Application of molecular markers in the study of genetic diversity of ethnomedicinal plants: A systematic mapping of genotype-to-culture intersections

Hongyu Che, Danning Pei, Yijin Liu, Lihao Guo, Yujie Liu, Shiying Chen, Haijun Li, Lin Guo, Yu Shi, Bo Liu

Correspondence

Hongyu Che[1], Danning Pei[1], Yijin Liu[†], Lihao Guo[†], Yujie Liu[†], Shiying Chen[2], Haijun Li[2], Lin Guo[3], Yu Shi[1], Bo Liu[1*]

[1]College of Life and Environmental Sciences, Minzu University of China, Beijing, China.
[2]College of Ethnology and Sociology, Minzu University of China, Beijing, China.
[3]Lomonosov Moscow State University, Anuchin Research Institute and Museum of Anthropology, Mokhovaya st., 11, Moscow, Russia.

*Corresponding Author: boliu@muc.edu.cn

Ethnobotany Research and Applications 35:40 (2026) - http://dx.doi.org/10.32859/era.35.40.1-26
Manuscript received: 20/05/2026 - Revised manuscript received: 15/08/2026 - Published: 16/08/2026

Review

Abstract

Background: Ethnomedicinal plants are biocultural resources whose identities, collection places, preparation practices, and therapeutic meanings are maintained within traditional knowledge systems. This review evaluates how molecular markers complement taxonomic, chemical, historical, and ethnographic evidence in authenticating materials, assessing genetic diversity, tracing place-based resources, and safeguarding medicinal knowledge.

Methods: We conducted a systematic evidence-mapping review of literature published from January 1980 to May 2026. CNKI, Wanfang Data, and PubMed were used for reproducible quantitative retrieval; Web of Science Core Collection and Google Scholar supported citation tracing and cross-checking. The counted databases yielded 4527 records. After removing 718 duplicates, 3809 records were screened, 1892 were retained at title/abstract stage, 61 full texts were assessed, and 55 sources entered qualitative synthesis.

Results: Molecular approaches progressed from fragment markers and standardized barcodes to genome-wide SNPs, organelle genomes, pangenomes, ancient DNA, and metabarcoding. Applications expanded from authentication to provenance, germplasm stewardship, functional variation, and market surveillance. Their ethnobotanical value was greatest when molecular data were linked to vouchers, vernacular names, preparation histories, chemical profiles, archaeological contexts, and knowledge holders.

Conclusions: Molecular markers provide reproducible evidence for plant identity and diversity, but do not independently establish therapeutic efficacy, cultural meaning, or historical use. Culturally annotated reference libraries, community-engaged interpretation, functional markers, and integrated genomic-metabolomic frameworks are priorities for biocultural conservation and transparent herbal supply chains.

Keywords: Molecular markers; Ethnomedicinal plants; Genetic diversity; DNA barcoding; DNA metabarcoding; Ancient DNA; Biocultural conservation

Background

Ethnomedicinal plants are embedded in relationships among plants, people, places, languages, and healing practices. A medicinal taxon may be recognized through a combination of morphology, habitat, harvesting season, plant part, processing method, vernacular name, and locally transmitted therapeutic experience. These culturally grounded categories do not always correspond one-to-one with Linnaean species, yet they are central to the continuity of traditional medicine and to the safety of present-day herbal supply chains. Botanical substitution, homonymous vernacular names, closely related taxa, and the loss of diagnostic characters after slicing, drying, or pulverization can therefore create both biological and cultural misidentification risks (Flora of China Editorial Committee 1993, Tineo *et al.* 2023).

During the past five decades, molecular markers have become an increasingly important complementary evidence system. RFLP was followed by PCR-based RAPD, ISSR, and AFLP markers, locus-specific SSRs and SCARs, standardized DNA barcodes, and more recently high-density SNP genotyping, RAD-seq, GBS, complete organelle genomes, WGS, and pangenomics. These approaches differ in cost, resolution, DNA-quality requirements, and interpretability. Their ethnobotanical value lies not merely in generating genetic distance matrices, but in testing botanical identities attached to named remedies, mapping the population histories of culturally valued resources, documenting the composition of traditional formulas, and supporting conservation decisions for plants under harvesting pressure.

Important conceptual limits accompany this technological expansion. Neutral genetic differentiation does not automatically predict the concentration of bioactive compounds, clinical efficacy, or the cultural status of a remedy. Likewise, a DNA sequence can indicate that a taxon is present, but it cannot by itself demonstrate why a community selected it, how it was prepared, or whether an archaeological specimen was used medicinally. Molecular findings must therefore be interpreted through an integrated framework that combines taxonomic vouchers, chemistry, ecology, historical records, archaeological context, and ethically documented traditional knowledge (Brown *et al.* 2015, Huang *et al.* 2015, Przelomska *et al.* 2020).

This review maps the technical development and ethnobotanical applications of molecular markers in medicinal plants. It asks four linked questions: (1) which marker systems are appropriate for particular sample types and research questions; (2) how molecular evidence contributes to authentication, genetic-diversity assessment, provenance, cultivation, and conservation; (3) how ancient DNA and DNA metabarcoding can illuminate the historical transmission and contemporary integrity of medicinal knowledge; and (4) which methodological and ethical safeguards are needed to prevent genetic evidence from being overinterpreted or detached from the people and cultural contexts to which it relates.

Materials and Methods

Review Design and Database Sources

This study used a systematic evidence-mapping design to synthesize the development and application of molecular markers and genomic technologies in ethnomedicinal-plant research. The search covered literature published from 1 January 1980 through the cut-off date of 31 May 2026. Three exportable primary repositories were used for quantitative counting: China National Knowledge Infrastructure (CNKI), Wanfang Data, and PubMed. Searches were run in the platform-supported theme, title, abstract, keyword, and topic fields, as applicable. Web of Science Core Collection and Google Scholar were used only for supplementary citation tracing, cross-checking of influential papers, and identification of foundational studies already represented in the counted corpus. This design prioritizes reproducible database-level counts while retaining the citation-chasing value of broader scholarly search tools. The reproducible database-specific search strings, date limits, fields, and search-set yields are reconstructed from the archived search-set records and provided in Supplementary Methods S1.

Search Architecture

Boolean operators (AND/OR) were used to combine three concept groups: molecular methods, plant or product types, and ethnobotanical applications. Search syntax was adapted to each database, and Chinese and English equivalents were used where available. Retrieval was intentionally broad; conference abstracts, editorials, short commentaries, and records lacking eligible DNA-level evidence were removed during screening rather than by restrictive publication-type filters. The reproducible database-specific strings reconstructed from the archived search sets are provided in Supplementary Methods S1.

The molecular-method group included molecular marker, DNA marker, RFLP, RAPD, ISSR, AFLP, SSR, microsatellite, SCAR, SNP, DNA barcoding, *ITS*, *ITS2*, *rbcL*, *matK*, *psbA-trnH*, RAD-seq, ddRAD, GBS, chloroplast genome, plastome, mitochondrial genome, WGS, pangenome, ancient DNA, *trnL P6 loop*, and DNA metabarcoding.

The plant and product group included medicinal plant, ethnomedicinal plant, traditional Chinese medicine, Ayurvedic medicine, herbal medicine, crude drug, decoction piece, herbal formula, polyherbal product, archaeological plant remains, and traditional pharmacopoeia.

The application group included genetic diversity, population structure, phylogeny, authentication, adulteration, vernacular name, provenance, Dao-di herb, geo-herbalism, germplasm conservation, seed purity, marker-assisted selection, traditional knowledge, cultural heritage, and market regulation.

Eligibility Criteria and Quantitative Screening

Studies were eligible when they were peer-reviewed original articles, methodologically important reviews, or representative case studies applying DNA-level markers, sequencing, or genomic analysis to medicinal or ethnomedicinal plants. Records had to contribute directly to at least one evidence-map function: authentication, genetic diversity, provenance, germplasm conservation or breeding, traditional-formula verification, market surveillance, or historical plant use. Studies relying exclusively on morphology or chemistry, addressing non-plant medicinal materials, lacking usable DNA-level evidence, or repeating evidence already represented by a more complete source were excluded. Foundational methodological papers, archaeobotanical frameworks, and editor-requested ethnobotanical or floristic studies identified through citation tracing were used to interpret the database-screened evidence but were not counted among the 55 included records.

Studies were excluded when they (1) relied exclusively on morphology or chemical extraction without DNA-level data; (2) lacked sufficient methodological information or accessible full text; (3) duplicated the evidence of a more complete report; or (4) addressed a biological system without a defensible connection to medicinal use, traditional knowledge, resource conservation, or herbal-product governance.

The documented retrieval workflow yielded 4527 records across the three counted databases (CNKI, 910; Wanfang Data, 2373; PubMed, 1244). Duplicate records were removed using DOI matching when available and normalized title-year-journal matching when DOI information was absent or inconsistent. This removed 718 duplicates and left 3809 unique records for title and abstract screening. Initial title/abstract screening excluded 1917 records and retained 1892 as potentially relevant. Secondary eligibility screening excluded 1831 records: 861 were not about medicinal or ethnomedicinal plants, 705 did not contribute to authentication, diversity, provenance, conservation, formula verification, or historical-use questions, 208 lacked usable DNA-level evidence, 35 were morphology-only or chemistry-only studies, and 22 were methodologically redundant reviews. Sixty-one records entered full-text assessment. Six were excluded at full text: three methodologically redundant reviews and three studies of animal-derived medicinal materials rather than medicinal or ethnomedicinal plants. The final database-screened evidence map therefore comprised 55 records: 10 indexed under CNKI and 45 under PubMed. No Wanfang record remained as an independent final record because cross-platform duplicates were retained under the CNKI or PubMed source record and the remaining Wanfang records did not meet the final eligibility criteria. Following a final reconciliation of eligibility, narrative use, and bibliographic records, Supplementary Table S1 lists the 55 included records; Supplementary Figure S1 and Supplementary Tables S2-S3 report the complete flow and exclusion counts.

Records and decisions were managed in a structured Microsoft Excel audit workbook. Deduplication was conducted first within each database and then across the merged database exports. The workbook recorded source database, bibliographic metadata, duplicate status, title/abstract decision, secondary-screening decision, full-text decision, standardized exclusion reason, and final inclusion status. Before finalization, the 61 full-text records, the six full-text exclusions, and the 55-record included set were rechecked against the prespecified eligibility criteria by the author team. The included set was then reconciled against Supplementary Table S1, the narrative synthesis, and Literature cited; each included record was required to appear in the reference list and to support at least one statement in the narrative or tables. Because marker systems, taxa, outcomes, and study designs were highly heterogeneous, no statistical meta-analysis or pooled effect estimate was attempted.

Ethical and Voucher Standards

This review was conducted in accordance with the ethical principles of the International Society of Ethnobiology Code of Ethics. Because the study synthesized published literature and did not involve new interviews, field collection, or biological sampling, no new participant consent or collection permit was required. This review underscores that all future ethnobotanical molecular research must include the deposition of voucher specimens in an internationally recognized herbarium and report the herbarium name and acronym, collector, collection number, voucher accession number, locality, and associated traditional-use metadata.

Data Extraction and Thematic Coding

For each included source, we extracted the plant or product studied, cultural or geographical context, sample type, molecular method, analytical scale, principal finding, limitation, and ethnobotanical contribution. Evidence was coded under four cross-cutting functions: identity and nomenclature; place, provenance, and resource stewardship; historical transmission and archaeological interpretation; and formula composition, market trust, and governance. The sources were then organized into nine thematic domains. Because marker systems, taxa, outcomes, and study designs were highly heterogeneous, no statistical meta-analysis was attempted; the review presents a structured qualitative synthesis.

The 55 records listed in Supplementary Table S1 constitute the counted database-screened evidence map. Each included record appears in Literature cited and is cited at least once in the narrative synthesis or a table-supported discussion. Additional foundational, conceptual, archaeobotanical, taxonomic, and editor-requested ethnobotanical sources were identified through citation tracing and are cited where they support technical interpretation or cultural context; these supplementary sources were not added to the quantitative count. For every included record, the audit workbook retained database, bibliographic metadata, screening decision, exclusion status, final inclusion status, evidence-map code, and narrative citation location.

