

Genomics, bioinformatics, and plant systematics

AARON LISTON

Department of Botany and Plant Pathology, Oregon State University, Corvallis, Oregon, 97331-2902 USA

(Received 24 April 2002)

ABSTRACT

Genomics and bioinformatics are emerging disciplines that have the potential to fundamentally alter current approaches to molecular systematics. The sequencing of eukaryotic genomes has demonstrated the ability to obtain and interpret nucleotide sequences at a scale and efficiency that was unimaginable just a few years ago. The immense amount of sequence information generated by genome projects has spurred the development of new analytical and data management tools. How can plant systematists take full advantage of these advances? Comparison of plant nuclear genomes will allow the identification of conserved nonrecombining and single copy loci that can serve as robust phylogenetic markers. The growth of molecular phylogenetic studies is dramatically altering organismal classification. The successful integration of these changes will require innovations in the communication of taxonomic information. Reductions in the cost of obtaining and analyzing nucleotide sequences should result in more time and resources being directed to the essential activities of systematists—the documentation and organization of biological diversity, and the interpretation of this diversity in the context of organismal evolution.

INTRODUCTION

The term bioinformatics is most commonly associated with the analysis of data generated by molecular biology. Specifically, genomics, the study of the nucleotide sequence of organismal genomes, and proteomics, the record of all proteins produced by a genome, are viewed as the frontiers of bioinformatics. The informatics challenge in these fields is turning the vast amounts of genomic and proteomic data into understandable and useful information. Systematics, the science of organismal diversity, has long faced a similar challenge: Documenting the earth's organisms, and placing this sometimes overwhelming diversity into a comprehensible and practical framework. Thus, it should not be surprising that systematists have made important contributions to the development of bioinformatic tools (e.g., Swofford, 1998; Maddison and Maddison, 2000), and are key providers of bioinformatic information (Bisby, 2000). Likewise, the practice of systematics may soon be transformed by advances in the fields of genomics and bioinformatics.

In this paper, I will explore the interface of genomics, bioinformatics, and systematics, focusing on the discipline of plant systematics.

GENOMICS

Genomics is dedicated to the understanding of biological phenomena utilizing the nucleotide sequence of organismal genomes. The current stage of genome sequencing has much in common with natural history endeavors of the 19th century: The science is focused on the discovery of genomic features, their description, and interpretation based on these observations. Experimental approaches are becoming widespread (e.g., gene expression studies using genomic microarrays), but still require technological and analytical advances before they can be considered well-established (Finkelstein et al., 2002; Li et al., 2002). As in natural history, comparisons among organisms provide significant insights into function and evolution. The comparative approach to

E-mail: liston@bcc.orst.edu

genomics is well-established in prokaryotes (Koonin et al., 2000), and developing rapidly in eukaryotes (Wood et al., 2002). In plants, comprehensive comparisons of the *Arabidopsis* and rice genomes are now possible (Bennetzen, 2002).

Although it has been frequently stated that a strength of molecular systematic approaches is the unlimited number of nucleotide sequence characters available (Avice, 1994), this potential has not been exploited. To date, molecular phylogenetics of plants has been based on a tiny number of loci. The vast majority of studies are based on either a limited number of chloroplast genes, or the nuclear regions that transcribe the ribosomes (Soltis and Soltis, 1998). The mitochondrial genome, of great utility in animal systematics, has a very low rate of sequence evolution in plants (see Palmer et al., 1999, for notable exceptions). Thus, its application has been restricted to the analysis of deep divergences (Qiu et al., 1999).

In recent years, there has been considerable interest in low-copy number nuclear genes for phylogenetic reconstruction. Typically, genes that have been well-characterized as a consequence of their biologically or economically important function are used. Such genes are prone to strong positive selection, and thus may not be appropriate for phylogenetic analysis. An analysis of floral homeotic genes in Gesneriaceae has uncovered incongruence between phylogenetic estimates based on the floral gene *cycloidea* and plastid and ribosomal genes (Smith et al., 2002). It is widely acknowledged that phylogenetic analysis of low-copy nuclear loci requires careful attention to gene orthology (Wendel and Doyle, 1998). Despite this awareness, the prevalence of naturally occurring and PCR-mediated recombination (Cronn et al., 2002) can impede accurate phylogenetic reconstruction. The widespread incorporation of nuclear loci into plant molecular systematists would be greatly facilitated by the identification of genes that are confirmed single copy and/or consistently found in regions of the genome that are not subject to recombination.

