

## Investigation of Genetic and Morphological Variation in the Sago Palm (*Metroxylon sagu*; Arecaceae) in Papua New Guinea

ANDERS KJÆR<sup>1,\*</sup>, ANDERS S. BARFOD<sup>1</sup>, CONNY B. ASMUSSEN<sup>2</sup> and OLE SEBERG<sup>3</sup>

<sup>1</sup>Department of Systematic Botany, University of Aarhus, Nordlandsvej 68, DK-8240 Risskov, Denmark, <sup>2</sup>Department of Ecology, Royal Veterinary and Agricultural University, Rolighedsvej 21, DK-1958 Frederiksberg C, Denmark and <sup>3</sup>Department of Evolutionary Botany, University of Copenhagen, Gothersgade 140, DK-1123 Copenhagen K, Denmark

Received: 16 June 2003 Returned for revision: 10 September 2003 Accepted: 9 March 2004 Published electronically: 21 May 2004

- **Background and Aims** The genetic and morphological variation in the sago palm (*Metroxylon sagu*, Arecaceae) in Papua New Guinea (PNG) was investigated.
- **Methods** Amplified fragment length polymorphism (AFLP) was used to investigate the genetic structure of 76 accessions of *M. sagu*, collected in seven wild and semi-wild stands in PNG.
- **Key Results** An analysis of ten quantitative morphological variables revealed that most of these were mutually correlated. Principal component analyses of the same morphological variables showed that neither armature (presence or absence of spines) nor geographical separation was reflected clearly in the quantitative morphological variation. Similarity matrices of genetic, quantitative morphological, geographical and armature data were tested for pair-wise correlations, using Mantel's test. The results only showed a significant correlation between genetic and geographical distances. Visual inspection of principal component analyses plots and a neighbour-joining dendrogram based on genetic distances supported this trend, whereas armature showed no relation with genetic distances.
- **Conclusions** Geographical distribution defines some weak patterns in the genetic variation, whereas the genetic variation does not reflect any patterns in the morphological variation, including armature. The present study supports the accepted taxonomy of *M. sagu*, recognizing only one species of *M. sagu* in PNG.

© 2004 Annals of Botany Company

**Key words:** *Metroxylon sagu*, sago palm, AFLP, genetic variation, morphological variation, Papua New Guinea.

### INTRODUCTION

The sago palm, *Metroxylon sagu*, is an increasingly socio-economically important crop in South-East Asia. Its centre of diversity is believed to be New Guinea (Rauwerdink, 1986; Flach, 1994) or the Moluccas (Ehara, 2002). The hapaxanthic (or monocarpic) palm (Corner, 1966) can accumulate up to several hundred kilograms of starch in the trunk, which, after a period of 5–15 years, is enzymatically released and a large terminal inflorescence and thousands of fruits are produced. Many lowland dwellers in New Guinea have used the sago palm as a source of starch for many generations. Through trade routes, the sago palm has been transplanted within New Guinea and to other parts of South-East Asia. *Metroxylon sagu* is grown commercially in plantations in, for example, Indonesia and Malaysia.

In New Guinea, sago palms grow in vast wild and semi-wild dominant stands in many swampy lowland regions. Palms with various morphological features (e.g. presence or absence of spines) grow in completely intermixed populations. Notable morphologically variable features are length and presence of spines, length of trunk at maturity, diameter of trunk, leaf morphology, and properties and quantity of starch (Ehara *et al.*, 2000). Local sago growers in Papua New Guinea typically recognize 5–15 different local

varieties, based on elusive morphological features, and Schuilling (1995) gives an extensive account of local taxonomy and variability. Sexual reproduction in *M. sagu* is relatively well-studied (Tomlinson, 1971; Flach and Schuilling, 1989; Jong, 1995), and several researchers report that germinating sago seeds are rare (Barrau, 1960; Rauwerdink, 1986; Flach, 1994). Reproduction is probably mostly asexual by suckers.

Beccari (1918) recognized two species, *M. sagu* without spines and *M. rumphii* with spines, in his treatment of *Metroxylon* section *Metroxylon*. Rauwerdink (1986) later merged the two species into *M. sagu*, based on the fact that seeds from spineless palms can produce spiny seedlings (Rauwerdink, 1986; Ehara *et al.*, 1998). Nevertheless, Rauwerdink recognized four forms based on the length of the spines, proposing that spine length is controlled by a two-allele system. The question remaining to be answered based on this treatment is whether morphological markers such as presence of spines and length of spines are correlated with genetic variation, and if these markers can consequently be used in an infraspecific classification of *M. sagu*.

Possible segregations in morphology, geography and genetics are essential for future germplasm propagation, conservation and breeding programs. Molecular markers may provide a reliable tool for measuring genetic divergence (Jones *et al.*, 1997), and Kjær *et al.* (2002) have

\* For correspondence. Present address Bramstrup 5, DK-5792 Årslev, Denmark. E-mail: anders.kjaer@biology.au.dk

already demonstrated the usefulness of AFLP at the subspecific level in *M. sagu*.

The aim of the present study was to investigate possible relations between morphology, genetic structure and geographical distribution of individuals of *M. sagu* from Papua New Guinea. The null hypotheses tested was that no overall association existed between armature on the one hand and genetic distance and quantitative morphological variation on the other. The analyses were broken down into three parts. First, pairs of quantitative morphological variables were tested for mutual correlation. Then it was hypothesized that the armature and geographical distribution of individuals did not define the underlying quantitative morphological variation found by Principal Component Analysis (PCA). Second, hypotheses of correlations between geographical, genetic, and morphological distances were tested. Third, genetic distances were used as independent evidence to test whether differently armed individuals from the same locality were more closely related than similarly armed individuals from different localities.

