

Genetic diversity of sago palm  
(Metroxylon sagu Rottb.) accessions  
based on plastid cpDNA matK gene as  
DNA barcoding

*By* Barahima Abbas

## Genetic diversity of sago palm (*Metroxylon sagu* Rottb.) accessions based on plastid cpDNA *matK* gene as DNA barcoding

BARAHIMA ABBAS<sup>1,2,\*</sup>, IHWAN TJOLLI<sup>1</sup>, MUNARTI<sup>3</sup>

<sup>1</sup>Faculty of Agriculture, Universitas Papua, Jl. Gunung Salju, Manokwari 98314, West Papua, Indonesia. \*email: b.abbas@unipa.ac.id

<sup>2</sup>Post Graduate Program, Universitas Papua, Jl. Gunung Salju, Amban, Manokwari 98314, West Papua, Indonesia

<sup>3</sup>Faculty of Teacher Training and Education Science, Universitas Pakuan, Jl. Pakuan, Bogor 16143, West Java, Indonesia

Manuscript received: xxx. Revision accepted: xxx December 2019.

**Abstract.** Abbas B, Tjolli I, Mumarti. 2020. Genetic diversity of sago palm (*Metroxylon sagu* Rottb.) accessions based on plastid cpDNA *matK* gene as DNA barcoding. *Biodiversitas* 21: xxxx. *Metroxylon sagu* is one of the plant species including in the Palmae family that deposits starch in the trunk. Sago palm was reported has high capability to produce a large amount of starch and large variation starch qualities which were mostly influenced by genetic factors. The objectives of this study are revealed sequence chloroplast DNA (cpDNA) associated with *matK* genes for the genetic diversity identification of sago palm accessions. Plant materials used in the studies were derived from Sago Research Center (SRC) collection. DNA extraction has adopted the procedure of Plant Genomic DNA Mini Kit. Polymerase chain reaction (PCR) were performed by using primer sets of MatK-1RKIM-f and MatK-3FKIM-r. DNA PCR product was sequenced by the 1<sup>st</sup> Base Asia, Singapore. Results of the study .Showed that the cpDNA sequence associated with *matK* genes in the genome of sago palm showed differences among accessions. Molecular diversities of sago palm accessions based on *matK* gene showed sago palm accessions separated into two genotypes. Genotype-1 incorporated of ten individuals of 15 accessions and genotype-2 incorporated five individuals of 15 accessions. A mutation site and deletion site occurred in the sequences of *matK* gene of the Genotype-2. Ten of the 15 sequences of the *matK* gene (belong to Genotype-1) were registered in the GenBank, NCBI as DNA barcoding and authenticity of sago palm germplasm from Papua islands, Indonesia territories.

**Keywords:** DNA barcoding, genetic diversity, *matK* gene, plastid cpDNA, sago palm

### INTRODUCTION

Sago palm (*Metroxylon sagu* Rottb.) is included in the family of Arecaceae (Palmae) and the genus of *Metroxylon*. Beccari (1918) divides the Genus *Metroxylon* into two groups namely eumetroxylon which has 3 species (*M. sagu* Rottb., *M. rumphii* Mart., and *M. quarossum* Becc.) And *Coelococcus* has 6 species (*M. warburgii* Heim, *M. Upoluense* Becc., *M. vitiense* Benth, Et Hook, *M. micarum* var. *Commune* Becc., And var. *Majus* Becc., *M. salmonense* Becc., And *M. bougainvilense* Becc.). Sago palm is a plant that produces large amounts of starch in the trunk. The ability of sago palm to produce starch is the highest compared to other starch producing plants. Karim et al. (2008) reported that sago starch production can reach 3 to 4 times higher than rice, corn, and wheat. The capabilities of sago palm accumulated large amounts of starch in the trunk were mostly influenced by genetic factors and sometimes influenced by environment. Yater et al. (2019) reported that sago palm growing on different ecological habitat showed morphological and starch production alike.

The genetic information of organisms in certain areas is very important to revealed for avoid conflicts of interest in germplasm ownership from a particular organization. Besides, the conservation of germplasm requires genetic diversity information to optimize and maintain germplasm efficiently. Genetic diversity information of organisms needs to be documented to maintain the sustainability of

biological wealth and the existence of certain types of organisms, including sago palm. In the previous study, genetic marker were subjected to distinguished intraspecies of sago palm such as RAPD markers (Abbas et al. 2009), chloroplast DNA markers (Abbas et al. 2010), and mitochondrial *nad2* gene markers (Abbas et al. 2019).

The important tools for recognizing genetic characteristic of plant and others organism is DNA barcoding. The DNA barcoding was reported as reliable tools for identification intraspecies level of rice (Singh and Banerjee 2018) and distinguished 54% of 286 species by using *matK* and *rbcL* as markers (Kuzmina et al. 2012). DNA barcoding is one or more short gene sequences taken from standard genome parts and used to identify species (Kress and Erickson 2008). Several genes that can be used for DNA coding of in the plants are genes in the chloroplast genome including the *accD*, *matK*, *ndhJ*, *rpoB2*, *rpoC1*, and *ycf5* genes (Chase et al. 2007; Lahaye et al. 2008); *rbcL* (Kress and Erickson 2007); *trnL* intron 3 (Aberlet et al. 2007); and *trnH-psbA* (Kress et al. 2005). The Consortium for the Barcoding of Life (CBOL) Plant Working Group (2009) recommends three genes, namely *rbcL*, *matK*, and ITS, whereas in animal DNA sequences that can be used for barcoding are the cytochrome C oxidation (COI) genes in the mitochondrial genome. Kress and Erickson (2008) revealed that DNA barcoding can be used in the field of plant taxonomy and phylogenetics to identify plants more accurately than morphological identification. The DNA coding used in this study is *matK*

(Maturase K), which is a chloroplast gene and which is around 1500 base pairs (bp) located on the *trnK* intron. Hollingsworth et al. (2011) stated that currently the *matK* gene has been used as an important tool for examining the genetic diversity of intra-species and inter-species. This study aims to reveal the genetic diversity of *Metroxylon sagu* and relationship of among sago palm accessions in intraspecies level based on plastid *matK* gene marker.

