Successful identification of the species of the semipetrified amber medicinal resin benzoin using DNA barcoding technology

Jian Feng
Chinese Academy of Medical Sciences and Peking Union Medical College
Qingqin He
Chinese Academy of Medical Sciences and Peking Union Medical College
Anzhen Xie
Chinese Academy of Medical Sciences and Peking Union Medical College
Yangyang Liu (yyliu@implad.ac.cn)
Chinese Academy of Medical Sciences and Peking Union Medical College

Article

Keywords:

Posted Date: August 24th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1959088/v1

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at Scientific Reports on February 20th, 2023. See the published version at https://doi.org/10.1038/s41598-023-30034-y.

Abstract

Benzoin, also known as semipetrified amber, is the resin secreted from the trunk of the benzoin plant. This medicinal material has blood circulation-promoting and pain-relieving properties. There are approximately 130 species of benzoin plants in the world, and the resin of many plants in the benzoin genus is used, so there is confounding of the species of origin. In this study, DNA was successfully extracted from the semipetrified amber medicinal material benzoin, and the species of benzoin in the medicinal herb market were evaluated by DNA barcoding. BLAST comparison of ITS2 primary sequences and homology prediction analysis of ITS2 secondary structures showed that the commercially available benzoin base originated from Styrax tonkinensis (Pierre) Craib ex Hart. and Styrax japonicus Sieb. et Zucc. of the genus Styrax Linn., and tissue from other genera was also detected to be mixed in with some samples, accounting for 29.6% of the samples. There is a risk that benzoin resin medical materials will be mixed with tissues from other plants. This is the first report on the subject, and the approach taken provides technical feasibility for the identification of the basis of the benzoin resin medical materials.

Introduction

Benzoin is the resin secreted from the trunk of the benzoin plant and is used as a medicine and spice[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 has anti-inflammatory, antipyretic, antitumor, analgesic and estrogen-promoting effects, protects against cerebral hypoxia, and promotes blood–brain barrier permeability, among other properties[7-10]. There are approximately 130 species of the genus *Styrax* Linn., both trees and shrubs, distributed worldwide, mainly in Asia (China), Southeast Asia (Vietnam, Thailand, Sumatra, Indonesia) and North America[11,12].

In ancient China during the Tang Dynasty, benzoin was recorded as a spice in the *Book of Jìn* and as a medicine in the *Tang Materia Medica* (*Xīnxiū Běn Cǎo)*[13,14]. It has a history of more than 700 years of medicinal use, and was used in approximately 93 prescriptions, such as in *zhìbǎo boluse*s and *storax pill* in the *Tàimín Huìmín and Pharmaceutical Bureau Formula* and in *dàhuóluò boluses* in *Sages' Salvation Records*[15,16]. Benzoin has well-established uses in traditional forms of medicine. Several national pharmacopoeias, including China, United States, Europe, and India, describe the specifications and tests for benzoin. The Chinese medicinal records benzoin as being 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 being 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 tincture[22]. The British Pharmacopoeia specifies the use of Sumatra benzoin in Benzoin Inhalation and Compound Benzoin Tincture[23]. The United States Pharmacopoeia 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 powders[24]. It is likewise used as an additive in aromatherapy[25]. 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 safety[28]. 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[5,29]. Thus, the chemical composition of the different sources of benzoin determines their organoleptic characteristics and also their uses. There are both similarities and differences in composition, which means that they are both used in fragrances and flavors. It has been reported that 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 extracts[30]. Therefore, there is a need to differentiate between the origin of the species and the traditional use of benzoin.

The resin of various plants in the genus *Styrax* Linn. have been used as "benzoin", and thus the confounding of the species is obvious. 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. Benzoin-producing plants are mainly wild and not managed through standardized cultivation, which are usually selected for benzoin production in North Sumatra, Indonesia. The benzoin resin production process involves first removing shrubs and parasitic plants from around the benzoin tree , then taping 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) and classified and sold according to quality[31]. Therefore, there is a risk of adulteration and counterfeiting during the harvesting process, either artificially or nonartificially. The confounding of the benzoin-producing species is evident[32], and thus it becomes necessary to study different identification methods.

In recent years, research on benzoin quality has mainly been based on the development of chemical-based volatile component analysis and aromatic lipids content determination methods, such as gas chromatography mass spectrometry, solid phase microextraction, headspace sampling, and high-performance liquid chromatography frit fast atombombardment mass spectrometry[33-34]. In addition, the existing quality standards for medicinal materials record benzoin quality control indicators, including identification, total ash, loss-on-drying, alcohol-insoluble matter and benzoic acid content. These routine indicators are simple to determine and can identify the authenticity of benzoin and evaluate its quality, but they cannot identify the species of benzoin.