## MATERIALS AND METHODS

### *Plant material and DNA extraction*

Field work in Papua New Guinea (PNG) was conducted from February to June 2000 at seven different localities (Fig. 1). A total of 76 individuals of *Metroxylon sagu* Rottb. (Rottböll, 1783) were sampled. To produce comparable results, individuals were only sampled if they were in the initial stage of flowering. This is considered by the farmers to be the appropriate harvest time. At each locality, the local villagers were interviewed to determine the number of locally distinguished varieties, and efforts were made to sample a representative selection of the local variation. From each palm ramet the following quantitative morphological variables were measured: trunk length as measured up to the lowermost leaf, diameter of the trunk at breast height (dbh), length of petiole, length of rachis, number of leaflets, length of longest leaflet, width of longest leaflet, thickness of petiole at first leaflet, width of petiole at first leaflet and number of green leaves in the crown. These variables were selected on the basis of a pilot study conducted in 1999 (Kjær, 2000). They represent vegetative characters only, as floral characters are too poorly represented due to the hapaxanthic life history. To ensure comparability between the sampled individuals, all leaf dimensions were measured on the oldest green leaf in the crown.

To ensure high quality of the DNA and to avoid contamination with DNA from epiphytic or endophytic organisms, the DNA samples were collected from the most recently developed leaf of a sucker adjoined to each of the sampled individuals. The leaf samples were dried in silica gel (Chase and Hills, 1991) in the field and frozen as soon as possible. DNA extractions were done using DNAeasy Plant Mini Kits (Qiagen, Germantown, MD, USA) according to the manufacturer's protocol. A total of 76 samples of *M. sagu* were extracted and included in the analysis.

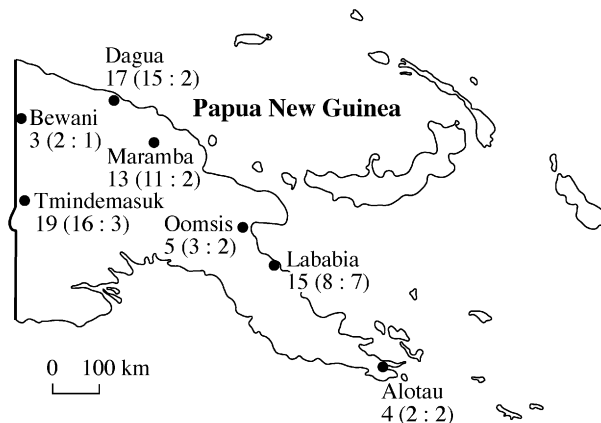


FIG. 1. Map of the seven collection localities in Papua New Guinea. Numbers indicate numbers of individuals collected and, in parentheses, the numbers of spined and non-spined individuals, respectively.

### *AFLP analysis*

In the present study, the AFLP procedure followed a protocol modified from that of Vos *et al.* (1995). The concentrations of double-stranded genomic DNA in the extracted samples were determined spectrophotometrically, and aliquots equivalent to 500 ng DNA were dried at 65 °C for approx. 2 h in 0.2 mL thin-walled tubes. The DNA was resuspended in 5.5 µL of double-distilled water. For each sample, master mix I contained 0.1 µL 10× T4 ligase buffer, 0.1 µL 0.5 M NaCl, 0.05 µL 1 mg mL<sup>-1</sup> BSA, 0.25 µL 4000 U mL<sup>-1</sup> *Mse*I restriction endonuclease, 0.25 µL 20 000 U mL<sup>-1</sup> *Eco*RI restriction endonuclease, 0.1 µL 10 U µL<sup>-1</sup> T4 DNA Ligase, and double-distilled water up to 1 µL. Master mix II contained 1 µL 10× T4 ligase buffer, 1 µL 0.5 M NaCl, 0.5 µL 1 mg mL<sup>-1</sup> BSA, 1 µL *Mse*I adaptor pair and 1 µL *Eco*RI adaptor pair. Prior to the preparation of master mix II, the two adaptor pairs were denatured at 95 °C for 5 min in a heat block and allowed to cool at room temperature for 10 min. Finally, 1.0 µL master mix I and 4.5 µL master mix II were added to each 5.5 µL DNA template. Digestion of the DNA and subsequent ligation of the enzyme specific adaptors were performed in a 2-h incubation at 37 °C in a PTC 200 thermocycler (MJ Research, Waltham, MA, USA). After the incubation, 189 µL TE<sub>0.1</sub> buffer was added to each sample.

For the preselective amplification, a master mix was prepared containing 7.5 µL AFLP Amplification Core Mix and 0.5 µL AFLP *Eco*RI and *Mse*I pre-amplification primers. In new 0.2 mL tubes, 0.8 µL master mix and 0.2 µL of the diluted restriction-ligation DNA were combined. The preselective amplification was done in a PTC 200 thermocycler using the program: 72 °C for 2 min, and 20 cycles of: 94 °C for 1 s, 56 °C for 30 s, 72 °C for 2 min and finally 60 °C for 30 min. From each sample, a 5 µL aliquot was run for 15 min at 150 V on a 1.5 % agarose gel to verify successful amplification of the target fragments. The remaining 5 µL pre-amplification product was diluted with 95 µL TE<sub>0.1</sub> buffer.

For the selective amplification, a master mix was prepared containing 7.5 µL AFLP Amplification Core Mix, 0.5 µL selective *MseI* primer and 0.5 µL selective *EcoRI* primer (each with three selective bases, see below) per sample. In 0.2 mL tubes, 8.5 µL master mix and 1.5 µL diluted pre-selective amplification product were combined. The selective amplification was done in a PTC 200 thermocycler using the program: one cycle of 94 °C for 2 min, 65 °C for 30 s, 72 °C for 2 min; eight cycles of 94 °C for 1 s, 64 °C for 30 s (the temperature was lowered 1 °C per cycle for this stage), 72 °C for 2 min; 23 cycles of 94 °C for 1 s, 56 °C for 30 s, 72 °C for 2 min; and finally 60 °C for 30 min.

