Annals of Microbiology, **59 (3)** 615-622 (2009)

Comparison of 16S rRNA, 23S rRNA and *gyrB* genes sequences in phylogenetic relationships of *Shigella* isolates from Iran

Zahra NOCHI[1,2], Navid SAHEBEKHTIARI[1,3], Pedram KHARAZIHA[1], Mohsen CHIANI[1], Kamyar MOTAVAZE[3], Hossein DABIRI[1], Fereshteh JAFARI[1], Haleh EDALATKHAH[1], Mohammad REZA ZALI[1], Babak NOORINAYER[1]

[1]The Research Center for Gastroenterology and Liver Diseases, Shahid Beheshti University, M.C., Tehran; [2]Department of Microbiology, Islamic Azad University, Science and Research Branch, Tehran; [3] Department of Microbiology, School of Basic Sciences, Islamic Azad University, North Tehran Branch, Tehran, Iran

Received 11 May 2009 / Accepted 4 August 2009

Abstract - Phylogenetic analysis of about twenty isolates of *Shigella* spp. was carried out using the nucleotide sequences of 16S rRNA, 23S rRNA and DNA gyrase B (*gyrB*) genes which was determined by directly sequencing PCR fragments. The results showed that the similarity based on 16S rRNA for specimens belonging to the same species of *Shigella flexneri*, *Shigella boydii*, *Shigella dysenteriae* and *Shigella sonnei* were 99.7, 99.6, 99.5 and 100%, the corresponding results based on 23S rRNA were 99.4, 99.7, 99.3 and 100%, and for *gyrB* were 99.4, 99.7, 99.1 and 100% respectively. Our results demonstrate that *gyrB* sequence analysis is a fruitful approach to determine the phylogenetic relationships of *Shigella* isolates and may be an alternative to 16S rRNA analysis. In particular, *gyrB* analysis of bacteria is an effective means to classify closely related species, while 23S rRNA gene sequence data may be used for phylogenetic analysis at the species level.

Key words: *Shigella*; phylogeny; 16S rRNA; 23S rRNA; *gyrB.*

INTRODUCTION

Shigella was first discovered by the Japanese scientist Shiga in the 1890s and was adopted as a genus in the 1950s (Hale, 1991). *Shigellae* are Gram-negative, non-motile and facultative interacellular pathogens (Murray *et al*., 2009), and cause bacillary dysentery or shigellosis, a significant threat to public health, particularly in less developed countries with poor sanitation. Based on phenotypic characteristics, four species are recognized in this genus: *Shigella dysenteriae*, *Shigella flexneri*, *Shigella boydii* and *Shigella sonnei* (Murray *et al*., 2009). Each of them has a number of serotypes, e.g., *S. dysenteriae* type 1-12, *S. sonnei* phase I and II, *S. boydii* type 1-18, *S. flexneri* type 1-6 (Sur *et al*., 2004).

In general, the chromosomes of *Shigella* comprise a single circular DNA molecule consisting 4×10^6 bp, with a relative molecular mass of 4×10^9 and a total length of about 1.4 mm (Fukushima *et al*., 2002).

Since 1980`s, microbiology has undergone a momentous shift from a determinative taxonomy to one based phylogeny (Woese *et al*., 1990). In the laboratories of Woese and other his colleagues, it was shown that phylogenetic relationships of bacteria, and, indeed, all life-forms, could be determined by comparing a stable part of the genetic code (Woese *et al*., 1985, Woese, 1987). Candidates for this genetic area in bacteria included the genes code for the 5S, the 16S, and the 23S rRNA and the spaces between these genes (Clarridge, 2004). The part of DNA that most commonly used for taxonomic purposes for bacteria, is the 16S rRNA gene (Bottger, 1989; Kolbert and Persing, 1999; Palys *et al*., 1997; Tortoli, 2003).

Bacterial analysis by 16S rRNA which is about 1550 bp long has become popular, because it is composed of both variable and conserved regions and also easy to sequence (Christensen *et al*., 1998; Clarridge, 2004). 16S rRNA gene analysis has proved useful for defining prokaryotic relationships from the species to phylum levels (Dewhirst *et al*., 2005).

The 23S rRNA gene offers the same advantages as the 16S rRNA gene (e.g., universal distribution, conserved function, and invariant and variable regions) (Ludwig and Schleifer, 1994). Comparison of 23S rRNA sequences allows phylogenetic analysis equivalent to 16S rRNA but due to it being twice as long and the presence of more variable regions, 23S rRNA gene sequence data is suggested for phylogenetic analysis at the species/subspecies level (Ludwig and Schleifer, 1994; Leffers *et al*., 1987; Anton *et al*., 1999).

Theoretically, any sequence which codes protein can be used for a phylogenetic analysis. However, many genes, particularly those for catabolism, are known to spread horizontally among various bacterial species, and they cannot be used to mark the evolutionary record of host bacteria. Selecting proper taxonomic

* Corresponding Author. Phone: + 98 21 2243 25 18; Fax: + 98 21 2243 25 17; E-mail: bnoorinayer@yahoo.com

marker genes may be because that, they are fundamentally necessary for housekeeping activities such as DNA replication, transcription, and translation, they are a single copy on each genome; and at least two regions are conserved for PCR amplification using primers complementary to the conserved regions. So, it was interested to apply *gyrB* sequences in the present study. *gyrB*, which is about 1.2 Kb long, is a single copy gene present in all bacteria which encodes the ATPase domain of DNA gyrase, an enzyme essential for DNA replication (Huang, 1996). *gyrB* partial sequence analysis seems to be more reliable and useful than 16S rRNA for determining the evolutionary relationships of *Shigella* species (Fukushima *et al*., 2002).

