Bioinformatics Application in Search of Phylogenetic Relationship on Eleven Species of *Phalaenopsis* Blume (*Orchidaceae*) Based on Gen Internal Transcribed Spacer (ITS1 AND ITS2), 5.8 DNA Ribosomal

Ridesti Rindyastuti

UPT Balai Konservasi Tumbuhan Kebun Raya Purwodadi
Jl. Raya Surabaya-Malang km.65 Purwodadi, Pasuruan
E-mail: ride17@gmail.com

ABSTRACT

Bioinformatics are a reliable method to trace biological information in the molecular level including the analysis of molecular phylogeny relationship of species. This paper discuss about the phylogeny relationship of eleven species of Phalaenopsis (Orchidaceae) based on internal transcribed spacer (ITS1 and ITS2) gene, 5.8 ribosomal DNA by bioinformatics applications. The method of this study is a database search in Gene Bank through the NCBI website (http://www.ncbi.nlm.nih.gov) to search nucleotide data and DDBJ (http//www.ddbj.nig.ac. jp) for the use of software BLAST and the EBI (http://www.ebi.ac.uk) for nucleotide alignment of similarity search. Cladogram shows that P. javanica and P. patherina have a very close relationship. Both species are also still closely related to P. cornu-cervii. Species P. modesta, P. tetraspis and P. lueddemannian are separated each other that can be revealed to have distant phylogenetic relationship than other species in the genus. Species P. amboinensis form a cluster with the species P. floresensis and P. mannii. However, P. sumatrana was closely related to P. corningiana. The cladogram showed that formation among Phalaenopsis species was not apparently grouping. Moreover, there is a large group consisting of P. amboinensis, P. Sumatrana, P. cornu-cervi, P. pantherina, P. floresensis, P. mannii and P. corningiana, while P. modesta, P. tetraspis and P. lueddemanniana were not clustered into the group.

Keywords: Bioinformatics, phylogenetic relationship, Phalaenopsis, databases, internal transcribed spacer (ITS) gene.

Introduction

Internal Transcribed Spacer (ITS) gene in ribosomal DNA is sequence gene from small subunit of rDNA, that essential for many organisms and highly conserved during evolution. For that reason, this gene is applicable to analyze the phylogeny relationship among species. In the Orchidaceae family, ITS gene has been sequenced from its isolates and deposited into bioinformatics database (Yokota *et al*., 2005). ITS genes of several orchid species have been sequenced. Tsai *et al*. (2006) reconstructed phylogeny tree of species in the genus of Phalaenopsis based on ITS sequence from nrDNA (nuclear ribosome DNA). This research result supported the systematic of generic level, however, did not support its sub-generic level based on the generic classification by Christensen.

Phylogeny relationship analysis of a species could be established by bioinformatics application. Bioinformatics are combination of two different disciplines which are information technology and computer science that integrated into molecular biology. Bioinformatics are firstly launched on 1979 by Paulien Hogeweg and rapidly developed into the use for genome information and DNA sequence. There are many advantages of bioinformatics, i. e mapping DNA and protein sequence analysis and also for DNA and protein alignment, and become the adequate database source for information searching of a species in the level of molecular. Bioinformatics were available for reconstructing Tree of life (phylogeny tree that including whole organism kingdom) (Anonim[1], 2010). Bioinformatics could be suggested as link between the genetic and phenotype that used for predicting the protein structure from nucleotide sequence so that it can perform the cellular condition that expressed particular phenotype characters (Bexevanis, 2001; Cohen, 2004).

Here, we discussed the application of bioinformatics for plant conservation by searching the phylogeny relationship of eleven species of *Phalaenopsis* Blume (*Orchidaceae*) based on Internal

Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal. We also relate the searching result with the distribution of orchid species to obtain preliminary evolution process information of orchid species.

METHOD

Internal Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal searching in the Gene Bank

One of the Phalaenopsis species that its nucleotide will be the reference for other species should be choosen first. In this research, we used *P. sumatrana*. The nucleotide of Internal Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal of *P. sumatrana* was searched in the Gene Bank (http://www.ncbi.nlm.nih.gov). The searching was started by typing keyword "Internal Transcribed Spacer in Phalaenopsis" in the homepage of NCBI with the searching type "nucleotide", thus we obtained the nucleotides list of ITS have been sequenced from *P. sumatrana.* One of the ITS gene appeared will be chosen for obtaining the sequence from Internal Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal completed by other information about that gene (Luscombe et al., 2001).

The searching of existing and nearness of Internal Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal in *Phalaenopsis* Blume

After the sequence of nucleotide Internal Transcribed Spacer (ITS1 and ITS2), 5.8 DNA ribosomal gene was obtained, the existing and nearness of ITS1 and ITS2, 5.8 DNA ribosomal in *Phalaenopsis* should be searched. The searching used the BLAST software that could be accessed from DDBJ (http//www.ddbj.nig.ac.jp). Sequence ITS genes of each species was searched in the next step by the searching method similar with the first step (Luscombe et al., 2001).

