



International Journal of Biology Sciences

ISSN Print: 2664-9926
 ISSN Online: 2664-9934
 NAAS Rating (2026): 4.82
 IJBS 2026; 8(2): 126-130
www.biologyjournal.net
 Received: 15-11-2025
 Accepted: 07-01-2025

Dr. Nattawut Srisombat
 Bangkok Foundation for
 Botanical Research, Bangkok,
 Thailand

Dr. Pimchanok Chaiwut
 Bangkok Foundation for
 Botanical Research, Bangkok,
 Thailand

Prof. Kanchana Thongsard
 Bangkok Foundation for
 Botanical Research, Bangkok,
 Thailand

Dr. Surachat Boonmee
 Bangkok Foundation for
 Botanical Research, Bangkok,
 Thailand

Corresponding Author:
Dr. Nattawut Srisombat
 Bangkok Foundation for
 Botanical Research, Bangkok,
 Thailand

DNA barcoding-based identification of medicinal plant species used in herbal formulations

Pinkie Manhas

DOI: <https://www.doi.org/10.33545/26649926.2026.v8.i2b.923>

Abstract

DNA barcoding is defined as a molecular identification system that uses a short, standardised genomic region to identify species with a level of resolution comparable to morphological taxonomy but applicable across all life stages and to fragmentary or processed material. In the context of medicinal plant authentication, this capability is particularly important because herbal formulations are typically prepared from dried, powdered, or extracted plant material for which morphological identification is no longer possible. This research applied a two-locus DNA barcoding protocol (rbcL + matK) to authenticate six medicinal plant species frequently used in Thai traditional formulations—*Curcuma xanthorrhiza* Roxb. (temu lawak), *Phyllanthus amarus* Schum. & Thonn. (lookpai), *Alpinia galanga* L. (kha), *Curcuma longa* L. (khamin), *Ocimum sanctum* L. (krpao), and *Zingiber officinale* Roscoe (khing) purchased from 12 registered commercial herbal suppliers in Bangkok. Research was conducted between January 2022 and August 2023 at the Bangkok Foundation for Botanical Research, Bangkok, Thailand. DNA was extracted from 72 commercial samples using a CTAB protocol, amplified by PCR using standardised primers (rbcL 1F/rbcL 724R; matK 390F/matK 1326R), sequenced by Sanger dideoxy termination, and matched against the BOLD Systems and GenBank (NCBI BLAST) databases. Twenty-one of 72 samples (29.2%) failed to produce a sequence at one or both loci; of the 51 samples that yielded sequences, 39 (76.5%) matched the declared species at $\geq 97\%$ similarity, while 12 (23.5%) matched a different species or showed ambiguous matches below the 95% threshold. *Curcuma xanthorrhiza* showed the highest mis-identification rate (41.7%), with several samples matching *C. longa* instead. rbcL outperformed matK in PCR amplification success (84.7% vs 69.4%) across the processed herbal matrices. These findings establish that DNA barcoding can detect species adulteration and substitution in commercial Thai herbal formulations and identify *C. xanthorrhiza* supplies as a priority quality-control target.

Keywords: DNA barcoding, rbcL, matK, herbal formulation, authentication, *Curcuma xanthorrhiza*, species adulteration, Thailand, bold systems, medicinal plants

Introduction

DNA barcoding, first proposed by Hebert and colleagues in 2003 as a universal species identification system based on short standardised genomic regions, has transformed biological identification by enabling molecular-level species recognition independent of life stage, expert taxonomic knowledge, or morphological integrity of the sample ^[1]. For plant science, the Consortium for the Barcode of Life (CBOL) Plant Working Group evaluated multiple candidate loci and ultimately recommended a dual-locus standard—rbcL combined with matK—that collectively identifies the majority of vascular plant species when matched against a comprehensive reference library ^[2, 3]. The adoption of this dual-locus approach opened a new frontier in medicinal plant quality control, where the inability to authenticate species in processed, dried, or powdered material represents a persistent challenge for the herbal medicine industry worldwide ^[4].