This report showed the homology rate of *Shigella* species in Iran using 16S rRNA, 23S rRNA and *gyrB* genes.

MATERIALS AND METHODS

Study bacterial strains and clinical samples. Twenty *Shigella* spp. included four serogroups, five isolates from each serogroup, were subjected for further characterization and phylogenetic relationships. The isolates had been selected from the microbial collection in the laboratory of food borne diseases in Research Center for Gastroenterology and Liver Diseases of Shahid Beheshti University of Medical Sciences. The samples subcultured on Xylose Lysine Deoxycholate medium (XLD) and McConkey agar. Non-fermenter colonies resembling *Shigella* spp. further confirmed using standard biochemical and microbiological methods (Downes *et al*., 2001). Serotyping of the isolates carried out using slide agglutination kit by specific poly and phase I and Π antisera (MAST House, Merseyside, UK) and the isolates classified within twenty serogroups based on antigenic differences.

Molecular identification of *Shigella* spp. The isolates positive for invasion plasmid antigen H (*ipaH*) and negative for invasion-associated locus (*ial*) considered as *Shigella* spp. and the isolates positive for invasion-associated locus (*ial*) considered as EIEC and excluded from the study (Sethabutr *et al*., 1993).

Primers design. The used primers in the PCR and in the sequencing protocols were designed based on the reported *Shigella boydii*, *Shigella dysenteriae*, *Shigella flexneri* and *Shigella sonnei* 16S rRNA, 23S rRNA and *gyrB* genes sequences (GenBank accession no. NC 007613, NC 007606, NC 008258, and NC 007384, respectively). 16S and 23S rRNA and *gyrB* genes sequences of *Escherichia coli* BW2952 (NC 012759), *Salmonella enterica* subsp. *enterica* serovar *paratyphi* A str. ATCC 9150 (NC 006511), *Klebsiella pneumoniae* NTHU-K2044 (NC 012731), *Yersinia enterocolitica* subsp. *enterocolitica* 8081 (NC 008800) and *Proteus mirabilis* HI4320 (NC 010554) were also used as outgroups to design appropriate primers. The query coverage for the blasted designed PCR and sequencing primers (Tables 1 and 2) for *Shigella* spp. was 100%.

16S rRNA. In order to obtain satisfactory results in DNA-sequencing and covering all the gene sequence, the 16S rRNA gene sequence was divided into two segments. The forward primer for the first segment was designed to amplify 67 bp upstream of the first nucleotide and the reverse primer for the second segment was designed to amplify 61 bp downstream of the last nucleotide to avoid missing the nucleotides in DNA sequencing.

23S rRNA. To prevent missing the nucleotides in DNA sequencing, the 23S rRNA gene sequence was divided into four segments. The forward primer for the first segment was designed to amplify 100 bp upstream of the first nucleotide and the reverse primer for the forth segment was designed to amplify 103 bp downstream of the last nucleotide.

gyrB. We have developed universal PCR primers (UP1fS and UP2Sr) with which the amplification and sequencing of *gyrB* genes from a variety of bacteria were possible (Yamamoto and Harayama, 1995). Sequencing primers were also designed to avoid missing nucleotides in DNA sequencing.

DNA extraction and amplification. To isolate chromosomal DNA, a big loopful of freshly grown colonies of bacteria were scrabed into a 1.5-ml micro centrifuge tube and resuspended in 700 µl of TE buffer (10 mM Tris-HCL; pH 7.5-8.0, 1 mM EDTA; pH 7.5-8.0). Chromosomal DNA was extracted according to standard phenol-chloroform method (Sambrook and Russell, 2001). PCR was performed using Taq DNA polymerase (Fermentas, Germany), 100 ng of bacterial DNA, 1.5 mM $MgCl_2$, 200 mM dNTPs and 10 pMol each primer mentioned in Table 1. The cycling conditions for 16S rRNA, 23S rRNA and *gyrB* genes were: initial denaturation at 95 °C for 4 min, following 30 cycles of 94 °C for 60 s, appropriate annealing temperature according to Table 1 for 30 s, 72 °C for 20 s and the final extension of 72 °C for 10 min.

Sequencing. The PCR-amplified products were subjected to direct sequencing, which was performed using a BigDye^{Tm} Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA) and a genetic Analyzer Abi Prism^{TM} 3130x1 (Applied Biosystems). All reactions were performed by sequencing with oligonucleotides mentioned in Table 2 and samples were sequenced in both directions and in case of any variations were found, results were confirmed by sequencing of at least two independent PCR products. All sequences were analyzed by Chromas version 2.22 (Queenlands, Australia).

Phylogenetic analysis. The nucleotide sequences of 16S rRNA, 23S rRNA and *gyrB* genes were aligned by the CLUSTAL W computer program (Thompson *et al*., 1994). A Maximum Parsimony analysis (Dayhoff, 1972) was used to reconstruct phylogenetic trees with the MEGA3 software (Kumar *et al*., 2004).

Nucleotide sequence accession numbers. The Nucleotide sequences of 16S rRNA, 23S rRNA and *gyrB* genes from twenty isolates of *Shigella* have been submitted to the GenBank and are available for public access under the accession numbers: EF643608, EF643609 and EU009177 to EU009199 (16S rRNA), EU855230 to EU855249 (23S rRNA), EU051355 to EU051374 (*gyrB*).

