CHAPTER **8**

Phylogenetic
Considerations

In developing the principles of numerical taxonomy, we have stressed repeatedly that phylogenetic considerations can have no part in taxonomy and in the classificatory process. Once a classification has been established, however, biologists will inevitably attempt to arrive at phylogenetic deductions from the evidence at hand. It will be our task in the present chapter to examine in some detail the nature of phylogenetic evidence and deductions. We shall, of course, restrict our efforts to phylogenetic considerations as they are related to numerical taxonomy; a comprehensive treatment of the subject of phylogeny is beyond the scope of this book.

8.1. THE RELATION BETWEEN PHENETIC AND PHYLOGENETIC EVIDENCE

8.1.1. The evolutionary basis of natural taxonomic groups

That evolution is responsible for the observed phenetic diversity of living creatures cannot be doubted. No other hypothesis can account for the organic world with the same cogency. The classificatory hierarchy was used by Darwin as one of his main arguments for natural selection (see Darwin, 1891, p. 340). The evolution of species (that is, genetic species) is intimately connected with mechanisms of genetic isolation, usually arising allopatrically, commonly culminating in the inability of the isolates to exchange genes even when sympatric. At this stage in their

evolution, the isolated lineages are not necessarily natural taxa (in the sense in which this term is used by us in this book), since they may have almost identical sets of genes and may differ only in their genetic isolating mechanisms (Section 2.4.2). Many cryptic species, sibling species, and some allopatric species are presumably of this kind.

The genomes of the taxa become more different with time owing to the accumulation of genetic differences as the result of mutation and selection. As evolution progresses, the possible evolutionary patterns become further restricted. This is because the genotypes are highly integrated into a functional whole, and most of the possible changes would lead to poorer adaptation. For example, whales have not evolved into sharks, nor are they ever likely to do so, since changes from lung-breathing to gill-breathing (and in thousands of other functional systems) would disrupt the functional organization of whales. Even small changes in the direction of similarity to sharks would be disadvantageous; large changes would occur with vanishingly small probability. Reversal of evolution to any marked degree is therefore impossible (Dollo's law). There is consequently a constant trend toward evolutionary divergence, or increased phenetic difference. This trend is likely to be reversed only rarely and for a relatively few characters.

Natural taxonomic groups are formed by the restriction of evolution to certain regions of the phenetic hyperspace and the accumulation with time of genetic differences. The small stepwise nature of genetic changes suggests that most characters of descendants are the same as those of their immediate ancestors. The taxa may be polythetic, for there is no assurance that any given character of an ancestor will persist in all its descendants. They are in any event operationally polythetic, since constant characters may not be among those available for our study. We do not yet know fine genetic structure sufficiently to say how frequently and to what degree the OTU's in natural taxa are genetically alike, but it seems probable that appreciable parts of the genotype are constant in all members of many natural taxa, at least in taxa of lower rank, so that these taxa are probably not fully polythetic.

Most conceivable intermediates between actual organisms are likely to be nonviable. The successful organisms will therefore possess complexes of correlated characters, and these correlations allow the recognition of distinct taxa. In addition, the constant evolutionary divergence leads to overall phenetic divergence, though at different rates in different lineages. Hence forms that are distantly related phylogenetically will in general be phenetically distant, and vice versa. However, these two fac-

tors do not of themselves lead to the formation of the phenetic hierarchy. They might lead to the production of a large number of distinct forms all more or less equidistant from each other phenetically. Such a "cloud" of points could not be satisfactorily divided hierarchically on phenetic grounds (though it could form a cladistic hierarchy—see Section 8.1.2). To make a well-defined phenetic hierarchy, the forms must be clustered into compact groups, and these groups clustered into higher groups, and so on (see Section 7.1). The fact that the natural classification clearly shows this phenomenon (though its quantitative study will be a main field of inquiry for numerical taxonomy, as discussed in Section 8.2.3) implies more than the simple fact of evolution. The phenetic clustering may have arisen because some lineages die out; if so, some parts of phenetic space will remain unoccupied because they could only have been occupied by descendants of the extinct lineages, owing to the stringent requirements of evolutionary pathways mentioned above. If sharks became extinct today, the same sharks would never evolve again, either from ancestral groups (which are in any event also extinct) or from teleosts, crocodiles, or whales. Nor would *any* shark evolve again; that is, no organism would arise which we would classify with living sharks on phenetic grounds. Such apparent repetitions of evolution amounting to allochronous overall convergence do not, we believe, occur at higher ranks. Any resemblance between, for example, sharks, whales, and ichthyosaurs would be in a minority of characters related to their aquatic habitus. Such repetitions at lower levels, such as races and subspecies, are, however, more likely. Overall convergence or parallelism of marked degree is very unlikely. This is well expressed by Stebbins (1950, p. 542):

> Species differences are based largely on systems of multiple factors, which are built up by the occurrence and establishment of large numbers of genetically independent mutations. Hence, the probability that two isolated populations will evolve in exactly the same way in all of their characteristics is astronomically low, and the convergence in every respect of previously dissimilar organisms is even less probable.

Phenetic clustering may also have arisen because of different rates of evolution, especially if a lineage evolves rapidly without branching and then breaks up into a cluster of more slowly evolving lines (see Figure 8-1).

The phenetic gaps between contemporaneous forms are closed, as one goes back in time, by the succession of phenotypes leading from their common ancestor, but this point should not be misunderstood: there is
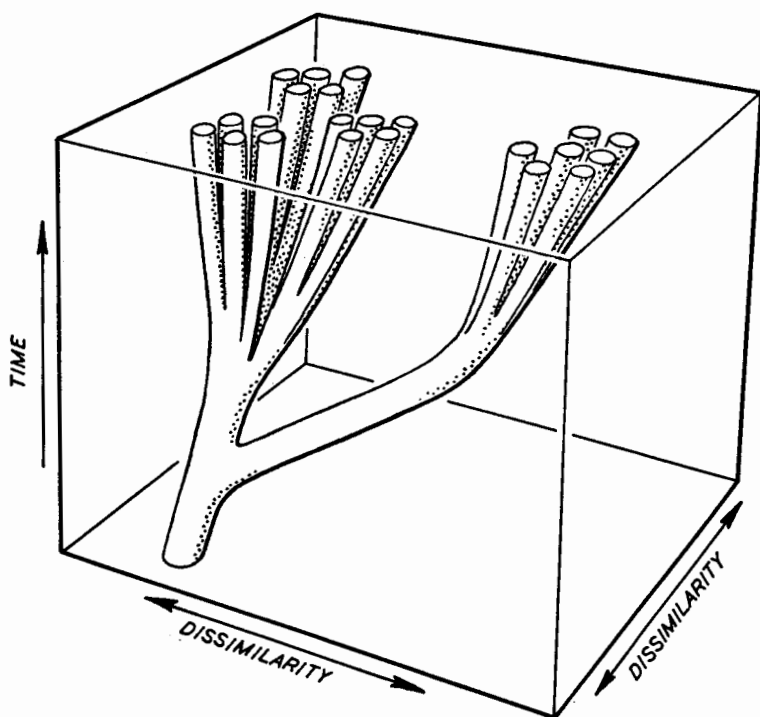
F I G U R E   8 - 1

*Phenetic clustering due to changes in rate of evolution. The form on the right at first
evolved rapidly and became phenetically very dissimilar from the remaining stocks; subse-
quently it broke into a cluster of slowly evolving lineages.*

no continuity between the phenotypes of the recent forms themselves.
For example, the phenotypes of mammals and birds merge only through
their lineages leading back to the early reptiles—that is, through a series
of phenotypes by no means intermediate between existing birds and
mammals. Chimeras, such as intermediates between mouse and sparrow,
with an equal number of features of both, do not occur and never have
occurred. The "intermediates" were reptiles, which were not pheneti-
cally intermediate between mice and sparrows.

Simpson (1961, p. 222) points out that most taxa arise as the result of
the invasion of a new ecological niche. This is likely to be a main cause
of a lineage's evolving rapidly for a time and hence yielding a new phe-
netic group. Though there may be some correspondence between the
phenetic differentiation of such taxa and the degree of ecological differ-
ence between the niches, this correspondence is by no means strict. To

explore these pathways of phyletics would require a way of measuring the ecological differences, which we do not at present have. Indeed, the concept of ecological niches has been distinctly vague; for example, a given habitat may be subdivided into a great many smaller habitats, using rather uncertain criteria, which turn in part on the phenotypes of the organisms themselves, since the organisms can change the habitat. Recent suggestions by Hutchinson (1957) to formulate ecological niches within the framework of a multidimensional system of environmental requirements may provide opportunities for fruitful development of an ecology-related phenetics. The concept of patterns in phenetic space evolving to fit configurations of ecological space suggests itself very readily.

