GCMTI RD-5:2020

GCMTI method publications

Generating DNA Barcodes

for Plant-derived Chinese Materia Medica (CMM)

Generating DNA Barcodes for Plant-derived Chinese Materia Medica (CMM)

Safety precautions: *This method involves the use of hazardous materials. It is the user's responsibility to apply appropriate precaution when handling such materials. Use eye and hand protection and where necessary carry out the work in a fume cupboard or biological safety cabinet.*

1. Introduction

DNA sequence analysis of phylogenetically meaningful DNA marker(s) is a useful tool for making taxonomic identification of biological materials. Internationally, some small and standard DNA segments called DNA barcodes gains wider recognition and accepts broadly in this application.

This method describes a general approach for generating DNA barcodes with pairs of universal primers for plant-derived Chinese Materia Medica (CMM). By selecting appropriate DNA markers, it is able to be used for plant species identification by matching with reference DNA sequence of organisms with known provenance to reveal taxonomic identity or to differentiate target species of interest from group of organisms.

A test protocol is included in Annex A & B and in the Supplementary Information as reference for ease of operation. This method can be applied to raw or moderately processed herbal material in whole or powdered form from the single species origin.

The analytical processes of this method can be broken down into the following steps: (1) sample preparation, (2) DNA extraction by manual operation or automated DNA workstation and evaluation of DNA quality and quantity, (3) polymerase chain reaction (PCR), PCR product check and cleanup, (4) Sanger DNA sequencing, and (5) post sequencing analysis.

This method provides a manual DNA workflow to enable minimal operation and investment. For more flexibility in dealing with medium to high throughput DNA workflow, automated tools were introduced to facilitate the routine steps in DNA extraction, DNA normalization, PCR setup and electrophoresis to achieve minimal manual operation.

2. Reagent and materials

All chemicals or reagents involved shall be of molecular biology or PCR grade and suitable for DNA analysis as a rule. The water used shall be double distilled or PCR grade water free from nuclease and nucleic acids. Chemicals, reagents and water shall be autoclaved prior to use, if feasible. Operator shall wear powder-free gloves for all operations. The use of aerosol-protected pipette tips as protection against cross contamination is recommended.

List of test reagents and materials for the method are listed for reference in the Supplementary Information.

3. Apparatus

General laboratory equipment and apparatus for this method are listed in the Supplementary Information. Critical equipment in the entire analytical processes are stated as follows:

3.1 **Laminar flow hood**, to provide unidirectional flow of HEPA-filtered air over the work area to prevent contamination.

3.2 **Temperature controlled incubator**, to provide experimental condition for nucleic acid extraction.

3.3 **Spectrophotometer**, to estimate the purity and yield of double-stranded DNA.

3.4 **PCR thermal cycler**, to use in PCR amplification of DNA.

3.5 **Electrophoresis system**, to check amplified PCR products.

3.6 **Automated liquid handling workstation**

3.7 **Automated capillary electrophoresis system** for DNA sequencing, to sequence nucleotide order of DNA barcodes.

4. General procedure

The test protocol described in Annex A and the Supplementary Information is for reference purpose only. The selection and suitability of method is an experience-based choice of the operator, taking into account the scope and given matrices. It is recommended operator to make reference to relevant international/ national standards if available. Having to guarantee the validity of test results, controls shall be run in parallel with test samples as described in Clause 5, where relevant, in order to detect DNA contamination of the analytical reagents or sample-to-sample contamination.

4.1 Preparation of test portion

Operator responsible for the preparation of test samples, such as reducing particle size, homogenization, and taking the test portion, shall have good laboratory practices and cross-contamination avoidance to preclude invalid test results.

4.2 DNA extraction

4.2.1 Acquisition of amplifiable DNA is a prerequisite for DNA method. The process in general can either be carried out by using a commercial DNA isolation kit or reagents

prepared in laboratory. Extraction methods validated by GCMTI are stipulated in Annex A for reference. Using commercial kits are much easier for quality control of extraction process and save time for troubleshooting as products' quality are controlled and guaranteed by manufacturers. However, the use of any other kits or any procedural modification from procedural stipulated in either A1 or A2 in Annex A shall be subjected to adequate validation by the user who has the responsibility to assess the suitability of the testing items when adopting.

