DOI:10.6661/TESFE.2001019 台灣昆蟲Formosan Entomol. 21: 237-245 (2001) 研究報告 Research report

Formosan Entomologist
Journal Homepage: entsocjournal.yabee.com.tw

One Approach to Insect Identification: Using DNA Characters to Identify Mixed Dipterans Simultaneously【Research report】

昆蟲鑑定新方法：運用 DNA 特徵同時鑑定數種混合的雙翅目昆蟲【研究報告】

Wen-Bin Yeh
葉文斌

*通訊作者E-mail : wbyeh@cc.kmu.edu.tw

Received: 2001/06/06 Accepted: 2001/08/17 Available online: 2001/09/01

Abstract

Molecular characters have recently been used to identify insect species. In this paper, I introduce an approach aimed directly at mixed samples, in which each species introduced into the sample can be identified simultaneously. If two species are mixed in a DNA template, it will generate two forms of DNA products of the two species after PCR amplification. The DNA products were cloned into a plasmid after PCR amplification and each DNA copy of the DNA products could be cloned and then analyzed to identify the species' origin. DNA extracted from different species of the dipteran was artificially mixed as a DNA template for PCR amplification. A primer set was designed from the conserved fragments of mitochondrial 16S rDNA for this study. Three different couplings of dipteran DNA mixtures were produced for this analysis. Results confirmed that mixed PCR products from mixed DNA templates can be differentiated by cloning. In addition, PCR-RFLP was employed to distinguish the mixed DNA product. PCR products from a mixture of two species were digested by specific restriction enzymes. Digestion maps of mixed products revealed the specific patterns of the species. Thus, based on an established data bank of nucleic acids, identification of the unknown insects could be made.

摘要

近年來常應用分子特徵於昆蟲的鑑定。本報告引介一種方法，可以將混合在一起的昆蟲物種分別鑑定出來。混有兩個物種的 DNA 模板，於 PCR 複製時會同時增殖出該 2 物種的產物，將產物送入質體後，每個質體僅可攜帶一條產物。不同質體選殖到不同物種的 DNA 產物。因此，分析這些質體的 DNA 特徵，即確知那些昆蟲物種存在。將不同蠅類的 DNA 混在一起，用一對從粒線體 16S rDNA 保守區域所設計之引子對，複製出 16S rDNA 片段。從三組混合兩種蠅類 DNA 的實驗顯示，透過選殖的步驟，的確可以將混合的 DNA 產物區分。另外，運用限制酵素分解 PCR 產物，也可以鑑定出混合樣品內的種類，限制圖譜顯示出不同的種類的圖譜。因此透過已建立的核酸資料庫，可以鑑定出混合標本內未知名的物種。

Key words: molecular identification, PCR-RFLP, dipterans, 16S rDNA.

關鍵詞: 分子鑑定、PCR-RFLP、雙翅目、16S rDNA

Full Text: PDF(0.45 MB)

下載其它卷期全文 Browse all articles in archive: http://entsocjournal.yabee.com.tw

台灣昆蟲 21: 237-245 (2001)
Formosan Entomol. 21:237-245 (2001)

One Approach to Insect Identification: Using DNA Characters to Identify Mixed Dipterans Simultaneously

Wen-Bin Yeh Department of Biology, Kaohsiung Medical University, 100 ShihChan 1st Rd., Kaohsiung 807, Taiwan, R.O.C.

ABSTRACT

Molecular characters have recently been used to identify insect species. In this paper, I introduce an approach aimed directly at mixed samples, in which each species introduced into the sample can be identified simultaneously. If two species are mixed in a DNA template, it will generate two forms of DNA products of the two species after PCR amplification. The DNA products were cloned into a plasmid after PCR amplification and each DNA copy of the DNA products could be cloned and then analyzed to identify the species' origin. DNA extracted from different species of the dipteran was artificially mixed as a DNA template for PCR amplification. A primer set was designed from the conserved fragments of mitochondrial 16S rDNA for this study. Three different couplings of dipteran DNA mixtures were produced for this analysis. Results confirmed that mixed PCR products from mixed DNA templates can be differentiated by cloning. In addition, PCR-RFLP was employed to distinguish the mixed DNA product. PCR products from a mixture of two species were digested by specific restriction enzymes. Digestion maps of mixed products revealed the specific patterns of the species. Thus, based on an established data bank of nucleic acids, identification of the unknown insects could be made.

