J. Biochem. 120, 1095-1103 (1996)

Complete Nucleotide Sequences of the Genes Encoding Translation Elongation Factors 1α and 2 from a microsporidian parasite, *Glugea plecoglossi*: Implications for the Deepest Branching of Eukaryotes[1]

Takashi Kamaishi,*[,†,2] Tetsuo Hashimoto,†[,3] Yoshihiro Nakamura,‡ Yutaka Masuda,§ Fuminori Nakamura,* Ken-ichi Okamoto,* Makoto Shimizu,‖ and Masami Hasegawa†

*Department of Medical Biology, School of Medicine, ‡Laboratory of Gene Manipulation, §Department of Biological Chemistry, School of Pharmaceutical Sciences, Showa University, 1-5-8 Hatanodai, Shinagawa-ku, Tokyo 142; †The Institute of Statistical Mathematics, 4-6-7 Minami-Azabu, Minato-ku, Tokyo 106; and ‖Department of Fisheries, Faculty of Agriculture, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113

Received for publication, June 21, 1996

Complete nucleotide sequences of the genes putatively encoding translation elongation factors 1α (EF-1α) and 2 (EF-2) from a mitochondrion-lacking protozoan, *Glugea plecoglossi*, that belongs to microsporidians were determined. The deduced amino acid sequences of the putative EF-1α and EF-2 of *Gl. plecoglossi* showed very unusual features compared with typical eukaryotic sequences. The degree of divergence was especially great in the EF-1α sequence, although it clearly showed a eukaryotic feature when aligned with homologs from the three primary kingdoms. Phylogenetic analyses of EF-1α and EF-2 on the basis of the maximum likelihood method of protein phylogeny clearly and consistently suggested that among eukaryotic species being analyzed, *Gl. plecoglossi* and another mitochondrion-lacking protozoan, *Giardia lamblia*, respectively represent the earliest and the second earliest offshoots of eukaryotes. When the EF-1α and EF-2 phylogenies were totally evaluated, the earliest divergence of *Gl. plecoglossi* in eukaryotes became more clearly confirmed. If the phylogenetic relationships inferred from the present analysis are correct, microsporidians might be extremely ancient eukaryotes that diverged before the occurrence of mitochondrial symbiosis.

Key words: ancient divergence of eukaryotes, elongation factor, *Glugea plecoglossi*, microsporidians, mitochondrion-lacking protozoa.

Most of the contemporary eukaryotes have mitochondria which are known to have descended from eubacteria by endosymbiosis (*1, 2*). The endosymbiotic origin of mitochondria and chloroplasts has been confirmed by many phylogenetic studies reported so far. The RNAs and proteins functional in mitochondria and chloroplasts are clearly of eubacterial origin. However, some protozoa are known to have no mitochondria, although they are eukaryotes (*3-5*). Phylogenetic placement of these mitochondrion-lacking protozoa among eukaryotes is important in understanding the early evolution of eukaryotic cells. If some of these protozoa turn out to represent deep branchings in the eukaryotic tree, they can be regarded as good candidates for living relics of the earliest phase of eukaryotic evolution before the endosymbiotic origin of proto-mitochondria occurred. If others turn out to have diverged later than mitochondrion-containing protozoa, they may have secondarily lost their mitochondria because of their lifestyle, which may have been, for example, parasitic.

Early studies on the phylogenetic positions of deep branching eukaryotes were carried out by sequence comparisons of small subunit ribosomal RNA (SrRNA). Sogin (*6, 7*), Sogin *et al.* (*8*), and Leipe *et al.* (*9*) proposed eukaryotic trees including mitochondrion-lacking protozoa. According to the up-to-date SrRNA tree by Leipe *et al.* (*9*), among mitochondrion-lacking protozoa, *Vairimorpha necatrix* (microsporidians), *Tritrichomonas foetus* (trichomonads), and the common ancestor of *Giardia lamblia* and *Hexamita inflata* (diplomonads) represent the earliest, the second earliest, and the third earliest offshoots, respectively, but the branching order of these lineages is not clearly confirmed. Since the SrRNA phylogeny may sometimes be unreliable due to drastic differences in G+C contents among species (*10*), and since a phylogenetic tree inferred from a single gene is sometimes fragile because of a lack of sufficient information, reexamination using alternative sequence data has been desirable (*11-16*).

Recently, we isolated a major part of the mRNA encoding translation elongation factor 1α (EF-1α) from a mitochondrion-lacking protozoan, *Glugea plecoglossi* (microspori-

[1] This work was carried out under the ISM Cooperative Research Program (95ISM-CRP-A70 and 96ISM-CRP-A93) and was supported by grants from the Ministry of Education, Science, Sports and Culture of Japan. Nucleotide sequence data reported here are available in the GSDB, DDBJ, EMBL, and NCBI databases under the accession numbers D84253 and D79220.

[2] On leave from Department of Fisheries, Faculty of Agriculture, The University of Tokyo, 1-1-1 Yayoi, Bunkyo-ku, Tokyo 113.

[3] To whom correspondence should be addressed. Phone: +81-3-5421-8773, FAX: +81-3-3446-1695, E-mail: hasimoto@eicsun2.ism.ac.jp

Abbreviations: EF-1α, translation elongation factor 1α; EF-2, translation elongation factor 2; SrRNA, small subunit ribosomal RNA; ML, maximum likelihood.

dians). From the protein phylogeny of EF-1α including the *Gl. plecoglossi* data, we suggested that mitochondrion-lacking protozoa, *Gl. plecoglossi* and *Giardia lamblia*, respectively represent the earliest and the second earliest offshoots of the eukaryotic tree, and that the EF-1α phylogeny is inconsistent with the SrRNA tree in several respects (*17*). Because there are few data of the genes completely sequenced from microsporidians, and because the position of microsporidians should not be concluded only from previous results, we here report complete nucleotide sequences of the genes putatively encoding EF-1α and EF-2 from *Gl. plecoglossi*, and further analyze the phylogenetic relationships among eukaryotes. Our ML analyses based both on the EF-1α and EF-2 data sets clearly and consistently demonstrated that *Gl. plecoglossi* represents the earliest offshoot of the eukaryotic tree among the protozoan lineages being analyzed.

Microsporidians are spore-forming unicellular parasites that infect many animal groups in vertebrates as well as invertebrates. Their cellular organization is very simple. The organelles present are rough endoplasmic reticulum, various vesicles of undetermined nature, and an atypical reticulate Golgi apparatus. There is no mitochondrion throughout the life cycle (*18, 19*). The ribosome of microsporidian species is 70S type (*20*), and there is no 5.8S ribosomal RNA gene between the small and large subunit ribosomal RNA genes (*21*). *Gl. plecoglossi*, isolated and identified as a microsporidian by Takahashi and Egusa in 1977 (*22*), is often found in cultures of *Plecoglossus altivelis* (Ayu fish) in Japan.

EF-1α and its eubacterial homolog EF-Tu promote GTP-dependent binding of aminoacyl-tRNA to the A site of ribosomes in a codon-dependent manner. EF-2 and its eubacterial homolog EF-G catalyze the GTP-dependent translocation of peptidyl-tRNA from the A site to the P site on ribosomes. Both the EF-1α/Tu and EF-2/G are essential for protein synthesis in all organisms (*23*), and because of their universal occurrence in all living organisms on earth and their relatively slow evolutionary rates (*24, 25*), they are useful for tracing the early evolution of life.

MATERIALS AND METHODS

Nucleotide Sequence Analyses—Host fish infected by *Gl. plecoglossi* in Lake Biwako were kindly provided by Dr. S. Takahashi. Spores of *Gl. plecoglossi* were isolated by hand and suspended in phosphate-buffered saline (PBS). Contaminating host tissues were digested by using 10% (w/v) trypsin at 37°C for 30 min. The spores were washed with PBS many times, dropped into liquid nitrogen in a mortar, and blended with a pestle until the spore wall was completely destroyed. The resultant mixture was subjected to the standard techniques of poly(A)$^+$-RNA and genomic DNA extractions (*26*).

