Online: ISSN 2771-5868
Print: ISSN 2771-585X

Original Research Article

Molecular Authentication of Indian Sandalwood (*Santalum album*) and Its Commercial Substitute *Gmelina arborea* Using trnH–psbA DNA Barcoding

Sabina Shirsekar* and Mahavir Gosavi

Department of Botany, S. I. E. S. College of Arts, Science and Commerce, Sion (West), Mumbai – 400 022

***Corresponding author:** sabina.shirsekar@gmail.com, Contact No.: 8779488952

Received: September 3, 2025; rvised: February 9, 2026; accepted: February 16, 2026.

https://doi.org/10.5281/zenodo.18800571

Published Online : February, 2026

Abstract: The demand for *Santalum album* (Indian sandalwood) has risen dramatically in recent decades, resulting in widespread and unsustainable harvesting. Due to its scarcity and high economic value, many manufacturers and traders substitute other timber species, such as *Gmelina arborea*, which may visually resemble sandalwood but differ greatly in aroma, density, and chemical composition. Such substitutions reduce product quality and mislead consumers and also undermine conservation initiatives aimed at protecting authentic sandalwood resources.

Traditional identification methods including morphological observation, anatomical characterization, and chemical profiling are effective for raw wood but lose reliability when the material is processed into powders or carvings, where diagnostic features are masked or destroyed. To overcome these limitations, molecular tools, particularly DNA fingerprinting and DNA barcoding, have emerged as powerful alternatives for species authentication.

This study used the **trnH–psbA intergenic spacer** to confirm the genetic identity of *S. album* and to distinguish it from *Gmelina arborea*, a common substitute in commercial markets. DNA sequence analysis revealed clear genetic separation between the two species, despite their similar appearance. The *S. album* sequences obtained showed a 100% match with reference accessions in GenBank, validating the accuracy of the identification and the reliability of the marker. Similarly, *G. arborea* sample grouped closely with published *Gmelina* sequences in the phylogenetic tree, confirming their correct classification.

The findings highlight the importance of integrating DNA-based identification into standard quality assurance protocols to protect endangered *S. album* populations and to strengthen traceability and transparency in the global sandalwood market.
Keywords: DNA barcoding, *Gmelina arborea,* Sandalwood, *Santalum album,* Sustainable trade, trnH–psbA intergenic spacer, White wood

Introduction:

Santalum album, commonly known as Indian sandalwood, is a highly valued tree species due to its aromatic heartwood, which is used in traditional medicine, perfumery, cosmetics, and religious rituals (Khan, Ikram and Faisal, 2021). *Santalum album* (Swetachandan or Srigandha in Sanskrit), belongs to family Santalaceae. It is an evergreen tree with drooping branches, where adult tree reaching the height of 20m and attaining girth of up to 2.4m. Young trees exhibit smooth, reddish-brown bark, whereas in older trees the bark becomes hard, rough, and deeply fissured, with colours ranging from brown and reddish to dark grey or nearly black (Rai and Sarma, 1990; Sharma and Kothari, 2021). The heartwood is

highly fragrant and golden to dark brown in colour while the sapwood is pale, white and notably odourless (Singh, Yadav and Anupam, 2022). Sandalwood is used traditionally as a **diuretic, expectorant, genitourinary disinfectant, cooling agent, sedative**, and **astringent,** offering a wide spectrum of health-supporting benefits (Kumar, Anjum and Tripathi, 2015).

Due to its high demand and decreasing natural population, unsustainable harvesting practices have critically imperilled sandalwood population across the globe (Bunney *et al.,* 2023). Various plant species are often marketed as "sandalwood" because their wood superficially resembles genuine sandalwood due to their outer look, whitish or light-yellow colour and soft texture. Such woods, especially from so-called "white wood trees" that are used for carving idols and art materials, can be immersed in or impregnated with Sandal wood oil or fragrances and dried. It can then easily be "substituted and marketed as Sandal Wood", (Fig. 1).