Key words: molecular identification, PCR-RFLP, dipterans, 16S rDNA.

Introduction

Adult morphological characters are most commonly used for insect identification. However, it is difficult to identify a species of insects if there is no clear identified key for that taxon. Usually, if a specimen can not be identified by an indigenous expert, it is sent to a foreign expert to be identified. Recently, molecular techniques have been introduced which are helpful in identifying species of insects, especially for determining population diversity (Mukha *et al.*, 2000; Palmer *et al.*, 2000), sibling species (Kelley *et al.*, 2000; Yeh *et al.*, 2000), and species complexes (Yeh *et al.*, 1997; Guillet *et al.*, 2000). These techniques provide precise and rapid methods to identify populations and

*Correspondence address
e-mail:wbyeh@cc.kmu.edu.tw

species. Consequently, the systematist does not have to understand detailed morphological, ecological, or behavioral information about the unknown species. Based on the established data bank of nucleic acids, the systematist can identify the relevant species.

There are four kinds of DNA characters commonly used to distinguish species or individuals: (1) RFLP (restriction fragment length polymorphism), (2) RAPD (random amplified polymorphic DNA), (3) satellite DNA, and (4) DNA sequences (Hoy, 1994; Yeh, 1999). DNA characters can be obtained from each population or species individually. Populations or species with high affinity have more DNA characters in common than do those of low affinity. However, most of these methods are based on detection of only one species, while different insects may occur in a given sample at the same time.

A further application is introduced here, aimed directly at mixed samples, by which each species can simultaneously be separated from mixed DNA samples. The presence of mixed DNA in a sample is detected by PCR amplification. Copies of DNA products are then cloned into plasmids, and each plasmid carries only one copy of DNA. Thus, mixed DNA copies are separated after the cloning. Nucleic acids of different dipterans were artificial mixed from their crude DNA. Conserved paired primers for samples were designed (Yeh *et al.*, 1997) and replicated by PCR. Results revealed that the mixed samples can be identified unambiguously after cloning and sequencing. In addition, a restriction map was analyzed directly to distinguish the mixed PCR products. PCR-RFLP characters identified mixed samples more rapidly than did DNA cloning and sequencing. Applications to the identification and classification of mixed samples are discussed in the text.

Materials and Methods

DNA Amplification and Direct Sequencing

Samples of four dipterans were used in this paper: *Bactrocera dosalis* (Hendel), *Gastrozona fasciventris* (Macquart), *Euthyplatystoma sauteri* Hendel, and *Homoneura* sp. DNA was obtained using phenol-chloroform extraction. Extraction of crude DNA was described by Yeh *et al.* (1997). Mid-legs from one individual of four different dipterans were separately homogenized. A part of the 3′ end of the 16S rDNA gene was amplified by polymerase chain reaction (PCR). Paired primers of 16SR21 and 16S22, and amplification procedures to amplify the target region were described by Yeh *et al.* (1998). The purification of double-stranded DNA products followed Jean *et al.* (1995), and DNA products were directly sequenced using S^{35} with a PCR sequencing kit (Perkin Elmer) for 29 cycles with the following temperature profile: 50 s each for denaturation at 95 °C, annealing at 50 °C, and extension at 72 °C.

DNA Cloning and Sequencing

An aliquot of DNA was mixed artificially using two of the four dipterans. The mixed PCR product was purified then ligated into the Advantage PCR cloning kit (Clontech). A ligated vector was transformed into *E. coli*. Plasmids of about ten clones were selected and extracted. DNA was sequenced directly using S^{35} with a PCR sequencing kit (Perkin Elmer) for 29 cycles with the following temperature profile: 50 s each for denaturation at 95 °C, annealing at 50 °C, and extension at 72 °C. Universal primers including forward and reverse primers on the plasmid were used in DNA sequencing.

