Glutamine Synthetase Gene Evolution in Bacteria

Graziano Pesole, Carmela Gissi,* Cecilia Lanave,† and Cecilia Saccone**

*Dipartimento di Biochimica e Biologia Molecolare, Universitá di Bari and †Centro Studi sui Mitocondri e Metabolismo Energetico, Bari, Consiglio Nazionale delle Ricerche, Italy

The evolution of the prokaryotic glutamine synthetase (GS) genes, namely the GSI and GSII isoforms, has been investigated using the second codon positions, which have previously proven to behave as a good molecular clock. Our data confirm the early divergence between prokaryotic and eukaryotic GSII before the splitting between plants and animals. The phylogenetic tree of the GSI isoforms shows Archaebacteria to be more closely related to Eubacteria than to Eukaryotes. This finding is confirmed by the phylogenetic analysis carried out on both large and small subunits of rRNA. However, differently from the rRNA analyses, *Crenarchaeota* and *Euryarchaeota* Archaebateria, as well as high- and low-GC gram-positive bacteria, appear to be polyphyletic. We provide evidence that the observed polyphyly of Archaebacteria might be only apparent, resulting from a gene duplication event preceding the split between Archaebacteria and Eubacteria and followed by the retention of only one isoform in the extant lineages. Both gram-negative bacteria and high-GC gram-positive bacteria, which appear closely related, have GS activity regulated by an adenylylation/deadenylylation mechanism. A lateral gene transfer from Archaebacteria to low-GC eubacteria is invoked to explain the observed polyphyly of gram-positive bacteria.

Introduction

Glutamine synthetase (GS), a pivotal enzyme for nitrogen metabolism, is found in at least three distinct forms. In the studies reported up to now, one of these forms, GSI, has been primarily associated with Prokaryotes, a second form, GSII, with Eukaryotes, and a third type, GSIII, has been recently found in the anaerobe *Bacteroides fragilis* (Hill et al. 1989) and *Butyrivibrio fibrisolvens* (Goodman and Woods 1993). These three types of enzymes are distinct in their primary as well as tertiary structures. GSI is made up of 12 identical subunits, arranged in two layers of 6. The active site, marked by a pair of Mn^{++} ions appears to be formed by two antiparallel β-structures, the one in the C domain of one subunit and the other in the N domain of the neighboring subunit (Yamashita et al. 1989). GSII has eight subunits and GSIII six subunits. A fourth type, glnT, found in *Rhizobium leguminosarum* (Chiurazzi et al. 1992), is more related to bacterial GSI. The GSIII isoenzyme, much longer than GSI and GSII, cannot be aligned with these isoform enzymes even if some common conserved boxes are found. In some bacteria both the prokaryotic enzyme GSI and the eukaryotic enzyme GSII have been found.

In a previous paper (Pesole et al. 1991), we demonstrated that GS behaves as an ideal molecular clock. By using the second codon positions of the GS enzymes, very reliable genetic distances between remotely related organisms can be obtained which also display ultrametricity. The sequences were analyzed by the stationary Markov clock (SMC) method devised in our laboratory (Lanave et al. 1984; Saccone et al. 1990) which allows a very accurate measurement of the distances between sequences. The calibration of the trees obtained in our previous studies showed that the divergence time between GSI and GSII is about 2,500 My. In addition, we demonstrated that the common ancestor of the GSII genes for plants and bacteria was more ancient than the divergence of animals and plants.

In this paper, we have extended our previous analyses; in particular, we have studied in detail the evolutionary behavior of the glutamine synthetase in Eubacteria and Archaebacteria in order to reach a deeper insight into the evolutionary history of prokaryotic organisms. The data obtained with the GS genes have been compared with those obtained using rRNA sequences. These studies relate to some important and controversial problems, in particular to the use of nucleic acid sequences for tracing phylogenetic relationships between organisms and the classification of life kingdoms.

Key words: tree of life, archaebacteria, phylogenetic trees, Markov process.

Address for correspondence and reprints: Cecilia Saccone, Dipartimento di Biochimica e Biologia Molecolare, Universitá di Bari, via Orabona, 4, 70126 Bari, Italy.

Mol. Biol. Evol. 12(2):189–197. 1995.

0737-4038/95/1202-0001$02.00

Methods

GS nucleotide sequences were extracted from the EMBL database and multialigned on the basis of the amino acid alignment previously proposed by us (Pesole et al. 1991). Large ribosomal subunit RNA sequences were extracted in the aligned format from the RDP database (Olsen et al. 1992). Small ribosomal subunit RNA sequences were extracted from the De Watcher compilation (De Rijk et al. 1992).

Distance measures between pairs of nucleotide sequences were carried out by using a stochastic model of gene evolution, the stationary Markov clock (Saccone et al. 1990). From pairwise distances, phylogenetic trees were calculated by using the programs NEIGHBOR, CONSENSE, and DRAWGRAM of the PHYLIP package (Felsenstein 1993). In the analysis of GS genes only second codon positions were used. Only the sites aligned in all sequences under examination, with no insertions or deletions, were considered in the evolutionary analyses. The rRNA genes have been analyzed by the maximum-likelihood method with the DNAMLK program of the PHYLIP package as they do not obey the stationary condition.

Results

All the GS sequences available in the database have been considered in this study. In particular, the phylogenetic analysis has focused on the 33 bacterial sequences and has used three eukaryotic sequences, namely human, drosophila, and alfalfa (a plant) for calibration purposes. The second codon positions of all genes obey the stationary condition (e.g., display the same base composition, within statistical fluctuations, at equivalent sites), which is a prerequisite for the applicability of the SMC model. This indicates that the second codon positions are not affected by the directional mutation pressure that, according to Sueoka (1988), determines the G+C content of the third codon positions. This is shown in figure 1 in which the extent of directional mutation pressure on the glutamine synthetase gene evolution can be measured by the plot of the G+C content at the first and second codon positions against that of the third codon positions. Figure 1 shows that, despite the high variability of the G+C content in the third codon positions, ranging from 13% to 98%, the G+C content of the second codon positions remained unchanged, showing no correlation with the G+C of the third codon positions. On the contrary, if we plot the G+C content of the first codon positions against that of the third ones, a significant correlation is observed which demonstrates that directional mutation pressures (Sueoka 1988), or more generally compositional constraints (Bernardi et al. 1985; Bernardi 1993), affect the base composition of the first codon positions. For this reason only the second codon positions of the GS genes were used for our analysis.

