Caryologia
International Journal of Cytology, Cytosystematics and Cytogenetics

Firenze University Press
www.fupress.com/caryologia

OPEN ACCESS

Citation: S.S. Sobieh, M.H. Darwish (2020) The first molecular identification of Egyptian Miocene petrified dicot woods (Egyptians' dream becomes a reality). *Caryologia* 73(2): 3-13. doi: 10.13128/caryologia-750

Received: December 1, 2019

Accepted: March 13, 2020

Published: July 31, 2020

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Competing Interests: The Author(s) declare(s) no conflict of interest.

The first molecular identification of Egyptian Miocene petrified dicot woods (Egyptians' dream becomes a reality)

Shaimaa S. Sobieh*, Mona H. Darwish

Botany Department, Faculty of Women for Arts, Science and Education, Ain Shams University, Cairo, Egypt
E-mail: shimaa.sobieh@women.asu.edu.eg; mona.darwish@women.asu.edu.eg
*Corresponding author

Abstract. This is the first work on Egyptian ancient DNA (aDNA) from plant fossil remains. Two aDNA extracts from Miocene petrified dicot woods were successfully obtained, amplified, sequenced and recorded for the first time in the world using a DNA barcoding technique. Internal transcribed spacers (ITS) barcoding is a technique for delimiting and identifying specimens using standardized DNA regions. The two Miocene dicot woods: *Bombacoxylon owenii* (Malvaceae/Bombacoideae) and *Dalbergioxylon dicorynioides* (Leguminosae/Papilionoideae) were collected from the Wadi Natrun area in Egypt and were identified by palaeobotanists on the basis of wood anatomy. The molecular identification by ITS region of *Bombacoxylon owenii* did not match the wood taxonomic assignation. The molecular identification of *Bombacoxylon owenii* suggested that it is more related to the extant genus *Ceiba* rather than to the extant genus *Bombax*. In contrast, the molecular identification by ITS of *Dalbergioxylon dicorynioides* matched the identification of the palaeobotanist (related to extant genus *Dalbergia*). Therefore, we suggest that this region should be used as a starting point to identify several plant fossil remains and this work will be helpful in solving problems related to the identification of plant fossils.

Keywords: Egyptian petrified woods, aDNA, DNA barcoding, ITS.

INTRODUCTION

Over the past twenty years, several ancient DNA studies have been published, but none has targeted ancient Egyptian DNA. Initial studies on ancient plant DNA were published in the mid-eighties (Golberg *et al.* 1991). Rogers and Bendich (1985) reported the extraction of nanogram amounts of DNA from plant tissues ranging in age from 22000 to greater than 44600 years old. DNA from fossils facilitates the calibration of mutation rates among related taxa (Poinar *et al.* 1993).

Ancient DNA (aDNA) is the most important and informative biological component that scientists can find in archaeological areas for identification purposes. Ancient DNA analysis is used synergistically with other identifi-

cation methods, such as morphological and anatomical observations and microscopic analyses. DNA barcoding complements the microscopic techniques used in archaeobotany. DNA analysis can be solely used for the identification of specimens when the morphological and anatomical characteristics are absent (Hamalton 2016). Ancient DNA may be used to reconstruct proximal histories of species and populations. Studies involving the extraction, sequencing, and verification of fossil DNA demonstrate the existence of material that can be useful to both palaeontologists and evolutionary geneticists. This opens the possibility for coordinated studies of macro- and microevolutionary patterns that directly approach the relationship between morphological changes on the one hand and genetic changes on the other. In addition, molecular evolutionary studies attempt to reconstruct relationships between concurrent taxa by deducing ancestral states and the genetic distances between them (Golenberg 1994).

Ancient wood is found in high abundance, and samples are usually large enough to be analysed. For that reason, wood is an ideal target for ancient plant DNA studies (Kim *et al.* 2004). However, three problems obstruct the isolation and amplification of DNA from any aDNA specimens (Nasab *et al.* 2010). The first is the presence of contamination. The second is the existence of inhibitors of *Taq* DNA polymerase in ancient samples, while the third is the small quantity and low quality of DNA that is regained from dead wood (Kaestle and Horsburgh 2002) and this is due to degradation of DNA into small fragments in dead tissue (Deguilloux *et al.* 2002). Nevertheless, there are several reports of molecular analyses of aDNA from plants. Ancient DNA was extracted from 1600 year-old millet (*Panicum miliaceum*) by Gyulai *et al.* (2006) and in 1993, aDNA was extracted from 600- year-old maize cobs (Goloubinoff *et al.* 1993). Wagner *et al.* (2018) characterized the aDNA preserved in subfossil (nonpetrified) and archaeological waterlogged wood from the Holocene age (550–9,800 years ago).

DNA barcoding is used to identify unknown samples, in terms of a pre-existing classification (Tripathi *et al.* 2013) or to assess whether species should be combined or separated. It is also used to establish a shared community resource of DNA sequences that can be used for organismal identification and taxonomic clarification (Tripathi *et al.* 2013). The nuclear ribosomal internal transcribed spacer (ITS) region is indicated as a plant barcoding region (Hollingsworth *et al.* 2011).

Miocene fossils are believed to be the best-preserved fossils of Egypt (El-Saadawi *et al.* 2014). These fossils are chemically well preserved because of the low oxygen content and cold temperatures of the water in which they were deposited (Kim *et al.* 2004). DNA sequences can be obtained from Miocene-age plant remains and the success rate is increased through the use of improved methods of DNA extraction and the amplification of small segments of the fossil DNA (Kim *et al.* 2004).

El-Saadawi *et al.* (2014) reported that Egypt contains the second largest deposit of Miocene dicot woods in Africa (containing 23 taxa) after Ethiopia that contains 55 taxa. Seven petrified dicot woods were collected from the Wadi Natrun area in Egypt by Prof. Wagih El-Saadawi and Prof. Marwa Kamal El-Din (Botany Department, Faculty of Science, Ain Shams University). They identified only three of them, namely (*Bombacoxylon owenii* (Leguminosae/Papilionoideae), *Dalbergioxylon dicorynioides* (Fabaceae/Faboideae) and *Sapindoxylon stromeri* (Sapindaceae) based on the wood anatomy (El-Saadawi *et al.* 2014). Therefore, the main purpose of the present study was to extract and amplify aDNA from these Egyptian Miocene petrified dicot woods to provide a complete identification. DNA was successfully isolated from the wood samples of *Bombacoxylon owenii* and *Dalbergioxylon dicorynioides*. We used molecular techniques to confirm the wood anatomy identification of the two Egyptian wood fossils using DNA barcoding method. In addition, we validated the relationship between the plant fossil woods and the nearest living relative (NLR) based on molecular data acquired from the ITS barcode.

