Molecular techniques for identification of rice insect pests

L Behera, R S Panda, S L Mohapatra, S K Mohanty and S D Mohapatra

Accurate identification of insect and pathogen species, especially those are difficult to differentiate is one of major objective of plant protection. DNA barcoding is one of the promising and cost effective molecular technique for rapid and accurate identification of insect pests. This technique uses a short DNA sequence from a standard and agreed-upon position in the genome. The technique includes four major steps:

A) Isolation of genomic DNA from insects:

Reagents and solutions

i) 1.0 M Tris-HCl (pH 8.0)

Trizma base = 12.11 g

Dissolve in sterile de-ionized water, adjust pH to 8.0 with conc. HCl, make up volume to 100 ml with de-ionized water and autoclave at 15 Psi for 15 min.

ii) 0.25 M EDTA (pH 8.0)

EDTA (disodium salt; Mw = 372.3) = 9.31 g

Dissolve, adjust pH to 8.0 with 5 N NaOH, make up volume to 100 ml with de-ionized water and autoclave at 15 Psi for 15 min.

iii) 5.0 M NaCl

NaCl = 29.2 gm

Dissolve, make up volume to 100 ml with de-ionized water and autoclave at 15 Psi for 15 min.

iv) Extraction buffer

1M Tris-HCl (pH 8.0) = 10 ml
0.25 M EDTA (pH 8.0) = 8 ml
5 M NaCl = 28 ml
1.5% CTAB (w/v) = 1.5 gm

Dissolve, make up to 100 ml with de-ionized water and autoclave at 15 Psi for 15 min.

v) Chloroform: iso-amyl alcohol Mixture (24:1)

Chloroform = 96ml
Iso-amyl alcohol = 4ml

vi) 70% Ethanol (100ml)

Absolute alcohol = 70 ml
Double distilled water = 30 ml

vii) **RNase stock**

1 M Tris-HCl (pH 8.0) = 100 µl
5 M NaCl = 300 µl
RNase = 10 mg (Sigma)

Adjust volume to 1 ml with de-ionized water and storing at -20^{0}C.

viii) **TE (10: 1) (pH 8.0)**

1 M Tris- HCl (pH 8.0) = 1 ml
0.25 M EDTA (pH 8.0) = 0.4 ml
Dissolve, make up volume to 100 ml with de-ionized water and autoclave at 15 Psi for 15 min.

ix) **10X TBE (pH 8.0)**

Trizma base = 108 gm
Boric acid = 55 g
EDTA = 9.3 gm
Dissolve, adjust pH with conc. HCL and make up volume to 1000 ml with double distilled water.

x) **Loading Gel Dye (Bromophenol Blue)**

0.5 M EDTA (pH 8) = 200 ìl
Bromophenol blue = 2 mg
Add 5 ml sterile water and store in freezer.

Protocol

High molecular weight genomic DNA is isolated from female moths of yellow stem borer as follows:

(i) Soak single adult females (YSB) in 50 *m*l of extraction buffer for 10 minutes.

(ii) Homogenize the females in a 1.5ml of eppendorf tube with sterilized polypropylene pestle gently and thoroughly.

(iii) Add 350 ml of extraction buffer slowly by rinsing the pestle.

(iv) Add 10 ml of 10 mg/ml Proteinase-K and incubate at 37^{0}C for 1 hr.

(v) Add 400 ml of equilibrated phenol to the tube, invert 25-30 times and centrifuge at 12,000-14,000 rpm at room temperature for 10 minutes.

(vi) Then transfer the supernatant to a new tube with a pipette avoiding protein and debris layer.

(vii) Add 10 *m*l of 10 mg/ml RNase and incubate for 1 hr at 37^{0}C.

(viii) Then add 400 ml phenol / chloroform- isoamyl alcohol, invert 25-30 times and spin at 12,000-14,000 rpm at 4^{0}C for 5 minutes and transfer the supernatant to a new tube with plastic pipette and add 9 ml of 5M NaCl.

(ix) Add 2X volume of ice-cold absolute ethanol.

(x) Mix gently and leave for 2 hr or over night at –20^{0}C .

(xi) Spin at 12,000-14,000 rpm at 4^{0}C for l0 min.

(xii) Remove the supernatant with pipette taking care not to dislocate the DNA pellet.