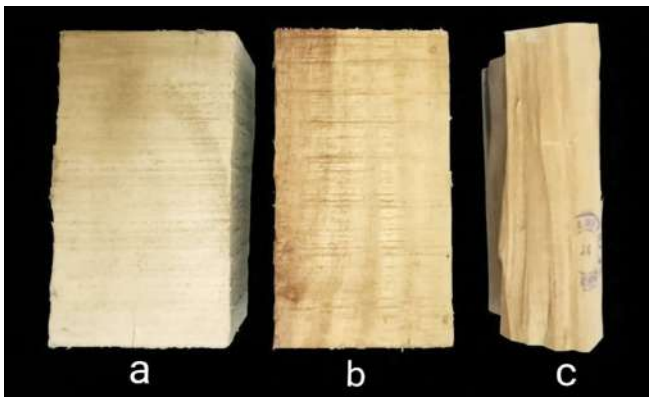

Fig. 1. Comparative visual observation of wood samples: (a) *White wood (Gmelina arborea)*, (b) *White wood soaked in sandalwood oil*, and (c) *Authentic sandalwood (Santalum album).*

Unscrupulous traders often mislabel or substitute inferior species, exploiting the high market demand for genuine *Santalum album*. Not only are these misleading, but they also take advantage of its high economic value (Shukla, Shashikala and Sujatha, 2021). One such wood, *Gmelina arborea* is often substituted in artisanal wood carving, especially in markets where authenticity is difficult to verify. This substitution misleads consumers and undermines conservation efforts.

Gmelina arborea (Gambhari/ Shivan), a tall deciduous tree of Lamiaceae family reaching up to height of 30 m or more and a diameter upto 4.5m. Its pale grey bark turns corky with maturity. It is used in Ayurveda and other traditional systems for various therapeutic purposes, including treatment for fever, headaches, abdominal pain and more. The wood is light yellow, coarse-grained, odourless with no essential oil. It contains lignans, alkaloids, flavonoids, phenolics (Waghchaure *et al*., 2022).

Morphological analysis alone is insufficient to detect substitution fraud, particularly when the wood is carved or processed. The key challenge lies in the inability of visual inspection or anatomical analysis to distinguish between processed sandalwood and similar species like *Gmelina arborea*. The latter is increasingly used due to its abundance, ease of carving and affordability (Sharma, Sharma and Lipoktila, 2013). However, from a biochemical and aromatic standpoint as well as chemistry and pharmacological benefits of sandal wood, it is an inferior alternative. DNA fingerprinting offers a scientific solution to this issue. DNA-based authentication is a critical step in ensuring traceability and maintaining the value chain integrity of sandalwood products. DNA barcoding, a method based on sequencing specific genomic regions provides a robust tool for species authentication (Sang, Crawford and Stuessy, 1997; Tate and Simpson, 2003).

Materials and methods

Sample Collection

Santalum album sample was procured from Royal Indian Sandalwood Private Limited, Mathikere, Bengaluru, Karnataka and authenticated by Authentic Botanicals Consultancy.
Wood blocks of white wood, (Fig. 2) used for wood carvings was collected from Kerala.

Principle of DNA Barcoding

DNA barcoding identifies species using short, standardized genomic regions. For plants, chloroplast markers such as rbcL, matK and the trnH–psbA intergenic spacer, together with nuclear ITS, are widely accepted core barcodes (Kress *et al*.,

Fig. 2. Blocks of *White wood (Gmelina arborea)* used for wood carvings.

2005; CBOL Plant Working Group, 2009). Among these, trnH–psbA shows high interspecific variability and short amplicon length, making it particularly suitable for degraded or processed materials such as wood (Sang *et al.*, 1997; Tate & Simpson, 2003; Nithaniyal *et al.*, 2014). In the present study, trnH–psbA was selected due to its proven discriminatory power in timber and medicinal plant authentication.

DNA Isolation and purification

The DNA of the plant materials were isolated by using DNeasy Plant Mini Kit. Using a mortar and pestle the dry samples were finely ground, then transferred into centrifuge tubes. The ground sample was incubated in lysis buffer containing RNase A at 65 °C to release nucleic acids. The lysate was clarified by centrifugation and subsequently passed through silica-based spin columns for purification. DNA was washed using proprietary buffers supplied in the kit and finally eluted in Buffer AE. The purified DNA was quantified and stored at "20 °C for subsequent PCR analysis (Qiagen, 2016).

PCR Amplification of trnH–psbA

PCR amplification was carried out in 20 µL reactions containing 1× PCR buffer (2 mM MgCl,), 200 µM of each dNTP, 0.5 µM forward and reverse primers, 1 U Taq DNA polymerase and ~5 ng template DNA.

Universal primers used for trnH–psbA: psbA3_f: 52 -GTTATGCATGAACGTAATGCTC-32 and truHf_05: 52 - CGCGCATGGTGGATTCACAATCC-32 (as described by Sang *et al.*, 1997; Tate & Simpson, 2003).