DNA barcoding has received much 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 can be selected to accurately identify most species of the entire genus and family, and it has high repeatability and stability[35-37]. Chen et al. (2010) 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 barcodes[38]. However, at present, molecular biology studies on benzoin are mainly focused on chromosome size, ploidy, chloroplast genome, phylogeny, and geographic evolution[39-42].

Previous reports have not addressed the species identification of benzoin resinous medicinal materials. The key to species identification using DNA barcoding is to extract DNA from the target. Benzoin is a resinous medicinal material secreted by injured *Styrax* trees and theoretically does not contain DNA. 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 or entrained tissue from other plants. This is similar to the process of amber formation, but the resin is not completely lithified, so it can be called semipetrified amber[43-45]. However, these residues stick to the resin, and the DNA is not easily extracted. Therefore, it is key to explore the DNA extracted from these residues and then use ITS2 as the DNA barcode to identify the species. This study provides a new way to solve the problem of identifying species producing the semipetrified amber benzoin medicinal material by obtaining original information on these residues. This is the first report that the correlation of different information can help determine the origin of benzoin.

Results

DNA concentration, sequence analysis and species identification. We anticipated extracting DNA directly 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 DNA barcode identification, which has the advantage of preventing sequencing failure in subsequent trials due to sample mixing (Fig. 1).

In total, there were 27 batches of benzoin samples containing 40 individual bark-like residues in this study. The DNA concentrations obtained from these residues ranged from 8.0-203.8 ng/μL, and the DNA purity ranged from 1.70-2.08. Although the AX151 sample had the lowest DNA concentration of 8.0 ng/μL, 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, considering that the samples originated from bark-like residues in benzoin resin, whose DNA was degraded despite the DNA concentration and purity satisfying the subsequent experiments, we chose the nuclear gene ITS2 universal barcode as primers for amplification, and all the sample DNA was

successfully amplified and sequenced. The complete ITS2 was obtained by removing the 5.8S and 26S regions from the sequences obtained by sequencing, which had a length of 217-230 bp.

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

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

Forty individual samples were found by 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 of over 92% (Table 2). The BLAST and secondary structure analysis 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 BLAST and secondary analysis results of sample AX174 were different, but both indicated that the sample belongs to species of *Trema* Lour. 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.

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

Sample No.	BLAST results				Secondary structures results	
	Styrax Linn.	Other species	Per. Ident (%)	E value	*Styrax* Linn.	Other species
AX011	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX012	*S. japonicus* *S. tonkinensis*		99.55%	1e-107	*S. tonkinensis*	
AX013	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX014	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX021	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX031	*S. tonkinensis*		94.20%	1e-93	*S. tonkinensis*	
AX041	*S. tonkinensis*		96.88%	3e-100	*S. tonkinensis*	
AX051	*S. tonkinensis*		92.40%	6e-82	*S. tonkinensis*	
AX061	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX062	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX071		*Aquilaria sinensis* (Lour.) Spreng. (*Aquilaria* Lam.)	99.13%	2e-112		*A. sinensis*
AX072	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX081	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX091	*S. japonicus*		99.11%	4e-108	*S. japonicus*	
AX092	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX093	*S. japonicus*		99.55%	9e-110	*S. japonicus*	

AX101	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX102		*Dimocarpus longan* Lour. (*Dimocarpus* Lour.)				*D. longan*
AX111	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX121	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX122		*Ageratum conyzoides* L. (*Ageratum* L.)	99.13%	2e-112		*A. conyzoides*
AX131	*S. japonicus*		98.66%	5e-107	*S. japonicus*	
AX132	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX141	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX151		*A. sinensis*	100%	4e-113		*A. sinensis*
AX161	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX171		*Glycine soja* Siebold & Zucc. (*Glycine* Willd.)	97.70%	9e-100		*G. soja*
		Glycine max (L.) Merr. (*Glycine* Willd.)	97.70%	9e-100		*G. max*
AX172		*Acronychia pedunculata* (L.) Miq. (*Glycine* Willd.)	99.55%	1e-108		*A. pedunculata*
AX173	*S. japonicus*		99.11%	1e-108	*S. japonicus*	
AX174		*Trema cannabinum* var. dielsianum (Hand.Mazz.) C. J. Chen (*Trema* Lour.)	100%	1e-107		*Trema micranthum* (L.) Blume
		Trema nitidum C. J. Chen	100%	1e-107		