4.2.2 DNA extraction kits or protocols should be selected depending on the chemical/ cellular composition of herbal materials. It is advisable, where relevant, to remove the following compositions in herbal materials in order to obtain sufficient quality and yield of DNA extracts for subsequent PCR amplification process:

4.2.2.1 polysaccharides such as pectin, cellulose, hemi-cellulose, and starch, which may be removed by the treatments with appropriate enzymes, e.g. pectinase, cellulose, α-amylase.

4.2.2.2 RNA and /or proteins, which may be removed by enzymatic treatment using RNase and proteinase, respectively.

4.2.2.3 Lipid fractions, which may be removed by enzymatic treatments or solvent extraction such as n-hexane.

4.2.2.4 Pigments such as phenolics compounds, tannins and other secondary metabolites, which may be removed using polyvinylpyrrolidone (PVP) and/or beta-mercaptoethanol.

4.2.3 Estimation of the quality of DNA extracts by means of

4.2.3.1 spectrophotometric analysis, and

4.2.3.2 amplification of endogenous gene of plant taxon.

4.3 Amplification of DNA barcodes by PCR

4.3.1 This method provides option of DNA barcodes, including internal transcribed spacer 2 (ITS2) of nuclear ribosomal RNA, chloroplast intergenic *psb*A-*trn*H spacer (*psb*A-*trn*H) and chloroplast ribulose bisphosphate carboxylase large chain (*rbc*L), for land plants. The locations of DNA barcodes selected in this method are stipulated as follows:

DNA barcode	**Location of DNA barcode**
ITS2	Internal transcribed spacer 2 (ITS2) located between nuclear 5.8S and 28S ribosomal RNA
*psb*A-*trn*H	*psb*A-*trn*H intergenic spacer (*psb*A-*trn*H) located between chloroplast photosystem II protein D1 and *trn*H gene
*rbc*L	Relative to 5' end of the ribulose bisphosphate carboxylase large chain gene between bp 1-599 (27-579 excluding primer sequences) of *Arabidopsis thaliana* chloroplast NC_000932

4.3.2 In general, amplification of DNA barcodes, and PCR product check and cleanup are conducted as below:

4.3.2.1 Select DNA barcode and amplify it by PCR. Suggested primers and corresponding PCR conditions for land plants are stated in Annex B.

4.3.2.2 Check the amplified PCR product with electrophoresis system and purify the product for subsequent DNA sequence analysis.

4.4 DNA sequencing

4.4.1 Purified PCR product and proceed to cycle sequencing reaction.

4.4.2 Purify the cycle sequencing reaction product.

4.4.3 Obtain the nucleotide sequence of DNA barcode with DNA sequencer.

4.5 Post sequencing analysis

4.5.1 For post sequencing analysis, check the quality of bi-directional sequences with appropriate types of sequencing editing software to perform analysis.

4.5.1.1 Remove low quality nucleotide(s) and primer-annealing regions for sequence assembly;

4.5.1.2 Assemble into a consensus sequence in FASTA file using both forward and reveres sequence;

4.5.1.3 Any ambiguous nucleotide (“N”) in one sequence direction can be now resolved using the complementary sequence of high quality. Base calls of high quality in one direction should be maintained over those with low quality in the other read.

5. Quality control parameters

The analytical performance of each run is checked for conformance with the acceptance criteria to determine if results of analyses are acceptable and able to meet the objective of the analytical method. In consideration of the compliance with QC plan, user should determine the number of samples in a batch to give a reasonable handling capacity for operator to conduct DNA analysis. For example, system controls should be performed for each batch of samples, or every 15 samples, whichever is the less.

5.1 Use of system controls

The system controls for the DNA test are tabulated as below. The arrows indicated that this control should be applied in the indicated analytical steps.