Genomic DNA and mRNA:cDNA duplex were used as templates for the polymerase chain reaction (PCR). The PCR-amplified fragments were purified and ligated into *Sma*I site of the plasmid pUC18 or *Eco*RV site of the plasmid pT7. *Escherichia coli* strain JM109 was transformed with the plasmid. Plasmid DNA was prepared by the alkaline method (*26*).

Ten micrograms of the genomic DNA was completely digested with each of *Eco*RI and *Bam*HI, and genomic DNA libraries were constructed respectively in the vectors λgt10 (Amersham) and λZAP *Express* (Stratagene). The fragments used as probes for screening were labeled with digoxigenin (DIG, Boehringer Mannheim). For the screening step, several reactive clones were isolated, purified, and subcloned into pUC18 or excised into pBK. Nucleotide sequences were determined either by a manual method (*Bca*Best Sequencing Kit, Takara) or with an ABI model 373A automated DNA sequencer using a *Taq* Dyedeoxy™ Terminator Cycle Sequencing Kit (Perkin-Elmer Cetus).

For Southern blot analyses, 10 μg of the genomic DNA was completely digested with *Eco*RI or *Bam*HI. The resultant fragments were separated by electrophoresis on a 1.0% agarose gel, denatured, transferred onto nylon membrane (Hybond-N, Amersham), and fixed with UV stratalinker (Stratagene). The fragment used as probe was α-^{32}P-labeled with a Random Primer Labeling Kit (Amersham) in the EF-1α analysis or DIG-labeled in the EF-2 analysis. Hybridization with the probe was carried out at 65°C overnight in a buffer containing 6×SSC, 5×Denhardt's solution, 0.5% SDS, and 50 μg/ml denatured salmon sperm DNA (*26*).

Amino Acid Sequence Alignments—All the sequence data used in the present analyses were collected from the EMBL and GenBank databases.

Because the EF-1α sequences of eukaryotes and archaebacteria were highly conserved, these sequences could be unambiguously aligned manually except for the N- and C-terminal portions. Among the eubacterial sequences including mitochondrial and chloroplast counterparts, an unambiguous alignment was also constructed with minor insertions/deletions except for the N-terminal transit peptide regions. However, when we aligned the eukaryotic/archaebacterial EF-1αs, the eubacterial EF-Tus, and the other paralogous protein sequences with each other, several insertions/deletions and N- and/or C-terminal extentions were necessary for these sequences.

To align EF-2 sequences, we first used the maximum likelihood alignment method (*27-29*), since we could not unambiguously determine gap positions between archaebacterial and eukaryotic sequences manually. All possible pairwise alignments among the EF-2 sequences used were inferred with the "statprot" program (kindly provided by Dr. J.L. Thorne). On the basis of the results, multiple alignment was performed by manually adjusting inconsistencies found among the pairwise alignments.

Phylogenetic Analyses—The maximum likelihood (ML) method of protein phylogeny (*30*) was used to infer phylogenetic relationships among related sequences. Models assumed for the amino acid substitution process were the JTT model (*31*).in the EF-1α analysis and the JTT-F model (*32*) in the EF-2 analysis, because each model was confirmed to best approximate the data among alternative models. When the number of operational taxonomic units (OTUs) to be analyzed becomes 9, we have to make the ML estimation for 135,135 possible trees. Although this number of trees can be analyzed, it represents a large computational burden. Ten or more OTUs are apparently impossible. In this situation, we used an approximate log-likelihood method (*33, 34*) to search for tree topologies. Up to 10,000 candidate tree topologies were selected by the approximate log-likelihood criterion, and these were examined for the exact ML method. The ML tree thus

selected was further analyzed by a local rearrangement method (*33, 34*) to confirm whether it really was an ML tree. If the OTUs became more than 11, we first inferred a neighbor-joining (NJ) tree (*35*) using a distance matrix estimated by the ML method. The NJ tree thus obtained was further analyzed by the local rearrangement method to search for the ML tree. All the phylogenetic analyses were performed with the most recent version of the PROTML program in MOLPHY Version 2.3 (*34*).

RESULTS AND DISCUSSION

Cloning and Nucleotide Sequence Analysis of Glugea EF-1α and EF-2 Genes—Restriction sites, open reading frames (ORF), and locations of the probes used are shown schematically in Fig. 1.

To isolate *Glugea* EF-1α gene, we screened an *Eco*RI library using fragment I that had been obtained previously by the RT-PCR method (see Fig. 1 in Ref. *17*). Two positive clones, A and B, in Fig. 1, were isolated. These respectively included upstream and downstream portions of the *Eco*RI site in the EF-1α ORF. By sequence analysis of these clones, more than 2,000 bases were determined. The putative EF-1α ORF was not interrupted by intron-like sequences, and it consisted of 1,398 bases corresponding to 465 amino acids. The ORF and deduced amino acid sequences showed only minor changes compared with our previous result based on the RT-PCR analysis. The heterogeneities probably depend on the original samples. The ORF showed 59% A+T, and a more marked preference for A+T of 73% was observed in its third codon position. The present sequence data with additional information on the putative N-terminal portion of the *Glugea* EF-1α revealed a high divergence, although it clearly showed a typical feature of eukaryotic EF-1α in the alignment including all the homologs and paralogous proteins. Genomic Southern analysis using fragment I as a probe revealed two bands (Fig. 2A) each in the digestions with *Eco*RI and *Bam*HI, whose restriction sites are located in the fragment, suggesting that the putative EF-1α gene is a single copy.

In the EF-2 analysis, although we tried various combinations of degenerated primers, only a 0.3-kb fragment (fragment II) that corresponds to the N-terminal GTP-binding portion of EF-2 could be obtained. Using fragment II as a probe, we screened the *Bam*HI library and isolated clone C in Fig. 1. The clone contained only the N-terminal one-third portion of the EF-2 gene. From the restriction map of clone C and the Southern analysis (Fig. 2B), it turned out that an *Eco*RI site existed 250 bp downstream of 3′ *Bam*HI site in the clone. To obtain a new probe for screening the downstream region of the *Bam*HI site, we rescreened the *Eco*RI library using the previous probe (fragment II) and isolated clone D. By digesting an insert of clone D with *Eco*RI and *Bam*HI, we acquired another probe (fragment III). Using fragment III as a probe, we further screened the *Bam*HI library and obtained a positive clone (clone E) which covered the downstream portion of EF-2 ORF. Sequence analysis of these clones, C, D, and E, identified an ORF with no intron-like sequence. The ORF consisted of 2,547 bases encoding 848 amino acids, which was clearly similar to typical eukaryotic EF-2s. The A+T contents for the whole ORF and its third codon position

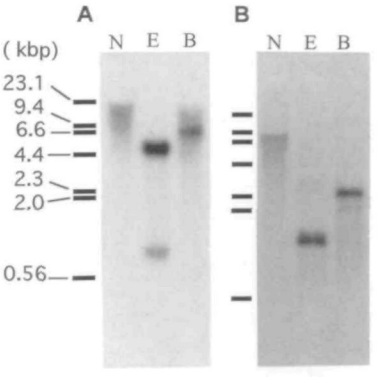

Fig. 2. **Genomic Southern analyses of *Gl. plecoglossi* genes putatively encoding EF-1α (A) and EF-2 (B).** N, not digested with restriction enzymes; E, digested with *Eco*RI; B, digested with *Bam*HI. Probes used for the analyses are fragment I (EF-1α) and fragment II (EF-2) shown in Fig. 1. λ DNA digested with *Hin*dIII was used as a size marker, and is indicated in the left side of each analysis (kilobases).