RESULTS

Serotyping

All isolated *Shigella* spp. subjected for serotyping using slide agglutination method by poly and Phase I and II antisera. The result of serotyping has been shown in Table 3.

Amplification and sequencing of 16S rRNA, 23S rRNA and gyrB genes

PCR-amplification of 16S and 23S rRNA genes has been performed using the two and four covering primer pairs to amplify complete sequence of 16S rDNA (1542 bp) and 23S rDNA (2904 bp) respectively (Table 1). Also a 1256 bp fragment of *gyrB* amplified, using universal primer pair in all twenty *Shigella* isolates (Table 1).

TABLE 1 - Primers for the PCR amplification of 16S rRNA, 23S rRNA and *gyrB* genes

Genes	Primer sequences (5'-3')	Position[a]	Product size (bp)	Annealing temp. (°C)
16S rRNA	TGAATACCAAGTCTCAAGAGTG GTCCACGCTGTAAACGAT	67 bp upstream start site 805-822	889	61
	GAAGACTGACGCTCAGGTG GATTGTCTGATAGAAAGTGAAAAGC	747-765 61 bp downstream stop site	856	62.5
23S rRNA	TCAAGCTGAAAATTGAAACACTG CAATCAAACCGGGAGATAGC	100 bp upstream start site 788-807	907	60.5
	AAGGGAAACCGAGTCTTAACTG CTGGTTTTCCAGGCAAATC	627-648 1481-1499	873	61.5
	GCGTAGTCGATGGGAAACAG GTTGGCTAATCCTGGTCGG	1370-1389 2291-2309	940	63
	TACCACCCTTTAATGTTTGATG GATAAAACAGAATTAGCCTGGC	2173-2194 103 bp downstream stop site	834	60
gyrB[b]	GAAGTCATCATGACCGTTCTGCA AGCAGGGTACGGATGTGCGAGCC	275-297 1508-1530	1256	67

[a] Positions correspond to the 16S (rrs H) and 23S rRNA (rrl H) nucleotide sequences of *Shigella flexneri* 5 str. 8401 (CP000266).
[b] UP1fS and UP2Sr Universal Primers (Yamamoto and Harayama, 1995).

TABLE 2 - Selected genes and sequencing oligonucleotides used in this study

Genes	Sequencing nucleotides	Position[a]
16S rRNA	TGAATACCAAGTCTCAAGAGTG	67 bp upstream start site
	CACACTGGAACTGAGACAC	314-332
	GAGGAAGGGAGTAAAGTTAA	449-468
	GTCCACGCTGTAAACGAT	805-822
	GAAGACTGACGCTCAGGTG	747-765
	CAACCCTTATCCTTTGTTG	1111-1129
	CTACACACGTGCTACAATG	1225-1243
	GATTGTCTGATAGAAAGTGAAAAGC	61 bp downstream stop site
23S rRNA	TCAAGCTGAAAATTGAAACACTG	100 bp upstream start site
	TACTCCTGACTGACCGATAG	432-451
	TACTCCTGACTGACCGATAG	432-451
	CAATCAAACCGGGAGATAGC	788-807
	AAGGGAAACCGAGTCTTAACTG	627-648
	AAAGCGTAATAGCTCACTGGTC	1089-1110
	GAAAGCGTAATAGCTCACTG	1088-1107
	CTGGTTTTCCAGGCAAATC	1481-1499
	GCGTAGTCGATGGGAAACAG	1370-1389
	TCGTGGAGCTGAAATCAGTC	1734-1753
	GTAACTATAACGGTCCTAAGG	1911-1931
	GTTGGCTAATCCTGGTCGG	2291-2309
	TACCACCCTTTAATGTTTGATG	2173-2194
	TGTTTGGCACCTCGATGTC	2490-2508
	GTGTTCGGGTTGTCATGCC	2687-2705
	GATAAAACAGAATTAGCCTGGC	103 bp downstream stop site
gyrB	GAAGTCATCATGACCGTTCTGCA[b]	275-297
	CTACTGCTTTACCAACAACATTC	799-821
	CAAACTGGTTTCTTCTGAGG	1015-1034
	AGCAGGGTACGGATGTGCGAGCC[b]	1508-1530

[a] Positions correspond to the 16S (rrs H) and 23S rRNA (rrl H) nucleotide sequences of *Shigella flexneri* 5 str. 8401 (CP000266).
[b] UP1fS and UP2Sr Universal Primers (Yamamoto and Harayama, 1995).

TABLE 3 - Serogroups of twenty *Shigella* isolates

Serotype	Isolates (No.)
Shigella sonnei Poly D Phase I & II	5
Shigella flexneri Poly B Types I , II , III , IV , V , VI Groups (3)4,6,7(8)	5
Shigella boydii Poly C Types 1,2,3,4,5,6,7 Poly C1 Types 8,9,10,11 Poly C2 Types 12,13,14,15 Poly C3 Types 16,17,18	 2 1 1 1
Shigella dysenteriae Poly A Types 1,2,3,4,5,6,7 Poly A1 Types 8,9,10,11,12	 3 2

PCR fragments were determined by the sequencing oligonucleotides (Table 2). Nucleotide sequences comparison in National Center of Biotechnology Information (NCBI) revealed high homology with other *Shigella* spp.

