9

Resolving Phylogeny: Evolutionary Archaeology's Fundamental Issue

MICHAEL J. O'BRIEN AND R. LEE LYMAN

An increasing number of archaeologists are showing interest in employing Darwinian evolutionary theory to explain variation in the material record. Several historical reviews of evolutionism in Americanist archaeology (Dunnell 1980; O'Brien 1996a, 1996b; O'Brien and Holland 1990) devote considerable space to the differences between Darwinian biological evolutionism and the cultural evolutionism of White (1949, 1959), Steward (1955), and others, but often ignored is the fact that some culture historians in the first half of the twentieth century (Colton and Hargrave 1937; Kidder 1915, 1932; Kroeber 1931) acknowledged, if only metaphorically, the applicability of Darwinian evolutionism to the study of prehistory (Lyman and O'Brien 1997, 2001b; Lyman et al. 1997; O'Brien and Lyman 1998, 1999b). Kidder (1932), a leading culture historian of the period 1910–1940, found some of the concepts embedded in Darwinian evolutionism—descent with modification and extinction—quite appropriate when discussing artifact lineages. Throughout his career Kidder focused on documenting variation—the singular issue that underlies any Darwinian study—and attempting to explain it. To Kidder, variation in such things as pottery was not fundamentally different from genetically controlled variation. Rather, there simply was more of it being produced as a result of the almost limitless imagination of humans and their enormous capacity for effecting change in their social and physical environment.

Kidder's views are important because they indicate how at least one archaeologist working in the early decades of the twentieth century conceived of the phenomena he was studying: Cultures evolved; a historically documented culture had a developmental heritage, or historical lineage, and it was the archaeologist's job to describe that lineage and to determine why it had taken the form that it did. But Kidder (1932) correctly indicated that archaeology lacked both the basic data and a theory consisting of cultural processes parallel to the biological ones of genetic inheritance and natural selection to help explain a culture's lineage in evolutionary terms. Without the means to document variation and then to link that variation to applicable theory, which dictates how and why we measure the variation, so-called explanations of how and why things change—the epitome of evolutionism—are simply untestable interpretations about the past.

Kidder realized that compiling the data and retooling Darwinian evolutionism so it was directly applicable to the archaeological record would take some concerted work, but few archaeologists were inspired by Kidder's call to action. Although chronologies of artifact types and larger units variously termed cultures, phases, complexes, and the like have been constructed, tested, refined, empirically verified, and are now available for many areas of the Americas, these have the status of being mere historical sequences. They may or may not comprise lineages, and thus explanations of them in the sort of evolutionary terms Kidder used are notably lacking. Any serious effort to make Darwinian evolutionism applicable to the archaeological record must be grounded in trying to resolve phylogeny. Our objective here is to outline a few of the many conceptual and methodological issues that must be addressed before such a resolution can occur.

MODERN EVOLUTIONARY ARCHAEOLOGY AND PHYLOGENY

Kidder's contemporaries, such as Gladwin (1936), argued that because cultural evolution occurred much more rapidly than biological evolution, Darwin's ideas were not applicable to the archaeological record—a point of view that reached its zenith with the oft-quoted statement by Brew (1946:53) that "phylogenetic relationships do not exist between inanimate objects." This pervasive sentiment effectively squelched any efforts to move beyond chronological ordering of artifacts and toward the development of a theoretical perspective on why some artifacts from different geographic areas closely resembled each other. In the absence of theory that might explain such similarities, culture historians borrowed common anthropological notions such as diffusion and trade to account for the similarities. As with their anthropological colleagues, culture historians viewed cultural evolution and biological evolution as entirely uncoupled phenomena. Darwinian evolutionism was viewed as applicable to the biological side of humans, but it was seen as irrelevant to the cultural side. Filling the void was White's (1949, 1959) brand of cultural evolutionism, with its emphasis on culture as humankind's extrasomatic means of adaptation. Whitean evolutionism became the cornerstone of processual archaeology in the 1960s. Although some processualists occasionally made reference to Darwinian evolutionism, it wasn't until publication of Dunnell's (1978a) "Style and Function: A Fundamental Dichotomy" that there was an incipient programmatic statement on how to write that particular kind of evolutionism in strictly archaeological terms. Few people paid serious attention to that article, primarily because it addressed two mainstays of archaeology—style and function—in terms that were uncommon in the archaeological literature. This was especially true with respect to how Dunnell treated style. We return to this point below.

Dunnell's early work (1978a, 1978b, 1980) was followed by occasional articles throughout the 1980s (e.g., Dunnell 1989; Leonard and Jones 1987; Meltzer 1981; O'Brien 1987; Rindos 1989c), but it was not until the 1990s that evolutionary archaeology attracted more than modest interest. During that decade the number of articles and books grew exponentially and continues to grow. Some of these focus on method or on rewriting evolutionary theory in archaeological terms (Bettinger et al. 1996; Hunt et al. 2001; Neff 1992; Neff and Larson 1997; O'Brien 1996a; O'Brien and Holland 1990, 1992; Teltser 1995a), but many are case studies of how Darwinian evolutionism can be applied to particular portions of the archaeological record (e.g., Allen 1996; Aranyosi 1999; Bettinger and Eerkens 1997, 1999; Cochrane 2001; Dunnell and Feathers 1991; Hamilton 1999; Hughes 1998; Leonard and Reed 1993, 1996; Lipo 2001; Lipo et al. 1997; Madsen et al. 1999; Maxwell 1995; McGimsey 1995; Neff et al. 1997; Neiman 1995; O'Brien et al. 1994; Pfeffer 2001; VanPool 2001; Vaughan 2001).

Despite the advances made in adapting Darwinian evolutionism to archaeology, we agree with Schiffer (1996) that more case studies are needed in which evolutionism has been shown to produce explanations that are on par with or superior to those produced by other intellectual programs. Further, until recently (e.g., Lipo and Madsen 1999; Lipo et al. 1997; Lyman and O'Brien 1998, 2000b; Neiman 1995; O'Brien and Lyman 2000b), evolutionary archaeologists have been largely silent about the processes that generate the enormous variation in the archaeological record. As we point out in detail elsewhere (O'Brien and Lyman 2000a, 2002a), for evolutionary archaeology to be more than an intellectual novelty—something ignored by the majority of the discipline (Preucel 1999)—those with an interest in the approach need to be clear about such basic issues as what it is that evolves (Dunnell 1995; Lyman and O'Brien 1998) and how evolutionary change is to be measured (Lyman and O'Brien 1999, 2000b; O'Brien and Lyman 2000b). And as has been made clear (Bettinger and Richerson 1996; Bettinger et al. 1996; Schiffer 1996, 2000), evolutionary archaeologists are going to have to be more accommodating of other intellectual approaches that incorporate elements of Darwinian theory, such as behavioralism (Schiffer 1996, 1999), human evolutionary ecology (Boone and Smith 1998; Broughton and O'Connell 1999; Winterhalder and Smith 2000), and even cognitive science (Mithen 1996).

These are a few of the challenges that evolutionary archaeology has ahead of it, but there is another challenge, and it comes from the inside. Lest it be assumed that evolutionary archaeology itself is a well-integrated approach over which there is universal agreement, we point out that considerable disagreement exists within evolutionary archaeology over fundamental issues such as whether artifactual units can be used to examine human fitness. Neff (1999, 2000, 2001), for example, is of the opinion that although artifacts can be examined in terms of artifact fitness—how well one kind of artifact outreplicates another kind—that particular kind of fitness may have no bearing on human reproductive success. We wonder if that view, which is simply a recasting of Gladwin's (1936) observation that culture change can be faster than genetic change, will be used as a reason to discard *all* aspects of Darwinism from archaeological research, as it was used to that end in the 1930s and 1940s (Lyman and O'Brien 1997). We therefore prefer to think that the "replicative success" (Leonard and Jones 1987) of an artifact type may be driven by natural selection working on the replicators (humans), by natural selection working on the interactors (artifacts), by the vagaries of transmission, or by some combination thereof such that one evolutionary process or the other applies at different times (Lyman and O'Brien 1998, 2001a; O'Brien and Lyman 2002b). We agree with Neff's central contention that the problem of which one works at which time is less straightforward than has previously been assumed.