Thermal cycling conditions consisted of initial denaturation at 94 °C (3 min), followed by 35 cycles of denaturation at 94 °C (30 s), annealing at 55 °C (30 s), extension at 72 °C (60 s) and a final extension at 72 °C (7 min). PCR provides exponential amplification of the target barcode region, enabling detection even from trace or degraded DNA typical of wood samples (Kress *et al.*, 2005).

Agarose Gel Electrophoresis and sequencing

PCR products were resolved on 1.8% agarose gels prepared in 1× TAE buffer and stained with ethidium bromide. Electrophoresis was performed at 70 V until clear band separation was achieved. Amplicon sizes were estimated using a 100 bp–3 kb DNA ladder. Agarose gel electrophoresis allows verification of successful amplification, assessment of fragment size (~300–450 bp for trnH–psbA) and detection of nonspecific products prior to sequencing. Purified PCR products were subjected to bidirectional Sanger sequencing. Raw chromatograms were edited using BioEdit v7.2.6 (Hall, 1999). Forward and reverse reads were aligned using CLUSTALW and merged into consensus contigs.

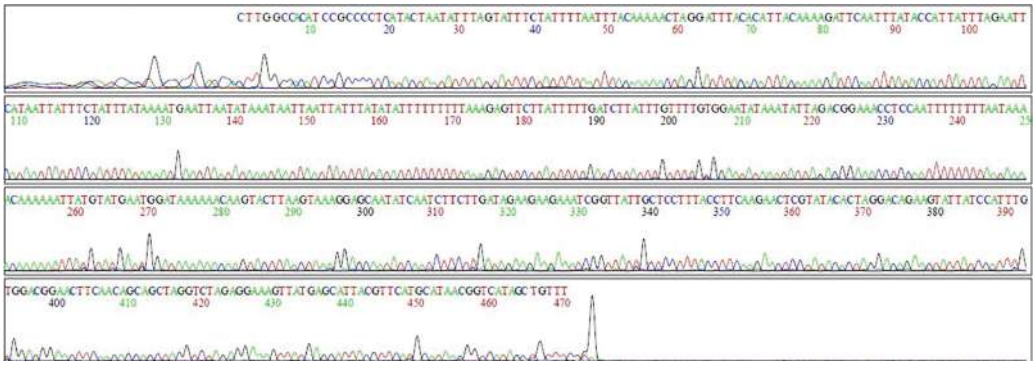

Fig. 3. Final contig sequence chromatogram for the trnH-psbA region of *Santalum album.*

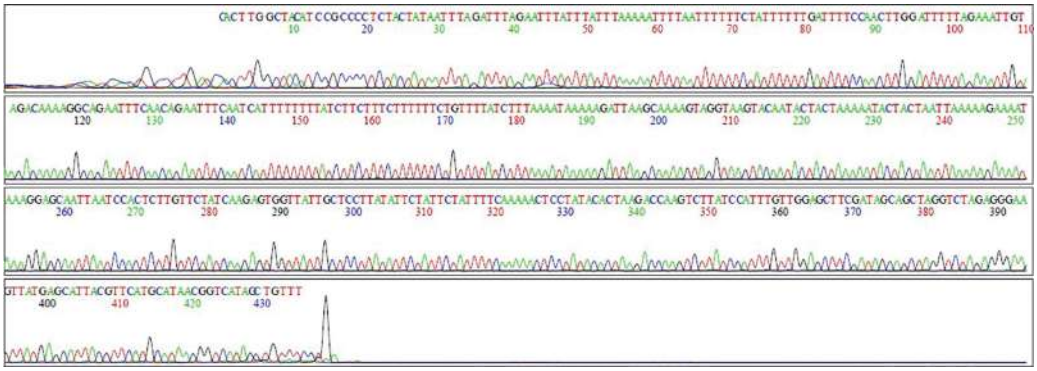

Fig. 4. Final contig sequence chromatogram for the trnH-psbA region of *Gmelina arborea.*

Sequences were compared against NCBI GenBank using BLASTn for preliminary identification. Genetic distances were calculated using the Kimura 2-parameter model (Kimura, 1980). Phylogenetic relationships were inferred in MEGA11 using Neighbor-Joining with 1000 bootstrap replicates (Tamura *et al*., 2021).