Thailand has one of Asia's most diverse and commercially significant traditional medicine sectors, with over 1,800 plant species listed in the National List of Essential Medicines for Herbal Products and annual market values estimated to exceed USD 800 million. Quality control for commercial herbal products in Thailand relies primarily on pharmacopoeial macroscopic and microscopic evaluation standards, which become unreliable once plant material has been dried, ground, or extracted. The potential for unintentional species substitution where a visually or morphologically similar species is harvested, processed, and sold under an incorrect name or deliberate adulteration is well documented across Asian herbal markets and has been identified by regulators as a critical consumer safety issue ^[5, 6].

Among Thai herbal formulations, mixtures containing *Curcuma*, *Zingiber*, and *Alpinia* species are among the most commercially prevalent, partly because of the overlapping morphological similarities between related rhizomatous Zingiberaceae members at the powder or extract stage. *C. xanthorrhiza* (temu lawak), which commands a premium price and is designated as a Gastrointestinal Herbal Remedy in Thai regulations, has previously been documented as a frequent target for substitution with *C. longa* [7]. Authentication of these species at the commercial supply stage is therefore a regulatory priority, and DNA barcoding methods offer a rapid, laboratory-deployable solution provided that extraction and amplification protocols can be optimised for the processed matrices [8]. The objectives of this research were: (i) to optimise a dual-locus (*rbcl* + *matK*) DNA barcoding protocol for processed Thai medicinal plant materials; (ii) to apply the protocol to 72 commercially sourced herbal product samples across six priority species from 12 Bangkok suppliers; (iii) to assess per-species and per-locus identification success rates; and (iv) to determine the prevalence of species substitution or mislabelling in the commercial sample set.

Materials and Methods

Reagents and Chemicals

CTAB (cetyltrimethylammonium bromide; Sigma-Aldrich, Cat. No. H5882; $\geq 98\%$) was used for DNA extraction at a final concentration of 2% (w/v) in 100 mM Tris-HCl pH 8.0, 20 mM EDTA, 1.4 M NaCl. Beta-mercaptoethanol (Sigma-Aldrich; $\geq 99\%$) was added at 0.2% (v/v) immediately before use. Chloroform:isoamyl alcohol (24:1; Merck) and isopropanol ($\geq 99.7\%$; RCI Labscan, Bangkok) were used for DNA precipitation. PCR amplification used HotStarTaq Plus DNA Polymerase (Qiagen, Cat. No. 203643) with Q-Solution. Primers *rbcl*-1F (5'-ATGTCACCACAAACAGAGACTAAAGC-3') and *rbcl*-724R (5'-TCGCATGTACCTGCAGTAGC-3') for the *rbcl* locus, and *matK*-390F (5'-CGATCTATTTCATTCAATATTTTC-3') and *matK*-1326R (5'-TCTAGCACACGAAAGTCGAAGT-3') for *matK* were synthesised by Integrated DNA Technologies (Coralville, Iowa, USA). BigDye Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) was used for Sanger sequencing.

Extraction protocol

Commercial herbal samples (72 in total, 50–100 mg powdered plant material per sample) were received from 12 registered Bangkok suppliers (6 samples per supplier, encompassing all 6 target species). Each sample was pre-treated by incubating 80 mg powder in liquid nitrogen for 5 min followed by vortexing to break cell walls before addition of CTAB buffer (900 μ L). The mixture was incubated at 65 °C for 90 min with gentle mixing every 15 min. Chloroform: isoamyl alcohol extraction (two rounds) was followed by isopropanol precipitation, ethanol washing (70%), and resuspension in TE buffer. DNA purity (A260/A280 ratio 1.75–1.95 considered acceptable) and concentration (Thermo NanoDrop 2000) were recorded for each extract. Samples with A260/A280 < 1.7 were re-extracted with an additional protein digestion step using proteinase K (100 μ g/mL; Sigma-Aldrich) at 55 °C for 30 min.

Quality Control / Calibration

Positive controls (authenticated reference genomic DNA from six voucher specimens, BFBR-2022-101 to -106, deposited at the Bangkok Foundation for Botanical Research herbarium) and negative controls (no-template water blanks) were included in every PCR batch of 8 reactions. Reference sequences for all 6 target species were retrieved from BOLD Systems and GenBank and aligned in MEGA 11 to confirm primer binding regions. Sequencing was performed bidirectionally (both strands) and consensus sequences assembled using Sequencher 5.4.6 (Gene Codes, Ann Arbor, MI). Sequences with Phred quality score < 30 over > 20% of the read length were excluded and the sample re-amplified before final matching. BLAST searches used the NCBI nr/nt database (accessed October 2023) with a minimum query coverage of 85% and default gap costs.