## MATERIALS AND METHODS

### Plant Materials

Plant materials used in the studies were fifteen Sago genotypes derived from the Sago Palm Research Center (SRC) University of Papua (UNIPA) collections. Leaf samples were taken from accessions of sago palm in a growth russet stage. The Accession names are Sagu1, Sagu2, Sagu3, Sagu4, Sagu5, Sagu6, Sagu7, Sagu8, Sagu9, Sagu10, Sagu11, Sagu12, Sagu13, Sagu14, and Sagu15. The surface of the young leaf samples of sago palm accessions were wiped with an alcoholic tissue and brought into the UNIPA's Biotechnology Laboratory for further use.

### DNA Extraction

The DNA extraction was done by following the procedure of Plant Genomic DNA Mini Kit (Geneaid 2012). The outlines of DNA extraction using Geneaid protocols are tissue dissociation, lysis, DNA binding, wash, and DNA elution. Tissue dissociation was done by grind the sample to a fine powder. As much as 20 mg fine powder of the sample was transferred to a 1.5 ml microcentrifuge tube, then following step 2, step 3, step 4, and step 5 in the protocols. The genomic DNAs were extracted and stored at -20 °C freezer until ready used.

### Polymerase Chain Reaction (PCR) and Sequencing

Design of *matK* primer sequences used in this study is adopted from Kuzmina et al. (2012) and synthesized by Integrated DNA Technology (IDT) Singapore 117610. The Primer sequences, MatK-1RKIM-f 5'-ACCCAGTCCATCTGGAAATCTTGG TTC-3 and MatK-3FKIM-r 5'-CGTACAGTACTTT TGTGTTACGAG-3', amplified chloroplast DNA associated with *matK* gene in the sago palm chloroplast genome. PCR mixtures were 42 µl total volume containing: 1 x PCR buffer contained 1.5 mM MgCl<sub>2</sub>, 10 mM dNTP mix, 5 µl genomic DNA, 2.5 µl forward and reverse primer, 1 µl BSA, 1 µl DMSO, and 25 µl GoTaq Green. PCR condition is as follows: initial denaturation for 10 second at 80 °C and for 5 minutes at 94 °C, followed by 40 cycles of denaturation for 30 second at 94 °C, annealing for 30 seconds at 50 °C, for 45 second extension at 72 °C, the end of extension for 7 minutes at 72 °C, and the end of PCR cycle for 1 minutes at 37 °C. PCR amplification fragments were separated on 1% agarose gels by electrophoresis, staining was done using Ethidium Bromide and visualization by using UV illumination apparatus. Sequencing and purification of

DNA PCR product were performed by 1st Base Asia, Singapore 117610.

### Data Analysis

DNA sequences in the form of electropherogram were edited and checked to obtain the correct DNA sequence. Editing and proofreading sequences were performed by matching the peak color of the electropherogram to the sequence of nucleotides produced using Molecular Evolutionary Genetics Analysis (MEGA) version 7.0 software (Kumar et al. 2016). Each sequence in this study was obtained from the forward and reverse sequences of each sample. The editing result of a nucleotide sequence is stored in the fasta file format. Cluster alignment performed based on Clustal W with MEGA7 software. The evolutionary history was inferred using the Minimum Evolution (ME) method (Rzhetsky and Nei 1992). The evolutionary distances were computed using the Maximum Composite Likelihood method (Tamura et al. 2004) and are in the units of the number of base substitutions per site. The ME tree was searched using the Close-Neighbor-Interchange (CNI) algorithm (Nei and Kumar 2000) at a search level of 40. The Neighbor-joining algorithm (Saitou and Nei 1987) was used to generate the initial tree and bootstrap consensus tree was performed 500 replications (Felsenstein 1985). The analysis involved nucleotide sequences of 15 sago palm accessions. Codon positions included were 1st+2nd+3rd+Noncoding. All positions containing gaps and missing data were eliminated. There were a total of 604 positions in the final dataset. Evolutionary analyses based on the *matK* gene of sago palm accessions were conducted in MEGA7 (Kumar et al. 2016).

## RESULTS AND DISCUSSION

### Nucleotide Sequence of *matK* Genes

Alignments of the chloroplast sequences associated with *matK* genes within species of sago palm were presented in Figure 1. The alignment in Figure 1 shows two genotypes based on the *matK* gene sequence. Genotype 1 was identified in Sagu01, Sagu02, Sagu08, Sagu12, and Sagu15 and Genotype-2 was identified in Sagu03, Sagu04, Sagu05, Sagu06, Sagu07, Sagu09, Sagu10, Sagu11, Sagu13, and Sagu14. Figure 1 showed that Genotype-1 was identified by Thymine (T) base in the position of nucleotide 5<sup>th</sup> base and 16<sup>th</sup> base. On the other hand, Genotype-2 was identified by adenine (A) base in the position of nucleotide base 5<sup>th</sup> and nucleotide deletion in the position of nucleotide base 16<sup>th</sup>. This features showed that *matK* gene in the chloroplast genome of sago palm is highly conserved because just two sites of nucleotide sequence differences of 604 bases length. Selvaraj et al. (2008) reported that *matK* in the chloroplast genome is highly conserved and Provan et al. (1999) stated very low mutation rates which range from 3.2 x 10<sup>-5</sup> to 7.9 x 10<sup>-5</sup>. Both of chloroplast genome and mitochondrial genome of sago palm belonging to highly conserved DNA sequences. Abbas et al. (2019) observed that no differentiation among



sago palm accessions in the mitochondrial genome based on *nad2* gene marker. However, the nuclear genome of sago palm from Papua islands highly varied identified by Random Amplified Polymorphic DNA (RAPD) marker (Riyanto et al. 2018; Abbas 2018; Abbas et al. 2009).