Results

From Molecular Technologies to Biocultural Evidence

Profile of the included evidence map. The reconciled 55-record database-screened corpus comprised 10 CNKI and 45 PubMed sources. Based on the principal function assigned during synthesis, 19 addressed DNA barcoding and species authentication, 10 addressed genetic diversity, germplasm, breeding, or provenance, six were methodological reviews or evidence syntheses, 11 addressed formula or product authentication and metabarcoding, and nine addressed genome-scale or other advanced molecular approaches. These categories organize the evidence map rather than imply mutually exclusive functions; several studies contributed to more than one discussion theme. Full bibliographic details, database assignments, identifiers, and classification codes are provided in Supplementary Table S1.

The historical development of molecular markers can be understood as an expansion in both analytical scale and cultural applicability. Early dominant markers were useful for detecting population differentiation in non-model plants but were sensitive to laboratory conditions. SSRs and sequence-characterized markers improved reproducibility and enabled germplasm fingerprinting. Standardized barcodes linked unknown materials to taxonomically verified reference sequences, while SNP-based and genome-wide approaches enabled fine-scale assignment and functional investigation. Ancient DNA mini-barcoding and metabarcoding extended these tools to archaeological remains and highly processed multi-ingredient products. Table 1 summarizes the practical trade-offs and the specific ethnobotanical questions each technology can address.

Across the included methodological reviews, a consistent conclusion was that marker choice should follow the biological question, DNA condition, taxonomic scale, and cultural context rather than technological novelty alone. These reviews also emphasized validated reference libraries, orthogonal chemical or morphological evidence, and explicit limits on inference from identity or neutral variation to efficacy and cultural meaning (Huang *et al.* 2015, Li *et al.* 2015, Pei *et al.* 2020, Przelomska *et al.* 2020, Raclariu *et al.* 2018a, Sarwat & Yamdagni 2016).

Standard DNA barcoding usually relies on nuclear ribosomal *ITS*/*ITS2* and plastid regions such as *rbcL*, *matK*, and *psbA-trnH*. Multi-locus combinations are preferable when a single region cannot resolve recent radiations, hybridization, or incomplete lineage sorting (CBOL Plant Working Group 2009, Chen *et al.* 2010, China Plant BOL Group 2011, Li *et al.* 2015). For ethnobotanical research, the reliability of a barcode depends not only on sequence quality but also on the quality of its voucher, locality, plant-part, vernacular-name, and use metadata. A sequence without a verifiable specimen or cultural context may reproduce taxonomic error rather than correct it.

The authentication subset provided the densest empirical support in the evidence map. Across medicinal genera, raw drugs, field-collected ethnomedicinal plants, and market materials, *ITS*/*ITS2* and plastid loci were most effective when several markers, verified vouchers, geographically representative sampling, and curated reference sequences were combined (Chen *et al.* 2010, Chen *et al.* 2019, Han *et al.* 2016, Intharuksa *et al.* 2020, Kim *et al.* 2016, Liu & Zhou 2024, Liu *et al.* 2019, Manzanilla *et al.* 2018, Mizukami *et al.* 1998, Moon *et al.* 2016, Mushwan *et al.* 2024, Nithaniyal *et al.* 2017, Phoolcharoen & Sukrong 2013, Shi *et al.* 2017, Veldman *et al.* 2020, Wang *et al.* 2018, Wong *et al.* 2024, Yuan *et al.* 2015, Zhu *et al.* 2017).

Table 1. Comparative characteristics and ethnobotanical contributions of molecular-marker and genomic approaches

Paradigm	Representative markers / platforms	Inheritance / data type	DNA-quality requirement	Throughput / data scale	Principal application	Principal limitations	Ethnobotanical contribution
Hybridization marker	RFLP	Codominant locus profiles	Abundant high-molecular-weight DNA; relatively high input	Low; tens to hundreds of loci	Baseline population differentiation and germplasm comparison	Laborious; low throughput; unsuitable for degraded material	Historical comparison of populations and germplasm resources
Random or semi-random PCR markers	RAPD, ISSR, AFLP	Predominantly dominant multilocus profiles	Moderate-quality DNA; low-to-moderate input	Low to moderate; dozens to hundreds of loci	Initial diversity and population-structure surveys in non-model taxa	Reaction-sensitive; limited cross-laboratory comparability; dominant scoring	Rapid surveys of under-studied medicinal populations and preliminary conservation units
Locus-specific markers	SSR, SCAR	SSR codominant; SCAR targeted presence/absence	Low-to-moderate DNA quantity and quality	Low to moderate; small validated panels	Germplasm fingerprinting, clonality, nursery purity, and targeted diagnostics	Marker development and population validation required	Maintenance of named landraces, cultivated stocks, and selected traits
Standard DNA barcodes	*ITS*/*ITS2*, *rbcL*, *matK*, *psbA-trnH*	Sequence haplotypes from nuclear or plastid loci	Low-to-moderate input; short loci may tolerate partial degradation	Low to moderate; one to several loci	Species authentication of raw, traded, or partially processed materials	Incomplete reference libraries; low resolution in some radiations, hybrids, or processed samples	Links named remedies and market materials to voucher-based botanical identities
Genome-wide SNP approaches	RAD-seq, ddRAD, GBS, SNP arrays or targeted panels	Codominant biallelic variants	Moderate-to-high DNA quality; protocol-dependent input	High; thousands to millions of variants	Fine-scale structure, AMOVA, assignment, provenance testing, and conservation-unit inference	Batch effects, missing data, ascertainment bias, and bioinformatic demands	Tests provenance claims and maps the diversity of culturally selected populations
Organelle genomics	Plastome super-barcoding; mitogenomics	Usually uniparentally inherited organelle haplotypes	Moderate-to-high DNA quality or sufficient organelle read depth	High; complete organelle genomes	Phylogenomics, maternal-lineage reconstruction, and high-resolution specimen identification	Represents one cytoplasmic history; limited power under introgression or recent radiation	Clarifies maternal lineages, plant movement, domestication, and difficult taxa

Paradigm	Representative markers / platforms	Inheritance / data type	DNA-quality requirement	Throughput / data scale	Principal application	Principal limitations	Ethnobotanical contribution
Whole-genome and pangenome analysis	WGS, resequencing, structural-variant analysis	Genome-wide SNPs, indels, SVs, CNVs, and PAVs	High-quality DNA; moderate-to-high input for long-read assembly	Very high; whole-genome scale	Functional-marker discovery, comparative genomics, and candidate chemotype analysis	High cost and computational demand; association does not establish functional or therapeutic causality	Connects lineage diversity with candidate chemotypes, cultivation traits, and underrepresented resources
Degraded and mixed-sample approaches	aDNA mini-barcoding; target capture; DNA metabarcoding	Ultra-short sequence haplotypes or mixed amplicon read sets	Very low quantity, highly degraded, inhibited, or mixed DNA	Moderate to high; targeted fragments or mixed libraries	Archaeological identification and multi-ingredient product or formula auditing	Contamination, primer bias, DNA loss, incomplete databases, qualitative read counts, and contextual dependence	Reconstructs historical plant presence and audits continuity of culturally transmitted formulas

RAD-seq, ddRAD, and GBS reduce genome complexity and recover thousands of SNPs without requiring a fully assembled reference genome (Andrews *et al.* 2016, Baird *et al.* 2008, Elshire *et al.* 2011, Peterson *et al.* 2012). These approaches are valuable for mapping population structure, identifying conservation units, and testing whether named regional resources correspond to genetically distinguishable populations. WGS and comparative genomics can further reveal structural variants and candidate biosynthetic genes in species such as *Panax ginseng*, *Dendrobium officinale*, *Artemisia annua*, *Taxus chinensis*, *Cannabis sativa*, and *Salvia miltiorrhiza* (Kim *et al.* 2018, Shen *et al.* 2018, Van Bakel *et al.* 2011, Xiong *et al.* 2021, Xu *et al.* 2016, Yan *et al.* 2015). However, genomic association with a metabolite remains a hypothesis until validated through expression, biochemical, environmental, and cultivation evidence.

Organelle genomes add complementary maternal-lineage information. Plastomes often provide higher discriminatory power than short chloroplast barcodes and can support phylogenomics, whereas plant mitochondrial genomes are structurally complex and generally evolve too slowly for routine species authentication. Their strongest contributions are to cytoplasmic-lineage reconstruction, organelle evolution, and the development of lineage-specific markers (Jang *et al.* 2021, Qian *et al.* 2013).

Genetic Diversity as a Basis for Biocultural Resource Conservation

Genetic diversity is the biological substrate that allows medicinal-plant populations to respond to environmental change, harvesting, cultivation, and domestication. Common population-genetic summaries include observed heterozygosity (H_O), expected heterozygosity (H_E), allelic richness, and the inbreeding coefficient F_{IS}. Differentiation among populations is commonly quantified using F_{ST} and analysis of molecular variance (AMOVA), while model-based population-structure analyses and principal coordinate analysis (PCoA) are used to examine clustering, admixture, and relationships among individuals or populations. These results should be interpreted with attention to representative sampling, missing data, marker ascertainment, and the biological and cultural definition of populations.

Early RAPD and ISSR studies documented population differentiation in *Ophiocordyceps sinensis*, *Siraitia grosvenorii*, *Rehmannia glutinosa*, and other non-model medicinal plants (Chen *et al.* 1997, Peng *et al.* 2005, Wu *et al.* 2007). Although these dominant-marker studies were technically limited, they established the need to conserve multiple populations rather than treating a named medicinal species as genetically uniform.

The diversity and germplasm subset likewise showed that dominant markers can reveal broad differentiation, whereas SSRs and sequence-based datasets provide stronger support for assignment, purity testing, and conservation-unit design. Collectively, the included studies covered wild and cultivated populations, endangered resources, named germplasm, and breeding materials (Chen *et al.* 1997, Hu *et al.* 2002, Peng *et al.* 2005, Shao *et al.* 2004, Song *et al.* 2021, Tian *et al.* 2022, Wen & Zimmer 1996, Wu *et al.* 2007, Xie *et al.* 2010, Zheng *et al.* 2010).

An ethnobotanical interpretation adds the human dimensions of selection and movement. Communities may maintain particular landraces, harvesting sites, or phenotypes because of taste, processing performance, ritual value, or perceived therapeutic quality. Genetic clusters can therefore help identify locally adapted or historically managed resource units, but they should be compared with oral histories, vernacular classifications, trade routes, and management practices before cultural significance is assigned. Conservation plans that protect only a taxonomic name while ignoring culturally preferred populations may preserve species presence yet erode biocultural diversity.

SSR and SNP datasets improve reproducibility and support individual assignment, clonality assessment, and the design of core germplasm collections. Reduced-representation sequencing further expands the number of loci available for non-model plants and can generate markers for endangered species, as demonstrated by RAD-seq-derived SSR development in *Tetracentron sinense* (Tian *et al.* 2022). These data are especially useful when wild collection, habitat fragmentation, or rapid commercialization threatens geographically restricted medicinal resources.

For high-value genera such as *Panax*, *Dendrobium*, *Glycyrrhiza*, and *Astragalus*, genome-scale data can distinguish lineages that short barcodes may collapse and can identify introgression or mixed ancestry. Such results can guide ex situ collections, cultivation-source documentation, and sampling strategies for pharmacological research. They should not, however, be converted into simple hierarchies of “superior” and “inferior” cultural groups without evidence linking genetic variants to quality traits and without consultation with knowledge holders.

A practical workflow for biocultural conservation is therefore: document local names and use categories; collect legally and ethically with voucher specimens; characterize population genetic structure; compare genetic patterns with ecological and cultural geography; and return interpretable results to participating communities and resource managers. This workflow treats molecular markers as tools for dialogue and stewardship rather than as replacements for local classification.
The main management outcome of diversity analysis is not a single diversity index, but a defensible set of conservation and cultivation units. Sampling should encompass multiple localities, habitats, named forms, and use categories, and should record whether materials are wild, cultivated, exchanged, or clonally propagated. Such metadata are necessary to distinguish natural population structure from human-mediated movement.

