JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Tau *et al.*, *Journal of Medical Microbiology* 2017;66:937–945

DOI 10.1099/jmm.0.000526

MICROBIOLOGY SOCIETY

Development and evaluation of a multiple-locus variable-number tandem-repeats analysis assay for subtyping *Salmonella* Typhi strains from sub-Saharan Africa

Nomsa P. Tau,[1,2,*] Anthony M. Smith,[1,2] John R. Wain,[3] Andrew Tarupiwa,[4] Kalpy J. Coulibaly,[5] Karen H. Keddy[1,2] for GERMS-SA

Abstract

Purpose. Molecular epidemiological investigations of the highly clonal *Salmonella enterica* subspecies *enterica* serovar Typhi (*S.* Typhi) are important in outbreak detection and in tracking disease transmission. In this study, we developed and evaluated a multiple-locus variable-number tandem-repeats (VNTR) analysis (MLVA) assay for characterization of *S.* Typhi isolates from sub-Saharan Africa.

Methodology. Twelve previously reported VNTR loci were evaluated and an MLVA assay consisting of five polymorphic loci was adopted. The MLVA assay was developed for use on capillary electrophoresis systems by testing a collection of 50 *S.* Typhi isolates. This *S.* Typhi strain panel consisted of six outbreak related isolates and 44 epidemiologically unlinked isolates. Amongst these were nine *S.*Typhi haplotype H58 isolates.

Results. The MLVA assay characterized the 50 isolates into 47 MLVA profiles while PFGE analysis of the same isolates revealed 34 pulsotypes. MLVA displayed higher discriminatory power (Simpson's index of diversity (DI) 0.998 [95 % confidence interval (CI) 0.995–1.000)] as compared to pulsed-field gel electrophoresis [Simpson's DI 0.984 (95 % CI 0.974–0.994)].

Conclusion. The MLVA assay presented in this study is a simple, rapid and more accessible tool that serves as a good alternative to other molecular subtyping methods for *S.* Typhi.

INTRODUCTION

Typhoid fever, a systemic infection caused by *Salmonella enterica* subspecies *enterica serovar* Typhi (*S.* Typhi) remains a significant public health problem worldwide [1, 2]. Recent global estimates indicate that typhoid fever causes approximately 26.9 million illnesses annually [1]. This disease is one of the most important causes of morbidity and mortality in Asia, Africa and Latin America where it is endemic. The spread of typhoid fever is aggravated by poor living conditions, substandard hygiene practices and unsafe drinking water [3–8]. Although *S.* Typhi has proven to be a significant public health problem in Africa [9], little is known about the continental-level molecular epidemiology and strain relatedness of this pathogen. Highly discriminatory molecular subtyping methods, which are accessible to strategic African laboratories, are essential to elucidate the epidemiology of *S.* Typhi;, which, in turn, would allow the implementation of appropriate control strategies in the sub-Saharan Africa (SSA) region; for many of these countries, the true burden of typhoid fever is unknown [1, 10].

Molecular subtyping techniques have enhanced the ability to discriminate bacterial strains [11, 12]. These methods allow for the examination of bacterial strain relatedness at the DNA level and as a result provide a powerful tool for surveillance and outbreak investigations [10–12]. The increasing importance of *S.* Typhi as well as the emergence and

Received 22 November 2016; Accepted 29 May 2017

Author affiliations: [1]Centre for Enteric Diseases, National Institute for Communicable Diseases, a division of the National Health Laboratory Service, Johannesburg, South Africa; [2]Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa; [3]University of East Anglia, Norwich, UK; [4]National Microbiology Reference Laboratory, Harare, Zimbabwe; [5]Unit of Chemistry and Environmental Microbiology, Pasteur Institute, Ivory Coast.

***Correspondence:** Nomsa P. Tau, nomsat@nicd.ac.za

Keywords: *Salmonella* Typhi; sub-Saharan Africa; MLVA; VNTR.

Abbreviations: CED, Centre for Enteric Diseases; CI, confidence interval; DI, diversity index; GERMS-SA, Group for Enteric, Respiratory and Meningeal Disease Surveillance in South Africa; MLST, multi-locus sequence-typing; MLVA, multiple-locus variable-number tandem-repeats analysis; NHLS, National Health Laboratory Services; NICD, National Institute for Communicable Diseases; PFGE, pulsed-field gel electrophoresis; PHE, Public Health England; SSA, sub-Saharan Africa; UPGMA, unweighted pair group method with arithmetic averages; VNTR, variable-number tandem-repeats; WGS, whole-genome sequencing.

dissemination of the multidrug-resistant *S.* Typhi haplotype H58 (*S.* Typhi H58) across SSA has made it imperative to develop new molecular subtyping methods that allow for sensitive strain discrimination [9, 13–15]. Current methodologies used for characterization of *S.* Typhi include multi-locus sequence-typing (MLST), pulsed-field gel electrophoresis (PFGE) and whole-genome sequencing (WGS) [15–17]; amongst these, only PFGE and WGS can discriminate *S.* Typhi strains [15, 18].

PFGE is widely used for subtyping of *S.* Typhi [10, 17]. This involves macro-restriction analysis of bacterial chromosomal DNA and discrimination of isolates based on the resulting banding patterns. The major drawbacks of PFGE are that it is time consuming, which can delay foodborne outbreak identification, and that subjective interpretation of PFGE patterns still remains even with the availability of standardized PulseNet protocols [10]. Furthermore, the discriminatory capacity of this methodology in strain typing *S.* Typhi is not absolute [10, 19]. *S.* Typhi is prone to chromosomal rearrangements, which involve the reshuffling of DNA throughout the genome from one location to another. This may result in alteration of PFGE patterns, which may lead to identification of multiple PFGE patterns from a single outbreak strain [19–21].

In recent years, several approaches have been made in an effort to improve molecular subtyping of *S.* Typhi using multiple-locus variable-number tandem-repeats (VNTR) analysis (MLVA) [22–25]. Even so, till today, a standardized set of VNTR loci for the typing of homologous *S.* Typhi strains has not been established. In this study, we evaluated 12 previously reported VNTR markers for epidemiological investigation of *S.* Typhi strains from SSA. Our aim was to develop an MLVA assay suitable for use on a capillary electrophoresis system that targeted five VNTR markers.