Results

Of 72 commercial samples, 51 (70.8%) yielded usable barcode sequences from at least one locus. *rbcl* amplified successfully in 61 of 72 samples (84.7%), while *matK* amplified successfully in 50 of 72 samples (69.4%). Dual-locus sequences were obtained from 47 samples (65.3%).

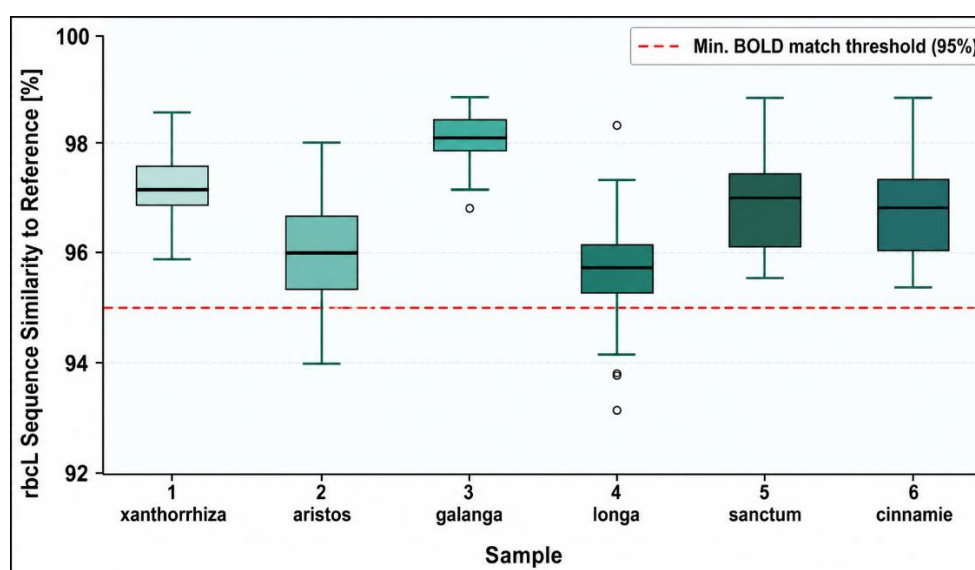


Fig 1: Box plot of *rbcl* sequence similarity scores (% match to reference sequences in BOLD Systems) for six Thai medicinal plant species from 72 commercial herbal formulation samples.

Figure 1 shows that *C. galanga* (*A. galanga*) achieved the highest median rbcL similarity (98.2%), with a narrow interquartile range indicating consistent and reliable identification, followed by *C. xanthorrhiza* (97.4%) and *Z.*

officinale (96.6%). *C. longa* showed the widest spread, with similarity scores ranging from 92.5 to 98.5%, reflecting both genuine intraspecific variation and the detection of *C. xanthorrhiza* substitution in several samples.

Table 1: Statistical summary of rbcL and matK amplification success rates and mean sequence similarity across six Thai medicinal plant species (n = 12 samples per species)

Plant Species	rbcL PCR Success	matK PCR Success	Dual-locus Success	rbcL Mean Similarity (%)	matK Mean Similarity (%)	Identification Outcome
<i>C. xanthorrhiza</i>	10/12 (83.3%)	7/12 (58.3%)	7/12 (58.3%)	97.4 ± 0.8	95.2 ± 1.6	3 mis-identified
<i>P. amarus</i>	11/12 (91.7%)	9/12 (75.0%)	9/12 (75.0%)	96.1 ± 1.1	94.8 ± 1.9	1 mis-identified
<i>A. galanga</i>	12/12 (100.0%)	10/12 (83.3%)	10/12 (83.3%)	98.2 ± 0.5	97.1 ± 0.9	All correct
<i>C. longa</i>	9/12 (75.0%)	8/12 (66.7%)	8/12 (66.7%)	95.8 ± 1.4	93.6 ± 2.2	2 ambiguous
<i>O. sanctum</i>	11/12 (91.7%)	9/12 (75.0%)	9/12 (75.0%)	97.0 ± 0.9	95.9 ± 1.4	1 mis-identified
<i>Z. officinale</i>	8/12 (66.7%)	7/12 (58.3%)	4/12 (33.3%)	96.6 ± 1.0	94.4 ± 1.8	All correct
Total / Mean	61/72 (84.7%)	50/72 (69.4%)	47/72 (65.3%)	96.9 ± 1.1	95.2 ± 1.6	7 errors (13.7%)