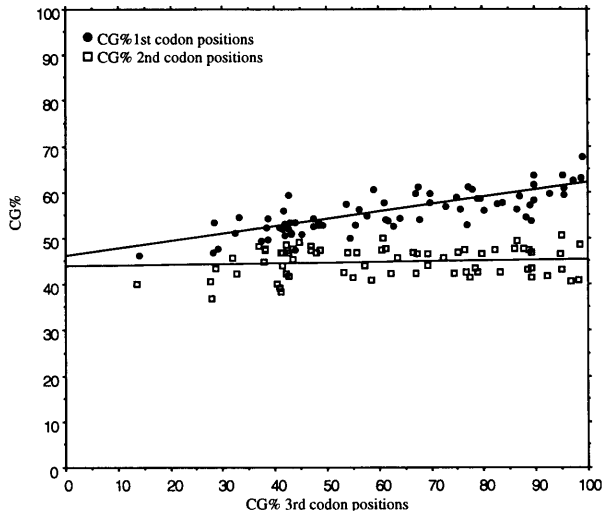

FIG. 1.—Plot of GC% of first and second codon positions against that of third codon positions for the 71 GS gene sequences available in the database (accession numbers are reported in parentheses): animals (GSII): *Cricetulus longicaudatus* (X03495), *Drosophila melanogaster* (X52759, X52760), *Gallus gallus* (M29076), *Homo sapiens* (X59834), *Mus musculus* (M60803), *Palinurus argus* (M96798), *Rattus norvegicus* (M91652), *Squalus acanthias* (U04617); plants (GSII): *Arabidopsis thaliana* (S69727), *Brassica napus* (X72751, X76736), *Glycine max* (S46513), *Hordeum vulgare* (X53580), *Lactuca sativa* (X60092), *Medicago sativa* (X03931), *Nicotiana plumbaginifolia* (M19055), *Nicotiana sylvestris* (X66940), *Oryza sativa* (X14246, X14245, X14244), *Phaseolus vulgaris* (X14605, X04001, X04002, X12738), *Pinus sylvestris* (X69822), *Pisum sativum* (X04763, M20664, M20663), *Saccharomyces cerevisiae* (M65157), *Vigna aconitifolia* (M94765), *Zea mays* (X65926, X65927, X65928, X65929, X65931, D14578); bacterial GSII: *Bradyrhizobium japonicum* (X04187), *Frankia sp.* (M58415), *Rhizobium meliloti* (X17523), *R. leguminosarum* (X67296), *Streptomyces hygroscopicus* (M33783), *S. viridochromogenes* (X52842); bacterial GSI: *Agmenellum quadruplicatum* (Z13965), *Anabaena* sp. (X00147), *Azospirillum brasilense* (M26107), *Azotobacter vinelandii* (M57275), *Bacillus cereus* (D00513), *B. subtilis* (D00854), *Calothrix* sp. (L05609), *Clostridium acetobutylicum* (M18966), *Escherichia coli* (X05173), *Frankia* sp. (L10631), *Haloferax volcanii* (U03029), *Lactobacillus delbrueckii* (D10020), *Methanococcus voltae* (X53509), *Methylococcus capsulatus* (M28472), *Neisseria gonorrhoeae* (M84113), *Proteus vulgaris* (X68129), *Pyrococcus furiosus* (L12410), *P. woesei* (X60161), *Rhizobium leguminosarum* (X04880), *Rhodobacter sphaeroides* (X71659), *Salmonella typhimurium* (M14536), *Streptomyces coelicolor* (M23172), *S. viridochromogenes* (X70924), *Sulfolobus solfataricus* (X53263), *Thermotoga maritima* (X60160), *Thiobacillus ferrooxidans* (M16626), *Vibrio alginolyticus* (L08499); *R. leguminosarum* glnT (S48357). The regression line equations are $y = 0.016x + 46.3$, $r^2 = 0.63$ (first codon position GC%) and $y = 0.02x + 43.9$, $r^2 = 0.013$ (second codon position GC%).

Figure 2 shows the multiple alignments of the consensus sequences representative of the GS genes considered in this analysis. Indeed, the reliability of any evolutionary analysis at the molecular level relies on the