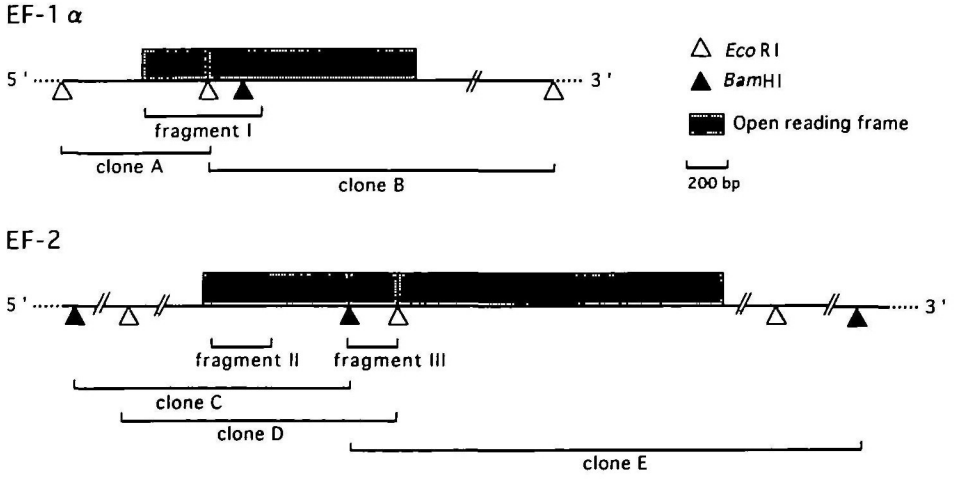

Fig. 1. **Schematic representation of the strategy to isolate putative EF-1α and EF-2 genes from *Glugea plecoglossi*.** The restriction map shows only *Eco*RI and *Bam*HI sites, respectively represented by open and closed triangles. Open reading frames are shown by hatched boxes. Fragments I, II, and III, were used as probes to screen genomic DNA libraries. Fragment I was previously obtained by the RT-PCR method (*17*). To obtain fragment II, the RT-PCR analysis was also used. Forward and reverse primers synthesized were degenerated for the highly conserved regions, IAHVDHGK and QTETVLRQ, respectively. These regions are located within the GTP-binding domain of the EF-2 alignment. Fragment III was derived by the digestion of clone D with *Eco*RI and *Bam*HI. Inserts of the positive clones isolated in the screening steps are also shown with their names.

were 62 and 69%, respectively. The EF-2 sequence also showed divergence in comparison with typical eukaryotic EF-2s, but the degree of divergence was not so great as that of EF-1α. In general, post-translational modification sites biochemically analyzed in the mammalian EF-2 are usually conserved among eukaryotic sequences with minor exceptions. These are phosphorylation of threonine residues 57 and 59 by Ca^{2+}/calmodulin-dependent protein kinase III (Fig. 3A) (*36*), oxidation of tryptophan-221 by *N*-bromosuccinimide (NBS) (Fig. 3B) (*37, 38*), and ADP-ribosylation of histidine-715 by diphtheria toxin (Fig. 3C) (*39-41*). The latter threonine, the tryptophan, and the histidine residues are also conserved in the *Glugea* EF-2, but the former threonine is replaced by methionine, as it is in the EF-2 sequences of *Trypanosoma* (*42*) and *Dictyostelium* (*43*). Genomic Southern analysis using fragment II, which contains neither *Eco*RI or *Bam*HI site, revealed one band (Fig. 2B), suggesting that the putative EF-2 gene is present as a single copy in the *Glugea* genome.

In the nucleotide sequences of 5′ and 3′ noncoding regions of the putative *Glugea* EF-1α and EF-2 genes, there are no consensus sequences clearly similar to those commonly found in higher eukaryotes, such as a TATA box or a polyadenylation signal. Although we did not carry out further characterizations, these unusual features suggest that the mechanism of transcription and poly(A) tail metabolism in *Glugea* is different from that of higher eukaryotes.

EF-1α Phylogeny—In the EF-1α analysis, 46 eukaryotic and 3 archaebacterial sequences were used, and an NJ tree was first estimated. By applying a local rearrangement method of the ML analysis to the NJ tree, we reached a final tree whose log-likelihood is the best among alternatives examined. This is shown in Fig. 4 (the JTT model). Assuming archaebacteria to be an outgroup, three mitochondrion-lacking protozoa, *Gl. plecoglossi*, *Gi. lamblia*, and *Entamoeba histolytica*, represent respectively the first, second, and third earliest offshoots of the eukaryotes. Although the branching order among *E. histolytica*, *Plasmodium falciparum*, and *Tetrahymena pyriformis* remained unclear, as it was in our previous analyses (*13, 15, 17*), early branchings of *Gl. plecoglossi* and *Gi. lamblia* were reconfirmed by the present analysis using more sequence data than before. The branching order among green plants, *Dictyostelium*, and *Trypanosoma-Euglena* was also unclear, but the close relationship between animals and fungi was confirmed. Duplicated genes are observed in several species, but most of these duplications seem to have occurred recently on the external branches as far as the present data set is concerned.

The estimated branch length leading to *Gl. plecoglossi* in Fig. 4 is far longer than those leading to any other species, indicating a high substitution rate on the lineage of *Glugea*. Because the ML method is generally robust against the violation of constant substitution rate among lineages (*44*), the placing of *Glugea* as the earliest offshoot among eukaryotes does not seem to be due to a long branch artifact. However, since the branch length of *Glugea* is extremely long, concern remains that it may be an artifact derived from a paralogous comparison.

In addition to the recently duplicated EF-1αs as shown in Fig. 4, there are several paralogous proteins that have been shown to be derived from the gene duplication events. Among these proteins, yeast SUP2 or GST1 (*45, 46*) and its human homolog (*47*), yeast HBS1 (*48*), ATP sulfurylase of *Escherichia coli* (*49*), and nodulation protein Q from two eubacterial species (*50, 51*) are apparently alignable with a major part of the EF-1α/Tus. Sequence alignment analyses of these proteins, however, clearly demonstrated significant differences of alignment patterns between these proteins and EF-1α/Tus. In addition, the N- and/or C-terminal parts of these proteins had long extension sequences compared with the EF-1α/Tus. In the alignment including these proteins, the *Glugea* sequence clearly showed common features of eukaryotic EF-1α, suggesting that it is not a paralogous protein but a constitutive EF-1α from *Glugea*.

To investigate the evolutionary relationship among the eukaryotic/archaebacterial EF-1αs, the eubacterial EF-Tus, and the paralogous proteins, 35 selected sequences were further analyzed. Figure 5 shows a final tree inferred by the local rearrangement method. The *Glugea* sequence

	A (57 59)	B (221)	C (715)
Homo	RF-TDTRKDE	GSGLHGWAFT	HADAIHRGGGQII
Saccharomyces	RF-TDTRKDE	GSGLHGWAFT	HADAIHRGGGQII
Chlorella	RL-TDTRADE	SAGLHGWAFT	HADAIHRGGGQII
Dictyostelium	RY-MSCRADE	GSGLHGWGFT	HTDAIHRGGGRII
Trypanosoma	RI-MDTRADE	GSGLQAWAFS	HADAIHRGGGQII
Entamoeba	RY-TDTRPDE	GSGLHGWAFT	HADAIHRGGAQMI
Giardia	RF-TDTRQDE	GSGLHGWGFT	HADAIHRGAGQLT
Glugea	RY-MDTREDE	CSGLQGWGFT	HSDAIHRTGNQIS
Sulfolobus	LA-LDYLSVE	GSAKDKWGFS	HEDPAHRGPAQLY
Thermoplasma	LV-LDYDEQE	GSAYNNWAIS	HEDSIHRGPAQVI
Methanococcus	LA-LDFDEEE	GSAYNNWAIS	HEDAIHRGPSQII
Halobacterium	LA-MDTEEDE	GSALYKWGVS	HEDAIHRGPAQVI
Escherichia	AT-MDWMEQE	GSAFKNKGVQ	HDVDS--SELAFK
Thermus	AT-MDFMEQE	GSALKNKGVQ	HEVDS--SEMAFK
Glycine chl.	AT-MDWMEQE	GSAFKNKGVQ	HDVDS--SVLAFQ
Saccharomyces mit.	GAKMDSMDLE	GSALANTGIQ	HAVDS--NELSFK

