Workshop of Diagnostic Training on Leafminers of Agriculture Importance

Bogor, 29 February-11March 2016

Guideline

Molecular identification of leafminers

Hari Sutrisno

Zoological Division, Research Center for Biology,

The Indonesian Institute of Science, Jl. Raya Jakarta- Bogor Km 46, Cibinong.

Telp. +62 21 8765056, Fax. +62 21 8765068, Email: hari.sutrisno@lipi.go.id

Organized by

In collaboration with

Contents

1. Collecting insect pests for molecular identification

DNA is repaired with a great efficiency in living cells (Stivers&Kuchta 2006), but this repair ceases upon the death of the organisms or preservations of a sample. The DNA in such samples degrades more or less strongly over time and often becomes inaccessible to genetic studies, depending on the conditions of the storage (Lindahl 1993; Lehmann &Kreipe 2001; Wandeler*et al*. 2003; Paabo*etal*. 2004).

It is often to keep the specimens in good condition for molecular identification, researchers preserving them into absolute ethanol (96%). The problem is that not all insect specimens can be preserved in ethanol especially for those which will loss their characters (scales or color patterns) such as moths and butterflies.Moreover, another disadvantage of using this method is that the specimens will damage and cannot be used for morphological studies. Therefore, we recommend preserving directly an apart of the body specimens into absolute ethanol for molecular studies and the restdeposited in Museum for morphological work or as voucher specimens. In the other hand, for specimens that will not damage when put into absolute ethanol such as dipteran and coleopteran can bi directly preserved within ethanol

2. DNA Extraction

DNA (deoxyribonucleic acid) extraction is a process with many vital scientific applications. In research and agriculture, its uses include DNA sequencing, detection of pathogen and pest. In the field of forensic science, it is used for identification of the dead and the living, as well as for crime scene analysis. Despite its sophisticated results, DNA extraction itself is actually quite simple, and basic extraction techniques work equally well in a research lab or at home.

2.1. Function

Extracting DNA is a three- to four-step operation, which involves:
1.Breaking the cells open, commonly referred to as cell disruption or cell lysis, to expose the DNA within. This is commonly achieved by grinding or sonicating the sample.
2.Removing membrane lipids by adding a detergent.
3.Removing proteins by adding a protease (optional but almost always done).
4.Precipitating the DNA with an alcohol, usually ice-cold ethanol or isopropanol.

Instruments that are commonly used in DNA extractions are a gel box, bead beater

and and a centrifuge. A gel box is used to separate DNA by sending charges through agarose gel. A bead beater is used to lyse or break apart cells to access DNA and a centrifuge spin at speeds of more than 15,000 rpm to help separate DNA in different phases of the extraction.

2.2. General Extraction

Once a sample is obtained, it must be broken down to obtain the genetic material inside the individual cells. In the lab, a sample is typically placed in a device called an Eppendorf tube. A special solution is then added to the tube, and the tube is placed in a warm water bath. The purpose of the solution is to lyse (break down) the material's cellular structure. It contains two key ingredients: a specially designed detergent and an enzyme called proteinase K. Once in the warm water bath, the detergent eats away the sample's cellular membrane as well as the nuclear membrane surrounding the cell's genetic material. Once these membranes are disrupted, the proteinase K breaks apart a protein called histone, which is wrapped around the DNA.

The Eppendorf tube is then removed from the water bath, and a concentrated salt solution is added to clump together the unwanted protein and cellular debris. The tube is then placed in a small centrifuge, where centrifugal force leaves the DNA diffused in a layer of solution above the heavier excess material. The DNA is then removed and placed by itself in another tube. Isopropyl alcohol is added to the DNA and mixed thoroughly. This process forces the DNA out of the solution, clumping it into visible strands. The material is then put through the centrifuge one more time to force the DNA strands together. The alcohol is removed, and the DNA is allowed to dry. Once this process is completed, the resulting DNA sample can be stored and used for any one of its many purposes. There are many Extraction KITs that are available in the market for insect, for an example: **DNeasy® Blood & Tissue**(for detail see Handbook)