Screening for usable primer combinations was done on a polyacrylamide gel in an ABI 377 sequencer (Applied Biosystems, Foster City, CA, USA), which can distinguish three different labels on the *EcoRI* primers. A total of 18 primer combinations were tested on the DNA of six individuals of *M. sagu*, and two combinations (*EcoRI*-ACT, *MseI*-CAT and *EcoRI*-ACC, *MseI*-CTT) were chosen for selective amplification of the DNA of all the 76 individuals of *M. sagu*. The primer labelling enables the laser in the sequencer to measure the strength of fluorescence at a given size class for each individual. For the purpose of further analysis, the fragments in a given size class were manually scored as either present (1) or absent (0). Peaks of less than 25 arbitrary fluorescence units were left out as background noise, as most peaks were several hundred or thousand units high. Furthermore, only fragments in the size range of 50–300 base pairs were included. The data from the two primer combinations were combined into one dataset.

#### Data analyses

Neighbour-joining was done in PAUP\* version 4.0b8 (Swofford, 2001). Mantel's correlation tests and principal coordinate (PCO) analyses were performed in R-Package version 4.0 (Casgrain and Legendre, 2001). Principal component analyses (PCA), normality tests, correlation tests, and some graphics were performed in SPSS 11.0.1 (SPSS, 2001).

In the first part of the analysis, only the quantitative morphological data were analysed. All variables were tested for normality: most showed normality within a 0.05 confidence limit. Only the values measured for dbh and number of leaflets fell outside the confidence limit. Visual inspection of normal distribution diagrams showed no signs of bimodality, thus normality was assumed for all variables in the remainder of the analyses. Using the 'bivariate correlation' option in SPSS, Pearson's correlation coefficients were computed, and a two-tailed test of significance was applied. PCAs were performed in SPSS using the 'factor analysis' option in the 'data reduction' option, and only components with eigenvalues of more than 1.0 were extracted, according to Manly (1986).

In the second part of the analysis, quantitative morphological, genetic, geographical and armature data were analysed. Prior to analysis, the quantitative morphological data were standardized to *z*-scores according to Legendre

and Legendre (1998). The Euclidian distance (*D*) matrix was computed and converted to a similarity (*S*) matrix by the formula  $S = 1 - D$ . A Nei genetic similarity matrix (Nei and Li, 1979) was computed for the genetic data. Using the geographical coordinates of the collection localities, a geographic distance matrix was computed and converted into a similarity matrix in R-package. A binary matrix was constructed to test hypotheses about armature. The distance was set to 0 between individuals with identical armature, and to 1 between individuals with different armature (spined *versus* non-spined). All matrices were tested for pair-wise correlation, using Mantel's non-parametric test (Mantel, 1967). The purpose of Mantel's test is to verify whether a correlation exists between two matrices by pair-wise comparison of the cells at corresponding positions. Correlation is expressed by the Mantel statistic. Under the null-hypothesis of no correlation, the value of this should not deviate significantly from the distribution of corresponding values obtained by repeatedly comparing one of the original matrices with 999 randomly generated matrices. The null hypothesis of no correlation is rejected when Mantel statistic falls outside the 0.05 confidence level. Pearson's *r*-value was used to measure linear correlation between two matrices. The datasets were further broken down into subsets according to individual populations. Legendre and Legendre (1998) state that Mantel's test is only reliable on datasets comprising more than four accessions. Consequently only the datasets containing individuals from the localities Tmindemasuk, Maramba, Lababia, Oomsis and Dagua were analysed. Only datasets showing significant correlations are presented here. The results were subjected to a Bonferroni correction (Sokal and Rohlf, 1995) *a posteriori*, which lowered the confidence level ( $\alpha$ ) in *k* consecutive analyses by the formula  $\alpha' = \alpha/k$ .

In the third part of the analysis, only the genetic data were analysed. Principal co-ordinate analysis (PCO) was performed using the R-package and SPSS for PC. An 'unrooted neighbour-joining phenogram' was constructed in PAUP\* using the 'Nei-Li restriction-site distances'.

## RESULTS

#### AFLP

The use of the selective primer combination *EcoRI*-ACT and *MseI*-CAT on the 76 DNA samples of *Metroxylon sagu* resulted in 41 bands, of which 16 were polymorphic. The use of *EcoRI*-ACC and *MseI*-CTT resulted in 26 bands, of which 16 were polymorphic. A total of 67 bands, including 32 polymorphic bands, were combined in one dataset and included in the data analyses.

#### Data analyses

Most of the ten quantitative morphological variables were highly intercorrelated (Table 1). The only exception was length of petiole. Dbh was well correlated with almost all the other variables (except length of trunk), whereas length of trunk only correlated well with three variables (number of leaves, length of petiole and width of petiole; Table 1).

TABLE 1. Two-tailed test of correlation between quantitative morphological variables

	dbh	Length of trunk	No. of leaves	No. of leaflets	Width of leaflets	Length of leaflets	Length of petiole	Length of rachis	Width of petiole	Thickness of petiole
dbh		-0.063	0.256	0.289	0.231	0.268	0.239	0.495	0.410	0.479
Length of trunk	0.586		0.319	0.028	0.118	0.149	-0.271	-0.143	0.232	0.097
No. of leaves	0.025*	0.005**		0.120	0.331	0.487	-0.035	0.352	0.481	0.412
No. of leaflets	0.011*	0.808	0.302		0.310	0.205	0.044	0.473	0.407	0.359
Width of leaflets	0.045*	0.310	0.004**	0.006**		0.588	0.075	0.474	0.557	0.441
Length of leaflets	0.019*	0.198	0.000**	0.075	0.000**		0.195	0.562	0.640	0.508
Length of petiole	0.038*	0.018*	0.764	0.706	0.520	0.092		0.408	-0.146	0.226
Length of rachis	0.000**	0.216	0.002**	0.000**	0.000**	0.000**	0.000**		0.454	0.544
Width of petiole	0.000**	0.044*	0.000**	0.000**	0.000**	0.000**	0.210	0.000**		0.671
Thickness of petiole	0.000**	0.406	0.000**	0.001**	0.000**	0.000**	0.049*	0.000**	0.000**	

Numbers above the diagonal indicate Pearson's correlation coefficient, and numbers below the diagonal indicate the probability of falsely rejecting the null hypothesis.

\*\* Probability within a 99 % confidence limit; \* probability within a 95 % confidence limit.