Fig. 3. **Comparisons of amino acid sequences around the post-translationally modified sites characterized in mammalian EF-2.** Sequences used and their accession numbers are as follows: eukaryotic EF-2s—*Homo sapiens* (X51466), *Saccharomyces cerevisiae* (M59369), *Chlorella kessleri* (M68064), *Dictyostelium discoideum* (M26017), *Trypanosoma cruzi* (D50806), *Entamoeba histolytica* (D21259), *Giardia lamblia* (D29835), *Glugea plecoglossi* (D79220); archaebacterial EF-2s—*Sulfolobus acidocaldarius* (X54972), *Thermoplasma acidophilum* (X56840), *Methanococcus vannielii* (X12384), *Halobacterium halobium* (X17148); eubacterial and organelle EF-Gs—*Escherichia coli* (X00415), *Thermus thermophilus* (X16278), *Glycine max* chloroplast (X71439), *Saccharomyces cerevisiae* mitochondrion (X58378). Numbers above the alignments indicate sequence positions of *Homo* EF-2. A, Threonine-57 and -59 phosphorylated by protein kinase. B, Tryptophan-221 oxidized by NBS. C, Histidine-715 ADP-ribosylated by diphtheria toxin.

was clearly included within the EF-1α lineages of eukaryotes, and all the possibilities that the *Glugea* sequence is included in the eukaryotic paralogous proteins were excluded by the criterion of 1SE of log-likelihood difference from the tree in Fig. 5 (data not shown). This clearly reconfirmed that the present sequence is a genuine EF-1α from *Glugea*. In Fig. 5, paralogous protein sequences form a monophyletic clade, and the clade is placed on the outer group to eukaryotic/archaebacterial EF-1αs. However, the log-likelihood difference of the tree, in which a common ancestor of eukaryotic paralogous proteins (GST1, SUP2, and HBS1) diverges immediately before the *Glugea* EF-1α, was 9.0$\pm$ 10.0 ($\pm$ denotes 1SE), and the possibility could not be ruled out. The archaebacterial EF-1αs did not form a monophyletic group, but *Sulfolobus* EF-1α showed a close relationship with eukaryotic EF-1αs, being consistent with the eocyte tree proposed by Rivera and Lake (*52*). The relationship of the eocyte tree could not be confirmed in the present analysis, but results favoring the eocyte tree have also been derived by Hashimoto and Hasegawa (*53*) and by Baldauf *et al.* (*54*) on EF-1α/Tus and EF-2/Gs. Unfortunately, because we could only use 229 positions in the present analysis, the branching patterns in Fig. 5 could not be confirmed in detail.

Eukaryotic EF-1α is one of the most abundant cytoplasmic proteins, constituting between 3 and 10% of the soluble proteins (*23*). In addition to its role in translation elongation, EF-1α appears to have a number of other possible functions (*55*), which might be related to the existence of multiple functional genes. More than one copy of EF-1α genes appears to be present in most higher eukaryotic species, as shown in Fig. 4. These multi-copy genes have been shown to undergo cell-type or stage-specific expression in *Mucor racemosus* (*56, 57*), *Drosophila melanogaster* (*58, 59*), and *Xenopus laevis* (*60-62*). Differential expression of statin and EF-1α during rat development (*63, 64*) may also be included in the same category. It is interesting to further investigate the evolutionary history of EF-1α sequences with special reference to the evolution of eukaryotic cellular functions.

EF-2 Phylogeny—The ML tree of EF-2 on the basis of the JTT-F model was selected from the 135,135 possible tree topologies for the nine lineages: archaebacteria, *Glugea, Giardia, Cryptosporidium parvum, Entamoeba, Trypanosoma, Dictyostelium, Chlorella kessleri*, and animals/fungi. Archaebacteria are used as an outgroup, and the relationship among them was assumed in advance as (*Sulfolobus*, (*Methanococcus, Halobacterium*)). The tree is shown in Fig. 6. Being consistent with the EF-1α analysis, *Glugea* and