Such internal wrangling should render as moot any claims that evolutionary archaeology is flawed because it is scientistic or because it makes a claim of corporate affiliation (Wylie 2000). The approach is not scientistic—it does not advocate one and only one approach to scientific investigation (here, the problem seems to reside in differences of opinion regarding the role of *theory* in scientific endeavors)—and neither is there a corporate affiliation. We believe evolutionary archaeology is pluralistic (O'Brien and Lyman 2000b), but we do not believe that all intellectual programs that exist in Americanist archaeology are of equal weight when the subject is Darwinian evolution (Lyman and O'Brien 1998; O'Brien et al. 1998). We believe there is much to be learned from the disciplines of evolutionary biology and paleobiology (Lyman and O'Brien 2000a, 2000b; O'Brien and Lyman 1999a, 2000b), and we have learned much from the successes and failures of the much-maligned paradigm of culture history (e.g., Lyman and O'Brien 2002; Lyman et al. 1997; O'Brien and Lyman 1998, 1999b), although some critics (e.g., Longacre 1999) have missed the latter point. We do, however, find metaphysical shortcomings in how evolution has generally been approached in archaeology, and we have offered alternatives (O'Brien and Lyman 2000b). Some archaeologists (e.g., Bettinger et al. 1996; Boone and Smith 1998; Schiffer 1996) have found what they consider to be weaknesses in our approach and have offered additional alternatives. Metaphysical, theoretical, and methodological differences of opinion are to be expected, if not demanded, in any scientific approach to problem solving (Bell 1982). That these differences over evolutionary archaeology exist and can be debated is a sure sign that whatever is going on, it certainly is not scientistic.

The attempt to integrate Darwinian evolutionism into archaeology reminds us in no small measure of how things were in evolutionary biology in the Synthesis days of the late 1930s and 1940s, when there was deep theoretical and methodological division among geneticists, neontologists, and paleontologist (see Mayr and Provine 1980). For example, Simpson worked for over two decades to show that paleontology could make significant contributions to the Darwinian theory emerging from the Synthesis rather than being simply a source of confirmation of that theory. Simpson's (1944) *Tempo and Mode in Evolution* provided both a statement on the applicability of the new evolutionism to the fossil record and a methodological synthesis of how paleontologists could examine that record and provide unique insights into the evolutionary process. Reviewers of the book (e.g., Hubbs 1945; Huxley 1945; Wright 1945), however, were not impressed, and most biologists and paleontologists ignored the implications of the differences between microevolution and macroevolution—paleontology being uniquely capable of revealing aspects of the latter—that Simpson underscored (Laporte 2000). Only in the last quarter century has Simpson's original vision been vindicated (Eldredge 1985).

The parallels between Americanist archaeology and paleontology are both remarkable and perhaps expectable. Like paleontology, Americanist archaeology has long struggled for legitimacy within its parent discipline. Since its birth in the late nineteenth century, and despite occasional pleas or encouragement for archaeologists to "shrive yourselves of the notion that the units which you seek to reconstruct must match the units in social organization which contemporary ethnographers have attempted to tell you exist" (Harris 1968:360), archaeology remained for the most part "the tail on an ethnological kite" (Steward 1942:341). To do more than serve as mere technicians who retrieve broken pots and arrowheads from layers of sediment in the service of historical reconstruction, as Service (1964) put it, archaeologists—in part because they were told to do so by anthropologists—adopted ethnological theory as the center piece of their explanatory tool kit. For their part, archaeologists were uneasy to make the break from anthropology, with most agreeing at least in principle with Taylor's (1948) detailed arguments that archaeology must strive to be less of a descriptive taxonomic enterprise and more anthropologically oriented and also agreeing with Phillips's (1955:246–247) admonition that "New World archaeology is anthropology or it is nothing." The latter became the rallying cry of processual archaeologists of the 1960s (e.g., Binford 1962; Flannery 1967).

Paleontology struggled to make itself heard throughout the years preceding and following the Synthesis, but it was in some ways more successful than archaeology has been. Paleontologists developed ways to adapt—not merely adopt—biologically based evolutionary theory to the unique aspects of the fossil record—its lack of direct evidence of genetic transmission, behavior, physiology, and so on—in the process developing what came to be known as "paleobiology." Archaeologists in the 1960s and 1970s followed exactly the opposite strategy and developed what came to be known as "anthropological archaeology." They transferred ethnological theories of various sorts to the archaeological record without any adaptation of those theories to the vagaries and nuances of the material record. This strategy required that the archaeological record be converted to an ethnological-like record, and thus ethnoarchaeology, formation-process studies, and middle-range theory grew in importance. As Watson (1995) documents, the culprit seems to reside in the adoption of "culture" as a keystone concept within anthropological archaeology (see Bennett [1998] for an outline of some of the attendant problems).

Evolutionary archaeology today in many respects is in a period similar to that of evolutionary biology in the 1970s when paleobiologists (e.g., Eldredge 1971; Eldredge and Gould 1972; Stanley 1975) proposed a new way to look at the fossil record that involved some retooling of biological evolutionary theory to fit the fossil record. Despite more than two decades of squabbling, the success of their endeavors is abundantly apparent (see review in Eldredge 1995). Evolutionary archaeology is a bit past the middle of the squabbling period, but certain challenges remain. Nonetheless, we suspect that in the next few years evolutionary archaeology's success (or failure) should be assured. The fits and starts that evolutionary archaeology is going through will, we believe, produce a synthesis of its own (O'Brien and Lyman 2002b). The theory requisite to archaeology, unabashedly conceived of as a science concerning the historical development of peoples and their cultures, has been around since 1859, the year Darwin penned *On the Origin of Species*. The basic problem for archaeology is that, as has been pointed out several times (O'Brien and Holland 1990; Rindos 1989c), Darwin did not write a theory that can be applied directly to the study of the archaeological record.

This is problematic for archaeology, but only if we fall into the trap of believing that precisely the same theoretical postulates, not to mention the same methods and techniques, that an evolutionary biologist employs must somehow be retrofitted for archaeological applicability. Such a belief is nonsensical, although this point was never made clear in the early evolutionary-archaeology literature. Despite differences between evolutionary biology and evolutionary archaeology, there are a few tenets that they undeniably hold in common: (1) variation exists, (2) that variation is heritable, and (3) some variants do better in certain environments than do other variants. Thus, in the simplest of terms, Darwinian evolutionism is a framework for explaining change as the differential persistence of variation (Campbell 1970; Endler 1986; Lewontin 1970, 1977). If the differential production of heritable variation, of whatever sort, is central to Darwinian evolutionism, then that particular evolutionism is a body of theory and method built around the subject of *change,* not simply difference and similarity:

> The continuity implied in the terms change and persistence bespeaks a fundamental assumption: the phenomena being examined are historically and empirically related to one another (Alland 1973:3). It is also critically important to note that evolution views change as a *selective,* and not as a *transformational,* process. Variability is conceived as discrete. Change is accomplished by alteration of the frequency of discrete variants rather than alterations in the form of a particular variant. This characteristic places rather severe constraints on the application of evolutionary theory, although perhaps not as severe as it may appear on first reading. (Dunnell 1980:38)

Evolutionists study populations of things; in biology the populations are organisms, and in archaeology the populations are artifacts. As Jones et al. (1995:28) put it, it is "the differential representation of variation at all scales among artifacts" for which evolutionary archaeology seeks explanations. Evolutionary archaeology rests on the premise that because they were parts of past phenotypes, objects occurring in the archaeological record were molded by the same evolutionary processes as were the somatic features of their makers and users (Dunnell 1989; Leonard and Jones 1987; O'Brien and Holland 1990, 1992, 1995a, 1995b). This premise has been viewed by some (Larson 2000; Maschner 1998) as problematic, but we do not share this view. That artifacts, along with the behaviors that created them, are phenotypic is nonproblematic to most biologists, who routinely view such things as a bird's nest and a beaver's dam as phenotypic traits (Dawkins 1982; Turner 2000; von Frisch and von Frisch 1974). Neither is it problematic to paleobiologists, who rely on the hard parts of phenotypes to study the evolution of extinct organisms and the lineages of which they were a part. Further, paleobiologists use trace fossils—casts, burrows, and the like—as proxies for phenotypic characters and behaviors. Human-manufactured tools can be used similarly.

Darwinian evolutionary theory concerns how and why particular variants look as they do and behave as they do, where behavior is manifest as a varying frequency distribution across space and time. From an evolutionary perspective, change is measured "in terms of frequency changes in analytically discrete variants rather than the transformation of a variant" into another variant (Teltser 1995b: 53). This perspective on change runs counter to the way change normally is viewed archaeologically—that is, as a gradual or sudden transformation of a variant from one state to another. The distinction between change, which is the more or less gradual *replacement* of one variant by a phylogenically related variant, and transformation is important to recognize. As Gould (1986) put it, in Darwinian evolutionism history matters. Selection, drift, gene flow, and all the other evolutionary processes are important factors in modern Darwinian explanations, but without history they are simply processes and mechanisms with little to do and nothing to produce. Darwin's theory of descent with modification was based solidly on the notion of common ancestry. What is "descent with modification" if not a historical statement? History, when applied to anything organismic, implies the existence of lineages. The cornerstone of evolutionary archaeology, then, is resolving lineages of artifacts—an old concept in archaeology but one that has had but limited success. The problem, as we will see, is in demonstrating heritable continuity between forms as opposed to simply historical continuity, the latter being merely a sequence of forms as opposed to a lineage of related forms. How does one demonstrate that two or more phenomena are indeed parts of a lineage?