Results

PCR amplification of the trnH–psbA region produced clear single bands of expected size in both wood samples, confirming successful recovery of chloroplast DNA from processed material.

The final edited contig sequences generated in this study are presented below:

In this study, *trnH–psbA* intergenic spacer sequences were generated for two wood samples: (1) *Santalum album*, and (2) *Gmelina arborea*. These sequences were then aligned with corresponding references from the CMPR-AVS database to confirm species identity. The final edited contigs for *S. album* (Fig. 3) showed complete sequence correspondence with the database reference entries, confirming both the authenticity of the specimen and the accuracy of the sequencing process. Similarly, the *G. arborea* samples (Fig. 4) displayed identical nucleotide patterns, reflecting genetic uniformity within this species for the *trnH–psbA* locus. Their sequences also matched perfectly with the *G. arborea* reference record in the CMPR-AVS database, validating the correctness of their taxonomic assignment.

Further verification through BLASTn analysis revealed that the *S. album* sequence showed 100% identity with authenticated *Santalum album* accessions (Table 2), while *G. arborea* samples corresponded most closely with *G. arborea* entries (Table 3). Although both woods may appear similar in colour and texture, their trnH–psbA sequences exhibited multiple nucleotide substitutions and indels, producing clear genetic divergence. This molecular separation demonstrates

Table 1. List of primers used for amplification of different loci and their PCR conditions.

S. No	DNA region	Primers Name	Sequence (5'-3')	Reference	PCR reaction conditions
1	*trnH-psbA*	psbA3_f	GTTATGCATGAACGTAATGCTC	Sang *et al.* 1997 Tate & Simpson, 2003	92 °C for 4 min
		truHf_05	CGCGCATGGTGGATTCACAATCC		94 °C for 1 min. } 35x
					52 °C for 1 min. } 35x
					64 °C for l min. } 35x
					64 °C for 8 min

Note: All the above primers had additional M13F and M13R sequences at the forward and reverse primers respectively at 5'ends which was used for sequencing.

Table 2. Blastn results of *Santalum album* (trnH-psbA) sequence in the NCBI database.

Scientific Name	Max. Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Santalum album	830	830	100%	0	100	144034	MK638825 1
Santalum album	830	830	100%	0	100	144034	MW464915.1
Santalum album	824	824	100%	0	99.78	144101	NC 048953.1
Santalum album	824	824	100%	0	99.78	144101	MN106256.1
Santalum album	824	824	100%	0	99.78	144101	MW464922.1
Santalum album	824	824	100%	0	99.78	449	LC435405 1
Santalum pyrularum	808	808	100%	0	99.11	143895	NC 077485 1
Santalum lanceolatum	806	806	100%	0	99.11	143801	NC 077483 1
Santalum boninense	802	802	100%	0	98.89	144263	NC 077481.1
Santalum boninense	802	802	100%	0	98.89	144263	LC522526.1

Table 3. Blastn results of *Gmelina arborea* 1 & *Gmelina arborea* 2 (trnH-psbA) sequence in the NCBI database.

Scientific Name	Max. Score	Total Score	Query Cover	E value	Per. Ident	Acc. Len	Accession
Gmelina villosa	678	678	100%	0	96.84	437	JQ618436.1
Gmelina arborea	667	667	100%	0	96.36	154369	MT473761 1
Gmelina arborea	667	667	100%	0	96.36	407	OQ148437 1
Gmelina arborea	667	667	100%	0	96.36	407	OQ148440 1
Gmelina arborea	667	667	100%	0	96.36	407	OQ148439 1
Gmelina arborea	667	667	100%	0	96.36	407	MW048000 1
Gmelina arborea	667	667	100%	0	96.36	407	OQ148438 1
Magnoliophyta environ. sample	667	667	100%	0	96.36	400	LN899135 1
Gmelina philippensis	614	707	100%	4e-171	94.2	154408	MT473763 1
Magnoliophyta environ sample	584	584	89%	3e-162	95.91	355	LN899050 1

that *S. album* and *G. arborea* are clearly distinguishable at the DNA level, even when morphological features are lost due to processing.

However, to confirm their identity a phylogenetic tree was constructed from the sequences downloaded from the NCBI database, for all the top matching sequences with maximum sequence coverage (Fig. 5). The tree clearly groups the sequence of *Santalum album* from the present study within the cluster consisting of many other *Santalum album* sequences from the NCBI including the one done by CMPR-AVS. Similarly, the *Gmelina arborea* samples formed a close-knit group consisting of other *Gmelina arborea* sequences from the NCBI along with the CMPR-AVS sequence.