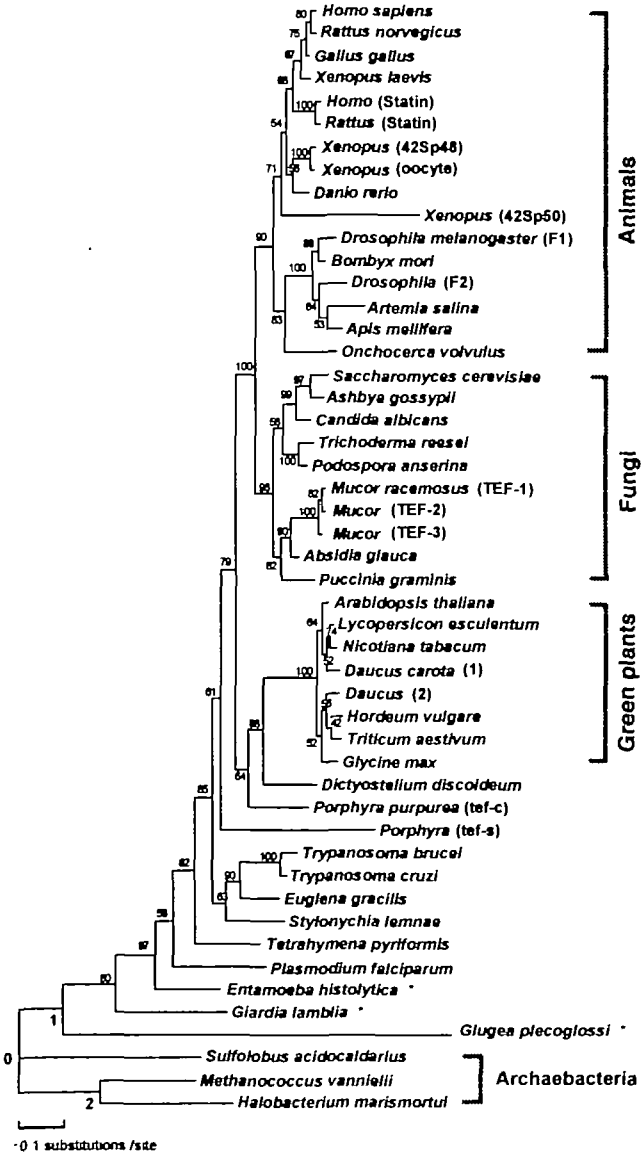

Fig. 4. **Eukaryotic tree of EF-1α with archaebacteria as an outgroup, by the ML method based on the JTT model.** The horizontal length of each branch is proportional to the estimated number of substitutions. Local bootstrap probabilities (*33, 34*) are attached to the internal branches. * denotes a mitochondrion-lacking protozoan species. The root of this tree is located either between nodes 0 and 1 (*24, 67*), or between nodes 0 and 2 (*52-54*). Sequences used in the analysis and their accession numbers are as follows: animals—*Homo sapiens* (X03558), *Rattus norvegicus* (X63561), *Gallus gallus* (L00677), *Xenopus laevis* (X55324), *Homo* (Statin) (X70940), *Rattus* (Statin) (M62751), *Xenopus* (42Sp48) (X56698), *Xenopus* (oocyte) (X52977), *Danio rerio* (L23807), *Xenopus* (42Sp50) (X52975), *Drosophila melanogaster* (F1) (X06869), *Bombyx mori* (D13338), *Drosophila* (F2) (X06870), *Artemia salina* (X03349), *Apis mellifera* (X52884), *Onchocerca volvulus* (M64333); fungi—*Saccharomyces cerevisiae* (X00779), *Ashbya gossypii* (X73978), *Candida albicans* (M29934), *Trichoderma reesei* (Z23012), *Podospora anserina* (X74799), *Mucor racemosus* (TEF-1) (J02605), *Mucor* (TEF-2) (X17476), *Mucor* (TEF-3) (X17475), *Absidia glauca* (X54730), *Puccinia graminis* (X73529); green plants—*Arabidopsis thaliana* (X16430), *Lycopersicon esculentum* (X53043), *Nicotiana tabacum* (U04632), *Daucus carota* (1) (X60302), *Daucus* (2) (D12709), *Hordeum vulgare* (Z23130), *Triticum aestivum* (M90077), *Glycine max* (X56856); other eukaryotes—*Dictyostelium discoideum* (X55972), *Porphyra purpurea* (tef-c) (U08844), *Porphyra* (tef-s) (U08841), *Trypanosoma brucei* (L25868), *Trypanosoma cruzi* (D29834), *Euglena gracilis* (X16890), *Stylonychia lemnae* (X57926), *Tetrahymena pyriformis* (D11083), *Plasmodium falciparum* (X60488), *Entamoeba histolytica* (M34256), *Giardia lamblia* (D14342), *Glugea plecoglossi* (D84253); archaebacteria—*Sulfolobus acidocaldarius* (X52382), *Methanococcus vannielii* (X05698), *Halobacterium marismortui* (X16677). On the basis of the alignment constructed, 369 positions were selected and used for the analysis. These are designated as the positions in the sequence of *Homo sapiens*: 21-120, 126-128, 131-157, 162-174, 176-182, 190-213, 226-331, 333-382, 385-390, and 395-427.

Giardia represented respectively the first and second earliest offshoots of eukaryotes, although the placement of *Cryptosporidium* was not conclusive. In the EF-2 analysis, we could not precisely resolve the relationship among *Dictyostelium*, *Chlorella*, *Entamoeba*, and *Trypanosoma*. There are many possible tree topologies for the relationship among these lineages, and most of them could not be discriminated from the ML tree with statistical significance. For example, a replacement of the positions of *Entamoeba* and *Chlorella* in the ML tree yields one of the likely relationships in the EF-1α analysis, and this tree could not be excluded. Therefore, the ML tree of EF-2 is not necessarily inconsistent with the results of EF-1α analysis. As in the EF-1α trees shown in Figs. 4 and 5, the branch length of *Glugea* EF-2 is also longer than those of other species, suggesting a high substitution rate. However, the degree of acceleration is not as great as in *Glugea* EF-1α. Gene duplications of eukaryotic EF-2s have also been identified in several species. As far as the present data are concerned, these duplications seem to have occurred recently on the branch leading to each species. For example, two genes, *EFT1* and *EFT2*, coding *Saccharomyces cerevisiae* EF-2, differ at only four third-codon positions in their coding regions (*65*).

Phylogenetic Position of Microsporidians and Early Evolution of Eukaryotes —To confirm the phylogenetic relationship among deep branching eukaryotes more precisely, the EF-1α and the EF-2 analyses were totally evaluated by adding the estimated log-likelihoods for each tree topology. Table I shows the results of the analysis. In this analysis, we considered 105 possibilities for the relationship among six lineages: archaebacteria, *Glugea*, *Giardia*, *Entamoeba*, *Trypanosoma*, and animals/fungi. Both the analyses consistently favored a tree in which three mitochondrion-lacking protozoa, *Glugea*, *Giardia*, and *Entamoeba*, branch off in this order (tree-1). When the independent data sets, EF-1α and EF-2, were totally considered, tree-1 became more likely, and the other trees could be ruled out by the 1SE criterion of log-likelihood differences. The deepest branching of *Gl. plecoglossi* in eukaryotes was also more clearly confirmed by the total evaluation. The ancient divergence of microsporidians in our analysis is also supported by the evidence that they have 70S ribosome lacking 5.8S ribosomal RNA, as mentioned in the Introduction.

Clark and Roger (*66*) recently obtained direct evidence that *Entamoeba* has secondarily lost its mitochondrion or still bears a relic of the mitochondrion by isolating two *E. histolytica* genes which encode proteins localized in the mitochondrion of higher eukaryotes. Analysis of the presence or absence of these mitochondrion-related genes in mitochondrion-lacking protozoa is very important to speculate which eukaryotic lineages diverged primitively before acquisition of the mitochondrion, but it has not been confirmed whether other mitochondrion-lacking protozoa also have such genes. In addition, Cavalier-Smith (*5*) proposed that *Trichomonas* might not be primitively amitochondrial, and their hydrogenosomes might have

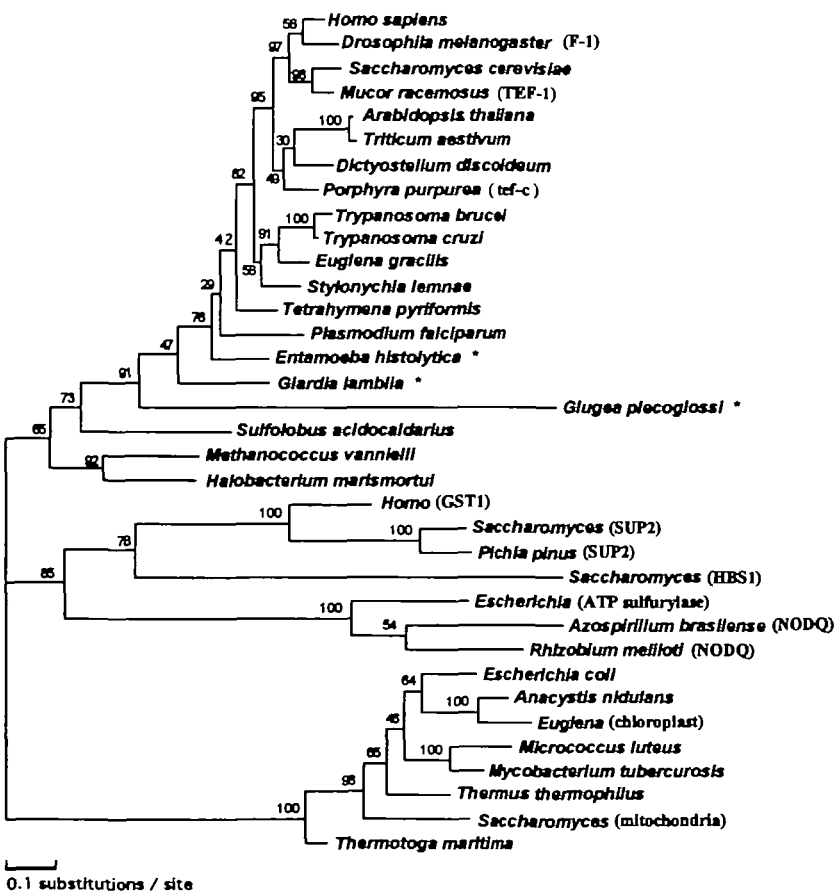

Fig. 5. **Unrooted tree showing the relationship among eukaryotic/archaebacterial EF-1αs, eubacterial EF-Tus, and paralogous sequences similar to the EF-1α/Tus.** The ML method based on the JTT model was used. The horizontal length of each branch is proportional to the estimated number of substitutions. Local bootstrap probabilities (*33, 34*) are attached to the internal branches. * denotes a mitochondrion-lacking protozoan species. Only representative sequences were used for eukaryotic EF-1αs. The paralogous sequences and eubacterial EF-Tu sequences used in the analysis and their accession numbers are as follows: paralogous sequences—*Homo* (GST1) (X17644), *Saccharomyces* (SUP2) (M21129), *Pichia pinus* (SUP2) (X56910), *Saccharomyces* (HBS1) (M98437), *Escherichia coli* (ATP sufurylase) (M74586), *Rhizobium meliloti* (NODQ) (X14809), *Azospirillum brasilense* (NODQ) (M94886); EF-Tus—*Escherichia coli* (J01690), *Micrococcus luteus* (M17788), *Thermotoga maritima* (M27479), *Thermus thermophilus* (X05977), *Mycobacterium tuberculosis* (X63539), *Anacystis nidulans* (SWISS-PROT P18668), *Euglena* (chloroplast) (X00044), *Saccharomyces* (mitochondrion) (K00428). On the basis of the alignment constructed, 229 positions were selected and used for the analysis. These are designated as the positions in the sequence of *Homo sapiens*: 21-32, 60-121, 127-128, 132-156, 177-182, 190-198, 210-213, 228-235, 238-255, 262-270, 276-278, 289-302, 305-308, 316-328, 340-349, 361-370, 394-404, 409-412, and 423-427.

