Nucleic Acids Research, 1993, Vol. 21, No. 21 ***4967–4974***

Codon reading scheme in *Mycoplasma pneumoniae* revealed by the analysis of the complete set of tRNA genes

Philippe Simoneau[1,+], Cheng-ming Li[1], Steve Loechel[1], Rainer Wenzel[4], Richard Herrmann[4] and Ping-chuan Hu[1,2,3,*]

Departments of [1]Pediatrics, [2]Microbiology and Immunology and [3]Center of Environmental Medicine, University of North Carolina, Chapel Hill, NC 27599, USA and [4]Department of Microbiology, University of Heidelberg, Im Neuenheimer Feld 282, D-6900, Heidelberg, Germany

Received May 5, 1993; Revised and Accepted September 17, 1993 Accession nos[§]

ABSTRACT

The 33 genes encoding the complete set of tRNA species in *Mycoplasma pneumoniae* have been cloned and sequenced. They are organized into 5 clusters in addition to 9 single genes. No redundant gene was found, indicating that 33 tRNAs correspond to 32 different anticodons and decode all 62 codons used in this organism. There is only one single tRNA for each of the Ala, Leu, Pro, and Val family boxes. Therefore, a simplified decoding system resembling that recently described for *Mycoplasma capricolum* (1) has to also exist in *M.pneumoniae*. However, analysis of the anticodon set and codon usage revealed features characteristic of the latter: (i) there is no obvious preference toward AT rich synonymous codons, (ii) CGG codons are assigned for arginine and are translated by tRNA ^{Arg}(UCG), and (iii) CNN or GNN anticodons are encountered in the Ser, Thr, Arg, and Gly family boxes. We thus propose that this codon-anticodon recognition pattern has emerged in the '*M.pneumoniae* cluster' under a genomic economization strategy but without the influence of AT pressure.

INTRODUCTION

In the highly degenerate universal genetic code, the reading of the third base of codon by the anticodon involves unconventional interactions, referred to as the wobble rules (2). Accumulated data demonstrate that *in vivo* reading of codons almost always follows this scheme. Possible pairing for the nucleotide in the wobble position are: G with C or U; U with G or A; I with U, C, or A; A with U, and C with G. However, exceptions have been found in mitochondria (3–5) where family codon boxes are each read by a single tRNA having an unmodified uridine at the wobble position. Recently, it has been shown that this lack of discrimination in the decoding of family codons is not restricted to eukaryotic organelles but also exists in the Mollicute *Mycoplasma capricolum* (6). Analysis of a complete set of tRNA in this derivative of gram-positive eubacteria revealed only 29 tRNA species with 28 different anticodons and a single isoacceptor tRNA for 6 of the 8 family boxes. The resemblance to mitochondria is enhanced by the fact that mycoplasmas also use both the UGG and UGA codons for tryptophan (7). These similarities in codon recognition patterns might be a sign of similar evolutionary constraints. Mycoplasmas possess the smallest genome of all free-living organisms so far examined (8,9). It is therefore tempting to speculate that, as has been proposed for mitochondria, the limited coding space of the mycoplasma genome has led to a paucity of tRNA genes achieved by the use of a simplified decoding method (4, 5, 10, 11). Moreover, it has been suggested that AT-pressure has concomitantly acted, during the evolution of mycoplasmas and mitochondria from lower eukaryotes, to eliminate non-obligate CNN anticodons and reassign the UGA stop codon (12). Accordingly, it seemed relevant to examine the codon reading scheme of *M. pneumoniae* to verify the existence of a simplified decoding system in this species and thereby aid in our understanding of the mechanisms governing the evolution of the genetic code in Mollicutes. In addition, the genome of this mycoplasma is only 840 kbp (13) vs. 1050 kbp for *M.capricolum* (14). They are also different in GC content, i.e. 40% for *M.pneumoniae* (15) vs. 25% for *M.capricolum* (16). Therefore, it is anticipated that a higher pressure on genome size and lower AT-pressure were likely to be applied during the evolution of *M.pneumoniae*.

We took advantage of the recent cloning of the complete *M.pneumoniae* genome (13) to select, by screening with tRNA probes, all the cosmid clones that carry tRNA genes. The present

* To whom correspondence should be addressed

+ Present address: Faculté des Sciences, Universite d'Angers, 2 Boulevard Lavoisier, 49045 Angers Cedex, France

§ EMBL accession no. X17113; GenBank accession nos M32340 and L15232–L15242 (incl.)

study describes the sequence analysis of 33 tRNA genes that represent the complete set of tRNA species in *M.pneumoniae* and correspond to the 32 different anticodons that translate all the sense codons in this microorganism.

MATERIALS AND METHODS

Preparation of nucleic acids

Crude tRNA preparations were obtained by direct phenol extraction of the cells followed by selective NaCl and isopropanol precipitations as described by Sambrook *et al.*, (17). Absence of cross-contamination with genomic DNA or high molecular weight RNAs was demonstrated by polyacrylamide gel electrophoresis. RNAs were deacylated, end-labeled with (5'-^{32}P) pCp using RNA ligase (New England Biolabs, Inc.) essentially as described by England et al (18), and used as hybridization probes. *M.pneumoniae* strain M129-B16 (ATCC 29342), was grown and harvested as previously described (19). DNA was extracted, digested with restriction enzymes, analyzed on agarose gels and blotted onto nitrocellulose by standard methods (20).

Cloning and sequencing procedures

Restriction fragments from cosmid clones, propagated in *E.coli* strain HB101, were subcloned by excision from low-melting agarose gels (Sea-plaque, FMC Corporation), ligated to restricted phages M13mp18/19 (21) or plasmid pIBI 24/25 (IBI) vectors and transformed into *E.coli* strain DH5αF'IQ (BRL) made competent by standard procedure (22). Hybridizations with labeled tRNAs were routinely performed in 50% formamide at 37°C with subsequent washings in 0.1 SSC −0.1% SDS at 55°C. Less stringent conditions, i.e. lower hybridization and washing temperatures, have also been used to ensure that all the hybridizing restriction fragments were detected. Nested deletions of recombinant clones were produced by the exonuclease III method of Henikoff (23) using the Erase-a-base system (Promega Biotec.) and their nucleotide sequence determined on both strands by the dideoxy chain-termination method (24) using the Taq track kit (Promega Biotec.). The data were compiled and analyzed with the Beckman Microgenie program.

