Cloning and Nucleotide Sequence Analysis of a *tuf* Gene from Loofah Witches' Broom Phytoplasma

Kuo-Chieh Ho(1)

(Manuscript received 25 September, 2001; accepted 4 November, 2001)

ABSTRACT: Using two primers and the enriched genomic DNA of a loofah witches' broom (LfWB) phytoplasma as the template, a 1651-bp DNA fragment was amplified in a PCR reaction. The DNA fragment contained two open reading frames (ORFs), one coding for part of the elongation factor EF-G (fus gene) and the other for the full-length EF-Tu (tuf gene). The tuf gene contained 1191 nucleotides encoding a 396-amino acid protein with a predicted molecular mass of 43.8 kDa. The sequence displayed a common characteristic of low G+C content of mollicutes and had a G+C content of 33.42%. It had a high degree of amino acid sequence homology, from 53.9 to 61.8%, with those of other bacteria except two archaea (lower than 26%). There was only one tuf gene locus on the genome. The phylogenetic relatedness of LfWB phytoplasma to other prokaryotes revealed by tuf DNA sequences was consistent with the result reported by 16S rDNA sequences.

KEY WORDS: Phytoplasma, tuf gene, Phylogenetic relatedness.

INTRODUCTION

Phytoplasmas previously were known as mycoplasma-like organisms, MLOs. In 1992, the Subcommittee on the Taxonomy of Mollicutes, the International Committee on Systemic Bacteriology suggested to change the name to phytoplasmas. The analyses of 16S rRNA and several ribosomal protein genes indicated that they were phylogenetically closer to acholeplasmas and anaeroplasmas, than to mycoplasmas (Tully, 1993). The organization of tRNA genes downstream the 5S rRNA gene also suggested that they were closer to acholeplasmas, and both evolved from *Bacillus subtilis* (Ho *et al.*, 2001).

Phytoplasmas are one group of plant bacterial pathogens. They are wall-less and obligate parasites that reside in the phloem of plants, and have a small genome of 500 - 1700 kbp and a relative low G+C content (Razin *et al.*, 1998). The failure of culturing these organisms *in vitro* makes it very difficult to study their biological and biochemical properties, and results in the uncertainty in their taxonomy. Although they are classified as a member of Class *Mollicutes*, they still cannot form a genus due to the lack of enough information (Sears and Kirkpatrick, 1994).

Elongation factor EF-Tu protein is considered to be one of the good molecular markers for studies of taxonomy and phylogeny because of its highly conservative nucleotide and amino acid sequences (Sela et al., 1989; Kamla et al., 1996). In addition, the copy number of the tuf gene is different in various bacteria. Gram-negative bacteria usually have two copies per genome, while most gram-positive bacteria such as B. subtilis and some Clostridium

^{1.} Department of Botany, National Taiwan University, 1, Sec. 4, Roosevelt Road, Taipei 106, Taiwan, ROC. Tel and Fax: +886-2-23637274; E-mail: kch@ccms.ntu.edu.tw

strains have only one copy (Sela et al., 1981). Exceptions are some Clostridium strains and Streptomyces ramocissimes. The former has more than one copy (Sela et al., 1989) and the latter has three copies (Vijgenboom et al., 1994). Mycoplasma galliseptium, Mycoplasma genitalium and Mycoplasma pneumoniae also only have one copy (Inamine et al., 1989; Loechel et al., 1989; Yogev et al., 1990). The context of the tuf gene on genome is also different in bacteria. Escherichia coli and most bacteria have a context of 5'-rps12-rps7-fus-tuf-3' (Cousineau et al., 1992). However, the tuf and fus genes of M. genitalium (Loechel et al., 1989), M. pneumoniae (Yogev et al., 1990), Sulfolous acidocaldarius, Pyrococcus woesei and Desulfuscoccus mobilis (Ceccarelli et al., 1995) are separated on genome.