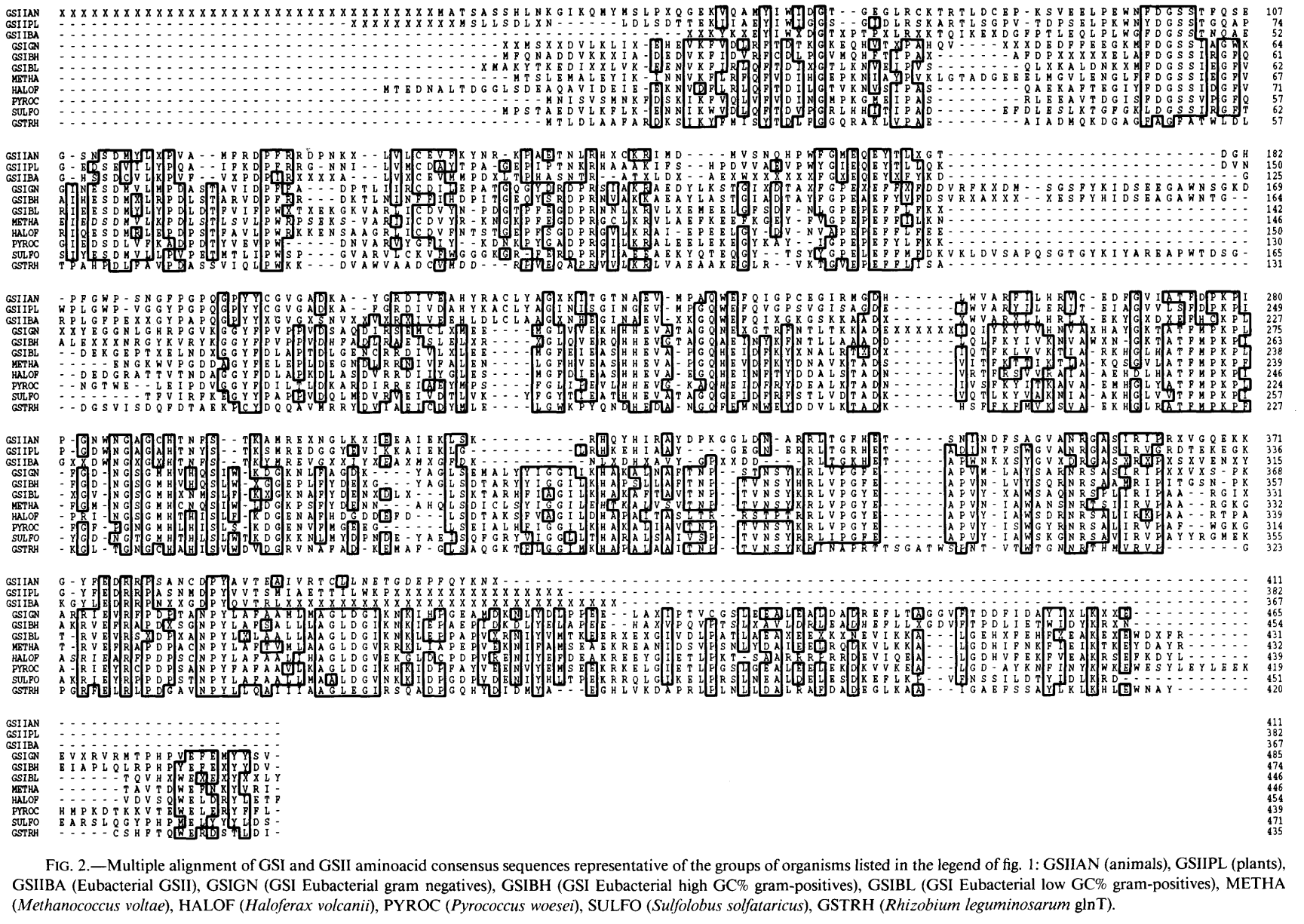

FIG. 2.—Multiple alignment of GSI and GSII aminoacid consensus sequences representative of the groups of organisms listed in the legend of fig. 1: GSIIAN (animals), GSIIPL (plants), GSIIBA (Eubacterial GSII), GSIGN (GSI Eubacterial gram negatives), GSIBH (GSI Eubacterial high GC% gram-positives), GSIBL (GSI Eubacterial low GC% gram-positives), METHA (*Methanococcus voltae*), HALOF (*Haloferax volcanii*), PYROC (*Pyrococcus woesei*), SULFO (*Sulfolobus solfataricus*), GSTRH (*Rhizobium leguminosarum* glnT).

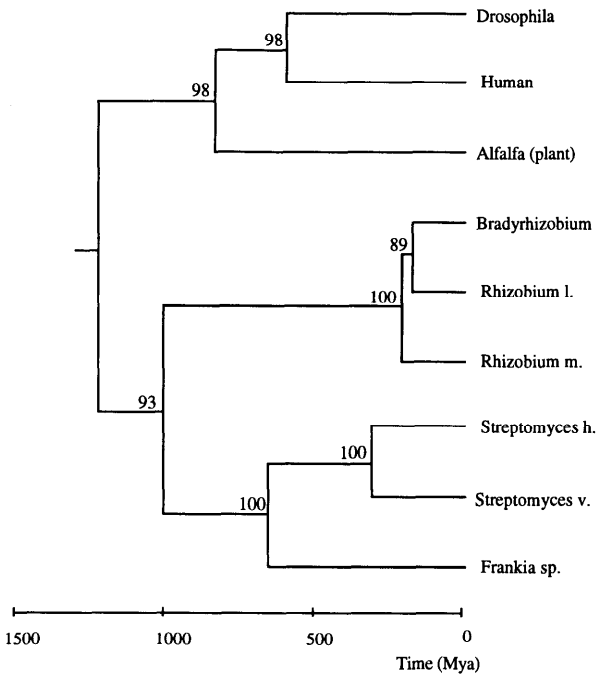

FIG. 3.—Phylogenetic tree calculated on Eukaryotic and Eubacterial GSII second codon positions by using SMC, NEIGHBOR, CONSENSE, and DRAWGRAM programs. The split between vertebrates and invertebrates at 600 Mya has been used as reference time. Bootstrap values out of 100 replicates are shown on each node.

accuracy of the multiple alignment. For this reason the multiple alignment generated by the computer has been carefully checked, by comparison with our previous alignment (Pesole et al. 1991) and with known structural data (Yamashita et al. 1989). It is striking to note that we obtained a degree of similarity between as distantly related genes as GSI and GSII significantly higher than reported by other authors (Tischer et al. 1986; Gupta and Golding 1993). For instance, the degree of aminoacid similarity between human and *Escherichia coli* increases from 11%, reported by Gupta and Golding (1993), to the 20.8%, based on our alignment.

Figure 3 shows the phylogenetic tree on the nine GSII isoform enzymes. The split between eukaryotic and prokaryotic GSII dates back at 1,260 ± 280 Mya. This result confirms our previous hypothesis (Pesole et al. 1991) of an early divergence between eukaryotic and prokaryotic GSII, about 300 My before the splitting between plants and animals. GSII genes from gram-positives and gram-negatives cluster into two distinct groups, diverging 1,000 ± 230 Mya.