8.1.2. The construction of phyletic classifications

Phenetic and phyletic classifications cannot be profitably compared until the meaning of phenetic and phyletic relationships has been made clear. Hennig (1957) and Cain and Harrison (1960b) have recently discussed this in some detail. Cain and Harrison divide relationship into phenetic and phyletic, and refer to phyletic relationship as "that which aims to show the course of evolution." They divide phyletic relationship into two kinds. First, two forms may be said to be closely related phyletically because they possess many characters which are derived from a common ancestor. They call the component of phyletic affinity which is due to such common ancestry (and not to convergence) *patristic* affinity. Second, the forms may be related closely through recency of common ancestry, without taking account of the number of characters derived from a common ancestor. This relationship in terms of phyletic lines is called by them *cladistic* (this is the type I phylogenetic relationship of Hennig, 1957). Cladistic relationship refers to the paths of the ancestral lineages and therefore describes the sequence of branching of the ancestral lines; it ignores evolutionary rates and therefore does not relate to phenetic similarity. This sequence of branching may be expressed by itself or in relation to time or number of generations.

While patristic affinity is a component of phyletic affinity, it is clear that our knowledge of it is likely to come from study of the phenetic affinities of the taxa concerned. In fact, patristic similarity comprises the major part of phenetic similarity, the other part being convergent similarity.

Since patristic and convergent relationships are both forms of phenetic

relationship, they can be measured by numerical taxonomic methods. Patristic relationship may be estimated by those characters which have the same character state in the ancestor and in both descendants. On the other hand, by using only those characters which have been independently acquired by the descendants, one may estimate *convergent* or *homoplastic* similarity, to use the term suggested by Lankester to distinguish independently acquired characters from those derived from a common ancestor (homologous ones in the usual sense; see Haas and Simpson, 1946). Patristic similarity is based upon homologous characters only. It is therefore in general less than total phenetic similarity, since it consists of the latter minus the convergent (homoplastic) component.

It is evident that cladistic affinity has nothing to do with the above types of relationship. They are phenetic, and measure phenotypic similarities between the forms. Cladistic affinity is concerned simply with the way in which the lineages branched and not with the degree of difference. A genealogy or pedigree is an example of a cladistic scheme.

The relation between phenetic, patristic, and convergent affinity is shown below in an example, using two-state characters for simplicity and the simple matching coefficient S_{SM}. The characters may for convenience be grouped into classes responding identically for the three OTU's.

Classes of Characters	Number of Characters in Each Class	OTU's		
		Ancestor	Descendants	
		1	2	3
A	*a*	+	+	+
B	*b*	−	+	+
C	*c*	+	−	+
D	*d*	−	−	+
E	*e*	+	+	−
F	*f*	−	+	−
G	*g*	+	−	−
H	*h*	−	−	−

The total number of characters is n.

Phenetic affinity between OTU's 2 and 3 is $(a + b + g + h)/n$, or the number of characters in classes **A**, **B**, **G** and **H** divided by the total number of characters, n. Patristic affinity between these two is given by

the proportion of characters in classes **A** and **H**, convergent affinity by those in classes **B** and **G**; that is, by $(a + h)/n$ and $(b + g)/n$, respectively. It will be noticed that we could also estimate patristic affinity by $(a + h)/[n - (b + g)]$ and convergent affinity by $(b + g)/[n - (a + h)]$, depending on whether we completely disregarded convergent or homologous characters, a point which emphasizes the lack of precision in these apparently simple concepts. Cladistic affinity is expresssed not by the characters, but by the ancestor-descendant scheme—that is, the fact that 2 and 3 are both descendants of 1, whatever their phenotypes may be, and by any data on the time relations between them. We may summarize the above by the following statements about relationships:

Phenetic Relationship = Homologous (patristic)
                        + Homoplastic (convergent) Relationships

Phyletic Relationship presents two aspects: (1) Patristic Relationship
                                              (2) Cladistic Relationship

The two aspects of phyletic relationship cannot be considered to be additive, as are the components of phenetic relationship, since they would have to be measured in different units. However, it is obvious that the close general correspondence between phenetic and phyletic relationships is caused by the almost invariable consideration of patristic affinity in judging phyletic affinity.

We do not know of any infallible criteria for overall phenetic convergence that may be obtained from a study of living forms of organisms alone. To detect convergence, we have to distinguish those features which do accurately reflect the phylogeny from those features which do not. This, however, is a question which can only be answered by knowing the phylogeny first. The problem therefore is insoluble within this logical framework, and one must have independent evidence (not derived from phenetic relations) in order to attack it. It can be seen from the example above that to distinguish patristic from convergent affinity we must be able to score the ancestor (OTU 1) for the characters we are studying; obviously we cannot score them if we only know the features of the descendants (OTU's 2 and 3). Cain and Harrison (1960b), in recommending the separation of convergent from patristic affinity, seem to us to gloss over the difficulties of deciding upon the status of the characters in question. They recommend that one should remove all enigmatic characters (those whose status is uncertain, unless it be so complex that this complexity suggests monophyly), to remove all but one of any com-

plex of characters which are of necessity functionally or ecologically related, and to remove the characters which are convergent or are likely to be convergent. If this could be easily and accurately done, it would indeed give an estimate of patristic affinity. We question, however, the feasibility of this procedure. Cain and Harrison also suggest that to obtain cladistic information one should now discard the characters which give only patristic affinity, but it is not clear to us just what information would be left, since cladistic relationship does not bear any necessary relation to the phenotype.

Simpson (1961, pp. 132–135 and 187–201) includes phylogenetic considerations as criteria for higher ranks. He uses this term for the ranks of genus and above—that is, for taxa of two or more contemporaneous species or two or more successional species (successive stages in an evolving lineage). From the criteria used by him (pp. 165, 189) the contemporaneous are separate lineages composed of entities distinguished on either a genetic or phenetic basis, while the successional are of necessity phenetic entities. He would employ criteria based on diversity (defined by him as the sum of dissimilarities in all the characteristics studied, and presumably equivalent to phenetic difference), provided it was consistent with the phylogeny. More specifically, he would, given consistency with phylogeny, admit the degree of similarity in homologous and to some extent in parallel characters, but not in others, as a secondary criterion for establishing higher categories. We have pointed out in many places in this book that there is no way of distinguishing homologous from homoplastic characters without knowing both the phylogeny (the branching of the lineages) and the phenotype of the ancestors. After mentioning that phenetic difference and recency of ancestry are usually proportional (in which case we would, of course, have no problem), Simpson (1961, p. 194) continues: "Having measured or estimated degrees of resemblance and sizes of gaps, the taxonomist should then bring into the picture qualitative judgment of homology, convergence and the like." This again assumes that the phylogeny is already known. He says that higher taxa should be monophyletic, at least in the broader definition of monophyly, which we have shown in Section 5.5.1 to be an unsatisfactory concept. Furthermore, we cannot deduce monophyly in contemporaneous forms except from their phenetic relations, since homologies as such are unrecognizable. Even fossil evidence is interpreted phenetically to yield the presumed phyletic lineages. The phylogenist cannot deny that phenetic similarity is generally a good indication of cladistic relationship, since he must use this principle to construct phylo-

genetic trees, though the time sequence of the fossils gives him assistance in this. The common occurrence of allopolyploidy in plants (Stebbins, 1950, p. 365) also makes monophyly difficult to apply as a general taxonomic criterion. Geographical data may sometimes give evidence that certain groups are monophyletic or polyphyletic, but again such evidence is available in only a minority of instances. Introduction of Huxley's concept of grades (Huxley, 1958) seems to us not to be helpful, as this is a somewhat vague idea of morphological or ecological advance, which need not reflect the whole phenotype and which we do not know how to measure.

Remane (1956) believes that the natural system is based upon the recognition of true homologies, though he does not use this term strictly in the phylogenetic sense. His main criteria for homology are morphological correspondence (including very detailed correspondences) and the occurrence of chains of intermediate forms of an organ in phenetically similar taxa. Other criteria are the greater constancy of homologous characters in taxonomic groups and the increased likelihood that simple structures will be homologous as we descend the taxonomic ranks. However, apart from morphological correspondence, which is measured phenetically, all these criteria require the prior recognition of taxonomic groups and therefore cannot be used for constructing such groups. His suggested method for recognizing cladistic relationships is to employ homologies. Remane's methods therefore cannot separate phenetic from cladistic relationship.

Hennig (1950, 1957) notes that the central problem is to distinguish cladistic relations from phenetic ones (although he does not employ this terminology). He argues that incongruence between larval and adult classifications shows that the cladistic relationships cannot be exactly proportional to the phenetic relationships, since the cladistic relationships of adults and larvae of the same species must be identical, while the phenetic relationships need not be. This statement is true, but the converse is not necessarily true; that is, congruence does not prove the exact correspondence of phenetic and cladistic relationships because both larval and adult features might have both undergone convergence to an equal degree, although this is unlikely.

