



OPEN

Successful identification of the species of the semipetrified amber medicinal resin benzoin using molecular diagnostic technology

Jian Feng¹, Qingqin He^{1,2}, Anzhen Xie¹ & Yangyang Liu^{1,2}✉

Benzoin is an incomplete lithified resin secreted from the trunk of the *Styrax* Linn. that is known as "semipetrified amber" and has been widely used in medicine due to its blood circulation-promoting and pain-relieving properties. However, the lack of an effective species identification method due to the numerous sources of benzoin resin and the difficulty of DNA extraction has led to the uncertainty of species of benzoin in the trade process. Here, we report the successful extraction of DNA from benzoin resin containing bark-like residues and the evaluation of commercially available benzoin species using molecular diagnostic techniques. By performing a BLAST alignment of ITS2 primary sequences and homology prediction analysis of ITS2 secondary structures, we found that commercially available benzoin species were derived from *Styrax tonkinensis* (Pierre) Craib ex Hart. and *Styrax japonicus* Sieb. et Zucc. of the genus *Styrax* Linn. In addition, some of the benzoin samples were mixed with plant tissues from other genera, accounting for 29.6%. Therefore, this study provides a new method to solve the problem of species identification of semipetrified amber benzoin using information from bark residues.

Benzoin is the resin secreted from the trunk of the benzoin plant and is used as a medicine^{1,2}. The main chemical components of benzoin are balsamic acids, lignans, terpenes, steroids, and other compounds^{3–6}. Pharmacological studies have shown that benzoin possesses anti-inflammatory, antipyretic, antitumor, analgesic and estrogen-promoting properties, protects against cerebral hypoxia, and promotes blood–brain barrier permeability, among other properties^{7–10}. Approximately 130 species of the genus *Styrax* Linn., both trees and shrubs, are distributed worldwide, mainly in Asia (China), Southeast Asia (Vietnam, Thailand, Sumatra, Indonesia) and North America^{11,12}.

During the Tang Dynasty in ancient China, benzoin was recorded as a spice in the *Book of Jin* and as a medicine in the *Tang Materia Medica (Xīnxiū Běncǎo)*^{13,14}. It has a history of more than 700 years of medicinal use and was used in approximately 93 prescriptions, such as in *zhībǎo boluses* and *storax pill* in the *Tàimín Huimín and Pharmaceutical Bureau Formula* and in *dàhuólùò boluses* in *Sages' Salvation Records*^{15,16}. Benzoin has well-established uses in traditional forms of medicine. Several national pharmacopoeias, including those from China, the United States, Europe, and India, describe the specifications and tests for benzoin. The Chinese medicinal records identify benzoin as derived from the species *Styrax tonkinensis* (Pierre) Craib ex Hart., whereas *S. tonkinensis* and *Styrax benzoin* Dryander are recorded in the European Pharmacopoeia standard. The United States Pharmacopoeia and the Indian Pharmacopoeia record benzoin as derived from *S. benzoin*, *S. paralleloneurus* Perkins (trade name Sumatran benzoin), *S. tonkinensis*, or other species of the genus *Styrax* Linn. (trade name Siam benzoin)^{17–21}. Benzoin is used in the form of a tincture²². The British Pharmacopoeia specifies the use of Sumatra benzoin in Benzoin Inhalation and Compound Benzoin Tincture²³. The United States Pharmacopoeia

¹Hainan Provincial Key Laboratory of Resources Conservation and Development of Southern Medicine, Hainan Branch of the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, No. 4, Yaogu 4th Road, Xiuying District, Haikou 570311, Hainan Province, China. ²Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education and National Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100193, China. ✉email: yliu@implad.ac.cn

also describes a compound benzoin tincture, although it does not specify which type of benzoin is to be used. The Swiss Pharmacopoeia and the Indian Pharmacopoeia describe a simple benzoin tincture using Siam benzoin. Benzoin is also used in other medicinal preparations (both official and proprietary), such as over-the-counter cough suppressants, cold and flu preventatives, lotions, mouthwashes, and antibacterial powders²⁴. It is likewise used as an additive in aromatherapy²⁵. In addition, benzoin is also widely used in food, flavoring, and daily chemical products. The main role of benzoin in food is as a flavoring agent, and it is used to create the flavor of chocolate. In Denmark and Switzerland, benzoin is used as a flavoring agent for baked goods, which can fix other flavors and add spiciness^{26,27}.

Benzoin is approved for use in food in the United States. As with all food additives, benzoin is subject to periodic review by the Joint FAO/WHO Expert Committee on Food Additives to assess its safety²⁸. In addition, the largest use of benzoin, in terms of comparative quantities, is for incense substitutes in religious ceremonies and for the manufacture of spices. These fragrances are then blended and used in a wide range of end products, such as personal hygiene and care products and household products^{5,29}. Siamese benzoin produces a lighter-colored extract than Sumatran benzoin for use in the fragrance, flavor, and pharmaceutical industries to prepare resins, tinctures, or other types of extracts³⁰. Therefore, we infer that species differences in benzoin may influence its traditional use.

The resin of one or various plants in the genus *Styrax* Linn. has been used as "benzoin", such as in the United States Pharmacopoeia collections. In China, this medicinal material was mainly imported from abroad through the Silk Road, a practice that continues to this day. Currently, the true identity of the benzoin species available in the Chinese domestic herbal market is uncertain, and the classification of medicinal plants and spices in general is unclear. In addition, in the North Sumatra region of Indonesia, the benzoin resin production process involves first removing shrubs and parasitic plants from around the benzoin tree, tapping the bark, and then cutting the opening with a knife. After an interval of 3–4 months, the resin is cut, usually in summer and autumn, dried for 1–2 weeks, and finally cleaned (removal of excess bark). The collected benzoin is pooled together, sorted according to quality, and sold³¹. The whole process is not strictly standardized and confusion between different species of benzoin resin is evident³². Therefore, a method for the identification of benzoin species must be established.

In recent years, research on benzoin quality has mainly been based on the development of a chemical-based volatile component analysis and aromatic lipid content determination methods, such as gas chromatography-mass spectrometry, solid phase microextraction, headspace sampling, and high-performance liquid chromatography-frit fast atom bombardment mass spectrometry^{33,34}. In addition, the existing quality standards for medicinal materials record benzoin quality control indicators, including identification, total ash content, loss-on-drying, alcohol-insoluble matter and benzoic acid contents. These conventional indicators are mainly used for benzoin quality conformity testing, but they do not identify the species of benzoin.

DNA barcoding has received increasing attention as a rapid and accurate method of species identification that compares a standard DNA fragment from the species of interest with a library of DNA barcodes of known taxonomy. It is characterized by the fact that only one or a few suitable gene segments are selected to accurately identify most species of the entire genus and family, and it has high repeatability and stability^{35–37}. Chen et al.³⁸ established a DNA barcode identification system for traditional Chinese medicinal materials. The ITS2 sequence is short in length and easy to amplify. It has been widely used in systematic research as a part of ITS and has been confirmed to have the highest resolution in DNA barcodes³⁸. However, at present, molecular biology studies examining benzoin are mainly focused on the chromosome size, ploidy, chloroplast genome, phylogeny, and geographic evolution^{39–42}.

The key to species identification using molecular diagnostic approaches is to extract DNA from the target. Previous reports have not addressed the species identification of benzoin resinous medicinal materials. Benzoin is a resinous medicinal material secreted by injured *Styrax* trees. Theoretically, these resins do not contain the tissues of the benzoin tree, and DNA is difficult to extract. However, when investigating the actual status of commercially available benzoin, we fortunately found that the resin contains a bark-like residue. This residue may be the bark of the *Styrax* tree. This residue is produced using a procedure similar to the process of amber formation, but the resin is not completely lithified, and thus it is called semipetrified amber^{43–45}. However, these residues stick to the resin, and the DNA is not easily extracted. Therefore, the DNA extracted from these residues must be explored and then ITS2 primary and secondary structures should be used as molecular diagnostic techniques to identify the species. This study provides a new method to solve the problem of identifying species producing the semipetrified amber benzoin medicinal material by obtaining original information on these residues. This report is the first to show that the correlation of different information may help determine the origin of benzoin.

Results

DNA concentration, sequence analysis and species identification. We anticipated the direct extracting of DNA from crushed benzoin (resin) samples after sampling in this investigation, but the extracted DNA concentration was insufficient for subsequent tests (results not shown). Thus, we investigated scraping bark-like residues off the benzoin resin with a knife and treating each residue as a separate sample with molecular diagnostic identification, which has the advantage of preventing sequencing failure in subsequent trials due to sample mixing (Fig. 1).

