BioFormosa（2008）43(1): 11-16

Phylogenetic Reconstruction Using a EST Encoding Ribosomal Protein Gene from a Wild Silkworm, *Eriogyna pyretorum* in Northern Taiwan

Mei-Mei Kuan[1,3*], Cheng-Ping Kuan[2], Ming-Liang Li[1]

[1]Department of Life Science, National Taiwan Normal University
Taipei, Taiwan.

[2]Department of Cosmetics Management, Carnical Tien Nursing Collage
Taipei, Taiwan

[3]Center of Disease Control, Department of Health, Executive Yuan
Taipei, Taiwan

(Received: 16 May 2007, accepted: 3 September 2007)

ABSTRACT

We cloned and characterized a cDNA encoded partial ribosomal protein (40 S RP) from a wild giant silk worm, *Eriogyna pyretorum*, collected from northern Taiwan. The molecular phylogeny using these expressed sequence tag (EST) sequences from *Eriogyna pyretorum* and other organisms are consist with phylogeny based on morphological traits.

Key words: 40S ribosomal protein, phylogenetic reconstruction, Lepidopteran

Introduction

There are about 150 000 described species in order Lepidoptera, which is one of the four major insect orders in terms of species number. In order to promote research on this important order, the study of Lepidopteran genomics was launched by international consortium since 2001. Aim to increase the amount of molecular data of Lepidoptera, hereby first time we analyzed one of the expressed sequence tag (EST) sequences encoding a 40S ribosomal protein (RP) from *Eriogyna pyretorum*, a giant wild silkworm.

The RP family contains about 55 proteins in prokaryotes and 80 proteins in eukaryotes (Doudna and Rath, 2002) which are generally well conserved, small and highly basic. Ribosome subunits (60S and 40S) are complex assemblies of four RNAs and about eighty proteins. Each of them is essential for the building of the ribosome. Specific RP proteine–rRNA interactions play a key role in the biogenesis and function of eukaryotic ribosome (Wool et al, 1990). Despite the importance of the RP family, the entire set of RP genes of eukaryote species is described in a limited number.

In this study, we characterized for the first time a cDNA cloned from the giant silkworm, *E. pyretorum*, which is clustered based on the 40S ribosomal protein p40, that is also similar to laminin receptor or a multidrug resistance-associated protein antigen. The phylogenetic tree using this EST and those corresponding sequences from other animal organisms are further investigated. Whether this ribosomal protein could serve as a good tool for molecule-based phylogeny is discussed.

Materials and Methods

RNA Extraction

Fat bodies of fifth instars' larvae of *Eriogyna pyretorum* were isolated first. Poly (A+) RNA was prepared from the fresh fat bodies using a Quick prep mRNA purification kit (Pharmacia LKB). RNA concentration was determined by measuring the optical density at 260 nm. Eight *u*g of poly (A+) RNA were used for cDNA synthesis with a cDNA synthesis kit (Pharmacia LKB).

Reverse Transcription, PCR

*Corresponding author: Mei-Mei Kuan; FAX: 886-2-23945359; E-mail: mmei8888@yahoo.com.tw or kuan@cdc.gov.tw

mRNA samples were heated at 70℃ for 5 min and reverse-transcribed at 37℃ for 1 h in a 20 *u*l reaction mix (for 1 *u*g of total RNA) containing 40 units RNNase inhibitor, 0.1*u*g oligo dT_{20} and l0.5 mM of each dNTPs, 50 mM Tris-HCl pH 8.3, 75 mM KCl, 3 mM $MgCl_2$, 10 mM DTT and 200 units Moloney murine leukemia virus reverse transcriptase. Samples were stored at -70℃ until subjected to polymerase chain reaction (PCR) amplification. Using the P1 primer (5´-ATHYTDWBCTTCGTSTTCGC-3´) and M4 primer(5´-TTTTTTTTTTTTTTTTTCGAGGACTC G AGCTCAAGC-3´), the first strand of cDNA was constructed with 8 *u*g of total RNA. The following cycling conditions were used for the PCR: 94 °C, 30 s; 55 °C, 30 s; and 72 °C, 1 min (35 cycles). The PCR products were cloned into vectors. DNA sequencing of selected clones was carried out using ABI PRISM Big Dye Terminator Kit (Applied Biosystems, France) on an automated sequencer (Applied Biosystems 373A DNA sequencer) and sequences were compared with the GeneBank database using BLAST algorithm.

RP Signatures

The Prosite Database (http://www.expasy.ch/prosite/) proposes an amino acid signature for most of the RPs, that was systematically compared with the *E. pyretorum* one.