METHODS

Bacterial isolates

The Centre for Enteric Diseases (CED) of the National Institute for Communicable Diseases, a division of the National Health Laboratory Services, serves as a reference centre for human enteric pathogens. *S.* Typhi cultures isolated from all body sites (i.e. normally sterile body sites as well as gastrointestinal sites) in both in-patients and out-patients are submitted to the CED from across the country for national surveillance through the Group for Enteric, Respiratory and Meningeal Disease Surveillance network in South Africa (GERMS-SA). A total of 50 *S.* Typhi isolates from the CED database were used to evaluate the MLVA assay. The selected isolates gave a good representation of diverse collection dates and good representation of geographic areas in South Africa. Amongst the strain collection were four isolates from known outbreaks and 9 *S.* Typhi H58 strains. *S.* Typhi isolates collected from Zimbabwe (n=2) and the Ivory Coast (n=1) were also included.

Pulsed-field gel electrophoresis

PFGE was performed as part of routine surveillance using the standardized PulseNet protocol for *Salmonella*, *Escherichia coli* and *Shigella sonnei* incorporating *Xba*I restriction enzyme for genomic digestion [26]. PFGE fingerprint patterns were analysed and compared using the BioNumerics software (Applied Maths, Belgium) with dendrograms of the fingerprint patterns created using the unweighted pair group method with arithmetic averages (UPGMA), and with analysis of banding patterns incorporating the Dice coefficient at an optimization setting of 1.5 % and a position tolerance setting of 1.5 %.

Bacterial DNA preparation

Bacterial DNA was prepared by inoculating half a loopful of bacterial culture in autoclaved 400 µl Tris-EDTA buffer (10 mM Tris: 1 mM EDTA, pH 8.0) and incubated for 25 min at 95 °C. The supernatant containing DNA was used as a PCR template.

PCR identification of *S.* Typhi H58

S. Typhi H58 strains in the validation panel were identified using previously described PCR [27]. *S.* Typhi NCTC 8385 was used as a positive control. PCR reactions were carried out in 25 µl volume with the forward primer (H58F: 5′-GCAGGCAAAATCGAAATCAG-3′) and reverse primer (H58R: 5′-CAAACCGTTGAATCGGAAGT-3′) at final concentrations of 0.4 µM as well as $MgCl_2$ and deoxynucleotide triphosphate at a final concentration of 1.5 mM and 200 µM, respectively. PCR thermal cycling included an initial denaturation at 94 °C for 2 min, followed by 30 cycles at 94 °C for 30 s, 60 °C for 30 s, 70 °C for 1 min, and a final extension at 72 °C for 5 min. PCR amplicons were analysed on a 1 % electrophoresis gel containing ethidium bromide solution.

MLVA

Selection of VNTR loci and PCR primers

Twelve previously reported VNTR were selected and used as markers to explore their potential in determining the strain relatedness of *S.* Typhi isolates from SSA [22–25].

Screening for length polymorphism of VNTR using simplex PCR

Evaluation of the presence of allelic variation at each VNTR locus in the 50 *S.* Typhi strain panel began with simplex PCRs. *S.* Typhi NCTC 8385 was used as a positive control. All PCR assays were performed using the Qiagen Multiplex PCR Kit. Each 25 µl reaction contained 12.5 µl of the Qiagen Master Mix, 2.5 µl Qiagen Q-solution, 1 µM each of the forward [labelled with a 6-FAM or NED fluorophore (Table 1)] and reverse primers. For PCR amplification, the initial denaturation at 95 °C for 15 min was followed by 25 cycles of a three-step cycle protocol: 94 °C for 60 s, 55 °C for 90 s and 72 °C for 90 s, and a final extension at 72 °C for 10 min. PCR amplicons were diluted 1 : 35 in sterile distilled water and 2 µl aliquots of the dilutions were mixed with 0.2 µl GeneScan 600 LIZ Size Standard v2.0 (Applied Biosystems) and 11 µl Hi-Di formamide solution (Applied Biosystems).

Table 1. Simpson's diversity indices and primers selected for the amplification of 12 *S.* Typhi VNTR

VNTR locus	Number of alleles	Diversity index	Confidence interval	Repeat sequence	VNTR primer sequences*	VNTR primer references
TR2	26	0.940	0.916–0.964	CCAGTTCC	Forward: CCCTGTTTTTCGTGCTGATACG Reverse: CAGAGGATATCGCAACAATCGG	[22]
TR4699	22	0.921	0.892–0.950	TGTTGG	Forward: CGGGCAATTCGAGATAGGTA Reverse: AACCTCCCTGTATCTACCAA	[23]
Sal02	16	0.916	0.896–0.936	TACCAG	Forward: CGATAGACAGCACCAGCAGA Reverse: TCGCCAATACCATGAGTACG	[24]
TR1	11	0.868	0.836–0.900	AGAAGAA	Forward: GCCAACGATCGCTACTTTTT Reverse: CAAGAAGTGCGCATACTACACC	[23, 25]
Sal16	10	0.839	0.805–0.874	ACCATG	Forward: TGCAGTTAATTTCTGCGATCA Reverse: CCTTCCGGATGTATGTGACC	[24]
Sal20	6	0.730	0.657–0.804	CAG	Forward: CAGCCGACACAACTTAACGA Reverse: ACTGTACCGTGCGCGTTT	[23]
TR4500	3	0.607	0.530–0.684	GGACTC	Forward: CGTTGCTGCTCCGAAAT Reverse: GCGGTGAAGTGGAAAAAG	[23]
Sal06	3	0.339	0.185–0.494	CTCAAT	Forward: TTGGTCGCGGAACTATAACTG Reverse: CTTCGTCTGATTGCCACTCC	[25]
TR5	2	0.039	0.000–0.114	CGTCACG	Forward: TGAAAACCGGCTCGTAGCAGTG Reverse: CATACGGTTACTGCGGATTGG	[22]
Sal15	1	0.000	0.000–0.132	No data published	Forward: GTGACCGGTTGAGTTTGCAT Reverse: GGCAGGTTGTACCAGTTCGT	[25]
Sal10	1	0.000	0.000–0.132	ACGCCGCTGCCG	Forward: AAGCGACGTTCTTCTGCAAC Reverse: TGGAATATGATGGCATGACG	[25]
TR4	1	0.000	0.000–0.132	GAAATAAAAAATG	Forward: AAAAGCCCGTCTAGTCTTGCAG Reverse: ATCCTTCGGTATCGGGGTATCC	[22]

*For simplex PCRs forward primers for VNTR locus TR2, TR4699, TR4500, Sal06, Sal15 and Sal10 were labelled with NED fluorophore, while VNTR locus Sal02, TR1, Sal16, Sal20, TR5 and TR4 were labelled with 6-FAM fluorophore. For the multiplex PCR, forward primers for VNTR locus TR1, TR2, Sal02, Sal20 and TR4699 were labelled with PET, 6-FAM, 6-FAM, VIC and NED fluorophores, respectively.