CONSTRUCTING ARCHAEOLOGICAL PHYLOGENIES

The place to begin is with systematics, which we take to be both the study of diversity of the phenomena of interest, irrespective of the scale or kind of phenomena, and the sorting of that diversity into sets such that like goes with like. The phenomena may comprise discrete objects such as projectile points or organisms; they may comprise the bits of temper in ceramics or the genes in organisms; or they may comprise assemblages of tools or populations of organisms variously termed faunas, floras, or communities. The goal of systematics in biology (Mayr 1942), paleobiology (Simpson 1961), and archaeology (Dunnell 1986; Ramenofsky and Steffen 1998) has always been to sort specimens into sets of individuals that are in some sense similar. Each set should be internally homogeneous such that within-group variation is analytically meaningless and between-group variation meaningful. Either the term *affinity* or the term *relationship* is typically used to indicate that some kind of relation exists between similar specimens within a group or between groups of specimens.

The kind of relation specified when one states that every specimen of kind A has an affinity with every other specimen of that kind, and that specimens of kind A have a different affinity with those of kind B, often is of a particular sort. Multiple kinds of things may be affines because they are close in time, in function, in symbolism, in ancestry, or in terms of something else. Specifying and measuring a *particular* kind of affinity is the ultimate goal of systematics, irrespective of discipline. This point is critical. If we are interested in ancestral-descendant affinity, this is quite different than functional affinity. We cannot assume that the kinds used to identify the latter will automatically tell us something about the former. In fact, there is every reason to suspect they will not (Allen 1996; Beck 1995b, 1998; Meltzer 1981). This was one of the points Dunnell (1978a) made in his seminal paper, but it was not made clearly enough and was lost in the ensuing debate over style and function.

The goal of systematics is realized by *classification,* which is the creation of new units and the modification and revision of old units by stipulating the necessary and sufficient conditions for membership within a unit. A *unit* is a conceptual entity that serves as a standard of measurement. An inch is a unit constructed explicitly to measure linear distance; the degrees found on a compass are conceptual units constructed explicitly for the measurement of orientation. As conceptual entities, units must be defined explicitly. Units can be specified at any scale. They can comprise one or multiple attributes of characteristics of a discrete object, of discrete objects, or of sets of discrete objects of various kinds. All sciences require that units for describing the phenomena under study be specified, but this requirement is not always readily met. Lewontin (1974:8) put it this way: "[T]he problem of theory building is a constant interaction between constructing laws and finding an appropriate set of descriptive [units] such that laws can be constructed. We cannot go out and describe the world any old way we please and then sit back and demand that an explanatory and predictive theory be built on that description." Hull (1970:32) used similar wording when he noted that the "two processes of constructing classifications and of discovering scientific laws and formulating scientific theories must be carried on together." Others (e.g., Hughes and Labandeira 1995) have recognized this critical point as well.

No one should argue with the statement that to be useful a classification must allow one to do some analytical work. The implications of this are several, and they are profound. One implication, noted above, is that a classification constructed for one purpose may not perform satisfactorily when used for another purpose (O'Brien and Lyman 2002a). A second implication is that a set of phenomena can be classified in a virtually infinite number of ways, although we are aware of very few examples of a collection of artifacts being classified in more than one way even when the collection is used to answer disparate analytical questions. A third implication is that the analytical validity of the units produced by classification must be testable. In particular, do they measure the kind of affinity sought?

Fifty years ago culture historians had as their main objective determining the chronology of occupation of a locality. Their classification procedure was to sort, by visual inspection, a pile of similar artifacts—usually sherds or projectile points—into smaller piles in which all items in a pile were more or less alike in terms of features considered chronologically diagnostic, and different than items in all the other piles in terms of the same features. One problem was the "more or less." When did the differences become important enough to start another pile? Arguments on this point raged for years in the literature, and similar problems arose in paleontology. Culture historians were well aware of the necessity of testing their hypothesized chronologies, recognizing that if their sorting did not measure the passage of time—that is, the individual piles of artifacts did not have particular temporal affinities—then they had to revise their classification. When the chronological test was passed, each pile—a *type*—was named, and a description of the type was written.

Although this classification method is suited for its purpose—to determine chronology—it more often than not fails when applied to any detailed investigation of small-scale changes over time. Not only is it completely subjective in that it places all decisions in the eyes of the classifier, it hides any small (and maybe important) changes in the "more or less alike." The type description is based either on an outstanding specimen or on an average specimen and, with discovery of more specimens within the "more or less" range, the description may require revision. Over the years, this has led to the inclusion of specimens in types that some archaeologists have argued should not be included and to the naming of new types that are identical to existing types. Units (types) constructed in this manner are descriptive in that the *general appearance* of a specimen is captured by the unit definition. These units are useful for communicative purposes so that when a particular object is said to be, say, a "pickup truck," everyone knows basically what is being discussed, but such descriptive units are oftentimes not useful for analytical purposes.

There must be some explanatory theory that guides analysis because it is theory and its derivative propositions that suggest which attributes should be measured and the requisite scale of resolution at which they should be measured. In paleontology an example would be the choice of millimeters not only to describe the size of a series of teeth of a mammalian species but to monitor variation in the size of those teeth over time and/or space in order to measure evolutionary change. One could choose color to describe a set of fossils, but color may not reflect the behaviors of the represented organisms. The construction of units useful for analysis, then, must consider the critical question: What must we measure in order to perform the required analysis?

Analytical units are *classes*—conceptual (as opposed to empirical) units of measurement. Combinations of character states define classes and are specified by the analyst. Classes can be univariate or multivariate. An advantage to multivariate classes attends the fact that morphospace is multidimensional—they allow simultaneous consideration of the values taken by "state, meristic, and metric variables" (Gould 1991:420) within a set of specimens, but *only* if they are constructed in a particular way. The procedure for constructing multivariate classes was spelled out independently by Shaw (1969) in paleontology and Dunnell (1971) in archaeology and termed "paradigmatic classification" by the latter. Morphospace is defined by a number of mutually exclusive characters, each with a set of states. Multivariate classes are defined by the intersection of characters—that is, by combinations of particular character states. The number of characters (and the number of states within a character) included in a particular classification is unrestricted.

To create a paradigmatic classification, the classifier lists all characters—length, width, color, and so forth—that he or she views as analytically important. Note the last part of that sentence. *Only* those characters that are viewed as analytically important are considered. This *might* mean that color is considered, but it might not. Color could be of analytical importance in one study and unimportant in another. In the latter case, it would be excluded from consideration. For each character used in the classification, the classifier lists the states in which each character might appear. It is the combination of character states, one state from each character, that creates classes.

Table 9.1. Shaw's (1969) Paradigmatic Classification of Conodont Morphology

Character 1: outer platform lobe development	
State L:	lobed
State U:	unlobed
Character 2: outer platform cross-sectional shape	
State F:	flat
State A:	arched
Character 3: blade and ridge (inner platform) plan	
State N:	nonparallel
State P:	parallel
State D:	divergent anteriorly
Character 4: inner platform profile	
State T:	triangular
State R:	rounded
State A:	alate

Classes can be constructed at various scales. A class can comprise a single character of a discrete object, a particular combination of character states of a discrete object, or multiple kinds of discrete objects (say, the taxa comprising a fauna). When constructing a classification with a paradigmatic structure it is important to ensure that all characters are of the same scale. This ensures that, for example, species are not being compared with families or projectile points with all chipped-stone artifacts. To repeat a point made earlier, paradigmatic classes can be constructed such that, using Gould's (1991) terms, one or more characters concern "meristic" variables (frequency), one or more characters concern "metric" variables (size), and one or more characters concern "state" variables (shape).