2.3. Non-destructive method

A new method of DNA extraction for museumspecimens has been reported to be successful to get DNAfrom only 50 bp to thousands bp without any morphological damages. This method is well known as anon destructive method. It is able to extract DNA from fossil specimens (Gilbert *et al*. 2007). One of the benefitsof using this method is saving funds for providing thefresh materials. Collecting fresh materials by doingexploration is not only time and budget consuming.Moreover, a lot of organism was not be able to berecovered again since their habitats have changed or loss.The other benefit of using this method is that it can beused to get the DNA from the extinct creature which isusually their type materials remain

in the museums. Onthe other hand, using conventional method, destructivemethod, usually need fresh materials and it does not workfor all specimen museums. It is often to keep the specimensin good condition for molecular study, researcherspreserving them into absolute ethanol (96%). The problemis that not all insect specimens can be preserved in ethanol especially for those which will loss their characters (scalesor color patterns). Adults of moths and butterflies alwayspreserve as dry collection by pinning them. Moreover, another disadvantage of using this method is that thespecimens will damage and cannot be used for morphological studies. Therefore, we recommendpreserving directly an apart of the body specimens intoabsolute ethanol for molecular studies and the restdeposited in Museum.

We used a nondestructive method which is modifications from QIAGEN animal tissue protocol kit using spin column. Firstly, the abdomens were removed from the body then was placed into a sterile 1.5 ml microtube and added 0.1 ml proteinase K (PK) 1% Buffer (1% PK buffer = 20 ul proteinase Ksolution (20 mg/ml) in 180 ul buffer ATL (Qiagen). This abdomen then was incubated at 55 oC for 2-4 hours andwas added a further 0.1 ml PK buffer and incubated at55 oC overnight. The next morning the abdomen was removed for morphologically work. The tube containingthe insect mixture was then treated by following the manualof QIAGEN animal tissue protocol kit using spin column.

2.4. Destructive method

For DNA extraction from each moth individual, a thorax was ground in a 1.5 ml microcentrifuge tube containing 600 ul CTAB buffer and incubated at 55 oC for 2 hours. This solution was extracted three times using phenol saturated with TE buffer (10mM Tris-HCL, pH 8.0, 1 mM EDTA); firstly with onevolume of phenol: Chloroform: iso-amyl alcohol (25:24:1).The solution was again extracted twice with chloroform:iso-amyl alcohol. The aqueous phase was transferred toa new tube, and then 1.5 volume of isopropanol was addedto precipitate DNA and left at -20 oC for more than 1 hour.

The DNA precipitant was pelleted by centrifugation at13.000 rpm for 20 minutes. The DNA pellet was washedwith 70% ethanol, air dried and dissolved in 50 Ul of TEbuffer.

Composition for reagen CTAB:

2x CTAB (cetyltrimethylammonium bromide) extraction buffer

100 mM Tris-HC1 (pH 8.0)

