Vol.2, No.6, 535-543 (2010)
http://dx.doi.org/10.4236/ns.2010.26067

Natural Science

Phylogeny of γ-proteobacteria inferred from comparisons of 3' end 16S rRNA gene and 5' end 16S-23S ITS nucleotide sequences

Sabarimatou Yakoubou[1,2], Jean-Charles Côté[1*]

[1]Agriculture and Agri-Food Canada, Research Centre, Gouin Blvd, St-Jean-sur-Richelieu, Québec, Canada; *Corresponding Author: Jean-Charles.Cote@agr.gc.ca
[2]Département des Sciences Biologiques, Université du Québec à Montréal, CP 8888, Succ. "Centre-Ville" Montréal, Québec, Canada

Received 27 January 2010; revised 25 March 2010; accepted 12 April 2010.

ABSTRACT

The phylogeny of γ-proteobacteria was inferred from nucleotide sequence comparisons of a short 232 nucleotide sequence marker. A total of 64 γ-proteobacterial strains from 13 Orders, 22 families, 40 genera and 59 species were analyzed. The short 232 nucleotide sequence marker used here was a combination of a 157 nucleotide sequence at the 3' end of the 16S rRNA gene and a 75 nucleotide sequence at the 5' end of the 16S-23S Internal Transcribed Spacer (ITS) sequence. Comparative analyses of the 3' end of the 16S rRNA gene nucleotide sequence showed that the last 157 bp were conserved among strains from same species and less conserved in more distantly related species. This 157 bp sequence was selected as the first part in the construction of our nucleotide sequence marker. A bootstrapped neighbor-joining tree based on the alignment of this 157 bp was constructed. This 157 bp could distinguish γ-proteobacterial species from different genera from same family. Closely related species could not be distinguished. Next, an alignment of the 16S-23S ITS nucleotide sequences of alleles from same bacterial strain was performed. The first 75 bp at the 5' end of the 16S-23S ITS was highly conserved at the intra-strain level. It was selected as the second part in the construction of our nucleotide sequence marker. Finally, a bootstrapped neighbor-joining tree based on the alignment of this 232 bp sequence was constructed. Based on the topology of the neighbour-joining tree, four major Groups, Group I to IV, were revealed with several sub-groups and clusters. Our results, based on the 232 bp sequence were, in general, in agreement with the phylogeny of γ-proteobacteria based on the 16S rRNA gene. The use of this 232 bp sequence as a phylogenetic marker presents several advantages over the use of the entire 16S rRNA gene or the generation of extensive phenotypic and genotypic data in phylogenetic analyses. First, this marker is not allele-dependant. Second, this 232 bp marker contains 157 bp from the 3' end of the 16S rRNA gene and 75 bp from the 5' end of the 16S-23S ITS. The 157 bp allows discrimination among distantly related species. Owing to its higher rate of nucleotide substitutions, the 75 bp adds discriminating power among closely related species from same genus and closely related genera from same family. Because of its higher percentage of nucleotide sequence divergence than the 16S rRNA gene, the 232 bp marker can better discriminate among closely related γ-proteobacterial species. Third, the method is simple, rapid, suited to large screening programs and easily accessible to most laboratories. Fourth, this marker can also reveal γ-proteobacterial species which may appear misassigned and for which additional characterization appear warranted.

Keywords: γ-Proteobacteria; 16S rRNA; 16S-23S ITS; Phylogeny

1. INTRODUCTION

The phylum proteobacteria or "purple bacteria and their relatives" encompasses bacteria with a wide variety of phenotype and physiological attributes and habitats [1,2]. Proteobacteria have been classified based on the homology of 16S ribosomal RNA or by hybridization of ribosomal DNA with 16S and 23S ribosomal RNA [3-6]. They have been subdivided in five major classes: α-, β-, γ-, δ- and ε- [7-9].

Most γ-proteobacteria are Gram-negative. This class

comprises 14 Orders and more than 200 genera. The γ-proteobacteria exhibit a wide range of metabolic diversity. Most are chemo-organotrophs, some are phototrophs or chemolitotrophs [1,2,10,11]. This class includes several medically and scientifically important bacteria. Some genera are human (*Klebsiella, Shigella, Salmonella, Yersina, Vibrio*), animal (*Pasteurella*) or plant pathogens (*Pseudomonas*, *Xanthomonas*, *Xylella*). Others are obligate endosymbionts (*Buchnera*, *Sodalis*, *Wigglesworthia* and *Coxiella*) [10-12]. Because of their biological importance, γ-proteobacteria are extensively studied.

The 16S ribosomal RNA (rRNA) gene has been established as the macromolecule of choice for phylogenetic analyses [5,13]. The current phylogeny of γ-proteobacteria is based on the homology of 16S rDNA nucleotide sequences [3-6,11,14].

The 16S-23S internal transcribed spacer (ITS) region is more variable than the 16S rRNA gene. It has been used, among others, in the study of specific γ-proteobacterial diversity at the species level, including *Escherichia*, *Haemophilus*, *Xanthomonas*, *Klebsiella* and *Pseudomonas* [15-19].

Additional approaches, based on different genes, have been used for the study of γ-proteobacterial phylogeny [12,20-25]. Very recently, Gao *et al*., [26] have used a combination of phylogenomic and comparative genomic approaches to reconstruct the phylogeny of γ-proteobacteria.

In an earlier work on the bacterial Gram-positive *Bacillus* genus and related genera [27], a short 220 bp nucleotide sequence "marker" was used to reconstruct their phylogeny. This 220 bp marker was a combination of a 150 bp sequence at the 3' end of the 16S rRNA gene and a 70 bp sequence at the 5' end of the 16S-23S ITS sequence. Owing to its higher rate of nucleotide substitution, the 70 bp sequence at the 5' end of the 16S-23S ITS sequence added a greater discriminatory power among closely related species than 16S rRNA gene nucleotide sequences alone. They showed that the phylogeny inferred from the 220 bp marker was in agreement with then current classifications based on phenetic and molecular data. The marker also indentified species which appeared misassigned. It also created new clusters suggesting the creation of new taxa levels. In a very recent study, we [28] have tested whether or not this marker could reconstruct the phylogeny of the bacterial Gram-positive Order of the *Bacillales*.