Phylogenetic analysis and genetic distance of 16S rRNA, 23S rRNA and gyrB genes

The results showed that every twenty isolates of *Shigella* have 99.7% similarity to each other based on 16S rRNA, while according to 23S rRNA; the homology rate was 99.6%. The similarity based on 16S rRNA for specimens belonging to the same species of *S. flexneri*, *S. boydii*, *S. dysenteriae* and *S. sonnei* were 99.7, 99.6, 99.5 and 100%, the corresponding

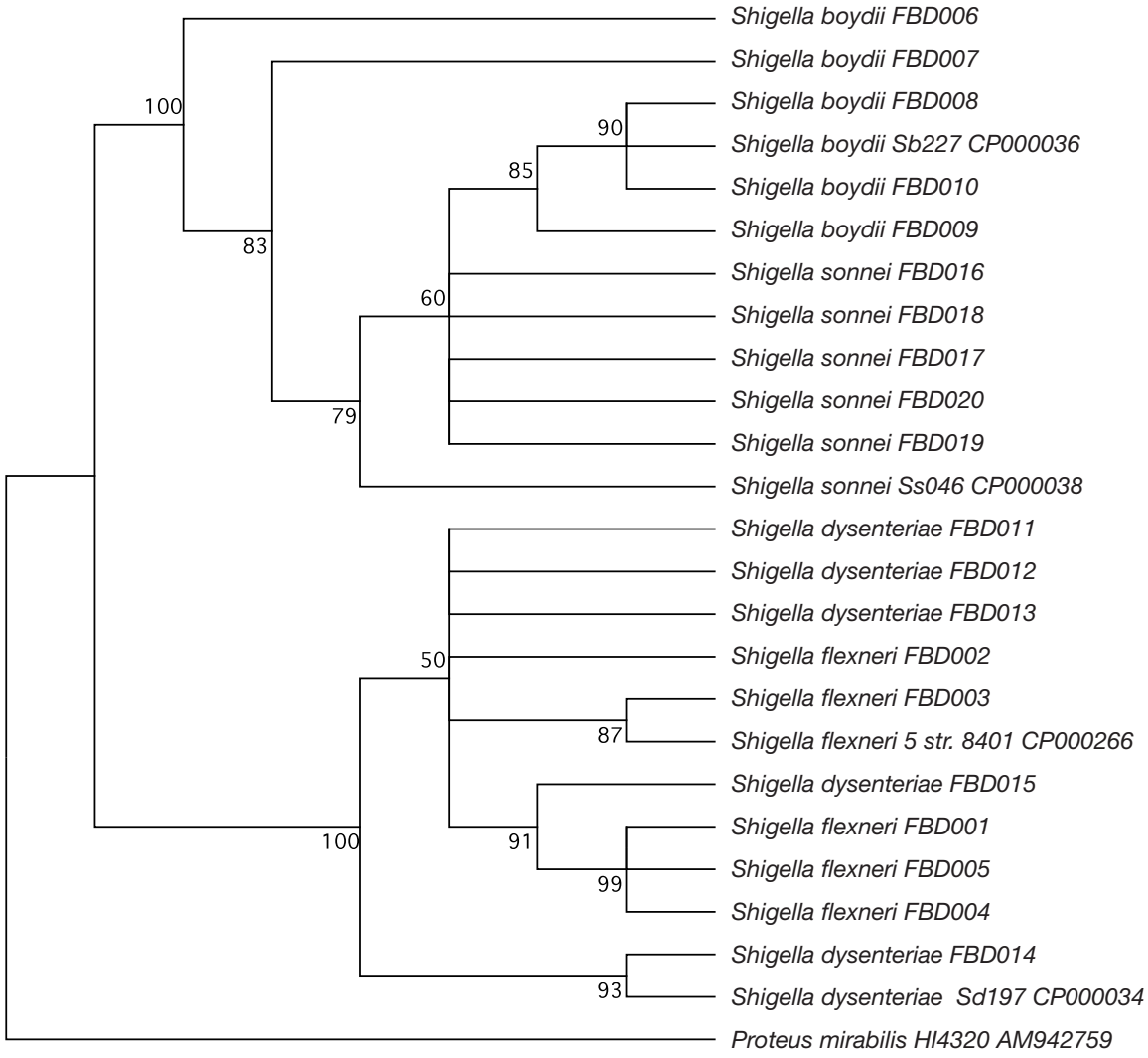

FIG. 1 - Phylogenetic tree based on the nucleotide sequences of 16S rRNA genes. The 16S rRNA sequences were adjusted to 1542 bases, and the tree was constructed by the Maximum Parsimony method, using the computer program MEGA3. The numbers shown next to the nodes indicate percent bootstrap values of the 1000 replicates. Genetic distances were computed using the Close-Neighbor-Interchange algorithm (Nei and Kumar, 2000). The sequence data for phylogenetic analysis were taken from the GenBank nucleotide sequence database for the following strains: *Shigella boydii* Sb227 (CP000036), *Shigella dysenteriae* Sd197 (CP000034), *Shigella flexneri* 5 str. 8401 (CP000266) and *Shigella sonnei* Ss046 (CP000038). *Proteus mirabilis* HI4320 (AM942759) was used as an outgroup.

results based on 23S rRNA were 99.4, 99.7, 99.3 and 100% respectively. The percent nucleotide divergence between *S. dysenteriae* and *S. flexneri* of 16S and 23S rRNA was 99.5 and 99.3% whereas between *S. boydii* and *S. sonnei* was 99.7 and 99.5%. Figures 1 and 2 show the phylogenetic trees for *Shigella* isolates on the basis of their 16S and 23S rRNA genes.