RESULTS

Screening the *M.pneumoniae* genome for tRNA genes

To detect all the restriction fragments that carry genes for tRNA, the 34 cosmids plus one plasmid and 2 lambda phage clones that span the complete genome of *M.pneumoniae* (13) were hybridized with labeled small stable RNA preparations (Fig. 1A). Based on the size of the hybridization signals together with the previously published *Eco*RI linkage map of the *M.pneumoniae* genome (25), thirteen *Eco*RI fragments, arbitrarily named E1 through E13, were selected as a potential source of tRNA genes: 28 kbp from pCosMP-A19/pCosMp-F11, 22.3 kbp from pCosMP-E11/pCosMP-H91, 19 kbp from pCosMP-A19, 15 kbp from pCosMP-A1/pCosMP-C9, 12.3 kbp from pCosMP-H10, 9.5 kbp from pCosMP-R1/pCosMP-R2, 9.2 kbp from pCosMP-H8/pCosMP-D7, 8 kbp from pCosMP-H3, 7.5 kbp from pCosMP-R2/pCosMP-A68, 7 kbp from pCosMP-G12, 6.8 kbp from pCosMP-A3/pCosMP-D12, 5.2 kbp from pCosMP-H8 and 2.7 kbp from pCosMP-H10. Two other *Eco*RI fragments produce a faint signal with the probe; a 6.6 kbp fragment from pCosMP-K5 and a fragment larger than 6.8 kbp from pCosMP-D12.

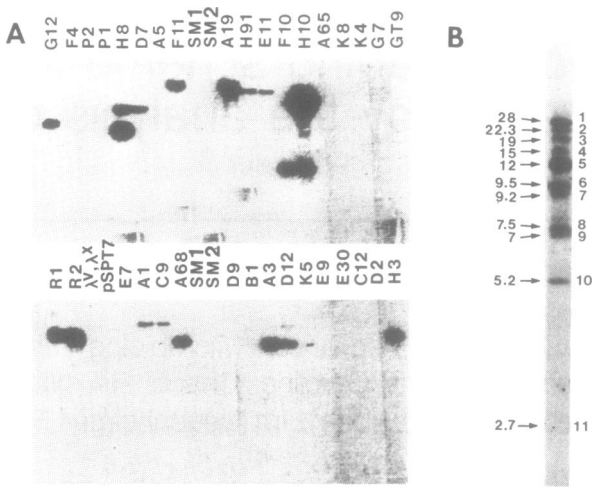

Figure 1. Detection of *Eco*RI fragments from *M.pneumoniae* genomic DNA carrying tRNA genes. Southern blot analysis of either DNA from the set of clones from an ordered library that cover the complete genome **(A)** or total genomic DNA **(B)** restricted with *Eco*RI, electrophoresed on agarose, transferred to nitrocellulose and hybridized with labeled tRNA probes. Letters in panel A refer to the clone names (reference 13); SM: size markers (SM1 = pUC18 polymers; SM2 = *Bst*EII digests from lambda phage;). The correspondence of *Eco*RI hybridizing bands in panel B and their attributed numbers and tRNA genes or clusters is explained in the text.

However, nucleotide sequence analysis of the former did not reveal any tRNA gene or related sequence and the latter did not correspond to any *Eco*RI fragment on the stained gel (data not shown), indicating it was likely to be the result of a partial hydrolysis of the clone with the restriction enzyme. Detection of additional putative hybridization fragments on genomic blots of *Eco*RI digested *M.pneumoniae* DNA revealed 11 discrete bands (Fig. 1B), accounting for all but one of the above mentioned cosmid fragments when assuming that the 6.8/7 kbp fragment pair co-migrated as a single band. The 8 kbp positive signal found in *Eco*RI digests from H3 was never observed on blots prepared from *M.pneumoniae* genomic DNA. However, tRNA gene screening of a plasmid *Eco*RI sublibrary constructed using genomic fragments larger than 12 kbp resulted in the selection of recombinant clones with a 12.4 kbp insert different from E5. This insert proved to be the sum of 2 *Eco*RI subfragments of 8 and 4.4 kbp. Restriction maps of the 8 kbp subfragment and the 8 kbp E8 fragment from H3 were found to be identical. Since similar clones were repeatedly isolated we assumed that methylation of an *Eco*RI site present in the mycoplasma genome prevented enzymatic digestion at this locus so that E8 in genomic blots appeared as a 12.4 kbp band which co-migrated with E5. The *Eco*RI map from H3 demonstrated a close linkage between E8 and a 4.4 kbp fragment thereby supporting our assumption. We therefore concluded that in *M.pneumoniae*, the entire set of tRNA genes mapped into a maximum of 13 different *Eco*RI fragments.

Sequence analysis and genomic organization of the tRNA

Fragments E1 to E11 were digested with restriction endonucleases and hybridized with labeled tRNAs; each of the positive fragments was subcloned into M13 phage vectors for sequencing. *Hind*III digestion of the E1 fragment from cosmid F11 generated three positive signals of 6.5, 4, and 2.2 kbp. Their nucleotide

