PCR PRIMER DESIGN METHOD FOR DIFFERENTIATING AMONG *SALMONELLA* SEROGROUPS BASED ON AN ALGORITHM TARGETING GENE-FLANKING REGIONS

Yasukazu Muramatsu[1], Yoshifumi Tsukaya[1], Tetsuya Ikeda[2], Leo Uchida[1], Yuichi Osa[3] and Daiji Endoh[4]

[1]Laboratory of Zoonotic Diseases and Laboratory of Veterinary Radiology, School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu; [2]Hokkaido Institute of Public Health, Sapporo; [3]Eastern Hokkaido Wildlife Station, Nature Conservation Division, Institute of Environmental Science, Hokkaido Research Organization, Kushiro; [4]Laboratory of Veterinary Radiology, School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu, Hokkaido, Japan

Abstract. We applied an algorithm targeting length polymorphisms of intergenic sequences between gene-flanking regions for constructing PCR primer pairs to distinguish among serogroups of *Salmonella*, a major pathogen of humans and animals. From 43 constructed primer pairs, a pair capable in a single-step conventional PCR to categorize five serogroups of *Salmonella enterica* subsp *enterica* into three classes according to amplicon lengths (400, 800, and 900 bp, respectively). Nucleotide sequences of the amplicons were those of flanking regions rfbH and rfbJ. No amplicon was generated in other bacterial genera examined, indicative of the high specificity of this PCR primer pair. As more genetic information becomes available, the smaller number of primer pairs will be required in multiplex-PCR for differentiating Salmonella microorganisms using the novel primer design method.

Keywords: *Salmonella*, algorithm, gene-flanking region, PCR, serogroup

INTRODUCTION

Several serovars of nontyphoidal *Salmonella enterica* are causative agents of gastroenteritis causing mild to moderate symptoms, such as fever, abdominal cramp, and diarrhea in humans and animals (Leader *et al*, 2009). These zoonotic pathogens primarily belong to *S. enterica* subsp. *enterica*, which is closely associated with livestock and poultry. There are more than 1,500 serovars of *S. enterica* subsp. *enterica*, divided into over 50 serogroups based on reactivity of somatic (O) antigen (Grimont and Weill, 2007). The majority of *S. enterica* subsp *enterica* isolates associated with humans and warm-blooded animals belong to O-antigen serogroups B (O:4), C1 (O:7), C2 (O:8), and D1 (O:9) (Iwen, 2015), although *S. enterica* subsp II 1, 4, 12, 27: b: [e, n, x] (*S.* Sofia) is predominant in chicken meat from Australia (Mellor *et al*, 2010; McAuley *et al*, 2015).

Correspondence: Dr Daiji Endoh, Laboratory of Veterinary Radiology, School of Veterinary Medicine, Rakuno Gakuen University, Ebetsu, Hokkaido 069-8501, Japan.
Tel/Fax: +81 11388 4847
E-mail: dendoh@rakuno.ac.jp

Multiplex-PCR has become a useful technique for serotyping *S. enterica* isolates (Liu *et al*, 2012; Karns *et al*, 2015) as there exists several genes responsible for O-antigen expression located within large regulon *rfb* clusters. The sequence information on these genes is suitable for serogroup-specific primer design (Fitzgerald *et al*, 2003; Herrera-León *et al*, 2007). Variation in *Salmonella* O-antigen mainly depends on genetic diversity within *rfb* clusters that contain three groups of genes, namely, those encoding enzymes related to sugar synthesis and formation of O-subunit, those encoding transferases that assemble sugar substituents into O-subunit, and those encoding proteins involved in O-antigen construction from O-subunit, such as O-antigen transporter (Fitzgerald *et al*, 2003). That genes are involved in sugar synthesis and O-antigen metabolism emphasize variations in sugar moieties of lipopolysaccharides (LPS) are important parameters for differences in the properties of O-antigens and serotypes (Curd *et al*, 1998; Wang *et al*, 2002). In addition to LPS and O-antigen capsule, there are operons encoding production of enterobacterial common antigen (*wec* operon), colanic acid (*wca* operon) and cellulose (*bcs* operon) (Marshall, 2013) as well as *cld* encoding a protein that determines O-antigen chain lengths (Thampapillai *et al*, 1994).

In general, a nucleotide sequence located in a single gene responsible for O-antigen expression is targeted in the design of each primer pair. Specificity of the target sequence allows for accuracy and precision of PCR method. On the other hand, more primer pairs are needed to differentiate isolates into serogroups with higher accuracy (Liu *et al*, 2011; Li *et al*, 2014); however, the less the numbers of primer pairs the better are the results due to reduction in complexity and in influence on amplification of target sequences.

The purpose of the study was to reduce number of PCR primer pairs used in multiplex-PCR for detecting and differentiating among Salmonella isolates by applying an unconventional PCR primer design methodology. We utilized a novel *in silico* approach based on an algorithm targeting polymorphic lengths of intergenic sequences in gene-flanking regions (GFRs). Successful results of this novel approach should enable a reduction in the numbers of PCR primers currently needed for serogrouping *Salmonella* isolates.

