

Supplementary text for: Defining Core Genome of *Salmonella enterica* serovar Typhimurium for genomic surveillance and epidemiological typing, Fu et al.

Gene content analysis

Presence of prophages from the sequenced strains was screened using PHAST (1). Antimicrobial resistance genes were identified using ResFinder (2). The number of coding sequences in the genomes was predicted based on RAST. One hundred and fifty known virulence determinants including *Salmonella* pathogenicity islands (SPIs), virulence gene encoding prophages/plasmids and fimbrial clusters were screened against each strain using BLASTn to determine their presence or absence(3). Unique genes were identified using progressiveMauve as they were not able to be aligned to LT2. To determine the function of the unique genes, contigs were searched against the GenBank non-redundant nucleotide database by using BLASTn (4). To determine the presence of pSLT from the 21 strains sequenced, reads and contigs were mapped onto the pSLT sequence in LT2 (NC_003277). Multiple plasmid comparisons were visualized using BRIG (version 0.95) (5).

Genome content variation among the 21 SARA strains

The 21 SARA strains exhibited differences in prophages, plasmids, and antibiotic resistance genes profiles. In total, 14 prophages or prophage remnants, including Fels-1, Fels-2, Gifsy-1, Gifsy-2, Gifsy-3, ST104, P22, ST64B, ST64T, PVP-SE1 Felix O1 (previously 2/2), RE-2010 P2, a SE-OLF-10058(3)-like phage, *sopE2* and *sspH2*, were identified among the 21 *S. Typhimurium* genomes (Figure S4, Table S7).

The Gifsy-3 phage remnant in SARA6 contained only first half of the 50,884 bp Gifsy-3 sequence. SARA19 contained a prophage with 85% DNA similarity with SE-OLF-10058 (3), which was first found in *S. Enteritidis* very recently (6). The SARA19 prophage consisted of 44 genes with five unique genes when compared with SE-OLF-10058 (3) (Table S8).

Fifteen *S. Typhimurium* strains carried the 90-kb virulence plasmid, pSLT (Table S7). Interestingly, SARA3, SARA17 and SARA18 which were all from the same lineage did not possess this plasmid (Figure S4). Moreover, SARA7 and SARA8 contained a 93 Kb F-plasmid with 90% DNA sequence similarity to *Escherichia coli* F-plasmid pEC_L8. This plasmid is a fusion of two replicons of type FII and FIA and may share the same ancestor with another *E. coli* F-plasmid R1 (7). In addition, SARA9 and SARA10 contained a 95 Kb plasmid with 90% and 91% DNA sequence similarities, to the *S. Heidelberg* plasmid pCFSAN002069_01 and *S. Heidelberg* plasmid pSH1148_107, respectively (Figure S5). We therefore named them plasmid pSARA9-2 (94 Kb) and pSARA10-2 (95 Kb). SARA17 had a plasmid (70 Kb) with 80% DNA similarity with pSTM2, which was first reported in an DT135 outbreak (8). Some small plasmids (< 10Kb) were also found in the genomes (Table S7). To investigate virulence profiles of the 21 SARA strains, 150 known virulence genes from *Salmonella* were screened by BLASTn searches (3). The difference in virulence profiles was mainly determined by virulence genes carried by prophages (Table S9,

Figure S6). The presence or absence of *rck* and *spvC*, located on the pSLT, was consistent with the presence or absence of the plasmid among the strains.

Antibiotic resistance genes were found in SARA7, SARA9 and SARA10 (Table S10). *aadA1* (encoding aryloxyalkanoate dioxygenase for streptomycin resistance) and *sulI* (sulphonamide resistance) were found in pSARA9-2 and pSARA10-2. Additionally SARA9 harboured *dfrA1* (associated with resistance to trimethoprim) on its plasmid pSARA9-2 and *tetA* (resistance to tetracyclines) on its chromosome. *tetA* was also found in SARA7 on its chromosome. The acquisition of these antibiotic resistance genes may be due to the extensive use of antibiotics in livestock since 1950s. Earlier studies in the 1980's found that the use of antibiotics in livestock increased by 5-fold in the U.S. between the 1950's and the 1980's (9) and subsequent studies found that antibiotic resistance of *Salmonella* infections in humans was likely to have been acquired from food animals (10).

References

1. **Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS.** 2011. PHAST: a fast phage search tool. *Nucl Acids Res* **39**:W347-352. doi: 10.1093/nar/gkr485.
2. **Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV.** 2012. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemoth* **67**:2640-2644.
3. **Suez J, Porwollik S, Dagan A, Marzel A, Schorr YI, Desai PT, Agmon V,**

- McClelland M, Rahav G, Gal-Mor O.** 2013. Virulence gene profiling and pathogenicity characterization of non-typhoidal *Salmonella* accounted for invasive disease in humans. PloS One. **8**:e58449. doi: 10.1371/journal.pone.0058449.
4. **Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ.** 1990. Basic local alignment search tool. J Mol Bio **215**:403-410.
5. **Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA.** 2011. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. BMC Genomics **12**:402.
6. **Ogunremi D, Devenish J, Amoako K, Kelly H, Dupras AA, Belanger S, Wang LR.** 2014. High resolution assembly and characterization of genomes of Canadian isolates of *Salmonella* Enteritidis. BMC Genomics. **15**:713.
7. **Boyd EF, Hartl DL.** 1997. Recent horizontal transmission of plasmids between natural populations of *Escherichia coli* and *Salmonella enterica*. J Bacteriol. **179**:1622-7.
8. **Hawkey J, Edwards DJ, Dimovski K, Hiley L, Billman-Jacobe H, Hogg G, Holt KE.** 2013. Evidence of microevolution of *Salmonella* Typhimurium during a series of egg-associated outbreaks linked to a single chicken farm. BMC Genomics **14**:800.
9. **O'Brien, T. Hopkins JD, Gillece ES, Medeiros AA, Kent RL, Blackburn BO, Holmes MB, Reardon JP, Vergeront JM, Schell WL, Christenson E, Bissett ML, Morse EV.** 1982. Molecular epidemiology of antibiotic resistance in *Salmonella* from animals and human beings in the United States. N Engl J Med. **307**:1-6.
10. **Oppegaard H, Steinum T M., Wasteson Y.** 2001. Horizontal transfer of a

multi-drug resistance plasmid between coliform bacteria of human and bovine origin in a farm environment. *Appl Environ Microbiol* **67**: 3732–3734.