MATERIAL AND METHODS

Population sampling

Fossil samples

Seven of the good quality Egyptian ancient Miocene petrified dicot wood specimens (23.03 to 5.33 Ma. years ago) were used to extract the aDNA. Only two specimens (*Bombacoxylon owenii* (Bombacaceae) and *Dalbergioxylon dicorynioides* (Fabaceae) (Fig. 1a, b) were successfully identified to the genus level by the analysis of the ITS of the nuclear ribosomal DNA and the other five samples gave negative results. These Miocene petrified dicot woods were found in the Wadi Natrun area in Egypt and were previously identified by palaeobotanists (El-Saadawi *et al.* 2014; Kamal EL-Din *et al.* 2015) on the basis of the wood anatomy. The wood specimens were housed in the palaeobotanical collection of the Botany Department, Faculty of Science, Ain Shams University, Cairo-Egypt.

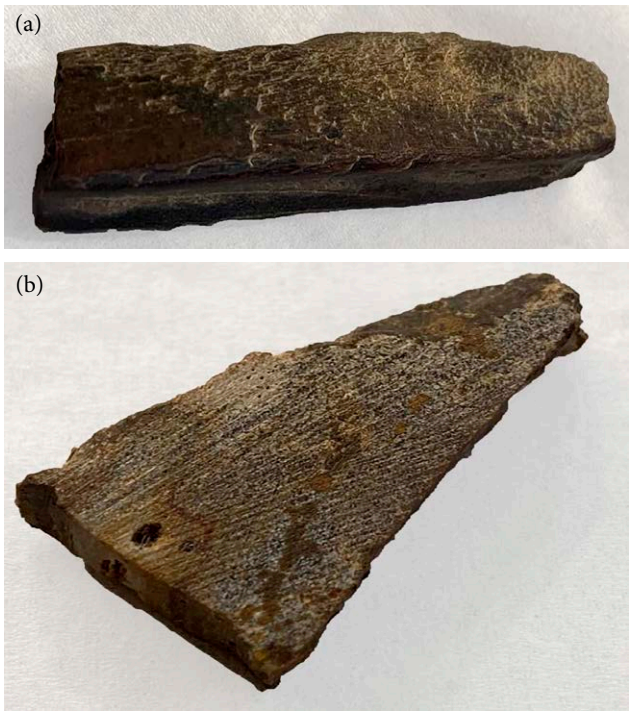


Fig. 1. Sections of *Bombacoxylon owenii* (a) and *Dalbergioxylon dicorynioides* (b).

Nearest Living Relative (NLR) samples

Living wood tissue from *Bombax ceiba* and *Dalbergia sissoo* was used in the present study as the NLR samples of *Bombacoxylon owenii* and *Dalbergioxylon dicorynioides,* respectively.

DNA Extraction, Amplification, and Sequencing

DNA Extraction

Total genomic DNA was extracted from the living woods and fossil wood using the cetyltrimethylammonium bromide method (CTAB) described by Doyle and Doyle (1987). As the extraction of aDNA in fossils is more difficult than the extraction of DNA from living wood several modifications were made. Layers of fossil surfaces were scraped with a sterile scalpel and were discarded under sterile conditions in order to remove any contamination, and mechanical disruption was used during the DNA extraction procedure. The original fossil samples were loose fragments scattered on the sand surface ranging between 10-50 cm in length and 5-20 cm in diameter (El-Saadawi *et al.*, 2014). They were very hard and difficult to break so they were cut by marble cutting machine into pieces and then those pieces were grinded mechanically into fine powder. The starting weight of the fossil sample was five times (5 g) higher than the living wood samples. Three volumes more of extraction buffer than the protocol suggested were added. Polyvinyl pyrrolidone was added to the lysis buffer. The quality of the DNA was estimated by checking the absorbance ratio at 260/280 nm using a Spectronic 21D spectrometer. The DNA samples from both the living and fossil samples were stored at -20°C for amplification and sequencing.

DNA Barcode

The internal transcribed spacers ITS of the nuclear ribosomal DNA was amplified using ITS4 and ITS5 primers with sequences of ITS4: TCC TCC GCT TAT TGA TAT GC and ITS5: GGA AGT AAA AGT CGT AAC AAG G (White *et al.* 1990). This region consists of a portion of 18S rDNA, ITS1, 5.8S rDNA, ITS2, and a portion of 28S rDNA (van Nues *et al.* 1994). The PCR mixture was a 25 µL solution containing 0.5 µL of dNTPs (10 mM), 0.5 µL of $MgCl_2$ (25 mM), 5 µL of 5× buffer, 1.25 µLof primer (10 pmol), 0.5 µL of template DNA (50 ng μL^{-1}), 0.1 µL of Taq polymerase (5 U μL^{-1}) and 17.15 µL of sterile ddH_2O. The amplification was carried out in a Techni TC-312 PCR, Stafford, UK system. The PCR cycles were programmed for the denaturation process for 4 min at 95°C (one cycle), followed by 30 cycles as follows: 94°C for 1 min; 53°C for 40 s;72°C for 1 min and finally one cycles extension of 72°C for 10 min and 4°C(infinite). The PCR products were run on 1.5% agarose gels, which were stained with ethidium bromide, at 120 V for 1 h. Successful PCR products were sent to LGC Genomics Sequencing (Germany) to be sequenced on a 3730xl DNA Analyzer (Applied Biosystems™/ Thermo Fisher Scientific).

Data analysis

The sequence identity was determined using the BlASTn algorithm available through the National Center for Biotechnology Information (NCBI) https://www.ncbi.nlm.nih.gov. The consensus sequences that showed a significant match with the earlier identified data in the NCBI were submitted to the Barcode of Life Data system (BOLD) v4 http://www.barcodinglife.org to identify each sequence sample to the genus and species level.