Principal component analysis of the ten quantitative morphological variables resulted in two components with eigenvalues larger than 1 (Table 2), which signifies that they explained more variation than the original variables (Manly, 1986). The first component, which explained 40.9 % of the total variation, received relatively high positive loadings from the variables thickness of petiole, width of petiole, length of leaflet and length of rachis, whereas the variables length of trunk and length of petiole only contributed with relatively low positive loadings. The second component, which explained an additional 16.5 % of the total variation, received the highest positive loading from the variable length of trunk, whereas the highest negative loading came from the variable length of petiole. In total the two components explained 57.4 % of the total variation in the dataset. The two components were applied as axes in two diagrams (Fig. 2A, B), and presence or absence of spines and collection localities were used as markers, respectively. Neither armature (Fig. 2A) nor geographical origin (Fig. 2B) defined any underlying quantitative morphological structure, since no clearly separated clusters were resolved.

The null hypothesis of no correlation between a similarity matrix based on the overall quantitative morphology and a genetic similarity matrix was not rejected by Mantel's test (Table 3). The tests of correlation between genetic distances and distances based on individual morphological characters revealed that thickness of petiole and width of petiole were correlated with the genetic distances (Table 3). Economically important characters such as dbh and length of trunk, on the other hand, were not significantly related to genetic variation (Table 3). Mantel's test did not reveal any correlation between armature, genetic distance and the overall quantitative morphological similarity (Table 3). Geographical and genetic distances were significantly correlated (Table 3). Mantel's test did not reject the null hypothesis of no correlation between geographical distance and the overall quantitative morphology (Table 3). In the data subsets representing single populations, no correlation between most of the variables was detected (Table 4). However, in Maramba genetic distance was correlated with

TABLE 2. Results of the principal component analysis

	Component I	Component II	Mean	s.d.
dbh	0.592	-0.307	46.2 cm	7.3 cm
Length of trunk	0.151	0.752	900 cm	300 cm
Number of leaves	0.604	0.381	14	3
Number of leaflets	0.526	-0.118	97	13
Width of leaflet	0.704	0.114	10 cm	2 cm
Length of leaflet	0.780	0.009	140 cm	17 cm
Length of petiole	0.222	-0.744	41 cm	21 cm
Length of rachis	0.779	-0.397	632 cm	108 cm
Width of petiole	0.820	0.313	10 cm	2 cm
Thickness of petiole	0.799	-0.005	4.3 cm	0.6 cm

The loadings of the individual morphological variables on the two components are listed. These two components explained 57 % of the total variation in the dataset. For each variable the mean values and standard deviations are given

the number of leaflets. In Dagua, the armature data was correlated with number of leaves. In Lababia, genetic distance was correlated with armature. Only in Tmindemasuk was armature weakly associated with both genetic distance and the morphological variables length of trunk and length of rachis (Table 4).

In the unrooted neighbour-joining network, the individuals of a specific locality or adjacent localities were clustered together in many instances, whereas spined and non-spined individuals were not clearly separated (Fig. 3). The first component of the PCO only explains 8 % of the total variation, whereas component two explains 6 % (Fig. 4A, B). No clear separation between accessions from spined and non-spined individuals were found on the the PCO plot based on genetic distance (Fig. 4A), whereas accessions from the same localities showed a weak tendency to group together (Fig. 4B).

## DISCUSSION

Amplified fragment length polymorphism (AFLP) was applied to reveal patterns of genetic variation in

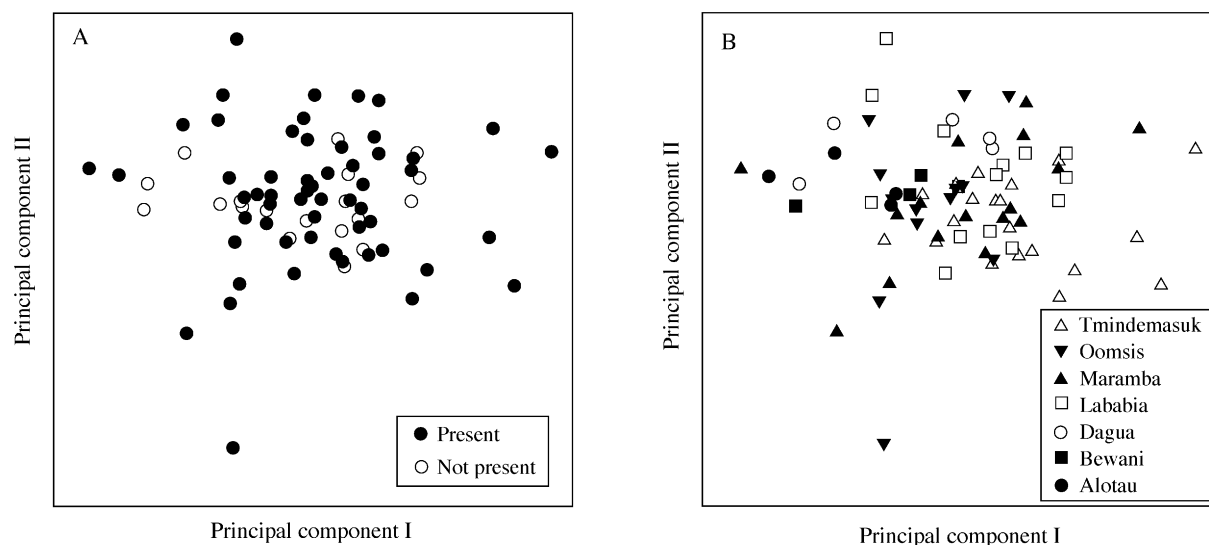


FIG. 2. Principal component analysis bi-plots based on the variation of ten vegetative morphological variables. The presence and absence of spines is indicated on (A), and collection localities on (B).