In order to test whether both GSI and GSII genes are suitable for measuring the genetic distance between bacterial species, we compared their nucleotide substitution rates by using organisms for which both isoform gene sequences were available, namely *Rhizobium leguminosarum, Frankia* sp., and *Streptomyces viridochromogenes.* The large overlapping observed between the rate values, as a result of their standard deviations, calculated on both GSI and GSII, as shown in table 1, confortably supports the rate constancy hypothesis for the GSI and GSII isoforms at the level of the second codon positions.

The phylogeny of GSI genes is shown in figure 4. The time scale has been calculated assuming the substitution rate, $2.2 \cdot 10^{-10}$ subs/site-year, observed at the second codon positions of GSII isoenzymes. It is noteworthy that all gram-negative bacteria cluster together whereas gram-positives are split in two distinct groups, the one more closely related to cyanobacteria, gram-negative bacteria and to the archaebacterium *Sulfolobus solfataricus,* containing high-GC species, the other containing low-GC species more akin to the remaining Archaebacteria, namely *Methanococcus voltae, Pyrococcus woesei,* and *Haloferax volcanii.*

It is interesting to note that the same tree topology shown in figures 3–4 has been obtained by analyzing the encoded protein sequences with the program PROTDIST (Felsenstein 1993). The tree depicting the relationships between GSI, GSII, and glnT is shown in figure 5 in which the divergence between human and drosophila at 600 My has been used as calibration time. The glnT sequence of *R. leguminosarum,* the only available, is clearly more closely related to the GSI genes from which it diverged 1,584 ± 348 Mya. The split be-

Table 1
Nucleotide Substitution Rates Calculated at the Second Codon Positions for Pairs of Organisms with Both GSI and GSII Isoform Gene Sequences Available

	NUCLEOTIDE SUBSTITUTION RATE (subs./site)	
ORGANISMS COMPARED	GSI	GSII
Rhizobium leguminosarum—Frankia sp.	0.35 ± 0.13	0.42 ± 0.19
R. leguminosarum—Streptomyces viridochromogenes	0.37 ± 0.13	0.37 ± 0.17
Frankia sp.—*S. viridochromogenes*	0.23 ± 0.10	0.30 ± 0.13

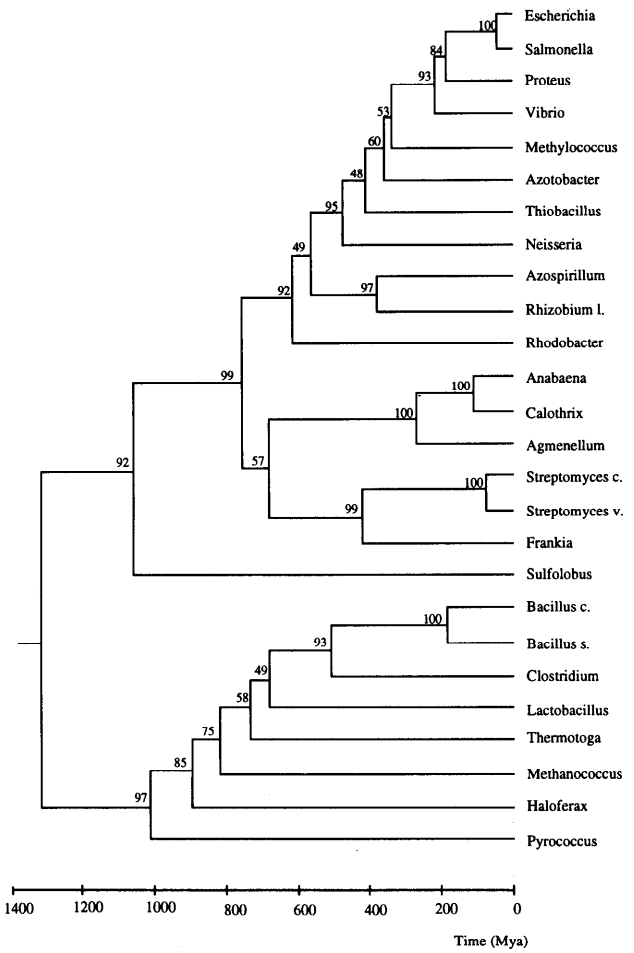

FIG. 4.—Phylogenetic tree calculated on Eubacterial and Archaebacterial GSI second codon positions by using SMC, NEIGHBOR, CONSENSE, and DRAWGRAM programs. The nucleotide substitution rate calculated on GSII second codon positions ($2.2 \cdot 10^{-10}$ subs./site.year) is assumed for the tree calibration. Bootstrap values out of 100 replicates are shown on each node.

tween GSI/glnT and GSII dates back at 2,484 ± 570 Mya, in agreement with our previous analyses (Pesole et al. 1991). This remote time might correspond either to the divergence between prokaryotes and Eukaryotes or to a gene duplication event preceding the divergence between prokaryotes and Eukaryotes.

The Archaebacteria, despite other molecular analyses (Woese 1987), appear to be polyphyletic (see below) but all more closely related to Eubacterial GSI genes. This closer relationship between Archaebacteria and Eubacteria is cosupported by the evolutionary analyses carried out on the large and small ribosomal RNA subunit. Since rRNA genes are not stationary in their base composition, the phylogenetic trees shown in figures 6*a–b* have been inferred by applying the maximum-likelihood method. It is striking to note that both SSU and LSU rRNA give the same tree topology, with organisms clearly split into eukaryotic and prokaryotic clusters.