The new fossil sequences were submitted to the NCBI to be listed and recorded in the GenBank database. The G+C content of the four samples were calcu-

lated online using the CG content calculator website https://www.biologicscorp.com/tools/GCContent#.WrSk5OhubIU.

The multiple DNA sequences alignments (MSA) were performed using the Molecular Evolutionary Genetics Analysis version 6 (MEGA 6) (Tamura *et al.* 2013), while double sequence alignment using the CLUSTAL W algorithm was performed according Thompson *et al.* (1994).

The genetic distances were computed using MEGA 6.06 according to the Kimura-2-Parameter (K2P) model (Kimura 1980).

Phylogenetic reconstruction

The aligned DNA sequences by the CLUSTAL W algorithm of MEGA 6 were trimmed online using the trimming website: http://users-birc.au.dk/biopv/php/fabox/alignment_trimmer.php. The final aligned sequences were used to construct the phylogenetic trees. Sixteen species with their accession numbers (Table 1) were used to construct the phylogenetic tree for cf. *Ceiba* sp., and 36 species with their accession numbers (Table 2) were used to construct the phylogenetic tree for cf. *Dalbergia* sp. Moreover, the sequences of *Persea pseudocarolinensis* (accession number. AY337335) and *Persea palustris* (accession number. AY3377330) from GenBank, were chosen as outgroup to root the trees.

Table 1. The eighteen species used for constructing the phylogenetic tree for cf. *Ceiba* sp. with their accession numbers.

Accession number	Corresponding species
MG603734	cf. *Ceiba* sp.
KM453172	*Ceiba ventricosa*
KM453167	*Ceiba erianthos*
KM453170	*Ceiba pubiflora*
HQ658387	*Ceiba crispiflora*
KM453171	*Ceiba rubriflora*
HQ658388	*Ceiba speciosa*
KM488629	*Ceiba insignis*
KM453168	*Ceiba jasminodora*
DQ284851	*Ceiba pentandra*
HQ658389	*Ceiba schottii*
HQ658384	*Ceiba aesculifolia*
HQ658385	*Ceiba acuminata*
HQ658376	*Bombax buonopozens*
KM453163	*Bombax ceiba*
DQ826447	*Bombax malabaricum*
AY337335	*Persea pseudocarolinensis*
AY3377330	*Persea palustris*

The maximum likelihood (ML) analysis was applied to construct the phylogenetic trees. The ML analysis was constructed in MEGA 6 using the K2P model, with 1,000 bootstrap replicates. The codon positions were combined as 1st+2nd+3rd+noncoding. All positions containing gaps and missing data were eliminated. The tree was drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to assume the phylogenetic tree.

Table 2. The thirty-eight species used for constructing the phylogenetic tree for cf. *Dalbergia* sp. with their accession numbers.

Accession number	Corresponding species
MG450751	cf. *Dalbergia* sp.
KM521409	*Dalbergia sissoo*
KP092712	*Dalbergia balansae*
KM521377	*Dalbergia odorifera*
AB828610	*Dalbergia assamica*
KM521378	*Dalbergia hupeana*
KM521413	*Dalbergia stipulacea*
AB828616	*Dalbergia bintuluensis*
AB828639	*Dalbergia hostilis*
KM521372	*Dalbergia dyeriana*
AB828619	*Dalbergia bracteolata*
AF068140	*Dalbergia congestiflora*
AB828632	*Dalbergia frutescens*
AB828633	*Dalbergia glomerata*
AB828649	*Dalbergia melanocardium*
KM276143	*Dalbergia melanoxylon*
KM276125	*Dalbergia latifolia*
AB828614	*Dalbergia benthamii*
AB828622	*Dalbergia canescens*
AB828608	*Dalbergia arbutifolia*
AB828626	*Dalbergia cultrate*
AB828605	*Dalbergia acariiantha*
AB828618	*Dalbergia bojeri*
AB828613	*Dalbergia baronii*
AB828640	*Dalbergia humbertii*
AB828635	*Dalbergia greveana*
AB828604	*Dalbergia abrahamii*
KM521415	*Dalbergia trichocarpa*
AB828648	*Dalbergia martini*
FR854138	*Dalbergia tonkinensis*
AB828653	*Dalbergia parviflora*
HG313773	*Dalbergia entadoides*
KM521404	*Dalbergia rimosa*
HG004883	*Dalbergia cf. kingiana*
HG313775	*Dalbergia dialoides*
KM521414	*Dalbergia subcymosa*
AY337335	*Persea pseudocarolinensis*
AY3377330	*Persea palustris*

RESULTS AND DISCUSSION

DNA isolation

As far as is known, this is the first time that DNA from ancient Egyptian wood samples was extracted. The absorbance ratios (A260/280 nm) of the DNA extracts ranged between 1.81- 1.94 (Table 3), indicating good quality of the DNA from both fossil and living specimens. The concentrations of the DNA extracts were 175,285, 375 and 470 ng/ µL for *Dalbergioxylon dicorynioides*, *Bombacoxylon owenii, Bombax ceiba* and *Dalbergia sissoo*, respectively, as given in Table 3.

At the present time, publications of aDNA from plant fossils are still relatively infrequent; however, there are many aDNA publications from animals and humans which make up most samples in this field (Gugerli *et al.* 2005).

Helentjaris (1988) indicated that plant material from archaeological sites may also be amenable to DNA analysis. Many researchers have explored the possibility of isolating DNA from ancient wood samples. DNA has been extracted from samples of modern papyri (writing sheets made with strips from the stem of *Cyperus papyrus*) varying in age from 0-100 years BP and from ancient specimens from Egypt, with an age-span from 1,300-3,200 years BP. The results showed that the DNA half-life in papyri is approximately 19-24 years. This means that the last DNA fragments will vanish within no more than 532-672 years from the sheets being manufactured (Marota *et al.* 2002). In the case of ancient wood, the risk of contamination during handling and analysis is lower than with human or microbial DNA (Gilbert *et al.* 2005). Earlier works on fresh wood by Asif and Cannon (2005), Deguilloux *et al.* (2006) and studies of aDNA from ancient wood from Quercus and Cryptomeria by Deguilloux *et al.* (2002) suggested the possibility of DNA survival in ancient wood remains, which was confirmed by the current work.