As an example of paradigmatic classification, we can examine Shaw's classification of

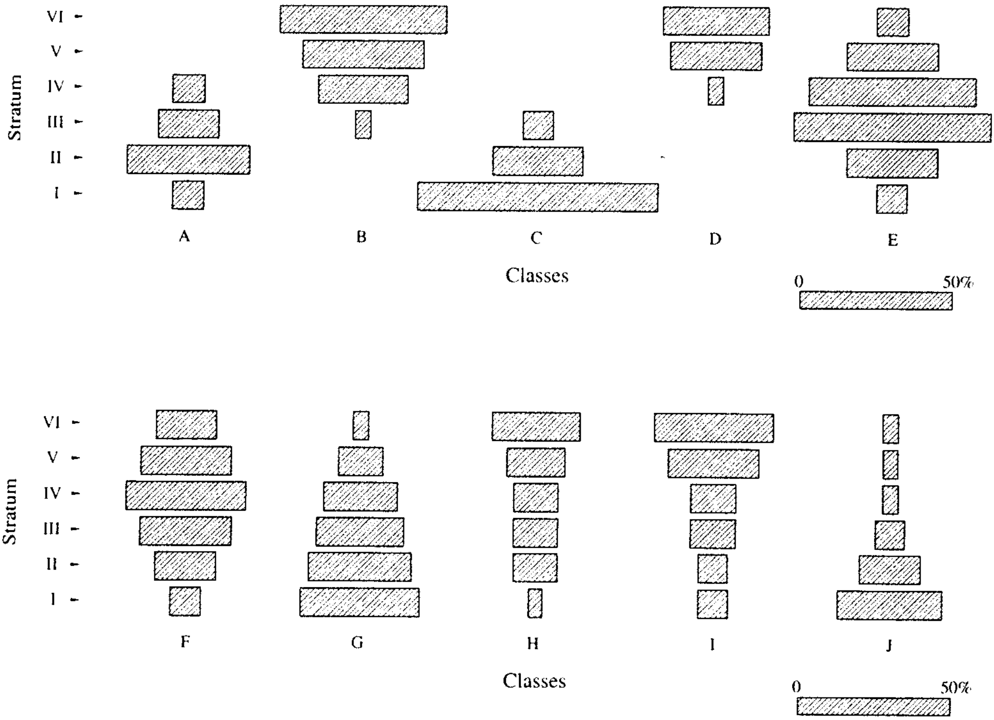


Figure 9.1. Examples of centered-bar graphs produced by percentage stratigraphy. There are six assemblages, one from each stratum (I–VI), in both the upper and lower examples. In this fictional example, change is faster and turnover more rapid among classes in the upper graph than among classes in the lower graph.

conodonts—toothlike phosphatic structures from now-extinct small, eel-like organisms. Shaw specified four characters, two of which included two states and two of which included three states (Table 9.1). Multiplication of the number of states per character (2 × 2 × 3 × 3) indicates 36 classes exist within Shaw's classification. For example, the class LFNT is a conodont with a lobed outer platform, outer platform cross section flat, nonparallel blade and ridge, and triangular inner platform profile. In paradigmatic classification it is unnecessary for all classes to have empirical members; empty morphospace is just as important analytically as filled morphospace is. How might such units be used analytically to monitor variation?

Archaeology has nearly a century-long history of graphing culture change (Lyman et al. 1998), rather consistently graphing it as change over time in terms of the relative frequencies of specimens within kinds, usually specified as *types*. Paleontology has a similar history. One graphic technique for monitoring change within lineages involves construction of a percentage-stratigraphy graph (Lyman et al. 1997, 1998), which comprises a set of columns of horizontal bars of various widths centered and stacked one atop another. Bar width signifies the relative abundance of a class within a set of stratigraphically associated specimens comprising multiple classes. Each row of bars represents a spatio-temporally unique assemblage, and each column represents a distinct class. The classes comprise variants of a more general category. Relative frequencies of classes are plotted against their stratigraphic proveniences. Character polarity (temporal order of character or character-state appearance) is dictated by superposition, with time passing from bottom to top along the vertical axis of

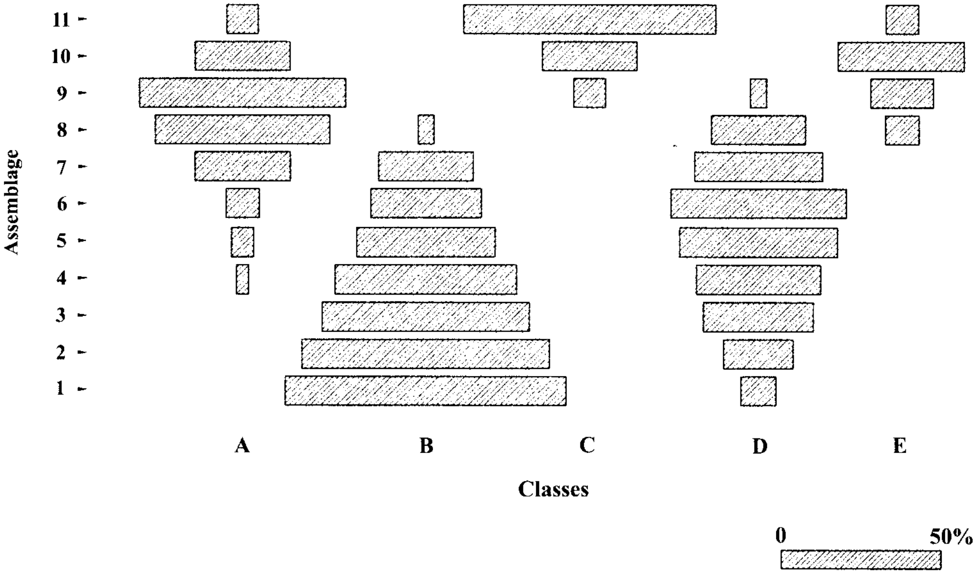


Figure 9.2. Hypothetical frequency seriation of 11 artifact assemblages using 5 artifact classes. Assemblages are ordered on the basis of artifact-class percentages, with bars summing to 100 percent for each assemblage. Only relative chronological ordering can be achieved through frequency seriation; further, time can run in either direction through the ordered assemblages.

the graph. What is shown in a percentage-stratigraphy graph is the history of the relative frequency of each of several classes of a particular category (e.g., conodonts, projectile points) over time.

Two examples of percentage-stratigraphy graphs (based on fictional data) employing classes are shown in Figure 9.1. As implied in this figure, relative rates of change can be monitored. Assuming the same strata are shown in both graphs, classes A–E (say, members of one taxonomic family) in the upper graph turn over much more rapidly than classes F–J (members of a second family) in the lower graph. Relative frequencies of classes within a family are plotted in both; values in each row of each graph sum to 100 percent. Say classes A–E represent classes of conodont morphology, as shown in Table 9.2. Because each class comprises an explicitly specified unique combination of character states, multiple classes may include the same character state, allowing the analyst to determine which classes have affinities on the basis of shared character states. In Table 9.2 all classes share the same state (N) of the third character (blade and ridge plan), classes 1 and 2 share three (of four) character states (L, F, N), classes 3 and 4 share three character states (U, F, N), classes 1 and 5 share two character states (N, T), and classes 3 and 5 share three character states (U,N,T). Changes (or lack thereof) in the allometric relations of various character states are revealed. For example, outer platform lobe development (character 1) and outer platform shape (character 2) do not appear to display any such relation.

The rate of change from one character state to another within a character can also be monitored, and, importantly, because each class is multidimensional, mosaic evolution can be monitored (Table 9.3). Blade and ridge plan (character 3) remains stable across the five taxa (it always is nonparallel [N in Table 9.1]) and thus is not shown in Table 9.3. The inner platform profile (character 4) changes from triangular (T) to rounded (R)

Table 9.2. Relative Frequencies of Classes and Character States Plotted in Upper Graph of Figure 9.1

	Classes[a]				
Stratum	1 (LFNT)	2 (LFNR)	3 (UFNT)	4 (UFNR)	5 (UANT)
VI	0	55	0	35	10
V	0	40	0	30	30
IV	10	30	0	5	55
III	20	5	10	0	65
II	40	0	30	0	30
I	10	0	80	0	10

[a] Class definitions from Table 9.1.

Table 9.3. Relative Frequencies of Character States for Conodont Characters 1, 2, and 4

	Character 1		Character 2		Character 4	
Stratum	L	U	F	A	T	R
VI	55	45	90	10	10	90
V	40	60	70	30	30	70
IV	40	60	45	55	65	35
III	25	75	35	65	95	5
II	40	60	70	30	100	0
I	10	90	90	10	100	0

rather rapidly as we move up through the column. The flat (F) cross-sectional shape of the outer platform (character 2) decreases and then increases in relative frequency, and the development of the outer platform (character 1) changes from a higher unlobed relative frequency in the lower strata to a higher lobed relative frequency in stratum VI. Deciding which class or combination of classes might represent a species or higher taxon, if such a result is desired, can be accomplished by statistical determination of which character-state combinations (classes) occur more frequently than random chance allows or by stasis of character-state combinations over time. The same procedure, of course, can be used for artifacts.

As important as percentage-stratigraphy graphs are for monitoring change, how can we be sure that that change is between specimens that are directly related? In other words, how can we be sure that we are seeing heritable continuity as opposed to merely historical continuity? The answer is, we can't. We might be more sure if the pattern of change were replicated at numerous locales, but all such replicability confirms is the temporal sequence, not an evolutionary sequence. One means of searching for the latter is by monitoring change through frequency seriation (Fig. 9.2), invented by Kroeber (1916) to examine culture change around Zuñi Pueblo, New Mexico. Paleobiologists who have studied evolutionary change in lineages of organisms have taken notice of these graphs (e.g., Gould et al. 1987; Uhen 1996), although it is clear they do not fully grasp the epistemological differences between them and similar graphs they themselves have constructed. Frequency seriation provides one means of building hypotheses of heritable continuity precisely because it measures transmission (Lipo 2001; Lipo and Madsen 2001; Lipo et al. 1997; Lyman 2001; Neiman 1995; O'Brien and Lyman 1999b, 2000b; Teltser 1995b).