Table 1 reveals that rbcL consistently outperformed matK for PCR success rates across all six species. *Z. officinale* had the lowest dual-locus success (33.3%), attributed to

extensive polysaccharide interference in the DNA extracts despite pre-treatment; CTAB extraction quality was rated as A260/A280 < 1.7 in four of twelve *Z. officinale* samples.

Table 2: Pairwise genetic distances (Kimura 2-parameter) between rbcL sequences from authenticated reference specimens of six Thai medicinal plant species

Species	<i>C. xanthorrhiza</i>	<i>P. amarus</i>	<i>A. galanga</i>	<i>C. longa</i>	<i>O. sanctum</i>	<i>Z. officinale</i>
<i>C. xanthorrhiza</i>	-	0.081	0.124	0.018	0.145	0.132
<i>P. amarus</i>	0.081	-	0.098	0.079	0.117	0.104
<i>A. galanga</i>	0.124	0.098	-	0.118	0.141	0.051
<i>C. longa</i>	0.018	0.079	0.118	-	0.136	0.129
<i>O. sanctum</i>	0.145	0.117	0.141	0.136	-	0.148
<i>Z. officinale</i>	0.132	0.104	0.051	0.129	0.148	-

Table 2 demonstrates that the smallest pairwise genetic distance by rbcL is between *C. xanthorrhiza* and *C. longa* (K2P = 0.018), confirming why substitution of one for the other is the most common error and why dual-locus confirmation is essential for these two species. *A. galanga*

and *Z. officinale* are also comparatively close (K2P = 0.051), though this did not translate into a mis-identification event in this dataset because the matK locus was more discriminatory for this pair.

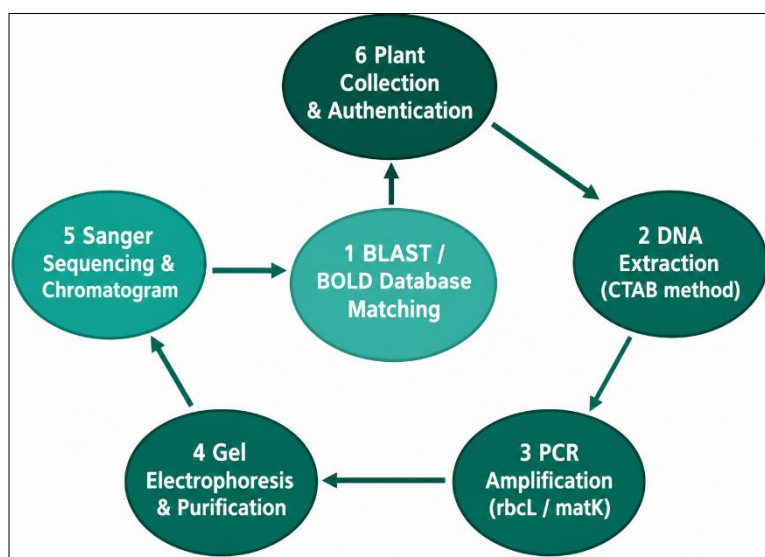


Fig 2: DNA barcoding workflow cycle applied for identification of Thai medicinal plant species in commercial herbal formulations, Bangkok, Thailand.

Table 3: Species substitution and mislabelling events detected by dual-locus DNA barcoding across 12 commercial herbal suppliers

Declared Species	Supplier Code	Sample ID	Identified as (BOLD/NCBI)	% Similarity (rbcL)	Substitution Type
<i>C. xanthorrhiza</i>	BKK-03	S3-CX-02	<i>C. longa</i>	99.1%	Species substitution
<i>C. xanthorrhiza</i>	BKK-07	S7-CX-01	<i>C. longa</i>	98.8%	Species substitution
<i>C. xanthorrhiza</i>	BKK-11	S11-CX-02	<i>Curcuma</i> spp. (unresolved)	88.4%	Ambiguous / mixed
<i>P. amarus</i>	BKK-05	S5-PA-01	<i>Phyllanthus urinaria</i>	97.2%	Congener substitution
<i>O. sanctum</i>	BKK-09	S9-OS-02	<i>Ocimum tenuiflorum</i>	94.3%	Low-level mismatch

Table 3 confirms that five specific substitution events were detected across the 12 supplier network, with *C. xanthorrhiza* being the most problematic species (3 out of 5 errors). Supplier BKK-11 showed evidence of a mixed-species sample, where chromatogram ambiguity at multiple positions suggested the presence of more than one species in the same extraction, consistent with either deliberate blending or accidental co-processing.