(xiii) Add 200 ml of 70% ethanol to the DNA pellet. Leave for 5 minutes, spin at 12,000-14000rpm at 4^{0}C for 5 minutes and remove the supernatant.

(xiv) Add finally with 200 *m*l 100% ethanol, spin at 12,000-14000 rpm at 4^{0}C for 5 minutes and remove the supernatant taking care of DNA pellet.

(xv) Dry the pellet under vacuum for 30 min.

(xvi) Add 50-100 *m*l of TE to each pellet, mix well and leave at 4^{0}C overnight.

(xvii) Store DNA samples at -20^{0}C till further analysis.

Measurement of DNA concentration and quality checking

Both agarose gel electrophoresis and spectrophotometric methods are used to assess the quantity and quality of genomic DNA

a) Agarose gel electrophoresis

(i) Prepare 0.8% agarose gel. Add 0.8 g of agarose into a conical flask containing 100 ml of 1X TBE buffer

(ii) Heat the agarose mixture in the microwave for about 4-5 min, swirl the mixture and check to make sure that all of the agarose has melted. There should be no lumps or particles.

(iii) Allow the agarose to cool for several minutes and add 5 µl of Ethidium bromide (10 mg/ml) to it and mix well.

(iv) Pour the agarose solution into a sealed gel casting tray, insert a comb at one end of the tray to form the sample wells and allow to solidify for about 30-60 min

(v) After the gel has completely hardened, carefully remove the comb

(vi) Place casting tray into the electrophoresis chamber

(vii) Pour 500 ml of 1X TBE buffer into gel running tray. Pipette 4 µl. of each sample with loading buffer into wells

(viii) If necessary, include a size standard (Lambda/*ECoRI*) mixed with the loading buffer at least in one lane

(ix) Connect the electrical leads to the electrophoresis chamber

(x) Turn on the power supply and adjust the voltage level to 60V and allow to run for 1 hour.

(xi) Turn off the power supply and disconnect the leads.

(xii) The gel is now ready to be photographed. Wear goggles to protect eyes from UV exposure

(xiii) Photgraph gel using Gel-Doc system

The quantity and quality of genomic DNA is determined by comparing with molecular weight markers.

b) Spectrophotometric *analysis*

The concentration of DNA is estimated by the measurement of the UV irradiation absorbed by nucleic acid bases. The steps involved are as follows:

(i) Calibrate spectrophotometer using 2000 µl of TE in a quartz cuvette at 260 nm and 280 nm.

(ii) Add five µl of DNA sample to 1995 µl of TE, mix well and take absorbance (O.D) at 260 nm and 280 nm.

(iii) The concentration of the DNA in the sample is estimated by multiplying O.D at 260 nm with dilution factor and 50 Concentration of DNA (mg/ml) = O.D at 260 x Dilution factor x 50

(iv) The ratio between readings at 260 nm and 280 nm (O.D260 / $O.D_{280}$) provides an estimate for the purity of nucleic acid. Any sample showing the ratio below 1.8 or above 2.0 is further subjected to purification.

B) PCR amplification of barcode region

(i) Perform PCR amplification in a 50µl reaction mixture volume containing 2 µl (~30 ng) of DNA, 1X PCR buffer, 200µM dNTP mix, 10 pico-mole of both forward (*Lep*F1 5'-ATTCAACCAATCATAAAGATATTGG-3') and reverse primers (*Lep*R1 5'-TAAACTTCTGGATGTCCAAAAAATCA-3'), 2 mM of magnesium chloride, 1 U of *Taq (Thermos aquaticus)* DNA polymerase enzyme.

Preparation of reaction mixture for amplification of barcode markers

Reagent	Final concentration	Vol. in µl			
Sterile de-ionized water	-	32.0 X 10		=	275.0
10 X PCR buffer		1 X5.0 X 10		=	50.0
10mM dNTP mix		200 µM of each	1.0 X 10	=	10.0
25 mM magnesium chloride	2 mM		4.0 X 10	=	40.0
Primer	(F)	10 pM	2.5 X 10	=	25.0
	(R)	10 pM	2.5 X 10	=	25.0
Taq DNA polymerase		1U	1.0 X 10	=	10.0
			48.0 X 10	=	480.0

The reaction mixture (master mix) for PCR is prepared as above.

(ii) Mix well reaction mixture, distribute 48 µl to each of 10 tubes.

(iii) Add 2 µl DNA to each tube, mix well and briefly centrifuged to collect drops from wall of tube.