In total, 27 batches of benzoin samples containing 40 individual bark-like residues were analyzed in this study. The DNA concentrations obtained from these residues ranged from 8.0 to 203.8 ng/ μ L, and the DNA purity ranged from 1.70 to 2.08. Although the AX151 sample had the lowest DNA concentration of 8.0 ng/ μ L, it had a purity of 1.89 and was able to be successfully amplified by PCR (Table 1). Therefore, treating the samples with 95% ethanol avoided the effect of resin on sample DNA extraction and amplification. In addition, because the samples originated from bark-like residues in benzoin resin, whose DNA was degraded, even though the DNA

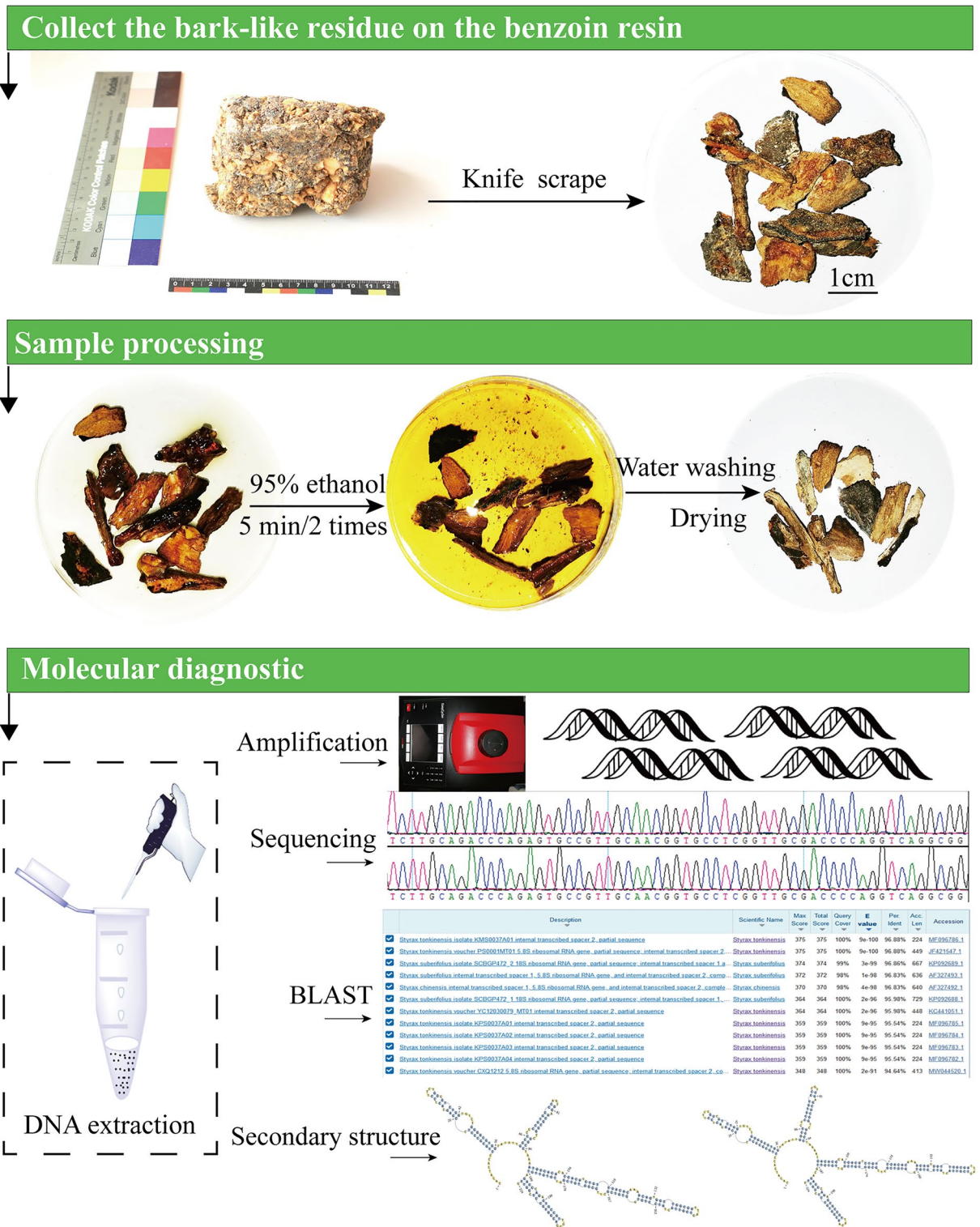


Figure 1. Process of pretreatment and molecular diagnostic identification of bark-like residues on benzoin.

concentration and purity satisfied the requirements for subsequent experiments, we chose primers for the nuclear gene ITS2 as the universal barcode as primers for amplification, and all DNA samples successfully amplified and sequenced. The complete ITS2 sequence was obtained by removing the 5.8S and 26S regions from the sequences obtained using sequencing, which had a length of 217–230 bp.

Forty individual samples were identified by a BLAST comparison with GenBank data from species of the genera *Styrax* Linn., *Dimocarpus* Lour., *Aquilaria* Lam., *Ageratum* L., *Glycine* Willd., *Acronychia* J. R. Forst. & G. Forst., *Trema* Lour., and *Musa* L., with a BLAST species identification rate > 92% (Table 2). The BLAST and

Sample no	DNA concentration (ng/μL)	A260	A280	DNA purity (A260/A280)	ITS2 Length (bp)	GC content (%)
AX011	16.4	0.329	0.166	1.98	224	55.80
AX012	25.6	0.514	0.260	1.98	224	55.80
AX013	101.7	2.034	1.038	1.96	224	55.61
AX014	73.5	1.470	0.764	1.92	225	56.00
AX021	39.7	0.796	0.393	2.03	224	60.71
AX031	52.4	1.047	0.531	1.97	224	60.27
AX041	62.8	1.256	0.664	1.89	224	60.27
AX051	135.2	2.703	1.377	1.96	226	59.73
AX061	140.8	2.816	1.407	2.00	225	60.71
AX062	51.7	1.034	0.530	1.95	224	60.71
AX071	41.4	0.829	0.459	1.81	230	57.39
AX072	51.3	1.027	0.526	1.95	224	55.80
AX081	53.1	1.061	0.550	1.93	224	55.80
AX091	52.0	1.042	0.500	2.08	224	55.36
AX092	61.5	1.230	0.597	2.06	224	55.80
AX093	33.6	0.672	0.365	1.84	224	55.80
AX101	15.4	0.308	0.154	2.00	224	55.80
AX102	18.3	0.037	0.020	1.85	217	64.98
AX111	75.4	1.509	0.770	1.96	224	60.71
AX121	68.9	1.380	0.723	1.91	224	60.71
AX122	48.3	0.965	0.472	2.04	230	51.74
AX131	13.1	0.261	0.142	1.84	224	55.36
AX132	87.6	1.752	0.927	1.89	224	55.80
AX141	99.0	1.981	1.075	1.84	224	55.80
AX151	8.0	0.160	0.082	1.95	230	57.39
AX161	53.1	1.062	0.593	1.79	224	60.71
AX171	31.5	0.631	0.317	1.99	217	56.22
AX172	12.2	0.246	0.133	1.85	222	58.56
AX173	19.6	0.393	0.205	1.92	224	55.36
AX174	17.1	0.342	0.187	1.83	217	65.90
AX181	38.4	0.768	0.411	1.87	224	60.71
AX191	95.6	1.912	0.998	1.92	224	60.71
AX201	31.6	0.633	0.348	1.82	222	58.56
AX211	53.5	0.654	0.347	1.88	224	60.71
AX221	52.8	1.056	0.557	1.90	224	61.16
AX231	82.7	1.655	0.877	1.89	218	68.35
AX241	80.3	1.607	0.828	1.94	224	60.27
AX251	43.0	0.859	0.506	1.70	217	64.98
AX261	57.8	1.156	0.630	1.83	224	55.8
AX271	203.8	4.075	2.157	1.89	224	60.71

Table 1. DNA concentration and purity of samples, and ITS2 length and GC content after sequencing.

secondary structure analyses showed that 30 samples were derived from the genus *Styrax* Linn., of which 16 were identified as *S. tonkinensis* and 14 were identified as *S. japonicus*. In addition, 10 samples were derived from other genera. The results from the BLAST search and secondary analysis results of sample AX174 were different, but both indicated that the sample belongs to the *Trema* Lour species. Samples AX071 and AX151 were identified as *Aquilaria sinensis* (Lour.) Spreng., AX102 and AX215 as *Dimocarpus longan* Lour., AX172 and AX201 as *Acronychia pedunculata* (L.) Miq., and AX171 as *Glycine soja* Siebold & Zucc. and *Glycine max* (Lour.) Merr. AX122 and AX231 were identified as *Ageratum conyzoides* L. and *Musa acuminata* Colla, respectively. These species are mainly found in tropical and subtropical regions, consistent with benzoin production areas. Benzoin is resinous, and it is inferred that the tissue of these plants may have been mixed with the harvested benzoin.

