

## Population Phenetics

Except for work in microbiology, where the application of population and species concepts is difficult and the OTU's have usually been individual strains or clones, the earlier development of numerical taxonomy dealt mainly with OTU's at the intermediate and higher categorical levels. Species were grouped into genera and genera into higher groups. The grouping of individuals into populations and species was not attempted until several years had passed, primarily for two reasons. First, the finest levels (the tips of the taxonomic hierarchy) comprise the greatest number of OTU's. In the early days of numerical taxonomy, computational difficulties deterred workers from analyzing such data. Second, while the arbitrariness of delimiting the higher taxa is easily recognized, it was believed that, at least in higher plants and animals, the definition and delimitation of species was based on objective grounds (that it was nonarbitrary).

It has taken some years for the lessons of phenetic taxonomy to penetrate to the species level. This required the confluence of three separate developments, in addition to the development of more powerful computers and computer programs. First, it became increasingly obvious, with the new emphasis on empiricism and operationism in taxonomy, that the biological species definition was difficult to apply and in practice was nonoperational. Section 7.1 deals with this aspect and also discusses the general problem of phenetic patterns in natural populations.

The second development stems from the consistent application of phenetic principles to taxonomy at all levels. Special problems of the methodology of applying numerical taxonomy to species and populations are discussed in Section 7.2. The conclusions about genetic and evolutionary structure of populations that can be obtained from phenetic evidence are discussed in 7.3. The third development was the realization that phenetics was not only a necessary approach to taxonomy at the finer hierarchic levels because of the difficulties of ascertaining breeding structure of populations, but also that the phenetic approach in population biology has considerable inherent interest for the evolutionist. This quite new approach is discussed in Section 7.4. The final section, 7.5, leads to a related topic, the biometric analysis of geographic variation of local populations or species.

## 7.1 PROBLEMS OF DEFINITION

When considering taxonomic studies at the species level we run head on into one of the most perplexing problems of taxonomy, the problem of what to designate as species. The origin of this problem stems from the recognition that there are relative discontinuities in nature among organisms and individuals, be these phenetic, geographic, or reproductive, or some combination. While the species concept has been a central tenet of biological belief since the early origins of biology as a science, the implications of this term have changed over the years. Different authors at different times have used the term species to denote different concepts and yet, because Linnean nomenclature applies a binomen to these different concepts, the reader is often uncertain about the nature of the entity to which a species name has been given.

At the base of the taxonomist's problems is his uncertainty of what kind of species concept he should employ and whether this should be one that is applicable through all biology. His views often depend on the kind of organisms he studies, but the use of widely differing concepts in different taxa is not conducive to clarity in biology as a whole. Many authors have pointed out that there are three major species concepts that depend on criteria differing in kind rather than in degree. These criteria are genetic, phenetic, and nomenclatural, and various separate terms have from time to time been suggested to distinguish species based upon them. Thus Ravin (1963) refers respectively to *genospecies* (a group of organisms exchanging genes), *taxospecies* (a phenetic cluster), and *nomenspecies* (organisms bearing the same binomen), while Blackwelder (1967a, p. 170) lists many similar terms. No consistent terminology has yet become widely accepted. The paleontologist views the species as an entity extended in time, yielding yet another kind of species concept, the "paleospecies," most fully considered in a symposium edited by Sylvester-Bradley (1956). As a segment of an evolving lineage the paleospecies obviously has no sharp borders in the time dimension, and though intended to be in principle a genetic entity, it is in practice phenetic.

Of these species concepts only the phenetic and nomenclatural are universally applicable, and the latter is readily seen to be secondary to the former. We may note, however, that the phenetic species concept itself is open to different definitions. This topic is taken up below.

A taxonomic species based on morphologically similar populations located in a definite geographic area and morphologically distinct from other populations assigned to different species is perhaps the most commonly employed concept by taxonomists in the plant and animal kingdoms (Davis and Heywood, 1963; Blackwelder, 1967a; Michener, 1970). Such a species concept is essentially a phenetic one. As employed in traditional taxonomy the recognition and delimitation of species have frequently been subjective, arbitrary procedures. But the introduction of numerical methods has enabled one to quantify these operations and to make them explicit.

Among species definitions the so-called biological species concept holds a special place. It is intimately tied up with the development of the New Systematics starting in the 1930's. The foremost advocate of the biological species concept, Ernst Mayr, has called it multidimensional (Mayr, 1963) because it deals with populations distributed through space and time, interrelated through mutual interbreeding, and distinguished from others by reproductive barriers. As defined by Mayr a biological species comprises "groups of actually or potentially interbreeding populations which are reproductively isolated from other such groups." In a more recent statement (Mayr, 1969a,b) he has dropped the much criticized phrase "potentially interbreeding" from the definition. Other biological species definitions are that of Emerson (1945): "... evolved (and probably evolving) genetically distinctive, reproductively isolated, natural population," or that of Grant (1957): "a community of cross-fertilizing individuals linked together by bonds of mating and isolated reproductively from other species by barriers to mating." Objections to these definitions have been made on several grounds. Blackwelder (1962) and Sokal (1962b) maintain that the employment of the biological species concept misleads us into viewing species described by conventional phenetic criteria as having the genetic properties ascribed to the biological species. This in turn leads to unwarranted implications about reproductive relationships from data about which little or no evidence on population structure or breeding structure is available. Sokal and Crovello (1970), in a detailed analysis of the operations required to determine whether a population is a biological species, show that the biological species concept as defined by Mayr or others is in practice nonoperational. Phenetic procedures must be resorted to in order to arrive at decisions on the status of the populations. These procedures in turn do not meet the definition as stated. In the conclusions of their paper they state that just as phenetics is necessary even when delimiting biological species, a biological species concept is not necessary even in evolutionary theory, since much of the theory is based on local interbreeding populations rather than on biological species. The

desirable fundamental taxonomic unit would seem to be the phenetic species of conventional taxonomy, the delimitation of which may be improved by methods of numerical taxonomy. Some considerations of a methodology for so doing are discussed by Rogers and Appan (1969).