TABLE I. **Phylogenetic relationships among *Glugea, Giardia, Entamoeba, Trypanosoma,* and animals and fungi by the ML analyses of EF-1α and EF-2, using archaebacteria as an outgroup.**[a]

Tree topology[b]	EF-1α JTT model		EF-2 JTT-F model		Total	
	Δl_i[c]	P_i[e]	Δl_i[c]	P_i[e]	Δl_i[c]	P_i[e]
1. (Gp,(Gl,(E,(T,AF))))	(−5975.2)[d]	.3873	(−8773.6)[d]	.4569	(−14748.8)[d]	.7677
2. (Gp,(Gl,(T,(E,AF))))	−6.5±12.9	.2355	−11.9±10.0	.0362	−18.3±16.3	.1152
3. (Gp,((Gl,E),(T,AF)))	−5.5± 8.5	.1603	−27.7± 9.7	.0000	−33.2±12.9	.0022
4. (Gl,(Gp,(E,(T,AF))))	−11.6± 7.0	.0071	−3.1± 6.4	.1792	−14.7± 9.5	.0367

[a]Of the 105 tree topologies for the six lineages, only four trees are shown. These could not be significantly discriminated from the ML tree on the basis of the criterion of 1SE of log-likelihood difference either in the EF-1α or EF-2 analysis. [b]Gp, *Glugea plecoglossi*; Gl, *Giardia lamblia*; E, *Entamoeba histolytica*; T, *Trypanosoma cruzi*; AF, animals and fungi. [c]Δl_i is the difference of log-likelihood of tree-*i* from that of the ML tree (tree-1), and ± is 1SE. [d]Log-likelihood of the ML tree. [e]Bootstrap probability estimated by 10^4 replications.

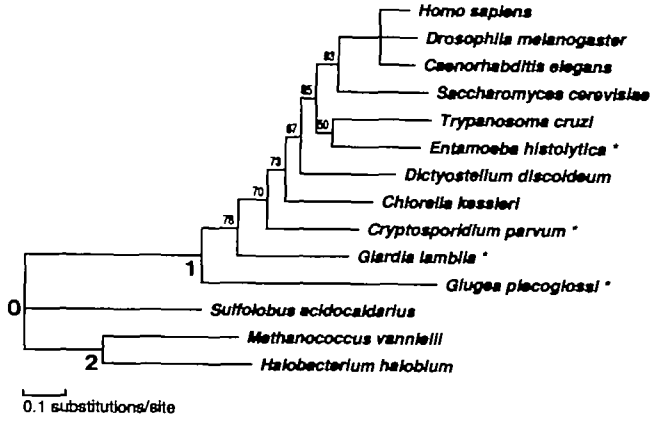

Fig. 6. **Eukaryotic tree of EF-2 with archaebacteria as an outgroup, by the ML method based on the JTT-F model.** The horizontal length of each branch is proportional to the estimated number of substitutions. Local bootstrap probabilities (*33, 34*) are attached to the internal branches. * denotes a mitochondrion-lacking protozoan species. The root of this tree is located either between nodes 0 and 1 (*24, 67*), or between nodes 0 and 2 (*52-54*). Sequences used in the analysis and their accession numbers are as follows: animals and fungi—*Homo sapiens* (X51466), *Drosophila melanogaster* (X15805), *Caenorhabditis elegans* (M72420); green algae—*Chlorella kessleri* (M68064); other eukaryotes—*Dictyostelium discoideum* (M26017), *Trypanosoma cruzi* (D50806), *Entamoeba histolytica* (D21259), *Cryptosporidium parvum* (U21667), *Giardia lamblia* (D29835), *Glugea plecoglossi* (D79220); archaebacteria—*Sulfolobus acidocaldarius* (X54972), *Methanococcus vannielii* (X12384), *Halobacterium halobium* (X17148). According to the alignment constructed, 542 positions were selected and used in the analysis. These are shown as follows by the positions for *Homo sapiens*: 30-52, 55-81, 98-188, 191-192, 210-224, 303-316, 344-401, 410-432, 447-476, 503-536, 539-564, 567-583, 605-622, 627-636, 640-649, 660-775, and 781-808.

evolved from mitochondria. Brown and Doolittle (*67*) further suggested by a phylogenetic analysis that the valyl-tRNA synthetase (ValRS) gene from *Trichomonas vaginalis* is clustered with other eukaryotic ValRS genes, which may have been transferred from the mitochondrial genome to the nuclear genome. This also supports the possibility that *Trichomonas* once harbored a mitochondrion-like organelle. Recent phylogenetic analyses by Yamamoto *et al.* including *Trichomonas tenax* EF-1α and EF-2 data revealed that the divergence of *T. tenax* is immediately next to *Giardia* (*68*). If the hypothesis that *Trichomonas* and *Entamoeba* are not primitively amitochondrial is true, and if the trees inferred in the present analyses reflect the true phylogenetic relationships, it would be reasonable to speculate that mitochondrial endosymbiosis occurred after *Glugea* and *Giardia* diverged successively from the lineage leading to higher eukaryotes both in the EF-1α (Fig. 4) and EF-2 (Fig. 6) trees.

On the other hand, *Cryptosporidium parvum* in the EF-2 tree (Fig. 6) has been considered to lack a mitochondrion, while *C. muris*, which has been shown to be close to *C. parvum* by the SrRNA analysis (*69*), has a mitochondrion. In this way, the secondary loss of the mitochondrion might have occurred in several lineages due, for example, to a parasitic lifestyle. Furthermore, the possibility remains that the mitochondrion has simply not yet been detected in some "mitochondrion-lacking" species. Therefore, it is difficult to determine solely from the phylogenetic analyses in which branch of the eukaryotic tree the mitochondrial endosymbiosis occurred.

Finally, many problems remain regarding the phylogenetic placement of protozoan species including mitochondrion-lacking ones. Further sequence data on the genes encoding various proteins from several protozoan species are highly desirable. Examinations of a deep branching phylogeny using these data will provide new and important insights into the origin and early evolution of eukaryotic cells.

We wish to thank Dr. S. Takahashi for providing us Ayu fish infected with *Gl. plecoglossi*; Dr. J. Adachi for expert technical support and discussion; and S. Kitahara, M. Nagashima, and A. Deguchi for technical assistance.