Discussion

In a previous paper we reported that GS enzymes behave as good clocks (Pesole et al. 1991). This stems from the fact that, according to the SMC method, by fixing as external input, a divergence time inferred from nonmolecular data between a couple of species, we were able to derive divergence times between distantly related organisms which agree with the morphological and paleontological data. For example, fixing the invertebrate-vertebrate distance at 600 Mya, we calculated the divergence between human and rodents at 75 Mya, which is remarkably in agreement with previous paleontological estimates. These results have been confirmed by other authors (Kumada et al. 1993).

This metronomic clocklike behavior is rather exceptional because normally it is limited to short time spans, as saturation phenomena or drift effects take place in longer divergence times. The genetic distances between GS enzymes were measured on the second codon

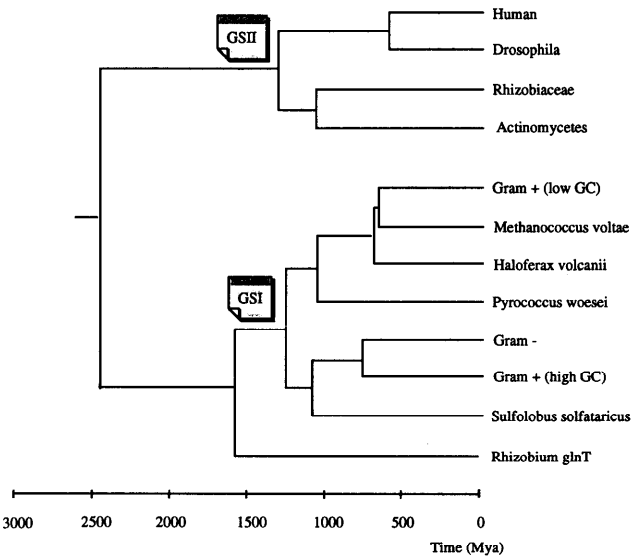

FIG. 5.—Phylogenetic tree calculated on the second codon positions of Eubacterial, Eukaryotic, and Archaebacterial GSI and GSII genes by using SMC, NEIGHBOR, and DRAWGRAM programs. For the sake of clarity, closely related groups of organisms have been reported under a single label: Rhizobiaceae: *Bradyrhizobium japonicum, Rhizobium leguminosarum* (GSII), *R. meliloti;* Actinomycetes: *Streptomyces hygroscopicus, S. viridochromogenes* (GSII), *Frankia* sp. (GSII); Gram + (low GC): *Bacillus cereus, B. subtilis, Clostridium acetobutylicum, Lactobacillus delbrueckii;* Gram −: *Agmenellum quadruplicatum, Azospirillum brasilense, Azotobacter vinelandii, Escherichia coli, Methylococcus capsulatus, Neisseria gonorrhoeae, Proteus vulgaris, Rhodobacter sphaeroides, Salmonella typhimurium, Thiobacillus ferrooxidans, Vibrio alginolyticus;* Gram + (high GC): *Streptomyces coelicor, S. viridochromogenes* (GSI), *Frankia* sp. (GSI).

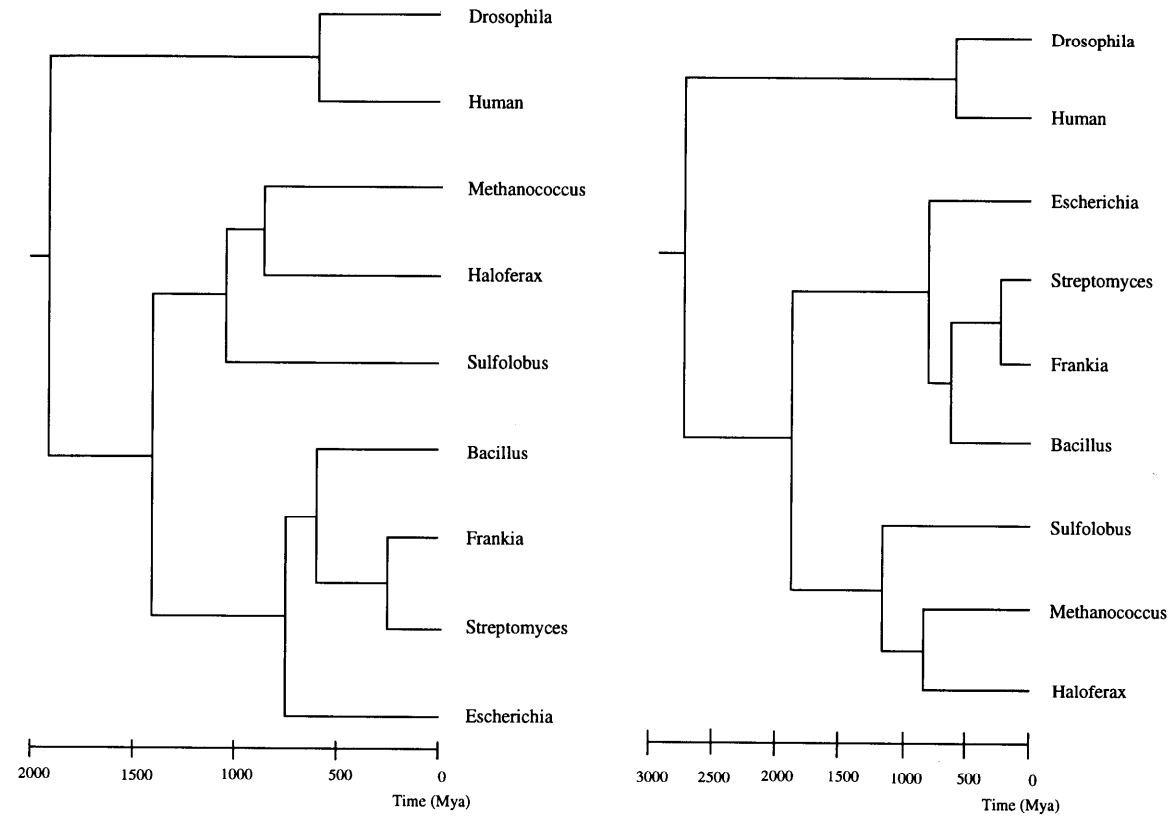

FIG. 6.—Phylogenetic tree calculated on lsu (*a*) and ssu (*b*) rRNA by using the DNAMLK (PHYLIP), NEIGHBOR, and DRAWGRAM programs. The split between vertebrates and invertebrates at 600 Mya has been used as reference time.

positions of the sequences which proved to be stationary (e.g., display the same base composition, within the statistical fluctuations) in all species for both GSI and GSII as shown in figure 1.