(iv) Load reaction tubes to thermocycler and carry out PCR amplification under following PCR conditions following temperature cycle:

Temperature cycling for amplification of barcode markers

Steps	Temperature (^{0}C)	Duration	Cycle (No.)
Initial denaturation	94	3 min	1
Denaturation	94	40 sec	40
Annealing	45	1 min	
Extension	72	1.5	
Final extension	72	5	1

(v) Briefly centrifuge to collect drops from wall of tube and store in the freezer till electrophoresis.

i) Agarose gel electrophoresis for separation for amplified products

The obtained PCR products are checked on 1% agarose gel before sequencing.

(i) Prepare 1% agarose gel. Add 1.8 g of agarose into a conical flask containing 180 ml of 1X TBE buffer

(ii) Heat the agarose mixture in the microwave for about 4-5 min, swirl the mixture and check to make sure that all of the agarose has melted. There should be no lumps or particles.

(iii) Allow the agarose to cool for several minutes and add 10 µl of Ethidium bromide (10 mg/ml) to it and mix well.

(iv) Pour the agarose solution into a sealed gel casting tray, insert a comb at one end of the tray to form the sample wells and allow to solidify for about 30-60 min

(v) After the gel has completely hardened, carefully remove the comb

(vi) Place casting tray into the electrophoresis chamber

(vii) Pour 500 ml of 1X TBE buffer into gel running tray. Load 10 µl of PCR products with 2 µl loading buffer into wells.

(viii) If necessary, include a size standard (100 bp ladder) mixed with the loading buffer at least in one lane.

(ix) Connect the electrical leads to the electrophoresis chamber.

(x) Turn on the power supply and adjust the voltage level to 80 V and allow to run for 3 hours.

(xi) Turn off the power supply and disconnect the leads.

(xii) The gel is now ready to be photographed. Wear goggles to protect eyes from UV exposure

(xiii) Photgraph gel using Gel-Doc system

(xiv) For samples showing clean, discrete PCR products proceed directly to sequencing

(xv) However, for samples with multiple bands or excessive smearing, excise the desired fragment from a gel under ultraviolet light and purify using a kit.

C) Sequencing Barcoding fragments

a) PCR product clean up

PCR products must be cleaned-up to remove salts, un-incorporated nucleotides and residual primers before sequencing. QIAEX II Agarose Gel Extraction Kit (Qiagen) is used to gel-purify PCR samples with modifications as follows:

(i) Electrophorese the remaining PCR products (~40 µl) on a subsequent gel.

(ii) Excise the desired band and place into a 1.5 ml microfuge tube.

(iii) Add 500 µl of Buffer QX1, vortex glassmilk suspension well (30seconds) and add 7 µl to samples.

(iv) Incubate samples at 55ºC for 1 hour to melt agarose.

(v) Spin samples at 13000 rpm for 1 minute.

(vi) Pour off supernatant. Add 500 µl of Buffer PE and re-suspend the pellets by vortexing.

(vii) Spin samples at 13000 rpm for 1 minute.

(viii) Pour off supernatant. Add 500 µl of Buffer PE and re-suspend the pellets again by vortexing.

(ix) Spin samples at 13000 rpm for 1 minute.

(x) Pour off supernatant. Invert tubes and air-dry pellets for 15-20 minutes.

(xi) Add 8 µl of ddH20, resuspend pellets and incubate at 55ºC for 10 minutes.

(xii) Spin samples at 13000 rpm for 1 minute. Pipet 7 µl of supernatant into clean microfuge tubes. Place aside.

(xiii) Add 7 µl of ddH20 to pellets, re-suspend and incubate at 55ºC for an additional 10 minutes.

(xiv) Spin samples at 13000 rpm for 1 minute. Pipet 6 µl of supernatant into the final microfuge tubes.

(xv) Electrophorese 3 µl of Qiaex cleaned products on a 1% TBE agarose gel to estimate its concentration.

(xvi) Use 1-5 µl (dependent on COI band intensity) of purified PCR product in the sequencing reaction (next section).

b) Sequencing Reaction(out sourcing)

The COX I bands of each sample are purified and bi-directionally sequenced following Sanger's Dideoxy method. Capillary sequencers (ABI PRISM® 377 and 373) provide a low-cost sequencing solution for labs that seek to analyze no more than 50 templates a day. However, for higher production goals and greater automation, a multicapillary instrument (3730 DNA analyzer, ABI, Hitachi) is critical.

