Evaluation of Molecular Methods for Serotyping *Shigella flexneri*

Amy Gentle, Philip M. Ashton, Timothy J. Dallman, Claire Jenkins

Gastrointestinal Bacteria Reference Unit, Public Health England, London, United Kingdom

Shigella flexneri can be phenotypically serotyped using antisera raised to type-specific somatic antigens and group factor antigens and genotypically serotyped using PCR targeting O-antigen synthesis or modification genes. The aim of this study was to evaluate a real-time PCR for serotyping *S. flexneri* and to use whole-genome sequencing (WGS) to investigate the phenotypic and genotypic serotype identifications. Of the 244 cultures tested retrospectively, 226 (92.6%) had concordant results between phenotypic serotyping and PCR. Seventy of the 244 isolates (including 15 of the 18 isolates where a serotype-PCR mismatch was identified) were whole-genome sequenced, and the serotype was derived from the genome. Discrepant results between the phenotypic and genotypic tests were attributed to insertions/deletions or point mutations identified in O-antigen synthesis or modification genes, rendering them dysfunctional; inconclusive serotyping results due to nonspecific cross-reactions; or novel genotypes. Phylogenetic analysis of the WGS data indicated that the serotype, regardless of whether it was phenotypically or genotypically determined, was a weak predictor of phylogenetic relationships between strains of *S. flexneri*. WGS data provided both genome-derived serotyping, thus supporting backward compatibility with historical data and facilitating data exchange in the community, and more robust and discriminatory typing at the single-nucleotide-polymorphism level.

The genus *Shigella* is divided into four species: *Shigella sonnei*, *S. boydii*, *S. dysenteriae*, and *S. flexneri*. All four species are primarily human pathogens, and there is no animal reservoir, with the exception of nonhuman primates. Transmission is by the fecal-oral route, most commonly by consumption of contaminated food and water or person-to-person contact. All four species of *Shigella* have a low infectious dose (10 to 100 organisms) and cause shigellosis (also known as bacillary dysentery) (1). Symptoms include watery and/or bloody diarrhea, abdominal cramps, fever, and malaise (1). Globally, there are an estimated 164.7 million cases of shigellosis each year, 163.2 million of which occur in low-income countries, with approximately 1.1 million deaths (2). In low-income countries, it is estimated that *S. flexneri* causes 60% of the cases of shigellosis, and it is associated with a higher mortality rate than other *Shigella* species (3, 4). It is estimated that 16% of the 1.5 million cases of shigellosis in middle- to high-income countries are caused by *S. flexneri* (3, 4).

S. flexneri can be divided into serotypes based on lipopolysaccharide (LPS) O-antigen structure. With the exception of serotype 6 (F6), all *S. flexneri* serotypes share a polysaccharide backbone composed of a linear tetrasaccharide of three L-rhamnose (Rha) residues and one N-acetyl-D-glucosamine (GlcNAc) residue (5). This basic polysaccharide structure is referred to as serotype Y, and the different serotypes arise as a result of glucosylation and/or O-acetylation of this polysaccharide backbone at various positions. The genes that confer O-antigen modification are carried on temperate bacteriophages (5). Bacteriophages infect *S. flexneri* strains, and the phage-borne genetic material is integrated into the bacterial chromosome through site-specific recombination, resulting in the addition of the O-antigen modification genes (6). Currently, there are 15 well-established *S. flexneri* serotypes (1a, 1b, 1c, 2a, 2b, 3a, 3b, 4a, 4b, 5a, 5b, X, Xv, Y, and F6), which are composed of type (I, II, III, IV, and V)-, group (3:4, 6, and 7:8)- and 1c-specific antigenic determinants (6). The O-antigen modification that results in the addition of phosphoethanolamine (PEtN) to either Rha II or Rha III, mediated by *opt*, is

found in serotypes 4a variant (4av), X variant (Xv), and Y variant (Yv) (7, 8).

