Supporting information

Studies contributing data and bacterial isolates

Data and isolates from several sequential studies were combined for the purposes of this investigation [1-3]. All bacterial isolates used in this study are shown in Table S1. All enrolled patients in the studies were treated as inpatients, were under the age of ten years, and there were no fatalities. The primary study (23 isolates, coded MS) was conducted at the HTD from January 1995 to August 1996. The enrollment and clinical observations for this randomized controlled trial (RCT) are as described previously [3]. Briefly, children that were aged >3 months and < 10 years, admitted to HTD with fever and bloody or mucoid diarrhea (bloody/mucoid diarrhea was defined >3 loose stools with obvious blood/mucus) for <5 days were eligible for enrolment. The second study (36 isolates, coded DE) was conducted at the HTD between March 2000 and December 2002. This study was a clinical and microbiological investigation of the etiology of diarrhea in the pediatric population admitted to the HTD [3]. Briefly, children who were aged >3 months and <10 years, admitted to HTD with acute diarrhea (defined as three or more loose stools or at least one bloody loose stool within a 24-hour period) for <3 days were eligible for enrollment. The third study was a RCT conducted at HTD between June 2006 and December 2008 (54 isolates, coded EG). This RCT compared ciprofloxacin and gatifloxacin for treating acute dysentery in Vietnamese children (controlled trials number ISRCTN55945881) [1]. The inclusion criteria for the RCT were as previously described and equivalent to those used in the primary study. The fourth study was a clinical and etiological investigation of diarrhea infections in HCMC (55 isolates, numerical codes in Table S1). All children who were

under 5 years of age with acute diarrheal disease admitted to the Gastrointestinal Department of CH1, CH2 and HTD between May 2009 and April 2010 in HCMC were eligible for participation [4]. Diarrhea was defined as three or more loose stools or at least one bloody loose stool within a 24-hour period. Children were excluded if they did not live within HCMC, had been pre-treated with antimicrobials, had multiple complications unrelated to diarrheal disease or did not provide informed consent. Bacterial isolates from Hue and Khanh Hoa were donated by these institutions for the purposes of this study and originated from the diagnostic microbiology laboratories; samples were collected as part of routine clinical care for diagnostic purposes. Also included in the study were the previously sequenced genomes of 19 isolates collected at the Khanh Hoa general hospital and community clinics during a diarrhea surveillance study conducted during 2001-2003 [5].

Detection of plasmids and resistance genes

To identify likely plasmids within the accessory genome, we performed a BLAST search of each accessory gene in each of the *S. sonnei* assemblies and generated a correlation table containing entries for each pair of genes co-occurring within 1000 bp of each other in a contig, weighted by the number of isolates in which this co-occurrence was detected. BioLayout 3D was used to infer a network from this data, and thereby identify groups of genes that likely occur in the same replicon. The RAST annotation of the genes was then used to determine whether each replicon was likely to be of plasmid, phage or other origin.

Resistance genes were identified as previously described [6]. Briefly, annotations of reference genomes or novel sequences were scanned for antimicrobial- or resistance-related

keywords. The presence of these genes in each isolate was extracted from the gene content data described above and compared to resistance phenotypes. For each isolate, a genetic basis for every resistance phenotype could be identified, in the form of a resistance gene or *gyrA* mutation (Table S1).

ESBL plasmids

Two large (> 70 kbp) plasmids were identified encoding extended spectrum beta-lactamase (ESBL) CTX-M genes. One of these was pEG356, a 71 kbp plasmid harbouring CTX-M-14, originally identified in a third-generation cephalosporin resistant *S. sonnei* from HCMC in 2001 [7]. Here, pEG356 was detected in one additional isolate from HCMC, one isolate from Hue and multiple strains from KH spanning both of the KH clones (KH1, N=4/11; KH2, N=11/16; see Fig. S6A, Table S1). The detection of pEG356 in different subclones within the same geographical location indicates that the genetically distinct *S. sonnei* clones co-circulating in KH have had access to a common gene pool.

We also detected an unrelated and novel ESBL encoding plasmid pKHSB1, 100 kbp in size, bearing a CTX-M-15 gene (submitted to GenBank under accession TBA). This was widespread among HCMC *S. sonnei* isolates, and became ubiquitous after 2006 (Fig. 2, Fig. S5a). pKHSB1 is an IncI1 plasmid and is nearly identical to pEK204, isolated from a uropathogenic *E. coli* strain in the United Kingdom (7 SNPs; Fig. S6a) [8]. All the plasmids share an insertion of Tn*3* carrying the $bla_{\text{TEM-1}}$ gene into the same site within the plasmid backbone (Fig. S6b), However, pEK204 carries a CTX-M-3 gene, inserted in a different position within Tn*3* to that of the CTX-M-15 gene in pKHSB1 (Fig. S6b). We also

identified a very similar plasmid in a 2003 Korean *S. sonnei* isolate within our previously sequenced global collection [6], which is closer to pKHSB1 than pEK204 (3 SNPs; Fig. S6a) and shares the same Tn3/$bla_{\text{TEM-1}}$ insertion but does not contain a CTX-M insertion. This confirms that the HCMC *S. sonnei* plasmid pKHSB1 has not been transferred directly from uropathogenic *E. coli* ST131, but descends from an IncI1 backbone that has been circulating in Asian *S. sonnei* for some time and has only recently acquired ESBL functionality, in a separate event from the ESBL acquisition occurring in pEK204.

Summary of colicin plasmid investigations

Within the VN clone we detected two small plasmids, pDPT1 and pDPT2, each encoding a colicin and a corresponding immunity protein (Fig. 3). Colicins are bacteriocidal toxins, produced by Gram-negative organisms to kill susceptible organisms, reducing potential competition for space and nutrients within their immediate environment [9]. Here, pDPT1, encoding an E5 type colicin and immunity gene, became fixed during the first selective sweep and has been universally maintained since (Fig. S1b). Plasmid pDPT2, encoding a JS-type colicin and immunity protein, concurrently entered the population with pDPT1, but has been sporadically lost following the second selective sweep and was almost undetectable after the fourth selective sweep (Fig. S1b, Fig. S5a).

To investigate the functional significance of the two colicin plasmids, we performed over 600 assays to assess colicin activity across 27 bacterial isolates, summarized in Fig 3. Detailed methods are provided below. Representative *S. sonnei* from each sweep, and *E. coli* isolates, were exposed to colicin extract from wild type isolates and laboratory

constructed strains containing the E5 operon (pDPT1), the JS operon (pDPT2) or both. Under experimental conditions, the JS colicin extract did not show any detectable inhibition of bacterial growth in any organism tested (Fig. 3). However, colicin extracts from wild-type *S. sonnei* carrying pDPT1 or pDPT1+pDPT2, and from *S. sonnei* or *E. coli* with pDPT1 or pDPT1+pDPT2 introduced in the laboratory, consistently demonstrated growth inhibition against hypothesized "non-immune" *E. coli* and *S. sonnei*, but not those containing pDPT1 (Fig. 3). These data demonstrate that the E5 colicin plasmid pDPT1, identified through sequencing, provides a competitive advantage over non-pDPT1 containing *S. sonnei* and *E. coli*. Given that the introduction and maintenance of the pDPT1 coincides precisely with the first selective sweep (Fig. 2, Fig. S1), it is likely that the activity of this colicin was a driver of the first selective sweep in Vietnamese *S. sonnei* and is linked to the increased contribution of *S. sonnei* to the overall dysentery burden in Vietnam.

Assembly of colicin plasmids

The colicin genes of plasmid pDPT1 were first identified during the gene content analysis described above, as a set of genes frequently detected in the VN clade, including three with >99% nucleotide sequence identity to the plasmid ColE5-099 colicin E5, immunity and lysis genes (X15857) [10] and one with 99% identity to the *mobA* plasmid mobilization gene from plasmid pO26-S4 (FJ004638). Searching for these genes in *de novo* assembled contigs by blastn search revealed these genes were adjacent to each other, usually on a single contig spanning ~6.8 kbp. Alignment of these colicin E5-containing contigs from different *S. sonnei* genomes revealed that they shared the same content, but different start and end

points, consistent with an identical circular molecule of 6,826 bp in each isolate. The consensus sequence was annotated and submitted to GenBank as plasmid pDPT1 (accession HF565446) (Fig. 3).

The colicin genes of pDPT2 were first identified as a set of genes frequently detected in the VN clade, with homology to Js colicin, immunity and lysis genes. Blast search of the sequences in the novel contig sets identified all three genes together in contigs ~5 kbp in size. Alignment of these contigs confirmed their similarity to each other and overlapping ends indicative of circular structure. A blastn search of the NCBI database identified a very similar plasmid (>99% nucleotide sequence identity across the full length of sequence) to *E. coli* plasmid pMG828-3 7 (DQ995353) and *S. sonnei* plasmid pScol7 (AB062753) which encode colicin type 7 (Js). pDPT2 was nearly identical to pMG828-3, hence the annotation from pMG828-3 was transferred to the consensus sequence for pDPT2, which was submitted to GenBank (accession HF565445) (Fig. 3).

Details of colicin assays

A selection of twenty *S. sonnei* spanning the selective sweeps, and two laboratory *E. coli* strains (DH5α and K12) were screened by PCR amplification for the presence of pDPT1 (E5) and pDPT2 (JS). All strains (containing either no colicin operon, E5, JS or E5 + JS) were exposed to crude colicin extract from wild type and laboratory-constructed strains of *S. sonnei* or *E. coli* containing pDPT1, pDPT2 or both, activity was determined by measuring growth inhibition (Fig. 3).