In this report, the nucleotide sequence of a loofah witches' broom (LfWB) phytoplasma tuf gene was presented and used as a basis to discuss the phylogenetic relatedness of this organism.

MATERIALS AND METHODS

Bacteria and plant

LfWB phytoplasmas were maintained by graft inoculation in periwinkle (Chen and Ho, 1997). Diseased periwinkle was provided by Dr. C.-P. Lin, Professor of Department of Plant Pathology, National Taiwan University, Taiwan.

Plant and phytoplasma DNA isolation

The healthy or diseased plant DNA was isolated using the procedures described previously by Ho et al (2001): Fifteen grams of plant branches were disinfected with 1% sodium hyperchloric acid, rinsed with distilled water, and then dried by kimwipes briefly. Liquid nitrogen was added to a pre-cooled motor containing plant tissues. The tissues were grounded by a pestle and then by a coffee grinder. The powder was suspended in 15 mL of extraction buffer (500 mM NaCl, 100 mM Tris-HCl, pH 8.0, 10 mM EDTA, 1% N-lauroyl sarcosine) and incubated at 55°C for 2 h. After centrifuging at 10,000 xg at 4°C for 20 min, the supernatant was mixed with 0.6 volumes of isopropanol. The mixture was chilled at -20°C for 30 min. The DNA was precipitated at 10,000 xg at 4°C for 20 min. The pellet was then resuspended in 5 mL of TE (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 0.5% SDS and 500 µg proteinase k. After incubation at 37°C for at least 1 h, 875 µL of 5 M NaCl and 700 µL of CTAB/NaCl solution (10% CTAB, 0.7 M NaCl) were added. The incubation was continued for another 10 min at 65°C. The mixture was extracted 3 times with chloroform/ isoamylalcohol (24:1), and then twice with phenol/chloroform/isoamlyalcohol (25:24:1). The supernatant was mixed with 0.6 volumes of isopropanol, chilled at -20°C for 1 h, and then spun at 12,000 xg at 4°C for 20 min. The pellet was rinsed with 75% alcohol, air dried and then resuspended in 1 mL of TE.

The cesium chloride (CsCl)-bisbenzimide density gradient centrifugation was used to separate phytoplasma DNA from host plant DNA as following: The gradient contained DNA solution (in TE), CsCl (gram) and bisbenzimide (0.5mg/mL in TE) in a ratio of 9.5:10.5:1. After centrifugation at 55,000 rpm in p65VT3 rotor (Hitachi Koki Co., Ltd., Japan) at 20°C for 16 h, the phytoplasma DNA was visualized under UV light as the uppermost band in the

gradient of DNA from the diseased plant and was collected with a syringe. This DNA fraction was purified one more time through the gradient and extracted 3 times by water-saturated butanol. After the addition of three volumes of sterile distilled water, eight volumes of absolute alcohol and chilled at -20°C for 1 h, the DNA was precipitated. The DNA pellet was rinsed with 75% alcohol, dried and then dissolved in 500 µL of TE.

Polymerase chain reaction (PCR)

Two primers (Fus-1: 5'-AGATGTTAGAGCTATTTTATTTGATG-3' and Tuf-1: 5'-TCATTCTAATATCTCGATAACTGTTC-3') were synthesized based on our primary result (Chen and Ho, 1998) to amplify a DNA fragment using LfWB phytoplasma DNA as the template in a PCR reaction: The first cycle of the 100 μL PCR reaction was as follows: denaturation at 95°C, 5 min; annealing at 50°C, 1 min; extension at 72°C, 2 min. The reaction continued for 35 cycles as follows: denaturation at 95°C, 1 min; annealing at 50°C, 1 min; extension at 72°C, 2 min plus 3 sec/cycle. Finally, the reaction was performed at 72°C for another 5 min.