		(*Trema* Lour.)				
		Trema orientale (L.) Blume (*Trema* Lour.)	100%	1e-107		
		Trema tomentosa (Roxb.) H.Hara (*Trema* Lour.)	100%	1e-107		
AX181	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX191	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX201		*Acronychia pedunculata* (L.) Miq. (*Acronychia* J. R. Forst. & G. Forst.)	99.10%	1e-107		*A. pedunculata*
AX211	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	
AX221	*S. tonkinensis*		96.43%	3e-99	*S. tonkinensis*	
AX231		*Musa acuminata* Colla *(Musa* L.)	96.33%	1e-97		*M. acuminata*
AX241	*S. tonkinensis*		97.32%	2e-101	*S. tonkinensis*	
AX251		*D. longan*	100%	1e-107		*D. longan*
AX261	*S. japonicus*		99.55%	9e-110	*S. japonicus*	
AX271	*S. tonkinensis*		96.88%	9e-100	*S. tonkinensis*	

We used the 14 samples identified by BLAST as *S. japonicus* as group 1 and the 16 samples identified as *S. tonkinensis* as 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, it was also found that there were different base sites 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 base sites, respectively. In addition, there are two base sites in group 2 that are identical to those of *S. japonicus* (base sites 32 and 209), and there are also two base variable sites (base sites 60 and 185).

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. Grpoup1 is the 14 samples identified as *S. japonicus* and Grpoup2 is the 16 samples identified as *S. tonkinensis*. Bolded black indicates differential loci for *S. japonicus* and *S. tonkinensis*.

Species	Base site											
	7	13	24	32	33	46	60	75	77	78	84	102
S. japonicus[1)]	A	**T**	**A**	T	T	**T**	A	**T**	T	**A**	T	**A**
Group1	A	**T**	**A**	T	T	**T**	A	**T**	T	**A**	C	**A**
S. tonkinensis[2)]	A	**C**	**C**	C	C/T	**C**	A	**C**	T/A/C	**G**	C	**G/A**
Group2	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	
S. japonicus[1)]	**G**	C	A	T	**G**	T	A	T	T	A	**T**	
Group1	**G**	C	A	T	**G**	T	A	T	T	A	**T**	
S. tonkinensis[2)]	**A**	T/C	G/A	T/C	**A**	G	A/G	T/C	T/C	G/A	**C**	
Group2	**A**	C	A	C	**A**	G/C	A	C	C	A	**C**	

ITS2 secondary structure analysis

The ITS2 database was used for homology prediction analysis 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 being the longest, 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 of 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 being more variable due to multiple unpaired bases and positional differences. In contrast, the shapes of helices II and IV, although basically the same, also have unpaired bases, leading to differences. *S. tonkinensis* has three slightly different secondary structures, as shown in Fig. 2A a), b), and c). The red dashed boxes are 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 these differences are due to differences in the position of helix III. *S. tonkinensis* has three ITS2 secondary structures with different base sites, such as bases A, C, and T at position 32 of helix I (black arrows). In addition, there are hCBCs in helix III (109/179: C-G→T-G, red arrows). We compared the DNA sequences of the two sets of basal samples obtained by BLAST identification. The ITS2 secondary structure sequence of group 1 was identical to that of *S. japonicus*, with the presence of hCBCs (73/84: G-C→G-T) only on helix II (Table 4). The results of the predicted secondary structure of ITS2 for group 2 homologs were the same as those of *S. tonkinensis*, but the differences in the position of helix III were clear, and there were also variations in base sites. For example, sample AX051 differed from the other 15 samples by the secondary structure at base 60 on the main loop (A) and the presence of hCBCs on helix III (103-185: C-C→C-G), but the rest of the sites were identical. Compared with *S. tonkinensis* a) and *S. tonkinensis* b), the bases at sites 7 and 60 of the main loop of group 2 are G, the bases at site 32 of the top asymmetric loop of 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. The samples compared to *S. tonkinensis* c) 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 of *S. tonkinensis* and *S. japonicus* differ significantly, and the species can be further precisely identified based on this information. In addition, the ITS2 secondary structures of the remaining 46 species of the genus *Styrax* Linn. are shown in Figure 2B. The size, position, and angle of the four helices and the bases on 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 homologous folding, but unfortunately, the ITS2 secondary structures of the four groups of species were also identical.

Table 4. Comparison of base differences on the helices of *S. tonkinensis* and *S. japonicus* with the samples. Grpoup1 was the 14 samples identified as *S. japonicus*, and Grpoup2 was the 16 samples identified as S*. tonkinensis* by BLAST. -: represents no differences in base sequences.

Species	Helix I	Helix II	Helix III	Helix IV	Main loop
S. japonicus	-	hCBCs (73/84: G-C→G-T)	-	-	-
Group1					
S. tonkinensis 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
S. tonkinensis b)					
Group2					
S. tonkinensis 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
Group2					

The species of the 27 batches of benzoin samples collected were inferred from a comprehensive BLAST analysis and by analysis of 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 does not mean that they 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.

