A Thesis for the Degree of Master of Science

Genome and Transcriptome Analyses of *Salmonella enterica* Virchow FORC_038 Isolated from Raw Chicken Meat

생닭에서 분리한 *Salmonella enterica* Virchow FORC_038에 대한 유전체 및 전사체 분석

February, 2018

Jin Su Song

Department of Agricultural Biotechnology
College of Agriculture and Life Sciences
Seoul National University

석사학위논문

Genome and Transcriptome Analyses of *Salmonella enterica* Virchow FORC_038 Isolated from Raw Chicken Meat

지도교수 최 상 호

이 논문을 석사학위논문으로 제출함

2018년 2월

서울대학교 대학원

농생명공학부

송 진 수

송진수의 석사학위논문을 인준함

2018년 2월

위원장 장 판 식 (인)

부위원장 최 상 호 (인)

위원 강 동 현 (인)

Abstract

Salmonella enterica is an ever-present worldwide public health threat and causes a range of diseases in humans such as gastroenteritis and invasive systemic infection. Especially, *Salmonella enterica* serovar Virchow shows antibiotic resistance and has been identified as an important cause of poultry meat outbreaks. The genome of *S.* Virchow FORC_038, which was obtained from raw chicken meat was completely sequenced using Illumina MiSeq and PacBio RS II platform. The genome consists of a circular chromosome of 4,938,076 bp with a GC content of 51.92%. The chromosome contains 4,632 open reading frames, 22 rRNAs, and 85 tRNA genes. BLAST analysis against VFDB (Virulence Factor Database) identified that the genome of FORC_038 contains the genes encoding SPI-1 (*Salmonella* pathogenic island 1) and SPI-2 effectors. Average nucleotide identity (ANI) analysis of FORC_038's genome with 40 other completely sequenced *Salmonella* genomes showed that the genome of FORC_038 is most closely related to those of *S.* Infantis 1326/28, isolated from chicken in 1973, UK. Comparative genome analysis of FORC_038 and *S.* Infantis 1326/28 revealed that the gene encoding to beta-lactamase, which contributes to resistance of beta-lactam antibiotics, and type III secretion system *Salmonella* outer protein (SopE), was also included on the genome of FORC_038 (FORC38_1065, FORC38_2974). To examine the role of SopE, the *sopE* mutant was constructed by integration of chloramphenicol resistance cassette into the chromosome of FORC_038 using red-mediated homologous recombination. Compared to the wild type, the *sopE* mutant displayed significantly reduced

invasiveness toward the HeLa cells, indicating that SopE is essential for the invasion of *S.* Virchow to the host cells. Additional genomic analysis revealed that genes related to resistance to β-lactams (FORC38_0096, 1541, 3615), tetracyclines (FORC38_1059), aminoglycosides (FORC38_1321), macrolides (FORC38_3090, 3091), and multiple antibiotic resistance protein (FORC38_1327, 1328, 2505, 2506, 2507, 2508) were widely disseminated on the chromosome of FORC_038, implying that the strain might be resistant to various antibiotics. *In silico* prediction was phenotypically confirmed by the subsequent Kirby-Bauer Disk Susceptibility Test, which indicated that FORC_038 is resistant to the antibiotics *bona fide*. To additional analyze total mRNA profiles of the strain upon contact with chicken breast in raw chicken part, which is highly consumed in Korea, RNA-Sequencing was used. After that, transcriptome analysis hinted that *S.* Virchow mainly uses the raw chicken meat as a reservoir for growth and survival according to up regulation of genes related to adhesion, iron uptake, amino acid metabolisms and down regulation of amino acid synthesis, carbohydrates metabolism. In addition, all of the SPI1 effector encoding genes that are major factors of *Salmonella* virulence and host cell invasiveness were upregulated including *sopE*. The results suggested the FORC_038 strain may be pathogenic and there is high risk of infection when humans consume chicken meat contaminated with FORC_038. Providing new insight on *S.* Virchow, this report will support further research in prevention and epidemiological investigation of *Salmonella* outbreaks in South Korea.

Key words: *Salmonella enterica* Virchow, Genomics, Transcriptomics, Whole genome sequencing, RNA sequencing

Student Number: 2016-21729

Contents

List of Figures

List of Tables

I. INTRODUCTION

The importance of studying *Salmonella* stems from the yearly recurrence of infections and the dangerous consequences of the infections. *Salmonella enterica* is an ever-present worldwide public health threat and causes a range of diseases in humans such as gastroenteritis and invasive systemic infection. Especially, *S. enterica* serovar Virchow shows antibiotic resistance and has been identified as an important cause of poultry meat outbreaks. In South Korea, the incidence of *S.* Virchow infections in humans has increased over the years, necessitating a nationwide survey of antimicrobial drug resistance in *S.* Virchow isolates. During 2005–2014 in South Korea, local public health laboratories participating in the national surveillance network isolated 68 *Salmonella* Virchow strains from feces samples from patients with acute diarrhea. Until 2010, <5 *S.* Virchow strains were isolated per year, but this number gradually increased to 17 in 2014. *S.* Virchow consistently ranked among the top 10 serotypes in prevalence during each study year in South Korea, accounting for ≈1.5%–2% of salmonellosis cases. (Kim *et al*., 2016). The regular occurrence of *S.* Virchow combined with the popularity of poultry meat especially raw chicken meat that have been consumed a lot is proving to be a threat to food safety and public health. There is room for improvement in terms of how the raw chicken meat has been prepared and served to customers. Needless to say, the process should be evidence-based and high throughput sequencing technologies can be a useful addition to the

toolbox. There are many benefits of the omics techniques in their applications to food safety. Genomics can be used to characterize pathogens, and transcriptomics can be used to characterize the pathogen's response to a stimulus such as stress and antimicrobial treatments (Bergholz *et al*., 2014). At the time of writing this manuscript, there are no genome sequences of *S.* Virchow on National Center for Biotechnology Information (NCBI), compared to the hundreds of other *Salmonella* whole genomes. More genome data of the species needs to be accumulated and analyzed in order to help prevent future outbreaks. Along with whole genome sequencing, RNA sequencing will be a useful method to characterize the transcripts of bacteria when in contact with model food raw chicken meat, because *S.* Virchow FORC_038 was isolated from raw chicken meat by Ministry of Food and Drug safety in South Korea. In this study, the bacterial genome was completely sequenced and analyzed, and its transcriptome was sequenced to uncover the genetic program that is set in motion upon contact with raw chicken meat. It is hoped that this work will further our understanding on how *S.* Virchow causes diseases and help prevent *S.* Virchow outbreaks in the future.

II. MATERIALS AND METHODS

Bacterial strains, plasmids, and growth condition

Table 1 shows the *Salmonella* and *E. coli* strains, and plasmids that were used in this study. The strains were both aerobically incubated at 37°C for 12 h in Luria-Bertani (LB) medium.

Genomic DNA extraction and identification

DNeasy Blood & Tissue Kit (QIAGEN, Valencia, California, USA) was used according to the manufacturers' protocol to extract genomic DNA. The 16S rRNA gene was amplified from the extracted genomic DNA and sequenced by an automated ABI3730XL capillary DNA sequencer (Applied Biosystems, Foster City, CA, USA) for taxonomic identification (Ku *et al*., 2014).

Virulence gene-specific PCR

Table 2 shows the primers of the genes (*invA, stn, hilA, sirA*) that were used for virulence gene-specific PCR. After resolving the amplified DNA fragments in 1.5% agarose gel, the electrophoresis results were visualized by using Geldoc™ EZ Image (Bio-Rad, Richmond, California, USA).

Invasion assay

To determine bacterial invasion of HeLa cells, an invasion assay was performed as described previously (Kim *et al.,* 2010). Briefly, bacteria were prepared by transferring a 1% inoculum from overnight cultures into fresh LB medium, followed by incubation for 2h at 37°C with constant shaking. *Salmonella* cells were collected by centrifugation at 10,000 x g, washed with phosphate-buffered saline (PBS) (pH 7.4), and resuspended in 1 ml of pre-warmed fresh Dulbecco's Modified Eagle's medium (DMEM) (Gibco). HeLa cells were seeded in DMEM supplemented with 20% fetal bovine serum (FBS) (Seradigm) in 24-well tissue culture plates with a cell density of 2.5 x 10^5 per well. The cell monolayers were incubated for 1 day, infected with bacteria at a multiplicity of infection (MOI) of 10 and 20, incubated for 30 min in the presence of 5% CO2. After the cells were washed once with PBS, fresh medium containing gentamicin (100 µg/ml; Sigma) was added, and the plates were further incubated for 30 min, followed by three washes with PBS. Then, 500 µl of 1% Triton X-100 was added, and the plates were incubated for a further 15 min before the bacteria were collected and plated on LB agar.

Transmission electron microscope

FORC_038 was negatively stained with Phosphotungstic acid (PTA) for a few seconds and then washed two to three times. The organism was observed using LIBRA 120 (Carl Zeiss, Oberkochen, Germany) transmission electron microscope (TEM) at 120 kV.

Genome sequencing and annotation

Whole genome sequencing and assembly were performed at ChunLab Incorporation (Seoul, South Korea). Dual platforms of Illumina MiSeq (Illumina, San Diego, California, USA), and PacBio (Pacific Biosciences, Menlo Park, California, USA) were used. The raw sequence reads from Illumina MiSeq were assembled with CLC Genomics workbench 7.5.1 (CLC Bio, Aarhus, Denmark), and those from PacBio were assembled with PacBio SMRT Analysis 2.3.0 (Pacific Biosciences). ORFs were predicted and annotated using GeneMarkS (Besemer, Lomsadze, and Borodovsky, 2001) and rapid annotations using subsytems technology (RAST) (Aziz *et al*., 2008). The ribosomal binding sites (RBSs) were predicted using RBS finder (J. Craig Venter Institute, Rockville, Maryland, USA). The existing annotations were modified after cross-checking with results from InterProScan 5 (Jones *et al*., 2014) and GAMOLA (Altermann and Klaenhammer, 2003). The end product was submitted to GenBank. Using virulence factor database (VFDB; http://www.mgc.ac/VFs/) as reference, putative virulence factors of FORC_038 were found by basic local alignment search tool (BLAST). The circular genome map of the chromosome was visualized by using cluster of orthologous group (COG)-based WebMGA and GenVision (DNASTAR, Madison, Wisconsin, USA).

Average nucleotide identity analysis and comparative genome analysis

To characterize the genetic relatedness amongst the currently sequenced *Salmonella* strains, average nucleotide identity (ANI) analysis was conducted. The completed

genome sequences of 40 strains (*S.* Newport FORC_020, USDA-1927, CVM21550, SL254; *S.* Typhimurium DT2, LT2, FORC_030, FORC_058; *S.* Heidelberg B182, SL476; *S.* Thompson ATCC 8391, RM6836; *S.* Anatum ATCC 1592, USDA; *S.* Gallinarum RK55078, CDC; *S.* Enteritidis FORC_075, FORC_052, FORC_019, FORC_050, FORC_051, P125109, FORC_007, FORC_074; *S.* Infantis 1326/28; *S.* Virchow FORC_038; *S.* Choleraesuis C500, SCB67; *S.* Tennessee TXSC; *S.* Mbandaka FORC_015; *S.* Agona 46000421, SL483; *S.* Montevideo 507440, USDA; *S.* Typhi CT18, Pstx12) were used for the analysis. The ANI values were calculated using Jspecies (Richter and Móra, 2009). The result was visualized with R. As *S.* Infantis 1326/28 was determined to be the most closely related to FORC_038, comparative genome analysis between FORC_038 and Infantis 1326/28 was carried out. Artemis comparison tool (ACT) was used (Carver *et al*., 2005).

Kirby-Bauer disk susceptibility test

The susceptibility of *S.* Virchow FORC_038 to various antibiotics was identified through the Kirby-Bauer disk susceptibility test. The diameters of clear zone formed by each antibiotic were interpreted as susceptible, intermediate, or resistant according to the recommendations of the Clinical and Laboratory Standards Institute (CLSI) as described by Bauer et al. Lawrence et al. (Lawrence *et al.*, 1972; Bauer *et al*., 1966) with Mueller-Hinton agar (Difco). Briefly, bacteria were prepared by transferring single colonies incubated 12 hours over from LB agar into 0.85% saline water, followed by measure turbidity of FORC_038 inoculum using turbidimeter (Mettler

Toledo) according to McFarland standards 0.5 (Absorbance OD_{600}; 0.063) (Mcfarland *et al*., 1907). 1ml inoculum of FORC_038 was added to Mueller-Hinton agar plate, and the 10 antimicrobial disks (ampicillin, cefazolin, cefotaxime, ceftazidime, streptomycin, tetracycline, erythromycin, nalidixic acid, ciprofloxacin, amoxicillin/clavulanic acid) (Oxoid, UK) were placed on the same agar plate on which *S.* Virchow FORC_038 cells were spread. *E. coli* ATCC 25922 served as a quality control strain in the disk susceptibility test. The plates were incubated for 18 hours over to confirm antibiotic resistance of the bacteria.

Construction of a *sopE* mutant using the Lambda-Red recombination method.

The whole genome sequence of *S.* Virchow FORC_038 was obtained and manipulated for construction of *sopE* mutant. Site-specific mutation of *S.* Virchow FORC_038 was generated by the Lambda-Red recombination method, as described by Kim et al. and Eatsenko and Wanner (Datsenko & Wanner, 2000; Kim *et al*., 2010) Briefly, the chloramphenicol resistance cassette from plasmid pKD3 was amplified using the following primers (Table 2): for *sopE* deletion mutant construction, *sopE*_F (5'-ATA TAT AAA TGA GTT ATG TAC ATA TAA AAG GAG CAT TAC CGT GTA GGC TGG AGC TGC TTC) and *sopE*_R (5'- GGA CCA TGG CTA ATT CCC ATA CAC AGA AAA ACC AAA AAA TAT GCG GAG CCT CTT CCT GAT). (The nucleotide sequences originating from pKD3 are underlined, and those from the *S.* Virchow FORC_038 gene of interest are shown in italics.) The PCR products were transformed into the wild-type (WT) strain, *S.* Virchow FORC_038 harboring the

pKD46 plasmid by electroporation and cells were selected for chloramphenicol-resistance transformants, *sopE*::*cat*.