The Gastrointestinal Bacterial Reference Unit (GBRU) at Public Health England (PHE) receives on average approximately 550 cultures of *S. flexneri* each year, most commonly serotypes 2a, 3a, and F6 (PHE, unpublished data). On average, 36% of the cases each year report travel in the week prior to the onset of symptoms (PHE, unpublished data), and 14% report that they have not traveled. Travel history data are unavailable for approximately 50% of cases. Traditionally, *S. flexneri* serotyping was performed by antigen-antibody agglutination reactions using in-house antisera to the LPS O-antigen raised in rabbits or commercially available antisera (9). Recently, a multiplex PCR approach to serotyping *S. flexneri* has been described (6). This PCR assay targets 10 O-antigen synthesis or modification genes: *wzx1-5* and *wzx6* are the *wzx* O-antigen synthesis genes specific to *S. flexneri* serotypes; *gtrI*, *gtrII*, *gtrIV*, *gtrV*, *gtrX*, and *gtr1c* encode serotype-specific glucosyltransferases; *oac* encodes an O-acetyltransferase that mediates the addition of an acetyl group to a specific sugar in the backbone structure; and *opt* encodes a phosphoethanolamine transferase involved in adding a phosphoethanolamine residue to the second L-rhamnose residue of the O-antigen polysaccharide backbone (7).

The aims of this study were to modify the gel-based PCR assay

Received 28 December 2015 Returned for modification 29 January 2016
Accepted 7 March 2016

Accepted manuscript posted online 16 March 2016

Citation Gentle A, Ashton PM, Dallman TJ, Jenkins C. 2016. Evaluation of molecular methods for serotyping *Shigella flexneri*. J Clin Microbiol 54:1456–1461. doi:10.1128/JCM.03386-15.

Editor: D. J. Diekema

Address correspondence to Claire Jenkins, claire.jenkins@phe.gov.uk.

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/JCM.03386-15>.

Copyright © 2016, American Society for Microbiology. All Rights Reserved.

TABLE 1 PCR targets and primers and probe sequences with serotype specificities^a

Multiplex reaction	Target gene	Sequence (5' to 3')	Serotype specificity
1	<i>wzx1-5</i>	GCCATCAGCAGCGAA	1-5, X, Xv, Y
		ACTTCGTACCGCTGGTATC	
		FAM-CAATCGGGTTACCCACGGAGC-BHQ1	
	<i>wzx6</i>	GAGCGATCATTCAACTTCA	F6
		TGTAATTGCGCTTAGTACAACAT	
		JOE-CGGTAATTCTAACTATATTGGGCTTGATTGG-BHQ1	
	<i>gtrV</i>	TTATTAATGTCATCGTCCATCC	5a, 5b
		CCACTCCCAGATTACGG	
		Cy5-GCAGGGTTGAACCTGAAAGAATACGAT-BHQ2	
2	<i>gtrIC</i>	ACCTTAGGTTCAAATGGGTTAC	1c
		GAAATAGCCGTCTCTCGAATA	
		FAM-TGTTTTACATTTAGTATTCCAACAAAACCAC-BHQ1	
	<i>gtrI</i>	F-AAATATGCCTCCATACAATTG	1a, 1b, 1c
		R-AGCATATGTATTAAACAATCAGCA	
	<i>gtrX</i>	Probe JOE-GCTGTAGCAACATCCGGTTCAAC-BHQ	2b, 3a, 5b, X, Xv
		AGCACCATCAAAAATCTTC	
3	<i>gtrIV</i>	CATACAATGATAAATACCAGTGAGCATT	4a, 4b
		Cy5-TATATTTAATTTGCATGCCCCGGGC-BHQ2	
	<i>oac</i>	TCTCATATGATGGCACATTA	1b, 3a, 3b, 4b, 5, 5a, 5b
		CCTAAGATCAAATGTGTGTGTGA	
		FAM-TTTATACCTGAAGGAAAATTTCAATTCAATTGT-BHQ1	
	<i>gtrII</i>	GCATAAGAGCAACTGCTTTG	2a, 2b
		CGCGTAGTGGTGACTG	
		JOE-OACGGCAAGGCTTGTGGCA-BHQ1	
		FCAAACGACTCAGGAAATATGC	
		AATTCATAAATGCAACCATCCT	
		Cy5-CTCCATGAGCGCAGACACTTTTG-BHQ1	

^a Primers and probe sequences for *opt* were not included in the real-time PCR assay.^b FAM, 6-carboxyfluorescein; BHQ1, Black Hole Quencher 1.

described by Sun et al. (6) to a real-time PCR format and to validate the assay by comparing phenotypic- and genotypic-serotyping results. Subsequently, a subset of the cultures used for this validation were whole-genome sequenced. Short-read sequences were mapped to the 10 target genes in the PCR assay, and this whole-genome sequencing (WGS) approach for serotyping *S. flexneri* was evaluated.

