Epidemiology of Nontyphoidal *Salmonella* Infections in Korean Children and Genetic Factors Associated with Extra-intestinal Invasion: A Whole-genome Sequencing Analysis

Hyun Mi Kang, M.D., Ph.D.[1,2,*], Jiyon Chu, M.S.[3,*], In Hyuk Yoo, M.D., Ph.D.[1], In Young Yoo, M.D., Ph.D.[4], Jeong-Ih Shin, Ph.D.[3,5], Mi-Ran Seo, Ph.D.[6], Yeun-Jun Chung, M.D., Ph.D.[3,5,7], Seung-Hyun Jung, Ph.D.[3,5,8], and Yeon Joon Park, M.D., Ph.D.[4]

[1]Department of Pediatrics, College of Medicine, The Catholic University of Korea, Seoul, Korea; [2]Vaccine Bio Research Institute, College of Medicine, The Catholic University of Korea, Seoul, Korea; [3]Department of Medical Sciences, Graduate School of The Catholic University of Korea, Seoul, Korea; [4]Department of Laboratory Medicine, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Seoul, Korea; [5]Catholic Research Institute for Human Genome Polymorphism, Precision Medicine Research Center, College of Medicine, The Catholic University of Korea, Seoul, Korea; [6]ConnectaGen, Hanam, Korea; [7]Department of Microbiology, College of Medicine, The Catholic University of Korea, Seoul, Korea; [8]Department of Biochemistry, College of Medicine, The Catholic University of Korea, Seoul, Korea

Supplemental Data

Supplemental Methods

Bacterial identification

Specimens were routinely inoculated onto appropriate media, including 5% sheep blood agar (Shin Yang Chemical, Seoul, Korea), MacConkey agar (Asan Pharmaceutical, Seoul, Korea), or Chocolate agar (Asan Pharmaceutical) for positive blood cultures, and Hektoen enteric agar (Asan Pharmaceutical), Campylobacter agar (Asan Pharmaceutical), thiosulfate–citrate–bile salts–sucrose agar (Asan Pharmaceutical), Cefsulodin-Irgasan-Novobiocin agar (Shin Yang Chemical) or Selenite F broth (Oxoid, Basingstoke, UK) for stool.

Isolates were tested using the ASTA MicroIDSys (ASTA, Suwon, Korea) MALDI-TOF MS system per the manufacturer's instruction. A single colony from a pure agar culture was smeared directly onto the target MALDI plate with a sterile wooden stick. Dried colonies were sequentially overlaid with 1.5 μL of 70% formic acid and 1.5 μL of α-cyano-4-hydroxycinnamic acid matrix solution and analyzed. The ASTA MicroIDSys CoreDB 1.29 (ASTA) was used for spectral analysis. The ASTA MicroIDSys standard isolate (*Escherichia coli* competent cells) was inoculated on the calibration spots for instrument calibration and quality control. The acceptable cutoff score for identification was set at 140 per the manufacturer's recommendation. Somatic (O) antigens of each *Salmonella* spp. strain were determined using the slide agglutination method, with polyvalent and monovalent O antigens provided by the Korea Centers for Disease Control and Prevention (Joongkyeom, Goyang-si, Korea).

Whole-genome sequencing analysis

Extracted DNA was evaluated using a NanoDrop instrument (Thermo Fisher Scientific, Waltham, MA, USA) and Qubit 2.0 Fluorometer (Thermo Fisher Scientific). Next-generation

sequencing libraries were prepared using a TruSeq Nano DNA sample preparation kit (Illumina, San Diego, CA, USA). The sequencing libraries were pooled and sequenced on an Illumina NovaSeq 6000 instrument for 300 cycles (2 × 150-bp, paired-end). Sequencing adapters and low-quality bases were trimmed using Trimmomatic [1]. The trimmed reads were assembled into contigs using SPAdes with default parameters [2], and the contigs were evaluated using Quast [3] and scaffolded using MeDuSa [4]. The obtained assemblies were annotated using Prokka (v1.14.6) with default parameters [5].