A phenetic definition of species does nevertheless raise certain other problems. Some of these are discussed in other sections (e.g., 5.11, 7.3, 9.2) but we may note that a phenetic definition can result from consideration of two principal alternatives. We may regard as a species (a) the smallest (most homogeneous) cluster that can be recognized upon some given criterion as being distinct from other clusters, or (b) a phenetic group of a given diversity somewhat below the subgenus category, whether or not it contains distinct subclusters. Whether genetic, phenetic, or nomenclatural concepts are employed, meanings of the term "species" have always been closer to the first alternative; it indicates a *distinct kind*, and by implication the smallest distinct kind. Yet in practice it is the second that is usually chosen if the two alternatives clash seriously, as they do in apomictic taxa containing great numbers of microspecies. Not only might a researcher be faced with naming innumerable microspecies, but one species might show more internal diversity than an adjacent polytypic genus. It is the traditional requirement that a species must bear a binomen that causes taxonomic difficulties at this level, for otherwise subspecific or *ad hoc* infrasubspecific designations could be used for the forms thought worthy of recognition. On the other hand, all categories above species are based on the degree of internal phenetic diversity, and consistency would require one to treat the species category in the same way. But then the species could not be equated with the smallest clusters, because these can be compact or diffuse, and moreover there might be no clear boundaries between them. The solution to this difficulty evidently depends on reforms in our systems of nomenclature (see Section 9.1). The worker who wishes to employ numerical taxonomic methods at the population level can postpone most of these problems until the final stage of a formal taxonomy, and will therefore be inclined to use the first alternative (smallest distinct homogeneous cluster) when sampling OTU's to be employed in the study.

The problems of definition multiply when applied to taxa in categories lower than species. The most common of these is subspecies, the existence, definition, and naming of which has been the source of considerable controversy in recent years. A definition such as that by Mayr (1969a) ("a subspecies is an aggregate of phenotypically similar populations of a species, inhabiting a geographic subdivision of the range of a species, and differing taxonomically from other populations of the species") is not operational in practice—it cannot be used to define critically any such unit in nature. Belief that subspecies exist widely as distinct and definable units and that their putative role is as incipient species in evolutionary processes has led to a description of subspecies that is widespread, especially in vertebrate zoology and the Lepidoptera. Such subdivisions are often based on few characteristics

that overlap considerably in their distributions between supposed subspecifically distinct populations. In order to overcome problems of overlapping unit characteristics, "overlap rules" for recognizing subspecies were established (Amadon, 1949); these have been criticized by a number of authors such as Pimentel (1959), Sokal and Rinkel (1963), and Sokal (1965). The fact that an assemblage of populations subdivided on the basis of one character need not always correspond to a similar subdivision of the same assemblage on the basis of a second character is by now evident (see Wilson and Brown, 1953; Gillham, 1956; Sokal and Thomas, 1965; Thomas, 1968a).

Blackwelder (1967a, p. 174) has usefully subdivided the problems encountered in defining the subspecies concept as "(1) whether there is in nature enough diversity to be usefully studied; (2) if so whether this diversity can be treated in the taxonomic system; and (3) if so whether the segregates should be named in the formal system of nomenclature..." He replies affirmatively to the first question and points out that the second question has scarcely been adequately investigated. But in saying that the diversity *might* be treated in the taxonomic system he implies that such a system permits arrangement of the individuals within populations and of populations within the species in hierarchic and mutually exclusive groups, for if one names the groups within the Linnean system of nomenclature one would in effect be implying an affirmative answer for the second question; Blackwelder has recognized this and correctly notes that his third question has "clouded the second."

This again demonstrates our tenet that naming organisms in a system with a predetermined structure may often preclude the correct description of the phenetic variation in nature. Certainly at the infraspecific level it is even less likely that tight hierarchically ordered phenetic clusters, separated by wide gaps, are the correct model for representing populations.

An analysis of the problems of definition of species and subspecies leads inevitably to an inquiry into the nature of populations. The existence of biological populations discrete in some sense from other such populations has been tacitly assumed in theory, and in fact whole branches of science, such as population ecology, population genetics, and now population biology imply that populations of this sort exist. However, the definition of populations and the demonstration of their existence is not easy. There are clearly nonuniform distributions of phenetically similar organisms over any sufficiently large environmental area or space. Note that we have had to include the qualification "phenetically similar" because without this the problem of definition of a population becomes almost insoluble. Sokal and Crovello (1970) distinguish between (1) "localized population samples," essentially a concept from statistical geography, which depends on the place where organisms are found and the ranges they are likely to attain during their life span, and (2) "interbreeding local populations," which are those generally of interest in population biology. A third criterion for delimiting populations is a phenetic one, aiming to establish clusters of phenetically similar individuals distinguished from others

by phenetic gaps. The three concepts of distribution, interbreeding, and phenetics are clearly interrelated, yet clusters established on the basis of each will not necessarily be coincident; differences between groupings based on all three should lead to important insights into evolutionary phenomena. Transformations of one set of relationships into the other may offer means of defining evolutionary forces leading to nonequivalence relationships. To assert that one or the other of these viewpoints regarding populations is paramount and that others should be subordinated to it is to exhibit an unwarranted prejudice in favor of one philosophical point of view in taxonomy and is not likely to lead to a deeper understanding of evolutionary mechanisms.