MATERIALS AND METHODS

Bacterial strains. Two hundred forty-four cultures of *S. flexneri* submitted to the GBRU between 2004 and 2013 and stored at room temperature on Dorset egg medium were selected from the GBRU archive to provide a comprehensive range of different serotypes. Each isolate was serotyped phenotypically and by real-time PCR following resuscitation from the archive. The strain set comprised the following serotypes (the number of isolates belonging to each serotype is in parentheses): 1a (7), 1b (26), 1c (25), 2a (24), 2b (20), 3a (20), 3b (23), 4a (17), 4av (20), 4b (1), 5a/b (4), F6 (19), X (19), Y (16), and “untypeable” (isolates exhibiting nonspecific agglutination reactions) (3) (see Data Set S1 in the supplemental material). Serotype 4av was previously described as E1037 and 4c (8, 10). Serotypes 4b, 5a, and 5b were underrepresented in the GBRU archive, and all available isolates of the serotypes were included in this study.

Of these 244 cultures, 70 were selected for WGS, i.e., phenotypic serotypes 1a (6), 1b (8), 1c (5), 2a (4), 2b (5), 3a (5), 3b (11), 4a (2), 4av (5), 5a/b (3), X (8), and Y (6) and 2 isolates designated untypeable (see Data Set S1 in the supplemental material). They included 15 of the 18 isolates that had a discrepancy between the phenotypic- and genotypic-testing results and a selection of at least one isolate of each serotype, as defined by phenotypic serotyping.

In addition, all isolates (no selection criteria were applied) submitted to GBRU during a 4-week period between July and August 2015 were prospectively genome sequenced, and the genome-derived serotype was compared to the phenotypic serotype. They included 54 isolates belonging to serotypes 1b (2), 1c (4), 2a (38), 2b (5), 3a (2), 3b (1) and 4av (2). Of the 38 isolates belonging to serotype 2a, 32 were associated with transmission in men who have sex with men (11).

Phenotypic serotyping and real-time PCR. The strains used in this study were serotyped by slide agglutination using both commercially available monovalent antisera (Denka Seiken, Japan) and monoclonal-antibody reagents (Reagensia AB, Sweden) and in-house antisera, raised in rabbits, to all type-specific somatic antigens and the group factor antigens (9). All the strains were tested using a modified real-time version of the PCR serotyping assay described by Sun et al. (6, 7). DNA was extracted by emulsifying a colony in 490 μ l distilled water and boiling for 15 min. The primers and probes designed for the real-time assay are shown in Table 1. Amplification was performed on a Rotorgene Q thermocycler 5-plex HRM (Qiagen, Hilden, Germany), and the amplification parameters were 95°C for 3 min, followed by 95°C for 3 s and 60°C for 30 s for 40 cycles, with the cycle threshold (C_T) set at 0.05. C_T values ranged between 18 and 24.

Whole-genome sequencing. Genomic DNA extracted, using the QiaSymphony DNA extraction platform (Qiagen), from 124 isolates (70 from the retrospective study and 54 from the prospective study) of *S. flexneri* was fragmented and tagged for multiplexing with Nextera XT DNA sample preparation kits (Illumina) and sequenced using the Illumina HiSeq 2500 at PHE. A reference database containing the gene sequences for the 10 O-antigen synthesis or modification genes described by Sun et al. (6, 7), i.e., *wzx1-5*, *wzx6*, *gtrI*, *gtrII*, *gtrIV*, *gtrV*, *gtrX*, *gtrIc*, *oac*,

TABLE 2 Numbers of isolates for each serotype where phenotype and genotype were concordant or discrepant

