The Journal of Microbiology (2012) Vol. 50, No. 6, pp. 962–969 DOI 10.1007/s12275-012-2547-x
Copyright ⓒ 2012, The Microbiological Society of Korea

Molecular Serotyping of *Salmonella enterica* by Complete *rpoB* Gene Sequencing

Won-Jin Seong[1†], Hyuk-Joon Kwon[2†], Tae-Eun Kim[3], Deog-Yong Lee[4], Mi-Sun Park[4], and Jae-Hong Kim[1,2*]

[1]*Laboratory of Avian Diseases,* [2]*Zoonotic Disease Institute (ZooDI), College of Veterinary Medicine and BK21 for Veterinary Science, Seoul National University, Seoul 151-742, Republic of Korea*
[3]*BioPOA Co., Yongin 446-599, Republic of Korea*
[4]*Division of Enteric Bacterial Infections, Center for Infectious Diseases, Korea Centers for Disease Control and Prevention, Osong 363-951, Republic of Korea*

(Received October 8, 2012 / Accepted November 26, 2012)

Serotyping has been the gold standard for identifying *Salmonella*, but it requires large amounts of standard antisera. Multilocus sequence typing (MLST) has been applied to identify *Salmonella* serovars, but the recombination of 4–7 housekeeping genes and multiple analytic steps diminish its applicability. In the present study, we determined the complete sequences of the RNA polymerase beta subunit gene (*rpoB*) and 7 housekeeping genes (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*) for 76 strains of 33 *Salmonella enterica* serovars and conducted phylogenetic analyses together with the corresponding gene sequences of 24 reference strains registered in the GenBank database. Based on the phylogenetic analyses, 100 strains from 40 serovars and 91 strains from 37 serovars were classified into 60 *rpoB* (RST) and 49 multilocus sequence types (ST), respectively. The nucleotide similarities were 98.8–100% and 96.9–100% for the complete *rpoB* gene and the seven concatenated housekeeping genes, respectively. The strains of 35 and 30 serovars formed serovar-specific branches or clusters in the *rpoB* and housekeeping gene phylogenetic trees, respectively. Therefore, complete *rpoB* gene sequencing and phylogenetic analysis may be a useful method for identifying *Salmonella* serovars that is a simpler, more cost-effective, and less time-consuming alternative or complementary method to MLST and conventional serotyping.

***Keywords*:** *Salmonella*, serotyping, *rpoB*, MLST, phylogenetic analysis

Introduction

Salmonella is one of the most frequent causative agents of food-borne illness, and the appearance of multidrug-resistant *Salmonella* serovars has become a threat to global public health. Thus, worldwide surveillance and serotyping of *Salmonella* have been conducted (Herikstad *et al.*, 2002). The genus *Salmonella* is composed of two species, *Salmonella enterica* and *Salmonella bongori*, and *S. enterica* is further divided into the six subspecies *S. enterica* subsp. *enterica*, *S. enterica* subsp. *salamae*, *S. enterica* subsp. *arizonae*, *S. enterica* subsp. *diarizonae*, *S. enterica* subsp. *houtenae*, and *S. enterica* subsp. *indica* based on phenotypic and genetic characteristics (Reeves *et al.*, 1989; Grimont and Weill, 2007). Based on the White-Kaufmann-Le Minor (WKL) scheme, *Salmonella* consists of 2,579 serovars, and large volumes of antisera are required for the typing of somatic (O), phase I and II flagellin (H), and capsular antigens (Grimont and Weill, 2007). Therefore, the complexity of the serotyping scheme, the preparation of a large number of standard antisera and antigen samples, the unavailability or variable quality of commercial antisera, and the time-consuming procedures in some serovars are disadvantages of the conventional serotyping method (Hendriksen *et al.*, 2009).

For these reasons, several molecular methods, such as the automated ribotyping system, 16S–23S rRNA spacer restriction fragment length polymorphism (RFLP), polymerase chain reaction (PCR)-RFLP, sequencing of H antigen genes (*fliC* and *fljB*), multiplex PCR of the O-antigen encoding genes (*tyv* and *wzx*), and phylogenetic analyses of the beta subunit of ATP synthase gene (*atpD*) have been developed for typing *Salmonella* serovars, but their discriminative ability is insufficient for practical applications (Christensen and Olsen, 1998; Dauga *et al.*, 1998; Oscar, 1998; Christensen *et al.*, 2000; Herrera-Leon *et al.*, 2007; Hendriksen *et al.*, 2009). Multilocus sequence typing (MLST) rather than multilocus enzyme electrophoresis has been developed to investigate the global epidemiology and population structure of infectious agents (Maiden *et al.*, 1998; Enright and Spratt, 1999). To date, several MLST schemes using different sets of housekeeping and virulence genes have been developed for phylogenetic analyses of *Salmonella* serovars, and MLST data have accumulated (Kidgell *et al.*, 2002; Kotetishvili *et al.*, 2002; Fakhr *et al.*, 2005; Sukhnanand *et al.*, 2005; Tankouo-Sandjong *et al.*, 2007; Achtman *et al.*, 2012). However, MLST analyses require the amplification and purification of 4–7 housekeeping genes and 8–14 sequencing reactions for each sample. Furthermore, relatively high recombination rates of the housekeeping genes can blur the interpretation of phylogenetic results (Kotetishvili *et al.*, 2002; Sukhnanand *et al.*, 2005; Torpdahl *et al.*, 2005; Falush *et al.*, 2006). Therefore, simpler and more economical phylogenetic schemes with high discriminative power that are free of recombination