These considerations lead to the conclusions reached by Ehrlich and Holm (1962) in their important paper that advocated study of patterns by which organisms are related in space and time rather than investigation of the concordance in nature between facts and preconceived concepts such as species, subspecies, communities, and even Mendelian populations.

## 7.2 NUMERICAL TAXONOMY AT THE POPULATION LEVEL

If the problems outlined in the last section are to be investigated fruitfully, comparisons between the phenetic, distributional, and reproductive properties of populations must be made. This will require phenetic analyses of populations at infraspecific levels. Numerical taxonomic work at this level, however, brings with it an entirely new series of problems.

As discussed in Sections 4.3, 5.6, and 8.5, a considerable literature in multivariate analysis of populations dating back into the 1930's is available for such work. Many of these are methods for establishing distances between populations and for ordinating these populations in a parsimonious manner based on as few dimensions as possible. Others attempt to find the best discriminant axes between populations. Relatively few methods concern clustering and grouping of similar populations. The early work in numerical taxonomy was largely at the supraspecific level, and hence did not need to rely on the assumptions of the multivariate methods (multivariate normality, homoscedasticity of variance-covariance matrices), but resorted to the by now familiar techniques of establishing similarity matrices and clustering these to form hierarchic groups. The increasing interest in the application of numerical taxonomic methods to problems in systematics led to the extension of these techniques at the infraspecific level without considering whether clustering methods are necessarily most suited for these problems. No clear answer can be given, even at this date. We are left with the problem of whether taxonomic methodology should primarily use ordination or clustering techniques. In the former case we might represent the phenetic relationships among populations and individuals within populations with a minimum of distortion, but if a large number of

populations is sampled and if many individuals are employed for each population sample the resulting dispersion of points in a hyperspace or even in a three-space would probably be unintelligible. Yet any attempt to make order out of such patterns by clustering them according to a given algorithm imposes a structure on these data that may not be justified. Fundamental problems in this regard have already been discussed (see Section 5.3) and need not be restated here.

Various methods of condensation would be appropriate for introducing structure in large ordinated samples, following lines suggested by Sneath (1966e). Not too much has been done about condensing points ordinated in hyperspace and we must therefore consider mostly instances of population phenetic analysis at the infraspecific level based on clustering approaches. If clustering and ordination approaches are to be integrated, problems of scale, character coding, and of dispersions of the characters at the various levels must be investigated. The development of scale free character coding methods (see Section 4.8) should be encouraged. Important also will be the problem of congruence (see Section 3.6) that, as we have seen, may present the investigator with different results at the infraspecific level than at higher taxonomic levels. Problems exist also in estimating the phenetic resemblance of very similar OTU's that may be identical in most of a large suite of character states. Thus the measure of their similarity may depend on differences in one or two character states, an unsatisfactory condition from the point of view of statistical reliability. The problem of resolving dimorphic or polymorphic populations into their components (discussed at the end of Section 5.2) is of great importance at low taxonomic levels.

There is some doubt even about whether meaningful hierarchic structures can be obtained below a given categorical level. Whether this level is below that of the species (whatever its exact definition may be), or whether nonhierarchic phenetic relations begin at a higher level needs further investigation. In fact, if it can be clearly shown over a wide range of organisms that phenetic structure changes markedly at a level commonly designated that of the species, this very fact would aid in the definition of the species category and might further the causal analysis of the phenomenon of speciation. We know that, at conventional specific levels, replicate individuals almost always cluster before they group with nonconspecific individuals (unless the problem is confounded by the introduction of different morphs or developmental stages as distinct OTU's in the same study, as for example, by Boyce, 1964). Thus, in a study of 55 conventionally accepted species of meliponine bees (da Cunha, 1969), each species based on three representatives serving as OTU's, only in 12 species did the conspecific triads not join before clustering with other species. A similar finding was reported by Funk (1964) with 25 apparent species of euzerconid mites and by Moss (1968b) on 15 species in the mite genus *Dermanyssus*.

Turning to results at lower categorical levels, an account of early investigations is furnished by Sokal and Sneath (1963, p. 243). Among the earliest applications of

cluster methods by numerical taxonomy at the population level was the analysis by Ehrlich (1961b) of the butterfly species *Euphydryas editha* and *E. chalcidonea*. This study showed that though supposedly conspecific populations clustered together, individuals proximate geographically did not emerge from a cluster analysis of individuals of *E. editha* based on 75 characters. This study was employed by Ehrlich to substantiate his views that patterns of relationship in nature may be inconsistent with their formal taxonomic placement and has been criticized by a number of systematists for being based on only 13 individuals. Although a greater number of OTU's would have been desirable, later studies with larger samples in other groups (reported below) seem to bear out Ehrlich's contentions that phenetic variation may not always be concordant with geography. Furthermore, there may perhaps be significant variation between samples from different years (Mason, Ehrlich, and Emmel, 1968). Among other early studies at the population level could be included a number of the microbiological studies in which the bacterial strains may correspond to individuals (see, for instance, the work of Liston, Weibe, and Colwell, 1963).