If frequency seriation in fact measures transmission, then the frequency distribution of each class over time will display a unimodal curve (Lipo and Madsen 2001; Lipo et al. 1997). *That* is the implication of frequency seriation as a test of heritable continuity (O'Brien and Lyman 2000b). The use of classes of artifacts representing what are referred to as "historical types" ensures heritable continuity at the scale of classes of artifacts because items are definitionally identical. The use of multiple classes of the same category of artifacts—projectile points, say—and their "overlapping," or occurrence in multiple assemblages, insures heritable continuity at the scale of the tradition of the artifact category (Lyman and O'Brien 2000a).

The phylogenetic implications of the hierarchical structure of the Linnaean taxonomy in biology are transferable to a similar hierarchical alignment of historical types of artifacts. For example, "pottery" can be aligned with a biological family, "types" of pottery with biological genera, and "varieties" of pottery with biological species. This is more or less what Gladwin (Gladwin and Gladwin 1930, 1934) and Colton and Hargrave (1937) proposed when they published major statements on the phylogenetic implications of pottery types. Gladwin (1936) shortly thereafter abandoned his scheme and mostly escaped criticism, but Colton and Hargrave did not (Lyman et al. 1997). Although Colton's knowledge of Darwinian evolution—he was a biologist by training—no doubt underpinned their scheme, the key to it was in the supposition that related forms were related because they were similar. As we discuss below, it should have been the other way around. The problem was that Colton and Hargrave offered no explicit theoretical argument and no empirical tests for their belief that the similarities were of the homologous sort. No one else at the time could offer an argument either, and Colton and Hargrave were blasted by archaeologists such as Brew (1946), who were not amused with their overtly biological scheme. Frequency seriation comprises a technique for testing hypotheses of relatedness.

A FURTHER WORD ON UNITS: HOMOLOGY AND ANALOGY

Evolutionary archaeologists have argued that, according to Darwinian theory, two kinds of units must be constructed so as to allow the measurement of two kinds of variation among artifacts (Dunnell 1978a; Lyman 2001; Lyman and O'Brien 1998; O'Brien and Holland 1990). Units that measure stylistic (adaptively neutral) variants allow the detection of transmission (e.g., Lipo and Madsen 2001; Lipo et al. 1997), a process that must be monitored to ensure that heritable continuity is being measured. As well, units that measure functional, or adaptive, variants must be constructed, as these contribute significantly to the writing and explanation of evolutionary history as it is driven by natural selection (e.g., O'Brien et al. 1994). The distribution of stylistic variants, because they are adaptively neutral and thus do not affect fitness, should, theoretically, oftentimes be different from the distribution of functional forms over time and space. Stylistic features measure interaction, transmission, and inheritance, whereas functional features will *sometimes* measure transmission as mediated by natural selection and other times measure adaptational difference alone (Beck 1995a; Meltzer 1981). As opposed to the continuous, unimodal frequency distribution exhibited by stylistic characters, functional characters can display one of several distributions. They might display a sharp rise in popularity followed by a steep decline (O'Brien and Holland 1990, 1992) as they are quickly replaced by other functional characters; they might display unimodal frequency distributions similar to those of stylistic characters; or they might display discontinuous, multimodal frequency distributions as a result of convergence or fluctuation in the selective environment. This brings us to an important point—the equation of style with homology and the equation of function with analogy. No other single issue has done more to confuse the style-function dichotomy in the archaeological literature than these seemingly straightforward equations.

In the case of style and homology, an equation is justified, but in the case of function and analogy it is not. Functional characters can be either analogous or homologous, which is why we noted above that functional forms *sometimes* measure transmission as mediated by natural selection and other times measure adaptational change alone. Homologous characters result from common ancestry. If we are sure that the characters with which we are dealing are homologous, then the phenomena exhibiting those characters are by definition related back through a common ancestor. In archaeology, the objects are related because of cultural transmission of various kinds (Boyd and Richerson 1985). But this tells us nothing about whether the homologous characters are functional or stylistic. O'Brien and Leonard (2001:5) state, "Do not be misled by the use of the term 'function' in defining what an analog is.... The key to whether a feature is homologous or analogous is strictly a matter of its history." A better way of putting it would be, "Although analogous characters can *always* be assumed to be functional, the reverse is not always true. Functional characters can be either homologous *or* analogous." In archaeology we assume that such things as decorations on pots are so complex that the probability of duplication by chance is small. If we find, say, two ceramic vessels containing identical decoration, we conclude that they are from the same tradition, or line of cultural heredity. They are homologous. We would normally not suspect that two completely independent groups of people arrived at *exactly* the same way of decorating their vessels, given the myriad possibilities available. There is no reason to suspect that we will *never* find such an example, but the more parsimonious explanation of such a phenomenon is that the vessels share a common developmental history and are from the same tradition.

Evolutionary archaeologists have emphasized the usefulness of stylistic traits for chronological purposes, making it appear as if those traits are the only kind that have such use, but this again masks the real issue: Is a character or set of characters homologous or analogous? If only stylistic characters or sets of characters (styles) can be used, then changes in, say, the hafting elements of projectile points, which we not only assume to be functional but can demonstrate empirically to be functional (Hughes 1998), would be useless as a basis for measuring the passage of time. This decidedly is not the case, as countless studies have shown (e.g., Beck 1995a, 1995b, 1998; Thomas and Bierwirth 1983; Wilhelmsen 2001).

We discuss elsewhere (Lyman 2001; O'Brien and Leonard 2001; O'Brien and Lyman 2000b) the hows and whys of constructing and testing stylistic and functional units, and a recent book on the subject (Hurt and Rakita 2001) adds substantially to the discussion. The matter, cast in slightly different terms, assumes center stage when we turn to another method of reconstructing phylogenetic histories, namely cladistics (see below). Here the focus is on first distinguishing between analogous and homologous characters, and second on distinguishing between two kinds of homologous characters: shared derived characters, or synapomorphies, and shared ancestral (primitive) characters, or symplesiomorphies. The easiest way of thinking about synapomorphies is to define them as homologous characters held in common by two or more taxa and their immediate ancestor but no other taxon. In contrast, symplesiomorphies are homologous characters held in common by an entire set of related lineages. Figure 9.3 illustrates the differences between ancestral and derived characters and how each occurs within a phylogenetic tree. If two phenomena share one or more synapomorphies, they are *by definition* phylogenetically related. The analytical challenge is to identify such characters. The distinction between homologous and analogous characters has long been noted by both processual archaeologists (Binford 1968b; Spencer 1992) and evolutionary archaeologists (Dunnell 1978a; Lyman 2001; O'Brien and Holland 1990, 1992), as it was by culture historians (Kroeber 1931; Steward 1929; Woods 1934), but only recently has interest been paid to the critical distinction between synapomorphies