1.4 M NaC1

20 mM EDTA

2% (w/v) CTAB

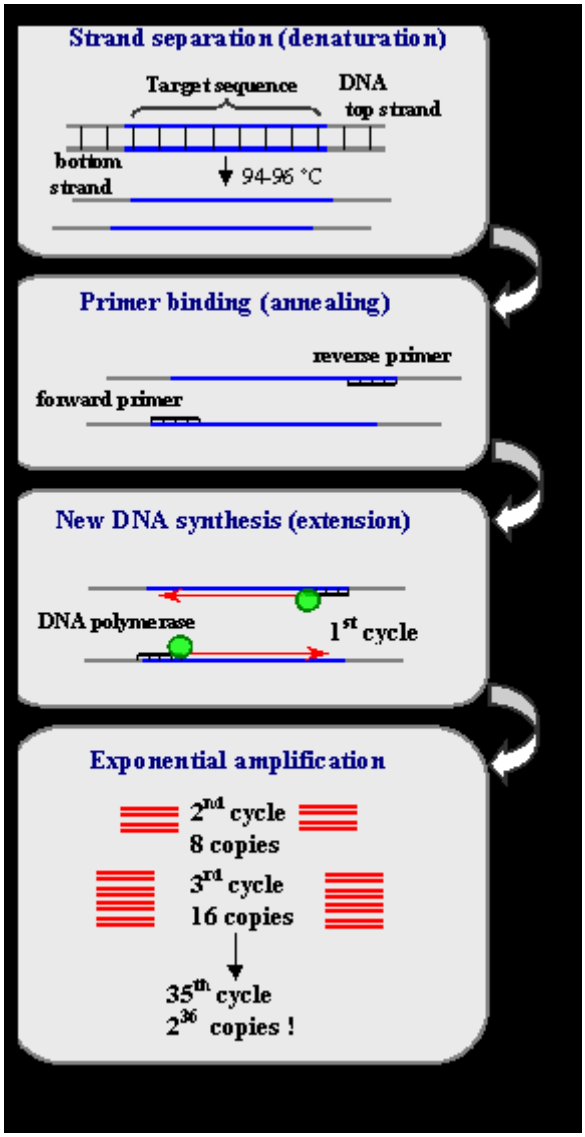
0.1% (w/v) PVPP (polyvinyl polypyrrolidine)

0.2% (v/v) B-mercaptoethanol (add directly before use, but do not store above with this).

3. PCR

PCR (Polymerase Chain Reaction) is a revolutionary method developed by Kary Mullis in the 1980s. PCR is based on using the ability of DNA polymeraseenzime to synthesize new strand of DNA complementary to the offered template strand. Because DNA polymerase can add a nucleotide only onto a preexisting 3'-OH group, it needs a primer to which it can add the first nucleotide. This requirement makes it possible to delineate a specific region of template sequence that the researcher wants to amplify. At the end of the PCR reaction, the specific sequence will be accumulated in billions of copies.

3.1. How It Works



(Sources: *www.ncbi.nlm.nih.gov*)

3.2. Components of the PCR reaction

DNA template: the sample DNA that contains the target sequence. At the beginning of the reaction, high temperature is applied to the original double-stranded DNA molecule to separate the strands from each other.

DNA polymerase: a type of enzyme that synthesizes new strands of DNA complementary to the target sequence. The first and most commonly used of these enzymes is *Taq* DNA polymerase (from *Thermis aquaticus*), whereas *Pfu* DNA polymerase (from *Pyrococcus furiosus*) is used widely because of its higher fidelity when copying DNA. Although these enzymes are subtly different, they both have two capabilities that make them suitable for PCR: 1) they can generate new strands of DNA using a DNA template and primers, and 2) they are heat resistant.

Primers: short pieces of single-stranded DNA that are complementary to the target sequence. The polymerase begins synthesizing new DNA from the end of the primer.

Lepidoptera

LepF1: 5'ATT CAA CCA ATC ATA AAG ATA TTG G 3', LepR1:5' TAA ACT TCT GGA TGT CCA AAA AATCA 3' (Hebert*et al*. 2010) and MLep F1: 5'GCT TTC CCA CGA ATA AAT AAT A 3' and MLep R1:5'CCT GTT CCA GCT CCA TTT TC 3'.

Diptera:

Nucleotides (dNTPs or deoxynucleotide triphosphates): single units of the bases A, T, G, and C, which are essentially "building blocks" for new DNA strands.

There are many PCR KITs that are available in the market for insect, for an example: **TopTaq™ PCR Qiagen** (for detail see Handbook).