Overall, genome-wide approaches increase resolution, but the quality of ethnobotanical inference remains constrained by sampling design and cultural documentation. More loci cannot compensate for poorly defined populations, undocumented vouchers, or unexamined local nomenclature.

Phylogenetics, Vernacular Nomenclature, and Knowledge Mapping
Phylogenetic analysis helps resolve taxonomic uncertainty and identify closely related lineages, but ethnobotanical names may refer to one species, several species, a plant part, a preparation, or a culturally recognized functional category. Molecular trees should therefore be used to map relationships among botanical specimens associated with names and practices, not to assume that folk classifications are erroneous whenever they differ from formal taxonomy.

Nuclear *ITS* sequences have clarified relationships within *Astragalus* and *Panax*, while the chloroplast *trnK* region has been informative in *Atractylodes* (Mizukami *et al.* 1998, Wen & Zimmer 1996, Wojciechowski *et al.* 1993). The *trnL* intron has also supported identification of *Stelechocarpus burahol* and related Annonaceae, demonstrating the value of short plastid regions for poorly studied or locally important taxa (Probojati *et al.* 2023). These studies provide evolutionary context for selecting diagnostic characters and reference taxa.

Complete plastomes and multi-locus phylogenomics can improve support within difficult species complexes. Phylogenomic analysis of *Panax* and plastome sequencing of *Salvia miltiorrhiza* illustrate how organelle-scale data can identify variable regions and clarify maternal relationships (Manzanilla *et al.* 2018, Qian *et al.* 2013). Because plastids are usually uniparentally inherited, plastome trees should be compared with nuclear evidence where hybridization or cultivation-mediated introgression is plausible.

Three applications are particularly relevant to ethnobotany. First, molecular vouchers can stabilize inventories of medicinal plants documented in community-based surveys. Second, phylogenetic placement can reveal whether similarly named remedies are close relatives or unrelated substitutes. Third, phylogenetic diversity can inform conservation priorities by identifying unique lineages within heavily harvested resource pools.

In northeastern Peru, the integration of molecular and morphological evidence improved the documentation of plants with ethnomedicinal uses and strengthened the taxonomic basis of local-use records (Tineo *et al.* 2023). Such integration is preferable to sequencing market names without retaining vouchers, because future researchers can re-examine both the specimen and its associated knowledge record.

In Bangladesh, partial sequencing of three plastid markers enabled molecular identification of eight plants used by Indigenous communities (Mushwan *et al.* 2024). The contribution of this type of work is not merely a list of sequences: it creates a traceable connection among a plant specimen, a community use report, and a reference identity that can support education, conservation, and future pharmacological study.

For rare plants or cross-border materia medica, phylogenetic evidence can also identify conservation-relevant relatives and clarify nomenclatural synonymy. Nevertheless, taxonomic revision should be communicated carefully because changes in scientific names can disrupt legal lists, pharmacopoeial monographs, labels, and local records even when the plant itself has not changed.

Accordingly, phylogenetic outputs should be accompanied by specimen identifiers, collection permits, locality, traditional name, language, plant part, preparation, and source of knowledge. These fields turn a tree from an abstract biological diagram into an auditable ethnobotanical knowledge map.

Authentication of Ethnomedicinal Crude Drugs and Market Materials

Drying, slicing, powdering, and extraction often remove the characters used in conventional botanical identification. This is particularly problematic for expensive or toxic materials and for genera such as Fritillaria, *Panax*, *Dendrobium*, *Glycyrrhiza*, Ephedra, and Asteraceae in which multiple species may circulate under related names. Molecular authentication can recover species-level evidence from tissues that retain amplifiable DNA, thereby protecting consumers and preserving the intended botanical identity of culturally established remedies.

Large-scale validation established *ITS2* as a useful barcode for medicinal plants, and case studies have differentiated *Dendrobium huoshanense* and *Chrysanthemum indicum* from related materials (Chen *et al.* 2010, Chen *et al.* 2019, Wang *et al.* 2018). Reliability is strengthened by using more than one locus, testing reference specimens from across the species range, and reporting failed amplification rather than excluding it silently.

Integrated authentication is especially important because DNA identity and pharmacognostic quality are different questions. Studies in Peru and Bangladesh linked molecular evidence to field vouchers and local uses, while an integrated framework for *Patrinia* combined macro- and micromorphology, HPLC fingerprints, and DNA barcodes (Mushwan *et al.* 2024, Tineo *et al.* 2023, Wong *et al.* 2024). These designs recognize that no single evidence stream captures botanical identity, chemical composition, and cultural appropriateness simultaneously.

A survey of 1436 market samples representing 295 medicinal species demonstrated the scalability of *ITS2*-based authentication and also highlighted the dependence of market surveillance on comprehensive reference databases (Han *et al.* 2016). In regulation, DNA results should be linked to transparent sampling, chain-of-custody records, validated reference material, and orthogonal chemical or microscopic confirmation when legal or safety decisions are involved.

Cultivation, Seed Systems, and Marker-Assisted Selection

The shift from wild collection to cultivation can reduce pressure on threatened populations, but it also creates risks of genetic bottlenecks, seed-lot mixtures, clonal uniformity, and replacement of locally valued forms. SSRs and SNPs allow seed or planting-material lots to be compared at an early stage and can document whether cultivated stocks represent the diversity of source populations.

SSR markers developed for *Dendrobium officinale* have been used for germplasm-purity identification (Xie *et al.* 2010). Similar marker panels can support nursery certification and the maintenance of named cultivars, provided that reference samples are verified and the marker set has sufficient discriminatory power. The cultural relevance is strongest when purity standards are tied to a clearly documented landrace, local production system, or pharmacopoeial identity rather than to genetic uniformity as an end in itself.

Markers linked to sex, disease resistance, flowering time, or other cultivation traits may reduce production costs and stabilize supply. However, dominant diagnostic markers require independent validation across populations, and traits important to growers may not be identical to the traits valued by healers, processors, or patients. Participatory priority-setting can prevent breeding programs from improving yield while inadvertently losing taste, aroma, processing behavior, or other culturally selected characteristics.

High-density linkage maps and association studies can accelerate marker-assisted selection, but neutral markers should not be presented as direct indicators of therapeutic quality. Functional-marker development requires replicated associations among genotype, gene expression, metabolite profiles, agronomic environment, and processing. The development of *Bupleurum chinense* cultivars illustrates the broader move toward standardized production, but molecular selection should remain connected to the historically recognized material and quality attributes of the drug (Zheng *et al.* 2010).

Geographical Provenance and the Evaluation of Dao-di Herbs

Dao-di herbs are place-based medicinal materials whose reputation is shaped by long-term interaction among genotype, climate, soil, cultivation, harvesting, processing, trade, and clinical tradition. Molecular markers can test whether production regions contain distinguishable lineages or whether market lots can be assigned to source populations. They cannot, on their own, prove superior efficacy or replace the historical and pharmacological criteria through which Dao-di status is recognized.

Early RAPD work generated fingerprints for *Alisma orientale*, and population-marker studies have examined cultivated *Panax ginseng* and other medicinal resources (Hu *et al.* 2002, Shao *et al.* 2004). Modern SSR and SNP panels can improve assignment

accuracy, but robust provenance testing requires representative sampling across candidate regions, blind validation, and explicit error rates. A marker that separates two study populations may fail when additional production areas or admixed planting materials are included.

A defensible provenance framework therefore combines nuclear markers, organelle haplotypes, ecological variables, metabolomics, processing records, and documentary evidence. The "bimolecular marking" concept similarly emphasizes that genetic and chemical evidence answer complementary questions (Huang *et al.* 2015). For geographical-indication protection, DNA can substantiate lineage or assignment, while cultural history and production practice establish the meaning and value of place.

Whole-Genome and Functional Markers: From Identity to Quality-Associated Variation

WGS has shifted medicinal-plant research from identifying taxa to investigating the genomic architecture of specialized metabolism. Reference genomes for *Panax ginseng*, *Dendrobium officinale*, *Artemisia annua*, *Taxus chinensis*, *Cannabis sativa*, and *Salvia miltiorrhiza* have provided candidates in pathways involving ginsenosides, polysaccharides, artemisinin, taxoids, cannabinoids, tanshinones, and phenolic acids (Kim *et al.* 2018, Shen *et al.* 2018, Van Bakel *et al.* 2011, Xiong *et al.* 2021, Xu *et al.* 2016, Yan *et al.* 2015). These resources can support functional-marker development, but candidate genes and gene families such as *CYP450*, *UGT*, *OSC*, *TS*, and *DBAT* require experimental validation before they are used to predict medicinal quality. The genome-scale subset demonstrates the transition from lineage identification to organelle reconstruction, reference-genome development, pathway discovery, and deep-time interpretation. Across these studies, genomic evidence was most informative when linked to population sampling, chemical validation, cultivation history, or archaeological context rather than treated as a stand-alone indicator of medicinal quality or use (Jang *et al.* 2021, Kim *et al.* 2018, Qian *et al.* 2013, Russo *et al.* 2008, Shen *et al.* 2018, Van Bakel *et al.* 2011, Xiong *et al.* 2021, Xu *et al.* 2016, Yan *et al.* 2015).

The strongest functional studies connect sequence variants to expression, enzyme activity, metabolite abundance, environment, and phenotype. Such integration can distinguish neutral population structure from variants that plausibly influence a chemotype. It can also reveal whether culturally preferred materials represent a unique genetic lineage, an environmentally induced phenotype, a processing tradition, or a combination of these factors.

Pangenomics can reduce reference bias by capturing structural variants, presence-absence variants, and copy-number variants that are absent from a single genome. The tomato pan-genome illustrates how dispensable genomic regions can contain rare alleles affecting complex traits (Gao *et al.* 2019). For medicinal plants, pangenomes should be designed to include wild populations, cultivated landraces, geographically named materials, and underrepresented community-managed resources so that genomic infrastructure does not reproduce existing sampling inequities.

Mitochondrial Genomics and Cytoplasmic Lineages

Plant mitochondrial genomes are large, recombinogenic, and structurally variable. Their low nucleotide-substitution rate generally limits routine species-level authentication, but rearrangements, repeats, and lineage-specific open reading frames can inform organelle evolution and maternal-lineage studies. Long-read sequencing and careful assembly validation are particularly important because alternative mitochondrial conformations may coexist.

The complete mitochondrial genome of *Panax ginseng* provides a reference for comparative analysis and the development of cytoplasmic markers (Jang *et al.* 2021). Such markers may help trace maternal lineages in cultivated material but claims about wild versus cultivated origin or cytoplasmic male sterility require population-level validation and direct functional evidence.

From an ethnobotanical perspective, mitogenomic data are most relevant when they clarify the history of cultivated lineages, seed exchange, or domestication. They should be integrated with nuclear and plastid data because a single cytoplasmic lineage represents only one component of a plant's ancestry.