In the current study, we further assess the usefulness of a similar marker among 13 of the 14 γ-proteobacterial Orders. The last 157 bp at the 3' end of the 16S rRNA gene was combined with the first 75 bp at the 5' end of the 16S-23S Internal Transcribed Spacer (ITS) to yield a single 232 bp DNA marker. This marker was used to reconstruct the phylogeny of γ-proteobacteria. A total of 64 γ-proteobacteria from 13 Orders, 22 families, 40 genera and 59 species was analyzed.

2. MATERIALS AND METHODS

2.1. Bacterial Species and Strains

A total of 64 γ-proteobacterial species and strains were analyzed. They were selected on the basis that their complete genome sequences were freely available in GenBank, at the National Center for Biotechnology Information (NCBI) completed microbial genomes database (http://www.ncbi.nlm.nih.gov/genomes/MICROBES/Complete.html). They encompassed 13 Orders, 22 families, 40 genera and 59 species. These 13 γ-proteobacterial Orders included six *Aeromonadales* families, one *Cardiobacteriales* family, two *Chromatiales* families, one *Enterobacteriales* family, two *Legionellales* families, one *Methylococcales* family, two *Oceanospirillales* families, one *Pasteurellales* family, two *Pseudomonadales* families, two *Thiotrichales* families, one *Vibrionales* family and one *Xanthomonadales* family. All bacterial species and strains and the GenBank accession number for their fully sequenced genome are listed in **Table 1**.

2.2. Sequences Analysis

The 16S rRNA gene nucleotide sequences were retrieved from GenBank (**Table 1**) for the 64 γ-proteobacteria species and strains under study. First, all 64 sequences were aligned using ClustalW [29] (data not shown). Next, the 3' end of the 16S rRNA gene nucleotide sequences of alleles from same bacterial strain, of alleles from different strains from same species, and of alleles from different species from same genus (data not shown) were aligned using ClustalW [29]. The length of the nucleotide sequence most conserved was determined at 157 bp. Likewise, the 16S-23S Internal Transcribed Spacer (ITS) nucleotide sequences of alleles from same bacterial strain were also aligned using ClustalW. The length of the nucleotide sequence most conserved was determined at 75 bp.

These two most conserved nucleotide sequences, the 157 bp at the 3' end of 16S, and the 75 bp at the 5' end of 16S-23S ITS were combined into a single 232 bp sequence for each bacterial species and strain under study. This 232 bp sequence will be used here as a phylogenetic marker for the γ-proteobacteria under study.

2.3. Phylogenetic trees

Two neighbor-joining trees were constructed [30], a first one based on the alignment of the last 157 bp at the 3' end of the 16S rRNA gene described above, a second

one based on the alignment of the 232 bp sequence also described above. Both trees were bootstrapped using 1,000 random samples of sites from the alignment, all using CLUSTAL W [29] at the DNA Data Bank of Japan (DDBJ) (http://clustalw.ddbj.nig.ac.jp/top-e.html), with the Kimura's parameter method [31]. The neighbor-joining tree was drawn using TreeView (version 1.6.6) [32,33].

Table 1. γ-proteobacteria strains used in this study.

Orders, Families, Genera, Species	Strain/Source	GenBank Accession no.
Aeromonadales		
Aeromonodaceae		
Aeromonas hydrophyla	ATCC 7966	NC_008570.1
Aeromonas salmonida	A449	NC_009348.1
Alteromonadales		
Alteromonadaceae		
Marinobacter aquaeolei	VT8	NC_008740.1
Saccharophagus degradans	2-40	NC_007912.1
Colwelliaceae		
Colwellia psychrerythrea	34H	NC_003910.7
Idiomarinaceae		
Idiomarina ihoihiensis	L2TR	NC_006512.1
Pseudoalteromonadaceae		
Pseudoalteromonas atlantica	T6c	NC_008228.1
Pseudoalteromonas haloplanktis	TAC125	NC_007481.1
Shewanellaceae		
Shewanella amazonensis	SB2B	NC_008700.1
Shewanella denitrificans	OS217	NC_007954.1
Shewanella frigidimarina	NCIMB 400	NC_008345.1
Shewanella oneidensis	MR-1	NC_004347.1
Cardiobacteriales		
Cardiobacteriaceae		
Dichelobacter nodosus	VCS1703A	NC_009446.1
Chromatiales		
Chromatiaceae		
Nitrosococcus oceani	ATCC 19707	NC_007484.1
Ectothiorhodospiraceae		
Alkalilimnicola ehrlichei	MLHE-1	NC_008453.1
Halorhodospira halophila	SL1	NC_008789.1
Enterobacteriales		
Enterobacteriaceae		
Buchnera aphidicola	APS	NC_002528.1
Citrobacter koseri	ATCC BAA-895	NC_009792.1
Enterobacter sakazakii	ATCC BAA-894	NC_009778.1
Enterobacter sp	638	NC_009436.1
Escherichia coli	CFT073	NC_004431.1
Escherichia coli	O157:H7 Sakai	NC_002695.1
Klebsiella pneumoniae	342	NC_011283.1
Photorhabdus luminescens	NTUH-K2044	NC_012731.1
Salmonella enterica	TT01	NC_005126.1
Salmonella enterica	Ty2	NC_004631.1
Salmonella enterica	arizonae	NC_010067.1
Shigella boydii	Sb227	NC_007613.1
Shigella dysenteriae	Sd197	NC_007606.1
Shigella flexineri	2457T	NC_004741.1
Sodalis glossinidus	morsitans	NC_007712.1
Wigglesworthia glossinidia		NC_004344.2
Yersinia enterocolitica	8081	NC_008800.1
Yersinia pestis	CO92	NC_003143.1
Yersinia pseudotuberculosis	IP31758	NC_009708.1
Legionellales		
Coxiellaceae		
Coxiella burneti	Dugway 5j108-111	NC_009727.1
Coxiella burnetii	RSA 493	NC_002971.3
Legionellaceae		
Legionella pneumophila	Lens	NC_006369.1
Legionella pneumophila	Corby	NC_009494.1
Methylococcales		
Methylococcaceae		
Methylococcus capsulatus	Bath	NC_002977.6
Pseudomonadales		
Moraxellaceae		
Acinetobacter baumannii	ATCC 17978	NC_009085.1
Acinetobacter sp.	ADP1	NC_005966.1
Psychrobacter arcticus	273-4	NC_007204.1
Psychrobacter cryohalolentis	K5	NC_007969.1
Pseudomonadaceae		
Pseudomonas fluorescens	Pf5	NC_004129.6
Pseudomonas syringea	DC3000	NC_004578.1
Thiotrichales		
Francisellaceae		
Francisella philomiragia	*philomiragia*	NC_010336.1
Francisella tularensis	*horlatica*	NC_007880.1
Piscirickettsiaceae		
Thiomicrospira crunogena	XCL-2	NC_007520.2
Vibrionales		
Vibrionaceae		
Photobacterium profundum	SS9	NC_006370.1
Vibrio cholerae	N16961	NC_002505.1
Vibrio parahaemolyticus	RIMD 2210633	NC_004603.1
Vibrio vulnificus	CMCP6	NC_004459.2
Xanthomonadales		
Xanthomonadaceae		
Xanthomonas axonopodis	306	NC_003919.1
Xanthomonas campestris	8004	NC_007086.1
Xylella fastidiosa	9a5c	NC_002488.3

