SAGO PALM 21: 14－19 (2013)

Molecular analyses of folk varieties of the sago palm (*Metroxylon sagu* Rottb.) using the internal transcribed spacer (ITS) region and nuclear microsatellite DNA

Yoshimasa Kumekawa[1], Agustinus Murjoko[2], Hiroshi Hayakawa[1, 3], Kyohei Ohga[1], Makito Mori[1], Akira Miyazaki[1], Katsura Ito[1], Ryo Arakawa[1], Tatsuya Fukuda[1], Hubertus Matanubun[2] and Yoshinori Yamamoto[1]

[1] Faculty of Agriculture, Kochi University, Monobe B200, Nankoku, Kochi 783-8502, Japan

[2] The State University of Papua

[3] National Institude for Agro-Environmental Sciences, 3-1-3 Kannondai, Tsukuba, 305-8604, Japan

Key words: internal transcribed spacer (ITS) region, microsatellite of nuclear DNA (nrDNA), sago palm

Introduction

It is difficult to classify taxa by using characters with little taxonomic information such as a part of the plant body and root, seedling, sterile branch, and so on. However, a diverse array of molecular approaches is now available that uses molecular markers to resolve some taxonomic problems in various plant taxa. In fact, some studies had been conducted that used molecular markers to identify indistinguishable crops (Fukuda et al., 2005b; Hayakawa et al., 2010; Hayakawa et al., 2011).

The sago palm, *Metroxylon sagu* Rottb. (Arecaceae), is an economically increasingly important crop in Southeast Asia, and can accumulate up to several hundred kilograms of starch in the trunk, which is enzymatically released and a large terminal inflorescence and thousands of fruits are produced (Yamamoto 2005). However, the morphological characteristics of the sago palm are quite simple. Most key characteristics to identify folk varieties of the sago palm exist in the vegetative parts above the ground, e.g. the spine, so that it is impossible to identify each folk variety by its leaves. Therefore, molecular markers are required to identify each folk variety of the sago palm. Some studies have been conducted using molecular markers to clarify important agronomic traits of the sago palm. For example, Ehara et al. (2003) investigated the relationship between geographical and genetic distance of the sago palm by using Random Amplified Polymorphic DNA (RAPD) analysis. Their results indicated that geographical distance is reflected in their genetic distance in the Malay Archipelago. Moreover, Barahima et al. (2005) reported a genetic relationship of sago palms in Indonesia on the basis of RAPD analysis. In general, RAPD analysis is an appropriate tool for resolving phylogenetic relationships in interspecific comparisons. RAPD polymorphisms result from nucleotide sequence mutations that prevent the amplification of a particular marker in some individuals, and this is the generally unstated assumption used when bands are scored as independent and binary characters in a data matrix (Rieseberg, 1996). However, there are some basic limitations of RAPD markers, i.e. problems of homology assessment and deviations from expectations of strict Mendelian inheritance, e.g. artifactual variations, epigenetic interactions, and so on (Wolfe and Liston, 1998). It is questionable whether trees obtained by RAPD analyses provided an appropriate resolution because many cases of RAPD analysis had low values of the consistency index (CI) due to parallelism of many characters.

In contrast to these studies, Kjær et al. (2004) investigated the genetic and morphological variations of the sago palm in Papua New Guinea by using amplified fragment length polymorphism (AFLP) analysis based on previously published reports of Jones et al. (1997). Their results indicated a significant correlation between genetic and geographical distances. Moreover, a similar study was conducted using sago palm cultivars (Celiz et al., 2004). In addition, Barahima et al. (2005) indicated that the analysis of simple sequence repeats (SSR) of chloroplast DNA (cpDNA) was available for the levels of populations and individuals of sago palm in Papua, suggesting that a molecular marker of sago palm was needed for sequences with high nucleotide substitution rate. Recently, Fukuda et al. (2012) revealed that internal transcribed spacer (ITS) sequences of nuclear DNA (nrDNA) were useful to analyse closely related taxa on the basis of phylogenetic relationships by comparison of nucleotide mutation rates of cpDNA, ITS sequences, and RAPD analyses. However, this study did not use sago palm but edible asparagus. Moreover, some microsatellite loci of nrDNA were isolated from species of Arecaceae (Choo et al. 2010). These nucleotide sequences allow us to identify molecular markers of the sago palm. Therefore, the aim of our study is to detect nucleotide mutations to distinguish sago palm cultivars by using preliminary application of highly substituted ITS sequences and previously published microsatellite loci of cpDNA (Nishizawa and Watano 2000) and nrDNA (Choo et al., 2010).