Phenotypic serotype	No. of isolates		
	Total typed	Concordant phenotype and genotype	Discrepant phenotype and genotype
1a	7	5	2
1b	26	26	0
1c	25	25	0
2a	24	24	0
2b	20	20	0
3a	20	20	0
3b	23	18	5
4a	17	17	0
4b	1	1	1
4av	20	20	0
5a/b	4	1	3
6	19	19	0
X	19	16	3
Y	16	15	1
Untypeable	3	0	3

and *opt*, was constructed. Using the GeneFinder tool (M. Doumith, unpublished data), FASTQ reads were mapped to the *S. flexneri* O-antigen synthesis or modification genes using Bowtie 2 (12), and the best match to each target was reported with metrics, including coverage, depth, mixture, and homology, in an XML format for quality assessment. Only *in silico* predictions of the serotype that matched a gene determinant at >80% nucleotide identity over >80% of the gene length were accepted.

Short reads were quality trimmed and mapped to the reference *S. flexneri* serotype 2a strain 2457T (GenBank accession no. AE014073.1) (13) using BWA v0.75 (14, 15). The sequence alignment map output from BWA was sorted and indexed to produce a binary alignment map (BAM) using Samtools (16). GATK v2.6.5 was used to create a variant call format (VCF) file from each of the BAMs, which were further parsed to extract only single-nucleotide-polymorphism (SNP) positions that were of high

quality (root mean square mapping quality [MQ], >30; read depth [DP], >10; genotype quality [GQ], >30; and variant ratio, >0.9) (17). Pseudo-sequences of polymorphic positions were used to create maximum-likelihood trees using RAxML v8.1.17 (18). *De novo* assembly was carried out with Spades 3.5.0 using “-careful” and “-k 21,33,55,65,77,83,91” options (19).

Nucleotide sequence accession numbers. The FASTQ sequences were deposited in the National Center for Biotechnology Information Short Read Archive, and the accession numbers can be found via BioProject record number PRJNA315192.

RESULTS

Of the 244 cultures tested retrospectively, 226 (92.6%) had concordant results with phenotypic serotyping and PCR (see Data Set S1 in the supplemental material). Seventy of the 244 isolates (including 15 of the 18 isolates where a serotype-PCR mismatch was identified) were whole-genome sequenced, and the serotype was derived from the genome. The WGS data correlated 100% with the PCR data. Serotypes 1b, 1c, 2a, 2b, 3a, and F6 showed 100% concordance using all three methods. The discordant results for serotypes 1a, 3b, 4a, 4b, 4av, and 5a/b are described in detail below (Tables 2 and 3). Prospectively, the phenotypic serotypes and the serotypes derived from the WGS data for an additional 54 isolates were compared and showed 100% concordance.

Ten of the mismatched results between the phenotypic and genotypic tests were attributed to insertions, deletions, or mutations identified in O-antigen synthesis or modification genes unexpectedly present in the genome (Table 3; see Data Set S1 in the supplemental material). Five of these cultures were typed phenotypically as serotype 3b but serotyped as either 1b (1 culture) or 5a (4 cultures) by PCR, most likely due to a deletion in *gtrI* or *gtrV*, respectively, rendering the type antigen modification gene inactive. Phylogenetic analysis demonstrated that the isolate that was serotyped as 1b by PCR (reference number 127) (Table 3 and Fig. 1; see Data Set S1 in the supplemental material) clustered on the same branch of the phylogeny as isolates belonging to serotypes

TABLE 3 Details of insertions, deletions and additional genes derived from WGS analysis and nonspecific phenotypic cross-reactions in isolates with discrepant phenotypic and genotypic serotypes and novel serotypes

