Mol. Cells, Vol. 10, No. 5, pp. 579–583

Communication

Molecules and Cells

Characterization of Geranium (*Pelargonium graveolens*) Chloroplast EF-Tu cDNA

Chul Joo Kang, Min Goo Lee, Young Sil Cho, Joong Won Lee, Yoon Jung Kyung, Jeong Sheop Shin[1], Eun-Soo Kim[2], and Jeong-Kook Kim*

Department of Biology, Korea University, Seoul 136-701, Korea;
[1] Graduate School of Biotechnology, Korea University, Seoul 136-701, Korea;
[2] Department of Biology, Konkuk University, Seoul 133-701, Korea.

(Received on March 28, 2000)

A geranium (*Pelargonium graveolens*) chloroplast translational elongation factor EF-Tu (*tufA*) cDNA was isolated. The geranium *tufA* cDNA is 1,584 bp long with 20 bp of 5′ untranslated region (UTR) and 139 bp of 3′ UTR. It encodes 474 amino acids including a putative chloroplast transit peptide of 65 amino acids. The deduced polypeptides of the geranium *tufA* cDNA contains four GTP binding sequences in its N-terminal region and two chloroplast EF-Tu signature regions in the C-terminal region. The predicted molecular weight of the mature geranium chloroplast EF-Tu protein was about 45,000 and its amino acid sequence identity with the chloroplast EF-Tu proteins of tobacco, pea, *Arabidopsis*, rice, and soybean ranges from 85% to 91%. The geranium *tufA* appears to exist as a single copy gene like Arabidopsis and rice, whereas other known dicot plants have more than one copy in their nuclear genomes.

Keywords: Chloroplast; Elongation Factor EF-Tu; Geranium.

Introduction

The translational elongation factor Tu (EF-Tu) plays a role in protein synthesis in elongation of the peptide chain by placing aminoacyl tRNAs on the ribosome with GTP.

Crystal structure analysis of bacterial EF-Tu showed that EF-Tu folded into three domains (Kjeldgaard and Nyborg, 1992; Kjeldgaard *et al*., 1993). Like many other GTP binding proteins, there are four specific amino acid sequences (GxxxGK, DxxG, NKxD, SAL) that interact with GTP, and these amino acid sequences are highly conserved throughout most translational elongation factor proteins (Bourne *et al*., 1990; 1991).

In the lower photosynthetic eukaryotes, including *Chlamydomonas reinhartii* (Baldauf and Palmer, 1990), *Euglena gracilis* (Montadon and Stutz, 1983), and probably most green algae (Baldauf *et al*., 1990), EF-Tus are encoded by the chloroplast genome. However, higher plant chloroplast EF-Tus are encoded by nuclear DNA (Baldauf and Palmer, 1990; Baldauf *et al*., 1990), synthesized as a precursor in the cytoplasm, and then imported into chloroplasts (Keegstra *et al*., 1989; Schmidt and Mishkind, 1986).

Several genomic or cDNA clones encoding chloroplast EF-Tu have been identified and characterized in a number of higher plants. A single nuclear *tufA* gene was identified in *Arabidopsis thaliana* (Baldauf and Palmer, 1990) and *Oryza sativa* (Lee *et al*., 1999). EF-Tu was encoded by two genes in the *Nicotiana sylvestris* nuclear genome (Murayama *et al*., 1993) and by four genes in *Glycine max* (Bonny and Stutz, 1993; Maurer *et al*., 1996).

In this study, we isolated geranium chloroplast EF-Tu cDNA using a method for performing rapid amplification of cDNA ends (RACE) (Chenchick *et al*., 1996) and examined its sequence similarity with the previously reported plant *tufA* genes. In addition, the copy number of the *tufA* gene in the geranium nuclear genome was examined.

Materials and Methods

Plant material and growth condition Rose geranium (*Pelargonium graveolens*) was purchased from Nong-Sim Won, Co. and grown in continuous light at 27°C. Young leaves were harvested, immediately frozen in liquid nitrogen and stored at −80°C.