REFERENCES

1. Margulis, L. (1970) *Origin of Eukaryotic Cells*, Yale Univ. Press, New Haven, CT
2. Schwartz, R.M. and Dayhoff, M.O. (1978) Origins of prokaryotes, eukaryotes, mitochondria, and chloroplasts. *Science* **199**, 395-403
3. Cavalier-Smith, T. (1987) Eukaryotes with no mitochondria. *Nature* **326**, 332-333
4. Cavalier-Smith, T. (1989) Archaebacteria and archaezoa. *Nature* **339**, 100-101
5. Cavalier-Smith, T. (1993) Kingdom protozoa and its 18 phyla. *Microbiol. Rev.* **57**, 953-994
6. Sogin, M.L. (1991) Early evolution and the origin of eukaryotes. *Curr. Opin. Genet. Dev.* **1**, 457-463
7. Sogin, M.L. (1994) The origin of eukaryotes and evolution into major kingdoms in *Early Life on Earth. Nobel Symposium No. 84* (Bengston, S., ed.) pp. 181-192, Columbia University Press, New York
8. Sogin, M.L., Gunderson, J.H., Elwood, H.J., Alonso, R.A., and Peattie, D.A. (1989) Phylogenetic meaning of the kingdom

concept: An unusual ribosomal RNA from *Giardia lamblia*. *Science* **243**, 75-77
9. Leipe, D.D., Gunderson, J.H., Nerad, T.A., and Sogin, M.L. (1993) Small subunit ribosomal RNA$^+$ of *Hexamita inflata* and the quest for the first branch in the eukaryotic tree. *Mol. Biochem. Parasitol.* **59**, 41-48
10. Hasegawa, M. and Hashimoto, T. (1993) Ribosomal RNA trees misleading? *Nature* **361**, 23
11. Hasegawa, M., Hashimoto, T., and Adachi, J. (1992) Origin and evolution of eukaryotes as inferred from protein sequence data in *The Origin and Evolution of the Cell* (Haryman, H. and Matsuno, K., eds.) pp. 107-130, World Scientific, Singapore
12. Hashimoto, T., Adachi, J., and Hasegawa, M. (1992) Phylogenetic place of *Giardia lamblia*, a protozoan that lacks mitochondria. *Endocytobiosis Cell Res.* **9**, 59-69
13. Hashimoto, T., Nakamura, Y., Nakamura, F., Shirakura, T., Adachi, J., Goto, N., Okamoto, K., and Hasegawa, M. (1994) Protein phylogeny gives a robust estimation for early divergences of eukaryotes: Phylogenetic place of a mitochondria-lacking protozoan, *Giardia lamblia*. *Mol. Biol. Evol.* **11**, 65-71
14. Shirakura, T., Hashimoto, T., Nakamura, Y., Kamaishi, T., Cao, Y., Adachi, J., Hasegawa, M., Yamamoto, A., and Goto, N. (1994) Phylogenetic place of a mitochondria-lacking protozoan, *Entamoeba histolytica*, inferred from amino acid sequences of elongation factor 2. *Jpn. J. Genet.* **69**, 119-135
15. Hashimoto, T., Nakamura, Y., Kamaishi, T., Adachi, J., Nakamura, F., Okamoto, K., and Hasegawa, M. (1995) Phylogenetic place of kinetoplastid protozoa inferred from a protein phylogeny of elongation factor 1α. *Mol. Biochem. Parasitol.* **70**, 181-185
16. Hashimoto, T., Nakamura, Y., Kamaishi, T., Nakamura, F., Adachi, J., Okamoto, K., and Hasegawa, M. (1995) Phylogenetic place of mitochondrion-lacking protozoan, *Giardia lamblia*, inferred from amino acid sequences of elongation factor 2. *Mol. Biol. Evol.* **12**, 782-793
17. Kamaishi, T., Hashimoto, T., Nakamura, Y., Nakamura, F., Murata, S., Okada, N., Okamoto, K., Shimizu, M., and Hasegawa, M. (1996) Protein phylogeny of translation elongation factor EF-1α suggests microsporidians are extremely ancient eukaryotes. *J. Mol. Evol.* **42**, 257-263
18. Lom, J. and Dyková, I. (1992) *Protozoan Parasites of Fishes*, Elsevier, Amsterdam
19. Takvorian, P.M. and Cali, A. (1994) Enzyme histochemical identification of the Golgi apparatus in the microsporidian, *Glugea stephani*. *J. Euk. Microbiol.* **41** (Suppl), 63S-64S
20. Ishihara, R. and Hayashi, Y. (1968) Some properties of ribosomes from the sporoplasm of *Nosema bombycis*. *J. Invert. Pathol.* **11**, 377-385
21. Vossbrinck, C.R. and Woese, C.R. (1986) Eukaryotic ribosomes that lack a 5.8S RNA. *Nature* **320**, 287-288
22. Takahashi, S. and Egusa, S. (1977) Studies on *Glugea* infection of the Ayu, *Plecoglossus altivelis*—I. Description of the *Glugea* and a proposal of a new species, *Glugea plecoglossi* (in Japanese). *Fish Pathol.* **11**, 175-182
23. Merrick, W.C. (1992) Mechanism and regulation of eukaryotic protein synthesis. *Microbiol. Rev.* **56**, 291-315
24. Miyata, T., Iwabe, N., Kuma, K., Kawanishi, Y., Hasegawa, M., Kishino, H., Mukohata, Y., Ihara, K., and Osawa, S. (1991) Phylogenetic relationships among archaebacteria, eubacteria, and eukaryotes in *Evolution of Life: Fossils, Molecules, and Culture* (Osawa, S. and Honjo, T., eds.) pp. 337-351, Springer-Verlag, Tokyo
25. Hasegawa, M., Hashimoto, T., Adachi, J., Iwabe, N., and Miyata, T. (1993) Early branchings in the evolution of eukaryotes: Ancient divergence of Entamoeba that lacks mitochondria revealed by protein sequence data. *J. Mol. Evol.* **36**, 380-388
26. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) *Molecular Cloning*, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor
27. Thorne, J.L., Kishino, H., and Felsenstein, J. (1991) An evolutionary model for maximum likelihood alignment of DNA sequences. *J. Mol. Evol.* **33**, 114-124
28. Thorne, J.L., Kishino, H., and Felsenstein, J. (1992) Inching toward reality: An improved likelihood model of sequence evolution. *J. Mol. Evol.* **34**, 3-16
29. Thorne, J.L. and Churchill, G.A. (1995) Estimation and reliability of molecular sequence alignments. *Biometrics* **51**, 100-113
30. Kishino, H., Miyata, T., and Hasegawa, M. (1990) Maximum likelihood inference of protein phylogeny and the origin of chloroplasts. *J. Mol. Evol.* **31**, 151-160
31. Cao, Y., Adachi, J., Yano, T., and Hasegawa, M. (1994) Phylogenetic place of guinea pigs: No support of the rodent-polyphyly hypothesis from maximum-likelihood analyses of multiple protein sequences. *Mol. Biol. Evol.* **11**, 593-604
32. Cao, Y., Adachi, J., Janke, A., Pääbo, S., and Hasegawa, M. (1994) Phylogenetic relationships among eutherian orders estimated from inferred sequences of mitochondrial proteins: instability of a tree based on a single gene. *J. Mol. Evol.* **39**, 519-527
33. Adachi, J. (1995) *Modeling of Molecular Evolution and Maximum Likelihood Inference of Molecular Phylogeny*. Ph D Thesis of the Graduate University for Advanced Studies, The Institute of Statistical Mathematics, Tokyo
34. Adachi, J. and Hasegawa, M. (1996) *Computer Science Monographs, No. 28, MOLPHY Version 2.3: Programs for Molecular Phylogenetics Based on Maximum Likelihood*. The Institute of Statistical Mathematics, Tokyo
35. Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406-425
36. Price, N.T., Redpath, N.T., Severinov, K.V., Campbell, D.G., Russell, J.M., and Proud, C.G. (1991) Identification of the phosphorylation sites in elongation factor-2 from rabbit reticulocytes. *FEBS Lett.* **282**, 253-258
37. Guillot, D., Penin, F., Di Pietro, A., Sontag, B., Lavergne, J.-P., and Reboud, J.-P. (1993) GTP binding to elongation factor eEF-2 unmasks a tryptophan residue required for biological activity. *J. Biol. Chem.* **268**, 20911-20916
38. Guillot, D., Lavergne, J.-P., and Reboud, J.-P. (1993) Trp^{22} is involved in the protective effect of elongation factor eEF-2 on the ricin/α-sarcin site of the ribosome. *J. Biol. Chem.* **268**, 26082-26084
39. Kohno, K., Uchida, T., Ohkubo, H., Nakanishi, S., Nakanishi, T., Fukui, T., Ohtsuka, E., Ikehara, M., and Okada, Y. (1986) Amino acid sequence of mammalian elongation factor 2 deduced from the cDNA sequence: Homology with GTP-binding proteins. *Proc. Natl. Acad. Sci. USA* **83**, 4978-4982
40. Van Ness, B.G., Howard, J.B., and Bodley, J.W. (1980) ADP-ribosylation of elongation factor 2 by diphtheria toxin. *J. Biol. Chem.* **255**, 10710-10716
41. Nilsson, L. and Nygård, O. (1985) Localization of the sites of ADP-ribosylation and GTP binding in the eukaryotic elongation factor EF-2. *Eur. J. Biochem.* **148**, 299-304
42. Nakamura, Y., Hashimoto, T., Kamaishi, T., Adachi, J., Nakamura, F., Okamoto, K., and Hasegawa, M. (1996) Phylogenetic position of kinetoplastid protozoa inferred from the protein phylogenies of elongation factors 1α and 2. *J. Biochem.* **119**, 70-79
43. Toda, K., Tasaka, M., Mashima, K., Kohno, K., Uchida, T., and Takeuchi, I. (1989) Structure and expression of elongation factor 2 gene during development of *Dictyostelium discoideum*. *J. Biol. Chem.* **264**, 15489-15493
44. Hasegawa, M. and Fujiwara, M. (1993) Relative efficiencies of maximum likelihood, maximum parsimony, and neighbor-joining methods for estimating protein phylogeny. *Mol. Phylogenet. Evol.* **2**, 1-5
45. Wilson, P.G. and Culbertson, M.R. (1988) SUF12 supressor protein of yeast: A fusion protein related to the EF-1 family. *J. Mol. Biol.* **199**, 559-573
46. Kikuchi, Y., Shimatake, H., and Kikuchi, A. (1988) A yeast gene required for the G_1-to-S transition encodes a protein containing an A-kinase target site and GTPase domain. *EMBO J.* **7**, 1175-1182
47. Hoshino, S., Miyazawa, H., Enomoto, T., Hanaoka, F., Kikuchi, Y., Kikuchi, A., and Ui, M. (1989) A human homologue of the yeast *GST1* gene codes for a GTP-binding protein and is express-