*Metroxylon sagu*. Several researchers (Cato *et al.*, 1999; Crouch *et al.*, 1999; Garcia-Mas *et al.*, 2000) recognize the strength and reproducibility of this method at the infraspecific level. Within the palm family *Arecaceae*, plastid DNA is known to evolve slowly (Wilson *et al.*, 1990), but many researchers (Kahn and Second, 1998; Perera *et al.*, 1998; Purba *et al.*, 2000; Teulat *et al.*, 2000; Matthes *et al.*, 2001) have found AFLP to be a valuable tool to distinguish palm populations and even palm individuals. In the present study, a relatively low number of polymorphic fragments (32) produced enough resolution to uniquely identify the large majority of the 76 investigated individuals (only five sets of duplicates and one set of triplicates were identified). Future studies should include more primer combinations, possibly resulting in more polymorphic fragments and better resolution.

Most of the morphological variables in the study were mutually correlated (Table 1). Only length of petiole and length of trunk seemed to be less associated with other variables. The length of petiole was highly variable. This was expected since all palms were sampled in the initial stages of flowering, when leaves are produced with progressively shorter petioles. Local sago farmers use this feature to identify harvest-ready palms. Interestingly, length of trunk is almost independent of the remaining variables (Table 1). This variable is possibly more strongly controlled by edaphic or climatic factors than by genetics, but this remains to be demonstrated.

The results discussed above were furthermore supported by the explanatory values of the two most important PCA components (Table 2). The highly intercorrelated variables (Table 1) contributed high positive variation loadings on principal component I, whereas the poorly intercorrelated variables length of petiole and length of trunk contributed high negative and high positive loadings, respectively, on principal component II (Table 2). The two principal

TABLE 3. Results of Mantel's test of the pair-wise correlations between similarity matrices

First matrix	Second matrix	Mantel's <i>r</i>	Probability
dbh	AFLP data	-0.023	0.337
Number of leaves	AFLP data	-0.007	0.47
Number of leaflets	AFLP data	0.047	0.483
Width of leaflets	AFLP data	0.025	0.288
Length of leaflets	AFLP data	0.035	0.23
Length of petiole	AFLP data	0.052	0.193
Length of rachis	AFLP data	0.007	0.445
Width of petiole	AFLP data	0.103	0.028*
Thickness of petiole	AFLP data	0.105	0.017*
All morphological data	AFLP data	0.047	0.228
Presence or absence of spines	AFLP data	0.002	0.481
Geographical distance	AFLP data	0.352	0.001* (B)
Presence or absence of spines	All morphological data	-0.046	0.22
Geographical distance	All morphological data	0.067	0.141

Rejection of the null hypothesis of no correlation within a 5 % confidence interval is marked '\*'. Rejection of the null hypothesis of no correlation using Bonferroni correction is marked 'B'.

components combined explained 57.4 % of the total variation in the dataset.

The distribution of spined and non-spined individuals plotted in the reduced two-way ordination space showed that armature was not linked to the quantitative morphological variation (Fig. 2A). This result supports the recent lumping of Beccari's (1918) two-species concept for *Metroxylon*, based on armature, into a single species (Rauwerdink, 1986). When origin of individuals was plotted on the ordination diagram, it became clear that the distances between populations were not reflected in the morphology in a readily interpretable manner (Fig. 2B). Ehara *et al.* (2000) performed an analysis of the relationship between

TABLE 4. Results of Mantel's test of correlation between similarity matrices for individual localities

Locality	First matrix	Second matrix	Mantel's <i>r</i>	Probability
Tmindemasuk	Presence or absence of spines	Length of trunk	-0.171	0.033*
Tmindemasuk	Presence or absence of spines	Length of rachis	-0.179	0.006*
Tmindemasuk	AFLP data	Presence or absence of spines	-0.759	0.003* (B)
Maramba	AFLP data	Number of leaflets	-0.367	0.014*
Dagua	Presence or absence of spines	Number of leaves	-0.185	0.039*
Lababia	AFLP data	Presence or absence of spines	-0.376	0.002* (B)

Only results indicating significant correlations are shown. Rejection of the null hypothesis of no correlation within a 5 % confidence interval is marked '\*'. Rejection of the null hypothesis of no correlation using Bonferroni correction is marked 'B'.