The samples were evaluated by capillary electrophoresis on the Applied Biosystems 3500 Genetic Analyzer (Applied Biosystems) and fragment sizes were analysed using the Gene-Mapper Software (Applied Biosystems). DNA fragments were automatically allocated to length bins and alleles were assigned based on the bin fragment sizes. The VNTR allele numbers were entered into the BioNumerics 6.5 software as character values and a dendrogram was constructed using a categorical coefficient with a 1.5 tolerance and UPGMA.

Diversity and discriminatory power

In order to determine the measure of diversity and the degree of polymorphism at each VNTR locus, Simpson's index of diversity (DI) and 95 % CIs were calculated using an online tool available at the Public Health England (PHE) website (www.hpa-bioinformatics.org.uk/cgi-bin/DICI/DICI.pl).

We used an online tool (www.comparingpartitions.info/) to determine Simpson's DI as well as the Wallace coefficient for the MLVA assay and PFGE analysis. Statistical comparisons of Simpson's DIs and the Wallace coefficients were also performed online.

Nucleotide sequencing

Nucleotide sequencing was performed in order to determine the size of the flanking region at each of the chosen VNTR loci and to also confirm that the variations in the length of amplicons were a result of variation in copy number of tandem repeats. For each VNTR locus, PCR amplicons of ten isolates representing various VNTR amplicon sizes served as template DNA in a PCR cycle sequencing reaction using the BigDye Terminator Cycle Sequencing Kit (Applied Biosystems, Foster City, CA, USA) and an Applied Biosystems 3500 Genetic Analyzer. DNA sequences were analysed using DNASTAR Lasergene (version 8.0) software (DNASTAR, Madison, WI, USA).

Multiplex PCR assay

VNTR loci that had the highest degree of polymorphism were selected for the development of the multiplex PCR assay. Each 25 µl reaction contained 12.5 µl of the Qiagen master mix, 2.5 µl Qiagen Q-solution, 1 µl crude DNA and primer concentrations as shown in Table 2. Forward primers for TR1, TR2, Sal02, Sal20 and TR4699 were labelled with PET, 6-FAM, 6-FAM, VIC and NED, respectively. PCR amplicons were diluted as described above and samples were evaluated by capillary electrophoresis on the Applied Biosystems 3500 genetic analyser (Applied Biosystems) and fragment sizes were analysed using the Gene-Mapper Software (Applied Biosystems).

RESULTS

Evaluation of VNTR loci polymorphism

We evaluated a total of 12 VNTR loci using a panel of 50 *S.* Typhi isolates from SSA. The primer sets for all 12 VNTR loci were able to produce PCR amplicons in all isolates at an annealing temperature of 55 °C. Of the 12

Table 2. Features of seven highly polymorphic *S.* Typhi VNTR loci

VNTR locus	Gene	Product	Repeat sequence	Unit length	Forward offset (bp)	Reverse offset (bp)	MLVA multiplex PCR primer concentrations	References
TR1	–	Intergenic region between *yedD* and *yedE*	AGAAGAA	7	39	116	1.2 µM	[22, 23]
TR2	–	Intergenic region between *arcD* and *yffB*	CCAGTTCC	8	191	105	1.2 µM	[22, 23]
TR4699	sefC	Outer membrane fimbral usher protein	TGTTGG	6	38	137	0.8 µM	[23]
Sal02	citT	Citrate carrier	TACCAG	6	136	59	1 µM	[23, 25]
Sal16	–	Intergenic region between *STY3169* (pseudogene) and *STY3172*	ACCATG	6	90	91, 97, 103, 109, 115, 127	–	[23]
Sal20	ftsN	Cell division protein	CAG	3	83	80	0.5 µM	[23, 25]
TR4500	STY4635	Hypothetical protein	GGACTC	6	76	195, 202	–	[23]

potential VNTR loci evaluated, seven (TR1, TR2, TR4500, TR4699, Sal02, Sal16 and Sal20) demonstrated the ability to discriminate *S.* Typhi strains. Simpson's DI for these loci ranged from 0.607 to 0.940 (Table 1). For the remaining five VNTR loci, two (TR5 and Sal06) demonstrated decreased ability to discriminate *S.* Typhi strains with lower diversity indices of 0.039 and 0.339 respectively, while the other three VNTR loci (Sal15, Sal10 and TR4) were homogenous and showed no variation amongst the strains.

MLVA validation by nucleotide sequencing

The seven most variable VNTR loci with high Simpson's DI were selected for nucleotide sequencing. VNTR loci with 100 % conserved flanking sequences were selected for inclusion in our MLVA assay. The flanking sequences in all the five VNTR loci, including TR1, TR2, Sal02, Sal20 and TR4699, were conserved and polymorphism was due to variation in the number of tandem-repeat units within these loci (Table 2). Nucleotide sequencing revealed that the flanking sequences at VNTR locus Sal16 were not conserved. Aside from the six-base-pair tandem-repeat sequence 'ACCATG' at VNTR locus Sal16, an additional 12-base-pair repeat sequence 'ACCACCATCACG' was identified. Therefore, polymorphism at this VNTR locus was due to variation at both tandem-repeat regions. Nucleotide sequencing also revealed that the flanking sequences at VNTR locus TR4500 were not conserved. A seven-base-pair insertion sequence 'TTGCCAC' was identified in six of the ten *S.* Typhi isolates that were subjected to nucleotide sequencing. The correct number of repeat units for each strain could not be accurately determined as not all strains harboured the seven-base-pair sequence. Redesigning primers would not have resolved the problem as the insertion sequence is located adjacent to the tandem-repeat region. For this reason, VNTR loci Sal16 and TR4500 were excluded from the MLVA assay.

Optimization of the MLVA multiplex PCR assay consisting of five VNTR loci

We developed an MLVA assay consisting of five highly polymorphic VNTR loci including VNTR loci TR1, TR2, Sal02, Sal20 and TR4699. Three of these VNTR loci (TR4699, Sal02 and Sal20) were located in gene regions and the other two VNTR loci (TR1 and TR2) were intergenic (Table 2). The five polymorphic VNTR loci were pooled into one PCR reaction for the development of the MLVA multiplex PCR. Capillary electrophoresis was used for the analysis of the MLVA assay in order to correctly determine the VNTR loci amplicon sizes. For easy identification of PCR amplicons, forward primers for each VNTR locus were labelled with fluorophores. The forward primers for VNTR loci TR1, TR2, Sal02, Sal20 and TR4699 were labelled with PET, 6-FAM, 6-FAM, VIC and NED, respectively (Fig. 1). Due to only four fluorophores being available for selection, this resulted in labelling of two VNTR loci (namely TR2 and Sal02) with the same fluorophore (6-FAM). PCR amplicons from these two VNTR loci could still be easily identified as their amplicon sizes did not overlap.