Comparisons of clusters based on phenetic analysis of geographic samples yield varying results. There is some correspondence between geographic proximity and phenetic similarity, but the correspondence is not complete. Such findings could be shown by Sneath (unpublished) in data on 20 populations of the house mouse published by Berry (1963) based on the incidence of 35 skeletal variants. Similar results were found by Fujii (1969) in eight populations of the Azuki bean weevil based on 13 ecological characteristics, by Thomas (1968a) for 16 characters of the rabbit-tick *Haemaphysalis leporispalustris*, and by Johnston (1969) in house sparrows based on morphological characteristics of the skins and skeletons.

A study by Soulé (1967a) on lizards showed that island populations were relatively homogeneous phenetically, and that they were most similar to the mainland populations that were closest to each island. A further study (Soulé, 1972) illustrates that general variability of local populations increases with the logarithm of the area of the island. The author believes this to be due to selection for stable genetic polymorphisms in the complex diverse communities of large islands as contrasted with the impoverished ecosystems of small islands. A similar study is that of Berry (1969) on field mice. A dimorphism observed subjectively by the collector (Sokal) in a field sample of 118 galls of the aphid *Pemphigus populi-transversus* could be shown to be supported but not entirely congruent with clusters based on external shape of galls and on the morphology and biology of the aphids (Takade, 1971). In plants, a broad correspondence between phenetics based on chemical characters and large scale geography has been noted by La Roi and Dugle (1968) and Thielges (1969). Similar findings are reported by Hickman and Johnson (1969), though they found that the effects of plant size introduced complications. It may be necessary to grow plants under identical conditions (e.g., Hubac, 1964) to avoid environmental effects.



In studies of population phenetics at the infraspecific level ample thought must be given to the allocation of samples at various hierarchic levels. It is almost always impossible to process as much material as is available, because it is generally easy to obtain population samples based on rather large numbers of individuals. This is complicated by doubt concerning the definition of a local population sample and how frequently such population samples should be taken for a given species. In clustering studies, computing time increases generally as the square of the number of OTU's, so one must do separate analyses for various subunits or take samples limited to the capacity of a program.

### 7.3 PHENETIC PATTERNS AND EVOLUTIONARY STRUCTURE

The comparison of phenetic relationship and interfertility is of considerable interest. Close correspondence between cross-fertility and phenetic resemblance may give one confidence in the phenetic method. We believe that these considerations should not be stressed too heavily. It is already clear that complete concordance can never be expected and discrepancies are of even greater interest. Some mechanisms that can explain discordance are already well known, such as the pollen incompatibility genes that prevent self-fertilization of many flowers. We may hope that others can be found through comparison of numerical taxonomic with biosystematic studies.

Genetic relationship, like taxonomic relationship, may refer to a number of different concepts. We only take up those that have been studied by numerical taxonomic methods. A few such studies have been made in anthropology, particularly in comparing identical twins with fraternal twins. Thus Vandenberg and Strandkov (1964) found that the within-pair variance of measurements on fraternal twins was several times greater than that of identical twins (and was somewhat greater for boys than for girls). The clear distinction between genetic identity and sibship shows the potential sensitivity of phenetic methods. Huizinga (1965), using head measurements, reports briefly on fathers and sons compared to unrelated persons; the phenetic differences here were less marked.

Other studies examining the phenetics of close blood relations include those of Grewal (1962), Berry (1963), and Berry and Searle (1963) on mice. One point that emerges from their work is that though highly inbred lines are less variable than wild populations, they nevertheless show much phenetic variability. Rhodes and Carmer (1966) in similar work on maize noted that those cultivars (inbred lines) that were closely related by pedigree were clustered together in phenograms. Four large phenons represented the major sweet corn stocks named "flint," "dent-flint," "dent," and "bantam," with some suggestion (as one might expect) that the first three consist of a partial continuum of phenetic forms. Goodman (1967a, 1968b) has made similar studies on maize. Oka (1964) reports some preliminary work of this kind on rice.

Several authors have investigated the effect of polyploidy on phenetics. The members of autoploid series are usually close phenetically (Hubac, 1964; Parups et al., 1966; Bidault and Hubac, 1967), while Heiser, Soria, and Burton (1965) noted that in such a series the hexaploid was closer to the tetraploid than to the diploid parent. Bidault (1968) studied the grass *Festuca ovina* and found the major clusters corresponded to ploidy; there was a cluster of diploid forms and another of tetraploid forms, and within each of these were smaller clusters corresponding to diploid and tetraploid parts of subspecies. For example, the first major cluster contained a subcluster of diploid forms of the subspecies *glauca*, and the second a subcluster of tetraploid forms of the *glauca* subspecies. These findings are probably not due to general size effects, inasmuch as Bidault used correlation coefficients. Such observations raise problems of formal taxonomic treatment similar to those raised by apomictic forms (see Section 9.2). The position with allopolyploids is less clear. Heiser et al. (1965) noted that experimental *Solanum* allopolyploids were often very different from both parents, and Katz and Torres (1965) reached similar conclusions on presumed allopolyploid species of *Zinnia*. This phenomenon is, of course, well known in a nonquantitative way, and relates to the phenetics of hybrids, further discussed later. The protein electrophoretic patterns of allopolyploids may be very close to the appropriately weighted sum of those of their parents (Johnson and Hall, 1965). Ising and Fröst (1969) found that clones of the same cytotype of *Cyrtanthus* species usually clustered together on numerical chemotaxonomy.

