Part VI

 \sim \sim

SCIENCE AS PROCESS

 $\boldsymbol{\dot{\gamma}}$

Phylogenetic Analogies in the Conceptual Development of Science¹

Brent D. Mishler

Duke University

David Hull's approach to science, which culminated in his important book *Science as a Process* (1988), represents an unprecedented conjunction of philosophy of science with the results and concepts of a particular science. Hull takes an evolutionary approach to the conceptual development of science, importing much of his ex-
planatory framework from comparative biology, the discipline where his empirical observations of scientists have been made. On the surface, such a cozy relationship between data, theory, and metatheory leads to worries about circular reasoning (Mishler 1989). Nevertheless, I will argue here that Hull's approach is basically sound, and that the strength of his arguments comes precisely from his recognition of key analogies between the evolution of organisms and the conceptual evolution of sci- entists; however, certain disanalogies must also be taken into account.

1. The Current Status of Systematic and Evolutionary Biology

Comparative biology can be divided into two distinct (although interrelated) gy focuses primarily on patterns in the history of life. Evolutionary biology focuses primarily on process explanations for these patterns. The two are interconnected in a feedback loop, because the choice of proper systematic methods depends on at least a tions depend on a sound systematic framework. Such an arrangement needs careful examination, to prevent it from drifting on the one hand into vicious circularity or on the other hand into a sterile separation of theory from data. The recent cladistic revo- lution in systematics was spurred by Hennig's (1966) insights on how to connect these two areas of comparative biology. He initiated a body of method and theory that does allow us, for really the first time in the history of systematics, to walk the thin line between the two undesirable alternatives: circularity and sterility.

Hennig had one brilliant, central insight, from which several corollaries were de-
rived. His key insight, sometimes known as the Hennig Principle, was that homolo-
gous similarities shared among a group of organisms are o similarity is a feature shared by all and only members of an assemblage of organisms due to inheritance from their immediate common ancestor. He termed such features

PSA 1990. Volume 2, pp. 225-235 Copyright © 1991 by the Philosophy of Science Association *synapomorphies* (i.e., shared, derived characters or "special" similarities). A second kind of similarity is a feature shared by all, but not only, members of an assemblage of organisms due to inheritance from a distant common ancestor. Hennig termed those features *symplesiomorphies* (i.e., shared, primitive characters or "general" similarities). His principle was that only synapomorphies are valid indicators of phylogenetic relationship (i.e., relative recency of common ancestry).

A major corollary to this principle is that only synapomorphies should be used as evidence for the naming of formal taxonomic groups (known in a general sense as "taxa" when the rank is unspecified). Hennig redefined *monophyly* so as to be compatible with this corollary: a monophyletic taxon is one that contains all and only descendants of a common ancestral species, recognized as such by the discovery of synapomorphies.

The meaning of *homology* has been clarified as well under Hennig's system. There are several distinct categories of homology, which when taken broadly, can be defined as a correspondence between two or more characteristics of organisms that is caused by a historical continuity of information (Roth 1988). *Iterative homology* is historical correspondence between different structures within a single organism. *Toxic homology* is a correspondence between features in different organisms due to inheritance from a common ancestor that possessed that feature (equivalent to synapomorphy; Patterson 1982). *Transformational homology* is a correspondence be-
tween two different features resulting from a historical modification of one into the other (equivalent to the relationship between an apomorphy and its plesiomorphy). Two or more features hypothesized to be related in this way are termed *transformation series.* Such transformationally related features are usually called "character states," whereas the whole series is called a "character."

The transformation series can be *polarized* if a defensible hypothesis can be made that one state is plesiomorphic (i.e., "primitive," or temporally prior). Evolutionary polarity of a transformation series can be hypothesized in several ways (Stevens 1980); the most favored criterion in cladistics is *outgroup comparison,* which is based on an examination of the distribution of character states outside the study group. That character state occurring widely outside the group is assumed to be plesiomorphic. Note that any state of a transformation series can be the plesiomorphic one, whether one of the "ends" of the series or an intermediate state.