MATERIALS AND METHODS

Bacterial strains and genomic DNA extraction

Of all the 19 *Salmonella* strains used in this study, 18 strains were *S. enterica* subsp *enterica* and one strain was *Salmonella enterica* subsp *salamae* (*S.* Sofia). The 19 strains were classified into 6 serogroups and 17 serovars (Table 1). Bacterial strains used as negative controls were *Escherichia coli* KI1225, KI677, KI680, and KI816 (kind gifts from the National Institute of Infectious Diseases, Japan) *Enterobacter aerogenes* (ATCC 13048), *E. cloacae* (ATCC 23355), *Klebsiella pneumoniae* (ATCC 13883), *Pseudomonas aeruginosa* (ATCC 27853), *Citrobacter freundii* (from the study), *Bacillus subtilis* (ATCC 6633), *Serratia marcescens* (ATCC 13880), and *Proteus vulgalis* (ATCC 13315). Bacterial isolates were incubated in Luria–Bertani broth (Formedium, Norfolk, UK) for 8-12 hours at 37°C, and then cultured on nutrient agar (Nissui Pharmaceutical, Tokyo, Japan). Genomic DNA extraction was performed using a QIAamp® DNA Mini Kit (QIAGEN, Tokyo, Japan) and DNA samples were stored at -30°C until used.

Table 1
Strains of *Salmonella enterica* subsp *enterica* and *salamae* used in the study.

Subgenera	Subspecies	O-antigen group (serogroup)	Serovar
I	*enterica*	O:4 (B)	Typhimurium
			O4:i:-
			Saintpaul
			Agona
			Derby
			Schwarzengrund
		O:7 (C1)	Infantis
			Infantis
			O7:-:-
			Montevideo
			Braenderup
			Braenderup
			Rissen
			Livingstone
			Mbandaka
		O:8 (C2)	Manhattan
		O:9 (D1)	Enteritidis
		O:1,3,19 (E4)	Senftenberg
II	*salamae*	1,4,12,[27]:b:[e,n,x]	II 1,4,12,[27]:b:[e,n,x] (Sofia)

Construction of PCR primer pairs

Complete genome sequences of *Salmonella* spp registered in DNA Data Bank of Japan (DDBJ; Mishima, Japan) were analyzed with focus on GFRs consisting of two adjacent genes and their intergenic sequences harboring length polymorphisms in rfb cluster of each genome (Fig 1). Based on sequence of the GFRs, a variety of primer pairs were constructed based on the following criteria: (i) a primer pair should target sequences of GFRs found in all *Salmonella* spp but not in other bacteria, (ii) length of target GFR should vary among *Salmonella* serogroups, and (iii) primer sets should amplify GFRs at 3′-region of the upstream gene, region between the upstream and downstream genes (intergenic sequence), and 5′ region of the downstream gene.

PCR and DNA sequencing

In order to determine whether amplicon was a single DNA fragment, all constructed primer pairs were subjected to touchdown PCR using *S*. Typhimurium DNA. The reaction mixture consisted of 1 μl of each primer, 2 μl of DNA template, 12.5 μl of a PCR pre-mixed reagent (2X Go Taq® Green Master Mix; Promega, Madison, WI), and DNA-free water

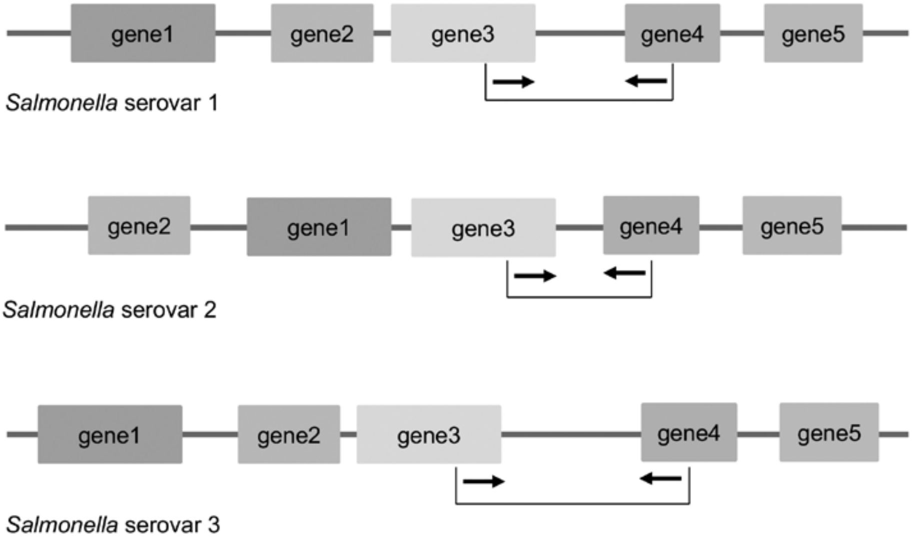

Fig 1-Primer design concept. Distance between each gene indicates length of gene-flanking region (GFR). Every gene in O-antigen coding region is numbered based on amino acid sequence. If the same gene is adjacent in all *Salmonella* strains and GFR length is different (*eg*, genes 3 and 4), the region is targeted for primer design. As all *Salmonella* strains have gene 4 adjacent to gene 5 and GFR lengths are similar, this region was not targeted for primer design.

