Manthan, International Journal, Vol. 12, June, 2011, Pages 31-32
ISSN No. 0974-6331 www.bbmanthan.info

Plant Taxonomy in the Current Scenario of Molecular Biology and Bioinformatics

M. A. Ali

Department of Botany and Microbiology
College of Science, King Saud University, Riyadh, Saudi Arabia

Plant taxonomy is the science that finds, describes, classifies, identifies, and names the plants. Plant identification is the determination of the identity of an unknown plant by comparison with previously collected specimens or with the aid of books or identification manuals. The process of identification connects the specimen with a published name. Once a plant specimen has been identified, its name and properties are known. Plant classification is the placing of known plants into groups or categories to show some relationship. Plant systematics is involved with relationships between plants and their evolution, whereas plant taxonomy deals with the actual handling of plant specimens.

Modern biological classification has its root in the work of Carolus Linnaeus, who grouped species according to shared physical characteristics. These groupings since have been revised to improve consistency with the Darwinian principle of common descent. The traditional classification of plants into respective classes, orders, families, genera and species has been based on shared morphologic, cytologic, biochemical and ecologic traits. There are several approaches have been put forward time to time for classification of plants viz., form or habit of plant (Theophrastus, Caesalpino), artificial (Tournefort, Linnaeus), natural (Bauhin, Ray, de Jussieu, de Candolle, Bentham & Hooker) and phylogenetic (Engler & Prantl, Bessey, Hutchinson, Cronquist, Takhtajan, Thorne, Dahlgren, Angiosperm Phylogeny Group).

Since the 1960s a trend called cladistic taxonomy (or cladistics or cladism) has emerged, arranging taxa in an evolutionary tree. Molecular systematics or molecular phylogenetics which is an essentially cladistic approach and is the use of the structure of molecules to gain information on an organism's evolutionary relationships was pioneered by Charles G. Sibley (birds), Herbert C. Dessauer (herpetology), and Morris Goodman (primates), followed by Allan C. Wilson, Robert K. Selander and John C. Avise. Early attempts of molecular systematics were also termed as chemotaxonomy and made use of proteins, enzymes, carbohydrates and other molecules which were separated and characterized using techniques such as chromatography. Beginning in the early 1980s and continuing to the present, the use of DNA has represented the "cutting edge" (glamour area) within the entire field of plant systematics. Our understanding of the relationships among organisms at various levels in the tree of life has been advanced greatly in the last two decades with the aid of DNA molecular systematic techniques and phylogenetic theory. A diverse array of molecular techniques are available to the plant systematist for use in phylogentic inference, including restriction site analysis, comparative sequencing, analysis of DNA rearrangements (e.g. inversions), gene and intron loss, and various polymerase chain reaction (PCR) based techniques. These are generally considered superior for evolutionary studies since the actions of evolution are ultimately reflected in the genetic sequences. At present it is still a long and expensive process to sequence the entire DNA of an organism, and this has been done for only a few species. However, it is quite feasible to determine the sequence of a defined area of a particular chromosome. Closely related organisms generally have a high degree of agreement in the molecular structure of these substances, while the molecules of organisms distantly related usually show a pattern of dissimilarity. Molecular phylogeny uses such data to build a relationship tree that shows the probable evolution of various organisms. The most common approach is the comparison of sequences for genes using sequence alignment techniques to identify similarity.

Plant molecular systematics has relied primarily on the chloroplast genome. Nuclear ribosomal DNA is arranged in tandem repeats in one or a few chromosomal loci. Each repeat consists of a transcribed region that comprises an external transcribed spacer (ETS) followed by the 18S gene, an internal transcribed spacer (ITS-1), the 5.8S gene, a second internal transcribed spacer (ITS-2), and finally the 26S gene. Each such repeat is separated from the next repeat by an intergenic spacer (IGS). The nuclear genes that code for rRNA are repeated thousands of times within the typical plant genome. In fact they can comprise as much as 10% of the total plant DNA. The most remarkable feature of rDNA is the overall sequence homogeneity among members of the gene family in a given species. The process by which this pattern of intraspecific homogeneity and interspecific

heterogeneity is maintained has been called concerted evolution. It is widely accepted that in the process of concerted evolution a single mutation can be fixed in a relatively short time period due to unequal crossing over or gene conversion. These homogenization processes have been described as molecular drive. The coding regions show little sequence divergence among closely related species, whereas the spacer regions exhibit higher rates of variability. Therefore, nuclear ribosomal ITS sequence data have a great potential to resolve plant phylogenies at various intrafamiliar levels in angiosperms. Despite the large size of the nuclear genome, most attempts to infer phylogeny with nuclear gene sequences have involved the nuclear ribosomal DNA cistron (rDNA). The approximate lengths of the three coding regions are very similar throughout plants. The 18S gene equals 1,800 bp, the 26S gene equals 3,300 bp, the 5.8S gene equals 160 bp. In contrast, the length of the IGS varies considerably (from 1 to 8 kb). This variation in IGS length is the major contributors to the large range of variation in total length of the repeat unit in plants, ranging from approximately 8-15 kb. Variation in the length of the ITS-1 and ITS-2 regions is also noteworthy. The external transcribed spacer (ETS) region (especially the 3'end of the 5'-ETS adjacent to 18S) has sometimes been exploited in lower-level phylogenetic analyses. The nuclear genome of plants consists of certain DNA sequences that are present once per genome. These are referred to as single copy or unique sequence DNA. The lengths of single copy sequences in plant genomes usually vary from 200 to several thousand bp. Single or low-copy nuclear genes have also great potential to elucidate phylogenetic relationships of plants. The advantages of nuclear genes include the availability of many genes, their overall faster rate of evolution, and their biparental inheritance.

As gene sequencing becomes easier and cheaper, molecular systematics is being applied to more and more groups, and in some cases is leading to radical revisions of accepted taxonomies. The term bioinformatics is most commonly associated with the analysis of data generated by molecular biology. 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. Development of phylogenetic theory and cladistic analysis of DNA sequences data has resulted into phylogenetic classification of the land plants. DNA barcoding -the use of short DNA sequences for biological identifications has gained worldwide attention in the scientific community which revolutionizes our knowledge of plant diversity and is on its way to being accepted as a global standard for the purpose of species identification.

India, with its wide range of physiographic and climatic conditions, has a rich varied flora, unparalleled in any other country in the world. The physiographic diversity of the country has produced all possible types and extremities of climatic conditions suitable for supporting wide varied types of ecosystems. It is estimated that about 45,000 species of plants which forms the conspicuous vegetal cover comprises about 6.8% of all known flowering plants of the world. With 2.46% of land area having 6.8% of flowering plants, India is recognized as one of the top 12 mega-biodiversity centers of the world. However, the Indian plant taxonomist is still cataloguing the life even in the era when Plant taxonomy is being practiced using tools and techniques of molecular biology and bioinformatics. The molecular systematic studies in India on Indian flora is in infancy due to lack of proper training of molecular biology and bioinformatics especially to taxonomist, thus a rich biodiversity of India has remained untouched from molecular systematic studies and in understanding the evolutionary relationship. There is thus an urgent need to review the status of taxonomic studies in India in context with latest development in the discipline. With the rich biological resources and many outstanding botanists who are familiar with the regional flora and interesting systematic questions, should initiate molecular systematic program to advance our understanding on the tree of life and to address new evolutionary questions.

Correspondence:
Phone: +966-4677561
Email: majmalali@rediffmail.com