Sequence and PCR-RFLP Analysis

Sequences were aligned using the Pileup program of the GCG software package (Genetic Computer Group, vers. 7.0) (Devereux *et al.*, 1991), then manually confirmed. Restriction maps of *Ase* I, *Swa* I, and *Pac* I of the four dipterans were simulated using the mapplot program of the GCG software package. The proportion of the nucleotide composition of each taxon was calculated using the MEGA program (Kumar *et al.*, 1993). DNA templates from *B. dosalis* and *E. sauteri* plasmids were mixed in successive ratios (0, 1:20, 1:15, 1:10, 1:5, and 1:2) in the amount of 100 ng. After amplification, PCR products of 16S rDNA from mixed templates of *B. dosalis* and *E. sauteri* were digested by restriction enzyme, then a 2% gel electrophoresis was run.

Results

16S rDNA Sequence Divergence among Four Dipterans

DNA sequences of the four dipteran species are shown in Fig. 1. Base compositions among the four species are similar. Average base compositions of G, A, T, and C are 16.4, 33.4, 39.2, and 11.1%, respectively. The results are similar to the compositions of other dipteran species (Xiong *et al.*, 1993; Tang *et al.*, 1995; McPheron and Han 1997). Sequence divergences among these four species ranged from 0.045 to 0.14. The DNA sequence of the oriental fruit fly in Hawaii has been reported (Han and McPheron, 1997). Sequence comparison of the two oriental fruit flies from different populations reveals that only two bases differ in the amplified region, which is reasonable for intra-specific divergence (Yeh *et al.*, 1997).

Identification of Species from Mixed Samples by Cloning and Sequencing

DNA templates were artificially mixed from the four samples to determine whether or not it is possible to clearly identify species from mixed samples. Three different treatments were conducted for this analysis. Mixed DNA from *E. sauteri*-*H.* sp., *B. dosalis*-*E. sauteri*, or *B. dosalis*-*G. fasiventris* was used as templates for PCR amplification. Six clones in *E. sauteri*-*H.* sp. were selected for sequencing. Figure 2 shows a portion of the image of the sequencing gel indicating that three clones belong to *E. sauteri* (clusters 2, 3, and 4), one clone (cluster 5) belongs to *H.* sp., and two clones (clusters 1 and 6) are unknown sequences. In addition, from the *B. dosalis*-*E. sauteri* samples, 15 clones were selected, of which 12 clones are from *B. dosalis* and 3 clones are from *E. sauteri*. Six clones of *B. dosalis*-*G. fasiventris* were selected, of which four clones belong to *B. dosalis*, one clone belongs to *G. fasciventris*, and one is an unknown sequence. Results confirm that mixed PCR products from mixed DNA templates can be differentiated after cloning.

Identifying Individual Species from Mixed Samples by PCR-RFLP

Mixed samples can also be detected directly by restriction enzyme digestion. Figure 3 shows the simulated restriction maps from sequences of the four species. Enzymes *Ase* I, *Swa* I, and *Pac* I digested the amplified products into distinguishable maps. For example, *Ase* I recognizes the ATTAAT sequence. In the species, *B. dosalis*, the PCR product was digested into the fragments of 78, 201, and 275 bp (Fig. 4A, lane 2). Digested bands in *E. sauteri* were in three fragments of 29, 172, and 349 bp (lane 12). The PCR products of both *B. dosalis* and *E. sauteri* were digested by *Swa* I into two different bands (Fig. 4B,