(Promega) to bring the final volume to 25 μl Touchdown PCR was performed in a TaKaRa PCR Thermal Cycler Dice™ (Takara Bio, Shiga, Japan) under the following conditions: 5 minutes at 94°C; 18 cycles of 1 minute each at 94°C, 69°C and 72°C with stepwise lowering of the annealing temperature from 69°C to 59°C by the 18th cycle; 22 cycles of 1 minute each at 94°C, 58°C, and 72°C; and a final incubation at 72°C for 5 minutes. Following 1.7% agarose gel-electrophoresis, amplicons were stained with ethidium bromide, visualized under UV illumination, then extracted and purified using NucleoSpin® Gel and PCR Clean-up kit (Macherey-Nagel, Düren, Germany) for nucleotide sequencing in both directions (FASMAC, Kanagawa, Japan). Similarity of the sequences obtained was analyzed using BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

RESULTS

Based on the novel algorithm for primer design, 43 primer pairs were constructed as candidates for specific amplification of *Salmonella* DNA (Table 2). All nucleotide sequences selected as primers were located on genes associated with sugar synthesis, O-antigen metabolism and production of enterobacterial common antigens, namely, *cld*, *galF*, *gnd*, *manC*, *rfb* cluster, *udg*, *wcaL*, and *wcaM*.

Touchdown PCR revealed six primer pairs (nos. 10, 12, 18, 27, 28, and 32) amplifying *Salmonella* DNA without non-specific amplicons (data not shown). Primer pair no. 18 resulted in amplification of all *Salmonella* serovars evaluated in the study, differentiating into three categories: serogroups B (O:4), D (O:8), E4 (O:1,3,19), and *S*. Sofia producing a 400-bp amplicon; serogroup C1 (O:7) a 800-bp amplicon; and

Table 2
Sequences of the fourty-three primer pairs used in the study.

	Forward primer		Reverse primer		Gene target related to Salmonella O-antigen biosynthesis*			
Primer pair no.	Name	Nucleotide sequence (5′ → 3′)	Name	Nucleotider sequence (5′ → 3′)	Gene	Encoded protein	Gene	Encoded protein
1	NC_003198_cld_wcaL_F	GTTATGTGATGAAGCCGACGC	NC_003198_cld_wcaL_R	CGTAGCGGGTAAAATTGAGCG	cld	polysaccharide chain length regulator	wcaL	glycosyltransferase STY2310
2	NC_003198_ugd_wcaM_F	AGAGCTTTTGGATGTCGCTGA	NC_003198_ugd_wcaM_R	GTCATCCGCCGGATAATCCTT	ugd	UDP-glucose 6-dehydrogenase	wcaM	colanic acid biosynthesis protein WcaM STY2309
3	NC_003198_gnd_galF_F	TGCTGTTAGCGCCTTACTTCA	NC_003198_gnd_galF_R	TTTTAGACGCGTGAGTCACCA	gnd	6-phosphogluconate dehydrogenase	galF	UTP-glucose-1-phosphate uridylyltransferase
4	NC_003198_rfbP_rfbB_F	GAAAGTTGTTTTGCGGCGAGA	NC_003198_rfbP_rfbB_R	AGGTGCCGACGATATTGGTTT	rfbP	undecaprenyl-phosphate galactosephosphotransferase	rfbB	dTDP-glucose 4,6-dehydratase
5	NC_003198_rfbK_rfbD_F	CTCCCAACTCGTTACACCCAT	NC_003198_rfbK_rfbD_R	GCTGCTGCGTTAACAATCACA	rfbK	phosphomannomutase	rfbD*	dTDP-4-dehydrorhamnose reductase
6	NC_003198_manC_rfbA_F	GGATACTATTCACGCGCTGGA	NC_003198_manC_rfbA_R	CATCCGGACTTGGCTGTACTT	manC	mannose-1-phosphate guanylyltransferase	rfbA	TDP-glucose pyrophosphorylase
7	NC_003198_rfbN_rfbC_F	CCGTTCTGGATTCCAAGAGCT	NC_003198_rfbN_rfbC_R	CGACCGCAACATCAAAAACCT	rfbN	rhamnosyltransferase	rfbC	dTDP-4-dehydrorhamnose 3,5-epimerase
8	NC_003198_rfbU_rfbI_F	TGTCGCACGGTATGACTTTGA	NC_003198_rfbU_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbU	glycosyltransferase	rfbI	reductase RfbI
9	NC_017623_rfbV_rfbC_F	AGAAAGGAAAGGCCAGACGAA	NC_017623_rfbV_rfbC_R	TGTAGCCCTCTGAGTACGTTC	rfbV	abequosyltransferase RfbV	rfbC*	dTDP-4-dehydrorhamnose 3,5-epimerase
10	NC_017623_rfbJ_rfbI_F	GCAATGGGCTGGAAACCAAAT	NC_017623_rfbJ_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbJ*	CDP-abequose synthase	rfbI	CDP-6-deoxy-delta-3,4-glucoseen reductase
11	CP000886_rfbH_rfbJ_F	TGGTTTGGCTTCCCTATCACC	CP000886_rfbH_rfbJ_R	TCCCAACTGCACCATCTAACA	rfbH*	Lipopolysaccharide biosynthesis protein	rfbJ	CDP-abequose synthase
12	CP000886_rfbJ_rfbX_F	GCAATGGGCTGGAAACCAAAT	CP000886_rfbJ_rfbX_R	GCCAAAATCAACAGCGCTACA	rfbJ	CDP-abequose synthase	rfbX*	O-antigen transporter
13	CP000886_rfbX_rfbV_F	GAATCAGTGGAGTGCTGCTTG	CP000886_rfbX_rfbV_R	TGGCCCAAGGTTAACGCTATT	rfbX*	O-antigen transporter	rfbV*	glycosyltransferase
14	CP003386_rfbU_rfbI_F	TGTCGCACGGTATGACTTTGA	CP003386_rfbU_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbU	glycosyltransferase	rfbI	reductase RfbI
15	CP003386_rfbV_rfbF_F	TGGGAAACTAGCATATGCAGTGA	CP003386_rfbV_rfbF_R	CAGACGACCACCAGTCATTGA	rfbV	glycosyltransferase	rfbF	glucose-1-phosphate cytidylyltransferase
16	CP003386_rfbX_rfbG_F	GAATCAGTGGAGTGCTGCTTG	CP003386_rfbX_rfbG_R	AGATAGGCGCACTAAAGGCTG	rfbX	O-antigen transporter	rfbG	CDP-glucose 4,6-dehydratase
17	CP003386_rfbJ_rfbH_F	GCAATGGGCTGGAAACCAAAT	CP003386_rfbJ_rfbH_R	GAGCTCGCTCGCCTAATTTTG	rfbJ	CDP-abequose synthase	rfbH	dehydratase RfbH
18	CP003386_rfbH_rfbJ_F	TGGTTTGGCTTCCCTATCACC	CP003386_rfbH_rfbJ_R	ACATCTCGAGTGATTGCGACA	rfbH	dehydratase RfbH	rfbJ	CDP-abequose synthase
19	CP003416_rfbV_rfbC_F	AGAAAGGAAAGGCCAGACGAA	CP003416_rfbV_rfbC_R	TGTAGCCCTCTGAGTACGTTC	rfbV*	glycosyltransferase	rfbC*	dTDP-4-deoxyrhamnose 3,5 epimerase SU5_02688