We assigned the 14 samples identified by BLAST as *S. japonicus* to group 1 and the 16 samples identified as *S. tonkinensis* to group 2. Twenty base variable loci were identified by ITS2 sequence alignment with *S. japonicus* and *S. tonkinensis* species in GenBank (Table 3). Base sites 13, 24, 46, 75, 78, 102, 104, 126, and 219 can be used as specific identification sites for these two species. However, different base sites were identified in the collected samples from these two species, such as G at base site 7 and C at base site 189 in group 2, while *S. japonicus* and *S. tonkinensis* species downloaded from GenBank had A and T for these base sites, respectively. In addition,

Sample no	BLAST results				Secondary structures results	
	<i>Styrax</i> Linn	Other species	Per. Ident (%)	E value	<i>Styrax</i> Linn	Other species
AX011	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX012	<i>S. tonkinensis</i>		99.55	1e-107	<i>S. tonkinensis</i>	
AX013	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX014	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX021	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX031	<i>S. tonkinensis</i>		94.20	1e-93	<i>S. tonkinensis</i>	
AX041	<i>S. tonkinensis</i>		96.88	3e-100	<i>S. tonkinensis</i>	
AX051	<i>S. tonkinensis</i>		92.40	6e-82	<i>S. tonkinensis</i>	
AX061	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX062	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX071		<i>Aquilaria sinensis</i> (Lour.) Spreng (<i>Aquilaria</i> Lam.)	99.13	2e-112		<i>A. sinensis</i>
AX072	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX081	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX091	<i>S. japonicus</i>		99.11	4e-108	<i>S. japonicus</i>	
AX092	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX093	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX101	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX102		<i>Dimocarpus longan</i> Lour (<i>Dimocarpus</i> Lour.)	100	1e-107		<i>D. longan</i>
AX111	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX121	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX122		<i>Ageratum conyzoides</i> L (<i>Ageratum</i> L.)	99.13	2e-112		<i>A. conyzoides</i>
AX131	<i>S. japonicus</i>		98.66	5e-107	<i>S. japonicus</i>	
AX132	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX141	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX151		<i>A. sinensis</i>	100	4e-113		<i>A. sinensis</i>
AX161	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX171		<i>Glycine soja</i> Siebold & Zucc (<i>Glycine</i> Willd.)	97.70	9e-100		<i>G. soja</i>
		<i>Glycine max</i> (L.) Merr (<i>Glycine</i> Willd.)	97.70	9e-100		<i>G. max</i>
AX172		<i>Acronychia pedunculata</i> (L.) Miq (<i>Acronychia</i> J. R. Forst. & G. Forst)	99.55	1e-108		<i>A. pedunculata</i>
AX173	<i>S. japonicus</i>		99.11	1e-108	<i>S. japonicus</i>	
AX174		<i>Trema cannabinum</i> var. <i>dielsianum</i> (Hand.Mazz.) C. J. Chen (<i>Trema</i> Lour.)	100	1e-107		<i>Trema micranthum</i> (L.) Blume
		<i>Trema nitidum</i> C. J. Chen (<i>Trema</i> Lour.)	100	1e-107		
		<i>Trema orientale</i> (L.) Blume (<i>Trema</i> Lour.)	100	1e-107		
		<i>Trema tomentosa</i> (Roxb.) H.Hara (<i>Trema</i> Lour.)	100	1e-107		
AX181	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX191	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX201		<i>Acronychia pedunculata</i> (L.) Miq (<i>Acronychia</i> J. R. Forst. & G. Forst.)	99.10	1e-107		<i>A. pedunculata</i>
AX211	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	
AX221	<i>S. tonkinensis</i>		96.43	3e-99	<i>S. tonkinensis</i>	
AX231		<i>Musa acuminata</i> Colla (<i>Musa</i> L.)	96.33	1e-97		<i>M. acuminata</i>
AX241	<i>S. tonkinensis</i>		97.32	2e-101	<i>S. tonkinensis</i>	
AX251		<i>D. longan</i>	100	1e-107		<i>D. longan</i>
Continued						

Sample no	BLAST results				Secondary structures results	
	<i>Styrax</i> Linn	Other species	Per. Ident (%)	E value	<i>Styrax</i> Linn	Other species
AX261	<i>S. japonicus</i>		99.55	9e-110	<i>S. japonicus</i>	
AX271	<i>S. tonkinensis</i>		96.88	9e-100	<i>S. tonkinensis</i>	

Table 2. Results of BLAST comparison and secondary structure analysis of ITS2 sequences of samples.

Species	Base site											
	7	13	24	32	33	46	60	75	77	78	84	102
<i>S. japonicus</i> ¹⁾	A	T	A	T	T	T	A	T	T	A	T	A
Group 1	A	T	A	T	T	T	A	T	T	A	C	A
<i>S. tonkinensis</i> ²⁾	A	C	C	C	C/T	C	A	C	T/A/C	G	C	G/A
Group 2	G	C	C	T	T	C	G/A	C	C	G	C	G
Species	104	105	106	109	126	185	113	189	200	209	219	
<i>S. japonicus</i> ¹⁾	G	C	A	T	G	T	A	T	T	A	T	
Group 1	G	C	A	T	G	T	A	T	T	A	T	
<i>S. tonkinensis</i> ²⁾	A	T/C	G/A	T/C	A	G	A/G	T/C	T/C	G/A	C	
Group 2	A	C	A	C	A	G/C	A	C	C	A	C	

Table 3. BLAST comparison of base difference sites of ITS2 sequences from *Styrax* Linn. samples with *S. japonicus* and *S. tonkinensis*. ¹⁾*S. japonicus* has the following accession numbers in Genbank: AB114900, LC534305, LC600960, and MF349070; ²⁾*S. tonkinensis* has the following accession numbers in Genbank: JF421547, MF096782, MF096783, MF096784, MF096785, and MF096786. Group 1 is the 14 samples identified as *S. japonicus* and group 2 is the 16 samples identified as *S. tonkinensis*. Bolded black indicates differential loci for *S. japonicus* and *S. tonkinensis*.

two base sites in group 2 were identical to those of *S. japonicus* (base sites 32 and 209), and two base sites were variable (base sites 60 and 185).

Analysis of the ITS2 secondary structure. The ITS2 database was used to predict the homology of the secondary structures of 48 major species of the genus *Styrax* Linn. The ITS2 secondary structures of all species were folded into the typical structure of a central loop with four helices (I, II, III, and IV), with helix III identified as the longest helix, followed by helices I, II, and IV. Helix II was the most conserved, with the central main loop relatively conserved and rich in purine bases^{46–48}. The ITS2 secondary structure predictions for 40 individual samples are shown in Table 4. By homology prediction, 30 samples were derived from species of the genus *Styrax* Linn., 16 of which were identified as *S. tonkinensis* and 14 as *S. japonicus*. In addition, 10 samples were derived from species of other genera. The ITS2 secondary structures of the two species identified as *S. tonkinensis* and *S. japonicus* in the benzoin samples are shown in Fig. 2A. The helix positions and angles of the secondary structures of these two species are completely different, with helices I and III identified as more variable due to multiple unpaired bases and positional differences. In contrast, the shapes of helices II and IV, although basically the same, also contain unpaired bases, leading to differences. *S. tonkinensis* contains three slightly different secondary structures, as shown in Fig. 2A a), b), and c). The red dashed boxes represent the base pairs between the base of helix III and the first unpaired loop on the helix arm of *S. tonkinensis* a) and *S. tonkinensis* b), and

Species	Helix I	Helix II	Helix III	Helix IV	Main loop
<i>S. japonicus</i>	–	hCBCs (73/84: G-C → G-T)	–	–	–
Group 1					
<i>S. tonkinensis</i> a)	32: T → C	77: C → A; C → T	Position difference; 100/188: → A-T hCBCs (103–185: C-C → C-G)	209: A → G	7: G → A; 60: G → A
<i>S. tonkinensis</i> b)					
Group 2					
<i>S. tonkinensis</i> c)	32/33:T-T → C-C	–	Position difference; 100–106/182–188: Base pairs difference	hCBCs (200–215: T-G → C-G)	7: G → A; 60: G → A
Group 2					

Table 4. Comparison of base differences on the helices of *S. tonkinensis* and *S. japonicus* with the samples. Group 1 was the 14 samples identified as *S. japonicus*, and group 2 was the 16 samples identified as *S. tonkinensis* by BLAST. –: represents no differences in base sequences.