To construct tractable plasmids containing the E5 and JS operons, DNA fragments containing each of these operons were amplified using Biotaq polymerase (Bioline, UK) and ligated with the kanamycin resistance cassette from pDK4 (accession number: AY048743.1) and the chloramphenicol resistance cassette from pKD3 (accession number: AY048742.1), respectively. The generated plasmids were designated as pDPT3(E5) and pDPT4(JS). Plasmids were visualized by agarose gel electrophoresis and purified. Strains *E. coli* DH5α and *S. sonnei* MS48 (without colicin plasmids) were electrotransformed with pDPT3 and pDPT4 using a Gene pulser II electroporation system (Biorad, USA). Transformants were selected on LB media supplemented with 100µg/ml kanamycin, 100µg/ml chloramphenicol or both. All potential transformants were screened by PCR amplification and sequencing to ensure the presence of the plasmid(s) (Fig. 3).

To extract crude colicin, 100µl of hypothetically colicinogenic and non-colicinogenic stationary phase bacteria were inoculated into 10ml LB broth and incubated at 37°C with aeration. After three hours, cultures were inoculated with 8 µl of mitomycin C (final concentration 0.4 µl/ml) or 8 µl of sterile LB broth. Bacterial cultures were incubated at 37°C with aeration for an additional two hours. Cultures were pelleted by centrifugation, the supernatant was harvested, sonicated and filter sterilized. Crude colicin extracts were stored -20°C until required for experimentation. Colicin activity was assessed by, inoculating 100µl of stationary phase bacterial isolate into TOP agar (0.7 % LB) onto LB agar plates. After equilibration at ambient temperature, 10µl of each of the colicin extracts was inoculated in a grid formation (4 x4) onto the LB plate. Again, the LB plates equilibrated at ambient temperature and incubated at 37°C overnight. Colicin activity was

assessed by calculated the size of the zone of growth inhibition in comparison to a negative control (phosphate buffered saline). Zone sizes we scored by diameter 0; no zone, 1, <2mm and 2, >2mm. All colicin assays were performed in triplicate and recorded independently by two individuals.

Supporting Figures

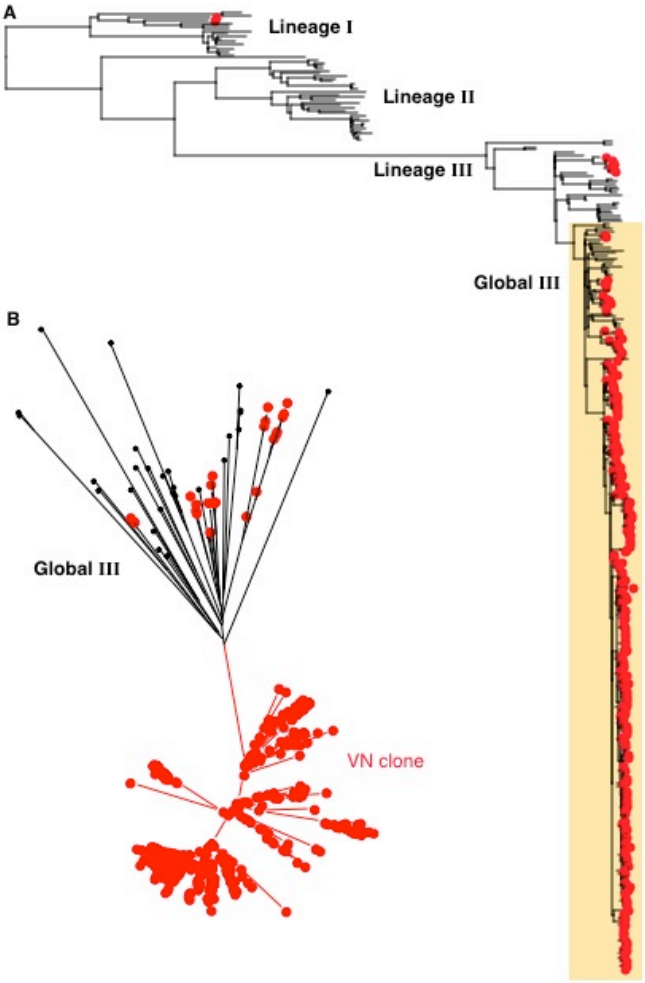


Figure S1. The VN clone of global *Shigella sonnei*

Maximum likelihood trees based on genome-wide chromosomal SNPs. Vietnamese isolates are highlighted in red. (A) Phylogenetic tree for the 263 Vietnamese *S. sonnei* sequenced in this study plus 113 non-Vietnamese *S. sonnei* sequenced previously (total N=376). (B) Phylogenetic tree for the Global III clade, which includes all but 8 of the Vietnamese *S. sonnei*. The Vietnam-specific VN clone, is labeled.

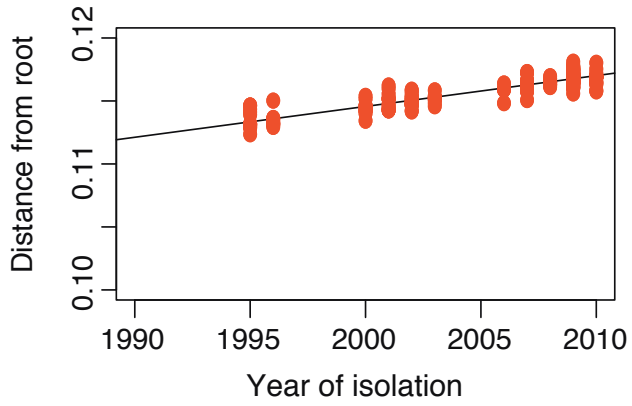


Figure S2. The mutation rate of the *Shigella sonnei* VN clone

Branch lengths (from maximum likelihood phylogenetic tree shown in Fig. 1) vs date of isolation for Vietnamese *S. sonnei*. Pearson correlation coefficient = 0.92, $p < 0.0001$.

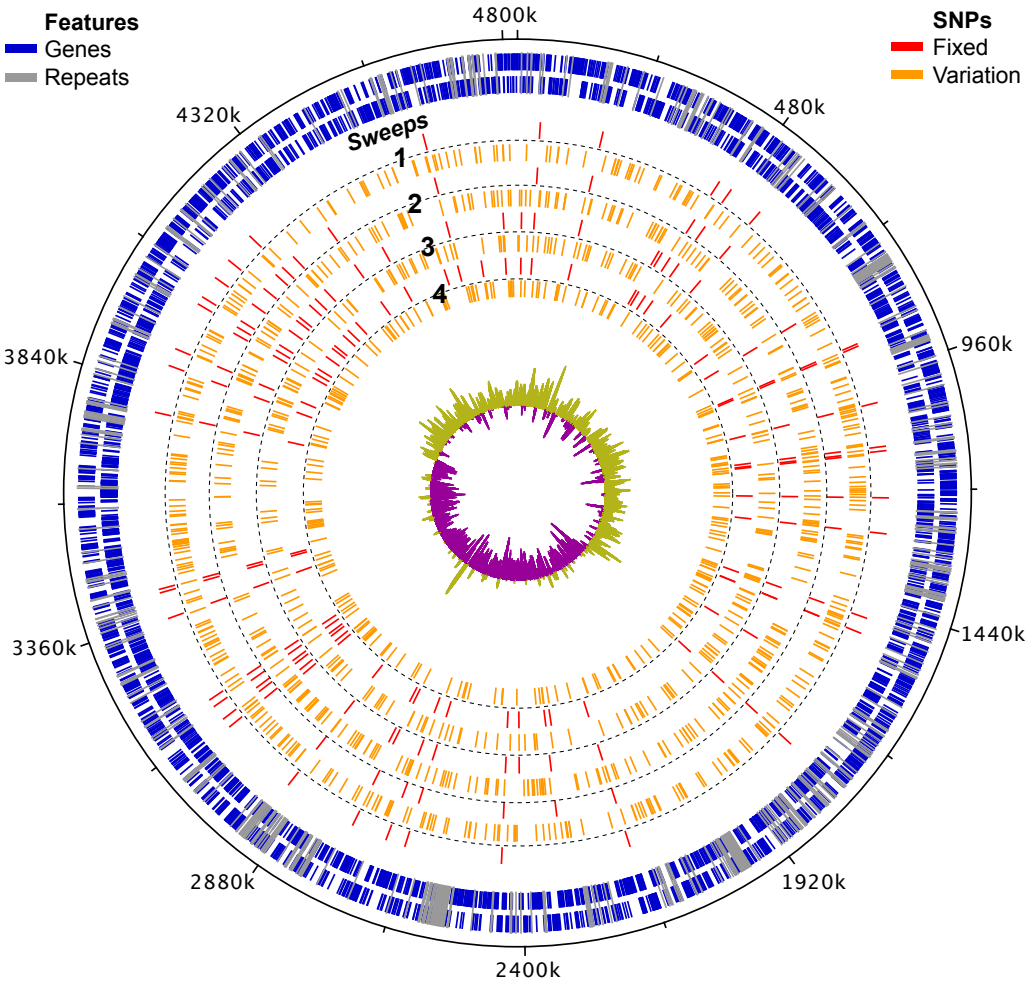


Figure S3. SNPs identified in the *Shigella sonnei* chromosome following the emergence of the VN clone. Circular map of the Ss046 reference chromosome, coordinates are given outside the circle, where k represents 1,000 bp. Outer rings indicate genes (blue) annotated on the forward and reverse strands, and repetitive or prophage regions (grey) excluded from SNP calling. Remaining rings represent SNPs that were fixed by (red) or arose following (orange) the four sweeps marked in Fig. 2. G+C skew is shown in the centre.