Blot hybridization analysis

Three micrograms of DNA were digested with different restriction enzymes including DraI, EcoRI, EcoRV and HindIII, and subjected to Southern blot analysis as previously described (Ho et al., 1992) using the full-length tuf DNA as probe. The hybridization was carried out at 42°C in the presence of 50% formamide and 0.1% SDS. Filter was washed in 0.1X SSC (1X SSC: 15 mM NaCl, 15 mM sodium citrate) containing 0.1% SDS at 50°C. The hybridized bands were detected by exposing the filter to a phosphorimager screen (PhosphoImager 425, Molecular Dynamics, USA).

DNA sequence determination and analysis

For DNA sequencing, the PCR product was cloned into pGEM-T easy vector (Promega). The sequences of DNA were determined by a DNA automated sequencer (ABI Prism Model 377, v. 3.0; Applied Biosystems) using a step-by-step procedure in which synthetic primers for forward sequencing were designed from sequences obtained previously. DNA sequence analysis was performed using DNAStar software (DNASTAR).

RESULTS AND DISCUSSION

Cloning strategy

A 1651-bp PCR DNA fragment was amplified on the LfWB phytoplasma DNA template using the primers Fus-1 and Tuf-1, and cloned into pGEM-T easy vector. The DNA fragment contained two open reading frames (ORFs), one coding for part of the elongation factor EF-G (fus gene) and the other for the full-length EF-Tu (tuf gene) (Fig. 1).

Analysis of the nucleotide sequence of tuf gene

The *tuf* gene contained 1191 nucleotides encoding 396 amino acids with a predicted isoelectric point of 6.42, and a molecular mass of 43.8 kDa. The gene had a G+C content of 33.42%, which was in agreement with the common characteristic of low G+C content of mollicutes, 31.13% to 44.6%, (Berg and Seemuller, 1999; Schneider *et al.*, 1997; Loechel *et al.*, 1989; Yogev *et al.*, 1990) and *B. subtilis*, 43.91%, the suspected ancestor of mollicutes (Weisburg *et al.*, 1989).

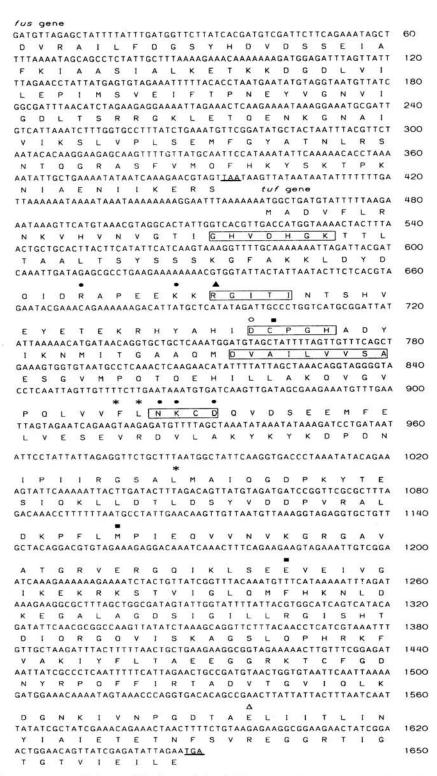


Fig. 1. Nucleotide sequences of the partial *fus* and the full-length *tuf* genes, and their deduced amino acid sequences. The sequences for GTP-binding were boxed. \bullet , the amino acid residues of phosphate group of GTP binding; \star , the amino acid residues of formation of a small pocket fitting the guanine ring; \circ , β -phosphate of GDP binding; \bullet , the amino acid residues interacting with tRNA; \bullet , the amino acid residues interacting with EF-Ts; \diamond , the amino acid residues of antibiotic kirromycin binding (Ludwig *et al.*, 1990; Yogev *et al.*, 1990). Two termination codons were underlined.