We used Kraken2 (v2.1.3) to confirm the species of each isolate [6]. *Salmonella* pathogenicity islands were explored using SPIFinder [7]. Serotypes were detected using SeqSero2 v1.2.1 [8]. Multilocus sequence typing (MLST) was conducted using the MLST software (https://github.com/tseemann/mlst) with allelic profiles of seven housekeeping genes (*aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA*) from the PubMLST database [9]. To assess possible evolutionary relationships among sequence types (STs), we compared the allelic profile of each ST using the PHYLOViZ v2.0 [10]. Core genome MLST was predicted using cgMLSTFinder [11] with EnteroBase [12]. The acquisition of and mutations in antimicrobial resistance genes were detected using ABRicate (https://github.com/tseemann/abricate) with the CARD database [13]. Genes associated with resistance to ampicillin (bla_{TEM}), fluoroquinolone (*parC*, gyrA, *qnr*A, *qnr*B, *qnr*S), trimethoprim-sulfamethoxazole (*dfr*A, *sul*1, *sul*2, sul3), third-generation cephalosporins ($bla_{\text{CTX–M}}$, bla_{DHA}), and chloramphenicol (*flo*R, *cml*A1) were investigated. Virulence factors were determined using ABRicate with the Virulence Factor Database (VFDB) [14].

SUPPLEMENTAL REFERENCES

1. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 2014;30:2114-20.

2. Prjibelski A, Antipov D, Meleshko D, Lapidus A, Korobeynikov A. Using SPAdes de

novo assembler. Curr Protoc Bioinformatics 2020;70:e102.

3. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics 2013;29:1072-5.

4. Bosi E, Donati B, Galardini M, Brunetti S, Sagot MF, Lió P, et al. MeDuSa: a multi-draft based scaffolder. Bioinformatics 2015;31:2443-51.

5. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics 2014;30:2068-9.

6. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the Kraken software suite. Nat Protoc 2022;17:2815-39.

7. Roer L, Hendriksen RS, Leekitcharoenphon P, Lukjancenko O, Kaas RS, Hasman H et al. Is the evolution of *Salmonella enterica* subsp. *enterica* linked to restriction-modification systems? mSystems 2016;1:e00009-16.

8. Zhang S, den Bakker HC, Li S, Chen J, Dinsmore BA, Lane C, et al. SeqSero2: rapid and improved *Salmonella* serotype determination using whole-genome sequencing data. Appl Environ Microbiol 2019;85:e01746-19.

9. Jolley KA and Maiden MCJ. BIGSdb: scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010;11:595.

10. Nascimento M, Sousa A, Ramirez M, Francisco AP, Carriço JA, Vaz C. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. Bioinformatics 2017;33:128-9.

11. Clausen PTLC, Aarestrup FM, Lund O. Rapid and precise alignment of raw reads against redundant databases with KMA. BMC Bioinformatics 2018;19:307.

12. Alikhan NF, Zhou Z, Sergeant MJ, Achtman M. A genomic overview of the population structure of *Salmonella*. PLoS Genet 2018;14:e1007261.

13. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, Edalatmand A, et al. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res 2020;48:D517-25.

14. Liu B, Zheng D, Jin Q, Chen L, Yang J. VFDB 2019: a comparative pathogenomic platform with an interactive web interface. Nucleic Acids Res 2019;47:D687-92.

Supplemental Data Table S1. Clinicopathologic and genomic characteristics of 58 patients with salmonellosis

The contents of Table S1 are provided in a separate Excel file.

Supplemental Data Table S2. Serogroup, serotype, and sequence type characterization by isolation year

Serogroup	Serotype	N isolates (N = 58)	Percent (%)	Sequence type (N isolates)	N isolates (%)		
					2019 N = 11	2020 N = 18	2021 N = 29
B	I 4,[5],12:i:-	5	8.6	19 (1), 2,379 (1), 36 (1), 34 (2)		1 (5.6)	4 (13.8)
B	I 4:b:-	2	3.4	42, 2,814			2 (6.9)
B	Saintpaul	2	3.4	27, 50		1 (5.6)	1 (3.4)
B	Typhimurium	2	3.4	19	1 (9.1)		1 (3.4)
B	Paratyphi B var. L(+) tartrate+	1	1.7	43	1 (9.1)		
B	Schwarzengrund	1	1.7	96			1 (3.4)
B	Stanley	1	1.7	29			1 (3.4)
C	Infantis	6	10.3	32	1 (9.1)	3 (16.7)	1 (3.4)
C	Bareilly	5	8.6	203	2 (18.2)	2 (11.1)	1 (3.4)
C	Oranienburg	2	3.4	23	2 (18.2)		