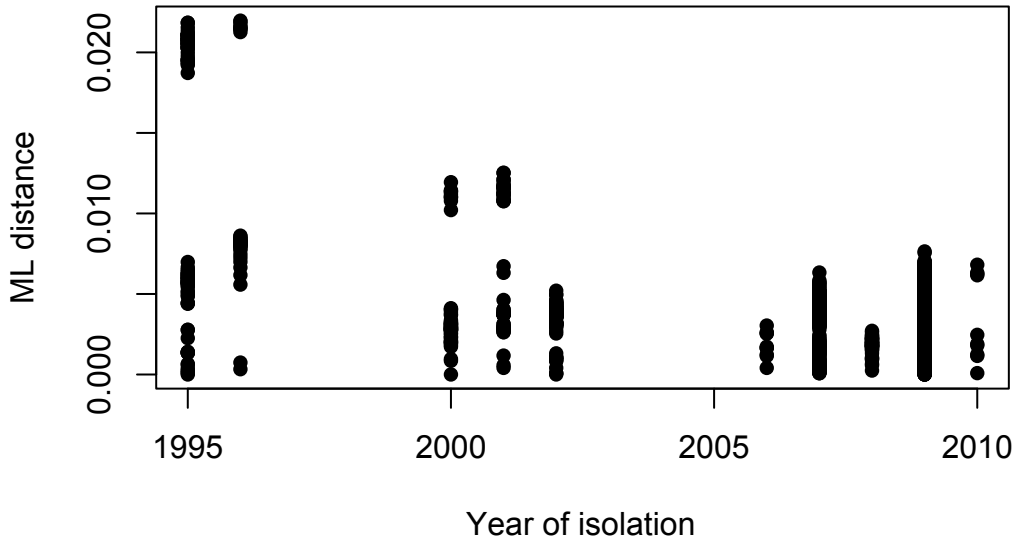
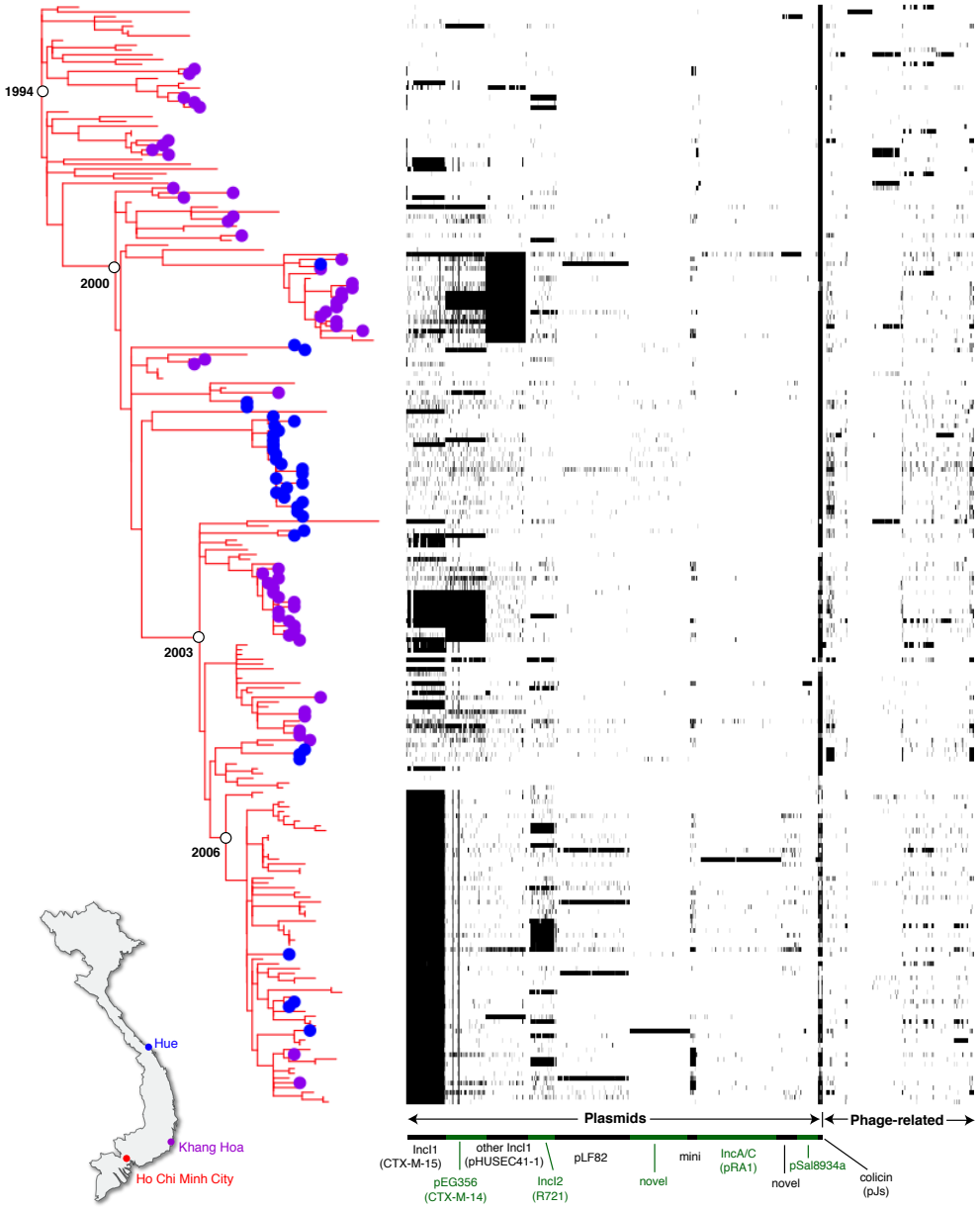


Figure S4. Genetic diversity within the *Shigella sonnei* VN clone

Declining genetic diversity within the HCMC *S. sonnei* population during the 16-year period of study. Pairwise genetic distances are based on the maximum likelihood phylogenetic tree shown in Figure 1, year indicates year of isolation.

a


1994
2000
2003
2006
Hue
Khang Hoa
Ho Chi Minh City
Plasmids
Phage-related
IncI1 (CTX-M-15)
other IncI1 (pHUSEC41-1)
pLF82
mini
IncA/C (pRA1)
pSal8934a
colicin (pJs)
pEG356 (CTX-M-14)
IncI2 (R721)
novel
novel

b

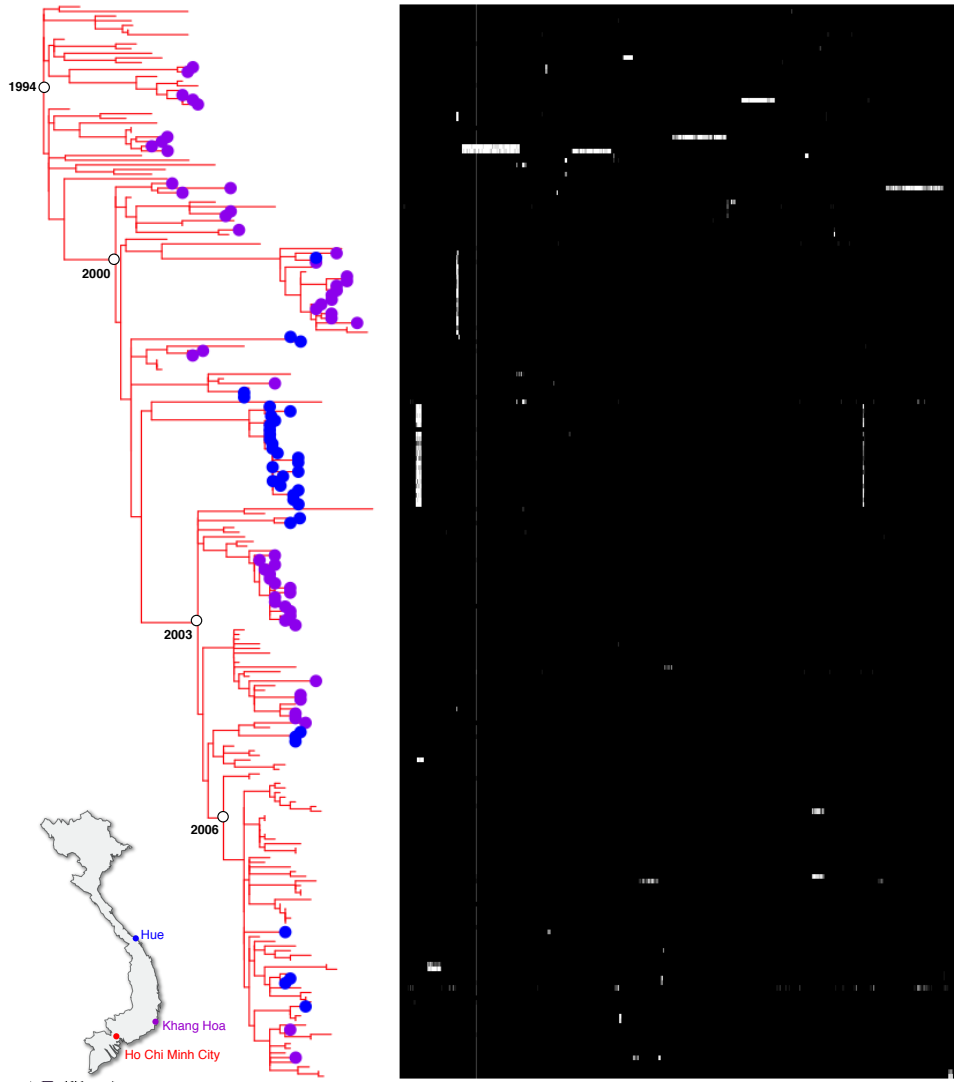


Figure S5. Gene content variation within the *Shigella sonnei* VN clone

Colours indicate the city of isolation, according to inset map; major sweeps are marked. **(a)** Novel genes acquired by one or more isolates. Nearest reference sequences are given for plasmids. **(b)** Conserved chromosomal genes subject to deletion in one or more isolates.