Control steps	Extraction blank control[1]	Extraction positive control[2]	Random sample duplicate control[3]	PCR negative control[4]	Cycle sequencing negative control[4]	Cycle sequencing positive control[5]
Nucleic acid extraction	↓	↓	↓			
Nucleic acid amplification	↓	↓	↓	↓		
Cycle sequencing amplification		↓	↓		↓	↓

1. At least one extraction blank control, preferably two, shall be included in each batch of DNA extraction. The control tubes shall always be the last in each series.

2. Extraction positive control is reference material of known taxonomic identity or the organisms of interest with taxonomic identity verified by appropriate channels such as identification or certification by recognized museums, national authorities, universities or institutes. An extraction positive control, preferably two, shall be included in each batch of DNA extraction. This control will reveal the validity of reagents and the performance of the extraction protocol.

3. At least one random sample duplicate control shall be included in each batch of DNA extraction. The control shall undergo the entire analysis and demonstrate the consistence of test result obtained from the test method.

4. PCR negative control and cycle sequencing negative control demonstrates the absence of contaminating nucleic acids in the PCR reagent and cycle sequencing reagent, respectively. Such negative controls, preferably two, shall be included in each batch of PCR and cycle sequencing reaction.

5. The cycle sequencing positive control, usually pGEM™-3Zf(+) double stranded DNA control template supplied in cycle sequencing kit, shall be included, preferably two, in each batch of cycle sequencing reaction. This control demonstrates the ability of the cycle sequencing amplification and electrophoresis separation to determine the accurate base pair composition of the target amplicon.

5.2 The following parameters shall be met:

	Parameter	Acceptance criteria
A	Extraction blank control	Negative findings of all target amplicons in PCR analysis.
B	Extraction positive control	● Positive findings of all target amplicons in entire analytical process from DNA extraction, PCR analysis to DNA sequencing. ● A band of the anticipated amplicon size shall be detected in PCR system. ● Give identity % ≥98 against reference DNA sequence in DNA sequence analysis.
C	Random sample duplicate control	Obtain consistent DNA findings in entire analytical process.
D	PCR negative control (i.e. prepared PCR or DNA sequencing master mix with no template DNA)	Give negative findings in electrophoresis analysis
E	Sequencing positive control (i.e. pGEM® - 3Zf(+) control template or equivalent)	Give positive sequencing data

6. References

Chinese Pharmacopoeia Commission,. (n.d.). 9107 Guidelines for molecular DNA barcoding of Chinese Materia Medica. In *Volume 4, Pharmacopoeia of the People's Republic of China* (2015 ed.). Beijing: China Medical Science Press.

PCR Amplification for Plants and Fungi. (n.d.). CCDB protocols. Retrieved from http://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_Amplification-Plants.pdf

Primet Sets for Plants and Fuingi. (n.d.). CCDB protocols. Retrieved from http://ccdb.ca/site/wp-content/uploads/2016/09/CCDB_PrimerSets-Plants.pdf

Watanabe T; Akiyama, H; Maleki S; Yamakawa H; Iijima K; Yamazaki F; Matsumoto T; Futo, S; Arakawa, F. ; Watai, M. and Maitani, T. (2006). A specific qualitative detection method for peanut (Arachis hypogaea) in foods using polymerase chain reaction. *Journal of Food Biochemistry, 30*, pp. 215-233.

Annex A

(Normative)

DNA Extraction Method

This test protocol has been in-housed validated by GCMTI for generating DNA barcodes for plant-derived Chinese Materia Medica (CMM). Any procedural modification from GCMTI method stipulated in Annex A shall be subjected to adequate validation by the user who has the responsibility to assess the suitability of the testing items when adopting.

A1 Preparation of PCR-quality DNA using magnetic bead based protocol with automated liquid handling workstation

A1.1 Spin down briefly the powdered sample material contained in grinding vial or microtube before open.

A1.2 Add 960 µL Buffer ATL and supplement with 40 µL proteinase K and 25 µL α-amylase into each sample and controls. Vortex sample to be a slurry.

Note: It should ensure that tissue particles are able to move freely in the lysis mix and do not remain stuck to the bottom of the tube. Additional volume of buffer may be required.

A1.3 Then place the tube in a thermomixer or heated orbital incubator, and incubate with shaking at 2,000 rpm at 56 °C for at least 3 hours or overnight.

