yet monographed would augment the total to 23. This remarkable reduction from more than 20 patterns in segmentation, tagmosis, and arrangement of appendages to but 4 markedly stable groups (despite the most impressive radiation of species in the entire animal kingdom) has broad implications for our general view of pattern and predictability in the history of life (Gould 1989).

I shall not here treat these implications, but wish to focus on the empirical claim that gives them substance: Does arthropod disparity (range of anatomical design), in opposition to Walcott's original argument and most received wisdom, reach a quick maximum early in the Phanerozoic and then notably decline (even while diversity, or number of species, markedly increases in the relatively few remaining *Baupläne*)? I argue in this article that Whittington's inversion is valid, that the critiques are logically misconstrued, but that the new iconography of life requires a quantitative apparatus and affirmation, not yet fully available, in order to merit general acceptance.

## II. Criticism of Whittington's Inversion

Reviews of my book (Gould 1989) by professional paleontologists and evolutionary biologists, while generally warm and positive, have presented three interrelated criticisms touching on one of the most fundamental is-

ssues in our interpretation of life's history and meaning: How can we measure and assess the direction of morphological change through time? The first of these criticisms (not the focus of anyone's serious rebuttal, but more a rhetorical point) is a linguistic fallacy, easily resolved; but the other two deserve serious consideration and debate.

1. *Taxonomic Rank*.—This false argument was advanced most succinctly by Ridley (1990: p. 11): "Gould suggests that the Burgess Shale fauna shows a relatively high diversity [disparity, in my terminology] because it contained representatives of more of the higher taxonomic groups than would a sample of the same number of organisms from the modern ocean" (see also Fortey, 1989: p. 303).

The basic argument runs: Supporters of greater Cambrian disparity base their claim upon the presence of many more taxa of higher rank (arthropod classes, for example) in the Burgess fauna than in modern oceans; since this pattern is an artifact of classificatory systems, the argument for greater disparity fails.

If the Burgess taxa were placed into a plethora of classes, and if such a classification formed the basis of our argument for greater disparity, then the critics would be right; for such a taxonomy would be an artifact, as correctly argued by Ridley (1990) and Bengtson (1990: p. 115). But no one has made such an argument; and no one would, for we all recognize the subjective and partly artificial character of higher taxonomic ranks. No one has formally placed the 23 lineages of "orphaned" Burgess arthropods into new classes, and no one, for that matter, has even established new classes for the unique and bizarre body plans that fit into no known phyla—for *Opabinia*, *Dinomischus*, and *Odontogriphus*, for example. At most, these have been listed as "Class Uncertain," surely a fair designation of current ignorance. New classes and other higher taxa have not been formally established because paleontologists are so rightly uncomfortable with higher taxa of such small membership, especially at a time when antiquity of monophyletic multicellular lineages was probably so insubstantial. (Traditional taxonomists might be comfortable with xiphosurids as a higher taxon, based on geo-

logical age, despite persistently low diversity, but would not establish *Hyracotherium* as a perissodactyl order during the Eocene. In any case, cladistic insights have rendered such taxonomic practices obsolete, making the entire issue irrelevant.) The problem has no good solution, even in a cladistic context. Geologically early and ultimately sterile monophyletic branches of major trunks must receive high rank in cladistic logic (Hennig 1966), even if they contain only one species, and even if (unlike many of the Burgess taxa, with their oddly disparate morphologies) they look scarcely different from contemporaries in lineages with significant patrimony. Perhaps the tactic of simply designating such taxa as plesions, and leaving them otherwise unranked, is the best solution to a logical dilemma. I, for one, support this solution.

In short, Burgess disparity presents a puzzle, not because systematists place the creatures into a large number of higher taxa, but because the animals seem, entirely *in se* and independent of their names, so damned curious, different one from the other and dissimilar from surviving lineages. At most, an author might say something like: "*Leancoilia* differs as much from any modern arthropod, as the living classes do among themselves." This may be an unfortunate shorthand, but such statements should not be misread as arguments for basing claims of enhanced disparity upon the rank of taxonomic names. In fact, our propensity to use such shorthand underscores the main point of this article (see section V): that we need a quantitative approach to the differential filling of morphospace (and that, in the absence of such procedures, we resort to linguistic analogies that may be misunderstood).

2. *The Retrospective Fallacy*.—Modern arthropod classes are defined by patterns of segmentation, tagmosis, and arrangement of appendages, features that seem architecturally "basic" in a metameric organization with, primitively, one pair of biramous appendages per segment. Uniramia have lost the gill branches of originally biramous limbs and build their body appendages from leg branches only. Chelicerata have six uniramous leg branches on the prosoma, and uni-

ramous appendages with gill branches on the opisthosoma. Crustacea, for all their overt ecological diversity, have (at some stage of ontogeny at least) five pairs of appendages on the head, two uniramous preoral (antenna and antennule), and three postoral, usually used in feeding. Trilobites have, on the cephalon, one uniramous preoral pair (antennae) and three postoral pairs of appendages.

