BMC Microbiology

RESEARCH ARTICLE   **Open Access**

Plasmids of *Shigella flexneri* serotype 1c strain Y394 provide advantages to bacteria in the host

Pawan Parajuli, Munazza I. Rajput and Naresh K. Verma*

Abstract

Background: *Shigella flexneri* has an extremely complex genome with a significant number of virulence traits acquired by mobile genetic elements including bacteriophages and plasmids. *S. flexneri* serotype 1c is an emerging etiological agent of bacillary dysentery in developing countries. In this study, the complete nucleotide sequence of two plasmids of *S. flexneri* serotype 1c strain Y394 was determined and analysed.

Results: The plasmid pINV-Y394 is an invasive or virulence plasmid of size 221,293 bp composed of a large number of insertion sequences (IS), virulence genes, regulatory and maintenance genes. Three hundred and twenty-eight open reading frames (ORFs) were identified in pINV-Y394, of which about a half (159 ORFs) were identified as IS elements. Ninety-seven ORFs were related to characterized genes (majority of which are associated with virulence and their regulons), and 72 ORFs were uncharacterized or hypothetical genes. The second plasmid pNV-Y394 is of size 10,866 bp and encodes genes conferring resistance against multiple antibiotics of clinical importance. The multidrug resistance gene cassette consists of tetracycline resistance gene *tetA*, streptomycin resistance gene *strA-strB* and sulfonamide-resistant dihydropteroate synthase gene *sul2*.

Conclusions: These two plasmids together play a key role in the fitness of Y394 in the host environment. The findings from this study indicate that the pathogenic *S. flexneri* is a highly niche adaptive pathogen which is able to co-evolve with its host and respond to the selection pressure in its environment.

Keywords: *Shigella flexneri*, Plasmids, Virulence, Multidrug resistance, Bacterial pathogen

Background

Shigella is a Gram-negative, non-motile, facultative anaerobic human enteric pathogen of the family *Enterobacteriaceae* that is closely related to *Escherichia coli* but has acquired specific traits of pathogenicity, physiology, and antigenic diversity [1]. The key to its virulence is the acquisition of a 210–230 kb virulence plasmid (VP) that enables the bacterium to invade and spread into the intestinal epithelial cells and induce apoptosis in the infected macrophages [2, 3]. There are at least 190 million shigellosis cases and 70,000 deaths annually; principally in developing countries [4]. *S. flexneri* is the primary cause of shigellosis in the developing countries (up to 62% of all *Shigella* spp. infections) [5]. *S. flexneri* has no

less than 19 different serotypes based on their antigenic determinants present on the O-antigen of the outer membrane lipopolysaccharide (LPS) [6]. Out of which, *S. flexneri* serotype 1c has emerged as a significant serotype in developing countries over the last decade [7].

Mobile genetic elements including bacteriophages, insertion sequences and plasmids play a very important role in pathogen evolution and genomic plasticity in several bacterial pathogens and *S. flexneri* is not an exception. The VP of *S. flexneri* encodes essential virulence factors regulated by several genes located on the VP as well as on its chromosome [8]. The plasmid comprises a conserved pathogenicity island of 31 kb region which encodes the Ipa-Mxi-Spa type III secretion system (T3SS) [9]. While performing the complete genome sequencing of *S. flexneri* serotype 1c strain Y394 ("Y394" hereafter) [7], we recovered two plasmids which provide pathogenicity advantage to Y394 in the host environment. The

* Correspondence: naresh.verma@anu.edu.au
Division of Biomedical Science and Biochemistry, Research School of Biology, The Australian National University, Canberra, ACT, Australia

first plasmid corresponds to the VP or invasion plasmid of Y394, hence named pINV-Y394. The second plasmid designated as pNV-Y394 possesses a multidrug-resistance cassette. This paper reports the complete nucleotide sequence and gene analysis of the two aforementioned plasmids to further facilitate the understanding of the evolution and pathogenic determinants of Y394, a newly emerged serotype, serotype 1c, of *S. flexneri*.

Results

General features of pINV-Y394

The large VP pINV-Y394 is of size 221,293 bp and comprises of 328 open reading frames (ORFs). The pINV-Y394 is a mosaic of essential virulence genes and their regulons, a significant number of IS elements and several hypothetical genes yet to be characterized (Fig. 1). The GC content of pINV-Y394 was found to be 45.9%, which is lower than that of the Y394 chromosome (50.9%).

The size and the composition, however, are similar to previously published *S. flexneri* serotype 2a VP, pCP301 (GenBank Accession number AL391753); *S. flexneri* serotype 5a VP, pWR100 (GenBank Accession number AF386526) and *S. flexneri* serotype 1a VP, unnamed plasmid1 (referred as "INV-0670" hereafter; GeneBank Accession number CP020087.1) (Table 1).

The pINV-Y394 has a similar arrangement of genes as seen in pCP301, pWR100 and INV-0670 with the exception of a large inversion and a few recombination events relative to the pINV-Y394 (Fig. 2). The 31 kb ipa-mxi-spa region is identical in all the compared VPs. We identified 27 unique genes in pINV-Y394 that are not present in three other compared plasmids (with 95% identity cut-off). These genes code for IS elements/transposases (21 ORFs), Osp2 domain protein (1 ORF), conjugation transfer protein (1 ORF) and uncharacterized or hypothetical proteins (4 ORFs).

IS elements

The sequence analysis of pINV-Y394 identified 159 ORFs related to IS elements, which accounts for 34 % (75,492 bp of 221,293 bp) of the total plasmid sequence. These 159 IS elements belong to 36 different IS families with IS3 being most dominant (76 ORFs) and encode for 52 complete IS elements, 106 partial IS elements and 1 novel putative IS element (Table 2). These IS elements spread all over the genome and usually enclose the virulence-associated genes as well as the uncharacterized ORFs [Fig. 1] with potential virulence function owing to its low GC [10].

Virulence associated genes

The 31 kb region of ipa-mxi-spa genes lies at position 100,933 bp-132,292 bp of pINV-Y394. This region comprises of 39 coding sequences clustered together and is transcribed in two directions. The 12 genes from *ipaJ* to *icsB* are transcribed in one direction whereas 27 genes from *ipgD* to *spaS/spa40* are in the opposite direction. The GC content of this region is 34.22% [range: 27% (ORF180; hypothetical gene)- 39% (Spa47/IpaC)] which is lower than the overall GC content of the plasmid. The 39 ORFs include five *ipa*, six *ipg*, twelve *mxi*, nine *spa*, *virB, ACP* (Acyl Carrier Protein), *icsB* and four uncharacterized genes. The ipa-mxi-spa region is flanked by truncated transposase on both the sides suggesting its origin from a single entity. The downstream to *spaS* lies a hypothetical gene (32% GC) containing conserved protein domain family *YmgB*. YmgB/AriR protein has been shown to be critical for biofilm formation and acid resistance in *E. coli* [11].