The nucleotide sequencing results of an organism will be accepted and registered in the GenBank after verification and showed clearly perfect. This study, the nucleotides sequence *matK* genes of sago palm genotypes-1 were accepted and registered in the GenBank, NCBI with the sequence number BankIt2219548 Sagu03 MK860160, BankIt2219548 Sagu04 MK860161, BankIt2219548 Sagu05 MK860162, BankIt2219548 Sagu06 MK860163, BankIt2219548 Sagu07 MK860164, BankIt2219548 Sagu09 MK860165, BankIt2219548 Sagu10 MK860166, BankIt2219548 Sagu11 MK860167, BankIt2219548 Sagu13 MK860168, BankIt2219548 Sagu14 MK860169. The sequences of *matK* genes of sago palm genotypes-two were verified through repetitions process of DNA

extraction, PCR, and DNA sequencing to ensure the sequence stability. After three times repetitions of the DNA sequencing process, the results showed the same basic arrangement as before. Therefore, based on the verification results, it is confirmed that the accession of Sagu01, Sagu02, Sagu08, Sagu12, and Sagu15 has mutated in the position of two nucleotides base in the DNA sequence arrangement, marked with black box sequences (Figure 2). The first nucleotide base mutation is thymine to the base of adenine and the second nucleotide base mutation is the deletion of thymine base, marked as gap or dashes (Figure 2). The chloroplast genome identification in the previous studies showed that plant chloroplast genome possesses a low mutation rate, a little compactness, large size, and high rearrange structure (Darracq et al. 2011). One of the reasons caused a low mutation rate of chloroplast genome is uniparentally inherited (Savolainen et al. 1995; Viard et al. 2001).

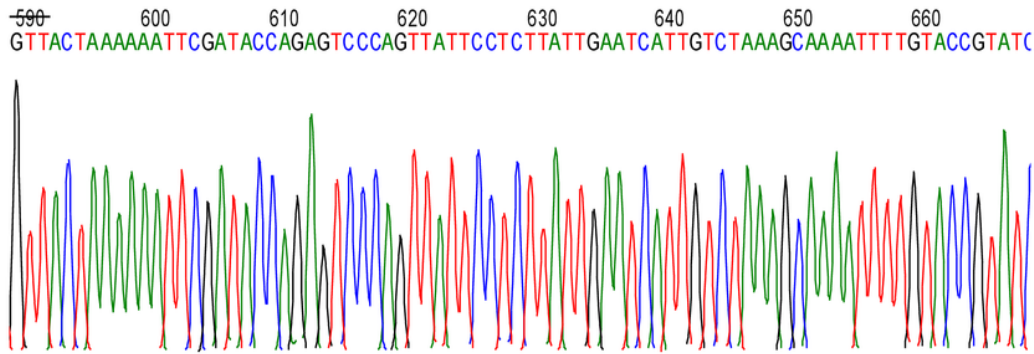


Figure 1. Performance of electropherogram sequencing result of sago palm accessions based on plastid *matK* gene.

SAGU01	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T	
SAGU02	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU03	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU04	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU05	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU06	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU07	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU08	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU09	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU10	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU11	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU12	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU13	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU14	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T
SAGU15	T	T	T	T	A	T	T	T	T	T	T	T	T	T	T	T	T	G	A	A	A	A	G	A	A	A	A	T	A	A	A	A	G	A	C	T	A	T	T	T	C	G	G	T	T	C	C	T

Figure 2. DNA sequences of sago palm genotype-1 and genotype-2 based on *matK* gene. The black boxes are segregation sites and deletion sites. Genotype-1 is Sagu03, Sagu04, Sagu05, Sagu06, Sagu07, Sagu09, Sagu10, Sagu11, Sagu13, and Sagu14. Genotype-2 is Sagu01, Sagu02, Sagu08, Sagu12, and Sagu15.

**Genetic Diversities of Sago Palm**

Nucleotide diversities of sago palm sequences based on the *matK* gene as barcoding were observed as low value that is 0.000788; only one number of segregation site and one number of deletion site (Table 1). In the previous

study using large amount of sago palm samples from around Indonesian territorial using cpDNA marker, a low variation was also reported and specific haplotype was found exist in the Papua island (Abbas et al. 2010).

**Table 1.** Tajima's neutrality test by using 15 nucleotide sequence of sago palm accessions based on plastid *matK* gene

m	S	ps	$\Theta$	$\Pi$	D
15	1	0.001656	0.000509	22.00788	1.122408

Notes: m = number of sequences, S = Number of segregating sites, ps = S/n,  $\Theta$  = ps/a,  $\pi$  = nucleotide diversity, and D is the Tajima test statistic (Tajima 1989).