In the phylogenetic-tree analysis of *gyrB*, homology in specimens of *S. flexneri, S. boydii, S. dysenteriae* and *S. sonnei* species were 99.4, 99.6, 99.1 and 100% respectively; also the homology of these twenty isolates was 99.5%. *Shigella boydii* and *S. dysenteriae* were found to have 99.2% similarity to each other, 99.6 and 99.1% similarity to *S. flexneri* respectively. The phylogenetic tree of twenty isolates of *Shigella* based on the nucleotide sequence of *gyrB* gene has been shown in Fig. 3.

DISCUSSION

Phylogenetic-tree analysis is frequently used as a method to classify organisms (Fukushima *et al*., 2002). Variant genes such as *thrB*, *ksgA*, *pabB*, *ipaB*, *ipaD* etc., have been examined for the analysis of phylogenetic relationships of *Shigella* (Pupo *et al*., 2000; Escobar-Paramo *et al*., 2003; Yang *et al*., 2007). In general, 16S rRNA is most often used for such analyses; however 23S rRNA sequences allows equivalent result compare to 16S rRNA, but due to it being twice as long, and the presence of more variable regions, 23S rRNA gene sequence data might be better marker for phylogenetic analysis at the species/subspecies level (Leffers *et al*., 1987). Also, the *gyrB* sequence analysis is a rapid and effective method for identifying bacterial species and for examining phylogenetic relationships (Yamamoto and Harayama, 1996).

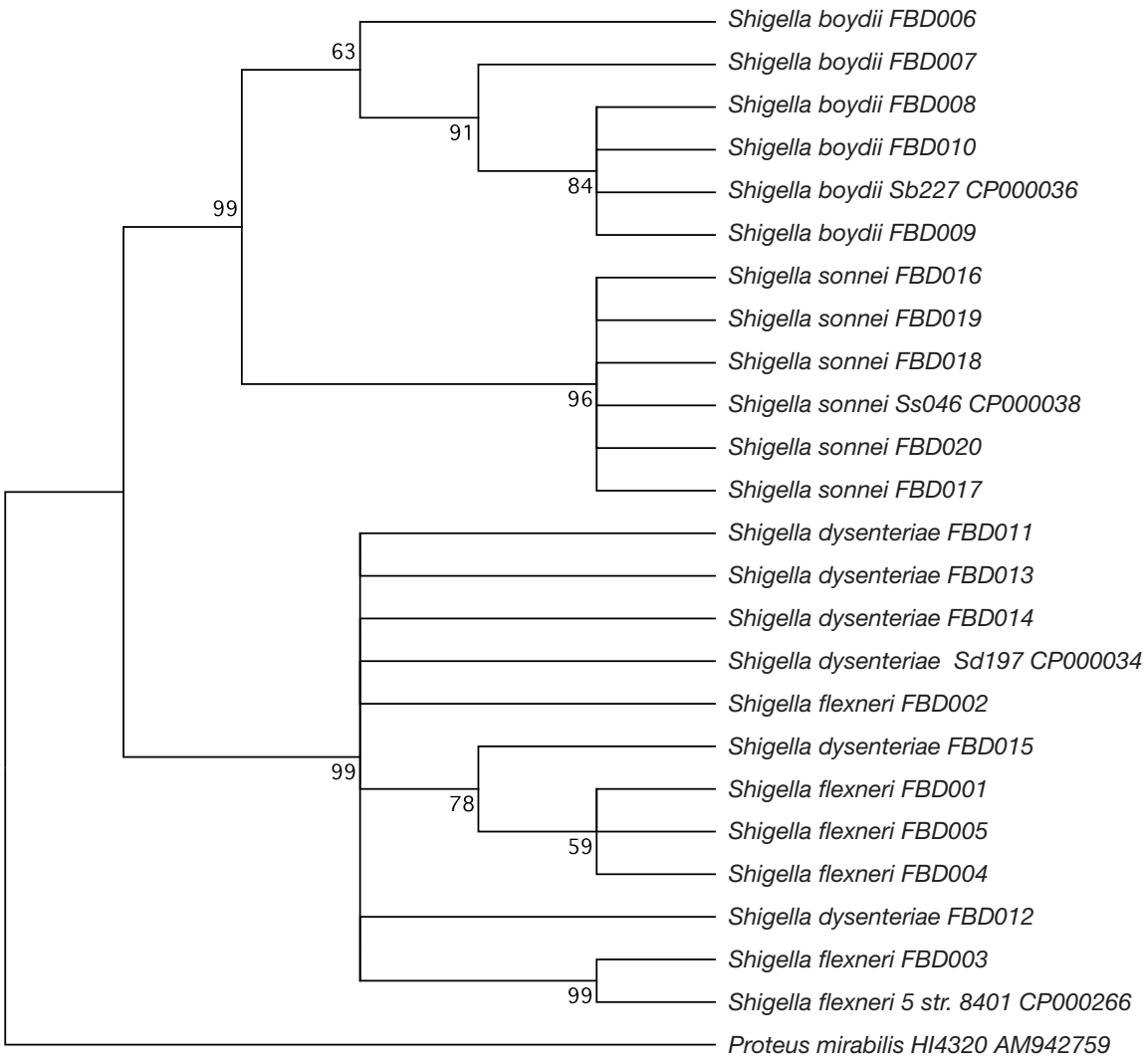

FIG. 2 - Phylogenetic tree based on the nucleotide sequences of 23S rRNA genes. The 23S rRNA sequences were adjusted to 2904 bases, and the tree was constructed by the Maximum Parsimony method, using the computer program MEGA3. The numbers shown next to the nodes indicate percent bootstrap values of the 1000 replicates. Genetic distances were computed using the Close-Neighbor-Interchange algorithm (Nei and Kumar, 2000). The sequence data for phylogenetic analysis were taken from the GenBank nucleotide sequence database for the following strains: *Shigella boydii* Sb227 (CP000036), *Shigella dysenteriae* Sd197 (CP000034), *Shigella flexneri* 5 str. 8401 (CP000266) and *Shigella sonnei* Ss046 (CP000038). *Proteus mirabilis* HI4320 (AM942759) was used as an outgroup.