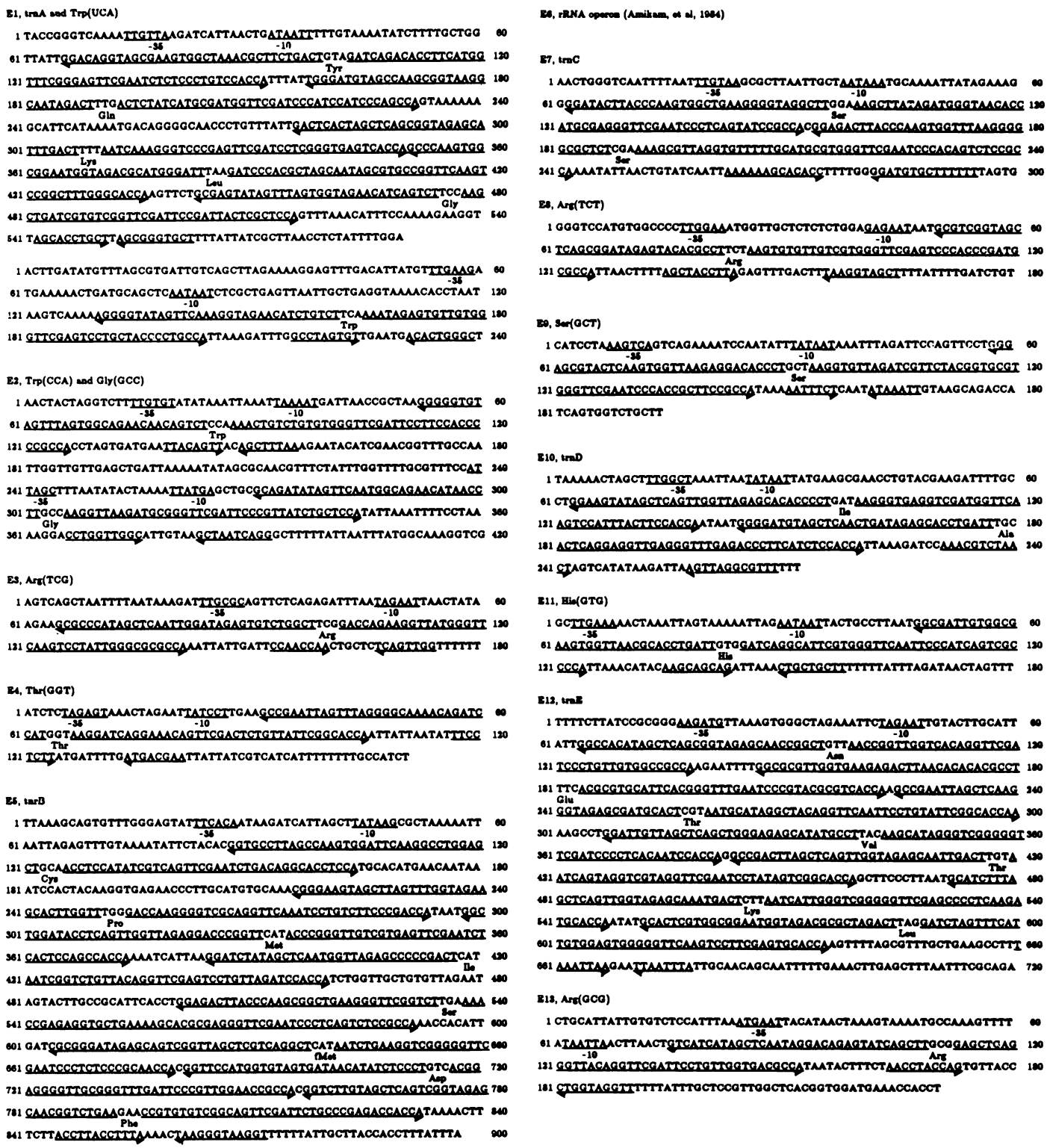

Figure 2. Nucleotide sequences of the 33 tRNA genes identified in the 13 *Eco*RI fragments. Individual tRNA genes are underlined and defined by opposite half-arrowheads. Putative promoter, −10 and −35 sequences are underlined and indicated. Potential terminator-like dyed-symmetrical sequences are indicated by arrows. Nucleotide sequences of cluster trnA and tRNA^{Trp}(UCA) were previously published (26, 28) and their EMBL/GenBank/DDBJ accession numbers are X17113 and M32340, respectively. Other nucleotide sequence data will appear in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases under the accession number from L15232 to L15242.

sequences have been described previously. They contain a cluster (refered to as trnA) of five tRNA genes including Tyr(GTA)-Gln(TTG)-Lys(TTT)-Leu(TAA)-Gly(TCC) (26), a gene for a 4.5S RNA homolog (27), and a gene for tRNA^{TRP}(UCA) (28). A remarkable feature of the cluster trnA is the absence of any spacer nucleotide between tRNA^{Lys}(TTT) and tRNA^{Leu}(TAA).