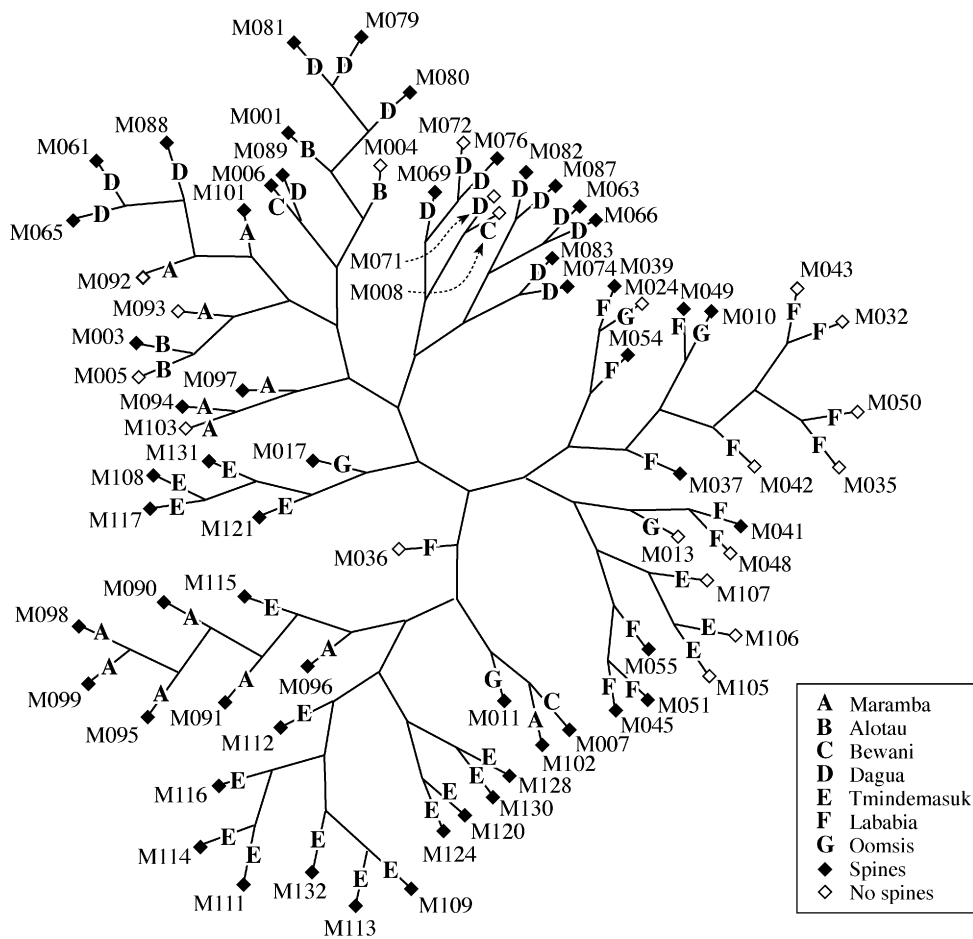


FIG. 3. Unrooted neighbour-joining hierarchy of genetic similarity as measured by using amplified fragment length polymorphism (AFLP). Collection localities are indicated in bold and armature (presence or absence of spines) with diamonds.

starch content and a number of morphological variables based on 11 accessions of *M. sagu* from Indonesia. The results showed that most variables (except, partly, dbh) were independent of starch content, excluding the measured variables as possible markers of this character. The combined results of the present study and the study of Ehara *et al.* (2000) indicate a high degree of independence between the geographical data, morphological data and starch content. More studies focusing on this particular problem are needed before a conclusion can be drawn.

The stepwise Mantel analyses of the relationships between the ten individual morphological variables and the genetic data showed that only thickness of petiole and width of petiole were significantly correlated with the AFLP data at the regional level (Table 3). The stepwise procedure, however, introduces an error that should be considered. In each of the ten tests of correlation there is a 5 % risk of making a Type I error, thereby rejecting a true hypothesis. Due to a cumulative effect, the risk of a Type I error is 32 % when ten consecutive experiments are performed. In order

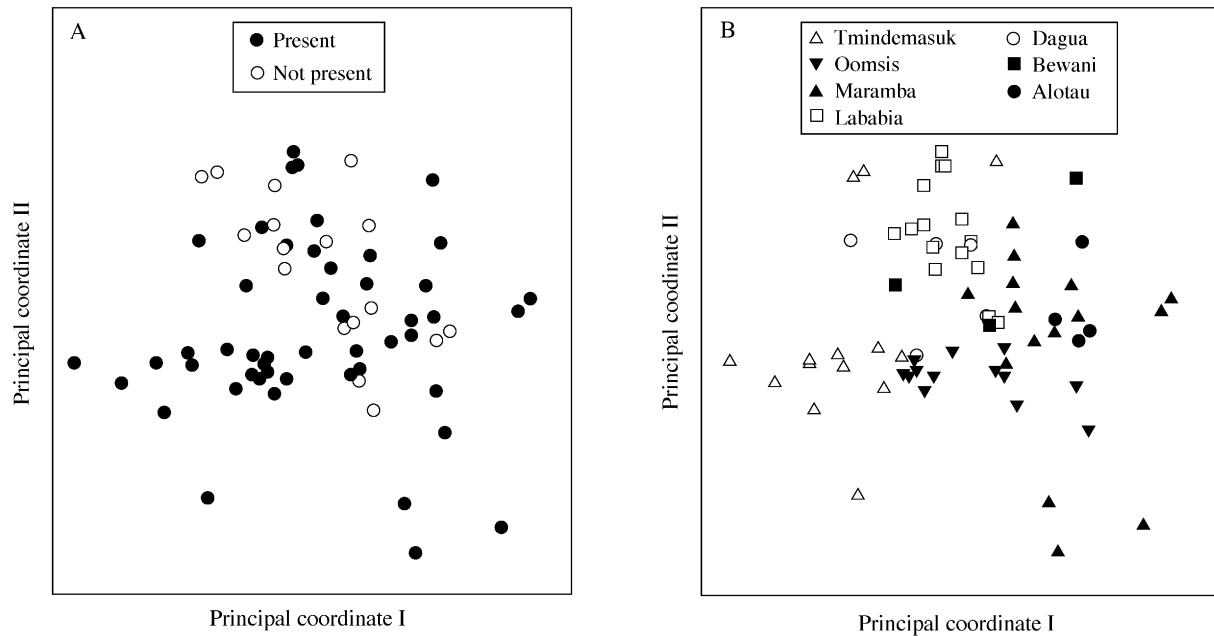


FIG. 4. Principal coordinate (PCO) bi-plots of genetic similarities of 76 individuals of *Metroxylon sagu* as measured by using amplified fragment length polymorphism (AFLP). The first principal coordinate represents 8 % of total variation, the second 6 %. Armature of the sampled individuals is indicated on (A), and the seven collection localities are indicated on (B).

to reduce this cumulative risk to 5 %, a Bonferroni correction was applied, which reduces the significance level to 0.005 in each test. None of the correlations between pairs of quantitative morphological variables that tested significant at first (Tables 3 and 4) remained significant when the overall experiment-wise error was Bonferroni-corrected (Sokal and Rohlf, 1995). Probably because of its conservative effect, Bonferroni correction is rarely used in connection with Mantel tests. The drawback of using it is the danger of a Type II error, thereby accepting a false hypothesis. Correlations in the zone between rejection and acceptance should be concluded upon with caution.

The results of the Mantel tests indicate that the genetic radiation detected by AFLP is not reflected in the quantitative morphological variables measured. Ehara *et al.* (2002) performed a RAPD (random amplified polymorphic DNA) analysis of 38 accessions of *M. sagu* from the Malay Archipelago, and found no relationship between genetic distances and pith colour or banding patterns on the back of the petiole and rachis. This information may prove valuable in the design of future breeding and propagation programs, as the results indicate that none of the morphological variables measured in the present study and in the Ehara *et al.* (2002) study were useful in the identification of specific genotypes.