The low G+C content of LfWB phytoplasma *tuf* gene led to a preference for A or T in the second and third positions of the codon: 64.73% (second position) and 85.39% (third position), respectively. Of the 61 codons, 13 codons were not used including the one (UGG) for Trp. The use of UGA as a tryptophan codon, an unusual feature of mycoplasmas (Razin *et al.*, 1998), was not found in the LfWB phytoplasma *tuf* gene and the genes encoding an ABC transporter (unpublished data). The LfWB phytoplasma *tuf* gene contained all the hallmark signatures of an EF-Tu protein, including the amino acid residue(s) interacting with GTP, GDP, tRNA, elongation factor Ts and antibiotic kirromycin (Ludwig *et al.*, 1990; Yogev *et al.*, 1990) (Fig. 1).

Comparing the amino acid and nucleotide sequences of LfWB phytoplasma *tuf* gene with those of other prokaryotic cells listed in the legend of Fig. 3, the results showed that it had high amino acid sequence homology with those of eubacteria (53.9 to 61.8%) and much lower homology with those of archaea (lower than 26%). The nucleotide sequences of *tuf* genes were somewhat divergent. LfWB phytoplasma *tuf* gene had 38.9 to 65.4% homology with those of other eubacteria, and 23.5 to 28% with those of archaea.

There were 63 nucleotides between the *tuf* and its proceeding *fus* gene with a Shine-Dalgarno (SD) sequence, AGGA, the ribosome binding site (Fig. 1).

Number of tuf gene locus in LfWB phytoplasma

In addition to the low G+C content, the copy number of the *tuf* gene on the genome has been used to support that mollicutes are phylogenetically related to gram-positive bacteria (Woese, 1987). Thus far, only one copy of the *tuf* gene has been found in the genome of mollicutes (Inamine *et al.*, 1989; Loechel *et al.*, 1989; Yogev *et al.*, 1990). In an attempt to find out the number of *tuf* gene locus, LfWB phytoplasma genomic DNA was digested with each of the following restriction enzymes known not to cleave DNA within the *tuf* gene: *DraI*, *EcoRI*, *EcoRV* and *HindIII*, and probed by a ³²P-labeled *tuf* DNA fragment. There was only one single band hybridized in each restricted DNA, suggesting the presence of only one gene locus for EF-Tu (Fig. 2).

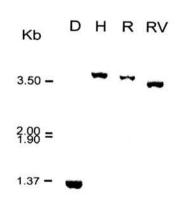


Fig. 2. Southern blot analysis of LfWB phytoplasma DNA. The DNA digested with DraI (D), EcoRI (R), EcoRV (RV) or HindIII (H) was probed with a ³²P-labelled tuf DNA fragment.

The gene context on the genome of mollicutes is variable. The *tuf* and *fus* genes are separated in the genomes of *M. genitalium* and *M. pneumoniae* (Inamine *et al.*, 1989; Yogev *et al*, 1990). These two genes were linked together in LfWB phytoplasma genome. The same result was reported for several other phytoplasmas (Berg and Seemuller, 1999).

Phylogenetic relatedness to other prokaryotes

Recently, the phylogeny of phytoplasmas and their relationship to other prokaryotes were actively investigated using the DNA sequence homology (Sela et al., 1989; Lee et al., 1993; Gunderson et al., 1994; Kamla et al., 1996; Schneider et al., 1997). Because of the nucleotide and amino acid sequence conservation, the tuf gene has been considered as one of the effective markers