3.3. PCR condition for Insect DNA amplification.

The amplification is conducted in thefollowing PCR conditions: one cycle of denaturation at94 oC for 10 min, followed by 35 cycles, with each cycleconsisting of denaturation at 92 oC for 30 sec., annealingat 47 oC for 30 sec, and extension at 72 oC for 1 min. 30 sec.These cycles were completed by final extension at 72 oCfor 10 min (Sutrisno 2012).

4. Electrophoresis

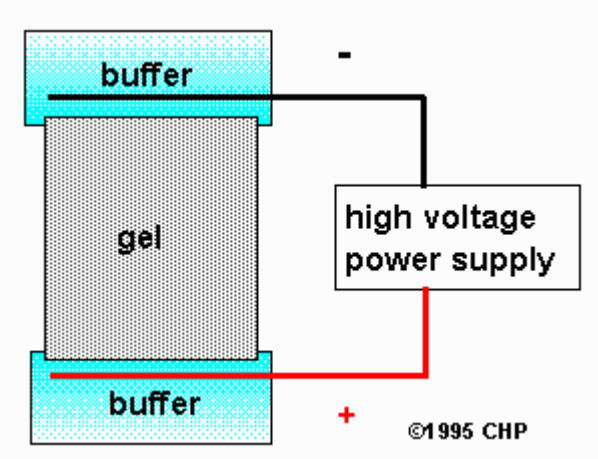
Electrophoresis is a separationstechnique that is based on the the mobility of ions in an electric field. Positively charged ions migrate towards a negative electrode and negatively-

charged ions migrate toward a positive electrode.For safety reasons one electrode is usually at ground and the other is biased positively or negatively. Ions have different migration rates depending on their total charge, size, and shape, and can therefore be separated.

4.1. Instrumentation

An electrode apparatus consists of a high-voltage supply, electrodes, buffer, and a support for the buffer such as filter paper, cellulose acetate strips, polyacrylamide gel, or a capillary tube. Open capillary tubes are used for many types of samples and the other supports are usually used for biological samples such as protein mixtures or DNA fragments. After a separation is completed the support is stained to visualize the separated components. Resolution can be greatly improved using isoelectric focusing. In this technique the support gel maintains a pH gradient. As a protein migrates down the gel, it reaches a pH that is equal to its isoelectric point. At this pH the protein is netural and no longer migrates, i.e, it is focused into a sharp band on the gel.

Schematic of zone electrophoresis apparatus



(Sources: http://elchem.kaist.ac.kr/vt/chem-ed/sep/electrop/electrop.htm)

4.2. Composition of reagent used

50x TAE (Tris-Acetate Buffer)

Dissolve 242 g of Tris base, 57.1 mL of glacial acetic acid, and 100 mL of 0.5 M EDTA (pH 8.0) in H_2O up to 1 liter.

The 50x TAE is the concentrated stock solution. Use lx TAE as working solution (0.04 M Tris-acetate, 0.001 M EDTA).