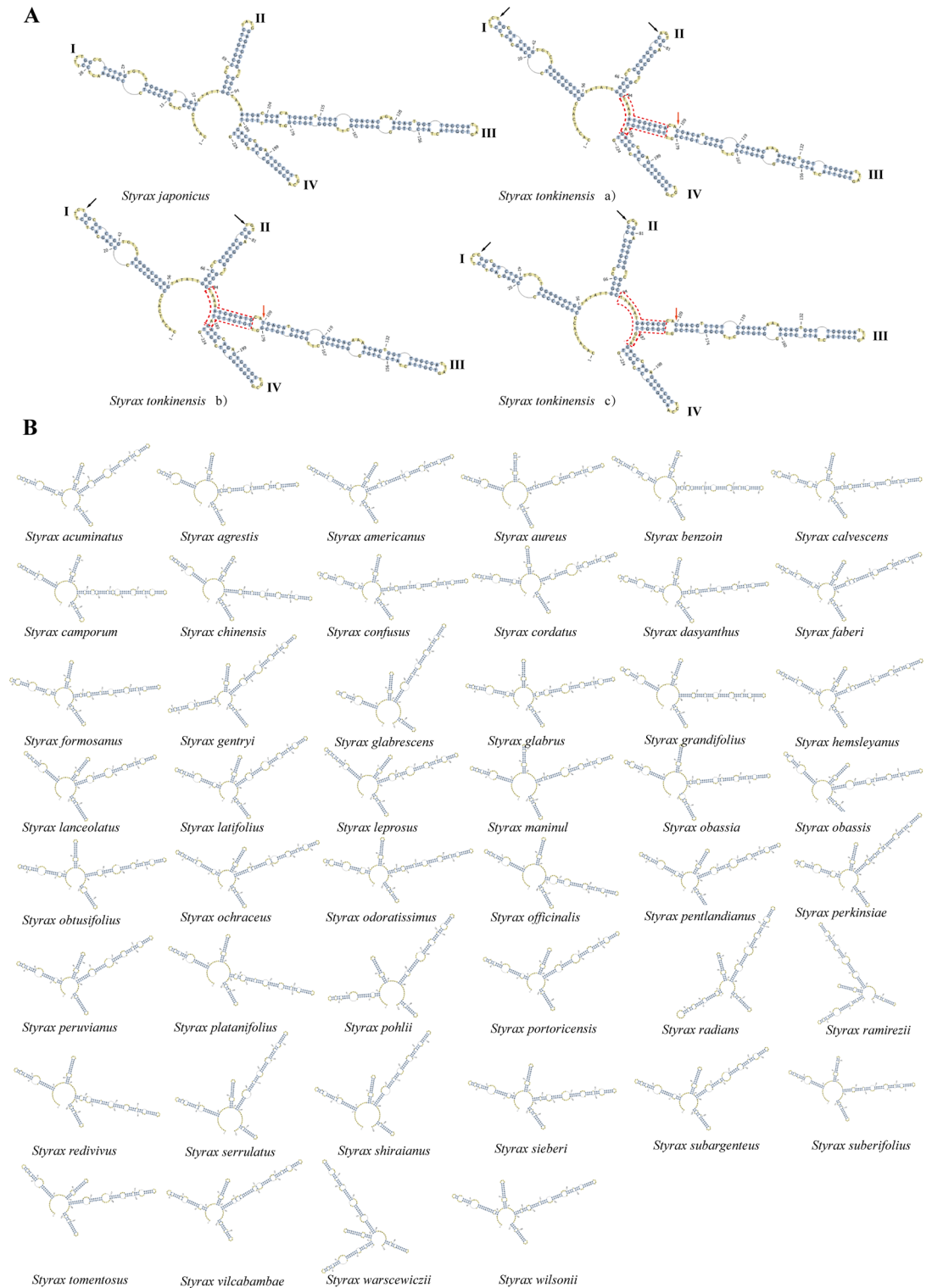


Figure 2. (A) Comparison of the secondary structure of ITS2 between benzoin samples from *S. tonkinensis* and *S. japonicus*. *S. tonkinensis* secondary structures a), b), and c) were all slightly different. Arrows indicate the sites of hCBCs (red) and different base sites (black), and the red dashed boxes represent the different base pairs in helix III of *S. tonkinensis*. (B) The secondary structures of ITS2 in 46 taxa analyzed in this study, except for *S. tonkinensis* and *S. japonicus*. The species name is marked below each structure.

these differences are due to differences in the position of helix III. *S. tonkinensis* contains three ITS2 secondary structures with different base sites, such as bases A, C, and T at position 32 in helix I (black arrows). In addition, hemi compensatory base changes (hCBCs) were identified in helix III (109/179: C-G → T-G, red arrows).

We compared the DNA sequences of the two sets of basal samples obtained using BLAST identification. The ITS2 secondary structure sequence in group 1 was identical to that of *S. japonicus*, with the presence of hCBCs (73/84: G-C→G-T) detected only on helix II (Table 4). The results for the predicted secondary structure of ITS2 in group 2 homologs were the same as those of *S. tonkinensis*, but the differences in the position of helix III were clear, and variations in base sites were also observed. For example, sample AX051 differed from the other 15 samples by the secondary structure at base 60 in the main loop (A) and the presence of hCBCs in helix III (103–185: C-C→C-G), but the remaining sites were identical. Compared with *S. tonkinensis* a) and *S. tonkinensis* b), the bases at sites 7 and 60 of the main loop in group 2 are G, the bases at site 32 of the top asymmetric loop in helix I are T, the bases at site 77 of the top asymmetric loop of helix II are C, while the base 100/188 of helix III lacks the A-T base pair and has hCBCs (103–185: C-C→C-G), and base 209 of the top asymmetric loop of helix IV is A. Compared to *S. tonkinensis* c), the sample in group 2 have G bases at sites 7 and 60 of the main loop, T-T bases at sites 32–33 of the top asymmetric loop of helix I, different base pairs at 100–106/182–188 of helix III, and the presence of hCBCs on helix IV (200–215: T-G→C-G), as shown in Table 4. Therefore, the secondary structures of ITS2 from *S. tonkinensis* and *S. japonicus* differ significantly, and the species may be further precisely identified based on this information. In addition, the ITS2 secondary structures from the remaining 46 species of the genus *Styrax* Linn. are shown in Fig. 2B. The sizes, positions, and angles of the four helices and the bases in the helix vary among the 38 species, except for the four species pairs of *S. calvescens* and *S. formosanus*, *S. formosanus* and *S. dasyanthus*, *S. formosanus* and *S. confusus*, and *S. confusus* and *S. hemsleyanus*. The primary sequences in these four groups of species are identical to each other. We tried to analyze the differences in their secondary structures by measuring homologous folding, but unfortunately, the ITS2 secondary structures of the four groups of species were also identical.

The species of the 27 batches of benzoin samples collected were inferred from a comprehensive BLAST analysis and by analyzing the ITS2 secondary structures of 40 individual samples, as shown in Table 5. The AX01 sample contained two species of the genus *Styrax* Linn., *S. japonicus* and *S. tonkinensis*. No species of the genus *Styrax* Linn were detected in the four batches of samples (AX15, AX20, AX23, and AX25), and only other plant species were detected, but this result does not mean that these samples were not benzoin. The contamination of commercially available benzoin samples was prominent, accounting for 29.6% of the samples, which is related to the harvesting and transportation processes at the site of origin.