Table 2. Representative genotype-to-culture case studies in ethnomedicinal-plant research

Context / tradition	Material	Molecular approach	Principal finding	Biocultural relevance	Reference
Yanghai, Xinjiang archaeological context	Approximately 2700-year-old *Cannabis* material	Botanical, phytochemical, and genetic analysis	Convergent evidence identified *Cannabis sativa* and preserved cannabinoid-related compounds	Shows how genetic evidence gains meaning through grave context and other archaeological data	Russo *et al.* 2008
Jiuwei Qianghuo Wan (traditional Chinese medicine)	Commercial multi-herb formula with a long documented history	*ITS2* + *psbA-trnH* metabarcoding with SMRT sequencing	Detected declared ingredients together with substitutions or contaminants	Audits continuity between a named historical formula and contemporary products	Xin *et al.* 2018
Four traditional Chinese medicine preparations	Processed multi-ingredient products	*ITS2* + *trnL* multi-barcode sequencing	Ingredient-detection sensitivity ranged from 77.8% to 100%	Demonstrates benefits and limits of complementary barcodes in formula verification	Yao *et al.* 2022
Chinese herbal markets	1436 samples representing 295 medicinal species	*ITS2* DNA barcoding	Large-scale authentication demonstrated feasibility and reference-library dependence	Supports nomenclatural consistency, consumer safety, and market transparency	Han *et al.* 2016
Ayurvedic products in Europe	79 commercial products	DNA metabarcoding	Usable data were obtained for only part of the sample set and ingredient fidelity was low	Highlights globalization, processing-related failure, and conservation concerns	Seethapathy *et al.* 2019
Echinacea products	38 commercial herbal products	DNA metabarcoding + HPTLC	Target *Echinacea* DNA was detected in most products, while overall ingredient fidelity was limited	Shows why genetic and chemical authentication should be interpreted together	Raclariu *et al.* 2018b
Northeastern Peru ethnomedicinal flora	Field-collected plants and vouchers	Morphology + molecular analysis	Improved identification of plants documented with ethnomedicinal uses	Anchors community-use records to re-examinable vouchers and molecular identities	Tineo *et al.* 2023
Indigenous communities of Bangladesh	Eight ethnomedicinal plant species	Partial sequencing of three plastid markers	Molecular data supported species identification	Creates traceable links among specimens, local uses, and conservation records	Mushwan *et al.* 2024
Chinese herbal products for endometriosis	Two multi-herb product categories	Dual-locus metabarcoding + HPTLC	High compositional variability and numerous unexpected taxa were detected	Demonstrates complementary evidence for global product authentication	Mück *et al.* 2023

Ancient DNA and Deep-Time Histories of Plant Use

Ancient plant DNA extends ethnobotanical inquiry beyond living collections by recovering genetic evidence from desiccated, waterlogged, charred, or otherwise degraded remains. Archaeological sequences can identify plants, track domestication and movement, and test continuity between past and present resources. Their interpretation, however, depends on archaeological context, chronology, taphonomy, contamination control, and independent botanical or chemical evidence (Brown *et al.* 2015, Przelomska *et al.* 2020, Schlumbaum *et al.* 2008).

Because ancient molecules are short and chemically damaged, analysis requires dedicated clean facilities, extraction blanks, independent replication or capture-based confirmation, and authentication criteria based on fragment length and damage patterns. The chloroplast *trnL P6 loop* is a short mini-barcode with broad taxonomic coverage and is useful for degraded plant mixtures, although its species-level resolution varies among groups (Taberlet *et al.* 2007). Mini-barcoding should not be described as ancient-DNA proof unless archaeological specimens and appropriate authentication procedures were actually analyzed.

The approximately 2700-year-old *Cannabis sativa* material from Yanghai in Xinjiang was investigated through a multidisciplinary combination of botanical, phytochemical, and genetic analyses (Russo *et al.* 2008). The case is important not because DNA alone proves medicinal use, but because genetic identity, preserved morphology, chemical constituents, grave context, and associated cultural materials jointly support interpretation. This illustrates a general principle: ancient DNA can establish biological presence and relatedness, whereas the inference of ritual, medicinal, dietary, or technological use must be built from convergent contextual evidence.

DNA Metabarcoding of Traditional Formulas and Complex Products

Traditional formulas are culturally transmitted compositions in which plant identity, plant part, proportion, processing, and preparation are interdependent. Once ingredients are pulverized, extracted, or combined, morphology becomes difficult or impossible to use. Early RAPD analysis of Yu-Ping-Feng San demonstrated the potential of DNA methods to detect prescription components (Cheng *et al.* 1998). Modern metabarcoding amplifies one or more standardized regions from total product DNA and assigns sequence reads against reference databases, enabling simultaneous detection of multiple taxa.

Using *ITS2* and *psbA-trnH* with single-molecule real-time sequencing, Xin *et al.* (2018) examined Jiuwei Qianghuo Wan, an approximately 800-year-old traditional formula, and detected prescribed plants as well as contaminants and substitutions. A multi-barcode strategy combining *ITS2* and *trnL* achieved 77.8-100% ingredient-detection sensitivity across four traditional Chinese medicine preparations (Yao *et al.* 2022). These studies show how metabarcoding can audit the continuity between a named formula, its declared materia medica, and the contents of contemporary products.

By detecting undeclared substitutions, omissions, and contaminants, molecular deconvolution can help protect culturally transmitted Indigenous and local recipes from dilution or misrepresentation in commercial markets. Such auditing should also recognize that living medical traditions may legitimately vary among communities and through time, rather than treating one formulation as culturally immutable.

Metabarcoding surveys have also revealed undeclared, substituted, or legally sensitive taxa in highly processed traditional medicines and global herbal markets (Coghlan *et al.* 2012, Mück *et al.* 2023, Raclariu *et al.* 2018b, Seethapathy *et al.* 2019). DNA and HPTLC were complementary in *Echinacea* and Chinese-formula products, reinforcing the need for orthogonal evidence. Important limitations include DNA degradation, primer bias, chloroplast copy-number differences, incomplete databases, contamination, and the fact that read abundance is not a reliable measure of ingredient quantity. A non-detection is therefore not proof of absence, and a detected taxon does not establish dose, chemical integrity, or therapeutic equivalence.

Taken together, the formula and product subset demonstrates that metabarcoding can detect declared ingredients, substitutions, contaminants, and unlisted taxa across teas, traditional Chinese medicines, Ayurvedic products, and polyherbal preparations, while also showing that detection depends on DNA survival, primer choice, reference coverage, and orthogonal confirmation (Cheng *et al.* 1998, Coghlan *et al.* 2012, Frigerio *et al.* 2021, Mück *et al.* 2023, Pandit *et al.* 2021, Raclariu *et al.* 2018b, Seethapathy *et al.* 2019, Travadi *et al.* 2023, Urumarudappa *et al.* 2022, Xin *et al.* 2018, Yao *et al.* 2022).

Discussion

Molecular Evidence as a Bridge, Not a Replacement

The central contribution of molecular markers to ethnobotany is their ability to connect a biological specimen with a traceable identity while retaining links to people, place, history, and practice. We propose a four-layer interpretive framework: (1) taxonomic evidence from vouchers, morphology, and nomenclature; (2) molecular evidence from barcodes, population markers, or genomes; (3) chemical and functional evidence concerning composition and quality; and (4) cultural-context evidence from ethnography, historical documents, archaeology, and community knowledge. Conclusions become stronger when these layers converge and more cautious when they conflict.

Recent ethnobotanical and floristic studies in northern Pakistan further demonstrate that medicinal-plant knowledge is organized through vernacular nomenclature, habitat, phenology, preparation practices, and community-specific use categories (Subhan *et al.* 2024, Ullah *et al.* 2024a, 2024b, Ullah *et al.* 2025a, 2025b). These editor-requested non-DNA sources were used only for cultural and sampling context and were not counted among the 55 database-screened molecular evidence records.

Methodological Limitations and Reporting Priorities

Cross-study comparability is reduced by variation in DNA extraction, primer choice, PCR conditions, sequencing platforms, filtering thresholds, and reference databases. Processed products and archaeological remains add fragmentation, inhibitors, and contamination. Neutral diversity may be unrelated to secondary-metabolite accumulation, while metabarcoding read counts are generally qualitative rather than quantitative. Minimum reporting should therefore include voucher identifiers, negative and positive controls, primer sequences, amplification failures, database version, taxonomic-assignment thresholds, raw-data accession numbers, and uncertainty at each analytical stage.

Biocultural Governance and Ethical Safeguards

Genetic documentation of ethnomedicinal resources can support conservation and recognition, but it may also expose sensitive locations, facilitate unapproved commercialization, or detach knowledge from its holders. Projects should obtain appropriate permits and prior informed consent, document provenance, and involve knowledge holders in deciding what can be published and reused. Where applicable, molecular research on ethnomedicinal plants should comply with the Nagoya Protocol on Access to Genetic Resources and the Fair and Equitable Sharing of Benefits Arising from their Utilization, together with relevant national legislation and mutually agreed terms. Benefits arising from genetic resources and associated traditional knowledge should be shared fairly and equitably with Indigenous peoples and local communities from whom the knowledge originated. Community taxonomies should be represented alongside scientific names rather than erased, and genetic clusters should not be used to essentialize ethnic identities or claim exclusive cultural ownership without historical and social evidence.

Future Directions: Linked Reference Systems, Functional Genomics, and Pangenomics

Future reference systems should link validated sequences to physical vouchers, collection locality, habitat, plant part, vernacular name and language, preparation, use category, chemical profile, and permissions governing reuse. Multi-omic studies should prioritize functional variants and replicate genotype-metabolite associations across environments. Pangenomes should sample wild, cultivated, and community-maintained diversity, while ancient-DNA studies should integrate damage authentication with archaeology and historical scholarship. Portable sequencing and CRISPR-based diagnostics may eventually support field or market testing, but only after assay validation, open reference standards, and clear decision thresholds are established.

Conclusion

The trajectory of molecular-marker research in ethnomedicinal plants has progressed from low-throughput fragment analysis to standardized barcoding, genome-wide variation, organelle genomics, ancient DNA, and metabarcoding. This expansion has created a reproducible toolkit for authenticating materials, describing population diversity, tracing place-based resources, auditing formulas, and investigating the long history of plant use.

Its greatest ethnobotanical value emerges when genotype is connected to culture. Molecular data can verify the botanical component of a knowledge record, but they cannot independently establish medicinal efficacy, cultural meaning, or historical intention. Voucher specimens, local nomenclature, ecological and chemical data, archaeological context, and

participation by knowledge holders are therefore not optional additions; they are the interpretive framework that makes genetic evidence relevant to ethnobotany.

Future priorities include culturally annotated reference libraries, functional markers linked to validated chemotypes, representative pangenomes, and ethical governance of genomic and traditional-knowledge data. Integrating these approaches can support wild-resource conservation, responsible cultivation, market transparency, and the intergenerational transmission of traditional medical knowledge while avoiding the reduction of complex biocultural systems to DNA sequences alone.

Declarations

List of abbreviations*:* aDNA, ancient DNA; AFLP, amplified fragment length polymorphism; AMOVA, analysis of molecular variance; CNKI, China National Knowledge Infrastructure; CRISPR, clustered regularly interspaced short palindromic repeats; DNA, deoxyribonucleic acid; FIS, inbreeding coefficient; FST, fixation index; GBS, genotyping-by-sequencing; HE, expected heterozygosity; HO, observed heterozygosity; HPLC, high-performance liquid chromatography; HPTLC, high-performance thin-layer chromatography; ISSR, inter-simple sequence repeat; *ITS*, internal transcribed spacer; PCoA, principal coordinate analysis; PCR, polymerase chain reaction; RAD-seq, restriction site-associated DNA sequencing; RAPD, random amplified polymorphic DNA; RFLP, restriction fragment length polymorphism; SCAR, sequence-characterized amplified region; SMRT, single-molecule real-time; SNP, single-nucleotide polymorphism; SSR, simple sequence repeat; WGS, whole-genome sequencing.

Ethics approval and consent to participate: Not applicable because this review analyzed published literature and public databases only and involved no new interviews, participant recruitment, field collections, or biological sampling. The review was conducted in accordance with the ethical principles of the International Society of Ethnobiology Code of Ethics.

Consent for publication: Not applicable.

Availability of data and materials: All data analyzed in this review were obtained from previously published literature and public databases. No new biological material or raw experimental dataset was generated. A concise quantitative summary is provided here, while the reconstructed database-specific search strings, the 55 included records, the PRISMA-style flow diagram, and stage-specific exclusion counts are supplied in Supplementary Methods S1, Supplementary Figure S1, and Supplementary Tables S1-S3. The documented workflow comprised 4527 retrieved records, 718 duplicates removed, 3809 unique records screened, 1917 title/abstract exclusions, 1831 secondary-screening exclusions, 61 full-text records assessed, six full-text exclusions, and 55 sources included in the final qualitative synthesis.

Competing interests: The authors declare that they have no competing financial or personal interests that could have appeared to influence the work reported in this paper.

Funding: This work was supported by the Major Project of the China Academy of History under the National Social Science Fund of China (Grant No. 24VLS005). It is also a phased research achievement of the National Undergraduate Innovation and Entrepreneurship Training Program at Minzu University of China (Project Nos. CXRP2025110025, URTP2025110153, and URTP2025110087). This work was also supported by the Lantai Youth Scholar Program of Chinese Academy of History (Grant number: 2025LTQN807). The funding bodies contributed to the study design and literature retrieval, and provided support for data analysis, interpretation of the results, and preparation of the manuscript.