[†]These authors contributed equally to this work.
*For correspondence. E-mail: kimhong@snu.ac.kr; Tel.: +82-2-880-1250; Fax: +82-2-885-6614

Table 1. ***S. enterica*** **strains used in this study and their multilocus sequence types (STs) and** ***rpoB*** **sequence types (RSTs)**

Strain	Serovar	Source	Accession No.	ST	RST
CVM29188	Kentucky	GenBank	ABAK_02000001	152	9-1
SNUS6D10	Kentucky	SNU	JQ728831	314	11-1
C1	Newport	KCDC	JQ728834	166	12-1
S50	Newport	KCDC	JQ728835	31	12-2
E4	London	KCDC	JQ728836	155	12-2
E5	London	KCDC	JQ728837	155	12-2
B13	Derby	KCDC	JQ728833	71	12-3
S3	Paratyphi B	KCDC	JQ728832	86	12-4
SPB7	Paratyphi B	GenBank	NC_010102	307	12-4
C2	Blockley	KCDC	JQ728838	52	12-5
C21	Blockley	KCDC	JQ728839	52	12-5
E1	Senftenberg	KCDC	JQ728840	14	12-6
E2	Senftenberg	KCDC	JQ728841	14	12-6
CT_02021853	Dublin	GenBank	NC_011205	10	12-7
C15	Thompson	KCDC	JQ728843	26	13-1
S40	Thompson	KCDC	JQ728842	26	13-1
SNUS7F	Thompson	SNU	JQ728844	26	13-1
C22	Newport	KCDC	JQ728845	166	13-2
SL254	Newport	GenBank	NC_011080	45	13-3
C7	Infantis	KCDC	JQ728846	32	13-4
D25	Dublin	KCDC	JQ728847	10	13-5
CDC07-0191	Tennesee	GenBank	NZ_ACBF01000033	nd*	13-6
C10	Rissen	KCDC	JQ728848	469	14-1
C11	Rissen	KCDC	JQ728849	469	14-1
D1	Enteritidis	KCDC	JQ728851	11	15-1
D5	Enteritidis	KCDC	JQ728852	11	15-1
D7	Enteritidis	KCDC	JQ728853	11	15-1
D8	Enteritidis	KCDC	JQ728854	11	15-1
D9	Enteritidis	KCDC	JQ728855	11	15-1
D10	Enteritidis	KCDC	JQ728856	11	15-1
D24	Enteritidis	KCDC	JQ728857	11	15-1
P125109	Enteritidis	GenBank	NC_011294	11	15-1
SNU1075/3	Enteritidis	SNU	JQ728858	11	15-1
SL483	Agona	GenBank	NC_011149	13	15-2
SNUS3D13	Agona	SNU	JQ728850	13	15-2
SL317	Newport	GenBank	NZ_ABEW01000017	nd	15-3
C19	Hadar	KCDC	JQ728863	33	16-1
RI_05P066	Hadar	GenBank	NZ_ABFG01000019	nd	16-1
B15	Agona	KCDC	JQ728859	13	16-2
C5	Braenderup	KCDC	JQ728862	22	16-3
E8	Anatum	KCDC	JQ728864	64	16-4
C4	Rissen	KCDC	JQ728861	469	16-5
C3	Litchfield	KCDC	JQ728860	214	16-6
B7	Typhimurium	KCDC	JQ728865	19	17-1
B21	Typhimurium	KCDC	JQ728866	34	17-2
B22	Typhimurium	KCDC	JQ728867	34	17-2
SARA23	Saintpaul	GenBank	ABAM_02000001	50	17-3
B4	Typhimurium	KCDC	JQ728869	19	18-1
B6	Typhimurium	KCDC	JQ728870	19	18-1
B8	Typhimurium	KCDC	JQ728871	19	18-1
B10	Typhimurium	KCDC	JQ728872	19	18-1
SNU1083/2	Typhimurium	SNU	JQ728873	19	18-1
SNU1090/6C	Typhimurium	SNU	JQ728874	19	18-1
BS240	Typhimurium	BS	JQ728875	19	18-1
S1	Paratyphi A	KCDC	JQ728868	85	18-2
RKS4594	Paratyphi C	GenBank	NC_012125	114	18-3