These fundamental features are remarkably stable through the Phanerozoic, and arthropods do not generate new basic patterns following the Cambrian explosion, as captured in Burgess Shale disparity. But most Burgess arthropods show a different arrangement of these defining features. *Marrella* (Whittington 1971) has two pairs of preoral antennae, and none postoral (though Walcott thought he had found the three postoral pairs and ranked *Marrella* as a trilobite), *Sidneyia* (Bruton 1981) has but one preoral pair of antennae (though Walcott considered this animal a merostome). *Yohoia* has three pairs of walking legs at the posterior part of the head shield and a large feeding organ in front, so peculiar and nonhomologous with others that Whittington simply dubbed it the "great appendage" (Whittington 1974). *Leancoilia* (Bruton and Whittington 1983) has two biramous pairs behind the mouth and the most bizarre of all appendages in front—a giant forward projection with a right-angle bend, composed of four stout basal segments with three long whiplike extensions; the entire "hinged" structure could be bent back against the body, presumably for streamlining in swimming. The list goes on and on.

By modern standards of defining arthropod classes, these Burgess taxa therefore possess differentia that would now make distinctions at the highest taxonomic level within the phylum. But does this mean that, in some absolute sense, Burgess disparity greatly exceeds modern limits? To say so on this account alone would be to commit a "retrospective fallacy" of the worst sort. Tagmosis and pattern of appendages may be characters that define classes today, but a Cambrian paleontologist would not have known this later history. He would have peered into the Burgess sea, found a bunch of arthropod species,

each with differing numbers and arrangements of segments and appendages, and concluded that such distinctions are superficial, easily made, and worthy only of defining, say, genera within families. Thus, Burgess arthropods were not so disparate in their own terms, but only by the invalid criterion of what happened to separate large groups later on.

This has been the major and recurring criticism of my claim for greater initial disparity. Ridley (1990: p. 12), for example, wrote:

If we did not know that two pairs of antennae was destined to become a highly significant character, we should not attach any great importance to the difference between the arthropods with two pairs and those with other numbers. . . . Their status as a higher taxon in the Burgess Shale stands only on their future importance.

In section IV, I present my rebuttal of the charge that claims for greater early disparity rest upon a retrospective fallacy. Moreover, I did not (as several reviews have implied) ignore the argument in my book, where I wrote (albeit in a footnote and for later confutation) (Gould 1989: p. 209):

If I wished to play devil's advocate against my own framework, I would argue that the criterion by which we make the claim of twenty losers and only four winners is falsely retrospective. By patterns of tagmosis, modern arthropods are surely strikingly less disparate than Burgess forebears. But why use patterns of tagmosis as a basis for higher-level classification of arthropods? A nearly microscopic ostracode, a terrestrial isopod, a planktonic copepod, a Maine lobster, and a Japanese king crab span more variety in size and ecological specialization than all the Burgess arthropods put together—though all these modern creatures are called Crustacea, and display the stereotyped tagmosis of this class. A paleontologist living during the Burgess might consider the arthropods as less varied because he had no reason to regard patterns of tagmosis as a particularly important character (for the utility of tagmosis in

distinguishing major genealogical lines only became apparent later, after most alternatives were decimated and stereotypy set in among the few surviving and highly disparate lines).

3. *Cladistic Analysis*.—Briggs and Fortey (1989) performed a cladistic analysis on 23 Cambrian arthropod species, including 19 from the Burgess Shale, plus 5 additional taxa (the well-preserved Ordovician trilobite *Triarthrus*, the eurypterid *Baltoeurypterus*, modern *Limulus*, and the two living but highly plesiomorphic crustaceans *Hutchinsoniella* and *Speleonectes*). Although the result included only a few small clades and many plesions (not a criticism, for this pattern expresses a reasonable genealogical reality for early radiations), the cladogram did converge to a root (with *Marrella* as sister to all other taxa), and did present the surprising arrangement of crustaceans at the primitive and trilobites at the most derived position.

Briggs and Fortey performed this study primarily to test (and, as it happened, to refute) the theory of arthropod polyphyly (supported by Whittington, among others), a question clearly suited for this form of analysis. But they also, if subsidiarily, used their cladogram to question the claim for greater early disparity. They argued, for example, on the basis of their cladistic ordering (1989: p. 241): "The arthropods that evolved during the Cambrian radiation show no more apparent morphological diversity than do the living groups."

I show in the next section that such a claim is both unsupported in its own terms and also illogical a priori. Nonetheless, this argument has been picked up and adumbrated by several reviewers of Gould (1989). Two of my otherwise favorite reviews repeat this error. Citing Briggs and Fortey's cladogram, Bard (1990: p. 305) wrote: "The body plans of the animals that survived thus show an artificial disparity which does not reflect their early evolution." While Rolfe (1990: p. 8) stated that the cladogram "showed the fossil forms to be less disparate than Whittington and Gould thought. To some extent therefore the disparity of these arthropods is an artifact of

their authors—the fossils can be grouped together into larger units.”

### III. Logical and Empirical Problems of the Burgess Cladogram

Most arguments presented by scientists eventually turn out to be wrong; all intellectual life presents this inherent danger. But the source of scientific error is usually empirical: new information refutes an old claim. Errors in logic, on the other hand, are much rarer.

Yet the argument that “greater early disparity is wrong because the cladogram of Burgess arthropods roots” is a logical error based on a conflation of categories. The logic of any mode of analysis prescribes limits to its field of operation. Cladistics is a powerful tool that has generated vital reforms in taxonomic practice, particularly by giving us the precious gift of an objective method for studying branching order in genealogy. But temporal branching order and morphological disparity are separate issues (empirically related only by the high probability of a rough correlation between cladistic and phenetic distance). Branching order is the cladistic component of “relationship” among taxa; disparity is a phenetic component of the same ambiguous term. And if we have learned anything in the past 25 years of intense taxonomic debate, we surely now understand that phenetics and cladistics give different results, and that “evolutionary systematics,” in trying to marry these two components of “relationship” into a consensus, could not succeed in principle. (If branching order were the same as disparity, then phenograms and cladograms would be identical and evolutionary taxonomy would work. But this identity can only occur if morphological similarity correlates perfectly with recency of common ancestry, an unrealized nirvana that would make our work much easier and life on this planet much less interesting.)