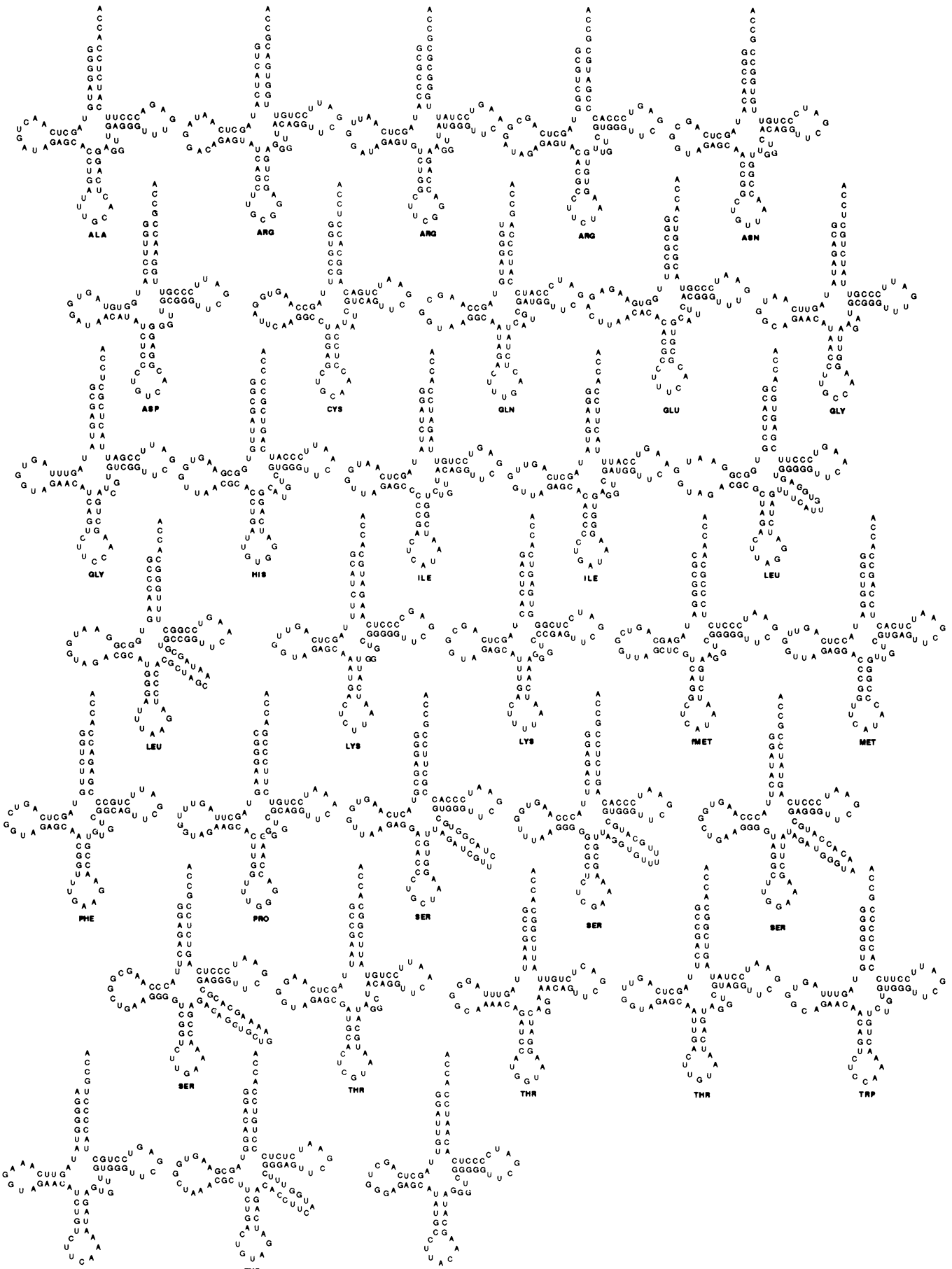

Figure 3. Clover-leaf structure of the 33 tRNA species from *M.pneumoniae* as deduced from their nucleotide sequences.

Table 1. Codon Usage of *M.pneumoniae* and *M.capricolum*

	M.p.	M.c.		M.p.	M.c.		M.p.	M.c.		M.p.	M.c.
Phe(UUU)	86	269	Ser(UCU)	20	107	Tyr(UAU)	33	171	Cys(UCU)	7	41
Phe(UUC)	39	28	Ser(UCC)	67	1	Tyr(UAC)	52	24	Cys(UGC)	3	10
Leu(UUA)	68	433	Ser(UCA)	26	169	Stop(UAA)	–	–	Trp(UGA)	29	42
Leu(UUG)	68	22	Ser(UCG)	44	5	Stop(UAG)	–	–	Trp(UGG)	27	7
Leu(CUU)	30	36	Pro(CCU)	27	66	His(CAU)	14	72	Arg(CGU)	33	54
Leu(CUC)	79	0	Pro(CCC)	71	3	His(CAC)	30	22	Arg(CGC)	48	4
Leu(CUA)	20	62	Pro(CCA)	48	120	Gln(CAA)	106	256	Arg(CGA)	9	2
Leu(CUG)	34	1	Pro(CCG)	40	3	Gln(CAG)	70	10	Arg(CGG)	23	0
Ile(AUU)	74	480	Thr(ACU)	36	206	Asn(AAU)	82	440	Ser(AGU)	87	117
Ile(AUC)	39	56	Thr(ACC)	134	8	Asn(AAC)	132	78	Ser(AGC)	39	20
Ile(AUA)	8	111	Thr(ACA)	33	151	Lys(AAA)	84	730	Arg(AGA)	9	189
Met(AUG)	29	153	Thr(ACG)	59	1	Lys(AAG)	117	71	Arg(AGG)	14	3
Val(GUU)	52	284	Ala(GGU)	89	228	Asp(GAU)	107	294	Gly(GGU)	92	180
Val(GUC)	54	8	Ala(GCC)	91	3	Asp(GAC)	85	24	Gly(GGC)	65	5
Val(GUA)	31	157	Ala(GCA)	31	140	Glu(GAA)	78	373	Gly(GGA)	28	199
Val(GUG)	90	16	Ala(GCG)	72	4	Glu(GAG)	40	29	Gly(GGG)	78	16

The 3331 codons of *M.pneumoniae* are from the P1 operon (41) and the *deo*C gene (42). The 6814 codons of *M.capricolum* are from published data (6).