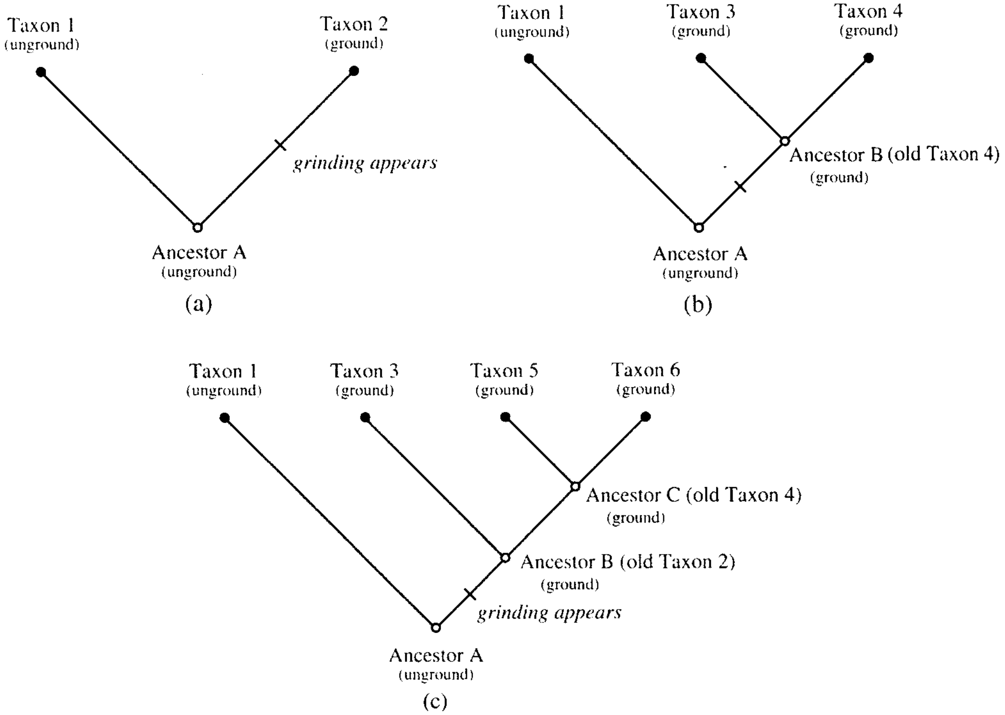


Figure 9.3. Phylogenetic trees showing the evolution of projectile-point taxa. In (a), basal grinding appears during the evolution of Taxon 2 out of its ancestral taxon. Its appearance in Taxon 2 is as an apomorphy, or derived character state. In (b), Taxon 2 has produced two taxa, 3 and 4, both of which contain basally ground specimens. The presence of grinding in those sister taxa and in their common ancestor makes grinding a synapomorphy, or *shared* derived character state. In (c), one of the taxa that appeared in the previous generation gives rise to two new taxa, 5 and 6, both of which contain ground specimens. If we focus attention *only* on those two new taxa, grinding is now a simple symplesiomorphy or *shared* ancestral (primitive) character state, because it is shared by more taxa than just sister taxa 5 and 6 and their immediate common ancestor. But if we include Taxon 3 in our focus, grinding is a synapomorphy because, following the definition, it occurs only in sister taxa, which in this case 3, 5, and 6 are, and in their immediate common ancestor, which is B (old Taxon 2).

and symplesiomorphies (Harmon et al. 2000; O'Brien and Lyman 2000b; O'Brien et al. 2001).

Identifying homologous traits is a significant analytical hurdle because a trait that is shared by two phenomena may be analogous —the result of convergence. Kroeber (1931: 151) suggested that "Where similarities are specific and structural and not merely superficial... has long been the accepted method in evolutionary and systematic biology." He was correct, for this was, and is, precisely how biologists distinguish between homologs and analogs, although such a simple statement belies the difficulties in so doing (Fisher 1994; Smith 1994; Szalay and Bock 1991). Kroeber (1931:151) pointed out that the "fundamentally different evidential value of homologous and analogous similarities for determination of historical relationship, that is, genuine systematic or genetic relationship, has long been an axiom in biological science. The distinction has been much less clearly made in anthropology, and rarely explicitly, but holds with equal force." He went on to imply that a "true homology" denoted "genetic unity," and he argued that "few biologists would doubt that sufficiently intensive

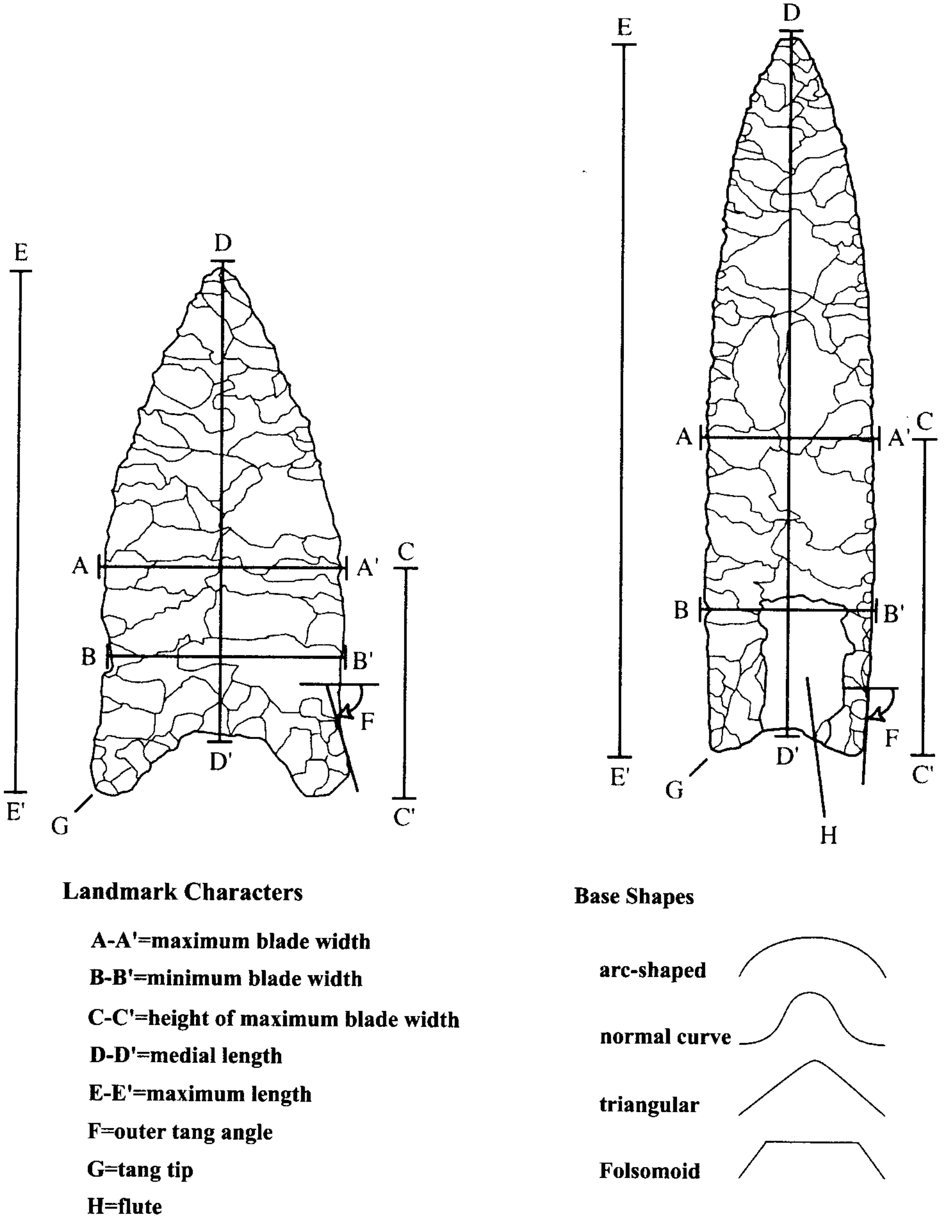


Figure 9.4. Locations of landmark characters used in the cladistical analysis of projectile points (see Table 9.4 for character states). Character states for base shape are shown at the lower right.

analysis of structure will ultimately solve such problems of descent.... There seems no reason why on the whole the same cautious optimism should not prevail in the field of culture" (Kroeber 1931:151).

Instead of implementing Kroeber's suggestions, archaeologists adopted the easily understood dictum "typological similarity is [an] indicator of cultural relatedness" (Willey 1953:363). This dictum was in fact central to the morphological-species concept of early twentieth-century biologists: Morphologically similar species were deemed members of the same taxon and thus seen as being

Table 9.4. System Used to Classify Projectile Points from the Southeastern United States

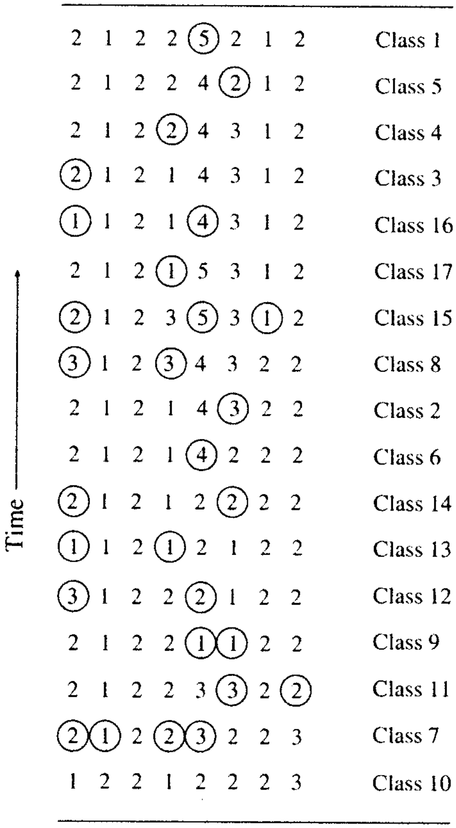
Character and Character State		Character and Character State	
I	Location of Maximum Blade Width	V.	Outer Tang Angle
1.	proximal quarter	1.	93°–115°
2.	secondmost proximal quarter	2.	88°–92°
3.	secondmost distal quarter	3.	81°–87°
4.	distal quarter	4.	66°–80°
		5.	51°–65°
		6.	≤ 50°
II.	Base Shape	VI.	Tang-Tip Shape
1.	arc-shaped	1.	pointed
2.	normal curve	2.	round
3.	triangular	3.	blunt
4.	folsomoid		
III.	Basal Indentation Ratio[a]	VII.	Fluting
1.	no basal indentation	1.	absent
2.	0.90–0.99 (shallow)	2.	present
3.	0.80–0.89 (deep)		
IV.	Constriction Ratio[b]	VIII.	Length/Width Ratio
1.	1.00	1.	1.00–1.99
2.	0.90–0.99	2.	2.00–2.99
3.	0.80–0.89	3.	3.00–3.99
4.	0.70–0.79	4.	4.00–4.99
5.	0.60–0.69	5.	5.00–5.99
6.	0.50–0.59	6.	≥ 6.00

[a] The ratio between the medial length of a specimen and its total length; the smaller the ratio, the deeper the indentation.