Materials and Methods

Plant materials

The plant materials contained 2 locations of sago palm in the Malay Archipelago. Fig. 1 shows the locations in the collection area used in this study. We used 11 folk varieties, i.e. Demai, Folo, Hobolo, Manno-Besar, Manno-Kecil, Osukulu, Panne, Para, Rondo, Ruruna, and Yepha. The sampling sites are shown in Table 1. We collected the young Demai of one year or less in a few places within the 1 km distance.

DNA analysis

Total DNA was isolated from approximately 200–300 mg of an air-dried leaf with a Plant

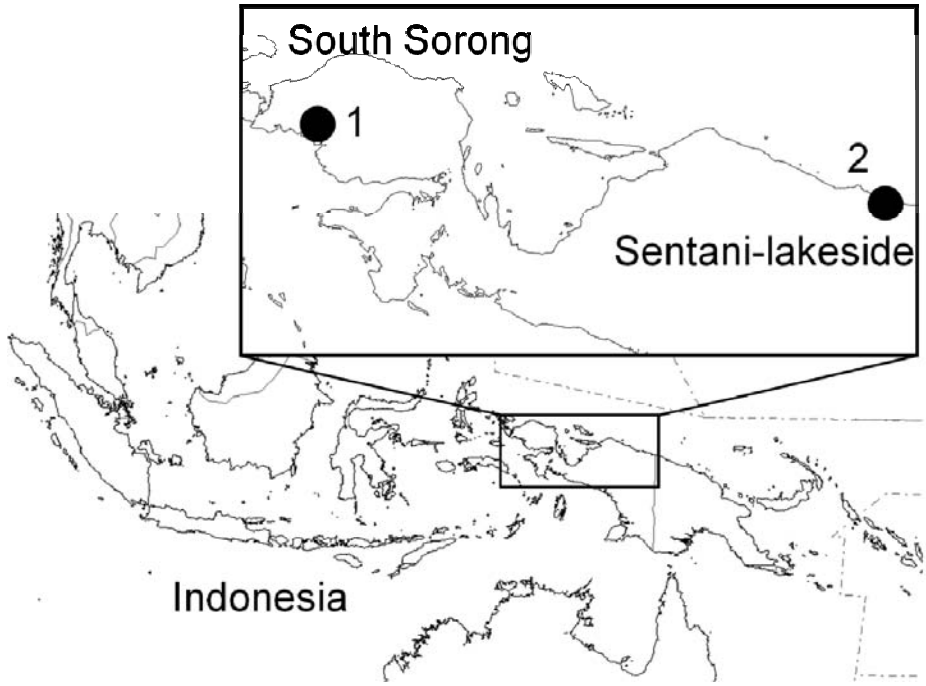

Fig. 1. Sampling areas of this study

Table 1. Plant materials used.

No.	Folk variety	Locality no.	Location for collection
1	Demai	1	West Papua Province, South Sorong
2	Demai	1	West Papua Province, South Sorong
3	Demai	1	West Papua Province, South Sorong
4	Demai	1	West Papua Province, South Sorong
5	Demai	1	West Papua Province, South Sorong
6	Demai	1	West Papua Province, South Sorong
7	Demai	1	West Papua Province, South Sorong
8	Demai	1	West Papua Province, South Sorong
9	Folo	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
10	Hobolo	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
11	Manno-Besar	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
12	Manno-Besar	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
13	Manno-Kecil	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
14	Osukulu	2	Papua Province, Jayapura, Sentani-lakeside, Yahim
15	Panne	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
16	Panne	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
17	Para	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
18	Para	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
19	Rondo	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
20	Rondo	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
21	Ruruna	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
22	Yepha	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar
23	Yepha	2	Papua Province, Jayapura, Sentani-lakeside, Ifar besar

Locality no. corresponds to that givin in Fig. 1.

Genomic DNA Mini Kit (VIOGENE, Sunnyvale, USA), according to the manufacturer's protocol. Isolated DNA was resuspended in Tris-EDTA (TE) buffer and stored at -20 °C until use.