Phylogenetic Analysis

We used ClustalX (Thompson, 1997) to compile and generate multiple alignments of the cDNA sequences. Phylogenic analysis was done using Phylogeny Inference Package (PHYLIP) version 3.6 (Felsenstein, 1993). Evolutionary trees were constructed with the neighbor-joining method and bootstrap analysis (1000 replications). In an alternative analytical strategy, the alignments were further analyzed by Maximum-Parsimony (MP) in DAMBE 9 (Xia, 2001).

Results

The EST Sequence Encoding a RP p40 from Larvae of Eriogyna pyretorum

We had cloned a RP p40 EST with 285 b.p. in nucleotide length from *E. pyretorum*. It displayed a high similarity with those 40S ribosomal proteins p40 from other eukaryotic organisms (Fig. 1). At nucleotide level, the *E. pyretorum* RP p40 EST was found to display 83-84% similarities with those from other Lepidoptera species e.g. *Lonomia oblique* (gb_AY829753.1) and *Bombyx mori* (gb_AY769314.1), and 74-77% similarities with those from other species e.g. *Apis mellifera* (XM_393965.3), *Culicoides sonorensis* (AY752834.1), *Diaphorina citri* (DQ673408.1), *Gallus gallus* (XM_418817.2), *Mus musculus* (BC084677.1), *Strongylocentrotus purpuratus* (XM_001177924.1), *Taeniopygia guttata* (DQ213582.1), *Xenopus laevis* (AY730625.1). At amino acids level, this *Eriogyna* EST was also found corresponding to 65-160th residues of RP p40, revealed 92-96% similarities with those from other Lepidoptera species e.g. *Lonomia oblique* (protein_id AAV91367.1) and *Bombyx mori* (protein_id AAV34856.1), and exhibited 77-92% similarities to those from other species e.g. *Mus musculus* (XP_484006), *Strongylocentrotus purpuratus* (XP_792396), *Apis mellifera* (P_989068), *Drosophila yakuba* (AB032437.1).

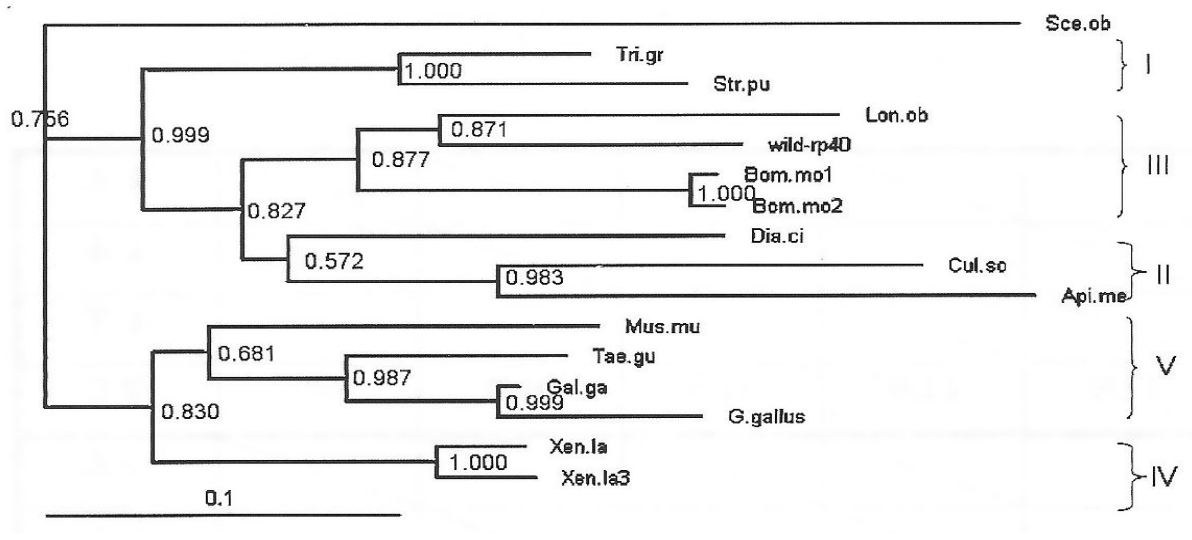
Figure 1. The EST sequence alignment of *Eriogyna pyretorum* RP P40 (wild-r), other Lepidopteran and distant species. Gray shadow amino acids indicate the residues of similarity.