Reference no.	Phenotypic serotype	Genotypic serotype	Explanation of discrepancy
6	1a	1b	Insertion in <i>oac</i> 804: + 1T
10	1a	1b	Large insertion in <i>oac</i> position 842 PCR primer binding sites not disrupted; F982-1001, Pr962-979, R929-944
122	1b and 3b ^c	1b	Isolate cross-reacted with both commercial and in-house antisera. Genotypic serotype was 1b.
126	3b	5a + <i>oac</i> ^a	Sequence not available
127	3b	1b	Deletion in <i>gtrI</i> position 1166 (1166:-1A)
128	3b	5a + <i>oac</i> ^a	Deletion in <i>gtrV</i> 646:-1T
129	3b	5a + <i>oac</i> ^a	Deletion in <i>gtrV</i> 646:-1T
133	3b	5a + <i>oac</i> ^a	Deletion in <i>gtrV</i> position 646 (646:-1T) and large indel at position 1039
170	1, 3, 4, 6, X and gp6 ^b	4a	Isolate cross-reacted with commercial and in-house antisera; sequence not available.
172	5b	5b + <i>oac</i> ^a	Sequence not available
175	5 ^b	5a + <i>oac</i> ^a	Deletion in <i>oac</i> (656:-2AT)
180	5 ^b	5a + <i>oac</i> ^a	Deletion in <i>oac</i> (656:-2AT)
181	5 ^b	5a + <i>oac</i> ^a	Deletion in <i>oac</i> (656:-2AT)
203	3, X and 1c ^c	3a	Isolate cross-reacted with both commercial and in-house antisera. Genotypic serotype was 1b.
209	X	3a	Premature stop codon in <i>oac</i> at position 195:S-A; substitution of G for T at position 583.
210	X	3a	Premature stop codon in <i>oac</i> at position 195:S-A; substitution of G for T at position 583.
219	X	3a	Premature stop codon in <i>oac</i> at position 195:S-A
223	Y	1c ^a	Premature stop codon in <i>gtrIC</i>

^a Novel combination of O antigen synthesis and/or modification genes.

^b Isolate reacted phenotypically with both X and Y antisera.

^c Designated “untypeable.”

1a, 1b, and 3b (Fig. 1, box A). The isolates that were serotype 5a by PCR (reference numbers 126, 128, 129, and 133) (Table 3 and Fig. 1, box B; see Data Set S1 in the supplemental material) were phylogenetically closely related to each other and most closely related to isolates belonging to serotype 3a, suggesting these isolates may have lost *gtrX*, as well as gaining a dysfunctional *gtrV*.

The remaining five mismatches were likely caused by the presence of a dysfunctional *oac* (Table 3). The three isolates that were serotyped as serotype X phenotypically but were serotype 3a by PCR all had a premature stop codon in *oac* (reference numbers 209, 210, and 219) (Table 3 and Fig. 1, box C; see Data Set S1 in the supplemental material). Two isolates (reference numbers 6 and 10) (Table 3 and Fig. 1; see Data Set S1 in the supplemental material) were serotype 1b by PCR but were serotyped phenotypically as serotype 1a. Each isolate had an insertion in *oac* that did not disrupt the primer binding sites but might have caused the inactivation of the O-acetyltransferase that mediates the addition of an acetyl group.

There were three isolates whose phenotypic serotypes were recorded as untypeable due to nonspecific cross-reactions and so could not be directly compared to the PCR data (reference numbers 122, 170, and 203) (Table 3; see Data Set S1 in the supplemental material). Nine isolates had novel O-antigen synthesis or modification gene combinations not described by Sun et al. (6). Of these, one isolate was serotyped phenotypically as serotype Y but by PCR had *gtr1c*. This *gtr1c* gene exhibited an early stop codon (reference number 223) (Table 3 and Fig. 1; see Data Set S1 in the supplemental material). Strains belonging to the conventional serotype 1c have both *gtrI* and *gtr1c* (6). This variant clustered phylogenetically with other strains of *S. flexneri* serotype 1c but did not agglutinate with 1c antisera.

The remaining eight isolates had both *gtrV* and *oac*. Four had a deletion in *gtrV* (reference numbers 126, 128, 129, and 133) (Table 3; see Data Set S1 in the supplemental material) and were serotyped phenotypically as 3b, as discussed above (three are shown in Fig. 1, box B). Of the remaining four phenotypic-serotyping results, three were recorded as 5a/b due to nonspecific cross-reactions with both X and Y antisera (shown in Fig. 1, box D) and one was serotyped phenotypically as 5b (reference numbers 172, 175, 180, and 181) (Table 3; see Data Set S1 in the supplemental material). The *S. flexneri* serotype 5a and 5b strains described by Sun et al. (6) did not have *oac*.