DNA amplification by polymerase chain reaction (PCR) was carried out in a 50-μL reaction volume containing approximately 50 ng total DNA, 10 mM Tris-HCl buffer (pH 8.3) with 50 mM KCl and 1.5 mM $MgCl_2$, 0.2 mM of each dNTP, 1.25 U *Taq* DNA polymerase (TOYOBO), and 0.5 μM of each primer. We used the following thermal cycle profile for amplification: 40 cycles of 1.5 min at 94 °C, 2 min at 48 °C, and 3 min at 72 °C, followed by a final extension step of 15 min at 72 °C. We amplified 13 microsatellite loci of cpDNA with primers designed by Nishizawa and Watano (2000) and Barahima et al. (2010), the ITS region reported by White et al. (1990), and a microsatellite loci of nrDNA reported by Choo et al. (2010). After amplification, the reaction products were run on a 2% agarose gel and digitally photographed.

We also confirmed the sequences of these regions. After amplification, the reaction mixtures were subjected to electrophoresis in 1–2% low-melting-temperature agarose gels for the purification of the amplified products. We sequenced the purified reaction products by using a BigDye-terminator Cycle Sequencing Kit (Applied BioSystems) and a Model 3730A automated sequencer (Applied BioSystems) according to the manufacturer's instructions. For sequencing, we used the same primers as those used for amplification. Sequences for each region were pre-aligned with the CLUSTAL X program (Thompson et al., 1997), and ambiguously aligned regions were manually corrected to minimise the number of indels. Alignment for all DNA regions required the inclusion of several indels.

Results and Discussion

Microsatellite loci are segments of DNA repeats of short repeat length, and are believed to be subject to mutational changes following a slippage model of duplication or deletion of repeat units (Jorde et al., 1995). Therefore, microsatellite loci are usually highly polymorphic and many authors have estimated the average diversity for microsatellite DNA loci (Bowcock et al., 1994; Deka et al., 1995). As for the microsatellite loci of cpDNA (Nishizawa and Watano 2000), sequences of *petB* intron and the intergenic region between *petD* and *rpoA* had higher substitution regions to reconstruct a phylogenetic tree using closely related taxa than those of the Restriction Fragment Length Polymorphism (RFLP) analysis of cpDNA (Fukuda et al. 2005a). However, our results indicated that approximately 2900 bp of 11 microsatellite loci of cpDNA (i.e. *petB* intron [300 bp] and the intergenic region between *petD* and *rpoA* [340 bp]) of *rps16* intron1 [350 bp], *rps16* intron2 [287 bp], *trnG* (UCC) intron1 [264 bp], *trnG* (UCC) intron2 [200 bp], *atpF* [270 bp], the intergenic region between *psbA* and *trnS* (UGA) [167 bp], the intergenic region between *trnG* (GCC) and *trnM* (CAU) [230 bp], the intergenic region between *trnW* (CCA) and *trnP* (UGG) [250 bp], and *rpl16* intron [240bp]) had no length mutations (indels) and that the sequences of all analysed samples were identical (Acc Nos. AB793576 - AB793586).

Our analyses indicated that a substitution site (9-bp) exists in the ITS region of the sago palm (Acc Nos. AB79290 and AB793291) (Table 2). According to this mutation, the folk varieties could be divided into 2 groups; 1 group comprises 8 individuals of the Demai variety and the other group comprises 16 individuals of the other varieties, i.e. Folo, Hobolo, Manno-Besar, Manno-Kecil, Osukulu, Panne, Para, Rondo, Ruruna, and Yepha. This result indicates that the Demai variety was different from the remaining folk varieties of the sago palm and could be distinguished from them by the ITS region. Moreover, there are 2 sites of length mutation in the microsatellite loci of nrDNA, i.e. 2-bp deletion and a 4-bp deletion (Acc Nos. AB793287 - AB793289) (Table 2). The 2-bp deletion in this region was found in the varieties Folo, Manno-Besar, Manno-Kecil,

Table 2. Differences found in the sequences obtained from *Metroxylon sagu* samples