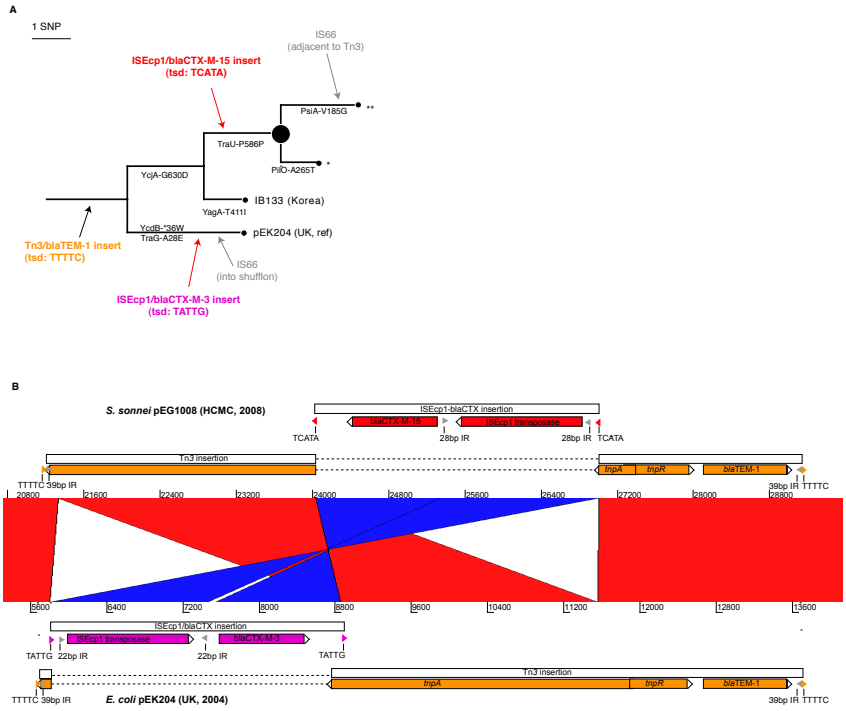


Figure S6. Novel ESBL plasmid identified in HCMC *Shigella sonnei*

(a) SNP-based phylogeny for the IncI1 plasmid backbone, for pEK204 (NC_013120) isolated from uropathogenic *E. coli* ST131 in the UK in 2004, *S. sonnei* IB133 isolated in Korea in 2003 [6] and variants of pKHSB1 from *S. sonnei* isolated in HCMC in 2007-2010 (Table S1). For branches defined by SNPs within coding regions, details of the SNP and its coding effects are provided. Insertions within the plasmid backbone are labelled, colours match those in (b); tsd=target site duplication. **(b)** Genes and repeats annotated around the CTX-M insertion sites from pKHSB1 (top) and pEK204 (bottom). These are joined by homology blocks visualised in ACT, where red = same orientation and blue = inverse orientation. Position and orientation of 5 bp direct repeats (target site duplications) are shown with coloured arrows, labelled with the repeat sequence; grey arrows show inverted repeats (IR) labelled with their size in bp.

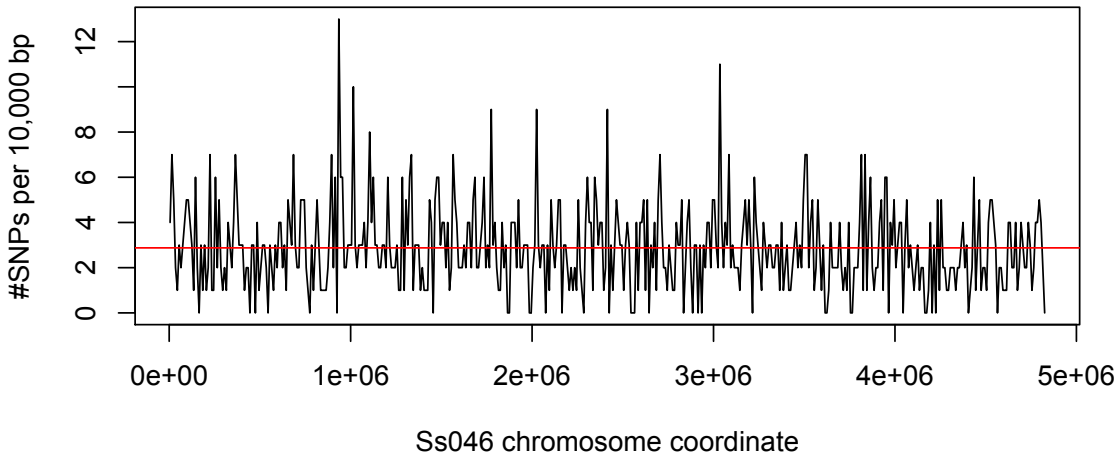


Figure S7. Genomic SNP distribution within the *Shigella sonnei* VN clone

Spatial distribution of SNPs arising in the VN clone, relative to the Ss046 chromosome reference sequence. The mean SNP rate, 3/10,000 bp (0.03%) is shown in red.

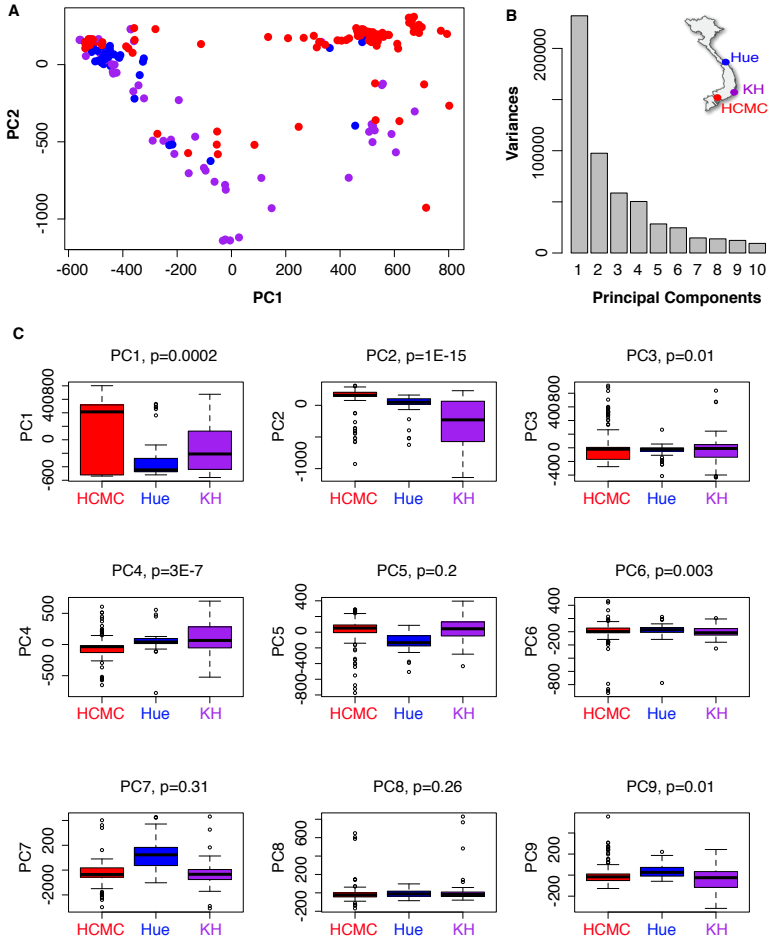


Figure S8. The accessory gene content within the *Shigella sonnei* VN clone

Principal components analysis of accessory gene content. **(a)** Clustering of isolates by first two accessory gene principal components, coloured by location as indicated in the map. **(b)** Scree plot showing the variance in gene content explained by each principal component. **(c)** Distribution of accessory gene principal components by city of isolation; p-values indicate strength of evidence for the association between location and the distribution of principal component values, using the Kolmogorov-Smirnov test