Discriminatory power of MLVA and PFGE analysis of *S.* Typhi isolates from SSA

Overall, PFGE distinguished the 50 *S.* Typhi isolates into 34 unique pulsotypes (Fig. 2). PFGE clusters were defined at 80 % similarity threshold. PFGE analysis revealed four clusters, with the largest cluster (cluster b) consisting of 34 isolates (Fig. 2). PFGE showed the ability to discriminate *S.* Typhi H58 isolates. The nine *S.* Typhi H58 strains were differentiated into seven pulsotypes. Unique pulsotypes were identified for isolates belonging to the Delmas, Mpumalanga 2005 outbreak and the Pretoria, Gauteng 2010 outbreak.

The MLVA assay characterized the 50 isolates into 47 unique MLVA profiles (Fig. 3). Only three clusters, consisting of two or more isolates with indistinguishable MLVA profiles, were identified. These clusters consisted of isolates from recent outbreaks in South Africa namely, the Delmas, Mpumalanga 2005 outbreak and the Pretoria, Gauteng 2010 outbreak. The MLVA assay was able to discriminate amongst *S.* Typhi H58 strains. The nine *S.* Typhi H58 strains in the validation panel were differentiated into nine MLVA profiles. MLVA concurred with PFGE analysis in clustering of outbreak isolates.

The discriminatory power of the MLVA assay was calculated using Simpson's DI applied to the *S.* Typhi strain panel. Simpson's DI does not only depend on the number of

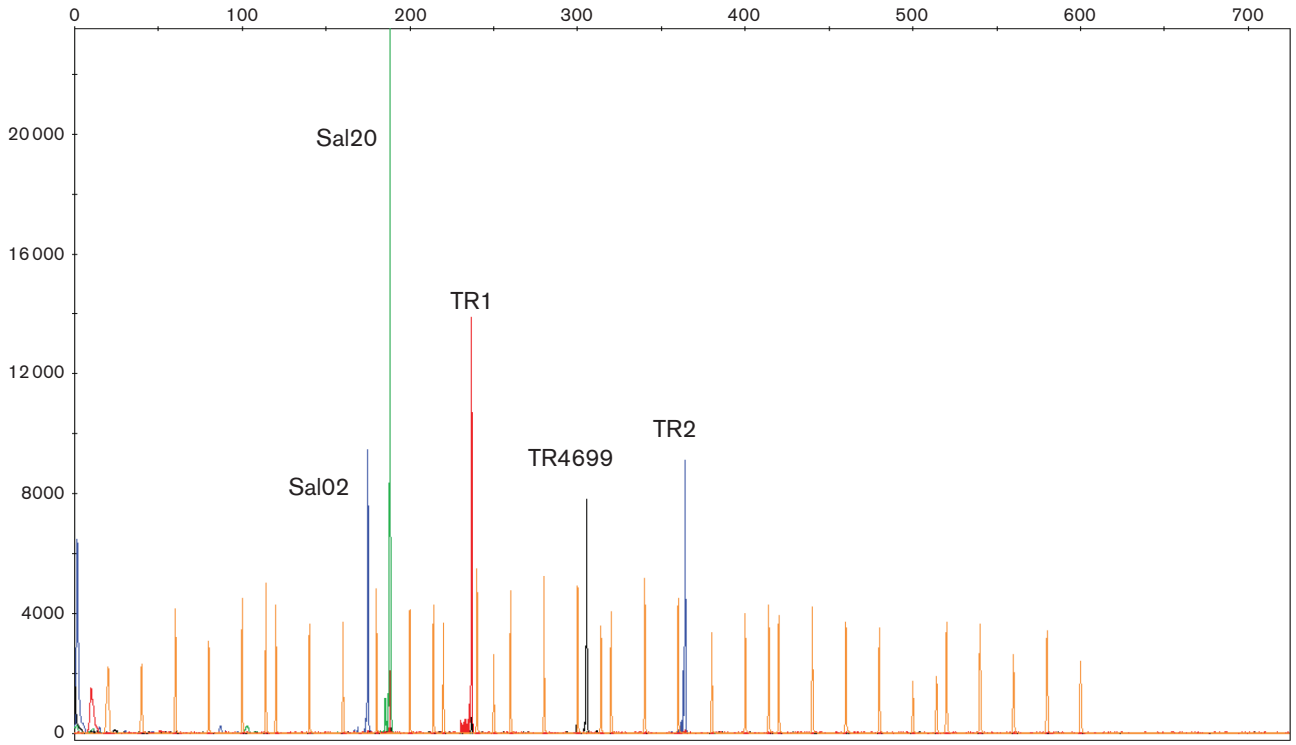


Fig. 1. Electropherogram showing PCR fragments of all five VNTR loci incorporated in a multiplex PCR for MLVA analysis of *S.* Typhi isolates. The fragments were resolved by capillary electrophoresis. VNTR loci TR1, TR2, Sal02, Sal20 and TR4699 were labelled with PET (red), 6-FAM (blue), 6-FAM (blue), VIC (green) and NED (black) fluorophores, respectively.

alleles present at each locus but it also takes into consideration the equitability with which the alleles are distributed at each locus [28, 29]. VNTR loci with a Simpson's DI value closer to 1 are better markers to differentiate the strains for epidemiological purposes. For the MLVA assay, Simpson's DI was calculated at 0.998 (95 % CI 0.995–1.000). Simpson's DI for PFGE analysis of these isolates was calculated at 0.984 (95 % CI 0.974–0.994). The difference in Simpson's DI for MLVA and PFGE were statistically significant (P=0.010).

The congruence between the MLVA assay and PFGE analysis was determined by calculating the Wallace coefficient. This coefficient indicates the probability that two isolates that cluster together by one subtyping method could also be clustered together using another subtyping method [30]. The Wallace coefficient between the MLVA assay and PFGE pulsotypes was 67 %. In contrast, the probability that two isolates with the same MLVA profile could have the same pulsotypes was 9 %. The MLVA assay showed a higher discriminatory power than PFGE analysis.

DISCUSSION

Epidemiological investigations are important for the control of the dissemination of typhoid fever. The ability to study the incidence and spread of *S.* Typhi in SSA relies on the selection and use of suitable and rapid molecular methods that are accessible to the many laboratories in this region. In order to study the dissemination of *S.* Typhi in SSA, highly discriminatory molecular methods are required for characterization of this pathogen. PFGE has been widely used for subtyping *S.* Typhi isolates; however, the suboptimal discriminatory power of this molecular method coupled with the intensive labour involved makes PFGE unsuitable. WGS has become the most commonly used molecular subtyping tool for human pathogens; however, the cost involved in using this molecular method for routine surveillance of *S.* Typhi infections in the developing countries of the SSA region makes WGS unfeasible, at present.