Arabidopsis and rice represent the two largest clades of flowering plants, the eudicots and monocots. Comparison of their genomes will, for the first time, open the entire nuclear genome to plant molecular systematists. The identification of genes in the centromeric regions of *Arabidopsis* (Copenhaver et al., 1999) was a surprising genomic discovery. If the same genes are present in centromeric locations in rice, and have an appropriate level of sequence variation, these could become ideal phylogenetic markers for angiosperm systematics. Comparison of the *Arabidopsis* and rice genome may also identify loci, or groups of loci, that are in conserved positions and single copy. These could then be surveyed for copy number in other flowering plants, in order to discover those that are

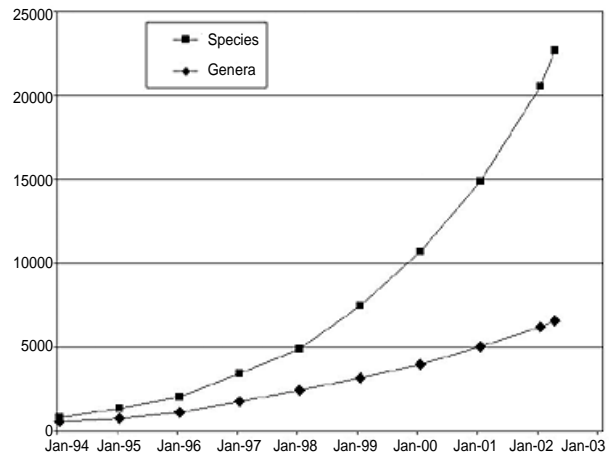


Fig. 1. Increase in the number of DNA sequences for species and genera of angiosperms. Data from National Center for Biotechnology Information (NCBI, 2002), accessed April, 2002. Approximately 7% of all angiosperm species, and 50% of all angiosperm genera, are represented.

consistently present. Conservative exons identified in this manner would be ideal for phylogenetic analyses in flowering plants at the genus level and above. With luck, some may be extended to gymnosperms, ferns, and bryophytes. However, to guarantee coverage of all land plants (embryophytes), sequencing of a genome from each of these major lineages will be required.

This suggestion, the sequencing of a gymnosperm, fern, and bryophyte genome, would have seemed whimsical just a few years ago. However, using today's technology, the speed and efficiency of genome sequencing is constantly increasing. Technological advances in DNA sequencing are likely to provide dramatic increases in efficiency and reductions in cost in the coming years. Leroy Hood (2002) has predicted that in "ten years the equivalent of the human genome can be sequenced in one day for \$10,000." For molecular systematists, this suggests that the amount of sequence information included in a typical study will increase 10–1000 fold, and taxonomic sampling will only be limited by access to material. The exponential growth in the number of sequences deposited in Genbank, the public DNA sequence database, shows no sign of abating. In parallel, the number of plant taxa represented by DNA sequences will continue to rise (Fig. 1).

PLANT SYSTEMATICS

Above the species level

The angiosperms comprise ca. 320,000 species (Prance, 2001), and from a molecular systematic perspective, are

2000?

perhaps the most thoroughly characterized large clade. The plastid gene encoding the large subunit of ribulose-bisphosphate carboxylase-oxygenase (*rbcL*) has been sequenced for over 5000 angiosperm species. Published analyses have included up to 2230 species (Kallersjö et al., 1998) and nearly complete representation of angiosperm families (Savolainen et al., 2000). A slightly less comprehensive study adds a second plastid gene, *atpB*, and the small subunit nuclear ribosomal DNA (Soltis et al., 2000). These studies are largely consistent with the landmark *rbcL* study of Chase et al. (1993), that formed the basis for a revised ordinal classification of the flowering plants (Angiosperm Phylogeny Group, 1998). A major result is the identification of 54 strongly and consistently supported monophyletic groups composed of 1–34 families each (Stevens, 2001). These have been classified as orders, however, relatively few directly correspond to orders recognized in previous classifications. For example, in comparing Cronquist (1981) to Stevens (2001), only Dipsacales, Zingiberales, and the monotypic Arecales are identical.