Table S1. *Shigella sonnei* isolates sequenced in this study

Strain	Source	GPS study	Location	Year	Depth	Lineage	Clone	pDPT1 (E5)	pDPT2 (Js)	gyrA Nal SNP	Nal	Cef	CTX-M	pKHSB1	pEG356	Other plasmids
HUE05	OUCRU	no	Hue	2008	82	III	Hue	1	0	-	S	S	-	no	no	
HUE02	OUCRU	no	Hue	2008	51.6	III	Hue	1	0	-	S	S	-	no	no	
HUE01	OUCRU	no	Hue	2008	57.9	III	Hue	1	0	-	S	S	-	no	no	
HUE48	OUCRU	no	Hue	2009	222	III	Hue	1	1	-	S	S	-	no	no	
HUE47	OUCRU	no	Hue	2009	158.1	III	Hue	1	1	-	S	S	-	no	no	
HUE42	OUCRU	no	Hue	2009	131.2	III	Hue	1	0	-	S	S	-	no	no	
HUE40	OUCRU	no	Hue	2009	233.8	III	Hue	1	0	-	S	S	-	no	no	
HUE34	OUCRU	no	Hue	2009	142.2	III	Hue	1	0	-	S	R	14	no	no	
HUE32	OUCRU	no	Hue	2009	197.5	III	Hue	1	0	-	S	S	-	no	no	
HUE31	OUCRU	no	Hue	2009	111.6	III	Hue	1	0	-	S	S	-	no	no	
HUE27	OUCRU	no	Hue	2009	94	III	Hue	1	1	-	S	S	-	no	no	
HUE26	OUCRU	no	Hue	2009	31.3	III	Hue	1	1	-	S	S	-	no	no	
HUE24	OUCRU	no	Hue	2009	81.5	III	Hue	1	1	-	S	S	-	no	no	
HUE23	OUCRU	no	Hue	2009	83	III	Hue	1	1	-	S	S	-	no	no	
HUE21	OUCRU	no	Hue	2009	223.8	III	Hue	1	1	-	S	S	-	no	no	
HUE16	OUCRU	no	Hue	2009	90.9	III	Hue	1	1	-	S	S	-	no	no	
HUE68	OUCRU	no	Hue	2010	105.9	III	Hue	1	1	-	S	S	-	no	no	
HUE67	OUCRU	no	Hue	2010	240.3	III	Hue	1	1	-	S	S	-	no	no	
HUE64	OUCRU	no	Hue	2010	158.9	III	Hue	1	1	-	S	S	-	no	no	
HUE62	OUCRU	no	Hue	2010	57	III	Hue	1	1	-	S	S	-	no	no	
HUE60	OUCRU	no	Hue	2010	129.8	III	Hue	1	1	-	S	S	-	no	no	
HUE57	OUCRU	no	Hue	2010	78.8	III	Hue	1	1	-	S	S	-	no	no	
EG0404	OUCRU	no	HCMC	2007	17	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
HUE58	OUCRU	no	Hue	2010	84.4	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1, LF82
KH41	OUCRU	no	Khanh Hoa	2009	98.6	III	KH1	1	0	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH37	OUCRU	no	Khanh Hoa	2009	100.3	III	KH1	1	0	gyrA-83L	R	S	-	no	no	pHUSEC41-1, R271
KH35	OUCRU	no	Khanh Hoa	2009	175.9	III	KH1	1	0	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH27	OUCRU	no	Khanh Hoa	2009	136.4	III	KH1	1	1	gyrA-83L	R	R	14	no	yes	pHUSEC41-1
KH26	OUCRU	no	Khanh Hoa	2009	115	III	KH1	1	1	gyrA-83L	R	R	14	no	yes	pHUSEC41-1
KH24	OUCRU	no	Khanh Hoa	2009	110.3	III	KH1	1	1	gyrA-83L	R	R	14	no	yes	pHUSEC41-1
KH23	OUCRU	no	Khanh Hoa	2009	140.6	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH20	OUCRU	no	Khanh Hoa	2009	174.1	III	KH1	1	1	gyrA-83L	R	R	14	no	yes	pHUSEC41-1
KH18	OUCRU	no	Khanh Hoa	2009	146	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH15	OUCRU	no	Khanh Hoa	2009	159	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1

KH11	OUCRU	no	Khanh Hoa	2009	239.4	III	KH1	1	0.627	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH09	OUCRU	no	Khanh Hoa	2009	177.2	III	KH1	1	0.536	gyrA-83L	R	S	-	no	no	pHUSEC41-1
KH53	OUCRU	no	Khanh Hoa	2010	134.6	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
EG0467	OUCRU	no	HCMC	2008	15	III	KH2	1	1	gyrA-87Y	R	S	-	no	no	
KH40	OUCRU	no	Khanh Hoa	2009	138.7	III	KH2	1	0	gyrA-87Y	R	S	-	no	no	
KH34	OUCRU	no	Khanh Hoa	2009	127.8	III	KH2	1	0	gyrA-87Y	R	R	14	no	yes	
KH33	OUCRU	no	Khanh Hoa	2009	179.6	III	KH2	1	0	gyrA-87Y	R	S	-	no	no	
KH32	OUCRU	no	Khanh Hoa	2009	175.1	III	KH2	1	0	gyrA-87Y	R	R	14	no	yes	
KH30	OUCRU	no	Khanh Hoa	2009	123.6	III	KH2	1	0.827	gyrA-87Y	R	R	14	no	yes	
KH25	OUCRU	no	Khanh Hoa	2009	126.8	III	KH2	1	0.961	gyrA-87Y	R	R	14	no	yes	
KH21	OUCRU	no	Khanh Hoa	2009	122.2	III	KH2	1	0.62	gyrA-87Y	R	S	-	no	no	
KH19	OUCRU	no	Khanh Hoa	2009	111.1	III	KH2	1	0.571	gyrA-87Y	R	R	14	no	yes	R271
KH17	OUCRU	no	Khanh Hoa	2009	190.2	III	KH2	1	0.6	gyrA-87Y	R	S	-	no	no	
KH14	OUCRU	no	Khanh Hoa	2009	158.4	III	KH2	1	0.998	gyrA-87Y	R	R	14	no	yes	
KH12	OUCRU	no	Khanh Hoa	2009	189.1	III	KH2	1	0.581	gyrA-87Y	R	R	14	no	yes	
KH07	OUCRU	no	Khanh Hoa	2009	118.8	III	KH2	1	0.407	gyrA-87Y	R	R	14	no	yes	
KH06	OUCRU	no	Khanh Hoa	2009	105.5	III	KH2	1	0.988	gyrA-87Y	R	R	14	no	yes	
KH05	OUCRU	no	Khanh Hoa	2009	89.9	III	KH2	1	0.986	gyrA-87Y	R	S	-	no	no	
KH04	OUCRU	no	Khanh Hoa	2009	74.9	III	KH2	1	1	gyrA-87Y	R	R	14	no	yes	
KH43	OUCRU	no	Khanh Hoa	2010	140.2	III	KH2	1	1	gyrA-87Y	R	R	14	no	yes	
DE0330	OUCRU	no	HCMC	2000	16	I	LinI	0.86	0.05	-	S	S	-	no	no	
EG0352	OUCRU	no	HCMC	2007	11	I	LinI	0.9	0.21	-	S	S	-	no	no	
MS0094	OUCRU	no	HCMC	1996	158.6	III	nonVN	0.067	0.678	-	S	S	-	no	no	
MS0122	OUCRU	no	HCMC	1996	28	III	nonVN	0.02	0.14	-	S	S	-	no	no	
MS0128	OUCRU	no	HCMC	1996	19	III	sweep1	1	1	-	S	S	-	no	no	
MS0110	OUCRU	no	HCMC	1996	24	III	sweep1	1	1	-	S	S	-	no	no	
DE0489	OUCRU	no	HCMC	2000	29	III	sweep1	1	1	-	S	S	-	no	no	
DE0477	OUCRU	no	HCMC	2000	8	III	sweep1	1	1	-	S	S	-	no	no	
DE0427	OUCRU	no	HCMC	2000	26	III	sweep1	1	1	-	S	S	-	no	no	
DE0306	OUCRU	no	HCMC	2000	20	III	sweep1	1	1	-	S	S	-	no	no	
DE0303	OUCRU	no	HCMC	2000	23	III	sweep1	1	1	-	S	S	-	no	no	
DE0295	OUCRU	no	HCMC	2000	18	III	sweep1	1	1	-	S	S	-	no	no	
DE0248	OUCRU	no	HCMC	2000	15	III	sweep1	1	1	-	S	S	-	no	no	novel plasmid
DE0199	OUCRU	no	HCMC	2000	11	III	sweep1	1	1	-	S	S	-	no	no	
DE0127	OUCRU	no	HCMC	2000	21	III	sweep1	1	1	-	S	S	-	no	no	
DE0115	OUCRU	no	HCMC	2000	16	III	sweep1	1	1	-	S	S	-	no	no	
DE0965	OUCRU	no	HCMC	2001	25	III	sweep1	1	1	-	S	S	-	no	no	
DE0900	OUCRU	no	HCMC	2001	34	III	sweep1	1	1	-	S	S	-	no	no	
DE0846	OUCRU	no	HCMC	2001	32	III	sweep1	1	1	-	S	S	-	no	no	
DE0816	OUCRU	no	HCMC	2001	17	III	sweep1	1	1	-	S	S	-	no	no	
DE0685	OUCRU	no	HCMC	2001	36	III	sweep1	1	1	-	S	S	-	no	no	