Besides the ipa-mxi-spa region, pINV-Y394 also possesses several other genes which might have a potential role in virulence. There are eleven putative *osp* genes (*ospB, ospC1, ospC2, ospC3, ospD1, ospD2, ospD3, ospE1, ospE2, ospF, ospG*) distributed across the plasmid genome. Several of these genes including *ospB, ospC1, ospE2*, and *ospF* have been shown to be regulated by the protein MxiE and is believed to have a role in post-invasion steps of infection [12].

Besides the invasion plasmid antigen (*ipa*) genes in the ipa-mxi-spa region, there are five more alleles of *ipaH* genes named as *ipaH1.4, ipaH2.5, ipaH4.5, ipaH7.8*, and *ipaH9.8*. These genes possess a novel E3 ligase (NEL) domain at the C-terminus and a series of leucine-rich repeats at the N-terminus which are found in many bacterial virulence factors. The leucine-rich repeats sequester a cysteine residue contained in the NEL domain until invasion has occurred allowing the release of NEL domain for site-specific function [13, 14].

The key regulatory gene for virulence function, *virF*, is located approximately 62 kb upstream of *virB* of the ipa-mxi-spa locus. The activation of the plasmid-encoded regulatory genes requires environmental stimuli such as temperature, osmolarity, and changes in DNA supercoiling [8, 15, 16]. The VirF protein binds to the promotor of *virB* which acts as a transcriptional activator of the genes of the T3SS [17]. Approximately 12-kb downstream of *spaS* lies the *virA* and *icsA/virG* with a distance of 0.5-kb between them. The product of these genes contributes to the intracellular spread and protection against autophagic degradation [18, 19]. The other *vir*-gene, *virK*, which is located 43-kb downstream to *virG* is required for post-transcriptional control of the *virG* gene [20].

Genes associated with plasmid maintenance and transfer functions

The sequence analysis also identified at least five sets of genes related to toxin-antitoxin (TA) systems. This includes
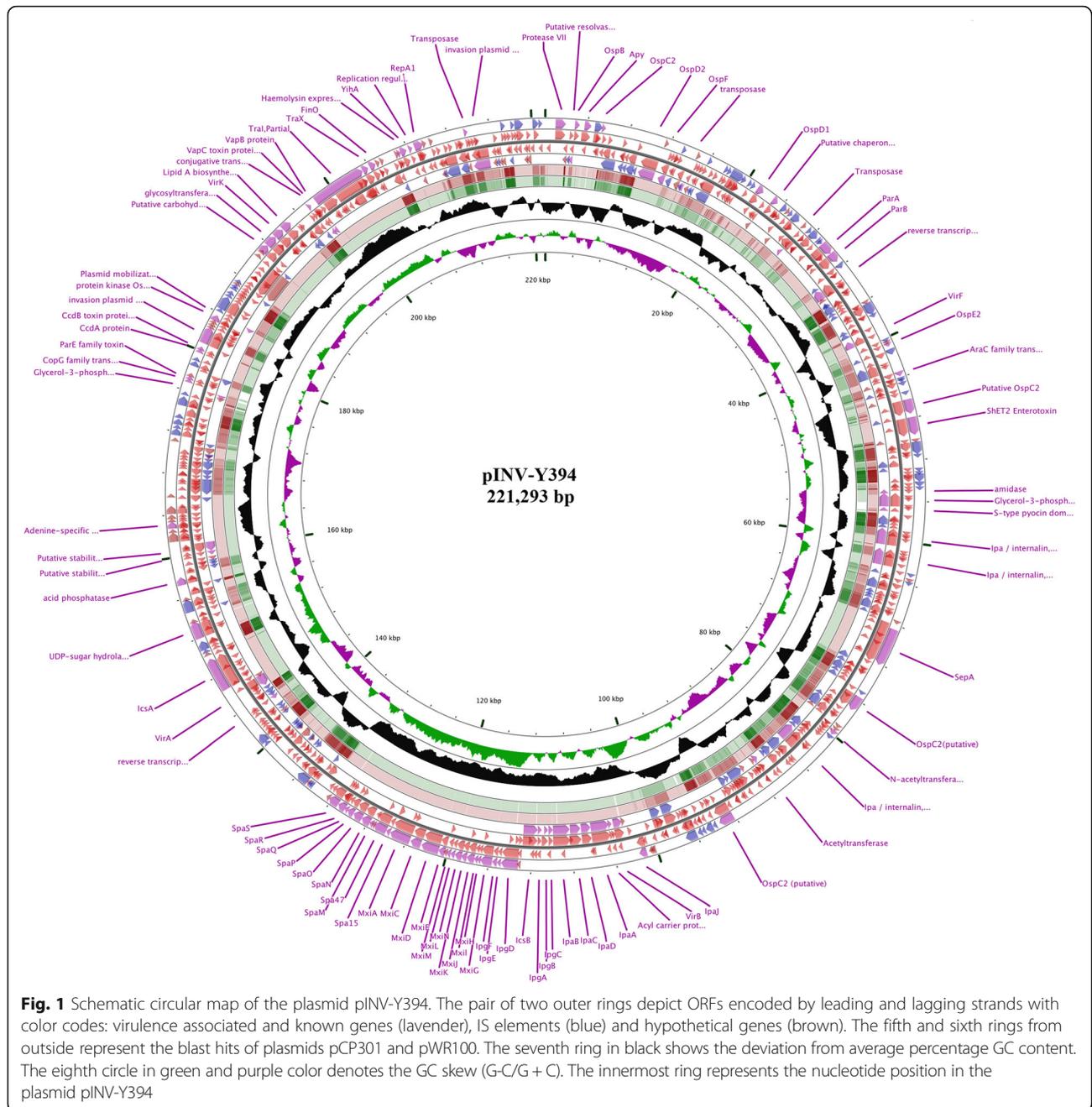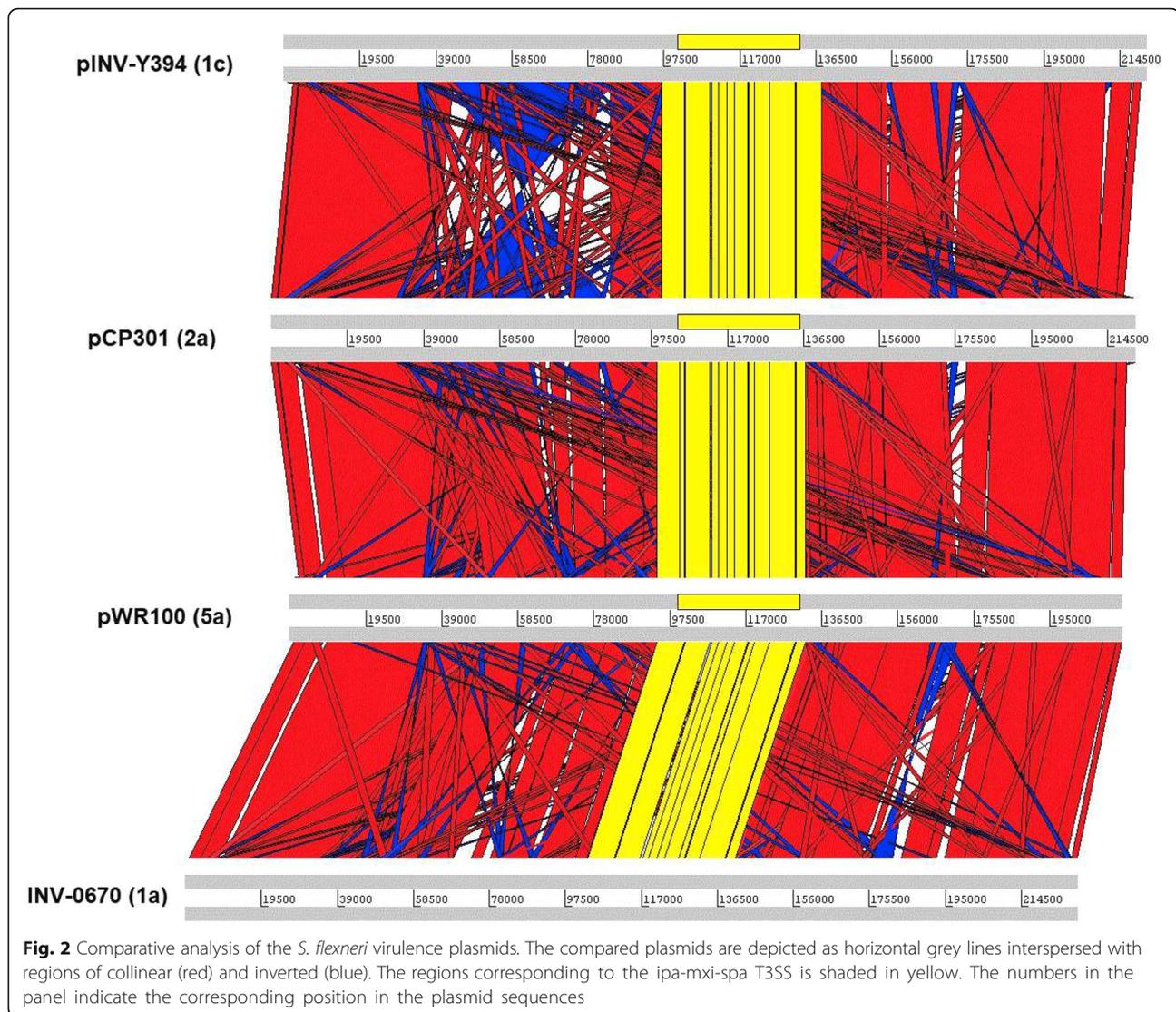
Fig. 1 Schematic circular map of the plasmid pINV-Y394. The pair of two outer rings depict ORFs encoded by leading and lagging strands with color codes: virulence associated and known genes (lavender), IS elements (blue) and hypothetical genes (brown). The fifth and sixth rings from outside represent the blast hits of plasmids pCP301 and pWR100. The seventh ring in black shows the deviation from average percentage GC content. The eighth circle in green and purple color denotes the GC skew (G-C/G + C). The innermost ring represents the nucleotide position in the plasmid pINV-Y394