Hennig (1957) bases the recognition of monophyletic taxa and the distinction between phenetic and cladistic relations very largely on recognition of original and derived states of characters. By implication these may be equated to ancestral and descendent states. He does not, however, give convincing arguments on how to recognize them (Figure 8-2).
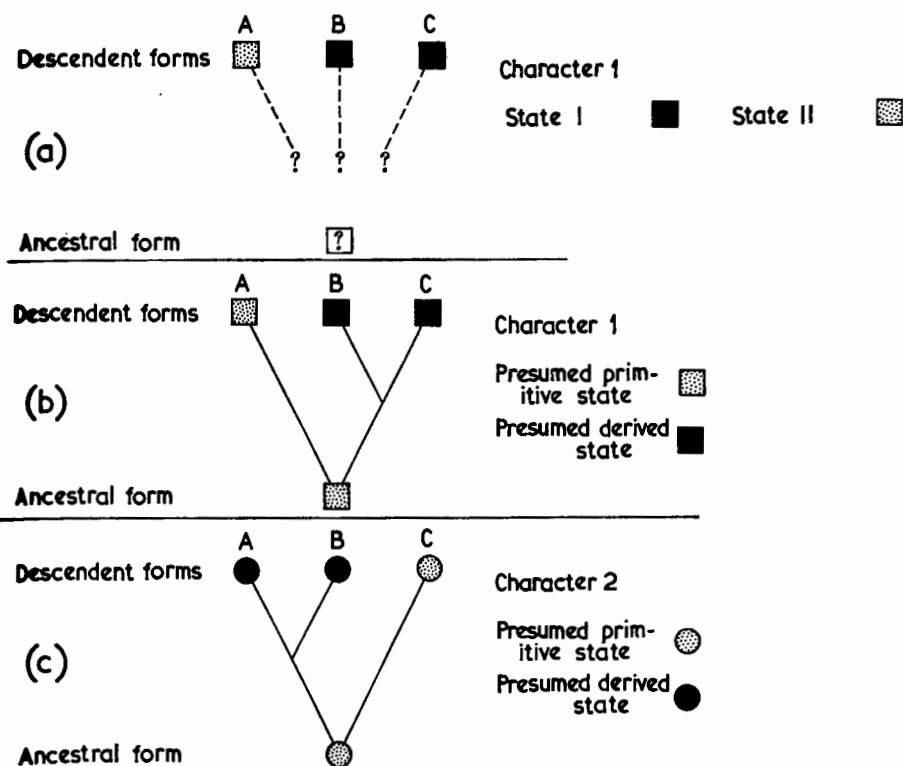
FIGURE 8-2

A diagrammatic critical reexamination of Hennig's (1957) criteria for cladistic relationships. (a) Two states of the first character are found, but it is not known which is the ancestral state. (b) State II of character 1 has been recognized as the ancestral one; forms B and C show the derived state. The cladistic relationships are then deduced as shown. (c) In a second character it is A and B which show what is believed to be the derived state. This would imply a different cladistic relationship from (b).

Even if we were able to recognize the ancestral state, shown as uncertain in Figure 8-2(a), we would be unable to accept the method he uses thereafter. Hennig would say that two forms possessing a derived (non-primitive) state were more closely related cladistically, as shown in Figure 8-2(b). However, it is common on examining another character, 2, to find that a different pair of forms possess the derived state—Figure 8-2(c)—giving a different cladistic scheme. To change our opinion about which state of character 2 is primitive means that our end result depends on which character we happened to study first. It is not clear to us how Hennig would handle contradictory data of this kind. If one counted the

characters indicating one phyletic arrangement and those indicating another and took the verdict of the majority, this would be in effect employing phenetic affinities to indicate phyletic relations.

Where fossils are few, the practical difficulty of making classifications based on phylogeny is acute. Stresemann (1959), in an illuminating article, points out that there has been little advance since 1888 in the classification of orders of birds, and the prospects for an improvement are not bright. This is despite the fact that there is little argument about the genera, which are presumably based mainly on phenetic criteria, since there can be even less phyletic evidence at this taxonomic level. The logical conclusion, that orders as well as genera should be phenetic taxa, is presumably unwelcome. Even when fossils are abundant the difficulty remains. For instance, Pichi-Sermolli (1959) concluded that the deplorable confusion in the classification of vascular cryptogams is chiefly due to the fact that the classification of ferns is based upon fossil plants rather than living ones. A similar problem exists with higher vertebrates, as Bigelow (1956) pointed out. Crocodiles are thought to be more closely related by ancestry (cladistically) to birds than to lizards or turtles, yet crocodiles are not classified with birds as the criterion of monophyly requires, but with other reptiles. It is evident that the phenetic groupings are too useful to be abandoned.

Even if we could make use of phylogeny to create classifications, we may still ask whether this is necessarily desirable. To do so would discard much important and interesting phenetic information. An allopolyploid might originate repeatedly, giving rise to phenetically identical new species each time. In some groups phyletic classifications might prove chaotic—for example in viruses and especially in bacteriophages. Recent work (see Morse et al., 1956; Jacob and Wollman, 1959; Lederberg, 1960) suggests that bacteriophages are being constantly derived from bacteria as genetic entities that acquire the properties of autonomy and parasitism. This work not only implies that bacteriophages that are identical (or almost so) may be polyphyletic, but, even more disconcerting, it suggests that they may be of composite origin. They are able to transfer genes from one bacterial form to another and to incorporate such genes into their own genomes; it is therefore possible that they derive some of their own genes from one host and some from another. Other viruses may be similar in these respects to bacteriophages, and this raises the disturbing possibility that, for example, an arthropod-borne virus may be part insect, part mammal, and part bird. The only way to bring order into such a system is by a phenetic classification. Such classifications

are now being attempted (Andrewes and Sneath, 1958; Sneath, 1962). While the phenomena discussed above do not really affect the taxonomy of most organisms, we feel that they are of considerable theoretical interest. There may also be borderline groups, such as the Rickettsiae, where the validity of a phylogenetic classification, even if possible, may be in serious doubt.

Cain and Harrison (1960b) pointed out that the principal disadvantage of phenetic classification is that some convergence and parallelisms may go unrecognized and that polyphyletic groups might be mistaken for monophyletic ones. However, the recognition of monophyly can scarcely be thought to be the only worthwhile or even the pre-eminent activity of systematists. If convergence is extreme, so that the convergent forms are practically identical, the phyletic classification only shows the fact that convergence had occurred, which could be expressed without the formal apparatus of a taxonomic and nomenclatural scheme.

8.1.3. What phylogenetic deductions can be made from a table of affinities?

Having made a phenetic classification, we will wish to know what deductions we may legitimately make about the phylogeny of the taxa. If the affinities are based on living organisms alone, we can only speculate on the phylogeny; to check our speculations we must have fossil evidence. Yet there are some conclusions which are more probable than others. We believe that these are as follows. (1) Phenetic clusters based on living organisms are more likely than not to be monophyletic. Thus phenetically adjacent taxa represent phyletic "twigs" which usually originate from the same branch; in other words, overall convergence is unlikely. (2) In the absence of direct evidence our best estimate of the attributes of a common ancestor of a cluster must be derived from the properties of the cluster itself. Whether this estimate should be based on a centroid or on a midrange value is a complicated problem (discussed in greater detail in connection with typology by Sokal, 1962b). In short, if we have no fossil evidence, the existing pattern is our best guide to the past history—though this may often be wrong. An argument similar to the first argument in favor of equal weighting of characters (Section 5.7) applies here: if we have no evidence that evolutionary rates differed, we must assume these to have been constant and equal in all the phyletic lines studied. If the reader thinks of a cross section through the top of a shrub with the vertical dimension representing time and the horizontal

representing phenetic dissimilarity, he will have a ready, though some-what inadequate, simile for the situation.

The two points mentioned can then be illustrated as follows. (1) Adjacent twigs will generally arise from the same branch. Admittedly it will be very difficult to detect overall convergence near the tips of the branches, and this may have occurred quite commonly, together with some reticulate evolution due to the fusing of phyletic lines; yet gross degrees of overall convergence between the tips of main branches will be very much less likely, and its improbability will increase with the taxonomic difference between the branches below the plane of section. (2) In the main, the branches from which the twigs arise will lie more or less directly below the twigs; but we will have no way of telling whether the twigs arose almost vertically or whether they came off at a pronounced slant, for we can have no confidence that the twigs will fill the phenetic space evenly in the way in which the branches of an actual shrub do in order to obtain adequate sunlight.

The above deductions on the phylogeny (which are made from organisms belonging to one point in time) cannot give any estimate of the rates of evolutionary change, which may have differed in different phyletic lines. To estimate rates we must have data from several points in time.

If we have a matrix of affinities which includes OTU's from different geological periods, we can attempt to measure rates of evolution (discussed in more detail in Section 8.2.1). However, we must first decide how the forms are related by ancestry (that is, cladistically) if we are to study the rate of change of one form into another. We shall sometimes have very adequate sequences of the evolutionary successions about which no reasonable doubt could arise. More often we will be faced with determining which form was the ancestor of which, or more probably (since the ancestral line itself may well be unrepresented and we may have only members of its side branches), which form was closest to the ancestral line. This must be solved on the same principles as those mentioned above: in the absence of other evidence, phenetic resemblance is the best indication of cladistic relationships.

The above guide line will hold only in a simple form where one or two geological levels are represented. When there are more than two, we will commonly be faced with the question: have rates of evolution been constant over the whole period? For example (see Figure 8-3), we may have an extant form **A** which is more similar to a slightly earlier form **B** than it is to a much earlier form **C**, though **B** and **C** are more similar. Did, then, **A** evolve recently from **B** (the more similar) or more slowly and

directly from **C** (the less similar)? In the absence of other evidence, such as geographical isolation or a fuller fossil record, no certain decision can be taken even in such a simple case. It is very easy to find cases which are more complex than the example given above. One lineage, for example, may evolve rapidly and another slowly, and it may be impossible or implausible to draw a phyletic tree in which all the lineages evolve at the same rate throughout the period. A diagram of the phenetic relations (expressed in one or two dimensions) versus time will usually make clear the degree to which we can safely reconstruct the phyletic tree, and where we must indicate by dotted and queried lines our uncertainty as to the course of the descent.