Thus, when Rolfe (1990) said that high disparity is an artifact because “the fossils can be grouped together into larger units,” and when Briggs and Fortey (1989) read “no more apparent morphological diversity” for Cambrian arthropods as the message of their

cladogram, they committed a logical error. The rooting of the cladogram only points to common ancestry within a monophyletic Arthropoda and does therefore argue against the polyphyletic theory. (This is a genealogical debate properly addressed by cladistics.) But this rooting simply does not speak to the phenetic issue of broader or diminished disparity. What, in phenetic terms, are we supposed to learn from the fact of rooting? Did any proponent of increased disparity ever doubt that a cladogram would root, if not in the Arthropoda at least at a more inclusive level? We are not, after all, creationists; and we do accept a monophyletic origin for life!

Consider Briggs and Fortey’s own method of eliminating autapomorphies in constructing their cladogram. This procedure is, obviously, correct for their goal of establishing branching order, for a unique character cannot in principle establish relative times of joining with other taxa. But autapomorphies lie at the heart of arguments about disparity. (Contemplate what the autapomorphy of human consciousness as expressed by language has done to this planet, despite our close cladistic affinity with chimps and gorillas). How can anyone make a claim about Burgess versus modern disparity with a chart that has eliminated all the unique tagmoses of the Burgess taxa and does not acknowledge the spiny carapace of *Marrella*, the three-pronged tail fluke of *Odaraia*, and the remarkable (and different) frontal appendages of *Leancoilia* and *Yohoia*?

Moreover, Briggs and Fortey’s cladogram does not inspire full confidence in its own terms, a fascinating expression of life’s actual history, not a failure of cladistic methodology or of Briggs and Fortey’s application. Its “rather low” (Briggs and Fortey 1989: p. 242) consistency index of 0.384 (for 28 taxa and 46 characters) reflects the unsatisfactory state of most characters as ripe for reversal and homoplasy. Most of the 46 features are defined as losses or reductions of parts or segments (“loss of labrum,” “loss of doublure,” “loss of cephalic gnathobases”), changes notoriously subject to homoplasy. Several other characters also seem architecturally or developmentally “simple” enough to occur over and over

again in independent lineages: "trilobation," "presence of unfurrowed plurae," "increase in trunk somites from 11 to 12-14."

In this light, the cladogram presents many joinings that do not make sense in the light of more detailed knowledge, but that arise from problems of characters so subject to reversal and homoplasy. To choose just two examples involving the two crucial taxa thought to represent first appearances of modern classes, the crustacean *Canadaspis* (Briggs 1978) and the chelicerate *Sanctacaris* (Briggs and Collins 1988). *Limulus* and *Baltoerypterus* properly form a miniclade, but the chelicerate *Sanctacaris* is far removed by six nodes, and the intervening genera (including *Sarotrocercus*, *Yohioia*, and *Leancoilia* among others) are decidedly nonchelicerate in anatomy. As a striking example of errors introduced by homoplasy, *Sarotrocercus* and *Yohioia* form a miniclade depicted as a sister group to the miniclade of *Limulus* and *Baltoerypterus*. The supposed synapomorphy defining their uniting node reads: "uniramous trunk appendages, reduced or absent inner ramus." Now modern chelicerates do possess gill branches only on appendages of the opisthosoma (book lungs of spiders and gill books of horseshoe crabs, for example). But *Sanctacaris*, as a basal chelicerate, retained biramous appendages on the body. Presumably, the reduction to uniramous gill branches came later; such losses are legion within all monophyletic groups. Yet, *Sanctacaris* is falsely distanced from modern chelicerates on Briggs and Fortey's cladogram because several other taxa independently eliminated walking legs on their body segments for reasons entirely divorced from genealogical position near the chelicerate line. *Sarotrocercus*, for example, is a tiny arthropod that swam on its back (Whittington 1981); reduction of body appendages to gill branches, in this case, is probably an adaptation for swimming. Thus, several nonchelicerate lineages that lost (or just reduced) leg branches for other reasons are falsely sandwiched between modern chelicerates and *Sanctacaris*. But what, in the evolution of arthropods, is more common than multiple, independent reduction of biramous limbs to uniramous state (consider the largest clade Uniramia, with its

elimination of gill branches on body segments). Similarly, the basal crustacean *Canadaspis* stands falsely far from *Hutchinsoniella* because another eminently convergeable character—"shield bivalved," in this case—characterizes both *Canadaspis* and several noncrustacean lineages, while *Hutchinsoniella* does not have a bivalved carapace.

Again, I am not criticizing Briggs and Fortey, but merely making a comment about the actual history of life. The two key properties that present cladistic problems—few clades with a plethora of plesions, and multiple reversals and homoplasy in labile characters—reflect a stunning and crucial reality of life's greatest radiation in the Cambrian explosion. This lability and interchangeability of characters that would later stabilize as differentia of higher taxa records the very phenomenon supporting my claim for higher early disparity, as the next section will argue.