A1.4 After incubation, centrifuge the sample tube at maximum speed for 15 min to pellet debris. Extension of centrifugation may be required for separation of supernatant from loose cell debris.

A1.5 Carefully transfer the supernatant to 2-mL sample tubes that are compatible with the sample rack of the instrument. A total of at least 600 µL sample lysate shall be recovered.

A1.6 Select appropriate extraction protocol with reference to the default protocol CW_500_ADV_HE_V8 for 500 µL input volume and retrieve 80 µL extracted DNA.

A1.7 The extracted DNA shall be ready for PCR or stored at -20 °C for long term storage, if necessary.

A2 Preparation of PCR-quality DNA using spin column DNA extraction protocol

A2.1 Spin down briefly the powdered sample material contained in grinding vial or microtube before open.

A2.2 Freshly prepare ProK-Lysis Buffer (PLB) by mixing Lysis Buffer (LB) with Proteinase K (20 mg/mL) at 10:1 ratio. For example, add 2 ml Proteinase K into 20 ml of LB and mix.

A2.3 Add 1 ml of PLB, 25 µL of α-amylase and 5 µL of RNase A into grinding vial or microtube.

A2.4 Close the cap and then mix the sample by vortexing.

Note: It should ensure that tissue particles are able to move freely in the lysis mix and do not remain stuck to the bottom of the tube. Additional volume of buffer may be required.

A2.5 If deem appropriate, put the microtube or grinding vial into the adaptor of Tissue Lyser and set frequency of 28 Hz for 30 sec. Then, disassemble adapters and rotate 180°, and apply 28 Hz for another 30 sec.

A2.6 Before incubation, shake each sample by hand.

A2.7 Incubate at 56 °C for 30 minutes and follow by 65 °C for 1 hour incubation in thermomixer.

A2.8 After incubation, centrifuge at maximum speed for 1 min to pellet the debris. Extension of centrifugation may be required for separation of supernatant from loose cell debris.

A2.9 Transfer 350 µL of lysate into a clean 1.5-mL tube and add 350 µL of Extraction Binding Buffer (EBB) to each sample.

Note: If any re-crystallization occurs in EBB, pre-warm at 56 °C to dissolve and mix well before use. If necessary, recovery of more volume of lysate for low DNA yield sample is feasible but EBB should be added in proportion.

A2.10 Mix the lysate with EBB thoroughly before loading to column.

A2.11 Transfer all lysate mixed with EBB from the tubes into the spin column placed on top of the collection tube.

A2.12 Centrifuge at 5000-6000 g for 2 mins to bind DNA to the membrane of spin column. After centrifuge, replace with a new collection tube for each sample.

A2.13 First wash step: Add 180 µL of Extraction Wash Buffer (EWB) to column. Centrifuge at 5000-6000 g for 2 mins.

A2.14 Second wash step: Add 600 µL of EWB to column. Centrifuge at 5000-6000 g for 4 mins. Replace a new collection tube for each sample and repeat once of this step.

Note: If necessary, an additional wash may need for sample rich in pigment or an extra volume of lysate has been recovered in Clause A2.9.

A2.15 After centrifuge, replace with a new collection tube for each sample and centrifuge at maximum speed for 4 mins to spin dry the column membrane.

A2.16 Replace collection tube with a new 1.5-mL microtube with lid opened. Open the column lid and incubate at 56 °C for 30 mins to remove residual ethanol in the column membrane.

A2.17 Add 80 µL pre-warmed water directly onto membrane of the column and incubate at room temperature for 1 min.

Note: Lower elution volumes may be used to result in higher DNA concentrations in comparison to the recommended volumes, but at the expense of DNA yield.

A2.18 Centrifuge at maximum speed for 4 mins to eluate the DNA.

A2.19 DNA extracts are ready for normalization to conduct PCR or keep at -20 ^{o}C for storage.

Annex B
(Normative)

Primer pairs and PCR conditions for this method

B1. Primer information

Oligo-nucleotide primers are detailed in below table. Primers may be ordered from a supplier who specializes in the synthesis of oligonucleotides. A standard desalted purification is sufficient for the use of DNA barcoding analysis.