We have seen that one cannot derive evolutionary rates from similarity coefficients among recent forms. This is shown by the "pregroup-exgroup problem" discussed by Michener and Sokal (1957). Is it possible to distinguish whether an aberrant member of a cluster of forms was derived phyletically from one of the members of the cluster, or from the

FIGURE 8-3

The time-rate problem. Organism C *is the ancestor of both* A *and* B. *With only the data shown one cannot tell whether* A *arose directly from* C *or via* B.

ancestral stem below the point at which the rest of the cluster arose (see Figure 8-4)? These authors suggested that if the aberrant member, **X**, showed approximately the same affinity to all the members of the cluster it was most likely "pregroup," or derived from the common stem; see Figure 8-4(a). If, however, the affinity with one member of the cluster was much greater than the mean of the affinities with the cluster, then the aberrant member, **Y**, was likely to be "exgroup," or derived from the member it most closely resembles—**D** in Figure 8-4(b). It is nevertheless possible to account for the observed affinities in the figure by means of a number of cladistic schemes, if the evolutionary rates differ in the lineages. At the time they wrote their paper, Michener and Sokal
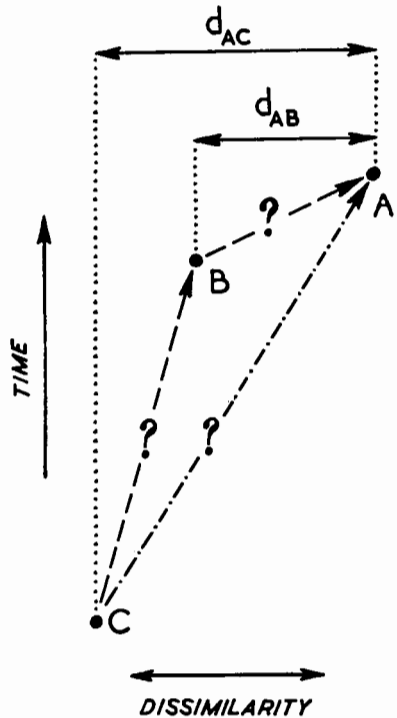
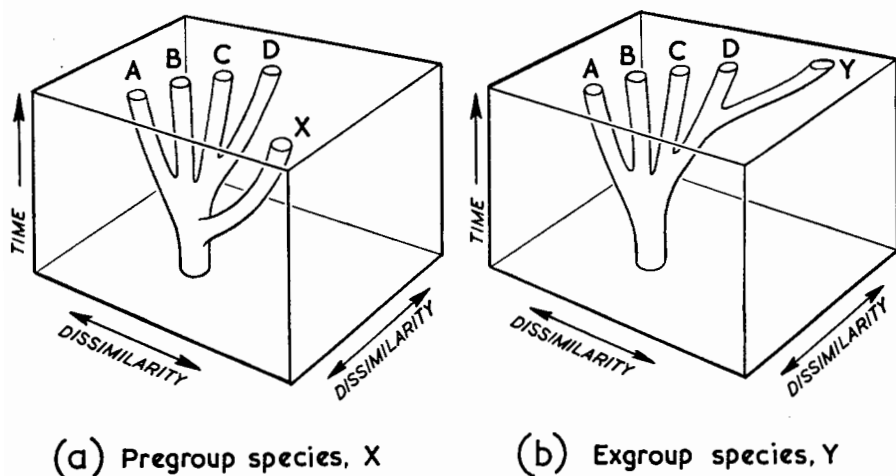(a) Pregroup species, X   (b) Exgroup species, Y

FIGURE 8-4

The pregroup-exgroup problem (for explanation, see text).

(1957) were still attempting to establish phylogenetic classifications. Only subsequently did they realize (see also Michener, 1957) that the only classifications with claims to consistency and objectivity are phenetic ones. Therefore, for classificatory purposes the pregroup-exgroup problem is of no consequence, and we would not now recommend modifying dendrograms to attach exgroup species to their presumed cladistic relatives, as was done by Michener and Sokal (1957).

8.1.4. "Primitive" characters and forms

A number of workers have proposed indices of primitiveness or advancement based on characters of living organisms. There is much confusion over the precise meaning of terms such as "primitive character" and "primitive form," as Sporne (1956) has shown. The term "primitive character" must be qualified to avoid ambiguity. Since at some sufficiently early stage the character will not yet have appeared, in practice we mean by "primitive character" a character complex in the early stages of its evolution. It is thus necessary to specify what stage we mean or the geological time of occurrence of the organism thought to possess it. It is also important to point out the taxonomic level relative to which a character is considered primitive. A thysanuran leg is a specialized arthropod appendage but a primitive insect leg.

Whether we mean by "primitive" the character state as it was when

it first appeared in the group concerned or as it was at the bifurcation of the lineages under study is seldom made clear. It is misleading to suggest that a structural or physiological simplicity is the essence of a primitive character. Evidence from a variety of groups (see, for example, Michener, 1949) points to structures appearing to be fully developed in ancestral or "primitive" forms and evolving by loss or reduction in number or size of constituent parts. Simplicity due to reduction is common in parasitism. It is also unjustified to assume that one or the other extreme of a series is necessarily primitive. It may happen that an organ evolves so as to become larger in one phyletic line and smaller in another.

Hennig (1950, pp. 172 et seq.) uses the occurrence of a given character state in a large proportion of the species of a taxon as his principal criterion for recognizing a character as ancestral. This however, implies the prior grouping of these species into the taxon. A second criterion is that agreement among members of a taxon in Bauplan of a complex structure is a good indication of monophyly, even if there are many simple characters which are inconsistent with the postulated phylogeny. However, we should point out that such a Bauplan is itself made up of a large number of characters. The decision will then depend on degree of correspondence among these characters, a procedure similar to numerical taxonomy.

The concept of the "primitive organism" is even more vague in most instances. "A primitive monocotyledon" may imply a plant either cladistically or phenetically close to the ancestor of the presumed monophyletic group *Monocotyledones*. A group is likely to be "primitive" in only some features. Furthermore, these features may be "advanced" with respect to a higher category. For example, a family of parasites might be "advanced" as a family with respect to mode of life, yet parasitism would be "primitive" within the family if some secondarily free-living species had later arisen within it. Thus the concept of primitiveness is relative.

"Overall primitiveness" would be the phenetic similarity between a given form and a specified (though possibly hypothetical) ancestor; and "overall advancement" would be the corresponding phenetic dissimilarity. Numerical taxonomy may be able to clarify some of these problems.

Frost (1930) and Sporne (1948) have developed numerical methods for estimating primitiveness. These are based on the hypothesis that, in general, features that are highly correlated in living forms are either all advanced or all primitive. If one can obtain from other sources an indication of the state in the ancestor of one of these characters, one can deduce which set of features is primitive and which advanced. Recently

Lowe (1961) has made a similar study of primitive and advanced features of monocotyledons, using Sporne's methods. Since fossil evidence is scanty, the studies of Frost on primitive features of woody tissues were employed to decide which characters were primitive. Bell (1956) has developed this idea further. He assumes that features common to large numbers of species within a genus are more likely to have been found in the ancestor of that genus than those features which occur in but few forms. The difficulty of this approach is that its validity depends on whether this assumption is correct in each instance and on the taxonomic validity of the taxa and how one decides their rank.

Another phylogenetic hypothesis is the "age and area" hypothesis of Willis (1922). This assumes that the age of a monophyletic group from the time of the common ancestor is proportional to the number of living taxa belonging to it and also is proportional to its present-day distributional area. There are so many obvious exceptions that this hypothesis is now generally discarded (see Stebbins, 1950, p. 531). The data of Willis, however, are of interest in connection with the branching pattern of dendrograms (see Section 8.2.3). A number of other generalizations—the center of origin hypothesis, trends in karotypes, Dollo's law, von Baer's law, Haeckel's biogenetic law, and others—are also very uncertain. They are well discussed by Stebbins (1950, pp. 445 ff., 448 ff., 532 ff.).

8.1.5. Can ancestral forms be included in phenetic classifications, and what is their rank?

It is plain that any division of continuous lineages is to some extent arbitrary. When two taxa ranked as classes are separated at a given line, the species on either side of this line, though both genetically and phenetically closely related, will be grouped in different classes. Remane (1956), in discussing this problem, points out that in lineages such as that in Figure 8-5 (representing the phylogeny of a family) the organisms **H** and **I** have a dual and partly contradictory relationship. They are very similar both in properties and closeness of ancestry and should therefore be placed in one genus. They are also ancestors respectively of the genus **A, B, C, D** and the genus **E, F, G**. Again, on their closeness of phenetic and phyletic relationship, they could legitimately be included in these two genera to give the genus **A, B, C, D, H** and the genus **E, F, G, I**.