In our study, we evaluated 12 VNTR loci and used them as molecular markers to discriminate amongst *S.* Typhi isolates from SSA. VNTR loci with shorter repeat sequences (≤8 bp in length) were included for selection, as these show more variation in copy number. In addition, VNTR loci and primers that harboured perfect homogenous repeat sequences and had 100 % conserved flanking sequences with no insertions or deletions in repeat sequences were selected. Five of the 12 VNTR loci evaluated in this study (TR1, TR2, Sal02, Sal20 and TR4699) proved to be polymorphic and showed the ability to discriminate *S.* Typhi isolates. VNTR loci TR1 and TR2 were first tested by Liu *et al.* [22] and were found to be highly

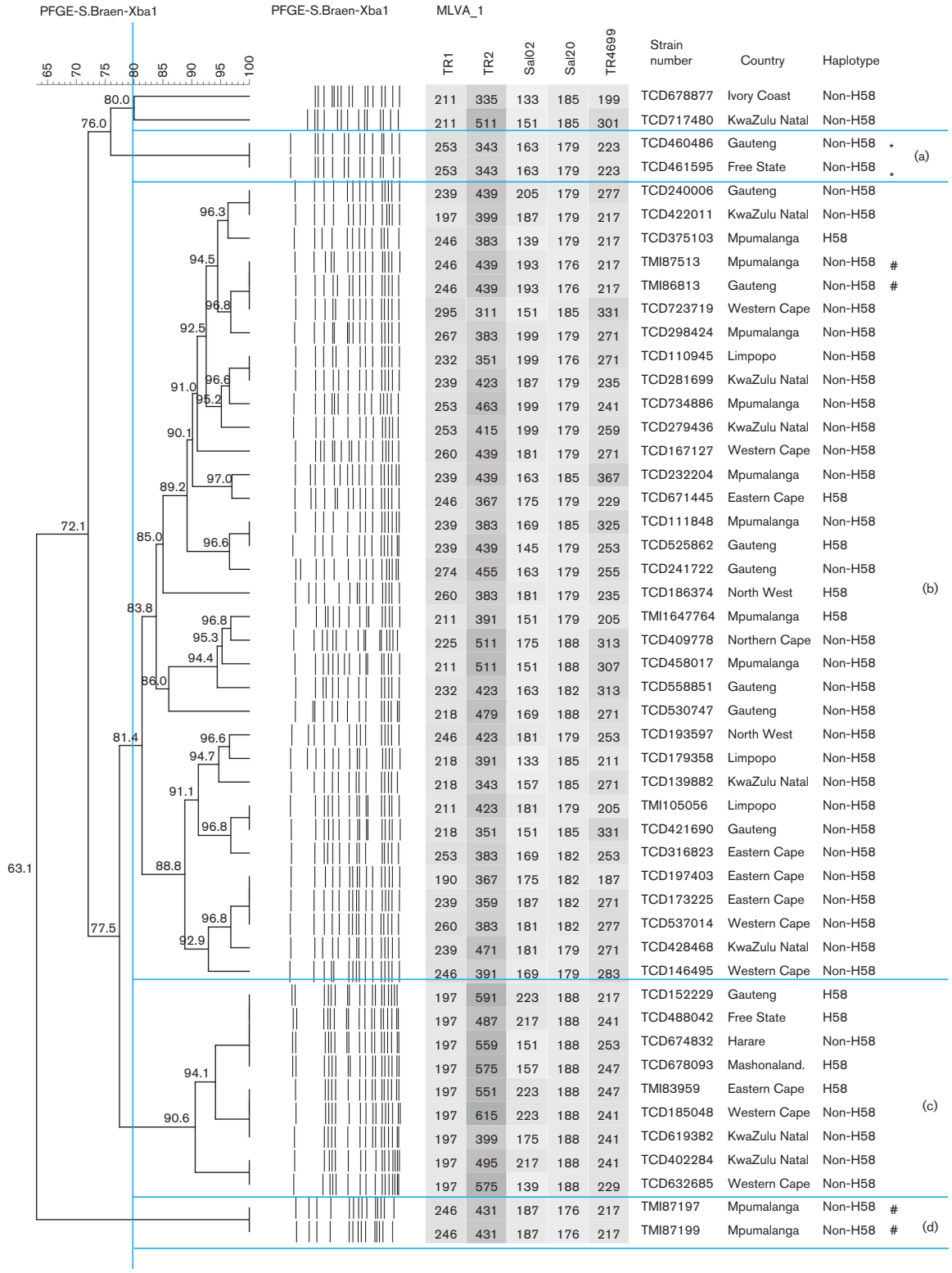
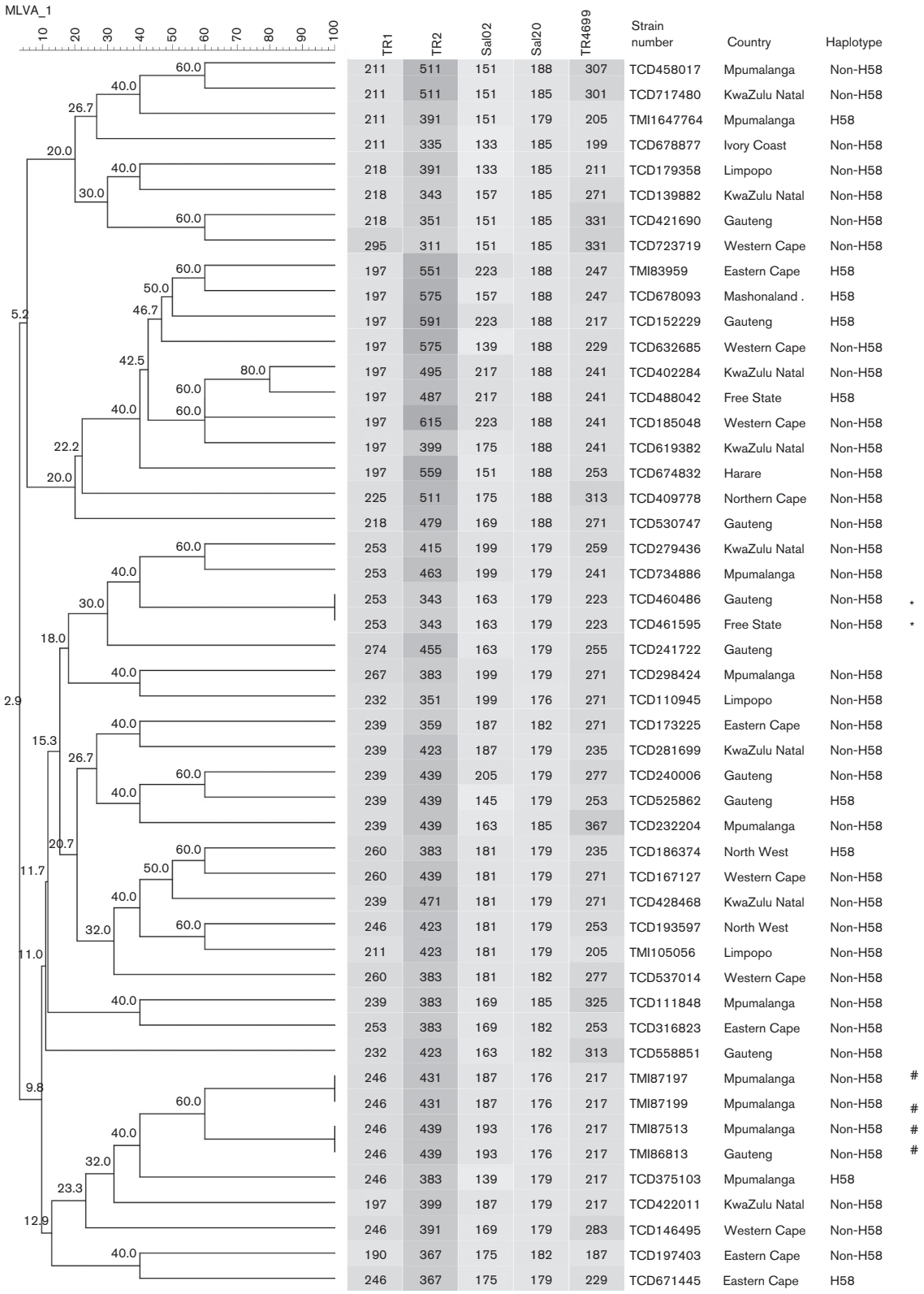