Similarity searching, alignment and phylogeny tree reconstruction

To search the similarity of each internal transcribed spacer (ITS1 and ITS2), 5.8 DNA ribosomal gene sequence analysis by *software Aligment* was established. Moreover, we used software namely ClustalW for searching the amino acid quantity, aligment and phylogeny tree reconstruction. Both software could be accessed through EBI (http://www. ebi.ac.uk) (Luscombe et al., 2001).

Data analysis

Gene sequence data of eleven orchid species were analyzed by ClustalW software for investigating the genetic variation of orchid and Tree View for searching the phylogeny relationship among species of Phalaenopsis.

Result and discussion

Phylogeny Relationship of eleven Phalaenopsis species

Phalaenopsis has approximately 60 species that widely distributed from Sri Lanka and Southern India to New Guinea and Northern Australia throughout the North and South of China. The distribution centre of genus Phalaenopsis are Philipine and Borneo. The species of Phalaenopsis are generally able to grow in the low land and hot-temper region. However, several species could adapt in the cooler temper and higher land region. The species of Phalaenopsis has beautiful flower and consequently yield the good quality hybrid from its parental. The growers developed the hybrid quality by enhancing the best shapes, sizes, textures and color variations of their flowers.

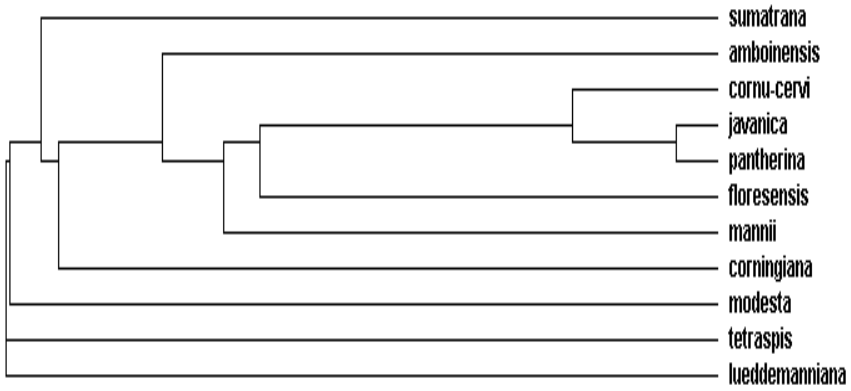
In this research, we conduct the searching of phylogeny realationship of internal transcribed spacer (ITS1 and ITS2), 5.8 DNA ribosomal on 11 Phalaenopsis species origin from Indonesia, they are *P. tetraspis, P. lueddemanniana, P. modesta, P. sumatrana, P. corningiana, P. amboinensis, P. mannii, P. floresensis, P. javanica, P. pantherina*, and *P. cornu-cervi*.

```
tetraspis           ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCCGACATCGGGTCTTCGTGGGGTGGA
lueddemanniana      ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCCGACATCGGGTCTTCGTGGGGTGGA
modesta             ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCCGACATCGGGTCTTCGTGGGGTGGA
sumatrana           ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCCGACATCGGGTCTTCGTGGGGTGGA
corningiana         ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCCGACATCGGGTCTTCGTGGGGTGGA
```

```
amboinensis          ATTGGCGCCAAGGGAACTTGTGAAAAACACGAGCCTGACATCGGGTCTTCGTGGGGTGGA
mannii               ATTGGCCCCATGGGAACTTGTGAAAAACACGAGCCCAATATCGGGTCTTCGTGGGGTGGA
floresensis          ATTGGCGACAAGGGAACTTGTGAAAAACACGAGCCTGACATCGGGTCTTCGGGGGGTGGA
javanica             ATTGGCGCCCAGGGAACTTGTGAAAAACACGAGCCAGACATCGGGTCTTTGTGGGGTGGA
pantherina           ATTGGCGCCCAGGGAACTTGTGAAAAACACGAGCCAGACATCGGGTCTTTGTGGGGTGGA
cornu-cervi          ATTGGCGCCCAGGGAACTTGTGAAAAACACGAGCCAGACATCGGGTCTTTGTGGGGTGGA
                    ******  *  ************************  * **********  * ********
```

Picture 1. Aligment of nucleotide arrangement internal transcribed spacer gene (ITS1 and ITS2), 5.8 DNA ribosomal on 11 *Phalaenopsis* spesies.

Based on the result of nucleotide base arrangement, the number 8, 9, 10, 12, 14, 39, 41, 43, 45, 46, 57, 58, 60, and 61 have distinctness among eleven species of Phalaenopsis. Cladogram that described the phylogeny relationship of eleven species of Phalaenopsis Blume is performed in Picture 2.