Liepelt *et al.* (2006) reported that, it was possible to isolate DNA from wood as old as 1000 years. Depending on the mode of conservation and the climate at the excavation site, as well older samples could be isolated and analysed successfully (Deguilloux *et al.* 2006).

DNA Barcoding by ITS

The DNA barcoding affords an important step for the molecular identification of aDNA from petrified woods. The amplification of genomic DNA uses the universal primers for the ITS region.

Two of seven aDNA extracts from the dicot wood fossil samples (*Bombacoxylon owenii* and *Dalbergioxylon dicorynioides*) were successfully used to amplify the ITS region. The PCR and sequencing success rates for the fossil and living samples were 100% (Table 4). The genus and species identity results of the query sequences were then determined using the BLAST and BOLD databases to estimate the reliability of the genus identification. The results of both databases showed that ITS was 100% correctly identified at the genus level, while the success rates for species identification were 50 and 25% for BLAST and BOLD respectively (Table 5).

Many studies have compared the discriminatory power revealed by the ITS region in its entirety with

Table 3. Optical densities and concentrations of the DNA isolated from fossil and living specimens.

Plant name	Optical density		Ratio 260/280 nm	DNA concentration (ng/µL)
	260 nm	280 nm		
Bombacoxylon owenii	0.057	0.032	1.84	285
Bombax ceiba	0.075	0.041	1.82	375
Dalbergioxylon dicorynioides	0.035	0.018	1.94	175
Dalbergia sissoo	0.094	0.052	1.81	470

Table 4. Success rates of the amplification and sequencing.

Barcode locus	Number of tested samples (fossil and living samples)	No of samples amplified and percentage of PCR success	Number and percentage of PCR failure	Number and percentage of sequencing success
ITS	4	4 (100%)	0 (0%)	4 (100%)

Table 5. Identification efficiency of the barcode loci using BLAST and BOLD.

Barcode Locus	No. of samples identified	Family level using BLAST	Family level using BOLD	Genus level using BLAST	Genus level using BOLD	Species level using BLAST	Species level using BOLD
ITS	4	100%	100%	100%	100%	50%	25%

Table 6. Identification matches of the ITS sequences using the BLAST and BOLD Databases.

Sample identification	Plant order	Plant family	Plant subfamily	BLAST search match	BLAST similarity (%)	BOLD search match	BOLD similarity (%)
cf. *Ceiba* sp. *(Bombacoxylon owenii)*	Malvales	Malvaceae	Bombacoideae	cf. *Ceiba* sp.	100	*Ceiba pantandra*	90.83
Bombax ceiba	Malvales	Malvaceae	Bombacoideae	*Bombax ceiba*	99	*Bombax malabaricum*	99.14
cf. *Dalbergia* sp. *(Dalbergioxylon dicorynioides)*	Fabales	Fabaceae	Papilionoideae	cf. *Dalbergia* sp.	100	*Dalbergia odorifera*	87.94
Dalbergia sissoo	Fabales	Fabaceae	Papilionoideae	*Dalbergia sissoo*	99	*Dalbergia sissoo*	98.57

ITS2, proposing the use of ITS2 as an alternative barcode to the entire ITS region (Han et al 2013). ITS2 was previously used as a standard DNA barcode to identify medicinal plants by Chen *et al.* (2010) and a barcode to identify animals (Li et al 2010). The length of the ITS2 region is sufficiently short to allow for the easy amplification of even degraded DNA, and the ITS2 region has enough variability to distinguish even closely related species and has conserved regions for designing universal primers (Yao *et al.* 2010). Therefore, it could be used as a DNA barcode for plant fossils in further investigations.

In addition, all 4 raw nucleotide sequences were verified with the other available sequences in GenBank using the BLASTn algorithm. The sequences of the two living samples of *Bombax ceiba* and *Dalbergia sissoo* showed an identity ratio of 99% with *Bombax ceiba* (accession no. KM453163) and *Dalbergia sissoo* (accession no. AB828659), respectively (Table 6).

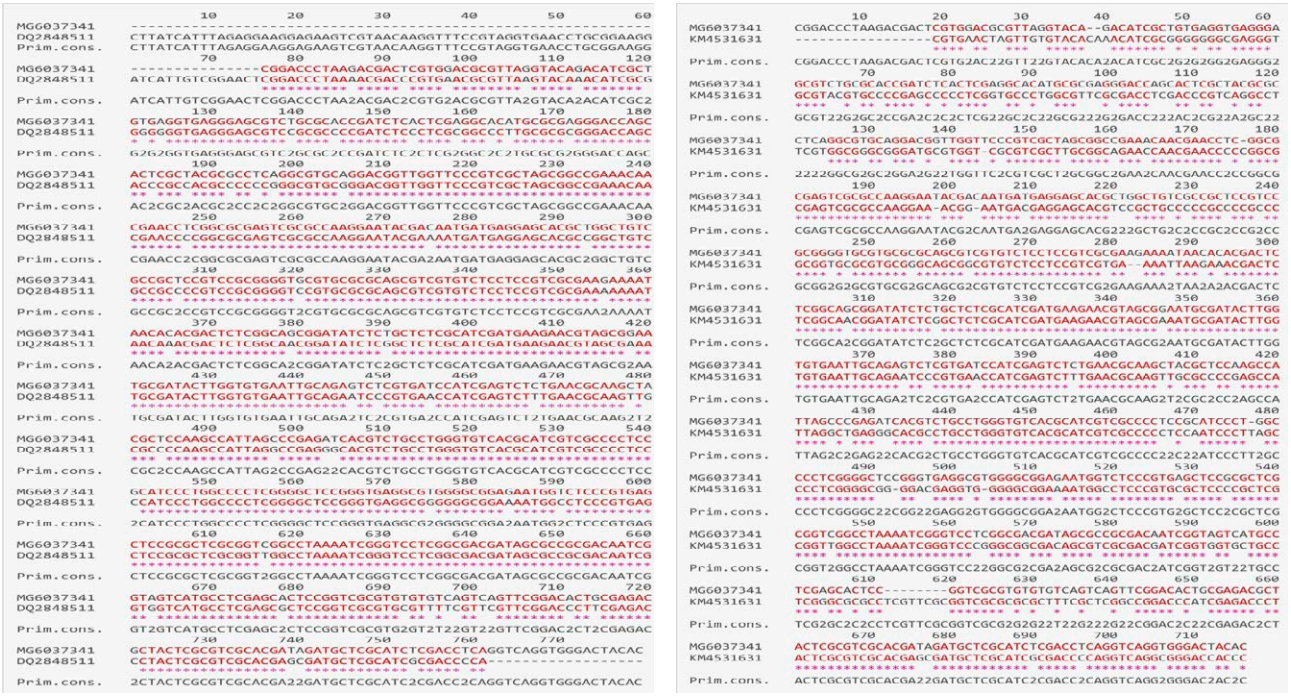
The identification of the fossil samples:

