DNA technology and studies in phylogenetic relationships of tropical plant: Prospect in Indonesia***

Topik Hidayat* and Adi Pancoro**

Department of Biology Education, Faculty of Mathematic and Natural Science Education, Indonesia University of Education (UPI), Bandung, Indonesia
**School of Life Science and Techonology, Institute Technology of Bandung (ITB), Bandung, Indonesia*
***Paper presented at International Conference on Mathematic and Natural Science, Faculty of Mathematic and Natural Science, Institute Technology of Bandung, 29-30 November 2006*

ABSTRACT

Phylogenetic is one of the most preferable methods in systematic to reconstruct evolutionary relationships of groups of biological organisms in order to understand the biodiversity. Until recently, morphological characters have long been preserved in most phylogenetic studies. With the recent advance of molecular techniques such as polymerase chain reaction (PCR) and DNA sequencing, the use of DNA sequences in phylogenetic studies has rapidly increased and been performed at all taxonomic levels. DNA sequences have drawn attention to many researchers due to in facts that, i.e.,: (1) they offer an increased precision by permitting better homology assessment of characters; (2) they have the large differences in substitution rates in different loci, providing the large number of characters; and (3) They provide an independent source with which to test hypotheses on the evolution of form, leading to a new and better supported-work of phylogeny of plant. In respect with many genes are recently available for such studies, the investigator is faced with a question which gene or genome must be selected; this should fit with phylogenetic problem of materials under study. Common choices in plants include the chloroplast genome (cpDNA) or components of the chloroplast genome, the nuclear ribosomal DNA genes (nrDNA), or nuclear-encoded, single-copy genes. Indonesia as a tropical country is inhabited by the large number and great diversity of plant. Many large and important tropical Indonesian plant group present challenges to systematists interested in understanding the evolution, and they, therefore, await exploration. However, phylogenetic studies dealing with these plant groups using molecular data are few. These include ongoing work by Adi Pancoro and Topik Hidayat and their colleagues at Institute Technology of Bandung (ITB), Indonesia University of Education (UPI), and Herbarium Bogoriense on various groups of angiospermae. Several molecular phylogenetic studies that have been done by the Pancoro group's research will be presented.

Keywords: Molecular Phylogenetic, DNA sequences, PCR, Angiospermae

Introduction

It has been long time plant systematists deduced relationships among plant groups based upon a wide variety of biological characters. These characters include morphology at both macro-morphology and micro-morphology, plant secondary metabolites, isozymes, and other characters. However, in some cases, these characters are inconsistent and subjected to parallelism.

In the last few years, many researchers have used molecular characters (read: DNA sequences) as a basis for inferring phylogenetic relationships. Our objectives in this paper are to provide a brief account of the use of DNA sequences in studies of plant phylogenetic and to promote as well as develop such studies in Indonesia. Indonesia, as a country that has a huge and diverse plant groups, has responsibility to conduct the study. However, it has ironically been a fact that Indonesian researchers who conducted phylogenetic study using molecular data are very few.

Goal of Systematics

Systematics plays a central role in the field of biology through providing the means for characterizing the organisms and the recognition in order to understand the biodiversity. Systematics is fundamentally aimed at discovering, describing and naming all the tips of the branches of the tree of life (Gravendeel, 2000) as well as documenting the changes on the branches occurred during evolution and transform these into a predictive classification system that reflect evolution (Systematics Agenda 2000). One of the most important tasks in systematics is, therefore, the reconstruction of the historical relationships of groups of biological organisms. A correctly inferred phylogeny may provide a solid basis of the knowledge of relationships, and is a prerequisite for comparative investigations within such the fields as ecology and biogeography.

Numerous methods for inferring evolutionary relationships have been established, of which cladistics, as outlined by Hennig (1966), Kluge and Farris (1969) and Fitch (1971), is the most common method. In this method, a group of organisms that shares many identical character states is considered to be closely related and is assumed to be derived from a common ancestor. The ancestor and all its descendants will form a monophyletic group. In practice, outgroups are required and cause polarization of characters: apomorphic state (a derived character states occur in the ingroup) and plesiomorphic state (a "primitive" character state occur in the outgroup). Synapomorphic states (shared derived character states) are used

to investigate the relationships between taxa in the ingroup.

To reconstruct phylogenetic relationships, several methods are available recently, of which maximum parsimony is the most optimal tree reconstruction. The principle of this method is that the differences observed among taxa under study are identified by minimization of character transformations on the branches of a tree (Li and Graur, 1991). In many cases, tree reconstructions are solely based on the parsimony criterion.

Further, the reconstructed skeleton of phylogeny can be translated into an unambiguous system of classification, and only monophyletic groups are subjected to this classification because the criterium of common descent is objective and makes the system defendable instead of intuitive (e.g., van Welzen, 1998; Gravendeel, 2000; Grant, 2003). Unlike traditional classification system (the Linnaean system), phylogenetic system classifies one taxon based upon the knowledge of phylogenetic relationships with other taxa (de Queiroz and Gauthier, 1990). In many cases, classifications are, therefore, strictly based on monophyletic groups. Figure 1 exhibits a phylogenetic tree as a result of phylogenetic analysis.