The lack of correlation between the genetic data and armature revealed by the Mantel test (Table 3), was largely confirmed by the neighbour-joining dendrogram (Fig. 3), and clearly so by the PCO analysis (Fig. 4A). Ehara *et al.* (2002) similarly found that presence or absence of spines did not correspond to genetic distance.

Figure 3 shows that the four individuals from Alotau (B) (upper left part of dendrogram) tend to group with individuals from Maramba (A), Bewani (C) and Dagua (D), although these are situated at the opposite end of the country (Fig. 1). The population collected in Alotau (B) appeared to be a planted population, and it is likely that plant material was transplanted from the Sepik region (represented by populations A, C and D) in the past. The only three spineless individuals (M105, M106 and M107) from Tmindemasuk (E) form a group with individuals from Lababia (F) and Oomsis (G) (lower right part of dendrogram, Fig. 3), which we again take as an indication of human transplanting. This observation was consistent with the Bonferroni-corrected significant correlation between the genetic data and the armature in Tmindemasuk (Table 4). The correlation between the genetic data and the armature in Lababia (Table 4) is not easily explained.

The significant correlation between the genetic data and the geographical data under the strict Bonferroni correction (Table 3), was partly supported by the neighbour-joining dendrogram (Fig. 3) and the PCO analysis (Fig. 4B). Both analyses revealed the same tendency for the individuals to group according to localities, although in a less convincing way than the Mantel test (Table 3). Ehara *et al.* (2002) also showed a strong connection between genetic distance and geographical distance. The genetic structure of *M. sagu* is apparently characterized by a relatively strong genetic isolation caused by geographical distances. At the same time the incompletely resolved results seen in Figures 3 and 4A indicate that, seen in an evolutionary scope, exchange of genetic material between even remotely situated

populations happens relatively frequently. Whether this exchange is dominated by human transplantsations or by natural dispersal mechanisms remains to be answered.

The results of the data analyses demonstrated overall independence between quantitative morphology and genetic distances as measured using AFLP analysis. Future studies should consider possible relations between evolutionary lineages and economically important features, such as quantity and quality of starch, age at maturation, number of suckers, trunk length at maturation, trunk diameter, etc. It is possible that the variation in these features does not have a well-defined historical component, but is caused by differences in allele frequencies, as suggested by Rauwerdink (1986) for the explanation of the presence and length of spines.

The results of the present study support the presently accepted taxonomy of Rauwerdink (1986), recognizing only one species of *Metroxylon* in PNG. Future studies should include DNA from accessions of *M. sagu* throughout the distribution area, and especially from members of the previously recognized species *M. squarrosus* Becc., which according to Rauwerdink (1986) is conspecific with *M. sagu*. Based on a phylogenetic analysis of morphological characters, the taxonomy of the section *Coelococcus* of the genus *Metroxylon* was recently revised by McClatchey (1998), and a new species was described. McClatchey's (1998) findings should be tested using independent molecular evidence.

### CONCLUSIONS

The present study indicates that variation in vegetative morphological characters is not related to underlying genetic variation in the sago palm. This applies particularly to the presence or absence of spines. Genetic and geographical distances are generally linked. Some of the exceptions may be interpreted as recent introductions. The present study supports the presently accepted taxonomy (Rauwerdink, 1986), recognizing only one species of *M. sagu* in PNG.

### ACKNOWLEDGEMENTS

The fieldwork in Papua New Guinea was partly funded by Frimodt-Heinecke Fonden, Denmark, and the Faculty of Natural Science, University of Aarhus, Denmark, Danish Research Council for Natural Sciences (grant no. 9600861 to A.S.B.) and the Carlsberg Foundation (grant no. 980298/10-1150 to A.S.B.). We thank the herbarium staff at the PNG Forest Research Institute, Lae, especially assistant curator Roy Banka, for providing help with the logistics and assistance in the field. During the fieldwork, Anders Kjær received help from many individuals, too numerous to mention here. Special thanks to the villagers in Maramba, Dagua, Tmindemasuk, Lababia, Oomsis, Alotau and Ituly for their indispensable assistance in the field and unconditional commitment to our research. Many thanks to laboratory technician Charlotte Hansen, Botanical Institute, University of Copenhagen, Denmark. Dr Hiroshi Ehara, University of Mie, Japan has been a very valued collaborator. We are thankful to the Japanese Society of

Sago Palm Studies for inviting A.K. to speak at SAGO 2001 in Japan. We are grateful to Dr Finn Borchsenius, Department of Systematic Botany, University of Aarhus for inspiration and valuable discussions.