Twenty-five years ago Vernon Heywood wrote: “Whilst there is a basic agreement between different authors as to the 200 or so ‘core’ families to be recognized, there is no such consensus as regards the next higher level, the order, and the different systems of classification differ widely not only in the number of orders ... recognized but also in their names and content. Fortunately, orders are not widely used ...” (Heywood, 1978, p. 11). As a result of the molecular phylogenetic studies described above, the situation today is completely reversed. I predict that in coming years the orders recognized in the Angiosperm Phylogeny Group classification (with some modifications, see Stevens, 2001) will be widely adopted as the “natural” units of flowering plant diversity. In contrast, many of Heywood’s “core” families are being challenged. In the flora of Israel, examples include the transfer of *Cyclamen* L. from Primulaceae to the Myrsinaceae (Kallersjö et al., 2000) and the removal of most Scrophulariaceae (except *Scrophularia* L. and *Verbascum* L.) to Plantaginaceae (Stevens, 2001) or Veronicaceae (Olmstead et al., 2001), depending on nomenclatural preference. As the comprehensive molecular sampling of flowering plants extends from nearly all 400 families (Savolainen et al., 2000) to the ca. 13,000 genera (Mabberley, 1997), additional recircumscription at the rank of family is expected.

Nomenclature

The nomenclatural instability anticipated from molecular phylogenetic advances may be alarming. This destabilization also runs counter to the prevailing sentiment

expressed in the current International Code of Botanical Nomenclature (ICBN, 2000). In the past decades, plant systematists have made a concerted effort to reduce the incidence of name changes, through modifications of the nomenclatural code. These changes include a relaxation of the constraints of the priority principle, which requires the use of the earliest published name for a taxon (McNeill, 2000). The “conservation” of later names is simplified, and a resolution advising against making changes based solely on bibliographic findings was accepted in 1993 (McNeill, 2000). These reforms are designed to lead to a more stable system of nomenclature. A proposal designed to further reduce nomenclatural clutter, a requirement for registration of new names, was defeated at the last Nomenclatural Session at the International Botanical Congress (ICBN, 2000).

The Code of Botanical Nomenclature has never legislated the taxonomic circumscription of names, that is, whether taxa should be recognized as distinct or synonymized. This is the area that stands to receive the brunt of the molecular systematic hurricane. I suspect that all systematists agree that when a taxon is determined to be polyphyletic, some renaming is inevitable. Whether this should be applied to paraphyletic lineages is a contentious issue (reviewed in Knox, 1998), and the debate shows no sign of abating. Nevertheless, the dominant trend today, especially among molecular systematists, is to only name monophyletic lineages. This is the rationale for the family name changes in Primulaceae and Scrophulariaceae described above.

Adding to the potential for nomenclatural confusion, an alternative set of nomenclatural rules, the PhyloCode, has been proposed (Cantino and de Queiroz, 2000). The PhyloCode implements an explicitly phylogenetic system of nomenclature, and is considered by its proponents as a successor to the principles in place for over two centuries (Cantino and de Queiroz, 2000). The PhyloCode has attracted significant attention and debate (reviewed in Bryant and Cantino, 2002). Implementations have begun to appear in the plant systematic literature (e.g., Olmstead et al., 2001), despite the fact that the PhyloCode remains incomplete, especially in regard to naming species (Cantino et al., 1999). Regardless of the ultimate fate of the PhyloCode, efforts to catalogue and communicate organismal diversity may need to maintain two alternative systems of nomenclature for some groups.