Table 2 (Continued)

	Forward primer		Reverse primer		Gene target related to Salmonella O-antigen biosynthesis*			
Primer pair no.	Name	Nucleotide sequence (5′ → 3′)	Name	Nucleotider sequence (5′ → 3′)	Gene	Encoded protein	Gene	Encoded protein
20	CP003416_rfbJ_rfbI_F	GCAATGGGCTGGAAACCAAAT	CP003416_rfbJ_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbJ	CDP-6-deoxy-delta-3,4-glucoseen reductase-like protein	rfbI	bactoprenol-linked glucose translocase
21	NC_006905_galF_galF_F	GCTGGATTCCGATTTGATGGC	NC_006905_galF_galF_R	CTTTTGGGATTGCCTTGGTGG	galF	UTP-glucose-1-phosphate uridylyltransferase		
22	NC_011080_rfbH_rfbB_F	GGAAATCTGATTCGCCAACCG	NC_011080_rfbH_rfbB_R	AGGTGCCGACGATATTGGTTT	rfbH	Lipopolysaccharide biosynthesis protein rfbH	rfbB	dTDP-glucose-4,6-dehydratase
23	NC_011080_rfbG_rfbD_F	ATGGATTCGCGGCGAAGATAT	NC_011080_rfbG_rfbD_R	TTTAGCGATGGCTTCCACACT	rfbG	CDP-glucose 4,6-dehydratase	rfbD	dTDP-4-dehydrorhamnose reductase
24	NC_011080_rfbF_rfbA_F	CTTTTCCACCTGGACGCTTTG	NC_011080_rfbF_rfbA_R	TGTGGCGTACTGATGATCAGG	rfbF	glucose-1-phosphate cytidylyltransferase	rfbA	glucose-1-phosphate thymidylyltransferase
25	NC_011080_rfbC_rfbC_F	GGAGTCGCCTACTTTTGGTCA	NC_011080_rfbC_rfbC_R	TTACCCTGTGCATTTGCTCCT	rfbC	dTDP-4-dehydrorhamnose 3,5-epimerase		
26	NC_011274_rfbH_rfbI_F	CGTGTAGTGGGTGAGTTGACA	NC_011274_rfbH_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbH	dehydratase RfbH	rfbI	bactoprenol-linked glucose translocase
27	NC_011274_rfbG_rfbF_F	ATGGATTCGCGGCGAAGATAT	NC_011274_rfbG_rfbF_R	CAGACGACCACCAGTCATTGA	rfbG	CDP-glucose 4,6-dehydratase	rfbF	glucose-1-phosphate cytidylyltransferase
28	NC_017046_rfbU_rfbI_F	TGTCGCACGGTATGACTTTGA	NC_017046_rfbU_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbU	glycosyltransferase	rfbI	reductase RfbI
29	NC_017046_rfbV_rfbF_F	TGGGAAACTAGCATATGCAGTGA	NC_017046_rfbV_rfbF_R	CAGACGACCACCAGTCATTGA	rfbV	glycosyltransferase	rfbF	glucose-1-phosphate cytidylyltransferase
30	NC_017046_rfbX_rfbG_F	GAATCAGTGGAGTGCTGCTTG	NC_017046_rfbX_rfbG_R	AGATAGGCGCACTAAAGGCTG	rfbX	O-antigen transporter	rfbG	CDP-glucose 4,6-dehydratase
31	NC_017046_rfbJ_rfbH_F	GCAATGGGCTGGAAACCAAAT	NC_017046_rfbJ_rfbH_R	GAGCTCGCTCGCCTAATTTTG	rfbJ	CDP-abequose synthase	rfbH	dehydratase RfbH
32	NC_017046_rfbH_rfbJ_F	TGGTTTGGCTTCCCTATCACC	NC_017046_rfbH_rfbJ_R	ACATCTCGAGTGATTGCGACA	rfbH	dehydratase RfbH	rfbJ	CDP-abequose synthase
33	AE017220_galF_galF_F	GCTGGATTCCGATTTGATGGC	AE017220_galF_galF_R	CTTTTGGGATTGCCTTGGTGG	galF	UTP-glucose-1-phosphate uridylyltransferase		
34	CP001113_rfbH_rfbB_F	GGAAATCTGATTCGCCAACCG	CP001113_rfbH_rfbB_R	AGGTGCCGACGATATTGGTTT	rfbH	lipopolysaccharide biosynthesis protein Rfb	rfbB	dTDP-glucose-4,6-dehydratase
35	CP001113_rfbG_rfbD_F	ATGGATTCGCGGCGAAGATAT	CP001113_rfbG_rfbD_R	TTTAGCGATGGCTTCCACACT	rfbG	CDP-glucose 4,6-dehydratase	rfbD	dTDP-4-dehydrorhamnose reductase
36	CP001113_rfbF_rfbA_F	CTTTTCCACCTGGACGCTTTG	CP001113_rfbF_rfbA_R	TGTGGCGTACTGATGATCAGG	rfbF	glucose-1-phosphate cytidylyltransferase	rfbA	glucose-1-phosphate thymidylyltransferase
37	CP001113_rfbC_rfbC_F	GGAGTCGCCTACTTTTGGTCA	CP001113_rfbC_rfbC_R	TTACCCTGTGCATTTGCTCCT	rfbC	dTDP-4-dehydrorhamnose 3,5-epimerase		