Discussion:

The present study demonstrates that the trnH–psbA intergenic spacer provides reliable distinction between *Santalum album* and *Gmelina arborea*, two species frequently confused in

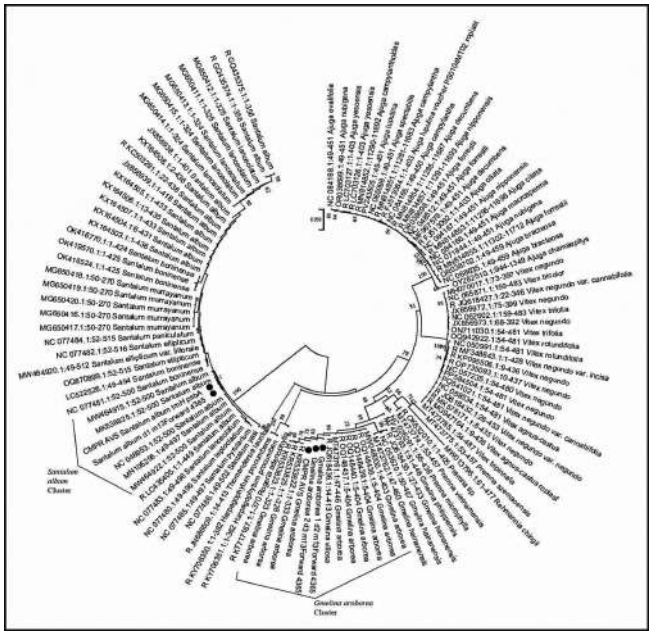

Fig. 5. The phylogenetic tree showing position of specific samples of this study relative to the matching sequences from the NCBI and inferred from the Neighbor-Joining method using the trnH-psbA sequence.

Phylogenetic analysis was conducted using the Neighbor-Joining (NJ) method implemented in MEGA11. Evolutionary distances between sequences were computed under the Kimura 2-Parameter (K2P) model, expressed as the number of base substitutions per site. The resulting tree was drawn to scale, with branch lengths proportional to these evolutionary distances. This analysis comprised 112 nucleotide sequences. To handle ambiguous or missing data, the pairwise deletion option was applied, ensuring that only unambiguous positions were included for each sequence pair. The final alignment consisted of 566 nucleotide positions. All evolutionary analyses were performed in MEGA11. (Saitou and Nei 1987; Kimura, 1980; Tamura *et al*., 2021).

commercial wood markets. Despite superficial morphological similarity, sequence analysis revealed distinct nucleotide patterns and phylogenetic separation, confirming their genetic independence.

The 100% match of *S. album* with GenBank references validates both sample authenticity and robustness of the method. Similarly, clustering of *G. arborea* with published Gmelina sequences confirms accurate taxonomic assignment. These results align with earlier reports highlighting trnH–psbA as a highly informative barcode for closely related or morphologically ambiguous plant taxa (Kress *et al*., 2005; CBOL Plant Working Group, 2009; Nithaniyal *et al*., 2014).

Conventional identification approaches fail when sandalwood is carved, powdered, or chemically treated. DNA barcoding overcomes this limitation by targeting stable genetic markers unaffected by processing. Our findings reinforce the value of molecular authentication for high-value and endangered species such as *S. album*, where substitution directly impacts conservation, trade transparency, and consumer confidence (Bunney *et al*., 2023).

Importantly, this study specifically addresses substitution of sandalwood with *G. arborea*, a common carving timber lacking sandalwood's characteristic aroma and bioactive profile. Integration of DNA barcoding into routine authentication protocols will significantly strengthen sustainable trade practices and contribute to long-term conservation of genuine Indian sandalwood.

Acknowledgments

The completion of this research paper would not have been possible without the invaluable support and guidance of Dr. DBA Narayana, CSO, Ayurvidye Trust, Bangalore. We are sincerely grateful to Ayurvidye Trust, Bangalore, for the financial support in procuring samples and DNA barcoding. We would also like to thank Dr. Suma Tagadur for her support in the collection and authentication of plant samples. Finally, we extend thanks to Dr. Rahul from Arya Vaidya Sala, Kottakkal – Centre for Medicinal Plants Research for his efforts and reliable testing results.