As we pointed out, several times in our previous papers (Saccone et al. 1993), the stationary condition of the nucleotide sequences is an important criterion to be fulfilled when nucleic acids are used as molecular clocks. It implies that no compositional bias arose during the evolution of the macromolecules in the different species under study, thus allowing us to assume that the rate matrix of nucleotide substitutions (i.e., the propensity of a given nucleotide to be changed into the other three) remained the same in all lineages. This is in turn a prerequisite for a molecular clock hypothesis. In our opinion, one of the reasons for such an exceptionally good clocklike behavior of these enzymes can reside in the lack of directional mutation pressure and in the homogeneous pattern of nucleotide substitutions along the entire gene at the level of the second codon positions as measured by our CONSTRAINT program (Pesole et al. 1992) (results not shown).

Our observation of the evolutionary rate constancy of the GS second codon positions was misinterpreted by Kumada et al. (1993), who inferred that we assumed it as being due to Darwinian selection. Actually, we ascribed such a phenomenon neither to neutral nor to the selective mutation forces. We only emphasized that the regular clocklike behavior occurred at the most constrained positions, whereas first and third codon positions evolved under clear directional mutation pressure (see fig. 1).

In this report, extending our previous studies to other GS enzymes, in particular to the GS isoforms found in Eubacteria and Archaebacteria, we found that Archaebacteria appear always more closely related to Eubacteria than to Eukaryotes. The two phyla of Archaebacteria, namely *Crenarchaeota* and *Euryarchaeota*, appear very distantly related, in both GS and rRNA subunit analysis. However, the two groups appear to be monophyletic in the analysis of both large and small rRNA subunits, whereas in GS analyses they appear to be polyphyletic, with *Euryarchaeota* clustering with the low GC gram-positives and *Crenarchaeota* with the gram-negatives.

The gram-negative bacteria cluster together and appear more closely related to the high-GC gram-positives; indeed, low-GC gram-positive bacteria are more closely related to *Euryarchaeota* Archaebacteria. Interestingly, both gram-negatives and high-GC gram posi-

tives GS are regulated by the adenylytation/deadenylylation system.

Moreover, the distance between GSI and GSII isoforms, the former found in Eubacteria and Archaebacteria and the latter found in Eukaryotes and also in some bacteria, supports our previous data which date the origin of the bacterial GSII-type isoform before 1,000 Mya. This time might correspond to a possible horizontal transfer from Eubacteria to Eukaryotes that occurred before the splitting between plants and animals (Pesole et al. 1991) or alternatively to the split between Prokaryotes and Eukaryotes (Kumada et al. 1993), which thus should have occurred 1,260 ± 280 Mya. We think the second hypothesis unlikely, as the divergence time between Eukaryotes and prokaryotes inferred from the SSU and LSU rRNA analysis is much higher, about 2,300 Mya, which in turn is in good agreement with the datation of the split between GSI and GSII at about 2,500 Mya and other nonmolecular considerations (Sogin 1991).

It is intriguing to note (see fig. 4) that the root of the two clusters of GSI, the one containing gram-negative, high-GC gram-positive and *Sulfolobus solfataricus* and the other containing low-GC bacteria and the other three Archaebacteria, have exactly the same datation, about 1,100 Mya, thus suggesting that these two gene clusters probably originated from a single gene duplication event. In such a case the polyphyly of Archaebacteria as well as of gram-positive bacteria might be only apparent due to the peculiar story of the GS gene. This hypothesis is schematically represented in figure 7. After the duplication event, the two GS genes diverged independently and the extant organisms retained only one of the two forms. It should be considered that GS activity is regulated by a reversible mechanism of deadenylylation (activation) adenylylation (deactivation) in both gram-negative and high-GC gram-positive bacteria involving a tyrosine residue which is conserved in all the members of this cluster including *S. solfataricus,* whereas the tyrosine is not always conserved in the members of the other cluster whose GS activity is not regulated by adenylylation. This hypothesis is further supported by a 25–30 aa region uniquely shared by gram-negative, high-GC gram-positive bacteria and *S. solfataricus* (aa 139–165 in *S. solfataricus;* see fig. 2). We suggest that the gene duplication, about 1,300 Mya, originated two GS isoforms, one of which acquired the adenylylation regulation mechanism. We then will refer to the two GS isoforms as GSI(A+) and GSI(A−) depending on the cluster they belong to and on whether they contain adenylylation regulated genes.

The unexpected close evolutionary relation between *Methanococcus voltae* and low-GC gram-positive bacteria can be explained by a single lateral gene transfer from *M. voltae* to the ancestor of low-GC gram-positive bacteria and *Thermotoga maritima.* This is in agreement with the data of Tiboni et al. (1993). The closer relationship between Thermotoga and gram-positive bacteria conflicts with other molecular data (Woese 1987) that show Thermotoga ancestral to both gram-positive and gram-negative bacteria. Indeed, morphological data (Cavalier-Smith 1992), in agreement with ours, support a closer relationship between Thermotoga and gram-positive bacteria than with gram-negative bacteria.