Primer Name	Primer direction	Oligonucleotide DNA Sequence (5'-3')	Amplicon size (bp)	Target amplicon
CP03F	Forward	5' – CGG ACG AGA ATA AAG ATA GAG T – 3'	About 123	Plant chloroplast DNA
CP03R	Reverse	5' – TTT TGG GGA TAG AGG GAC TTG A – 3		
ITS2F	Forward	5' – ATG CGA TAC TTG GTG TGA AT – 3'	About 500	Plant internal transcribed spacer 2 of nuclear ribosomal RNA (ITS2)
ITS3R	Reverse	5' – GAC GCT TCT CCA GAC TAC AAT – 3'		
psbAF	Forward	5' – GTT ATG CAT GAA CGT AAT GCT C – 3'	About 500	Plant chloroplast intergenic *psb*A-*trn*H spacer (*psb*A-*trn*H)
trnHR	Reverse	5' – CGC GCA TGG TGG ATT CAC AAT CC – 3'		
rbcLaF	Forward	5' – ATG TCA CCA CAA ACA GAG ACT AAA GC – 3'	About 600	Plant ribulose bisphosphate carboxylase large chain (*rbc*L)
rbcLaR	Reverse	5' – GTA AAA TCA AGT CCA CCR CG – 3'		

B2. PCR conditions for different primer pairs

PCR components[1]	**CP03F/ CP03R**	**ITS2F/ ITS3R**	**psbAF/ trnHR**	**rbcLaF/ rbcLaR**
DNA template[2]	About 20 ng	About 30 ng	About 30 ng	About 30 ng
10X High Fidelity PCR Buffer (Minus Mg)	1 X	1 X	1 X	1 X
$MgSO_4$	1.5 mM	2.0 mM	2.0 mM	2.5 mM
Trehalose	-	-	-	5 %
dNTPs	0.2 mM @ dNTP	0.2 mM @ dNTP	0.2 mM @ dNTP	0.05 mM @ dNTP
Forward primer	0.2 μM	0.1 μM	0.1 μM	0.1 μM
Reverse primer	0.2 μM	0.1 μM	0.1 μM	0.1 μM
High Fidelity *Taq* DNA Polymerase	0.625 U per reaction	1.0 U per reaction	1.0 U per reaction	0.6 U per reaction
Nuclease-free water	An appropriate amount to make up to 25 μL	An appropriate amount to make up to 25 μL	An appropriate amount to make up to 25 μL	An appropriate amount to make up to 25 μL

1. The PCR components listed the final concentration for PCR reaction in each primer pair.
2. In most circumstance, the recommended input DNA amount is sufficient to obtain good PCR result for downstream direct sequencing. Depending on sample nature, required input DNA amount may vary to improve PCR result from PCR inhibition or interference.

B3. PCR cycling condition for amplification of target amplicons

A. Plant DNA QC test with CP03F/ CP03R primer pair

Temperature	Time	No. of cycles
95°C	10 min	1
94°C 55°C 72°C	30 sec 30 sec 30 sec	40
72°C	7 min	1
4°C	∞	--

B. Chloroplast ribulose bisphosphate carboxylase large chain (*rbc*L) with rbcLaF/ rbcLaR primer pair

Temperature	Time	No. of cycles
94°C	4 min	1
94°C 55°C 72°C	30 sec 30 sec 1 min	35
72°C	10 min	1
4°C	∞	--

C. Internal transcribed spacer 2 (ITS2) with ITS2F / ITS3R primer pair

Temperature	Time	No. of cycles
94°C	5 min	1
94°C 56°C 72°C	30 sec 30 sec 45 sec	40
72°C	10 min	1
4°C	∞	--

D. Chloroplast intergenic *psb*A-*trn*H spacer (*psb*A-*trn*H) with psbAF/ trnHR primer pair

Temperature	Time	No. of cycles
94°C	5 min	1
94°C 56°C 72°C	1 min 1 min 1.5 min	30
72°C	7 min	1
4°C	∞	--