Taxic homology (synapomorphy) is initially postulated based on detailed similari- ty between corresponding features in different organisms (the "similarity test" of Patterson 1982), but it is not accepted until the distributions of all putative homologies are compared using a parsimony criterion (see Sober 1988 for a recent discussion). Those putative homologies that are congruent wit branching diagram representing a hypothesis of relative recency of common ancestry) based on the most parsimonious arrangement of all putative homologies are accepted as homologies (the "congruence test" of Patterson 1982); those that are incongruent are termed *homoplasies.*

Such character incongruencies can be due either to simple mistakes in character study, or to various biological processes that are of evolutionary interest (even though they cause epistemological problems for systematics). Two identical features may arise independently in two different groups, a phenomenon known as parallelism or convergence. As long as sufficient evidence (in the form of other, independent synapomorphies) exists, parallelism can be discovered.

Homoplastic character distributions can also arise via "horizontal" transmission. Hennig assumed an underlying model of diverging evolution coupled with "vertical" descent with modification. Various reticulating processes, however, including hybridization, introgression, and even lateral transmission of genetic material by viruses are known to occur. The epistemological effects of horizontal transmission on biological systematics are as yet incompletely explored; however, the consensus among cladists seems to be that rampant reticulation can destroy our ability to reconstruct relationships (and should, because complete reticulation is an "information destroying" process in the sense of Sober 1988), but that partial or occasional reticulation can be discovered and accounted for in cladistic analysis by means of careful study of organisms.

These various concepts can be illustrated using Figure 1, which is a hypothetical cladogram of six species. The topology of this cladogram is assumed to be based on other characters than just the one discussed here (since a fully resolved cladogram must have synapomorphies supporting all nodes). The distribution of a single character is mapped onto the cladogram. This character is a three state transformation series between $\mathbf Q$ (which is primitive, widely distributed outside this group of six species), \odot , and \bullet . The latter two character states are transformational homologs of \bigcirc is hypothesized to be transformed from \bigcirc because elements of \bigcirc are seen in \bullet (perhaps in ontogeny), yet the latter has elements unique to itself. Given this scenario, \bullet is a synapomorphy of species 5 & 6, relative to the plesiomorphy \mathbb{G} . Likewise \mathbb{G} is a synapomorphy of species 3, 4, 5, & 6, relative to the plesiomorphy O .

Figure 1

This example illustrates two important points: (1) The relational nature of the concept of synapomorphy; © can be referred to either as a symplesiomorphy or a synapomorphy, depending on the phylogenetic level (one reason why the apomorphy/plesiomorphy distinction is not identical to the older, derived/primitive distinction). (2) The abstract nature of synapomorphy; not all members of a taxon (e.g., a group composed of 3,4,5, & 6) need actually possess the synapomorphic feature. Several exceptions are possible. The original synapomorphy may have been transformed into a distinctive new feature in some derivative lineage (evidence for such transformation can often be discovered through ontogenetic studies), or it may even have been "lost" in a derivative lineage, giving rise to a character state that cannot be distinguished from the plesiomorphic state (such character "loss" or "reversal" is con-
fusing epistemologically, but can be discovered if enough other synapomorphies are present to reconstruct the true relationships). The synapomorphy for a group may happen to only occur in adult females; the lack of the character in males and juveniles can be explained genetically and developmentally and thus is not taken as evidence to exclude such individuals from the group. Therefore, groups diagnosed using synapo-
morphies are not necessarily monothetic in an observational sense, but they are mono-
thetic in a theoretical sense.

Continuing the example (Fig. 1), of the possible higher taxa shown, only C would be monophyletic. If named, A and B would be *paraphyletic* (i.e., a group of organisms including only some of the decendants of a common ancestor, thus having some included organisms that are actually more closely related to organisms outside the group). Such taxa are considered unnatural and therefore prohibited in Hennigian cladistics. Ontologically, groups such as C are monophyletic if they are hypothesized to include all and only descendants of a common ancestor. Synapomorphic characters have primarily an epistemological role; they do not, strictly speaking, *define* groups, but rather *diagnose* them. They are not the taxon, they are the evidence for the taxon.

The nature of species, as the basal taxon and perhaps a basic unit of evolution, is a subject of considerable controversy both within and without the cladistic school (see recent discussion by Mishler 1990). For the purposes of the present paper, it is partic- ularly necessary to come to grips with the biological meaning of "basal" phylogenetic taxa, in order to investigate the model for Hull's classification of basal scientific re- search groups.