Fig. 2. Dendrogram analysis of PFGE patterns (*Xba*I digestion) showing genetic relationships amongst *S.* Typhi isolates from SSA. Percentage relatedness is indicated at nodes. Dendrogram analysis revealed 34 pulsotypes amongst 50 *S.* Typhi isolates. # denotes Mpumalanga, 2005 outbreak isolates and * denotes Gauteng, 2010 outbreak isolates. The blue blocks denote strains belonging to the same PFGE cluster at 80 % similarity threshold (clusters a–d). MLVA profiles are provided for direct comparison.



TR1	TR2	Sal02	Sal20	TR4699	Strain number	Country	Haplotype	
211	511	151	188	307	TCD458017	Mpumalanga	Non-H58	
211	511	151	185	301	TCD717480	KwaZulu Natal	Non-H58	
211	391	151	179	205	TMI1647764	Mpumalanga	H58	
211	335	133	185	199	TCD678877	Ivory Coast	Non-H58	
218	391	133	185	211	TCD179358	Limpopo	Non-H58	
218	343	157	185	271	TCD139882	KwaZulu Natal	Non-H58	
218	351	151	185	331	TCD421690	Gauteng	Non-H58	
295	311	151	185	331	TCD723719	Western Cape	Non-H58	
197	551	223	188	247	TMI83959	Eastern Cape	H58	
197	575	157	188	247	TCD678093	Mashonaland .	H58	
197	591	223	188	217	TCD152229	Gauteng	H58	
197	575	139	188	229	TCD632685	Western Cape	Non-H58	
197	495	217	188	241	TCD402284	KwaZulu Natal	Non-H58	
197	487	217	188	241	TCD488042	Free State	H58	
197	615	223	188	241	TCD185048	Western Cape	Non-H58	
197	399	175	188	241	TCD619382	KwaZulu Natal	Non-H58	
197	559	151	188	253	TCD674832	Harare	Non-H58	
225	511	175	188	313	TCD409778	Northern Cape	Non-H58	
218	479	169	188	271	TCD530747	Gauteng	Non-H58	
253	415	199	179	259	TCD279436	KwaZulu Natal	Non-H58	
253	463	199	179	241	TCD734886	Mpumalanga	Non-H58	
253	343	163	179	223	TCD460486	Gauteng	Non-H58	*
253	343	163	179	223	TCD461595	Free State	Non-H58	*
274	455	163	179	255	TCD241722	Gauteng		
267	383	199	179	271	TCD298424	Mpumalanga	Non-H58	
232	351	199	176	271	TCD110945	Limpopo	Non-H58	
239	359	187	182	271	TCD173225	Eastern Cape	Non-H58	
239	423	187	179	235	TCD281699	KwaZulu Natal	Non-H58	
239	439	205	179	277	TCD240006	Gauteng	Non-H58	
239	439	145	179	253	TCD525862	Gauteng	H58	
239	439	163	185	367	TCD232204	Mpumalanga	Non-H58	
260	383	181	179	235	TCD186374	North West	H58	
260	439	181	179	271	TCD167127	Western Cape	Non-H58	
239	471	181	179	271	TCD428468	KwaZulu Natal	Non-H58	
246	423	181	179	253	TCD193597	North West	Non-H58	
211	423	181	179	205	TMI105056	Limpopo	Non-H58	
260	383	181	182	277	TCD537014	Western Cape	Non-H58	
239	383	169	185	325	TCD111848	Mpumalanga	Non-H58	
253	383	169	182	253	TCD316823	Eastern Cape	Non-H58	
232	423	163	182	313	TCD558851	Gauteng	Non-H58	
246	431	187	176	217	TMI87197	Mpumalanga	Non-H58	#
246	431	187	176	217	TMI87199	Mpumalanga	Non-H58	#
246	439	193	176	217	TMI87513	Mpumalanga	Non-H58	#
246	439	193	176	217	TMI86813	Gauteng	Non-H58	#
246	383	139	179	217	TCD375103	Mpumalanga	H58	
197	399	187	179	217	TCD422011	KwaZulu Natal	Non-H58	
246	391	169	179	283	TCD146495	Western Cape	Non-H58	
190	367	175	182	187	TCD197403	Eastern Cape	Non-H58	
246	367	175	179	229	TCD671445	Eastern Cape	H58	

Fig. 3. Dendrogram based on MLVA profiles of 50 *S.* Typhi isolates. Percentage relatedness is indicated at nodes. The dendrogram was generated using categorical coefficient with a 1.5 tolerance and UPGMA clustering. A total of 47 unique MLVA profiles were identified amongst 50 *S.* Typhi isolates. # denotes Mpumalanga, 2005 outbreak isolates and * denotes Gauteng, 2010 outbreak isolates.