Substitution patterns and rates were estimated under the Tamura and Nei (1993) model (+G). A discrete Gamma distribution was used to model evolutionary rate differences among sites (5 categories, [+G]). The mean value of the evolutionary rates in these categories were 0.90, 0.96, 1.00, 1.04, 1.10 substitutions per site. The nucleotide frequencies are A = 29.8%, T = 36.8%, C = 17.5%, and G = 15.9%. The highest nucleotide frequencies are Guanine (G) and the lowest nucleotide frequencies are Guanine (G) (Table 2).

The probability of rejecting the null hypothesis of the strict-neutrality ( $dN = dS$ ) is shown below diagonal, while

the test statistic ( $dN - dS$ ) is shown above the diagonal, in which dS and dN are the numbers of synonymous and nonsynonymous substitutions per site, respectively (Table 3). The variance of the difference was computed using the analytical method. Analyses were conducted using the Nei-Gojobori method (Nei and Gojobori (1986). The probability of 1.000 values (below diagonal) indicated rejecting the null hypothesis that mean plastid *matK* loci of the sago palm in the intra-species level is distinctly different. On the other hand, the probability of 0.319 values represented the null hypothesis. Therefore, the *matK* gene loci can be used as DNA barcoding for sago palm. The previous studies reported that DNA barcoding studies on the sand rice (*Agriophyllum squarrosum* (L.) Moq) identified 1.8% variabilities by using *matK* marker (Genievskaia et al. 2017), intra-specific mangroves were identified had 0.2% variabilities by using *matK* marker (Saddhe et al. 2016), and vascular plant were identified had 0.04% variabilities by using *matK* marker (Kuzmina et al. 2012).

**Table 2.** Nucleotide composition of 15 sago palm accessions based on plastid *matK* gene

	T	C	A	G	Total	T-1	C-1	A-1	G-1	Pos #1	T-2	C-2	A-2	G-2	Pos #2	T-3	C-3	A-3	G-3	Pos #3
Sagu01	36.6	17.5	30.0	15.9	604.0	33	22.3	26.7	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.5	15.5	200.0
Sagu02	36.6	17.5	30.0	15.9	604.0	33	22.3	26.7	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.5	15.5	200.0
Sagu03	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu04	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu05	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu06	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu07	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu08	36.6	17.5	30.0	15.9	604.0	33	22.3	26.7	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.5	15.5	200.0
Sagu09	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu10	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu11	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu12	36.6	17.5	30.0	15.9	604.0	33	22.3	26.7	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.5	15.5	200.0
Sagu13	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu14	36.9	17.5	29.8	15.9	605.0	33	22.3	26.2	18.3	202.0	36	18.3	31.7	13.9	202.0	41	11.9	31.3	15.4	201.0
Sagu15	36.6	17.5	30.0	15.9	604.0	33	22.3	26.7	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.5	15.5	200.0
Avg.	36.8	17.5	29.8	15.9	604.7	33	22.3	26.4	18.3	202.0	36	18.3	31.7	13.9	202.0	41	12.0	31.4	15.4	200.7

**Table 3.** Codon-based Test of Neutrality for analysis between sequences of sago palm intra-species by using Z-Test of Neutral evolution

Acces.	Sagu01	Sagu02	Sagu03	Sagu04	Sagu05	Sagu06	Sagu07	Sagu08	Sagu09	Sagu10	Sagu11	Sagu12	Sagu13	Sagu14	Sagu15
Sagu01	1.000	0.000	1.001	1.001	1.001	1.001	1.001	0.000	1.001	1.001	1.001	0.000	1.001	1.001	0.000
Sagu02	1.000	1.000	1.001	1.001	1.001	1.001	1.001	0.000	1.001	1.001	1.001	0.000	1.001	1.001	0.000
Sagu03	0.319	0.319	1.000	0.000	0.000	0.000	0.000	1.001	0.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu04	0.319	0.319	1.000	1.000	0.000	0.000	0.000	1.001	0.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu05	0.319	0.319	1.000	1.000	1.000	0.000	0.000	1.001	0.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu06	0.319	0.319	1.000	1.000	1.000	1.000	0.000	1.001	0.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu07	0.319	0.319	1.000	1.000	1.000	1.000	1.000	1.001	0.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu08	1.000	1.000	0.319	0.319	0.319	0.319	0.319	1.000	1.001	1.001	1.001	0.000	1.001	1.001	0.000
Sagu09	0.319	0.319	1.000	1.000	1.000	1.000	1.000	0.319	1.000	0.000	0.000	1.001	0.000	0.000	1.001
Sagu10	0.319	0.319	1.000	1.000	1.000	1.000	1.000	0.319	1.000	1.000	0.000	1.001	0.000	0.000	1.001
Sagu11	0.319	0.319	1.000	1.000	1.000	1.000	1.000	0.319	1.000	1.000	1.000	1.001	0.000	0.000	1.001
Sagu12	1.000	1.000	0.319	0.319	0.319	0.319	0.319	1.000	0.319	0.319	0.319	1.000	1.001	1.001	0.000
Sagu13	0.319	0.319	1.000	1.000	1.000	1.000	1.000	0.319	1.000	1.000	1.000	0.319	1.000	0.000	1.001
Sagu14	0.319	0.319	1.000	1.000	1.000	1.000	1.000	0.319	1.000	1.000	1.000	0.319	1.000	1.000	1.001
Sagu15	1.000	1.000	0.319	0.319	0.319	0.319	0.319	1.000	0.319	0.319	0.319	1.000	0.319	0.319	1.000