#### IV. Contemporary Fact versus Retrospective Fallacy

The resolution of the retrospective fallacy is uncomplicated once we turn to the right issue, that is, away from branching order and toward developmental architecture. Applied only to the Cambrian, the retrospective argument is correct: we cannot designate tagmosis patterns as intrinsic differentia of large groups; a Cambrian paleontologist might have viewed such features as trivial markers of minor distance. But the point is that *all* characters are similarly labile in the Cambrian (hence, the cladistic difficulties discussed in the last section). Nothing defined stable clades back then because nothing had stabilized in this great era of experimentation and lability. Beyond the trilobites, and the probable affinity of *Actaeus* and *Alalcomenaeus* (Whittington 1981), the Burgess arthropods do not form subgroups based on shared derived characters, hence so many plesions in Briggs and Fortey's cladogram. This is a biological phenomenon of great import (and current mystery), not merely a formal difficulty for taxonomists.

If such lability had persisted throughout time, then the retrospective fallacy would be entirely justified. For the modern gaps be-

tween groups would only record the extinction of intermediate forms; and the differentia of higher taxa would emerge passively from this weeding, thus contributing no active component to their persistence. These differentia, in other words, would be contingent criteria of the moment, subject to annulment and change at any time. But the differentia are actively stable. They form *Baupläne* that, for whatever reason, developed a remarkable constancy, in particular, an apparent power to prevent any subsequent evolution of additional basic architectures.

In other words, *the key to disparity is not Cambrian lability, but later and active stabilization.* The tagmosis patterns congealed and served, forever after, as foci of genealogical groups, great monophyletic trunks of the arthropod tree. Other character sets did not so stabilize. Four complexes (all represented in the Burgess Shale) congealed; all others disappeared. For 500 million years, despite the most impressive radiation of species in the entire history of animals, these complexes have been stable, and actively so (we must assume), with one further loss in extinction of trilobites. This stability is not an artifact of naming; it is, in many ways, the most astonishing fact in natural history. It begs for explanation, if we are to have any hope of unraveling the history of morphology. The best candidates for good ideas at the moment (though we are all horribly plesiomorphic with respect to final solutions) lie in developmental concepts of locking pathways (or perhaps in genetic notions of evolution toward systems less capable of substantial change), as in Riedl's idea of "burden" (1978), or Wimsatt's of "generative entrenchment" (Schank and Wimsatt 1986; Wimsatt and Schank 1988). But we need data and experiments to get beyond these metaphors.

This cardinal point has been best expressed by Jonathan Bard, the only developmental biologist who wrote a major commentary on my book. Bard both acknowledged the geometry of greater early disparity and located the rationale in subsequent "turning off" of a capacity to generate fundamentally new body architectures. Bard wrote (1990: pp. 304–306):

Clearly the arthropods were a far wider phylum then than now. . . . By the time of these [Burgess] animals, the problems of pattern-formation, differentiation and morphogenesis had all been solved and a wide range of solutions found, some of which have since been lost. There is thus a strong case for saying that all developmental and physiological processes evolved in a relatively short time and that, since then, changes have merely been rung on an extant repertoire. How could the problems of development have been solved so fast and why has subsequent evolution been so limited? . . . Why has evolution varied extant themes rather than created new melodies? . . . Perhaps the answer is that, over long periods of evolution, the simplicity of the original pattern-formation process has been lost, possibly because of mutation-generated intragenomic feed-back mechanisms which stabilized existing form-generating mechanisms and so inhibited their ability to form new types of patterns.

Evolutionists who do not grasp Bard's point, and who think that greater early disparity is a retrospective fallacy, usually accept Darwin's central argument (1859) that later differences among groups in genealogical trees are caused simply by the extinction of intermediate forms and not by active stabilization. Ridley (1990: p. 12), for example, wrote: "Today, two pairs of antennae just happens to be a reliable indicator of a major branch in the tree of life. This is a mere evolutionary fluke: in a million years time, a crustacean may evolve with one, or three, pairs of antennae, and the successful Crustacea of the future may be descended from a three-antennaed, rather than a two-antennaed shrimp." In this statement, Ridley only shows his lack of familiarity with the fossil record, and his continuing commitment to gradualism. Two pairs may be contingent in the sense that a rerun of life's tape would have yielded different results, but two pairs is not a transitory fluke, arising capriciously in the recent past and subject to annulment in the near future. Two pairs of antennae characterized Burgess crustaceans (*Canadaspis*), and have continued to

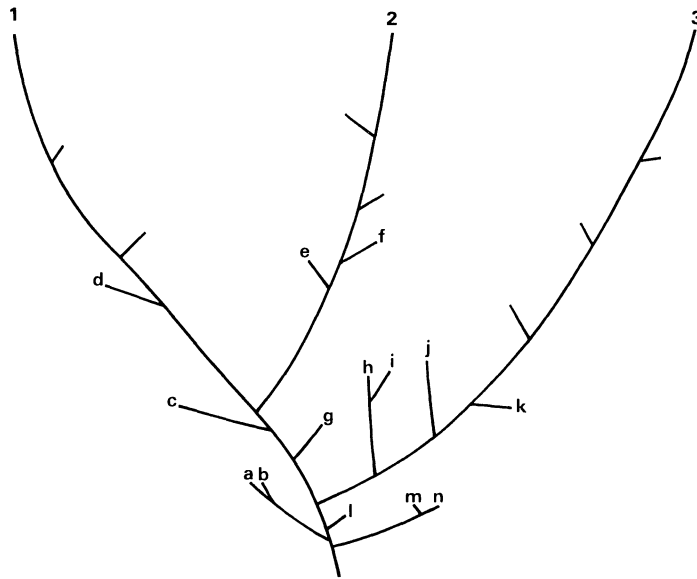


FIGURE 1. Continual increase in disparity under conventional "cone" model; short-lived lineages arising outside the morphological bounds of long-lived branches do not surpass modern disparity because surviving branches continue to diverge through time. Vertical axis is time; horizontal is morphological distance.

mark crustacean affinity ever since, despite the radiation of the group into forms as disparate as barnacles and lobsters (see Cisne's important paper [1974] on retained flexibility within the fixed designs). I think it fair to assume that this and other features of the crustacean developmental *Bauplan* represent active stabilizations. The fact that tagmosis was so labile in Burgess times, and that only four patterns both survived and congealed, points to marked reduction in disparity for characters of special interest by virtue of their power to constitute active archetypes.

