A fluoroquinolone-sensitive *Shigella sonnei* UKMCC1015 downregulates the expression of ribosomal proteins in response to sub-lethal concentration of norfloxacin

Wong Bao Chi[1], Muhammad Zarul Hanifah Md Zoqratt[1, 2], Gan I-Ning[1], Qasim Ayub[1, 2], Tan Hock Siew[1, 3*]

[1] School of Science, Monash University Malaysia, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

[2] Monash University Malaysia Genomics Facility, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

[3] Tropical Medicine and Biology Multidisciplinary Platform, 47500 Bandar Sunway, Selangor Darul Ehsan, Malaysia

*Corresponding author: Tan Hock Siew, tan.hocksiew@monash.edu

Abstract

Shigella spp. commonly causes gastroenteritis known as shigellosis, and severe cases require antibiotic treatments. However, the increased presence of antibiotics in the environment in recent years can lead to antibiotic resistance. Sub-lethal antibiotic concentrations can promote resistance, but its effect on sensitive bacteria prior to resistance development is unknown. In this study, the complete genome of *Shigella sonnei* UKMCC1015, a sensitive strain from Malaysia is presented. Subsequently, the transcriptomic response of this strain towards sub-MIC levels of norfloxacin, a commonly used fluoroquinolone is investigated. This study reveals that the expressions of the antimicrobial resistance genes and antibiotic targets (DNA gyrase and topoisomerase IV) are not affected. Interestingly, multiple ribosomal protein operons were downregulated while a small regulatory RNA (sRNA), GcvB was upregulated in response to constant exposure to sub-MIC norfloxacin.. To our knowledge, association between sub-lethal fluoroquinolone and expression of ribosomal protein operons as well as GcvB sRNA have never been reported. Collectively, these findings suggest that there are novel pathways possibly related to ppGpp levels that are involved in the adaptive response of *S. sonnei* towards fluoroquinolones.

Keywords

Shigella sonnei UKMCC1015, whole genome sequencing, comparative phylogenomics, transcriptome analysis, sub-MIC

Impact Statement

In this study, the genome of a sensitive *Shigella sonnei* UKMCC1015 and the transcriptomic analysis of its response towards sub-lethal norfloxacin is presented. This study brings new information towards the multitudes of responses *S. sonnei* can have against a low concentration of antibiotics, including the involvement of the stringent response and a small non-coding RNA called GcvB. This study also highlights the idea that the response of a sensitive strain towards antibiotics are different compared to a resistant strain. Future studies elucidating the exact pathways involving the stringent response and GcvB can lead to the discovery of potential targets as antibiotics or as anti-resistance agents.

Data Summary

The complete genome of *Shigella sonnei* UKMCC1015 and its plasmids has been deposited at GenBank under accession numbers CP060117, CP060118 and CP060119. RNA-seq data are accessible through BioSample accession number SAMN16871737.

Introduction

Shigella species is a common causal agent of global diarrhoea episodes, affecting mainly children younger than 5 years old (Baker & The 2018). Among the four species of *Shigella, Shigella dysenteriae* and *Shigella boydii* are infrequently isolated, while *Shigella flexneri* and *Shigella sonnei* are more commonly encountered in clinical settings. *S. flexneri* is commonly associated in resource-poor countries, while *S. sonnei* is associated with the disease in high-income countries. Due to the improving economic prosperity of many countries, *S. sonnei* has recently replaced *S. flexneri* as the dominant species to cause bacillary dysentery globally (Thompson, Duy & Baker 2015). The current population of *S. sonnei* descended from a common ancestor that led to distinct lineages and spread to other parts of the world, with strains in Asia that are mostly descendants of one lineage from Europe (Holt *et al.* 2012). This etiological shift of shigellosis cause to *S. sonnei* is also reported in Malaysia (Banga Singh *et al.* 2011; Koh *et al.* 2012). Most studies on *Shigella* in Malaysia focused on their antimicrobial susceptibility profiles and highlight the lack of research done on their genetics and phylogeny. Additionally, among all 36 *S. sonnei* complete genomes provided in the NCBI database (last accessed on 09/2020), there are no genomes of *S. sonnei* strains from Southeast Asia.

Due to the misuse and overuse of antibiotics, the increased concentration of antibiotics in the environment has exacerbated the antimicrobial resistance crisis (Gonzalez-Plaza *et al.* 2019). Among all antimicrobial-resistant bacteria, fluoroquinolone-resistant *Shigella spp.* are among the priority 3 (medium) category for new antibiotics research and development. Norfloxacin is one of the first fluoroquinolones synthesised. Although it is not the recommended drug of treatment for shigellosis, its mechanism of action is similar to the recommended drug ciprofloxacin. All fluoroquinolones act by inhibiting DNA gyrase and topoisomerase IV by binding at the enzyme-DNA interface, blocking ligation of the DNA and leads to lethal chromosomal breaks (Alfred, Kerns & Osheroff 2014). By inhibiting bacterial DNA synthesis, it will ultimately cause bacterial cell death.

Studies had shown that low antibiotic concentrations may contribute to antimicrobial resistance. While antibiotic concentration that is higher than the minimum inhibitory concentration (MIC) leads to antimicrobial effects on susceptible cells, concentrations lower than the MIC (also known as sub-MIC) can instead act as a signal for diverse responses that ultimately induce tolerance (Bernier & Surette 2013). As antibiotic pollution leads to sub-MIC levels of antibiotics in the environment, this condition can not only enrich pre-existing resistant variants but also cause a selective pressure for the development of *de novo* resistance (Li *et al.* 2018). Thus, the effect of sub-MIC towards antibiotic tolerance in sensitive bacterial strains needs to be well understood to counteract the inevitable presence of resistant strains.

Although exposure to sub-MIC fluoroquinolone can promote resistance via mutations in the DNA gyrase gene (Ranjbar & Farahani 2019), it is not known what are the transcriptional responses of a sensitive bacterium to low levels fluoroquinolone prior to the development of resistance. This study first presents the complete genome of a *S. sonnei* strain UKMCC1015 from Malaysia, providing the first reference genome for *S. sonnei* from South East Asia. Next, RNA-seq was used to investigate how this fluoroquinolone-sensitive bacterium responds to sub-lethal levels of norfloxacin (a second generation fluoroquinolone). Interestingly, this study revealed some novel observations: 1) none of the gyrase or antibiotic resistance genes were differentially regulated and 2) multiple ribosomal protein operons and a small regulatory RNA, GcvB are differentially expressed, suggesting new unreported pathways are involved in the response of *S. sonnei* towards fluoroquinolone.

Methods and Materials

Strain growth and culture conditions

Shigella sonnei UKMCC1015 strain is a clinical isolate purchased from the University Kebangsaan Malaysia Culture Collection (UKMCC). The strain was cultured in Mueller-Hinton broth (Oxoid, UK) and incubated at 37°C with shaking at 200 rpm overnight. Total genomic DNA was extracted from the mid-log culture using Wizard® Genomic DNA Purification Kit from Promega according to the manufacturer's protocol.

Genome sequencing, assembly and annotation

To obtain a closed genome, we used a hybrid approach employing long-read Nanopore and short-read Illumina MiSeq sequencing. Briefly, for long-read sequencing DNA libraries were prepared using the Ligation Sequencing Kit (SQK-LSK109) according to the manufacturer's protocol and sequenced on a MinION FLOW-MIN106 flowcell and MinION MK1B sequencing device (Oxford Nanopore Technologies). Base calling was conducted with Guppy v3.2.10 on MinKnow 3.6.17 using a fast base calling configuration. For the genome assembly, multiple long-read assemblers were used: Flye v2.6 (Kolmogorov *et al.* 2019), Raven v1.1.10 (https://github.com/lbcb-sci/raven) and Miniasm-Minipolish v0.3 (https://github.com/rrwick/Minipolish). Then, Trycycler (https://github.com/rrwick/Trycycler) was used to obtain the consensus long-read assemblies from different long-read assemblers. The long reads were also used for error correction using Medaka v1.0.3 (https://github.com/nanoporetech/medaka). For Illumina sequencing, DNA libraries were prepared using Nextera XT library preparation kit and sequenced using a 2 X 250 bp paired-end configuration. The Illumina reads were trimmed and used for error correction using Pilon v1.23 (Walker *et al.* 2014). The closed genome was annotated using PGAP (Tatusova *et al.* 2016).

Core genes phylogenetic analysis among all *shigella* spp.

To establish the phylogenetic relatedness of UKMCC1015 strain within all *Shigella*, a total of 122 *Shigella* complete genomes from NCBI, including UKMCC1015, with one *Escherichia coli* K-12 substrain MG 1655 were obtained (Figure 2). The complete genomes of *Shigella* species obtained consist of 19 *S. boydii,* 23 *S. dysenteriae,* 42 *S. flexneri* and 37 *S. sonnei* including UKMCC1015 (Supplementary Table S2). The core genome sequences were determined using Roary v3.13.0 (Page *et al.* 2015), which were then used for multiple sequence alignment using PRANK (Löytynoja 2014). The phylogenetic tree was inferred using RAxML (Stamatakis 2014) with 100 bootstrap replicates and visualized using FigTree v1.4.3 (http://tree.bio.ed.ac.uk/software/figtree). The tree was rooted using *E. coli* K12 strain.