(i) Set up a sequencing reaction according to the table below. Use 8 ng/100 basesof DNA for Qiaex cleaned or 1-2 µl (depending on intensity of the band) for directly sequenced PCR products.

Reagent

Dye terminator mix v3.1	= 0.25 µl
5 X Sequencing Buffer *	= 1.875 µl
10% trehalose	= 5 µl
10 µM Primer	= 1 µl
Deionised, sterile water	= 0.875 µl
DNA	= 1µl
Final Volume	= 10 µl

*5X Sequencing buffer is: 400 mM Tris-HCl pH 9 + 10 mM $MgCl_2$ or 5X ABI sequencing buffer.

(ii) Run the sequencing reactions in a thermocycler under the conditions at 96 °C for 2 min followed by 25 cycles of 30 second denaturation at 96ºC, a 15 second annealing period at 50ºC, and a 60ºC extension for 2minutes and 30 seconds. The final extension is at a temperature of 60ºC for 5 minutes.

(iii) After sequencing reaction is over, the labeled DNA fragments are run out on capillary gel to determine the sequence of the bases of the DNA. Since the DNA fragments have been terminated at different bases along the chain, they will run out on a gel and form bands of different lengths.

(iv) ABI (Applied Biosystems) Automated Sequencer performs electrophoresis inside capillary tubes and records the bases using a laser beam that detects fluorescently-labeled ddNTPs.

c) Sequence Editing

(i) The raw sequence should be filtered as per standard QC parameters. Based on the quality report, raw sequences are trimmed to remove ambiguous base calls in order to retain high quality sequence reads for further analysis.

(ii) The forward and the reverse Sequences are edited and assembled using software like SeqMan (DNA STAR package, DNAStar Inc., Madison, WI, USA), SeqScape® (Applied Biosystems), ChromasPro etc.

D) Sequence Analysis

(i) FASTA format of these sequences are used for species identification using BLAST search at NCBI(http://blast.ncbi.nlm.nih.gov) and species identification tool at BOLD(www.boldsystems.org).

(ii) Paste your DNA sequence into the search box, ***Enter Query Sequence*** or upload a file containing the sequence (.txt files).

(iii) Click the ***BLAST*** button to start your search

(iv) Each BLAST result is a pairwise comparison between your DNA sequence and the DNA sequences in an NCBI database.

(v) Select your first search result, either by clicking on the first/top red line, or scrolling down and clicking on the ***Max Score*** to the right of the ***Description***.

(vi) The parameters used for BLAST were search in nucleotide collection database with MEGABLAST search, which is the more appropriate for comparing a query to closely related sequences.

(vii) The Reference data sets are stored in a BLAST database for subsequent matching with query sequences. Up to 100 hits with at least 80% identity are returned for each query, which is identified as the species associated with its best hit (highest bit score). In case more than one species are associated the query's identification is considered uncertain.

(viii) Additionally, to check the performance of DNA barcoding, downloaded sequences of the same species from NCBI and BOLD database, submitted by others from different geographical areas.

(ix) Submit all sequences to GenBank.

(x) **DNA-BAR**. Reference data sets are converted to a matrix comprising presence/absence of distinguishers (sequence substrings) using the software 'degenbar'. Input parameters are as follows: distinguishers of length 5–50 nucleotides ('l-min 5', 'l-max 50'), up to 100 redundant distinguishers ('Redundancy 100'), GC content 0–100% ('MinCandidGC 0', 'MaxCandidGC100'), annealing temperature 0–100uC ('MinCandidTemp 0','MaxCandidTemp 100'), salt and DNA concentration 50 nM ('SaltConc 50', 'DNAconc 50'), and a maximum common substring weight of 100 ('MaxCommSubstrWt 100') (note that degenbar was originally designed to pick DNA probes).

In case of multi-locus DNA barcodes (i.e. Inga data set) loci in the reference alignment are separated by 50 'N' positions. The presence/absence matrix of distinguishers is then used as reference data set. Each query sequence is scored for presence/absence of distinguishers and identified as the species associated with the reference sequence with the greatest number of matching presence/absences. In case more than one reference sequence of the same species membership shared the greatest number of matches the query was identified as that species. In case reference sequences associated with different species shared the greatest number of matches identification was considered uncertain.

❑❑❑