Table 2 (Continued)

	Forward primer		Reverse primer		Gene target related to Salmonella O-antigen biosynthesis*			
Primer pair no.	Name	Nucleotide sequence (5′ → 3′)	Name	Nucleotider sequence (5′ → 3′)	Gene	Encoded protein	Gene	Encoded protein
38	AM933173_rfbH_rfbI_F	CGTGTAGTGGGTGAGTTGACA	AM933173_rfbH_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbH	dehydratase RfbH	rfbI	bactoprenol-linked glucose translocase CDP-6-deoxy-delta-3,4-glucoseen reductase
39	AM933173_rfbG_rfbF_F	ATGGATTCGCGGCGAAGATAT	AM933173_rfbG_rfbF_R	CAGACGACCACCAGTCATTGA	rfbG	CDP-glucose 4,6-dehydratase	rfbF	glucose-1-phosphate cytidylyltransferase
40	NC_011294_rfbH_rfbI_F	CGTGTAGTGGGTGAGTTGACA	NC_011294_rfbH_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbH	dehydratase RfbH	rfbI	bactoprenol-linked glucose translocase CDP-6-deoxy-delta-3,4-glucoseen reductase
41	NC_011294_rfbG_rfbF_F	ATGGATTCGCGGCGAAGATAT	NC_011294_rfbG_rfbF_R	CAGACGACCACCAGTCATTGA	rfbG	CDP-glucose 4,6-dehydratase	rfbF	glucose-1-phosphate cytidylyltransferase
42	AM933172_rfbH_rfbI_F	CGTGTAGTGGGTGAGTTGACA	AM933172_rfbH_rfbI_R	GTTGGTGGTGTGCGTAACTTC	rfbH	dehydratase RfbH	rfbI	CDP-6-deoxy-delta-3,4-glucoseen reductase; bactoprenol-linked glucose translocase
43	AM933172_rfbG_rfbF_F	ATGGATTCGCGGCGAAGATAT	AM933172_rfbG_rfbF_R	CAGACGACCACCAGTCATTGA	rfbG	CDP-glucose 4,6-dehydratase	rfbF	glucose-1-phosphate cytidylyltransferase

*GenBank accession number: Salmonella enterica subsp enterica, NC_003198; S.Typhi strain CT18, NC_017623; S. Heidelberg strain B182, CP000886; S. Paratyphi B strain SPB7, CP003416; S. Heidelberg strain B182, NC_006905; S. Cholerasuis strain SC-B7, NC_011080; S. Newport strain SL254, NC_011274; S. Gallinarum strain 287/91, NC_017046; S. Typhimurium strain 798, AE017220; S. Choleraesuis strain SC-B67, CP001113; S. Newport strain SL254, AM933173; S. Gallinarum strain 287/91, NC_011294; S. Enteritidis strain P125109, AM933172; S. Enteritidis strain P125109.

C2 (O:8) a 900-bp amplicon (Fig 2A). The primer pair did not produce amplicons with any non-*Salmonella* bacterial strains examined (Fig 2B).

Sequences of the amplicons generated by primer pair no. 18 were identified as sequences of *Salmonella* spp (Table 3). Of the 19 *Salmonella* strains used in the study, 17 strains demonstrated high similarity (99.4-100%) to sequences of *Salmonella* serovars corresponding to each strain (Table 3). However, amplicon of *S*. Manhattan showed highest similarity (99.8%) to *S*. Senftenberg (GenBank accession no. LN868943.1) and *S*. Sofia amplicon highest similarity (96.9%) to *S. enterica* subsp *enterica* (group B) *rfb* cluster (GenBank accession. no. X56793.1).