RNA extraction

FORC_038 was first prepared so that it can be added to the control and the experimental conditions: FORC_038 was grown to mid-log phase (OD_{600}; 0.8), and centrifuged at 10,000 x g for 1 min. After removing the supernatant, the pellet was washed with 1x phosphate buffered saline (PBS), and the resulting solution was centrifuged at 10,000 x g for 1 min. The wash and centrifugation step was repeated two or three times. Lastly, the pellet was re-suspended with M9 minimal medium containing histidine 2% (Sirsat *et al*., 2011). Prepared FORC_038 was incubated at 37°C for 2 h in M9 minimal medium either with or without raw chicken meat. After incubation, the culture was centrifuged at 10,000 x g at 4°C for 1 min. The pellet was re-suspended with 5 ml of diethylpyrocarbonate (DEPC) treated cold 1x PBS and the solution was quickly mixed with 10 ml of RNAprotect® Bacteria Reagent (RBR) (QIAGEN). Total RNAs were extracted from the RBR-treated samples under 4°C using RNeasy Mini Kit (QIAGEN) according to the manufacturers' protocol. To prevent DNA contamination, TURBO DNase (AMbion, Austin, Texas, USA) was used. The quality of RNAs were confirmed by ChunLab Incorporation, with Agilent 2100 Bioanalyzer and Agilent RNA 6000 Nano reagents (Agilent Technologies, Waldbronn, Germany)

Strand-specific cDNA library construction and RNA sequencing

Strand-specific complementary DNA (cDNA) library construction and RNA sequencing were performed by ChunLab Incorporation. rRNAs were depleted by using Ribo-Zero™ rRNA Removal Kit (Epicentre, Madison, Wisconsin, USA). Enriched mRNA and TruSeq Stranded mRNA Sample Preparation kit (Illumina) was used to construct cDNA library. The quality of cDNA libraries was checked by ChunLab Incorporation, with Agilent 2100 Bioanalyzer and Agilent DNA 1000 reagents (Agilent Technologies). Strand-specific paired-ended 100 nucleotide reads from each cDNA library were obtained using Illumina HiSeq2500. For biological replication, three cDNA libraries were constructed and sequenced from three independent extracted RNAs of FORC_038.

Transcriptomic data analysis

The reads obtained from RNA sequencing were mapped to the reference genome of FORC_038 by using CLC Genomics Workbench 5.5.1 (CLC Bio). The GenBank (http://www.ncbi.nlm.nih.gov) accession numbers of the reference genome are CP015574 (Chromosome). The relative transcript abundance was measured by the reads per kilobase (kb) of transcript per million mapped sequence read (RPKM) (Mortazavi *et al*., 2008). The genes were considered to be differentially expressed when the fold change was equal to 2 or greater and *P*-value was < 0.05. The volcano plot and the heat map were created to visualize the RNA sequencing result using CLC Genomics Workbench 5.5.1 (CLC Bio) and Gitools (Biomedical Research Park,

Barcelona, Spain; Perez-Llamas and Lopez-Bigas, 2011).

RNA purification and transcript analysis

Total RNA from FORC_038 was isolated using RNeasy Mini Kit (QIAGEN). The culture condition was the same as above (See the 'RNA extraction' part). cDNA was synthesized using iScripit™ cDNA Synthesis Kit (Bio-Rad). The quantitative real-time PCR (qRT-PCR) of the synthesized cDNA was conducted using iQ™ SYBR® Green Supermix (Bio-rad) and the Chromo 4 Real-Time PCR detection system (Bio-rad) as described previously (Lim and Choi, 2014). The genes and their specific primers used are listed in Table 2. The relative expression levels of the specific transcripts were calculated by using the 16S rRNA expression level as the internal reference for normalization. The experiment was conducted with biological duplicates and for each biological sample, technical triplicate wells were used.

Growth kinetics of FORC_038

The growth condition and bacteria preparation step was the same as described above (See the 'RNA extraction' part), except that the incubation times were 0 h, 1 h, 2 h. During incubation at each time points, the incubated culture was collected, and serially diluted by 1X PBS and plated on LB agar. The plates were incubated in 37°C for 12 h and the colonies were counted. The experiments were run with biological duplicates and for each biological sample, technical triplicates were used.

Table 1. Bacterial strains and plasmids used in this study

Strain or plasmid	Relevant characteristics	Reference or Source
Bacterial strains		
Salmonella		
Virchow FORC_038	Wild type, environmental isolate	Ministry of Food and Drug Safety
JS171	FORC_038 harboring pKD46	This study
JS172	FORC_038 with *sopE*::*cat*	This study
Typhimurium SL1344	Virulent strain, Sm^R	Laboratory collection
Typhimurium LT2	Principal strain for cellular and molecular biology	Laboratory collection
Enteritidis 2000672	Environmental isolate	Ministry of Food and Drug Safety
E. coli		
ATCC 25922	CLSI control strain for antimicrobial susceptibility testing	Laboratory collection
DH5α	*supE44 DlacU169 (f80 lacZ DM15) hsdR17 recA1 endA1 gyrA96 relA1 thi-1 relAl*; plasmid replication	Laboratory collection
Plasmid		
pKD46	*oriR101 repA101(*Ts) Amp^R *ara BADpgam-bet-exo*	(Datsenko & Wanner, 2000)
pKD3	*oriR6K* Amp^R FRT Cm^R FRT	(Datsenko & Wanner, 2000)

Table 2. Oligonucleotides used in this study

Oligonucleotide name	Oligonucleotide sequence (5` → 3`)	Purpose	Predicted size (bp)
[a] **For virulence gene-specific PCR screening**			
*invA*_F	ACAGTGCTCGTTTACGACCTGAAT	PCR analysis of *invA*	244
invA _R	AGACGACTGGTACTGATCGATAAT		
*stn*_F	CTTTGGTCGTAAAATAAGGCG	PCR analysis of *stn*	260
*stn*_R	TGCCCAAAGCAGAGAGATTC		
*hilA*_F	CTGCCGCAGTGTTAAGGATA	PCR analysis of *hilA*	497
*hilA*_R	CTGTCGCCTTAATCGCATGT		
*sirA*_F	GCCGTACTAACGCCGTTGAC	PCR analysis of *sirA*	430
*sirA*_R	TAGCGATAGCTGTTCACCGT		
Mutant construction			
*bet*_F	AACTCGCGAAGGCAGAGAAA	JS171 confirmation	332
*bet*_R	AGAACTGACACAGGCCGAAG		

Oligonucleotide name	Oligonucleotide sequence (5` → 3`)	Purpose	Predicted size (bp)
*sopE*_F	ATATATAAATGAGTTATGTACATATAA AAGGAGCATTACCGTGTAGGCTGG AGCTGCTTC	JS172 construction	1113
*sopE*_R	GGACCATGGCTAATTCCCATACAC AGAAAAACCAAAAAATATGCGGAG CCTCTTCCTGAT		
[b] For qRT-PCR			
FORC38_0328_qRT_F	AACGCAAATCAGCCATACGC	qRT-PCR of FOR38_0328	106
FORC38_0328_qRT_R	GTCCGGTTCTTCCGGGATAC		
FORC38_0959_qRT_F	TGGGGCGGAATATCATGACG	qRT-PCR of FOR38_0959	130
FORC38_0959_qRT_R	GTTTGGCGGCACTACGTTTT		
FORC38_0966_qRT_F	GAACAGACGCTGCCTTTTGG	qRT-PCR of FOR38_0966	133
FORC38_0966_qRT_R	AAACTTCGCCATACCAGCCA		
FORC38_0969_qRT_F	AATGCGTAAGGCGGCAAAAG	qRT-PCR of FOR38_0969	128
FORC38_0969_qRT_R	TAGCGCCTCCAGATAGACCA		
FORC38_0970_qRT_F	GCTGATTTTATGCTCGCCCG	qRT-PCR of FOR38_0970	101
FORC38_0970_qRT_R	AGTTCCGCCGTTACCTTCTG		

FORC38_1331_qRT_F	TAAGGCATGGGCGGAGAAAA	qRT-PCR of FOR38_1331	108
FORC38_1331_qRT_R	TGACCCTGAACATCGGCTTC		
FORC38_1332_qRT_F	CCCGCGTTTGTCTATCTGGT	qRT-PCR of FOR38_1332	184
FORC38_1332_qRT_R	GTACAATCGGCGGTAGAGCA		
FORC38_1372_qRT_F	GCTACGGGTGGTACGAATGG	qRT-PCR of FOR38_1372	110
FORC38_1372_qRT_R	GCATCAGCACCAGTAAAGCC		
FORC38_1395_qRT_F	AGTTCGGCGTGGAAGTGATT	qRT-PCR of FOR38_1395	146
FORC38_1395_qRT_R	GGTACGACCGTTCAGCGTAT		
FORC38_2209_qRT_F	GCCCTGAATCCTGGGTTGAA	qRT-PCR of FOR38_2209	131
FORC38_2209_qRT_R	GGACAATAGTGGTGGTGCGA		
FORC38_2463_qRT_F	CTGGGAGGATAGACAACGGC	qRT-PCR of FOR38_2463	131
FORC38_2463_qRT_R	GTCCCATGCCCTCAAACGTA		
FORC38_2974_qRT_F	TCTGCCCGGAACAACCAAAA	qRT-PCR of FOR38_2974	206
FORC38_2974_qRT_R	GAGTCGGCATAGCACACTCA		
FORC38_3311_qRT_F	ATGACGTGACCACTATCCGC	qRT-PCR of FOR38_3311	144
FORC38_3311_qRT_R	CGTTCGGCGCTTTATGGAAG		

FORC38_3801_qRT_F	GATGTGGGTCTTCGCACTGA	qRT-PCR of FOR38_3801	144
FORC38_3801_qRT_R	GCCGTCGCTAAAATCGAACC		
FORC38_4005_qRT_F	GAGCTGGAAGGCCGTATCTC	qRT-PCR of FOR38_4005	143
FORC38_4005_qRT_R	GCCAGTACCCATAGCCACTC		
FORC38_4155_qRT_F	TCGGTCTGAATCTCCCGTCT	qRT-PCR of FOR38_4155	140
FORC38_4155_qRT_R	GTCGTAGTTGTCGATGCGGT		
FORC38_4358_qRT_F	GACTATGCTGGCCGATGTGA	qRT-PCR of FOR38_4358	80
FORC38_4358_qRT_R	GGACATCCAGCTTAACGGCT		
FORC38_4369_qRT_F	TATGACGACGCGACTAACGG	qRT-PCR of FOR38_4369	166
FORC38_4369_qRT_R	CGCGCGCTTTCACTATTTCA		
FORC38_4533_qRT_F	GTTCCATTTTCCCGCACCAC	qRT-PCR of FOR38_4533	103
FORC38_4533_qRT_R	GTTCCATTTTCCCGCACCAC		
FORC38_rrs_qRT_F	TGCCTGATGGAGGGGGATAA	qRT-PCR of FOR38_rrs (16S rRNA)	189
FORC38_rrs_qRT_R	TTCCAGTGTGGCTGGTCATC		

[a] The oligonucleotides were designed using *S.* Typhimurium LT2 genome sequence (GenBank™ accession numbers AE006468).

[b] The oligonucleotides were designed using *S.* Virchow FORC_038 genome sequence (GenBank™ accession numbers CP015574).

III. RESULTS

Virulence gene-specific PCR screening

To identify the virulent properties of FORC_038, which is an environmental strain, major virulence genes of *Salmonella* spp. were amplified using PCR. *S.* Typhimurium LT2 were used as positive controls. The target genes and their PCR product sizes are listed in Table 2. All genes were detected in 3 strains (Fig. 1)

Invasion activity of S. Virchow FORC_038

As FORC_038 is an environmental strain, it is necessary to estimate its virulence. Therefore, invasion assay was performed using HeLa cell lines. The HeLa cells were infected with FORC_038 and, the virulent strain, *S.* Typhimurium SL1344 was used as the positive control at various multiplicity of infections (MOIs) for 30 min. The result showed that FORC_038 has relative high invasion activity than SL1344 (Fig. 2).

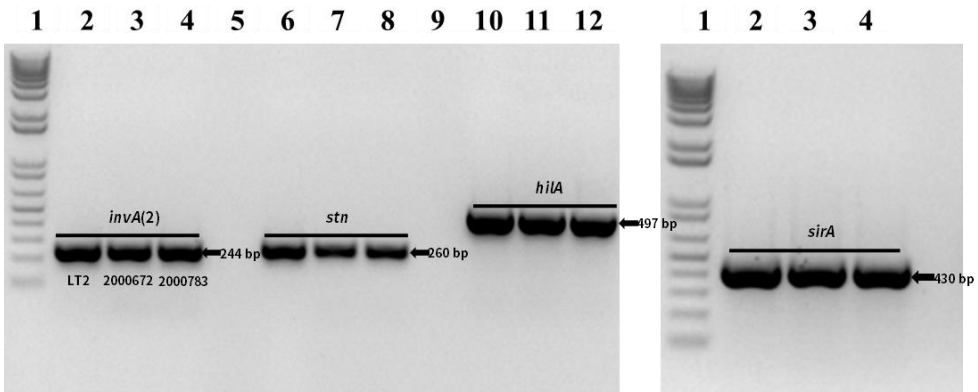


Fig 1. Virulence gene-specific PCR screening of *invA, stn, hilA, sirA* in *Salmonella* strains. 4 virulence genes of *Salmonella* were amplified using PCR. Each band sizes are specified as part of the figure. *invA*, invasion gene A encoding gene; *stn*, enterotoxin encoding gene; *hilA*, SPI1 virulence correlated gene; *sirA*, a global regulator of invasion genes. Molecular weight marker, lane 1; control, *S.* Typhimurium LT2, lane 2, 6, 10; *S.* Virchow FORC_038, lane 3, 7, 11; *S.* Enteritidis 2000672, lane 4, 8, 12.