Growth curve under different antibiotic concentrations

The MIC of *S. sonnei* UKMCC1015 was determined to be 0.08125 µg/ml using broth dilution method. To determine the antibiotic concentration used for transcriptome analysis, the growth curve of *S. sonnei* UKMCC1015 was determined using 0xMIC, 0.1xMIC (0.008µg/ml), 0.25xMIC (0.02 µg/ml) and 0.5xMIC (0.04 µg/ml) of norfloxacin (TCI, Japan). Briefly, the overnight bacterial culture was aliquoted to 50ml of broth containing the different antibiotic concentrations. The OD at 600nm was obtained every 15 minutes for 5.5 hours. The growth curve was performed in triplicates. As 0.1xMIC was shown to have the least growth defect on the bacteria, further analysis was performed using 0.1xMIC

Total RNA extraction

RNA isolation of the samples was performed in duplicates. S. sonnei UKMCC1015 were grown overnight at 37°C with shaking at 200 rpm. Four 250ml conical flasks containing MH broth were inoculated with the overnight culture to an initial optical density of 0.05 at 600nm. 0.1xMIC of norfloxacin were added to two of the experimental cultures. The culture was then incubated 37°C with shaking at 200 rpm until it reached mid-log phase (OD 600nm of 0.5). As a control, another set of duplicates was done without exposure towards norfloxacin. The bacterial cell pellets obtained were frozen before RNA extraction. Total RNA extraction was done using Trizol™ reagent (Invitrogen, US) according to the manufacturer's instructions. Briefly, the frozen bacterial cells were defrosted in ice at room temperature, and the cells were collected by centrifuging at 4°C, 4000 rpm for 10 minutes. The obtained pellet was resuspended with Trizol and incubated at room temperature for 5 minutes for complete dissociation of nucleoprotein complexes. After adding chloroform and incubating at room temperature for 5 minutes, the mixture was centrifuged at 13000 rpm at 4°C for 15 minutes. The colourless aqueous phase was obtained and an equal amount of isopropanol was added. The mixture was incubated in ice for 10 minutes, then centrifuged at 13000rpm for 10 minutes. To wash the RNA from the salt residue, 75% ethanol was added to the pellet. After centrifuging at 13000 rpm at 4°C for 5 minutes, the supernatant was completely removed. The pellet was air-dried for around 10 minutes and resuspended in sterile Milli-Q water.

cDNA library preparation and sequencing

rRNA depletion was performed using the MICROB*Express* Bacterial mRNA Enrichment kit (Thermo Fisher Scientific, Massachusetts, USA) following the manufacturer's protocol, followed by cDNA library preparation using the NEBNext Ultra RNA Library Prep Kit for Illumina (New England Biolabs, Massachusetts, USA). DNA library quality control was done by measuring concentration using Qubit 2.0 fluorometer (Invitrogen, California, USA) and assessing the size distribution using TapeStation 2200 (Agilent Technologies,California, USA). All libraries were pooled equimolar and 15pM was loaded into the MiSeq cartridge. Sequencing was performed on the MiSeq system using a 150 bp run configuration.

Bioinformatics analyses

The *Shigella sonneii* UKMCC genome (GenBank assembly accession: GCA_014217935.1) was used as the reference genome to map the RNA sequencing data. Genome annotation was obtained from the NCBI PGAP annotation. To find additional ncRNA genes in the genome, *Shigella sonneii* Ss046 ncRNA sequences from the BSRD database (Li et al., 2013) were used to BLAST against *Shigella sonnei* genome; BLAST hits were then appended into the annotation. To properly merge the two annotation sources if annotations of the two sources overlap, PGAP annotation takes precedence over BLAST BSRD hits.

Illumina adapter sequences and low quality bases were trimmed off using Trimmomatic v0.39 using parameters ILLUMINACLIP:adapters.fa:2:30:10:1:true, SLIDINGWINDOW:5:20, LEADING:20, TRAILING:20 (Bolger et al., 2014). The trimmed, quality-filtered reads were mapped to the reference genome using Bowtie2 v2.3.5 and samtools v1.9 (Langmead and Salzberg, 2012; Li et al., 2009). The counting of the alignments on annotated genes was done using htseq-count v0.11.2 using parameters mode = union, order = pos and minaqual = 0 (Anders et al., 2015). Features included into htseq counting were CDS, rRNA, ncRNA, tRNA, antisense RNA and riboswitch. Differential abundance analyses and visualizations were conducted on Degust web platform (Powell, 2017) using the edgeR quasi-likelihood method. The clusters of orthologous groups (COGs) of the genes were determined using eggNOG (evolutionary genealogy of genes: Non-supervised Orthologous Groups) web-based server (Huerta-Cepas *et al.* 2018).

Results

Genome analysis reveals AMR genes, prophages and genome islands

Using hybrid assembly, a closed chromosome and two plasmids were obtained for *S. sonnei* strain UKMCC1015 (Figure 1, constructed using BRIG (Alikhan *et al.* 2011)). The chromosome has a size of 4,926,332 bp, while the two plasmids (termed UKMCC1015_2 and UKMCC1015_3) have a size of 7,231 and 2,101 bp, respectively, with an average GC content of 51.03%. Annotation with PGAP predicted a total of 4808 coding sequences (CDS), 103 tRNA, 22 rRNA and 11 ncRNA in the UKMCC1015 genome. The UKMCC1015 strain has two plasmids: UKMCC1015_2 (7k bp) and UKMCC1015_3 (2k bp). We performed sequence comparison using *sourmash* (Pierce *et al.* 2019) against the PLSDB database (Galata *et al.* 2018). UKMCC105_2 has the highest similarity (34%) with NZ_CP027259.1 *E. coli* strain EC11 plasmid unnamed4. UKMCC1015_3 has the highest similarity (98.6%) with CM009398.1 *S. sonnei* strain 4351 plasmid p_C.

UKMCC1015 has 45 annotated AMR genes that are involved in the resistance of fluoroquinolone, macrolides, tetracycline and many others (Figure 1a. Full list in Supplementary Table 1). Interestingly, there were no AMR genes detected on the two plasmids, even though numerous resistance genes are often of plasmid origins, such as plasmid-mediated quinolone resistance (PMQR) genes namely *qnr,* or extended spectrum beta-lactamase (ESBLs) (Ranjbar & Farahani 2019). Other than that, phage profiling by PHASTER (Arndt *et al.* 2016) showed 17 prophage regions, including 10 intact regions (Figure 1a). These regions showed homology towards enterobacteria phages including SflV, phiP27, DE3 and Gifsy-2, and they carry multiple transposase and integrase genes. Insertion elements (IS) are associated with genome degradation as they cause loss of gene function by disrupting coding sequences, or genome rearrangement. *Shigella* genomes often contain a high number of IS which leads to formation of pseudogenes (inactivated genes) (Feng, Chen & Liu 2011). There are a total of 404 insertion sequences in UKMCC1015 detected using ISEScan (Xie & Tang 2017) from numerous IS families such as IS1, IS3, and IS4. This genome also contained 662 pseudogenes, most of which are frameshifted, incomplete or contain an internal stop, or multiple problems in one gene. Almost all pseudogenes in *Shigella* spp. are directly caused by IS, demonstrating that IS can have a significant effect on the Shigella genome.

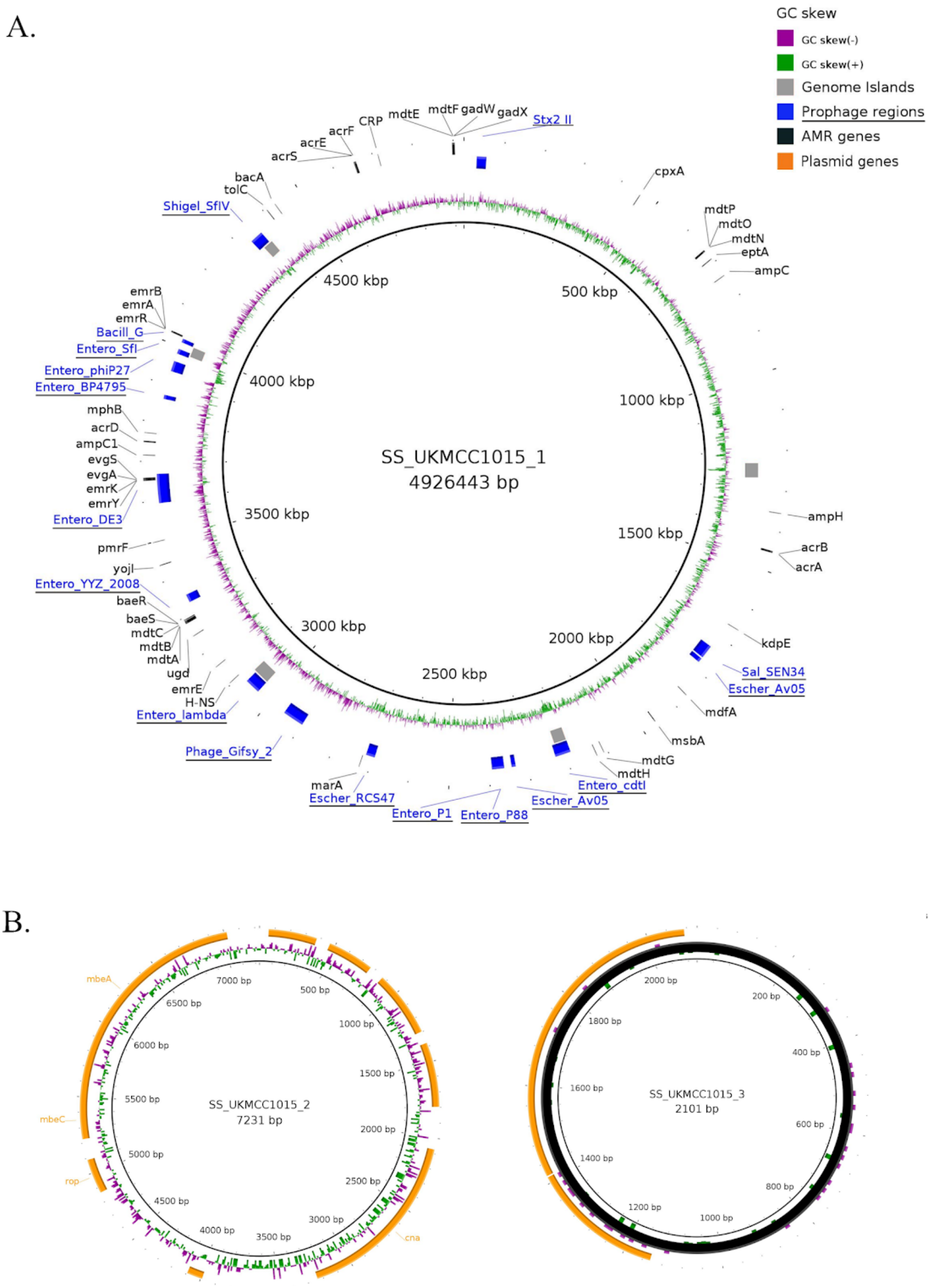
A.
GC skew
GC skew(-)
GC skew(+)
Genome Islands
Prophage regions
AMR genes
Plasmid genes
SS_UKMCC1015_1
4926443 bp
500 kbp
1000 kbp
1500 kbp
2000 kbp
2500 kbp
3000 kbp
3500 kbp
4000 kbp
4500 kbp
acrS
acrE
acrF
CRP
mdtE
mdtF
gadW
gadX
Stx2 II
cpxA
mdtP
mdtO
mdtN
eptA
ampC
ampH
acrB
acrA
kdpE
Sal_SEN34
Escher_Av05
mdfA
msbA
mdtG
mdtH
Entero_cdtI
Escher_Av05
Entero_P88
Entero_P1
marA
Escher_RCS47
Phage_Gifsy_2
Entero_lambda
H-NS
emrE
ugd
mdtA
mdtB
mdtC
baeS
baeR
Entero_YYZ_2008
yojI
pmrF
Entero_DE3
emrY
emrK
evgA
evgS
ampC1
acrD
mphB
Entero_BP4795
Entero_phiP27
Entero_Sfl
Bacill_G
emrR
emrA
emrB
Shigel_SfIV
tolC
bacA
B.
SS_UKMCC1015_2
7231 bp
500 bp
1000 bp
1500 bp
2000 bp
2500 bp
3000 bp
3500 bp
4000 bp
4500 bp
5000 bp
5500 bp
6000 bp
6500 bp
7000 bp
mbeA
mbeC
rop
cna
SS_UKMCC1015_3
2101 bp
200 bp
400 bp
600 bp
800 bp
1000 bp
1200 bp
1400 bp
1600 bp
1800 bp
2000 bp