Hennig himself (1966), as well as many of his followers (e.g., Wiley 1981; Nixon and Wheeler 1990) argued that the concepts of synapomorphy and monophyly do not, arid should not, extend to the species level. These authors have variously argued for the application of other criteria at the species level (e.g., reproductive compatibility or diagnosability using unpolarized similarities). Mishler and Donoghue (1982) and Mishler and Brandon (1987), however, developed a case for treating species like cladistic taxa at all other levels. The primary argument for this is logical consistency; the basal units in a phylogenetic system should be phylogenetic units.

In this phylogenetic approach, species are viewed as basal monophyletic taxa, into which organisms are grouped because of the presence of synapomorphies. Hennig's definition of monophyly was broadened and clarified by Mishler and Brandon (1987) to include "all and only descendants of a common ancestor, originating in a single event," where "ancestor" refers not to an ancestral *species* (as in Hennig's definition), but to a lower-level entity that is fully an *individual* in the sense of Hull (1976). The "event" referred to in this definition is the spatiotemporally localized action of one of a number of possible causes, including hybridization.

It was further argued in Mishler and Brandon (1987) that whereas all formally named taxa should be monophyletic, not all monophyletic groups should be formally named. Formal taxonomy should extend "down" (i.e., towards less inclusive groups) only so far as distinct, "important" lineages can be discovered. It is counterproductive (not to mention impossible) to attempt to name all monophyletic groups. Small lineages are constantly being produced, existing for a geographically localized kin groups in organisms with limited dispersability, multiple

origins of sterile hybrids between two widespread parental types, and even cell lin- eages within clonal plants). Only a tiny fraction of these go on to be recognized as important lineages and to be formally named as species (e.g., those that have acquired a distinctive evolutionary novelty that alters selective regimes or patterns of inter- breeding). Often this recognition is post facto, because it may not be clear early on that an important new lineage is arising.

Evolutionary theory, particularly the nature of natural selection, has also under- gone considerable clarification and expansion in recent years. A recent book by Brandon (1990) codifies and integrates a rational approach to the related processes of selection and adaptation, that can be outlined briefly as follows. An adaptation, in a loose sense, is a match between some feature of an organism and a "problem" posed by its environment. The only known evolutionary process producing adaptations is natural selection. Several conditions are necessary for selection to occur: (1) There must be variation in some feature within a population (i.e., a spatiotemporally local- ized group of genealogically related individuals). (2) Variation in that feature must be heritable to some degree (i.e., offspring must preferentially resemble parents). (3) Some variants must endow their organisms with a propensity to leave more offspring in future generations (i.e., to have higher fitness) than other organisms. If these conditions hold, then competition in a common selective environment will lead to evolution of the population by natural selection, producing adaptations in the strict sense (Brandon 1990). Such adaptations are a match between organism and environment produced through a historical process of selection for certain heritable characters in the context of that environment.

The process of natural selection is a hierarchical one; the most familiar example of selection among individual organisms, given above, is by no means the only possibili- ty. To understand, however, the nature of the "levels of selection" problem, one must realize that two hierarchies are involved (Brandon 1990).

Hull (1980) developed an important distinction between *replicators* and *interactors* in selection processes. The replicator is the unit of heredity and reproduction: an entity of which direct copies are made. Replicators thus form lineages; a population is a spatio-temporally isolated, integrated and/or cohesive, genealogically related, col- lection of replicator-lineages. The interactor is the unit involved in the competition process: an entity whose interaction with the environment causes replication to be dif- ferential, lnteractors, unlike replicators, are not defined genealogically.

Brandon (1990) has shown that both sorts of entities are organized into hierar-
chies, and that these two hierarchies are not necessarily congruent. In other words,
several levels of replication may be in operation in a p (whether organisms, populations, species, or higher taxa), but at least the properties necessary for a level to participate causally in a selection process have been clarified.
For one of these higher-level entities to participate, it must either replicate itelf direct-
ly, or act as an interactor. For the l when the fitness of an organism in a particular selection process depends on its group membership rather than on its genotype; Brandon 1990).