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.

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

Discussion

The core of Chinese medicine identification is to conduct variety authenticity and quality merit evaluation, which are related to drug safety and the healthy development of the Chinese medicine industry. Traditional identification, 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 imperfect[49]. In addition, microscopic identification, physicochemical identification, and other mainstream methods for the identification of Chinese medicinal materials have some limitations in the identification of species with multiple primordia[50]. The molecular identification method, represented by DNA barcoding, can accurately identify and characterize species and is a simple and precise method for identification of Chinese medicinal materials with clear judgment criteria; it is thus suitable for the accurate identification of Chinese medicinal materials and their multisubstrate species without requiring professional knowledge of taxonomy or of ambiguous morphological characters[36].

Despite the original processing of the benzoin resin, bark or plant tissue fragments remain on the resin. However, most Chinese herbs and tablets undergo further processing and preparation steps, where DNA degradation is more severe, especially 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]. These resins are insoluble in water and increase the difficulty of DNA extraction. Therefore, the most critical step for DNA barcode identification is to extract higher quality DNA from the samples. In pre-experiments, the benzoin was directly ground to extract DNA (the purity of most of the DNA concentration was low and could not meet the experimental requirements), but the amplification results were unsatisfactory (most of the amplification failed, and some sequences were heterozygous). This could be due to low DNA yield because of less retention or resin inhibition. Therefore, we changed our approached 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 much of the substances that interfere with DNA release as possible (including other impurities). Previous scholars have mostly explored DNA extraction methods from dried wood (Lobeliaceae), archaeological wood remains, ancient wood (oak), endangered wood (prismatic wood), etc. 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 DNA extraction method improvement[58,59]. The present study is also the first to report the successful extraction of DNA from resin samples exuded from excised trees that can meet the requirements of subsequent experiments. Therefore, this study shows that it is necessary to pretreat resinous materials used for DNA extraction.

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 *mat*K, *rbc*L, *psb*A-*trn*H, *trn*L, *trn*K, the nuclear genes ITS, ITS2, and their corresponding barcode combinations[38,60–62]. However, for some samples with severe DNA degradation and fragmented DNA, using barcodes with longer amplification products as primers is unsuccessful, especially for medical materials such as resinous and stemmed wood, and selecting barcodes with moderate amplification lengths is essential for the study. ITS2, as a

DNA barcode, generally has a length of 200–300 bp, which is suitable for identification to the genus and Species level. 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 can 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 structure[65]. In this study, 40 individual bark-like residues from 27 batches of benzoin were analyzed by BLAST of their ITS2 primary structure and secondary structure analysis, and 30 were found to be from the benzoin genus, identified as *S. tonkinensis* and *S. japonicus*. Ten samples were identified as species of other genera. Using the GenBank database, we compared the ITS2 sequences of the two species identified as benzoin, and 9 base sites could be used as specific identification sites for these two species. This study also found that benzoin resinous medicinal materials may be mixed with other plant tissues during production and transportation. The contamination of commercially available benzoin samples was prominent, accounting for 29.6% of the samples, which was related to the collection and transportation of the resin at the original growing site. Overall, the ITS2 primary and secondary structure analysis reported in this study allowed for the preliminary identification of benzoin medicinal material species and potential contaminants.

Given the development of the molecular biology field, molecular biology techniques can be applied to the genetic identification of traditional Chinese medicine. We hope that, on the basis of this study, barcodes with high specificity for identifying the basal origin of all species of the genus *Styrax* Linn. will be developed.

Materials And Methods

Sample collection

Twenty-seven batches of benzoin medicinal materials were collected from medicinal material markets and drugstores in China. Those 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 is orange–yellow or yellowish white, and the material is fragrant, with a gritty feel when chewed. It is very brittle and can be broken or cracked by hand, and the resin contains bark residues inside. Two to four replicates were added during the experiment for batches with high bark residues. Therefore, a total of 40 individual bark-like residues from final experimental materials were obtained, and the details are shown in Table 6. All voucher specimens of bark-like residues are kept 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).