FIG 1 Phylogenetic analysis of 69 isolates from the retrospective part of the study (10 isolates of F6 are not included, as the serotype was incorrectly identified to the species level historically; isolates belonging to F6 are not genetically related to *S. flexneri* and do not cluster with other serotypes phylogenetically) and 54 isolates from the prospective part of the study. Serotype 1a, pale green; 1b, yellow; 1c, pale blue; 2a, pink; 2b, red; 3a, brown; 3b, purple; 4a, pale gray; 4av, dark green; 5, turquoise; X, orange; Y, dark blue; untypeable, black. The isolates labeled 6 and 10 were serotyped phenotypically as 1a but genotypically as 1b and have an insertion in *oac*; isolate 127 was serotyped phenotypically as 3b but genotypically as 1b and has an insertion in *gtrI*; isolate 223 was serotyped phenotypically as Y but had a novel combination of O-antigen synthesis/modification genes (*gtr1c* only). Box A is highlighted to show serotype variation in phylogenetically related isolates. Box B highlights three isolates that were typed phenotypically as serotype 3b but were serotype 5a by PCR. Box C shows the three isolates that were serotyped as serotype X phenotypically but were serotype 3a by PCR. Box D comprises three isolates recorded as 5a/b due to nonspecific cross-reactions with both X and Y antisera.

DISCUSSION

Retrospective and prospective comparisons of the phenotypic and genotypic *S. flexneri* serotyping results showed a robust correlation for the common *S. flexneri* serotypes, specifically, 1b, 1c, 2a, 2b, 3a, and F6 (Table 2). The most common cause of the discrepant results between the phenotypic and genotypic data was O-antigen synthesis or modification genes present in the genome but phenotypically inactive. It was not possible to resolve all the mismatches because either the serological cross-reactions observed during the phenotypic-serotyping tests made it impossible to determine the serotype or the PCR identified a novel combination of O-antigen synthesis or modification genes not yet been assigned a universally agreed upon serotype.

The discrepant results were associated with phenotypic serotypes 1a (2), 3b (5), 5a/b (4), X (3), and Y (1) (Table 2) (plus three untypeable isolates). The WGS analysis provided insight into the mismatches by identifying potential mechanisms for the inactivation of the O-antigen modification genes, *gtrI*, *gtrV*, and *oac* (Table 3). These data revealed the effects of insertions, deletions, and mutations on O-antigen expression and highlighted the impact that these genetic events have on comparisons between phenotypic and PCR-based serotyping schemes.

Phylogenetic analysis of the whole-genome sequences of 124 isolates indicated that the serotype, regardless of whether it was phenotypically or genotypically determined, was a weak predictor of phylogenetic relationships between strains of *S. flexneri*. Loss or gain of the O-antigen synthesis or modification genes, carried for the most part on bacteriophages, can result in serotype switching (Fig. 1). For example, loss of *gtrX* or *oac* results in the modification of serotype 3a to serotype 3b or X, respectively. Connor et al. (20) observed evidence of high levels of recombination among the genes responsible for serotypes. When cross comparing the classical and *in silico* molecular-serotyping schemes, they found good correlation between the two methods but also showed that all of the phylogroups in their study contained multiple serotypes (20).

Phenotypic serotyping is expensive, labor-intensive, and difficult to quality control and can result in errors in interpretation due to unresolvable cross-reactions or novel combinations of antigens. The PCR approach to serotyping described by Sun et al. (6), for the most part, resolves these issues. However, while serotyping greatly facilitates detection and investigation of outbreaks that occur locally and have a short duration, serotype switching undermines the principle that the serotype can be used either as a basis for describing the properties of a lineage or for epidemiological surveillance (20). Colleagues in the field have questioned the efficiency of vaccines targeting a particular phylogenetic backbone based on serotyping (20).

The use of WGS data for public health surveillance of *S. flexneri* provides both a serotype, thus supporting backward compatibility with historical data and facilitating data exchange in the community, and a robust and highly discriminatory single-nucleotide-polymorphism-based typing approach to outbreak detection and investigation (PHE, unpublished data). This approach has been used to track a large, globally disseminated outbreak of *S. flexneri* 3a associated with transmission between men who have sex with men (21) and local outbreaks associated with consumption of contaminated food (PHE, unpublished data). Analysis of WGS data also enables us to monitor the emergence of novel serotypes. Describing and establishing novel serotypes using these data are

much less demanding and operationally complex than producing and verifying new antisera for the phenotypic schemes. Furthermore, this approach presents the opportunity to examine the effects of mobile genetic elements, such as prophage and plasmids, on O-antigen expression and evolution.