In our phylogenetic-tree analysis of 16S and 23S rRNA, the percent divergences of *S. flexneri* from *S. dysenteriae* and *S. sonnei* from *S. boydii* were 0.5, 0.6, 0.2 and 0.4% respectively. These data are in accordance with the previous study using the 16S rRNA gene sequence (Wang *et al*., 1997; Fukushima *et al*., 2002) indicating close relationship among these species.

In Fig. 1 and 2 *S. flexneri* and *S. dysenteriae* are grouped together, while these phylogenetic trees included *S. sonnei* and *S. boydii* in one group. However, the results of the phylogenetic analysis of the twenty isolated *Shigella* based on 23S rRNA nucleotide sequences confirmed the 16S rRNA phylogenetic analysis, the divergence values between species are smaller in 16S rRNA analysis than in 23S rRNA and previous studies using several housekeeping and invasion genes (Boyd *et al*., 1994, 1997; Li *et al*., 1995; Nelson *et al*., 1991; Nelson and Selander, 1994).

In the phylogenetic-tree analysis of *gyrB*, the percent divergences of *S. boydii* from *S. dysenteriae* and *S. flexneri* were 0.7 and 0.8% respectively, also *S. dysenteriae* and *S. flexneri* had 0.8% divergence. The divergence values are significantly improved in *gyrB* analysis compared with 16S and 23S rRNA analysis which is in accordance with the previous study (Fukushima *et al*., 2002). It is reported that when 20 *Pseudomonas* strains were analyzed using the nucleotide sequences of *gyrB* and the genes for 16S rRNA and RNA polymerase σ^{70} factor (*rpoD*), the percent divergences of *gyrB* were larger than those of the other genes (Yamamoto and Harayama, 1998).

Comparisons of Fig. 1 and 2 with Fig. 3 highlight the different pattern of *Shigella* divergence which is determined that *S. sonnei* isolates are located in the same cluster; whereas, *S. boydii*, *S. flexneri* and *S. dysenteriae* are gathered in one group.