Batch no	Included species	Judgment results
AX01	<i>S. japonicus</i> ; <i>S. tonkinensis</i>	Y
AX02	<i>S. tonkinensis</i>	Y
AX03	<i>S. tonkinensis</i>	Y
AX04	<i>S. tonkinensis</i>	Y
AX05	<i>S. tonkinensis</i>	Y
AX06	<i>S. tonkinensis</i>	Y
AX07	<i>S. japonicus</i> ; <i>A. sinensis</i>	Y/N
AX08	<i>S. japonicus</i>	Y
AX09	<i>S. japonicus</i>	Y
AX10	<i>S. japonicus</i> ; <i>D. longan</i>	Y/N
AX11	<i>S. tonkinensis</i>	Y
AX12	<i>S. tonkinensis</i> ; <i>A. conyzoides</i>	Y/N
AX13	<i>S. japonicus</i>	Y
AX14	<i>S. japonicus</i>	Y
AX15	<i>A. sinensis</i>	-/N
AX16	<i>S. tonkinensis</i>	Y
AX17	<i>S. japonicus</i> ; <i>G. soja</i> ; <i>G. max</i> ; <i>A. pedunculata</i> ; <i>Trema Lour</i>	Y/N
AX18	<i>S. tonkinensis</i>	Y
AX19	<i>S. tonkinensis</i>	Y
AX20	<i>A. pedunculata</i>	-/N
AX21	<i>S. tonkinensis</i>	Y
AX22	<i>S. tonkinensis</i>	Y
AX23	<i>M. acuminata</i>	-/N
AX24	<i>S. tonkinensis</i>	Y
AX25	<i>D. longan</i>	-/N
AX26	<i>S. japonicus</i>	Y
ZX27	<i>S. tonkinensis</i>	Y

Table 5. Comprehensive determination of species in 27 batches of benzoin samples. Y represents *Styrax* Linn. species, N represents other genus species, and – represents *Styrax* Linn. species not detected.

Discussion

The core of Chinese medicine identification is to conduct variety authenticity and quality evaluation, which are related to drug safety and the quality assurance of the Chinese medicine industry. Traditional identification, as represented by the application of classical morphological classification to study the origin of Chinese herbs, is based on the description of individual traits and macroscopic observation, and the conclusions it yields are often imperfect⁴⁹. This limitation is attributed to the need for specialized taxonomic knowledge, which is mostly based on an empirical judgment. In addition, microscopic identification, physicochemical identification, and other mainstream methods for the identification of Chinese medicinal materials have some limitations in the identification of genera with various species. For example, microscopic identification requires specialized expertise, and physical and chemical identification is time-consuming⁵⁰. Molecular diagnostic technology, as represented by DNA barcoding, can accurately identify and characterize species and is a simple and precise method for the identification of Chinese medicinal materials with clear judgment criteria; it is thus suitable for the accurate identification of Chinese medicinal materials and their various species without requiring professional knowledge of taxonomy or ambiguous morphological characters³⁶.

Most Chinese herbs and tablets undergo further processing and preparation steps, where DNA degradation is more severe, particularly for resinous herbs such as benzoin, which is formed by the solidification of secretions after the injury of trees of the genus *Styrax* Linn., and these resins mainly contain balsamic acids, lignans, and terpenoids^{1,5,51–53}. Despite the original processing of the benzoin resin, bark or plant tissue fragments remain on the resin. However, these resins are insoluble in water, which increases the difficulty of DNA extraction. Therefore, the most critical step for molecular diagnostic techniques is to extract higher quality DNA from the samples. In preliminary experiments, benzoin was directly ground to extract DNA (the purity and concentration of most of the DNA samples were low and did not meet the experimental requirements), but the amplification results were unsatisfactory (most of the amplifications failed, and some sequences were heterozygous). This failure might be due to low DNA yield because of lower retention or resin inhibition. Therefore, we changed our approach and collected the bark-like residues directly from the benzoin resin and used each piece of residue as a separate sample. These bark-like residues were treated with 95% ethanol, and fortunately, DNA was successfully extracted and purified from them using a plant kit. The purpose of the 95% ethanol treatment is to remove as many of the substances that interfere with DNA release as possible (including other impurities). Previous scholars have mainly explored methods for extracting DNA extraction methods from dried wood (Lobeliaceae), archaeological wood remains, ancient wood (oak), endangered wood (prismatic wood), and other samples. Although the DNA of wood may be severely degraded, DNA can be extracted through method optimization^{54–57}. To date, resinous medicinal materials have been studied in *Draconis sanguis* (resin exuded from fruits) and agarwood (wood-containing resin) by improving DNA extraction methods^{58,59}. The present study is also the first to report the successful extraction of DNA from resin samples exuded from excised trees that meets the requirements for subsequent experiments. Therefore, this study shows that resinous materials used for DNA extraction must be pretreated.

In recent years, DNA barcodes have been widely used in species identification studies of medicinal materials, and the candidate barcodes are mostly the chloroplast markers *matK*, *rbcL*, *psbA-trnH*, *trnL*, and *trnK*, the nuclear genes ITS and ITS2, and their corresponding barcode combinations^{38,60–62}. However, for some samples with severe DNA degradation and fragmented DNA, the use barcodes with longer amplification products as primers is unsuccessful. The selection of barcodes with appropriate PCR amplification product lengths is crucial for molecular diagnostic studies, especially for DNA degradation or fragmentation in medical materials such as resinous or stemmed wood. ITS2, a nuclear gene sequence, is generally 200–300 bp in length and is suitable for identification at the genus and species levels. Moreover, the secondary structure is more conserved, which is a neck-loop structure formed by the single-strand DNA folding back on itself, with paired bases forming the stem and unpaired bases forming the loop, and a mature ITS2 database and structure prediction software are available^{63,64}. In species identification studies, the prediction of secondary structure may be a useful complement to the primary structure phylogenetic tree by avoiding or excluding the misleading effect of paralogous homologs or pseudogenes on the primary structure⁶⁵. Although the assessment of the ITS2 secondary structure was an early technique used for molecular diagnostics, it is still valid for species identification as a complement to DNA barcoding.

In the present study, 40 individual bark-like residues from 27 batches of benzoin were analyzed by performing BLAST searches of their ITS2 primary structure and secondary structure, and 30 were found to be from the benzoin genus and identified as *S. tonkinensis* and *S. japonicus*. Ten samples were identified as species of other genera, perhaps possibly related to mixing during resin collection and transport at the original cultivation site. However, uncertainty also exists about the results of identifying species of other genera in benzoin samples because only ITS2 was used as the barcode sequence in this study, and the reliable information obtained was limited. This study only presents the possibility of this potential contamination. Therefore, some additional primers (e.g., sequences of the chloroplast genome) are needed at a later stage to obtain more genetic information and improve the reliability of the identification results. In addition, morphological, taxonomic, or anatomical methods must be combined with comparative studies of the tissues of these species growing in the field to more accurately confirm the species information contained in the bark-like residues.

In this study, using the GenBank database, we compared the ITS2 sequences of the two species identified as benzoin, and 9 base sites were useful as specific identification sites for these two species. Therefore, this study is expected to develop specific DNA probes for the accurate identification of species of the genus *Styrax* Linn., which provides a good basis for determining the authenticity of benzoin products. Overall, this study provides a new method to solve the problem of species identification of the semipetrified amber medicinal resin benzoin by obtaining original information on these residues.

Materials and methods

Sample collection. Twenty-seven batches of benzoin medicinal materials were collected from medicinal material markets and drugstores in China. These collections are permitted and legal. All benzoin samples were identified by Professor Yangyang Liu of the Hainan Branch of the Institute of Medicinal Plants, Chinese Academy of Medical Sciences. We observed the visual properties of the collected benzoin samples and found that the samples were irregular and often agglomerated into clumps. The surface was orange–yellow or yellowish white, and the material was fragrant, with a gritty feel when chewed. It was very brittle and could be broken or cracked by hand; and the resin contained bark residues inside. Two to four replicates were added during the experiment for batches with high bark residues. Therefore, 40 individual bark-like residues from the final experimental materials were obtained, and the details are shown in Table 6. All voucher specimens of bark-like residues are maintained in the herbarium of the Aromatic Southern Medicine Identification Center, Hainan Branch, Institute of Medicinal Plants, Chinese Academy of Medical Sciences. In addition, 56 ITS2 sequences of 48 species of *Styrax* Linn. were downloaded from GenBank for comparative analysis (Table 7).

