Development of Sequence-Based Microsatellite Marker for Phalaenopsis Orchid

FATIMAH^{1*}, DEWI SUKMA²

¹Molecular Biology Division, Indonesian Center for Agricultural Biotechnology and Genetic Resources Research and Development (ICABIOGRAD), Jalan Tentara Pelajar No.3A Cimanggu, Bogor 16111, Indonesia ²Department of Agronomy, Faculty of Agriculture, Bogor Agricultural University, Darmaga Campus, Bogor 16680, Indonesia

Received November 12, 2010/Accepted June 13, 2011

Phalaenopsis is one of the most interesting genera of orchids due to the members are often used as parents to produce hybrids. The establishment and development of highly reliable and discriminatory methods for identifying species and cultivars has become increasingly more important to plant breeders and members of the nursery industry. The aim of this research was to develop sequence-based microsatellite (eSSR) markers for the *Phalaenopsis* orchid designed from the sequence of GenBank NCBI. Seventeen primers were designed and thirteen primers pairs could amplify the DNA giving the expected PCR product with polymorphism. A total of 51 alleles, with an average of 3 alleles per locus and polymorphism information content (PIC) values at 0.674, were detected at the 16 SSR loci. Therefore, these markers could be used for identification of the *Phalaenopsis* orchid used in this study. Genetic similarity and principle coordinate analysis identified five major groups of *Phalaenopsis* sp. the first group consisted of *P. amabilis*, *P. fuscata*, *P. javanica*, and *P. zebrine*. The second group consisted of *P. amabilis*, *P. violaceae sumatra*, *P. modesta*. The forth group consisted of *P. cornucervi* and *P. lueddemanniana*, and the fifth group was *P. amboinensis*.

Key words: microsatellite marker, Phalaenopsis, polymorphism, genetic relationship

INTRODUCTION

Phalaenopsis is one of the most interesting genera of orchids, due to their members are often used as parents to produce hybrids. Most species of this genus occur in the tropics, whereas about 45 species distributes in India, China, South-East Asia, Indonesia, Philiphines, New Guinea, and Australia. Twenty one species of Phalaenopsis distributes in Indonesia, including Sumatera, Java, Kalimantan, Nusa Tenggara, Sulawesi, Moluccas, and Papua. The Bogor Botanical Garden has collected 15 species which were collected from the Indonesian archipelago. Many species of Phalaenopsis have become rare or threatened because of over exploitation, extensive destruction of habitat or other environmental disturbances. Bogor Botanic Garden has taken role in the ex situ conservation of Phalaenopsis species (Puspaningtyas & Sofi 1999).

The germplasm management of these orchids is particular challenging because the relatively long phase in the juvenile phase. The establishment and development of highly reliable and discriminatory methods for identifying of *Phalaenopsis* species and cultivars has become increasingly more important to plant breeders and members of the nursery industry who need sensitive and accurate tools to differentiate and identify cultivars for the purpose of plant patent protection (Phuekvilai 2009).

*Corresponding author. Phone: +62-251-8337975, Fax: +62-251-8338820, E-mail: fathyapril@yahoo.com In general, the breeding programs are designed to improve the size and color of the flowers as well as other characteristics such as, longevity, stalk length, leaf shape, ease of cultivation, disease resistance and the number of viable seeds through the selection of parents for hybridization. Through tremendous efforts in breeding, various types of *Phalaenopsis* varieties with attractive color and graceful appearance have been developed and the success of the development has made Taiwan one of the most important producers of *Phalaenopsis* in the world (Tang 2007).

Microsatellite or simple sequence repeat (SSR) are short (2-6 bp) tandemly repeated DNA sequences and have been demonstrated to be a powerful tool in the studies of characterization, conservation and improvement of important agricultural crops. The principal reason for the increase success of SSRs as a molecular tool is that it provides a higher incidence of detectable polymorphism than other molecular techniques such as RFLPs and RAPDs. The great value of these markers arises because they are hypervariable, codominant, abundant, reproducible and easy to detect by PCR method. Microsatellite (SSR) fingerprinting has been successfully used for the construction of genetic linkage maps and for the identification and tagging of economically important genes in a large number of plant species (Mc Couch SR 2002).