3. RESULTS AND DISCUSSION

In a previous study [27], on the bacterial genus *Bacillus* and closely-related genera, we reported the development of a short DNA marker that could be used to reconstruct their phylogeny. This marker was a combination of the last 150 bp at the 3' end of the 16S rRNA gene and the first 70 bp at the 5' end of the 16S-23S rRNA internal transcribed spacer (ITS) into a single 220 bp "marker". It could cluster *Bacillus* species and species from closely related genera into taxa akin to genera and could also distinguish closely related species. The 3' end of the 16S rRNA gene contained three regions that were known to be highly conserved among bacteria [34]. The 5' end of the 16S-23S rRNA ITS was conserved among alleles from same strains [27].

In the current study on γ-proteobacteria, we further assessed the usefulness of this marker. The sizes of the 3' end of the 16S rRNA gene and the 5' end of the 16S-23S rRNA ITS retained here for the construction of our phylogenetic marker for γ-proteobacteria are slightly different at 157 and 75 bp, respectively, for a total marker size of 232 bp. These sizes were selected as follows: first, an alignment of the 16S rRNA gene nucleotide sequence of alleles from same strain showed that these sequences were highly conserved. The intra-strain alleles shared 99% nucleotide sequence identities. Alleles from species from same genus, however, covered a wider spectrum of nucleotide sequence identities. Whereas alleles from *Pseudomonas* (*Ps*.) *fluorescens* Pf-5 and from *Ps. syringae* pv. tomato strain DC3000 share 98% nucleotide sequence identities, alleles from *Vibrio* (*V.*) *cholerae* and from *V. parahaemolyticus* RIMD 2210633 share 92% nucleotide sequence identities. Comparative analyses of the 3' end of the 16S rRNA gene nucleotide sequence showed that the last 157 bp were in many cases highly conserved among strains from same species. This is exemplified by *Salmonella* (*Sal*.) *enterica* arizonae and *Sal. enterica* Ty2 which share 99% nucleotide sequence identities over the last 157 bp at the 3' end of the 16S rRNA gene. Species from same genus share lower nucleotide sequence identities. This is exemplified by *Shewanella* (*She*) *amazonensis* and *She. denetrificans* which share 94% nucleotide sequence identities over the last 157 bp at the 3' end of the 16S rRNA gene.

A bootstrapped neighbor-joining tree based on the alignment of this 157 bp located at the 3' end of the 16S rRNA gene was constructed (**Figure 1**). Although, in most cases, this 157 bp could distinguish species from different genera from same family, in some cases, closely related species from different genera from same family appeared undistinguishable. This is the case for *Mannheimia succiniciproducens* and both *Actinobacillus* (*Act*), species, *Act. pleuropneumonia* and *Act. succinogenes*. This is also true for *Xylella* (*Xy*.) *fastidiosa* and both *Xanthomonas* (*X*.) species, *X. axonopodis pv. citri* str. 306 and *X. campestris pv. campestris* str.8004. And this is true for the human and animal pathogenic *Enterobacteriaceae*: the *Yersinia*, *Shigella*, *Klebsiella*, *Escherichia*, *Enterobacter* and *Citrobacter* species. In some cases, this 157 bp could distinguish species from same genus as exemplified by the *Psychrobacter*, *Pseudomonas*, *Shewanella* and *Vibrio* species. In other cases, closely related species from same genus could not be distinguished as exemplified by the *Francisella*, *Aeromonas*, *Actinobacillus*, *Haemophilus*, *Acinetobacter* and *Xanthomonas* species.

In all cases, this 157 bp could distinguish species from different families, with one exception: *She. amazonensis* and *Photobacterium* (*Ph*.) *profundum*, members of the *Shewanellacea* and *Vibrionacea* family, respectively. Both appear undistinguishable. Clearly, this 157 bp seq uence cannot distinguish closely-related species. An additional DNA sequence appears necessary to better distinguish closely-related species.