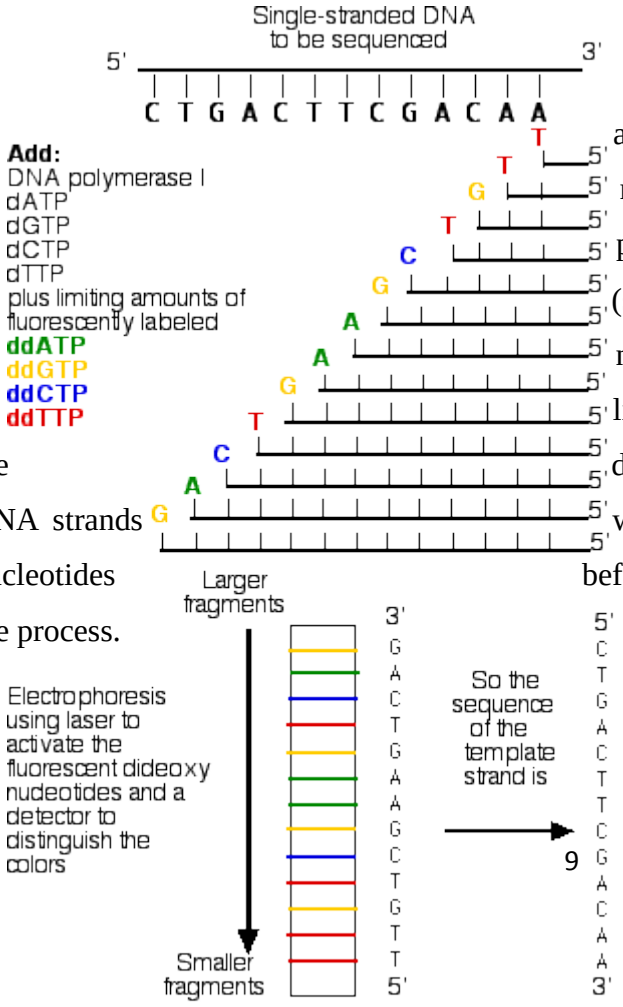
5. DNA Sequencing

DNA sequencing is the determination of the precise sequence of nucleotides in a sample of DNA. The most popular method for doing this is called the **dideoxy method** or Sanger method (named after its inventor, Frederick Sanger, who was awarded the 1980 Nobel prize in chemistry [his second] for this achievment).

The DNA to be sequenced is prepared as a single strand.

This template DNA is supplied with

- a mixture of all four **normal** (deoxy) nucleotides in ample quantities
  - dATP
  - dGTP
  - dCTP
  - dTTP
- a mixture of all four **dideoxy**nucleotides, each present in limiting quantities and each labeled with a "tag" that fluoresces a different color:
  - **ddATP**
  - **ddGTP**
  - **ddCTP**
  - **ddTTP**
- DNA polymerase I



Because all four normal nucleotides are present, chain elongation proceeds normally until, by chance, DNA polymerase inserts a dideoxy nucleotide (shown as colored letters) instead of the normal deoxynucleotide (shown as vertical lines). If the ratio of normal nucleotide to dideoxy versions is high enough, some DNA strands will succeed in adding several hundred nucleotides before insertion of the dideoxy version halts the process.

(Sources:

http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/D/DNAsequencing.html)

6. Alignment data

Data from the squence machine that we get usually is as a raw data squence. It isoften computer is confused to translate a certain peak in the graph when there is not clear peak or a double peak. More over, it is often that forward and revers ssquence is not complementary. Therefore, alignment is necesary to be edited manually.

After getting the data sequence from the machine, we can start to open and edit the data by using aligment software. There many avaliable software that is available free download that runing in windows computers such as Bioedit, Clustal W, Mega 5.1 and others. Each software has advantages and disadvantages. The followings are editing step that should be conducted after getting the squence data from the computer

1. Open the data, check the graph and the base squence to make sure that the squence is going well which is indicated by minimum number of untranslated peak of graph into base squence (the number of N is small) both on revers and forward data.
2. 0-30 base first, should not be used and omitted when not showing clear squence
3. Elignment editing can be started by checking the base sequence of the forward squence and than the revers complementary (please be carefuly when you alignment the pair wise squence, you will see there are many gaps that are not always indicate indels.
4. CO I is protein coding gen, each three bases indicated a amino acid. It means that within this squen there is no stop codon when the squence is translated into amino acid. Please check carefully when you find a stop codon in the midle of squence.
5. Please keep the original squence data before editing and save to a new file for ones which has finished to be edited.