* To whom correspondence should be addressed.
Tel: 82-2-3290-3157; Fax: 82-2-923-9522
E-mail: jkkim@kuccnx.korea.ac.kr

RNA isolation Total cellular RNA of geranium leaves was prepared using the method of Schultz *et al.* (1994). Poly (A) mRNAs were isolated using DYNAL beads (DYNAL, USA) according to the manufacturer's instructions. From isolated poly(A) mRNA, adapter-ligated cDNAs were synthesized using a Marathon cDNA Amplification kit (Clontech, USA) according to the manufacturer's instructions.

cDNA isolation To perform 3′-RACE, a degenerate primer was designed for gene specific primer 2 (GSP2) at a conserved region of other plant *tufA* genes (5′-AAC ACC/T GCC/T ACC/T GTG/T GAG TAC GAG AC-3′) (Lee, 1998). The 3′-RACE conditions were: 1 cycle of 1 min at 94°C; 5 cycles of 30 s at 94°C and 4 min at 72°C; 5 cycles of 30 s at 94°C and 4 min at 70°C; 25 cycles of 20 s at 94°C and 4 min at 68°C. The 3′-RACE product was sequenced directly with GSP2 as a sequencing primer. After confirming that the *tufA* genes were amplified, 3′-RACE products were ligated into TA cloning vector, pCR-Script Amp (Stratagene, USA). The cloned 3′-RACE products were completely sequenced using Big Dye (Perkin Elmer, USA) on a ABI Prism 310 genetic analyzer (PE Applied Biosystems, USA). For the 5′-RACE reaction, the rice *tufA* gene-specific primer 1 (GSP1) was used (Lee, 1998), although it has 4 nucleotide mismatches to the geranium *tufA* gene sequence. The 5′-RACE condition was the same as the 3′-RACE condition. The 5′-RACE product was ligated into a TA cloning vector, pGEM EASY T-vector (Promega, USA) and sequenced.

Nucleotide and predicted peptide sequences were analyzed using DNASIS (Hitachi, Japan) and the BLAST programs in non-redundant databases of the National Centre for Biotechnology Information (Altschul *et al.*, 1990). Multiple alignment of *tufA* amino acid sequences was analyzed by the CLUSTAL W program of EMBL and the unrooted phylogenetic tree was constructed using a treeview program.

Northern blot analysis Twenty micrograms of the denatured total cellular RNAs from leaves were separated on 1% agarose-formaldehyde gel and transferred onto a Hybond-N membrane (Amersham, USA) according to the capillary method (Sambrook *et al.*, 1989), and then the blots were fixed by UV irradiation. The blot was probed with the 3′-RACE product that was labeled with [α-^{32}P] dCTP (Amersham, USA) using a ladderman labeling kit (Takara, Japan). Hybridization was performed as previously done (Lee *et al.*, 1999). After hybridization, the membrane was scanned with a Phosphoimage analyzer (Fujifilm, Japan).

Preparation of geranium genomic DNA and Southern blot analysis The geranium genomic DNA was prepared from leaves according to the methods described by Porebski *et al.* (1997). The twenty micrograms of genomic DNA were digested with several restriction enzymes and separated on 1% agarose gel. The gel was transferred onto a Hybond-N membrane (Amersham, USA) and then blots were fixed by UV irradiation. The blot was probed with the 3′-RACE product that was used in the Northern analysis. Hybridization was performed as previously done (Lee *et al.*, 1999). After hybridization, the membrane was scanned with a Phosphoimage analyzer (Fujifilm, Japan).

Results and Discussion

Isolation of the geranium *tufA* cDNA and its comparison to other plant *tuf* genes The rapid amplification of cDNA ends (RACE) technique was applied to isolate a geranium chloroplast translation elongation factor EF-Tu (*tufA*) cDNA. The cDNA clone (Genbank accession number: AF234537) obtained is 1,584 bp long with 20 bp of the 5′ untranslated region (UTR) and 139 bp of the 3′ UTR and one open reading frame (ORF) of 474 codons. There are two ATG codons which could be translational start codons; however, the latter ATG codon (GAATGGC) is in the plant consensus initiation context for the flanking regions of ATG (Joshi, 1987a). Therefore, the latter ATG codon was considered as the translational start codon of our geranium chloroplast EF-Tu cDNA clone. A consensus polyadenylation signal (AATAAA) did not exist in the 3′ UTR of our clone. In plants, however, repeated AT-rich sequences (ATTTAT, TAATAA, ATTAAA and TTAAAT) were regarded as alternative polyadenylation signals in the nuclear genes of higher plants (Joshi, 1987b). Therefore, it is likely that one of several AT-rich sequences existing in the 3′ UTR of our clones would act as a polyadenylation signal.