Table 1. Continued

Strain	Serovar	Source	Accession No.	ST	RST
C16	Mbandaka	KCDC	JQ728877	413	18-4
C9	Hartford	KCDC	JQ728876	405	18-5
287/91	Gallinarum	GenBank	NC_011274	331	18-6
B5	Typhimurium	KCDC	JQ728879	19	19-1
D13	Typhi	KCDC	JQ728880	1	19-2
D22	Typhi	KCDC	JQ728881	1	19-2
D23	Typhi	KCDC	JQ728882	1	19-2
Ty2	Typhi	GenBank	NC_004631	1	19-2
D14	Typhi	KCDC	JQ728883	2	19-2
D15	Typhi	KCDC	JQ728884	2	19-2
D16	Typhi	KCDC	JQ728885	2	19-2
D18	Typhi	KCDC	JQ728886	2	19-2
A3	Paratyphi A	KCDC	JQ728878	85	19-3
AKU_12601	Paratyphi A	GenBank	NC_011147	nd	19-3
ATCC 9150	Paratyphi A	GenBank	NC_006511	85	19-3
SP12	Pullorum	GenBank	FJ797633	nd	19-4
B1	Typhimurium	KCDC	JQ728888	19	20-1
LT2	Typhimurium	GenBank	NC_003197	19	20-2
CT18	Typhi	GenBank	NC_003198	2	20-3
S57	Typhi	KCDC	JQ728890	2	20-4
B16	Paratyphi B var java	KCDC	JQ728887	28	20-5
C18	Ohio	KCDC	JQ728889	329	20-6
C12	Virchow	KCDC	JQ728891	16	21-1
SL491	Virchow	GenBank	NZ_ABFH02000001	nd	21-1
E7	Weltevreden	KCDC	JQ728892	365	21-2
HI_N05-537	Weltevreden	GenBank	NZ_ABFF01000047	nd	21-2
S76	London	KCDC	JQ728893	504	24-1
C8	Montevideo	KCDC	JQ728895	4	25-1
BS291	Montevideo	BS	JQ728896	4	25-1
SNUS10B	Montevideo	SNU	JQ728897	4	25-1
B23	Brandenburg	KCDC	JQ728894	65	25-2
SNUS15B	Oranienburg	SNU	JQ728898	23	25-3
S87	Senftenberg	KCDC	JQ728899	290	25-4
GA_MM04042433	Javiana	GenBank	ABEH02000013	nd	25-5
SL476	Heidelberg	GenBank	NC_011083	15	26-1
SL486	Heidelberg	GenBank	NZ_ABEL01000038	nd	26-1
B14	Derby	KCDC	JQ728900	40	26-2
D26	Panama	KCDC	JQ728902	48	26-3
B19	Schwarzengrund	KCDC	JQ728901	96	26-4
CVM19633	Schwarzengrund	GenBank	NC_011094	322	27-1
C13	Bareilly	KCDC	JQ728903	362	28-1
SNUS12D5	Bareilly	SNU	JQ728904	203	28-1
E9	Give	KCDC	JQ728905	516	28-2
B24	Schleissheim	KCDC	JQ728906	280	30-1

* nd, not determined.

bias are preferable.

Among the housekeeping genes, the partial or complete RNA polymerase beta subunit gene (*rpoB*) has been applied to phylogenetic analyses of various bacteria and to distinguish serovars of *S. enterica* (Mollet *et al.*, 1997; Kim *et al.*, 1999, 2003; Kwon *et al.*, 2001; Korczak *et al.*, 2004; La Scola *et al.*, 2006; Marianelli *et al.*, 2006; Adékambi *et al.*, 2008, 2009). In previous studies, partial *rpoB* sequencing and phylogenetic analysis were commonly used to identify pathogenic bacteria only at the species level but could not distinguish some *S. enterica* serovars (Kwon *et al.*, 2001). In the present study, we determined the complete *rpoB* sequences and MLST of 76 strains of 33 *S. enterica* serovars and conducted phylogenetic analyses together with 24 reference strains registered in the GenBank database.

Materials and Methods

Bacterial strains

The 76 strains of 33 *S. enterica* serovars were collected from the National Salmonella Reference Laboratory at the Division

of Enteric Bacterial Infections (Center for Infectious Diseases, Korea Centers for Disease Control and Prevention, KCDC), the Laboratory of Avian Diseases, College of Veterinary Medicine, Seoul National University (SNU), and Bansuk livestock technical center (BS). The 24 reference strains were selected from the GenBank database (Benson *et al.*, 2011) for the phylogenetic analyses of the complete *rpoB* gene (24 strains) and *in silico* MLST (15 strains) (Table 1).

Serotyping

The 65 strains of *S. enterica* serovars from KCDC were identified using the WKL scheme (Grimont and Weill, 2007), and the O-antigens of the remaining 11 strains from SNU and BS were determined using commercial antisera (BD Difco, USA) according to the manufacturer's instructions.

Primers

The upstream and downstream regions of the *rpoB* genes from the reference strains were compared using the Bioedit software (ver. 5.0.9.1.) (Hall, 1999). A primer set to amplify the complete rpoB coding region and seven sequencing primers were designed. The primer sequences will be provided to researchers by request.

PCR and sequencing

The *Salmonella* strains were grown overnight in tryptic soy broth at 37°C, and genomic DNA was extracted using a genomic DNA extraction module in a *Sal*-ID kit (BioPOA, Korea), according to the manufacturer's instructions. The *rpoB* PCR reaction was composed of 5 μl of 10× buffer, dNTPs (10 mM, 1 μl), forward and reverse primers (10 pmol/μl, 1 μl of each), Taq DNA polymerase (5 U/μl; Macrogen, Korea; 0.5 μl), distilled water (40.5 μl), and template DNA (50 ng/μl, 1 μl). The mixture was incubated at 94°C for 5 min; 40 cycles at 94°C for 30 sec, 50°C for 30 sec, and 72°C for 4 min; and at 72°C for 5 min. The PCR and purification of the amplicons for the seven MLST genes (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*) were performed as described previously (Kidgell *et al.*, 2002; Achtman *et al.*, 2012). The PCR amplicons were purified using the purification module in the *Sal*-ID kit and sequenced with the sequencing primers in the *Sal*-ID kit using an ABI3711 automatic sequencer (Macrogen).

Sequence analysis

The overlapping sequences of *rpoB* and the seven housekeeping genes were assembled into single complete sequences using ChromasPro version 1.5 (Technelysium Pty Ltd., Australia). The seven housekeeping gene sequences were aligned with allele templates from the MLST database (http://mlst.ucc.ie/mlst/dbs/Senterica) and trimmed using the BioEdit software. Then, the sequences were queried to the MLST website for allele number assignment. The *in silico* MLST was performed by querying the retrieved housekeeping genes of 15 reference strains in the MLST database. Nucleotide similarity, variable nucleotide comparisons, and translation of the *rpoB* nucleotides were performed with Bioedit. Phylogenetic analyses with the complete *rpoB* gene and seven concatenated housekeeping genes were conducted using the MEGA software (ver. 5.0.5, neighbor-joining method with Tamura-Nei distance and 1000 bootstrapping replicates) (Saitou and Nei, 1987; Tamura *et al.*, 2011).