Authors' contributions: Y.J.L. & D.N.P.: Literature survey and macro-data analysis; H.Y.C. & Y.J.L.: Drafting and writing the initial manuscript; H.J.L. & L.H.G.: Phylogenetic lineage analysis and table organization; S.Y.C. & Y.S.: Assisting with literature retrieval, as well as the standardization and verification of citations; B.L. & L.G.: Formulating the overall scientific framework, supervising and coordinating the project, critical revision, and final polishing of the English manuscript. All authors have read and formally approved the final version of the submitted manuscript.

Acknowledgements

We highly appreciate the College of Life and Environmental Sciences and the College of Ethnology and Sociology at Minzu University of China for providing comprehensive technical and administrative support, as well as excellent management of laboratory resources. Furthermore, our sincere gratitude goes to the global researchers whose pioneering scientific foundations and raw data made this study possible.

Literature cited

Andrews KR, Good JM, Miller MR, Luikart G, Hohenlohe PA. 2016. Harnessing the power of RADseq for ecological and evolu tionary genomics. Nature Reviews Genetics 17:81-92. doi: 10.1038/nrg.2015.28.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA. 2008. Rapid SNP discovery and genetic mapping using sequenced RAD markers. PLoS ONE 3:e3376. doi: 10.1371/journal.pone.0003376.

Brown TA, Cappellini E, Kistler L, Lister DL, Oliveira HR, Wales N, Schlumbaum A. 2015. Recent advances in ancient DNA research and their implications for archaeobotany. Vegetation History and Archaeobotany 24:207-214. doi: 10.1007/s00334-014-0489-4.

CBOL Plant Working Group, Hollingsworth PM, Forrest LL, Spouge JL, Hajibabaei M, Ratnasingham S, van der Bank M, Chase MW, Cowan RS, Erickson DL, Fazekas AJ, Graham SW, James KE, Kim KJ, Kress WJ, Schneider H, van Alphen Stahl J, Barrett SC, Barrett CF, de Vere N, Fritsch C, Guisinger MM, Barrett M, Justodo M, Little DP, Meyer CP, Nabhan D, Novoselov MS, Pozzi R, Sanderson I, Soltis DE, Soltis PS, Sytsma J, Wilkinson MJ, Damon PL. 2009. A DNA barcode for land plants. Proceedings of the National Academy of Sciences 106:12794-12797. doi: 10.1073/pnas.0905845106.

Chen FR, Wang T, Guo QS, Zhu ZB, Zhou QJ, Gui SQ, Zhao SY. 2019. Identification of *Chrysanthemum indicum* and its adulterants based on *ITS2* barcode. China Journal of Chinese Materia Medica 44:654-659. Chinese. doi: 10.19540/j.cnki.cjcmm.20181204.010.

Chen S, Yao H, Han J, Liu C, Song J, Shi L, Zhu Y, Ma X, Gao T, Pang X, Luo K, Li Y, Li X, Jia X, Lin Y, Leon C. 2010. Validation of the *ITS2* region as a DNA barcode for medicinal plants. PLoS ONE 5:e8613. doi: 10.1371/journal.pone.0008613.

Chen YJ, Wang W, Yang YX, Su B, Zhang YP, Xiong LY, He ZF, Shu C, Yang DR. 1997. Genetic divergence of *Cordyceps sinensis* as estimated by random amplified polymorphic DNA analysis. Acta Genetica Sinica 24:410-416.

Cheng KT, Tsay HS, Chen CF, Chou TW. 1998. Determination of the components in a Chinese prescription Yu-Ping-Feng San by RAPD analysis. Planta Medica 64:563-565. doi: 10.1055/s-2006-957515.

China Plant BOL Group, Li DZ, Gao LM, Li HT, Wilkinson M, Yang JB, Jiao BH, Zhou H, Fan PY, He Q, Liang HQ, Han J, Wang FX, Xun WF, Zhao WL, Jin WT, Liu Y, Zhou SL, Zhang W, Han CY, Lin JP, Lu K, Shi SF, Wang WT, Wang ZZ, Yu L. 2011. Comparative analysis of a large dataset indicates that internal transcribed spacer should be incorporated into the core barcode for seed plants. Proceedings of the National Academy of Sciences 108:19641-19646. doi: 10.1073/pnas.1104551108.

Coghlan ML, Haile J, Houston J, Murray DC, White NE, Moolhuijzen P, Bellgard MI, Bunce M. 2012. Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns. PLoS Genetics 8:e1002657. doi: 10.1371/journal.pgen.1002657.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES. 2011. A robust, simple genotyping-by-sequencing approach for high diversity species. PLoS ONE 6:e19379. doi: 10.1371/journal.pone.0019379.

Flora of China Editorial Committee. 1993. Flora of China. Science Press, Beijing, China.

Frigerio J, Agostinetto G, Mezzasalma V, De Mattia F, Labra M, Bruno A. 2021. DNA-Based Herbal Teas' Authentication: An *ITS2* and *psbA-trnH* Multi-Marker DNA Metabarcoding Approach. Plants 10:2120. doi: 10.3390/plants10102120.

Gao L, Gonda I, Sun H, Ma Q, Bao K, Tieman DM, Burzynski-Chang EA, Fish TL, Stromberg KA, Sacks GL, Thannhauser TW, Foolad MR, Diez MJ, Blanca J, Canizares J, Xu Y, van der Knaap E, Huang S, Klee HJ, Giovannoni JJ, Fei Z. 2019. The tomato pan-genome uncovers new genes and a rare allele regulating fruit flavor. Nature Genetics 51:1044-1051. doi: 10.1038/s41588-019-0410-2.

Han J, Pang X, Liao B, Yao H, Song J, Chen S. 2016. An authenticity survey of herbal medicines from markets in China using DNA barcoding. Scientific Reports 6:18723. doi: 10.1038/srep18723.

Hu SM, Zhou HT, Zhang QG, Yuan YB. 2002. Studies on DNA fingerprint of genuine Chinese herb *Alisma orientale* by RAPD. Chinese Traditional and Herbal Drugs 33:161.

Huang LQ, Qian D, Deng C. 2015. Hypothesis and application of bimolecular marking methods in Chinese materia medica. China Journal of Chinese Materia Medica 40:165-168. PMID: 26080538.

Intharuksa A, Sasaki Y, Ando H, Charoensup W, Suksathan R, Kertsawang K, Sirisa-ard P, Mikage M. 2020. The combination of *ITS2* and *psbA-trnH* region is powerful DNA barcode markers for authentication of medicinal *Terminalia* plants from Thailand. Journal of Natural Medicines 74:282-293. doi: 10.1007/s11418-019-01365-w.

Jang W, Lee HO, Lee JW, Kwon N, Kim DH, Bang KH, Jo IH. 2021. The complete mitochondrial genome of *Panax ginseng* (Apiales, Araliaceae): an important medicinal plant. Mitochondrial DNA Part B 6:3080-3081. doi: 10.1080/23802359.2021.1981167.

Kim NH, Jayakodi M, Lee SC, Choi BS, Jang W, Lee J, Kim HH. 2018. Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*. Plant Biotechnology Journal 16:1904-1917. doi: 10.1111/pbi.12926.

Kim WJ, Ji Y, Choi G, Kang YM, Yang S, Moon BC. 2016. Molecular identification and phylogenetic analysis of important med icinal plant species in genus *Paeonia* based on *rDNA-ITS*, *matK*, and *rbcL* DNA barcode sequences. Genetics and Molecular R esearch 15:gmr.15038472. doi: 10.4238/gmr.15038472.

Li X, Yang Y, Henry RJ, Rossetto M, Wang Y, Chen S. 2015. Plant DNA barcoding: from gene to genome. Biological Reviews 9 0:157-66. doi: 10.1111/brv.12104.

Liu ZW, Gao YZ, Zhou J. 2019. Molecular authentication of the medicinal species of *Ligusticum* (*Ligustici Rhizoma et Radix*, " Gao-ben") by integrating non-coding internal transcribed spacer 2 (*ITS2*) and its secondary structure. Frontiers in Plant Scie nce 10:429. doi: 10.3389/fpls.2019.00429.

Liu ZW, Zhou J. 2024. DNA barcoding of *Notopterygii Rhizoma et Radix* (Qiang-huo) and identification of adulteration in its medicinal services. Scientific Reports 14:2879. doi: 10.1038/s41598-024-53008-0.

Manzanilla V, Kool A, Nhat LN, Van HN, Thu HL, Boer HJ. 2018. Phylogenomics and barcoding of *Panax*: toward the identific ation of ginseng species. BMC Evolutionary Biology 18:44. doi: 10.1186/s12862-018-1160-y.

Mizukami H, Shimizu R, Kohjyouma M. 1998. Phylogenetic analysis of *Atractylodes* plants based on chloroplast *trnK* sequen ce. Biological & Pharmaceutical Bulletin 21:474-478. doi: 10.1248/bpb.21.474.

Moon BC, Kim WJ, Ji Y, Lee YM, Kang YM, Choi G. 2016. Molecular identification of the traditional herbal medicines, *Arisae matis Rhizoma* and *Pinelliae Tuber*, and common adulterants via universal DNA barcode sequences. Genetics and Molecula r Research 15:gmr.15017064. doi: 10.4238/gmr.15017064.

Mushwan K, Uddin SB, Siddiki AMAM, Faruque MO. 2024. Molecular identification of eight ethnomedicinal plants used by t he indigenous communities of Bangladesh through the partial genome sequencing of three plastid markers. Genetic Resour ces and Crop Evolution 71:2117-2134. doi: 10.1007/s10722-023-01781-8.

Mück F, Scotti F, Mauvisseau Q, Raclariu-Manolică AC, Schrøder-Nielsen A, Wangensteen H, de Boer HJ. 2023. Complement ary authentication of Chinese herbal products to treat endometriosis using DNA metabarcoding and HPTLC shows a high le vel of variability. Frontiers in Pharmacology 14:1305410. doi: 10.3389/fphar.2023.1305410.

Nithaniyal S, Vassou SL, Poovitha S, Raju B, Parani M. 2017. Identification of species adulteration in traded medicinal plant r aw drugs using DNA barcoding. Genome 60:139-146. doi: 10.1139/gen-2015-0225.

Pandit R, Travadi T, Sharma S, Joshi C, Joshi M. 2021. DNA meta-barcoding using *rbcL* based mini-barcode revealed presenc e of unspecified plant species in Ayurvedic polyherbal formulations. Phytochemical Analysis 32:804-810. doi: 10.1002/pca.3 026.

Pei YF, Zhang QZ, Wang YZ. 2020. Application of authentication evaluation techniques of ethnobotanical medicinal plant ge nus *Paris*: a review. Critical Reviews in Analytical Chemistry 50:405-423. doi: 10.1080/10408347.2019.1642734.

Peng YT, Tang SQ, Li BL, Liu YH. 2005. Genetic diversity of *Siraitia grosvenorii* detected by ISSR markers. Biodiversity Science 13:36-42. doi: 10.1360/biodiv.040079.

Peterson BK, Weber JN, Kay EH, Fisher HS, Hoekstra HE. 2012. Double digest RADseq: an inexpensive method for de novo S NP discovery and genotyping in model and non-model species. PLoS ONE 7:e37135. doi: 10.1371/journal.pone.0037135.

Phoolcharoen W, Sukrong S. 2013. Molecular analysis of *Vitex* species using candidate DNA barcoding and PCR-RFLP of the *matK* gene for authentication of *Vitex glabrata*. Natural Product Communications 8:125-128.

Probojati RT, Hadiyanti N, Hapsari L. 2023. Assessment for identification of *Stelechocarpus burahol* and sister species compl ex of Annonaceae family using *trnL* intron sequences. Jurnal Biologi Tropis 23:643-649. doi: 10.29303/jbt.v23i4.5777.

Przelomska NAS, Armstrong CG, Kistler L. 2020. Ancient plant DNA as a window into the cultural heritage and biodiversity o f our food system. Frontiers in Ecology and Evolution 8:74. doi: 10.3389/fevo.2020.00074.