Although the transcriptional and processing analysis of the *M.pneumoniae* tRNA gene sets has not been done, the fact that some of the tRNA structural genes found in this cluster are separated from one another by one or even zero nucleotides leads us to assume that these genes themselves may act in part as processing signals for enzymes. Four other tRNA gene clusters, referred to as trnB to E, were also found in the *M.pneumoniae* genome. Their nucleotide sequences are presented in Figure 2. Cluster trnB, located on E5, contains eight tRNAs: Cys(GCA)-Pro(TGG)-Met(CAT)-Ile(CAT)-Ser(TGA)-fMet(CAT)-Asp(TGC)-Phe(GAA). Cluster trnE, harbored within fragment E12, consists of seven tRNA genes: Asn(GTT)-Glu(TTC)-Thr(CGT)-Val(TAC)- Thr(TGT)-Lys(CTT)-Leu(TAG). Both clusters trnC (from E7) and trnD (from E10) carry tandem genes for Ser(GGA)-Ser(CGA) and Ile(GAT)-Ala(TGC), respectively. In every case, putative promoter sequences similar to those of *E.coli* (29) and *B.subtilis* (30), as well as terminator hairpin loops have been found, although the identity of the promoter sequences has not been experimentally established in mycoplasma genes. No such characteristic features could be detected in the spacer sequences of these clusters, therefore indicating that they are probably single transcription units. Each of the remaining hybridizing *Eco*RI fragments, except E6 which carries the single rRNA operon from *M.pneumoniae* (31), contains one or two tRNA genes: Trp(CCA) and Gly(GCC) on E2, Arg(TCG) on E3, Thr(GGT) on E4, Arg(TCT) on E8, Ser(GCT) on E9, His(GTG) on E11 and Arg(GCG) on E13. Together, these genes can potentially be transcribed into 33 tRNA species corresponding to 32 different anticodons; 24 of them are grouped into 5 clusters. Concentration of tRNA genes to a few regions in the genome has been previously described for *Bacillus subtilis* (32). An interesting observation is that substantial similarities in the tRNA gene organization exist between *M.pneumoniae* and other mycoplasma species as well as *B.subtilis*. For example, the order of the tRNA genes in cluster trnA is identical to that of a part of a '5 tRNA gene cluster' from *M.capricolum* (6). Similarly cluster trnD corresponds exactly to the two spacer tRNA genes positioned in between the two rRNA gene sets *rrn*A and *rrn*O from *B.subtilis* (33), while the '4 tRNA gene clusters' from *M.mycoides* (34) and *M.capricolum* (35) constitute a portion of cluster trnE. Cluster trnE also ends by the same 4 tRNA genes, Val(TAC)-Thr(TGT)-Lys(CTT)-Leu(TAG), that start the ribosomal gene set *rrn*B sequence from *B.subtilis* (36). Three tRNA genes in cluster trnB have the same CAT sequence in the region which defines the anticodon. By analogy with the 21 tRNA gene clusters of *B.subtilis* (37), we concluded that the CAT anticodons code for methionine, isoleucine, and the initiating formyl methionine. Further support for such amino acid assignment came from the comparison of the cloverleaf structure of the entire set of tRNA (Figure 3) and those proposed for *B.subtilis*. For example, the A-U base pair at the third position in the aminoacyl stem of *M.pneumoniae* $tRNA^{Ile}$(NAT), where N is probably the lysidine form from C (38), is also present in the *B.subtilis* homologous tRNA and may prevent its action as a substrate for methionyl tRNA synthase (39). Secondly, as in other eubacteria, the $tRNA^{fMet}$(CAT) identified in *M.pneumoniae* also possesses a C1–A72 mispair in the acceptor stem (40). Nucleotide mispairs were also found in the following *M.pneumoniae* tRNAs: $tRNA^{Ile}$(CAT) in the anticodon stem (C27–U43); $tRNA^{Trp}$(CCA) in the T C stem (U50–U64). In almost all the other tRNAs the consensus sequences depicted for eubacterial tRNAs are conserved as it has been previously demonstrated in *M.capricolum* (6). The unusual feature of $tRNA^{Thr}$(GGU), an A at position 32 instead of the classical pyrimidine, and of $tRNA^{His}$(GUC), an extra G at the 5′ end of the gene, have been described for *B.subtilis* (32). However, a striking difference between tRNA genes from *B.subtilis* and those from *M.pneumoniae* is that in the latter all the tRNA genes possess a 3′CCA terminus. The distribution of tRNA genes throughout the *M.pneumoniae* chromosome is represented in Figure 4 and constitutes another point of divergence between the two organisms. In contrast to *B.subtilis* (41), the transcription units for tRNAs from *M.pneumoniae* are scattered on the genome and none contain rRNA genes.

DISCUSSION

Number of anticodons and decoding system used in *M.pneumoniae*

We have described the cloning and sequencing of 33 different tRNA genes of *M.pneumoniae* in the present study. No other tRNA gene was found upon hybridization of cloned genomic DNA from this organism with labeled small stable RNAs. The possibility that additional tRNA genes may have escaped detection

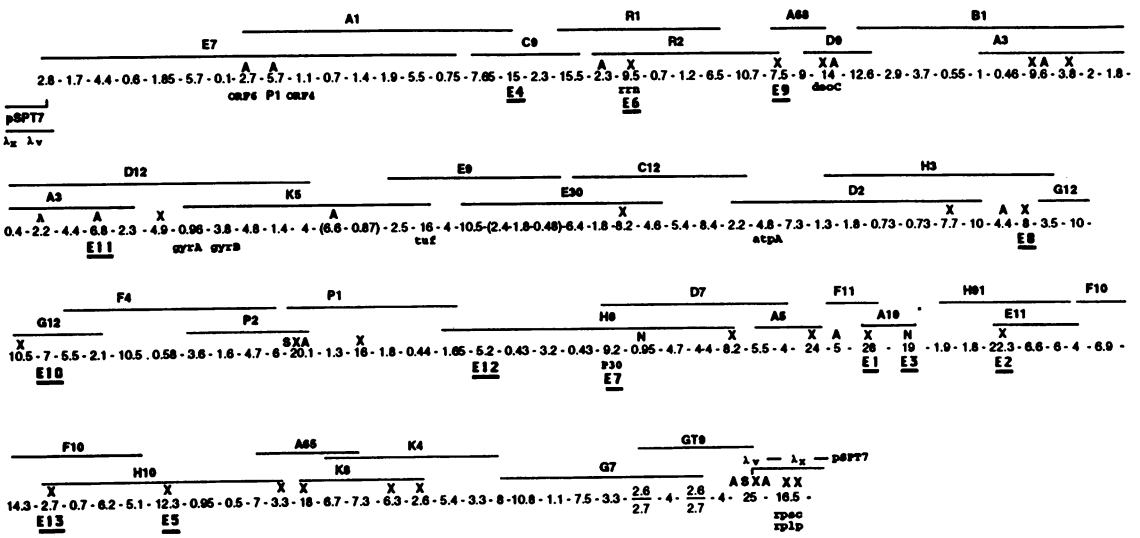

Figure 4. Location of the tRNA genes in the *M.pneumoniae* chromosome. Symbols E1 to E13, that indicate the loci for tRNA clusters are described in the text. The map was constructed by complementing linearized *Eco*RI map of the *M.pneumoniae* genome (25). The numbers indicate the size of individual fragments in kbp. Restriction sites for endonucleases *Xho*I (X), *Apa* I (A), *Not*I (N) and *Sfi*I (S) are indicated.