7. Analysis data for identification

The final step in process of molecular identification is comparing the alignment sequence with the library of squence data or genbank by using BLAST. BLAST (Basic Local Alignment Search Tool)is an algorithm for comparing primary biological sequence information, such as the amino-acid sequences of different proteins or the nucleotides of DNA sequences. A BLAST search enables a researcher to compare a query sequence with a library or database of sequences, and identify library sequences that resemble the query

sequence above a certain threshold. Different types of BLASTs are available according to the query sequences. The BLAST program was designed by Altschul*et al*. (1990)

BLAST output can be delivered in a variety of formats. These formats include HTML, plain text, and XML formatting. For NCBI's web-page, the default format for output is HTML. When performing a BLAST on NCBI, the results are given in a graphical format showing the hits found, a table showing sequence identifiers for the hits with scoring related data, as well as alignments for the sequence of interest and the hits received with corresponding BLAST scores for these. The easiest to read and most informative of these is probably the table.

BLAST can be used for several purposes. These include identifying species, locating domains, establishing phylogeny, DNA mapping, and comparison.

Identifying species

With the use of BLAST, you can possibly correctly identify a species and/or find homologous species. This can be useful, for example, when you are working with a DNA sequence from an unknown species.

Locating domains

When working with a protein sequence you can input it into BLAST, to locate known domains within the sequence of interest.

Establishing phylogeny

Using the results received through BLAST you can create a phylogenetic tree using the BLAST web-page. Phylogenies based on BLAST alone are less reliable than other purpose-built computational phylogenetic methods, so should only be relied upon for "first pass" phylogenetic analyses.

DNA mapping

When working with a known species, and looking to sequence a gene at an unknown location, BLAST can compare the chromosomal position of the sequence of interest, to relevant sequences in the database(s).

Comparison

When working with genes, BLAST can locate common genes in two related species, and can be used to map annotations from one organism to another.

References

Altschul S, Warren Gish, Webb Miller, Eugene Myers, and David J. Lipman 1990. Basic Local Alignment Search Tool.*Journal of Molecular Biology* 1990.

Gilbert MTP, Moore W, Melchior L, Worobey M. 2007. DNAextraction from dry museum beetles without conferringexternal morphological damage. *PLoS One* 2:e72. http://dx.doi.org/10.1371/journal.pone.0000272

Hajibabaei M, Janzen DH, Burns JM, Hallwachs W, Hebert PDN.2006. DNA barcodes distinguish species of tropicalLepidoptera. *Proc Nat Acad Sci USA* 103:968-971. http://dx.doi.org/10.1073/pnas.0510466103

Hebert PDN, de Waard JR, Landry JF. 2010. DNA barcodes for 1/1000 of the animal kingdom. *Biol Let* 6:359-362. http://dx.doi.org/10.1098/rsbl.2009.0848

Hebert PDN, Gregory TR. 2005. The promise of DNA barcodingfor taxonomy. *Syst Biol* 54:852-859. http://dx.doi.org/10. 1080/10635150500354886

Lehmann U, Kreipe H. 2001. Real-time PCR analysis of DNA and RNA extracted fromformalin-fixed and paraffin-embeddedbiopsies. *Methods* 25:409-418.http://dx.doi.org/10.1006/meth.2001.1263

Lindahl T. 1993. Instability and decay of the primary structure ofDNA. *Nature* 362:709-715. http://dx.doi.org/10.1038/362709a0

Paabo S, Poinar H, Serre D, Jaenicke-Despres V, Hebler J, RohlandN, Kuch M, Krause J, Vigilant L, Hofreiter M. 2004. Geneticanalyses from ancient DNA. *Annu Rev Genet* 38:645-679.http://dx.doi.org/10.1146/annurev.genet.37.110801.143214

Altschul S, Warren Gish, Webb Miller, Eugene Myers, and David J. Lipman 1990.*Journal of Molecular Biology*.

Stivers JT, Kuchta RD. 2006. DNA damage and repair. *Chem Rev*106:213-214. http://dx.doi.org/10.1021/cr0404949

Wandeler P, Smith S, Morin PA, Pettifor RA, Funk SM. 2003.Patterns of nuclear DNA degeneration over time-a case studyin historic teeth samples. *Mol Ecol* 12:1087-1093. http://dx.doi.org/10.1046/j.1365-294X.2003.01807.x