Results

Complete *rpoB* sequences of the *Salmonella* strains

The complete coding region of *Salmonella rpoB* was successfully amplified by the PCR primer set. The complete coding region of *Salmonella rpoB* was 4,029 nucleotides long including the stop codon. The nucleotide similarities between *S. enterica* serovars were 98.4–100%. The number of variable nucleotide loci was 198 (4.9% of the complete sequence), and synonymous and nonsynonymous nucleotide changes were 181 and 17, respectively. Among the tested *S. enterica* strains, SNUS6D10 (serovar Kentucky), C22 (serovar Newport), D25 (serovar Dublin), B15 (serovar Agona), C4 (serovar Rissen), B7 (serovar Typhimurium), B4 (serovar Typhimurium), B6 (serovar Typhimurium), B8 (serovar Typhimurium), B10 (serovar Typhimurium), SNU1083/2 (serovar Typhimurium), SNU1090/6C (serovar Typhimurium), BS240 (serovar Typhimurium), SP12 (serovar Gallinarum biovar Pullorum), LT2 (serovar Typhimurium), CT18 (serovar Typhi), S57 (serovar Typhi), and B23 (serovar Brandenburg) had one amino acid changes, whereas 287/91 (serovar Gallinarum biovar Gallinarum), B5 (serovar Typhimurium) and C13 (serovar Bareilly) had two and B1 (serovar Typhimurium) had three amino acid changes (Table 2).

The MLST and phylogenetic analyses of *S. enterica* strain serovars

The length of the seven concatenated housekeeping genes was 3,336 bp, and the nucleotide sequence variabilities of the 91 strains of *S. enterica* serovars were 96.9–100%. The number of variable nucleotide loci was 264 (7.9% of the complete sequence) in the concatenated genes. The 91

Table 2. ***RpoB* amino acid changes in *S. enterica* serovars**

Amino acid changes	*S. enterica* strain (serovar)
D222V	C22 (serovar Newport)
R247L	D25 (serovar Dublin)
D281Y	SP12 (serovar Gallinarum biovar Pullorum)
A312P	SNUS6D10 (serovar Kentucky)
K422E	B15 (serovar Agona)
S917F	CT18 (serovar Typhi)
G982S	B7 (serovar Typhimurium), B4 (serovar Typhimurium), B6 (serovar Typhimurium), B8 (serovar Typhimurium), B10 (serovar Typhimurium), SNU1083/2 (serovar Typhimurium), SNU1090/6C (serovar Typhimurium), BS240 (serovar Typhimurium), LT2 (serovar Typhimurium)
T1163S	B23 (serovar Brandenburg)
K1178I	S57 (serovar Typhi)
D1341N	C4 (serovars Rissen)
Y47C, Q1038K	287/91 (serovar Gallinarum biovar Gallinarum)
Q798H, G982S	B5 (serovar Typhimurium)
I873L, G907R	C13 (serovar Bareilly)
T207P, I208F, G982S	B1 (serovar Typhimurium)

strains of *S. enterica* serovars were classified into 49 STs, and the serovars of 11 strains of *Salmonella* from SNU and BS were predicted based on their STs (Table 1).

The strains of Agona, Blockley, Dublin, Enteritidis, Montevideo, Paratyphi A, Rissen, and Thompson serovars were classified into single STs of 13, 52, 10, 11, 4, 85, 469, and 26, respectively, which formed separate clusters in the phylogenetic tree constructed with the seven concatenated housekeeping gene sequences (Fig. 1). The serovars Schwarzengrund (two strains in two STs, 96 and 322), Typhi (10 strains in two STs, 1 and 2), and Typhimurium (13 strains in two STs, 19 and 34) contained multiple STs (Table 1), but all strains of each serovar clustered together in the MLST tree (Fig. 1).

In contrast, seven serovars, including Bareilly (two strains in two STs, 203 and 362), Derby (two strains in two STs, 40 and 71), Kentucky (two strains in two STs, 152 and 314), London (three strains in two STs, 155 and 504), Newport (four strains in three STs, 31, 45, and 166), Paratyphi B (three strains in three STs, 28, 86, and 307), and Senftenberg (three strains in two STs, 14 and 290) contained multiple STs that formed separate branches or clusters in the MLST tree (Fig. 1). The numbers of variable nucleotides between strains of the same serovar ranged from 0 to 3, but the numbers of different nucleotides between B13 (serovar Derby) and B14 (serovar Derby), C1 (serovar Newport) and S50 (serovar Newport), C1 (serovar Newport) and SL254 (serovar Newport), SL254 (serovar Newport) and S50 (serovar Newport), C13 (serovar Bareilly) and SNUS12D5 (serovar Bareilly), S87 (serovar Senftenberg) and E1 (serovar Senftenberg), S3(serovar Paratyphi B) and B16 (serovar Paratyphi B), E4 (serovar London) and S76 (serovar London), and CVM29188 (serovar Kentucky) and SNUS6D10 (serovar Kentucky) were 38, 40, 33, 17, 49, 43, 31, 41, and 32, respectively.

In the case of the serovars Anatum, Braenderup, Brandenburg, Gallinarum, Give, Hadar, Hartford, Heidelberg, Infantis, Litchfield, Mbandaka, Ohio, Oranienburg, Panama, Paratyphi C, Saintpaul, Schleissheim, Virchow, and Weltevreden, only one strain for each serovar was available, and they were classified into STs 64, 22, 65, 331, 516, 33, 15, 32, 214, 413, 329, 23, 48, 114, 50, 280, 16, and 365, respectively, which formed separate branches in the MLST tree (Fig. 1).

