Multiple locus variable number of tandem repeat analysis of *Salmonella enterica* serovar Typhi

Sophie Octavia and Ruiting Lan*

School of Biotechnology and Biomolecular Sciences, University of New South Wales, Sydney, NSW 2052, Australia

*For correspondence
Email r.lan@unsw.edu.au
Phone 61-2 9385 2095
Fax 61-2 9385 1591

Running title: MLVA for typing of global Typhi isolates
Abstract

Multilocus variable number of tandem repeats (VNTRs) analysis (MLVA) is widely used as molecular markers to differentiate isolates of homogenous pathogenic clones. We explored the genomes of *Salmonella enterica* serovar Typhi strains CT18 and Ty2 for potential VNTRs. Among the forty-three identified, two were found to be polymorphic. Together with seven polymorphic VNTRs from previous studies, they were used to type 73 global Typhi isolates. A total of 70 MLVA profiles were found, distinguishing all except three pairs of isolates into individual profiles. The discriminatory power was 0.999. Phylogenetic analysis showed that the MLVA profiles can be divided into seven clusters. However, except for the closely related isolates the relationships derived were in conflict with that inferred from Single Nucleotide Polymorphism (SNP) typing using 38 SNPs done previously. We concluded that MLVA can only resolve the relationships among closely related isolates. A combination of SNP typing and MLVA typing offers the best approach for local and global epidemiology and the evolutionary analysis of Typhi. We suggest seven of the nine most polymorphic VNTRs be used as a standardized typing scheme for epidemiological typing.
Introduction

Typhoid fever remains as a devastating disease in developing countries, and is endemic in areas with inadequate sanitation and poor hygiene. It is a serious systemic disease, spread via faecal-oral route. Annually, there are more than 16 million cases of typhoid fever with 600,000 deaths reported worldwide (www.who.int). The aetiological agent of typhoid fever is *Salmonella enterica* serovar Typhi (or simply Typhi) which is highly homogenous (13, 33). The genetic homogeneity of Typhi has significantly impeded the development of suitable typing methods to differentiate Typhi isolates, for both phylogenetic and epidemiological purposes.

Single nucleotide polymorphisms (SNPs) have recently been shown to be useful markers for typing Typhi isolates (23, 29). SNP typing can resolve the relationships among global Typhi isolates and more discriminating than widely accepted population genetic methods including Multilocus Enzyme Electrophoresis (28) and Multilocus Sequence Typing (13). However, some haplotypes or SNP profiles contained many isolates which could not be further differentiated (23, 29). In Roumagnac et al. (29) study, 88 SNPs differentiated 481 global Typhi isolates into 85 haplotypes. The majority of the isolates belonged to H58 (35%), H50 (12%) and H52 (11%). In the study by Octavia et al. (23), 38 SNPs distinguished 73 global Typhi isolates into 23 SNP profiles and the majority of these isolates had SNP profile 10 (32%) and SNP profile 2 (16%). Clearly, SNP typing still has a limited discriminatory power.

Variable number of tandem repeats (VNTRs) have the potential to be more discriminating than SNPs and also to be used to establish the evolutionary relationships of the isolates. VNTRs are short sequence repeats which are unique DNA elements repeated in tandem. The polymorphisms in VNTRs are believed to be resulted from slippage strand misalignment (17). Therefore, isolates may contain different copy numbers for a repeat locus, allowing differentiation between isolates. Multilocus VNTRs analysis (MLVA) involves determination of the number of repeats at multiple VNTR loci and the number of loci required varies depending on the diversity of the organism studied. MLVA has been particularly effective to type homogeneous clones including *Yersinia pestis* (1, 14, 21, 25), *Bacillus anthracis* (8, 11, 12, 34) and *Mycobacterium tuberculosis*.
In *Salmonella*, a few serovars including Enteritidis, Typhimurium and Typhi have been studied by MLVA (3, 4, 18-20, 27).

Two MLVA studies of Typhi showed different levels of variation of VNTR loci analysed (20, 27). Liu *et al.* (20) found five potential VNTR loci designated as TR1 to TR5, with the first three showing variation among 59 Typhi isolates from several Asian countries studied. Ramisse *et al.* (27) found five new polymorphic VNTRs (SAL02, SAL06, SAL10, SAL15 and SAL20). Together with two markers from previous Typhi and Typhimurium studies, TR1 (20) and STTR5/Sal16 (18), a total of seven VNTRs distinguished 27 Typhi isolates from France into 25 MLVA profiles (27). In these two studies, VNTR PCR products were resolved on standard agarose gels. However, the resolution of agarose gels is known to be low, which is especially difficult to resolve short repeat unit such as SAL10 with a 2 bp repeat unit. In this study we used seven published VNTRs, including SAL02, SAL06, SAL10, SAL16, SAL20, TR1 and TR2, and two new VNTRs uncovered in this study as markers to explore their potential in studying the molecular evolution of global Typhi isolates. Our MLVA assay employed universal M13-tailed primer tagged with a different fluorescent dye to resolve the tandem repeats on capillary electrophoresis. We combined the more rapidly evolving VNTR markers with the slower evolving SNPs to achieve an optimal resolution for typing global Typhi isolates.