Bioinformatics and databases

Communication is essential to the mission of systematics. A central question that systematists address is “what are the world’s species?” (SA2000, 1994). To answer this question, systematists have developed a number of

international efforts and accompanying databases (Heywood, 2001). The success of these efforts is dependent on the existence of local and regional enumerations. Clara Heyn played a leading role in two such works, the Med-Checklist (Greuter et al., 1984–1989) covering the flora of the Mediterranean Basin, and the *Conspectus florae orientalis: An annotated catalogue of the flora of the Middle East* (Zohary et al., 1980–1994). The importance of these and related publications is in synthesizing and reconciling the information dispersed in geopolitically-based floristic treatments.

An impending challenge for all such endeavors, from local to global, is assimilating the classification changes resulting from the ever-increasing number of molecular phylogenetic studies. The National Center for Biotechnology Information, the home of Genbank, has perhaps done the best job of maintaining an up-to-date taxonomic database, incorporating the most recent classifications (Wheeler et al., 2000). A staff of taxonomists and bioinformatics experts maintains the “Taxonomy Browser”, essentially, the taxonomic scaffolding for the millions of sequences added annually. As of early 2002, over 117,000 species are included. A limitation of the Taxonomy Browser is a lack of synonymy and retrospective information. One cannot follow the history of nomenclatural changes, nor trace an older name. With their experience with these issues, systematists can make important contributions to taxonomic databases that track and archive changes. With the anticipated increase in nomenclatural changes due to molecular phylogenetic results, such features will go from desirable to essential.

with?

Another area in need of development is associating the names used for organisms directly to the phylogenetic hypotheses (trees) these names are based on. Databases of phylogenetic trees exist (TreeBASE, 2002; Maddison, 2002), as do various nomenclatural databases (Bisby, 2000). It may be possible to integrate these using the architecture of genomics databases. For example, the virtual representation of a genome is linked to its constituent sequences, and tools allow for viewing of the genome at different scales of resolution (Kent et al., 2002). An analogous system for systematics would allow the user to move from a taxonomic name to its position in one or more phylogenetic trees, to the underlying character data, and provide for visualizing the trees and data at different levels of resolution. Needless to say, the above scenarios assume that the majority of taxonomic information will be disseminated and accessed electronically (Godfray, 2002).

Prospects for the future

The growth of DNA sequence information has been, and will continue to be, characterized by increased automa-

tion. Each advance in sequencing technology further reduces the demand for human involvement. Currently, an increasing amount of the systematist’s time is devoted to the analysis of the resulting data. Improvements in the methods of phylogenetic analysis have facilitated the analysis of large data sets; nevertheless, serious methodological and theoretical challenges remain (Sytsma and Pires, 2001), especially as the size of data matrices continues to expand. As the methods become even more computationally intensive, it seems likely that innovations will be made by computer scientists and data processing experts, and not systematists. If the data gathering is automated, and the data analyses are best done by others, what is left for the systematist? In my view, this situation frees the systematist to do what s/he is uniquely qualified for—the documentation and organization of biological diversity and the interpretation of this diversity in the context of organismal evolution.

Recent decades have seen a decline in “experimental studies on the nature of species” (Clausen et al., 1940) conducted by plant systematists. Among the majority of systematists, the focus has apparently shifted from investigating the process of evolution to documenting its resulting pattern (Knox, 1998). By integrating the experimental approaches of the biosystematists (Stebbins, 1970), the tools of molecular biology, and the methods of phylogenetic analysis, systematists can make important contributions to our understanding of the process and pattern of organismal evolution (Sytsma and Pires, 2001).

Large gaps still remain in our knowledge of the morphological (and functional) features of many taxa (Stevens, 2000). These lacunae become apparent when molecular phylogenetic results “change” the long-accepted position of a taxon (e.g. *Platanus* L. and *Simmondsia* Nutt.), forcing a reevaluation of past morphological interpretations (Stevens, 2000). In *Garrya* Lindley (Garryaceae), the molecular confirmation of a sister-group relationships with *Aucuba* Thunb. (Aucubaceae) led to the clarification of its floral morphology, and the resolution of ambiguities that existed in the descriptive literature for over 150 years (Liston, 2002). This example demonstrates how molecular phylogenetic results can contribute to a better understanding of morphological features.