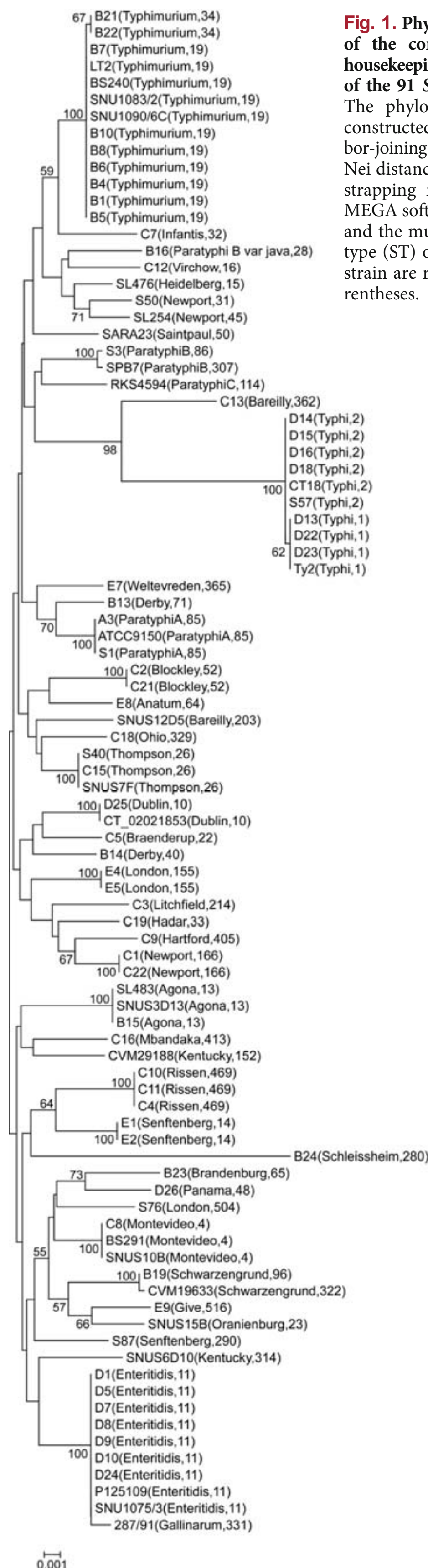

Fig. 1. Phylogenetic analysis of the concatenated seven housekeeping gene sequences of the 91 *S. enterica* strains. The phylogenetic tree was constructed using the neighbor-joining method (Tamura-Nei distance and 1,000 bootstrapping replicates) in the MEGA software. The serovar and the multilocus sequence type (ST) of each *S. enterica* strain are represented in parentheses.

The *rpoB* gene sequence typing and phylogenetic analysis of strains of *S. enterica* serovars

The consensus *rpoB* gene sequence was generated based on the most common nucleotide sequences of the 100 strains of *S. enterica* serovars. The complete *rpoB* gene sequences of 100 strains of 40 serovars were compared with the consensus sequence, and they were classified into 60 *rpoB* sequence types (RSTs) (Table 1). The RST is represented by two numbers separated by a hyphen. The first and the second numbers designate the number of different nucleotides from the consensus sequence and the identifier of the different nucleotide sequences, respectively. The strains of each serovar, including Bareilly, Blockley, Enteritidis, Hadar, Heidelberg, Montevideo, Thompson, Virchow, and Weltevreden, had identical nucleotide sequences to each other and were classified into the unique RSTs of 28-1, 12-5, 15-1, 16-1,

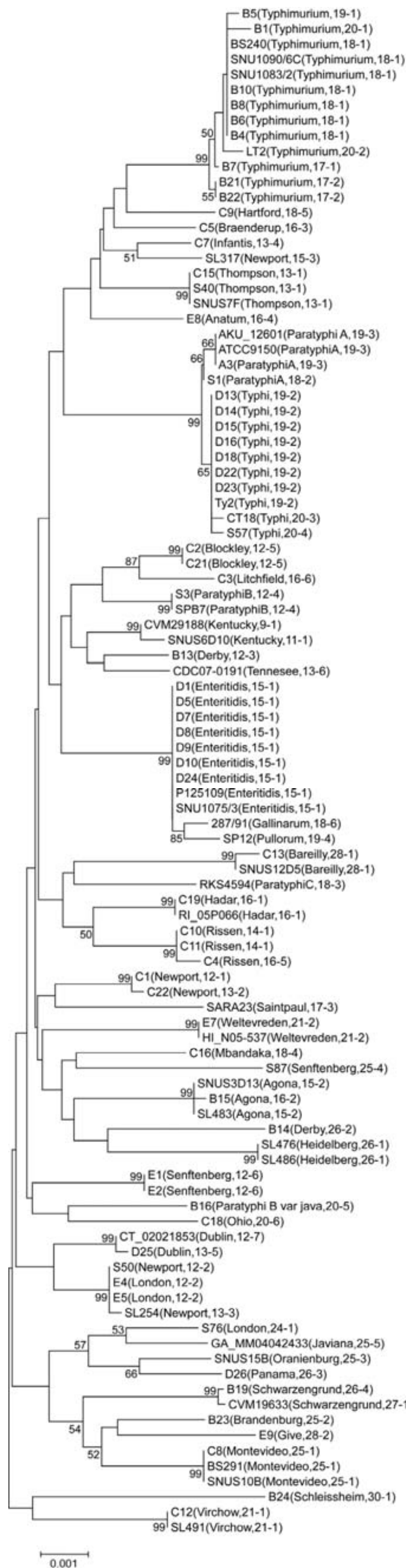

Fig. 2. Phylogenetic analysis of the complete RNA polymerase beta subunit gene (*rpoB*) sequence of 100 *S. enterica* strains. The phylogenetic tree was constructed using the neighbor-joining method (Tamura-Nei distance and 1,000 bootstrapping replicates) in the MEGA software. The serovar and the *rpoB* gene sequence type (RST) of each *S. enterica* strain are represented in parentheses.