Qian J, Song J, Gao H, Zhu Y, Xu J, Pang X, Yao H, Sun C, Li X, Li C, Liu J, Xu H, Chen S. 2013. The complete chloroplast genom e sequence of the medicinal plant *Salvia miltiorrhiza*. PLoS ONE 8:e57607. doi: 10.1371/journal.pone.0057607.

Raclariu AC, Heinrich M, Ichim MC, de Boer HJ. 2018a. Benefits and limitations of DNA barcoding and metabarcoding in her bal product authentication. Phytochemical Analysis 29:123-128. doi: 10.1002/pca.2732.

Raclariu AC, Ţebrencu CE, Ichim MC, Ciupercă OT, Brysting AK, de Boer HJ. 2018b. What's in the box? Authentication of *Echi nacea* herbal products using DNA metabarcoding and HPTLC. Phytomedicine 44:32-38. doi: 10.1016/j.phymed.2018.03.05.

Russo EB, Jiang HE, Li X, Sutton A, Carboni A, del Bianco F, Mandolino G, Potter DJ, Zhao YX, Bera S, Zhang YB, Lü EG, Fergus on DK, Hueber F, Zhao LC, Liu CJ, Wang YF, Li CS. 2008. Phytochemical and genetic analyses of ancient cannabis from Centra l Asia. Journal of Experimental Botany 59:4171-4182. doi: 10.1093/jxb/ern260.

Sarwat M, Yamdagni MM. 2016. DNA barcoding, microarrays and next generation sequencing: recent tools for genetic dive rsity estimation and authentication of medicinal plants. Critical Reviews in Biotechnology 36:191-203. doi: 10.3109/073885 51.2014.947563.

Schlumbaum A, Tensen M, Jaenicke-Després V. 2008. Ancient plant DNA in archaeobotany. Vegetation History and Archaeo botany 17:233-244. doi: 10.1007/s00334-007-0125-7.

Seethapathy GS, Raclariu-Manolică AC, Anmarkrud JA, Wangensteen H, de Boer HJ. 2019. DNA metabarcoding authenticati on of Ayurvedic herbal products on the European market raises concerns of quality and fidelity. Frontiers in Plant Science 1 0:68. doi: 10.3389/fpls.2019.00068.

Shao AJ, Li X, Huang LQ, Wei JH, Lin SF. 2004. Genetic analysis of cultivated ginseng population with the assistance of RAPD technology. China Journal of Chinese Materia Medica 29:1033-1036. PMID: 15656130.

Shen Q, Zhang L, Liao Z, Brodelius PE, Rose J, Tang K. 2018. The genome of *Artemisia annua* provides insight into the evoluti on of Asteraceae and artemisinin biosynthesis. Molecular Plant 11:776-788. doi: 10.1016/j.molp.2018.03.015.

Shi Y, Zhao M, Yao H, Yang P, Xin T, Li B. 2017. Rapidly discriminate commercial medicinal *Pulsatilla chinensis* (Bge.) Regel fr om its adulterants using *ITS2* barcoding and specific PCR-RFLP assay. Scientific Reports 7:40000. doi: 10.1038/srep40000.

Song MF, Guan YH, Zhang Y, Li HT, Zhang LX, Zhang ZL. 2021. Genetic diversity assessment of a Chinese medicinal endemic species, *Aspidopterys obcordata var. obcordata*, by combined molecular marker methods (ISSR and SRAP). Biochemical Gen etics 59:283-299. doi: 10.1007/s10528-020-10001-2.

Subhan G, Ullah S, Ali S, Zainab R, Sohail M, Shakir L, Khan I, Ali M, Ali S, Khan H, Khan WH, Ullah I. 2024. Floristic compositi on and phenology of vegetation in the district Dir Lower, Khyber Pakhtunkhwa, Pakistan. Plant Protection 8:469-486. doi: 1 0.33804/pp.008.03.5264.

Taberlet P, Coissac E, Pompanon F, Gielly L, Miquel C, Valentini A, Vermat T, Corthier G, Brochmann C, Willerslev E. 2007. P ower and limitations of the chloroplast *trnL* (UAA) intron for plant DNA barcoding. Nucleic Acids Research 35:e14. doi: 10.1 093/nar/gkl938.

Tian JQ, Duan F, Li S, Gan XH. 2022. The development of SSR markers from the endangered plant *Tetracentron sinense* Oliv. (Tetracentraceae) based on RAD-seq technique. Biologia 78:15-22. doi: 10.1007/s11756-022-01109-4.

Tineo D, Calderon MS, Bustamante DE, Oliva M. 2023. Molecular and morphological analyses of plants with ethnomedicinal uses in northeastern Peru. Ethnobotany Research and Applications 25:58. doi: 10.32859/era.25.8.1-7.

Travadi T, Shah AP, Pandit R, Sharma S, Joshi C, Joshi M. 2023. A combined approach of DNA metabarcoding collectively en hances the detection efficiency of medicinal plants in single and polyherbal formulations. Frontiers in Plant Science 14:1169 984. doi: 10.3389/fpls.2023.1169984.

Ullah R, Rahim F, Sajid M, Ullah S, Ali S, Shakir L, Sohail M, Subhan G. 2024a. Ethnobotanical study of Munda Khazana, Distri ct Dir Lower, Khyber Pakhtunkhwa. Pakistan Journal of Weed Science Research 30:105-120. doi: 10.17582/journal.PJWSR/2 024/30.3.105.120.

Ullah S, Saeed M, Sohail M, Ali S, Aziz L, Noreen A, Subhan G, Khan I, Shakir L. 2025a. Indigenous knowledge of medicinal pl ants used by the tribal communities of Hasham Valley, District Dir Lower, Khyber Pakhtunkhwa, Pakistan. Pakistan Journal o f Weed Science Research 31:37-55. doi: 10.17582/journal.PJWSR/2025/31.1.37.55.

Ullah S, Shakir L, Asif S. 2025b. Ethnobotanical survey of wild medicinal plants in Tahsil Arrang, Bajaur District, Pakistan. Jou rnal of Plant Health, Ecology and Evolution 1:1-27. doi: 10.17582/journal.jphee/2025/01.1.01.27.

Ullah S, Shakir L, Subhan G, Sohail M, Ali S, Ali S, Ali A. 2024b. An ethnobotanical survey in urban areas of Dir Lower Bar Mal akand, Khyber Pakhtunkhwa, Pakistan. Plantarum 6:108-119. doi: 10.46662/plantarum.v6i2.112.

Urumarudappa SKJ, Tungphatthong C, Jaipaew J, Pornputtapong N, Pakdeesattayapong D, Vimolmangkang S, Sukrong S. 20 22. Development of a DNA barcode library of plants in the Thai Herbal Pharmacopoeia and Monographs for authentication of herbal products. Scientific Reports 12:9624. doi: 10.1038/s41598-022-13287-x.

Van Bakel H, Stout JM, Cote AG, Tallon CM, Sharpe AG, Hughes TR, Page JE. 2011. The draft genome and transcriptome of *C annabis sativa*. Genome Biology 12:R102. doi: 10.1186/gb-2011-12-10-r102.

Veldman S, Ju Y, Otieno JN, Abihudi S, Posthouwer C, Gravendeel B, van Andel TR, de Boer HJ. 2020. DNA barcoding augme nts conventional methods for identification of medicinal plant species traded at Tanzanian markets. Journal of Ethnopharm acology 250:112495. doi: 10.1016/j.jep.2019.112495.

Wang H, Shi LL, Zhou J, Zhu GP. 2018. DNA barcoding identification of *Dendrobium huoshanense* and its adulterants. China J ournal of Chinese Materia Medica 43:4055-4061. doi: 10.19540/j.cnki.cjcmm.2018.0107.

Wen J, Zimmer EA. 1996. Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): Inferences from *ITS* sequ ences of nuclear ribosomal DNA. Molecular Phylogenetics and Evolution 6:167-77. doi: 10.1006/mpev.1996.0069.

Wojciechowski MF, Sanderson MJ, Baldwin BG, Donoghue MJ. 1993. Monophyly of aneuploid *Astragalus* (Fabaceae): Evide nce from nuclear ribosomal DNA internal transcribed spacer sequences. American Journal of Botany 80:711-722. doi: 10.10 02/j.1537-2197.1993.tb15241.x.

Wong KH, Zheng T, Yue GGL, Li MC, Wu HY, Tong MH, Zhao XL, Chen HB, Lau CBS, Shaw PC, Lau DTW. 2024. A systematic ap proach for authentication of medicinal *Patrinia* species using an integration of morphological, chemical and molecular met hods. Scientific Reports 14:6566. doi: 10.1038/s41598-024-57115-w.

Wu ZG, Wang M, Huang LQ, Liu HY, Wang F, Yuan QJ. 2007. RAPD analysis of genetic relationship among different varieties of *Rehmannia glutinosa*. China Journal of Chinese Materia Medica 32:1865-1869.

Xie ML, Hou BW, Han L. 2010. Development of SSR markers and germplasm purity identification for *Dendrobium officinale*. Acta Pharmaceutica Sinic 45:667-672.

Xin T, Xu Z, Jia J, Leon C, Hu S, Lin Y, Ragupathy S, Song J, Newmaster SG. 2018. Biomonitoring for traditional herbal medicin al products using DNA metabarcoding and single molecule, real-time sequencing. Acta Pharmaceutica Sinica B 8:488-497. d oi: 10.1016/j.apsb.2017.10.001.

Xiong X, Gou J, Liao Q, Li Y, Zhou Q, Bi G, Li C, Du R, Wang X, Sun T, Guo L, Liang H, Lu P, Wu Y, Zhang Z, Ro DK, Shang Y, Hua ng S, Yan J. 2021. The *Taxus* genome provides insights into paclitaxel biosynthesis. Nature Plants 7:1026-1036. doi: 10.1038 /s41477-021-00963-5.

Xu H, Song J, Luo H, Zhang Y, Li Q, Zhu Y, Xu J, Li Y, Song C, Wang B, Sun W, Shen G, Zhang X, Qian J, Ji A, Xu Z, Luo X, He L, Li C, Chen S. 2016. Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*. Molecular Plant 9:949-952. doi : 10.1016/j.molp.2016.03.010.

Yan L, Wang X, Liu H, Tian Y, Lian J, Yang R, Hao S, Wang X, Yang S, Li Q, Qi S, Kui L, Okpekum M, Ma X, Zhang J, Ding Z, Zhan g G, Wang W, Dong Y, Sheng J. 2015. The genome of *Dendrobium officinale* illuminates the biology of the important traditio nal Chinese orchid herb. Molecular Plant 8:922-934. doi: 10.1016/j.molp.2014.12.011.

Yao Q, Zhu X, Han M, Chen C, Li W, Bai H, Ning K. 2022. Decoding herbal materials of traditional Chinese medicine preparati ons with a multi-barcode sequencing approach. Scientific Reports 12:5988. doi: 10.1038/s41598-022-09979-z.

Yuan QJ, Zhang B, Jiang D, Zhang WJ, Lin TY, Wang NH, Chiou SJ, Huang LQ. 2015. Identification of species and materia medi ca within *Angelica* L. (Umbelliferae) based on phylogeny inferred from DNA barcodes. Molecular Ecology Resources 15:358-371. doi: 10.1111/1755-0998.12296.

Zheng TT, Sui C, Wei JH. 2010. Breeding study of new second-generation varieties "Zhongchai No. 2" and "Zhongchai No. 3" of *Bupleurum chinense*. China Journal of Chinese Materia Medica 35:1931-1934.

Zhu RW, Li YC, Zhong DL, Zhang JQ. 2017. Establishment of the most comprehensive *ITS2* barcode database to date of the t raditional medicinal plant *Rhodiola* (Crassulaceae). Scientific Reports 7:10051. doi: 10.1038/s41598-017-09769-y.

Supplementary Materials

Supplementary Methods S1. Reproducible database-specific search strategies

Search cut-off: 31 May 2026. Date range: 1 January 1980 to 31 May 2026. CNKI and Wanfang Data searches used the platform "Theme" field, which searches title, abstract, keywords, and indexed subject terms. PubMed searches used Title/Abstract terms and the publication-date limit. The platforms did not preserve a complete export of the exact query-history syntax; therefore, the strings below are reproducible reconstructions based on the archived concept combinations, search-set structure, and recorded hit counts. Counts are the archived yields at the cut-off date; later reruns may differ because database indexing is dynamic.