RP Signature

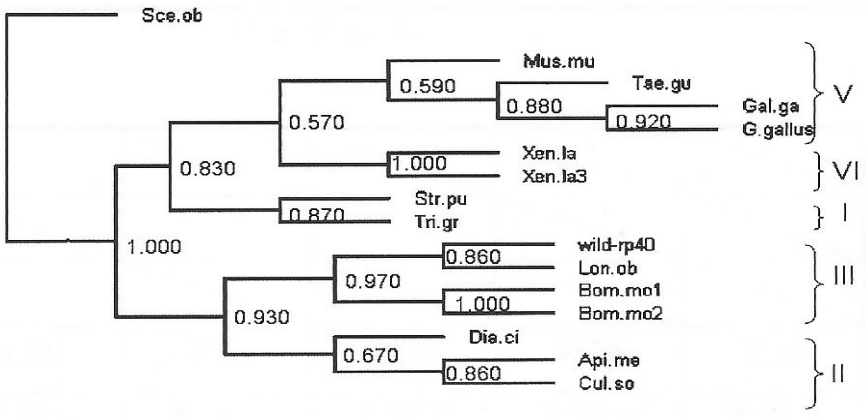
We have aligned the *E. pyretorum* EST encoded RP p40 with those of other eukaryotic species to investigate whether it can be a potential classification features. In Figure 1, conserved residues shared by all species are shaded in gray. By analyzing protein motifs, we detected one canonical signature of a ribosomal protein S2 in this EST encoded RP p40, i.e. signature : PrlLIVlDpaqDhqpItEasnyvNIP (55-79th amino acids).

The Phylogenetic Reconstruction Using EST Encoding 40S Ribosomal Proteins p40 Gene

Using the nucleotide sequences of a *E. pyretorum* RP p40 EST and those from other insects, invertebrate and vertebrate, the phylogenetic analysis had been performed. Both constructed neighboring and maximum parsimony trees displayed similar topological features: the divergences are observed not only between invertebrates, but also among vertebrates (Fig. 2). In the neighbor joint tree, the Lepidoptera including wild giant silkworm (*E. pyretorum*) could be grouped into a clade with a high bootstrap value (bootstrap value = 0.877) which moreover splits the Bombycoidea Saturniiforms e.g. *E. pyretorum, L. obliqueand* apart from the Bombycoidea



(A)



(B)

Figure 2. Phylogenic analysis using ribosomal protein p40 EST from *E. pyretorumh* and those from other Lepidopteran as well distant species by (A) Neighbor–Joining / UPGMA method version 3.65. (B) Maximum-Parsimony method. Bootstrap values are shown on branches, proportions based upon 1000 replicates. GeneBank Accession Nos. for the nucleotide sequences used, are: *Apis mellifera* XM_393965.3, *Bombyx mori* NM_001043415.1, *Bombyx mori* AY769314.1, *Culicoides sonorensis* AY752834.1, *Diaphorina citri* DQ673408.1, *Gallus gallus* XM_418817.2, *G.gallus* X94368.1, *Lonomia obliqua* AY829753.1, *Mus musculus* BC084677.1, *Scenedesmus obliquus* EC189174.1, *Strongylocentrotus purpuratus* XM_001177924.1, *Taeniopygia guttata* DQ213582.1, *Tripneustes gratilla* U02371.1, *Xenopus laevis* AY730625.1, *Xenopus laevis* BC130138.1. *Scenedesmus obliquus* is posited as a outgroup.

Bombyciformes e.g. *B .mor*. Overall, in RP p40 EST tree, the vertebrate species formed as a clade (bootstrap value = 0.830) that is well distinct from those of invertebrate also formed as a clade (bootstrap value = 0.999). Each of them might be further split into several sub clades e.g. insect (flies: *Diptera*, *Culicoides*, moths: Lepidoptera, beetles: *Coleoptera*, bees: *Hymenoptera*), urchin (*Echinoidea, Strongylocentrotus, Tripneustes*), amphibians (*Xenopus*), bird (*Gallu*, *Taeniopygia*) and mammal (*Mus*).

Discussion

By sequence alignment analysis, the *E. pyretorum* RP p40 EST sequences are more conserved than that at animo acids level. It implied that synomous variations occurred within P40 ribosomal proteins. In RP p40 EST tree, the clade Insecta in whole exhibits with a good robustness (bootstrap value = 0.827), and apart from other vertebrate species clade (bootstrap value = 0.999). The *E. pyretorum* is seen fall into Bombycoidea species clade under Lepidoptera. Besides, the invertebrate species clade splits apart from other vertebrate species clade with a good robustness (bootstrap value = 0.756) and does fit exactly with the morphological based phylogeny. We further reconstructed this phylogenic tree using parsimony method, the consistent relationship and similar topological pattern of tree is shown. We therefore suggest the RP p40 can be a potential phylogenetic marker.