Figure 1: A. Closed chromosome of *Shigella sonnei* UKMCC1015. UKMCC1015_1 (chromosome), from inner ring outwards: GC skew, Genome Islands (Grey, obtained from IslandViewer, predicted with at least 2 prediction methods), Prophage regions (Blue, underlined, obtained from Phaster), AMR genes (Black, obtained from Abricate using databases CARD). B.

Plasmids of UKMCC1015. UKMCC1015_2 on the left, UKMCC1015_3 on the right, from inner ring outwards: GC content, GC skew, genes in black.

To establish the phylogenetic relatedness of UKMCC1015 strain within all *Shigella*, a total of 122 *Shigella* complete genomes from NCBI, including UKMCC1015, with one *Escherichia coli* K-12 substrain MG 1655 were obtained (Figure 2). Compared to other species, *S. sonnei* has a monophyletic clade, indicating their clonal nature, and this is supported by other phylogenetic studies of *Shigella* (Pupo, Lan & Reeves 2000; Sahl *et al.* 2015). Other species of *Shigella,* however, did not have monophyletic genomic distribution. From the phylogenetic analysis, while *S. sonnei* are in the same clade except for one species *S. sonnei* SE6-1, other *Shigella* species are not monophyletically distributed. This is supported by other phylogenetic studies of *Shigella* (Pupo, Lan & Reeves 2000; Sahl *et al.* 2015). All *Shigella* spp. and enteroinvasive *E. coli* are derived from *E. coli* several times independently and formed a single pathovar (Yang *et al.* 2005). Traditionally, to differentiate *Shigella* from *E. coli* as well as to identify their species, biochemistry and serotyping was used to differentiate the pathogenic *Shigella* from the previously known to be non-pathogenic *E. coli* (Chattaway *et al.* 2017). However, the recurring problem with differentiating *Shigella* from *E. coli* demonstrated that they are closely related to the extent that they should be classified as the same species (Pettengill, Pettengill & Binet 2016). Additionally, species designation among *Shigella* which is done through biochemistry and serotyping does not support the clustering using phylogeny, as can be seen from the mixing of species in a cluster (Figure 3). Thus, the phylogenetic analysis demonstrates that *S. sonnei* is monophyletic according to core genes comparison, unlike other *Shigella* species.

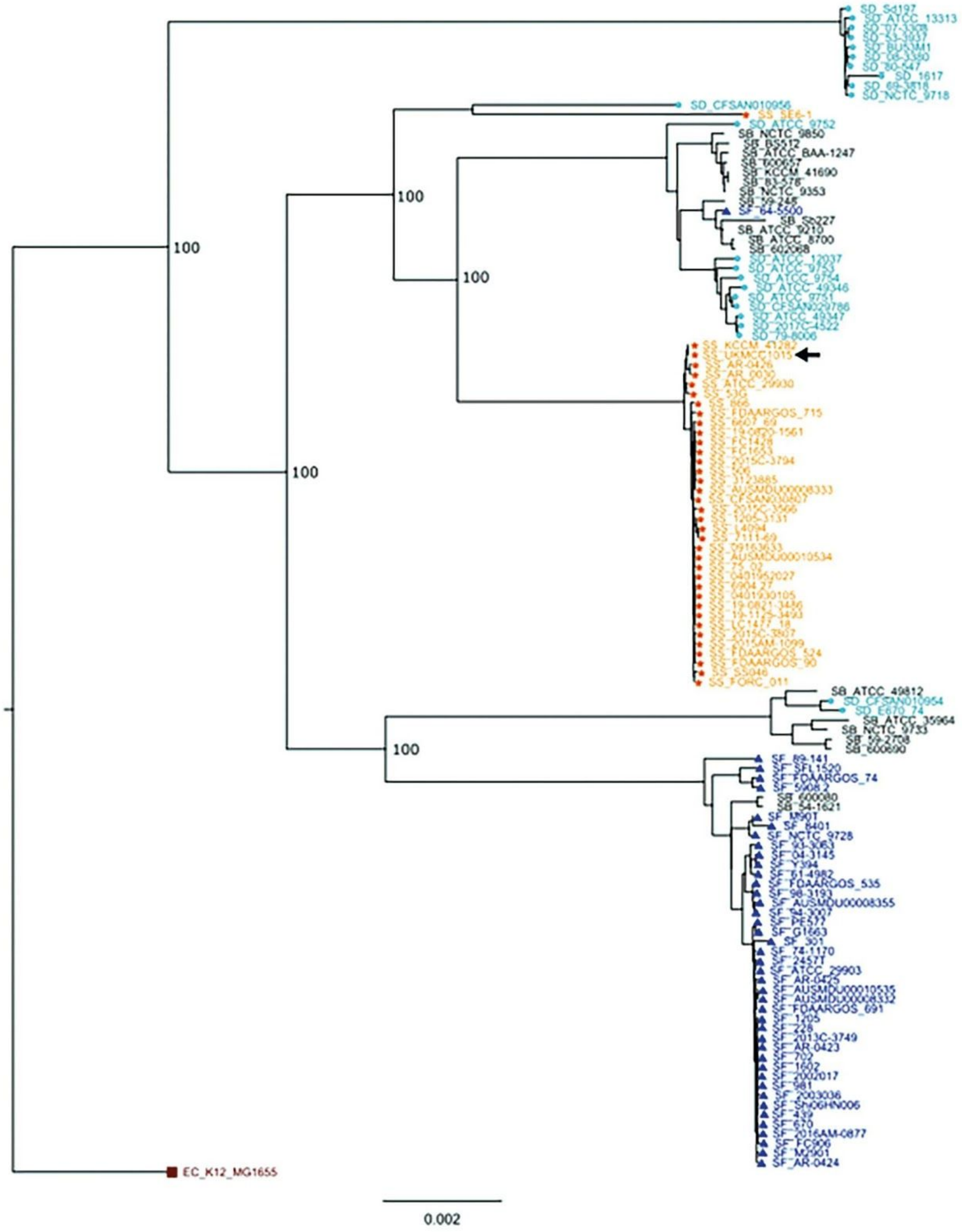
SD_Sd197
SD_ATCC_13313
SD_07-3308
SD_53-3937
SD_BU53M1
SD_08-3380
SD_80-547
SD_1617
SD_69-3818
SD_NCTC_9718
SD_CFSAN010956
SS_SE6-1
SD_ATCC_9752
SB_NCTC_9850
SB_BS512
SB_ATCC_BAA-1247
SB_600657
SB_KCCM_41690
SB_83-578
SB_NCTC_9353
SB_59-248
SF_64-5500
SB_Sb227
SB_ATCC_9210
SB_ATCC_8700
SB_602068
SD_ATCC_12037
SD_ATCC_9753
SD_ATCC_9754
SD_ATCC_49346
SD_ATCC_9751
SD_CFSAN029786
SD_ATCC_49347
SD_2017C-4522
SD_79-8006
SS_KCCM_41282
SS_UKMCC1015
SS_AR-0426
SS_AR_0030
SS_ATCC_29930
SS_53G
SS_866
SS_FDAARGOS_715
SS_6607_69
SS_19-0820-1561
SS_FC1428
SS_FC1653
SS_2015C-3794
SS_506
SS_3123885
SS_AUSMDU00008333
SS_CFSAN030807
SS_2015C-3566
SS_1205-3131
SS_L4094
SS_7111-69
SS_09163633
SS_AUSMDU00010534
SS_75_02
SS_0401952027
SS_6904_27
SS_0401930105
SS_19-0821-3456
SS_19-1117-3493
SS_LC1477_18
SS_2015C-3807
SS_2015AM-1099
SS_FDAARGOS_524
SS_FDAARGOS_90
SS_SS046
SS_FORC_011
SB_ATCC_49812
SD_CFSAN010954
SD_E670_74
SB_ATCC_35964
SB_NCTC_9733
SB_59-2708
SB_600690
SF_89-141
SF_SFL1520
SF_FDAARGOS_74
SF_59082
SB_600080
SB_54-1621
SF_M90T
SF_8401
SF_NCTC_9728
SF_93-3063
SF_04-3145
SF_Y394
SF_61-4982
SF_FDAARGOS_535
SF_98-3193
SF_AUSMDU00008355
SF_94-3007
SF_PE577
SF_G1663
SF_301
SF_74-1170
SF_2457T
SF_ATCC_29903
SF_AR-0425
SF_AUSMDU00010535
SF_AUSMDU00008332
SF_FDAARGOS_691
SF_1205
SF_228
SF_2013C-3749
SF_AR-0423
SF_702
SF_1602
SF_2002017
SF_981
SF_2003036
SF_Shi06HN006
SF_439
SF_670
SF_2016AM-0877
SF_FC906
SF_M2901
SF_AR-0424
EC_K12_MG1655
100
100
100
100
100
100
0.002

Figure 2: Phylogeny of complete *Shigella* genomes with an *E. coli* representative strain. Genomes of all complete *Shigella* (*S. boydii,* SB, black without symbol, *S. dysenteriae,* SD, cyan circle, *S. flexneri,* SF, blue triangle, *S. sonnei,* SS, orange star) genomes including UKMCC1015 (black arrow) and *E. coli* K12 MG1655 (maroon square) were used.