The predicted amino acid sequences of our geranium cDNA clone were compared to those of the other plant chloroplast and mitochondrial EF-Tus from *O. sativa*, *G. max*, *A. thaliana*, *N. sylvestris*, and *Pisum sativum* (Fig. 1). In geranium *tufA* amino acid sequences, there are four GTP binding sites (GHVDHK, DCPG, NKQD and SAL) and two chloroplast translational elongation factor Tu signature regions (ALMANPAIKR and KDEES) which are well conserved characteristics of other known chloroplast EF-Tus (Delwiche *et al.*, 1995; Kuhlman and Palmer, 1995). Second, the overall amino acid identity of our geranium clone excluding the putative transit peptides with the chloroplast EF-Tu of *O. sativa*, *G. max, A. thaliana*, *N. sylvestris*, and *P. sativum* are 85%, 91%, 88%, 90%, and 88%, respectively. However, our geranium clone shares only 65% and 27% amino acid identity, respectively, with *Arabidopsis* mitochondrial EF-Tu and cytosolic EF-1α. These results indicate that our geranium clone encodes the chloroplast EF-Tu (*tufA*), but neither the mitochondrial EF-Tu nor its cytosolic counterpart EF-1α.

The first 65 amino acid residues of our geranium cDNA clone were rich in serine, threonine, and basic amino acids suggesting that this region might be a transit peptide. As expected, these regions have no specific homology to other transit peptides of the known chloroplast elongation factor Tu (Fig. 1). Any highly conserved motif is not known for chloroplast targeted transit peptide, so several prediction programs are

Fig. 1. Multiple alignment of geranium chloroplast EF-Tu (Genbank accession number: AF234537) with other EF-Tus of various plants [*A. thaliana tufA* (X52256) and *tufM* (X89227), *G.max tufA1* (X66062), *P. sativum tufA* (Y14561), *N. tabacum tufA* (M94204), *O. sativa tufA* (AF145053)] by CLUSTAL W. Transit peptides were not aligned and are printed in italic. GTP binding sequences were shaded and chloroplast EF-Tu signature sequences were boxed.

protein import to the chloroplasts. The predicted transit peptide cleavage site (CGTVR↓ AARG) of this geranium *tufA* gene is also the same as those of other *tufA* genes.

The unrooted phylogenetic tree inferred from the amino acid sequences of chloroplast EF-Tu is shown in Fig. 2. The tree was obtained by the Neighbor-Joining method (Saitou and Nei, 1987) based on the amino acid alignments with gaps included using the CLUSTAL W program (Thompson *et al.*, 1994). In this phylogenetic tree, all chloroplast elongation factor Tu (*tufA*) genes were clustered together and show an evolutionary distance from the mitochondrial elongation factor Tu (*tufM*) gene of *Arabidopsis*. There were some nodes in the phylogenetic tree that tell us about the evolutionary meaning of the gene transfer to the nuclear genome from the chloroplast genome, in both dicots and monocots. The divergence of the monocot *tufA* gene occurred at an earlier stage of eukaryotic evolution than that of dicot *tufA* genes (Lee *et al.*, 1999). The geranium

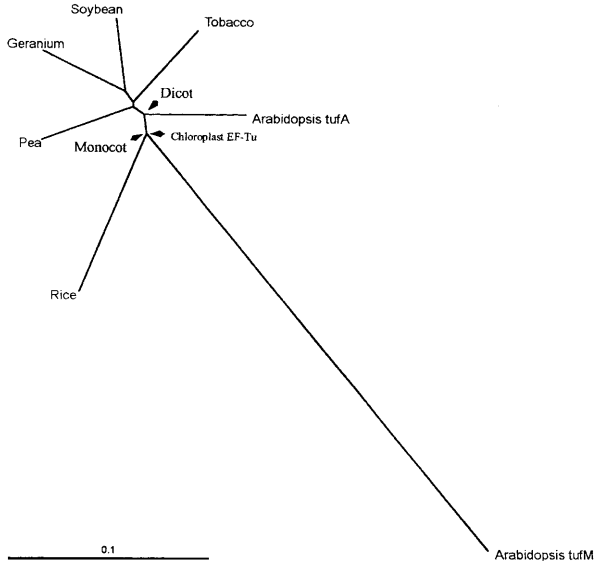

Fig. 2. The unrooted phylogenetic tree inferred from the amino acid sequences of elongation factor Tu. The tree was obtained by the Neighbor-Joining method based on the amino acid alignment using the Clustal W program. The scale bar indicates a divergence of 0.1 amino acid substitutions per site. Arrows indicate nodes.