26-1, 25-1, 13-1, 21-1, and 21-2, respectively, which formed separate clusters in the phylogenetic tree constructed with the complete *rpoB* nucleotide sequences (Fig. 2). The serovars Agona (three strains in two RSTs, 15-2 and 16-2), Dublin (two strains in two RSTs, 12-7 and 13-5), Kentucky (two strains in two RSTs, 9-1 and 11-1), Paratyphi A (four strains in two RSTs, 18-2 and 19-3), Rissen (three strains in two RSTs, 14-1 and 16-5), Schwarzengrund (two strains in two RSTs, 26-4 and 27-1), Typhi (10 strains in three RSTs; 19-2, 20-3, and 20-4), and Typhimurium (13 strains in six RSTs; 17-1, 17-2, 18-1, 19-1, 20-1, and 20-2) contained multiple RSTs (Table 1), but all strains of each serovar clustered together in the *rpoB* tree (Fig. 2).

In contrast, the serovars Derby (two strains in two RSTs, 12-3 and 26-2), London (three strains in two RSTs; 12-2 and 24-1), Newport (five strains in five RSTs; 12-1, 12-2, 13-2, 13-3, and 15-3), Paratyphi B (three strains in two RSTs, 12-4 and 20-5), and Senftenberg (three strains in two RSTs, 12-6 and 25-4) contained different RSTs that formed separate branches or clusters in the phylogenetic tree (Fig. 2). The numbers of variable nucleotides between strains of the same serovar ranged from 0 to 7, but the numbers of different nucleotides between B13 (serovar Derby) and B14 (serovar Derby), SL317 (serovar Newport), C1 (serovar Newport), and S50 (serovar Newport), S87 (serovar Senftenberg) and E1 (serovar Senftenberg), S3(serovar Paratyphi B), and B16 (serovar Paratyphi B), and E4 (serovar London) and S76 (serovar London) were 36, 29, 27, 32, and 20, respectively.

In the case of the serovars Anatum, Braenderup, Brandenburg, Give, Hartford, Infantis, Javiana, Litchfield, Mbandaka, Ohio, Oranienburg, Panama, Paratyphi C, Saintpaul, Schleissheim, and Tennessee, only one strain for each serovar was available, but they were classified into unique RSTs, including 16-4, 16-3, 25-2, 28-2, 18-5, 13-4, 25-5, 16-6, 18-4, 20-6, 25-3, 26-3,18-3, 17-3, 30-1, and 13-6, respectively, and formed separate branches in the *rpoB* tree (Fig. 2). The complete *rpoB* sequences of several strains of serovar Gallinarum biovar Gallinarum and biovar Pullorum were available in the GenBank database. The biovars Gallinarum and Pullorum were classified into RSTs 18-6 and 19-4, respectively, but they formed a cluster together with serovar Enteritidis (Fig. 2).

Discussion

Various molecular methods targeting various genes have been applied to replace the conventional serotyping method for *S. enterica* serovars, but their discriminative ability was insufficient to distinguish different *Salmonella* serovars. Pulsed field gel electrophoresis and amplified fragment length polymorphism have been applied to type *S. enterica* strains due to their high discriminative abilities, but too much genetic diversity of strains in the same serovar was a drawback of serotyping (Torpdahl *et al.*, 2005; Harbottle *et al.*, 2006). MLST might be more suitable for typing *Salmonella* serovars because of its high discriminative ability, and the MLST database has been increasing (Torpdahl *et al.*, 2005; Harbottle *et al.*, 2006; Achtman *et al.*, 2012). MLST is costly and requires multiple steps and reactions for analysis (Kidgell *et al.*,

2002; Kotetishvili *et al*., 2002; Fakhr *et al*., 2005; Sukhnanand *et al*., 2005; Tankouo-Sandjong *et al*., 2007).

In previous studies, partial or complete *rpoB* genes have been used for species classification and identification of clinical bacteria, and the partial hypervariable region (612–937 bp) of the *rpoB* gene has been applied to distinguish *S. enterica* serovars (Mollet *et al*., 1997; Kwon *et al*., 2001; Adékambi *et al*., 2008, 2009). Partial *rpoB* gene sequencing has been suggested to be able to identify clinical bacteria based on simplicity and cost-effectiveness, but partial *rpoB* gene sequencing failed to differentiate serovar Gallinarum from Enteritidis, serovar Hadar from Newport, serovar Schotmalleri from Thomson, and serovar Paratyphi A from Typhi (Kwon *et al*., 2001; Adékambi *et al*., 2009). In the present study, complete *rpoB* gene sequencing and MLST could successfully differentiate serovar Gallinarum from Enteritidis, serovar Hadar from Newport, and serovar Paratyphi A from Typhi, with the exception of serovar Schotmalleri from Thomson due to the absence of the Schotmalleri serovar strain.

