helpful in identifying critical factors for population management. Tony Starfield (University of Minnesota) described another population simulation model and showed how a different approach to modelling is required when many parameters and their ranges are unknown. Unlike engineering applications, models should not be used to seek optimal solutions, but instead to identify and minimize risks. A policy of 'minimum regret' may be a good one for conservation biology.

A second theme was 'altered states', a term introduced by William Conway (New York Zoological Society) in his opening remarks. The world is changing rapidly, especially for wild species in tropical areas, and the conservation community will soon face a variety of new problems arising from the need to manage many species, even those in reserves and protected areas. Long-term survival of many species will depend upon the development of principles very different from those presently adopted in wildlife management—principles that may conflict with our more traditional concepts of wildness and freedom. As intensive population management techniques start to incorporate both captive and reintroduced populations, the distinction between captive population management and management in reserves and protected areas breaks down. There are aesthetic and moral issues to be resolved in this area that are even more demanding than the biological ones.

Acknowledgements

The meeting was sponsored by the National Zoo, American Association of Zoological Parks and Aquariums, Institute of Museum Studies and Chicago Zoological Society. GMM was also supported by a grant from the Wellcome Trust.

References

3 Rails, K. and Ballou, J.D. (1986) Zoo Biol. 5, 81–238
7 Haig, S., Ballou, J.D. and Derrickson, S. Conserv. Biol. (in press)

DNA and Morphology: Inference of Plant Phylogeny

Kenneth J. Sytsma

In recent years, there has been an increasing number of reports of discrepancies between analyses of DNA and morphology in the estimation of phylogeny. In plants, the discrepancies can be attributed to procedural problems (apparent discrepancies) or to biological attributes of the organisms (real discrepancies). The problems can arise from within both morphological and molecular aspects of the study. A better understanding of both kinds of problem permits a more thorough synthesis of DNA and morphology in the inference of plant phylogeny, and can result in the further refinement or resolution of a morphologically based phylogeny by molecular evidence, and of a molecularly based phylogeny by morphological evidence.

In the past decade, DNA-based inference of plant phylogeny has developed into a multi-faceted systematic endeavor. Although this approach is recent, it has already had a significant impact on many levels of plant systematics. The great interest in, and major movement toward, the use of DNA in phylogenetic studies of plants raises the question: Why is molecular evidence so appealing in plants? The appeal comes from several beliefs, justified or not, concerning DNA and morphology: (1) that molecular analysis provides numerous and independent characters whereas morphological analysis provides fewer characters, often of questionable homology; (2) that morphology, unlike most DNA regions, is prone to considerable convergence; (3) that the genetic basis of convergence in molecules is better known; and (4) that molecular analysis is free of the subjective kinds of character analysis and outgroup selection that might mislead morphological analysis. As Patterson aptly queried, have molecules superseded morphology as guides to the history of life, or are the approaches sides of the same coin, with same problems and limitations? Do molecules and morphology give the same picture of the history of life, or two more or less distorted views of the same picture, or two quite different pictures? These questions have been recently examined for animals, but the situation in plants is unresolved.

The primary goal in phylogenetic studies is to infer the single historical genealogy—i.e. the true phylogeny—of a group of organisms. It might thus be expected (1) that systematic studies of any set of genetically determined characters within a group of organisms should be congruent with other such studies based on different suites of characters, (2) that congruency provides strong evidence that the true phylogeny has been inferred, and (3) that conflict among the results may indicate low resolution power of the data sets, invalid and inappropriate
underlying assumptions, or procedural problems in some or all of the analyses. Although these expectations often hold, the situation for plants appears to be even more complex.

Most studies of plant groups that have been examined by both DNA and traditional approaches indicate close congruence in many of the resulting phylogenies. This congruence is reassuring and indicates that both approaches are important in inferring plant phylogeny. In this review, however, I examine DNA-based inferences of plant phylogeny that apparently conflict with morphologically based inferences. These examples can be placed in seven categories, each characterized by features generating the apparent or real lack of congruency between the DNA and morphological data sets; the seven categories fall loosely into two groups (Box 1).