Picture 2. Cladogram of 11 Phalaenopsis species based on internal transcribed spacer (ITS1 and ITS2), 5.8 DNA ribosomal.

The cladogram showed that *P. javanica* has close phylogeny relationship with *P. patherina*. The both species were still closely related to *P. cornu-cervi*. Species *P. modesta*, *P. tetraspis* and *P. lueddemanniana* were separately each other and did not have close relationship with other species in genus Phalaenopsis. Species *P. amboinensis* formed the cluster with *P. floresensis* and *P. mannii*. While *P. sumatrana* had close relationship with *P. Corningiana.* Moreover, there was slightly grouping formed among the species targets, and there was one big group that consist of *P. amboinensis P. Sumatrana, P. cornu-cervi, P. pantherina, P. floresensis, P. mannii* and *P. corningiana,* while *P. modesta, P. tetraspis* and *P. lueddemanniana* were out of the group.

According to Tsai *et al.* (2006), *P. javanica, P. lueddemanniana, P. amboinensis* and *P. floresensis* were included into similar section amboinenses. *P. sumatrana* and *P. tetraspis* were clasified into section zebrinae. *P. pantherina* was clasified into section polychilos together with *P. cornu-cervi.* Section in the genus Phalaenopsis described category level of subgenus based on its geographic distribution. Subgenus of Polychilos only had diminish species that distributed in India and commonly found in Indonesia and Philipines. Based on its subgenus, the cladogram showed minor diference from previous research especilaly for three species *P. patherina, P. cornu-cervi,* and *P. javanica.*

Bioinformatics Aplication

Three main basic data support the nucleic acid data sequence for bioinformatics, they are NCBI (United State of America), EMBL (Eropa), and DDBJ (Japan). In the establishment of searching process,

researcher need BLAST and ClustalW software. BLAST (*Basic Local Alignment Search Tool*) is the software that relate to use of biological sequence basic data. BLAST search on the sequence basic data is used for searching the nucleic acid and protein that similar to the sequence target. BLAST searching are usefull for finding the similar gene in several organism or for testing the validity of sequence result or for verifying the function of gene. ClustalW can be used to determine the alignment, reconstruct phylogeny tree and predict the number of amino acids from amino acid sequences were compared (Gerstein, 2009; Lesk, 2002).

Bioinformatics is useful because through this application many biological problems can be solved using DNA sequence and amino acids data. In bioinformatics, sequence alignment to predict the structure of protein and form of RNA secondary structure, phylogeny analysis, and gene expression analysis can be performed. Bioinformatics offers several advantages such as application of bioinformatics for genetic information searching on orchid species is quite easy, because the availability of data from the genome research results were adequately support the information needed. Data sequence of orchid is adequate but is not the same as other plant species, particularly native plants to Indonesia. Genetic data on nucleotide composition of native species in Indonesia such as *Mangifera* spp., *Syzygium* spp., *Diospyros* spp., *Musa* spp. and others are still limited in the database. This showed that there was still limited research on Indonesian plants or tropics plant, particularly in the field of genetic diversity. Therefore, basic research and molecular genetic diversity of organisms is still very necessary.

Acknowledgement

I am grateful Agista Zaziroh, S. Si for being a discussion partner during the research.

References

[1] Baxevanis A. D. 2001. *Bioinformatics*. 2nd ed. John Wiley & Sons, Inc. New York.

[2] Cohen, J. 2004. Bioinformatics-An Introduction for Computer Scientists. ACM Computing Surveys, Vol. 36, No. 2, June 2004, pp. 122–158.

[3] Gerstein, M. 2009. Bioinformatics Introduction. http://www.imb-jena.de/~rake/PDF/gerstein.pdf.

[4] Leks A. M. 2002. *Introduction to Bioinformatics*. Oxford University Press Inc. New York.

[5] Luscombe, N.M., D. Greenbaum, M. Gerstein. 2001. What is bioinformatics? An Introduction and Overview. Yearbook of Medical Informatics. Pp: 83-100.

[6] Tsai, C. C., S. C. Huang and C. H. Chou. 2006. Molecular phylogeny of Phalaenopsis Blume (Orchidaceae) based on the internal transcribed spacer of the nuclear ribosomal DNA. Pl. Syst. Evol. 256: 1–16.

[7] Yokota, Y., T. Kawata, Y. Iida, A. Kato, S. Tanifuji. 2005. Nucleotide sequences of the 5.8S rRNA gene and internal transcribed spacer regions in carrot and broad bean ribosomal DNA. *Journal of Molecular Evolution* (Preview): vol 29, 4.

[8] Zaziroh, A. 2009. Hubungan Kekerabatan Jenis-Jenis Anggota Genus *Rasbora* Berdasarkan Gen *Cytochrome Oxidase Subunit I* (COI) Mitokondria Menggunakan Bioinformatika. Makalah Seminar. Fakultas Biologi UGM. Yogyakarta.