	A	B	C		A	B	C		A	B	C		A	B	C
UUU F	G	G	G	UCU S	G	U	U	UAU Y	G	G	G	UGU C	G	G	G
UUC F				UCC S				UAC Y				UGC C			
UUA L	U	U	U	UCA S	U			UAA End				UGA W	U	U	U
UUG L			C	UCG S	C			UAG End				UGG W	C		C
CUU L	U	U	U	CCU P	U	U	U	CAU H	G	G	G	CGU R	G	U	I
CUC L				CCC P				CAC H				CGC R			
CUA L				CCA P				CAA Q	U	U	U	CGA R	U		
CUG L				CCG P				CAG Q				CGG R			na
AUU I	G	G	G	ACU T	G	U	A	AAU N	G	G	G	AGU S	G	G	G
AUC I				ACC T				AAC N				AGC S			
AUA I	C		C	ACA T	U		U	AAA K	U	U	U	AGA R	U	U	U
AUG M	C	C	C	ACG T	C			AAG K	C		C	AGG R			
GUU V	U	U	U	GCU A	U	U	U	GAU D	G	G	G	GGU G	G	U	U
GUC V				GCC A				GAC D				GGC G			
GUA V				GCA A				GAA E	U	U	U	GGA G	U		
GUG V				GCG A				GAG E				GGG G			

Figure 5. Comparison of tRNA genes in *M.pneumoniae* (**A**), human mitochondria (**B**) and *M.capricolum* (**C**). The letters in boxes denote the wobble nucleotide in each tRNA anticodon. Data for mitochondria and *M.capricolum* are from references 3 and 6 respectively. Bold letters refer to amino acids according to the simplified international rules.

cannot be completely ruled out. However, this is not a likely possibility because the method employed, hybridization of cosmid cloned DNA, is expected to be much more sensitive than regular genome blot probing. Hence, it appears that the 33 tRNA species represent the complete set of *M.pneumoniae* tRNAs. Although this number of anticodons is higher than the 32 required by wobble hypothesis (2), careful examination of Figure 5 shows that their distribution in the different boxes of the genetic code is not consistent with the rules established for this decoding system. For instance, in family boxes for Leu, Val, Pro, and Ala only a single isoacceptor tRNA (UNN) is present. This feature resembles that found in mitochondria where tRNAs for unmixed families are unique and possess an unsubstituted U in the first position of the anticodon (5). These tRNAs are likely to translate all four family codons either by a two-out-of-three method (42) or by four-way wobbling (5). Our data suggest that such a simplified decoding system could also exist in *M.pneumoniae*. This would allow the 33 anticodons to translate all the 62 codons used in *M.pneumoniae* (43, 44) and support, in part, the expanded codon-anticodon recognition rules proposed for *M.capricolum* (33). In that bacterium, only 29 tRNA species were found, with a single anticodon in all but one of the family boxes (Thr). Similarly, at least one of the family codons ACN for threonine is read by more than one anticodon in *M.pneumoniae*. However, in this case, it is the ACG rather than the ACU codon (1). The difference in the number of anticodons between the two *Mycoplasma* species is accounted for by redundant *M.pneumoniae* tRNAs in the Ser box, as well as in the Gly and Arg boxes if tRNA (UCC) and tRNA (UCG) are unmodified in the wobble position. Two other non-obligate tRNAs, $tRNA^{Lys}$(CUU) and $tRNA^{Trp}$(CCA), were also present in both organisms. The discovery of the latter contradicts previous assumptions that *M.pneumoniae* decodes both UGG and UGA codons for tryptophan using a single UCA anticodon (28). The failure to detect the $tRNA^{Trp}$(CCA) in our earlier study might be due to the lower sensitivity of the screening method.

It is notable that, in sharp contrast with eubacterial code where anticodons ICG and CCG are used to translate Arg (CGN) codons, *M.pneumoniae* seems to read the Arg family box using two tRNAs with anticodons UCG and GCG. To our knowledge, no other $tRNA^{Arg}$(GCG) or (UCG) has been described to date in Eubacteria while they are present in some Archaebacteria (43).

Evolutionary significance

Mycoplasmas, whose genome sizes range from 600 to 1800 kbp (8, 46, 47), are thought to have evolved from Gram-positive bacteria (48) probably by both deletions of non-essential genes and reduction of the size of essential genes (27). The comparison of the tRNA gene organization and redundancy within