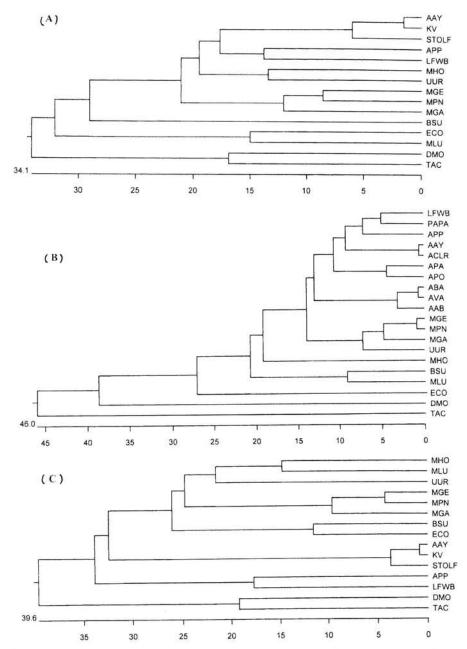


Fig. 3. Phylogenetic trees constructed based on (A) nucleotide sequence of tuf gene, (B) nucleotide sequence of 16S rRNA and (C) amino acid sequence of tuf gene using clustal method with weight table (DNASTAR). The scales in the bottom represent the branch lengths. The abbreviations for organisms are as followings: (a) tuf genes: AAY (L46368); APP (AJ011104); KV (L46369); STOLF (L46370); BSU, B. subtilis (D64127); DMO, D. mobilis (X73582); ECO, E. coli (U00006); MGA, M. gallisepticum (X16462); MGE, M. genitalium (X16463); MHO, M. hominis (X57136); MLU, M. luteus (M17788); MPN, M. pneumoniae (X55768); UUR, U. urealyticum (Z34275); TAC, T. acidophilum (X53866). (b) 16S DNA: AAY (X68373); ACLR (X68338); APP (AF248958); LfWB (AF251150); PAPAYA (Y08173); STLL (Y17055); AAB, Anaeroplasma abactoclasticum (M25050); ABA, Anaeroplasma bactoclasticum (M25049); APA, Acholeplasma palmae (L33734); APO, Acholoplasma polakii (AF031479); AVA, Anaeroplasm varium (M23934); BSU, B. subtilis (M10606); DMO, D. mobilis (M36474); ECO, E. coli (V00348); MGA, M. gallisepticum (L08896); MGE, M. genitalium (U39693, L43967); MHO, M. hominis (M96660); MLU, M. luteus (U85892); MPN, M. pneumoniae (M21374); UUR, U. urealyticum (L08642); TAC, T. acidophilum (M38637). AAY, ACLR, APP, KV, LfWB, PAPAYA, STOLF and STLL are the strains of phytoplasmas. The numbers in the parentheses are the DNA sequence accession number of the genes used.

in the study of bacterial phylogeny (Ludwig et al., 1993; Yogev et al., 1988). Figure 3A is a phylogenetic tree constructed by DNA sequences of tuf genes of LfWB phytoplasma and several other prokaryotes. It showed that the five phytoplasmas could be closely grouped together, but were separated from other mollicutes (M. galliseptium, M. genitalium, M. hominis, M. pneumoniae and Ureaplasma urealyticum). The data also showed that the mollicutes were closer to B. subtilis than to M. luteus and E. coli, and the two archaea (Thermoplasma acidoplilum and D. mobilis) were more diverged. Figure 3B was the phylogenetic tree based on the 16S rDNA sequences. It also showed that the phytoplasmas were closely grouped together, and that phytoplasmas were closer to acholeplasmas and anaeroplasmas than to other mollicutes. However, it was surprised that the tuf phylogenetic tree showed that mollicutes were more closely related to B. subtilis. But in the 16S rDNA phylogenetic tree, B. subtilis was grouped closely with M. luteus and E. coli, not with the mollicutes. This may be due to the fact that 16S rDNA sequences are more conserved among the organisms. When the amino acid sequences of the tuf genes were used, a different tree structure was obtained. The LfWB phytoplasma was grouped with apple proliferation (APP) phytoplasma, but not with other phytoplasmas (Fig. 3C). The amino acid sequence of the tuf gene may be too conservative to be used for phylogenetic study on the genus level, although it is still good for reflection of phenotypic and metabolic characteristics within the phytoplasmas (Kamla et al., 1996).

ACKNOWLEDGEMENTS

Author would like to thank Dr. Te-Chang Lee of the Institute of Biomedical Science, Academia Sinica, Taiwan, ROC for providing aid in DNA autosequencing. The work was supported by a grant (NSC 85-2311-B-002-042) from National Science Council, Taiwan.