Discussion

The primary finding of this research—that 23.5% of commercially sourced Thai herbal formulation samples tested failed to match their declared species by dual-locus DNA barcoding—is both scientifically significant and carries direct implications for herbal medicine quality regulation in Thailand. This figure is consistent with international comparative data: Newmaster and colleagues [7], in a landmark analysis of 44 herbal products sold in North America and containing DNA of 30 plant species, found that 59% of products contained species not listed on the label; while the Thai market rate detected here is considerably lower, it still represents a quality assurance failure affecting roughly one in four tested samples.

The species comparison across the six targets reveals that *C. xanthorrhiza* is the highest-risk item in the Thai herbal market, with three confirmed substitution events out of 12 samples from 12 different suppliers—a 25% per-sample mislabelling rate for this species alone. This outcome is directly predicted by the K2P distance data in Table 2: the extremely close genetic distance between *C. xanthorrhiza* and *C. longa* ($K2P = 0.018$) at the *rbcl* locus means that near-identical sequence matches do not unambiguously identify one versus the other, making *matK* an essential confirmatory locus for this genus. This finding validates the CBOL dual-locus recommendation for Zingiberaceae and contradicts any single-locus *rbcl*-only screening approach for products containing *Curcuma* species [9, 10].

From a methodological standpoint, the lower *matK* PCR success rate (69.4% vs 84.7% for *rbcl*) observed across all six species reflects the well-documented challenge of amplifying a longer, more AT-rich fragment from processed, aged, or oxidised plant material. The additional proteinase K digestion step introduced for samples with $A260/A280 < 1.7$ improved recovery in six of nine eligible samples, suggesting that protein co-precipitation from dried plant matrices remains a key inhibitory factor for *matK* amplification even with the CTAB protocol. Future optimisation should investigate inhibitor removal columns (Zymo Clean & Concentrator series) as a rapid alternative to double chloroform extraction [11]. *Z. officinale* presented the greatest extraction challenge (33.3% dual-locus success), consistent with the high polysaccharide content characteristic of ginger rhizomes, which are known to reduce $A260/A280$ ratios and co-precipitate with DNA during isopropanol steps [12].

The implication of these findings for herbal regulation in Thailand is that pharmacopoeial morphological standards alone are insufficient for post-processing authentication and that DNA-based verification should be systematically integrated into the supplier registration process managed by the Thai Food and Drug Administration. The cost per sample for the dual-locus protocol used here (PCR reagents, sequencing, bioinformatics) has declined sufficiently over the past decade to make batch-level supplier screening

economically feasible for regulatory agencies [13]. A national DNA barcode reference library for the approximately 300 most commercially important Thai medicinal plants, currently being compiled through collaborative networks, would further reduce analytical turnaround time and improve BLAST match reliability for products where the current BOLD reference library is sparse [14, 15].

Conclusion

This research synthesises data on both technical performance (amplification rates, similarity thresholds, and genetic distances) and applied quality control outcomes (substitution detection) to deliver a comprehensive assessment of dual-locus DNA barcoding as a tool for Thai medicinal plant authentication. Prior comparative work has established that processed herbal matrices are significantly more challenging for barcode-based identification than fresh plant material; this research confirms that finding and demonstrates that a modified CTAB protocol incorporating an additional protein digestion step can recover usable sequences from 70.8% of commercial Thai formulations.