Recent study in Indonesia for genetic relationship among 13 sub tribe Sarcanthinae orchids based on 38 phenotypic data and 14 RAPD primers resulted in consistent relationships among orchid based those two markers (Kartikaningrum 2002). RAPD marker cannot identify any specific band for any particular character or genotype among 19 genotypes of Phalaenopsis orchid (Dwiatmini et al. 2003). That study revealed a small correlation coefficient hence indicated that the relationship between variables is weak, meaning that average taxonomic distance could not be used to estimate the genetic similarity. Other study in phylogenetic analyses using two DNA sequence data sets derived from the maturase-coding (matK) gene located in an intron of the plastid gene trnK, and the internal transcribed spacer region of 18S-26S nuclear ribosomal DNA to examine relationships in subtribe Aeridinae (Orchidaceae) resulted inconsistent with previous classifications of the subtribe caused by homoplasy of these characters (Topik 2005, 2008).

Considering the previous studies, the objectives of this study were to (i) develop microsatellite markers for *Phalaenopsis* orchids, (ii) evaluate a substantial number of these markers for their ability to detect polymorphism of several diverse genotypes, (iii) evaluate the results with respect to phylogenetic relationships and classifications.

The markers developed in this study will not only provide valuable tools for molecular breeding of *Phalaenopsis* orchids, but may also have applications in related taxa. This PCR-fingerprinting technique based on orchid microsatellites would be very useful for the documentation of genetic resources of *Phalaenopsis* cultivars and the protection of the orchid breeder's right.

MATERIALS AND METHODS

Plant Material. A total of 17 species of *Phalaenopsis* orchid representing major in Indonesia (Table 1) were analyzed with selected SSR markers, which obtained from the genetic resources units of The Bogor Botanical Garden and other nursery. Total genomic DNA was isolated from young leaves collected, following the procedure described by Doyle (1991), using the cetyl trimethyl ammonium bromide (CTAB) method with minor modifications.

Microsatellite Development. All microsatellite related to the *Phalaenopsis* sp. listed in the National Center for Biotechnology Information (NCBI, http://www.ncbi. nlm.nih.gov) were downloaded in FASTA format and analyzed. Designed 17 primers from unique SSRs for repeat patterns using defining an SSR with all repeats (dimmers and trimers) using BatchPrimer3 program (http:// probes.pw.usda.gov/cgi-bin/batchprimer3/ batchprimer3.cgi.), with following parameters defined: product size, min 100 bp, max 200 bp; primer length min 18 bp, opt 20 bp, max 24 bp; GC content min 40%, opt 50%, max 55%; melting temperature min 50 °C, opt 55 °C, max 60 °C and combined with three primer pairs from research paper (Rodrigues 2009).

SSR-PCR Amplification and Detection. The PCR reaction mixtures contained 10 mM Tris-HCl (pH 8.3), 50 mM KCl, 50 ng genomic DNA, 10 pmol of each primer, 1.5 mM MgCl2, 0.2 mM of each dNTP, and 0.2 units of Taq DNA polymerase (Kapa Biosystems). Amplification was performed in 20 µl volumes using MJ Research Thermal cycler with an initial denaturation at 94 °C for 4 min followed by 34 cycles of 1 min at 94 °C, 1 min at the specific annealing temperature, 1 min at 72 °C and a final extension step of 10 min at 72 °C. Each pair of SSR primers was initially screened for amplification of a specific product from genomic DNA of two species one from P. cornucervi and the other from P. lueddemanniana. The PCR products were separated on 2% agarose gels followed by ethidium bromide staining and visualized by UV light. SSR loci that gave amplification products in *Phalaenopsis* sp. were used in a second round of PCR on 17 diverse Indonesia orchid cultivars. The PCR products were separated on 8% denaturing polyacrylamide gels with Dual Triple-Wide mini-Vertical Kit (CBS Scientific) and stained with ethidium bromide and visualised by UV light. The electrophoretic patterns of the PCR products were recorded digitally using a Gel-Doc image analysis system.

SSR Fragment Analysis. The summary statistics including the number of alleles per locus, major allele frequency, gene diversity (He), polymorphism information content (PIC) values were determined using POWERMARKER version 3.25 (http://www.powermarker.