Not all evolutionary change, at any hierarchical level, is due to the process of nat- ural selection/ adaptation (Gould and Lewontin 1979). In finite populations (defined as above in a genealogical sense), *random drift* will occur, its importance depending

upon population size and the strength of selection. Furthermore, natural selection/ adaptation is not the only constraint on observed phylogenetic patterns of character distribution. Not all conceivable variants are available to be input to selection. Some conceivable variants violate basic physical laws; such prohibitions are called *physical constraints.* Some features are physically possible, yet are never produced in a particular lineage because complex developmental processes give rise to an epigenetic homeostasis that biases against (or even prohibits) certain character combinations. Such historically contingent homeostatic mechanisms are called *developmental constraints.*

2. Hull's Evolutionary Approach to the Process of Science

The view of scientific process laid out by Hull, most fully in his 1988 book, relies heavily on analogies derived from systematic and evolutionary theory. His view can be summarized as follows (see also Mishler 1987, 1989). Scientists are grouped into basic lineages, called research groups, just as organisms are grouped into species. These research groups are bound together by various sociological ties, including: stu-
dent/teacher and collegial relationships, shared language, peer reviewing of papers
and grants, mutual use and citation of each other's ogous to the way in which species are bound together by interbreeding, ecology, and/or developmental constraints. Specific ideas function as traits; an analog of natural selection occurs as scientists use other scientists' uncommitted scientists, graduate students), "speciating" sometimes as they give rise to splinter groups, and eventually going extinct.

A clear understanding of these putative analogies, and their efficacy, requires de- tailed comparisons between the biological concepts and scientific process. A summa- ry of some proposed analogies is given in Table 1.

Any close analogy between the biological process of natural selection and the sci- entific process depends on the nature of replicators and interactors in the latter. The replicator, in the case of science, could be either the scientist or the idea, depending on the nature of transmission. Sometimes, scientific ideas are passed on wholesale, as in many professor/ graduate student relationships. In such a case, the "structure" of the scientist is being passed on directly. In many other cases (e.g., professional inter- action among colleagues), individual ideas are taken up and added to an existing mix of other ideas.

Just as in biology, a distinction should be made between "genotype" and "pheno-
type" (see also Griesemer 1988). In terms of the evolution of scientific ideas, a distinc-
tion is necessary between the "meme" (sensu Dawkins ed, applied, and understood in the scientific community. The "developmental" connection between individual memes and their assembly into functioning, public ideas can be complex, as is often the case in biology (and may lead to a sort of "intellectual inertia" — an analogy of developmental constraints in biological evolution). Unlike biology, however, the transmission of memes is via their colleagues do not have direct access to the memes in the brain of a scientist, only indi- rect access through writings, discussion, and observation of behavior.

Table 1. Proposed analogies between concepts derived from systematic and evolutionary biology and components of the scientific process. See text for discussion and explanation.

Interaction in science clearly occurs at many levels of inclusion, thus presenting a pronounced "levels of selection" problem. Competition between ideas can occur within the mind of an individual scientist (analogous to somatic selection in biology;

Brandon 1990). Given that scientists are "fixed" in their ideas, competition can occur among them for converting or recruiting other scientists to their viewpoint. Higher fitness for a scientist would imply a propensity for propagating a relatively high number of copies of ideas to later "generations" of scientists. A group of scientists (Hull's "research groups") sharing homologous ideas can function as an interactor relative to other such groups (analogous to the models of group selection in biology discussed by Sober 1984; Brandon 1990), if the fitness of scientists is affected by group membership irrespective of their personal set of ideas. It is also at least conceivable that high-
er-order groups of scientists (e.g., entire research traditions, say systematic biology) could function as interactors in a selection process relative to other such groups (say molecular biology).

As emphasized by Brandon (1990), natural selection among several variants must occur within a common selective environment. The common environment in which the putative selection process among scientific ideas takes place, is composed of both sociological and empirical factors. This is analogous to the usual situation in biology, where the relevant environment consists of both biotic and physical factors (e.g., competitors, parasites, and climate). Other organi part) of the environment that needs to be adjusted to. An adaptive idea is one that is perceived to "solve" a scientific problem (in the sense of Laudan 1977). This implies both "progress" in the eyes of other scientists, and "progress" in increasing our under- standing of the natural world.