Based on the author's knowledge, thus far, there has been no published work on aDNA from petrified wood. Therefore, this is considered the first molecular identification of Egyptian plant fossil remains and of petrified wood (*Bombacoxylon owenii* and *Dalbergioxylon dicorynioides*) worldwide. Meanwhile, the authors hope that many other fields (anatomy and morphology) besides the molecular field will contribute to determining the relationship between living plants and their fossil remains.

Bombacoxylon owenii (cf. *Ceiba* sp. accession no.: MG603734)

The ITS sequence from the fossil specimen was amplified and produced a 704 bp fragment. The sequence was uploaded to the NCBI database and was documented, for the first time with accession number MG603734.

Bombacoxylon owenii was listed in the NCBI database as cf. *Ceiba* sp. because the GenBank policy is not to add fossil taxa to the taxonomy database, since it is a database of living or recently extinct organisms. *Bombacoxylon* is a fossil genus for woods with features characteristic of the Bombacoideae, not a whole plant. Moreover, the molecular identification revealed a close resemblance of the submitted sequence to *Ceiba pentandra* (the commercial kapok tree) rather than *Bombax* as was expected by Kamal El-Din *et al.* (2015). This identification is not surprising since the two living genera (*Bombax* and *Ceiba*) are grouped in the same subfamily Bombacoideae and have very few differences between them. Moreover, the wood anatomy of both genera reveals the high resemblance between them, and they can be only distinguished by a combination of macroscopic characteristics, which are the shape of the vessel-ray pit, the ray width, the sheath cells and mineral inclusion (Nordahlia *et al.,* 2016). The NLR of some fossil wood taxa might be wrong, *Bombacoxylon* shares characters with Sterculiaceae and Bombacaceae rather than only with *Bombax*, *Grewioxylon* with other members of the Malvaceae with tile cells, (e.g., *Craigia)* instead of only *Grewia* (Skala 2007). In addition, Wickens (2008) stated that it must



(a) (b)

Fig. 2. (a) Sequence alignment between cf. *Ceiba* sp. and *Ceiba pentandra* (accession no. DQ284851) using CLUSTAL W, Identity (*): 625 is 80.23 %. (b) Sequence alignment between cf. *Ceiba* sp. and *Bombax ceiba* (accession no. KM453163) using CLUSTAL W, Identity (*): 548 is 76.54 %.

not be assumed that the names of fossil wood necessarily represent species close to modern genera.

The sequence of cf. *Ceiba* sp. was compared with other available sequences in GenBank using the BLASTn algorithm. The results showed that the sequences belonged to the homologous sequences of the genus *Ceiba*. The sequence of cf. *Ceiba* sp. showed identities with several living *Ceiba* species rather than *Bombax*. The identity ratios among the *Ceiba* species indicated that the *Ceiba pentandra* ITS nucleotide sequence (accession no. DQ284851) was the nearest related ITS sequence for *Bombacoxylon owenii* (cf. *Ceiba* sp.).

The sequence of cf. *Ceiba* sp. was aligned with both *Bombax ceiba* (accession no. KM453163) and *Ceiba pentandra* (accession no. DQ284851) using CLUSTAL W (Thompson *et al.* 1994). The identity between the cf. *Ceiba* sp. ITS sequence and that of *Ceiba pentandra* was 625 (80.23%) (Fig. 2a), while the identity between the cf. *Ceiba* sp. ITS sequence and that of *Bombax ceiba* was 548 (76.54%) (Fig. 2b).

The final aligned sequences obtained by sequence trimming revealed that G+C content was obviously higher than of A+T content (Table 7). Genetic distances were calculated by the Kimura-2-Parameter (K2P) model (Kimura 1980).

Moreover, both *Bombacoxylon owenii* (cf. *Ceiba* sp.) and *Ceiba pentandra* shared similarities in the wood anatomy characteristics, with the presence of diffuse to semiring porous wood in both of them. *Bombacoxylon owenii* (cf. *Ceiba* sp.) and *Ceiba pentandra* contain solitary vessels and have radial multiples of 2 to 4 and medium to large vessels that are often filled with tyloses. The growth rings in both are distinct or absent and the vessel frequency is 5 to 20 per mm^2. The perforation plates are simple, and the intervessel pits are alternate. The vessel-ray parenchyma pits are like the intervessel pits and the fibres are nonseptate with thick-walls and diffuse to diffuse-in-aggregate axial parenchyma (Table 8) (inside wood 2013; Kamal EL-Din *et al.* 2015; Nordahlia *et al.* 2016).

Dalbergioxylon dicorynioides (cf. *Dalbergia* sp.accession no.: MG450751)

The ITS sequence (610 bp) was amplified and recorded in the NCBI database with GenBank accession no. MG450751. *Dalbergioxylon dicorynioides* was recorded as cf. *Dalbergia* sp. in the NCBI, since it is a database of living or recently extinct organisms. *Dalbergioxylon dicorynioides* is a fossil genus for woods not a whole plant.

Table 7. Sequence length and GC and AT content.