Transcriptome analysis

Growth curve of *S. sonnei* UKMCC1015 under different concentrations of norfloxacin indicated that 0.1xMIC norfloxacin causes the least growth defect (Figure 3a). Thus the transcriptome analysis was performed using 0.1xMIC. The total RNA of *S. sonnei* UKMCC1015 grown on normal broth (control) and under 0.1xMIC norfloxacin (experimental) was extracted, sequenced and analysed. The total read counts of all the genes were used in Degust (Powell 2017) to determine the differentially expressed genes (DEGs) (Figure 3). Relative to the control, there are a total of 151 DEGs in the experimental group. 52 genes show >2-fold upregulation, while 99 genes show >2-fold downregulation (Figure 3b. Full list of DEGs in Supplementary Table 3). Among these 151 genes, 4 are ncRNAs and 9 are pseudogenes. Thus, the remaining 138 genes were categorized using the COG database to further elucidate their role in *S. sonnei* response towards 0.1xMIC norfloxacin (Figure 3c). The top five categories are J (translation, ribosomal structure and biogenesis) with 54 genes, K (transcription) with 14 genes, S (function unknown) with 14 genes, C (energy production and conversion) with 11 genes, and F (nucleotide transport and metabolism) with 9 genes.

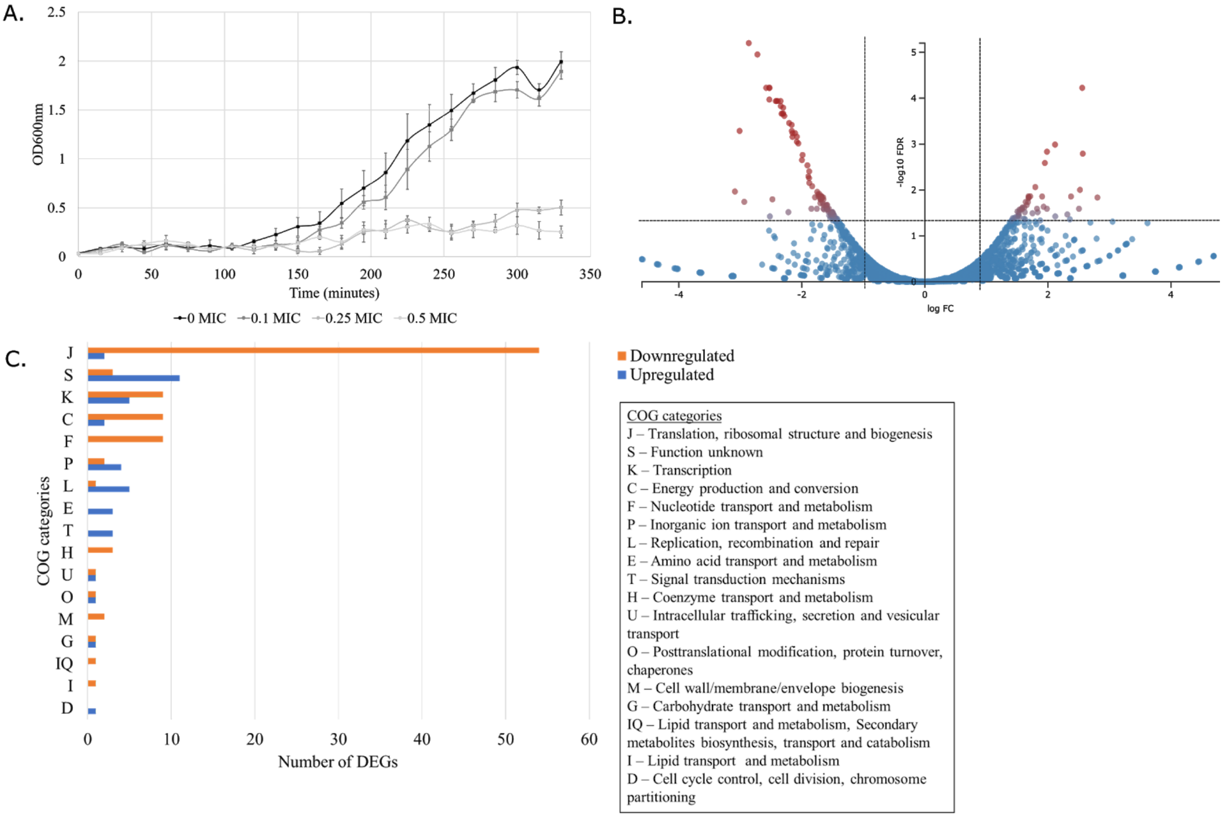

Figure 3: A. Growth curve of *S. sonnei* UKMCC1015 under 0xMIC, 0.1xMIC, 0.25xMIC and 0.5xMIC of norfloxacin. **B.** Volcano plot of transcriptome analysis. **C.** COG categorization of the 136 genes annotated by the non-supervised Orthologous Group (eggNOG).

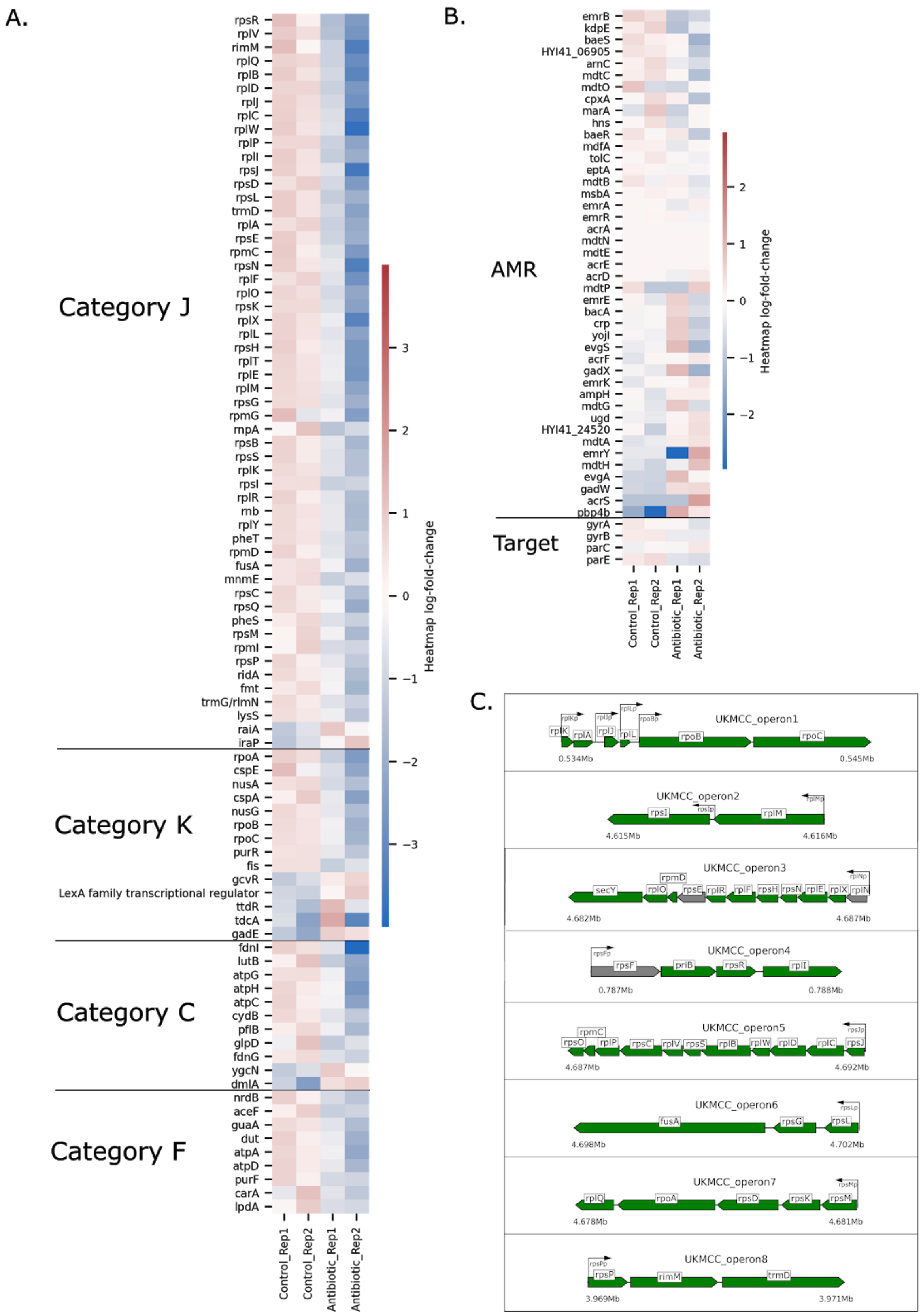

Figure 4: A. Heatmap of DEGs of the four top categories - Category J (translation, ribosomal structure and biogenesis), Category K (transcription), Category C (energy production and conversion), and Category F (nucleotide transport and metabolism). **B.** Heatmap of AMR genes and norfloxacin

targets. **C.** Ribosomal proteins operons of *S. sonnei* UKMCC1015. Green indicates genes that are differentially expressed, grey indicates genes that are not differentially expressed.