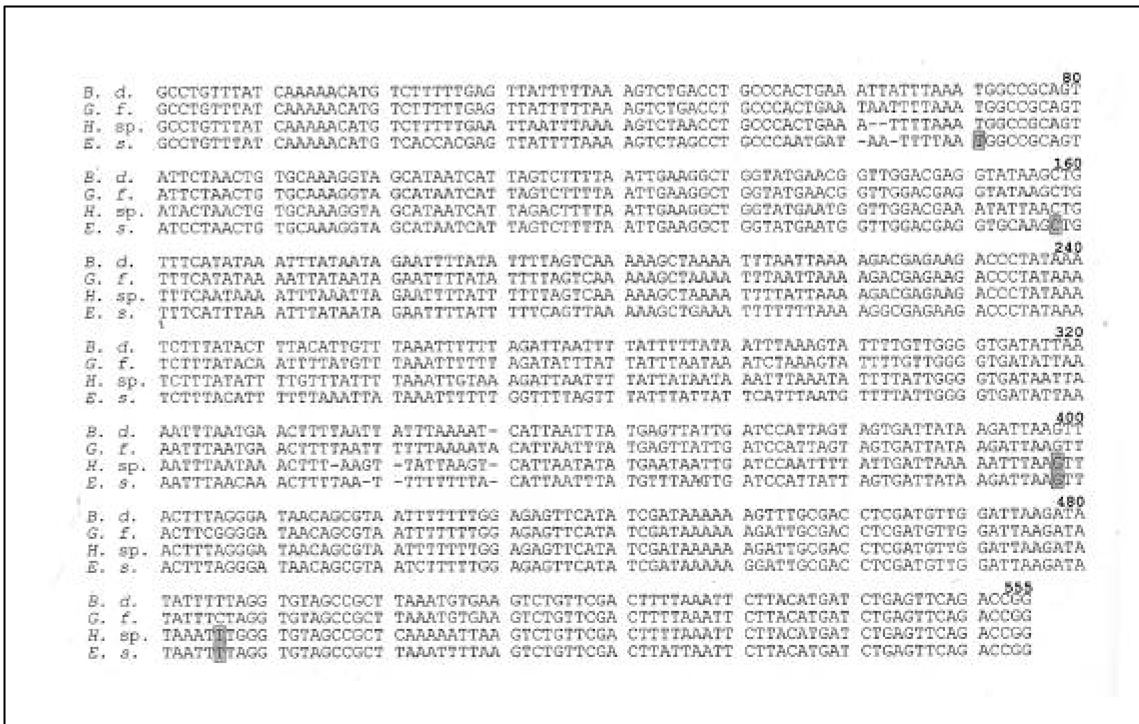

Fig. 1. Alignment of 16S rDNA sequences. Gaps inserted to improve alignment are indicated by hyphens according to the secondary structure (Davis *et al.*, 1994). Abbreviated taxa names are given in "Materials and Methods". Shaded nucleotides are in the relative position as shown in Figure 2.

lanes 2 and 12). *Pac* I digested the product from *B. dosalis* into two fragments (Fig. 4C, lane 12), but due to the lack of a recognition site, *E. sauteri* was not digested. Restriction maps from *Ase* I, *Swa* I, and *Pac* I reveal that the restriction fragment length can indeed be used as a discrete character to identify species. Furthermore, the amount of PCR product is directly related to the amount of mixed DNA template. DNA templates from *B. dosalis* and *E. sauteri* were mixed artificially in successive ratios of from 0 and 1:20 to 1:2. Restriction maps following PCR amplification reveal that *B. dosalis* and *E. sauteri* were amplified at the same time, and that the product increased in amplification as the DNA template ratio increased (Figs. 4A-C).

Discussion

Results from the DNA pools of different species reveal that cloning can be used to separate each species. However, cloning efficiency is not constant for each mixed sample. For example, 15 clones were selected in mixed samples of *B. dosalis* and *E. sauteri*, of which 12 clones belong to *B. dosalis* and three clones belong to *E. sauteri*. The unequal number of clones of each species in the mixed DNA samples may be due to the fact that the amount of DNA template differed. The optical density (OD) value reveals that the relative DNA amount of *B. dosalis*: *H.* sp.: *G. fasciventris*: *E. sauteri* was 6 : 1 : 1 : 2. Furthermore, there was a correlation between the amplification of the mixed DNA with different ratios of mixed DNA (Figs.