Hence, Hull's view of science does not boil down to sociological relativism (even though he seems to have missed the nice analogy with adaptation). Some ideas *are* intrinsically better than others in their fit to the empirical world (as it is currently known). The inescapable sociological milieu of science functions in both replication and interaction, yet competing scientists do continually attempt to match their con- cepts up with the real world. Just as in biological evolution, however, many ideas (traits) evolve by drift, rather than by natural selection. In addition, ideas that are adaptive are "tracking" an "environment" that may well be labile. Increasing fit between ideas and the real world occurs over time, but this progress is neither linear, precise, nor perfect. All this illustrates the way in which scientific evolution is and is not "progressive," and may help to assuage the concerns of Donoghue (1990), who rightly pointed out that Hull (1988) neglected the importance of "the worth of ideas" to scientists.

What is the nature of the patterns produced as a result of these various selection processes in science? Replicators, by their very definition, produce lineages. The biological concepts of taxic homology (synapomorphy), transformational homology, speciation, and monophyly have their analogs in recognizing and understanding lineages in the the history of science.

To count as the "same" (i.e., a taxic homology) an idea shared among several scientists must be identical by descent. As emphasized by Hull (1988), independent origins of an idea by an unappreciated precurser or by a conte versions of old ideas, connected by descent, count as transformational homologies.

As with organisms, informative classifications of scientists should be based on ho- mologies, take into account the polarity of transformation series of ideas, and attempt to group by synapomorphy. New research groups can be produced either by fission of a pre-existing one ("speciation"), or by fusion of two or more pre-existing ones ("hy-

bridization"). The meaning of monophyly may seem even more problematic in reference to lineages of scientists than to organisms, because it is clear than an exceptionally large amount of horizontal transmission occurs in t to adopt ideas from diverse research groups (although the idea-flow among groups is much less than within a group), in a way that has no close analogy among the more complex organisms such as mammals or land plants.

Nonetheless, considerable horizontal transmission is suspected to have occurred in the phylogeny of simpler organisms such as the fungi or bacteria. As discussed above for organismal systematics, horizontal transmission causes epistemological problems for phylogeny reconstruction, yet not ontological problems so long as it is relatively infrequent compared to vertical transmission. Recognizable lineages can be maintained in the face of considerable horizontal gene-flow. If sufficient horizontal transmission occurs, then two lineages may blend to produce o

The application of the concept of monophyly to research groups must involve a synchronic approach, called by Sober (1988) the "cut method" (similar to the approach taken for species monophyly by Mishler and Brandon 1987). At any given moment in time, one can make a horizontal slice across lineages and define monophyletic groups. At a later time, a horizonal slice might well imply a different decision. Furthermore, the synapomorphous characters marking a single lineage will be different (both in number and description) at different time-slices.

3. Conclusions

It appears that most of the relevant analogies that can be made between organis- mal and scientific evolution are sound. Given the complexity of comparative biology and its theoretical structure (which has only been rigorously formalized in the last two decades), it is not surprising that a rigorous view of evolution of science has been long in coming. This view, as it has developed in the works of Kuhn and Hull, among many others, is in need of further theoretical and empirical work.

In the theoretical realm, careful consideration is needed of the effects of various complicating factors known from the study of biological evolution. The models of selection discussed above are rather simplistic as compared to real situations known in biology wherein several levels of selection operate simultaneously and often in opposing directions. The presence of possible analogies to developmental or physical constraints should be investigated.

The meaning of heritability as applied to possible replicators in science similarly needs to be addressed, as does the process of "reproduction" itself. One disanalogy between science and evolutionary biology (mentioned above) is the lack of compara- ble transmission mechanisms for memes and genes. A second disanalogy is that in biology, organismic reproduction occurs only via production of new individual organ- isms, while in science, scientists reproduce either via production of "new" scientists (i.e., training of students) or via conversion of an existing scientist. The effects of such conversion on standard models of selection may be interesting; one could well imagine an accelerating effect on evolutionary rates.

In standard formulations of selection processes in evolutionary biology the raw ficial traits (Sober 1984). In the conceptual evolution of science, however, the origin of memes often occurs in a directed, "Lamarckian" manner (Boyd and Richerson

1985). Such a bias in production of raw memetic variation towards adaptive solutions would also be expected to have a accelerating effect on evolutionary rates.