discriminatory in typing 59 *S.* Typhi isolates from Asia. These VNTR loci were also explored in several other MLVA assays and were found to be highly polymorphic, exhibiting diversity indices ranging from 0.87 to 0.90 [23–25]. Similarly, in our study, VNTR loci TR1 and TR2 were highly polymorphic and exhibited Simpson's DI calculated at 0.87 and 0.94, respectively. VNTR loci Sal02 and Sal20 were first explored by Ramisse *et al.* [25]. In their study, they identified VNTR loci Sal02 and Sal20 as two of the most variable VNTR loci exhibiting diversity indices calculated at 0.87 and 0.81, respectively [25]. Other studies have also shown the ability of VNTR loci Sal02 and Sal20 in discriminating *S.* Typhi isolates [23, 24]. In our study, VNTR loci Sal02 and Sal20 were highly polymorphic and exhibited Simpson's DI calculated at 0.92 and 0.73, respectively. VNTR locus TR4699 was first described by Octavia and Lan [23]. This VNTR locus was identified as one of the highly polymorphic loci in differentiating 73 global *S.* Typhi isolates and exhibited a diversity index calculated at 0.95. Tien *et al.* [24] also explored the use of VNTR locus TR4699 in an MLVA assay and found this locus to be highly polymorphic, exhibiting a diversity index of 0.92. The findings in our study concur with those of previous studies as VNTR locus TR4699 exhibited Simpson's DI calculated at 0.92. The high Simpson's DI values calculated at each locus indicated that the selected loci are of highly polymorphic nature and have greater discriminatory power sufficient to differentiate epidemiologically unrelated strains.

A multiplex PCR assay containing primers targeting the five highly polymorphic VNTR loci was developed and proved to be rapid and highly discriminatory in characterizing *S.* Typhi isolates from SSA. Simpson's DI for the MLVA assay was calculated at 0.998, indicating high discriminatory abilities. The combination of the five VNTR loci showing high-diversity levels enabled differentiation of closely related and unrelated *S.* Typhi isolates from SSA. MLVA showed higher discrimination of *S.* Typhi (including *S.* Typhi H58) and would be an effective molecular subtyping tool for epidemiological investigation of *S.* Typhi outbreaks in SSA.

Conclusion

Typhoid fever surveillance is of crucial importance in the SSA region where there is a high burden of disease. MLVA has a great ability to differentiate *S.* Typhi strains from SSA. Being PCR-based, MLVA can be used in many laboratories in developing countries which may not have access to WGS.

Our study describes an MLVA assay consisting of five VNTR markers for analysis of *S.* Typhi strains from SSA. This assay can assist in rapid identification of *S.* Typhi strain relatedness as well as typhoid fever outbreak detection. To the best of our knowledge, capillary electrophoresis equipment is available in four other laboratories in the SSA region to which the MLVA assay can be adopted. These include Kenya, Gambia, the Ivory Coast and Uganda. The implementation of MLVA analysis in these laboratories could be helpful in monitoring the spread of typhoid fever across the continent and highlight the role of specific clones in disease causation. The MLVA assay could also assist in identifying the role of imported cases of typhoid fever into South Africa and their contribution to the burden of disease. This will promote effective and appropriate disease intervention strategies, including prevention and treatment.

Funding information
This work was supported in part by funding allocated through a Co-operative Agreement from the Global Disease Detection Co-operative agreement with the NHLS, grant 1U19GH000571-02. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of NICD/NHLS or CDC. This study was undertaken as part of a Master of Science (MSc) project by Nomsa P. Tau.

Acknowledgements
The authors would like to thank their colleagues in Africa for submitting *S.* Typhi strains for inclusion in this work. We would also like to thank our partners in the GERMS-SA surveillance programme for submission of isolates to the Centre for Enteric Diseases (CED). GERMS-SA members include Vanessa Quan – lead author (vanessaq@nicd.ac.za), Ananta Nanoo, Anne von Gottberg, Andries Dreyer, Anthony Smith, Arvinda Sooka, Cecilia Miller, Charlotte Sriruttan, Cheryl Cohen, Chikwe Ihekweazu, Claire von Mollendorf, Frans Radebe, Genevie Ntshoe, Gillian Hunt, Karen Keddy, Linda de Gouveia, Linda Erasmus, Marshagne Smith, Martha Bodiba, Mbhekiseni Khumalo, Motshabi Modise, Nazir Ismail, Nelesh Govender, Nicola Page, Olga Perovic, Oliver Murangandi, Penny Crowther-Gibson, Portia Mutevedzi, Riyadh Manesen, Ruth Mpembe, Samantha Iyaloo, Sarona Lengana, Shabir Madhi, Sibongile Walaza, Sonwabo Lindani, Susan Meiring, Thejane Motladiile, Verushka Chetty (NICD); Carel Haumann, Patricia Hanise; Sandeep Vasaikar, John Black, Vanessa Pearce (Eastern Cape); Anwar Hoosen, Vicky Kleinhans (Free State); Alan Karstaedt, Caroline Maluleka, Charl Verwey, Charles Feldman, David Moore, David Spencer, Gary Reubenson, Khine Swe Swe Han, Jeannette Wadula, Jeremy Nel, Kathy Lindeque, Maphoshane Nchabeleng, Nicolette du Plessis, Norma Bosman, Ranmini Kularatne, Ruth Lekalakala, Sharona Seetharam, Theunis Avenant, Trusha Nana, Vindana Chibabhai (Gauteng); Adhil Maharj, Asmeeta Burra, Fathima Naby, Halima Dawood, Koleka Mlisana, Lisha Sookan, Praksha Ramjathan, Prasha Mahabeer, Romola Naidoo, Sumayya Haffejee, Yacoob Coovadia (Kwa-Zulu Natal); Ken Hamese, Ngoaka Sibiya (Limpopo); Greta Hoyland, Jacob Lebudi (Mpumalanga); Eunice Weenink; Riezaah Abrahams, Sindiswa Makate (Northern Cape); Ebrahim Variava, Erna du Plessis (North West); Andrew Whitelaw, Catherine Samuel, Mark Nicol, Preneshni Naicker, Shareef Abrahams (Western Cape); Adrian Brink, Elizabeth Prentice, Inge Zietsman, Maria Botha, Peter Smith, Xoliswa Poswa (AMPATH); Chetna Govind, Keshree Pillay, Suzy Budavari (LANCET); Catherine Samuel, Marthinus Senekal (PathCare); Cynthia Whitney (CDC); Keith Klugman (Emory). Special thanks go to the team at the Gastrointestinal Bacteria Reference Unit of Public Health, England for training in performing MLVA analysis.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References

1. **Buckle GC, Walker CL, Black RE.** Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010. *J Glob Health* 2012;2:1–9.
2. **Connor BA, Schwartz E.** Typhoid and paratyphoid fever in travellers. *Lancet Infect Dis* 2005;5:623–628.
3. **Feasey NA, Archer BN, Heyderman RS, Sooka A, Dennis B *et al.*** Typhoid fever and invasive nontyphoid salmonellosis, Malawi and South Africa. *Emerg Infect Dis* 2010;16:1448–1451.
4. **Kanungo S, Dutta S, Sur D.** Epidemiology of typhoid and paratyphoid fever in India. *J Infect Dev Ctries* 2008;2:454–460.