Batch no	Sample no	Label	Species	Producing area	Collection location	Purchase time
AX01	AX011	Benzoin	Unknown	Myanmar	Medicinal materials market of Anguo, China	2019.08.23
	AX012					
	AX013					
	AX014					
AX02	AX021	Benzoin	Unknown	Myanma		
AX03	AX031	Benzoin	Unknown	Myanma		
AX04	AX041	Benzoin	Unknown	Thailand		
AX05	AX051	Benzoin	Unknown	Iran		
AX06	AX061	Benzoin	Unknown	Iran		
	AX062					
AX07	AX071	Benzoin	Unknown	Vietnam		
	AX072					
AX08	AX081	Benzoin	Unknown	Myanmar	Medicinal materials market of Bozhou, China	2019.08.20
AX09	AX091	Benzoin	Unknown	Sumatra		
	AX092					
	AX093					
AX10	AX101	Benzoin	Unknown	Thailand		
	AX102					
AX11	AX111	Benzoin	Unknown	Indonesia		
AX12	AX121	Benzoin	Unknown	Indonesia		
	AX122					
AX13	AX131	Benzoin	Unknown	Laos	Medicinal materials market of Yulin, China	2019.09.26
	AX132					
AX14	AX141	Benzoin	Unknown	Laos		
AX15	AX151	Benzoin	Unknown	Laos		
AX16	AX161	Benzoin	Unknown	Uncertainty		
AX17	AX171	Benzoin	Unknown	Nepal	Medicinal materials market of Hehuachi, China	2019.12.14
	AX172					
	AX173					
	AX174					
AX18	AX181	Benzoin	Unknown	Indonesia		
AX19	AX191	Benzoin	Unknown	Uncertainty		
AX20	AX201	Benzoin	Unknown	India		
AX21	AX211	Benzoin	Unknown	Uncertainty		
AX22	AX221	Benzoin	Unknown	India		
AX23	AX231	Benzoin	Unknown	Uncertainty		
AX24	AX241	Benzoin	Unknown	Indonesia		
AX25	AX251	Benzoin	Unknown	Uncertainty		
AX26	AX261	Benzoin	Unknown	Indonesia		
AX27	AX271	Benzoin	Unknown	Uncertainty	Drugstore of Haikou, China	2019.6.20

Table 6. Information on samples of benzoin was collected from medicinal markets and drugstores.

No	Species	Genus	Collection site	Accession ID
1	<i>Styrax acuminatus</i> Pohl	<i>Styrax</i> Linn	Genbank	KP793409
2	<i>Styrax agrestis</i> (Lour.) G. Don	<i>Styrax</i> Linn	Genbank	AF327483
3	<i>Styrax americanus</i> Lam	<i>Styrax</i> Linn	Genbank	AF327472
4	<i>Styrax aureus</i> (Mart.) Miers	<i>Styrax</i> Linn	Genbank	AY143577
5	<i>Styrax benzoin</i> Dryander	<i>Styrax</i> Linn	Genbank	AF327494
6	<i>Styrax calvescens</i> Perk	<i>Styrax</i> Linn	Genbank	AF327468
7	<i>Styrax camporum</i> Pohl	<i>Styrax</i> Linn	Genbank	AF327504
8	<i>Styrax chinensis</i> Hu et S. Y. Liang	<i>Styrax</i> Linn	Genbank	AF327492
9	<i>Styrax confusus</i> Hemsl	<i>Styrax</i> Linn	Genbank	AF327470
10	<i>Styrax cordatus</i> A.D C	<i>Styrax</i> Linn	Genbank	AF327506
11	<i>Styrax dasyanthus</i> Perk	<i>Styrax</i> Linn	Genbank	AF327467
12	<i>Styrax faberi</i> Perk	<i>Styrax</i> Linn	Genbank	AF327485
13	<i>Styrax formosanus</i> Matsum	<i>Styrax</i> Linn	Genbank	AF327466
14	<i>Styrax gentryi</i> P.W. Fritsch	<i>Styrax</i> Linn	Genbank	AF327502
15	<i>Styrax glabrescens</i> Benth	<i>Styrax</i> Linn	Genbank	AF327475
16	<i>Styrax glabrus</i> Sw	<i>Styrax</i> Linn	Genbank	AY143582
17	<i>Styrax grandifolius</i> Ait	<i>Styrax</i> Linn	Genbank	AF327476
18	<i>Styrax hemsleyanus</i> Diels	<i>Styrax</i> Linn	Genbank	AF327477
19	<i>Styrax japonicus</i> Sieb. et Zucc	<i>Styrax</i> Linn	Genbank	AB114900; LC534305; LC600960; MF349070
20	<i>Styrax lanceolatus</i> P.W. Fritsch	<i>Styrax</i> Linn	Genbank	AF327501
21	<i>Styrax latifolius</i> Pohl	<i>Styrax</i> Linn	Genbank	AY143585
22	<i>Styrax leprosus</i> Hook. & Arn	<i>Styrax</i> Linn	Genbank	KP793405
23	<i>Styrax maninul</i> Hook. & Arn	<i>Styrax</i> Linn	Genbank	AY143576
24	<i>Styrax obassia</i> Siebold & Zucc	<i>Styrax</i> Linn	Genbank	AF327479
25	<i>Styrax obassia</i> Siebold et Zucc	<i>Styrax</i> Linn	Genbank	MF349081
26	<i>Styrax obtusifolius</i> Griseb	<i>Styrax</i> Linn	Genbank	AY143589
27	<i>Styrax ochraceus</i> Urb	<i>Styrax</i> Linn	Genbank	AY143580
28	<i>Styrax odoratissimus</i> Champ. ex Benth	<i>Styrax</i> Linn	Genbank	AF327460
29	<i>Styrax officinalis</i> L	<i>Styrax</i> Linn	Genbank	AF327489
30	<i>Styrax pentlandianus</i> J. Rémy	<i>Styrax</i> Linn	Genbank	AF327507
31	<i>Styrax perkinsiae</i> Rehd	<i>Styrax</i> Linn	Genbank	AF327459
32	<i>Styrax peruvianus</i> Zahlbr	<i>Styrax</i> Linn	Genbank	KP793406
33	<i>Styrax platanifolius</i> Engelm. ex Torr	<i>Styrax</i> Linn	Genbank	AF327486
34	<i>Styrax pohlil</i> A. D C	<i>Styrax</i> Linn	Genbank	AY143581
35	<i>Styrax portoricensis</i> Krug & Urban	<i>Styrax</i> Linn	Genbank	AF327505
36	<i>Styrax radians</i> P.W. Fritsch	<i>Styrax</i> Linn	Genbank	AF327497
37	<i>Styrax ramirezii</i> Greenm	<i>Styrax</i> Linn	Genbank	AF327499
38	<i>Styrax redivivus</i> (Torr.) L.C. Wheeler	<i>Styrax</i> Linn	Genbank	AF327490
39	<i>Styrax serrulatus</i> Roxb	<i>Styrax</i> Linn	Genbank	AF327481
40	<i>Styrax shiratanus</i> Makino	<i>Styrax</i> Linn	Genbank	AF327480
41	<i>Styrax sieberi</i> Makino	<i>Styrax</i> Linn	Genbank	AY143587
42	<i>Styrax subargenteus</i> Sleumer	<i>Styrax</i> Linn	Genbank	KP793407
43	<i>Styrax suberifolius</i> Hook. et Arn	<i>Styrax</i> Linn	Genbank	AF327493
44	<i>Styrax tomentosus</i> Bonpl	<i>Styrax</i> Linn	Genbank	AY143590
45	<i>Styrax tonkinensis</i> (Pierre) Craib ex Hartw	<i>Styrax</i> Linn	Genbank	MF096782; MF096783; MF096784; MF096785; JF421547; MF096786;
46	<i>Styrax vilcabambae</i> (D.R. Simpson) B. Walln	<i>Styrax</i> Linn	Genbank	AF327496
47	<i>Styrax warszewiczii</i> Perkins	<i>Styrax</i> Linn	Genbank	AF327500
48	<i>Styrax wilsonii</i> Rehd	<i>Styrax</i> Linn	Genbank	AF327462

Table 7. ITS2 sequences of species of the genus *Styrax* Linn. in Genbank.

DNA extraction, amplification, and sequencing. A knife was used to scrape and collect the bark-like residue in the benzoin resin, which was placed in disposable petri dishes. An appropriate amount of 95% ethanol was added twice and incubated for 5 min each, and the residue was finally washed with sterile water and dried (Fig. 1). The sample was cut into pieces and placed in a 2 mL centrifuge tube to which two steel balls were

added. Then, it was frozen in liquid nitrogen and ground in a high-throughput tissue grinder for 5 min. DNA was extracted with the modified HP Plant DNA kit (OMEGA Bio-tek), where 1000 μ L of buffer CPL and 10 μ L of β -mercaptoethanol were added, and the remaining steps were performed according to the manufacturer's instructions. DNA samples were purified using the MicroElute[®] DNA Clean-Up Kit. The concentration of total DNA was measured, the DNA purity (the ratio of A260 to A280) was calculated, and the extracted DNA was stored in a -20 °C freezer until use. Amplification was performed using primers for ITS2 (ITS2-F: ATGCGA TACTTGGTGTGAAT; ITS2-R: GACGCTTCTCCAGACTACAAT)³⁹. The PCR amplification program was as follows: pre-denaturation at 94 °C for 5 min; 40 cycles of denaturation at 94 °C for 30 s, annealing at 56 °C for 30 s, and extension at 72 °C for 45 s; extension at 72 °C for 10 min; and storage at -20 °C. PCR products were sequenced bidirectionally at Guangzhou Aike Biotechnology Co., Ltd. (Guangzhou, China).