Sample name	Full length	G+C	G+C%	A+T	A+T%
cf. *Ceiba* sp.	704	221+219	62%	134+130	38%
Bombax ceiba	692	231+234	66%	113+114	34%
cf. *Dalbergia* sp.	610	185+194	61%	135+96	39%
Dalbergia sissoo	610	189+211	64%	127+83	36%

Table 8. Comparison of anatomical features between *Bombacoxylon owenii & Ceiba pentandra.*

Species Feature	*Bombacoxylon owenii*	*Ceiba pentandra* (L.)
Growth ring	Distinct	Distinct, indistinct or absent
Porosity	Diffuse to semiring-porous	Diffuse-porous
Perforation plates	Simple	Simple
Intervessel pits	Alternate	Alternate
Radial diameter	240 μm (220 to260)	350 to 800 μm
Vessels groupings	Solitary and in radial multiples of 2 to 4	Restricted to marginal rows
Tyloses	Common	Common
Vessel/mm^2	5 to 15(8)	5 to 20
Vessel element length μm	335 μm	350 to 800 μm
Axial Parenchyma	Diffuse, diffuse-in-aggregates, scanty, narrow vasicentric paratracheal and in narrow bands or lines	Diffuse, diffuse-in-aggregates, scanty, narrow vasicentric paratracheal and in narrow bands or lines
Rays	1 to 3 cells, seriate	Larger rays commonly 4 to 10 seriate
Fibers	Nonseptate with very thick walls	Nonseptate with thin- to thick-walled

The total sequence length of ITS in the *Dalbergia* genus ranged from 600 to 800 bp as reported by several records in the NCBI database for ITS in the *Dalbergia* genus.

The sequence was tested with other available sequences in GenBank using the BLASTn algorithm. The results showed that the sequences belonged to the homologous sequences of the genus *Dalbergia*. The sequence of cf. *Dalbergia* sp. showed identities with several living *Dalbergia* species, but when we compared the identity ratios among them we found that the *Dalbergia sissoo* ITS nucleotide sequences (accession no. AB828659.1) were the nearest ITS sequence for *Dalbergioxylon dicorynioides* (cf. *Dalbergia* sp.), with an identity ratio of 91%.

The final aligned sequences obtained by sequence trimming revealed that the G+C content was obviously higher than the A+T content (Table 7). Genetic distances for *Dalbergia* sequences alignment were calculated by the Kimura-2-Parameter (K2P) model (Kimura 1980).

The comparison of the wood anatomy characteristics of *Dalbergioxylon dicorynioides* (cf. *Dalbergia* sp.) with those of living *Dalbergia* species revealed that *Dalbergia sissoo* was most closely related to *Dalbergioxylon dicorynioides* (Table 9) because both contained diffuse-porous wood, solitary vessels and radial multiples of 2 to 3, indistinct or absent growth rings, exclusively simple perforation plates, alternate and vestured intervessel pits, vessel-ray pits similar to intervessel pits in size and shape throughout the ray cell, combinations of aliform, confluent and irregular banded (1 to 4 cells wide) axial parenchyma, 1-3 seriate rays up to 20 cells high, and thick-walled non-septate fibers (inside wood 2013; El-Saadawi *et al.* 2014).

Phylogenetic analysis

The phylogenetic analyses were conducted in MEGA6 (Thompson *et al.* 1994) and the phylogenetic trees were inferred with the ML based on the Kimura model (Kimura 1980). Nowadays, several programs can be used to construct maximum likelihood phylogenetic tree. The fastest ML-based phylogenetic programs that differ in implementations of rearrangement algorithms are PhyML (Guindon *et al.* 2010) and RAxML/ExaML (Stamatakis 2014).

The topologies of the phylogenetic trees were evaluated using the bootstrap resampling method of Felsenstein (1985) with 1000 replicates. The analysis involved

Table 9. Comparison of anatomical features between *Dalbergioxylon dicorynioides & Dalbergia sissoo*.

Species Feature	*Dalbergioxlon dicorynioides*	*Dalbergia sissoo*
Growth ring	Absent.	Distinct, indistinct or absent
Porosity	Diffuse- porous	Diffuse- porous
Perforation plates	Simple	Simple
Intervessel pits	Alternate	Alternate
Tangential diameter µm	170 µm (range 100 to 210 µm)	100 to 200
Vessels groupings	Solitary and in radial multiples 2 to 3	Solitary or grouped in radial multiples of 2 to3 cells.
Vessel groupings /mm^2	8/mm^2 (range 5 to13/ mm^2)	5 to 20
Vessel element length µm	330 µm (range 280 to 410 µm)	<= 350
Axial Parenchyma	Aliform, confluent and irregular banded (1 to 4 celled wide)	Aliform, confluent and irregular banded, 4 (3to 4) cells per parenchyma strand
Rays	1to 3 seriate	1 to 3 cells
Fibers	Thick-walled, nonseptate	Very thick-walled, nonseptate

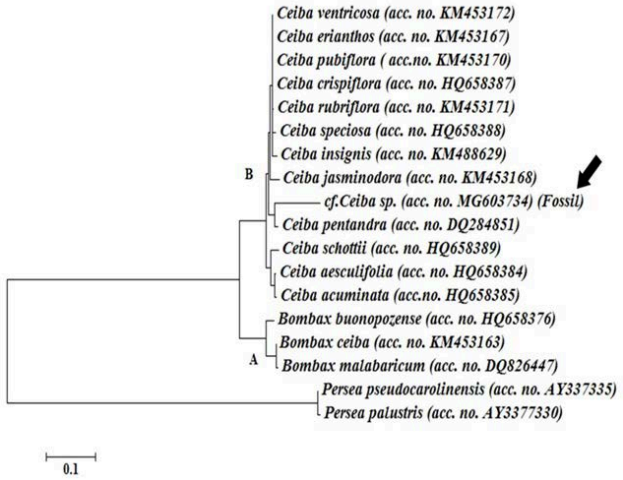


Fig. 3. Maxim- Likelihood (ML) cladogram showing the relationships of the ITS gene from cf. *Ceiba* sp. in relation to its relatives. All analyses were performed with 1000 bootstrap replicates (arrow: fossil specimens, acc. no.: accession number).