Our data which show Eubacteria more closely related to Archaebacteria place the root of the universal tree of life in the eukaryotic line of descent and are in agreement with other evolutionary analyses carried out on DNA topoisomerase II (Wyckoff et al. 1989), citrate synthase (Sutherland et al. 1990), HSP70 (Gupta and Golding 1993), and glutamate dehydrogenase (Benachenhou-Lahfa et al. 1993). However, the two other possibilities are supported by other genes. In particular, EF-Tu and EF-G elongation factors (Iwabe et al. 1989), α- and β-subunits of ATPase (Gogarten et al. 1989; Tsutumi et al. 1991), and DNA-dependent RNA poly-

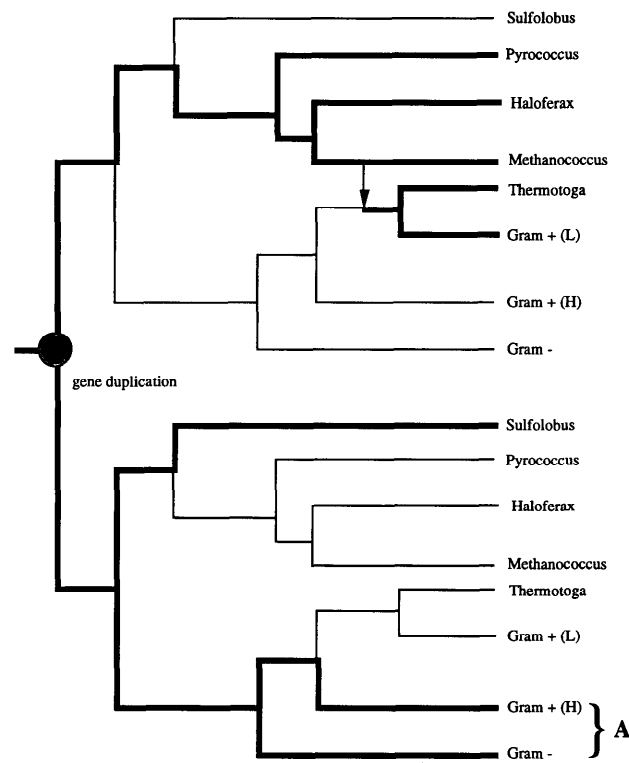

FIG. 7.—Schematic representation of the gene duplication hypothesis put forward for explaining the phylogenetic pattern of prokariotic GSI genes. Thick and thin lines represent extant and extinct lineages. A+ indicates adenylylation regulated lineages; Gram + (L), gram positives with low GC%; gram + (H), gram positives with high GC%; Gram −, gram negatives.

merases (Zillig et al. 1989) show a closer relationship between Eukaryotes and Archaebacteria, whereas glyceraldehyde-3-phosphate dehydrogenase (Hensel et al. 1989) and malate dehydrogenase (Honka et al. 1990) show higher similarity between Eukaryotes and Eubacteria.

This discrepancy can be explained by the hypothesis put forward by Sogin (1991), who proposed that after the separation between the protoeukaryote and the protobacterium, from which Archaebacteria and Eubacteria arose, the protoeukaryote engulfed an archebacterium to form its nucleus. The subsequent endosymbiosis of an eubacterium to form eukaryotic organelles can explain the mosaic structure of eukaryotic genomes having both Eubacterial and Archaebacterial contributions.

We would like to stress that when a phylogenetic tree is constructed from one gene for each species, the inferred phylogeny does not necessarily agree with species phylogeny. This is especially true if multiple gene transfers or gene duplications occur in the evolutionary history of the gene. In this context caution should be paid and as many genes as possible should be used before drawing definite conclusions on the phylogenetic relationships of species. On the other hand, it should be stressed that a clocklike behavior is necessary for a correct reconstruction of the story of the genes and then, if a correct evolutionary model is hypothesized, of the story of the organisms.

The classification we present here is based on the analyses of a pivotal enzyme, GS, as well as on the sequences of the two rRNA species. The results obtained are in a surprisingly good agreement and, under the molecular clock assumption, all support the classification of the living organisms into two superkingdoms.

Acknowledgments

This work was partially financed by Ministero Universitá Ricerca Scientifica Technologica, Italy, and Progetto Finalizzato Ingegneria Genetica, Consiglio Nazionale delle Ricerche, Italy.

LITERATURE CITED

BENACHENHOU-LAHFA, N., P. FORTERRE, and B. LABEDAN. 1993. Evolution of glutamate dehydrogenase genes: evidence for two paralogous protein families and unusual branching patterns of the archaebacteria in the universal tree of life. J. Mol. Evol. **36**:335–346.

BERNARDI, G. 1993. The vertebrate genome: isochores and evolution. Mol. Biol. Evol. **10**:186–204.

BERNARDI, G., B. OLOFSSON, J. FILIPSKI, M. ZERIAL, J. SALINAS, G. CUNY, M. MEUNIER-ROTIVAL, and F. RODIER. 1985. The mosaic genome of warm-blooded vertebrates. Science **228**:953–958.

CAVALIER-SMITH, T. 1992. Origin of secondary metabolites: their function and evolution. Pp. 64–87 *in* Ciba Foundation Symposium 171. Wiley, Chichester.

CHIURAZZI, M., R. MEZA, M. LARA, A. LAHM, R. DEFEZ, M. IACCARINO, and G. ESPÌN. 1992. The Rhizobium leguminosarum biovar phaseoli glnT gene, encoding glutamine synthetase III. Gene **119**:1–8.

DE RIJK, P., J. M. NEEFS, V. VAN DE PEER, and R. DE WATCHER. 1992. Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res. **20**:2075–2089.

FELSENSTEIN, J. 1993. PHYLIP—phylogeny inference package. Department of Genetics, University of Washington, Seattle.

GOGARTEN, J. P., H. KIKAB, P. DITTRICH, L. TAIZ, E. M. BOWMAN, B. J. BOWMAN, M. F. MANOLSON, R. J. POOLE, T. DATE, T. OSHIMA, J. KONISHI, K. DENDA, and M. YOSHIDA. 1989. Evolution of the vacuolar H+-ATPase: implications for the origin of eukaryotes. Proc. Natl. Acad. Sci. USA **86**:6661–6665.