Data analysis. Sequence splicing was performed using CodonCode Aligner V3.7.1 (CodonCode Co., USA) to remove primer regions and low-quality sequences. The sequences obtained by sequencing and those downloaded from GenBank were annotated and cropped. Based on hidden Markov models (HMMs), the 5.8S and 26S regions were removed to obtain the complete ITS2 sequence⁶³. The complete ITS2 sequence obtained through sequencing was used for the BLAST alignment analysis in GenBank, and MEGA 6.0 was used for alignment. The variation sites were recorded, and the species information of the samples was analyzed. Moreover, the ITS2 secondary structures of all species were predicted using the ITS2 database homology model⁶⁴. Additionally, hCBCs (such as C-G \rightarrow C-A or T-T \rightarrow T-C) were calculated.

Data availability

All the data generated during this study are included in this published article and its supplementary information files. ITS2 sequences generated in this study are available at GenBank (accession numbers: OP787212-OP787251). The 56 ITS2 sequences of *Styrax* Linn. were downloaded from GenBank for comparative analysis, accession no: KP793405-KP793407, KP793409, MF096782-MF096786, MF349070, MF349081, AY143580-AY143582, AY143576, AY143577, AY143585, AY143587, AY143589, AY143590, JF421547, AB114900, LC534305, LC600960, AF327459, AF327460, AF327462, AF327466-AF327468, AF327470, AF327472, AF327475-AF327477, AF327479-AF327481, AF327483, AF327485, AF327486, AF327489, AF327490, AF327492-AF327494, AF327496, AF327497, AF327499, AF327500-AF327502, AF327504-AF327507.

Received: 13 August 2022; Accepted: 14 February 2023

Published online: 20 February 2023

References

- Hacini, Z., Khedja, F., Habib, I., Kendour, Z. & Debba, Z. Evaluation of antibacterial and antioxidant activities of three types of benzoin resin. *Eur. J. Chem.* **9**, 408–411. <https://doi.org/10.5155/eurjchem.9.4.408-411.1755> (2018).
- Hu, P. *Study on the Quality of Benzoin Based on Coniferyl Benzoate and Pharmacokinetics* 1–90 (Chengdu Univ. TCM, 2016).
- Chen, Q. F., Chen, X. Z., Li, G. Y., Wang, C. & Zhang, G. L. Two new 2-phenylbenzofurans from the bark of *Styrax perkinsiae*. *Chin. J. Nat. Med.* **10**, 92–97. <https://doi.org/10.3724/SP.J.1009.2012.00092> (2012).
- Min, B. S., Cao, T. Q., Hung, T. M. & Kim, J. A. Inhibitory effect on no production of compounds from *Styrax obassia*. *Planta Med.* **81**, PX60. <https://doi.org/10.1055/s-0035-1556504> (2015).
- Burger, P. *et al.* New insights in the chemical composition of benzoin balsams. *Food Chem.* **210**, 613–622. <https://doi.org/10.1016/j.foodchem.2016.05.015> (2016).
- El-Razek, M. H. A. Triterpenes from *Styrax benzoin*. *Der Pharma Chemica.* **10**, 30–34 (2018).
- Kim, M. R. *et al.* Matrix metalloproteinase-1 inhibitor from the stem bark of *Styrax japonica* S. et Z. *Chem. Pharm. Bull.* **52**, 1466–1469 (2004).
- Kwon, J. H. *et al.* Triterpenoids and a sterol from the stem-bark of *Styrax japonica* and their protein tyrosine phosphatase 1B inhibitory activities. *Phytother. Res.* **22**, 1303–1306 (2008).
- Lee, S. J., Lee, J., Song, S. & Lim, K. T. Glycoprotein isolated from *Styrax japonica* Siebold *et al.* Zuccarini inhibits oxidative and pro-inflammatory responses in HCT116 colonic epithelial cells and dextran sulfate sodium-treated ICR mice. *Food Chem. Toxicol.* **87**, 12–22. <https://doi.org/10.1016/j.fct.2015.11.004> (2016).
- Gu, A. T., Zhang, Q. & Wang, F. A preliminary study on the active ingredient of *Styrax tonkinensis* to protect endothelial injury. *Gd Chem.* **44**, 11–12 (2017).
- IBCAS. STYRAX Linnaeus. *Flora China.* **15**, 253–263 (1996).
- Liu, Z. *Revision of the Morphological Taxonomy of the Genus Styrax L. (Styracaceae) in China* 24–132 (Nanjing Forest. Univ, 2019). <https://doi.org/10.27242/d.cnki.gnjlu.2019.000265>.
- Su, J. Tang materia medica. *Anhui Sci. Tech.* **2**, 193 (2004).
- Fang, X. L. The Book of Jin. *Zhonghua Book Publi. Press* **95**, 4922–4923 (2008).
- Ji, S. Y. Formulaology. *China TCM. Press* **2**, 242 (2018).
- Chen, S. W. Tàimín Huimín and Pharmaceutical Bureau Formula. *Shanghai People's Publi. Press* **1**, 106–107 (2005).
- National Medical Products Administration. *Notice on Promulgating Quality Standards for 43 Kinds of Imported Medicinal Materials, Including Ercha*. 13. <https://www.nmpa.gov.cn/xxgk/fgwj/gzwwj/gzwwjyp/20040508010101798.html> (2004).
- EDQM Council of Europe. *European Pharmacopoeia EP10.0* 1341–1342 (2008).
- Indian Pharmacopoeia Commission. *Indian Pharmacopoeia* 888–889 (2010).
- Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China. China Med. Sci. Press* **1**, 154–155 (2020).
- Kharidia, S. The United States Pharmacopoeial. NF38. <https://online.uspnf.com/uspnf/> (2020).
- Bedigian, D. Monograph on Benzoin (Balsamic resin from *Styrax* species). *Econ. Bot.* **57**, 427–428 (2003).
- Medicines & healthcare products regulatory agency. *British Pharmacopoeia*, Vol. IV. TSO. 46–48. www.Pharmacopoeia.co.uk (2013).
- Sabbah, D. A. *et al.* Benzoin schiff bases: Design, synthesis, and biological evaluation as potential antitumor agents. *Med. Chem.* **14**, 695–708. <https://doi.org/10.2174/1573406414666180412160142> (2018).
- Scardamaglia, L., Nixon, R. & Fewings, J. Compound tincture of benzoin: A common contact allergen?. *Australas. J. Dermatol.* **44**, 180–184. <https://doi.org/10.1046/j.1440-0960.2003.00674.x> (2003).