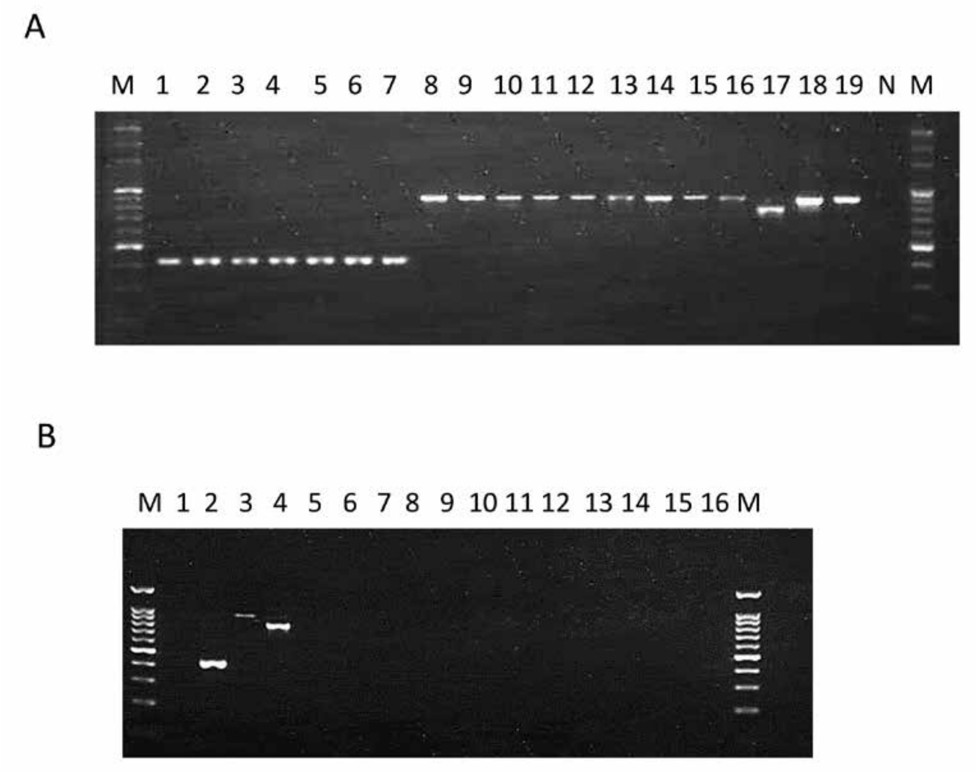

Fig 2-Amplicons generated by primer pair no. 18. Following standard PCR amplification protocol, amplicons were analyzed by 1.7% agarose gel-electrophoresis and ethidium bromide staining. Details of *Salmonella* serovars and serogroups are shown in Table 1. Panel A. Lanes 1-7, *S*. serogroup B samples; lanes 8-16; *S*. serogroup C1 samples; lane 17, *S*. serogroup C2 sample; lane 18, *S*. serogroup D sample; lane 19, *S*. serogroup E sample. Panel B. Lane *S*. Typhimurium, *S*. Typhimurium (serogroup B); lane *S*. Infantis, *S*. Infantis (serogroup C1); lane *S*. Manhattan, *S*. Manhattan (serogroup C2); lanes 1, *Escherichia coli* KI1225; lane 2, *E. coli* KI677; lane 3, *E. coli* KI680; lane 4, *E. coli* KI816; lane 5, *Enterobacter aerogenes*; lane 6, *E. cloacae*; lane 7, *Klebsiella pneumoniae*; lane 8, *Pseudomonas aeruginosa*; lane 9, *Citrobacter freundii*; lane 10, *Bacillus subtilis*; lane 11, *Serratia marcescens*; lane 12, *Proteus vulgaris*. Lane M, 100-bp DNA size markers; lane N, negative control (DNA-free water).

Table 3
Salmonella spp identified by primer pair no. 18.

Serovar	Serogroup	Salmonella spp[a]	GenBank acc. no.	% identity	Nt position	Nt position
Typhimurium	B	*Salmonella enterica* subsp *enterica* serovar Typhimurium strain SL1344RX	CP011233.1	100	2106119	2106451
O4:i:-(Typhimurium?)	B	*Salmonella enterica* subsp *enterica* serovar Typhimurium strain SL1344RX	CP011233.1	100	2106120	2106451
Agona	B	*Salmonella enterica* subsp *enterica* serovar Agona strain 460004 2-1	CP011259.1	100	3960653	3960985
Schwarzengrund	B	*Salmonella enterica* subsp *enterica* serovar Schwarzengrund strain CVM19633	CP001127.1	100	2192396	2192752
Infantis	C1	*Salmonella enterica* subsp *enterica* serovar Infantis genome assembly SINFA	LN649235.1	99.7	1051405	1052062
Infantis	C1	*Salmonella enterica* subsp *enterica* serovar Infantis genome assembly SINFA	LN649235.1	99.4	1052172	1051691
Montevideo	C1	*Salmonella enterica* subsp *enterica* serovar Montevideo strain USDA-ARS-USMARC-1921	CP007540.1	99.7	2765914	2765218
Manhattan[b]	C2	*Salmonella enterica* subsp *enterica* serovar Senftenberg NCTC10384	LN868943.1	99.8	2213778	2214396
Enteritidis	D	*Salmonella enterica* subsp *enterica* serovar Enteritidis strain EC2011022	CP007323.2	100	1025607	1026264
Senftenberg	E4	*Salmonella enterica* subsp *enterica* serovar Senftenberg NCTC10384	LN868943.1	99.9	2214578	2213907
Sofia	1,4,12,[27]:b:[e,n,x]	*Salmonella enterica* subsp *enterica* (group B) rfb gene cluster	X56793.1	96.9	11836	11481