Next, an alignment of the 16S-23S ITS nucleotide sequences of alleles from same bacterial strain was carried with a subset of the bacteria under study: *Xanthomonas campestris* pv. campestris str. 8004, *Ps. syringae* pv. tomato str. DC3000, *Act. succinogenes* 130Z, *E. coli* K12, *V. parahaemolyticus* RIMD 2210633 and *Shigella flexneri* 2a str. 301 (**Figure 2**). The total number of alleles vary from two to ten for *X. campestris* pv. campestris str. 8004 and *V. parahaemolyticus* RIMD 2210633, respectively. The allelic sequences were highly homologous for some species and highly hererologous for others. *Xanthomonas campestris* pv. campestris str. 8004 and *Ps. syringae* pv. tomato str. DC3000 carry two and five identical alleles, respectively. *Actinobacillus succinogenes* 130Z, *E. coli* K12, *V. parahaemolyticus* RIMD 2210633 and *Shigella flexneri* 2a str. 301 carry five, eight, ten and six alleles respectively, with varying level of heterogeneity, where highly homologous alleles are grouped together and can be distinguised from different alleles in same strain (**Figure 2**). Alleles carry from zero to four tRNA genes. An alignment of the nucleotide sequences among alleles at the intra-strain level required the introduction of several gaps. The first 75 bp at the 5' end of the 16S-23S ITS, however, was highly conserved at the intra-strain level. It was retained here for the construction of our phylogenetic marker. The two conserved nucleotide sequences identified above, the 157 bp at the 3' end of 16S rRNA gene and the 75 bp at the 5' end of 16S-23S ITS, were combined into a single 232 bp sequence. This will be used here as a phylogenetic marker for the γ-proteobacteria under study.

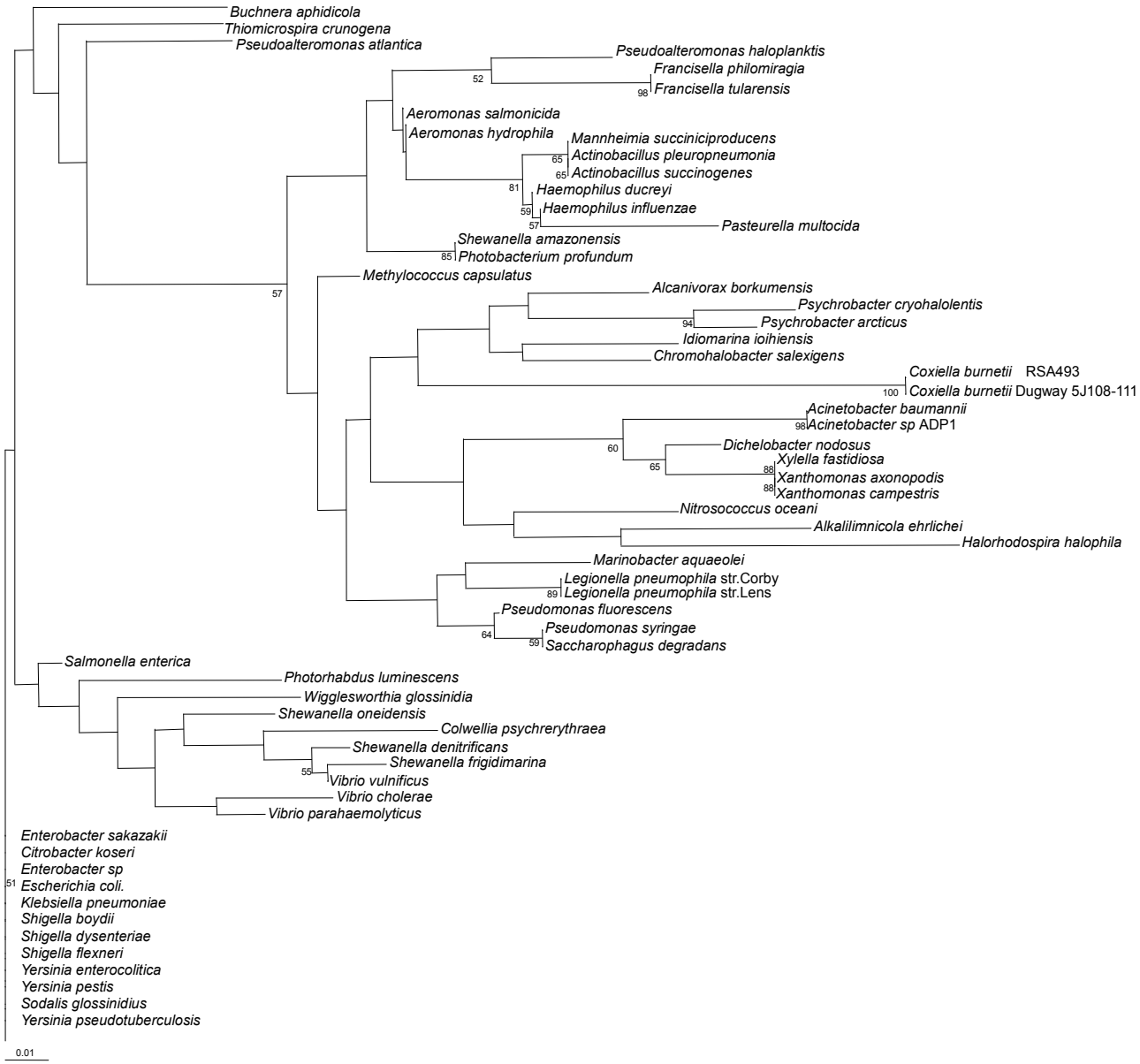

Figure 1. Bootstrapped neighbor-joining tree of γ –proteobacteria species inferred from the alignment of a 157 nucleotide sequence at the 3' end of the 16S rRNA gene. Bootstrap values higher than 50% are indicated (expressed as percentage of 1000 replication). The horizontal bar represents 1% nucleotide difference.