LITERATURE CITED

- Berg, M. and E. Seemuller. 1999. Chromosomal organization and nucleotide sequence of the genes coding for elongation factors G and Tu of the apple proliferation phytoplasma. Gene 226: 103-109.
- Ceccarelli, E., M. Bocchetta, R. Creti, A. M. Sanangelantoni, O. Tiboni and P. Cammarano. 1995. Chromosomal organization and nucleotide sequence of the genes for elongation factors EF-1α and EF-2 and ribosomal protein S7 and S10 of the hyperthermophilic archaeum *Desulfurococcus mobilis*. Mol. Gen. Genet. **246**: 687-696.
- Chen, L.-M. and K.-C. Ho. 1997. Characterization of a phytoplasma-related ribonuclease in *Catharantus roseus*. Taiwania **42**: 8-16.
- Chen, Y.-H. and K.-C. Ho. 1998. Isolation and characterization of *tuf* gene of phytoplasma associated with loofah witches' broom. The 12th International Organization of Mycoplasmology Conference. 22-28 July, Sydney, Australia.
- Cousineau, B., C. Cerpa, J. Lefebrre and R. Cedergren. 1992. The sequence of the gene encoding elongation factor Tu from *Chlamydia trachomatis* compared with those of other organisms. Gene **120**: 33-41.
- Gundersen, D. E., I.-M. Lee, S. A. Rehner, R. E. Davis and T. Kingsbury. 1994. Phylogeny of mycoplasmalike organisms (phytoplasmas): a basis for their classification. J. Bacteriol. 176: 5244-5254.

- Ho, K.-C., C.-C. Tsai and T.-L. Chung. 2001. Organization of ribosomal RNA from a loofah witches' broom phytoplasma. DNA and Cell Biol. 20: 115-122.
- Ho, K.-C., V. E. Quarmby, F. S. Frenchand E. M Wilson. 1992. Molecular cloning of rat prostate transglutaminase complementary DNA: The major androgen-regulated protein DP1 of rat dorsal prostate and coagulating gland. J. Biol. Chem. **267**: 12660-12667.
- Inamine, J. M., S. Loechel and P-C. Hu. 1989. Nucleotide sequence of the *tuf* gene from *Mycoplasma gallisepticum*. Nucleic Acids Res. 17: 10126.
- Kamla, V., B. Henrich and U. Hadding. 1996. Phylogeny based on elongation factor Tu reflects the pheno-typics features of mycoplasmas better than that based on 16S rRNA. Gene 171: 83-87.
- Lee, I.- M., W. Hammond, R. E. Davis and D. E. Gundersen 1993. Universal amplification and analysis of pathogen 16S rDNA for classification and identification of mycoplasmalike organisms. Phytopathol. 83: 339-354.
- Loechel, S., J. M. Inamine and P-C. Hu. 1989. Nucleotide sequence of tuf gene from *Mycoplasma genitalium*. Nucleic Acids Res. 17: 10127.
- Ludwig, W., J. Neumaier, N. Klugbauer, E. Brockmann, C. Roller, S. Jilg, K. Reetz, I. Schachtner, A. Ludvigsen, M. Bachleitner, U. Fischer and K. H. Schleifer. 1993. Phylogenetic relationships of Bacteria based on comparative sequence analysis of elongation factor Tu and ATP synthase β -subunit genes. Antonie Leeuwenhoek **64**: 285-305.
- Ludwig, W., M. Weizenegger, D. Betzl, E. Leidel, T. Lenz, A. Ludvigsen, D. Mollenhoff, P. Wenzig and K. H. Schleifer. 1990. Complete nucleotide sequences of seven eubacterial genes coding for the elongation factor Tu: functional, structural and phylogenetic evaluations. Arch. Microbiol. 153: 241-247.
- Razin, S., D. Yogev and Y. Naot. 1998. Molecular biology and pathogenicity of mycoplasmas. Microbiol. Mole. Biol. Rev. 62: 1094-1156.
- Schneider, B., K. S. Gibb and E. Seemuler. 1997. Sequence and RFLP analysis of the elongation factor Tu gene used in differentiation and classification of phytoplasmas. Microbiol. 143: 3381-3389.
- Sears, B. B. and B. C. Kirkpatrick. 1994. Unveiling the evolutionary relationships of plant-pathogenic mycoplasmalike organisms: phylogenetic insights may provide the key to culturing phytoplasmas. ASM-News 60: 307-312.
- Sela, S., D. Yogev, S. Razin and H. Bercovier. 1981. The conservation of DNA sequence over very long periods of evolutionary time: Evidence against intergeneric chromosomal transfer as an explanation for the presence of *Escherichia coli tuf* gene sequences in taxonomically-unrelated prokaryotes. Eur. J. Biochem. **120**: 69-77.
- Sela, S., D. Yogev, S. Razin and H. Bercovier. 1989. Duplication of the *tuf* gene: a new insight into the phylogeny of eubacteria. J. Bacteriol. 171: 581-584.
- Tully, J. G. 1993. International committee on systemic bacteriology subcommittee on the taxonomy of Mollicutes, minutes of interim meeting, 1 and 2 Auguest, 1992, Ames, Iowa. Inter. J. System. Bacteriol. **43:** 394-397.
- Vijgenboom, E., L. P. Woudt, W. H. Heinstra, K. Rietveld, J. van Haarlem, G. P. van Wezel, S. Shochat and L. Bosch. 1994. Three tuf-like genes in the kirromycin producer Streptomyces ramocissimus. Microbiol. 140: 983-998.
- Weisburg, W. G., J. G. Tully, D. L. Rose, J. P. Petzel, H. Oyaizu, D. Yang, L. Mandelco, J. Sechrest, T. G. Lawrence, J. van Etten, J. Maniloff and C. R. Woese. 1989. A