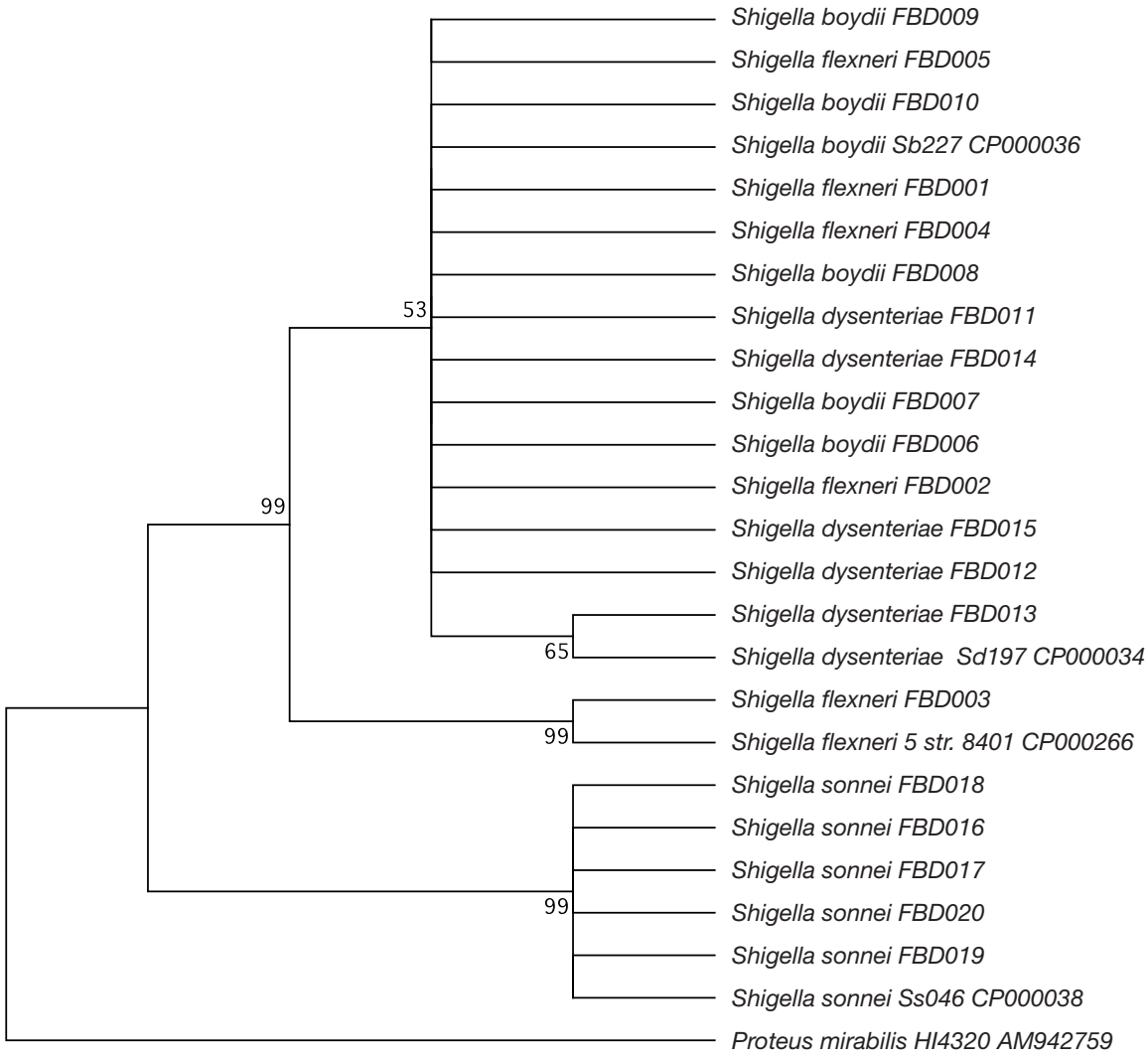

FIG. 3 - Phylogenetic tree based on the nucleotide sequences of *gyrB* genes. The *gyrB* sequences were adjusted to 1256 bases, and the tree was constructed by the Maximum Parsimony method, using the computer program MEGA3. The numbers shown next to the nodes indicate percent bootstrap values of the 1000 replicates. Genetic distances were computed using the Close-Neighbor-Interchange algorithm (Nei and Kumar, 2000). The sequence data for phylogenetic analysis were taken from the GenBank nucleotide sequence database for the following strains: *Shigella boydii* Sb227 (CP000036), *Shigella dysenteriae* Sd197 (CP000034), *Shigella flexneri* 5 str. 8401 (CP000266) and *Shigella sonnei* Ss046 (CP000038). *Proteus mirabilis* HI4320 (AM942759) was used as an outgroup.

Although the 16S rRNA gene is preferred for phylogenetic studies because of its attributes, such as little evidence for its lateral transfer (Sneath, 1993), the sequence variation of 16S rRNA is not as high as those of other genes in numerous studies. In addition, the rate of base substitution is greater for the *gyrB* sequence than for the 16S rRNA sequence (Dams *et al*., 1988; Yamamoto and Harayama, 1995, 1998) same as the results of this study.

The study shows that 16S and 23S rRNA sequence analyses are complementary because 23S rRNA was able to verify conclusions drawn from 16S rRNA analysis and it provided phylogenetic information at the species level; while, *gyrB* analysis generally shows more precise separations of *S. sonnei* isolates.

Acknowledgments

The authors are grateful to M. Azimi Rad, N. Hosseinkhan, M. Bolfion, M. Yaghoobi, P. Torabi, M. Tajbakhsh and M. Hamidian for their sincere cooperation and their assistance in the laboratory. Thanks are also due to M.M. Feizabadi, K. Azadmanesh and Pham Hong Nhung for helpful discussions.

REFERENCES

Anton A.I., Martinez-Murcia A.J., Rodriguez-Valera F. (1999). Intraspecific diversity of the 23S rRNA gene and the spacer region downstream in *Escherichia coli*. J. Bacteriol., 181: 2703-2709.

Bottger E.C. (1989). Rapid determination of bacterial ribosomal RNA sequences by direct sequencing of enzymatically amplified DNA. FEMS Microbiol. Lett., 53: 171-176.

Boyd E.F., Nelson K., Wang F.S., Whittam T.S., Selander R.K. (1994). Molecular genetic basis of allelic polymorphism in malate dehydrogenase (mdh) in natural populations of *Escherichia coli* and *Salmonella enterica*. Proc. Natl. Acad. Sci USA, 91: 1280-1284.

Boyd E.F., Li J., Ochman H., Selander R.K. (1997). Comparative genetics of the inv-spa invasion gene complex of *Salmonella enterica*. J. Bacteriol., 179: 1985-1991.

Christensen H., Nordentoft S., Olsen J.E. (1998). Phylogenetic relationships of *Salmonella* based on rRNA sequences. Int. J. Syst. Bacteriol., 48: 605-610.

Clarridge J.E., 3rd (2004). Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases. Clin. Microbiol. Rev., 17: 840-862.

Dams E., Hendriks L., Van De Peer Y., Neefs J.M., Smits G., Vandenbempt I., De Wachter R. (1988). Compilation of small ribosomal subunit RNA sequences. Nucleic Acids Res., 16 Suppl, r87: 173.

Dayhoff M.O. (1972). Atlas of Protein Sequence and Structure. National Biomedical Research Foundation, Silver Spring, Maryland.

Dewhirst F.E., Shen Z., Scimeca M.S., Stokes L.N., Boumenna T., Chen T., Paster B.J., Fox J.G. (2005). Discordant 16S and 23S rRNA gene phylogenies for the genus *Helicobacter*: implications for phylogenetic inference and systematics. J. Bacteriol., 187: 6106-6118.

Downes F.P., Itåo K., American Public Health Association (2001). Compendium of Methods for the Microbiological Examination of Foods, American Public Health Association, Washington, DC.

Escobar-Paramo P., Giudicelli C., Parsot C., Denamur E. (2003). The evolutionary history of *Shigella* and enteroinvasive *Escherichia coli* revised. J. Mol. Evol., 57: 140-148.