In the empirical realm, well worked-out examples are needed. Despite Hull's descriptive efforts in the field of systematic biology itself, there remain no explicit applications of the analytical tools of cladistics or population biology to the evolution of science (or, for that matter, to other academic disciplines, where the same tools and concepts may be relevant). Such applications will present novel problems and opportunities. We have better sources of evidence for cohesion and interaction within and between lineages than are usually available for organisms (e.g., the written historical record), yet synapomorphous traits (ideas) marking lineages may be difficult to discover and describe given the speed with which transformation occurs in scientific
ideas and the significant amount of inter-lineage borrowing that goes on in science. Selection processes in science should be relatively easy to discover, as long as care is taken to untangle causation at specific levels in the hierarchies of replicators and in- teractors.

It appears that a successful investigator of the process of science is going to need extensive training in the theory and methodology of comparative biology as well as in more standard training in philosophy and historical methods. Perhaps no one current- ly has sufficient strengths in all these areas to produce a rigorous study; extensive inter-lineage borrowing is clearly indicated. I just happen to have a place in my lab for any philosopher of science who wants to try.

Note

 JI thank **V.** Albert, R. Brandon, A. Gutierrez, C. Horvath, D. Hull, and S. Rice for comments on the manuscript and discussion.

References

- Boyd, R. and Richerson, PJ. (1985), *Culture and the Evolutionary Process.* Chicago: University of Chicago Press.
- Brandon, R.N. (1990), *Adaptation and Environment.* Princeton: Princeton University Press.

Dawkins, R. (1976), *The Selfish Gene.* Oxford: Oxford University Press.

- Donoghue, MJ. (1990), "Sociology, Selection, and Success: A Critique of David Hull's Analysis of Science and Systematics", *Biology and Philosophy* 5: 459- 72.
- Gould, S J. and Lewontin, R. (1979), "The Spandrels of San Marco and the Panglossian Paradigm: A Critique of the Adaptationist Programme", *Proceedings of the Royal Society of London, B* 205: 581-98.
- Griesemer, J.R. (1988), "Genes, Memes, and Demes", *Biology and Philosophy* 3: 179-84.

Hennig, W. (1966). *Phylogenetic Systematics*. Urbana: University of Illinois Press.

Hull, D.L. (1976), "Are Species Really Individuals?", *Systematic Zoology* 25: 174-91.

- . (1980), "Individuality and Selection", *Annual Review of Ecology and Systematics* 11: 311-32.
- . (1988), *Science as a Process: An Evolutionary Account of the Social and Conceptual Development of Science.* Chicago: Chicago University Press.
- Laudan, L. (1977), *Progess and its Problems: Towards a Theory of Scientific Growth.* Berkeley: University of California Press.
- Mishler, B.D. (1987), "Sociology of Science and the Future of Hennigian Phylogenetic Systematics", *Cladistics* 3:55-60.

... (1989), [Untitled review of Hull, D.L. 1988, Science as A Process.] *Systematic Botany* 14: 266-268.

- ... (1990), "Species, Speciation, and Phylogenetic Systematics", *Cladistics* 6: 205-209.
- _ and Brandon, R.N. (1987), "Individuality, Pluralism, and the Phylogenetic Species Concept", *Biology and Philosophy 2:* 397-414.

and Donoghue, MJ. (1982), "Species Concepts: A Case For Pluralism", *Systematic Zoology* 31: 491-503.

- Nixon, K.C. and Wheeler, Q.D. (1990), "An Amplification of the Phylogenetic Species Concept", *Cladistics* 6: 211-23.
- Patterson, C. (1982), "Morphological Characters and Homology", in *Problems of Phylogenetic Reconstruction,* K.A. Joysey and A.E. Friday (eds.). London: Academic Press, pp. 21-74.
- Roth, V.L. (1988), "The Biological Basis of Homology", in *Ontogeny and Systematics,* C J. Humphries (ed.). New York: Columbia University Press, pp. 1-26.
- Sober, E. (1984), *The Mature of Selection: Evolutionary Theory in Philosophical Focus.* Cambridge, MA: MIT Press.
- . (1988), *Reconstructing the Past: Parsimony, Evolution, and Inference.* Cambridge, MA: MIT Press.
- Stevens, P.F. (1980), "Evolutionary Polarity of Character States", *Annual Review of Ecology and Systematics* 11: 333-58.
- Wiley, E.O. (1981), *Phylogenetics: The Theory and Practice of Phylogenetic Systematics.* New York: John Wiley.