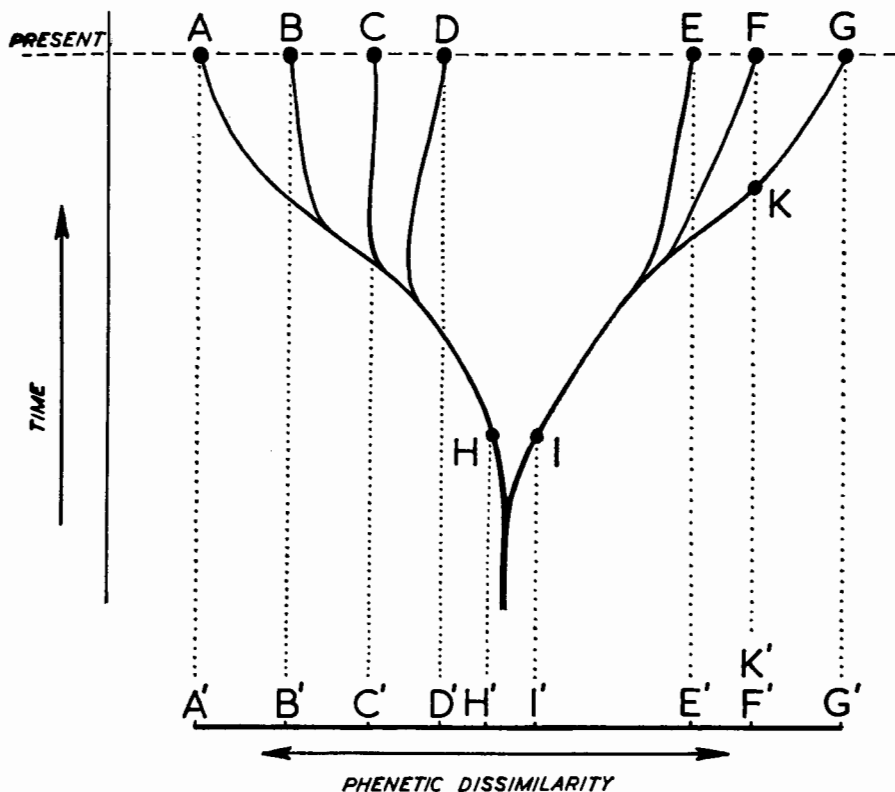This problem occurs wherever the dividing line is drawn. Some taxon-

A phylogenetic tree as commonly represented, with a time dimension and one phenetic dimension (for explanation, see text).

omists would divide the lineages to give monophyletic groups wherever possible, since this would involve the least number of such arbitrary divisions. Similarly, when we employ phenetic criteria for classification, we would divide the genera (or other OTU's) so as to give the most cohesive clusters. In general the phenetic divisions would, we believe, be mainly monophyletic, but they would be based on observation of creatures and their characters, rather than on more or less speculative phylogenies inferred from phenetic data.

One misleading point about figures, such as Figure 8-5, is that they do not represent the phenetic relations at all accurately. In Figure 8-5, for example, the organism **K** is directly below **F**; it is, however, exceedingly unlikely that it would be phenetically identical with **F**. The phenetic relations of the forms in this diagram are their relations in the
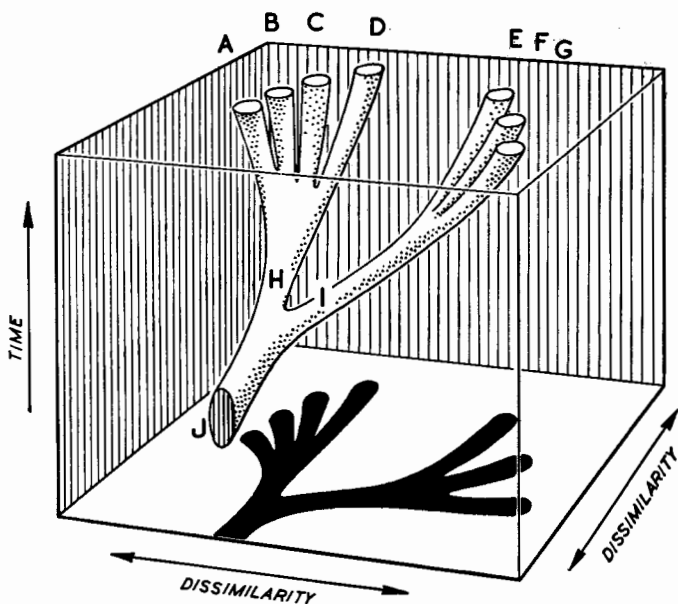
FIGURE 8-6

A phylogenetic tree in three dimensions, one of time and two of phenetic dissimilarity. The "shadow" of the tree on the base indicates the purely phenetic relationships (for explanation, see text).

horizontal plane, as would be shown by throwing a shadow of them onto the base line. This is equivalent to making projections of all the points on the dendrogram onto the abscissa, as is shown for a few points in Figure 8-5. It must be added that, of course, a single dimension is in principle insufficient for representing phenetic relationships but is employed here to convey understanding of relations which are equally applicable to multidimensional models. In this figure these projections would form a single line from the point **A′** to the point **G′**. The phenetic relations are better represented (though still inadequately) in Figure 8-6, where a three-dimensional model of a substantially similar diagram is shown. Here the horizontal plane shows phenetic dissimilarity in two dimensions, while the vertical dimension represents time, as before. The phenetic relations would then be shown by the shadow of the phylogenetic "tree" projected onto the base plate. This shadow is shown again in Figure 8-7(a). It is a fronded figure in which the organisms of all time periods are shown without overlapping one another (on this scale of taxonomic discrimination). If we wished to divide it phenetically, we

would divide it into two main groups, genera perhaps, roughly as shown. The exact place of the division line could be determined mathematically, if we wished, though for most purposes a division at the point of branching would suffice. We would, however, rarely have enough fossil data to have a complete shadow; we would be more likely to have an incomplete set of organisms, giving a shadow such as in Figure 8-7(b), and in most cases we would be lucky to get this amount of information. The exact position of the dividing line would then be not worth much argument.

We might have had some subsidiary branches (say, **L** and **M**) near the common ancestor, as shown in Figure 8-7(c). If so, one would, on phenetic grounds, divide it more or less as shown into the three genera **U**, **V**, and **W**. Note that in all cases the phenetic divisions—divisions made on the basis of the shadows—are, as we would expect, fairly close approximations of monophyletic groups or single lineages, though divergent clusters of branches may be excluded from the basal taxon. This seems to us the only honest thing to do, since we believe that phylogenies are deduced substantially from the phenetic relations. In addition, the shadows in this model are not as unsubstantial as they may seem. They represent a great many attributes of the organism. If convergence should give phenetic groupings that are not monophyletic, the groups nevertheless would contain far more information than the monophyletic groups and would be far more useful and "natural" to everyone except the phylogenist. What is likely to be troublesome, we believe, is not so much discrepancies between the phyletic and phenetic arrangements as discrepancies between taxonomies based on different sets of characters or on alternative methods of cluster analysis (see Chapters 5 and 7).

8.2. POSSIBLE CONTRIBUTIONS OF NUMERICAL TAXONOMY TO PHYLOGENETIC, GENETIC, AND PALEONTOLOGICAL PROBLEMS

In the discussion that follows, such contributions have been arranged under these headings: (1) rates of evolution; (2) studies of speciation, and the correlation of phenetic groups with genetic relations among the organisms; (3) the pattern of branching of taxonomic dendrograms; (4) paleontological problems.

In Section 8.1.2 we discussed the problems of deducing from phenetic classifications which fossil organisms are the most probable ancestors of others and of reconstructing the most probable phylogenetic lineages. If

monophyly and phenetic affinity are in fact largely overlapping if not coincident properties, affinity values will determine very largely the form of the phylogenetic trees by showing which fossil organisms are the most probable ancestors of other organisms, as judged by their overall resemblance. This will aid taxonomists in deciding between alternative choices for the ancestor of a taxon.

Although the concept of the direction of evolutionary change has been repeatedly used by students of evolution, particularly in hypotheses such as orthogenesis and parallel evolution, it has never been clearly defined or measured. If the organisms at successive periods of an evolving lineage are treated as points in a phenetic hyperspace, it is in principle possible to draw lines through this hyperspace from each organism to its descendant. The directions of these lines, their curvature and the convolutions which they display can then be evaluated by standard methods of analytical geometry. In Euclidean hyperspace a straight line from point A to C is one such that for any intermediate point on this line, B, the equation $AB + BC = AC$ holds, or, in the convention used for taxonomic distance,

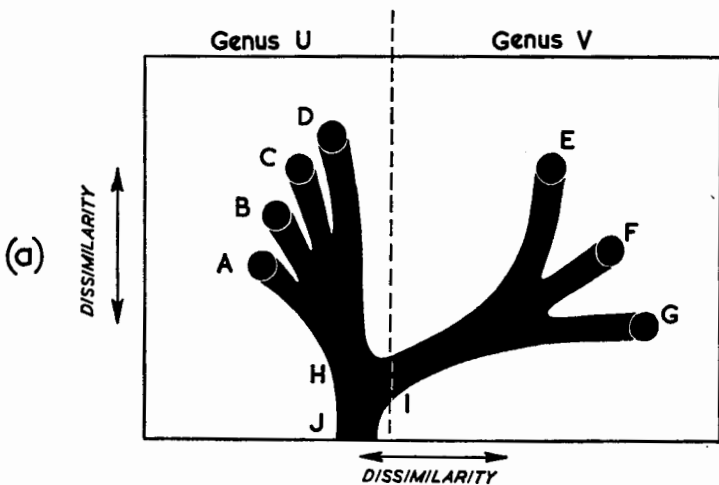$$\sqrt{d_{AB}^2} + \sqrt{d_{BC}^2} = \sqrt{d_{AC}^2}.$$