The first group consists of categories that deal with procedural problems in data analysis or interpretation (i.e. the data might in fact provide the true phylogeny if correctly analysed or interpreted). The second group consists of categories in which inherent biological or evolutionary features of the plants, either in DNA or morphology, yield the discrepancies (i.e. the available data are such that the true phylogeny is not possible to obtain without other knowledge). These categories and the groups within which they fall are admittedly artificial, in that many examples of morphological and molecular inconsistency result from combinations of both these problems.

Equating overall morphological similarity with phylogenetic relationships

Reliance on overall morphological similarity (the phenetic approach) as an estimate of phylogenetic relationships, rather than on shared derived character states (the cladistic approach), can generate misleading estimates of phylogenetic relationships (Box 2). The phenetic approach (usually using clustering or principal components analyses) often explicitly assumes the operation of some type of clock-like rate of change in morphology, whereas the cladistic approach does not. A more subtle aspect of this problem is that much of the extant taxonomic and evolutionary literature is viewed as reflecting cladistic relationships, but may in fact be based on the traditional 'phenetic/evolutionary' approach, with varying degrees of reliance upon morphological divergence as the main criterion for establishing relationships.

Although molecular data sets can be, and often are, treated in a phenetic manner and are thus open to the same criticism, the phenetic and cladistic approaches when used on the same molecular data sets have often shown remarkable congruence. Additionally, overall molecular divergence data are often analysed with 'phenetic' algorithms (the Fitch and Margoliash algorithm for example) that do not assume equal rates of divergence and thus are not strictly phenetic. One spectacular example will serve to illustrate the problem of relying on the traditional 'phenetic/evolutionary' approach rather than on a cladistic approach, although many other studies undoubtedly will fall into this category as rigorous cladistic techniques are increasingly applied to both morphological and DNA data sets.

During the past 30 years, six different schemes of phylogenetic relationships have been proposed for the Asteraceae (Compositae), with no consensus regarding the number and relationships of monophyletic tribes. One of the reasons suggested as contributing to the
confusion was the lack of application of cladistic principles to morphological characters for phylogenetic reconstruction within the family, at least until recently. A recent morphology-based cladogram for the Asteraceae (Fig. 1a) is startlingly similar to a cladogram based on chloroplast DNA (cpDNA) (Fig. 1b). The cpDNA evidence is based on both comparative restriction-site analysis and the distribution of a major inversion found only in the Asteraceae, excluding the subtribe Mutisieae–Barnadesiinae. The morphological cladistic analysis unfortunately also uses the presence or absence of this inversion, thus making the two approaches not completely independent. Nevertheless, the DNA and morphological cladograms have produced some important complementary results: (1) the small and monophyletic subtribe Mutisieae–Barnadesiinae is placed as the most basal lineage in the family, (2) the tribe Mutisieae and subfamily Cichorioideae are paraphyletic, and (3) the subfamily Asteroideae is monophyletic.

The basal status within the Asteraceae of a number of different tribes has been suggested over the years, but these suggestions lacked rigorous outgroup analysis of all characters. The application of strict cladistic analysis on these morphological characters, rather than using ad hoc assumptions about primitive and advanced character states, allowed for a more objective approach. The close congruence of the DNA and morphological cladograms suggests that a better estimate of the true phylogeny for the Asteraceae is now becoming available. These studies indicate that the failure of applying cladistic methods to morphological characters can be an important reason for the apparent discrepancy of morphology and DNA. This example underscores the need for more detailed cladistic analysis of groups with both morphological and molecular characters.

Equating crossing relationships with phylogenetic relationships

The ability of two plant species to cross as a function of their genetic relationship has long been accepted in the biosystematic literature. It has been argued, and demonstrated, however, that the lack of ability of two species to cross does not necessarily reflect lesser genetic relationship. The ability to cross can in fact be considered a shared primitive character at some taxonomic level, and a barrier to crossing can be considered, in some instances, to be a shared derived character between two closely related species. Equating crossing relationships with phylogenetic relationships can thus potentially yield misleading estimates of the true phylogeny. Two examples, in Clarkia (Onagraceae) and in Eleusine (Poaceae), demonstrate this kind of inconsistency between traditional and molecular data sets.