A bootstrapped neighbor-joining tree based on the alignment of a 232 bp sequence was constructed (**Figure 3**). Based on the topology of the neighbor-joining tree, four major Groups, Group I to IV are revealed (**Figure 3**). Based on nucleotide sequence identities, sub-groups and clusters can be formed. Group I contains seven Orders and nine families. Group II contains eight Orders and eleven families. Group III contains three Orders and five families. Group IV contains one Order and one family. Of the 13 Orders under study, four are present in more than Group one. The *Thiotrichales* are present in Groups I and III. The *Alteromonadales* are present in Groups I, II and III. The *Legionellales* are present in Groups I and II. The *Vibrionales* are present in Groups I and III. All other nine Orders are present in a single Group. Of the 22 families under study, four are present in more than one Group. The *Alteromonadaceae* are present in Groups I and II. The *Pseudoalteromonadaceae* (*Pse*) are present in Groups I and III. The *Vibrionaceae* are present in Groups I and III. The *Shewanellaceae* are present in Groups I and III. All other 18 families are present in a single Group. All species from same genus are present in

same Group with the exception of *Pse. haloplanktis* and *Pse. atlantica* present in Group I and III, respectively.

Group I can be sub-divided into six sub-groups, sub-group I-1 to I-6. Sub-group I-1 contains both species of the *Francisellaceae* family. Sub-group I-2 contains three species, members of two families: *Pseudomonadaceae* and *Alteromonadaceae*. Sub-group I-3 comprises *Pse. haloplanktis* and both *Legionella pneumophila* strains. Although in the same sub-group, *Pse. haloplanktis* shows 20% nucleotide sequence divergence with the two *Legionella* strains. Both *Legionella* strains are tightly grouped together and form cluster I-3-1. Sub-group I-4 contains *Ph. profundum* and *She. amazonensis*, two species from two different families, *Vibrionaceae* and *Shewanellaceae*, respectively. Although they appear very similar on the neighbor-joining tree, both sequences show 15% nucleotide divergence. Sub-group I-5 contains both *Aeromonas* species, tightly grouped together. Sub-group I-6 contains all six *Pasteurellaceae* species. The *Mannheimia* species is tightly grouped with the two *Actinobacillus* species and form cluster I-6-1. Both *Haemophilus* species form cluster I-6-2. In Group I, clusters comprised species from same genus or closely related species from different genera from same family. All other branches corresponded to families.

Group II can be sub-divided into four sub-groups, sub-group II-1 to II-4, and one ungrouped species, *Methylococcus capsulatus*. Sub-group II-1 comprises members of four families from four Orders. Members of sub-group II-1 share up to 30% nucleotide sequence divergences. Closely related species can be further grouped together. This is the case for the two *Acinetobacter* species and all three *Xanthomonadaceae* species which form cluster II-1-1 and II-1-2, respectively. Sub-group II-2 contains *Nitrosoccocus oceani* and *Halorhodospira halophila*, member of the *Chromatiaceae* and *Ectothiorhodospiracea* family respectively. Both species show 42% nucleotide sequence divergences. Both families belong to the *Chromatiales* Order. Sub-group II-3 contains six species, from five families and three Orders. Members of sub-group II-3 share up to 24% nucleotide sequence divergences. Two closely related species, *Psychrobacter* (*Psy*) *cryohalolentis* and *Psy. arcticus* form cluster II-3-1. Sub-group II-4 comprises both strains of *Coxiella burnetti*, members of the *Coxelliaceae* family. In Group II, clusters comprised species from same genus or closely related species from different genera from same family. All other branches corresponded to families.

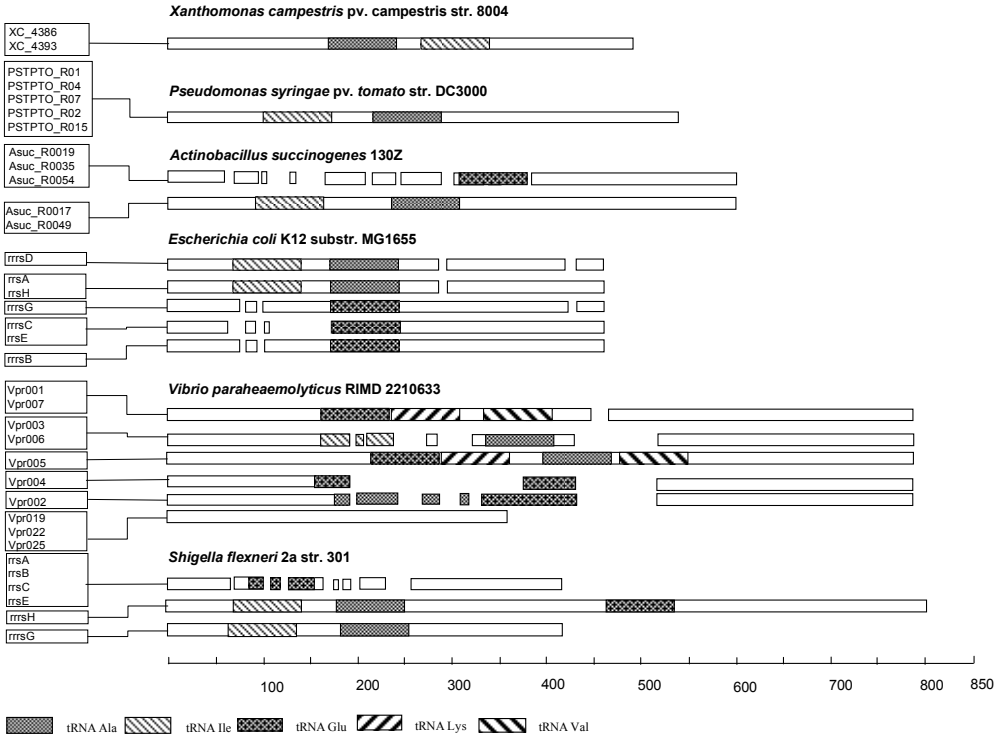

Figure 2. Schematic representation of allelic 16S-23S rDNA Internal Transcribed Spacer of γ –proteobacteria.The non filed boxes represent regions of homologus nucleotide sequences between allelic ITS of the same bacteria. Filed boxes represent tRNA. The blank spaces between boxes represent non conservation regions between allelic ITS of the same bacteria.