Table 1. Phalaenopsis sp. orchid based	on Christenson (2001) classification	and clustering in this study
--	--------------------------------------	------------------------------

Name	Collection	Classification	Cluster
P. amabilis	Bogor Botanic Garden	Phalaenopsis	Ι
P. fuscata	Bogor Botanic Garden	Fuscatae	Ι
P. javanica	Bogor Botanic Garden	Amboinensis	Ι
P. zebrina	Bogor Botanic Garden	Zebrinae	Ι
P. amabilis	Permata Anggrek	Phalaenopsis	II
P. amboinensis	Indoflowers	Amboinensis	II
P. mannii	Indoflowers	Polychilos	II
P. floresensis	TAIP Jakarta	Amboinensis	II
P. bellina	TAIP Jakarta	Amboinensis	II
P. cornucervi	Permata Anggrek	Polychilos	III
P. violaceae sumatra	Sandra Collection	Amboinensis	III
P. modesta	Sandra Collection	Amboinensis	III
P. bellina	Sandra Collection	Amboinensis	III
P. cornucervi	Sandra Collection	Polychilos	III
P. lueddemaniana	Indoflowers	Amboinensis	IV
P. cornucervi	Bogor Botanic Garden	Polychilos	IV
P. amboinensis	Bogor Botanic Garden	Amboinensis	V

net) based on evaluation from the 17 *Phalaenopsis* spesies. Based on these allele sizes, statistical analysis software, program NTSYSpc 2.11p (Exeter Software, Setauket, USA) was used for SSR analysis. Providing the binary data, presence or absence of alleles at each locus, was coded in binary form 1 or 0, respectively. The dice coefficient (SIMQUAL) and unweighted pair group method aritmatic (UPGMA) method used to cluster the varieties and visualize their genetic relatedness to each other. Principal coordinate analysis (PCoA) was conducted to create a scatter plot showing genetic relationship among accession. Genetic distance was calculated using *Dc* (Cavalli-Sforza & Edwards 1967).

RESULTS

SSR Polymorphism in the Entire Sample. Molecular analysis of the 17 selected species of *Phalaenopsis* and 16 SSR primers generated strong and reproducible amplification products, all of which displayed polymorphism among the species. As highly polymorphic primers were used for analysis, a relatively large number of polymorphic SSR markers were detected by these primers. Examples of amplification patterns for molecular characterization obtained by SSR in selected species of *Phalaenopsis* are shown in Figure 1.

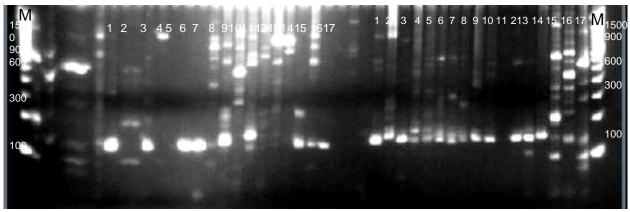


Figure 1. SSR profile in acrylamide gel from 17 species of *Phalaenopsis* using primer pair IPS1 and IPS3 respectively. M: molecular size ladder x 100 bp. Numbers correspond to the genotypic number (please refer to Table 1).