Although active stabilization is a proper criterion for investing claims of disparity in characters that define *Baupläne*, two additional arguments for greater early disparity should be mentioned.

1. *The Architectural Depth of Characters.*—Suppose that we are building freight trains from numerous cars, each with two sets of wheels. We construct hundreds of trains, not haphazardly, but following strictly limiting rules of order. Cars must be arrayed in a certain sequence, but the total number can change either by fusion of adjacent cars or by duplication of particular cars; numbers of wheels per car may also augment or decrease. We look at the resulting collection and try to

order the variation in terms of fundamental and superficial properties.

We would all, I think, share an intuition that fusions, duplications, and wheel arrangements might be deemed more fundamental than changes in paint or exterior paneling, because these "deeper" characters vary the basic rules of modular building, whereas paint and paneling are literally superficial and easy to alter piecemeal. Similarly, since arthropods are modular animals with a pair of appendages on each original segment, we might view changes in segment numbers, and alteration in basic form and number of appendages, as architecturally more fundamental—a priori and before any observation of what stabilizes in geological time—than, say, color of the carapace or use of a gill branch to swim or filter. Since the characters that do stabilize to form *Baupläne* have this apparent architectural depth, they become especially appropriate for arguments about disparity. I do not regard this point as retrospectively fallacious; we would, I think, be intensely puzzled if a small orange spot on the elytra formed the only morphological synapomorphy of Coleoptera (while molecular evidence clearly asserted their monophyletic nature).

I recognize that this argument is meta-

phorical and more than a bit fuzzy. Such verbal difficulties and subjectivities plague the study of morphological pattern through time, and evoke my plea for a quantitative apparatus, the primary point of this paper (see section V).

2. "Insiders" and "Outsiders" as a Logical Issue.—One argument against greater early disparity makes an intriguing point, but ultimately fails by a logical gap. One might claim that an impression of greater early disparity only confuses the necessary oddness of early forms (versus modern representatives) with a truly expanded range of *Bauplane* characters. As we go back in time, we will encounter the intermediates that evolution eliminated in forming the gaps between modern higher taxa; these creatures will look peculiar, if only in their amalgams of characters later sorted into separate lines. But they are not more disparate, only somewhat discomfoting in oddness. In fact, if the morphospace of life expands in conformity with the conventional cone of increasing diversity, then disparity should now be maximal, even if fewer main lines survive (see Fig. 1).

But if—as I claim for the Burgess and later history of arthropods, and for life in general—*Bauplane* congeal and stabilize, with most eventually decimated and few surviving, then the argument that early oddballs are intermediates, and that weeding has not limited the total range of morphospace (but only thinned it from within), almost surely cannot hold. Let all early lineages within the bounds of morphospace set by surviving *Bauplane* be termed "insiders," with all beyond the current borders called "outsiders" (see Fig. 2). Disparity will be constant or increase in time only if all peripheries of the original distribution yield survivors, a most unlikely proposition given the tenuous nature of edges. Almost surely, some peripheries will be abandoned, leaving numerous outsiders among early forms, and leading to an ultimate and marked reduction of early disparity.

Consider the Burgess arthropods in particular. All four basic architectures of later Phanerozoic higher taxa have Burgess representatives. Is it plausible to think that the other 23 Burgess designs are all insiders within the space marked by these four survivors?

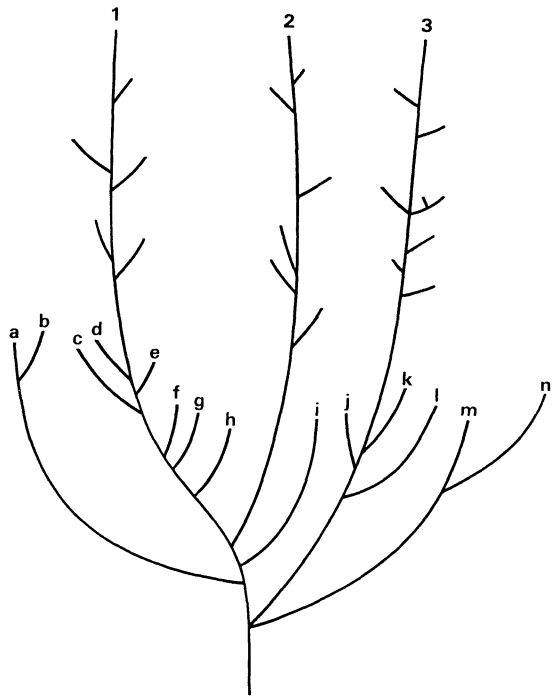


FIGURE 2. Reduction of early disparity under the argument that differentia of *Bauplane* actively stabilize. Surviving lineages show less disparity than the early era of substantial branching. It is most improbable that the few survivors will occupy all the morphological peripheries. Therefore, many early lineages will be "outsiders" (a-d and k-n) with respect to modern survivors, leaving e-j as "insiders." (In Fig. 1, by contrast, all early lineages are inside the peripheries set by continual expansion of disparity through time.)