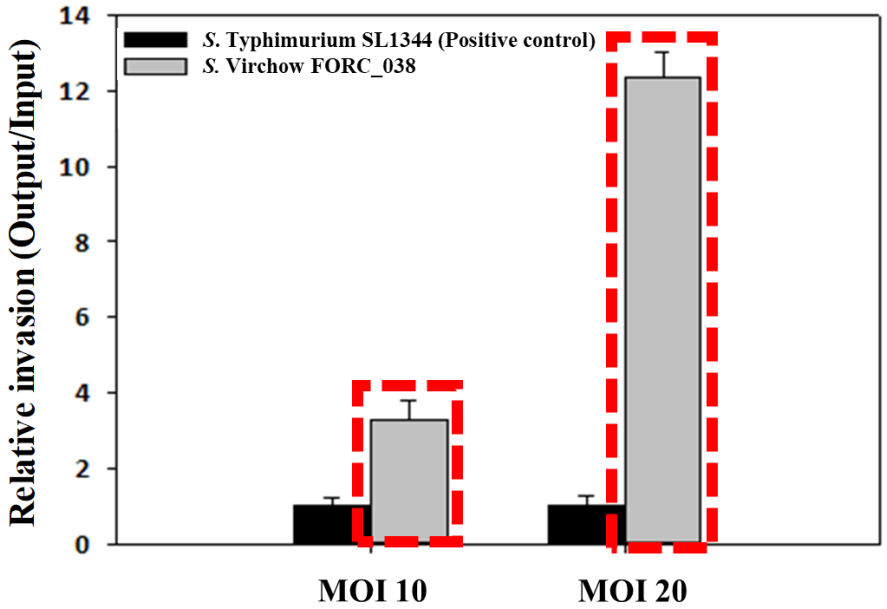


Fig 2. Invasion activity of *S.* Virchow FORC_038. Invasion activities of *Salmonella* strains were determined by the invasion assy. HeLa cells were infected with *S.* Virchow FORC_038 and *S.* Typhimurium SL1344 at multiplicity of infection (MOI) of 10 and 20 for 30 min were lysed with 1% (v/v) Triton X-100 to recover the intracellular *Salmonella*. Relative invasion of *S.* Typhimurium SL1344 was presented as 1. The bars represent the standard deviation.

Transmission electron microscopy (TEM) image of *S.* Virchow FORC_038

The conformation of FORC_038 was identified through transmission electron microscope (TEM). FORC_038 has curved and rod-shape with peritrichous flagella. (Fig. 3).

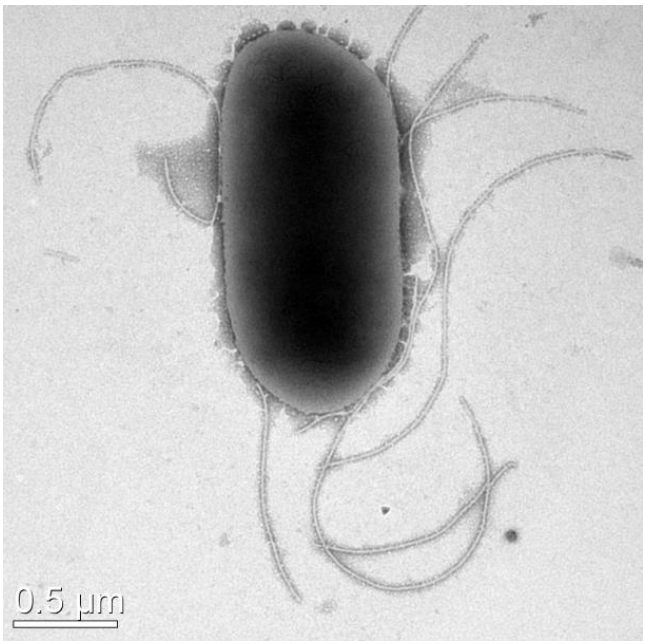


Fig 3. Transmission electron microscopy image of *S.* Virchow FORC_038. The cells were negatively stained with PTA. Bars, 0.5 μm.

Genome properties of *S.* Virchow FORC_038

The whole genome sequencing properties of *S.* Virchow FORC_038 were summarized in Table 3. The genome consists of a circular chromosome of 4,938,076 bp with a GC content of 51.92%, containing 4,735 predicted ORFs, 85 tRNA genes and 22 rRNA genes. Among the predicted ORFs, 4093 ORFs (86.44%) were predicted to encode the functional protein and 642 ORFs (13.56%) were expected to encode hypothetical proteins. The genome annotation information including a chromosome has been deposited under the GenBank accession number, CP015574. All the results are summarized in Table 4. Based on the bioinformatics analysis of the chromosome, circular genome map was drawn (Fig. 4) and genes with the specialized function such as putative virulence factors were noted.

Table 3. Summary of *S.* Virchow FORC_038 genome sequencing

Property	Term
Finishing quality	Finished
Libraries used	Illumina 300 base pair paired-end library
	PacBio SMRTbell™ library (> 10 kb) for draft assembly
Sequencing platform	Illumina MiSeq
	PacBio RS II
Assembler	CLC Genomics Workbench 7.5.1
	PacBio SMRT Analysis 2.3.0
Gene calling method	GeneMarkS
	RAST server
Average genome coverage	498.01X
Contig length (bp)	4,938,076
Contig No.	1
Scaffold No.	1
N50	4,938,076
Locus tag	FORC38
GenBank accession No.	CP015574
GenBank release date	2017-05-09
BioProject No.	PRJNA320341
Source material identifier	FORC_038
Project relevance	Agricultural

Table 4. Chromosomal properties of *S.* Virchow FORC_038

Property	Term
Sample	*S.* Virchow FORC_038
Classification	Chromosome
GenBank accession No.	CP015574
Genome size (bp)	4,938,076
Protein coding genes	4,735
Annotated genes	4,093
Hypothetical genes	642
tRNA	85
rRNA	22

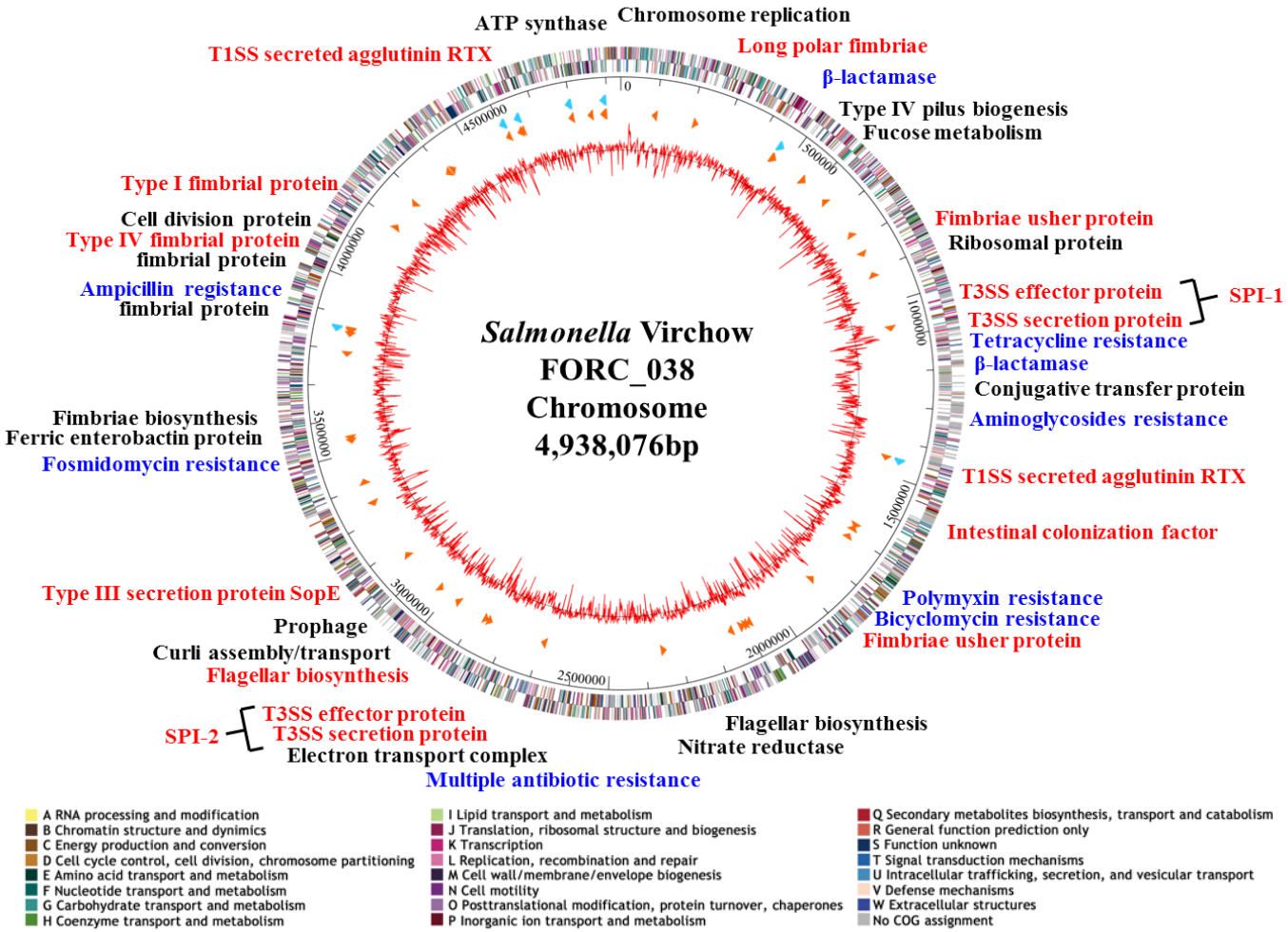

Salmonella Virchow
FORC_038
Chromosome
4,938,076bp
ATP synthase
Chromosome replication
T1SS secreted agglutinin RTX
Long polar fimbriae
β-lactamase
Type IV pilus biogenesis
Fucose metabolism
Type I fimbrial protein
Cell division protein
Type IV fimbrial protein
fimbrial protein
Ampicillin registance
fimbrial protein
Fimbriae usher protein
Ribosomal protein
T3SS effector protein
T3SS secretion protein
SPI-1
Tetracycline resistance
β-lactamase
Conjugative transfer protein
Aminoglycosides resistance
Fimbriae biosynthesis
Ferric enterobactin protein
Fosmidomycin resistance
T1SS secreted agglutinin RTX
Intestinal colonization factor
Type III secretion protein SopE
Prophage
Curli assembly/transport
Flagellar biosynthesis
Polymyxin resistance
Bicyclomycin resistance
Fimbriae usher protein
SPI-2
T3SS effector protein
T3SS secretion protein
Electron transport complex
Flagellar biosynthesis
Nitrate reductase
Multiple antibiotic resistance
0
500000
1000000
1500000
2000000
2500000
3000000
3500000
4000000
4500000
A RNA processing and modification
B Chromatin structure and dynimics
C Energy production and conversion
D Cell cycle control, cell division, chromosome partitioning
E Amino acid transport and metabolism
F Nucleotide transport and metabolism
G Carbohydrate transport and metabolism
H Coenzyme transport and metabolism
I Lipid transport and metabolism
J Translation, ribosomal structure and biogenesis
K Transcription
L Replication, recombination and repair
M Cell wall/membrane/envelope biogenesis
N Cell motility
O Posttranslational modification, protein turnover, chaperones
P Inorganic ion transport and metabolism
Q Secondary metabolites biosynthesis, transport and catabolism
R General function prediction only
S Function unknown
T Signal transduction mechanisms
U Intracellular trafficking, secretion, and vesicular transport
V Defense mechanisms
W Extracellular structures
No COG assignment

Fig 4. Genome map of *S.* Virchow FORC_038. The outer circle indicates the locations of all annotated open reading frames (ORFs), and the inner circle with the red peaks indicates GC content. Between these circles, sky blue and orange arrows indicate the rRNA operons and the tRNAs, respectively. All annotated ORFs are colored differently according to the clusters of orthologous groups (COG) assignments as indicated. Genes with specialized functions labeled with different colors as follows; virulence-related genes in red, antibiotic resistance related genes in blue, and prophage-related and other functional genes in black. SPI-1, *Salmonella* pathogenic island 1; T3SS, type III secretion system; T1SS, type I secretion system; RTX, repeats in toxin.

Pathogenesis and virulence factor

According to VFDB, virulence factors of *S.* Virchow FORC_038 can be categorized as those related to SPI1, SPI2 type III secretion systems, which are major factors for invasion of eukaryotic cells and ability of *S. enterica* to cause systemic infections and for intracellular pathogenesis (Velge *et al*., 2012, Hensel *et al*., 2000). Also, *S.* Virchow FORC_038 had genes related to macrophage inducible gene, magnesium uptake, fimbrial adherence, non-fimbrial adherence and two-component system. The macrophage inducible gene, *mig-14, i*s an inner membrane-associated protein, necessary for bacterial proliferation in the liver and spleen. (Valdivi *et al*., 2000). Magnesium uptake is essential for intracellular survival or for virulence (Moncrief *et al*., 1998). Fimbrial and non-fimbrial adherence related genes are required for intestinal colonization of *Salmonella*. Two component systems, PhoPQ regulon, are required for intracellular survival, cantionic antimicrobial peptides (CAMPs) resistance, and stimulation of cytokine secretion. (Gunn *et al.,* 1996, Guo *et al.,* 1997). Virulence factors of *S.* Virchow FORC_038 are specified in Table 5.