Table 1 General features of *Shigella flexneri* virulence plasmids

Plasmid	Host Strain	Plasmid Size (bp)	GC (%)	ORFs Total	ORFs related to IS	ORFs Uncharacterized	Unique ORFs[a]
pINV-Y394	*Shigella flexneri* 1c strain Y394	221,293	45.9	328	159	72	27
pCP301	*Shigella flexneri* 2a strain 301	221,618	45.8	316	147	78	24
pWR100	*Shigella flexneri* 5a strain M90 T	213,494	45.7	311	138	82	23
unnamed1	*Shigella flexneri* 1a strain 0670	228,834	46.1	301	145	93	35

[a]Unique ORFs refer to ORFs which are present in corresponding plasmid only

Fig. 2 Comparative analysis of the *S. flexneri* virulence plasmids. The compared plasmids are depicted as horizontal grey lines interspersed with regions of collinear (red) and inverted (blue). The regions corresponding to the ipa-mxi-spa T3SS is shaded in yellow. The numbers in the panel indicate the corresponding position in the plasmid sequences

parAB, ccdAB, stbAB, vapBC and *parE/copG.* These genes are essential for maintaining the plasmids in bacteria through post-segregational killing mechanisms which eliminate bacterial cells lacking a plasmid after cell division. Besides, these genes are also involved in multiple cellular functions associated with survival under stress conditions [21].

pINV-Y394 consisted of several truncated conjugative transfer genes suggesting that it had conjugative transfer system but had lost its essential components over time. The plasmid possessed only 6 % (45 AA) of the fully functional TraD (723 AA) which is an essential component of DNA transfer in bacterial conjugation system [22]. The pINV-Y394 lacks the essential region of TraD, the 38 amino acids at the very C-terminal end of the protein. The *traI* gene in pINV-Y394 has a deletion leading to a frameshift mutation. The two other genes present in pINV-Y394 related to conjugation are *traX*

and *finO* which code for pilus acetylation protein and fertility inhibition protein. The FinO is one of the components of FinOP fertility inhibition complex. This protein inhibits the expression of *traJ* gene which in turn regulates the expression of some 20 other transfer genes. However, the remainder of the transfer genes was absent in pINV-Y394. The repeated conjugation experiments with pINV-Y394 were unsuccessful and are consistent with the previous observation [23].

Plasmid pNV-Y394

The plasmid pNV-Y394 is 10,866 bp in size and has GC content of 61.54% which is well above the 51% GC content of the host bacteria Y394 (Fig. 3). The plasmid encodes multiple antibiotic resistance genes of clinical importance (Table 3). The plasmid is novel in *S. flexneri* but has partial homology of multidrug-resistance genes cassette (from nucleotides 5091 to 10,715) with other

Table 2 Predicted Insertion Sequence (IS) family in pINV-Y394

IS Family	Total ORFs	Complete IS	Different IS(s)	Total Size
IS1	19	9	4	4806
IS110	3	2	2	2568
IS110 ssgr[a] IS1111	2	0	2	1140
IS21	4	2	1	4971
IS3 ssgr IS2	11	3	4	4953
IS3 ssgr IS3	32	11	4	11,718
IS3 ssgr IS51	28	8	1	11,124
IS3 ssgr IS150	5	1	3	1731
IS4 ssgr IS4	3	3	1	3960
IS4 ssgr IS10	2	1	1	1527
IS4 ssgr IS50	1	0	1	1116
IS5 ssgr IS427	7	3	1	2547
IS66	18	8	4	12,699
IS91	20	0	5	8376
IS630	2	1	1	1617
ISL3	2	0	1	639
Total	159	52	36	75,492

[a]ssgr -subgroups

bacterial genomes and plasmids. The BLASTN result showed 100% query cover and 99% nucleotide identity with genomes and plasmids of several other clinical isolates including *Salmonella enterica* (GenBank accession numbers CP022498.1 and CP004059.1), *Citrobacter freundii* (GenBank accession number KY986974.1), *Escherichia coli* (GenBank accession number CP007137.1), *Klebsiella pneumoniae* (GenBank accession number CP021960.1) and *Vibrio parahaemolyticus* (GenBank accession number KY014465.1). The arrangement of genes in 12 kb region of the aforementioned genomes/plasmids along with 5 kb region of homology is shown in Fig. 4.