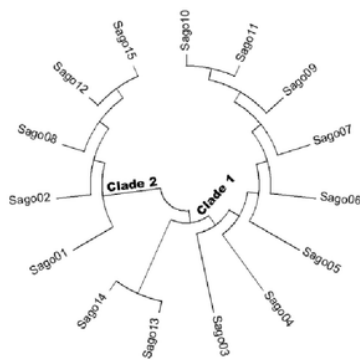
**Genetic Relationship**

The evolutionary distances of 15 sago palm accessions based on *matK* genes showed that the genotype-1 and genotype-2 were different with amount 0.002 evolutionary distances (Table 4). The relationships among 15 sago palm accessions were separated into two clades and the optimal tree with the sum of branch length = 0.00165716 is shown (Figure 3). Clade-1 is incorporated sago palm accessions number Sagu03, Sagu04, Sagu05, Sagu06, Sagu07, Sagu09, Sagu10, Sagu11, Sagu13, and Sagu14. Clade-1 which were identified as genotype-1. The clade-2 is incorporated sago palm accessions number Sagu01, Sagu02, Sagu08, Sagu12, and Sagu15, identified as Genotype-2 (Figure 3). The differentiation between Genotype-1 and Genotype-2 may occur in the process of

cpDNA replication for a long time. In the previous studies showed that the interspecific sequence divergence of *Astragalus* was reached 3.92% (Liston 1992). The specific chloroplast genes of Conifers were detected at 23, 26, 38, 48, 67, and 25 site changes in the *frxC*, *rbcL*, *psbA*, *psbD*, *trnK*, and 16S respectively among species of Conifer (Tsumura et al. 1995). Segregations in the DNA plastid of *matK* gene in the sago palm genotype-2 were in harmony with phylogenetic of Dipterocarpaceae based on the *matK* gene (Hamelly et al. 2018), and corresponds to the DNA barcode by using the *matK* gene in Pandanus (Zebua et al. 2019). Therefore, the DNA Barcode by using plastid *matK* gene can be used to determine intra-species of the sago palm. This finding.

**Table 4.** The evolutionary distances of sago palm based on *matK* gene were computed using the Maximum Composite Likelihood method.

Acces.	Sagu01	Sagu02	Sagu03	Sagu04	Sagu05	Sagu06	Sagu07	Sagu08	Sagu09	Sagu10	Sagu11	Sagu12	Sagu13	Sagu14	Sagu15
Sagu01	0.000														
Sagu02	0.000	0.000													
Sagu03	0.002	0.002	0.000												
Sagu04	0.002	0.002	0.000	0.000											
Sagu05	0.002	0.002	0.000	0.000	0.000										
Sagu06	0.002	0.002	0.000	0.000	0.000	0.000									
Sagu07	0.002	0.002	0.000	0.000	0.000	0.000	0.000								
Sagu08	0.000	0.000	0.002	0.002	0.002	0.002	0.002	0.000							
Sagu09	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.002	0.000						
Sagu10	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000					
Sagu11	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000				
Sagu12	0.000	0.000	0.002	0.002	0.002	0.002	0.002	0.000	0.002	0.002	0.002	0.000			
Sagu13	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.002	0.000		
Sagu14	0.002	0.002	0.000	0.000	0.000	0.000	0.000	0.002	0.000	0.000	0.000	0.002	0.000	0.000	
Sagu15	0.000	0.000	0.002	0.002	0.002	0.002	0.002	0.000	0.002	0.002	0.002	0.000	0.002	0.002	0.000



**Figure 3.** Evolutionary relationships of sago palm accessions based on *matK* genes in the plastid genome

In conclusions, this study indicated that cpDNA sequences associated with *matK* genes in the genome of

sago palm were shown low differences among accessions. Nucleotide diversities of sago palm sequences based on the *matK* gene as barcoding were observed at 0.000788 value. Molecular phylogenetic of sago palm based on the *matK* gene markers showed that the sago palm intra-species incorporated into two clades. DNA barcoding using *matK* gene might be used as a tool for intra-species identification of the sago palm.

**ACKNOWLEDGMENTS**

The research was financially funded by the Directorate General Strengthen Research and Development Republic Indonesia with Project title "Transformation of the invention based on sago palm to become innovation for entrepreneurship orientation to accelerate food competitiveness and independent of the Nation" under contract No.198/SP2H/LT/DRPM/2019. Authors thank to



the Research Development Project Manager and Food Security Service of West Papua Province.