mycoplasmas and *B.subtilis*, suggests that tRNA genes in mycoplasmas originated from larger ancestral clusters by deletions under genome economization constraints. Pressure on genome size was probably one of the essential evolutionary constraints that also led to the appearance of a simplified decoding system in eukaryotic organelles. Along the same lines, in mycoplasmas, one would expect to find a similar codon-anticodon recognition pattern. This was strongly suggested by the lack of discrimination in the reading of certain codons found in *M.mycoides* (34) and recently demonstrated by the analysis of the complete set of tRNA species from *M.capricolum* (6). We have extended these results to *M.pneumoniae*. This unconventional decoding method may therefore be ubiquitous among mycoplasmas with the probable exception of the *Acholeplasma* species. In *A.laidlawii*, partial analysis of tRNA genes revealed that the anticodon composition in family boxes is similar to that of *B.subtilis* rather than that of *M.capricolum* (49). These findings are consistent with the evolutionary position of this bacterium in the mycoplasmal lineage (48) and its relatively large genome. Accordingly, if genomic economization constitutes the most significant factor that led to the emergence of the simplified decoding system, we should expect to find a further reduction of redundant anticodons in *M.pneumoniae* compared to *M.capricolum* due to its smaller genome. Surprisingly, the situation is reversed: 33 tRNA gene species in the former vs 29 tRNA species corresponding to 30 genes in *M.capricolum* (35). It seems therefore relevant to look for additional forces that might influence the disappearance or conservation of some of the non-obligate anticodons. Based on the extremely low G+C content (25%) of *M.capricolum*, a codon capture theory has been proposed by Osawa and Jukes (50) to account for the reduction in the number of tRNA gene species as well as the preferential use of AT-rich codons by this species. This theory stipulates that under strong AT-pressure: (i) the universal opal codon UGA has been reassigned to tryptophan (12), (ii) most of NNC and NNG codons have been converted to the synonymous NNU and NNA codons (51), and (iii) all the CGG codons have been converted to other arginine codons by silent mutations leading to the disappearance of the unnecessary CCG anticodons (52). The reading of UGA as a sense codon for tryptophan has been previously demonstrated in *M.pneumoniae* (28) and suggests that both mycoplasmas may have followed a similar evolutionary pathway. However, the analysis of codon usage in *M.pneumoniae* (Table 1) seems to contradict the above mentioned hypothesis because there is no apparent preferential use of AT rich synonymous codons in its coding sequences. The conservation of CNN and GNN anticodons in family boxes is most likely due to a high demand of GC-rich codons. From these observations, it is likely that the evolution model proposed for *M.capricolum* may not be fully applicable to all mycoplasmas. Recent phylogenetic analysis, based upon small-subunit rRNA sequences, indicated that mycoplasmas can be divided into five groups, namely, the pneumoniae group, the hominis group, the spiroplasma group, the anaeroplasma group, and the asteroleplasma group (53). *M.pneumoniae* was assigned to the pneumoniae group, while both *M.capricolum* and *M.mycoides* were assigned to the spiroplasma group. Some of the results presented here support this classification: (i) codons ACN for threonine are read by the anticodons UGU, CGU and GGU in *M.pneumoniae* while AGU and UGU are used in both *M.capricolum* (1) and *M.mycoides* (10); (ii) anticodon GCC for glycine is present in *M.pneumoniae* while UCC alone was found sufficient in *M.capricolum* (6) and *M.mycoides* (34) to decode all four GGN codons; (iii) genes for $tRNA^{Trp}$(CCA) and $tRNA^{Trp}$(UCA) are present in different transcription units in *M.pneumoniae* and in the closely related *M.gallisepticum* (28) while they are tandemly organized in *M.capricolum* (54); (iv) the physical map for *M.capricolum* revealed that 18 of the 30 tRNA genes are grouped within a 168 kbp segment covering only 16% of the genome (14) while those from *M.pneumoniae* are dispersed. A possible scenario for the evolution of the decoding system and assignment of the termination signal UGA in mycoplasmas of the 'pneumoniae group' would involve a mechanism driven essentially by a 'genome minimization strategy' (55) in the absence of any appreciable mutational bias influences. This could explain the apparent disappearance of the non-obligate $tRNA^{Trp}$(CCA) in closely related mycoplasmas with an even smaller genome such as *Mycoplasma genitalium* (28), despite its lower GC content. Since Acholeplasmas, which are found within two distinct clusters (53), are presumed to recognize UGA as a conventional stop codon (49), it would seem to suggest that the $tRNA^{Trp}$(UCA) has independently emerged in mycoplasmas after their separation into 5 phylogenetic branches. Therefore it is conceivable that 2 different mechanisms, the codon capture under AT pressure in the 'capricolum lineage' and the genome minimization strategy without AT-biased directional mutation pressure in the 'pneumoniae lineage,' have both led to a similar UGA reassignment but with different codon-anticodon recognition patterns. However, we recognize that such an event may occur very rarely during evolution. An alternate scenario is one in which the emergence of $tRNA^{Trp}$(UCA) occurred in a common ancestor due to strong AT-pressure before the divergence of these groups. This is based on the observation that most mycoplasma genomes are AT-rich; it is possible that the common ancestor of mycoplasmas may have also been AT-rich. If this assumption holds true, then the exceptionally high GC-content in *M.pneumoniae* may represent a recent event in the evolution of this species.

ACKNOWLEDGEMENTS

This work was supported in part by National Institutes of Health grant HL19171 and US Environmental Protection Agency Cooperative Agreement CR807392. The authors would like to thank S.-C.Hu for photographic assistance and D.Tulloch for secretarial assistance.

REFERENCES

1. Andachi, Y., Yamao, F., Iwami, M., Muto, A., and Osawa, S. (1987) Proc. Natl. Acad. Sci. USA **84**,7398–7402.
2. Crick, F.H.C. (1966) J. Mol. Biol. **19**,548–555.
3. Barrel, B.G., Anderson, S., Bankier, A.T., de Bruijn, M.H.L., Chen, E., Coulson, A.R., Drouin, J., Eperon, I.C., Nierlich, D.P., Roe, B.A., Sanger, F., Schreier, P.H., Smith, A.J.H., Staden, R., and Young, I.G. (1980) Proc. Natl. Acad. Sci. USA **77**,3164–3166.
4. Bonitz, S.G., Berlani, R., Coruzzi, G., Li, M., Macino, G., Nobrega, F.G., Nobrega, M.P. Thalenfeld, B.E., and Tzagoloff, A. (1980) Proc. Natl. Acad. Sci. USA **77**,3167–3170.
5. Heckman, J.E., Sarnoff, J., Alzer Deweerd, B., Yin, S., and RajBhandary, U.L. (1980) Proc. Natl. Acad. Sci. USA **77**,3159–3163.
6. Andachi, Y., Yamao, F., Muto, A., and Osawa, S. (1989) J. Mol. Biol. **209**,37–54.
7. Yamao, F., Muto, A., Kawauchi, Y., Iwami, M., Iwagami, S., Azumi, Y., and Osawa, S. (1985) Proc. Natl. Acad. Sci. USA **82**,2306–2309.
8. Colman, D.D., Hu, P.C., Litaker, W., and Bott, K.F. (1990). Mol. Microbiol. **4**,683–687.
9. Razin, S. (1985) Microbiol. Rev. **49**,419–455.