[a]From BLAST search. [b]Complete genome sequence not determined.

DISCUSSION

Nontyphoidal *Salmonella* are important foodborne pathogens that cause gastroenteritis, bacteremia, and subsequent focal infection (Hohmann, 2001). The majority of these pathogens are zoonotic, transmitted from healthy carrier animals to humans via contaminated food (Kornschober *et al*, 2009). *S*. Enteritidis and *S*. Typhimurium are the major causative agents of food-borne salmonellosis in humans worldwide (Baay and Veld, 1993; Tan and Shelef, 1999). *S*. Infantis also is a frequent source of foodborne infections in most industrialized countries, primarily causing gastroenteritis (Schroeder *et al*, 2016). Large outbreaks of *S*. Infantis infection associated with contaminated food products (*eg*, raw or undercooked chicken meat, egg, pork, and beef) were reported (Schroeder *et al*, 2016).

Performing conventional PCR with a single primer pair (no. 18) targeting the length polymorphisms of intergenic sequences in GFRs could successfully categorized *Salmonella* isolates into three amplicon groups. *S*. Typhimurium was distinguishable from *S*. Enteritidis and *S*. Infantis on the basis of their respective amplicon size. *S*. Typhimurium, Enteritidis and Infantis have been reported as the major serovars to be detected from cattle, pigs, and broiler and layer chickens in Japan (Ishiguro *et al*, 1980; Asai *et al*, 2006; Kobayashi *et al*, 2014). Feces is one of the most common samples used to detect *Salmonella* DNA, and contains a large amount and a variety of the intestinal bacteria other than *Salmonella*. However, primer pair no. 18 would not detect any other Enterobacteriaceae species DNA except that of *Salmonella*. This notion is supported by sequencing results. Therefore, primer pair no. 18 is suitable for screening major serovars of *Salmonella* isolates. Other serovars tested are also of interest because they have been detected from food-producing animals and pet animals in Japan (Asai *et al*, 2006; Ishihara *et al*, 2009; Suzuki and Yamamoto, 2009; Kato *et al*, 2015; Kuroki *et al*, 2019; Yukawa *et al*, 2019). In addition, the primer pair generated *S*. Sofia amplicon with the same size as *S*. Typhimurium, suggesting this primer pair might be useful for detecting the former serovar, the most common found in chicken meat in Australia (Mellor *et al*, 2010; McAuley *et al*, 2015).

Of the 43 primer sets constructed for *Salmonella* GFRs, only one primer set efficiently and specifically detected the microorganisms tested. *Salmonella* genus consists of more than 2,600 serovars and, thus, nucleotide sequence variations are expected to be higher than those of other bacterial genera (Grimont and Weill, 2007). Due to these variations, it is difficult to detect a gene region that specifically exists in all *Salmonella* serovars. A successfully designed primer set that fulfills the requirements of specificity for *Salmonella* serogroups indicates the utility of this novel algorithm.

Although the complete genome sequence for *S*. Manhattan has not yet been determined, primer pair no. 18 amplicon of *S*. Manhattan demonstrated highest nucleotide similarity to that of *S*. Senftenberg. In addition, amplicon of *S*. Sofia showed highest sequence similarity to *S*. *enterica* subsp *enterica* (group B) *rfb* cluster. This is probably because *S*. Sofia was formerly categorized in *S*. *enterica* subsp *enterica*. serogroup B (Grimont and Weill, 2007). However, genome sequence data for *S*. Sofia could not be found in GenBank database (at least at the time of the present study). Although knowledge of genomes of the various bacterial strains is increasing, in the meantime the novel primer design method developed in the present study provides a useful tool for differentiating *Salmonella* serogroups, and can also be applied to other genera.

ACKNOWLEDGEMENTS

The work was supported by JSPS KAKENHI Grant nos. 20380163 and JP24580430, the Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT)-Supported Program for Strategic Research Foundation at Private Universities, 2013-2017 and in part by the priority research (2502-11E004) of Hokkaido Research Organization.

CONFLICTS OF INTEREST

The authors declare no conflicts of interest.

REFERENCES

Asai T, Esaki H, Kojima A, *et al*. Antimicrobial resistance in *Salmonella* isolates from apparently healthy food-producing animal from 2000 to 2003: the first stage of Japanese veterinary antimicrobial resistance monitoring (JVARM). *J Vet Med Sci* 2006; 68: 881-4.

Baay MF, Huis in't Veld JH. Alternative antigens reduce cross-reactions in an ELISA for the detection of *Salmonella enteritidis* in poultry. *J Appl Bacteriol* 1993; 74: 243-7.