The multidrug cassette consists of genes *sul2, strA-strB* and *tetA* which confers resistance against sulfonamide, streptomycin, and tetracycline, respectively. The phenotype was consistent with the genotype of pNV-Y394. The phenotypic expression of these antibiotic resistance genes was also confirmed with *E. coli* JM109 cells transformed with pNV-Y394.

Although pNV-Y394 harbored genes *mobA* and *mobC* associated with plasmid transfer and *repA* and *repC* associated with plasmid replication, we were unable to show the plasmid transfer by conjugation experiments, suggesting these genes are not sufficient for conjugative transfer.

Discussion

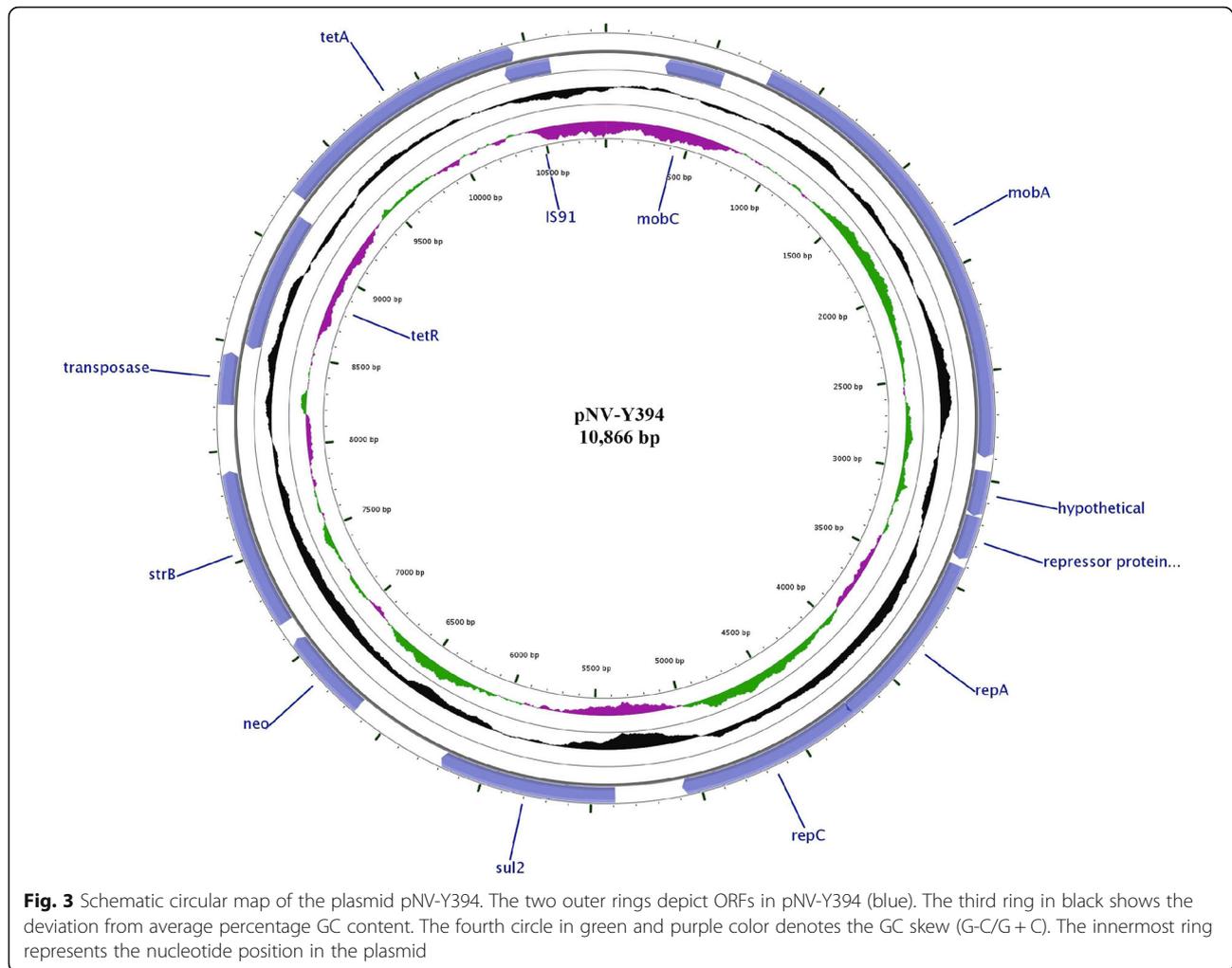The large VP pINV-Y394 in *S. flexneri* serotype 1c strain Y394 encodes a 30-kb ipa-mxi-spa locus consisting of type III secretion system (T3SS) and an array of virulence