4A-C). However, there is a maximum limit to the PCR product. The PCR product increased as the template concentration increased to 50 ng, which was the maximum ratio, and higher template concentrations did not affect PCR production.

Cloning isolation has commonly been used in the analysis of viral quasispecies populations or persistent infections (Miralles *et al.*, 1999; Collins *et al.*, 1999; Zhou *et al.*, 2001). However, it has not been used for insect identification. In systematics, identification of mixed samples can be applied to pest quarantine, sympatric morphospecies, and incomplete specimens. DNA characters obtained either from each insect stage or each portion of that stage are the same. Properties of DNA characters are useful in pest quarantine. For example at present, if a larval form is found in imported agricultural products, quarantine officials have to wait until the larva develops into an adult. It may require several days or weeks to observe the emergence of the larva, or determine the precise name of the pest from an expert. Based on established DNA characters, quarantine officials do not have to understand detailed morphological, ecological, or behavioral information of the quarantined pest. If several quarantined pests are found in agricultural products, the DNA of these species can be extracted together in either the adult or larval stage. After PCR amplification and cloning, these species can be

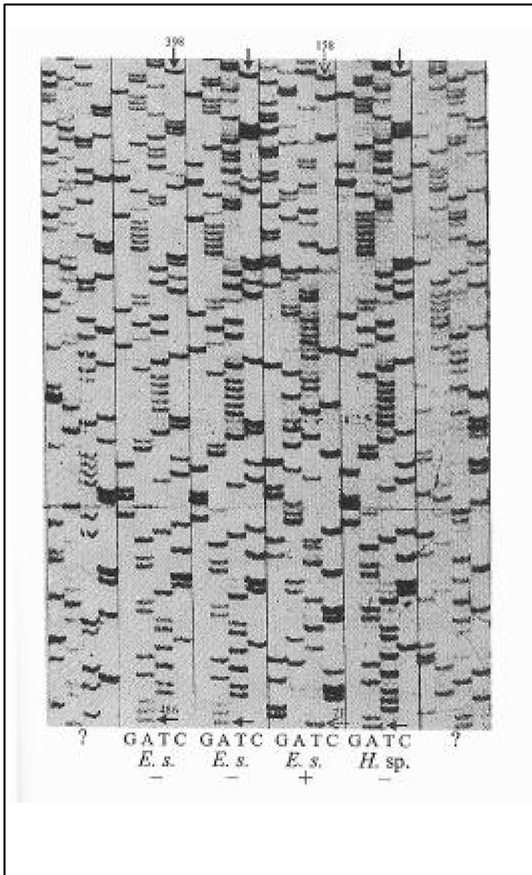

Fig. 2. Clones from *Euthyplatystoma sauteri* and *Homoneura* sp. mixed samples. Clusters 2, 3, and 4 are from *E. sauteri*, and cluster 5 is from *H.* sp., whereas clusters 1 and 6 are unknown clones (?). Signals "+" and -" indicate the upstream and downstream direction of the gene, respectively. The arrows indicate the relative position according to Figure 1.

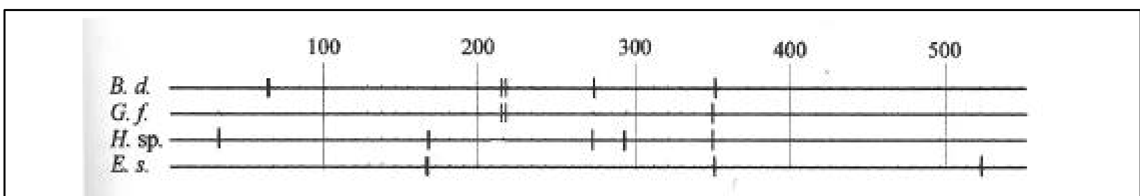

Fig. 3. Simulated restricted sites of the four dipterans. The mapplot program in the GCG software package was used to simulate the restricted sites. Three enzymes were selected to present the different restriction maps of the four dipterans. The bold bar is the *Swa* I site, the thin bar is the *Ase* I site, and the dashed bar is the *Pac* I site.