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

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

No.	Species	Genus	Collection site	Accession ID
1	*Styrax acuminatus* Pohl	*Styrax* Linn.	Genbank	KP793409
2	*Styrax agrestis* (Lour.) G. Don	*Styrax* Linn.	Genbank	AF327483
3	*Styrax americanus* Lam.	*Styrax* Linn.	Genbank	AF327472
4	*Styrax aureus* (Mart.) Miers	*Styrax* Linn.	Genbank	AY143577
5	*Styrax benzoin* Dryander	*Styrax* Linn.	Genbank	AF327494
6	*Styrax calvescens* Perk.	*Styrax* Linn.	Genbank	AF327468
7	*Styrax camporum* Pohl	*Styrax* Linn.	Genbank	AF327504
8	*Styrax chinensis* Hu et S. Y. Liang	*Styrax* Linn.	Genbank	AF327492
9	*Styrax confusus* Hemsl.	*Styrax* Linn.	Genbank	AF327470
10	*Styrax cordatus* A.D C.	*Styrax* Linn.	Genbank	AF327506
11	*Styrax dasyanthus* Perk.	*Styrax* Linn.	Genbank	AF327467
12	*Styrax faberi* Perk.	*Styrax* Linn.	Genbank	AF327485
13	*Styrax formosanus* Matsum.	*Styrax* Linn.	Genbank	AF327466
14	*Styrax gentryi* P.W. Fritsch	*Styrax* Linn.	Genbank	AF327502
15	*Styrax glabrescens* Benth.	*Styrax* Linn.	Genbank	AF327475
16	*Styrax glabrus* Sw.	*Styrax* Linn.	Genbank	AY143582
17	*Styrax grandifolius* Ait.	*Styrax* Linn.	Genbank	AF327476
18	*Styrax hemsleyanus* Diels	*Styrax* Linn.	Genbank	AF327477
19	*Styrax japonicus* Sieb. et Zucc.	*Styrax* Linn.	Genbank	AB114900; LC534305; LC600960; MF349070
20	*Styrax lanceolatus*P.W. Fritsch	*Styrax* Linn.	Genbank	AF327501
21	*Styrax latifolius* Pohl	*Styrax* Linn.	Genbank	AY143585
22	*Styrax leprosus* Hook. & Arn.	*Styrax* Linn.	Genbank	KP793405
23	*Styrax maninul* Hook. & Arn.	*Styrax* Linn.	Genbank	AY143576
24	*Styrax obassia* Siebold & Zucc.	*Styrax* Linn.	Genbank	AF327479

25	*Styrax obassis* Siebold et Zucc.	*Styrax* Linn.	Genbank	MF349081
26	*Styrax obtusifolius Griseb.*	*Styrax* Linn.	Genbank	AY143589
27	*Styrax ochraceus* Urb.	*Styrax* Linn.	Genbank	AY143580
28	*Styrax odoratissimus* Champ. ex Bentham	*Styrax* Linn.	Genbank	AF327460
29	*Styrax officinalis* L.	*Styrax* Linn.	Genbank	AF327489
30	*Styrax pentlandianus* J. Rémy	*Styrax* Linn.	Genbank	AF327507
31	*Styrax perkinsiae* Rehd.	*Styrax* Linn.	Genbank	AF327459
32	*Styrax peruvianus* Zahlbr.	*Styrax* Linn.	Genbank	KP793406
33	*Styrax platanifolius* Engelm. ex Torr.	*Styrax* Linn.	Genbank	AF327486
34	*Styrax pohlii* A. D C.	*Styrax* Linn.	Genbank	AY143581
35	*Styrax portoricensis* Krug & Urban	*Styrax* Linn.	Genbank	AF327505
36	*Styrax radians* P.W. Fritsch	*Styrax* Linn.	Genbank	AF327497
37	*Styrax ramirezii* Greenm.	*Styrax* Linn.	Genbank	AF327499
38	*Styrax redivivus* (Torr.) L.C. Wheeler	*Styrax* Linn.	Genbank	AF327490
39	*Styrax serrulatus* Roxb.	*Styrax* Linn.	Genbank	AF327481
40	*Styrax shiraianus* Makino	*Styrax* Linn.	Genbank	AF327480
41	*Styrax sieberi* Makino	*Styrax* Linn.	Genbank	AY143587
42	*Styrax subargenteus* Sleumer	*Styrax* Linn.	Genbank	KP793407
43	*Styrax suberifolius* Hook. et Arn.	*Styrax* Linn.	Genbank	AF327493
44	*Styrax tomentosus* Bonpl.	*Styrax* Linn.	Genbank	AY143590
45	*Styrax tonkinensis* (Pierre) Craib ex Hartw.	*Styrax* Linn.	Genbank	MF096782; MF096783; MF096784; MF096785; JF421547; MF096786;
46	*Styrax vilcabambae* (D.R. Simpson) B. Walln.	*Styrax* Linn.	Genbank	AF327496
47	*Styrax warscewiczii* Perkins	*Styrax* Linn.	Genbank	AF327500
48	*Styrax wilsonii* Rehd.	*Styrax* Linn.	Genbank	AF327462