factors which enable the bacteria to initiate invasion and subsequent interaction with the host epithelial cells and adapt intracellular lifestyle [2]. While the *ipa*, *mxi* and *spa* genes are mostly involved in early stage of infection, the *osp* and *ipaH* genes are involved in post-invasion events [12]. The *ipaH* family genes are present on both the VP and the *Shigella* chromosome [24] and are capable of subverting the host's ubiquitination pathway [13, 14].
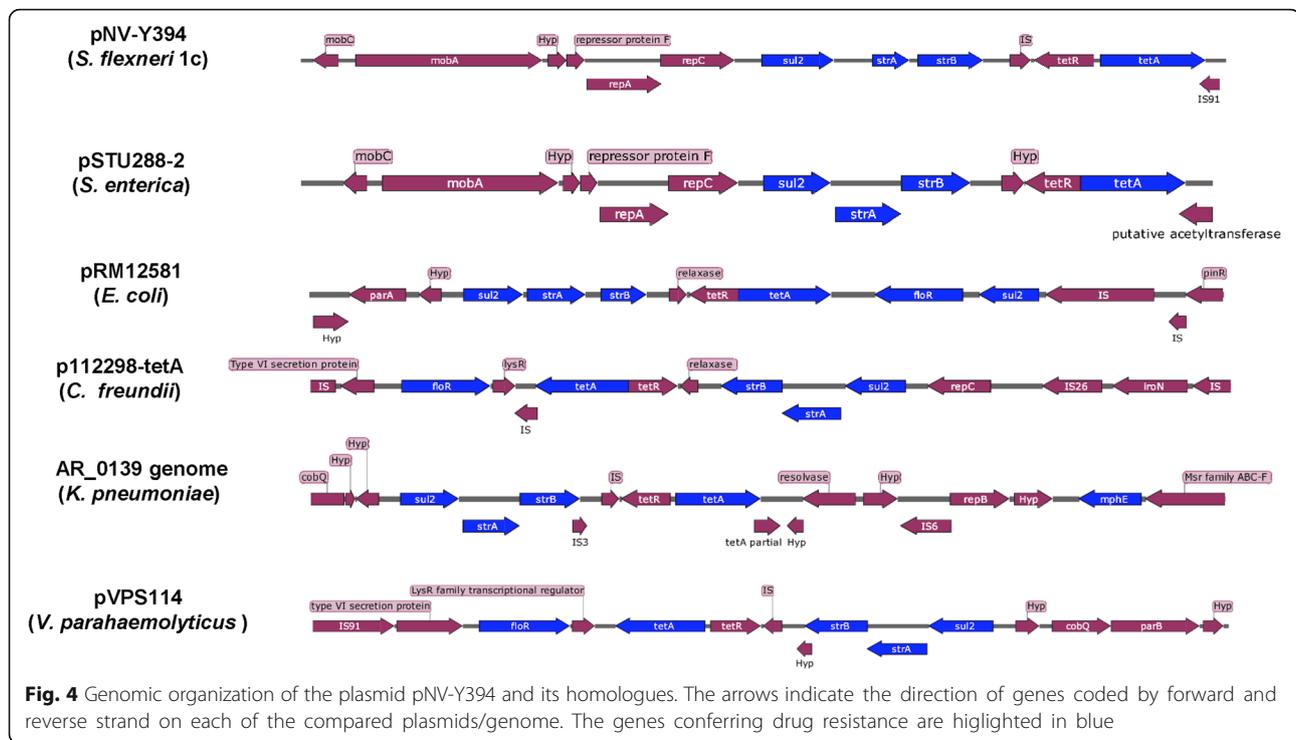
The low GC percentage of pINV-Y394 suggests that the pathogen has acquired the plasmid later as a response to adaptive evolution. There are a large number of IS elements in pINV-Y394 which accounts over one-third of the total plasmid genome. Most of the genes associated with virulence have lower GC content and are flanked by the IS elements suggesting that these genes have been brought as the result of numerous horizontal gene transfer via IS-mediated recombination events. The comparison of VP sequences from three other serotypes of *S. flexneri* showed a high level of similarity with some recombination and a large inversion [23, 25]. The ipa-mxi-spa region was highly conserved in all the compared plasmids and is flanked by the IS elements at both the ends suggesting that these genes have a common origin and was acquired by IS-mediated recombination. Further, the few differences in gene content of the compared plasmids were mostly in the sequence of IS elements; suggesting all the compared plasmids had a common backbone. Out of 328 ORFs, about 26 % of them are uncharacterized in pINV-Y394. Many of which have low GC content and flanked by IS elements, suggesting that these genes are foreign and might have brought to the plasmid by several IS-mediated recombination events. These uncharacterized genes warrant further studies on their function(s).

The lack of conjugative transfer of pINV-Y394 is consistent with previous studies [23, 25] and can be explained by the loss of intact *tra* locus which is essential for the exchange of plasmids between two bacteria by conjugation. The TraD protein which acts as a coupling protein and connects the DNA-processing machinery to the mating pair-forming transfer apparatus is truncated in pINV-Y394 with the loss of 38 amino acid C-terminal region essential for TraM interaction [22]. The genes *traD, traM,* and *traY* are involved in the *oriT* nicking, strand displacement, and DNA transfer during conjugation [26].

The other important gene of the *tra* locus, *traI*, has a frameshift mutation in its coding sequence. The protein TraI is a bifunctional protein which acts as a sequence-specific DNA trans-esterase, providing the site- and strand-specific nick required to initiate DNA transfer and a processive 5′ to 3′ helicase reaction that provides the motive force for strand transfer [27]. The loss of transfer function in VP has stabilized the virulence genes

Fig. 3 Schematic circular map of the plasmid pNV-Y394. The two outer rings depict ORFs in pNV-Y394 (blue). The third ring in black shows the deviation from average percentage GC content. The fourth circle in green and purple color denotes the GC skew (G-C/G + C). The innermost ring represents the nucleotide position in the plasmid

Table 3 Features of pNV-Y394

Feature/gene	Start	Stop	Length (bp)	Function
mobC	569	285	285	Mobilization protein C
mobA	768	2897	2130	Conjugal transfer protein TraA
Hypothetical protein	2958	3170	213	Hypothetical protein (Multispecies of Bacteria)
Repressor protein F	3172	3378	207	Major facilitator superfamily MFS_1
repA	3408	4247	840	IS, phage, Tn; Transposon-related functions
repC	4234	5085	852	Plasmid replication protein C
sul2	5393	6208	816	Dihydropteroate synthase (EC 2.5.1.15)
strA/neo	6653	7072	420	Aminoglycoside 3′-phosphotransferase (EC 2.7.1.95) Streptomycin 3′-kinase StrA (EC 2.7.1.87)
strB	7164	7907	744	Aminoglycoside 3′-phosphotransferase 2 (EC 2.7.1.95) Streptomycin 3′-kinase StrB (EC 2.7.1.87)
Putative transposase	8213	8455	243	Putative transposase/relaxase /helicase
tetR	9164	8487	678	Transcriptional regulator, TetR family
tetA	9243	10,442	1200	Tetracycline efflux protein TetA
IS91	10,373	10,597	225	IS91 family transposase

Fig. 4 Genomic organization of the plasmid pNV-Y394 and its homologues. The arrows indicate the direction of genes coded by forward and reverse strand on each of the compared plasmids/genome. The genes conferring drug resistance are higlighted in blue

alongside allowing parallel evolution of the VPs with the chromosomes [28].