As one would expect from the existence in plants of mechanisms to prevent close inbreeding, there is often little congruence between ease of hybridization and phenetic resemblance, and it seems unwise to make hybridization the main test of the validity of a numerical taxonomy, as some authors have done. There is nevertheless some congruence as shown by the studies of Morishima and Oka (1960), Soria and Heiser (1961), Heiser et al. (1965), and Rhodes et al. (1968). For example, Morishima and Oka found that the species of rice in a cluster designated "sativa" all crossed easily, and Soria and Heiser had similar findings on *Solanum*. Rhodes and his coworkers found correlation coefficients of over 0.8 between phenetic similarities and cross-fertility scores among species of *Cucurbita*. Some information is available for bacteria. The relations between genetic behavior of bacterial genera and their phenetic resemblances are reasonably congruent (see reviews by Sneath, 1964a, and Jones and Sneath, 1970) including bacteriophage host range. Dr. H. Morishima (personal communication) tells us that she is at present comparing phenetic resemblance with the pattern of cross-fertility in rice, as judged by behavior of OTU's on crossing with a standard series of tester stocks. This pattern of cross-fertility is in some respects like a phenetic pattern.

Much work is now being done on the application of phenetics to plant breeding. Some of this is related also to the study of character variability and the effects of the environment discussed in the next section (e.g., Goodman, 1968a, 1969; Goodman and Paterniani, 1969; Bhatt, 1970). Other work is concerned with studying the phenetic positions of hybrids in relation to their parents. Hybrids are

usually phenetically close to one or both parents, as one would expect, although there are occasional exceptions (e.g., Ramon, 1968). It is often possible to guess the identity of one or other parent from the phenetic position of the hybrid (Bemis et al., 1970; Kaltsikes and Dedio, 1970b). Numerical methods are also powerful in detecting hybrids (e.g., Smith, 1969; see also Rising, 1968, on birds) and this continues the tradition associated with the use of the Hybrid Index (Anderson, 1936), which is still often useful for simple hybrid identification problems (e.g., Goodman, 1967b).

Hybrids are commonly thought to be morphologically intermediate between their parents (a belief based largely on early work with the Hybrid Index), but numerical taxonomic studies on this show that they are not usually intermediate in the sense of lying on the line in phenetic hyperspace that joins the two parents (Cousin, 1956b; Heiser et al., 1965; Katz and Torres, 1965; Wirth, Estabrook, and Rogers, 1966; Whitehouse, 1969). They appear more often to be well to the side of this line. The same was true of a graft chimera studied by Sneath (1968a) that, though not a hybrid, is analogous in combining two genotypes. Allopolyploids appear to behave in much the same way. This displacement from the line indicates that some genes from one parent are dominant and others are recessive, or it may be the result of overdominance. Using canonical analysis on barley and beans, Whitehouse (1969) found that  $F_1$  hybrids lay on the average midway between their parents but displaced laterally about 35 percent of the interparental distance; the offset was rather less for  $F_2$  hybrids. The scatter in hyperspace of independently formed hybrids of the same parents has not yet been much investigated, but they are apparently usually fairly close phenetically (Heiser et al., 1965; Ramon, 1968; Casas, Hanson, and Wellhausen, 1968). Hubac (1969) observed that hybrids of *Campanula* were closer to the seed parent than the pollen parent, but no sex influence has been noted by other workers. These findings on the position of hybrids have been generally supported by parallel work in chemotaxonomy (e.g., Olsson, 1967; Ising and Fröst, 1969; Dedio, Kaltsikes, and Larter, 1969b). Certain problems in such work when two-state characters are used have been discussed by Sneath (1968a). In particular, considered in Euclidean space, the hybrid cannot lie on the line joining the parents, because it must occupy a corner of a hypercube (this has not been studied for angular measures of resemblance). This may distort the relationships that might have been obtained if the characters could be measured quantitatively and is particularly important with chemical characters that are commonly codominant (see B. L. Turner, 1969) and usually scored only qualitatively. This distorting effect rapidly diminishes as the average number of states per character is increased.

The application of this new knowledge of phenetics of hybrids is being used in plant breeding in several ways. Whitehouse (1969, 1971) suggests it can help in the choice of appropriate parents to produce hybrids with desired properties (the scaling of character axes to reflect economic importance might be also considered). Bhatt (1970) suggests that phenetically dissimilar parents are most likely to yield

useful novelties. Errors in pedigree records may often be located by phenetic studies (Rhodes and Carmer, 1966; Rhodes et al., 1970). Such work may also assist in the choice of stocks for preservation in banks of germ plasm (Rhodes, Carmer, and Courter, 1969; Whitehouse, 1969), since in general plant breeders require the greatest variety, that is, stocks with the greatest atypicality values (see Section 4.11). Finally this work may help the understanding of selection in crop plants (Eshbaugh, 1970; Vaughan, Denford, and Gordon, 1970).

The resemblance detected by nucleic acid pairing is in a sense a genetic relationship because it is close to being an expression of the genetic message, though in a sense it is also phenetic; the very good congruence in bacteria between the degree of nucleic acid pairing and the resemblance from numerical taxonomic studies has been discussed in Sections 3.5 and 5.12.

The patterns of distribution of organisms in phenetic hyperspace is no doubt related to the patterns of speciation in different groups. It is possible that the phenetic distribution of individuals is different in sexually reproducing and in apomictic groups of organisms (see Sokal and Sneath, 1963, p. 244). At present there is little pertinent information available about sexually reproducing forms, but some is at hand for apomictic organisms, chiefly bacteria (see Sneath, 1968c; Jones and Sneath, 1970). In bacteria it appears likely that the phenetic pattern consists of compact clusters of individuals (representing a clone or a number of extremely similar clones), with isolated individuals or small clusters (representing uncommon clones) scattered between them. These isolated forms may perhaps originate from occasional episodes of hybridization, or from unusual occurrences of mutation and selection. It is beginning to be recognized that some means of gene recombination is almost universal in living creatures, and in this sense the analogue of the sexual species may be seen in most groups of organisms. However, the mechanisms are often so unlike the well-known sexual mechanisms that it is no easy matter to define and delimit the populations that are undergoing gene exchange. Heslop-Harrison (1962) has discussed this at some length and has suggested some reasons for these taxonomic patterns. This is an area of phenetics in which we expect to see rapid progress very soon.