DE0654	OUCRU	no	HCMC	2001	23	III	sweep1	1	1	-	S	S	-	no	no	
DE0611	OUCRU	no	HCMC	2001	26	III	sweep1	1	1	-	S	R	14	no	no	
DE0579	OUCRU	no	HCMC	2001	27	III	sweep1	1	1	gyrA-83L	R	S	-	no	no	
DE1486	OUCRU	no	HCMC	2002	44	III	sweep1	1	1	-	S	S	-	no	no	
DE1336	OUCRU	no	HCMC	2002	18	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	pHUSEC41-1
DE1318	OUCRU	no	HCMC	2002	17	III	sweep1	1	1	-	S	S	-	no	no	
DE1256	OUCRU	no	HCMC	2002	11	III	sweep1	1	1	-	S	S	-	no	no	
DE1209	OUCRU	no	HCMC	2002	18	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	
DE1198	OUCRU	no	HCMC	2002	30	III	sweep1	1	1	-	S	S	-	no	no	
DE1165	OUCRU	no	HCMC	2002	11	III	sweep1	1	1	-	S	S	-	no	no	
DE1150	OUCRU	no	HCMC	2002	19	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	
DE1140	OUCRU	no	HCMC	2002	15	III	sweep1	1	1	-	S	S	-	no	no	
IB1970	IVI	no	Khanh Hoa	2001	21	III	sweep1	1	1	-	S	S	-	no	no	
IB2024	IVI	no	Khanh Hoa	2002	52	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	R271
IB2015	IVI	no	Khanh Hoa	2002	18	III	sweep1	1	1	-	S	S	-	no	no	
IB2026	IVI	no	Khanh Hoa	2003	33	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	
IB2009	IVI	no	Khanh Hoa	2003	11	III	sweep1	1	1	-	S	S	-	no	no	
IB2008	IVI	no	Khanh Hoa	2003	25	III	sweep1	1	1	-	S	S	-	no	no	mini
IB2004	IVI	no	Khanh Hoa	2003	33	III	sweep1	1	1	-	S	S	-	no	no	
IB2000	IVI	no	Khanh Hoa	2003	29	III	sweep1	1	1	-	S	S	-	no	no	mini
IB1995	IVI	no	Khanh Hoa	2003	13	III	sweep1	1	1	gyrA-87G	R	S	-	no	no	R271
DE1404	OUCRU	no	HCMC	2002	15	III	sweep2	1	1	-	S	S	-	no	no	
DE1208	OUCRU	no	HCMC	2002	18	III	sweep2	1	1	-	S	S	-	no	no	
DE1191	OUCRU	no	HCMC	2002	16	III	sweep2	1	1	-	S	S	-	no	no	
DE1063	OUCRU	no	HCMC	2002	21	III	sweep2	1	1	-	S	S	-	no	no	
EG0304	OUCRU	no	HCMC	2006	18	III	sweep2	1	1	-	S	S	-	no	no	
EG0410	OUCRU	no	HCMC	2007	10	III	sweep2	1	1	-	S	S	-	no	no	
EG0392	OUCRU	no	HCMC	2007	8	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
EG0386	OUCRU	no	HCMC	2007	15	III	sweep2	1	1	-	S	S	-	no	no	
EG0385	OUCRU	no	HCMC	2007	14	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	R271
EG0379	OUCRU	no	HCMC	2007	14	III	sweep2	1	1	-	S	S	-	no	no	
EG0373	OUCRU	no	HCMC	2007	37	III	sweep2	1	1	-	S	R	15	yes	no	
EG0372	OUCRU	no	HCMC	2007	15	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
EG0159	OUCRU	no	HCMC	2007	7	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
HUE43	OUCRU	no	Hue	2009	106.8	III	sweep2	1	0	-	S	S	-	no	no	
HUE33	OUCRU	no	Hue	2009	102.5	III	sweep2	1	0	-	S	S	-	no	no	
HUE30	OUCRU	no	Hue	2009	182	III	sweep2	1	0	-	S	S	-	no	no	
HUE25	OUCRU	no	Hue	2009	190.3	III	sweep2	1	1	-	S	S	-	no	no	
IB2018	IVI	no	Khanh Hoa	2002	58	III	sweep2	1	1	-	S	S	-	no	no	
IB1987	IVI	no	Khanh Hoa	2002	43	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
IB1997	IVI	no	Khanh Hoa	2003	37	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	

IB1993	IVI	no	Khanh Hoa	2003	34	III	sweep2	1	1	-	S	S	-	no	no	
IB1990	IVI	no	Khanh Hoa	2003	17	III	sweep2	1	1	-	S	S	-	no	no	
KH02	OUCRU	no	Khanh Hoa	2009	186.5	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
KH55	OUCRU	no	Khanh Hoa	2010	160	III	sweep2	1	0	gyrA-83L	R	S	-	no	no	
KH54	OUCRU	no	Khanh Hoa	2010	90.1	III	sweep2	1	1	gyrA-83L	R	S	-	no	no	
KH42	OUCRU	no	Khanh Hoa	2010	146.3	III	sweep2	1	1	-	S	R	14	no	no	
EG0318	OUCRU	no	HCMC	2006	19	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0315	OUCRU	no	HCMC	2006	37	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0313	OUCRU	no	HCMC	2006	8	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0309	OUCRU	no	HCMC	2006	23	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0401	OUCRU	no	HCMC	2007	13	III	sweep3	1	0.18	gyrA-87Y	R	S	-	no	no	
EG0395	OUCRU	no	HCMC	2007	26	III	sweep3	1	1	gyrA-87Y	R	R	15	yes	no	
EG0394	OUCRU	no	HCMC	2007	7	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	R271
EG0393	OUCRU	no	HCMC	2007	17	III	sweep3	1	0.5	gyrA-87Y	R	S	-	no	no	
EG0390	OUCRU	no	HCMC	2007	34	III	sweep3	1	1	gyrA-87Y	R	R	15	yes	no	
EG0388	OUCRU	no	HCMC	2007	19	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0383	OUCRU	no	HCMC	2007	10	III	sweep3	1	0.2	gyrA-87Y	R	S	-	no	no	
EG0375	OUCRU	no	HCMC	2007	19	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0369	OUCRU	no	HCMC	2007	22	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0365	OUCRU	no	HCMC	2007	13	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0362	OUCRU	no	HCMC	2007	21	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0357	OUCRU	no	HCMC	2007	28	III	sweep3	1	0.8	gyrA-87Y	R	S	-	no	no	
EG0129	OUCRU	no	HCMC	2007	9	III	sweep3	1	0.16	gyrA-87Y	R	S	-	no	no	
EG1001	OUCRU	no	HCMC	2008	14	III	sweep3	1	0.31	gyrA-87Y	R	R	15	yes	no	
EG0451	OUCRU	no	HCMC	2008	17	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0430	OUCRU	no	HCMC	2008	28	III	sweep3	1	0.78	gyrA-87Y	R	R	15	yes	no	R271, pSal8934a
EG0425	OUCRU	no	HCMC	2008	14	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
EG0255	OUCRU	no	HCMC	2008	14	III	sweep3	1	0.23	gyrA-87Y	R	S	-	no	no	
EG1017	OUCRU	no	HCMC	2009	111.1	III	sweep3	1	0	gyrA-87Y	R	R	15	yes	no	
EG1014	OUCRU	no	HCMC	2009	588	III	sweep3	1	0	gyrA-87Y	R	R	15	yes	no	
EG1015	OUCRU	no	HCMC	2009	1165.6	III	sweep3	0	0	gyrA-87Y	R	R	15	yes	no	
HUE50	OUCRU	no	Hue	2009	139	III	sweep3	1	1	gyrA-87Y	R	R	14	no	yes	
HUE46	OUCRU	no	Hue	2009	175.4	III	sweep3	1	0	gyrA-87Y	R	S	-	no	no	
HUE29	OUCRU	no	Hue	2009	190	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
HUE55	OUCRU	no	Hue	2010	171.6	III	sweep3	1	0.999	gyrA-87Y	R	S	-	no	no	
HUE53	OUCRU	no	Hue	2010	95.4	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
KH29	OUCRU	no	Khanh Hoa	2009	144.7	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
KH28	OUCRU	no	Khanh Hoa	2009	191.2	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
KH16	OUCRU	no	Khanh Hoa	2009	175.8	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
KH13	OUCRU	no	Khanh Hoa	2009	240.1	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	
KH10	OUCRU	no	Khanh Hoa	2009	110	III	sweep3	1	0.566	gyrA-87Y	R	S	-	no	no	