Fukushima M., Kakinuma K., Kawaguchi R. (2002). Phylogenetic analysis of *Salmonella, Shigella*, and *Escherichia coli* strains on the basis of the gyrB gene sequence. J. Clin. Microbiol., 40: 2779-2785.

Hale T.L. (1991). Genetic basis of virulence in *Shigella* species. Microbiol. Rev., 55: 206-224.

Huang W.M. (1996). Bacterial diversity based on type II DNA topoisomerase genes. Annu. Rev. Genet., 30: 79-107.

Kolbert C.P., Persing D.H. (1999). Ribosomal DNA sequencing as a tool for identification of bacterial pathogens. Curr. Opin. Microbiol., 2: 299-305.

Kumar S., Tamura K., Nei M. (2004). MEGA3: Integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. Brief Bioinform, 5: 150-163.

Leffers H., Kjems J., Ostergaard L., Larsen N., Garrett R.A. (1987). Evolutionary relationships amongst archaebacteria. A comparative study of 23 S ribosomal RNAs of a sulphur-dependent extreme thermophile, an extreme halophile and a thermophilic methanogen. J. Mol. Biol., 195: 43-61.

Li J., Ochman H., Groisman E.A., Boyd E.F., Solomon F., Nelson K., Selander R.K. (1995). Relationship between evolutionary rate and cellular location among the Inv/Spa invasion proteins of *Salmonella enterica*. Proc. Natl. Acad. Sci. USA, 92: 7252-7256.

Ludwig W., Schleifer K.H. (1994). Bacterial phylogeny based on 16S and 23S rRNA sequence analysis. FEMS Microbiol. Rev., 15: 155-173.

Murray P.R., Rosenthal K.S., Pfaller M.A. (2009). Medical Microbiology, Philadelphia, Mosby/Elsevier.

Nei M., Kumar S. (2000). Molecular Evolution and Phylogenetics. Oxford University Press, New York.

Nelson K., Selander R.K. (1994). Intergeneric transfer and recombination of the 6-phosphogluconate dehydrogenase gene (gnd) in enteric bacteria. Proc. Natl. Acad. Sci. USA, 91: 10227-10231.

Nelson K., Whittam T.S., Selander R.K. (1991). Nucleotide polymorphism and evolution in the glyceraldehyde-3-phosphate dehydrogenase gene (gapA) in natural populations of *Salmonella* and *Escherichia coli*. Proc. Natl. Acad. Sci. USA, 88: 6667-6671.

Palys T., Nakamura L.K., Cohan F.M. (1997). Discovery and classification of ecological diversity in the bacterial world: the role of DNA sequence data. Int. J. Syst. Bacteriol., 47: 1145-1156.

Pupo G.M., Lan R., Reeves P.R. (2000). Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics. Proc. Natl. Acad. Sci. USA, 97: 10567-10572.

Sambrook J., Russell D.W. (2001). Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.

Sethabutr O., Venkatesan M., Murphy G.S., Eampokalap B., Hoge C.W., Echeverria P. (1993). Detection of Shigellae and enteroinvasive *Escherichia coli* by amplification of the invasion plasmid antigen H DNA sequence in patients with dysentery. J. Infect. Dis., 167: 458-461.

Sneath P.H. (1993). Evidence from *Aeromonas* for genetic crossing-over in ribosomal sequences. Int. J. Syst. Bacteriol., 43: 626-629.

Sur D., Ramamurthy T., Deen J., Bhattacharya S.K. (2004). Shigellosis: challenges and management issues. Indian J. Med. Res., 120: 454-462.

Thompson J.D., Higgins D.G., 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 Res., 22: 4673-4680.

Tortoli E. (2003). Impact of genotypic studies on mycobacterial taxonomy: the new mycobacteria of the 1990s. Clin. Microbiol. Rev., 16: 319-354.

Wang R.F., Cao W.W., Cerniglia C.E. (1997). Phylogenetic analysis and identification of *Shigella* spp. by molecular probes. Mol. Cell Probes, 11: 427-432.

Woese C.R. (1987). Bacterial evolution. Microbiol. Rev., 51: 221-271.

Woese C.R., Stackebrandt E., Macke T.J., Fox G.E. (1985). A phylogenetic definition of the major eubacterial taxa. Syst. Appl. Microbiol., 6: 143-151.

Woese C.R., Kandler O., Wheelis M.L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. Proc. Natl. Acad. Sci. USA, 87: 4576-4579.

Yamamoto S., Harayama S. (1995). PCR amplification and direct sequencing of gyrB genes with universal primers and their application to the detection and taxonomic analysis of *Pseudomonas putida* strains. Appl. Environ. Microbiol., 61: 1104-1109.

Yamamoto S., Harayama S. (1996). Phylogenetic analysis of *Acinetobacter* strains based on the nucleotide sequences of gyrB genes and on the amino acid sequences of their products. Int. J. Syst. Bacteriol., 46: 506-511.

Yamamoto S., Harayama S. (1998). Phylogenetic relationships of *Pseudomonas putida* strains deduced from the nucleotide sequences of gyrB, rpoD and 16S rRNA genes. Int. J. Syst. Bacteriol., 48: 813-819.

Yang J., Nie H., Chen L., Zhang X., Yang F., Xu X., Zhu Y., Yu J., Jin Q. (2007). Revisiting the molecular evolutionary history of *Shigella* spp. J. Mol. Evol., 64: 71-79.