Ribosomal protein genes are downregulated in response to sub-lethal norfloxacin

56 DEGs are in the COG category J (translation, ribosomal structure and biogenesis). Two genes, *iraP* and *raiA* are upregulated, while the rest of the genes are downregulated. The downregulated genes encode for ribosomal proteins (r-proteins), which are commonly expressed in operons. Among all the DEGs, 42 genes are expressed in 8 operons, suggesting that any regulation towards these genes are done through the promoters of these 8 operons. To further investigate these operons and their gene order, the locus of these r-proteins genes are identified (Figure 4c) and compared to *Escherichia coli* K12 (Supplementary Figure S1). Except for rpsP operon, the r-proteins operons of both this strain and *E. coli* K12 have the same orientation and the gene positions, indicating their similarity. The downregulation of r-proteins may be a result of a common regulator, and one potential regulator is the alarmone nucleotides, guanosine tetraphosphate and guanosine pentaphosphate, collectively referred to as (p)ppGpp (ppGpp from here on). ppGpp is involved in the stringent response, which is triggered during amino acid or fatty acid starvation, or other environmental stress such as osmotic shock, heat shock and oxygen variation (Pacios *et al.* 2020). Some roles of ppGpp include growth rate control by inhibiting production of ribosomal RNA, and regulation of general metabolism through expression of genes in amino acid biosynthesis, and nucleotide biosynthesis and uptake (Hauryliuk *et al.* 2015). The elevated ppGpp levels leads to this molecule binding to DksA to downregulate the transcription of multiple r-proteins operons. In the stringent response, this downregulation would conserve energy during nutrient starvation. Hence, these results suggest that *S. sonnei* increases the levels of intracellular ppGpp as a response towards low levels of norfloxacin.

The gene *iraP* encodes for anti-adapter protein IraP that increases the stability of the signal stress factor RpoS (σS) and inhibits its proteolysis (Bougdour *et al.* 2008). In particular, *iraP* promoter is positively regulated by ppGpp, leading to the rapid accumulation of the stress factor. Along with other genes regulated by ppGpp, the anti-adapter protein IraP can affect the σS response. The gene *raiA* encodes for ribosome-associated inhibitor A, which role is to interfere with translation by blocking the dissociation of 70S ribosome (Agafonov & Spirin 2004). The upregulation of this inhibitor will cause a disruption in translation, thereby downregulating translation. In general, genes in this category are regulated to ultimately reduce translation.

Downregulation of cold shock proteins and transcription factors

Among the 14 DEGs in the category K (transcription), four genes involved in transcription termination and antitermination are downregulated (*cspA, cspC, nusA* and *nusG).* The two cold-shock domain family proteins (Csps) genes, *cspA* and *cspC* can function as RNA chapenrones to maintain the single-stranded state of target RNA, allowing efficient transcription and translation (Keto-Timonen *et al.* 2016). CspA is inducible by cold, while CspC is constitutively expressed during growth at 37C (Phadtare & Inouye 2001). Both CspA and CspC, were shown to be transcription antiterminators, and their overexpressions are capable of inducing transcription of genes downstream of multiple transcription terminators (Bae *et al.* 2000). Other than the Csps, two transcription termination/antitermination proteins expressions, *nusA* and *nusG* are downregulated. These two transcription factors can affect the speed of RNA chain elongation and termination, and they are both involved in Rho-dependent termination (Burns *et al.* 1998). NusA both promotes intrinsic transcription termination (Gusarov & Nudler 2001; Mondal *et al.* 2016) and antitermination factor in the presence of P7 (Esyunina *et al.* 2015; You *et al.* 2019). Additionally, NusA is demonstrated to be crucial in stress-induced mutagenesis (Cohen & Walker 2010), and it is involved in the cellular response to DNA damage through coupling nucleotide excision repair (NER) and translesion synthesis (TLS) processes to transcription (Cohen *et al.* 2010). Similarly, NusG is essential for transcription termination and antitermination, both enhancing the RNA polymerase processivity (Burova *et al.* 1995) and regulating the Rho-dependent termination (Pasman & von Hippel 2000). The effects of the downregulation of these two transcription factors remains unclear in terms of *S. sonnei* response towards norfloxacin, and needs further investigation.

Downregulation of genes involved in energy production and nucleotide metabolism

All the DEGs in category C (energy production and conversion) and F (nucleotide transport and metabolism) are downregulated except for *dmlA* and *ygcN* from category C. *dmlA* encodes for the d-malate dehydrogenase which allows for aerobic growth on d-malate as the sole carbon source (Stern & Hegre 1966). *ygcN* encodes for the probable electron transfer flavoprotein-quinone oxidoreductase YgcN. It is thus indicated that 0.1xMIC norfloxacin exposure had led the bacteria to downregulate energy production and nucleotide metabolism. This is interesting because ppGpp is shown to positively regulate energy production/conversion (Yang *et al.* 2020). Additionally, ppGpp plays a role in maintaining normal cellular energy homeostasis (Roghanian *et al.* 2019). However, it is unclear whether it is the downregulation of energy production leading to a decrease in ATP production that ultimately leads to the stringent response, or if ppGpp is the one responsible to repress energy production. Other than that, nucleotide transport and metabolism are repressed. All nine DEGs are involved in the synthesis of nucleotides. Nucleotide metabolism has also been shown to be affected by ppGpp (Traxler *et al.* 2008; Liam *et al.* 2018). It is possible that the stringent response has downregulated nucleotide synthesis in order to conserve energy.

Common AMR genes and norfloxacin targets (gyrase and topoisomerase IV) are not differentially expressed during 0.1xMIC exposure

In the *S. sonnei* UKMCC1015 genome, CARD has annotated 45 AMR genes, and 19 of these genes confer resistance to fluoroquinolones by antibiotic efflux. However, none of these genes are differentially expressed during 0.1xMIC norfloxacin exposure. Several studies on fluoroquinolone-resistant *Shigella* indicate that the overexpression of genes encoding efflux pumps contributes to their resistance towards fluoroquinolones (Kim *et al.* 2008; Yang *et al.* 2008; Taneja *et al.* 2015). This suggests that the initial transcriptional response of a fluoroquinolone-sensitive bacteria is different from resistant ones. Instead of overexpressing the efflux pumps available, the sensitive strain perhaps uses other pathways as a response towards 0.1xMIC norfloxacin.

There are 5 DEGs that are commonly regarded as SOS genes (Simmons *et al.* 2008). The genes that are upregulated are *dinB, sulA, recN* and LexA family transcriptional regulator. *nrdB* and *purF* which are LexA-independent genes are downregulated. Norfloxacin and other fluoroquinolones commonly induce the SOS response through production of DNA double-stranded breakage (DSBs) by blocking DNA gyrase and affecting the DNA replication (Pohlhaus & Kreuzer 2005; Qin *et al.* 2015). However, there is little evidence that supports the suggestion that SOS response is induced as only these 5 genes are differentially regulated. No late SOS genes such as *umuCD* were differentially regulated. *umuCD* encodes for error-prone DNA polymerase V, and they are induced only when the damage is extensive (Maslowska *et al.* 2019). Under 0.1xMIC norfloxacin, the damage done by the antibiotic is not sufficient to induce the SOS response.

There are also no changes in the expression of *gyrA, gyrB, parC* and *parE* (Figure 4B). These genes encode for DNA gyrase and topoisomerase IV, which are the targets of norfloxacin. This is interesting because commonly the regulation of antibiotic target genes are often differentially regulated in response to antibiotic exposure. Machuca *et al.* (2017) reported that low-level quinolone-resistant *E. coli* downregulates the genes of DNA gyrase subunits. In contrast, exposure towards norfloxacin induced the expression of *gyrA* and *gyrB* in *Acinetobacter baumannii* (Heo *et al.* 2014). Similarly, *E. coli* overexpresses *gyrA* as a response towards DNA gyrase inhibitors such as norfloxacin and aminocoumarin (Neumann & Quinones 1997; Anderle *et al.* 2008). Even though the main fluoroquinolone resistance mechanism is linked to chromosomal mutations of *gyrA* and *parC* (Fabrega *et al.* 2009) and not their regulation, it is possible that the upregulation of the drug target is an attempt to oversaturate the antibiotics. This increases the levels of DNA gyrase not bound to the drug, thus reducing the amount of double-stranded DNA breaks that the antibiotics can cause. In our case, the normal expression of these target genes suggest that the bacteria has other methods of responding to the effects of norfloxacin.

The increase in the levels of a small regulatory RNA GcvB suggest possible role in antibiotic response

S. sonnei UKMCC1015 differentially expressed a small regulatory RNA GcvB. GcvB is upregulated during exposure to 0.1xMIC norfloxacin, which is interesting because there has not been any correlation between this small regulatory RNA with norfloxacin or fluoroquinolone response. In fact, there is only one study linking GcvB to D-cycloserine, an antibiotic that inhibits cell wall biosynthesis. GcvB in *E. coli* negatively regulates *cycA,* a gene that encodes for a permease that can transport D-cycloserine, thus providing resistance (Pulvermacher *et al.* 2009). However, GcvB can promote mutagenic break repair (MBR) via the general σS stress response by indirectly promoting membrane integrity and suppressing the competing σE membrane stress response (Barreto *et al.* 2016). As norfloxacin causes DSBs, GcvB may be upregulated in response to norfloxacin to fix the damages through MBR. The mutations caused by these repair systems might lead to the chromosome-mediated quinolone resistance, including the mutations in the *gyrA* genes and *parC* genes (Fabrega *et al.* 2009). Thus, the upregulation of GcvB to indirectly promote mutagenesis may be one of the methods used by the sensitive *S. sonnei* to develop resistance.

The targets of GcvB such as *oppA* and *dppA* are not differentially regulated in *S. sonnei* during 0.1xMIC exposure. GcvB, like any other small non-coding RNAs (sRNAs) regulate multiple mRNAs at the post-transcriptional level. For example, GcvB regulates *phoPQ* by binding to the translation initiation region (TIR) of the *phoPQ* mRNA, thereby blocking ribosome binding and translation initiation (Coornaert *et al.* 2013). The mRNA will thus be degraded over time. However, it has been shown that oppA and dppA are not involved in σS stress response, while GcvB does not regulate *rpoS* at the transcriptional level (Jin *et al.* 2009). This may explain why these two genes are not differentially expressed in response to sub-lethal norfloxacin.