FIGURE 8-7

The "shadow" from Figure 8-6. The phenetic relationships are represented as shadows on the horizontal plane. (a) The shadow from Figure 8-6 with a dashed line dividing it into two phenetic taxa, such as two genera. (b) A patchy shadow. This is a more realistic representation because of the usual scarcity of fossils. (c) Division into three phenetic taxa (such as genera) when branches L and M have been added.
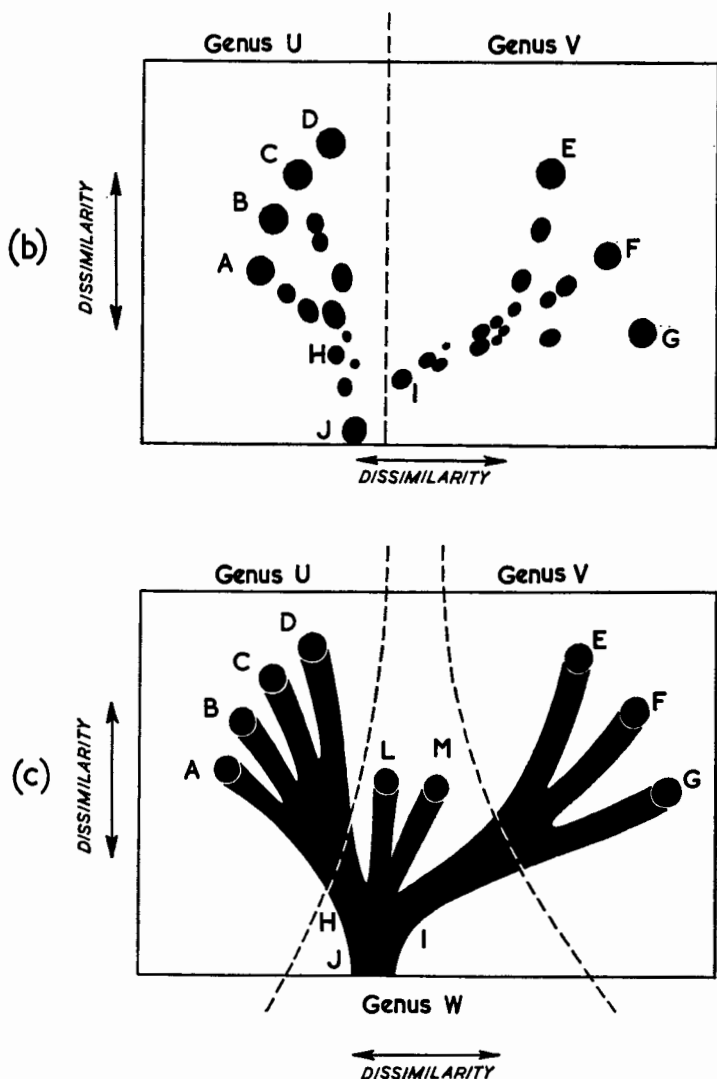
For any profitable work of this kind, a considerable number of fossil specimens from many time horizons and offering many characters for analysis will, of course, be necessary.

8.2.1. Rates of evolution

When it is possible to study fossil material and thus obtain data from several known points of time, the affinity coefficients will allow estimates of overall evolutionary rates. The dissimilarity between ancestral and

descendant forms will be the measure of the overall evolution which has occurred in the intervening period. Simpson (1944) has discussed the great advantages of measuring the overall rate of evolution (what he calls the "organism rate"), as well as the rate of evolution in one or a few characters ("character" and "character complex" rates such as those studied by Haldane, 1949; Kurtén, 1958, 1959; and Buzzati-Traverso, 1959). Numerical taxonomy therefore offers a solution to many of the problems propounded by Simpson and by Huxley (1957). Simpson (1944) uses the terms tachytelic, horotelic, and bradytelic to describe rates which are respectively rapid, moderate, and slow. Little is known about tachytelic evolution, since the changes are so rapid that there is small chance of finding fossils of the relevant period. Bradytelic evolution is the kind shown by "living fossils" such as *Lingula, Gingko, Metasequoia,* and the coelacanth, *Latimeria.*

8.2.1.1. Character rates

Haldane (1949) has suggested a measure of the rate of evolution of a single character (for example, the length of an organ), such that the unit rate, the *darwin*, corresponds to a change by a factor of *e* in one million years—that is,

$$\frac{\log_e x_t - \log_e x_0}{t} = 1,$$

when the character has the value x_0 at time 0 and x_t at time t, with t measured in *crons*. [The word *cron* (Huxley, 1957) is a convenient term for one million years.] This is approximately equivalent to a change by a factor of $1/1{,}000$ in 1,000 years. If the allometry equation $\log y = \log a + b \log x$ is used, the constant b should be used without transformation into logarithms since it is itself effectively a logarithm. In the examples studied by Haldane, the rate of change in horotelic evolution was around 0.04 darwins (40 millidarwins), but he noted that domestic animals have changed at rates of kilodarwins, so that increased selection can evidently greatly increase the usual rate of evolution.

In a recent study of fossil horses, Downs (1961) found rates of 12.3 to 124.3 millidarwins for tooth characters, their mean being about 57 millidarwins. These rates seem fairly typical of horotelic evolution, though Simpson (1953) notes that there is considerable variation in horotelic lines. The rate may, of course, vary within any one line as well. It may be that in the work to date there has been a tendency to select for measurement the more rapidly changing characters.

The darwin cannot be thought of as an absolute measure of evolution, since its value depends on the manner of scaling of the character. If a bacterium, for instance, is selected for drug resistance, the resistance may increase (by successive mutations and selection) by a factor of one thousand in the course of a few days, and the precise rate in darwins will depend on how the character is measured, whether in terms of the inhibitory concentration of the drug or, as is common, in terms of the logarithm of the concentration. The rates in microorganisms can be very great, as in this instance, where it is of the order of 10^8 darwins.

Boyden and Gemeroy (1950) have suggested that some serological characters, notably those of serum globulins, have changed at a slower rate than those of serum albumins, and quantitative data are available here. However, there are, of course, no data on the proteins of ancestral forms, and these findings could be attributed to the different antigenicity of some classes of proteins as compared with others.

8.2.1.2. *Organism rates*

When we come to a discussion of evolutionary rates in organisms, we enter a subject fraught with many pitfalls. In a simple, superficial sense it seems quite clear that different groups have evolved at different rates. For example, the statement that coelacanths have evolved more slowly than horses since the Eocene appears self-evident and is not likely to be challenged. However, when we attempt to analyze in detail the meaning of statements such as these, we run into considerable difficulty. Should we measure overall change of all the characters of each organism? Or should we only use those characters known to have changed within the forms during the periods under study? If we use only characters that change, we face a problem if we analyze groups in which only a small proportion of the characters have changed. These changes might appear quite appreciable. They might equal the changes in another lineage where most of the observed characters had changed but where the change per character was on the average the same as that of the changing characters of the first lineage. Yet we would intuitively feel that the first lineage was evolving more slowly.

We think that we can make two definite statements. First, no group should be investigated from the point of view of absolute evolutionary rates until its characters and phenetic resemblances have been investigated and evaluated with relation to the higher ranking groups to which they belong and also to neighboring taxa. If this has not been done, we cannot be sure that the taxonomic group under study is really a natural

taxon. We cannot define a taxon as natural unless all possible contenders for affiliation have been examined and included or rejected. Once such a group has been established, the relative changes within it can be measured with consistency. How such relative changes are to be compared among groups remains problematical.

Second, a paleontologist, in comparing the rates of evolution of coelacanths and horses, has in the back of his mind an idea of the range of characters within vertebrates which have to his knowledge changed and the kinds and degrees of change which have happened in all the vertebrate classes. He is evaluating the changes in the coelacanths, and in horses, against this unexpressed standard. In this connection a possible mode of evaluation may be the following: once a large group, such as the mammals, for instance, has been sufficiently studied by means of numerical taxonomy, a series of marker taxa may be chosen, with which evolutionary change can be compared. Thus a single representative of each order of mammals might be appropriately included in the matrix of affinities and serve to furnish the proper scale for a group such as the horses, for example. A fairly comprehensive standard set of taxa of this sort would comprise a reasonably stable standard of comparison.

An advantage of organism rates is that they are likely to be more steady than character rates, since bursts of rapid change in individual characters will tend to be smoothed out. Simpson (1944) has estimated organism rates (the "taxonomic rates" of Kurtén, 1959) by measuring the time for a phyletic lineage to change morphologically (phenetically) from one genus to another. From the data of Simpson (1944, 1961) and Kurtén (1959) and a consideration of the time of appearance of different taxa in vertebrate evolution, we may estimate that the time corresponding to change in rank in horotelic evolution in vertebrates is approximately as follows: morphospecies, 0.5 crons; genus, 7 crons; family, 20 crons; order, 45 crons; class, 80 crons. Myers (1960) discusses the rate of evolution of fishes after their introduction into lakes. With the exception of one lake in the Philippines (where very rapid evolution of several genera may have occurred in as little as 10,000 years), the usual pattern is the evolution of a few new species and subspecies after about 1/10 cron, many new species and some new genera after 1/2 cron, and many new genera and some new families after 1 to 2 crons. These rates are somewhat faster than those given above.

