GCMTI RD-6:2020

GCMTI method publications

Generating DNA Barcodes

for Animal-derived Chinese Materia Medica (CMM)

Generating DNA Barcodes for Animal-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 animal-derived Chinese Materia Medica (CMM). By selecting appropriate DNA markers, it is able to be used for animal 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 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 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 enables the 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, cleanup 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 capillary electrophoresis system** for DNA sequencing, to sequence nucleotide order of DNA barcodes.

4. General procedure

The test protocol described in 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.

The protocol, reagents, materials and apparatus described in the Supplementary Information is for reference purpose only. Any modification of the protocol, including but not limited to the use of other kits or changes of procedures, shall be subjected to adequate validation by the user who has the responsibility to assess the suitability of the testing items when adopting.

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. 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.

4.2.2 DNA extraction kits or protocols should be selected depending on the chemical/ cellular composition of CMM. 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 RNA and /or proteins, which may be removed by enzymatic treatment using RNase and proteinase, respectively.

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

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 animal taxon.

4.3 Amplification of DNA barcodes by PCR

4.3.1 This method provides option of DNA barcodes, including mitochondrial cytochrome c oxidase subunit I (COI), mitochondrial cytochrome b (CYTB) and mitochondrial 16S ribosomal RNA (16SrRNA), for animals. The locations of DNA barcodes selected in this method are stipulated as follows:

DNA barcode	Location of DNA barcode
COI	Relative to 5' end of the cytochrome c oxidase subunit I gene between bp 23–731 (48–705 excluding primer sequences) of *Mus musculus* mitochondrion AP013031.
CYTB	Relative to 5' end of the cytochrome b gene between bp 398–869 (423–843 excluding primer sequences) of *Mus musculus* mitochondrion AP013031.
16SrRNA	Relative to 5' end of the16S ribosomal RNA gene between bp 923–1434 (951–1407 excluding primer sequences) of *Mus musculus* mitochondrion AP013031.

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 animals are stated in Annex A.

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

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

4.5.3 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.

	Parameter	Acceptance criteria
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.

Hedges SB. (1994). Molecular evidence for the origin of birds. *Proceedings of the National Academy of Sciences of the United States of America, 91*, pp. 2621-2624.

Sawyer J; Wood C; Shanahan D; Gout S and McDowell D. (2003). Real-time PCR for quantitative meat species testing. *Food Control, 14*, pp. 579-583.

Verma SK and Singh L. (2003). Novel universal primers establish identity of an enormous number of animal species for forensic application. *Molecular Ecology Notes, 3*, pp. 28-31.

Annex A

(Normative)

Primer pairs and PCR conditions for this method

A1. 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
UnivP	Forward	GGTTTACGACCTCGATGTTG	About 104	Animal mitochondrial DNA
UnivQ	Reverse	CGGGTCTGAACTCAGATCAC		
Invertebrate COI-F	Forward	GGTCAACAAATCATAAAGATATTGG	About 709	Animal mitochondrial cytochrome c oxidase subunit I (COI)
Invertebrate COI-R	Reverse	TAAACTTCAGGGTGACCAAAAAATCA		
MCB398	Forward	TACCATGAGGACAAATATCATTCTG	About 472	Animal mitochondrial cytochrome b (CYTB)
MCB869	Reverse	CCTCCTAGTTTGTTAGGGATTGATCG		
16L1	Forward	CTGACCGTGCAAAGGTAGCGTAATCACT	About 505	Animal mitochondrial 16S ribosomal RNA (16SrRNA)
16H1	Reverse	CTCCGGTCTGAACTCAGATCACGTAGG		

A2. PCR conditions for different primer pairs

PCR components[1]	**UnivP/UnivQ**	**Invertebrate COI-F/ Invertebrate COI-R**	**MCB398/ MCB869**	**16L1/16H1**
DNA template[2]	At least 20 ng	At least 30 ng	At least 20 ng	At least 20 ng
10X High Fidelity PCR Buffer (Minus Mg)	1 X	1 X	1 X	1 X
$MgSO_4$	2.0 mM	2.0 mM	2.0 mM	1.5 mM
dNTPs	0.2 mM @ dNTP	0.2 mM @ dNTP	0.2 mM @ dNTP	0.1 mM @ dNTP
Forward primer	0.4 μM	0.4 μM	0.2 μM	0.2 μM
Reverse primer	0.4 μM	0.4 μM	0.2 μM	0.2 μM
High Fidelity *Taq* DNA Polymerase	1.25 U per reaction	1.0 U per reaction	1.0 U per reaction	1.25 U per reaction
Sterilized 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.

A3. PCR cycling condition for amplification of target amplicons

A. Animal DNA QC test with UnivP/UnivQ primer pair.

Temperature	Time	No. of cycles
95°C	5 min	1
94°C 55°C 72°C	30 sec 30 sec 1 min	30
72°C	7 min	1
4°C	∞	--

B. Mitochondrial cytochrome c oxidase subunit I with Invertebrate COI-F/Invertebrate COI-R primer pair.

Temperature	Time	No. of cycles
94°C	1 min	1
94°C 45°C 72°C	1 min 1.5 min 1.5 min	5
94°C 50°C 72°C	1 min 1.5 min 1 min	35
72°C	5 mins	
4°C	∞	--

C. Mitochondrial cytochrome b with MCB398/MCB869 primer pair.

Temperature	Time	No. of cycles
95°C	5 min	1
94°C 52°C 72°C	30 sec 30 sec 1 min	35
72°C	7 min	1
4°C	∞	--

D. Mitochondrial 16S ribosomal RNA with 16L1/16H1 primer pair.

Temperature	Time	No. of cycles
94°C	5 min	1
94°C 50°C 72°C	45 sec 45 sec 1 min	35
72°C	5 min	1
4°C	∞	--