And even if one is tempted to make such a claim for arthropods, what about the numerous Burgess taxa so peculiar that they cannot be placed in any modern phylum? Is it plausible that all these forms, including morphologies that seem more appropriate for a science-fiction set than an earthly sea (*Opabinia* with five eyes and a vacuum-cleaner hose for a frontal projection), are insiders within the space of modern phyla?

## V. Quantifying Morphospace and Its Differential Filling

I have argued that the critics of greater early disparity are wrong in their central claims about cladistics and retrospective fallacies. But these critics, while choosing false targets, are motivated by a serious and entirely legitimate malaise. The problem, however, is caused by absent technique, not incorrect argument or even inherent ambiguity.

The vague concept of evolutionary "relationship" includes both cladistic (branching order) and phenetic (morphological distance) components; and these are not the same, either logically or empirically. A major triumph of evolutionary studies in our generation has been, through the development of cladistics, the codification of a methodology for objective definition and determination of branching order. (The soundness of the logic does not guarantee a resolution in every, or even in most, cases; for our world of rampant homoplasy and imperfect data often frustrates hope for confident answers.)

Success often breeds both overconfidence and overextension. Many evolutionists, intoxicated with the victory of cladistics in its own sphere, have assumed that it must represent a panacea for all questions about "relationships." (Or, if not quite so imperialistic, proponents of cladistics might just claim that other questions, if not resolvable by branching order, are therefore neither interesting nor important.) Thus we note, as in the debate about Burgess disparity, a false extension of cladistic relationships to questions about the differential filling of morphospace.

The phenetic question of defining a morphospace and mapping the distribution of actual organisms is much more difficult than the cladistic problem of determining branching order. Branching order unfolds in Newtonian time and in a space even simpler than Euclid's—the two-dimensional world of Abbott's *Flatland*. A morphospace, on the other hand, is not only richly multidimensional but seemingly idiosyncratic for each kind of organism. How can it be defined with anything like the rigor of cladistic space and time?

In the light of this frustrating difference between rigor for branching order and verbal vagueness for morphospace, who can blame critics for their lack of sympathy with the phenetic question? Bengtson (1990: p. 115) referred to the notion of body plans as "a seedy concept if there ever was any"; while Fortey (1989: p. 303) wrote: "The problem with 'disparity' is that its estimation depends on the authority of the expert: how is he to know what makes—what is 'worth'—a phylum? Or what a class?"

We cannot assuage this frustration by importing a logically inappropriate technique like cladistics. The resolution can only arise within the structure of phenetic methods. Either we develop a set of quantitative techniques for the definition and differential filling of morphospace, or we are condemned to vagueness.

What, then, do we need? Not simply a good method for the multivariate description of organisms, for such we have (Sneath and Sokal 1973; Bookstein 1977a,b). And not even a proper multivariate description of morphological transformation—whether by D'Arcy Thompsonian (1917) coordinate transformations, Huxleyan (1932) allometric growth gradients, or more modern methods like trend surface analysis (Sneath 1967). We need, instead, to define a full range of the abstract (and richly multivariate) space into which all organisms may fit (the morphospace). We must then be able to characterize individual organisms and plot them within this encompassing space. Finally, we need to measure density, range, clumping, and a host of other properties that determine differential filling of this totality; and we must be able to assess the variation in this differential filling through time. (The claim for greater initial disparity is, effectively, a statement that the Cambrian range exceeds modern occupation.)

These questions are dauntingly difficult, and I do not pretend to have a solution. Consider just a random sample of problems: If we define the morphospace empirically by a range of organisms, how do we plot other creatures not of the original set and outside its bounds (by dimensionality, not just extrapolation of distance)? How do we deal, theoretically, with the fact that potential measurements are infinite? How do we treat correlation, covariation, the simultaneous consideration of state, meristic, and metric variables? How do we express quantitatively the vague but vital concept of relative importance or developmental depth, when superficial characters may display vast metrical differences among species, while small changes in deeper characters may trigger much more fundamental divisions? (This last is akin, of course, to the old issue of character

weighting in phenetics; such difficulties led the original numerical taxonomists to advocate equal weighting of characters, and this patently unbiological procedure helped to seal their defeat. Perhaps we could develop some hierarchical or "flow chart" or treelike picture of development and the cascading consequences of changes, and then measure sequelae rather than simple metric differences.)

Questions of this type underlie almost all interesting issues in the macroevolutionary history of morphology. Think how much ink has been spilled in analyzing and arguing Simpson's famous claim (1953), and imperfect demonstration (as he admitted himself) through the inadequate surrogate of taxonomic ranks and numbers, of faster evolution in mammals than in clams? So many vital questions in evolution run up against these problems of differential filling of morphospace and relative rates of change among groups. As I write this article, this week's *Nature* (Eckhardt 1991: p. 112) reports the discovery of less mitochondrial difference among Lake Victoria cichlids than among human populations, but mentions the apparent anomaly of greater age for the fish radiation than for the presumed African ancestor of all modern humans. The anomaly arises from our perception that morphological differences among the fishes are greater than among humans. So are the dates wrong? Or does fish morphology evolve faster? Or is the apparently greater difference among fishes an artifact of large change in superficial characters? Allan Wilson was so frustrated by such questions as he published his pioneering work on "genomic clocks" in the calibration of evolutionary pattern that he devoted considerable attention to quantifying morphospace for purposes of comparison with genetic distances (Cherry et al. 1979; Wyles et al. 1983; Wilson et al. 1984). His attempt may have been crude and preliminary, but his questions were acute, and his work should be furthered.