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 two times for 5 min each time, and the residue was finally washed with sterile water and dried (Fig. 1). The sample was cut into pieces, placed in a 2 mL centrifuge tube, and 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 μL of buffer CPL and 10 μL of β-mercaptoethanol were added, and the rest of the steps were performed according to the kit 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 ℃ refrigerator for later use. Amplification was performed using the primer ITS2 (ITS2-F: ATGCGATACTTGGTGTGAAT; ITS2-R: GACGCTTCTCCAGACTACA

AT)[39]. The PCR amplification program was as follows: predenaturation at 94 ℃ for 5 min; 40 cycles of denaturation at 94 ℃ for 30 s, annealing at 56 ℃ for 30 s, extension at 72 ℃ for 45 s, and extension at 72 ℃ for 10 min; and storage at -20 ℃. 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 sequence[63]. The complete ITS2 sequence obtained by sequencing was used for 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 by the ITS2 database homology model[64]. And hemi-compensatory base changes (hCBCs, such as C-G→C-A or T-T→T-C) were calculated.

Declarations

Data availability

All data generated during this study are included in this published article and its supplementary information files.

Acknowledgments

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 number 2060302-1906-10], and

the project of medical and health science and technology innovation engineering of Chinese Academy of Medical Sciences [grant number 2021-1-I2M-032], and the Major Science and Technology Project of Hainan Province [grant number 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 that they have no known competing financial interests or personal relationships.

Additional information

Supplementary Information. Raw sequence and secondary structure results of the 40 individual bark-like residues in this study.

References

1. Leffemberg, C., Gautier, A. & Ohleyer, E. Enzymatic preparation of coniferaldehyde from coniferyl benzoate ex. Siam benzoin. *Appl. biochem. biotechnol.* **37,** 43-52. http://doi.10.1007/BF02788856 (1992).
2. Hu, P. Study on the quality of benzoin based on coniferyl benzoate and pharmacokinetics. Chengdu Univ. TCM. 1-90 (2016).
3. 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).
4. 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. 10.1055/s-0035-1556504 (2015).
5. 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).
6. El-Razek, M.H.A. Triterpenes from *Styrax benzoin*. *Der Pharma Chemica.* **10,** 30-34 (2018).
7. 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).
8. 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).
9. 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).

10. 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).
11. IBCAS. STYRAX Linnaeus. *Flora of China.* **15,** 253-263 (1996).
12. Liu, Z. Revision of the morphological taxonomy of the genus *Styrax* L. (Styracaceae) in China. *Nanjing Forest. Univ.* 24-132. https://doi.10.27242/d.cnki.gnjlu.2019.000265 (2019).
13. Su, J. Tang materia medica. *Anhui Sci. Tech.* **1,** 338 (1981).
14. Fang, X.L. The Book of Jìn. *Zhonghua Book Publi. Press* **95,** 4922-4923 (2008).
15. Zhao, J.Sages' Salvation Records. *People's Sanit. Publi. Press* **1,** 245 (1982).
16. Chen, S.W. Tàimín Huìmín and Pharmaceutical Bureau Formula. *Shanghai People's Publi. Press* **1,** 106-107 (2005).
17. 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/gzwj/gzwjyp/20040508010101798.html (2004).
18. EDQM Council of Europe. European Pharmacopoeia EP10.0. 1341-1342 (2008).
19. Indian Pharmacopoeia Commission. Indian Pharmacopoeia. 888-889 (2010).
20. Chinese Pharmacopoeia Commission. Pharmacopoeia of the people´s republic of China. *China Med. Sci. Press* **I,** 154-155 (2020).
21. Kharidia, S. The United States Pharmacopeial. NF38. https://online.uspnf.com/uspnf/ (2020).
22. Bedigian, D. Monograph on Benzoin (Balsamic resin from Styrax species). *Economic Botany* **57,** 427-428 (2003).
23. Anon. British Pharmacopoeia. Her Majesty's Stationery Office, London **2,** 791 (1993).
24. Anon. Pharmacopoea Helvetica, 7th edition. Benzoe Tonkinensis = Benjoin du Laos. Departamente Federal de l'Interieur, Berne, Switzerland 2 (1995).
25. 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.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.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. *Portland, Cambridge, Timber Press* 586 (2003).
30. Coppen, J.JW. Benzoin: production, uses and international trade. *Perfumer & Flavorist.* **24,** 11-24 (1999).