References

Addgene. 2024. Protocols: Agarose Gel Electrophoresis. https://www.addgene.org/protocols/ gel-electrophoresis/.

Bunney E, McInerney FA, Dormontt E, Malik A, Welti N, Wilkins D, Plant M, Hettiarachchi DS, Thomas D, Dowell A, Hamalton T and Lowe AJ. 2023. Safeguarding sandalwood: A review of current and emerging tools to support sustainable and legal forestry. Plants, People, Planet. 5(2): 190-202. https://doi.org/10.1002/ ppp3.10349

CBOL Plant Working Group. 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences of the United States of America. 106(31): 12794-12797. https://doi.org/10.1073/pnas.0905845106

Hall TA. 1999. A user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acid Symposium Series. 41: 95-98.

Khan S, Ikram M and Faisal M. 2021. Commercial, cosmetic, and medicinal importance of sandal (*Santalum album*): A valuable forest resource. In: Non-Timber Forest Products: Food, Healthcare and Industrial Applications. Springer International Publishing. Pp.: 129-144.

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

Kress WJ, Wurdack KJ, Zimmer EA, Weigt LA and Janzen DH. 2005. Use of DNA barcodes to identify flowering plants. *Proc. Natl. Acad. Sci. USA.* 102(23): 8369-8374. https://doi.org/10.1073/pnas.0503123102.

Kumar R, Anjum N and Tripathi Y. 2015. Phytochemistry and pharmacology of *Santalum album* L.: A review. World Journal of Pharmaceutical Research. 4: 1842-1876.

Letsiou S, Madesis P, Vasdekis E, Montemurro C, Grigoriou ME, Skavdis G, Moussis V, Koutelidakis AE and Tzakos AG. 2024. DNA barcoding as a plant identification method. Applied Sciences. 14(4): 1415. https://doi.org/10.3390/app14041415.

Nithaniyal S, Newmaster SG, Ragupathy S, Krishnamoorthy D and Vassou SL. 2014. DNA barcode authentication of wood samples of threatened and commercial timber trees within the tropical dry evergreen forest of India. PLoS ONE. 9(9): e107669. https://doi.org/10.1371/journal.pone.0107669.

Qiagen. 2016. Handbook: DNeasy Plant Mini/Maxi Kit (Cat. nos. 69104, 69106, 68163 and 68165). Qiagen.https://s8439dac10fdb3f99.jimcontent.com/download/version/1477120800/ module/12695627832/name/Dneasy+miniplant+handbook.pdf.

Rai SN and Sarma CR. 1990. Sandalwood: Its cultivation and utilization. Indian Council of Forestry Research and Education.

Saitou N and Nei M. 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Molecular Biology and Evolution. 4: 406-425.

Sang T, Crawford DJ and Stuessy TF. 1997. Chloroplast DNA phylogeny, reticulate evolution and biogeography of Paeonia (Paeoniaceae). American Journal of Botany. 84(8): 1120-1136.

Sharma A and Kothari SL. 2021. A comprehensive review on *Santalum album* L. (Indian sandalwood): Botany, propagation, phytochemistry, and industrial applications. Journal of Phytology. 13(3): 132-141.

Sharma C, Sharma M and Lipoktila A. 2013. Timber species used in traditional wood carving by Ao Tribe of Nagaland. Life Sciences Leaflets. 4: 31-40.

Shukla SR, Shashikala and Sujatha. 2021. Anatomical, physical and mechanical properties of *Santalum album* Linn. (Sandalwood) and its various uses. Wood Is Good: Grow More, Use More. 2(2).

Singh SP, Yadav B and Anupam K. 2022. *Madhuca longifolia* (Mahuwa) and *Santalum album* (Indian sandalwood). In: Herbs, Shrubs, and Trees of Potential Medicinal Benefits. CRC Press. Pp.: 461-470.

Tamura K, Stecher G and Kumar S. 2021. MEGA 11: Molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution. https://doi.org/10.1093/molbev/msab1.

Tate JA and Simpson BB. 2003. Paraphyly of Tarasa (Malvaceae) and diverse origins of the polyploid species. Systematic Botany. 28: 723-737.

Waghchaure AG, Jadhav RS, Vikhe DN and Ghule, YL. 2022. A review on pharmacognostic, phytochemical and biological potential of *Gmelina arborea* Roxb (Shivan plant). International Journal of Research and Review. 9(3): 237-248. https://doi.org/10.52403/ijrr.20220328.