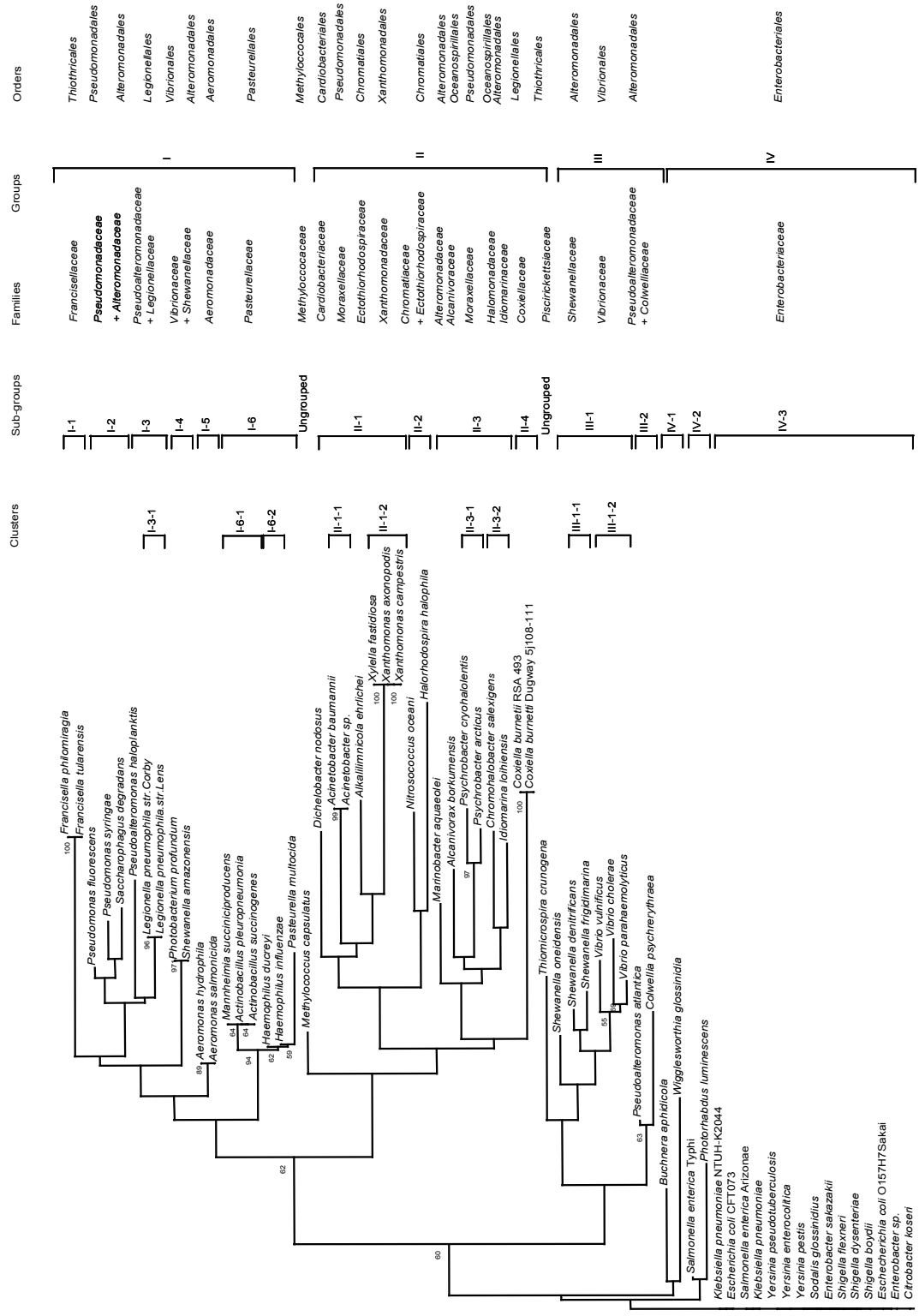

Figure 3. Bootstrapped neighbor-joining tree of γ-proteobacteria species inferred from the alignment of 232 nucleotide sequence marker. This 232 nucleotide sequence marker is a combination of a 157 nucleotide sequence at the 3' end of the 16S rRNA gene and a 75 nucleotide sequence at the 5' end of the 16S-23S Internal Transcribed Spacer (ITS) sequence. Major Groups are indicated in capital roman numerals. Sub-groups and clusters are indicated in arabic numbers. Bootstrap values higher than 50% are indicated (expressed as percentage of 1000 replication). The horizontal bar represents 1% nucleotide difference.

Group III can be sub-divided into two sub-groups, sub-groups III-1 and III-2, and one ungrouped species, *Thiomicrospira crunogena*. Sub-group III-1 contains three *Shewallenaceae* and three *Vibrionaceae* species. *She. denitrificans* and *She. frigidimarina*, and the three *Vibrio* species form two clusters, III-1-1 and III-1-2, respectively. Sub-group III-2 contains two genera from two families of the same Order. Both species, on two separate branches, show 23% nucleotide sequence divergence. In Group III, both clusters comprised species from same genus. All other branches corresponded to families.

Group IV contains all the *Enterobacteriaceae* species under study. Three sub-groups can be revealed: sub-group IV-1 to IV-3. Sub-group IV-1 contains two insect obligate endosymbionts. Sub-group IV-2 contains *Salmonella enterica* Ty2 and *Photorhabtus luminescens*. All other *Enterobacteriaceae* species are in sub-group IV-3. The latter are closely related to each others.