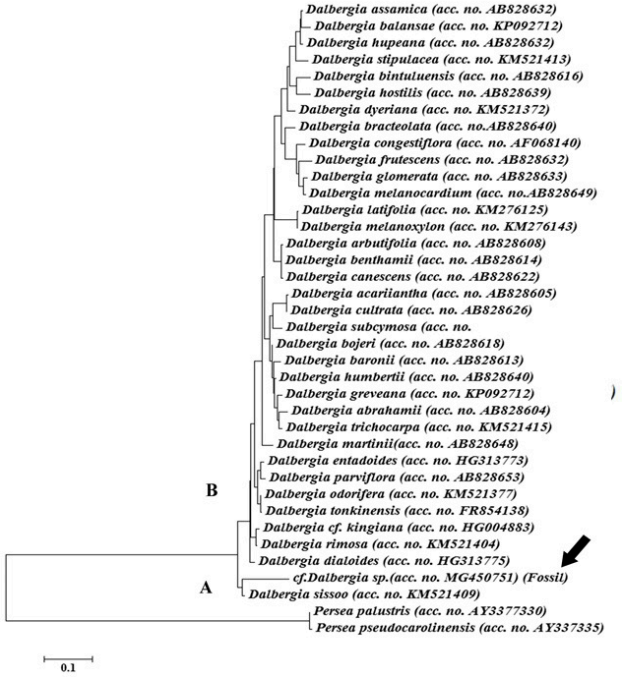


Fig. 4. Maxim- Likelihood (ML) cladogram showing the relationships of the ITS gene from cf. *Dalbergia* sp. in relation to its relatives. All analyses were performed with 1000 bootstrap replicates (arrow: fossil specimens, acc. no.: accession number).

18 nucleotide sequences (cf. *Ceiba* sp., 12 species of *Ceiba* and 3 species of *Bombax* which were downloaded from the NCBI database), and *Persea pseudocarolinensis* and *Persea palustris* were used as outgroups. There was a total of 1374 positions in the final dataset, and the ambiguous positions were completely eliminated for each sequence pair.

The ML tree was divided into two clades, namely A and B. Clade A included *Bombax* members, while clade B included the *Ceiba* species in addition to cf. *Ceiba* sp. (*Bombacoxylon owenii*). Both cf. *Ceiba* sp. and *Ceiba pentandra* were on the same branch. Therefore, the phylogenetic tree showed that *Bombacoxylon owenii* (cf. *Ceiba* sp.) was very similar to the *Ceiba* genus, which previously was thought to resemble the *Bombax* genus (Kamal EL-Din *et al.* 2015) (Fig. 3).

In the ML tree, all the *Dalbergia* species were divided into two clades, namely clade A and clade B (Fig. 4). Clade A includes cf. *Dalbergia* sp. and *Dalbergia sissoo.* The second group (clade B) was subdivided into many subclades that contained the other species of *Dalbergia*. Therefore, the present work matches the palaeobotanist

assumption that there is a close relationship between *Dalbergioxylon dicorynioides* (cf. *Dalbergia* sp.) and *Dalbergia sissoo*.

CONCLUSION

The DNA barcoding dataset in the present study provides an important first step towards establishing an effective molecular tool for the identification of aDNA from petrified woods. We hope that these results will encourage reliable aDNA studies of other petrified woods. The further studies of ancient wood DNA from the abundant store of fossil plant remains will rely on this study and by the intensive works of researchers from different fields, and these findings could provide a powerful tool to increase world knowledge about the history of forests, plant evolution and historical biogeography.

AUTHOR CONTRIBUTIONS

Both authors suggested the point of the work and Dr. Shaimaa S. Sobieh planned the experimental design to achieve this point. Both authors supplied the financial support for the work. Prof. Mona Darwish shared other palaeobotanists in the identification of dicot woods (see El-Saadawi *et al.* 2014). The experimental part was done by Dr/Shaimaa S. Sobieh. The writing of the manuscript was done by both authors.

ACKNOWLEDGEMENTS

The authors would like to thank Prof. Wagih El-Saadawi and Prof. Marwa Kamal El-Din (Profs. of Palaeobotany, Botany Department, Faculty of Science, Ain Shams University) for supplying them the fossil specimens. In addition, they thank Fatma Abdel Naby Mursi and Aya Abdel Gawad (MSc. students, Botany Department, Faculty of Women for Art, Science and Education, Ain Shams University) for helping them in the extraction of aDNA. Finally, the authors would like to thank Dr Enas Hamdy Ghallab (Lecturer of Medical Entomology, Entomology Department, Faculty of Science, Ain Shams University) and Mohamed Emad El-din Elsaid (Biotechnology Bachelors, Misr University for Science and technology) for helping the authors in understand many points in the bioinformatics programs.

REFERENCES

Asif MJ, Cannon CH. 2005. DNA extraction from processed wood: a case study for the identification of an endangered timber species (*Gonstylus bancanus*). Plant Mol Biol Rep. 23:1–8.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, et al .2010. Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. PLoS ONE 5: e8613. doi.org/10.1371/journal.pone.0008613

Deguilloux MF, Bertel L, Celant A, Pemonge MH, Sadori L, Magri D, Petit RJ. 2006. Genetic analysis of archaeological wood remains: first results and prospects, J Archaeol Sci. 33: 1216–1227.

Deguilloux MF, Petit MH, Pemonge RJ. 2002. Novel perspectives in wood certification and forensics: dry wood as a source of DNA. Proc R Soc London. 269: 1039-1046.

Doyle JJ, Doyle JL. 1987. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytoch Bull. 19: 11-15.

El-Saadawi W, Kamal El-Din MM, Darwish MH, Osman R. 2014. African Miocene dicot woods with two new records for this epoch from Egypt. Taeckholmia. 34:1-2.

Felsenstein J. 1985. Confidence limits on phylogenies: An approach using the bootstrap. Evol. 39: 783-791.

Gilbert MTP, Bandelt HJ, Hofreiter M, Barnes I. 2005. Assessing ancient DNA studies. Tren Ecol Evol. 20: 541–544.

Golberg EM, Brown TA, Bada JL, Westbroek P, Bishop MJ, Dover GA. 1991. Amplification and analysis of Miocene plant fossil DNA. Phil Trans R Soc Lond B. 333: 419-427.

Golenberg EM. 1994. Fossil samples DNA from plant compression fossils. In: Herrmann B, Hummel S, (eds) Ancient DNA recovery and analysis of genetic material from paleontological, archaeological, museum, medical, and forensic specimens. New York: Springer-Verlag Inc pp 237-256.