Curd H, Liu D, Reeves PR. Relationships among the O-antigen gene clusters of *Salmonella enterica* groups B, D1, D2, and D3. *J Bacteriol* 1998; 180: 1002-7.

Fitzgerald C, Sherwood R, Gheesling LL, *et al*. Molecular analysis of the *rfb* O antigen gene cluster of *Salmonella enterica* serogroup O:6,14 and development of a serogroup-specific PCR assay. *Appl Environ Microbiol* 2003; 69: 6099-105.

Grimont P, Weill FX. Antigenic formulae of the *Salmonella* serovars. 9th ed. Geneva: WHO Collaborating Centre for Reference and Research on Salmonella, 2007.

Herrena-León S, Ramiro R, Arroyo M, *et al*. Blind comparison of traditional serotyping with three multiplex PCRs for the identification of *Salmonella* serotypes. *Res Microbiol* 2007; 158: 122-7.

Hohmann EL. Nontyphoidal salmonellosis. *Clin Infect Dis* 2001; 32: 263-9.

Ishiguro N, Hirose K, Sato G. Distribution of citrate utilization plasmid in *Salmonella* strains of bovine origin in Japan. *Appl Environ Microbiol* 1980; 40: 446-51.

Ishihara K, Takahashi T, Morioka A, *et al*. National surveillance of *Salmonella enterica* in food-producing animals in Japan. *Acta Vet Scand* 2009; 51: 35.

Iwen PC. *Salmonella* typing. Omaha: Nebraska Public Health Laboratory, 2014. [Cited 2018 Dec 4]. Available from: http://www.nphl.org/salmonellatyping-iwen.pdf.pdf

Karns JS, Haley BJ, van Kessel JAS. Improvements to a PCR-based serogrouping scheme for *Salmonella enterica* from dairy farm samples. *J Food Protect* 2015; 78: 1182-5.

Kato R, Matsushita S, Shimojima Y, *et al*. *Kansensyogakuzasshi* 2015; 89: 46-52 (in Japanase with English abstract).

Kobayashi A, Takahashi S, Ono M, *et al*. Molecular typing of *Salmonella enterica* serovar Enteritidis isolates from food-producing animals in Japan by multilocus variable-number tandem repeat analysis: evidence of clonal dissemination and replacement. *Acta Vet Scand* 2014; 56: 31.

Kornschober C, Mikula C, Springer B. Salmonellosis in Austria: situation and trends. *Wien Klin Wochenschr* 2009; 121: 96-102.

Kuroki T, Ishihara T, Nakajima N, *et al*. Prevalence of *Salmonella enterica* subspecies *enterica* in red-eared sliders *Trachemys scripta elegans* retailed in pet shops in Japan. *Jpn J Infect Dis* 2019; 72: 38-43.

Leader BT, Frye JG, Hu J, *et al*. High-throughput molecular determination of *Salmonella* *enterica* serovars by use of multiplex PCR and capillary electrophoresis analysis. *J Clin Microbiol* 2009; 47: 1290-9.

Li R, Wang Y, Shen J, *et al*. Development of a novel hexa-plex PCR method for identification and serotyping of *Salmonella* species. *Foodborne Pathog Dis* 2014; 11: 75-7.

Liu B, Zhang L, Zhu X, *et al*. PCR identification of *Salmonella* serogroups based on specific targets obtained by comparative genomics. *Int J Food Microbiol* 2011; 144: 511-8.

Marshall JM. 2013: [Cited 2018 Dec 4]. Available from: URL: https://etd.ohiolink.edu/rws_etd/document/get/osu1385998769/inline

McAuley CM, Duffy LL, Subasinghe N, Hogg G, Coventry J, Fegan N. *Salmonella* Typhimurium and *Salmonella* Sofia: growth in and persistence on egg under production and retail conditions. *BioMed Res Int* 2015; 2015: 914987.

Mellor GE, Duffy LL, Dykes GA, *et al*. Relative prevalence of *Salmonella* Sofia on broiler chickens pre- and postprocessing in Australia. *Poult Sci* 2010; 89:1544-8.

Schroeder S, Harries M, Prager R, *et al*. A prolonged outbreak of *Salmonella* Infantis associated with pork products in central Germany, April–October 2013. *Epidemiol Infect* 2016; 144: 1429-39.

Suzuki H, Yamamoto S. A literature survey of *Salmonella* contamination in shell eggs and liquid eggs in the world. *Bull Natl Inst Health Sci* 2009; 127: 74-83. (In Japanase with Abstract and Tables in English).

Tan W, Shelef LA. Automated detection of *Salmonella* spp. in foods. J *Microbiol Methods* 1999; 37: 87-91.

Thampaillai G, Lan R, Reves PR. Molecular evolution in the *gnd* locus of *Salmonela enteria. Mol Biol Evol* 1994; 11: 813-28.

Wang L, Andrianopoulos K, Liu D, *et al.* Extensive variation in the O-antigen gene cluster within one *Salmonella enterica* serogroup reveals an unexpected complex history. *J Bacteriol* 2002; 184:1669-77.

Yukawa S, Uchida I, Tamura Y, *et al*. Characterisation of antibiotic resistance of *Salmonella* isolated from dog treats in Japan. *Epidemiol Infect* 2019; 147: e102, 1-6.