No.	Folk variety	Locus site (bp)																															
		ITS																microsatellite																	
		3	3	3	3	3	3	3	3	3	3	3	3	3	3	3																		
		1	1	1	1	1	2	2	2	2	2	2	2	2	2	2		1	1	2	2	2	2	2	2	3	3	3	3	3	4	4	4	4
		5	6	7	8	9	0	1	2	3	4	5	6	7	8	9		8	9	0	1	2	3	4	5	6	5	6	7	8	9	0	1	2	3
1	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
2	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
3	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
4	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	N	A	N	A	N	A	T	A	C	A	T	A	T	A
5	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	-	-	-	-	T	A
6	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
7	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	N	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
8	Demai	G	G	G	T	G	G	C	G	G	G	G	G	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
9	Folo	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	T	A	C	A	T	A	-	-	-	-	T	A
10	Hobolo	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
11	Manno-Besar	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
12	Manno-Besar	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	T	A	C	A	T	A	-	-	-	-	T	A
13	Manno-Kecil	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	T	A	C	A	T	A	-	-	-	-	T	A
14	Osukulu	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	N	A	N	A	N	A	-	-	-	-	T	A
15	Panne	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	-	-	-	-	T	A
16	Panne	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	-	-	-	-	T	A
17	Para	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	C	A	T	A	T	A
18	Para	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	N	A	C	A	T	A	C	A	T	A	-	-	-	-	T	A
19	Rondo	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	T	A	C	A	T	A	-	-	-	-	T	A
20	Rondo	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	N	A	C	A	T	A	C	A	T	A	T	A
21	Ruruna	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	-	-	T	A	C	A	T	A	-	-	-	-	T	A
22	Yepha	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	N	A	N	A	N	A	N	A	-	-	-	-	T	A
23	Yepha	G	G	G	-	-	-	-	-	-	-	-	-	G	C	G		C	A	C	A	C	A	T	A	C	A	T	A	-	-	-	-	T	A

Osukulu, Rondo, and Ruruna. Among them, Manno-Besar and Rondo were included in the group of individuals without deletion at this site. Although the 4-bp deletion in the same microsatellite DNA locus was detected in the varieties Demai, Folo, Manno-Besar, Manno-Kecil, Osukulu, Panne, Para, Rondo, Ruruna, and Yepha, non-deletion individuals were also identified in the varieties Demai, Manno-Basar, Para, and Rondo, respectively. Therefore, our results indicated that the ITS region could be used as molecular marker to distinguish the Demai variety from the other folk varieties. However, we could not gain molecular information to identify folk varieties of the sago palm by using microsatellite loci of nrDNA because the same folk variety had different sequencing results. Moreover, although ITS sequences could clarify the difference between the Demai variety and the remaining folk varieties, this may reflect the geographical distance between them. The sampling area of the Demai variety is far from those of the remaining folk varieties used in this study, and this long distance could act as a barrier to limit the gene flow between them. In fact, previous studies of Ehara et al. (2003) and Barahima et al. (2005) suggested that the geographical distance between folk varieties of the sago palm correlated with genetic relationships. Therefore, whether the ITS sequence is suitable to identify the Demai variety will be tested in future studies by analysing samples of other folk varieties collected from neighbour areas of the Demai sample used in this study. In addition, detailed analyses of dates of sago palm populations originating from different geographic locations will promote the understanding of their genomes and will reveal the true extent of gene flow between populations.

In summary, we provided molecular markers for folk varieties of the sago palm by analysing combined data of the ITS region and microsatellite loci of nrDNA. Further studies will determine whether more comprehensive sampling and additional genetic evidence of microsatellite loci of nrDNA help to understand folk varieties of the sago palm.

Acknowledgements

This research has been supported by a research grant from The Mitsui & Co., Ltd. Environment Fund.

References

Barahima, A., M. H. Bintoro, H. Sudarsono, M. Surahman and H. Ehara 2005 Haplotype diversity of sago palm in Papua based on chloroplast DNA. *In*: Sago Palm Development and Utilization (Karafir Y. P., F. S. Jong and V. E. Fere eds.), Proceeding of the 8th International Sago Symposium in Jayapura, Indonesia. Japan Society for the Promotion Science, Jayapura, 4-6 August 2005.

Barahima, A., M. H. Bintoro, H. Sudarsono, M. Surahman and H. Ehara 2009 Biodiversitas 10: 168-174.

Barahima, A., Y. Renwarin, M. H. Bintoro, H. Sudarsono, M. Surahman and H. Ehara 2010 Biodiversitas 11: 112-117.