Our results, based on the 232 bp phylogenetic marker described here are, in general, in agreement with the phylogeny of γ-proteobacteria based on the 16S rRNA gene with some exceptions. In the neighbor-joining tree, clusters comprised species from same genus or closely related species from different genera from same family. All other branches corresponded to families. As indicated above, of the 22 families under study, 18 are present in a single Group and four are present in more than one Group. These latter four families encompass the marine bacteria [35]. They are *Vibrionaceae*, in Groups I and III; and the *Alteromonas*-related protobacteria (Ivanova *et al*., 2004): *Alteromonadaceae*, in Groups I and II; *Pseudoalteromonadaceae*, in Groups I and III; and *Shewanellaceae*, in Groups I and III. Interestingly, within a Group, these marine bacteria are found in close proximity of one another. It reflects the varying level of heterogeneity among *Alteromonas*-related protobacteria. The grouping showed here is based on a 232 bp marker. The biological significance of this grouping is unknown. Clearly, however, the phylogenetic analyses of these related marine heterotrophic bacteria is a work in progress [36].

4. CONCLUSIONS

In conclusion, the use of this 232 bp marker presents several advantages over the use of the entire 16S rRNA gene or the generation of extensive phenotypic and genotypic data in phylogenetic analyses. First, this marker is not allele-dependant. The 3' end of the 16S rRNA gene is highly conserved at the intra-strain level. We have shown here that although the16S-23S ITS allelic sequences can be very heterogeneous within a strain, the first 75 bp, however, are conserved among alleles from same strain in γ-proteobacteria. Clearly, any allele would generate the same results. Second, this 232 bp marker contains 157 bp from the 3' end of the 16S rRNA gene and 75 bp from the 5' end of the 16S-23S ITS. The 157 bp is highly conserved among closely related species. Owing to its higher rate of nucleotide substitutions, the 75 bp adds discriminating power among closely related species from same genus and closely related genera from same family. Because of its higher percentage of nucleotide sequence divergence than the 16S rRNA gene, the 232 bp marker can better discriminate among closely related γ-proteobacteria species. Third, the method is simple, rapid, suited to large screening programmes and easily accessible to most laboratories. More importantly, however, this 232 bp marker can group γ-proteobacteria families and genera in accordance with established phylogenies, with the exceptions indicated above. It can also reveal γ-proteobacteria species which may appear misassigned and for which additional characterization appear warranted.

REFERENCES

[1] Stackebrandt, E., Murray, R.G.E. and Trüper, H.G. (1988) Proteobacteria classis nov., a name for the phylogenetic taxon that includes the "purple bacteria and their relatives". *International Journal of Systematic Bacteriology*, **38(3)**, 321-325.

[2] Gupta, R.S. (2000) The phylogeny of proteobacteria: relationships to other eubacterial phyla and eukaryotes. *FEMS Microbiology Reviews*, **24(4)**, 367-402.

[3] Fox, G.E., Stackebrandt, E., Hespell R.B., Gibson J., Maniloff, J., Dyer, T.A., Wolfe, R.S., Balch, W.E., Tanner, R.S., Magrum, L.J., Zablen, L.B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B.J., Stahl, D.A., Luehrsen, K.R., Chen, K.N. and Woese, C.R. (1980) The phylogeny of prokaryotes. *Science*, **209(4455)**, 457-463.

[4] Woese, C.R., Stackebrandt, E., Macke, T.J. and Fox, G. E. (1985a) A phylogenetic definition of the major eubacterial taxa. *Systematic and Applied Microbiology*, **6**, 143-151.

[5] Woese, C.R. (1987) Bacterial Evolution. *FEMS Microbiology Reviews*, **51(2)**, 221-271.

[6] De Ley, J. (1992) The proteobacteria: Ribosomal RNA cistron similarities and bacterial taxonomy. In: by Balows, A., Trüper, H.G., Dworkin, M., Harder, W., and Schleifer, K.H., Eds., *The Prokaryotes*, 2nd Edition, Springer-Verlag, Berlin, 2111-2140.

[7] Woese, C.R., Stackebrandt, E., Weisburg, W.G., Paster, B. J., Madigan, M.T., Fowler, V.J., Hahn, C.M., Blanz, P., Gupta, R., Nealson, K.H. and Fox, G.E. (1984a) The phylogeny of purple bacteria: The alpha subdivision. *Systematic and Applied Microbiology*, **5**, 315-326.

[8] Woese, C.R., Weisburg, W.G., Paster, B.J., Hahn, C.M., Tanner, R.S., Krieg, N.R., Koops, H.-P., Harms, H. and Stackebrandt, E. (1984b) The phylogeny of purple bacteria: The beta subdivision. *Systematic and Applied Micro-*

biology, **5**, 327-336.

[9] Woese, C.R., Weisburg, W.G., Hahn, C.M., Paster, B.J., Zablen, L.B., Lewis, B.J., Macke, T.J., Ludwig, W. and Stackebrandt, E. (1985b) The phylogeny of purple bacteria: The gamma subdivision. *Systematic and Applied Microbiology*, **6**, 25-33.

[10] Brenner, D.J., Krieg, N.R., Staley, J.T. and Garrity, G.M., (Eds.) (2005) Bergey's manual of systematic bacteriology, 2nd Edition, Vol. 2 (The *Proteobacteria*), part C (The *Alpha-*, *Beta-*, *Delta-*, and *Epsilonproteobacteria*), Springer, New York.

[11] Kersters, K., Devos, P., Gillis, M., Swings, J., Vandamme, P. and Stackebrandt, E. (2006) Introduction to the *Proteobacteria*. In: Dworkin, M., Falkow, S., Rosenberg, E. Schleifer, K.H. and Stackebrandt, E., Eds., *The Prokaryotes*: *A Handbook on the Biology of Bacteria*. Springer, New York, 3-37.

[12] Belda, E., Moya, A. and Silva, F.J. (2005) Genome rearrangement distances and gene order phylogeny in γ-proteobacteria. *Molecular Biology and Evolution*, **22(6)**, 1456-1467.

[13] Woese, C.R., Kandler, O. and Wheelis, M.L. (1990) Towards a natural system of organisms: Proposal for the domains archaea, bacteria, and eucarya. *Proceedings of the National Academy of Sciences*, USA, **87(12)**, 4576-4579.