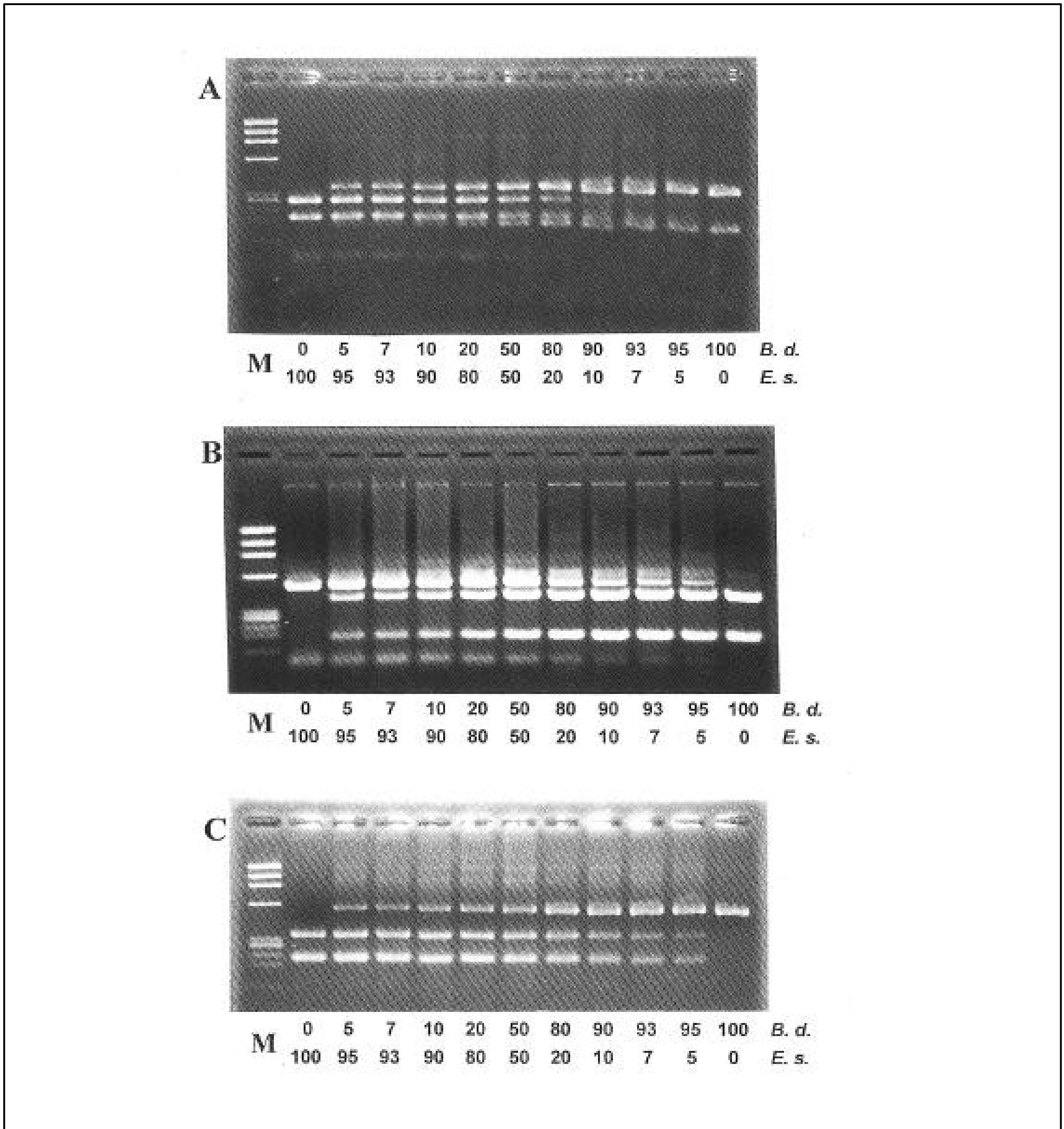

Fig. 4. Digestion maps of *Bactrocera dosalis* (*B. d.*) and/or *Euthyplatystoma sauteri* (*E. s.*) from PCR products, A (*Ase* I), B (*Swa* I), and C (*Pac* I). Relative template amounts of the mixed products from *B. d.* and *E. s.* are shown in the figure. The product increased with amplification as the DNA template ratio increased. M, marker; *ÖX*174 RF *Hae* III.

separated.