KH57	OUCRU	no	Khanh Hoa	2010	162.1	III	sweep3	1	0	gyrA-87Y	R	S	-	no	no	
EG1008	OUCRU	no	HCMC	2008	41	III	sweep4	1	0.41	gyrA-87Y	R	R	15	yes	no	
EG1004	OUCRU	no	HCMC	2008	23	III	sweep4	1	0.54	gyrA-87Y	R	R	15	yes	no	
EG0472	OUCRU	no	HCMC	2008	15	III	sweep4	1	1	gyrA-87Y	R	R	15	yes	no	
EG0204	OUCRU	no	HCMC	2008	15	III	sweep4	1	0.2	gyrA-87Y	R	R	15	yes	no	
10102	OUCRU	no	HCMC	2009	106.4	III	sweep4	1	0.408	gyrA-87Y	R	R	15	yes	no	
EG1029	OUCRU	no	HCMC	2009	146.1	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1028	OUCRU	no	HCMC	2009	185.7	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1027	OUCRU	no	HCMC	2009	170	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
EG1026	OUCRU	no	HCMC	2009	150.6	III	sweep4	1	0.915	gyrA-87Y	R	R	15	yes	no	
EG1025	OUCRU	no	HCMC	2009	224.1	III	sweep4	1	0.6	gyrA-87Y	R	R	15	yes	no	
EG1024	OUCRU	no	HCMC	2009	109.6	III	sweep4	1	0.444	gyrA-87Y	R	R	15	yes	no	
EG1023	OUCRU	no	HCMC	2009	29.5	III	sweep4	1	0.35	gyrA-87Y	R	R	15	yes	no	
EG1022	OUCRU	no	HCMC	2009	32	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1021	OUCRU	no	HCMC	2009	75.9	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1020	OUCRU	no	HCMC	2009	75.1	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
EG1019	OUCRU	no	HCMC	2009	31.9	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1018	OUCRU	no	HCMC	2009	45.3	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
EG1016	OUCRU	no	HCMC	2009	87.6	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
HUE22	OUCRU	no	Hue	2009	142	III	sweep4	1	1	gyrA-87Y	R	R	15	yes	no	
HUE20	OUCRU	no	Hue	2009	91.6	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
HUE19	OUCRU	no	Hue	2009	294	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
HUE17	OUCRU	no	Hue	2009	155.8	III	sweep4	1	1	gyrA-87Y	R	R	15	yes	no	
KH38	OUCRU	no	Khanh Hoa	2009	229	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	mini
KH45	OUCRU	no	Khanh Hoa	2010	65.2	III	sweep4	1	1	gyrA-87Y	R	R	15	yes	no	mini
MS0004	OUCRU	no	HCMC	1995	22	III	VN	0.85	0.17	-	S	S	-	no	no	
MS0039	OUCRU	no	HCMC	1995	19	III	VN	0.84	0.16	-	S	S	-	no	no	
MS0034	OUCRU	no	HCMC	1995	14	III	VN	0.84	0.14	-	S	S	-	no	no	
MS0032	OUCRU	no	HCMC	1995	15	III	VN	0.83	0.15	-	S	S	-	no	no	
MS0043	OUCRU	no	HCMC	1995	134.8	III	VN	0.077	0	-	S	S	-	no	no	
MS0063	OUCRU	no	HCMC	1995	18	III	VN	0.05	0.18	-	S	S	-	no	no	
MS0069	OUCRU	no	HCMC	1995	26	III	VN	0.02	0.11	-	S	S	-	no	no	
MS0065	OUCRU	no	HCMC	1995	26	III	VN	0.02	0.12	-	S	S	-	no	no	
MS0048	OUCRU	no	HCMC	1995	17	III	VN	0.02	0.13	-	S	S	-	no	no	
MS0035	OUCRU	no	HCMC	1995	21	III	VN	0.02	0.11	-	S	S	-	no	no	
MS0011	OUCRU	no	HCMC	1995	13	III	VN	0.02	0.15	-	S	S	-	no	no	
MS0080	OUCRU	no	HCMC	1995	13	III	VN	0.01	0.11	-	S	S	-	no	no	
MS0070	OUCRU	no	HCMC	1995	18	III	VN	0.01	0.12	-	S	S	-	no	no	
MS0042	OUCRU	no	HCMC	1995	17	III	VN	0.01	0.12	-	S	S	-	no	no	
MS0083	OUCRU	no	HCMC	1996	25	III	VN	0.88	0.17	-	S	S	-	no	no	
MS0102	OUCRU	no	HCMC	1996	23	III	VN	0.83	0.16	-	S	S	-	no	no	

MS0119	OUCRU	no	HCMC	1996	28	III	VN	0.04	0.15	-	S	S	-	no	no	
MS0111	OUCRU	no	HCMC	1996	20	III	VN	0.02	0.13	-	S	S	-	no	no	
MS0127	OUCRU	no	HCMC	1996	17	III	VN	0.01	0.16	-	S	S	-	no	no	
DE0490	OUCRU	no	HCMC	2000	20	III	VN	0.05	0.19	gyrA-83L	R	S	-	no	no	
DE0655	OUCRU	no	HCMC	2001	15	III	VN	0.07	0.16	-	S	S	-	no	no	
DE0891	OUCRU	no	HCMC	2001	26	III	VN	0.02	0.13	gyrA-83L	R	S	-	no	no	
DE0885	OUCRU	no	HCMC	2001	35	III	VN	0.02	0.12	gyrA-83L	R	S	-	no	no	
IB2013	IVI	no	Khanh Hoa	2001	51	III	VN	0.08	0.31	-	S	S	-	no	no	
IB2012	IVI	no	Khanh Hoa	2001	17	III	VN	0.05	0.32	gyrA-87Y	R	S	-	no	no	none
IB1985	IVI	no	Khanh Hoa	2002	20	III	VN	0.23	0.4	-	S	S	-	no	no	
IB1976	IVI	no	Khanh Hoa	2002	11	III	VN	0.04	0.17	gyrA-87G	R	S	-	no	no	
IB1980	IVI	no	Khanh Hoa	2002	48	III	VN	0.02	0.16	-	S	S	-	no	no	
30450	OUCRU	yes	HCMC	2010	188.3	III	KH1	1	0	gyrA-83L	R	S	-	no	no	pHUSEC41-1
30010	OUCRU	yes	HCMC	2009	111.6	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
20261	OUCRU	yes	HCMC	2009	252.4	III	KH1	1	1	gyrA-83L	R	S	-	no	no	pHUSEC41-1
10073	OUCRU	yes	HCMC	2009	126.8	III	KH1	1	0	gyrA-83L	R	S	-	no	no	
30366	OUCRU	yes	HCMC	2009	147.4	III	KH2	1	0	gyrA-87Y	R	S	-	no	no	
10021	OUCRU	yes	HCMC	2009	13.1	III	KH2	1	0	gyrA-87Y	R	S	-	no	no	
30293	OUCRU	yes	HCMC	2009	156.3	III	sweep2	1	0	gyrA-83L	R	R	14,15	yes	yes	
10071	OUCRU	yes	HCMC	2009	35.8	III	sweep2	1	0	-	S	S	-	no	no	
10031	OUCRU	yes	HCMC	2009	12.8	III	sweep2	1	0	-	S	S	-	no	no	
30451	OUCRU	yes	HCMC	2010	260.7	III	sweep3	1	0	gyrA-87Y	R	S	-	no	no	
30073	OUCRU	yes	HCMC	2009	252.8	III	sweep3	1	1	gyrA-87Y	R	S	-	no	no	pSal8934a
30003	OUCRU	yes	HCMC	2009	328.5	III	sweep3	1	0	gyrA-87Y	R	S	-	no	no	
20094	OUCRU	yes	HCMC	2009	100.8	III	sweep3	1	1	gyrA-87Y	R	R	14	no	yes	
30387	OUCRU	yes	HCMC	2010	114.2	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
30371	OUCRU	yes	HCMC	2009	116.6	III	sweep4	1	0.98	gyrA-87Y	R	R	15	yes	no	
30233	OUCRU	yes	HCMC	2009	182.8	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271, mini
30174	OUCRU	yes	HCMC	2009	105.6	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
30172	OUCRU	yes	HCMC	2009	201.7	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
30169	OUCRU	yes	HCMC	2009	153.5	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
30164	OUCRU	yes	HCMC	2009	112.6	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
30162	OUCRU	yes	HCMC	2009	139.5	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	R271
30124	OUCRU	yes	HCMC	2009	31.8	III	sweep4	1	0.25	gyrA-87Y	R	R	15	yes	no	
30112	OUCRU	yes	HCMC	2009	38.3	III	sweep4	1	1	gyrA-87Y	R	R	15	yes	no	R271, mini
30100	OUCRU	yes	HCMC	2009	224	III	sweep4	1	0.92	gyrA-87Y	R	R	15	yes	no	
30071	OUCRU	yes	HCMC	2009	133.1	III	sweep4	1	0.901	gyrA-87Y	R	R	15	yes	no	
30059	OUCRU	yes	HCMC	2009	152.9	III	sweep4	1	0.765	gyrA-87Y	R	R	15	yes	no	R271
30054	OUCRU	yes	HCMC	2009	88.9	III	sweep4	1	0.584	gyrA-87Y	R	R	15	yes	no	R271
30037	OUCRU	yes	HCMC	2009	155.4	III	sweep4	1	0.843	gyrA-87Y	R	R	15	yes	no	R271
30008	OUCRU	yes	HCMC	2009	42.4	III	sweep4	1	0.58	gyrA-87Y	R	R	15	yes	no	R271