## REFERENCES

- Abbas B, Bintoro MH, Sudarsono, Surahman M, Ehara H. 2009. Genetic Relationship of Sago Palm (*Metroxylon Sagu* Rottb.) in Indonesia Based on RAPD Markers. *Biodiversitas, Journal of Biological Diversity* 10 (4): 168–174. Abbas B, Renwarin Y, Bintoro MH, Sudarsono, Surahman M, Ehara H. 2010. Genetic Diversity of Sago Palm in Indonesia Based on Chloroplast DNA (CpDNA) Markers. *Biodiversitas, Journal of Biological Diversity* 11 (3): 112–117. <https://doi.org/10.13057/biodiv/d110302>.
- Abbas B, Tjollil I, Dailami M, Munarti. 2019. Phylogenetic of Sago Palm (*Metroxylon Sagu*) and Others Monocotyledon Based on Mitochondrial Nad2 Gene Markers. *Biodiversitas, Journal of Biological Diversity* 20(8): 2249–2256. <https://doi.org/10.13057/biodiv/d200820>.
- Abbas B. 2018. Sago Palm Genetic Resource Diversity in Indonesia. In: Ehara H, Toyoda Y, Johnson D. (eds) *Sago Palm: Multiple Contributions to Food Security and Sustainable Livelihoods*. Springer, Singapore. <https://rd.springer.com/book/10.1007/978-981-10-5269-9>.
- Beccari O. 1918. Asiatic palms. *Lepidocarpaceae*. *Ann. Roy. Bot. Gard. Calcutta*. 12:156-195.
- Chase MW, Cowan RS, Hollingsworth PM, Berg CVD, Madrin'an S, Petersen G, Seberg O, Jorsensen T, Cameron KM, Carine M, Pedersen N, Hedderson TAJ, Conrad F, Salazar GA, Richardson JE, Hollingsworth ML, Barraclough TG, Kelly L, Wilkinson M. 2007. A proposal for a standardized protocol to barcode all land plants. *Taxon* 56: 295–299.
- Darracq A, Varre JS, Drouard LM, Courseaux A, Castric V, Laprade PS, Ozias S, Lenoble P, Barbe B, Touzet P. 2011. Structural and content diversity of mitochondrial genome in beet: A comparative genomic analysis. *Genome Biol Evol* 3: 723–736.
- Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution* 39:783-791.
- Genievskaia Y., Abugalieva S., Zhubanysheva A. Turuspekov Y. 2017. Morphological description and DNA barcoding study of sand rice (*Agriophyllum squarrosum*, Chenopodiaceae) collected in Kazakhstan. *BMC Plant Biology*. 17(1):177-185. DOI 10.1186/s12870-017-1132-1.
- Hamelly E, Thomy Z, Fathiya N. 2018. Phylogenetic analysis of Dipterocarpaceae in Ketambe Research Station, Gunung Leuser National Park (Sumatra, Indonesia) based on rbcL and matK genes. *Biodiversitas* 19: 1074-1080. <https://doi.org/10.13057/biodiv/d190340>.
- Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and Using a Plant DNA Barcode. *PLoS ONE*. 6(5):1-13.
- Karim AA, Pei-Lang Tie A, Manan DMA, Zaidul ISM. 2008. Starch from the sago (*Metroxylon sago*) palm tree-properties, prospects and challenges as a source for food and other uses. *Compr Rev Food Sci Food Saf* 7: 215-228.
- Kress WJ, Erickson DL. 2008. DNA DNA barcoding: Genes, genomics, and bioinformatics. *Proceedings of the National Academy of Sciences USA*, 105:2761–2762.
- Kress W J, Erickson DL. 2007. A two-locus global DNA barcode for land plants: The coding rbcL gene complements the noncoding trnH-psbA spacer region. *PLoS ONE* 2: e508. doi: 10.1371/journal.pone.0000508.
- Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA, Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *Proceedings of the National Academy of Sciences USA*, 102: 8369–8374.
- Kumar S, Stecher G, Tamura K. 2016. MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular Biology and Evolution* 33:1870-1874.
- Kuzmina ML, Johnson KL, Barron HR, Herbert PDN. 2012. Identification of vascular plants of Chure, Manitoba, using a DNA barcode library. *BMC Ecology* 12(25):1-11. Doi:10.1186/1472-6785-12-25.
- Lahaye R, Bank MVD, Bogarin D, Wamer J, Pupulin F, G. Maurin O, Duthoit S, Barraclough TG, Savolainen V. 2008. DNA barcoding the floras of biodiversity hotspots. *Proceedings of the National Academy of Sciences USA*, 105: [www.pnas.org/cgi/doi/10.1073/pnas.0709936105](http://www.pnas.org/cgi/doi/10.1073/pnas.0709936105).
- Liston A. 1992. Variation in the chloroplast genes RPOC1 and RPOC2 of the genus *Astragalus* (Fabaceae): Evidence from restriction site mapping of a PCR amplified fragment. *Am J Bot* 79: 953-961.
- Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* 3:418-426.
- Nei M, Kumar S. 2000. *Molecular Evolution and Phylogenetics*. Oxford University Press, New York.
- Provan J, Soranzo N, Wilson NJ, Goldstein DB, Powell W. 1999. A low mutation rate for chloroplast microsatellites. *Genetics* 153: 943-947.
- Riyanto R, Widodo I, Abbas B. 2018. Morphology, growth and genetic variations of sago palm (*Metroxylon sago*) seedlings derived from seeds. *Biodiversitas* 19: 602-608.
- Rzhetsky A, Nei M. 1992. A simple method for estimating and testing minimum evolution trees. *Molecular Biology and Evolution* 9:945-967.
- Saddhe AA, Jamdade AR, Kumar K. 2016. Assessment of mangroves from Goa, west coast India using DNA barcode. *SpringerOpen* 5:1554-1564. DOI 10.1186/s40064-016-3191-4.
- Saitou N, Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution* 4:406-425.
- Savolainen V, Corbaz R, Moncousin C, Spchiger R, Manen JF. 1995. Chloroplast DNA variation and parentage analysis in 55 apples. *Theor Appl Genet* 90: 1138-1141.
- Selvaraj D, Sarma RK, Sathishkumar R. 2008. Phylogenetic analysis of chloroplast matK gene from Zingiberaceae for plant DNA barcoding. *Bioinformation* 3(1): 24-27.
- Singh J, Banerjee S. 2018. Utility of DNA barcoding tools for conservation and molecular identification of intraspecies of rice genotypes belonging to Chhattisgarh using rbcL and matK gene sequences. *Plant Archives* 18: 69-75.
- Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier, Brochmann C, Willerslev E. 2007. Power and limitations of the chloroplast trnL (UAA) intron for plant DNA barcoding. *Nucleic Acids Research*. 35(3):1-8. doi:10.1093/nar/gkl938
- Tajima F. 1989. Statistical methods to test for nucleotide mutation hypothesis by DNA.
- Tamura K, Nei M. 1993. Estimation of the number of nucleotide substitutions in the control region of mitochondrial DNA in humans and chimpanzees. *Molecular Biol and Evolution* 10:512-526.
- Tamura K, Nei M, Kumar S. 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proceedings of the National Academy of Sciences (USA)* 101:11030-11035.
- Tsumura Y, Yoshimura K, Tomaru N, Ohba K. 1995. Molecular phylogeny of conifer using RFLP analysis of PCR-amplified specific chloroplast genes. *Theor Appl Genet* 91: 1222-1236.
- Viard F, El-Kassaby YA, Ritland A. 2001. Diversity and genetic structure in populations of *Pseudotsuga menziesii* (Pinaceae) at chloroplast microsatellite loci. *Genome* 44: 336-344.
- Yater T, Tubur HW, Meliala C, Abbas B. 2019. Short Communication: A comparative study of phenotypes and starch production in sago palm (*Metroxylon sago*) growing naturally in temporarily inundated and non-inundated areas of South Sorong, Indonesia. *Biodiversitas* 20: 1121-1126. <https://doi.org/10.13057/biodiv/d200425>.
- Zebua LI, Gunaedi T, Budi IM, Lunga N. 2019. The DNA barcode of red fruit pandan (*Pandanaceae*) cultivar from Wamena, Papua Province, Indonesia based on matK gene. *Biodiversitas* 20: 3405-3412. <https://doi.org/10.13057/biodiv/d201138>.