26. Castel, C. *et al.* Volatile constituents of benzoin gums: Siam and Sumatra, part 2. Study of headspace sampling methods. *Flavour Fragr. J.* **21**, 59–67. <https://doi.org/10.1002/ffj.1502> (2006).
27. Abdulmumeen, H. A., Risikat, A. N. & Sururah, A. R. Food: Its preservatives, additives, and applications. *Int. J. Chem. Biol. Sci.* **1**, 36–47. <https://doi.org/10.13140/2.1.1623.5208> (2012).
28. Joint FAO/WHO Expert Committee on Food Additives. *Evaluation of Certain Food Additives and Contaminants: Seventy-Fourth Report of the Joint FAO/WHO Expert Committee on Food Additives. WHO Technical Report Series 966* (2011).
29. Langenheim, J. H. *Plant Resins: Chemistry, Evolution Ecology and Ethnobotany* 586 (Timber Press, 2003).
30. Murthy, H. N. *Gums, Resins and Latexes of Plant Origin, Reference Series in Phytochemistry* 559–579 (Springer, 2022).
31. Harada, K., Wiyono, X. X. & Munthe, L. Production and commercialization of benzoin resin: Exploring the value of benzoin resin for local livelihoods in North Sumatra, Indonesia. *Trees For. People.* **7**, 100174 (2022).
32. Sun, Y. Y., Feng, J., Wu, M. S. & Liu, Y. Y. Herbal Textual research on benzoinum in famous classical formulas. *Chin. J. Exp. Tradit. Med. Formulae* <https://doi.org/10.13422/j.cnki.syfjx.20220456> (2022).
33. Mao, D. *et al.* Authenticity assessment of (E)-cinnamic acid, vanillin, and benzoic acid from Sumatra benzoin balsam by gas chromatography combustion/pyrolysis isotope ratio mass spectrometry. *Int. J. Anal. Chem.* <https://doi.org/10.1155/2022/1747053> (2022).
34. Fernandez, X. *et al.* Volatile constituents of benzoin gums: Siam and Sumatra, part 3. Fast characterization with an electronic nose. *Flavour Fragr. J.* **21**, 439–446. <https://doi.org/10.1002/ffj.1675> (2006).
35. Gregory, T. R. DNA barcoding does not compete with taxonomy. *Nature* **434**, 1067 (2005).
36. Michel, C. I., Meyer, R. S., Taveras, Y. & Molina, J. The nuclear internal transcribed spacer (ITS2) as a practical plant DNA barcode for herbal medicines. *J. Appl. Res. Med. Aromat. Plants* **3**, 94–100. <https://doi.org/10.1016/j.jarmap.2016.02.002> (2016).
37. Veldman, S. *et al.* DNA barcoding augments conventional methods for identification of medicinal plant species traded at Tanzanian markets. *J. Ethnopharmacol.* **250**, 112495. <https://doi.org/10.1016/j.jep.2019.112495> (2020).
38. Chen, S. L. *et al.* Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS ONE* **5**, e8613. <https://doi.org/10.1371/journal.pone.0008613> (2010).
39. Fritsch, P. W. Multiple geographic origins of Antillean *Styrax*. *Syst. Bot.* **28**, 421–430. <https://doi.org/10.1043/0363-6445-28.2.421> (2003).
40. Tong, L. L. *et al.* The complete chloroplast genome of *Styrax faberi* Perk. (Styracaceae). *Mitochondrial DNA Part B Resour.* **5**, 98–99. <https://doi.org/10.1080/23802359.2019.1698351> (2020).
41. Cai, X. L. *et al.* Plastome structure and phylogenetic relationships of Styracaceae (Ericales). *BMC Ecol. Evol.* **21**, 103. <https://doi.org/10.1186/s12862-021-01827-4> (2021).
42. Song, Y. *et al.* Chloroplast genome evolution and species identification of *Styrax* (Styracaceae). *BioMed. Res. Int.* <https://doi.org/10.1155/2022/5364094> (2022).
43. Lobdell, M. S. & Shearer, K. Genome sizes, ploidy levels, and base compositions of *Styrax* species and cultivars. *HortScience* **57**, 478–484. <https://doi.org/10.21273/HORTSCI16294-21> (2022).
44. Valderrama, N. D. L., Schaeffer, P., Leprince, A., Schmitt, S. & Adam, P. Novel oxygenated fossil nor-diterpenoids from Cretaceous amber (South-Western France) as potential markers from Cupressaceae and/or Cheirolepidiaceae. *Org. Geochem.* **167**, 104372. <https://doi.org/10.1016/j.orggeochem.2022.104372> (2022).
45. Gomez, B., Martínez-Delclos, X., Bamford, M. & Philippe, M. Taphonomy and palaeoecology of plant remains from the oldest African Early Cretaceous amber locality. *Lethaia* **35**, 300–308. <https://doi.org/10.1111/j.1502-3931.2002.tb00090.x> (2002).
46. Zhang, Z. Q. Spectral characteristics of amber—their application in provenance determination, and study on fluorescent components. *China Univ. Geosci.* <https://doi.org/10.27492/d.cnki.gzdz.2021.000248> (2021).
47. Schultz, J. *et al.* A common core of secondary structure of the internal transcribed spacer 2 (ITS2) throughout the Eukaryota. *RNA* **11**, 361–364. <https://doi.org/10.1261/rna.7204505> (2005).
48. Wolf, M. *et al.* Homology modeling revealed more than 20,000 rRNA internal transcribed spacer 2 (ITS2) secondary structures. *RNA* **11**, 1616–1623. <https://doi.org/10.1261/rna.2144205> (2005).
49. Ankenbrand, M. J. *et al.* ITS2 database V: Twice as much. *Mol. Biol. Evol.* **32**, 3030–3032. <https://doi.org/10.1093/molbev/msv174> (2015).
50. Huang, L. Q. & Yuan, Y. Molecular identification of traditional Chinese medicine (TCM). *Shanghai Sci. Tech. Publ.* **1**, 3–20 (2014).
51. Ma, S. J. *Molecular Identification of Herbal Medicines: From Barcode to Genomes* 1–9 (Chinese Academy of Medical Sciences & Peking Union Medical College, 2017).
52. Li, Q. L. *et al.* Four new benzofurans from seeds of *Styrax perkinsiae*. *Planta Med.* **71**, 847–851. <https://doi.org/10.1055/s-2005-871226> (2005).
53. Park, S. Y. *et al.* Benzofurans from the seeds of *Styrax obassia*. *Bull. Korean Chem. Soc.* **28**, 1874–1876 (2007).
54. Son, N. T. *et al.* Genus *Styrax*: A resource of bioactive compounds. *Stud. Nat. Prod. Chem.* **69**, 299–347. <https://doi.org/10.1016/B978-0-12-819487-4.00008-2> (2021).
55. Schwörer, C., Leunda, M., Alvarez, N., Gugerli, F. & Sperisen, C. The untapped potential of macrofossils in ancient plant DNA research. *New Phytol.* **235**, 391–401. <https://doi.org/10.1111/nph.18108> (2022).
56. Asif, M. J. & Cannon, C. H. DNA extraction from processed wood A case study for the identification of an endangered timber species (*Gonystylus bancanus*). *Plant Mol. Biol. Rep.* **23**, 185–192. <https://doi.org/10.1007/BF02772709> (2005).
57. Liepelt, S. *et al.* Authenticated DNA from ancient wood remains. *Ann. Bot.* **98**, 1107–1111. <https://doi.org/10.1093/aob/mcl188> (2006).
58. Tnah, L. H. *et al.* DNA extraction from dry wood of *Neobalanocarpus heimii* (Dipterocarpaceae) for forensic DNA profiling and timber tracking. *Wood Sci. Technol.* **46**, 813–825. <https://doi.org/10.1007/s00226-011-0447-6> (2012).
59. Wang, J. *et al.* DNA barcodes identification of one rare traditional Chinese medicine *Draconis sanguis*. *Chin. Pharm. J.* **50**, 1261–1265. <https://doi.org/10.11669/cpj.2015.15.001> (2015).
60. Jiao, L. C., Yin, Y. F., Cheng, Y. M. & Jiang, X. M. DNA barcoding for identification of the endangered species *Aquilaria sinensis*: Comparison of data from heated or aged wood samples. *Holzforchung* **68**, 487–497. <https://doi.org/10.1515/hf-2013-0129> (2014).
61. Nishizawa, T. & Watano, Y. Primer pairs suitable for PCR-SSCP analysis of chloroplast DNA in angiosperms. *J. Phytogeogr. Taxon* **48**, 63–66 (2000).
62. Chung, S., Decker-Walters, D. S. & Staub, J. E. Genetic relationships within the Cucurbitaceae as assessed by consensus chloroplast simple sequence repeats (ccSSR) marker and sequence analyses. *Can. J. Bot.* **81**, 814–832. <https://doi.org/10.1139/B03-074> (2003).
63. Hollingsworth, M. L. *et al.* Selecting barcoding loci for plants: Evaluation of seven candidate loci with species-level sampling in three divergent groups of land plants. *Mol. Ecol. Resour.* **9**, 439–457. <https://doi.org/10.1111/j.1755-0998.2008.02439.x> (2009).
64. Keller, A. *et al.* 5.8S–28S rRNA interaction and HMM-based ITS2 annotation. *Gene* **430**, 50–57. <https://doi.org/10.1016/j.gene.2008.10.012> (2009).
65. Koetschan, C. *et al.* ITS2 database IV: Interactive taxon sampling for internal transcribed spacer 2 based phylogenies. *Mol. Phylogenet. Evol.* **63**, 585–588. <https://doi.org/10.1016/j.ympev.2012.01.026> (2012).

Acknowledgements

This work was supported by grants from the project of major increase and decrease of expenditure at the central level of the Chinese Academy of Chinese Medical Sciences [Grant No. 2060302-1906-10], and the project of medical and health science and technology innovation engineering of Chinese Academy of Medical Sciences [Grant No. 2021-1-I2M-032], and the Major Science and Technology Project of Hainan Province [Grant No. ZDKJ2021034].

Author contributions

J.F. and Y.Y.L. designed the work and wrote the paper. Q.Q.H. collection and processing of samples. A.Z.X. was involved in the extraction of sample DNA. Y.Y.L. provided the idea for this study. All authors read and approved the final manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-30034-y>.

Correspondence and requests for materials should be addressed to Y.L.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023