## 7.4 PHENETICS AND ENVIRONMENT

In this section we discuss the relation between phenetics and environmental factors. The way in which the environment affects the phenotype is considered first, followed by a discussion of classification of environments on the basis of the responses of phenetic characters to them. The effect of environment on choice of characters has already been discussed in Section 3.6. A good general discussion of this field in plants is given by Davis and Heywood (1963, pp. 335–416).

We should make clear a basic difference between phenetic studies and the more usual studies of environmental effects. Usually one studies the effect of one variable

of the environment upon one character of the organism. But in phenetics we study the overall response of the phenotype to different environments. It is thus necessary to calculate the resemblances between OTU's under different environmental conditions. As in more usual taxonomic applications we will wish on occasion to separate environmental effects on the size of the organisms from the effects on shape and to distinguish other appropriate components of phenetic resemblance.

These applications of numerical taxonomic methods can form an experimental science, for which various experimental designs can be adopted. Thus we might change one variable of an environment at a time to find one with the greatest effect on phenetics (or more explicitly, to find the magnitude of the environmental changes that give unit phenetic change). The phenotypic responses form a multi-dimensional response system; this is an area that awaits exploration by mathematical methods. Factor analysis is an obvious technique for discovering the most important environmental factors and the major components of the phenetic response; analysis of variance and covariance is another important method. In other experimental designs one might not be able to control the environment directly but would be able to observe and analyze "experiments" made by nature. This field seems particularly important for future work on the plasticity of the phenotype. It may also be noted that the phenotype can affect the environment—most obviously perhaps in the case of vegetation, where the macroenvironment determines phenetics, which in turn determines microenvironments.

There has been little work so far in these areas. Phenotypic plasticity is very marked in plants, and studies of clones and purebred seed lines are obvious ones. It would be of special interest to quantify the phenetic responses to environment and ecotypes by following up, for example, the classic work of Clausen and his colleagues (Clausen, Keck, and Heisey, 1940; Clausen and Heisey, 1958; and see discussion by Davis and Heywood, 1963, pp. 390–398). Goodman and Paterniani (1969) have investigated the plasticity of different characters in maize and adduce evidence that some (but not all) characters associated with reproduction are less affected by the environment than are morphological characters. Abou-El-Fittouh, Rawlings, and Miller (1969) used cluster analysis to classify cotton-growing areas into regions where environmental effects on cotton are similar. In using  $D^2$  to study locusts Gillett (1968) found some marked phenetic changes due to changing environmental conditions. A little work has been done in microbiology. Melville (1965) investigated the same bacteria under aerobic and anaerobic growth conditions. Although the direct effects of environment on phenotype were not recorded, the two different environments gave the same phenons. Small changes in phenetic relationships were found by Davis et al. (1969) on altering the growth temperature of certain bacteria. It would be of interest to know whether the relative similarities between OTU's were more stable in the changed environment than the actual similarity values between given OTU's under the different conditions. One can envisage classificatory investigations where the OTU's would be very different

phenetically in two environments, yet classifications made from the different environments would be highly congruent, and there may be general biological grounds for thinking this likely. Some notes on the effect of growth temperature and age of culture of bacteria will be found in Sneath (1968d); these two variables have effects analogous to a general size factor.

The use of phenetics to discover the optimal environment for crop plants could be of economic potential. Such an application is not the same as controlling the environment to maximize crop yield, because one might not be able to achieve the ideal state of every attribute of economic importance—higher yields might be of poorer quality, for example. Where numerous attributes of economic significance were involved one might be able to make a numerical taxonomic approach in the following way. An ideal but hypothetical OTU is constructed and represented as a point in attribute space, whose dimensions are the relevant attributes suitably scaled (not necessarily linearly) to reflect their economic importance. The search then would be for that environment giving OTU's as close as possible to the ideal one, with the hope of finding the best compromise. Such an approach is very close to the rationale of operations research (see Section 11.5) and is similar to the way in which the best genotype (in breeding programs) can be found. Such approaches differ from the more familiar multiple regression and variance-covariance analyses in that the phenetic response is not assumed to be linear. Also, if we consider hill-climbing techniques we find they suggest the best experiments to perform next. For instance, we may consider the change in environment to be the difference between two environmental vectors  $\mathbf{v}_1$  and  $\mathbf{v}_2$ , which correspond to two phenetic vectors  $\boldsymbol{\phi}_1$  and  $\boldsymbol{\phi}_2$ . The search procedure would then be to discover the relations between these vectors, in order to discover the required environment giving the optimal phenetic vector. The technique of response surface analysis (e.g., see Peng, 1967) seems appropriate to obtain the desired results for such problems.

Hutchinson (1957) has suggested that ecological niches should be formulated within the framework of a multidimensional system of environmental requirements. The concept of patterns in phenetic space associated with configurations of ecological space suggests itself very readily (see, for example, Maguire, 1967 and Wuen-scher, 1969). Hutchinson (1968) himself has pursued this line of reasoning by illustrating patterns of ecological space occupied by strains or species of rotifers. Thus one might measure the diversification of ecological niches by the phenetics of the organisms that inhabit them. Notable in this connection is a study by Fujii (1969) defining niches from biotic parameters of strains of the bean weevil *Callosobruchus*. Other studies are underway.