20343	OUCRU	yes	HCMC	2009	106	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	pRA1, novel
20263	OUCRU	yes	HCMC	2009	127.1	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	pHUSEC41-1
20228	OUCRU	yes	HCMC	2009	153.8	III	sweep4	1	0.945	gyrA-87Y	R	R	15	yes	no	R271, mini
20070	OUCRU	yes	HCMC	2009	101.6	III	sweep4	1	0.715	gyrA-87Y	R	R	15	yes	no	LF82
20037	OUCRU	yes	HCMC	2009	39.9	III	sweep4	1	0.524	gyrA-87Y	R	R	15	yes	no	
20023	OUCRU	yes	HCMC	2009	137.7	III	sweep4	1	0.841	gyrA-87Y	R	R	15	yes	no	min
20021	OUCRU	yes	HCMC	2009	145.6	III	sweep4	1	0.776	gyrA-87Y	R	R	15	yes	no	
20006	OUCRU	yes	HCMC	2009	152.8	III	sweep4	1	0.991	gyrA-87Y	R	R	15	yes	no	
10365	OUCRU	yes	HCMC	2010	59	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10320	OUCRU	yes	HCMC	2010	104	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10263	OUCRU	yes	HCMC	2009	125.2	III	sweep4	1	0.374	gyrA-87Y	R	R	15	yes	no	
10188	OUCRU	yes	HCMC	2009	26.2	III	sweep4	1	0.801	gyrA-87Y	R	R	15	yes	no	
10159	OUCRU	yes	HCMC	2009	114.1	III	sweep4	1	0.369	gyrA-87Y	R	R	15	yes	no	
10152	OUCRU	yes	HCMC	2009	22.7	III	sweep4	1	0.222	gyrA-87Y	R	R	15	yes	no	
10135	OUCRU	yes	HCMC	2009	23.1	III	sweep4	1	0.158	gyrA-87Y	R	R	15	yes	no	
10134	OUCRU	yes	HCMC	2009	14.6	III	sweep4	1	0.151	gyrA-87Y	R	R	15	yes	no	
10115	OUCRU	yes	HCMC	2009	9.4	III	sweep4	1	0.098	gyrA-87Y	R	R	15	yes	no	
10111	OUCRU	yes	HCMC	2009	12.9	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10093	OUCRU	yes	HCMC	2009	66.3	III	sweep4	1	0.159	gyrA-87Y	R	R	15	yes	no	
10083	OUCRU	yes	HCMC	2009	47.6	III	sweep4	1	0.245	gyrA-87Y	R	R	15	yes	no	
10063	OUCRU	yes	HCMC	2009	30.4	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10060	OUCRU	yes	HCMC	2009	16.2	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10035	OUCRU	yes	HCMC	2009	33.7	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	
10014	OUCRU	yes	HCMC	2009	13.1	III	sweep4	1	0	gyrA-87Y	R	R	15	yes	no	

Table S2. Genera contributing >1% of accessory genes annotated by MG-RAST within the pan-genome of the *S. sonnei* VN clone.

Genus	Accessory genes
Escherichia	46.8%
Salmonella	9.2%
Shigella	7.4%
Yersinia	6.7%
Klebsiella	2.3%
Citrobacter	2.2%
Aeromonas	2.0%
Vibrio	2.0%
Photobacterium	1.3%
Serratia	1.2%
Enterobacter	1.1%
Pseudomonas	1.0%
Other	16.6%

Table S3. Genes displaying evidence of convergent evolution

Convergent evolution defined as independent mutations in phylogenetically distinct backgrounds resulting in changes to identical or neighbouring amino acids or diversifying selection (a higher rate of non-synonymous mutations over synonymous mutations, dN/dS>1.5).

Gene ID	Nonsynonymous SNPs	Synonymous SNPs	dN/dS	Homoplasic SNPs	Gene symbol	Product	Function	Note
SSON_2289	3	0	Inf	1	gyrA	DNA gyrase subunit A	antimicrobial resistance	resistance to quinolone-based antimicrobials
SSON_3392	4	0	Inf	0	mreB	rod shape-determining protein	antimicrobial resistance	mecillinam resistance; cell envelope
SSON_2743	1	0	Inf	1	smpB	SsrA-binding protein	antimicrobial resistance	aminoglycoside resistance
SSON_0691	4	0	Inf	0	tolA	cell envelope integrity inner membrane protein	colicin sensitivity	colicin interation; cell envelope
SSON_0836	5	1	1.54	0	nfsA	nitroreductase A	metal resistance	chromium tolerance
SSON_0131	3	0	Inf	1	yacK	multicopper oxidase	metal resistance	copper detoxification
SSON_3707	1	0	Inf	1	zntA	zinc/cadmium/mercury/lead-transporting ATPase	metal resistance	-
SSON_2085	1	0	Inf	1	yeeY	putative transcriptional regulator LYSR-type	regulator	transcriptional regulator, unknown function
SSON_2224	1	0	Inf	1	fruK	1-phosphofructokinase	starvation response	carbon metabolism
SSON_0087	1	0	Inf	1	fruR	DNA-binding transcriptional regulator	starvation response	regulation of carbon usage pathways
SSON_0935	4	0	Inf	0	pepN	aminopeptidase N	starvation response	response to phosphate starvation
SSON_4271	2	0	Inf	1	phnM	phosphonate metabolism	starvation response	-
SSON_0891	1	1	0.31	1	ftsK	DNA translocase	stress response	stress response
SSON_0979	4	1	1.23	0	hyaB	hydrogenase 1 large subunit	stress response	acid tolerance
SSON_0646	4	2	0.62	0	kdpD	sensor protein	stress response	response to osmotic shock
SSON_0857	3	0	Inf	1	poxB	pyruvate dehydrogenase	stress response	rpoS regulated; response to chemicals
SSON_2889	5	2	0.77	0	rpoS	RNA polymerase sigma factor	stress response	master regulator of the general stress response
SSON_3060	5	0	Inf	0	ubiH	2-octaprenyl-6-methoxyphenyl hydroxylase	stress response	response to oxidative stress
SSON_0788	1	0	Inf	1	ybiO	hypothetical protein	stress response	rpoS regulated transporter
SSON_2853	2	1	0.617	1	ygaA	anaerobic nitric oxide reductase transcription regulator	stress response	response to reactive nitrogen species
SSON_1122	4	1	1.23	0	fhuE	ferric-rhodotorulic acid outer	transporter	iron acquisition

						membrane transporter		
SSON_2575	1	0	Inf	1	perM	putative permease	transporter	inner membrane transport
SSON_3332	3	0	Inf	1	yhbE	hypothetical protein	transporter	-
SSON_0773	3	0	Inf	1	ybhF	putative ATP-binding component of a transport system	transporter	-
SSON_3441	4	0	Inf	0	secY	preprotein translocase subunit	type II secretion	type II secretion
SSON_4063	4	0	Inf	0	fdoG	formate dehydrogenase-O, major subunit	anaerobic respiration	anaerobic respiration
SSON_2567	1	1	0.31	1	hyfF	hydrogenase 4 subunit F	anaerobic respiration	anaerobic respiration
SSON_3022	2	0	Inf	1	ygeX	diaminopropionate ammonia-lyase	metabolism	-
SSON_4426	1	0	Inf	1	pyrB	aspartate carbamoyltransferase catalytic subunit	DNA metabolism	-
SSON_3017	1	1	0.31	1	xdhA	xanthine dehydrogenase subunit	DNA metabolism	-
SSON_3100	4	1	1.23	0	rsmE	16S ribosomal RNA methyltransferase	RNA modification	RNA modification
SSON_2196	1	0	Inf	1	yohI	tRNA-dihydrouridine synthase C	RNA modification	RNA modification
SSON_2555	3	2	0.46	1	ypfI	hypothetical protein	RNA modification	RNA modification
SSON_0033	2	0	Inf	1	fkpB	FKBX-type 16KD peptidyl-prolyl cis-trans isomerase	repair	protein repair
SSON_0374	4	1	1.23	0	sbcC	exonuclease subunit	repair	phage resistance; DNA repair
SSON_2310	1	0	Inf	1		competence damage-inducible protein A	repair	damage-inducible protein
SSON_4205	1	0	Inf	1	yjbF	hypothetical protein	Unknown	extracellular polysaccharide biosynthetic process
SSON_0550	2	0	Inf	1	ybdD	hypothetical protein	Unknown	-
SSON_2657	0	2	0	1		hypothetical protein	Unknown	-

References

1. Vinh, H. *et al.* A multi-center randomized trial to assess the efficacy of gatifloxacin versus ciprofloxacin for the treatment of shigellosis in Vietnamese children. *PLoS Negl Trop Dis* **5**, e1264 (2011).
2. Vinh, H. *et al.* A changing picture of shigellosis in southern Vietnam: shifting species dominance, antimicrobial susceptibility and clinical presentation. *BMC Infect Dis* **9**, 204 (2009).
3. Vinh, H. *et al.* Treatment of bacillary dysentery in Vietnamese children: two doses of ofloxacin versus 5-days nalidixic acid. *Trans Royal Soc Trop Med Hyg* **94**, 323-6 (2000).
4. Thompson, C.N. *et al.* Epidemiological features and risk factors of Salmonella gastroenteritis in children resident in Ho Chi Minh City, Vietnam. *Epidemiol Infect*, 1-10 (2012).
5. von Seidlein, L. *et al.* A multicentre study of Shigella diarrhoea in six Asian countries: disease burden, clinical manifestations, and microbiology. *PLoS Med* **3**, e353 (2006).
6. Holt, K. *et al.* Shigella sonnei genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. *Nat Genet* **44**, 1056-9 (2012).
7. Nguyen, N.T. *et al.* The sudden dominance of blaCTX-M harbouring plasmids in Shigella spp. Circulating in Southern Vietnam. *PLoS Negl Trop Dis* **4**, e702 (2010).
8. Woodford, N. *et al.* Complete nucleotide sequences of plasmids pEK204, pEK499, and pEK516, encoding CTX-M enzymes in three major Escherichia

coli lineages from the United Kingdom, all belonging to the international O25:H4-ST131 clone. *Antimicrob Agents Chemother* **53**, 4472-82 (2009).

9. Cascales, E. *et al.* Colicin biology. *Microbiol Mol Biol Rev* **71**, 158-229 (2007).

10. Lau, P.C. & Condie, J.A. Nucleotide sequences from the colicin E5, E6 and E9 operons: presence of a degenerate transposon-like structure in the ColE9-J plasmid. *Mol General Genet* **217**, 269-77 (1989).