Table 5. Virulence factors of *S.* Virchow FORC_038

Virulence factor	Annotation	Location	Function
Fimbrial adherence determinants			
csgG, csgF, csgE, csgD, csgB, csgA, csgC	Curli fimbriae (*csg*)	2964613..2969061 (FORC38_2882 ~ FORC38_2888)	Adhesion, attachment
bcfA, bcfB, bcfC, bcfD, bcfE, bcfF, bcfG	Fimbrial protein / fimbrial chaperon (*bcf*)	4117730..4124470 (FORC38_3979 ~ FORC38_3985)	Adhesion, attachment
fimA, fiml, fimC, fimD, fimH, fimF, fimZ, fimY, fimW	Type 1 fimbriae (*fim*)	3538437..3547623 (FORC38_3434 ~ FORC38_3443)	Adhesion, attachment
lpfE, lpfD, lpfC, lpfB, lpfA	Long polar fimbriae (*lpf*)	214565..220053 (FORC38_0203 ~ FORC38_0207)	Adhesion, attachment
safA, safB, safC, safD	Outer membrane fimbrial usher protein / chaperone (*saf*)	3815012..3819359 (FORC38_3701 ~ FORC38_3704)	Adhesion, attachment
stbE, stbD, stbC, stbB, stbA	Putative usher fimbriae (*stb*)	3766891..3772849 (FORC38_3651 ~ FORC38_3655)	Adhesion, attachment
stcD, stcC, stcB, stcA	Fimbriae usher protein / chaperone (*stc*)	1945473..1950268 (FORC38_1852 ~ FORC38_1855)	Adhesion, attachment
stdC, stdB, stdA	Fimbriae usher protein / chaperone (*std*)	852151..856125 (FORC38_0810 ~ FORC38_0812)	Adhesion, attachment
steA, steB, steC, steD, steE, steF	Fimbriae usher protein / periplasmic fimbriae chaperone (*ste*)	942348..948044 (FORC38_0888 ~ FORC38_0893)	Adhesion, attachment
stfA, stfC, stfD, stfE, stfF, stfG	Fimbrial subunit / fimbrial chaperon (*stf*)	3909158..3914765 (FORC38_3793 ~ FORC38_3798)	Adhesion, attachment
sthE, sthD, sthC, sthB, sthA	Putative fimbrial protein / fimbrial assembly chaperone (*sth*)	4152983..4158566 (FORC38_4012 ~ FORC38_4015)	Adhesion, attachment
stiH, stiC, stiB, stiA	Fimbriae usher protein / chaperone (*sti*)	3943764..3948692 (FORC38_3825 ~ FORC38_3828)	Adhesion, attachment

stjB, stjC	Fimbriae usher protein / chaperone (*stj*)	4175287..4180489 (FORC38_4032 ~ FORC38_4036)	Adhesion, attachment
tcfA, tcfB, tcfC, tcfD	unidentified transcellular chemotactic factor (*tcf*)	3804620..3809881 (FORC38_3691 ~ FORC38_3694)	Adhesion, attachment
Macrophage inducible genes			
mig-14	Putative transcriptional regulator	1386846..1387742 (FORC38_1362)	Survival in the host
Magnesium uptake			
mgtB, mgtC	Magnesium transporter	77548..81171 (FORC38_0075 ~ FORC38_076)	Magnesium uptake
Nonfimbrial adherence determinants			
misL	Autotransporter	85085..87952 (FORC38_0081)	Adherence, intestinal colonization
ratB, shdA, sinH	Putative intestinal colonization factor	1559471..1568406 (FORC38_1501 ~ FORC38_1503)	Intestinal colonization, persistence
Secretion system *sprB, hilC, orgC, orgB, orgA, prgK, prgJ, prgI, prgH, hilD, hilA, iagB, sicP, iacP, sipD, sicA, spaS, spaR, spaQ, spaP, spaO, invJ, invI, invC, invB, invA, invE, invG, invF, invH*	SPI-1 encoded Type III secretion system component (T3SS-1)	1027712..1038762 (FORC38_0979 ~ FORC38_0990) 1027712..1028194 (FORC38_0979) 1025648..1026040 (FORC38_0976) 1025069..1025317 (FORC38_0974) 1021943..1022974 (FORC38_0972) 1004897..1018831(FORC38_0955 ~ FORC38_0969)	Secretion of effector proteins, invasion

ssrB, ssrA, ssaC, ssaD, ssaE, sseA, sseB, sscA, sseC, sseD, sseE, sscB, ssaG, ssaH, ssaI, ssaJ, ssaK, ssaL, ssaM, ssaV, ssaN, ssaO, ssaP, ssaQ, ssaR, ssaS, ssaT, ssaU	SPI-2 encoded Type III secretion system component (T3SS-2)	2719908..2723165 (FORC38_2632 ~ FORC38_2633) 2711373..2718930 (FORC38_2621 ~ FORC38_2630) 2708306..2709795 (FORC38_2615 ~ FORC38_2618) 2697843..2707743 (FORC38_2602 ~ FORC38_2613)	Secretion of effector proteins, replication of intracellular bacteria
slrP	T3SS effectors translocated via both systems	3286503..3288800 (FORC38_3201)	Translocated into eukaryotic host cells
sptP, sipA, sipC, sipB, sopA, sopB/sigD, sopD, sopE, sopE2, avrA	T3SS-1 translocated effectors	1026051..1026410 (FORC38_0977) 1023038..1025050 (FORC38_0973) 1018834..1021872 (FORC38_0970 ~ FORC38_0971) 2045460..2047809 (FORC38_1936 ~ FORC38_1937) 3016807..3018492 (FORC38_2938) 961122..962204 (FORC38_0905) 3055382..3056104 (FORC38_2974) 2239971..2240693 (FORC38_2151) 1038989..1039831 (FORC38_0991)	Actin cytoskeleton rearrangement, nuclear response
spiC/ssaB, sseF, sseI/srfH, sseJ, sseL, sifA, sifB, pipB, pipB2, sopD2, sspH2, sseK1	T3SS-2 translocated effectors	2718932..2719315 (FORC38_2631) 2709889..2711357 (FORC38_2619 ~ FORC38_2620) 2470320..2471546 (FORC38_2379) 1806788..1807810 (FORC38_1727) 2886233..2887243 (FORC38_2802) 2504440..2505390 (FORC38_2412) 3019398..3020273 (FORC38_2941) 1389436..1390536 (FORC38_1365) 3143283..3144242 (FORC38_3062) 1859285..1861651 (FORC38_1773) 4622609..4623637 (FORC38_4455)	Salmonella-containing vaculoles (SCVs) modification

Two-component system			
phoP, *phoQ*	Sensor / transcriptional regulator	2877724..2879861 (FORC38_2794 ~ FORC38_2795)	Transcriptional regulation

ANI analysis

ANI analysis was performed using the chromosomal sequences of total 40 *Salmonella* species: *S.* Newport FORC_020, USDA-1927, CVM21550, SL254; *S.* Typhimurium DT2, LT2, FORC_030, FORC_058; *S.* Heidelberg B182, SL476; *S.* Thompson ATCC 8391, RM6836; *S.* Anatum ATCC 1592, USDA; *S.* Gallinarum RK55078, CDC; *S.* Enteritidis FORC_075, FORC_052, FORC_019, FORC_050, FORC_051, P125109, FORC_007, FORC_074; *S.* Infantis 1326/28; *S.* Virchow FORC_038; *S.* Choleraesuis C500, SCB67; *S.* Tennessee TXSC; *S.* Mbandaka FORC_015; *S.* Agona 46000421, SL483; *S.* Montevideo 507440, USDA; *S.* Typhi CT18, Pstx12. The result showed that FORC_038 had the highest value of ANI (98.87) with *S.* Infantis 1326/28, isolated from chicken in UK, 1994 (Ferenc *et al*., 2015) (Fig. 5). Therefore, *S.* Infantis 1326/28 was selected for additional comparative genome analysis.

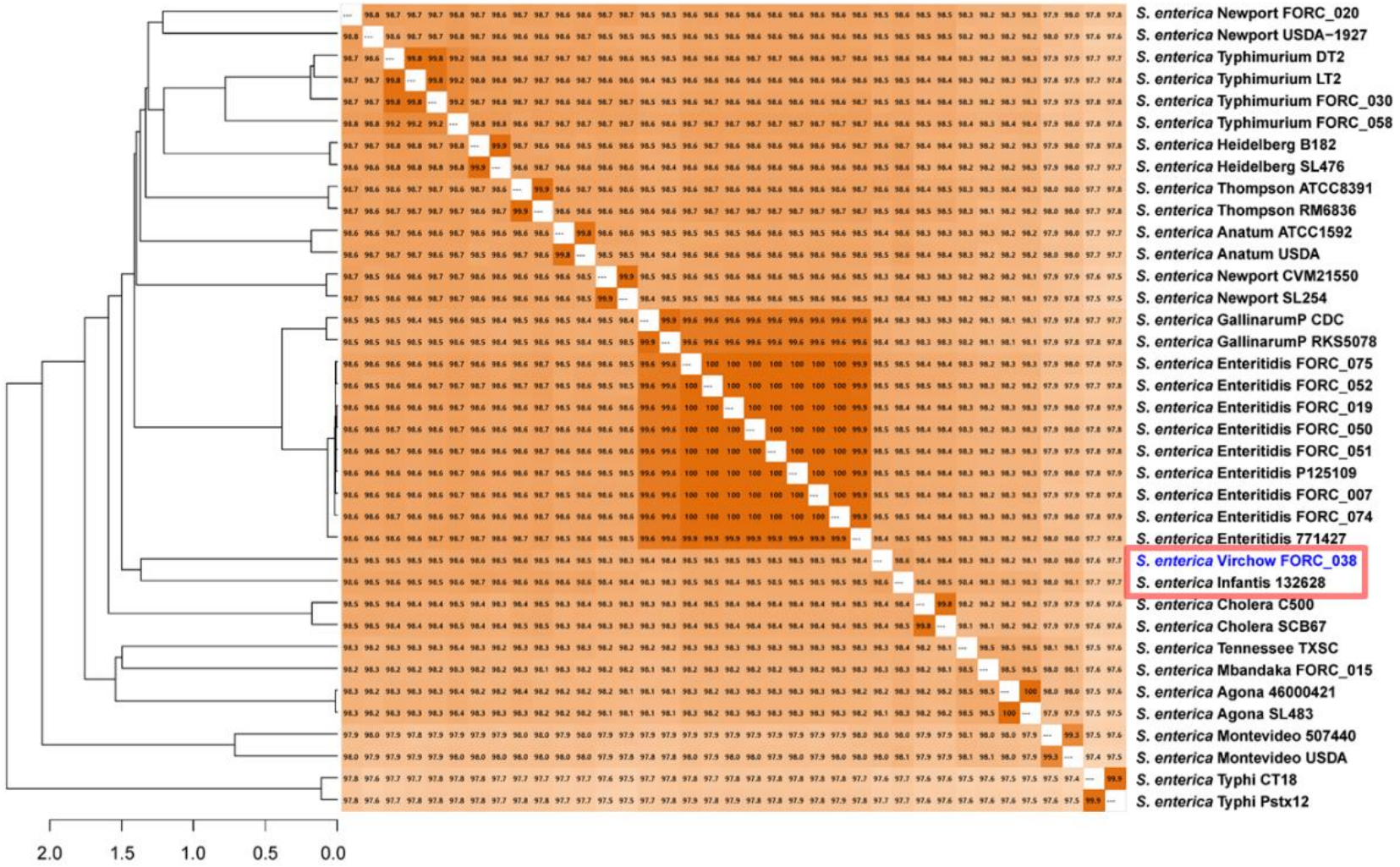
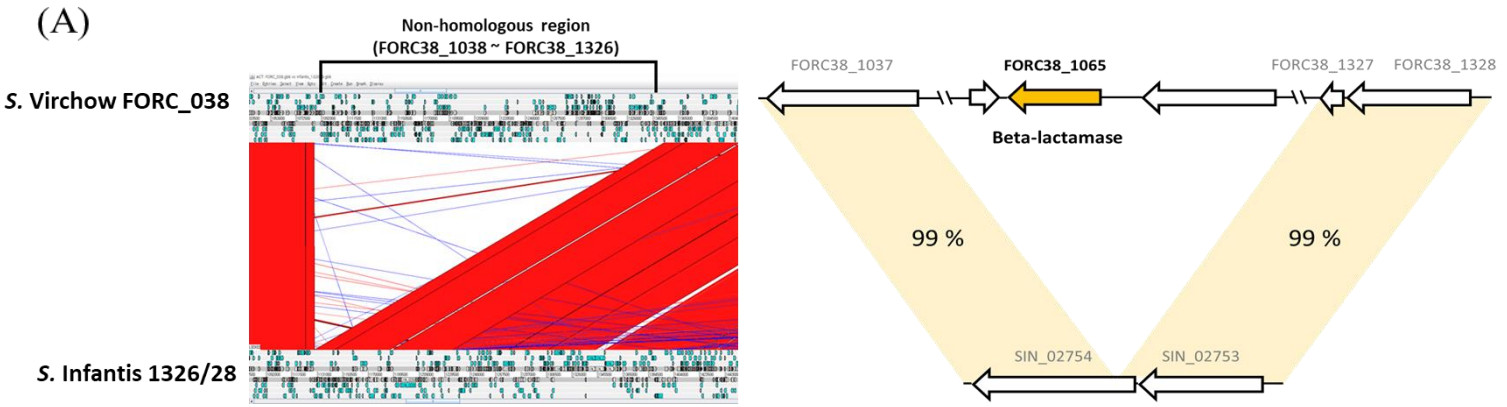

S. enterica Newport FORC_020
S. enterica Newport USDA−1927
S. enterica Typhimurium DT2
S. enterica Typhimurium LT2
S. enterica Typhimurium FORC_030
S. enterica Typhimurium FORC_058
S. enterica Heidelberg B182
S. enterica Heidelberg SL476
S. enterica Thompson ATCC8391
S. enterica Thompson RM6836
S. enterica Anatum ATCC1592
S. enterica Anatum USDA
S. enterica Newport CVM21550
S. enterica Newport SL254
S. enterica GallinarumP CDC
S. enterica GallinarumP RKS5078
S. enterica Enteritidis FORC_075
S. enterica Enteritidis FORC_052
S. enterica Enteritidis FORC_019
S. enterica Enteritidis FORC_050
S. enterica Enteritidis FORC_051
S. enterica Enteritidis P125109
S. enterica Enteritidis FORC_007
S. enterica Enteritidis FORC_074
S. enterica Enteritidis 771427
S. enterica Virchow FORC_038
S. enterica Infantis 132628
S. enterica Cholera C500
S. enterica Cholera SCB67
S. enterica Tennessee TXSC
S. enterica Mbandaka FORC_015
S. enterica Agona 46000421
S. enterica Agona SL483
S. enterica Montevideo 507440
S. enterica Montevideo USDA
S. enterica Typhi CT18
S. enterica Typhi Pstx12
2.0
1.5
1.0
0.5
0.0

Fig 5. ANI analysis of *Salmonella* species. The phylogenetic tree was created with R using ANI values, which were computed according to the chromosomal sequence identities of the selected *Salmonella* species. Sequence identity percentages of *Salmonella* spp. were painted with different colors depending on the high and low figures; ANI values are visualized by a color gradient, where the color intensifies with increasing ANI value.

Comparative genome analysis between FORC_038 and *S.* Infantis 1326/28

Comparative genome analysis between FORC_038 and *S.* Infantis 1326/28 was conducted. It is identified that FORC_038 has additional genes, which may be related to virulence and antibiotic resistances (Fig. 6A, B). FORC38_1065 and FORC38_2974 are the regions encoding beta-lactamase and type III secretion system effector SopE, respectively (Fig. 6B).