The documentation of earth’s biota is not keeping pace with the rate of habitat destruction (Myers et al., 2000; Pimm and Raven, 2000). There is an urgent need to determine what exists before it is gone. This crisis situation makes especially appealing the prospect of weaning molecular systematists from the DNA sequencer to the field and to museum collections. This does not assume a reduced interest in the results of

molecular phylogenetics, but rather a recognition that these efforts can be conducted by others, provided that systematists provide the appropriately collected and documented starting materials. If ongoing efforts to increase the financial support for taxonomic activities are successful (All Species Foundation, 2002), we need to ensure that the personnel are available to “discover, describe, and classify” (SA2000, 1994) organismal diversity. Furthermore, the obligation to train a new generation of taxonomists, particularly in biologically-rich regions of the world, cannot be neglected.

ACKNOWLEDGMENTS

I thank Uzi Plitmann for inviting me to contribute to this issue dedicated to Clara Heyn, and for his helpful comments on an earlier version of the manuscript. Portions of this paper were presented at the XV Congreso Mexicano de Botánica, Queretaro, Mexico in October, 2001. I would also like to acknowledge Clara Heyn, who was responsible for my career in plant systematics. In 1982, she took a chance, hiring an American undergraduate with limited botanical experience for a technician position in The Hebrew University Herbarium. In addition to introducing me to herbarium curation and specimen-based research, Clara supervised my first botanical field work; collecting seeds of native Leguminosae for a plant genetic resources project. Through this project I discovered the joys of botanical field work, a passion that has continued to this day. The germplasm project also encompassed a pioneering databasing effort led by Shulamith Weitz-Ingber. Clara was an important supporter of this database, which was, in many ways, years ahead of its time (Weitz-Ingber, 1983). The efficient documentation, organization, and dissemination of this germplasm information was a central concern of Clara and Shula, and today their efforts would be classified as bioinformatics.

REFERENCES

- All Species Foundation. 2002. <www.all-species.org>
- Angiosperm Phylogeny Group. 1998. An ordinal classification for the families of flowering plants. *Ann. Mo. Bot. Gard.* 85: 531–553.
- Avise, J.C. 1994. *Molecular markers, natural history and evolution*. Chapman and Hall, New York.
- Bennetzen, J. 2002. The rice genome: opening the door to comparative plant biology. *Science* 296: 60–63.
- Bisby, F.A. 2000. The quiet revolution: biodiversity informatics and the internet. *Science* 289: 2309–2311.
- Bryant, H.N., Cantino, P.D. 2002. A review of criticisms of phylogenetic nomenclature: is taxonomic freedom the fundamental issue? *Biol. Rev.* 77: 39–55.
- Cantino, P.D., Bryant, H.N., de Queiroz, K., Donoghue, M.J., Eriksson, T., Hillis, D.M., Lee, M.S.Y. 1999. Species names in phylogenetic nomenclature. *Syst. Biol.* 48: 790–807.
- Cantino, P.D., de Queiroz, K. 2000. PhyloCode: a phylogenetic code of biological nomenclature. <www.ohiou.edu/phylocode/>
- Chase, M.W., et al. (41 co-authors) 1993. Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. *Ann. Mo. Bot. Gard.* 80: 528–580.
- Clausen, J., Keck, D.D., Hiesey, W.M. 1940. *Experimental studies on the nature of species*. Washington, D.C. publisher?
- Copenhaver, G.P., Nickel, K., Kuromori, T., Benito, M.I., Kaul, S., Lin, X., Bevan, M., Murphy, G., Harris, B., Parnell, L.D., McCombie, W.R., Martienssen, R.A., Marra, M., Preuss, D. 1999. Genetic definition and sequence analysis of Arabidopsis centromeres. *Science* 286: 2468–2474.
- Cronn, R., Cedroni, M., Haselkorn, T., Grover, C., Wendel, J. 2002. PCR-mediated recombination in amplification products derived from polyploid cotton. *Theor. Appl. Genet.* 104: 482–489.
- Cronquist, A. 1981. *An integrated system of classification of flowering plants*. Columbia University Press, New York.
- Finkelstein, D., Ewing, R., Gollub, J., Sterky, F., Cherry, J.M., Somerville, S. 2002. Microarray data quality analysis: lessons from the AFGC project. *Arabidopsis Functional Genomics Consortium. Plant Mol. Biol.* 48: 119–31.
- Godfray, H.C.J. 2002. Challenges for taxonomy. *Nature* 417: 17–19.
- Greuter, W. Burdet, H.M., Long, G., eds. 1984–1989: *Med-Checklist. A critical inventory of vascular plants of the circum-mediterranean countries*. OPTIMA, Geneva and Berlin.
- Heywood, V.H. 1978. *Flowering plants of the world*. Lexington Books, New York.
- Heywood, V.H. 2001. Floristics and monography—an uncertain future? *Taxon* 50: 361–380.
- Hood, L. 2002. *Seminar at Oregon State University*. February, 2002. on what?
- ICBN. 2000. *International code of botanical nomenclature (St Louis Code)*. Regnum Vegetabile 138. Koeltz Scientific Books, Königstein, Germany, 474 pp.
- Kallersjo, M., Farris, J.S., Chase, M.W., Bremer, B., Fay, M.F., Humphries, C.J., Petersen, G., Seberg, O., Bremer, K. 1998. Simultaneous parsimony jackknife analysis of 2538 *rbcL* DNA sequences reveals support for major clades of green plants, land plants, seed plants and flowering plants. *Plant Syst. Evol.* 213: 259–287.
- Kallersjo, M., Bergqvist, G., Anderberg, A.A. 2000. Generic realignment in primuloid families of the Ericales s.l. (Angiosperms): a phylogenetic analysis based on DNA sequences of *rbcL* and *ndhF*. *Am. J. Bot.* 87: 1325–1341.
- Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., Haussler, D. 2002. The human genome browser at UCSC. *Genome Research*, in press. update?
- Knox, E.B. 1998. The use of hierarchies as organizational models in systematics. *Biol. J. Linn. Soc.* 63: 1–49.