Clarkia rostrata has been placed within the subsection Sympherica, based on both the overwhelming phenetic similarity and, and the ability to produce fertile hybrids with, the other two species of the subsection (C. lewisi and C. cylindrica) (see Refs 7 and 12). Both isozyme gene duplication data and cpDNA restriction-site analysis indicate, however, that C. rostrata is in fact more closely related to C. epilobioides of subsection Micranthae than it is to the other two species with which it can cross (Fig. 2). The ability to cross within the paraphyletic subsection Sympherica is thus best considered as a shared primitive state, and the lack of crossability between the closely related C. rostrata and C. epilobioides

---

**Fig. 1.** Cladograms of selected tribes and subtribes within the Asteraceae: (a) based on morphological characters and one molecular character (redrawn with permission of the author); and (b) based on restriction-site mutations and one inversion in the chloroplast genome (redrawn with permission of the authors). The shaded bars in (a) represent unresolved nodes.

**Fig. 2.** Cladogram of Clarkia and Heterogaura (Onagraceae), based on cpDNA restriction-site analysis and gene duplications and subsequent losses of activity of orth. Restriction-site changes are indicated by solid vertical bars. The two gene duplications (+1) and subsequent losses (-1) are indicated by open bars (C. cytoplasmic P. plastid).
is best considered as a derived state.

Incongruities between very detailed cytological/crossing studies and DNA results have been demonstrated in domesticated and wild tetraploid finger millets (Eleusine coracana subsp. coracana and africana) and a putative diploid progenitor (E. indica)14. Eleusine indica was first considered to be one of the diploid progenitors for the tetraploid finger millets on the basis of morphology and biogeography. It was later concluded, however, that E. indica could not have contributed any of the genomes of finger millet, due to the almost complete lack of chromosome pairing in synthetic hybrids between subspp. coracana or africana and indica. The cDNA's of the three, however, are identical and differ substantially from other diploid species, suggesting that E. indica served as the maternal progenitor for the tetraploids15. Chromosome pairing behavior was argued as not accurately reflecting phylogenetic affinities in these species.

Inappropriate molecule(s) for taxonomic level or question addressed

The choice of molecule or DNA sequence to be examined is an important first step in any evolutionary or systematic study. A priori knowledge of rates of change, amounts of variation at different taxonomic levels, considerations of homology, and degree of expected homoplasy, is essential before and after data are collected. As more and more studies involving a given molecule or sequence are done, the value and limitations of each molecule or sequence at different taxonomic levels become more evident. Thus, for example, cpDNA restriction-site analysis in angiosperms has been occasionally useful at the species level, extremely useful at the generic and familial level, but seldom used at the ordinal level or above1. Similarly, analysis of restriction-site and fragment-length variation in the non-coding spacer of nuclear ribosomal DNA (rDNA) is possible within and among populations and among congeneric species16,18, but it is analysis in the coding portions of rDNA that is useful only at higher levels18,19.

Perhaps more of a problem in estimating the true phylogeny is that certain portions of a given molecule or nucleic acid sequence evolve at quite different rates. One portion of the sequence will exhibit too few shared mutations between species, while other portions will have multiple substitutions and thus show 'evolutionary noise' between species. The use of 5s rRNA sequence data to reconstruct relationships in green plants demonstrates these problems.

Phenetic analysis using an evolutionary distance modified for transversion: transversion bias, resulted in the phenogram presented in Fig. 3a (Ref. 18). This phenogram has some unusual relationships relative to currently accepted ideas based on traditional kinds of data. The two most unusual include the placement of the Bryophyta (mosses, hornworts and liverworts) within the lower vascular plants rather than basal to all vascular plants, and wheat with tomato rather than with duckweed, the other monocot.

Cladistic analysis of the same data set generated numerous equally parsimonious cladograms, one of which is presented in Fig. 3b (Ref. 19). This cladogram shows unusual placements of the Bryophyta (now as the sister group to the gymnosperms and flowering plants), wheat, as well as the green alga Chlamydomonas. The latter occupies the most basal position in the phenogram, but is nested within the angiosperms in the cladogram. The wide departure of all most parsimonious cladograms from expected relationships and from the phenogram was attributed in part to a high level of homoplasy (43%) - 'evolutionary noise' - in some portions of the 5s rDNA data set.