Figure 1 A phylogenetic tree and polarization of characters

The use DNA sequences in phylogenetic analysis

Until recently, morphological characters have long been preserved in most phylogenetic studies. There are two basic methods in performing phylogenetic analysis at DNA level: restriction analysis and DNA sequences analysis. The former is the easiest and most widely employed method. With the recent advances of molecular techniques such as polymerase chain reaction (PCR) and DNA sequencing, however, the use of DNA sequences in

phylogenetic studies has rapidly increased and been performed at all taxonomic levels.

DNA sequences have drawn attention to many researchers to be chosen, perhaps because of offering an increased precision by permitting better homology assessment of characters (Baldwin et al., 1995) and of the large differences in substitution rates in different loci (Moritz and Hillis, 1996). On the other word, DNA character provides the large number of characters for such studies. Further, molecular data provides an independent source with which to test hypotheses on the evolution of form (Endress et al., 2000) and has led to a new and better supported-framework of angiospermae phylogeny (e.g., Chase et al., 1993). However, several criteria of the DNA region should be considered before putting it into practice, for example, the DNA sequences must be orthologous and have an appropriate substitution rate to which the phylogenetic problem at hand.

In plant, there are three genomes (nucleus, chloroplast, and mitochondria) with several genes available for the study recently. Consequently, the researchers are faced with a major question: what appropriate genome or gene that must be chosen to solve the specific phylogenetic question at hand? This is because the different genes evolve at markedly different rates and therefore provide varying degrees of genetic resolution among plant groups. The researchers should also consider which molecular method must be selected; this is related with technical difficulty. If one works with a large number of plant lineages, it is suggested to choose a genome or gene which can be easily assayed.

DNA chloroplast (cpDNA) is the most chosen genome for plant phylogenetic investigation. The reasons are: (1) it is relatively easy to purify, characterize, clone, and sequence (Clegg and Durbin, 1990); (2) it evolves at a conservative rate (Zurawski and Clegg, 1987), providing an excellent window of genetic resolution for phylogenetic studies among major taxonomic groups; and (3) it has a large body of information on the molecular structure and organization to study evolution of chloroplast itself (Clegg et al., 1990).

In contrast with the cpDNA, the use of DNA mitochondria (mtDNA) for the investigation in plant is very limited due to the fact that it is large in size, so that it is more difficult to isolate and purify. Despite protein-coding genes on the mtDNA are well known to evolve at a conservative rate, thus far, very few applications of mtDNA sequence data to plant phylogenetic problems have been published. In the future, the plant systematists should also consider this genome to be applied in their phylogenetic analysis.

A third genome is nuclear DNA (nDNA) that is frequently used in plant phylogenetic

investigation. The main component of this genome is that of nuclear-encoded ribosomal RNA gene family (nrDNA) including internal transcribed spacer (ITS) region. The ITS region includes 5.8S subunit and two spacers (ITS-1 and ITS-2). nrDNA itself is consisted of three genes that encoded small subunit rRNA (16S to 18S), large subunit rRNA (26S to 28S) and 5.8S gene. An external transcribed spacer (ETS) is situated just outside the ITS region. Along with the other components of the nrDNA, the ITS region is highly repeated of up to many thousands copies in plant nuclear genome (Fig. 2). Non-transcribed spacer (NTS) separates from one to another copy.

Nuclear ribosomal DNA

Figure 2 Organization of ITS region of nrDNA

Evolution of nrDNA is complicated because each repeating unit evolves at different rates. The NTS and ITS tend to evolve relatively rapid whereas the coding regions of 18S, 5.8S, and 26S evolve much more slowly (e.g., Hamby and Zimmer, 1988). As consequent, these two rate differentials are used to address different phylogenetic problems. While the NTS and ITS address evolutionary relationships among closely related taxa (intrageneric and intraspecific levels), the coding regions are used to clarify relationships of more distantly related groups.

Further, unlike cpDNA and mtDNA, it is often difficult to obtain an orthologous sequence from nrDNA due to its properties. As mentioned above that this gene family is arranged in several hundreds to thousands or more tandem repeat copies in which these copies do dependently evolve in concert. However, when concerted evolution fails to homogenize the sequences, possibility of several different sequences, which have different evolutionary histories, will be produced. This failure is primarily caused when the arrays of copies occur at

different chromosomal loci. In recent years, however, several phylogenetic studies of plant have applied single- or low-copy nuclear sequences from which an orthologous sequences can be obtained easily.

Prospect of molecular phylogenetic study in Indonesia

Among tropical regions, Indonesia is the best place for many plant groups to live in respect with topography, geography, climate, temperature, ecological interaction, etc. These circumstances have caused remarkable diversification of morphological characters of the plants themselves. Owing to this remarkable diversification and possible parallelism of vegetative and reproductive features, phylogenetic relationships of many tropical plant groups are ill defined, and they, therefore, await exploration. To get comprehensive knowledge on the relationship, comparative phylogenetic study incorporated with molecular data is much needed. It has proven that molecular characters (read: DNA sequences) can reconstruct a more natural scheme of phylogenetic relationship (see e.g., Topik, 2005; Yukawa et al., 2005)

It has been a fact that phylogenetic study based upon molecular characters is less attention to many Indonesian plant systematists, while scheme of phylogenetic relationship revealed from molecular data can be used as a basis of the knowledge of relationships. This deals with question whether previous scheme of relationship is consistent with present molecular tree.