### LITERATURE CITED

- Barrau J. 1960.** The sago palm. *Principes* 4: 44–54.
- Beccari O. 1918.** Asiatic palms – Lepidocaryaceae. *Annals of the Royal Botanical Garden, Calcutta* 12: 156–195.
- Casgrain P, Legendre P. 2001.** *The R-package for multivariate and spatial analysis, version 4.0.* Canada: Departement de Sciences Biologiques, Université de Montréal.
- Cato SA, Corbett GE, Richardson TE. 1999.** Evaluation of AFLP for genetic mapping in *Pinus radiata* D. Don. *Molecular Breeding* 5: 275–281.
- Chase MW, Hills HG. 1991.** Silica gel: an ideal dessicant for preserving field-collected leaves for use in molecular studies. *Taxon* 41: 215–220.
- Corner EJH. 1966.** *Natural history of palms.* London: Weidenfeld & Nicolson.
- Crouch JH, Crouch HK, Constandt H, Van Gysel A, Breyné P, Van Montuga M, Jarret RL, Ortiz R. 1999.** Comparison of PCR-based molecular marker analyses of *Musa* breeding populations. *Molecular Breeding* 5: 233–244.
- Ehara H, Komada C, Morita O. 1998.** Germination characteristics of sago palm seeds and spine emergence in seedlings produced from spineless palm seeds. *Principes* 42: 212–217.
- Ehara H, Susanto S, Mizota C, Hirose S, Matsuno T. 2000.** Sago palm (*Metroxylon sagu*, *Arecaceae*) production in the eastern archipelago of Indonesia: variation in morphological characteristics and pith dry-matter yield. *Economic Botany* 54: 197–206.
- Ehara H, Kosaka S, Shimura N, Matoyama D, Morita O, Mizota C, Naito H, Susanto S, Bintoro MH, Yamamoto Y. 2002.** Genetic variation of sago palm (*Metroxylon sagu* Rottb.) in the Malay Archipelago. In: Kainuma K, Okazaki M, Toyada Y, Cecil JE, eds. *New frontiers of sago palm studies.* Tokyo: Universal Academy Press, 93–100.
- Flach M. 1994.** *Sago palm, Metroxylon sagu Rottb.* Promoting the conservation and use of underutilized and neglected crops, 15. Rome: IPGRI.
- Flach M, Schuiling DL. 1989.** Revival of an ancient starch crop: a review of the agronomy of the sago palm. *Agroforestry Systems* 7: 259–281.
- Garcia-Mas J, Olivier M, Gómez-Paniagua H, de Vicente MC. 2000.** Comparing AFLP, RAPD and RFLP markers for measuring genetic diversity in melon. *Theoretical and Applied Genetics* 101: 860–864.
- Jones CJ, Edwards KJ, Castaglione S, Winfeld MO, Sala F, van de Wiel C et al. 1997.** Reproducibility testing of RAPD, AFLP and SSR markers in plants by a network of European laboratories. *Molecular Breeding* 3: 381–390.
- Jong FS. 1995.** *Research for the development of sago palm (Metroxylon sagu Rottb.) cultivation in Sarawak, Malaysia.* PhD thesis, Wageningen Agricultural University, Netherlands.
- Kahn F, Second G. 1998.** The genus *Astrocaryum* (Palmae) in Amazonia: classical taxonomy and DNA analysis (AFLP). In: Andrew Henderson ed. *Evolution, variation, and classification of palms.* Memoirs of the New York Botanical Garden 83. New York Botanical Garden Press, 179–184.
- Kjær A. 2000.** *Morphological variation between spined and non-spined sago palms, Metroxylon sagu of Papua New Guinea.* Bachelor thesis, University of Aarhus, Denmark.
- Kjær A, Barfod AS, Asmussen CB, Seberg O. 2002.** Genetic and morphological variation in the sago palm *Metroxylon sagu* Rottb. (*Arecaceae*) in Papua New Guinea. In: Kainuma K, Okazaki M, Toyada Y, Cecil JE, eds. *New frontiers of sago palm studies.* Tokyo: Universal Academy Press, 100–101.
- Legendre P, Legendre L. 1998.** *Numerical ecology.* Amsterdam: Elsevier Science B.V.
- McClatchey WC. 1998.** Phylogenetic analysis of morphological characters of *Metroxylon* section *Coelococcus* (Palmae) and resulting implications for studies of other *Calamoideae* genera. In:



- Andrew Henderson, ed. *Evolution, variation, and classification of palms*. Memoirs of the New York Botanical Garden 83. New York Botanical Garden Press, 285–306.
- Manly BFJ. 1986.** *Multivariate statistical methods. A primer*. London: Chapman & Hall.
- Mantel N. 1967.** The detection of disease clustering and a generalized regression approach. *Cancer Research* **27**: 209–220.
- Matthes M, Singh R, Cheah SC, Karp A. 2001.** Variation in oil palm (*Elaeis guineensis* Jacq.) tissue culture-derived regenerants revealed by AFLPs with methylation-sensitive enzymes. *Theoretical and Applied Genetics* **102**: 971–979.
- Nei M, Li W. 1979.** Mathematical model for studying genetic variation in terms of restriction endonucleases. *Proceedings of the National Academy of Sciences of the USA* **76**: 5269–5273.
- Perera L, Russell JR, Provan J, McNicol JW, Powell W. 1998.** Evaluating genetic relationships between indigenous coconut (*Cocos nucifera* L.) accessions from Sri Lanka by means of AFLP profiling. *Theoretical and Applied Genetics* **96**: 545–550.
- Purba AR, Noyer JL, Baudouin L, Perrier X, Hamon S, Lagoda P.JL. 2000.** A new aspect of genetic diversity of Indonesian oil palm (*Elaeis guineensis* Jacq.) revealed by isoenzyme and AFLP markers and its consequences for breeding. *Theoretical and Applied Genetics* **101**: 956–961.
- Rauwerdink JB. 1986.** An essay on *Metroxylon*, the sago palm. *Principes* **30**: 165–180.
- Rottböll CF. 1783.** *Metroxylon*. *Nye Samling Kongelige Danske Videnskabers Selskabsskrifter, Anden Deel*. 527–529.
- Schuilling DL. 1995.** The variability of the sago palm and the need and the possibilities for its conservation. *Acta Horticulturae* **389**: 41–66.
- Sokal RR, Rohlf FJ. 1995.** *Biometry*, 3d edn. New York: W.H. Freeman and Company.
- SPSS. 2001.** *SPSS for Windows, Release 11.0.1*. Chicago: SPSS Inc.
- Swofford DL. 2001.** *PAUP. Phylogenetic Analysis Using Parsimony*, version 4.0b8. Washington: Laboratory of Molecular Systematics, Smithsonian Institution.
- Teulat B, Aldam C, Trehin R, Lebrun P, Barker JHA, Arnold GM, Karp A, Baudouin L, Rognon F. 2000.** An analysis of genetic diversity in coconut (*Cocos nucifera*) population from across the geographic range using sequence-tagged microsatellites (SSRs) and AFLPs. *Theoretical and Applied Genetics* **100**: 764–771.
- Tomlinson PB. 1971.** Flowering in *Metroxylon* (the Sago Palm). *Principes* **15**: 49–62.
- Vos P, Hogers R, Bleeker M, Reijans M, Van de Lee T, Hornes M et al. 1995.** AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Research* **23**: 4407–4414.
- Wilson MA, Gaut B, Clegg MT. 1990.** Chloroplast DNA evolves slowly in the palm family (Arecaceae). *Molecular Biology and Evolution* **7**: 303–314.