(A)


Non-homologous region
(FORC38_1038 ~ FORC38_1326)
S. Virchow FORC_038
S. Infantis 1326/28
FORC38_1037
FORC38_1065
FORC38_1327
FORC38_1328
Beta-lactamase
99 %
99 %
SIN_02754
SIN_02753

(B)

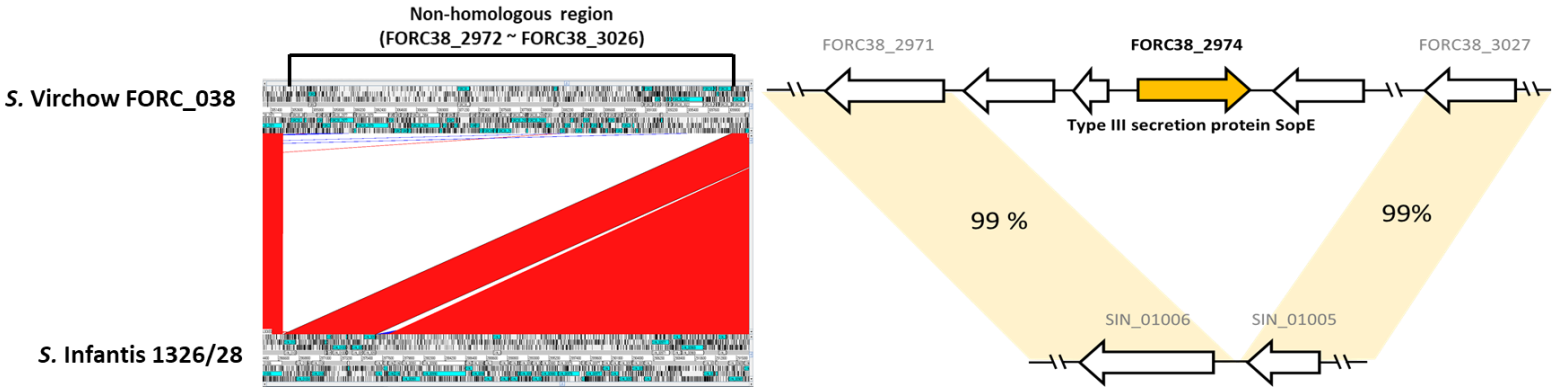

Non-homologous region
(FORC38_2972 ~ FORC38_3026)
S. Virchow FORC_038
S. Infantis 1326/28
FORC38_2971
FORC38_2974
FORC38_3027
Type III secretion protein SopE
99 %
99%
SIN_01006
SIN_01005

Fig 6. Comparative genome analysis between FORC_038 and *S.* Infantis 1326/28.

The regions included (A) the beta-lactamase (FORC38_1065), and (B) the type III secretion system SopE (FORC38_2974) gene were expressed respectively. The red regions are the homologue regions and the white regions are the additional region in FORC_038. The direction of each arrow corresponds to the directionality of each corresponding gene. The orange arrows represent the functionally important genes in FORC_038. The percentage in the figure was indicated the nucleotide sequence similarity of the homologue regions.

Identification of antimicrobial resistances of *S.* Virchow FORC_038.

Additional genomic analysis revealed that genes related to resistance to β-lactams (FORC38_0096, 1541, 3615), tetracyclines (FORC38_1059), aminoglycosides (FORC38_1321) and macrolides (FORC38_3090, 3091) and multiple antibiotic resistance protein (FORC38_1327, 1328, 2505, 2506, 2507, 2508) were widely disseminated on the chromosome of FORC_038, implying that the strain might be resistant to various antibiotics (Table 6). *In silico* prediction was phenotypically confirmed by the subsequent Kirby-Bauer Disk Susceptibility Test. According to Clinical and Laboratory Standards Institute (CLSI) standards, it was indicated that FORC_038 is resistant to 8 antibiotics (ampicillin, cefazolin, cefotaxime, ceftazidime, streptomycin, tetracycline, erythromycin, nalidixic acid) *bona fide*, and susceptible to amoxicillin/clavulanic acid, but the results were not definitive for ciprofloxacin (figure 7) (Table 7). *E. coli* ATCC 25922 served as a quality control strain in the Kirby-Bauer disk susceptibility test.

Table 6. Antibiotic resistance-related genes of *S.* Virchow FORC_038

Locus tag	Annotation	Location	Function
FORC38_0096	Putative beta-lactamase	1107594..1108469	[a]
FORC38_1059	Tetracycline efflux protein TetA	11022129..1103328	Tetracycline resistance
FORC38_1065	Beta-lactamase	1107594..1108469	β-lactam antibiotics resistance
FORC38_1321	Aminoglycoside 3-N-acetyltransferase	1349862..1350722	[b] Aminoglycosides resistance
FORC38_1541	Beta-lactamase class C and other penicillin binding proteins	1617650~1618948	β-lactam antibiotics resistance
FORC38_1327	Multiple antibiotic resistance protein B (ErmB)	1354154~1354369	Multiple antibiotic resistance
FORC38_1328	Multiple antibiotic resistance protein A (ErmA)	1354386~1355558	Multiple antibiotic resistance
FORC38_1713	Polymyxin resistance protein ArnT, undecaprenyl phosphate-alpha-L-Ara4N transferase; Melittin resistance protein PqaB	1791922..1793568	Polymyxin resistance (4-amino-4-deoxy-L-arabinose lipid A transferase)
FORC38_1782,	Bcr/CflA family drug resistance efflux transporter	1869429..1870619	Bicyclomycin resistance protein in *E. coli*

FORC38_2505	Multiple antibiotic resistance protein MarC	2600506..2601171	Multiple antibiotic resistance
FORC38_2506	Multiple antibiotic resistance protein MarR	2601430..2601864	Multiple antibiotic resistance
FORC38_2507	Multiple antibiotic resistance protein MarA	2601884..2602267	Multiple antibiotic resistance
FORC38_2508	Multiple antibiotic resistance protein MarB	2602269..2602511	Multiple antibiotic resistance
FORC38_3090	Macrolide export ATP-binding protein MacB	3176841..3178787	[c] Macrolides antibiotic resistance
FORC38_3091	Macrolide-specific efflux protein MacA	3178784..3179902	Macrolides antibiotic resistance
FORC38_3493	Fosmidomycin resistance protein	3598261..3599481	Fosmidomycin resistance

[a] β-lactam antibiotics; penicillin derivatives (penams), cephalosporins , monobactams, and carbapenems

[b] Aminoglycoside antibiotics; streptomycin, kanamycin, tobramycin, gentamicin, and neomycin

[c] Macrolides antibiotics; erythromycin, azithromycin, fidaxomicin, josamycin

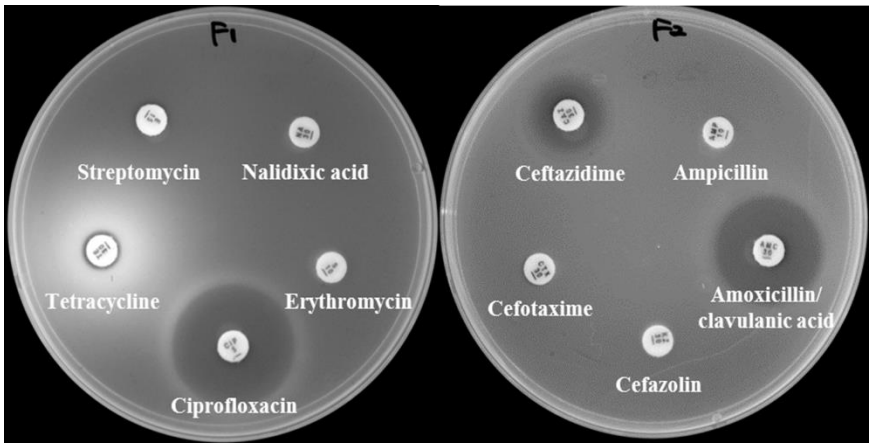


Figure 7. Kirby-Bauer disk susceptibility test to identify antibiotic resistances of *S.* Virchow FORC_038. The susceptibility of *S.* Virchow FORC_038 to various antibiotics was identified through Kirby-Bauer disk susceptibility test and the diameters of clear zone formed by each antibiotic were interpreted as susceptible (S), intermediate (I), or resistant (R) according to the recommendations of the Clinical and Laboratory Standards Institute. The upper panel image indicated the representative image from three individual experiments. The 10 antimicrobial disks (Oxoid, UK), as listed in the table, were placed on the same Muller-Hinton Agar plate on which *S.* Virchow FORC_038 cells were spread.

Table 7. Antibiotic resistances of *S.* Virchow FORC_038 according to CLSI standards

Antimicrobial agent	[a] Disk content (µg)	[b] Interpretive criteria (nearest whole mm)			[c] Zone diameter (nearest whole mm)	
		S	I	R		
Ampicillin	10	≥17	14-16	≤13	-	R
Cephazolin	30	≥15	-	≤14	-	R
Cefotaxime	30	≥26	23-25	≤22	-	R
Ceftazidime	30	≥21	18-20	≤17	17	R
Streptomycin	30	≥15	12-14	≤11	-	R
Tetracycline	15	≥15	12-14	≤11	-	R
Erythromycin	30	≥23	14-22	≤13	-	R
Nalidixic acid	10	≥19	14-18	≤13	-	R
Ciprofloxacin	30	≥31	21-30	≤20	24.5	I
Amoxicillin/ clavulanic acid	5	≥18	14-17	≤13	23.5	S

[a] Antibiotic-impregnated discs (6 mm) with amount, µg shown in bracket

[b] Interpretive criteria according to the recommendations of the Clinical and Laboratory Standards Institute

[c] Diameter of inhibition from three individual experiments. **S, susceptible; I, intermediate; R, resistant**

SopE is essential for the invasion of *S.* Virchow to the host cells

It was identified that a mutant, which lack a gene encoding type III secretion system effector SopE, exhibited significantly lower invasion activity than the wild type, especially in *S.* Dublin and *S.* Typhimurium (Wood *et* al., 1996, Susanne *et* al., 1999). The gene encoding to SopE, which is effector protein of type III secretion system, was also included on the genome of FORC_038. To examine the role of SopE, the *sopE* mutant was constructed by integration of chloramphenicol resistance cassette into the chromosome of FORC_038 using red-mediated homologous recombination. Invasion activity of the *sopE* mutant was measured by invasion assay, which quantifies the proportion of bacteria that invade the eukaryotic cell, HeLa. HeLa cells were infected with the wild type and the *sopE* mutant at MOI of 10 (A) and 30 (B) for 30 minutes. Compared to the wild type, the *sopE* mutant displayed significantly reduced invasiveness toward the HeLa cells (figure 8), indicating that SopE is essential for the invasion of *S.* Virchow to the host cells.

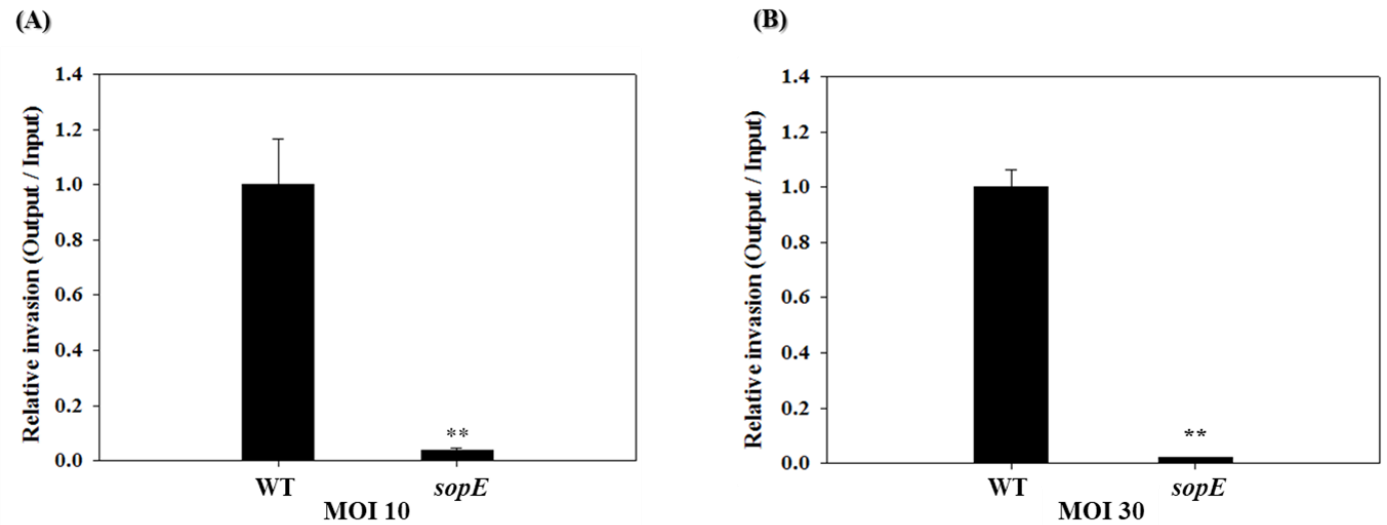


Fig. 8. Effect of *sopE* mutation on the invasion activity of *S*. Virchow FORC_038. Invasion of the *S*. Virchow FORC_038 wild type and its isogenic *sopE* mutant toward HeLa cells were determined by using gentamicin protection assay. The HeLa cells were infected with the *S*. Virchow strains at MOI of 10 (A) and 30 (B) and then were lysed with 1% (v/v) Triton X-100 to recover the intracellular *Salmonella*. Relative invasion of the wild type was presented as 1. The bars represent the means and SEM from 3 independent experiments. **, $p < 0.001$ relative to the wild type. WT, wild type; *sopE*, *sopE* mutant.

Identification of DEGs (Differentially Expressed Genes) of FORC_038 under exposure to raw chicken meat

Transcriptome analysis was performed to compare the transcriptional profile of FORC_038, when exposed to raw chicken meat. RNA-sequencing was conducted after total RNAs was extracted from FORC_038 incubated in M9 minimal media with presence or absence of raw chicken meat. Average RPKM values from the biological triplicate samples were used to represent the expression level of genes. Figure 9 is a volcano plot, which represent the overall transcriptome profile. Genes that are differently expressed significantly with statistical significance (2 fold threshold, P-value < 0.05) were identified. A total of 693 genes were differentially expressed, where 340 genes were up-regulated and 353 genes were down-regulated (figure 10). Genes related to carbohydrates metabolism and amino acids synthesis were downregulated. The genes related to adherence, amino acid uptake, iron uptake and host cell invasion associated type III secretion system were up-regulated (Table 8). Also, heat map was used to visualize the RNA sequencing results (Fig. 11). To identify the result of RNA sequencing, qRT-PCR was conducted (Fig. 12).