Bowcock, A. M., A. Ruiz-Linares, J. Tomfohrde, E. Minch, J. R. Kidd and L. L. Cavalli-Sforza 1994 Nature 368: 455- 457.

Celiz, L. L., K. Toyota, M. Okazaki and A. Power 2004 genetic characteristics of sago palm (*Metroxylon sagu*) cultivars using AFLP (Amplified Fragment Length Polymorphism) markers. Abstract of the 13th Conference of Japanese Society Sago Palm Studies: 61-65.

Deka, R., L. Jin, M. D. Shriver, L. M. Yu, S. DeCroo, J. Hundrieser, C. H. Bunker, R. E. Ferrell and R. Chakraborty 1995 The American Journal of Human Genetics 56: 461-474.

Ehara, H., S. Kosaka, N. Shimura, D. Matoyama, O. Morita, H. Naito, C. Mizota, S. Susanto, M. H. Bintoro and Y. Yamamoto 2003 Sago Palm 11: 8-13.

Fukuda, T., H. Ashizawa, T. Nakamura, T. Ochiai, A. Kanno, T. Kameya and J. Yokoyama 2005a Plant Species Biology 20: 123-134.

Fukuda, T., I. J. Song, T. Nakamura, M. Nakada, A. Kanno, T. Kameya, H. Yamaji, A. Terabayashi, S.

Takeda, M. Aburada and J. Yokoyama 2005b Natural Medicines 59: 91-94.

Fukuda, T., I. J. Song , T. Ito, H. Nakayama, H. Hayakawa, Y. Minamiya, R. Arakawa, A. Kanno and J. Yokoyama 2012 Environmental Control in Biology 50: 13 – 18.

Jones, C. J., K. J. Edwards, S. Castaglione, M. O. Winfield, F. Sala, C. van de Wiel, G. Bredemeijer, B. Vosman, M. Matthes, A. Daly, R. Brettschneider, P. Bettini, M. Buiatti, E. Maestri, A. Malcevschi, N. Marmiroli, R. Aert, G. Volckaert, J. Rueda, R. Linacero, A. Vazquez and A. Karp 1997 Molecular Breeding 3: 381-390.

Jorde, L. B., M. J. Bamshad, W. S. Watkins, R. Zenger, E. Fraley, P. A. Krakowiak, K. D. Carpenter, H. Soodyall, T. Jenkins and A. R. Rogers 1995 The American Journal of Human Genetics 57: 523-538.

Kjær, A., A. S. Barfod, C. B. Asmussen and O. Seberg 2004 Annals of Botany 94: 109-117.

Hayakawa, H., T. Kobayashi, Y. Minamiya, K. Ito, A. Miyazaki, T. Fukuda and Y. Yamamoto 2010 The Journal of Japanese Botany 85: 263-269.

Hayakawa, H., T. Kobayashi, Y. Minamiya, K. Ito, A. Miyazaki, T. Fukuda and Y. Yamamoto 2011 American Journal of Plant Sciences 2: 15-26.

Nishizawa, T. and Y. Watano, 2000 Journal of Phytogeography and Taxonomy 48: 67-70.

Rieseberg, L. H. 1996 Homology among RAPD fragments in interspecific comparisons. Molecular Ecology 5: 99-105.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin and D. G. Higgins 1997 Nucleic Acids Research 24: 4876-4882.

White, T. J., T. Bruns, S. Lee and J. Taylor 1990 Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. *In*: PCR protocols: a guide to methods and application (Innis M, D. Gelfan, J. Sninsky and T. J. White eds.). Academic Press, San Diego 315-322.

Wolfe, A. D. and A. Liston 1998 Contribution of PCR-based methods to plant systematics and evolutionary biology. *In*: Molecular systematics of plants II (Soltis, D. E., P. S. Soltis and J. J. Doyle eds.). Kluwer Academic Publishing, London 43-86.

Yamamoto, Y., T. Yoshida, A. Miyazaki, F. S. Jong, Y. B. Pasolon and H. Matanubun 2005 Biodiversity and productivity of several sago palm varieties in Indonesia. *In*: Sago Palm Development and Utilization (Karafir, Y. P., F. S. Jong and V. E. Fere eds.), Proceeding of the 8th International Sago Symposium in Jayapura, Indonesia. Japan Society for the Promotion Science, Jayapura, 4-6 August 2005.