Most of the same serovar strains had similar nucleotide sequences and clustered together in the phylogenetic trees constructed using the *rpoB* and seven concatenated housekeeping genes, but the clustering patterns of the *rpoB* tree were slightly different from those of the MLST tree. The strains of serovar Typhimurium clustered together in the *rpoB* and MLST trees, but *rpoB* was more variable than the seven housekeeping genes, six RSTs vs two STs. Additionally, the strains of serovar Typhi clustered together in both trees but they were composed of more variable RSTs than STs, three RSTs vs. two STs. The serovars Typhi and Paratyphi A formed a cluster in the *rpoB* tree, but serovar Typhi formed a cluster with the C13 strain of serovar Bareilly. Considering human restrictions and that a systemic infection is caused only by serovars Typhi and Paratyphi A, which are genetically similar, the *rpoB* tree may be more informative than that of the MLST tree (McClelland *et al*., 2004). In the same context, the clustering of serovar Enteritidis which is most frequent cause of poultry products-related food-poisoning in human with avian-restricted and systemic infection-causing serovar Gallinarum biovar Gallinarum, and biovar Pullorum may give a hint of its high prevalence and, to some extent, its virulence in poultry (Hopper and Mawer, 1988; Suzuki, 1994).

Some strains of serovars Derby, London, Newport, Paratyphi B, and Senftenberg had apparently different nucleotide sequences and formed different branches and clusters in the *rpoB* and MLST trees. In contrast to the *rpoB* tree, the strains of serovars Bareilly and Kentucky did not form a cluster in the MLST tree. The presence of several distinct STs and RSTs in some serovars might reflect convergent evolution of the serovars but indicates that a more extended database of the complete *rpoB* gene or the seven housekeeping genes to cover all genetic variability of *S. enterica* serovars is required for nucleotide sequence-based serotyping. In addition, the conventional serotyping method is difficult to perform correctly; in practice, only 33% of the EQAS (External Quality Assurance System)-participating laboratories serotyped all eight strains correctly in 2008 (Hendriksen *et al*., 2009). Therefore, a complete *rpoB* gene or seven housekeeping gene database should be constructed, using a sufficient number of confirmed reference *Salmonella* serovars under collaboration with the International and National *Salmonella* Reference Laboratories.

In conclusion, the phylogenetic analysis of complete *rpoB* gene sequence in the present study employed simpler, more cost-effective, and less time-consuming procedures and might be more informative than that of MLST. Therefore, this procedure can be a promising alternative or complementary method to MLST and conventional serotyping for identifying *S. enterica* serovars in humans and animals.

Acknowledgements

This research was supported by iPET (Korea Institute of Planning and Evaluation for Technology in Food, Agriculture, Forestry and Fisheries), Ministry for Food, Agriculture, Forestry and Fisheries, Republic of Korea.

References

Achtman, M., Wain, J., Weill, F.X., Nair, S., Zhou, Z., Sangal, V., Krauland, M.G., Hale, J.L., Harbottle, H., Uesbeck, A., Dougan, G., Harrison, L.H., Brisse, S., and S.E.M.S. Group. 2012. Multilocus sequence typing as a replacement for serotyping in *Salmonella enterica*. *PLoS Pathog.* **8**, e1002776.

Adékambi, T., Drancourt, M., and Raoult, D. 2009. The *rpo*B gene as a tool for clinical microbiologists. *Trends Microbiol.* **17**, 37–45.

Adékambi, T., Shinnick, T.M., Raoult, D., and Drancourt, M. 2008. Complete *rpoB* gene sequencing as a suitable supplement to DNA-DNA hybridization for bacterial species and genus delineation. *Int. J. Syst. Evol. Microbiol.* **58**, 1807–1814.

Benson, D.A., Karsch-Mizrachi, I., Lipman, D.J., Ostell, J., and Sayers, E.W. 2011. GenBank. *Nucleic Acids Res.* **39**, D32–37.

Christensen, H., Moller, P.L., Vogensen, F.K., and Olsen, J.E. 2000. 16S to 23S rRNA spacer fragment length polymorphism of *Salmonella enterica* at subspecies and serotype levels. *J. Appl. Microbiol.* **89**, 130–136.

Christensen, H. and Olsen, J.E. 1998. Phylogenetic relationships of *Salmonella* based on DNA sequence comparison of *atpD* encoding the beta subunit of ATP synthase. *FEMS Microbiol. Lett.* **161**, 89–96.

Dauga, C., Zabrovskaia, A., and Grimont, P.A. 1998. Restriction fragment length polymorphism analysis of some flagellin genes of *Salmonella enterica*. *J. Clin. Microbiol.* **36**, 2835–2843.

Enright, M.C. and Spratt, B.G. 1999. Multilocus sequence typing. *Trends Microbiol.* **7**, 482–487.

Fakhr, M.K., Nolan, L.K., and Logue, C.M. 2005. Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing *Salmonella enterica* serovar Typhimurium. *J. Clin. Microbiol.* **43**, 2215–2219.

Falush, D., Torpdahl, M., Didelot, X., Conrad, D.F., Wilson, D.J., and Achtman, M. 2006. Mismatch induced speciation in *Salmonella*: model and data. *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* **361**, 2045–2053.

Grimont, P.A.D. and Weill, F.X. 2007. Antigenic formulae of the *Salmonella* serovars. WHO Collaborating Centre for Reference and Research on Salmonella, Institut Pasteur, Paris, France.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symp. Ser.* **41**, 95–98.

Harbottle, H., White, D.G., McDermott, P.F., Walker, R.D., and

Zhao, S. 2006. Comparison of multilocus sequence typing, pulsed-field gel electrophoresis, and antimicrobial susceptibility typing for characterization of *Salmonella enterica* serotype Newport isolates. *J. Clin. Microbiol.* **44**, 2449–2457.

Hendriksen, R.S., Karlsmose, S., Jensen, A.B., and Aarestrup, F.M. 2009. The External Quality Assurance System of the WHO Global Salm-Surv Year 2008, 2nd ed. The National Food Institute, Technical University of Denmark, Copenhagen, Denmark.