Although, no other *S. flexneri* strain has been previously reported with multidrug resistance region as found in pNV-Y394, the 5-kb antibiotic resistance cassette of plasmid pNV-Y394 which provides resistance to the bacteria against sulfonamide, streptomycin and tetracycline antibiotics of clinical importance, has been isolated from several bacterial pathogens including *S. sonnei* [29], *Salmonella enterica* [30] and *Klebsiella pneumoniae* [31]. The antibiotic resistance cassette in pNV-Y394 is flanked by IS91 family transposases, which are linked to active rolling-circle transposition of several virulence and toxin genes and play an important role in pathogen evolution [32, 33]. Although, pNV-Y394 lacks complete conjugative transfer elements and is non-conjugative, the antibiotic gene cassette might have brought by IS91-mediated gene dissemination into Y394. The horizontal acquisition of multiple drug resistance genes has been suggested to play an important role in driving the global spread of a pathogen as a new epidemic clone [34, 35]. The increase in multiple drug resistance strains of *S. flexneri,* endemic in developing countries, is of major concern due to limited resource settings for antibiotic susceptibility testing and limited therapeutic options. The antibiotic resistance surveillance studies across multiple *Shigella* endemic regions including China, India, Bangladesh, Pakistan, Nepal, Gabon, Kenya, Senegal, Iran and Egypt showed high percentage

of *S. flexneri* isolates being resistant to tetracycline, trimethoprim-sulfamethoxazole, ampicillin and streptomycin [36–48]. This warrants the choice of antibiotic therapy be based on recent antibiotics susceptibility data of *Shigella* isolated from nearby geographical regions.

Conclusions

The comparison of pINV-Y394 with pCP301, pWR100 and INV-0670 showed that all four plasmids were highly conserved despite they were derived from different *S. flexneri* serotypes. This suggests that VP is a prerequisite for *S. flexneri* virulence and might have been acquired from a common ancestor. Further functional characterization of the conserved hypothetical genes would increase our understanding of virulence gene regulatory networks. The VP pINV-Y394 has the essential genetic element for its virulence and pNV-Y394 provides resistance to Y394 against an array of antibiotics of clinical importance. Hence, the two plasmids isolated from Y394 have an important role in the fitness of Y394 in the host environment. Furthermore, proper surveillance of antimicrobial resistance of this globally important pathogen is warranted, as this information is crucial for effective outbreak management and control of the disease.

Methods
Bacterial strains and plasmids
The two plasmids pINV-Y394 and pNV-Y394 presented in this study were derived from *S. flexneri* serotype 1c

strain Y394 which was kindly provided to us by Nils I. A. Carlin [49]. The Y394 was grown aerobically (180 rpm) at 37 °C in Luria Bertani broth (LB). The total bacterial DNA was prepared using the Genomic Tip 100/G (Qiagen) according to the manufacturer's instructions. The LB was supplemented with antibiotics (Sigma-Aldrich, Germany) where needed. The antibiotic susceptibility pattern of Y394 was determined as described previously [7].

Plasmid DNA sequencing and assembly

The complete genome of *S. flexneri* serotype 1c strain Y394 was sequenced using the hybrid method of single--molecule real-time (SMRT) sequencing technology and short-read MiSeq (Illumina) sequencing technology as described previously [7]. Briefly, The SMRTbell Template Prep Kit 1.0 (PacBio) was used for SMRT sequencing library preparation of about 20 kb insert size. The sequencing was performed using PacBio RSII sequencing system. The Nextera XT DNA library preparation kit (Illumina) was used for MiSeq v3 300bp paired-end sequencing. A de novo assembly of these reads was performed with HGAP.3 (Pacific Biosciences) on the SMRT Analysis Pipeline version 2.3.0 [50]. Following de novo assembly of the SMRT reads, a single contig of 4,584,634 bp (203X coverage) representing the Y394 chromosome was obtained (GenBank Accession number CP020753). The sequence reads that did not map to the chromosome were then independently assembled into two contigs of size 221,307 bp (250X coverage) representing pINV-Y394 and 10,873 bp (30X coverage) representing pNV-Y394. The circularization of the plasmid was obtained using a package Circlator [51]. To improve the quality of the SMRT sequences, the miseq reads were quality trimmed using Trimmomatic v0.36 [52] and used for error correction by using BWA-MEM and Pilon [53, 54]. After circularization and Pilon improvement the final size of pINV-Y394 was 221,293 bp and pNV-Y394 was 10,866 bp.

DNA sequence annotation and analysis

The automated annotation of plasmid sequences was performed using Rapid Annotation using Subsystem Technology (RAST) version 2.0 [55] followed by manual curation using BLASTP searches [56]. The IS elements were identified using ISsaga [57]. The antibiotic profiles of the plasmids were examined with ResFinder [58]. The genes that were common and unique in each of the compared plasmids were analysed using Roary [59].

The image files showing the synteny map of the plasmids were obtained using ACT (Artemis Comparison Tool) [60], CG View Server [61] and Snap Gene Viewer (Version 3.3.1).

Plasmid characterization and tagging

The plasmid pINV-Y394 was first tagged with *kan* gene for antibiotic selection using λ-red recombinase system as described previously [62]. Briefly, the Y394 strain was first transformed with pKD46 plasmid encoding λ-red recombinase system. Fifty-base pair overhang primers (Sigma, Australia) were designed corresponding to the putative reverse transcriptase gene of pINV-Y394 and used for amplification of *kan* gene using pKD4 miniprep DNA as a template (Additional file 1: Table S1). The PCR amplicons were then DpnI-treated; column purified and transformed with Y394 + pKD46 electrocompetent cells. The transformants were selected using LB agar plates containing kanamycin (50 μg/ml) and confirmed by Sanger sequencing.

To examine the phenotypic expression of antibiotics resistant genes in pNV-Y394, miniprep DNA was extracted using Wizard® Plus SV Minipreps DNA purification system (Promega, USA) and the plasmid preparations were used to transform electrocompetent *Escherichia coli* JM109 cells and the resultant clones were recovered using LB-agar plates containing either tetracycline (10 μg/ml) and/or streptomycin (10 μg/ml).

Conjugation

To determine the ability of conjugative transfer of the two plasmids, mating experiments were performed. The mating experiment was carried out by mixing the exponential growth phase cultures of donor *S. flexneri* serotype 1c strain Y394 and recipient *E. coli* JM109 transformed with pBCSK (+) containing chloramphenicol-resistance gene (*Cm^r*) to aid in antibiotic selection. The mating mixture was incubated in a 37 C water bath for 1 h and resuspended in PBS. The serial dilutions were then plated on LB plate containing either kanamycin (50 μg/ml) and chloramphenicol (25 μg/ml) or tetracycline (10 μg/ml) and chloramphenicol (25 μg/ml) to discriminate the conjugative transfer of plasmids pINV-Y394 and pNV-Y394, respectively.

Additional file

Additional file 1: Table S1. Primers used in this study. (PDF 89 kb)

Parajuli *et al. BMC Microbiology*   (2019) 19:86

Page 9 of 10

The corresponding author had full access to all the data in the study and final decision to submit for publication.