Primer name	Repeat motif	Size range alleles	Sequence	No.of alleles	Major allele freq.(%)	Но	He	PIC	GenBank accession number
IPS1	(AG)7	129-140	GAGAGAGAAAGAGAGAGGTACG	3	69	0.0000	0.4615	0.4042	AJ575763
	. ,		CTACGGCTGATTTGATTTCTA						(P. violaceae)
IPS2	(AG)7	129-139	GCTAGAGATAGAGAGAGAAAGAG	2	76	0.0000	0.3599	0.2951	AJ575759
		CTACGGCTGATTTGATTTCTA						(P. lueddemanniana)	
IPS4 (AG)6	100-120	GAGAGAGAAAGAGAGAGAAACAT	3	73	0.0000	0.4298	0.3855	AJ575750	
		CTACGGCTGATTTGATTTCTA						(P. amboinensis)	
IPS5	(AG)6	110-140	GAGAGAGAAAGAGAGAGAAACAT	2	54	0.0000	0.4970	0.3735	AJ566347
			CTACGCCTGATTTGATTTCTA						(P. amboinensis)
IPS6	(AG)7	150-160	GCTAGAGATAGAGAGAGAAAGAG	5	72	0.2500	0.4531	0.4215	AJ575759
			CTACGGCTGATTTGATTTCTA						(P. lueddemanniana)
IPS7	(GA)7	137-160	GTTTGTTGTCAGAGAAGCATT	4	36	0.0000	0.7245	0.6740	AJ888295
			AATTGAGACCTTCACTTCCTC						(P. amabilis)
IPS9	(AG)7	120-170	GAGAGAGAAAGAGAGAGAAACAT	4	58	0.0000	0.5972	0.5524	AJ575752
			CTACGGCTGATTTGATTTCTA						(P. cornucervi)
IPS10	(AG)7	120-170	AGAGAGAGAAAGAGAGAGATGC	5	50	0.0000	0.6667	0.6221	AJ566356
			CTACGCCTGATTTGATTTCTA						(P. venosa)
IPS13	(AG)6	137-160	GCTAGAGATAGAGAGAGAAAGAG	3	60	0.0000	0.5511	0.4846	AJ566353
			CTACGGCTGATTTGATTTCTA						(P. lueddemanniana)
IPS51	(GAA)4	160-200	TTGGCAGCGTTAAAGATATAG	2	50	0.0000	0.5000	0.3750	AY378151
			CATCCTCTGCATAAAAATCAC						(P. equistris)
IPS52	(TTG)6	110-180	GCAATGGAGAAAAAGGATTTA	3	54	0.3077	0.5562	0.4652	AY378151
			GCTCCACTCACCTGTTAGTTA						(P. equistris)
PGA06	(CA)10	125-130	GTGAAAGACACACACACACACA	4	58	0.3846	0.5828	0.5249	EF462862 (P. gigantea)
			GGTTGTACGCCTTTGTCGAT						(Rodriguez, 2009)
PGVB11	1 (AT)4	127-130	TGCATTCTTGCTTCGTTTTG	2	73	0.0000	0.3911	0.3146	EF462843 (P. gigantea)
			CGAGAAGCATTAAGCTGCATT						(Rodriguez, 2009)
PGCT13	3 (CT)15	135-160	GGAAGGGCTCTCTCTCTCTCT	3	90	0.0667	0.1844	0.1754	EF545013 (P. gigantea)
			TTGCCCTCCTCTCTTTCTTT						(Rodriguez, 2009)
IPS14	(GTT)4	100-180	TAGTTTGTTTCGTGGTTCAGT	3	61	0.3333	0.5494	0.4890	AY954515
			TGGTGTGAATTATCTCGTTCT						(P. hybrid cultivar)
IPS15	(AT)6	147-230	CGTTTCACTCAAAGTTTTGTT	3	50	0.0000	0.6111	0.5355	EU023907
			TCCTGACGTAATATTTGTTGG						(P. bellina)

Table 2. Data summary for 16 microsatellite markers across 17 Phalaenopsis sp. Orchids

A total of 51 alleles were detected at sixteen microsatellite markers across 17 *Phalaenopsis* sp. The number of alleles per locus average 3 and ranged from 2 (IPS3, IPS5, IPS51, PGVB11) to as many as 5 (IPS6 and 1PS10 respectively). Polymorphism information content (PIC) values, a reflection of allele diversity and frequency among the species, were high for most of the microsatellites, the average of 0.50 and range from a low of 0.17 (PGCT13) to a high of 0.67 (IPS7). PGCT13 was present the highest frequency allele (90%) of the 17 *Phalaenopsis* sp. used in this study. IPS7 showed the lowest frequency alleles to 36% (Table 2). Matrices of

coephenetic values correlated between the dendrograms was obtained at r = 0.72, generated from SSR data using the Mantel test.

Clustering of *Phalaenopsis* **sp. with Microsatellite Markers.** Genetic similarity values between the various *Phalaenopsis* **sp.** were used to produce an UPGMA dendogram (unrooted cluster tree analysis) which clearly explained the relationship among 17 orchid accessions. UPGMA clustering distinguished five major groups (Figure 2). The first group consisted of *P. amabilis, P. fuscata, P. javanica,* and *P. zebrine.* The second group consisted of *P. amabilis, P. amboinensis, P. bellina, P.*

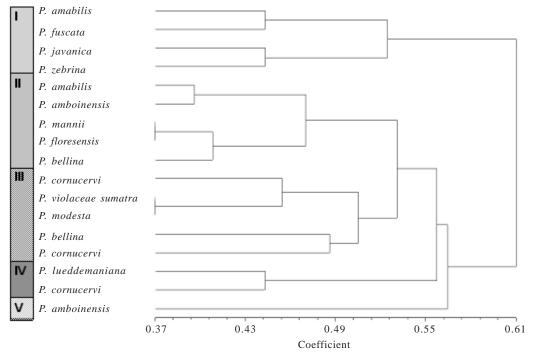


Figure 2. UPGMA tree showing the genetic relationships between the 17 *Phalaenopsis* sp. based on 16 microsatellite markers distinct five groups of *Phalaenopsis*.