5. **Mweu E, English M.** Typhoid fever in children in Africa. *Trop Med Int Health* 2008;13:532–540.
6. **Beyene G, Asrat D, Mengistu Y, Aseffa A, Wain J.** Typhoid fever in Ethiopia. *J Infect Dev Ctries* 2008;2:448–453.
7. **Kariuki S.** Typhoid fever in sub-Saharan Africa: challenges of diagnosis and management of infections. *J Infect Dev Ctries* 2008; 2:443–447.
8. **Crump JA, Mintz ED.** Global trends in typhoid and paratyphoid fever. *Clin Infect Dis* 2010;50:241–246.
9. **Wain J, Hendriksen RS, Mikoleit ML, Keddy KH, Ochiai RL.** Typhoid fever. *Lancet* 2015;385:1136–1145.
10. **Wattiau P, Boland C, Bertrand S.** Methodologies for *Salmonella enterica* subsp. *enterica* subtyping: gold standards and alternatives. *Appl Environ Microbiol* 2011;77:7877–7885.
11. **Sabat AJ, Budimir A, Nashev D, Sá-Leão R, van Dijl J *et al.*** Overview of molecular typing methods for outbreak detection and epidemiological surveillance. *Euro Surveill* 2013;18:20380.
12. **Baker S, Hanage WP, Holt KE.** Navigating the future of bacterial molecular epidemiology. *Curr Opin Microbiol* 2010;13:640–645.
13. **Lunguya O, Lejon V, Phoba MF, Bertrand S, Vanhoof R *et al.*** *Salmonella* Typhi in the democratic republic of the congo: fluoroquinolone decreased susceptibility on the rise. *PLoS Negl Trop Dis* 2012;6:e1921.
14. **Kariuki S, Revathi G, Kiiru J, Mengo DM, Mwituria J *et al.*** Typhoid in Kenya is associated with a dominant multidrug-resistant *Salmonella enterica* serovar Typhi haplotype that is also widespread in Southeast Asia. *J Clin Microbiol* 2010;48:2171–2176.
15. **Wong VK, Baker S, Pickard DJ, Parkhill J, Page AJ *et al.*** Phylogeographical analysis of the dominant multidrug-resistant H58 clade of *Salmonella* Typhi identifies inter- and intracontinental transmission events. *Nat Genet* 2015;47:632–639.
16. **Kidgell C, Reichard U, Wain J, Linz B, Torpdahl M *et al.*** *Salmonella* Typhi, the causative agent of typhoid fever, is approximately 50,000 years old. *Infect Genet Evol* 2002;2:39–45.
17. **Thong KL, Cheong YM, Puthucheary S, Koh CL, Pang T.** Epidemiologic analysis of sporadic *Salmonella* Typhi isolates and those from outbreaks by pulsed-field gel electrophoresis. *J Clin Microbiol* 1994;32:1135–1141.
18. **Fakhr MK, Nolan LK, Logue CM.** Multilocus sequence typing lacks the discriminatory ability of pulsed-field gel electrophoresis for typing *Salmonella enterica* serovar Typhimurium. *J Clin Microbiol* 2005;43:2215–2219.
19. **Echeita MA, Usera MA.** Chromosomal rearrangements in *Salmonella enterica* serotype Typhi affecting molecular typing in outbreak investigations. *J Clin Microbiol* 1998;36:2123–2126.
20. **Liu SL, Sanderson KE.** Highly plastic chromosomal organization in *Salmonella* Typhi. *Proc Natl Acad Sci USA* 1996;93:10303–10308.
21. **Matthews TD, Rabsch W, Maloy S.** Chromosomal rearrangements in *Salmonella enterica* serovar Typhi strains isolated from asymptomatic human carriers. *MBio* 2011;2:e00060-11.
22. **Liu Y, Lee MA, Ooi EE, Mavis Y, Tan AL *et al.*** Molecular typing of *Salmonella enterica* serovar Typhi isolates from various countries in Asia by a multiplex PCR assay on variable-number tandem repeats. *J Clin Microbiol* 2003;41:4388–4394.
23. **Octavia S, Lan R.** Multiple-locus variable-number tandem-repeat analysis of *Salmonella enterica* serovar Typhi. *J Clin Microbiol* 2009;47:2369–2376.
24. **Tien YY, Ushijima H, Mizuguchi M, Liang SY, Chiou CS.** Use of multilocus variable-number tandem repeat analysis in molecular subtyping of *Salmonella* enterica serovar Typhi isolates. *J Med Microbiol* 2012;61:223–232.
25. **Ramisse V, Houssu P, Hernandez E, Denoeud F, Hilaire V *et al.*** Variable number of tandem repeats in *Salmonella enterica* subsp. enterica for typing purposes. *J Clin Microbiol* 2004;42: 5722–5730.
26. **Ribot EM, Fair MA, Gautom R, Cameron DN, Hunter SB *et al.*** Standardization of pulsed-field gel electrophoresis protocols for the subtyping of *Escherichia coli* O157:H7, *Salmonella*, and *Shigella* for PulseNet. *Foodborne Pathog Dis* 2006;3:59–67.
27. **Murgia M, Rubino S, Wain J, Gaind R, Paglietti B.** A novel broadly applicable PCR-RFLP method for rapid identification and subtyping of H58 *Salmonella* Typhi. *J Microbiol Methods* 2016;127:219–223.
28. **Hunter PR, Gaston MA.** Numerical index of the discriminatory ability of typing systems: an application of Simpson's index of diversity. *J Clin Microbiol* 1988;26:2465–2466.
29. **Simpson EH.** Measurement of diversity. *Nature* 1949;163:688.
30. **Wallace DL.** A method for comparing two hierarchical clusterings: comment. *J Am Stat Assoc* 1983;78:569–576.