- Koonin, E.V., Aravind, L., Kondrashov, A.S. 2000. The impact of comparative genomics on our understanding of evolution. *Cell* 101: 573–576.
- Li, X., Gu, W., Mohan, S., Baylink, D.J. 2002. DNA microarrays: their use and misuse. *Microcirculation* 9: 13–22.
- Liston, A. 2002. A new interpretation of floral morphology in the genus *Garrya* (Garryaceae). *Botany 2002 Abstract*. www.botany2002.org/section12/abstracts/94.shtml
- Mabberley, D.J. 1997. *The plant-book*, 2nd ed. Cambridge University Press, Cambridge.
- Maddison, D.R. 2002. The tree of life web project. tolweb.org/tree/phylogeny.html
- Maddison, W.P., Maddison, D.R. 2000. *MacClade 4: analysis of phylogeny and character evolution*. Sinauer, Sunderland, Mass.
- McNeill, J. 2000. Naming the groups: developing a stable and efficient nomenclature. *Taxon* 49: 705–720.
- Myers, N., Mittermeier, R.A., Mittermeier, C.G., da Fonseca, G.A.B., Kent, J. 2000. Biodiversity hotspots for conservation priorities. *Nature* 403: 853–858.
- NCBI. 2002. National center for biotechnology information taxonomy database. www.ncbi.nlm.nih.gov/Taxonomy/tax.html
- Olmstead, R.G., dePamphilis, C.W., Wolfe, A.D., Young, N.D., Elisons, W.J., Reeves, P.A. 2001. Disintegration of the Scrophulariaceae. *Am. J. Bot.* 88: 348–361.
- 1999? Palmer, J.D., Adams, K.L., Cho, Y., Parkinson, C.L., Qiu, Y.L., Song, K. 2000. Dynamic evolution of plant mitochondrial genomes: mobile genes and introns and highly variable mutation rates. *Proc. Natl. Acad. Sci. USA* 97: 6960–6966.
- Pimm, S.L., Raven, P.H. 2000. Biodiversity: extinction by numbers. *Nature* 403: 843–844.
- Prance, G.T. 2001. Discovering the plant world. *Taxon* 50: 345–360.
- Qiu, Y.-L., Lee, J., Bernasconi-Quadroni, F., Soltis, D.E., Soltis, P.S., Zanis, M., Zimmer, E.A., Chen, Z., Savolainen, V., Chase, M.W. 1999. The earliest angiosperms: evidence from mitochondrial, plastid and nuclear genes. *Nature* 402: 404–407.
- SA2000. 1994. *Systematics agenda 2000: charting the biosphere*. American Museum of Natural History, New York.
- Savolainen, V., Fay, M.F., Albach, D.C., Backlund, A., van der Bank, M., Cameron, K.M., Johnson, S.A., Lledo, M.D., Pintaud, J.-C., Powell, M., Sheahan, M.C., Soltis, D.E., Soltis, P.S., Weston, P., Whitten, W.M., Wurdack, K.J., Chase, M.W. 2000. Phylogeny of the eudicots: a nearly complete familial analysis based on *rbcL* gene sequences. *Kew Bull.* 55: 257–309.