Discussion

In this study, the effects of sub-lethal norfloxacin towards the gene expression of the fluoroquinolone-sensitive *S. sonnei* were demonstrated. In response to 0.1xMIC of norfloxacin, *S. sonnei* differentially expressed 148 genes. The results show that the exposure towards sub-MIC of norfloxacin causes *S. sonnei* to: 1) downregulate the ribosomal proteins expression, indirectly indicating that the bacteria increases the production of ppGpp; 2) upregulate a small regulatory RNA GcvB; and 3) trigger the SOS response. More surprisingly, even though this strain has AMR genes in its genome, they are not differentially expressed. Additionally, the targets of the norfloxacin, DNA gyrase and topoisomerase IV are also not differentially expressed. In general, *S. sonnei* downregulates translation, transcription, energy production and nucleotide production in order to cope with the sub-lethal norfloxacin, possibly through a never-before reported pathway that includes increasing ppGpp and GcvB levels in the cell.

Currently, there is no link between ppGpp and GcvB nor with norfloxacin. But there are studies showing that ppGpp can lead to antibiotic tolerance and formation of persister cells, thus contributing to antibiotic resistance (Das & Bhadra 2020). Additionally, ppGpp can inhibit the gene expression of DNA gyrase (*gyrA*) and topoisomerase IV (*parC*), which are targets of norfloxacin (Fernandez-Coll *et al.* 2020). DNA replication is inhibited by ppGpp to help the bacteria stabilize the genome DNA during starvation as well as conserving energy (Fernandez-Coll *et al.* 2020; Yang *et al.* 2020). Yang *et al.* 2020 also showed that ppGpp can suppress SOS system, which is contradictory to another paper stating that the stalling of transcription elongation complexes and the arrest of DNA replication forks by ppGpp can expose the single-stranded DNA and thus induce the SOS response (Srivatsan & Wang 2008).

In this study, both stringent response and SOS response are activated by 0.1xMIC norfloxacin as seen from the downregulation of the r-proteins genes and the upregulation of SOS genes. Comparing the number of genes involved in each of these two systems, the response of the fluoroquinolone-sensitive *S. sonnei* towards norfloxacin at sub-MIC seems to lean towards the stringent response. The bacteria enters a semi-dormant state with the repression of the pertinent genes by ppGpp, with the end result of decreased ribosomal proteins and production, and the decrease in 70S ribosome due to the upregulation of *iraP* by ppGpp. By decreasing the DNA replication, the bacteria decreases the effect of norfloxacin on the DNA gyrase and topoisomerase IV. This is also supported by how *gyrA* and *parC,* the targets of norfloxacin are not being upregulated. Instead of increasing the amount of targets and DNA replication to overwhelm the concentration of antibiotics like others (Heo *et al.* 2014), this fluoroquinolone-sensitive *S. sonnei* decreases its activities and becomes semi-dormant.

Exposure to sub-lethal antibiotics are known to be able to induce tolerance and even resistance in sensitive bacteria (Gullberg *et al.* 2011; Chow, Waldron & Gillings 2015; Wistrand-Yuen *et al.* 2018). This is also seen in our study where exposure to 0.1xMIC and 0.5xMIC of norfloxacin for 250 generations can lead to an increased number of cells able to grow at certain concentrations of antibiotics. Studies on the genome of bacteria that were grown under sub-lethal concentration of antibiotics for more than 100 generations showed that the weaker selective pressure can induce and select for small resistance mutations. This suggests that the low concentration of antibiotics can induce mutations in the genome through the process of inducing ppGpp and GcvB. And therefore the bacteria develops the small resistance mutations that ultimately are selected by the presence of antibiotics in the environment.

In conclusion, 0.1xMIC of norfloxacin triggers SOS response as well as the stringent response in fluoroquinolone-sensitive *S. sonnei* UKMCC1015 as seen through the upregulation of *lexA* and downregulation of r-proteins genes expression. This is the first study to link norfloxacin with the downregulation of r-proteins genes through a possible increase in the cellular ppGpp levels. This is also the first study to observe the increase of the small regulatory RNA GcvB during norfloxacin exposure. The exposure towards sub-MIC of norfloxacin causes the downregulation of transcription, translation, energy production, and nucleotide metabolism in the fluoroquinolone-sensitive *S. sonnei.*

Acknowledgements

This work was supported by the Monash University Malaysia 2019 Early Career Research Seed Grant, Tropical Medicine and Biology Multidisciplinary Platform Monash University Malaysia and the School of Science, Monash University Malaysia.

Conflict of Interest

The authors have no conflict of interests to declare.

References

Alikhan, NF, Petty, NK, Ben Zakour, NL & Beatson, SA 2011, 'BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons', *BMC Genomics,* vol. 12, no. 402.

Allué-Guardia, A, Koenig, SSK, Quirós, P, Muniesa, M, Bono, JL & Eppinger, M 2018, 'Closed genome and comparative phylogenetic analysis of the clinical multidrug resistant *Shigella sonnei* Strain 866', *Genome Biology and Evolution,* vol. 10, no. 9, pp. 2241-2247.

Arndt, D, Grant, JR, Marcu, A, Sajed, T, Pon, A, Liang, Y & Wishart, DS 2016, 'PHASTER: a better, faster version of the PHAST phage search tool', *Nucleic Acids Research,* vol. 44, no. W1, pp. W16-21.

Baker, S & The, HC 2018, 'Recent insights into *Shigella*', *Current opinion in infectious diseases,* vol. 31, no. 5, pp. 449-454.

Banga Singh, KK, Ojha, SC, Deris, ZZ & Rahman, RA 2011, 'A 9-year study of shigellosis in Northeast Malaysia: Antimicrobial susceptibility and shifting species dominance', *Zeitschrift fur Gesundheitswissenschaften* vol. 19, no. 3, pp. 231-236.

Bernier, S & Surette, M 2013, 'Concentration-dependent activity of antibiotics in natural environments', *Frontiers in Microbiology,* vol. 4, no. 20.

Chattaway, MA, Schaefer, U, Tewolde, R, Dallman, TJ & Jenkins, C 2017, 'Identification of *Escherichia coli* and *Shigella* species from whole-genome sequences', *Journal of Clinical Microbiology,* vol. 55, no. 2, pp. 616.

Danner, M-C, Robertson, A, Behrends, V & Reiss, J 2019, 'Antibiotic pollution in surface fresh waters: Occurrence and effects', *Science of The Total Environment,* vol. 664, pp. 793-804.

Feng, Y, Chen, Z & Liu, S-L 2011, 'Gene decay in *Shigella* as an incipient stage of host-adaptation', *PLoS One,* vol. 6, no. 11, pp. e27754-e27754.

Galata, V, Fehlmann, T, Backes, C & Keller, A 2018, 'PLSDB: a resource of complete bacterial plasmids', *Nucleic Acids Research,* vol. 47, no. D1, pp. D195-D202.

González-Plaza, JJ, Blau, K, Milaković, M, Jurina, T, Smalla, K & Udiković-Kolić, N 2019, 'Antibiotic-manufacturing sites are hot-spots for the release and spread of antibiotic resistance genes and mobile genetic elements in receiving aquatic environments', *Environment International,* vol. 130, pp. 104735.

Holt, KE, Baker, S, Weill, FX, Holmes, EC, Kitchen, A, Yu, J, Sangal, V, Brown, DJ, Coia, JE, Kim, DW, Choi, SY, Kim, SH, Da Silveira, WD, Pickard, DJ, Farrar, JJ, Parkhill, J, Dougan, G & Thomson, NR 2012, '*Shigella sonnei* genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe', *Nature genetics,* vol. 44, no. 9, pp. 1056-1059.

Jia, B, Raphenya, AR, Alcock, B, Waglechner, N, Guo, P, Tsang, KK, Lago, BA, Dave, BM, Pereira, S, Sharma, AN, Doshi, S, Courtot, M, Lo, R, Williams, LE, Frye, JG, Elsayegh, T, Sardar, D, Westman, EL, Pawlowski, AC, Johnson, TA, Brinkman, FS, Wright, GD & Mcarthur, AG 2017, 'CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database', *Nucleic Acids Research,* vol. 45, no. D1, pp. 566-573.

Koh, XP, Chiou, CS, Ajam, N, Watanabe, H, Ahmad, N & Thong, KL 2012, 'Characterization of *Shigella sonnei* in Malaysia, an increasingly prevalent etiologic agent of local shigellosis cases', *BMC Infectious Diseases,* vol. 12, no. 1, pp. 122.

Kolmogorov, M, Yuan, J, Lin, Y & Pevzner, PA 2019, 'Assembly of long, error-prone reads using repeat graphs', *Nature Biotechnology,* vol. 37, no. 5, pp. 540-546.

Li, Z, Hu, Y, Yang, Y, Lu, Z & Wang, Y 2018, 'Antimicrobial resistance in livestock: antimicrobial peptides provide a new solution for a growing challenge', *Animal Frontiers,* vol. 8, no. 2, pp. 21-29.

Löytynoja, A 2014, 'Phylogeny-aware alignment with PRANK', *Methods in Molecular Biology,* vol. 1079, pp. 155-70.

Mattock, E & Blocker, AJ 2017, 'How do the virulence factors of *Shigella* work together to cause disease?', *Frontiers in Cellular and Infection Microbiology,* vol. 7, no. 64.

Page, AJ, Cummins, CA, Hunt, M, Wong, VK, Reuter, S, Holden, MTG, Fookes, M, Falush, D, Keane, JA & Parkhill, J 2015, 'Roary: rapid large-scale prokaryote pan genome analysis', *Bioinformatics (Oxford, England),* vol. 31, no. 22, pp. 3691-3693.

Pettengill, EA, Pettengill, JB & Binet, R 2016, 'Phylogenetic analyses of *Shigella* and enteroinvasive *Escherichia coli* for the identification of molecular epidemiological markers: whole-genome comparative analysis does not support distinct genera designation', *Frontiers in Microbiology,* vol. 6, no. 1573.