# Genetic diversity of sago palm (*Metroxylon sagu* Rottb.) accessions based on plastid cpDNA matK gene as DNA barcoding

---

ORIGINALITY REPORT

---

# 19%

SIMILARITY INDEX

---

PRIMARY SOURCES

---

- |   |  |                 |
|---|--|-----------------|
| 1 | <a href="https://hdl.handle.net">hdl.handle.net</a><br>Internet  | 88 words — 2%   |
| 2 | Farshid GOLSHANI, Barat Ali FAKHERI, Mahmood SOLOUKI, Nafiseh MAHDINEZHAD, Majid Reza KIANI FERIZ. "Study of phylogenic relationships of some wild and crop species of Iranian Crocus by ITS nuclear loci", BIOCELL, 2019<br>Crossref                      | 71 words — 1%   |
| 3 | Lina Maloukh, Alagappan Kumarappan, Mohammad Jarrar, Jawad Salehi, Houssam El-wakil, T. V. Rajya Lakshmi. "Discriminatory power of rbcL barcode locus for authentication of some of United Arab Emirates (UAE) native plants", 3 Biotech, 2017<br>Crossref | 63 words — 1%   |
| 4 | <a href="https://repository.up.ac.za">repository.up.ac.za</a><br>Internet  | 58 words — 1%   |
| 5 | <a href="https://journals.plos.org">journals.plos.org</a><br>Internet  | 49 words — 1%   |
| 6 | <a href="https://www.spygen.com">www.spygen.com</a><br>Internet  | 43 words — 1%   |
| 7 | Leandro A. Sánchez. "Cold-adapted microorganisms as a source of new antimicrobials", Extremophiles, 01/2009<br>Crossref  | 28 words — < 1% |



8	<a href="http://www.lseb.cn">www.lseb.cn</a> Internet	27 words — < 1%
9	<a href="http://jmd.amjpathol.org">jmd.amjpathol.org</a> Internet	27 words — < 1%
10	Jim Provan. "Restricted gene flow in fragmented populations of a wind-pollinated tree", <i>Conservation Genetics</i> , 12/2008 Crossref	26 words — < 1%
11	<a href="http://www.public.asu.edu">www.public.asu.edu</a> Internet	26 words — < 1%
12	Hai -Yan Xu, Hua-Chuan Zheng, Hui -Wen Zhang, Jin-Yu Zhang, Chao-Mei Ma. " Comparison of Antioxidant Constituents of Seed with Conventional Crop Seeds ", <i>Journal of Food Science</i> , 2018 Crossref	26 words — < 1%
13	Svetlana V. Nikiforova, Duccio Cavalieri, Riccardo Velasco, Vadim Goremykin. "Phylogenetic Analysis of 47 Chloroplast Genomes Clarifies the Contribution of Wild Species to the Domesticated Apple Maternal Line", <i>Molecular Biology and Evolution</i> , 2013 Crossref	25 words — < 1%
14	<a href="http://gemi.mpl.ird.fr">gemi.mpl.ird.fr</a> Internet	25 words — < 1%
15	Jun An, Jun-Cheol Moon, Cheol Seong Jang. "Markers for distinguishing <i>Orostachys</i> species by SYBR Green-based real-time PCR and verification of their application in commercial <i>O. japonica</i> food products", <i>Applied Biological Chemistry</i> , 2018 Crossref	24 words — < 1%
16	<a href="http://perpustakaan.unitomo.ac.id">perpustakaan.unitomo.ac.id</a> Internet	24 words — < 1%
17	<a href="http://medinalab.org">medinalab.org</a> Internet	