10. Samuelsson, T., Guindy, Y.S., Lustig, F., Boren, T., and Lagerkvist, U. (1987) Proc. Natl. Acad. Sci. USA **84**,3166–3170.
11. Hu, P.C., Inamine, J.M., Loechel, S., Huang, C.H., Collier, A.M., and Clyde, W.A., Jr. (1990) Zent. Bakteriol. Suppl. **20**,183–187.
12. Osawa, S. and Jukes, T.H. (1988) Trends Genet. **4**,191–198.
13. Wenzel, R. and Herrmann, R. (1989) Nucl. Acids Res. **17**,7029–7043.
14. Whitley, J.C., Muto, A., and Finch, L.R. (1991) Nucl. Acids Res. **19**,399–400.
15. Razin, S., and Freundt, E.A. (1984) In: N.R. Krieg and J.G. Holdt (ed.). Bergey's Manual of Systemic Bacteriology. Vol.1, The Williams and Wilkins Co. Baltimore, pp.740–793.
16. Jones, A.S. and Walker, R.T. (1963) Nature (London) **198**,588–589.
17. Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. 2nd edition. Cold Spring Harbor Laboratory Press.
18. England, T.E., Bruce, A.G., and Uhlenbeck, O.C. (1980). Methods Enzymol. **65**,65–74.
19. Hu, P.C., Collier, A.M., and Baseman, J.B. (1977) J. Exp. Med. **145**,1328–1343.
20. Ausubel, F.M., Brent, R., Kingston, R.E., Moore, D.D., Seidman, J.G., Smith, J.A., and Struhl, K. (eds.). (1987) Current protocols in molecular biology, Wiley, N.Y.
21. Norrander, J., Kempe, T., and Messing, J. (1983) Gene **26**,101–106.
22. Hanahan, D. (1985) Techniques for transformation of *E.coli* In Glover, D.M. (ed.) DNA cloning, Voll. IRL Press, Oxford, pp109–135.
23. Henikoff, S. (1984) Gene **28**,351–359.
24. Sanger, F., Nicklen, S., and Coulson, A.K. (1977) Proc. Natl. Acad. Sci. USA **74**,5463–5467.
25. Wenzel, R., Pirkl, E., and Herrmann, R. (1992) J. Bacteriol. **174**,7289–7296.
26. Simoneau, P., Wenzel, R., Herrmann, R., and Hu, P.C. (1990) Nucl. Acids Res. **18**,2814.
27. Simoneau, P. and Hu, P.C. (1992) J. Bacteriol. **174**,627–629.
28. Inamine, J.M., Ho, K.C., Loechel, S., and Hu, P.C. (1990) J. Bacteriol. **172**,504–506.
29. Mawley, D.K., and McClure, W.R. (1983) Nucl. Acids Res. **11**, 2237–2255.
30. Moran, C.P., Lang, N., LeGrice, S.F.J., Lee, G., Stephens, M., Sonenstein, A.L., Pero, J., and Losick, R. (1982) Mol. Gen. Genet. **186**, 339–346.
31. Amikam, D., Glaser, G., and Razin, S. (1984) J. Bacteriol. **158**,376–378.
32. Vold, B.S. (1985) Microbiol. Rev. **49**,71–80.
33. Oqasawara, N., Moriya, S., and Yoshikawa, H. (1983) Nucl. Acids Res. **11**,6301–6318.
34. Samuelsson, T., Axberg, T., Boren, T., and Lagerkvist, U. (1983) J. Biol. Chem. **258**,13178–13184.
35. Muto, A., Andachi, Y., Yuzawa, H., Yamao, F., and Osawa, S. (1990) Nucl. Acids Res. **18**,5037–5043.
36. Wawrousek, E.F., and Hansen, J.N. (1983) J. Biol. Chem. **258**,291–298.
37. Green, C.L. and Vold, B.S. (1983) Nucl. Acids Res. **11**,5763–5774.
38. Muramatsu, T., Yokoyama, S., Horic, N., Matsuda, A., Ueda, T., Yamaizumi, Z., Kuchino, Y., Nishimura, S., and Miyazawa, T. (1988) J. Biol. Chem. **263**,9261–9267.
39. Schulman, L. (1979) In: Schimmel, P.R., Soll, D. and Abelson, J.N. (eds.) Transfer RNA: structure, properties, and recognition. Cold Spring Harbor Laboratories. Cold Spring Harbor, N.Y., pp311–324.
40. Scong, B.L., and RajBhandary, U.L. (1987) Proc. Natl. Acad. Sci. USA **84**,8859–8863.
41. Ozeki, H. and Inokuchi, H. (1986) Adv. Biophys. **21**,35–47.
42. Lagerkvist, U. (1981) Cell **23**,305–306.
43. Inamine, J.M., Loechel, S., and Hu, P.C. (1988) Gene **73**,175–163.
44. Loechel, S., Inamine, J.M., and Hu, P.C. (1989) Nucl. Acids. Res. **17**,801.
45. Gupta, R. (1984) J. Biol. Chem. **259**,9461.
46. Bak, A.L., Black, F.T., Christiansen, C., and Freundt, E.A. (1969) Nature **224**,1209–1210.
47. Ye, F., Laigret, F., Whitley, J.C., Citti, C., Finch, L.R., Carle, P., Renandin, J., and Bove, J.M. (1992) Nucl. Acids Res. **20**,1559–1565.
48. Rogers, M.J., Simmons, J., Walker, R.T., Weisburg, W.G., Woese, C.R., Tanner, R.S., Robinson, I.M., Stahl, D.A., Olsen, G., Leach, R.H., and Maniloff, J. (1985) Proc. Natl. Acad. Sci. USA **82**,1160–1164.
49. Tanaka, K., Andachi, Y., Muto, A., and Osawa, S. (1990) IOM Letters. **1**,268–269.
50. Osawa, S. and Jukes, T.H. (1989) J. Mol. Evol. **28**,271–278.
51. Ohkubo, S. Muto, A., Kawauchi, Y., Yamao, F., and Osawa, S. (1987) Mol. Gen. Genet. **210**,314–322.
52. Oba, T., Andachi, Y., Muto, A., and Osawa, S. (1991) Proc. Natl. Acad. Sci. USA **88**,921–925.
53. Weisburg, W.G., Tully, J.G., Rose, D.L., Petzel, J.P., Ovaizu, H., Yang, D., Mandelco, L., Sechrest, J., Lawrence, T.G., Van Etten, J., Maniloff, J. and Woese, C.R. (1989) J. Bacteriol. **171**,6455–6467.
54. Yamao, F., Iwagami, S., Azumi, Y., Muto, A., Osawa, S., Fujita, N., and Ishihama, A. (1988). Mol. Gen. Genet **212**,364–369.
55. Andersson, S.G.E., and Kurland, C.G. (1990) Microbiol. Rev. **54**,198–210.