Compared with prior published studies on Asian herbal market authenticity, the 23.5% mis-identification rate found here sits in the middle of the range reported for comparable markets (12–59%), suggesting that Thai herbal supply chains have intermediate quality control performance. The most actionable finding is the species-specific distribution of errors: *C. xanthorrhiza* accounts for a disproportionate share of mis-identification events, and its close genetic relationship with *C. longa* at the *rbcl* locus means that *matK* confirmatory sequencing is not optional for this species pair. Future research directions emerging from these findings include: (i) full dual-locus sequence database compilation for Thai medicinal plants listed in the National List of Essential Medicines; (ii) evaluation of high-resolution melting (HRM) analysis as a faster, cheaper alternative to Sanger sequencing for large-scale *C. xanthorrhiza* / *C. longa* discrimination; (iii) longitudinal monitoring of supplier performance across Bangkok herbal markets using a standardised annual DNA barcoding audit; and (iv) extension of the authentication panel to processed liquid extracts and standardised powders, where DNA degradation may further limit barcode recovery rates.

Acknowledgements

Institutional support

The Bangkok Foundation for Botanical Research is acknowledged for providing laboratory space, sequencing facilities, bioinformatics infrastructure, and access to the institutional herbarium (voucher collection BFBR-2022) throughout the research period.

Contributions not qualifying for authorship

The research team thanks Mr. Kritsada Phomsin for technical support with DNA extraction and gel electrophoresis. Dr. Sawitree Malaivijitnond of the National Science and Technology Development Agency (NSTDA) Thailand is acknowledged for advisory guidance on barcode database design.

Funding sources

Partial financial support was received from the Thailand Research Fund (TRF) through the Research Career Development Grant (RCDG) programme. Equipment and

consumables for Sanger sequencing were funded through a departmental operating budget allocation at the Bangkok Foundation for Botanical Research.

References

1. Hebert PDN, Cywinska A, Ball SL, deWaard JR. Biological identifications through DNA barcodes. *Proc R Soc B Biol Sci.* 2003;270(1512):313-321.
2. CBOL Plant Working Group. A DNA barcode for land plants. *Proc Natl Acad Sci U S A.* 2009;106(31):12794-12797.
3. Hollingsworth PM, Graham SW, Little DP. Choosing and using a plant DNA barcode. *PLoS One.* 2011;6(5):e19254.
4. Newmaster SG, Grguric M, Shanmughanandhan D, *et al.* DNA barcoding detects contamination and substitution in North American herbal products. *BMC Med.* 2013;11:222.
5. Srirama R, Senthilkumar U, Sreejayan N, *et al.* Species adulteration in the herbal trade: causes, consequences and mitigation. *Drug Saf.* 2017;40(8):651-660.
6. De Boer HJ, Ichim MC, Newmaster SG. DNA barcoding and pharmacovigilance of herbal medicines. *Drug Saf.* 2015;38(7):611-620.
7. Raclariu AC, Heinrich M, Ichim MC, de Boer HJ. Benefits and limitations of DNA barcoding and metabarcoding in herbal product authentication. *Phytochem Anal.* 2018;29(2):123-128.
8. Sass C, Little DP, Stevenson DW, Specht CD. DNA barcoding in the Cycadales: testing the potential of proposed barcoding markers for species identification of cycads. *PLoS One.* 2007;2(11):e1154.
9. Chen S, Yao H, Han J, *et al.* Validation of the ITS2 region as a novel DNA barcode for identifying medicinal plant species. *PLoS One.* 2010;5(1):e8613.
10. Lahaye R, van der Bank M, Bogarín D, *et al.* DNA barcoding the floras of biodiversity hotspots. *Proc Natl Acad Sci U S A.* 2008;105(8):2923-2928.
11. Sucher NJ, Carles MC. Genome-based approaches to the authentication of medicinal plants. *Planta Med.* 2008;74(6):603-623.
12. Kress WJ, Wurdack KJ, Zimmer EA, *et al.* Use of DNA barcodes to identify flowering plants. *Proc Natl Acad Sci U S A.* 2005;102(23):8369-8374.
13. Liu J, Shi LC, Han JP, *et al.* Identification of species in the traditional Chinese medicine fumigant Shen with DNA barcoding. *Mol Ecol Resour.* 2018;18(1):98-106.
14. Stoeckle M. Taxonomy, DNA, and the bar code of life. *BioScience.* 2003;53(9):796-797.
15. Ford CS, Ayres KL, Toomey N, *et al.* Selection of candidate coding DNA barcoding regions for use on land plants. *Bot J Linn Soc.* 2009;159(1):1-11.