The concept of ecological niches is vague. A given habitat may be subdivided into a great many smaller habitats using rather uncertain criteria, which depend in part on the phenotypes of the organisms themselves, since the organisms can change the habitats. Indeed one might say that this method is the only way we recognize one niche as being distinct from any other; if there were no difference in

phenetics we would consider them the same, despite differences in physical variables. For example, one school of ecological thought would not consider designating the north and south sides of sand dunes in a desert as separate niches if no differences could be found in the organisms, even though the temperature might be noticeably different. There have been a number of publications noting the difficulty of measuring such ecological entities (e.g., Ehrlich and Holm, 1962; Bock, 1963).

Studies of different environments can be of two kinds. In some, the different environments will affect the phenetics of the same species of organisms (for example, by affecting size, color, frequency, etc.). In other studies the environments will contain different taxa of organisms and the phenetics of these taxa will be an object of study. Studies of this second kind border closely on ecological classification (Section 11.1).

## 7.5 ANALYSIS OF GEOGRAPHIC VARIATION

Readers of the previous sections in this chapter will have realized that many of the techniques for numerical taxonomy at the population level grade directly into those phenetic methods commonly referred to as geographic variation analysis. These methods have in the past been almost exclusively univariate; that is, investigators would describe the variation over space of a single characteristic at one time, although more than one characteristic might have been investigated. The study of numerous characteristics provided new problems inasmuch as the large amount of data collected needed to be simplified and made interpretable. Although the methods of geographic variation analysis *sensu strictu* do not fall within the province of numerical taxonomy (since they relate primarily to the variation of characters in a geographic 2-space, or 3-space if topography is included), the points of contact and transition between the two methodologies are so numerous that a brief discussion seems appropriate here. The purposes, problems, and methodologies of testing and of representation will briefly be outlined and suitable references for further study and detailed exposition of the methods will be provided.

The purposes of geographic variation analysis are first of all the description and summarization of patterns of variation and covariation of characteristics of organisms distributed over an area. Any phenetic characteristic regardless of its nature may be so studied. Geographic variation analysis may be carried out on continuous as well as discontinuous variables and also on categorical ones (attributes). Although in general variation patterns have been stressed in relation to spatial distribution whenever the material justifies it, variation in time, as in fossil deposits in any one locality, may be similarly analyzed. A step beyond mere description of variation patterns is categorization. This refers to the grouping together of localities (perhaps by cluster analysis) that are geographically adjacent and whose populations are similar in their characteristics. This may be desired merely for purposes of simplification and summarization, or for the formal or

semiformal recognition of a population or a series of populations in terms of the Linnean system. Both aims are consonant with the general aims of taxonomy as we have outlined them in this book. Other purposes of geographic variation analysis include causal analysis of the geographic variation patterns in order to interpret these as adaptations to variation in known environmental factors, such as climatic, topographic, or edaphic variables, or differences in the distributional, reproductive, or ecological patterns within the populations. Finally, geographic analysis may lead to the allocation of unknown specimens to a given population or a geographic locality with a stated probability of success.

In studying geographic variation, problems arise due to errors of three kinds: (1) sampling error due to natural variability of organisms at a given locality; (2) measurement error for the sampled organisms; and (3) errors in representation of the entire area of study by means of a particular set of localities. These errors are discussed in somewhat greater detail by Gabriel and Sokal (1969). The third source of error has been least investigated, but it is of the most profound importance, since very often trend lines bounding contours of equal phenetic values are established without any evidence that the contour lines they imply have any statistical validity. The geographic variation literature abounds with so-called *isophenes* and *isarithmetic lines* whose validity is only as good as the validity of the individual points on which they are based and the degree to which the sampling of points over a geographic area represents the trends actually occurring in the area. In recent years statistical geographers have begun to concern themselves with this question (e.g. Haggett, 1965, p. 214; Stearns, 1968), but the field remains essentially open.

The approaches to testing the significance of the results obtained relate also to the method of representation. If the data are plotted as continuous trends with emphasis on the construction of isophenes then the method of *trend surface analysis* will be preferred. By this method a polynomial or Fourier surface is fitted to the local observations for each character by least squares, and contour lines are plotted on a map for easy visualization of the trends. These contour lines estimate the loci of all points with a given value of the character. A first study of geographic variation by such means is due to Marcus and Vandermeer (1966) and an application to zoogeographic data has been made by Fisher (1968). A general discussion of trend surface analysis can be found in Krumbein and Graybill (1965), and in Harbaugh and Merriam (1968). Local errors of estimating parameters of characteristics for any locality can have serious effects on the confidence that may be placed in the fitted trend surfaces. Errors due to nonregular distribution of localities over the geographical area studied can also be serious; this is discussed by Mandelbaum (1963), who also suggests criteria for deciding what is the highest order of the polynomial in the trend surface that should be accepted as "significant."