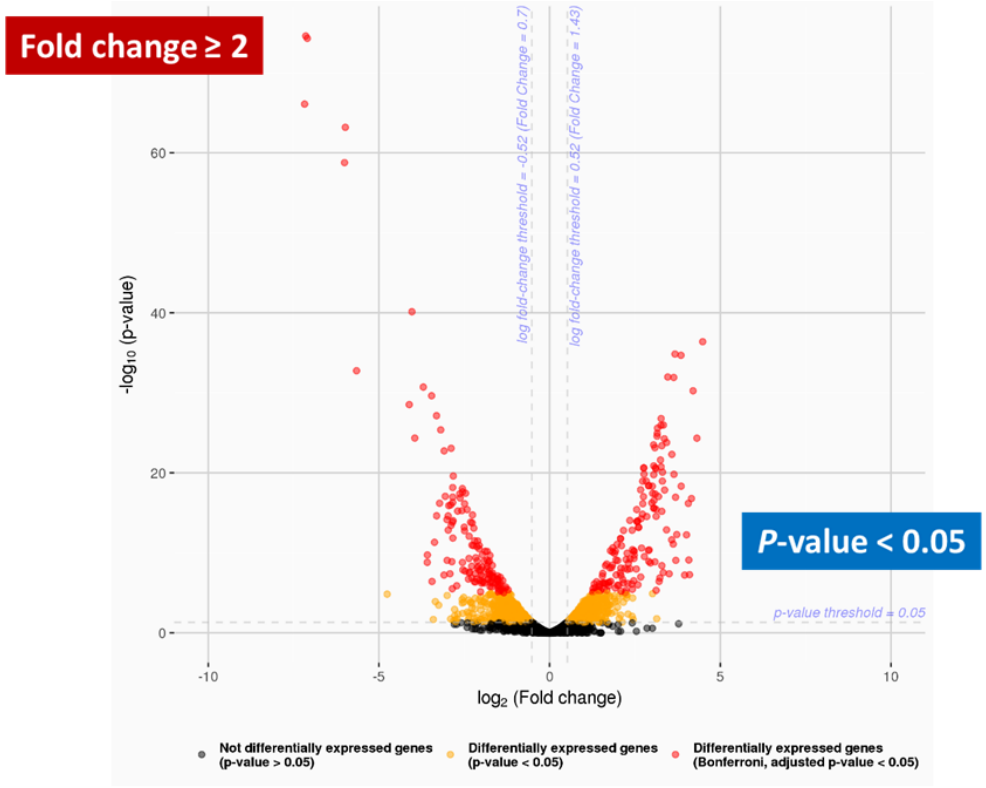


Fig 9. Transcriptome comparison of the RNA sequencing samples. The volcano plot was expressed to visualize differentially total expressed genes when FORC_038 was exposed to raw chicken meat.

Table 8. List of genes differentially expressed by exposure raw chicken meat

Locus tag	Annotation	Fold change [a]	*P*-value [b]
Adherence			
FORC38_0032	Mediator of hyperadherence YidE	1.010657318	0.03453575
FORC38_0232	GGDEF/EAL domain protein YhjH	4.306383883	2.45E-25
FORC38_0811	Fimbriae usher protein StdB	1.236588631	2.43446E-06
FORC38_1182	Pilus assembly protein TraU	2.236518633	3.07E-05
FORC38_1372	Flagellar biosynthesis protein FliC	3.4138548	5.33E-13
FORC38_1373	Flagellar synthesis: repressor of fliC	2.287733943	0.012176663
FORC38_1395	co-chaperone GrpE	1.067649398	1.27E-05
FORC38_2043	Flagellar biosynthesis protein FliC	3.275329941	1.90E-15
FORC38_3374	Chaperone protein TorD	1.919002514	3.07E-06
FORC38_3500	Chaperone protein HtpG	1.090700842	0.000646501
FORC38_4356	Putative type-1 secretion protein	1.117021665	0.00607728
FORC38_4358	HlyD family type I secretion periplasmic adaptor	3.323950508	3.36E-08
FORC38_4359	Agglutination protein	2.58295677	6.62E-06

Host cell invasion associated type III secretion system			
FORC38_0955	Invasion protein invH precursor	1.901818563	0.000157
FORC38_0956	Type III secretion thermoregulatory protein (LcrF,VirF,transcription regulation of virulence plasmid)	4.146946092	3.37E-17
FORC38_0957	EscC/YscC/HrcC family type III secretion system outer membrane ring protein	3.848345062	1.87E-20
FORC38_0958	Type III secretion outermembrane contact sensing protein (YopN,Yop4b,LcrE)	3.719318255	1.03E-12
FORC38_0959	Type III secretion inner membrane channel protein (LcrD,HrcV,EscV,SsaV)	2.913692096	9.18E-11
FORC38_0960	Type III secretion system protein BsaR	3.497365605	5.79E-08
FORC38_0961	Type III secretion cytoplasmic ATP synthase (YscN,SpaL,MxiB,HrcN,EscN)	2.534787457	4.23E-08
FORC38_0962	Surface presentation of antigens protein SpaM	3.218572391	7.44E-07
FORC38_0963	Type III secretion host injection and negative regulator protein (YopD); Surface presentation of antigens protein SpaN (Invasion InvJ)	2.946684509	3.85E-09
FORC38_0964	Type III secretion inner membrane protein (YscQ,flagellar export component-like protein); Surface presentation of antigens protein SpaO	2.162650796	1.20E-06
FORC38_0965	Type III secretion inner membrane protein (YscR,SpaR,HrcR,EscR,flagellar export	3.571208146	2.33E-12

	component-like protein)		
FORC38_0966	type III secretion system protein SpaQ	3.206286333	7.52E-07
FORC38_0967	Type III secretion inner membrane protein (YscT,HrcT,SpaR,EscT,EpaR1,flagellar export component-like protein)	2.418383282	6.10E-06
FORC38_0968	Type III secretion inner membrane protein (YscU,SpaS,EscU,HrcU,SsaU,flagellar export component-like protein)	1.890403699	2.95E-05
FORC38_0969	Type III secretion chaperone protein for YopD (SycD)	1.64305982	0.007002158
FORC38_0970	cell invasion protein SipB	1.7460095	0.000160063
FORC38_0971	Type III secretion negative modulator of injection (YopK,YopQ,controls size of translocator pore)	1.59685975	0.000340171
FORC38_0974	Acyl carrier protein	1.909102629	0.008146659
FORC38_0975	hypothetical protein	1.53354596	0.040868806
FORC38_0976	secretion chaparone	2.127483944	5.76E-05
FORC38_0979	Invasion protein IagB precursor	2.428408869	0.000134168
FORC38_0980	Type III secretion transcriptional activator HilA	2.070763616	4.63E-05
FORC38_0982	MxiG protein; Pathogenicity 1 island effector protein	4.055866525	2.13E-16
FORC38_0985	Type III secretion bridge between inner and	2.466951519	3.44701E-09

	outermembrane lipoprotein (YscJ,HrcJ,EscJ, PscJ)		
FORC38_0986	Oxygen-regulated invasion protein OrgA	1.201439474	0.00020239
FORC38_0989	Type III secretion transcriptional regulator HilC	1.179384934	0.005070568
FORC38_0990	SPI1-associated transcriptional regulator SprB	2.214072118	2.6183E-05
FORC38_2974	Type III secretion protein SopE	2.286109273	2.53E-06
Amino acid metabolism			
FORC38_0043	Acetolactate synthase large subunit	-2.034126123	0.001104146
FORC38_0044	Acetolactate synthase small subunit	-1.576661474	0.00972483
FORC38_0126	L-threonine 3-dehydrogenase	1.059130749	0.000282754
FORC38_0500	Glutamate synthase [NADPH] small chain	-1.540859894	6.40216E-10
FORC38_0501	glutamate synthase	-1.297664672	3.25781E-10
FORC38_0579	Threonine dehydratase, catabolic	1.704799536	0.001065737
FORC38_0787	Glycine dehydrogenase (glycine cleavage system P protein)	-1.431722998	0.000195611
FORC38_0851	N-acetylglutamate synthase	-1.176368185	0.000584419
FORC38_0871	L-serine dehydratase, beta subunit / L-serine dehydratase, alpha subunit	3.630899679	3.82E-54
FORC38_0872	Serine transporter	3.146726317	1.02E-38

FORC38_1331	L-proline glycine betaine binding ABC transporter protein ProX	-1.986476674	2.75257E-07
FORC38_1332	L-proline glycine betaine ABC transport system permease protein ProW	-1.191350761	0.003658977
FORC38_1458	Lysine decarboxylase, inducible	1.030028331	2.75E-06
FORC38_1462	Serine hydroxymethyltransferase	-1.836455332	1.1662E-20
FORC38_1683	Aspartate aminotransferase	1.629911286	4.30E-13
FORC38_1927	Imidazole glycerol phosphate synthase amidotransferase subunit	-3.066260857	9.98806E-18
FORC38_1929	Histidinol-phosphate aminotransferase	-3.320334525	1.53056E-14
FORC38_1930	Histidinol dehydrogenase	-2.853758814	3.83368E-14
FORC38_2044	Lysine-N-methylase	2.157845914	3.47E-17
FORC38_2155	serine/threonine protein phosphatase	-1.700116537	1.95533E-07
FORC38_2277	Tryptophan synthase beta chain	-2.978292765	2.28008E-15
FORC38_2278	Indole-3-glycerol phosphate synthase / Phosphoribosylanthranilate isomerase	-2.890364744	6.6287E-14
FORC38_2279	Anthranilate synthase, amidotransferase component / Anthranilate phosphoribosyltransferase	-4.11958413	2.62815E-34
FORC38_3109	Arginine ABC transporter, periplasmic arginine-binding protein ArtJ	1.157748268	0.000908925
FORC38_3209	Histidine ammonia-lyase	-1.100188963	0.000119278

FORC38_3210	Urocanate hydratase	-1.409488004	0.001186879
FORC38_3311	Asparagine synthetase	1.488015673	7.16E-06
FORC38_4004	Threonine synthase	-1.043890703	0.00031636
FORC38_4005	Homoserine kinase	-1.471153144	1.12742E-07
FORC38_4090	Tryptophanyl-tRNA synthetase	-2.368313839	2.61741E-07
FORC38_4153	Arginine pathway regulatory protein ArgR, repressor of arg regulon	1.731071016	0.000431304
FORC38_4154	Aspartate carbamoyltransferase	1.585540874	1.12E-06
FORC38_4155	Aspartate carbamoyltransferase regulatory chain (PyrI)	1.373729595	0.000390128
FORC38_4285	Aspartate ammonia-lyase	2.887948275	2.25E-10
FORC38_4369	Biosynthetic Aromatic amino acid aminotransferase alpha	-1.527540575	5.80327E-10
FORC38_4429	5-methyltetrahydrofolate--homocysteine methyltransferase	-1.528687216	1.9407E-09
FORC38_4435	Homoserine O-succinyltransferase	-1.456369276	0.001268787
FORC38_4478	Argininosuccinate lyase	-1.575787033	0.001564979
FORC38_4479	Acetylglutamate kinase	-1.199543832	0.014282405
FORC38_4495	Aspartokinase / Homoserine dehydrogenase	-1.476295671	0.003640636

FORC38_4699	Aspartate--ammonia ligase	1.740064894	5.74E-07
Carbohydrate metabolism			
FORC38_1545	Transketolase	-1.983378612	3.48111E-07
FORC38_1864	Fructose-bisphosphate aldolase class I	-2.507360058	4.56E-14
FORC38_1922	UDP-glucose 6-dehydrogenase	-1.525633193	1.54214E-06
FORC38_2209	Trehalase; Periplasmic trehalase precursor	-1.039101662	0.001589785
FORC38_2461	Malto-oligosyltrehalose trehalohydrolase	-1.902486563	0.000370891
FORC38_2462	Malto-oligosyltrehalose synthase	-1.49520897	3.97765E-05
FORC38_2463	Glycogen debranching enzyme	-1.087181973	7.81213E-06
FORC38_2504	Sugar efflux transporter SotB	-1.159824698	2.76597E-05
FORC38_3097	Pyruvate oxidase	-1.384725538	0.001160318
FORC38_3833	glucose dehydrogenase	-2.063243526	1.72056E-06
FORC38_3850	Dihydrolipoamide dehydrogenase of pyruvate dehydrogenase complex	-1.027941358	0.004416877
FORC38_3852	Pyruvate dehydrogenase E1 component	-1.20958984	1.45449E-05
FORC38_4482	Phosphoenolpyruvate carboxylase	-1.139379123	6.29625E-05
FORC38_4533	Fructokinase	-1.064279984	0.014793573

Iron uptake			
FORC38_0328	Ferrous iron-sensisng transcriptional regulator FeoC	1.987386569	0.001748123
FORC38_1762	heme ABC transporter permease	2.556350833	4.24136E-17
FORC38_2823	Putative OMR family iron-siderophore receptor precursor	2.000042629	0.000382507
FORC38_3171	Non-specific DNA-binding protein Dps / Iron-binding ferritin-like antioxidant protein / Ferroxidase	4.218042541	3.42E-08
FORC38_3298	Ferric uptake regulation protein FUR	1.107693971	0.000543485
FORC38_3391	Enterobactin exporter EntS	1.173658398	0.006849658
FORC38_3398	Enterobactin esterase	1.978069135	0.000260706
FORC38_3800	Fe3+-hydroxamate ABC transporter permease FhuB	1.179255152	0.025823369
FORC38_3801	Ferric hydroxamate ABC transporter, periplasmic substrate binding protein FhuD	1.090463469	0.005483567
FORC38_3803	Ferric hydroxamate outer membrane receptor FhuA	1.21913575	0.037609495

[a] The $\log_2$ mRNA expression level of *S.* Virchow FORC_038 after two hours exposure to raw chicken meat with or without exposure to raw chicken meat.

[b] The *P*-value less than six decimal places were denoted as zero.