CNKI-1 (n = 217): SU=(("DNA 分子标记" OR "分子标记" OR RFLP OR RAPD OR ISSR OR AFLP OR SSR OR "微卫星" OR SCAR OR SNP) AND ("药用植物" OR "中药材" OR "民族药用植物" OR "中药") AND ("遗传多样性" OR "群体结构" OR "种质" OR "种质资源" OR "保护"))

CNKI-2 (n = 677): SU=(("DNA 条形码" OR "DNA barcoding" OR ITS OR ITS2 OR *rbcL* OR *matK* OR *psbA-trnH*) AND ("药用植物" OR "中药材" OR "中药" OR "草药") AND ("鉴定" OR "真伪鉴别" OR "掺伪" OR "混伪品" OR "物种鉴定"))

CNKI-3 (n = 14): SU=((SNP OR "RAD-seq" OR ddRAD OR GBS OR "叶绿体基因组" OR plastome OR "全基因组测序" OR WGS OR pangenome OR "泛基因组") AND ("药用植物" OR "道地药材" OR "中药材") AND ("产地溯源" OR "地理溯源" OR "道地性" OR "质量" OR "种质"))

CNKI-4 (n = 2): SU=(("古 DNA" OR aDNA OR "DNA 宏条形码" OR "DNA metabarcoding" OR *trnL* OR "mini-barcode") AND ("中药复方" OR "中成药" OR "传统方剂" OR "考古植物遗存") AND ("成分鉴定" OR "配方验证" OR "传统知识" OR "历史利用" OR "考古植物学"))

Wanfang-1 (n = 319): 主题:(("DNA 分子标记" OR "分子标记" OR RFLP OR RAPD OR ISSR OR AFLP OR SSR OR SCAR OR SNP) AND ("药用植物" OR "中药材" OR "民族药用植物") AND ("遗传多样性" OR "群体结构" OR "种质资源" OR "保护"))

Wanfang-2 (n = 1871): 主题:(("DNA 条形码" OR "DNA barcoding" OR ITS OR ITS2 OR *rbcL* OR *matK* OR *psbA-trnH*) AND ("药用植物" OR "中药材" OR "中药" OR "草药") AND ("鉴定" OR "掺伪" OR "真伪鉴别" OR "混伪品"))

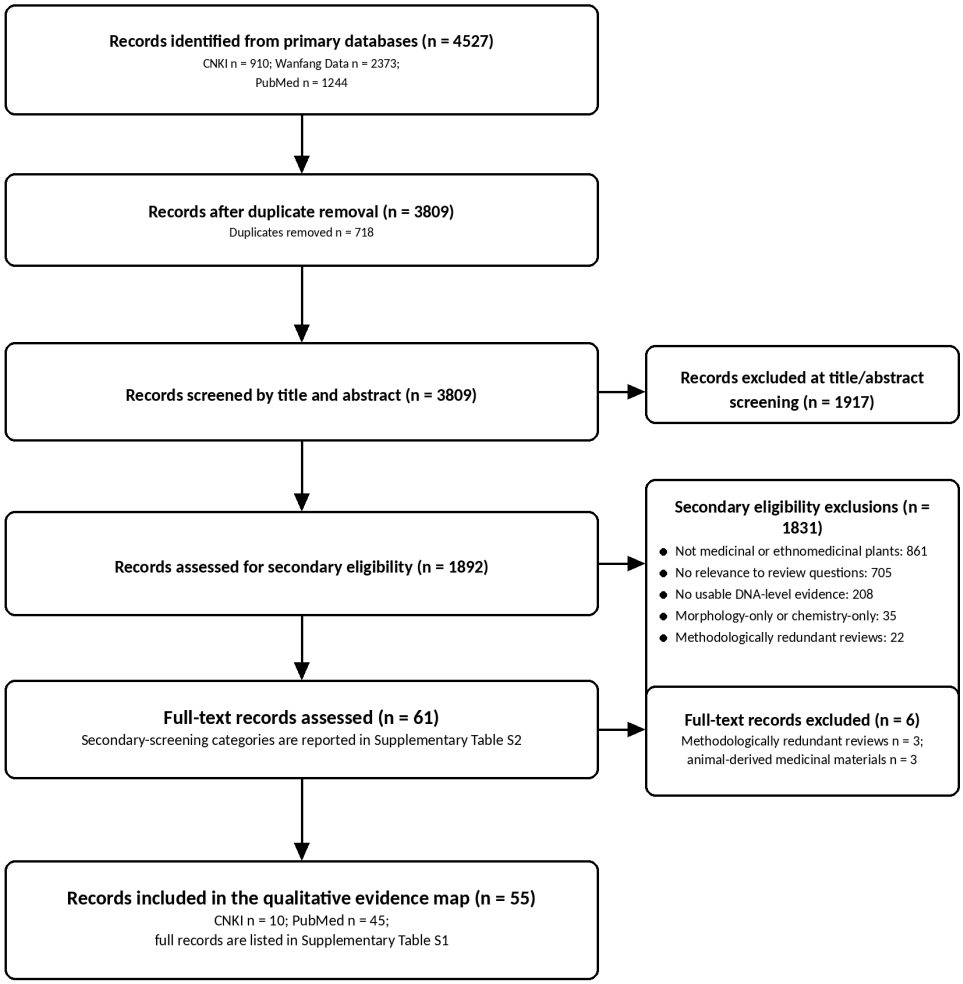
Wanfang-3 (n = 67): 主题:((SNP OR "RAD-seq" OR ddRAD OR GBS OR "叶绿体基因组" OR plastome OR "全基因组测序" OR WGS OR pangenome OR "泛基因组") AND ("药用植物" OR "道地药材" OR "中药材") AND ("产地溯源" OR "地理溯源" OR "道地性" OR "质量" OR "种质"))

Wanfang-4 (n = 10): 主题:(("古 DNA" OR aDNA OR "DNA 宏条形码" OR "DNA metabarcoding" OR *trnL* OR "mini-barcode") AND ("传统方剂" OR "中药复方" OR "中成药" OR "考古植物遗存") AND ("考古植物学" OR "传统知识" OR "历史利用" OR "成分鉴定"))

Wanfang-5 (n = 106): 主题:(("分子鉴定" OR "DNA 条形码" OR "DNA 宏条形码" OR metabarcoding OR "mini-barcode") AND ("中成药" OR "中药复方" OR "传统方剂") AND ("处方药味" OR "组方成分" OR "成分鉴定" OR "质量控制" OR "配方验证"))

PubMed (n = 1244): (("molecular marker"[Title/Abstract] OR "DNA marker"[Title/Abstract] OR RFLP[Title/Abstract] OR RAPD[Title/Abstract] OR ISSR[Title/Abstract] OR AFLP[Title/Abstract] OR SSR[Title/Abstract] OR microsatellite[Title/Abstract] OR SCAR[Title/Abstract] OR SNP[Title/Abstract] OR "DNA barcoding"[Title/Abstract] OR ITS[Title/Abstract] OR ITS2[Title/Abstract] OR *rbcL*[Title/Abstract] OR *matK*[Title/Abstract] OR *psbA-trnH*[Title/Abstract] OR "RAD-seq"[Title/Abstract] OR ddRAD[Title/Abstract] OR GBS[Title/Abstract] OR "chloroplast genome"[Title/Abstract] OR plastome[Title/Abstract] OR "mitochondrial genome"[Title/Abstract] OR WGS[Title/Abstract] OR pangenome[Title/Abstract] OR "ancient DNA"[Title/Abstract] OR "DNA metabarcoding"[Title/Abstract] OR *trnL*[Title/Abstract]) AND ("medicinal plant"[Title/Abstract] OR "ethnomedicinal plant"[Title/Abstract] OR "traditional Chinese medicine"[Title/Abstract] OR Ayurveda[Title/Abstract] OR "herbal medicine"[Title/Abstract] OR "crude drug"[Title/Abstract] OR "herbal formula"[Title/Abstract] OR "polyherbal product"[Title/Abstract] OR "archaeological plant remains"[Title/Abstract]) AND ("genetic diversity"[Title/Abstract] OR "population structure"[Title/Abstract] OR phylogeny[Title/Abstract] OR authentication[Title/Abstract] OR adulteration[Title/Abstract] OR provenance[Title/Abstract] OR "germplasm conservation"[Title/Abstract] OR "seed purity"[Title/Abstract] OR "traditional knowledge"[Title/Abstract] OR "market regulation"[Title/Abstract] OR "historical use"[Title/Abstract])) AND ("1980/01/01"[Date - Publication] : "2026/05/31"[Date - Publication])

Supplementary Figure S1. PRISMA-style literature identification and screening flow



Supplementary Table S1. Database-screened records included in the 55-record evidence map

The table reproduces the reconciled final included set from the screening audit workbook: 10 CNKI records and 45 PubMed records. Every record is cross-referenced in Literature cited and supports at least one statement in the narrative synthesis or table-supported discussion. Foundational, conceptual, archaeobotanical, taxonomic, and editor-requested contextual references identified through citation tracing are discussed in the manuscript but are not included in this quantitative count. Function codes: A, DNA barcoding and species authentication; B, genetic diversity, germplasm, breeding, and provenance; C, methodological reviews and evidence syntheses; D, formula/product authentication and metabarcoding; E, genome-scale and advanced molecular approaches.

DB	First author / year	Title	Journal	DOI / identifier	Code
CNKI	Chen FR (2019)	Identification of *Chrysanthemum indicum* and its adulterants based on *ITS2* barcode.	China Journal of Chinese Materia Medica 44:654-659	10.19540/j.cnki.cjcmm.20181204.010	A
CNKI	Chen YJ (1997)	Genetic divergence of *Cordyceps sinensis* as estimated by random amplified polymorphic DNA analysis.	Acta Genetica Sinica 24:410-416	—	B
CNKI	Hu SM (2002)	Studies on DNA fingerprint of genuine Chinese herb *Alisma orientale* by RAPD.	Chinese Traditional and Herbal Drugs 33:161	—	B
CNKI	Huang LQ (2015)	Hypothesis and application of bimolecular marking methods in Chinese materia medica.	China Journal of Chinese Materia Medica 40:165-168	PMID: 26080538	C
CNKI	Peng YT (2005)	Genetic diversity of *Siraitia grosvenorii* detected by ISSR markers.	Biodiversity Science 13:36-42	10.1360/biodiv.040079	B
CNKI	Shao AJ (2004)	Genetic analysis of cultivated ginseng population with the assistance of RAPD technology.	China Journal of Chinese Materia Medica 29:1033-1036	PMID: 15656130	B
CNKI	Wang H (2018)	DNA barcoding identification of *Dendrobium huoshanense* and its adulterants.	China Journal of Chinese Materia Medica 43:4055-4061	10.19540/j.cnki.cjcmm.2018.0107	A
CNKI	Wu ZG (2007)	RAPD analysis of genetic relationships among different varieties of *Rehmannia glutinosa*.	China Journal of Chinese Materia Medica 32:1865-1869	—	B
CNKI	Xie ML (2010)	Development of SSR markers and germplasm purity identification for *Dendrobium officinale*.	Acta Pharmaceutica Sinica 45:667-672	—	B
CNKI	Zheng TT (2010)	Breeding study of second-generation varieties “Zhongchai No. 2” and “Zhongchai No. 3” of *Bupleurum chinense*.	China Journal of Chinese Materia Medica 35:1931-1934	—	B
PubMed	Chen S (2010)	Validation of the *ITS2* region as a DNA barcode for medicinal plants.	PLoS ONE 5:e8613	10.1371/journal.pone.0008613	A
PubMed	Cheng KT (1998)	Determination of the components in a Chinese prescription Yu-Ping-Feng San by RAPD analysis.	Planta Medica 64:563-565	10.1055/s-2006-957515	D
PubMed	Coghlan ML (2012)	Deep sequencing of plant and animal DNA contained within traditional Chinese medicines reveals legality issues and health safety concerns.	PLoS Genetics 8:e1002657	10.1371/journal.pgen.1002657	D
PubMed	Han J (2016)	An authenticity survey of herbal medicines from markets in China using DNA barcoding.	Scientific Reports 6:18723	10.1038/srep18723	A