[b] The ratio between the minimum blade width (proximal to the point of maximum blade width) and the maximum blade width as a measure of "waistedness"; the smaller the ratio, the higher the amount of constriction.

phylogenetically related. This dictum was also axiomatic in culture history, and in both contexts it put the cart before the horse. It began to die in biology shortly after the Synthesis, and Simpson (1961:69) penned the obituary when he noted that individuals do not belong in the same taxon because they are similar; rather, they are similar *because they belong to the same taxon*. The dictum is, however, still very much alive among many modern archaeologists. They note similarities between artifact types, assemblages of artifacts, and the like, and then *assume* that those similarities are the result of cultural transmission among the individuals who created the artifacts, thus making the latter homologs. But such similarities might not be of the homologous sort required to determine phylogenetic relations and write phylogenetic histories. That is, they might be symplesiomorphies instead of synapomorphies.

Cladistics, which uses only synapomorphies in making ancestor-descendant determinations, is a powerful method, although certainly not the only one, for constructing phylogenetic histories of anything that evolves over time, including material remains found in the archaeological record. It contrasts with phenetics and other grouping methods that use any character, regardless of origin, to make determinations of affinity. To date, the major use of cladistics has been in the biological realm, but the approach is identical in logic and similar in method to tracing historical patterns of descent in languages (Platnick and Cameron 1977; Ross 1997). Recently, archaeologists have begun to explore the use of cladistics in creating phylogenetic histories of artifacts (Harmon et al. 2000; O'Brien et al. 2001). The logical basis for extending cladistics into archaeology is the same as it is in biology: Artifacts are complex systems, comprising any number of replicators, units analogous to genes (Hull 1988a, 1988b; Lyman and O'Brien 1998). The kinds of changes that occur over generations of tool production are constrained, meaning that new structures and functions almost always arise through modification of existing structures and functions as opposed to arising de novo. The history of these changes, which include additions, losses, and transformations, is recorded in the similarities and differences in the complex characteristics of related objects—that is, in objects that have common ancestors (Hennig 1950).



2	1	2	2	⑤	2	1	2	Class 1
2	1	2	2	4	②	1	2	Class 5
2	1	2	②	4	3	1	2	Class 4
②	1	2	1	4	3	1	2	Class 3
①	1	2	1	④	3	1	2	Class 16
2	1	2	①	5	3	1	2	Class 17
②	1	2	3	⑤	3	①	2	Class 15
③	1	2	③	4	3	2	2	Class 8
2	1	2	1	4	③	2	2	Class 2
2	1	2	1	④	2	2	2	Class 6
②	1	2	1	2	②	2	2	Class 14
①	1	2	①	2	1	2	2	Class 13
③	1	2	2	②	1	2	2	Class 12
2	1	2	2	①	①	2	2	Class 9
2	1	2	2	3	③	2	②	Class 11
②	①	2	②	③	2	2	3	Class 7
1	2	2	1	2	2	2	3	Class 10

Figure 9.5. Arrangement of the 17 projectile-point classes ordered by the least number of character changes needed to create the ordering (28 steps). In effect, this is an occurrence seriation of the classes. If we have the classes in correct order, we can determine the evolution of characters through time, with each circled character state denoting a change from the immediately earlier state. Evolution, of course, is primarily a branching process as opposed to an anagenetic one; thus our interest is in knowing not only which character state is ancestral to another but also the pattern of evolving characters. Cladistics offers a solution to this problem.

We used cladistics to create a phylogenetic ordering of projectile points from the southeastern United States that date roughly 9250–8500 B.C. (O'Brien et al. 2001). Projectile point form varied considerably during that period, and although much of the variation undoubtedly is temporally related, there is no agreed-on historical ordering of traditional point types. There undoubtedly was heritable continuity between at least some sequent forms, but this has never been documented. Instead, common archaeological practice is to tie changes in form to mechanisms such as diffusion and population movement. We used paradigmatic classification to create our taxa, or classes, as shown in Figure 9.4 and Table 9.4.

Classes circumvent a problem that cladistics has faced since its inception—the use of

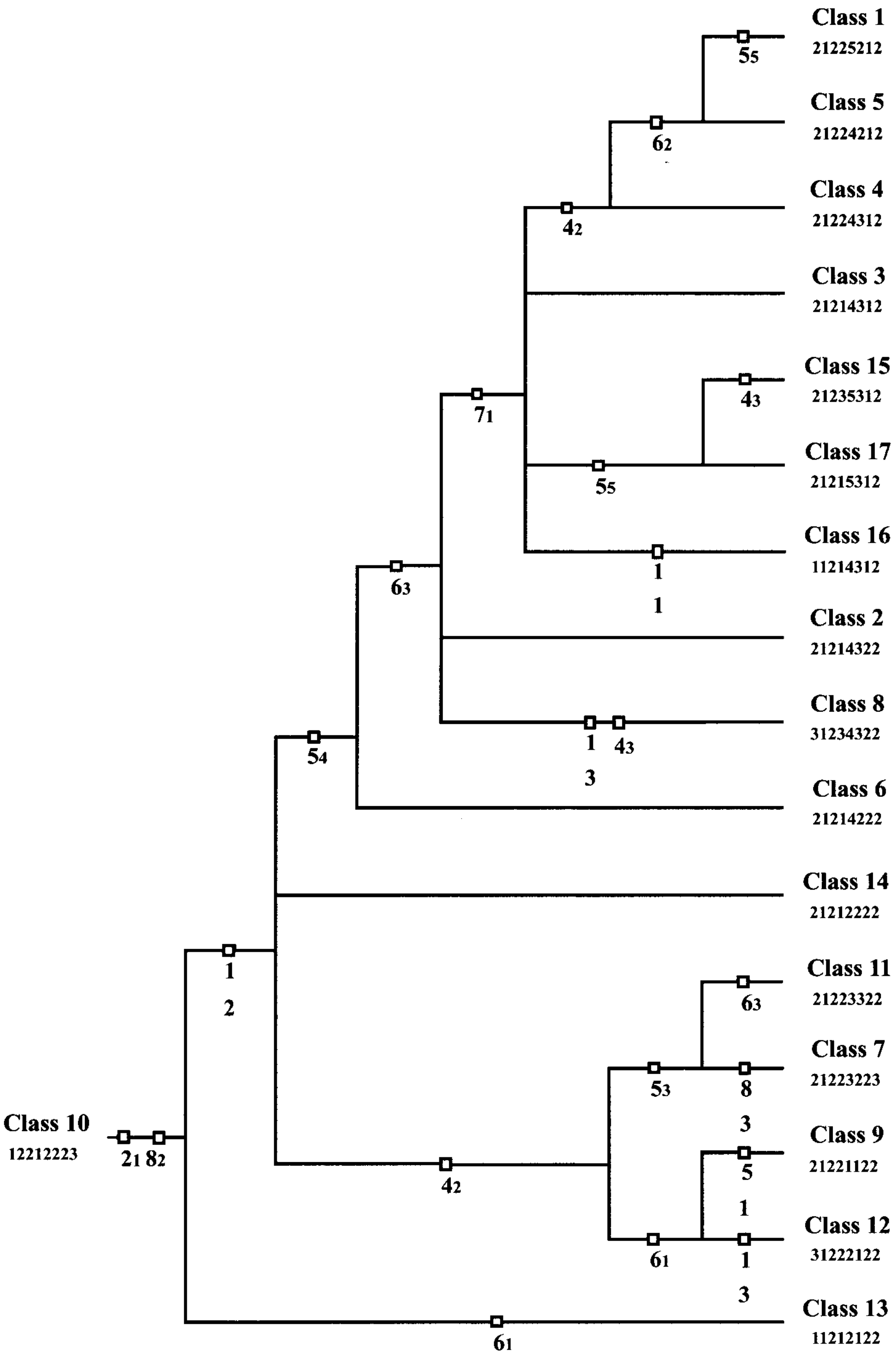


Figure 9.6. Phylogenetic tree of 17 projectile-point classes. The tree has been slightly simplified from the original (see O'Brien et al. [2001] for details). Class 10, shown here as an ancestor to all other classes, is the outgroup against which all other classes were compared. In a strict cladogram, it would be shown as a terminal taxon (located at branch tips), as are all the other taxa.

the term *transformation* or *transformation series*. Rarely is consideration given to the fact that characters do not transform from one state to another. Rather, the frequency of a particular state of a given character changes within a population relative to the frequencies of other states of that character, as shown in Table 9.3. Plotting such changes at the level of individual character offers a means of monitoring changes in character states throughout a lineage. In Figure 9.5, the 17 projectile point classes we used in the cladistical analysis are arranged in an order determined by the least number of character changes needed to create the ordering (28 steps). In effect, this is an occurrence seriation (O'Brien and Lyman 1999b, 2000b). If we have the classes in correct order, we can determine the evolution of characters through time, with each circled character state denoting a change from the immediately earlier state. But biological evolution is primarily a cladogenetic (branching) process as opposed to an anagenetic (linear) one; thus our interest is in knowing not only which character state is possibly ancestral to another but also the pattern of diversification. Cladistics, through its use of only shared *derived* characters (synapomorphies) as opposed to both shared derived and shared ancestral characters (symplesiomorphies), offers a solution to this problem.