Goloubinoff P, Pääbo S, Wilson A. 1993. Evolution of Maize Inferred from Sequence Diversity of an *adh*2 Gene Segment from Archaeological Specimens. Proc Natl Acad Sci USA. 90:1997-2001.

Gugerli F, Parducci L, Petit RJ. 2005. Ancient plant DNA: review and prospects, New Phytologist. 166: 409–418.

Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. 2010. New algorithms and methods to estimate maximumlikelihood phylogenies: assessing the performance of PhyML 3.0. Syst Biol. 59(3): 307–321.

Gyulai G, Humphreys M, Lagler R, Szabó Z, Tóth Z, Bittsánszky A, Gyulai F, Heszky L. 2006. Seed

remains of common millet from the 4th (Mongolia) and 15th (Hungary) Centuries: AFLP, SSR and mtDNA sequence recoveries. Seed Sci Res. 16: 179-191.

Han J, Zhu Y, Chen X, Liao B, Yao H, Song J, Chen S, *et al.* 2013. The short ITS2 sequence serves as an efficient taxonomic sequence tag in comparison with the full-length ITS. Biomed Res Int. 2013:741476. doi.org/10.1155/2013/741476

Hamalton T. 2016. DNA from ancient wood. Van Sangyan. 3: 27-30.

Helentjaris T. 1988. Maize Genet Coop. News Lett. 62: 104-105.

Hollingsworth PM, Graham SW, Little DP. 2011. Choosing and using a plant DNA barcode. PLoS ONE. 6: e19254. doi.org/10.1371/journal.pone.0019254

Kaestle AF, Horsburgh KA. 2002. Ancient DNA in anthropology: methods, applications, and ethics. Am J Phys Anthropol. 35: 92-130.

Kamal EL-Din MM, Darwish MH, EL-Saadawi W. 2015. Novelties on Miocene woods from Egypt with a summary on African fossil woods of Fabaceae, Malvaceae and Dipterocarpaceae. Palaeontographica Abt B. 292:173-199.

Kim S, Soltis DE, Soltis PS, Suh Y. 2004. DNA sequences from Miocene fossils: an *ndhF s*equence of *Magnolia latahensis* (Magnoliaceae) and an *rbcl* sequence of *Persea pseudocarolinensis* (Lauraceae). Am J Bot . 91: 615–620.

Kimura M. 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. J Mol Evol. 16: 111-120.

Liepelt S, Sperisen C, Deguilloux MF, Petit RJ, Kissling R, Spencer M, De Beaulieu J, Taberlet P, Gielly l, Ziegenhagen B. 2006. Authenticated DNA from Ancient Wood Remains. Ann Bot. 98: 1107–1111.

Li YW, Zhou X, Feng G, Hu HY, Niu LM, Hebert PD, *et al.* 2010. COI and ITS2 sequences delimit species, reveal cryptic taxa and host specificity of fig-associated Sycophila (*Hymenoptera, Eurytomidae*). Mol Ecol Resour 10: 31–40.

Marota I, Basile C, Ubaldi M, Rollo F. 2002. DNA decay rate in Papyri and human remains from Egyptian archaeological sites. Am j phys anthropol. 117: 310–318.

Nasab HM, Mardi M, Talaee H, Nashli HF, Pirseyedi SM, Nobari AH, Mowla SJ. 2010. Molecular analysis of ancient DNA extracted from 3250-3450 year-old plant seeds excavated from Tepe Sagz Abad in Iran. J Agr Sci Tech. 12: 459-470.

Nordahlia AS, Noraini T, Chung RCK, Lim SC, Nadiah I, Azahana NA, Solihani NS. 2016. Comparative wood anatomy of three *Bombax* species and *Ceiba pentandra* (Malvaceae: Bombacoideae) in Malaysia. Mal Nat J. 68: 203-216.

Poinar HN, Cano RJ, Poinar GO. 1993. DNA from an extinct plant. Nature 363: 677.

Rogers SO, Bendich AJ. 1985. Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. Pl molec Biol. 5: 69-76.

Sakala, J. 2007. The potential of fossil angiosperm wood to reconstruct the palaeoclimate in the Tertiary of Central Europe (Czech Republic, Germany). Acta Palaeobotanica.. 47: 127–133 (2007).

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 30(9):1312–1313.

Tamura K, Stecher G, Peterson D, Filipski A, Kumar S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol Biol Evol. 30: 2725-2729.

Thompson JD, Higgins DG, Gibson TG. 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, positions-specific gap penalties and weight matrix choice. Nucl Aci Res. 22: 4673-4680.

Tripathi AM, Tyagi A, Kumar A, Singh A, Singh S, Chaudhary LB, Roy S. 2013. The internal transcribed spacer (ITS) region and trnhHpsbA are suitable candidate loci for DNA barcoding of tropical tree species of India. PloS ONE. 8: e57934. doi.org/10.1371/journal.pone.0057934.

van Nues R W, Rientjes J M J, van der Sande C A F M., Zerp S F, Sluiter C, Venema J, Planta R J, Raue´ HA. 1994. Separate structural elements within internal transcribed spacer 1 of Saccharomyces cerevisiae precursor ribosomal RNA direct the formation of 17S and 26S rRNA. Nucl Aci Res. 22: 912–919.

Wagner S, Lagane F, Seguin-Orlando A, *et al.* 2018 High-Throughput DNA sequencing of ancient wood. Mol Ecol. 27: 1138-1154.

White TJ, Bruns TD, Lee SB, Taylor JW. 1990. Amplification and direct sequencing of fungal ribosomal RNAgenes for phylogenetics. In: Innis MA, Gelfard H, Sninsky JS, WhiteTJ (eds) PCR-protocols and applications. A laboratory manual. New York: Academic Press, pp 315–322.

Wickens GE. 2008. The Baobabs: Pachycauls of Africa, Madagascar and Australia Springer Science & Business Media

Wood data base available at inside wood home page. 2013. Online search of fossil and modern.

Yao H, Song J, Liu C, Luo K, Han J, Li Y, *et al.* 2010. Use of IRS2 region as the universal DNA barcode for plants and animals. PLoS ONE 5: e13102. doi.org/10.1371/journal.pone.0013102