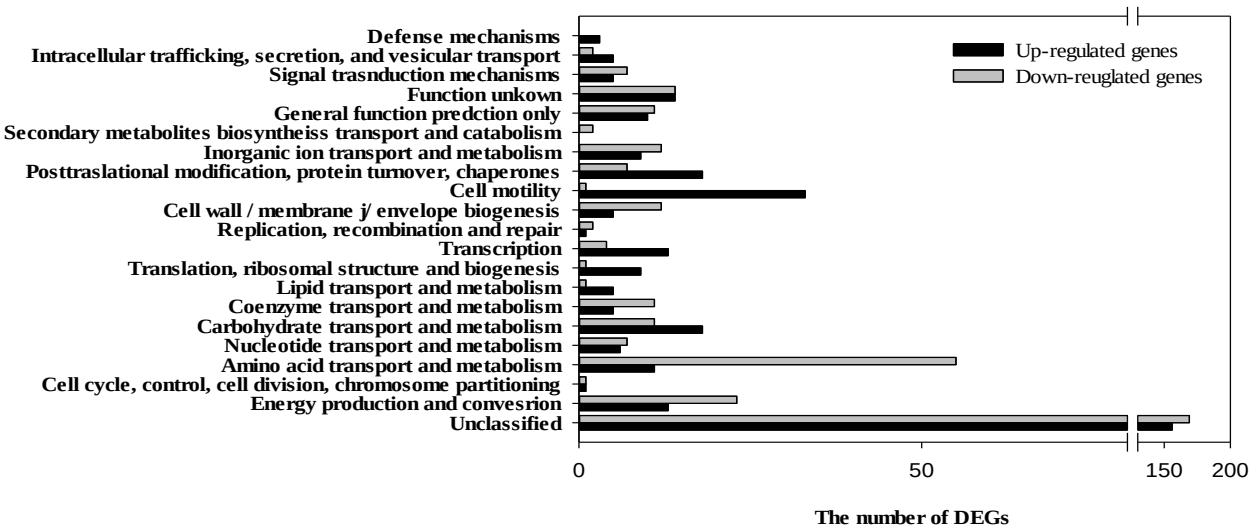


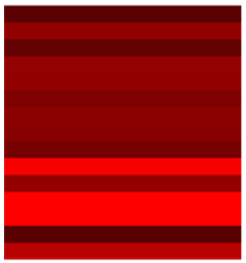
Fig 10. Functional categorization of genes differentially expressed upon exposure to raw chicken meat. Gene expressions based on the RNA sequencing results were considered when exposed to raw chicken meat according to the functional categories. (*P*-value < 0.05, fold change ≥ 2 or fold change ≤ -2)

❖ Adherence

FORC38_0032 – Hyper adherence mediator
FORC38_0811 – Fimbriae usher protein
FORC38_1182 – Pilus assembly protein
FORC38_0232, FORC38_1372, FORC38_1373, FORC38_2043 – Biofilm formation associated protein
FORC38_1395, FORC38_3374, FORC38_3500 – Chaperon-usher assembled fimbrial protein
FORC38_4356, FORC38_4358 – Type I secretion protein
FORC38_4359 – Agglutination protein

❖ Amino acid

FORC38_0043, FORC38_0044 – Acetolactate synthase
FORC38_0500, FORC38_0501, FORC38_0851, FORC38_2278, FORC38_2279, FORC38_4429, FORC38_4479 – Glutamate synthase
FORC38_0787 – Glycine synthase
FORC38_1331, FORC38_1332 – Proline synthase
FORC38_1462, FORC38_4005, FORC38_4435, FORC38_4495 – Serine synthase
FORC38_1927, FORC38_1929, FORC38_1930, FORC38_3209, FORC38_3210 – Histidine synthase
FORC38_2277, FORC38_4090 – Tryptophan synthase
FORC38_2155, FORC38_4044 – Threonine synthase
FORC38_4369 – Aromatic amino acid synthase
FORC38_4478 – Arginine synthase



FORC38_0126, FORC38_0579 – Threonine dehydrogenase
FORC38_3109, FORC38_4153 – Arginine ABC transporter
FORC38_0733, FORC38_3311 – Asparagine metabolism
FORC38_1683, FORC38_4154, FORC38_4155, FORC38_4285, FORC38_4699 – Aspartate metabolism
FORC38_0871, FORC38_0872 – Serine transporter
FORC38_1458, FORC38_2044 – Lysine decarboxylase

❖ Iron uptake

FORC38_0328 – Iron sensing regulator
FORC38_1762, FORC38_3391, FORC38_3398 – Iron-siderophore receptor
FORC38_2823, FORC38_3171 – Iron uptake regulator
FORC38_3298, FORC38_3800, FORC38_3801, FORC38_3803 – Iron transporter

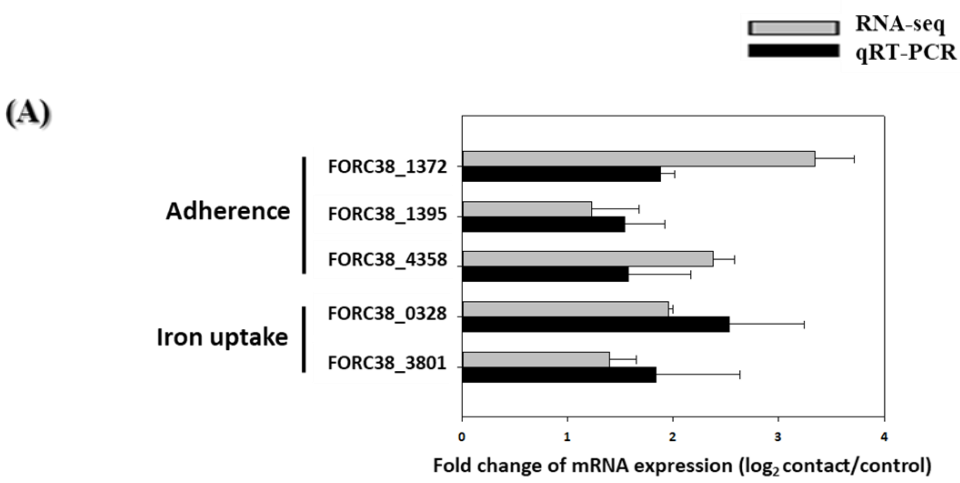
❖ Carbohydrate

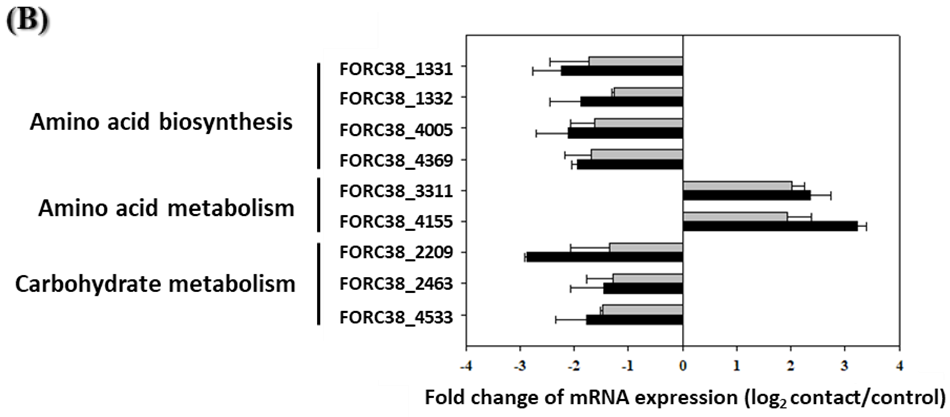
FORC38_1864, FORC38_4533 – Fructose metabolism
FORC38_1545, FORC38_2209 – Glucose metabolism
FORC38_1922, FORC38_2463, FORC38_3833 – Glucose dehydrogenase
FORC38_2461, FORC38_2462 – Trehalose metabolism
FORC38_3097 – Pyruvate oxidase
FORC38_3850, FORC38_3852 – Pyruvate dehydrogenase
FORC38_4482 – Phosphoenolpyruvate carboxylase
FORC38_2504 – Sugar efflux transporter

❖ Type III secretion system (T3SS)

FORC38_0955, FORC38_0956, FORC38_0957, FORC38_0958, FORC38_0959, FORC38_0960, FORC38_0961, FORC38_0962, FORC38_0963, FORC38_0964, FORC38_0965, FORC38_0966, FORC38_0967, FORC38_0968, FORC38_0969, FORC38_0970, FORC38_0971, FORC38_0974, FORC38_0975, FORC38_0976, FORC38_0979, FORC38_0980, FORC38_0982, FORC38_0985, FORC38_0986, FORC38_0989, FORC38_0990, FORC38_2974 – Host cell invasion associated type III secretion system

Fig 11. Heat map of selected genes' transcriptome after 2 h exposure to raw chicken meat. The red bars represented the up-regulated genes and the green bars indicated that the down-regulated genes. The scale bar is above the heat map. Heat map was constructed using the Gitools.


RNA-seq
qRT-PCR
(A)
Adherence
FORC38_1372
FORC38_1395
FORC38_4358
Iron uptake
FORC38_0328
FORC38_3801
0
1
2
3
4
Fold change of mRNA expression (log2 contact/control)


(B)
Amino acid biosynthesis
FORC38_1331
FORC38_1332
FORC38_4005
FORC38_4369
Amino acid metabolism
FORC38_3311
FORC38_4155
Carbohydrate metabolism
FORC38_2209
FORC38_2463
FORC38_4533
-4
-3
-2
-1
0
1
2
3
4
Fold change of mRNA expression (log2 contact/control)

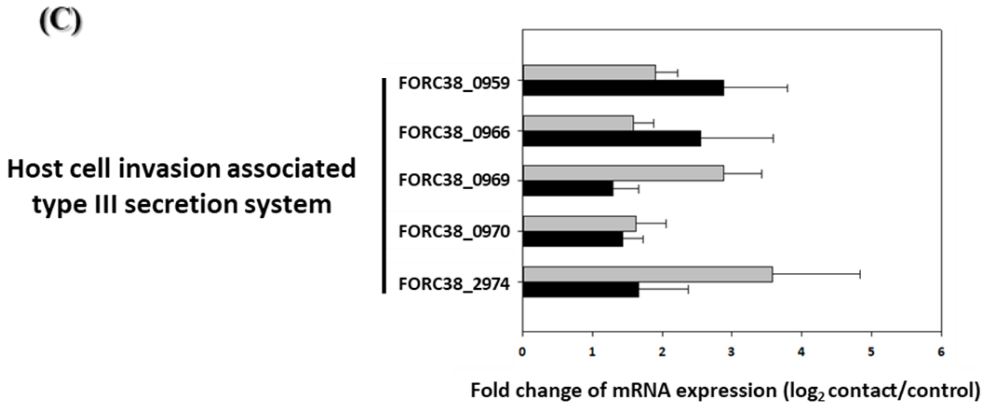

(C)
Host cell invasion associated type III secretion system
FORC38_0959
FORC38_0966
FORC38_0969
FORC38_0970
FORC38_2974
0
1
2
3
4
5
6
Fold change of mRNA expression (log2 contact/control)

Fig 12. Expression comparison between RNA sequencing and qRT-PCR. The results of RNA sequencing were confirmed using qRT-PCR targeting the represented genes. (A) adherence and iron-uptake, (B) amino acid and carbohydrate metabolism, (C) host cell invasion associated type III secretion system were identified, respectively. The products information of the represented genes was listed in Table 8. The error bars represent the standard deviations.

Growth kinetics of FORC_038 exposed to raw chicken meat

Because the protein is one of the major components of domestic raw chicken meat (Koh *et al*., 2015) and iron is an essential component for the survival of bacteria, FORC_038 may utilize iron and protein in the form of amino acid for growth and survival when exposed to raw chicken meat. The kinetics of growth were compared between three samples which were incubated in M9 minimal media with presence or absence of the model food (raw chicken meat), respectively. FORC_038 under exposure to the raw chicken meat grew faster and the colony forming unit (CFU) was about 4 fold higher (Fig. 12).

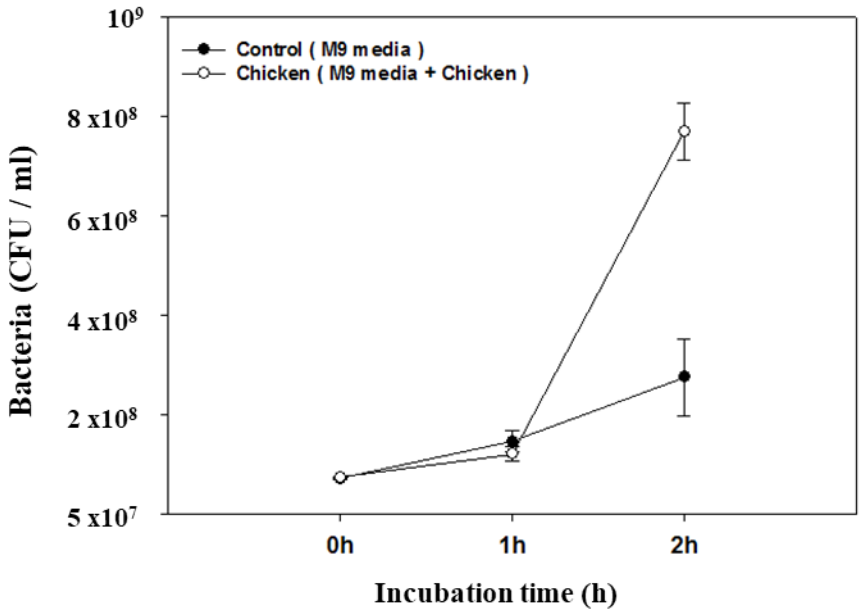


Fig 13. Growth kinetics of FORC_038 incubated in M9 minimal media either in the presence or absence of raw chicken meat. At the specific time points (0 h, 1 h, 2 h), the CFUs were counted. The experiments were conducted with biologically and technically triplicated for each sample. Open circle (○), the raw chicken meat-exposed sample; closed circle (●), control sample. The error bars represented the standard deviations.

IV. DISCUSSION

Salmonella enterica is an ever-present worldwide public health threat, related to antibiotic resistance and human diseases such as gastroenteritis and invasive systemic infections. Especially *S*. Virchow is one of the increasing global foodborne pathogens isolated from poultry meat products, which has contained slightly less susceptibility to antibiotics. Identified as an important cause of poultry meat outbreaks, *S*. Virchow infections in humans has also increased over the years including South Korea. Preventing future outbreaks can be aided by studying the genome and by analyzing the change in transcriptome of the pathogen, upon contact with food. In this study, FORC_038, a *S.* Virchow isolated from a raw chicken meat, was analyzed.