The computer program PAUP* (version 4) (Swofford 1998) was used to generate the tree shown in Figure 9.6. That program, like other algorithms used in cladistics, uses a search routine to cull through the myriad possible trees, of which there may be thousands or even millions, to find those that require the fewest number of steps to produce, with steps defined as character-state changes. Depending on the number of classes and the number of characters, the program can produce multiple trees with the same tree length (number of steps), and it is up to the analyst to make the case for one tree over another. Most programs will also produce consensus trees in an attempt to resolve some of the disparity among the multiple trees. With respect to the 17 classes of projectile points, however, only a single tree was produced. Notice that it contains several polytomies, or points at which the program cannot make a simple dichotomous split—for example, a trichotomous branching that produces classes 8, 2, and the clade comprising classes 1 + 5 + 4 + 3 + 15 + 17 + 16. The first characters in the ancestral taxon (class 10, defined as 12212223)[1] to change were (1) character II—base shape—which changed from state 2 (normal curve) to state 1 (arc-shaped), and (2) character VIII—length/width ratio—which changed from state 3 (3.00–3.99) to state 2 (2.00–2.99). This produced (1) an ancestor (11212222) that in turn produced class 13 after an additional change in character VI from state 2 to state 1 and (2) an ancestor to all the other taxa. The latter ancestor underwent a change in character I—location of maximum blade width—from state 1 (0–0.25) to state 2 (0.26–0.50) and produced an ancestor (21212222) that in turn produced (1) class 14 (with no modification), (2) the clade comprising classes 11 + 7 + 9 + 12, and (3) the clade comprising all remaining taxa.

Homoplasy, including functional convergence, is as problematic in reconstructing phylogenetic histories of artifacts as it is in reconstructing the histories of organisms. The tree shown in Figure 9.6 has a retention index of 0.7000 and a consistency index of 0.5909 —the latter much higher than we expected based on our review of cladistical analyses in biology and paleobiology, but still low enough to indicate that considerable homoplasy is present. Homoplasy obviously is present in the occurrence seriation as well (Fig. 9.5). For example, reading from bottom to top for character I, notice the reappearance of character states 1, 2, and 3 at different points in time. The phylogenetic tree, which has a length of 22, in essence requires six fewer steps to create than the 28 required for the occurrence seriation, suggesting that cladogenesis is as important a process in artifact evolution as it is in organismic evolution.

As pointed out in more detail elsewhere (O'Brien et al. 2001), we expect several objections to the use of cladistics in archaeology. First, it might be argued that artifacts do not breed—akin to Brew's (1946:53) statement quoted above that "phylogenetic relationships do not exist between inanimate objects." This statement is false. Tools (interactors) certainly do not breed, but neither do the teeth and bones (interactors) studied by paleobiologists. But tool makers *do* breed, and they *do* pass on information (replicators) to other tool makers. Cultural transmission, both vertical and horizontal (Boyd and Richerson 1982, 1983, 1985), creates *tool traditions,* or *lineages.* Second, it might be argued that although there are tool lineages, they are impossible to discover archaeologically because of the rapidity with which cultural evolution produces variation. The tempo of cultural evolution most certainly *is* much faster than that of biological evolution, but this hardly means that we cannot see change and track its manifestations. It might also be argued that the mode of cultural evolution can be different than the modes of biological evolution—reticulate and branching, respectively—although we note that a growing body of evidence indicates reticulate evolution occurs with some regularity in the biological realm (Arnold 1997; Doolittle 1999; Endler 1998).

The claim has been made (Dewar 1995; Moore 1994a, 1994b; Terrell 1988, 2001; Terrell and Stewart 1996; Terrell et al. 1997; Welsch and Terrell 1994; Welsch et al. 1992) that diffusion results in reticulation, thus swamping all traces of phylogenetic history and reducing the cultural landscape to little more than a blur of hybrid forms. This line of reasoning is not new (e.g., Kroeber 1948; Steward 1944), and it is just as incorrect today as it was over half a century ago. We agree that some cultural evolution is reticulate, but we do not view that as being particularly problematic to archaeological analysis. It again is a matter of scale, and the same principles apply regardless of whether one is studying the phylogenetic history of fossils, of cultural practices (Mace and Pagel 1994), or of languages (Platnick and Cameron 1977). Goodenough (1997:178) makes an excellent point with respect to language: "Contact between Japan and the United States has resulted in considerable borrowing in language and culture by Japan and some reverse borrowing by the United States, but their languages and cultures retain their respectively distinct phylogenetic identities." Linguists do not guess as to whether two or more languages share a phylogenetic history. Innumerable case studies have provided the basis for deciding which linguistic characters might be apomorphic—bound morphemes and vocabulary, for example—and which might be homoplasious—syntax, for example (Nichols 1996). Thus the comparative method groups languages not on the basis of shared similarities but according to the distribution of shared innovations relative to a reconstructed protolanguage ancestral to the whole family (Ross 1997). This is nothing more than separating apomorphic from plesiomorphic characters and using the former to construct a phylogeny. The same principle applies in archaeology.

CONCLUDING REMARKS

Darwinian evolutionism is a theory about history—specifically, how and why genealogically connected things change over time. Evolutionists write historical narratives and then attempt to explain why those narratives look the way they do (Lyman and O'Brien 1998). Both steps employ concepts deeply embedded within Darwinian evolutionary theory, such as (1) lineage, which is a line of development owing its existence to heritability; (2) natural selection, which is a mechanism of change; (3) innovation, which is another source of new variants; (4) a transmission mechanism, which itself is a source of new variants; and (5) heritability. The fourth produces the fifth, and both ensure that we are examining change within a lineage rather than an instance of convergence.

Not surprisingly, evolutionary archaeology has numerous parallels to modern paleobiology. It is geared toward providing Darwinian-like explanations of the archaeological record, just as paleobiologists explain the paleontological record. Perhaps of more surprise is the fact that evolutionary archaeology is not really so different from a lot of what the discipline has been doing throughout the twentieth century (Lyman et al. 1997; O'Brien and Lyman 1998). Evolutionary archaeology *does* require a different perspective on how change is measured and on what the change means, but one does not have to proclaim an allegiance to Darwinian evolutionism in order to actually *do* evolutionary archaeology. All one need do is first, construct hereditarily based lineages of artifacts; second, monitor change over time in those lineages; and third, offer theoretically based explanations for the lineages looking the way they do. The key term is "theoretically based," meaning that the cause of change is lodged in theory as opposed to in the things being explained. Another way of looking at this is to say that cause is *external* to the things being examined.

Darwin did not understand inheritance the way we do today, but he understood the connection between inheritance and reproductive success. We prefer the term "replicative success" (Leonard and Jones 1987) rather than reproductive success, since it is portions of the human cultural phenotype—artifacts—that are evident in the archaeological record. The relation between the replicative success of artifacts, at least those that influence adaptive fitness, and the reproductive success of the organisms that bear them is clear: "The replicative success of artifact traits depends, at least in part, upon their contribution to the fitness of individuals possessing those traits as part of their phenotype but almost certainly also to the effectiveness of transmission mechanisms" (Jones et al. 1995:19).

If we define evolution as "any net directional change or any cumulative change in the characteristics of [genetically related] organisms or populations over many generations" (Endler 1986:5), then we are talking about descent with modification. There is a critical question here that must be addressed: If two things are similar in form and different in age, do they indicate that change has somehow taken place? From a modern Darwinian viewpoint, they represent change only if they are genetically—in a metaphorical sense with respect to artifacts—related, in which case the similarity of form and difference in age signify inheritance and thus continuity—an ancestor-descendant lineage. The issue of lineage construction must be explicitly considered if we ever hope to place archaeology on a sound scientific footing.

Notes

1. Technically, class 10 is referred to as *outgroup* —the taxon against which all other taxa are compared in order to determine which characters are ancestral and which are derived.