If one cannot assume that the characters are continuously distributed one may consider that the area consists of separate homogeneous regions differing from each other. This *categorization* approach leads to significance testing by multiple



comparisons techniques (see Sokal and Rohlf, 1969, for a discussion of these approaches). Gabriel and Sokal (1969) have explored this approach in depth using Gabriel's simultaneous test procedures, which can be applied to continuous univariate data, nonparametric tests (ranked variables), or categorical data, as well as multivariate continuous data. The method consists of finding subsets homogeneous for the characters in question that belong as well to regions that are contiguous in a geographic sense. Those sets of localities that are statistically homogeneous and geographically contiguous (as defined in graph theoretical terms by Gabriel and Sokal) are categorized as being biologically homogeneous. The difficulty with Gabriel and Sokal's approach, as with other multiple comparisons testing, is that often a multiplicity of results emerges, yielding overlapping homogeneous sets of localities rather than a simple categorization of mutually exclusive groups. However, while this may offend one's customary sense of order and neatness and run counter to the common trend of establishing allopatric, named populations, it is undoubtedly more representative of the true relations in nature, where complex breeding patterns exist among various populations distributed in an area. Sokal and Rinkel (1963), Sokal and Thomas (1965), Rinkel (1965), and Sokal, Heryford, and Kishpaugh (1971) applied multiple comparison techniques to geographic variation in the aphid *Pemphigus populi-transversus*, and Thomas (1968a) applied them to the rabbit tick *Haemaphysalis leporispalustris*. Johnston (1969) and Flake, von Rudloff, and Turner (1969) have employed simultaneous test procedures in a geographic variation study of the European sparrow, and a study of clinal variation in junipers, respectively.

Difficulties will arise when clines or time trends occur in the data. If pure clines for any one character are found, the method of trend surface analysis is clearly superior to any categorical method. However, usually either because of inadequacy of sampling or because of the known discontinuity of environmental or distributional factors, breaks in such trends occur, making the categorization approach advisable. Difficulties also arise when the phenetic distribution pattern is mosaic or crazy quilt, in which case again the categorization approach is likely to be more representative of the true pattern of variation. Biogeographical data will rarely be of the quantity and quality needed for simple contouring of the kind commonly employed in geography and related disciplines. Indeed this is a major reason for using the techniques discussed above. Computer methods for contouring are reviewed by Crain (1970).

Representation of the results such as the geographic variation analysis depends on the method employed. When multiple comparisons approaches are used, various shading techniques to indicate the magnitude of a given variable have been employed by Sokal and Rinkel (1963) and Sokal and Thomas (1965), as well as by Thomas (1968a). Gabriel and Sokal (1969) superimposed upon these the requirement of geographic contiguity, and readers are referred to their paper for methods of representation. Examples of trend surface analysis can be found in the publica-

tions referred to earlier. Frequently, three-dimensional models represented in A-space (the ordination techniques discussed in Section 7.2 and described in detail in Section 5.6) are constructed. These models are an attempt to summarize geographic variation patterns described for numerous characters  $n$  in a smaller dimensional system ( $k < n$  dimensions). Such work becomes necessary when dealing with a large number of characters. Sokal and Rinkel (1963), Sokal and Thomas (1965), Rinkel (1965, in work on the aphid *Pemphigus populi-transversus*), and Thomas (1968a, in work on the rabbit tick *Haemaphysalis leporispalustris*) all plotted factor scores of the character correlation matrices over the geographic area of study and performed various multiple comparison tests on the data. In a similar manner Fisher (1968) plotted trend surfaces of zoogeographic factors. Thomas (1968a) and Gould (1969b, on the pulmonate snails *Cerion* and *Tudora*) ordinated the localities in a factor space summarizing the attributes. In Gould's case, factors were based on only four characters and it is dubious whether the extraction of three factors from such a correlation matrix is worthwhile. When localities are ordinated in A-space the comparison of the two-dimensional ordination graphs or the three-dimensional models with geographic distribution patterns of the localities may frequently prove of interest. Quite commonly at least one of the axes of variation in the model gives a geographic direction (either longitude or latitude, or possibly topography). One would like to be able to transform the phenetic ordination into a geographic ordination and to impart some biological and evolutionary significance to the entries in a transformation matrix between these two models. Holloway and Jardine (1968) have measured the goodness of fit between a two-dimensional ordination and geographic position in an ecological study, and their methods could be readily applied to phenetic data. Johnston (1969), in a study of the geographic distribution of the European sparrow, has attempted to correlate phenetic dispersion with geographic distribution pattern and has shown that while no 1:1 correspondence exists, there is sufficient correspondence to engender useful hypotheses about the evolution of these sparrows.

Only a few general conclusions have yet emerged from this work. Over large geographical areas there is commonly good correspondence between phenetics and geography or climate (e.g., Thomas, 1968a; Power, 1969; Morishima, 1969b; Hickman and Johnson, 1969; Banks and Hillis, 1969) and in some cases the variation is sufficiently discontinuous to give reasonably distinct phenetic clusters. If variation consists of even clines, however, such clusters may not be found, and on a small geographic scale the pattern of variation appears to be extremely diverse (e.g., Berry, 1963; Petras, 1967). Size and shape coefficients may give notably different results in work of this kind (Rees, 1969a,b). In a microgeographic study meaningful clusters did not emerge (Sokal, Heryford, and Kishpaugh, 1971), and it is possible that such absence of structure may be used to indicate an important taxonomic rank level. It may well indicate panmictic populations or a homogeneous environment. For a comprehensive review see Gould and Johnston (1972).

As computer technology increases, automated mapping of operational biogeographical units (Soper, 1964) will become commonplace and the development of optimal representation techniques for summarizing geographic variation is therefore an important goal of systematic research. In a few years we also should have a sufficient study of various models that will permit us to generate evolutionary hypotheses about the origins of those variation patterns. Undoubtedly this will require substantial advances in statistical methodology as well as considerable theoretical work in the form of model building and computer simulation.