The separation of 'noisy' characters from those providing meaningful estimates of the true phylogeny is no easy task. It requires a considerable amount of a priori knowledge about the structure and evolution of 5s rRNA, weighting of characters or character states, elimination of hypervariable sites, and/or some amount of (necessary?) circular reasoning involving secondary analysis of the data by a priori defining of already accepted groups.

It is noteworthy that phenetic cluster analysis resulted in less distorted relationships (in reference to generally accepted relationships) than did cladistic analysis of the same data. Perhaps at great evolutionary divergences (e.g. all green plants) and with data sets exhibiting considerable levels of convergence (e.g. 5s rRNA), distances rather than cladistic procedures provide better estimates of true relationships. In simulation studies involving amino acid sequence data at great evolutionary distances, a genetic distance analysis that did not assume a molecular clock resulted in better phylogenetic approximations of the known true phylogeny than did the more 'cladistic' ancestral sequence method20. It has been demonstrated that when large differences in rates of evolutionary change exist between characters, cladistic methods

**Fig. 3. (a) Phenogram of green plants based on 5s rRNAs** (redrawn with permission of the authors). (b) Cladogram of green plants based on 5s rRNAs (redrawn with permission of the authors).
that attempt to minimize the total number of character state changes (parsimony) do not work as well as other methods such as character compatibility or likelihood estimation.

Unequal rates of morphological evolution

As described earlier, reliance on phenetic similarity rather than on shared derived character states, whether morphological or molecular, can give rise to instances of apparent conflict between the two types of data set. However, in cases where wholesale changes have occurred rapidly in one lineage, it is possible that both phenetic and cladistic analyses of morphological data would have difficulty in aligning this lineage correctly with sister lineages. A number of recent molecular studies have demonstrated such discrepancies.

One dramatic increase in rate of morphological divergence giving rise to a discrepancy between traditional and molecular techniques is seen in Heterogaura (Onagraceae). The monotypic genus is closely related to Clarkia based on floral morphology, stamen surface, seed coat structure, anther anatomy and flavonoids, but differs markedly from Clarkia in having only four fertile anthers (four are sterile), an unlobed stigma, and a round nut-like indehiscent fruit with one or two seeds (Fig. 4). The floral and fruit differences between the two genera are so distinctive that they have been maintained as separate genera since 1866. Chloroplast and nuclear rDNA analyses, however, indicated that Heterogaura is actually derived from within Clarkia subsection Lautiflorae, with C. dudleyana as its sister species (Fig. 2). The extreme floral and fruit reduction seen in Heterogaura relative to Clarkia masked the close phylogenetic relationship of Heterogaura to an advanced subsection within Clarkia.

Cladistic analysis of morphological characters would have difficulty in aligning Heterogaura properly not only because of its lack of great modification of some characters important in defining monophyletic groups within Clarkia (e.g. capsular and anther characters), but also because there is no a priori reason not to accept Heterogaura as the sister group for all of Clarkia. Synapomorphic (shared derived) morphological characters unifying Heterogaura with these Clarkia species might exist, but they will be found a posteriori and as a direct result of the compelling isozyme and cpDNA data. To date, however, no morphological character has yet been identified that unites Heterogaura solely with Clarkia subsection Lautiflorae, let alone C. dudleyana.

Rapid morphological change in other monotypic genera that has obscured close phylogenetic relationships to species within more species-rich genera has also been documented in the Saxifragaceae.

The genera Bensoniella and Comiteilla, like Heterogaura, have undergone unusually rapid floral and fruit evolution relative to their closest relatives in the progenitor genus Mitella. Evidence from cpDNA has also indicated similar kinds of rapid but asymmetric rates of morphological evolution within monophyletic lineages of the Oncidinae (Orchidaceae), again demonstrating a surprising degree of plasticity in floral morphology. The 'mule-ear succulent vegetative condition seen in the genus Trichocentrum and one series within the genus Oncidium has been viewed as arising by parallelism, because the two groups are radically different in floral morphology. Trichocentrum, however, has a chloroplast genome that is closest to this 'mule-ear' series within the larger Oncidium. Thus, it is the vegetative features previously considered to be convergent that provide the meaningful phylogenetic relationships in the group, and not the 'conservative' floral features previously emphasized in cladistic studies of orchids.