The rates appear to be much slower in some other phyla, such as many lines of molluscs. In flowering plants, Stebbins (1950, pp. 529, 547 ff.) states that there has been little change during the Tertiary period; many

genera of woody plants have evolved new species in the last 50 crons, and a few families have evolved new genera. Many genera and species of tropical trees have been almost unchanged for 20 crons. Herbaceous plants have evolved at about the same rate as trees, though there has been more rapid evolution of some groups during the Pleistocene, with new species arising within 1 cron, including some arising in historic times.

Westoll (1949) tried to estimate by a simple arithmetic technique the overall evolutionary rate in the lungfishes. Few characters were used and these were chosen to show regular trends; they were thus not representative of all the characters available, which were unfortunately rather few in number.

It should be made clear that by the phenetic change corresponding to a genus (for example), we mean the minimum phenetic difference between two forms which would just necessitate the placing of the two forms in different but closely similar genera instead of placing both in one genus (according to the criteria established by the investigator).

Kurtén (1958) suggested that the percentage of significantly differing allometric growth gradients between two populations could be used as a measure of the organism rate. He calls this the "differentiation index." The index increases as a geometric series, with the limit value of 100; for example, the steps from 0 to 50, from 50 to 75, and from 75 to 87.5 are all equivalent. It runs parallel with taxonomic (phenetic) change but has the disadvantage of not taking into account the magnitude of the differences in the gradients (except so far as the magnitudes make the differences statistically significant). In most instances in mammals the rate of change was about 0.2% per thousand years, but periods of more rapid evolution also occurred. He found that the morphological difference between two subspecies was equivalent to an index of about 50% and that between two species was about 75% (Kurtén, 1959). It is clear from the context that Kurtén here uses subspecies to indicate a major morphological subdivision of a species rather than a trivial variant, and that by species he means a category approximating a morphospecies.

We should emphasize that all the above considerations are based on conventional judgments of taxonomic rank and are only as precise as these evaluations. Lacking better ones, we cite them to give a general indication of the nature of the problem. It should be evident that the methods of numerical taxonomy will have a considerable contribution to make to this field. Present techniques in numerical taxonomy will yield affinity values between chronologically successive organisms, which can be used as measures of evolutionary rate. Over small ranges the change

in the affinity values compared with time will be satisfactory expressions of the rate. Over larger ranges this may be unsatisfactory since the affinity values may be in a scale (such as the index of Kurtén, described above) in which equal affinity differences do not have equal significance with regard to taxonomic rank at all parts of the affinity-value scale. For the present, therefore, it may be wisest in such cases to divide the affinity-value scale to define different ranks (see Section 7.5) and to express the evolutionary rate as the time taken for a phyletic line to pass through the degrees of affinity values applicable to these ranks, as Simpson suggested. Sneath (1963) has speculated on the relations between phenetic change and genetic change in evolution, particularly with respect to changes in DNA.

Should we attempt to express organism rates in darwins? In principle it would be possible to measure the mean of a large number of character rates expressed in darwins, which would approximate to measuring organism rates in darwins. We feel that the use of this term would be confusing, in view of the conceptual difference between similarity in a single character and overall phenetic similarity, and we therefore suggest that other terms should be employed when necessary.

8.2.2. Speciation and the correlation of numerical taxonomy with genetic relations

Little has yet been done with numerical taxonomic methods in studies of the formation of species and of the correspondence between genetic relationships and phenetic groupings. Ehrlich (1961c) recently made a start on the problem of finding whether phenetic similarity is concordant with the currently accepted taxonomies in certain species of butterflies. He did not find a very close correspondence; in fact his work suggests that there may be many phenetic groupings of low rank (possibly of the nature of subspecies or local races) that can be recognized by numerical taxonomic methods. Some of these phenetic groups, while keying out to certain species by the usual identification methods, appear to be different from those species with which they are commonly identified. Such work may lead to considerable revision of the detailed classification of certain genera and species. In the material studied by Ehrlich, not many data were available on the genetic relationships among the populations, but this will obviously become one important aspect of such studies, together with geographical and phenetic investigations. Morishima and Oka (1960) found reasonable agreement between phenetic relations and the

ease of hybridization. For example, the species of the "sativa" cluster of rice are all easily crossed. Soria and Heiser (1961) found similar agreement in *Solanum* species.

It must be remembered that at low taxonomic levels, such as the species and below, there is a sharp distinction between phenetic groupings and genetic groupings. This has been discussed in Section 2.4.2; it was there pointed out that if a sterility barrier arose within a phenetic group, there would be for a period of time two distinct genetic groups within a single phenetic group. Phenetic, genetic, and geographical groupings cannot always be equated with one another. The point of interest is how far these different ways of grouping the organisms are concordant and how far they are not, and why this is so. We believe that phenetically distinct groups based on numerical taxonomy will usually be found to be genetically distinct, just as the experience of conventional taxonomy shows.

Heincke (1898) investigated the distribution of individuals of races of herrings in what was in effect a taxonomic character space and obtained results giving the following generalizations for any one homogeneous race: if the mean of each character is calculated and a hypothetical individual is postulated whose characters have these mean values, one obtains a sort of "average individual," analogous to the "average man" of anthropologists. In a homogeneous population such as this, we would expect the several characters of the race to be distributed approximately normally. The sum of the squares of deviations of all characters from this "average individual" is, according to Heincke, approximately the same for every member of the race. He formulated this as a law, sometimes called "Heincke's law." Zarapkin (1934), however, has reservations on the validity of this concept.

Little additional data are yet available in numerical taxonomy to test "Heincke's law," but it may be noted that the similarity values found by Sneath (1957b) for strains of each of two species of bacteria, considered in turn, showed approximately a normal distribution. Smirnov (1925) did not expect the "law" to hold at levels higher than the race, and our knowledge of quantitative genetics and speciation would support his views.

The pattern of speciation emerging from numerical taxonomic studies in microorganisms is one of many slightly different forms (usually clones) grouped into rather ill-defined "series" or "species-groups" (see Sneath, 1962, for details). This pattern seems to be common, if one may accept orthodox and intuitive taxonomic studies in groups which are largely

apomictic, as is well illustrated by many plants (for example, dandelions, blackberries, and many of the fungi). It is not confined to plants, similar phenomena seeming to occur among the annelid worms, for instance. This pattern is also found among the protozoa, as discussed in an excellent review by Sonneborn (1957). The pattern may be likened to an ill-made brush, in which the clones are represented by the bristles, and the bristles are clustered erratically into tufts of various shapes and sizes and various degrees of compactness. The phenetic differences are represented by horizontal distances and time by vertical distances among the bristles. The pattern in sexually reproducing organisms is more like a bundle of ropes, where each rope represents a single and well-defined sexual species, within which the fibers represent the interwoven lineages. These two phenetic patterns presumably depend on different genetic patterns; in bisexual species the lineages are continually meeting and blending in a reticulate fashion and it is the rope as a whole which is the main unit of evolutionary change. Sometimes two such ropes will fuse wholly or partially, as when hybridization occurs between these sexual species, from time to time, in nature.

Clonal reproduction, however, will produce many minor variants, as mutations occur in the stocks, and most of these variants will in all probability die out. But a few of them persist and may undergo rare sexual recombination with other clones, leading to new bursts of variation due to reassortment of the genes of the parental clones (and not simply to their content of mutants). The bristles representing the clones will form a diffuse array within any tuft, and adjacent tufts will often be connected by intermediates. Sometimes the tufts will be tightly clustered into masses of bristles, and sometimes they will be straggly and widely spaced. The smallest practicable unit of taxonomic nomenclature will often be the tuft or perhaps the whole brush, although for certain special purposes it may be useful to label the bristles with some number or other designation. Numerical taxonomy, by locating the bristles in the horizontal plane which represents the phenetic resemblance in our analogy, will help systematists to recognize these tufts and will assist in their description. It will also help the student of evolution to see whether sexual recombination occurs between adjacent (phenetically similar) bristles or between bristles which are further apart (more dissimilar) and thus to gain some understanding of the relation between phenotype and the ability to hybridize.

Intermediate patterns may be met with in fungi, where Pontecorvo and his colleagues have shown the occurrence of another mechanism of

gene exchange known as the parasexual cycle, operating through hetero-karyons and rare diploids with occasional segregation of genes (Ponte-corvo, 1956). Heterokaryosis is just as efficient in storing genetic variation as the normal diploid sexual condition, but genetic recombination will generally be less frequent than with sexual reproduction. It is beginning to be recognized that some means of gene recombination is almost uni-versal in living creatures, and in this sense the analogue of the sexual species may be seen in most groups of organisms. However, the mechan-isms are so unlike the well-known sexual mechanisms that it is no easy matter to define and delimit the populations which are undergoing gene exchange. Heslop-Harrison (1962) has discussed this at some length and has suggested some reasons for these taxonomic patterns.