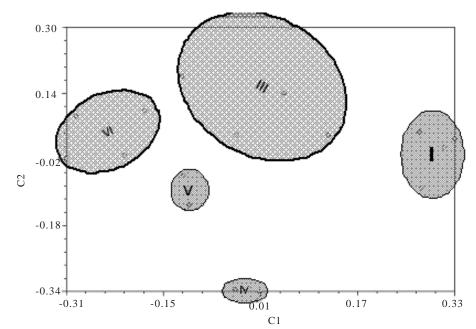


Figure 3. Genetic relationship depicted among 17 Phalaenopsis with five groups outlined derived from Principal coordinate analysis.

floresens, and P. mannii. The third group consisted of P. bellina, P. cornucervi, P. cornucervi, P. violaceae sumatra, P. modesta. The forth group consisted of P. cornucervi and P. lueddemanniana, and the fifth group was P. amboinensis. Principle - coordinate and cluster analysis (PCoA) (Figure 3) separation was similar to those identified by the UPGMA tree cluster analysis (Figure 2).

DISCUSSION

This study evaluated marker polymorphism, emphasizing the differences in genetic variability of microsatellite sequences with different SSR motifs originating from GenBank accessions. However GenBankderived microsatellites had lower levels of polymorphism compare to SSR marker from research paper (Rodrigues 2009). In study by Chin *et al.* (1996) on maize and Becker and Heun (1995) on barley, relatively low levels of polymorphism were detected for markers developed from the GenBank sequence information.

Temnykh *et al.* (2000) reported in rice, GenBank-derived microsatellites had low variability values (number of repeat units, number of alleles, allele size range, and expected genetic diversity) than microsatellites isolated from genomic libraries. This difference reflected the constraints on DNA sequence variation in transcriptionally active portions of the genome. From this point of view, screening of genomic libraries was more efficient than the GenBank search, since more informative markers capable of detecting more genetic differences were developed based on genomic clones. On the other hand, the GenBank-derived markers produced a high proportion of intrasubspecifically conserved microsatellite markers with distinct allele patterns these can be useful for evolutionary studies and applications in breeding.

The results from the genetic distance-based analysis of 17 *Phalaenopsis* sp. with 16 microsatellite markers reveal five major groups. The results of this study are consistent with an earlier study of species of *Phalaenopsis* for RAPD marker (Goh *et al.* 2005). Based on the dendogram, seven main clusters were obtained. However, from the seven clusters only cluster 3 contains subgenus *Polychilos*, sections *Amboinenses* and *Zebrinae*. In our study using microsatelite marker have distinguished this cluster 3 into four sub clusters (group II, III, IV, and V). Niknejad *et al.* (2009) reported that the third group of generated RAPD data consisted of *P. mannii, P. Modesta*, and *P. cornucervi*.

In our study, *P. amabilis* and *P. cornucervi* from Bogor Botanic Garden is not clustered together with from Permata Anggrek. The Bogor Botanic Garden collection of orchids focuses primarily on wild species and accessions native to Indonesia that arriving after collecting expedition. It should be noted that Permata Anggrek is a commercial orchid nursery. Therefore, the authenticity of the plant materials that they worked on is doubtful since *Phalaenopsis* is a genus that is commonly utilized for breeding purposes.

These results are quite different from the previously analyzed with RAPD data (Fu et al. 1997; Been et al. 2002)

since the microsatellite marker used in this study provides a higher incidence of detectable polymorphism than RAPDs. In other hand, most of the *Phalaenopsis* used in this study was the only representative taxon, the interpretation of this placement due to different sampling strategies. More sampling of the plants from the various localities should be undertaken to resolve this apparent inconsistency with previous reports.

The results clearly indicated that these microsatellite markers applied in this economically important group of orchids for the study of molecular characterization and relationships have high potential in cultivar identification, the evaluation of cultivar purity in commercial samples and other applications.

ACKNOWLEDGEMENT

We appreciate Yupi for the *Phalaenopsis* Orchid from The Botanic Garden, Bogor and Masumah for technical assistance. This work was financially supported by Hibah Strategis Nasional from Directorate Generale of Higher Education (DGHE), Indonesian Ministry of Education, No. 550/SP2H/PP/DP2M/PP/2010 for Dewi Sukma, Institut Pertanian Bogor.