In order to characterize the potential virulence of FORC_038, virulence gene-specific PCR screening and invasion assay were performed. FORC_038 had various virulence genes and showed a high level of invasiveness toward HeLa cell lines (Figures 1 and 2). The result suggested that this strain may be pathogenic toward human cells. FORC_038's genomic DNA was extracted and sequenced, and its complete genome sequence was analyzed. The ANI analysis suggested that FORC_038 had the highest ANI value (98.87) with *S.* Infantis 1326/28, isolated from chicken in UK, 1994. In order to take a closer look at the FORC_038 genome, BLAST searches were performed. The goal was to predict whether the genome had the virulence factors that are present in genomes of virulent *Salmonella* spp. (Table 5). The major virulence factors of *S.* Virchow such as a invasion protein (InvA) (Galan

et al., 1992) and transcriptional regulator (HilD) (Petrone *et al*., 2014) were located in the genome of this strain. Besides, FORC_038's genome harbored various antibiotic resistance related genes encoding such as tetracycline efflux protein TetA and aminoglycoside 3-N-acetyltransferase. In short, although FORC_038 is not a clinical isolate, its genome harbors the various virulence factors and antibiotic resistances. Also, the invasion activity of FORC_038 was similar to that of a virulent strain, *S.* Typhimurium SL1344. This is an example of an environmental isolate which shows a potential to be as virulent as a clinical isolate or virulent strains. Furthermore, comparative genome analysis was carried out between FORC_038 and *S.* Infantis 1326/28, it's the most closely related strain according to ANI analysis. FORC_038 has non-homologous regions, encoding β-lactamase (FORC38_1065), which has been reported to play a major role of β-lactam antibiotics resistance, and type III secretion system effector SopE (FORC38_2974), which stimulates cytoskeletal reorganization and JNK activation in a CDC42- and Rac-1-dependent manner. (Galan *et al*., 1998). This suggests that FORC_038 may have an advantage from antibiotics to survive, and relative high invasion activity to host cells.

To examine the behavior of *S.* Virchow FORC_038 when this strain is exposed to raw chicken meat, RNA sequencing was performed using three RNA samples which were extracted from the bacterium incubated in M9 minimal media either with or without raw chicken meat. Genes related to carbohydrate metabolism and amino acid biosynthesis were down-regulated but the genes related to host cell invasion associated type III secretion system, adherence, amino acid uptake, and iron uptake

were up-regulated. *Salmonella* species adhere to a number of different mammalian cell lines. During adherence, bacteria trigger a complex set of signaling events in the host cell which cause cytoskeletal rearrangements and concomitant changes in the morphology of the host plasma membrane, ultimately leading to the entry of *Salmonella* into the eukaryotic cell (Suarez *et al*., 1998). So, FORC_038 may have high invasion activity after contact from raw chicken meat according to up-regulation of adherence and host cell invasion associated type III secretion system related genes. This type III secretion system (T3SS) was required to inject virulence proteins into their hosts and initiate infections (Wang *et al*., 2010). As protein is a major component of the raw chicken meat (Koh *et al*., 2015), this expression change indicates that *S.* Virchow FORC_038 may be utilizing amino acid from protein for survival and growth when in contact with a raw chicken meat. Lastly, various iron uptake related genes were upregulated. Iron acquisition is required for the growth of the pathogen in this fruit, as it is known to be in animal tissues (Staci *et al*., 2015). In addition, the loss of iron by the pathogen (in the case, *S.* Virchow) gives negative influence to the host cells. As *S.* Virchow FORC_038 can uptake free iron from the host cell using these proteins, it can survive and be pathogenic in host environment.

These results described above suggested the behavior of the FORC_038 when it was exposed to the raw chicken meat; (1) increase of adherence, (2) up-regulation of host cell invasion associated type III secretion system, (3) utilization of amino acid from protein of raw chicken meat, and (4) increased iron uptake. This suggests that the raw chicken meat may play a role of reservoir, as increase invasiveness activity, nutrients

for growth and survival as long as *S.* Virchow will be exposed to raw chicken meat. Thus studying the genetic program set in motion upon contact with raw chicken meat suggests that better practices in processing poultry meat should be recommended, given that the bacterium seems to exhibit increase of the invasiveness activity, motility, survival and growth, which may contribute to virulence and pathogenic characteristics of *Salmonella*. This research supports the accumulation of database about the domestic pathogens and helps to deal with the coming *Salmonella* outbreak from poultry meat in South Korea.

V. REFERENCES

Altermann, E., & Klaenhammer, T. R. (2003). GAMOLA: a new local solution for sequence annotation and analyzing draft and finished prokaryotic genomes. *Omics A Journal of Integrative Biology*, 7(2), 161-169.

Aziz, R. K., Bartels, D., Best, A. A., DeJongh, M., Disz, T., Edwards, R. A., ... & Meyer, F. (2008). The RAST Server: rapid annotations using subsystems technology. *BMC genomics*, 9(1), 1.

Besemer, J., Lomsadze, A., & Borodovsky, M. (2001). GeneMarkS: A self-training method for prediction of gene starts in microbial genomes. Implications for finding sequence motifs in regulatory regions. *Nucleic acids research*, 29(12), 2607-2618.

Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., & Parkhill, J. (2005). ACT: the Artemis comparison tool. *Bioinformatics*, 21(16), 3422-3423.

P Vaudaux., F A Waldvogel. (1979). Gentamicin antibacterial activity in the presence of human polymorphonuclear leukocytes. *Antimicrobial agents and chemotherapy*, 16(6), 743-749.

Michael McClelland, Kenneth E. Sanderson, John Spieth, Sandra W. Clifton, Phil Latreille, Laura Courtney, Steffen Porwollik, Johar Ali, Mike Dante, Feiyu Du, Shunfang Hou, Dan Layman, Shawn Leonard, Christine Nguyen, Kelsi Scott, Andrea Holmes, Neenu Grewal, Elizabeth Mulvaney, Ellen Ryan, Hui Sun, Liliana Florea, Webb Miller, Tamberlyn Stoneking, Michael Nhan, Robert Waterston & Richard K. Wilson (20001). Complete genome sequence of *Salmonella enterica* serovar Typhimurium LT2. *Nature*, 413, 852-856.

Costa, T. R., Felisberto-Rodrigues, C., Meir, A., Prevost, M. S., Redzej, A., Trokter, M., & Waksman, G. (2015). Secretion systems in Gram-negative bacteria: structural and mechanistic insights. *Nature Reviews Microbiology*, 13(6), 343-359.

Bergholz, T. M., Switt, A. I. M., & Wiedmann, M. (2014). Omics approaches in food safety: fulfilling the promise? *Trends in Microbiology*, 22(5), 275-281.

Carsten Kröger, Shane C. Dillon, Andrew D. S. Cameron, Kai Papenfort, Sathesh K. Sivasankaran, Karsten Hokamp, Yanjie Chao, Alexandra Sittka, Magali Hébrard, Kristian Händler, Aoife Colgan, Pimlapas Leekitcharoenphon, Gemma C. Langridge, Amanda J. Lohan, Brendan Loftus, Sacha Lucchini, David W. Ussery, Charles J. Dorman, Nicholas R. Thomson, Jörg Vogel, and Jay C. D. Hinton. (2012). The transcriptional landscape and small RNAs of *Salmonella enterica* serovar Typhimurium. *Proceedings of National Academy of Science*, 109(20),

E1277-E1286

Howard Ochman, Eduardo A. Groisman (1996). Distribution of Pathogenicity Islands in *Salmonella* spp. *Infection and immunity,* 64(12), 5410-5412.

Wolf-Dietrich Hardt, Li-Mei Chen, Kornel E. Schuebel, Xose R Bustelo, Jorge E. Galan (1998). S. Typhimurium Encodes an Activator of Rho GTPases that Induces Membrane Ruffling and Nuclear Responses in Host Cells. *Cell,* 93, 815-826.

H. Rahman, W. Streckel, R. Prager, H. Tschape. (2004). Presence of *sopE* gene & its phenotypic expression among different serovars of *Salmonella* isolated from man & animals. *Indian Journal of Medical Research*, 120(1), 35-38.

Jones, P., Binns, D., Chang, H. Y., Fraser, M., Li, W., McAnulla, C, Pesseat, S. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics*, 30(9), 1236-1240.

Anne-Marie S, Gail L, Georgios N, Anja K, Paul W. (2014). *Salmonella* Virchow infection of the chicken elicits cellular and humoral systemic and mucosal responses, but limited protection to homologous or heterologous re-challenge. *Frontiers in Veterinary Science,* 1

JR McQuiston, S Herrera-Leon, BC Werthem, J Doye, PI Fields, RV Tauxe, JM Logsdon Jr. (2008). Molecular Phylogeny of the Salmonellae: Relationships among *Salmonella* Species and Subspecies Determined from Four Housekeeping Genes and Evidence of Lateral Gene Transfer Events. *Journal of bacteriology*, 190(21), 7060-7067.

Ku, H., & Lee, J. (2014). Development of a Novel Long-Range 16S rRNA Universal Primer Set for Metagenomic Analysis of Gastrointestinal Microbiota in Newborn Infants. *Journal of Microbiology and Biotechnology,* 24, 812-822.

Kumar, S., Stecher, G., & Tamura, K. (2016). MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets. *Molecular biology and evolution*, msw054.

David HSU, Liu-Mei S, Yuan CZ. (1994). Degradation of rRNA in *Salmonella* Strains: a Novel Mechanism To Regulate the Concentrations of rRNA and Ribosomes. *Journal of bacteriology*, 176(15), 4761-4765.

Dieye Y, Ameiss K, Mellata M, Curtiss R. (2009). The *Salmonella* Pathogenicity Island (SPI) 1 contributes more than SPI2 to the colonization of the chicken by *Salmonella enterica* serovar Typhimurium. BMC Microbiology, 9(1), 3

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., & Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature methods*, 5(7), 621-628.

Perez-Llamas, C., & Lopez-Bigas, N. (2011). Gitools: analysis and visualisation of genomic data using interactive heat-maps. *PloS one*, 6(5), e19541.

Richter, M., & Rosselló-Móra, R. (2009). Shifting the genomic gold standard for the prokaryotic species definition. *Proceedings of the National Academy of Sciences*, 106(45), 1912-1913.

Staib, L., & Fuchs, T. M. (2015). Regulation of fucose and 1, 2-propanediol utilization by *Salmonella enterica* serovar Typhimurium. *Frontiers in microbiology*, 6.

Csordas AT, Barak, JD, Delwiche MJ. (2004). Comparison of primers for the detection of *Salmonella enterica* serovars using real-time PCR. *Letters in applied microbiology*, 39(2), 187-193.

Kotrba P, Masayuki I, Hideaki Y. (2001). Bacterial phosphotransferase system (PTS) in carbohydrate uptake and control of carbon metabolism. *Journal of Bioscience and Bioengineering*, 92(6), 502-517.

VI. 국문초록

*Salmonella enterica*는 전 세계적으로 공중 보건상의 위협이 되는 균주이며 위장질환, 설사 혹은 침습성 전신감염과 같은 질병을 일으킬 수 있는 식중독균이다. 특히, *Salmonella enterica* Virchow는 다양한 항생제 내성과 높은 독성을 나타내며, 가금류로부터 유래한 살모넬라 식중독 발생의 주 원인균 중 하나이다. 그러나 높은 병원성에도 불구하고 *S.* Virchow의 독성 및 항생제 저항성 인자들에 관한 연구는 전무한 실정이다. 본 연구에서는 국내 생닭에서 분리한 *S.* Virchow 'FORC_038' 균주에 대한 유전체 분석과 전사체 분석을 통해 *S.* Virchow에 대한 병원성 및 모델식품에서의 생리적 특성을 유전자 수준에서 파악하였다. 전장 유전체 분석 결과 FORC_038은 하나의 크로모좀을 가지고 있고, 4,735개의 개방형 해독틀, 85개의 운반 RNA와 22개의 리보솜 RNA를 지니고 있었다. 또한 병원성에 기여하는 다양한 독성인자 및 항생제 저항성 유전자들을 가지고 있었으며 세포침투 실험을 통해 FORC_038의 높은 독성을 확인하였다. 또한, 유전체가 완전히 밝혀져 NCBI에 등재된 총 40개의 타 살모넬라균과의 종 간 유사도를 확인한 결과, FORC_038은 영유아 살모넬라 감염의 주 원인균으로써 실제 식중독을 일으켰던 *Salmonella enterica* Infantis 1326/28과 계통학적으로 가장 유사함을 알 수

있었다. *S.* Infantis 1326/28과의 비교유전체 분석을 통해 FORC_038이 다양한 항생제 저항성 관련 유전자 뿐만 아니라 beta-lactamase (CTX-M-15 extended spectrum beta-lactamase) (FORC38_1065) 와 type III secretion system effector SopE (FORC38_2974) 를 암호화하는 유전자를 추가적으로 가지고 있음이 확인되었다. *sopE* 돌연변이 균주는 세포주를 이용한 침투능력 실험에서 정상 균주에 비해 침투능력이 현저하게 낮아지는 것이 확인되었으며, 이를 통해 SopE가 *S.* Virchow의 독성에 있어 중요한 역할을 할 것으로 예측된다. 또한 항생제 저항성 실험을 진행한 결과 beta-lactam 항생제뿐만 아니라 tetracyclines, aminoglycoside, macrolides, quinolones 계열의 선정된 10종의 항생제 중 8종의 항생제에 저항성을 나타내는 것이 확인되었으며, 추가적으로 더 많은 항생제에 대한 내성을 가지고 있을 것으로 예상된다.

생닭은 한국을 포함하여 전세계적으로도 소비량이 많은 식품이다. 살모넬라 감염의 주요 원인일 뿐만 아니라 FORC_038의 분리원이기 때문에 생닭을 전사체 분석을 위한 모델 식품으로 선정하였다. FORC_038을 생닭에 노출시켜 유전자 발현양상의 변화를 RNA sequencing을 통해 분석한 결과 부착능력, 침투독성 관련 type III secretion system, 아미노산 흡수, 철 흡수 관련 유전자는 발현이 증가했지만 탄수화물 대사, 아미노산 생합성 관련된 유전자들의 발현은 감소하였다.

생닭으로부터 분리된 FORC_038은 여러 독성인자와 다양한 항생제 저항성이 존재하여 병원성을 나타낼 가능성이 높고, 생닭에 노출되었을 시 부착하여 철과 아미노산을 이용하여 성장과 생존을 이어 나아갈 것으로 예상된다. 또한 type III secretion system에 의한 침투독성의 증가로 인체에 감염 시, 질병 발생의 가능성을 높일 것으로 예측된다. 본 연구를 통해 국내에서 처음으로 *S.* Virchow에 대한 데이터베이스를 축적함으로써, 이후 가금육 유래 살모넬라균에 의한 질병 발생을 예방할 수 있을 뿐만 아니라 이를 제어하기 위한 추가적인 연구에도 도움이 될 것이라 기대한다.

주요어: *Salmonella enterica* Virchow, 생닭, 유전체, 전사체, 전장유전체 시퀀싱, 리보 핵산 시퀀싱, 항생제 저항성

학 번: 2016-21729