		24 words — < 1%
18	<a href="http://edepot.wur.nl">edepot.wur.nl</a> Internet	24 words — < 1%
19	<a href="http://estudogeral.sib.uc.pt">estudogeral.sib.uc.pt</a> Internet	23 words — < 1%
20	<a href="http://fsd.unsyiah.ac.id">fsd.unsyiah.ac.id</a> Internet	23 words — < 1%
21	Livia Lucentini, Manuela Rebora, Maria Elena Puletti, Lilia Gigliarelli, Diego Fontaneto, Elda Gaino, Fausto Panara. "Geographical and seasonal evidence of cryptic diversity in the Baetis rhodani complex (Ephemeroptera, Baetidae) revealed by means of DNA taxonomy", <i>Hydrobiologia</i> , 2011 Crossref	22 words — < 1%
22	<a href="http://bmcgenomics.biomedcentral.com">bmcgenomics.biomedcentral.com</a> Internet	22 words — < 1%
23	"Research Developments in Saline Agriculture", Springer Science and Business Media LLC, 2019 Crossref	20 words — < 1%
24	<a href="#">Biotechnology for Medicinal Plants, 2013.</a> Crossref	18 words — < 1%
25	<a href="http://www.ncbi.nlm.nih.gov">www.ncbi.nlm.nih.gov</a> Internet	17 words — < 1%
26	<a href="http://www.frontiersin.org">www.frontiersin.org</a> Internet	16 words — < 1%
27	<a href="http://www.pnas.org">www.pnas.org</a> Internet	14 words — < 1%
28	<a href="http://medcraveonline.com">medcraveonline.com</a> Internet	13 words — < 1%

- 
- 29 [media.neliti.com](http://media.neliti.com) Internet 13 words — < 1%
- 
- 30 S. Winter. "The analysis of Cassava brown streak viruses reveals the presence of distinct virus species causing cassava brown streak disease in East Africa", *Journal of General Virology*, 01/13/2010  
Crossref 13 words — < 1%
- 
- 31 [wwwnc.cdc.gov](http://wwwnc.cdc.gov) Internet 12 words — < 1%
- 
- 32 Karmakar Ruma, Kumar Sunil, Kukkundoor R. Kini, Harischandra Sripathy Prakash. "Genetic diversity and antimicrobial activity of endophytic *Myrothecium* spp. isolated from *Calophyllum apetalum* and *Garcinia morella*", *Molecular Biology Reports*, 2015  
Crossref 12 words — < 1%
- 
- 33 Vidushi Mahajan, Rekha Chouhan, Surinder Kitchlu, Kushal Bindu, Sushma Koul, Bikarma Singh, Yashbir S. Bedi, Sumit G. Gandhi. "Assessment of chemical and genetic variability in *Tanacetum gracile* accessions collected from cold desert of Western Himalaya", *3 Biotech*, 2018  
Crossref 12 words — < 1%
- 
- 34 I Damopolii, B Kurniadi. "Training students metacognitive skill using mobile learning", *Journal of Physics: Conference Series*, 2019  
Crossref 11 words — < 1%
- 
- 35 [biovalentia.mipa.unsri.ac.id](http://biovalentia.mipa.unsri.ac.id) Internet 11 words — < 1%
- 
- 36 [link.springer.com](http://link.springer.com) Internet 10 words — < 1%
- 
- 37 [webshop.coreit.se](http://webshop.coreit.se) Internet 10 words — < 1%
- 
- 38 [ejournal.undip.ac.id](http://ejournal.undip.ac.id) Internet



10 words — < 1%

39 [www.pecc.org](http://www.pecc.org)  
Internet

9 words — < 1%

40 J. A. Green. "Characterization of the Peptidase Activity of Recombinant Porcine Pregnancy-Associated Glycoprotein-2.", *Journal of Biochemistry*, 10/03/2008  
Crossref

9 words — < 1%

41 [www.ecmj.org.cn](http://www.ecmj.org.cn)  
Internet

9 words — < 1%

42 [www.jimmunol.org](http://www.jimmunol.org)  
Internet

9 words — < 1%

43 Suzette B. Lina, Masanori Okazaki, Dorothea Sonoko Kimura, Yoshiharu Yano et al. " Nitrogen uptake by sago palm ( Rottb.) in the early growth stages ", *Soil Science and Plant Nutrition*, 2009  
Crossref

9 words — < 1%

44 [doaj.org](http://doaj.org)  
Internet

8 words — < 1%

45 Denis V. Axenov-Gribanov, Irina V. Voytsekhovskaya, Yuriy V. Rebets, Bogdan T. Tokovenko et al. "Actinobacteria possessing antimicrobial and antioxidant activities isolated from the pollen of scots pine (*Pinus sylvestris*) grown on the Baikal shore", *Antonie van Leeuwenhoek*, 2016  
Crossref

8 words — < 1%

46 R. Lahaye. "From the Cover: DNA barcoding the floras of biodiversity hotspots", *Proceedings of the National Academy of Sciences*, 02/25/2008  
Crossref

7 words — < 1%

47 Rhishikesh S. Dhanve. "Coordinate action of exiguobacterial oxidoreductive enzymes in biodegradation of reactive yellow 84A dye",

7 words — < 1%

---

48 Haiyan Guan. "RNA Isolation and Real-Time Quantitative RT-PCR", *Methods in Molecular Biology*™, 2008 7 words — < 1%

Crossref

---

49 Jie Liu, Richard I. Milne, Michael Möller, Guang-Fu Zhu et al. " Integrating a comprehensive DNA barcode reference library with a global map of yeasts ( L.) for forensic identification ", *Molecular Ecology Resources*, 2018 6 words — < 1%

Crossref

---

50 Thomas J. Lyimo, Arjan Pol, Mike S. M. Jetten, Huub J. M. Op den Camp. " Diversity of methanogenic archaea in a mangrove sediment and isolation of a new strain ", *FEMS Microbiology Letters*, 2009 6 words — < 1%

Crossref

---

EXCLUDE QUOTES OFF  
EXCLUDE BIBLIOGRAPHY OFF

EXCLUDE MATCHES OFF