[14] Ludwig, W. and Klenk, H.P. (2005) Overview: A phylogenetic backbone and taxonomic framework for prokaryotic systematics. In: Brenner, D.J. Krieg, N.R. Staley J.T. and Garrity, G.M., Eds., *Bergey's Manual of Systematic Bacteriology*, Springer-Verlag, Berlin, 49-65.

[15] Anton, A.I., Martinez-Murcia, A.J. and Rodriguez-Valera, F. (1998) Sequence diversity in the 16S-23S intergenic spacer region (ISR) of the rRNA operons in representatives of the *Escherichia coli* ECOR collection. *Journal of Molecular Evolution*, **47(1)**, 62-72.

[16] Giannino, V., Rappazzo, G., Scuto, A., Di Marco, O., Privitera, A., Santagati, M. and Stefani, S. (2001) rrn operons in *Haemophilus parainfluenzae* and mosaicism of conserved and species-specific sequences in the 16S-23S rDNA long spacer. *Research in Microbiology*, **152(5)**, 461-468.

[17] Goncalves, E.R. and Rosato, Y.B. (2002) Phylogenetic analysis of *Xanthomonas* species based upon 16S-23S rDNA intergenic spacer sequences. *International Journal of Systematic and Evolutionary Microbiology*, **52(Pt 2)**, 355-361.

[18] Wang, M., Cao, B., Yu, Q., Liu, L., Gao, Q., Wang, L. and Feng, L. (2008) Analysis of the 16S-23S rRNA gene internal transcribed spacer region in *Klebsiella* species. *Journal of Clinical Microbiology*, **46(11)**, 3555-3563.

[19] Tambong, J.T., Xu, R. and Bromfield, E.S.P. (2009) Intercistronic heterogeneity of the 16S-23S rRNA spacer region among *Pseudomonas* strains isolated from subterranean seeds of hog peanut (*Amphicarpa bracteata*). *Microbiology*, **155(Pt 8)**, 2630-2640.

[20] Kunisawa, T. (2001) Gene arrangements and phylogeny in the class *Proteobacteria*. *Journal of Theoretical Biology*, **213(1)**, 9-19.

[21] Lerat, E., Daubin, V. and Moran, N.A. (2003) From gene trees to organismal phylogeny in prokaryotes: The case of the γ-proteobacteria. *Public Library of Science Biology*, **1(1)**, e19.

[22] Brown, J.R. and Volker, C. (2004) Phylogeny of γ –proteobacteria: Resolution of one branch of the universal tree? *Bioessays*, **26(5)**, 463-468.

[23] Ciccarelli, F.D., Doerks, T., von Mering, C., Creevey, C. J., Snel, B. and Bork, P. (2006) Toward automatic reconstruction of a highly resolved tree of life. *Science*, **311 (5765)**, 1283-1287.

[24] Mrazek, J., Spormann, A.M. and Karlin, S. (2006) Genomic comparisons among γ-proteobacteria. *Environmental Microbiology*, **8(2)**, 273-288.

[25] Lee, H.Y. and Côté, J.-C. (2006) Phylogenetic analysis of γ-proteobacteria inferred from nucleotide sequence comparisons of the house-keeping genes *adk*, *aroE* and *gdh*: comparisons with phylogeny inferred from 16S rRNA gene sequences. *Journal of General and Applied Microbiology*, **52(3)**, 147-158.

[26] Gao, B., Mohan, R. and Gupta, R.S. (2009) Phylogenomics and protein signatures elucidating the evolutionary relationships among the *Gammaproteobacteria. International Journal of Systematic and Evolutionary Microbiology*, **59(2)**, 234-247.

[27] Xu, D. and Côté, J.-C. (2003) Phylogenetic relationships between *Bacillus* species and related genera inferred from comparison of 3' end 16S rDNA and 5' end 16S-23S ITS nucleotide sequences. *International Journal of Systematic and Evolutionary Microbiology*, **53(Pt 3)**, 695-704.

[28] Yakoubou, S., Xu, D. and Côté, J.-C. Phylogeny of *Bacillales* inferred from 3' 16S rDNA and 5' 16S-23S ITS nucleotide sequences. *Nature Science*, in Press.

[29] Thompson, J.D., Higgins, D.G. and Gibson, T.J. (1994) Clustal W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research*, **22(22)**, 4673-4680.

[30] Saitou, N. and Nei, M. (1987) The neighbor-joining method: A new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, **4(4)**, 406-425.

[31] Kimura, M. (1983) The neutral theory of molecular evolution. Cambridge University Press, Cambridge.

[32] Page, R.D.M. (1996) TreeView: An application to display phylogenetic trees on personal computers. *Computer Application in the Biosciences*, **12(4)**, 357-358.

[33] Page, R.D.M. (2000) TreeView—tree drawing software for apple macintosh and windows. http://taxonomy.zoology.gla.ac.uk/rod/treeview.html

[34] Gürtler, V. and Stanisich V.A. (1996) New approaches to typing and identification of bacteria using the 16S-23S rDNA spacer region. *Microbiology*, **142(Pt 1)**, 3-16.

[35] Suen, G., Goldman, B.S. and Welch, R.D. (2007) Predicting prokaryotic ecological niches using genome sequence analysis. *Public Library of Science Biology ONE*, **2(8)**, e743.

[36] Ivanova, E.P., Flavier, S. and Christen, R. (2004) Phylogenetic relationships among marine *Alteromonas*-like proteobacteria: Emended description of the family *Alteromonadaceae* and proposal of *Pseudoalteromonadaceae* fam. nov., *Colwelliaceae* fam. nov., *Shewanellaceae* fam. nov., *Moritellaceae* fam. nov., *Ferrimonadaceae* fam. nov., *Idiomarinaceae* fam. nov. and *Psychromona-daceae* fam. nov. *International Journal of Systematic and Evolutionary Microbiology*, **54(Pt 5)**, 1773-1788.