REFERENCES

- Becker J, Heun M. 1995. Barley microsatellites: allele variation and mapping. *Plant Mol Biol* 27:835-845. http://dx.doi.org/ 10.1007/BF00020238
- Been CO, Kim JB, Kim HY. 2002. Random amplified polymorphic DNA (RAPD) for genetic analysis of *Phalaenopsis* species. J Kor Hort Sci 43:387-391.
- Cavalli-Sforza LL, Edwards AWF. 1967. Phylogenetic analysis: model and estimation procedures. *Am J Hum Gen* 19:233-257.
- Chin ECL, Senior ML, Shu H, Smith JSC. 1996. Maize simple repetitive DNA sequences: abundance and allele variation. *Genome* 39:866-873. http://dx.doi.org/10.1139/g96-109
- Christenson EA. 2001. Phalaenopsis a monograph. Oregon, Timber Pr.
- Doyle J. 1991. DNA protocols for plants-CTAB total DNA isolation. In: Hewitt GM, Johnston A (ed). Molecular Techniques in Taxonomy. Berlin: Springer. p 283-293.
- Dwiatmini K, Nurhayati AM, Hajrial A, Nurita LTM. 2003. Analisis pengelompokan dan hubungan kekerabatan spesies anggrek *Phalaenopsis* berdasarkan kunci determinasi fenotipik dan marka molekuler RAPD. J Hort 13:16-27.
- Fu YM, Chen WH, Tsai WT, Lin YS, Chyou MS, Chen YH. 1997. Phylogenetic studies of taxonomy and evolution among wild species of *Phalaenopsis* by random amplified DNA markers. *Report Taiwan Sugar Res Inst* 157:27-42.
- Goh MWK, Kumar PP, Lim SH, Tan HTW. 2005. Random amplified polymorphic DNA analysis of the moth orchid, *Phalaenopsis* (Epidendroideae: Orchidaceae). *Euphytica* 141: 11-22. http://dx.doi.org/10.1007/s10681-005-4814-y
- Kartikaningrum S, Nani H, Achmad B, Murdaningsih HK, Nurita TM. 2002. Kekerabatan antar genus anggrek sub tribe Sarcanthinae berdasarkan data fenotip dan pola pita DNA. *Zuriat* 13:1-10.
- McCouch SR. 2002. Development and mapping of 2240 new SSR markers for rice (*Oryza sativa* L.). *DNA Research* 9:199-207. http://dx.doi.org/10.1093/dnares/9.6.199
- Niknejad A, Kadir MA, Kadzimin SB, Abdullah NAP, Sorkheh K. 2009. Molecular characterization and phylogenetic relationships among and within species of *Phalaenopsis* (Epidendroideae: Orchidaceae) based on RAPD analysis. *African J Biotechnol* 8:5225-5240.

76 FATIMAH AND SUKMA

- Phuekvilai P, Pradit P, Surin P. 2009. Development of microsatellite markers for *Vanda* orchid kasetsart. *J Nat Sci* 43:497-506.
- Puspitaningtyas DW, Sofi M. 1999. Koleksi Anggrek Kebun Raya Bogor. Bogor: LIPI. p 72.
- Rodrigues KF, Kumar SV. 2009. Isolation and characterization of microsatellite loci in *Phalaenopsis gigantean*. Conserv Genet 10:559-562. http://dx.doi.org/10.1007/s10592-008-9569-2
- Tang CY, Wen HC. 2007. Breeding and development of new varieties in *Phalaenopsis*. In: Wen-Huei Chen, Hong-Hwa Chen (eds). Orchid Biotechnology. World Scientific Publishing Co. Pte. Ltd. p 1-20. http://dx.doi.org/10.1142/97898127 75900_0001
- Temnykh S, Park WD, Ayres N, Cartinhour S, Hauck N, Lipovich L, Cho YG, Ishii T, McCouch SR. 2000. Mapping and genome organization of microsatellite sequences in rice (*Oryza sativa* L.). *Theor Appl Genet* 100:697-712. http://dx.doi.org/10.1007/s001220051342
- Topik H, Yukawa T, Ito M. 2005. Molecular phylogenetics of subtribe Aeridinae (Orchidaceae): insights from plastid matK and nuclear ribosomal ITS sequences. J Plant Res 118:271-284. http://dx.doi.org/10.1007/s10265-005-0217-3
- Topik H, Adi P. 2008. Kajian filogenetika molekuler dan peranannya dalam menyediakan informasi dasar untuk meningkatkan kualitas sumber genetik anggrek. J Ag Bio 4:35-40.