Convergence of morphological and molecular characters

Rapid and plastic changes in (supposedly) the same (and thus taxonomically important) floral or fruit characters is only one inherent problem in estimating phylogenetic relationships with morphological characters. Morphological convergence or parallelism is another problem that can result in misleading phylogenetic relationships. Two recent studies, in the subtribe Oncidinae (Orchidaceae) and in Polystichum (Dryopteridaceae), exemplify this problem.

A cpDNA phylogeny of the subtribe Oncidinae based on comparative cpDNA restriction-site mapping reveals that strong regional evolution in this neotropical group of orchids has produced quite similar floral morphologies in parallel among relatively distantly related species. For example, several sections of Oncidium in Brazil are phylogenetically more closely related, based on DNA, to other genera in the same region than they are to other sections of Oncidium with similar floral morphologies, both gross and detailed, but occurring in
other geographical areas. The similarity of floral morphologies seen in unrelated sections of Oncidium is a striking case of floral parallelism that had previously misled taxonomists into assuming that these sections, although geographically widespread, were closely related on the basis of floral synapomorphies. A cpDNA phylogeny for Polystichum and related fern genera indicates that frond venation, traditionally important in these genera, is fraught with convergences, but that rhizome characters provide results congruent with cpDNA.

Although convergence in molecular characters occurs, in some instances to a considerable extent (as noted above in 5S sequence data for all plants), the molecular data sets are more amenable to dealing with convergences than are their morphological counterparts. This difference is due to our knowledge of the genetic basis for convergences in molecules as opposed to in morphology. For example, the relative probabilities of different types of restriction-site convergence are well known, allowing for further discrimination among equally short cladograms. Also, the bias towards transitions rather than transversions in most DNA sequence data is documented, permitting additional analysis on sequence data with high levels of convergences.

Hybridization and/or introgression

A number of cases of discrepancies between morphology and DNA have been attributed to hybridization and/or introgression between two species. These gene movements can lead to situations in which maternally inherited genomes (chloroplast and mitochondrion in most cases) of one species are more closely related to those of a distinct species than to those of other conspecifics. Thus, contradictory cpDNA and nuclear phylogenies obtained in two accessions of Brassica napus led to the hypothesis that introgressive hybridization has been involved in their recent evolution. In this case the second species providing the cytoplasm (including the chloroplast genome) could not be identified. Likewise, one of 13 examined cultivated lines of the garden pea Pisum sativum did not exhibit the chloroplast genome of northern populations of P. humile, the putative wild progenitor. Rather, this line accession had a chloroplast genome similar to P. elatius and southern populations of P. humile, indicating that secondary hybridization may have occurred in the domestication of the garden pea.

In a recent cpDNA and rDNA phylogenetic analysis of Populus (Salicaceae), the chloroplast genomes of all varieties of the European black poplar (P. nigra) were distinctly different from those of the American cottonwoods (P. deltoides and P. fremontii), all placed in section Algeiros. However, the black poplar chloroplast genome was most similar to that of the European white poplar (P. alba) of section Populus. The nuclear genome of the black poplars, however, was distinct from both the cottonwoods and the white poplar, suggesting that the origin of the black poplar included hybridization with the white poplar as the maternal parent. It is noteworthy that the black and white poplars can be forced to cross, but only with the latter as the maternal parent.

Surprisingly, molecular data contradict a classic example of introgression based on morphological evidence – the recent origin of weedy races of Helianthus bolanderi by the introgression of genes from H. annuus into a serpentine race of H. bolanderi. Isozyme, rDNA and cpDNA data strongly suggest that the weedy race of H. bolanderi is relatively ancient in origin and did not arise by introgression. It should be noted that phylogenetic sorting of matrarchial lineages during speciation, independent of hybridization, can in theory also give rise to discrepancies between species boundaries and chloroplast or mitochondrial genealogies.