Publisher's Note
Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References
1. Killackey SA, Sorbara MT, Girardin SE. Cellular aspects of Shigella pathogenesis: focus on the manipulation of host cell processes. Front Cell Infect Microbiol. 2016;6:38.
2. Sansonetti PJ, Kopecko DJ, Formal SB. Involvement of a plasmid in the invasive ability of Shigella flexneri. Infect Immun. 1982;35(3):852–60.
3. Zychlinsky A, Prevost MC, Sansonetti PJ. Shigella flexneri induces apoptosis in infected macrophages. Nature. 1992;358(6382):167–9.
4. Pires SM, Fischer-Walker CL, Lanata CF, Devleesschauwer B, Hall AJ, Kirk MD, Duarte AS, Black RE, Angulo FJ. Aetiology-specific estimates of the global and regional incidence and mortality of Diarrhoeal diseases commonly transmitted through food. PLoS One. 2015;10(12):e0142927.
5. Gu B, Cao Y, Pan S, Zhuang L, Yu R, Peng Z, Qian H, Wei Y, Zhao L, Liu G, et al. Comparison of the prevalence and changing resistance to nalidixic acid and ciprofloxacin of Shigella between Europe-America and Asia-Africa from 1998 to 2009. Int J Antimicrob Agents. 2012;40(1):9–17.
6. Sun Q, Lan R, Wang J, Xia S, Wang Y, Wang Y, Jin D, Yu B, Knirel YA, Xu J. Identification and characterization of a novel Shigella flexneri serotype Yv in China. PLoS One. 2013;8(7):e70238.
7. Parajuli P, Adamski M, Verma NK. Bacteriophages are the major drivers of Shigella flexneri serotype 1c genome plasticity: a complete genome analysis. BMC Genomics. 2017;18(1):722.
8. Maurelli AT, Sansonetti PJ. Identification of a chromosomal gene controlling temperature-regulated expression of Shigella virulence. Proc Natl Acad Sci U S A. 1988;85(8):2820–4.
9. Sasakawa C, Kamata K, Sakai T, Makino S, Yamada M, Okada N, Yoshikawa M. Virulence-associated genetic regions comprising 31 kilobases of the 230-kilobase plasmid in Shigella flexneri 2a. J Bacteriol. 1988;170(6):2480–4.
10. Lawrence JG, Ochman H. Amelioration of bacterial genomes: rates of change and exchange. J Mol Evol. 1997;44(4):383–97.
11. Lee J, Page R, García-Contreras R, Palermino J-M, Zhang X-S, Doshi O, Wood TK, Peti W. Structure and function of the Escherichia coli protein YmgB: a protein critical for biofilm formation and acid-resistance. J Mol Biol. 2007;373(1):11–26.
12. Kane CD, Schuch R, Day WA Jr, Maurelli AT. MxiE regulates intracellular expression of factors secreted by the Shigella flexneri 2a type III secretion system. J Bacteriol. 2002;184(16):4409–19.
13. Quezada CM, Hicks SW, Galan JE, Stebbins CE. A family of Salmonella virulence factors functions as a distinct class of autoregulated E3 ubiquitin ligases. Proc Natl Acad Sci U S A. 2009;106(12):4864–9.
14. Rohde JR, Breitkreutz A, Chenal A, Sansonetti PJ, Parsot C. Type III secretion effectors of the IpaH family are E3 ubiquitin ligases. Cell Host Microbe. 2007;1(1):77–83.
15. Bernardini ML, Fontaine A, Sansonetti PJ. The two-component regulatory system ompR-envZ controls the virulence of Shigella flexneri. J Bacteriol. 1990;172(11):6274–81.
16. Dorman CJ, Ni Bhriain N, Higgins CF. DNA supercoiling and environmental regulation of virulence gene expression in Shigella flexneri. Nature. 1990;344(6268):789–92.
17. Adler B, Sasakawa C, Tobe T, Makino S, Komatsu K, Yoshikawa M. A dual transcriptional activation system for the 230 kb plasmid genes coding for virulence-associated antigens of Shigella flexneri. Mol Microbiol. 1989;3(5):627–35.
18. Campbell-Valois FX, Sachse M, Sansonetti PJ, Parsot C. Escape of actively secreting Shigella flexneri from ATG8/LC3-positive vacuoles formed during cell-to-cell spread is facilitated by IcsB and VirA. MBio. 2015;6(3):e02567–14.
19. Ogawa M, Yoshimori T, Suzuki T, Sagara H, Mizushima N, Sasakawa C. Escape of intracellular Shigella from autophagy. Science. 2005;307(5710):727–31.
20. Nakata N, Sasakawa C, Okada N, Tobe T, Fukuda I, Suzuki T, Komatsu K, Yoshikawa M. Identification and characterization of virK, a virulence-associated large plasmid gene essential for intercellular spreading of Shigella flexneri. Mol Microbiol. 1992;6(16):2387–95.
21. Fernandez-Garcia L, Blasco L, Lopez M, Bou G, Garcia-Contreras R, Wood T, Tomas M. Toxin-antitoxin Systems in Clinical Pathogens. Toxins (Basel). 2016;8(7).
22. Beranek A, Zettl M, Lorenzoni K, Schauer A, Manhart M, Koraimann G. Thirty-eight C-terminal amino acids of the coupling protein TraD of the F-like conjugative resistance plasmid R1 are required and sufficient to confer binding to the substrate selector protein TraM. J Bacteriol. 2004;186(20):6999–7006.
23. Buchrieser C, Glaser P, Rusniok C, Nedjari H, D'Hauteville H, Kunst F, Sansonetti P, Parsot C. The virulence plasmid pWR100 and the repertoire of proteins secreted by the type III secretion apparatus of Shigella flexneri. Mol Microbiol. 2000;38(4):760–71.
24. Venkatesan MM, Buysse JM, Hartman AB. Sequence variation in two ipaH genes of Shigella flexneri 5 and homology to the LRG-like family of proteins. Mol Microbiol. 1991;5(10):2435–45.
25. Zhang J, Liu H, Zhang X, Yang J, Yang F, Yang G, Shen Y, Hou Y, Jin Q. Complete DNA sequence and gene analysis of the virulence plasmid pCP301 of Shigella flexneri 2a. Sci China C Life Sci. 2003;46(5):513–21.
26. Raleigh EA, Low KB. Conjugation A2 - Maloy, Stanley. In: Hughes K, editor. Brenner's Encyclopedia of Genetics. 2nd ed. San Diego: Academic Press; 2013. p. 144–51.
27. Matson SW, Ragonese H. The F-plasmid TraI protein contains three functional domains required for conjugative DNA strand transfer. J Bacteriol. 2005;187(2):697–706.
28. Yang J, Nie H, Chen L, Zhang X, Yang F, Xu X, Zhu Y, Yu J, Jin Q. Revisiting the molecular evolutionary history of Shigella spp. J Mol Evol. 2007;64(1):71–9.
29. Yang F, Yang J, Zhang X, Chen L, Jiang Y, Yan Y, Tang X, Wang J, Xiong Z, Dong J, et al. Genome dynamics and diversity of Shigella species, the etiologic agents of bacillary dysentery. Nucleic Acids Res. 2005;33(19):6445–58.
30. Hooton SP, Timms AR, Cummings NJ, Moreton J, Wilson R, Connerton IF. The complete plasmid sequences of Salmonella enterica serovar typhimurium U288. Plasmid. 2014;76:32–9.
31. Sonnevend A, Yahfoufi N, Ghazawi A, Jamal W, Rotimi V, Pal T. Contribution of horizontal gene transfer to the emergence of VIM-4 carbapenemase producer Enterobacteriaceae in Kuwait. Infect Drug Resist. 2017;10:469–78.
32. del Pilar Garcillan-Barcia M, Bernales I, Mendiola MV, de la Cruz F. Single-stranded DNA intermediates in IS91 rolling-circle transposition. Mol Microbiol. 2001;39(2):494–501.
33. Garcillan-Barcia MP, de la Cruz F. Distribution of IS91 family insertion sequences in bacterial genomes: evolutionary implications. FEMS Microbiol Ecol. 2002;42(2):303–13.
34. Holt KE, Baker S, Weill FX, Holmes EC, Kitchen A, Yu J, Sangal V, Brown DJ, Coia JE, Kim DW, et al. Shigella sonnei genome sequencing and phylogenetic analysis indicate recent global dissemination from Europe. Nat Genet. 2012;44(9):1056–9.
35. Holt KE, Wertheim H, Zadoks RN, Baker S, Whitehouse CA, Dance D, Jenney A, Connor TR, Hsu LY, Severin J, et al. Genomic analysis of diversity, population structure, virulence, and antimicrobial resistance in Klebsiella