Herikstad, H., Motarjemi, Y., and Tauxe, R.V. 2002. *Salmonella* surveillance: a global survey of public health serotyping. *Epidemiol. Infect.* **129**, 1–8.

Herrera-Leon, S., Ramiro, R., Arroyo, M., Diez, R., Usera, M.A., and Echeita, M.A. 2007. Blind comparison of traditional serotyping with three multiplex PCRs for the identification of *Salmonella* serotypes. *Res. Microbiol.* **158**, 122–127.

Hopper, S.A. and Mawer, S. 1988. *Salmonella enteritidis* in a commercial layer flock. *Vet. Rec.* **123**, 351.

Kidgell, C., Reichard, U., Wain, J., Linz, B., Torpdahl, M., Dougan, G., and Achtman, M. 2002. *Salmonella typhi*, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect. Genet. Evol.* **2**, 39–45.

Kim, K.S., Ko, K.S., Chang, M.W., Hahn, T.W., Hong, S.K., and Kook, Y.H. 2003. Use of *rpoB* sequences for phylogenetic study of *Mycoplasma* species. *FEMS Microbiol. Lett.* **226**, 299–305.

Kim, B.J., Lee, S.H., Lyu, M.A., Kim, S.J., Bai, G.H., Chae, G.T., Kim, E.C., Cha, C.Y., and Kook, Y.H. 1999. Identification of mycobacterial species by comparative sequence analysis of the RNA polymerase gene (*rpoB*). *J. Clin. Microbiol.* **37**, 1714–1720.

Korczak, B., Christensen, H., Emler, S., Frey, J., and Kuhnert, P. 2004. Phylogeny of the family *Pasteurellaceae* based on *rpoB* sequences. *Int. J. Syst. Evol. Microbiol.* **54**, 1393–1399.

Kotetishvili, M., Stine, O.C., Kreger, A., Morris, Jr., J.G., and Sulakvelidze, A. 2002. Multilocus sequence typing for characterization of clinical and environmental *Salmonella* strains. *J. Clin. Microbiol.* **40**, 1626–1635.

Kwon, H.J., Park, K.Y., Kim, S.J., and Yoo, H.S. 2001. Application of nucleotide sequence of RNA polymerase beta-subunit gene (*rpo*B) to molecular differentiation of serovars of *Salmonella enterica* subsp. *enterica*. *Vet. Microbiol.* **82**, 121–129.

La Scola, B., Gundi, V.A., Khamis, A., and Raoult, D. 2006. Sequencing of the *rpoB* gene and flanking spacers for molecular identification of *Acinetobacter* species. *J. Clin. Microbiol.* **44**, 827–832.

Maiden, M.C., Bygraves, J.A., Feil, E., Morelli, G., Russell, J.E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D.A., and *et al.* 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc. Natl. Acad. Sci. USA* **95**, 3140–3145.

Marianelli, C., Ciuchini, F., Tarantino, M., Pasquali, P., and Adone, R. 2006. Molecular characterization of the *rpoB* gene in *Brucella* species: new potential molecular markers for genotyping. *Microbes Infect.* **8**, 860–865.

McClelland, M., Sanderson, K.E., Clifton, S.W., Latreille, P., Porwollik, S., Sabo, A., Meyer, R., Bieri, T., Ozersky, P., McLellan, M., and *et al.* 2004. Comparison of genome degradation in Paratyphi A and Typhi, human-restricted serovars of *Salmonella enterica* that cause typhoid. *Nat. Genet.* **36**, 1268–1274.

Mollet, C., Drancourt, M., and Raoult, D. 1997. *rpoB* sequence analysis as a novel basis for bacterial identification. *Mol. Microbiol.* **26**, 1005–1011.

Oscar, T.P. 1998. Identification and characterization of *Salmonella* isolates by automated ribotyping. *J. Food Prot.* **61**, 519–524.

Reeves, M.W., Evins, G.M., Heiba, A.A., Plikaytis, B.D., and Farmer, 3rd., J.J. 1989. Clonal nature of *Salmonella typhi* and its genetic relatedness to other salmonellae as shown by multilocus enzyme electrophoresis, and proposal of *Salmonella bongori* comb. nov. *J. Clin. Microbiol.* **27**, 313–320.

Saitou, N. and Nei, M. 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **4**, 406–425.

Sukhnanand, S., Alcaine, S., Warnick, L.D., Su, W.L., Hof, J., Craver, M.P., McDonough, P., Boor, K.J., and Wiedmann, M. 2005. DNA sequence-based subtyping and evolutionary analysis of selected *Salmonella enterica* serotypes. *J. Clin. Microbiol.* **43**, 3688–3698.

Suzuki, S. 1994. Pathogenicity of *Salmonella enteritidis* in poultry. *Int. J. Food Microbiol.* **21**, 89–105.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., and Kumar, S. 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol. Biol. Evol.* **28**, 2731–2739.

Tankouo-Sandjong, B., Sessitsch, A., Liebana, E., Kornschober, C., Allerberger, F., Hachler, H., and Bodrossy, L. 2007. MLST-v, multilocus sequence typing based on virulence genes, for molecular typing of *Salmonella enterica* subsp. *enterica* serovars. *J. Microbiol. Methods* **69**, 23–36.

Torpdahl, M., Skov, M.N., Sandvang, D., and Baggesen, D.L. 2005. Genotypic characterization of *Salmonella* by multilocus sequence typing, pulsed-field gel electrophoresis and amplified fragment length polymorphism. *J. Microbiol. Methods* **63**, 173–184.