GOODMAN, H. J. K., and D. R. WOODS 1993. Cloning and nucleotide sequence of the Butyrivibrio fibrisolvens gene encoding a type III glutamine synthetase. J. Gen. Micro. **139**:1487–1493.

GUPTA, R. S., and G. B. GOLDING. 1993. Evolution of HSP70 gene and its implications regarding relationships between Archaebacteria, Eubacteria and Eukaryotes. J. Mol. Evol. **37**:573–582.

HENSEL, R., P. ZWICKL, S. FABRY, J. LANG, and P. PALM. 1989. Sequence comparison of glyceraldehyde-3-phosphate dehydrogenases from the three urkingdoms: evolutionary implication. Can. J. Microbiol. **35**:81–85.

HILL, R. T., J. R. PARKER, H. J. K. GOODMAN, D. T. JONES and D. R. WOODS. 1989. Molecular analysis of a novel glutamine synthetase of the anaerobe Bacteroides fragilis. J. Gen. Microb. **135**:3271–3279.

HONKA, E., S. FABRY, T. NIERMANN, P. PALM, and R. HENSEL. 1990. Properties and primary structure of the L-malate dehydrogenase from the extremely thermophilic archaebacterium *Methanothermus fervidus*. Eur. J. Biochem. **188** 623–632.

IWABE, N., K. KUMA, M. HASEGAWA, S. OSAWA, and T. MIYATA. 1989. Evolutionary relationship of archaebacteria eubacteria and eukaryotes inferred from phylogenetic trees of duplicated genes. Proc. Natl. Acad. Sci. USA **86**:9355–9359.

KUMADA, Y., D. R. BENSON, D. HILLEMANN, T. J. HOSTED D. A. ROCHEFORT, C. J. THOMPSON, W. WOHLLEBEN, and Y. TATENO. 1993. Evolution of the glutamine synthetase gene, one of the oldest existing and functioning genes. Proc Natl. Acad. Sci. USA **90**:3009–3013.

LANAVE, C., G. PREPARATA, C. SACCONE, and G. SERIO. 1984. A new method for calculating evolutionary substitution rates. J. Mol. Evol. **20**:86–93.

OLSEN, G. J., R. OVERBEEK, N. LARSEN, T. L. MARSH, M. J MCCAUGHEY, M. A. MACIUKENAS, W. KUAN, T. J MACKE, Y. XING, and C. R. WOESE. 1992. The ribosomal database project. Nucleic Acids Res. **20**:2199–2200.

PESOLE, G., M. ATTIMONELLI, G. PREPARATA, and C. SACCONE. 1992. A statistical method for detecting regions with

different evolutionary dynamics in multialgned sequences. Mol. Phyl. Evol. **1**:91–96.

PESOLE, G., M. P. BOZZETTI, C. LANAVE, G. PREPARATA, and C. SACCONE. 1991. Glutamine synthetase gene evolution: a good molecular clock. Proc. Natl. Acad. Sci. USA **88:** 522–526.

SACCONE, C., C. LANAVE, and G. PESOLE. 1993. Time and biosequences. J. Mol. Evol. **37**:154–159.

SACCONE, C., C. LANAVE, G. PESOLE, and G. PREPARATA. 1990. Influence of base composition on quantitative estimates of gene evolution. Meth. Enzymol. **183**:570–583.

SOGIN, M. L. 1991. Early evolution and the origin of eukaryotes. Curr. Opin. Genet. Dev. **1**:457–463.

SUEOKA, N. 1988. Directional mutation pressure and neutral molecular evolution. Proc. Natl. Acad. Sci. USA **85**:2653–2657.

SUTHERLAND, K. J., C. M. HENNEKE, P. TOWNER, D. W. HOUGH, and M. J. DANSON. 1990. Citrate synthase from the thermophilic archaebacterium *Thermoplasma acidophilum:* Cloning and sequencing of the gene. Eur. J. Biochem. **194**:839–844.

TIBONI, O., P. CAMMARANO, and A. M. SANANGELANTONI. 1993. Cloning and sequencing of the gene encoding glutamine synthetase I from the archaeum *Pyrococcus woesei:* anomalous phylogenies inferred from analysis of archaeal and bacterial glutamine synthetase I sequences. J. Bacteriol. **175**:2961–2969.

TISCHER, E., S. DASSARMA, and H. M. GOODMAN. 1986. Nucleotide sequence of an alfalfa glutamine synthetase gene. Mol. Gen. Genet. **203**:221–229.

TSUTUMI, S., K. DENDA, K. YOKOYAMA, T. OSHIMA, T. DATE, and M. YOSHIDA. 1991. Molecular cloning of genes encoding major two subunits of a eubacterial V-type ATPase from *Thermus thermophilus.* Biochim. Biophys. Acta **1098:** 13–20.

WOESE, C. R. 1987. Bacterial evolution. Microbiol. Rev. **51:** 221–271.

WYCKOFF, E., E. NATALIE, J. M. NOLAN, M. LEE, and T. S. HSIEH. 1989. Structure of the *Drosophila* DNA topoisomerase II gene nucleotide sequence and homology among topoisomerases II. J. Mol. Biol. **205**:1–13.

YAMASHITA, M. M., R. J. ALMASSY, C. A. JANSON, D. CASCIO, and D. EISENBERG. 1989. Refined atomic model of glutamine synthetase at 3.5 Å resolution. J. Biol. Chem. **264:** 17681–17690.

ZILLIG, W., H. P. KLENK, P. PALM, H. LEFFERS, G. PUHLER, G. GROPP, and R. GARRETT. 1989. Did eukaryotes originate by a fusion event? Endocytobiosis Cell. Res. **6**:1–25.

MANOLO GOUY, reviewing editor

Received April 4, 1994

Accepted September 29, 1994