There are also other conditions in which specimens are not easy to identify. (1) If specimens are from sympatric morphospecies, their morphological characters are likely indistinguishable. More refined methods are essential to identify these kinds of specimens. (2) Improper storage in an alcohol solution may render it nearly impossible to identify which

portion belongs to each species. (3) Many insects are tiny and two similar species could easily be in the sample pool.

Cloning and sequencing indeed constitute an efficient method to identify mixed samples; however, it is time-consuming and expensive. PCR-RFLP is a more rapid and inexpensive analysis thereby reducing the cost and time for identification. PCR-RFLP is recommended for application as nucleic acid information is acquired.

According to a previously established data bank, either from sequencing or RFLP, DNA characters can be applied in the identification of one-species samples. However, this paper provides a further application. Results from either cloned DNA sequences or a map of PCR-RFLP reveal that DNA characters can be used to identify individual components of mixed-species samples simultaneously. Furthermore, successive ratios in mixed DNA templates reveal that the relative amount of DNA in the mixed sample correlates with the amount of PCR products.

Acknowledgments

Scientific names of the four dipterans were provided by Horng-Yih Chang of the Plant Protection Department, Bureau of Animal and Plant Health Inspection and Quarantine, Council of Agriculture, Taipei.

References

Collins, M. E., M. Desport, and J. Brownlie. 1999. Bovine viral diarrhea virus quasispecies during persistent infection. Virology 259: 85-98.

Davis, R. E., T. J. Kelly, E. P. Master, B. S. Thyagaraja, C. A. Rote, and R. B. Imberski. 1994. Complete base sequence for the mitochondrial large subunit ribosomal RNA of the gypsy moth *Lymantria dispar* (L.) Insect Mol. Biol. 3: 219-228.

Devereux, J., P. Haeberli, and P. Marquess. 1991. Genetic computer group manual, vers. 7.0. Univ. Wisconsin, Madison, WI.

Guillet, S., A. Guiller, J. Deunff, and M. Vancassel. 2000. Analysis of a contact zone in the *Forficula auricularia* L. (Dermaptera: Forficulidae) species complex in the Pyrenean Mountains. Heredity 85: 444-449.

Han, H.-Y. and B. A. McPheron. 1997. Molecular phylogenetic study of Tephritidae (Insecta: Diptera) using partial sequences of the mitochondrial 16S ribosomal DNA. Mol. Phyl. Evol. 7: 17-32.

Hoy, M. A. 1994. DNA amplification by the polymerase chain reaction: molecular biology mode accessible. pp. 203-244 *In*: M. A. Hoy, ed. Insect Molecular Genetics. Academic Press, London.

Jean, C. T., C. F. Hui, S. C. Lee, and C. T. Chen. 1995. Variation in mitochondrial DNA and phylogenetic relationships of fishes of the subfamily Sparinae (Perciformes: Sparidae) in the coastal waters of Taiwan. Zool. Stud. 34: 270-280.

Kelley, S. T., B. D. Farrell, and J. B. Mitton. 2000. Effects of specialization on genetic differentiation in sister species of bark beetles. Heredity 84: 218-227.

Kumar, S., K. Tamura, and M. Nei. 1993. MEGA: Molecular Evolutionary Genetics Analysis, vers. 1.01. Pennsylvania State Univ., University Park, PA.

McPheron, B. A. and H. Y. Han. 1997. Phylogenetic analysis of North American *Rhagoletis* (Diptera: Tephritidae) and related genera using mitochondrial DNA sequence data. Mol. Phyl. Evol. 7: 1-16.