- Smith, J.F., Baum, D.A., Hileman, L. 2002. Floral evolution and phylogenetic analysis of tribes Gloxinieae and Gesnerieae (subfamily Gesnerioideae: Gesneriaceae) with an emphasis on GCYC (a Gesneriaceae CYCLOIDEA homolog). *Botany 2002 Abstract*. www.botany2002.org/section12/abstracts/3.shtml
- Soltis, D.E., Soltis, P.S. 1998. Choosing an approach and an appropriate gene for phylogenetic analysis. In: Soltis, D.E., Soltis, P.S., Doyle, J.J., eds. *Molecular systematics of plants II: DNA sequencing*. Kluwer Academic, Norwell, Mass. pp. 1–42.
- Soltis, P.S., Soltis, D.E., Chase, M.W., Mort, M. E., Albach, D.C., Zanis, M., Savolainen, V., Hahn, W.H., Hoot, S.B., Fay, M.F., Axtell, M., Swensen, S.M., Price, L.M., Kress, W.J., Nixon, K.C., Farris, J.S. 2000. Angiosperm phylogeny inferred from 18S rDNA, *rbcL*, and *atpB* sequences. *Bot. J. Linn. Soc.* 133: 381–461.
- Stebbins, G.L. 1970. Biosystematics: an avenue towards understanding evolution. *Taxon* 19: 205–214.
- Stevens, P.F. 2000. Botanical systematics 1950–2000: change, progress, or both? *Taxon* 49: 635–660.
- Stevens, P.F. 2001. Angiosperm phylogeny website. Version 2 August 2001. www.mobot.org/MOBOT/research/APweb/
- Swofford, D.L. 1998. *PAUP* 4: phylogenetic analysis using parsimony (and other methods)*. Sinauer, Sunderland, Mass.
- Sytsma, K.J., Pires, J.C. 2001. Plant systematics in the next 50 years—re-mapping the new frontier. *Taxon* 50: 713–732.
- TreeBASE. 2002. A database of phylogenetic knowledge. www.treebase.org/treebase/
- Wendel, J.F., Doyle, J.J. 1998. Phylogenetic incongruence: window into genome history and molecular evolution. In: Soltis, D.E., Soltis, P.S., Doyle, J.J. eds. *Molecular systematics of plants II: DNA sequencing*, 265–296. Kluwer Academic, Norwell, Mass., USA.
- Wheeler D.L., Chappey, C., Lash, A.E., Leipe, D.D., Madden, T.L., Schuler, G.D., Tatusova, T.A., Rapp, B.A. 2000. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* 28: 10–14.
- Weitz-Ingber, S. 1983. Documentation of intraspecific plant character diversity. *Isr. J. Bot.* 32: 67–74.
- Wood, V., et al. (131 co-authors) 2002. The genome sequence of *Schizosaccharomyces pombe*. *Nature* 415: 871–880.
- Zohary, M., Heyn, C.C., Heller, D. 1980–1994. *Conspectus florum orientalis: an annotated catalogue of the flora of the Middle East*. Israel Academy of Sciences, Jerusalem.