31. Harada, K., Wiyono & 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. Experim. Tradit. Med. Formulae*, https://doi.org/10.13422/j.cnki.syfjx.20220456 (2022).
33. Pastorova, I., de Koster, C.G., Boon, J.J. Analytical study of free and ester bound benzoic and cinnamic acids of Gum Benzoin resins by GC-MS and HPLC-frit FA-MS. *Phytochem. Anal.* **8,** 63-73 (1997).
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.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. http://dx.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.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.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.* 5364094. 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. Anderson, K.B., Winans, R.E. & Botto, R.E. The nature and fate of natural resins in the geosphere II. Identification, classification and nomenclature of resinites. *Org. Geochem.* **18,** 829-841. https://doi.org/10.1016/0146-6380(92)90051-X (1992).
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.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.* 2-5.

https://doi.10.27492/d.cnki.gzdzu.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. http://www.rnajournal.org/cgi/doi/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.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. *Chinese academy of medical sciences & Peking union medical college* 1-9 (2017).
52. Li, Q.L. *et al*. Four new benzofurans from seeds of *Styrax perkinsiae*. *Planta Med.* **71,** 847-851. https://doi.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. Dumolin-Lapègue, S. *et al*. Amplification of oak DNA from ancient and modern wood. *Mol. Ecol.* **8,** 2137-2140. http:// doi.10.1046/j.1365-294x.1999.00788.x (1999).
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. http://doi.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. Tech.* **46,** 813-825. http://doi.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. http:// doi.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. *Holzforschung* **68,** 487-497. http://doi.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. http://ci.nii.ac.jp/naid/10011443152 (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. http://doi.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).
66. Prasad, P.K. *et al*. Phylogenetic reconstruction using secondary structures and sequence motifs of ITS2 rDNA of *Paragonimus westermani* (Kerbert, 1878) Braun, 1899 (Digenea: Paragonimidae) and related species. *BMC Genomics* **10,** S25. https://doi.10.1186/1471-2164-10-S3-S25 (2009).