DB	First author / year	Title	Journal	DOI / identifier	Code
PubMed	Jang W (2021)	The complete mitochondrial genome of *Panax ginseng* (Apiales, Araliaceae): an important medicinal plant.	Mitochondrial DNA Part B 6:3080-3081	10.1080/23802359.2021.1981167	E
PubMed	Kim NH (2018)	Genome and evolution of the shade-requiring medicinal herb *Panax ginseng*.	Plant Biotechnology Journal 16:1904-1917	10.1111/pbi.12926	E
PubMed	Li X (2015)	Plant DNA barcoding: from gene to genome.	Biological Reviews 90:157-166	10.1111/brv.12104	C
PubMed	Manzanilla V (2018)	Phylogenomics and barcoding of *Panax*: toward the identification of ginseng species.	BMC Evolutionary Biology 18:44	10.1186/s12862-018-1160-y	A
PubMed	Mizukami H (1998)	Phylogenetic analysis of *Atractylodes* plants based on chloroplast *trnK* sequence.	Biological & Pharmaceutical Bulletin 21:474-478	10.1248/bpb.21.474	A
PubMed	Mück F (2023)	Complementary authentication of Chinese herbal products to treat endometriosis using DNA metabarcoding and HPTLC shows a high level of variability.	Frontiers in Pharmacology 14:1305410	10.3389/fphar.2023.1305410	D
PubMed	Mushwan K (2024)	Molecular identification of eight ethnomedicinal plants used by the indigenous communities of Bangladesh through the partial genome sequencing of three plastid markers.	Genetic Resources and Crop Evolution 71:2117-2134	10.1007/s10722-023-01781-8	A
PubMed	Przelomska NAS (2020)	Ancient plant DNA as a window into the cultural heritage and biodiversity of our food system.	Frontiers in Ecology and Evolution 8:74	10.3389/fevo.2020.00074	C
PubMed	Qian J (2013)	The complete chloroplast genome sequence of the medicinal plant *Salvia miltiorrhiza*.	PLoS ONE 8:e57607	10.1371/journal.pone.0057607	E
PubMed	Raclariu AC (2018b)	What's in the box? Authentication of *Echinacea* herbal products using DNA metabarcoding and HPTLC.	Phytomedicine 44:32-38	10.1016/j.phymed.2018.03.058	D
PubMed	Russo EB (2008)	Phytochemical and genetic analyses of ancient cannabis from Central Asia.	Journal of Experimental Botany 59:4171-4182	10.1093/jxb/ern260	E
PubMed	Seethapathy GS (2019)	DNA metabarcoding authentication of Ayurvedic herbal products on the European market raises concerns of quality and fidelity.	Frontiers in Plant Science 10:68	10.3389/fpls.2019.00068	D
PubMed	Shen Q (2018)	The genome of *Artemisia annua* provides insight into the evolution of Asteraceae and artemisinin biosynthesis.	Molecular Plant 11:776-788	10.1016/j.molp.2018.03.015	E
PubMed	Tian JQ (2022)	The development of SSR markers from the endangered plant *Tetracentron sinense* Oliv. (Tetracentraceae) based on RAD-seq technique.	Biologia 78:15-22	10.1007/s11756-022-01109-4	B
PubMed	Van Bakel H (2011)	The draft genome and transcriptome of *Cannabis sativa*.	Genome Biology 12:R102	10.1186/gb-2011-12-10-r102	E
PubMed	Wen J (1996)	Phylogeny and biogeography of *Panax* L. (the ginseng genus, Araliaceae): inferences from *ITS* sequences of nuclear ribosomal DNA.	Molecular Phylogenetics and Evolution 6:167-177	10.1006/mpev.1996.0069	B

DB	First author / year	Title	Journal	DOI / identifier	Code
PubMed	Wong KH (2024)	A systematic approach for authentication of medicinal *Patrinia* species using an integration of morphological, chemical and molecular methods.	Scientific Reports 14:6566	10.1038/s41598-024-57115-w	A
PubMed	Xin T (2018)	Biomonitoring for traditional herbal medicinal products using DNA metabarcoding and single molecule, real-time sequencing.	Acta Pharmaceutica Sinica B 8:488-497	10.1016/j.apsb.2017.10.001	D
PubMed	Xiong X (2021)	The *Taxus* genome provides insights into paclitaxel biosynthesis.	Nature Plants 7:1026-1036	10.1038/s41477-021-00963-5	E
PubMed	Xu H (2016)	Analysis of the genome sequence of the medicinal plant *Salvia miltiorrhiza*.	Molecular Plant 9:949-952	10.1016/j.molp.2016.03.010	E
PubMed	Yan L (2015)	The genome of *Dendrobium officinale* illuminates the biology of an important traditional Chinese orchid herb.	Molecular Plant 8:922-934	10.1016/j.molp.2014.12.011	E
PubMed	Yao Q (2022)	Decoding herbal materials of traditional Chinese medicine preparations with a multi-barcode sequencing approach.	Scientific Reports 12:5988	10.1038/s41598-022-09979-z	D
PubMed	Frigerio J (2021)	DNA-Based Herbal Teas' Authentication: An *ITS2* and *psbA-trnH* Multi-Marker DNA Metabarcoding Approach.	Plants 10:2120	10.3390/plants10102120	D
PubMed	Raclariu AC (2018a)	Benefits and Limitations of DNA Barcoding and Metabarcoding in Herbal Product Authentication.	Phytochemical Analysis 29:123-128	10.1002/pca.2732	C
PubMed	Sarwat M (2016)	DNA barcoding, microarrays and next generation sequencing: recent tools for genetic diversity estimation and authentication of medicinal plants.	Critical Reviews in Biotechnology 36:191-203	10.3109/07388551.2014.947563	C
PubMed	Veldman S (2020)	DNA barcoding augments conventional methods for identification of medicinal plant species traded at Tanzanian markets.	Journal of Ethnopharmacology 250:112495	10.1016/j.jep.2019.112495	A
PubMed	Intharuksa A (2020)	The combination of *ITS2* and *psbA-trnH* region is powerful DNA barcode markers for authentication of medicinal *Terminalia* plants from Thailand.	Journal of Natural Medicines 74:282-293	10.1007/s11418-019-01365-w	A
PubMed	Nithaniyal S (2017)	Identification of species adulteration in traded medicinal plant raw drugs using DNA barcoding.	Genome 60:139-146	10.1139/gen-2015-0225	A
PubMed	Urumarudappa SKJ (2022)	Development of a DNA barcode library of plants in the Thai Herbal Pharmacopoeia and Monographs for authentication of herbal products.	Scientific Reports 12:9624	10.1038/s41598-022-13287-x	D
PubMed	Travadi T (2023)	A combined approach of DNA metabarcoding collectively enhances the detection efficiency of medicinal plants in single and polyherbal formulations.	Frontiers in Plant Science 14:1169984	10.3389/fpls.2023.1169984	D
PubMed	Liu ZW (2024)	DNA barcoding of *Notopterygii Rhizoma et Radix* (Qiang-huo) and identification of adulteration in its medicinal services.	Scientific Reports 14:2879	10.1038/s41598-024-53008-0	A
PubMed	Yuan QJ (2015)	Identification of species and materia medica within *Angelica* L. (Umbelliferae) based on phylogeny inferred from DNA barcodes.	Molecular Ecology Resources 15:358-371	10.1111/1755-0998.12296	A

DB	First author / year	Title	Journal	DOI / identifier	Code
PubMed	Shi Y (2017)	Rapidly discriminate commercial medicinal *Pulsatilla chinensis* (Bge.) Regel from its adulterants using *ITS2* barcoding and specific PCR-RFLP assay.	Scientific Reports 7:40000	10.1038/srep40000	A
PubMed	Moon BC (2016)	Molecular identification of the traditional herbal medicines, *Arisaematis Rhizoma* and *Pinelliae Tuber*, and common adulterants via universal DNA barcode sequences.	Genetics and Molecular Research 15:gmr.15017064	10.4238/gmr.15017064	A
PubMed	Zhu RW (2017)	Establishment of the most comprehensive *ITS2* barcode database to date of the traditional medicinal plant *Rhodiola* (Crassulaceae).	Scientific Reports 7:10051	10.1038/s41598-017-09769-y	A
PubMed	Song MF (2021)	Genetic Diversity Assessment of a Chinese Medicinal Endemic Species, *Aspidopterys obcordata var. obcordata*, by Combined Molecular Marker Methods (ISSR and SRAP).	Biochemical Genetics 59:283-299	10.1007/s10528-020-10001-2	B
PubMed	Liu ZW (2019)	Molecular Authentication of the Medicinal Species of *Ligusticum* (*Ligustici Rhizoma et Radix*, "Gao-ben") by Integrating Non-coding Internal Transcribed Spacer 2 (*ITS2*) and Its Secondary Structure.	Frontiers in Plant Science 10:429	10.3389/fpls.2019.00429	A
PubMed	Phoolcharoen W (2013)	Molecular analysis of *Vitex* species using candidate DNA barcoding and PCR-RFLP of the *matK* gene for authentication of *Vitex glabrata*.	Natural Product Communications 8:125-128	—	A
PubMed	Kim WJ (2016)	Molecular identification and phylogenetic analysis of important medicinal plant species in genus *Paeonia* based on *rDNA-ITS*, *matK*, and *rbcL* DNA barcode sequences.	Genetics and Molecular Research 15:gmr.15038472	10.4238/gmr.15038472	A
PubMed	Pei YF (2020)	Application of Authentication Evaluation Techniques of Ethnobotanical Medicinal Plant Genus *Paris*: A Review.	Critical Reviews in Analytical Chemistry 50:405-423	10.1080/10408347.2019.1642734	C
PubMed	Pandit R (2021)	DNA meta-barcoding using *rbcL* based mini-barcode revealed presence of unspecified plant species in Ayurvedic polyherbal formulations.	Phytochemical Analysis 32:804-810	10.1002/pca.3026	D

Supplementary Table S2. Exclusion-category counts by screening stage

Screening stage	Exclusion category	Count
Duplicate removal	Duplicate record	718
Title/abstract screening	Not medicinal or ethnomedicinal plants	597
Title/abstract screening	No DNA-level evidence	474
Title/abstract screening	No relevance to authentication, diversity, provenance, conservation, formula verification, or historical use	436
Title/abstract screening	Morphology-only or chemistry-only study	361
Title/abstract screening	Methodologically redundant review	37
Title/abstract screening	Conference abstract / editorial / short commentary	12
Secondary eligibility screening	Not medicinal or ethnomedicinal plants	861
Secondary eligibility screening	No DNA-level evidence	208
Secondary eligibility screening	No relevance to authentication, diversity, provenance, conservation, formula verification, or historical use	705
Secondary eligibility screening	Morphology-only or chemistry-only study	35
Secondary eligibility screening	Methodologically redundant review	22
Full-text assessment	Methodologically redundant review	3
Full-text assessment	Not medicinal or ethnomedicinal plants	3

Supplementary Table S3. Records excluded after full-text assessment

database	Title	Year / journal	DOI	Reason
CNKI	中药材质量分析研究进展	2025 / 化工与医药工程	—	Methodologically redundant review
CNKI	DNA条形码技术在中药领域的应用研究进展	2024 / 现代中药研究与实践	10.13728/j.1673-6427.2024.01.016	Methodologically redundant review
CNKI	种子果实类中药DNA条形码鉴定的研究进展	2021 / 光明中医	—	Methodologically redundant review
CNKI	DNA条形码分子鉴定在动物药中的研究进展	2018 / 中国中药杂志	10.19540/j.cnki.cjcmm.20181025.008	Not medicinal or ethnomedicinal plants
CNKI	DNA条形码鉴定技术在动物类中药材鉴定领域的研究进展	2017 / 吉林中医药	10.13463/j.cnki.jlzyy.2017.04.016	Not medicinal or ethnomedicinal plants
PubMed	The metabarcoding of Grubs: Traditional herbal medicine of Scarabaeidae larvae.	2024 / Gene	10.1016/j.gene.2024.148303	Not medicinal or ethnomedicinal plants