- phylogenetic analysis of the mycoplasmas: basis for their classification. J. Bacteriol. 171: 6455-6467.
- Woese, C. R. 1987. Bacterial evolution. Microbiological Rev. 51: 221-271.
- Yogev, D., S. Levisohn, S. H. Kleven, D. Halachmi and S. Razin. 1988. Ribosomal RNA gene probes to detect intraspecies heterogeneity in *Mycoplasma galliseptium* and *Mycoplasma synoviae*. Am. Assoc. Avian Pathologist 32: 220-231.
- Yogev, D., S. Sela, H. Bercovier and S. Razin. 1990. Nucleotide sequence and coden usage of the elongation factor Tu (EF-Tu) gene from *Mycoplasma pneumoniae*. Mol. Microbiol. 4: 1303-1310.

絲瓜簇葉病植物菌質體 tuf 基因之選殖及序列分析

何國傑(1)

(收稿日期:2001年9月25日;接受日期:2001年11月4日)

摘 要

利用絲瓜簇葉病植物菌質體 DNA 為模板及兩個引作 PCR 反應,產生一段具 1651 個鹽基對的 DNA 片段。經核甘酸序列分析發現此 DNA 片段含有兩個譯讀框架,其中一個為合成部分之延長因子 EF-G (fus 基因)的譯讀框架,另一個為合成全長之延長因子 EF-Tu (tuf 基因)的譯讀框架。tuf 基因含有 1911 個核甘酸,能合成一個具有 396 個氨基酸,分子量為 43.8 千道爾頓的蛋白質。此基因之核甘酸序列的 G+C 百分比為 33.42,合乎膜柔綱之生物的特性。與其他細菌比較,在氨基酸序列上有高達 53.9 到 61.8%的相似性,而與古生菌的相似性則低於 26%。基因組上,只存有一個 tuf 基因版本。在親緣關係方面,由 tuf DNA 序列分析所構成的親緣樹或是利用 16S rDNA 序列分析構成的親緣樹,顯示絲瓜簇葉病植物菌質體與其他菌種的親緣關係是相當一致的。

關鍵詞: 植物菌質體、tuf 基因、親緣關係。