pneumoniae, an urgent threat to public health. Proc Natl Acad Sci U S A. 2015;112(27):E3574–81.

36. Soltan Dallal MM, Omidi S, Douraghi M, Haghi Ashtiani MT, Sharifi Yazdi MK, Okazi A. Molecular analysis of integrons and antimicrobial resistance profile in Shigella spp. isolated from acute pediatric diarrhea patients. GMS Hyg Infect Control. 2018;13:Doc02.

37. Afroze F, Ahmed T, Sarmin M, Smsb Shahid A, Shahunja KM, Shahrin L, Chisti MJ. Risk factors and outcome of Shigella encephalopathy in Bangladeshi children. PLoS Negl Trop Dis. 2017;11(4):e0005561.

38. Shakya G, Acharya J, Adhikari S, Rijal N. Shigellosis in Nepal: 13 years review of nationwide surveillance. J Health Popul Nutr. 2016;35(1):36.

39. Schaumburg F, Alabi AS, Kaba H, Lell B, Becker K, Grobusch MP, Kremsner PG, Mellmann A. Molecular characterization of Shigella spp. from patients in Gabon 2011-2013. Trans R Soc Trop Med Hyg. 2015;109(4):275–9.

40. Connor TR, Barker CR, Baker KS, Weill FX, Talukder KA, Smith AM, Baker S, Gouali M, Pham Thanh D, Jahan Azmi I, et al. Species-wide whole genome sequencing reveals historical global spread and recent local persistence in Shigella flexneri. Elife. 2015;4:e07335.

41. Zhang N, Lan R, Sun Q, Wang J, Wang Y, Zhang J, Yu D, Hu W, Hu S, Dai H, et al. Genomic portrait of the evolution and epidemic spread of a recently emerged multidrug-resistant Shigella flexneri clone in China. J Clin Microbiol. 2014;52(4):1119–26.

42. Zhang J, Jin H, Hu J, Yuan Z, Shi W, Yang X, Xu X, Meng J. Antimicrobial resistance of Shigella spp. from humans in Shanghai, China, 2004-2011. Diagn Microbiol Infect Dis. 2014;78(3):282–6.

43. Sambe-Ba B, Seck A, Wane AA, Fall-Niang NK, Gassama-Sow A. Sensitivity to antibiotics and genetic support to resistance of Shigella flexneri strains isolated in Dakar from 2001 to 2010. Bull Soc Pathol Exot. 2013;106(2):89–94.

44. Njuguna HN, Cosmas L, Williamson J, Nyachieo D, Olack B, Ochieng JB, Wamola N, Oundo JO, Feikin DR, Mintz ED, et al. Use of population-based surveillance to define the high incidence of shigellosis in an urban slum in Nairobi, Kenya. PLoS One. 2013;8(3):e58437.

45. Das SK, Ahmed S, Ferdous F, Farzana FD, Chisti MJ, Latham JR, Talukder KA, Rahman M, Begum YA, Qadri F, et al. Etiological diversity of diarrhoeal disease in Bangladesh. J Infect Dev Ctries. 2013;7(12):900–9.

46. Ahmed SF, Klena J, Husain T, Monestersky J, Naguib A, Wasfy MO. Genetic characterization of antimicrobial resistance of Shigella flexneri 1c isolates from patients in Egypt and Pakistan. Ann Clin Microbiol Antimicrob. 2013;12:9.

47. Ghosh S, Pazhani GP, Chowdhury G, Guin S, Dutta S, Rajendran K, Bhattacharya MK, Takeda Y, Niyogi SK, Nair GB, et al. Genetic characteristics and changing antimicrobial resistance among Shigella spp. isolated from hospitalized diarrhoeal patients in Kolkata, India. J Med Microbiol. 2011;60(Pt 10:1460–6.

48. Wilson G, Easow JM, Mukhopadhyay C, Shivananda PG. Isolation & antimicrobial susceptibility of Shigella from patients with acute gastroenteritis in Western Nepal. Indian J Med Res. 2006;123(2):145–50.

49. Wehler T, Carlin NI. Structural and immunochemical studies of the lipopolysaccharide from a new provisional serotype of Shigella flexneri. Eur J Biochem. 1988;176(2):471–6.

50. Chin CS, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013;10(6):563–9.

51. Hunt M, Silva ND, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator: automated circularization of genome assemblies using long sequencing reads. Genome Biol. 2015;16(1):294.

52. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.

53. Li H, Durbin R. Fast and accurate long-read alignment with burrows-wheeler transform. Bioinformatics. 2010;26(5):589–95.

54. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. PLoS One. 2014;9(11):e112963.

55. Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, et al. The RAST server: rapid annotations using subsystems technology. BMC Genomics. 2008;9:75.

56. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215(3):403–10.

57. Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. Genome Biol. 2011;12(3):R30.

58. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, Aarestrup FM, Larsen MV. Identification of acquired antimicrobial resistance genes. J Antimicrob Chemother. 2012;67(11):2640–4.

59. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–3.

60. Carver TJ, Rutherford KM, Berriman M, Rajandream MA, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21(16):3422–3.

61. Grant JR, Stothard P. The CGView server: a comparative genomics tool for circular genomes. Nucleic Acids Res. 2008;36(Web Server):W181–4.

62. Datsenko KA, Wanner BL. One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products. Proc Natl Acad Sci U S A. 2000; 97(12):6640–5.