ACKNOWLEDGMENTS

The concept of the PCR validation part of the study was discussed and developed at an International Shigella meeting in Buenos Aires, Argentina, in May 2012, funded by the Gates Foundation Program for Appropriate Technology in Health (since 2014, known as PATH) and the Pan American Health Organization (PAHO). The meeting was organized and coordinated by Norma Bernstein, Alejandro Cravioto, Mariana Pichel, and Josefina Campos. We are grateful to all our international colleagues who attended the meeting for their insight, expertise, and support, especially Silvina Brengi and Kaisar Ali Talukder. We also thank Vivienne do Nascimento and Dawn Hedges for curation of the serotyping scheme in the GBRU.

This work was supported by the National Institute for Health Research Health Protection Research Unit in Gastrointestinal Infections.

The views expressed are ours and not necessarily those of the NHS, the NIHR, the Department of Health, or Public Health England.

FUNDING INFORMATION

This work, including the efforts of Amy Gentle, Philip Ashton, Timothy J. Dallman, and Claire Jenkins, was funded by DH | National Institute for Health Research (NIHR) (#109524).

REFERENCES

- Niyogi SK. 2005. Shigellosis. *J Microbiol* 43:133–143.
- Kotloff KL, Winickoff JP, Ivanoff B, Clemens JD, Swerdlow DL, Sansonetti PJ, Adak GK, Levine MM. 1999. Global burden of *Shigella* infections: implications for vaccine development and implementation of control strategies. *Bull World Health Organ* 77:651–666.
- Jennison AV, Verma NK. 2004. *Shigella flexneri* infection: pathogenesis and vaccine development. *FEMS Microbiol Rev* 28:43–58. <http://dx.doi.org/10.1016/j.femsre.2003.07.002>.
- Kotloff KL, Nataro JP, Blackwelder WC, Nasrin D, Farag TH, Panchalingam S, Wu Y, Sow SO, Sur D, Breiman RF, Faruque AS, Zaidi AK, Saha D, Alonso PL, Tamboura B, Sanogo D, Onwuchekwa U, Manna B, Ramamurthy T, Kanungo S, Ochieng JB, Omoro R, Oundo JO, Hossain A, Das SK, Ahmed S, Qureshi S, Quadri F, Adegbola RA, Antonio M, Hossain MJ, Akinsola A, Mandomando I, Nhamposha T, Acácio S, Biswas K, O'Reilly CE, Mintz ED, Berkeley LY, Muhsen K, Sommerfelt H, Robins-Browne RM, Levine MM. 2013. Burden and aetiology of diarrhoeal disease in infants and young children in developing countries (the Global Enteric Multicenter Study, GEMS): a prospective, case-control study. *Lancet* 382:209–222. [http://dx.doi.org/10.1016/S0140-6736\(13\)60844-2](http://dx.doi.org/10.1016/S0140-6736(13)60844-2).
- Allison GE, Verma NK. 2000. Serotype-converting bacteriophages and O-antigen modification in *Shigella flexneri*. *Trends Microbiol* 8:17–23. [http://dx.doi.org/10.1016/S0966-842X\(99\)01646-7](http://dx.doi.org/10.1016/S0966-842X(99)01646-7).
- Sun Q, Lan R, Wang Y, Zhao A, Zhang S, Wang J, Wang Y, Xia S, Jin D, Cui Z, Zhao H, Li Z, Ye C, Zhang S, Jing H, Xu J. 2011. Development of a multiplex PCR assay targeting O-antigen modification genes for molecular serotyping of *Shigella flexneri*. *J Clin Microbiol* 49:3766–3770. <http://dx.doi.org/10.1128/JCM.01259-11>.
- Sun Q, Knirel YA, Lan R, Wang J, Senchenkova SN, Jin D, Shashkov AS, Xia S, Perepelov AV, Chen Q, Wang Y, Wang H, Xu J. 2012. A novel plasmid-encoded serotype conversion mechanism through addition of phosphoethanolamine to the O-antigen of *Shigella flexneri*. *PLoS One* 7:e46095. <http://dx.doi.org/10.1371/journal.pone.0046095>.
- Sun Q, Lan R, Wang J, Xia S, Wang Y, Wang Y, Jin D, Yu B, Knirel YA, Xu J. 2013. Identification and characterization of a novel *Shigella flexneri* serotype Yv in China. *PLoS One* 8:e70238. <http://dx.doi.org/10.1371/journal.pone.0070238>.
- Gross RJ, Rowe B. 1985. Serotyping of *Escherichia coli*, p 345–360. In Sussman M (ed), *The virulence of Escherichia coli: reviews and methods*.