When the time since separation of two species is relatively short, the probability is highest that some chloroplast or mitochondrial lineages from one species will be more closely related to lineages from another species than they are to other lineages from the same species (i.e. the first species appears either paraphyletic or polyphyletic). No examples of this phenomenon are yet known in plants.

Polyplody

Polyplody can show several kinds of discordance between morphology and molecular results. One kind of discrepancy arises when reciprocal crosses between two diploid species generate the allopolyploid. Aegilops triuncialis thus appears diphylectic with two quite different maternal chloroplast genomes, due to independent and reciprocal crosses between the progenitor species A. caudata and A. umbellulata, each acting as the chloroplast donor in turn. A second kind of discrepancy arises when a suite of closely related diploid species can generate the allopolyploid. The tetraploid Andean Solanum tuberosum subsp. andigena complex displays several chloroplast genomes due to crossings of numerous cultivated diploid species. A third kind of discrepancy arises when allo- or autopolyploids originate independently numerous times and in different places, such as in the allopolyploid moss Plagiomnium medium and in autopolyploid populations of Heuchera micrantha (Saxifragaceae).

Conclusion

For most studies in which both DNA and morphology have been examined, the majority of the results are encouragingly congruent. Many of these studies, however, show some instances of discrepancy. These discrepancies in most cases can be attributed to certain problems with techniques and procedures, or to biological phenomena that provide alternative perceptions on how evolution has progressed. In many of the examples cited here, the discovery of the discrepancies has actually initiated studies to address new questions and avenues of research. For example, the robustness with which various molecular data place the genus Heterogaura within Clarkia now provides a model system to re-examine the rate and type of, and genetic basis for, morphological evolution in plants.

Where is plant phylogenetics heading? Undoubtedly, the new molecular evidence is forcing a serious reassessment of relationships based on the traditional systematic approaches. The examples
reviewed here, in which problems can arise from strict dependency on either molecular or morphological characters. Characterization indicates that we should cease the debate about "molecules versus morphology" and begin to cooperate on the basis of molecules and morphology. An interesting by-product of the emergence of molecular studies is their impact on morphological studies. Because of both the presence and the absence of congruence of phylogenetic results in the two kinds of study, there has been a resurgence of interest in more critical cladistic analysis of morphological characters, and a movement towards studies that involve both morphology and molecules. It is just such a kind of multidisciplinary study that ultimately will have the most impact and benefit for the issue of molecules and morphology.

The utilization of DNA approaches will become greater and even more powerful as rapid sequencing via polymerase chain reaction mediated amplification of specific DNA segments becomes routine in phylogenetic studies. The phylogenetic potential of immense molecular data sets will be realized, however, only if the theoretical bases for phylogenetic analysis are refined, incorporation of molecular evolution, and the development of sophisticated computer programs to handle the data can keep pace with the actual data collection. The utilization of DNA approaches to clarify relationships that are difficult to ascertain with the actual data development of sophisticated computer programs to handle the data will also increase, especially in multidisciplinary studies that ultimately will have the most impact and benefit for the issue of molecules and morphology.

A final aspect of the molecular and morphology dilemma that will have to be addressed is how to deal with both molecular and morphological characters in the inference of plant phylogeny. Four possibilities exist and merit examination, although in practice only the first three have been considered. (1) Combine the two sets of characters with equal weighting for each character. In this way the morphological characters might provide strength to branches that are weak based on molecular characters, and vice versa. A variant on this approach, but perhaps difficult to justify and implement, is to provide equal weighting for each set of characters, thus not discriminating against the (usually) fewer molecular characters. (2) Analyze the two sets of characters independently and construct a consensus cladogram that depicts only relationships recognized by both approaches. (3) Use only the molecular characters to generate a cladogram and secondarily overlay the morphological characters onto the molecular cladogram. (4) Use only the morphological characters to generate a cladogram and secondarily overlay the molecular characters onto the morphological cladogram.

These issues related to handling both or many kinds of data set in inferences of plant phylogeny must be dealt with if we are going to cooperate successfully on the basis of molecules and morphology.

References