Pierce, N, Irber, L, Reiter, T, Brooks, P & Brown, C 2019, 'Large-scale sequence comparisons with sourmash', *F1000Research,* vol. 8, no. 1006.

Pupo, GM, Lan, R & Reeves, PR 2000, 'Multiple independent origins of *Shigella* clones of *Escherichia coli* and convergent evolution of many of their characteristics', *Proceedings of the National Academy of Sciences,* vol. 97, no. 19, pp. 10567-10572.

Puzari, M, Sharma, M & Chetia, P 2018, 'Emergence of antibiotic resistant Shigella species: A matter of concern', *Journal of Infection and Public Health,* vol. 11, no. 4, pp. 451-454.

Ramisetty, BCM & Sudhakari, PA 2019, 'Bacterial ‘Grounded’ prophages: hotspots for genetic renovation and innovation', *Frontiers in Genetics,* vol. 10, no. 65.

Ranjbar, R & Farahani, A 2019, '*Shigella:* antibiotic-resistance mechanisms and new horizons for treatment', *Infection and drug resistance,* vol. 12, pp. 3137-3167.

Sahl, JW, Morris, CR, Emberger, J, Fraser, CM, Ochieng, JB, Juma, J, Fields, B, Breiman, RF, Gilmour, M, Nataro, JP & Rasko, DA 2015, 'Defining the phylogenomics of *Shigella* species: a pathway to diagnostics', *Journal of Clinical Microbiology,* vol. 53, no. 3, pp. 951-960.

Stamatakis, A 2014, 'RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies', *Bioinformatics (Oxford, England),* vol. 30, no. 9, pp. 1312-1313.

Tatusova, T, Dicuccio, M, Badretdin, A, Chetvernin, V, Nawrocki, EP, Zaslavsky, L, Lomsadze, A, Pruitt, KD, Borodovsky, M & Ostell, J 2016, 'NCBI prokaryotic genome annotation pipeline', *Nucleic Acids Research,* vol. 44, no. 14, pp. 6614-6624.

Thompson, CN, Duy, PT & Baker, S 2015, 'The rising dominance of *Shigella sonnei:* an intercontinental shift in etiology of bacillary dysentery', *PLoS neglected tropical diseases,* vol. 9, no. 6, pp. e0003708-e0003708.

Walker, BJ, Abeel, T, Shea, T, Priest, M, Abouelliel, A, Sakthikumar, S, Cuomo, CA, Zeng, Q, Wortman, J, Young, SK & Earl, AM 2014, 'Pilon: An integrated tool for comprehensive microbial variant detection and genome assembly improvement', *PLoS One,* vol. 9, no. 11, pp. e112963.

Xie, Z & Tang, H 2017, 'ISEScan: automated identification of insertion sequence elements in prokaryotic genomes', *Bioinformatics (Oxford, England),* vol. 33, no. 21, pp. 3340-3347.

Yang, F, Yang, J, Zhang, X, Chen, L, Jiang, Y, Yan, Y, Tang, X, Wang, J, Xiong, Z, Dong, J, Xue, Y, Zhu, Y, Xu, X, Sun, L, Chen, S, Nie, H, Peng, J, Xu, J, Wang, Y, Yuan, Z, Wen, Y, Yao, Z, Shen, Y, Qiang, B, Hou, Y, Yu, J & Jin, Q 2005, 'Genome dynamics and diversity of *Shigella* species, the etiologic agents of bacillary dysentery', *Nucleic Acids Research,* vol. 33, no. 19, pp. 6445-6458.

Supplementary Data

Start	End	Gene	Resistance against	Resistance mechanism
454192	455565	cpxA	aminocoumarin;aminoglycoside	Antibiotic efflux
661624	663090	mdtP	acridine_dye;nucleoside	Antibiotic efflux
663087	665138	mdtO	acridine_dye;nucleoside	Antibiotic efflux
665138	666169	mdtN	acridine_dye;nucleoside	Antibiotic efflux
688875	690518	eptA	peptide	Antibiotic target alteration
740154	741287	ampC	cephalosporin;penam	Antibiotic inactivation
1354182	1355339	ampH	cephalosporin;penam	Antibiotic inactivation
1448897	1452046	acrB	cephalosporin;fluoroquinolone;glycylcycline;penam;phenicol;rifamycin;tetracycline;triclosan	Antibiotic efflux
1452069	1453262	acrA	cephalosporin;fluoroquinolone;glycylcycline;penam;phenicol;rifamycin;tetracycline;triclosan	Antibiotic efflux
1660042	1660719	kdpE	aminoglycoside	Antibiotic efflux
1862971	1864193	mdfA	benzalkonium_chloride;rhodamine;tetracycline	Antibiotic efflux
1962383	1964131	msbA	nitroimidazole	Antibiotic efflux
2108876	2109922	mdtG	fosfomycin	Antibiotic efflux

2127351	2128559	mdtH	fluoroquinolone	Antibiotic efflux
2724968	2725351	marA	carbapenem;cephalosporin;cephamycin;fluoroquinolone;glycylcycline;monobactam;penam;penem;phenicol;rifamycin;tetracycline;triclosan	Antibiotic efflux, reduced permeability
3104078	3104491	H-NS	cephalosporin;cephamycin;fluoroquinolone;macrolide;penam;tetracycline	Antibiotic efflux
3155558	3155890	emrE	macrolide	Antibiotic efflux
3248822	3249988	ugd	peptide	Drug target alteration
3283808	3285049	mdtA	aminocoumarin	Antibiotic efflux
3285049	3288171	mdtB	aminocoumarin	Antibiotic efflux
3288172	3291249	mdtC	aminocoumarin	Antibiotic efflux
3292662	3294065	baeS	aminocoumarin;aminoglycoside	Antibiotic efflux
3294062	3294784	baeR	aminocoumarin;aminoglycoside	Antibiotic efflux
3439794	3441437	yojI	peptide	Antibiotic efflux
3499016	3499984	pmrF	peptide	Drug target alteration
3652181	3653719	emrY	tetracycline	Antibiotic efflux
3653719	3654774	emrK	tetracycline	Antibiotic efflux
3655298	3655912	evgA	fluoroquinolone;macrolide;penam;tetracycline	Antibiotic efflux

3655917	3659510	evgS	fluoroquinolone;macrolide;penam;tetracycline	Antibiotic efflux
3714636	3715940	ampC1	cephalosporin;penam	Antibiotic inactivation
3747259	3750372	acrD	aminoglycoside	Antibiotic efflux
3770604	3771080	mphB	macrolide	Antibiotic inactivation
4026263	4026793	emrR	fluoroquinolone	Antibiotic efflux
4026920	4028092	emrA	fluoroquinolone	Antibiotic efflux
4028109	4029647	emrB	fluoroquinolone	Antibiotic efflux
4407734	4409221	tolC	aminocoumarin;aminoglycoside;carbapenem;cephalosporin;cephamycin;fluoroquinolone;glycylcycline;macrolide;penam; penem; peptide; phenicol; rifamycin; tetracycline; triclosan	Antibiotic efflux
4432735	4433554	bacA	peptide	Drug target alteration
4651319	4651981	acrS	cephalosporin;cephamycin;fluoroquinolone;glycylcycline;penam;phenicol;rifamycin;tetracycline;triclosan	Antibiotic efflux
4652380	4653537	acrE	cephalosporin;cephamycin;fluoroquinolone;penam	Antibiotic efflux
4653549	4656653	acrF	cephalosporin;cephamycin;fluoroquinolone;penam	Antibiotic efflux
4713203	4713835	CRP	fluoroquinolone;macrolide;penam	Antibiotic efflux

4897300	4898457	mdtE	fluoroquinolone;macrolide;penam	Antibiotic efflux
4898482	4901595	mdtF	fluoroquinolone;macrolide;penam	Antibiotic efflux
4901958	4902686	gadW	fluoroquinolone;macrolide;penam	Antibiotic efflux
4903054	4903878	gadX	fluoroquinolone;macrolide;penam	Antibiotic efflux

Table S1: List of antimicrobial resistance (AMR) genes in the genome of UKMCC1015, identified by Abricate using CARD database.

Species	**Name of strain**	**Genome assembly identifier**
SS	**UKMCC1015**	GCA_014217935.1
SS	0401930105	GCA_013146805.1
SS	0401952027	GCA_013146745.1
SS	09163633	GCA_013146765.1
SS	1205.3131	GCA_013147185.1
SS	19.0820.1561	GCA_013147065.1
SS	19.0821.3486	GCA_013146845.1
SS	19.1125.3493	GCA_013146785.1
SS	2015AM-1099	GCA_002142635.1
SS	2015C-3566	GCA_002224625.1
SS	2015C-3794	GCA_002224605.1
SS	2015C-3807	GCA_002224465.1
SS	3123885	GCA_013146725.1
SS	506	GCA_013201485.1
SS	53G	GCA_000283715.1
SS	6607.69	GCA_009931515.2
SS	6904.27	GCA_009931495.2

SS	7111.69	GCA_013147165.1
SS	75/02	GCA_002150905.1
SS	866	GCA_003265245.1
SS	AR-0426	GCA_008727215.1
SS	AR_0030	GCA_003606425.1
SS	ATCC 29930	GCA_002950395.1
SS	AUSMDU00008333	GCA_900659655.1
SS	AUSMDU00010534	GCA_009664515.1
SS	CSFAN030807	GCA_002442535.1
SS	FC1428	GCA_006843305.1
SS	FC1653	GCA_002735385.2
SS	FDAARGOS_524	GCA_003855115.1
SS	FDAARGOS_715	GCA_009729995.1
SS	FDAARGOS_90	GCA_001558295.2
SS	FORC_011	GCA_001518855.1
SS	KCCM 41282	GCA_007106385.1
SS	L4094	GCA_013147205.1
SS	LC1477/18	GCA_004328985.1