C	Montevideo	2	3.4	4			2 (6.9)
C	Menston	1	1.7	7,031		1 (5.6)	
C	Newport	1	1.7	45		1 (5.6)	
C	Ohio	1	1.7	329			1 (3.4)
C	Singapore	1	1.7	501			1 (3.4)
D	Enteritidis	21	36.2	11	3 (27.3)	8 (44.4)	10 (34.5)
D	Berta	1	1.7	435			1 (3.4)
D	Javiana	1	1.7	1547	1 (9.1)		
E	Give	1	1.7	516		1 (5.6)	
O	Alachua	1	1.7	2061			1 (3.4)

Supplemental Data Table S3. Genetic factors associated with an increased risk of extra-intestinal invasion

Sample ID	Extra-intestinal invasion	Serotype	MLST	No. VFs	SPI-1 (U16278)	SPI-2 (JN673269)	SPI-5 (AF323077)	SPI-13 (AY956833)	*cdtB-pltA-pltB*	*fepC*	*tcpC*	*STM0271*	*STM0275*
KPNTS36	Yes	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS14	Yes	Give	516	7	.	Yes	Yes	Yes	Yes	Yes	.	Yes	Yes
KPNTS37	Yes	I 4:b:-	2814	5	Yes	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS42	Yes	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS01	Yes	Javiana	1547	5	.	Yes	.	Yes	Yes	Yes	.	Yes	.
KPNTS49	Yes	Montevideo	4	8	Yes	Yes	Yes	Yes	Yes	Yes	.	Yes	Yes
KPNTS05	Yes	Oranienburg	23	9	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KPNTS08	Yes	Oranienburg	23	9	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
KPNTS02	Yes	Paratyphi B var. L(+) tartrate+	43	5	Yes	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS17	Yes	Saintpaul	50	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS54	Yes	Schwarzengrund	96	7	.	Yes	Yes	Yes	Yes	Yes	.	Yes	Yes
KPNTS51	Yes	Singapore	501	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS48	No	Alachua	2061	5	.	Yes	Yes	.	.	Yes	.	Yes	Yes
KPNTS09	No	Bareilly	203	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS10	No	Bareilly	203	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS21	No	Bareilly	203	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS26	No	Bareilly	203	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS34	No	Bareilly	203	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS31	No	Berta	435	5	.	Yes	.	Yes	.	Yes	.	Yes	Yes
KPNTS04	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS07	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS11	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS12	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS13	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS15	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS20	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS24	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS25	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.

KPNTS27	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS28	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS40	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS41	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS43	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS44	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS45	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS46	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS50	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS55	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS58	No	Enteritidis	11	0	.	.	.	.	.	.	.	.	.
KPNTS29	No	I 4,[5],12:i:-	19	5	Yes	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS30	No	I 4,[5],12:i:-	2379	5	Yes	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS47	No	I 4,[5],12:i:-	36	3	.	.	.	Yes	.	Yes	.	Yes	.
KPNTS53	No	I 4,[5],12:i:-	34	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS57	No	I 4,[5],12:i:-	34	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS38	No	I 4:b:-	42	5	Yes	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS06	No	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS16	No	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS19	No	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS22	No	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS52	No	Infantis	32	4	.	Yes	.	Yes	.	Yes	.	Yes	.
KPNTS23	No	Menston	7031	6	Yes	Yes	Yes	.	.	Yes	.	Yes	Yes
KPNTS33	No	Montevideo	4	8	Yes	Yes	Yes	Yes	Yes	Yes	.	Yes	Yes
KPNTS18	No	Newport	45	3	.	.	.	Yes	.	Yes	.	Yes	.
KPNTS32	No	Ohio	329	5	Yes	Yes	Yes	.	.	.	.	Yes	Yes
KPNTS39	No	Saintpaul	27	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS56	No	Stanley	29	2	.	.	.	Yes	.	Yes	.	.	.
KPNTS03	No	Typhimurium	19	4	.	.	.	Yes	.	Yes	.	Yes	Yes
KPNTS35	No	Typhimurium	19	4	.	.	.	Yes	.	Yes	.	Yes	Yes

Abbreviations: MLST, multilocus sequence type; VFs, virulence factors.