Figures

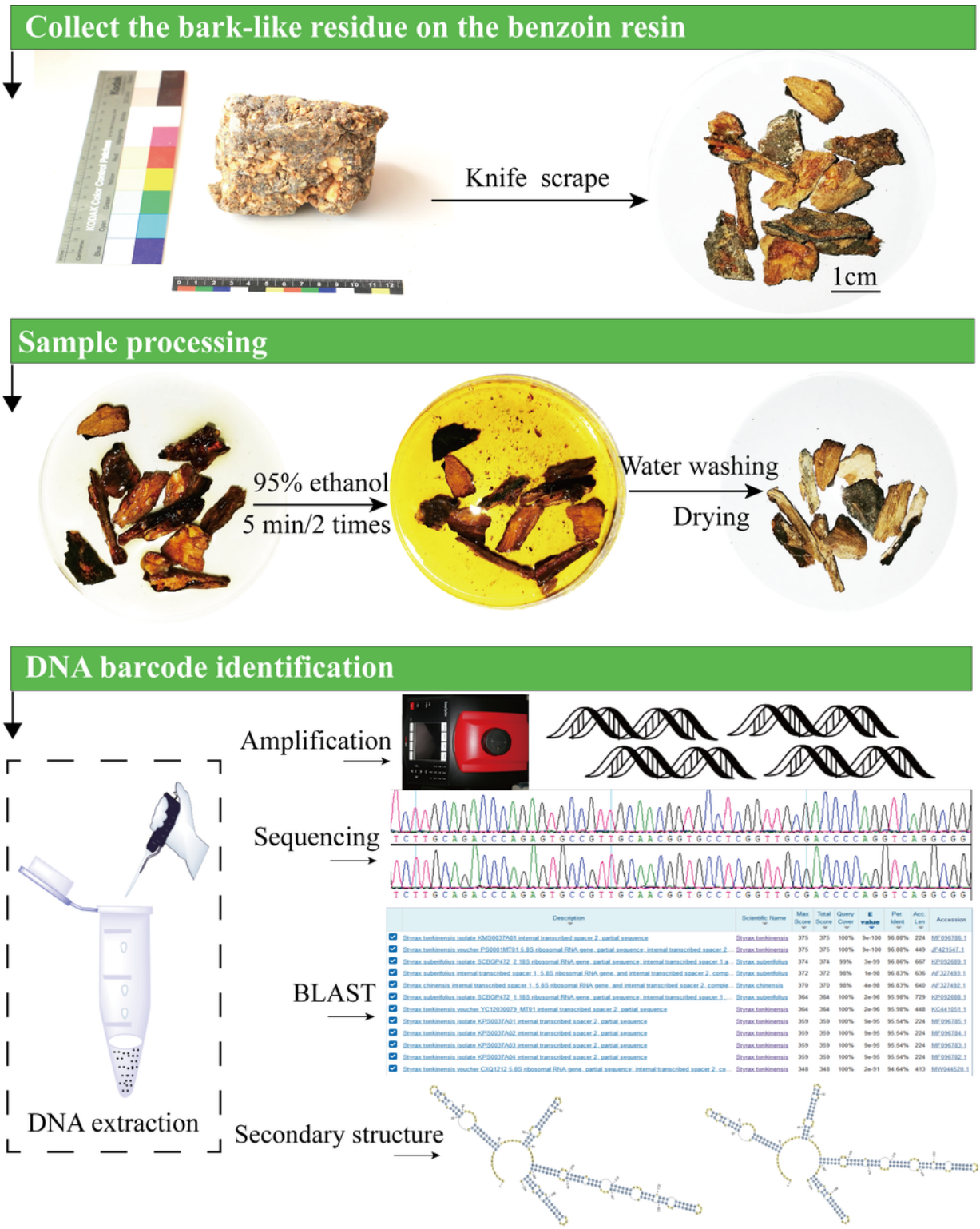

Figure 1

Process of pretreatment and DNA barcode identification of bark-like residues on benzoin.

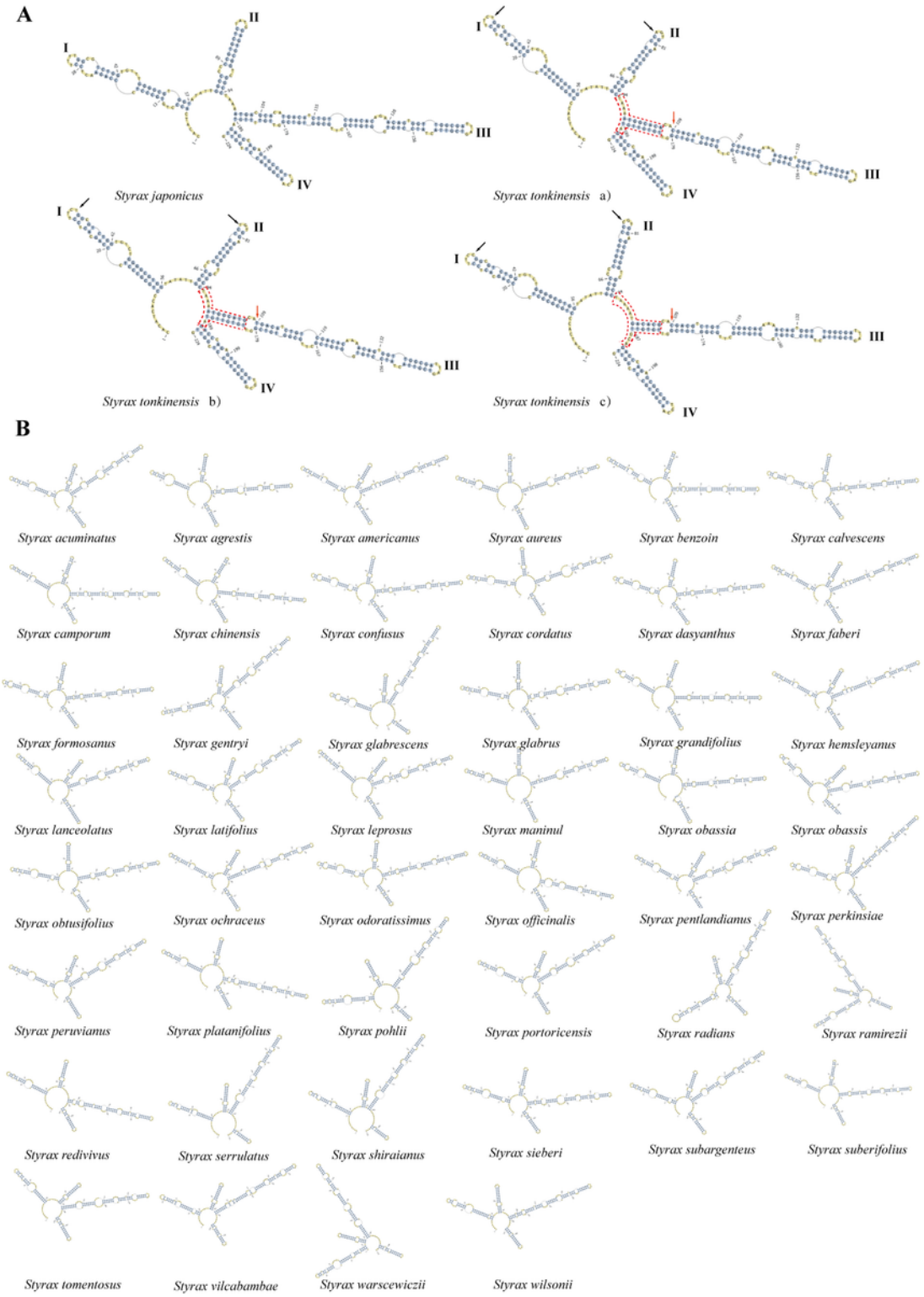

Figure 2

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

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementaryinformation.docx