I do confess some fears that, *in toto*, the question of morphospace may be logically intractable, not merely difficult. Consider the two obvious impediments: (1) All successful methods of comparison (from transformed

coordinates by Albrecht Dürer to modern factor analysis) work with homologous points. How can differences be quantified if such points of comparison do not exist. Is a coral more different from an oak tree than a cow? What is the distance between a rock and a hard place? Between Scylla and Charybdis? (2) How can an objective morphospace be determined if organisms have infinite numbers of potential characters? We are not, after all, simply placing creatures into an exterior and objective Euclidian world. We are defining their morphospace in terms of their characters.

Such issues may make a general solution intractable. We may not be able to answer, with satisfying rigor, the question of whether mammals evolve faster than clams: for how can teeth be compared with pallial sinuses? But science usually proceeds by resolving smaller puzzles and then moving on toward more general formulations. We should be able to establish adequate morphospaces for creatures with comparable body plans and joint possession of sufficient homologies. The issue of humans and cichlid fishes should be resolvable. In this light, whatever the general problems throughout the animal kingdom, our specific issue of disparity among Burgess arthropods should fall into the domain of tractable questions. We should be able to describe an arthropod morphospace and to learn whether Burgess animals inhabited a wider portion than do modern forms.

Some satisfying preliminary studies in differential filling of morphospace have been made. Raup (1965, 1966, 1967, 1968) attracted much attention and inspired much fruitful work by defining a simple morphospace for coiled shells (based admittedly on the unrealistic premises of isometric growth and construing the shape of the generating curve as a circle) and then plotting the domains and differential fillings of various molluscan taxa. He posed and discussed with great insight the alternative explanations for clumping, patchiness, and empty portions: adaptation, developmental constraint, and historical drift (availability for occupation, but no entrance as yet). I defined the morphospace of spire shape for *Cerion* and found inhomogeneities

in occupation by taxa of different sizes (Gould 1984). In a most promising study, using more complex multivariate techniques and therefore a greater range and number of characters, Foote (1988, in press) defined the morphospace of trilobite cranidia and successfully resolved an old question—heretofore only discussed in the unsatisfactory and subjective terminology of “intuition” by taxonomic experts—a question squarely in the tradition of change in differential filling of morphospace through time. Foote found that Cambrian trilobites do occupy a smaller area more homogeneously, whereas Ordovician forms develop greater total range and larger gaps between groups. The issue of Burgess arthropods should be resolvable by similar techniques extended to a wider domain of homologous body parts.

I have written this paper because I believe that the question of defining morphospaces and mapping their differential filling through time is so vital to our understanding of life's history, particularly to the potential contribution of paleontologists. Yet relatively little has been done in this area, despite promising starts, and despite a near assurance that smaller subdivisions of the problem are tractable (for creatures with sufficient homology), whatever the status of the full generality. I believe that a serious attack on this problem would be well worth the concentrated attention of paleontologists with quantitative skills and evolutionary interests. Usually, when we are stymied in science, our impediment lies with missing data, thus breeding frustration; for, in such an empirical field, supply of absent data often involves an inevitable component of waiting and hoping. But, in the case of morphospace and its differential filling, we face the much happier and resolvable issue of abundant data waiting for a technique.