However, several research groups in Indonesia have conducted phylogenetic analysis of tropical plant using DNA sequences. Adi Pancoro, Topik Hidayat, and their colleagues at Institute Technology of Bandung (ITB), Indonesia University of Education (UPI), and Herbarium Bogoriense have worked a lot of molecular phylogenetic analysis on such important tropical Asian plant groups as family Anacardiaceae, family Orchidaceae, genus *Mangifera* (family Anacradiaceae), and genus *Piper* (family Piperaceae). This also includes ongoing and future work on genus *Mangifera* and genus *Phyllantus* (family Euphorbiaceae), and family Pandanaceae. This group's research has made a major contribution to the understanding of the taxonomy, evolutionary relationships and biogeographic origins of these previously poorly understood groups. Further, the molecular phylogenetic analysis that they have commenced represents one of the few systematic studies dedicated to understand evolutionary relationships in a large, complex, tropical plant group using modern molecular techniques.

In order to stimulate the studies of plant phylogenetic using molecular characters in

Indonesia, the Pancoro group's research recently has organized a short course on application of software PAUP and MrBayes in molecular phylogenetic study. Participants of this course included, i.e., taxonomists/systematists, researchers from relevant field, lecturers, that came from various regions of Indonesia. There is a hope that after this meeting, some collaborative activities on molecular phylogenetic can be implemented in the future. In next stage, it is expected that phylogenetic network in Indonesia can be established.

Literature Cited

Baldwin BG, MJ Sanderson, JM Porter, MF Wojciechowski, CS Campbell, and MJ Donoghue (1995) The ITS region of nuclear ribosomal DNA: A valuable source of evidence on Angiosperm phylogeny. Ann Missouri Bot Gard 82:247-277

van Welzen PC (1998) Phylogenetic versus Linnaean taxonomy, the continuing story. Taxon 47:413-423

Chase MW, DE Soltis, and RG Olmstead (1993) Phylogenetics of seed plants: an analysis of nucleotide sequences from the plastid gene *rbcL*. Ann Missouri Bot Gard 80:528-580

Clegg MT and ML Durbin (1990) Molecular approaches to the study of plant biosystematics. Paper presented at Symposium of the Australian Systematic Botany Society, CSIRO, Australia

Clegg MT, GH Learn, and EM Golenberg (1990) Molecular evolution of chloroplast DNA. In: RK Selander, AG Clark and TS Whittam [Eds] Evolution at the molecular level. Sinauer Associate, Sunderland, MA, USA, in press

de Queiroz K and J Gauthier (1990) Phylogeny as a central principle in taxonomy: phylogenetic definitions of taxon names. Syst Zool 39:307-322

Endress PK, P Baas, and M Gregory (2000) Systematic plant morphology and anatomy-50 years of progress. Taxon 49:401-434

Fitch WM (1971) Toward defining the course of evolution. Minimal change for a specific tree topology. Syst Zool 20:406-416

Hamby RK and EA Zimmer (1988) Ribosomal RNA sequences for inferring phylogeny within the grass family (Poaceae). Plant Syst Evol 160:29-37

Grant V (2003) Incongruence between cladistic and taxonomic systems. Am J Bot 90:1263-1270

Gravendeel B (2000) Reorganising the orchid genus *Coelogyne*: a phylogenetic classification

based on morphology and molecules. Nationaal Herbarium Nederland

Hennig W (1966) Phylogenetic systematics. University of Illinois press, Urbana

Kluge AG and JS Farris (1969) Quantitative phyletics and the evolution of anurans. Syst Zool 18:1-32

Li W-H and D Graur (1991) Fundamentals of molecular evolution. Sinauer Associates, Sunderland, MA, USA, pp 111

Moritz C and DM Hillis (1996) Molecular systematics: context and controversies. In: DM Hillis, C Moritz, BK Mable [Eds.] Molecular Systematics, 2nd edition, Sinauer Associate, Sunderland, MA, USA, pp 1-13

Systematics Agenda 2000: Charting the Biosphere. Technical Report. Herbarium, New York Botanical Garden, USA

Topik H. (2005) Systematic study of subtribe Aeridinae (Orchidaceae). Ph.D. thesis, The University of Tokyo, Japan

Yukawa T, K Kita, T Handa, Topik H and M Ito (2005) Molecular phylogenetics of *Phalaenopsis* (Orchidaceae) and allied genera: re-evaluation of generic concepts. Acta Phytotax Geobot 56:141-162

Zurawski G and MT Clegg (1987) Evolution of higher-plant DNA-encoded genes: implications for structure, function and phylogenetic studies. Ann Rev Plant Physsiology 38:391-418