Miralles, R., A. Moya, and S. F. Elena. 1999. Effect of population patchiness and migration rates on the adaptation and divergence of vesicular stomatitis virus quasispecies populations. J. Gen. Virol. 80: 2051-2059.

Mukha, D. V., A. P. Sidorenko, I. V. Lazebnaya, B. M. Wiegmann, and C. Schal. 2000. Analysis of intraspecies polymorphism in the ribosomal DNA cluster of the cockroach *Blattella germanica*. Insect Mol Biol. 9: 217-222.

Palmer, M. R., D. R. Smith, and O. Kaftanoglu. 2000. Turkish honeybees: genetic variation and evidence for a fourth lineage of *Apis mellifera* mtDNA. J. Hered. 91: 42-46.

Tang, J., L. Toe, C. Back, P. A. Zimmerman, K. Pruess, and T. R. Unnasch. 1995. The *Simulium damnosum* species complex: phylogenetic analysis and molecular identification based upon mitochondrially encoded gene sequences. Insect Mol. Biol. 4: 79-88.

Xiong, B. and T. D. Kocher. 1993. Phylogeny of sibling species of *Simulium venustrum* and *S. verecundum* (Diptera: Simuliidae) based on sequences of the mitochondrial 16S rRNA gene. Mol. Phyl. Evol. 2: 293-303.

Yeh, W. B., C. T. Yang, and S. C. Kang. 1997. Identification of two sibling species, *Ephemera formosana* and *E. sauteri* (Ephemeroptera: Ephemeidae), based on mitochondrial DNA sequence analysis. Chin. J. Entomol. 17: 257-268.

Yeh, W. B., C. T. Yang, and C. F. Hui. 1998. Phylogenetic relationships of the Tropiduchidae-group (Homoptera: Fulgoroidea) of planthoppers inferred through nucleotide sequences. Zool. Stud. 37: 45-55.

Yeh, W. B. 1999. Identification of mites by molecular methods. pp. 49-62 *In*: Chin. J. Entomol., Special Publ. No. 12. The Entomol. Soc. ROC, Taipei, Taiwan, ROC, pp. 175.

Yeh, W. B., C. L. Ho, C. F. Hui, and C. C. Ho. 2000. Genetic diversity of *Amblyseius longispinosus* and *A. womersleyi* (Acari: Phytoseiidae) using RAPD analysis. Chin. J. Entomol. 20: 335-345.

Zhou, H., G. J. Rainey, S. K. Wong, and J. M. Coffin. 2001. Substrate sequence selection by retroviral integrase. J. Virol. 75: 1359-1370.

Received June 6, 2001
Accepted Aug. 17, 2001

昆蟲鑑定新方法：運用 DNA 特徵同時鑑定數種混合的雙翅目昆蟲

葉文斌 **高雄醫學大學生物系**

摘　要

近年來常應用分子特徵於昆蟲的鑑定。本報告引介一種方法，可以將混合在一起的昆蟲物種分別鑑定出來。混有兩個物種的 DNA **模板，於** PCR **複製時會同時增殖出該** 2 **物種的產物，將產物送入質體後，每個質體僅可攜帶一條產物。不同質體選殖到不同物種的** DNA **產物。因此，分析這些質體的** DNA **特徵，即確知那些昆蟲物種存在。將不同蠅類的** DNA **混在一起，用一對從粒線體** 16S rDNA **保守區域所設計之引子對，複製出** 16S rDNA **片段。從三組混合兩種蠅類** DNA **的實驗顯示，透過選殖的步驟，的確可以將混合的** DNA **產物區分。另外，運用限制酵素分解** PCR **產物，也可以鑑定出混合樣品內的種類，限制圖譜顯示出不同的種類的圖譜。因此透過已建立的核酸資料庫，可以鑑定出混合標本內未知名的物種。**

關鍵字：分子鑑定、PCR-RFLP**、雙翅目、**16S rDNA