### Literature Cited

- BARD, J. 1990. The fifth day of creation. *Bioessays* 12:303-306.
- BENGTSON, S. 1990. Wonderful life, lovely story. *Lethaia* 23:115-116.
- BOOKSTEIN, F. L. 1977a. Orthogenesis of the hominids: an exploration using biorthogonal grids. *Science* 197:901-904.
- BOOKSTEIN, F. L. 1977b. The study of shape transformation after D'Arcy Thompson. *Mathematical Biosciences* 34:177-219.
- BRIGGS, D.E.G. 1978. The morphology, mode of life, and affinities of *Canadaspis perfecta* (Crustacean: Phyllocarida), Middle Cambrian, Burgess Shale, British Columbia. *Philosophical Transactions of the Royal Society, London B* 281:439-487.
- BRIGGS, D.E.G., AND D. COLLINS. 1988. A Middle Cambrian chelicerate from Mount Stephen, British Columbia. *Palaeontology* 31:779-798.
- BRIGGS, D.E.G., AND R. A. FORTEY. 1989. The early radiation and relationships of the major arthropod groups. *Science* 246:241-243.
- BRUTON, D. L. 1981. The arthropod *Sidneyia inexpectans*, Middle Cambrian, Burgess Shale, British Columbia. *Philosophical Transactions of the Royal Society, London B* 295:619-656.
- BRUTON, D. L., AND H. B. WHITTINGTON. 1983. *Emeraldella* and *Leanchoilia*, two arthropods from the Burgess Shale, British Columbia. *Philosophical Transactions of the Royal Society, London B* 300:553-585.
- CHERRY, L. M., S. M. CASE, J. G. KUNKEL, AND A. C. WILSON. 1979. Comparisons of frogs, humans, and chimpanzees. *Science* 204:435.
- CISNE, J. L. 1974. Evolution of the world fauna of aquatic free-living arthropods. *Evolution* 28:337-366.
- DARWIN, C. 1859. *On the Origin of Species*. John Murray; London.
- ECKHARDT, R. B. 1991. Fish scales for human origins. *Nature* 349:112.
- FOOTE, M. 1988. Survivorship analysis of Cambrian and Ordovician trilobites. *Paleobiology* 14:258-271.
- FORTEY, R. A. 1989. The collection connection. *Nature* 342:303.
- GOULD, S. J. 1984. Morphological channeling by structural constraint: convergence in styles of dwarfing and gigantism in *Cerion*, with a description of two new fossil species and a report on the discovery of the largest *Cerion*. *Paleobiology* 10:172-194.
- GOULD, S. J. 1989. *Wonderful Life*. W. W. Norton and Company; New York.
- HENNIG, W. 1966. *Phylogenetic Systematics*. University of Illinois Press; Urbana, Illinois.
- HUXLEY, J. S. 1932. *Problems of Relative Growth*. MacVeagh; London.
- RAUP, D. M. 1965. Theoretical morphology of the coiled shell. *Science* 147:1294-1295.
- RAUP, D. M. 1966. Geometric analysis of shell coiling: general problems. *Journal of Paleontology* 40:1178-1190.
- RAUP, D. M. 1967. Geometric analysis of shell coiling: coiling in ammonoids. *Journal of Paleontology* 41:43-65.
- RAUP, D. M. 1968. Theoretical morphology of echinoid growth. *Journal of Paleontology* 42:50-63.
- RIDLEY, M. 1990. Dreadful beasts. *The London Review of Books*, June 28, pp. 11-12.
- RIEDL, R. 1978. *Order in Living Organisms*. John Wiley and Sons; New York.
- ROBISON, R. A. 1985. Affinities of *Aysheaia* (Onychophora) with description of a new Cambrian species. *Journal of Paleontology* 59:226-235.
- ROLFE, W.D.I. 1990. *Life: a loser's manual*. The Scotsman, February 10, p. 8.
- SCHANK, J. C., AND W. C. WIMSATT. 1986. Generative entrenchment and evolution. *Philosophy of Science Association* 2:33-60.
- SIMPSON, G. G. 1953. *The Major Features of Evolution*. Columbia University Press; New York.
- SNEATH, P.H.A. 1967. Trend-surface analysis of transformation grids. *Proceedings of the Zoological Society of London* 151:65-122.
- SNEATH, P.H.A., AND R. R. SOKAL. 1973. *Numerical taxonomy*. W. H. Freeman and Company; San Francisco.
- STORMER, L. 1959. Trilobitoidea. Pp. 23-37. *In* R. C. Moore (ed.),

- Treatise on Invertebrate Paleontology, Arthropoda I. Geological Society of America; New York.
- THOMPSON, D. W. 1917. On Growth and Form. Macmillan; London.
- WALCOTT, C. D. 1912. Middle Cambrian Branchiopoda, Malacostraca, Trilobita and Merostomata. Cambrian Geology and Paleontology, II. Smithsonian Miscellaneous Collections 57: 145-228.
- WHITTINGTON, H. B. 1971. Redescription of *Marrella splendens* (Trilobitoidea) from the Burgess Shale, Middle Cambrian, British Columbia. Geological Survey of Canada Bulletin 209:1-24.
- WHITTINGTON, H. B. 1974. *Yohoia* Walcott and *Plenocaris* n. gen., arthropods from the Burgess Shale, Middle Cambrian, British Columbia. Geological Survey of Canada Bulletin 231:1-21.
- WHITTINGTON, H. B. 1975. Trilobites with appendages from the Middle Cambrian, Burgess Shale, British Columbia. Fossils and Strata (Oslo) 4:97-136.
- WHITTINGTON, H. B. 1977. The Middle Cambrian trilobite *Naraoia*, Burgess Shale, British Columbia. Philosophical Transactions of the Royal Society, London B 280:409-443.
- WHITTINGTON, H. B. 1978. The lobopod animal *Aysheaia pedunculata* Walcott, Middle Cambrian, Burgess Shale, British Columbia. Philosophical Transactions of the Royal Society, London B 284:165-197.
- WHITTINGTON, H. B. 1981. Rare arthropods from the Burgess Shale, Middle Cambrian, British Columbia. Philosophical Transactions of the Royal Society, London B 292:329-357.
- WHITTINGTON, H. B. 1985. The Burgess Shale. Yale University Press; New Haven.
- WHITTINGTON, H. B., AND D.E.G. BRIGGS. 1985. The largest Cambrian animal, *Anomalocaris*, Burgess Shale, British Columbia. Philosophical Transactions of the Royal Society, London B 309: 569-609.
- WHITTINGTON, H. B., AND S. CONWAY MORRIS. 1985. Extraordinary fossil biotas: their ecological and evolutionary significance. London: Royal Society. Published originally in Philosophical Transactions of the Royal Society, London B 311:1-192.
- WILSON, A. C., J. G. KUNKEL, AND S. J. WYLES. 1984. Morphological distance: an encounter between two perspectives in evolutionary biology. *Evolution* 38:1156-1159.
- WIMSATT, W. C., AND J. C. SCHANK. 1988. Two constraints on the evolution of complex adaptations and the means for their avoidance. Pp. 231-273. In Nitecki, M. H. (ed.), *Evolutionary Progress*. University of Chicago Press; Chicago.
- WYLES, J. S., J. G. KUNKEL, AND A. C. WILSON. 1983. Birds, behavior, and anatomical evolution. *Proceedings of the National Academy of Sciences* 80:4394-4397.