- Special publication of the Society for General Microbiology no. 13. Academic Press, London, United Kingdom.
10. Pryamukhina NS, Khomenko NA. 1988. Suggestion to supplement *Shigella flexneri* classification scheme with the subserovar *Shigella flexneri* 4c: phenotypic characteristics of strains. *J Clin Microbiol* 26:1147–1149.
 11. Simms I, Field N, Jenkins C, Childs T, Gilbert VL, Dallman TJ, Mook P, Crook PD, Hughes G. 2015. Intensified shigellosis epidemic associated with sexual transmission in men who have sex with men: *Shigella flexneri* and *S. sonnei* in England, 2004 to end of February 2015. *Euro Surveill* 20:21097. <http://dx.doi.org/10.2807/1560-7917.ES2015.20.15.21097>.
 12. Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. <http://dx.doi.org/10.1038/nmeth.1923>.
 13. Wei J, Goldberg MB, Burland V, Venkatesan MM, Deng W, Fournier G, Mayhew GF, Plunkett G III, Rose DJ, Darling A, Mau B, Perna NT, Payne SM, Runyen-Janecky LJ, Zhou S, Schwartz DC, Blattner FR. 2003. Complete genome sequence and comparative genomics of *Shigella flexneri* serotype 2a strain 2457T. *Infect Immun* 71:2775–2786. <http://dx.doi.org/10.1128/IAI.71.5.2775-2786.2003>.
 14. Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. <http://dx.doi.org/10.1093/bioinformatics/btu170>.
 15. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.
 16. Li H, Durbin R. 2010. Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26:589–595. <http://dx.doi.org/10.1093/bioinformatics/btp698>.
 17. McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, Kernytsky A, Garimella K, Altshuler D, Gabriel S, Daly M, DePristo MA. 2010. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20:1297–1303. <http://dx.doi.org/10.1101/gr.107524.110>.
 18. Stamatakis A. 2014. RAXML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30:1312–1313. <http://dx.doi.org/10.1093/bioinformatics/btu033>.
 19. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 19:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
 20. Connor TR, Barker CR, Baker KS, Weill FX, Talukder KA, Smith AM, Baker S, Gouali M, Pham Thanh D, Jahan Azmi I, Dias da Silveira W, Semmler T, Wieler LH, Jenkins C, Cravioto A, Faruque SM, Parkhill J, Wook Kim D, Keddy KH, Thomson NR. 2015. Species-wide whole genome sequencing reveals historical global spread and recent local persistence in *Shigella flexneri*. *eLife* 4:e07335.
 21. Baker KS, Dallman TJ, Ashton PM, Day M, Hughes G, Crook PD, Gilbert VL, Zittermann S, Allen VG, Howden BP, Tomita T, Valcanis M, Harris SR, Connor TR, Sintchenko V, Howard P, Brown JD, Petty NK, Gouali M, Thanh DP, Keddy KH, Smith AM, Talukder KA, Faruque SM, Parkhill J, Baker S, Weill FX, Jenkins C, Thomson NR. 2015. Intercontinental dissemination of azithromycin-resistant shigellosis through sexual transmission: a cross-sectional study. *Lancet Infect Dis* 15:913–921. [http://dx.doi.org/10.1016/S1473-3099\(15\)00002-X](http://dx.doi.org/10.1016/S1473-3099(15)00002-X).