ed in a proliferation-dependent manner in mammalian cells. *EMBO J.* **8**, 3807-3814
48. Nelson, R.J., Ziegelhoffer, T., Nicolet, C., Werner-Washburne, M., and Craig, E.A. (1992) The translation machinary and 70 kd heat shock protein cooperate in protein synthesis. *Cell* **71**, 97-105
49. Leyh, T.S., Vogt, T.F., and Suo, Y. (1992) The DNA sequence of the sulfate activation locus from *Escherichia coli* K-12. *J. Biol. Chem.* **267**, 10405-10410
50. Cervantes, E., Sharma, S.B., Maillet, F., Vasse, J., and Truchet, G. (1989) The *Rhizobium meliloti* host-range nodQ gene encodes a protein which shares homology with translation elongation and initiation factors. *Mol. Microbiol.* **3**, 745-755
51. Vieille, C. and Elmerich, C. (1990) Characterization of two *Azospirillum brasilense* Sp7 plasmid genes homologous to *Rhizobium meliloti* nodPQ. *Mol. Plant Microbe. Interact.* **3**, 389-400
52. Rivera, M.C. and Lake, J.A. (1992) Evidence that eukaryotes and eocyte prokaryotes are immediate relatives. *Science* **257**, 74-76
53. Hashimoto, T. and Hasegawa, M. (1996) Origin and early evolution of eukaryotes inferred from the amino acid sequences of translation elongation factors 1α/Tu and 2/G. *Adv. Biophys.* **32**, 73-120
54. Baldauf, S.L., Palmer, J.D., and Doolittle, W.F. (1996) The root of the universal tree and the origin of eukaryotes based on elongation factor phylogeny. *Proc. Natl. Acad. Sci. USA* **93**, 7749-7754
55. Condeelis, J. (1995) Elongation factor 1α, translation and the cytoskeleton. *Trends Biochem. Sci.* **20**, 169-170
56. Sandstrom, P., Lira, L.M., Choi, D., Linz, J.E., and Sypherd, P.S. (1987) Sequence analysis of the EF-1α gene family of *Mucor racemosus*. *Nucleic Acids Res.* **23**, 9997-10005
57. Linz, J.E. and Sypherd, P.S. (1987) Expression of three genes for elongation factor 1α during morphogenesis of *Mucor racemosus*. *Mol. Cell Biol.* **7**, 1925-1932
58. Walldorf, U., Hovemann, B., and Bautz, E.K.F. (1985) F1 and F2: Two similar genes regulated differently during development of *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* **82**, 5795-5799
59. Hovemann, B., Richter, S., Walldorf, U., and Cziepluch, C. (1988) Two genes encode related cytoplasmic elongation factors 1α (EF-1α) in *Drosophila melanogaster* with continuous and stage specific expression. *Nucleic Acids Res.* **16**, 3175-3194
60. Krieg, P.A., Varnum, S.M., Wormington, W.M., and Melton, D.A. (1989) The mRNA encoding elongation factor 1-α (EF-1α) is a major transcript at the midblastula transition in Xenopus. *Dev. Biol.* **133**, 93-100
61. Djé, M.K., Mazabraud, A., Viel, A., le Marie, M., Denis, H., Crawford, E., and Brown, D.D. (1990) Three genes under different developmental control encode elongation factor 1-α in *Xenopus laevis*. *Nucleic Acids Res.* **18**, 3489-3493
62. Coppard, N.J., Poulsen, K., Madsen, H.O., Frydenberg, J., and Clark, B.F.C. (1991) 42Sp48 in previtellogenic *Xenopus* oocytes is structurally homologous to EF-1α and may be a stage-specific elongation factor. *J. Cell Biol.* **112**, 237-243
63. Ann, D.K., Lin, H.H., Lee, S., Tu, Z.-J., and Wang, E. (1992) Characterization of the statin-like S1 and rat elongation factor 1α as two distinctly expressed messages in rat. *J. Biol. Chem.* **267**, 699-703
64. Lee, S., Wolfraim, L.A., and Wang, E. (1993) Differential expression of S1 and elongation factor-1α during rat development. *J. Biol. Chem.* **268**, 24453-24459
65. Perentesis, J.P., Phan, L.D., Gleason, W.B., LaPorte, D.C., and Livingston, D.M. (1992) *Saccharomyces cerevisiae* elongation factor 2: genetic cloning, characterization of expression, and G-domain modeling. *J. Biol. Chem.* **267**, 1190-1197
66. Clark, C.G. and Roger, A.J. (1995) Direct evidence for secondary loss of mitochondria in *Entamoeba histolytica*. *Proc. Natl. Acad. Sci. USA* **92**, 6518-6521
67. Brown, J.R. and Doolittle, W.F. (1995) Root of the universal tree of life based on ancient aminoacyl-tRNA synthetase gene duplications. *Proc. Natl. Acad. Sci. USA* **92**, 2441-2445
68. Yamamoto, A., Hashimoto, T., Asaga, E., Hasegawa, M., and Goto, N. (1996) Phylogenetic position of mitochondrion-lacking protozoan, *Trichomonas tenax*, based on amino acid sequences of elongation factors 1α and 2. *J. Mol. Evol.*, in press
69. Cai, J., Collins, M.D., McDonald, V., and Thompson, D.E. (1992) PCR cloning and nucleotide sequence determination of the 18S rRNA genes and internal transcribed spacer 1 of the protozoan parasites *Cryptosporidium parvum* and *Cryptosporidium muris*. *Biochim. Biophys. Acta* **1131**, 317-320