available for searching this specific feature. A computer analysis using neural network method (ChloroP) (Emanuelsson *et al.*, 1999) was used for searching transit peptide of geranium *tufA*. This program predicted that this region is the transit peptide necessary for

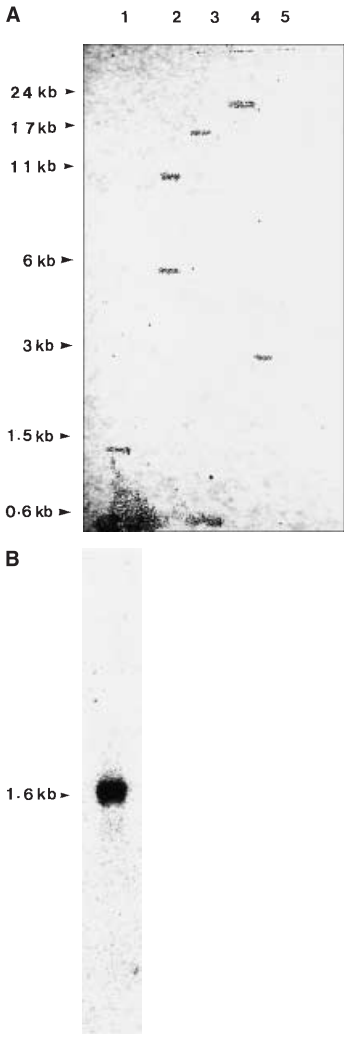

Fig. 3. A. Geranium genomic Southern hybridization. Genomic DNA from geranium, *P. graveolens,* was digested with *Eco*RI(E), *Eco*RV(R), *Bam*HI(B), *Xho*I(X), and *Hin*dIII(H). The 3′-RACE product was used as a probe. Approximate sizes (kb) of the signal are shown on the left. The λ DNA/*Hin*dIII fragment size marker was used. **B.** Northern Hybridization of geranium total RNA. Twenty micrograms of geranium total RNA were loaded. The 3′-RACE product was used as a probe.

tufA gene revealed higher homology to the *tufA* gene of soybean than other plant *tufA* genes.

The copy number and transcript size of the geranium *tufA* gene In *Arabidopsis* and rice, there is one *tufA* gene in the nuclear genome; but in other known plants there are more than one in their nuclear genomes. To determine the copy number of the *tufA* gene in the geranium genome, Southern blot analysis was performed (Fig. 3A). Genomic DNAs digested with enzymes (*Bam*HI, *Xho*I and *Hin*dIII) that have no restriction site on the geranium *tufA* cDNA showed one signal, but enzymes (*Eco*RI and *Eco*RV) that have one restriction site produced two signals. In addition, all 18 clones of the RACE-PCR products have the same DNA sequences. Data from both genomic Southern hybridization and DNA sequencing of the RACE-PCR products suggest that a single copy of the *tufA* gene may exist in the geranium nuclear genome. Considering that in the geranium chloroplast genome, there is a duplication of 10 protein genes present only once in the chloroplast genomes of all other land plants, it is interesting that there is only one copy of the *tufA* gene. Northern hybridization data suggested that the transcript size of *tufA* was about 1.7 kb (Fig. 3B), as predicted from its cDNA sequence.

Acknowledgments The authors wish to acknowledge the financial support of the Korean Research Foundation in the 1997 program year.

References

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990) Basic local alignment search tool. *J. Mol. Biol.* **215**, 403–410.

Baldauf, S. L. and Palmer, J. D. (1990) Evolutionary transfer of the chloroplast *tufA* gene to the nucleus. *Nature* **344**, 262–265.

Baldauf, S. L., Manhart, J. R., and Palmer, J. D. (1990) Different fates of the chloroplast *tufA* gene following the transfer to the nucleus in green algae. *Proc. Natl. Acad. Sci. USA* **87**, 5317–5321.

Bonny, C. and Stutz, E. (1993) Soybean (*Glycine max* L.) nuclear DNA contains four *tufA* genes coding for the chloroplast-specific translation elongation factor EF-Tu. *Chimia* **47**, 247–249.