8.2.3. The branching of dendrograms

Willis (1922) noted that in many families of animals and plants a histogram recording the number of genera containing respectively one, two, three species (and so on) gave a "hollow curve"; that is, there was a marked excess over expectation of monotypic genera and also an excess of genera containing a great many species. Willis and Yule (1922) found that plotting the logarithm of frequency against the logarithm of the number of species these genera contained yielded straight lines with slopes close to 1.5. At high species numbers the number of genera fell off (possibly because taxonomists tend to subdivide large, unwieldy gen-era into several genera). A similar pattern was found in higher ranks, as when the number of genera per family was studied. The interpretation of these curves has been much debated (see Wright, 1941; Stebbins, 1950, pp. 531–532), but in the absence of objective criteria for what should be a family or a genus, and so on, the problem has been difficult to study, and the regularities have never been satisfactorily explained.

Recently Walters (1961) has pointed out that historical factors may partly explain these curves. With the development of numerical taxon-omy, objective tests of the number of subgroups extant per phenon will become possible, together with a study of the mode of branching of the dendrograms. It should be remembered, however, that the dendrograms are not phylogenies. They only indicate phenetic relations existing among forms and do not indicate how many forms have become extinct. Since they do not have simple evolutionary interpretations, the mathe-matical study of Yule (1924) does not seem applicable.

If the branching in the dendrograms—not of the phylogenies—occurs

at random (so that there is an equal chance of every stem giving off one or more side branches in a given interval of the phenon scale), the number of side branches per stem in a given interval will obey a Poisson distribution (the number of OTU's being the total number of side branches plus one).

The data of Michener and Sokal (1957) on bees show that there is, at most levels, an excess over expectation (on the basis of the Poisson distribution) of stems which do not branch and also of stems which branch many times. This gives an excess of phenons with only one species and of phenons with many species. For example, at the 80-phenon level (800 level in the original paper) the distribution of the numbers of branches (which are always one less than the number of OTU's in the phenon) are as follows: phenons with 0 branches, 16; with 1 branch, 4; with 2 branches, 1; with 3 branches, 2; and 6 phenons with, respectively, 4, 5, 7, 8, 12, and 20 branches each. The variance, 19.06, is much greater than the mean number of branches per phenon, 2.31, and the expected number with no branch (monotypic phenons) is 3.8 instead of the observed 16. Only a few other studies have been made which include a good cross section of the species of one taxon, but a similar if less marked pattern is seen in the dendrograms of species of *Oryza* (Morishima and Oka, 1960) and *Bacillus* (Sneath, 1962).

The general form of the "hollow curves" is a clustered distribution, in which the variance is higher than that expected for a random distribution. This has a bearing on the construction of hierarchies, as has been discussed in Section 7.1, and it offers an explanation for the validity and usefulness of hierarchic classifications in systematics.

8.2.4. Paleontological problems

The most obvious application of numerical taxonomy in paleontology is to fairly complete and well-preserved fossil material, in which the hypothesis of nonspecificity is likely to hold well enough for us to obtain reasonably good estimates of overall phenetic similarity. In paleontological studies, the importance of exact and numerical methods is even greater than it is with living material, since in extinct forms there can be no appeal to genetic data, and unsuspected heterogeneities may complicate what at first sight seem to be single phyletic lines. Robinson's method (1951) may assist in determining chronological sequences in

phyletic lines. It would therefore be difficult to overestimate the importance of numerical taxonomy in paleontology.

It is of great importance in paleontological work to make use of allometry, where this is possible (see Section 5.3.7). This is because one may have no way of knowing (or estimating) the age of the specimens at death, and the crude character sizes or ratios will sometimes be partly dependent on age and other factors. This has recently been well discussed by Joysey (1956). In many instances a fossil form is known from only a single specimen, which may represent a young or an old individual, and it may then be very difficult to know how to code its characters. The same problem, of course, occurs in orthodox taxonomic studies

An example of the application of numerical taxonomy to paleontology, which raises points that are too often glossed over in paleontological taxonomy, was given by Sneath (1961). Four samples of the fossil fish, *Knightia* (from the Paleocene and Eocene), which probably represent one phyletic line, were studied, using data given by Olson and Miller (1958). The number of features, fourteen, was unfortunately small—fewer than we would like to employ—but they could each be scored for many states, and as a pilot study it proved of some interest. It was noted that some individuals were more similar phenetically to individuals from different strata than they were to their contemporaries in their own stratum, but in the main fishes from any one stratum were most similar. Several aberrant individuals were noticed in some of the samples. Two coefficients of affinity gave very similar results, and the fish were then roughly clustered in two dimensions. After adding a third dimension, to represent time, a schematic diagram of their possible evolution was drawn (see Figure 8-8). It may be noted that the comparison of individuals considered as geometric shapes was fairly straightforward; when, however, one wishes to draw conclusions about the populations to which the fishes belonged, many more data are required. Allometric transformations are needed, as well as some assurance that only one population is represented in each sample. These are severe limitations on the use of fossil material in numerical taxonomy. They are, however, equally severe for orthodox taxonomy.

Sneath's study also showed that the phenetic means did not indicate a regular displacement with the time sequence. This raised the question of whether there had been a small degree of reversal of evolution. Alternatively, the ancestors of each group might have been an unrepresentative section of their contemporaries, so that there had perhaps been
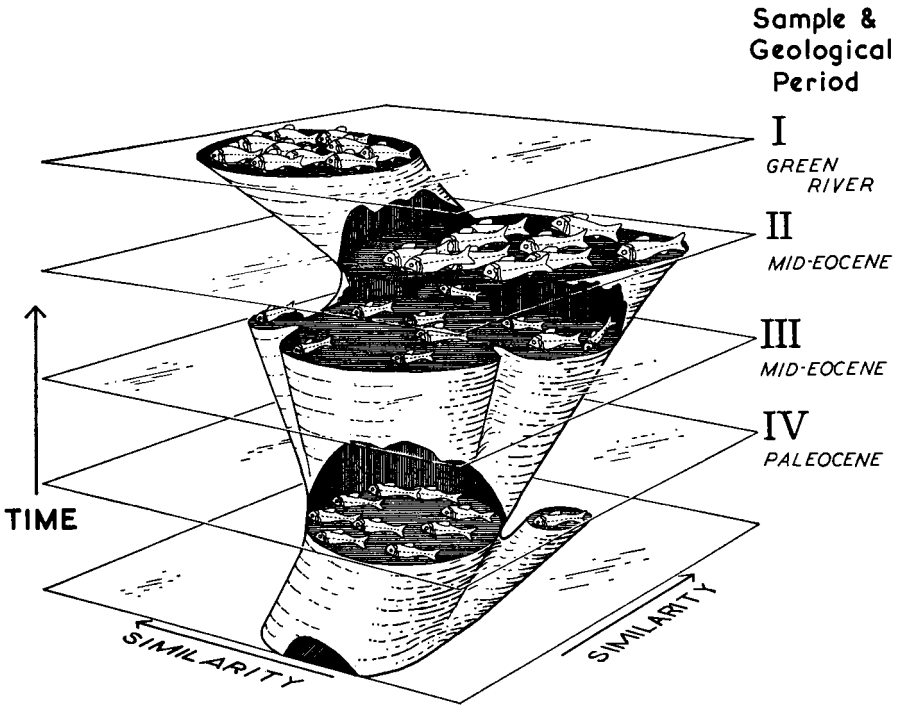
Sample &
Geological
Period

I

GREEN
RIVER

II

MID-EOCENE

III

MID-EOCENE

IV

PALEOCENE

TIME

SIMILARITY

SIMILARITY

FIGURE 8-8

Schematic and speculative diagram of phylogeny in Knightia. *The position in the horizontal plane indicates phenetic relationships among individual specimens. The time cuts are equally spaced although the actual time intervals are not equal. The axes labeled "similarity" approximate principal component factor axes. They should properly be labeled "dissimilarity." [From Sneath (1961). Reproduced by permission of the editors of* Systematic Zoology.]

repeated burgeonings of forms well adapted to the prevailing conditions but rapidly dying out and contributing little to the succeeding part of the phyletic line. This study was not suitable for estimating evolutionary rates.

It is perhaps worth noting that a large amount of the fossil record is based on fragmentary material—not merely on skeletal remains, but only on certain bones or teeth. Even complete fossil remains may yield very few characters because of the simplicity of their structure—for example mollusc and brachiopod shells and even the tests of echinoderms. It is especially unfortunate that many well-documented lineages showing gradual evolutionary changes are based on material of this kind (ammonites, oysters, sea urchins, and others), so that the interpretation of

apparent convergence, parallelism, saltations, and the like is often in some doubt.

Stratigraphy is an activity on the periphery of systematics and might therefore be discussed in Section 10.6 with "other applications of numerical taxonomic methods." Thus study of the chemical similarities between strata, for example, would be such work. However, the comparison of fossils in different strata involves numerical taxonomy of the fossils, even though the results of the taxonomy may be intended to enable the strata to be identified in different localities, across geological faults, and so on. Such work is akin to ecological studies in that many different kinds of organisms may be included as components of one OTU (the OTU's would here be strata instead of organisms). Such studies would therefore involve a new element: strata would be similar or dissimilar not only in the number of species of fossils which they shared but also in the degree to which the fossils of one higher taxon were similar in the two rocks. For example, pertinent evidence on the degree of similarity between the strata A, B, and C might be obtained from the degree of similarity between three species of a given genus, species **X**, **Y**, and **Z**, each characteristic, respectively, of strata A, B, and C.