Recently, molecule-based phylogeny most relies on ribosomal RNA (Adoutte *et al.* 2000 ; Whiting, 2002), or mitochondrial DNA sequences (Castro and Dowton, 2005) has been performed. The small subunit ribosomal RNA sequence remains the best choice for comparison owing to its enormous available database. In contrast, the use of ribosomal proteins for molecule-based phylogeny may suffer from some limitations due to their small size and possible structural rearrangements of these proteins (Liao and Dennis, 1994; Muller and Wittmann-Liebold, 1997). However, using ribosomal proteins as tools for analyzing evolutionary relationships between distantly related species still present several advantages. They are very conserved molecules occur in all eukaryotes (Wittmann-Liebold *et al*., 1990), and a growing number of them is being characterized among very divergent species (see GeneBank database). Despite analyses based on morphological, or individual molecular markers such as ribosomal RNA, or mitochondrial DNA sequences, some hardly elucidated relationships with sufficient confidence existed. Herein, we suggest the ribosomal protein could be a potential phylogenetic maker and can also be combined with other molecular markers including ribosomal RNA or mitochondrial DNA sequences to help for confirming and/or resolving particular nodes of the species phylogenetic tree. Results of this study suggests RP p40 EST may serve as an additionally available marker to elucidate the phylogenetic uncertainties.

References

Adoutte, A., Balavoine G., Lartillot N., Lespinet O., Prud´homme B., De Rosa R. 2000. The new animal phylogeny: reliability and implications. *Proc. Natl. Acad. Sci.* USA, 97: 4453-4456

Castro, L.R. and Dowton, M. 2005. The position of the Hymenoptera within the Holometabola as inferred from the mitochondrial genome of *Perga condei* (Hymenoptera: Symphyta: Pergidae). *Mol. Phylogenet. Evol.* 34: 469–479.

Doudna,J.A. and Rath,V.L. 2002. Structure and function of the eukaryotic ribosome: the next frontier. *Cell*, 109, 153–156.

Felsenstein, J. 1993. *PHYLIP* (Phylogeny Inference Package) version 3.5c. Seattle: University of Washington.

Liao, D., Dennis P.P. 1994. Molecular phylogenies based on ribosomal protein L11, L1, L10, and L12 sequences. *J. Mol. Evol.*, 38: 405-419.

Muller, E.C., Wittmann-Liebold B. 1997. Phylogenetic relationship of organisms obtained by ribosomal protein comparison. *Cell. Mol. Life Sci*., 53: 34-50

Philippe, H. 1993 MUST, a computer package for management utilitarians for sequences and trees. *Nucleic Acids Res.*, 21: 5264-5272

Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F, Higgins DG. 1997. The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. Dec 15; 25: 4876-82

Whiting, M.F. 2002. Phylogeny of the holometabolous insect orders: Molecular evidence. Zool. Scr. 31: 3–15.

Wittmann-Liebold, B., Köpke A.K.E., Arndt E., Krömer W., Hatakeyama T., Wittmann H.G. 1990. Sequence comparison and evolution of ribosomal proteins and their genes. Hill W.E., (Ed.) *The Ribosome: Structure, Function, and Evolution,* 598-616.

Wool, I.G., Endo Y., Chan Y.L., Gluck A. 1990. Structure, function, and evolution of mammalian ribosomes. Hill W.E., (Ed.). *The Ribosome: Structure, Function, and Evolution,* 203-214.

X. Xia, and Z. Xie. 2001. DAMBE: Software Package for Data Analysis in Molecular Biology and Evolution. The Journal of Heredity 92: 371-373.

生物學報（2008）43(1)：11-16

利用 40S 核醣體蛋白質基因片段進行臺灣野生蠶（*Eriogyna pyretorum*）之親源分析

關嬡嬡[1,3]* 關政平[2] 李銘亮[1]
[1]國立臺灣師範大學生命科學系
[2]耕莘健康管理專科學校化妝品應用與管理科
[3]衛生署疾病管制局

（收稿日期：2007.5.16，接受日期：2007.9.3）

摘　　要

我們首次報告有關採自臺灣北部山區的野生蠶(*Eriogyna pyretorum*)之核醣體蛋白質 40S RP40 的基因片段選殖，並以其序列與其他動物之同源基因序列進行分子親源性分析。

關鍵詞：40S 核醣體蛋白質、親源分析、鱗翅目

*通信作者：關嬡嬡（Mei-Mei Kuan）；FAX：886-2-23945359；E-mail：mmei8888@yahoo.com.tw or kuan@cdc.gov.tw