Bourne, H. R., Sanders, D. A., and McCormick, F. (1990) The GTPase superfamily: a conserved switch for diverse cell function. *Nature* **348**, 125–132.

Bourne, H. R., Sanders, D. A., and McCormick, F. (1991) The GTPase superfamily: a conserved structure and molecular mechanism. *Nature* **349**, 117–127.

Chenchik, A., Mogadam, F., and Siebert, P. (1996) A new method for full-length cDNA cloning by PCR.; in *A Laboratory Guide to RNA: Isolation, Analysis, and Synthesis*, Krieg, P. A. (ed.), pp. 273–321, Wiley-Liss, Inc.

Delwiche, C. F., Kuhsel, M., and Palmer, J. D. (1995) Phylogenetic analysis of *tufA* sequence indicates a cyanobacterial origin of all plastids. *Mol. Physiol. Evol.* **4**, 110–128.

Emanuelsson, O., Nielsen, H., and Heijne, G. (1999) ChloroP, a neural network-based method for predicting chloroplast transit peptides and their cleavage sites. *Protein Sci.* **8**, 978–984.

Joshi, J. P. (1987a) An inspection of the domain between putative TATA box and translation start site in 79 plant genes. *Nucleic. Acids Res.* **15**, 6643–6653.

Joshi, J. P. (1987b) Putative polyadenylation signals in nuclear genes of higher plants: a comparison and analysis. *Nucleic Acids Res.* **15**, 9627–9640.

Keegstra, K., Olsen, L. J., and Theg, S. M. (1989) Chloroplast precursors and their transport across the envelope membranes. *Annu. Rev. Plant Physiol.* **40**, 471–501.

Kjeldgaard, M. and Nyborg, J. (1992) Refined structure of elongation factor EF-Tu from *Escherichia coli. J. Mol. Biol.* **223**, 721–742.

Kjeldgaard, M., Nissen, P., Thirup, S., and Nyborg, J. (1993) The crystal structure of elongation factor EF-Tu from *Thermus aquaticus* in the GTP conformation. *Structure* **15**, 35–50.

Kuhlman, P. and Palmer, J. D. (1995) Isolation, expression, and evolution of the gene encoding mitochondrial elongation factor Tu in *Arabidopsis thaliana*. *Plant Mol. Biol.* **29**, 1057–1070.

Lee, J. H., Lee, J. W., Chung, Y. Y., Paek, K. H., Shin, J. S., Yun, C. H., and Kim, J.-K. (1999) Cloning and characterization of the chloroplast elongation factor EF-Tu cDNA of *Oryza sativa* L. *Mol. Cells* **9**, 484–490.

Lee, J. W. (1998) The cloning of chloroplast EF-Tu cDNA in *Oryza sativa* L. M. S. thesis, Korea University, Seoul.

Maurer, F., Murone, M., and Stutz, E. (1996) The *tuf* gene family of soybean: structure and differential transcription. *Plant Sci.* **117**, 83–93.

Montadon, P. E. and Stutz, E. (1983) Nucleotide sequence of a *Euglena gracilis* chloroplast genome region coding for the elongation factor Tu: evidence for a spliced mRNA. *Nucleic Acids Res.* **11**, 5877–5892.

Murayama, Y., Matsubayashi, T., Sugita, M., and Sugiura, M. (1993) Purification of chloroplast elongation factor Tu and cDNA analysis in tobacco: the existence of two chloroplast elongation factor Tu species. *Plant Mol. Biol.* **22**, 767–774.

Porebski, S. L., Basiley, G., and Baum, B. R. (1997) Modification of a CTAB DNA extraction protocol for plants containing high polysaccharides and polyphenol components. *Plant Mol. Biol. Rep.* **15**, 8–15.

Saitou, N. and Nei, M. (1987) The Neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Sambrook, J., Fritch, J. E., and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Schmidt, G. W. and Mishkind, M. L. (1986) The transport of proteins into chloroplasts. *Annu. Rev. Biochem.* **55**, 879–912.

Schultz, D. J., Craig, C., Cox-Foster, D. L., Mumma, R. O., and Medford, J. I. (1994) RNA isolation from recalitrant plant tissue. *Plant Mol. Biol. Rep.* **12**, 310–316.

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids. Res.* **22**, 4673–4680.