SS	SE6-1	GCA_013374815.1
SS	SS046	GCA_000092525.1
SB	54-1621	GCA_002949455.1
SB	59-248	GCA_002949475.1
SB	59-2708	GCA_002949495.1
SB	600080	GCA_013394495.1
SB	600657	GCA_013394535.1
SB	600690	GCA_013394515.1
SB	602068	GCA_013394555.1
SB	83-578	GCA_002949535.1
SB	ATCC 35964	GCA_002950035.1
SB	ATCC 49812	GCA_002950135.1
SB	ATCC 8700	GCA_002946735.1
SB	ATCC 9210	GCA_001027225.1
SB	ATCC BAA-1247	GCA_002950275.1
SB	BS512	GCA_000020185.1
SB	KCCM 41690	GCA_007106565.1
SB	NCTC 9353	GCA_002950315.1

SB	NCTC 9733	GCA_002950375.1
SB	NCTC 9850	GCA_002949175.1
SB	Sb227	GCA_000012025.1
SD	07-3308	GCA_002949755.1
SD	08-3380	GCA_002949815.1
SD	1617	GCA_000497505.1
SD	2017C-4522	GCA_002949415.1
SD	53-3937	GCA_002949775.1
SD	69-3818	GCA_002949715.1
SD	79-8006	GCA_010231685.1
SD	80-547	GCA_002949855.1
SD	ATCC 12037	GCA_002950015.1
SD	ATCC 13313	GCA_002949675.1
SD	ATCC 49346	GCA_002950075.1
SD	ATCC 49347	GCA_002950095.1
SD	ATCC 9751	GCA_002949915.1
SD	ATCC 9752	GCA_002950115.1
SD	ATCC 9753	GCA_002950155.1

SD	ATCC 9754	GCA_002950175.1
SD	BU53M1	GCA_002741615.1
SD	CSFAN010954	GCA_002949875.1
SD	CSFAN010956	GCA_002949935.1
SD	CSFAN029786	GCA_009662135.1
SD	E670/74	GCA_002943855.1
SD	NCTC 9718	GCA_002949835.1
SD	Sd197	GCA_000012005.1
SF	0228	GCA_001578125.1
SF	04-3145	GCA_002949255.1
SF	0439	GCA_002240115.1
SF	0670	GCA_002240135.1
SF	0702	GCA_002240095.1
SF	1205	GCA_001579965.1
SF	1602	GCA_002240075.1
SF	2002017	GCA_000022245.1
SF	2003036	GCA_000743955.1
SF	2013C-3749	GCA_010231485.1

SF	2016AM-0877	GCA_003719775.1
SF	2457T	GCA_000007405.1
SF	301	GCA_000006925.2
SF	5908.2	GCA_009931475.1
SF	61-4982	GCA_002950235.1
SF	64-5500	GCA_002949515.1
SF	74-1170	GCA_002950255.1
SF	8401	GCA_000013585.1
SF	89-141	GCA_002946695.1
SF	93-3063	GCA_002949575.1
SF	94-3007	GCA_002741635.1
SF	98-3193	GCA_002949695.1
SF	981	GCA_001580175.1
SF	AR-0423	GCA_008727275.1
SF	AR-0424	GCA_008727255.1
SF	AR-0425	GCA_008727235.1
SF	ATCC 29903	GCA_002950215.1
SF	AUSMDU00008332	GCA_900659665.1

SF	AUSMDU00008355	GCA_900659645.1
SF	AUSMDU00010535	GCA_009664535.1
SF	FC906	GCA_002174695.2
SF	FDAARGOS_535	GCA_003855135.1
SF	FDAARGOS_691	GCA_013364115.1
SF	FDAARGOS_74	GCA_000783735.2
SF	G1663	GCA_001021855.1
SF	M2901	GCA_013402835.1
SF	M90T	GCA_004799585.1
SF	NCTC 9728	GCA_002950335.1
SF	PE577	GCA_008370775.1
SF	SFL1520	GCA_004171185.1
SF	Shi06HN006	GCA_000743995.1
SF	Y394	GCA_002442995.2
EC	K12_MG1655	GCA_000005845.2

Table S2: List of strains used in phylogenetic tree, including 19 *Shigella boydii,* 24 *Shigella dysenteriae,* 43 *Shigella flexneri,* 36 *Shigella sonnei,* and *Escherichia coli* K-12 substrain MG 1655

Table S3: Total list of genes and non-coding RNAs annotated in the transcriptomic study. Bold indicates the differentially expressed genes (DEGs) in *Shigella sonnei* UKMCC1015 in response to 0.1xMIC of norfloxacin.

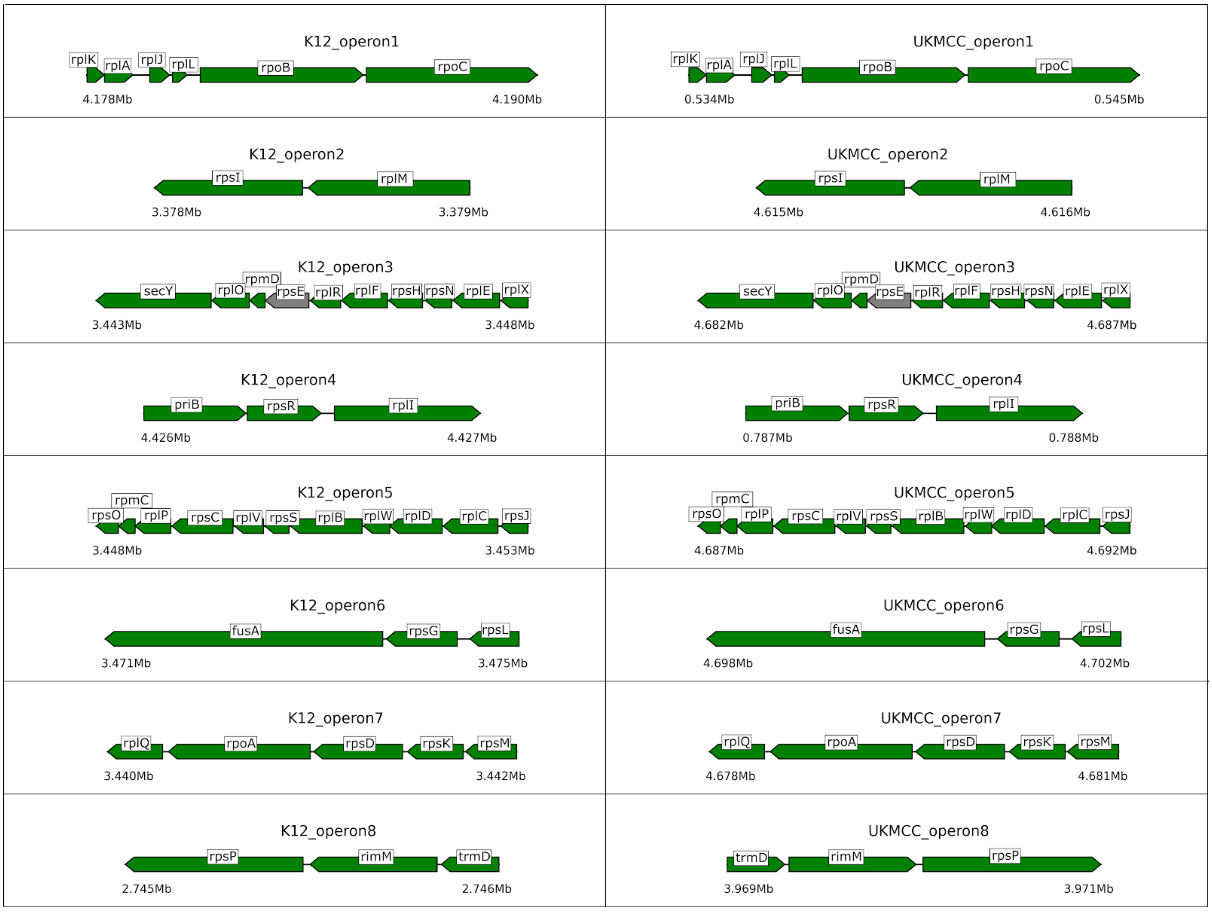
K12_operon1
rplK
rplA
rplJ
rplL
rpoB
rpoC
4.178Mb
4.190Mb
UKMCC_operon1
rplK
rplA
rplJ
rplL
rpoB
rpoC
0.534Mb
0.545Mb
K12_operon2
rpsI
rplM
3.378Mb
3.379Mb
UKMCC_operon2
rpsI
rplM
4.615Mb
4.616Mb
K12_operon3
rpmD
secY
rplO
rpsE
rplR
rplF
rpsH
rpsN
rplE
rplX
3.443Mb
3.448Mb
UKMCC_operon3
rpmD
secY
rplO
rpsE
rplR
rplF
rpsH
rpsN
rplE
rplX
4.682Mb
4.687Mb
K12_operon4
priB
rpsR
rplI
4.426Mb
4.427Mb
UKMCC_operon4
priB
rpsR
rplI
0.787Mb
0.788Mb
K12_operon5
rpmC
rpsO
rplP
rpsC
rplV
rpsS
rplB
rplW
rplD
rplC
rpsJ
3.448Mb
3.453Mb
UKMCC_operon5
rpmC
rpsO
rplP
rpsC
rplV
rpsS
rplB
rplW
rplD
rplC
rpsJ
4.687Mb
4.692Mb
K12_operon6
fusA
rpsG
rpsL
3.471Mb
3.475Mb
UKMCC_operon6
fusA
rpsG
rpsL
4.698Mb
4.702Mb
K12_operon7
rplQ
rpoA
rpsD
rpsK
rpsM
3.440Mb
3.442Mb
UKMCC_operon7
rplQ
rpoA
rpsD
rpsK
rpsM
4.678Mb
4.681Mb
K12_operon8
rpsP
rimM
trmD
2.745Mb
2.746Mb
UKMCC_operon8
trmD
rimM
rpsP
3.969Mb
3.971Mb

Figure S1: Comparison of ribosomal protein operons between *E. coli* K12 (left) and *S. sonnei* UKMCC1015 (right). Other than operon 8 where the orientation of the operon in *S. sonnei* is the opposite of *E. coli* K12, the other 7 operons are similar to each other.