**Materials and Methods**

**Bacterial strains.** The 73 global Typhi isolates analysed in this study were from a collection of isolates that have been previously SNP typed (23).

**Identification of new VNTR markers.** A publicly available database ([http://minisatellites.u-psud.fr](http://minisatellites.u-psud.fr)) was used to identify tandem repeats in the genomic sequence of Typhi strain CT18. The search for VNTRs was only run for Typhi strain CT18 under the parameters: total length more than or equal to 20 bp, unit length more than or equal to 3 bp and the repeat number of more than or equal to three. The potential VNTR markers were named according to their chromosomal locations.
**MLVA typing.** MLVA typing was done by PCR amplification of each VNTR locus separately. Each VNTR PCR reaction contained a pair of VNTR-specific primers which amplify the region flanking the potential VNTR and a fluorescent dye labeled M13 sequence universal primer with one of the four dyes; FAM (blue), VIC (green), NED (yellow) and PET (red). The primers were listed in Table 1. The forward VNTR-specific primer has a M13 sequences attached to the 5’ end. This allowed incorporation of M13 sequence to a newly synthesised strand in the initial PCR cycles and thus, the dye-labeled M13 universal primer would prime the PCR product synthesis in subsequent cycles. Therefore the PCR product was labeled with a fluorescent dye utilising a dye labelled M13 universal primer.

The PCR was done in 20 µl reaction containing 0.2 µl DNA template (~10 ng), 50 nM, 200 nM and 250 nM for M13-tailed forward, M13 dye-labelled and reverse primers respectively, 0.2 µl 10 mM dNTPs, 2 µl 10xTaq polymerase PCR buffer (New England Biolabs), 0.125 µl (1.25 U) Taq polymerase (New England Biolabs) and MilliQ water to adjust to the final volume. The PCR conditions included a touch down cycling profile as the following: Initial denaturation 95°C for 5 min; 96°C for 1 min, 68°C for 5 min (-2°C/cycle, a decrease of 2°C after each cycle) and 72°C for 1 min for 5 cycles; 96°C for 1 min, 58°C for 2 min (-2°C/cycle) and 72°C for 1 min for 5 cycles; 96°C for 1 min, 50°C for 1 min and 72°C for 1 min for 25 cycles; and final extension at 72°C for 5 min.

The PCR products for four VNTRs labeled with different fluorescent dyes, FAM (blue), VIC (green), NED (yellow) and PET (red), were pooled and run as one sample on an Automated GeneScan Analyser ABI3730 (Applied Biosystem) at the sequencing facility of the School of Biotechnology and Biomolecular Sciences, the University of New South Wales. The fragment size was determined using the LIZ600 size standard (Applied Biosystem) and data were analysed using GeneMapper v 3.7 software (Applied Biosystem). The dyes affect mobilities differently. VIC and NED labelled products give identical size calling while FAM labelled product is 1 bp shorter and PET labelled 2 bp longer. Therefore, corrections were done for the sizing data. When using the same dye, the fragment sizes (allowing 1 bp difference amongst different runs) of each VNTR matched to the expected number of repeats. There was no ambiguity when converting
Bioinformatic analysis. A dendrogram was constructed using unweighted pair group method with arithmetic means (UPGMA), available in PHYLIP package, to represent the genetic relationships of the MLVA profiles. This package is accessible through the Australian National Genomic Information Service (ANGIS). Minimum spanning tree (MST) was generated using Arlequin v. 3.1 available from http://cmpg.unibe.ch/software/arlequin3. The Simpson’s index of diversity (D value) was calculated using an in-house program, MLEECOMP package (26).

Results and Discussion

Universal M13 primer hemi-nested PCR for MLVA typing of nine VNTR loci

We explored the genome of Typhi strain CT18 for potential VNTR loci additional to those reported previously (20, 27), aiming to obtain a set of most variable VNTR loci for epidemiological typing. Forty-three potential VNTRs were identified using the web-based resource tool (6) (Supplementary Table 1). The primers were tested using a panel of 12 Typhi isolates. Five primer sets failed to produce any PCR product and were not investigated further. Two of the 38 potential VNTRs with successful PCR amplification, designated as TR4500 and TR4699, were polymorphic. These two new VNTRs are located in genic regions each with a repeat unit size of 6 bp. These two, together with seven VNTRs from previous studies (20, 27), were used to type the 73 global Typhi isolates. The seven VNTRs (SAL02, SAL06, SAL16, SAL20, TR1 and TR2) were selected as they showed three or more alleles in one or both studies (20, 27). TR3 was excluded because of presence of deletions immediately upstream of the repeat region (20). The primer pairs for the nine selected VNTRs could efficiently amplify a unique DNA fragment from all 73 isolates.

MLVA requires accurate sizing of VNTR PCR products (24). We employed fluorescent dye labeled universal M13 primer in a hemi-nested touch-down PCR (31) to obtained dye labeled VNTR product for accurate sizing on capillary electrophoresis. Each VNTR was amplified
separately and therefore four VNTR loci each with a different fluorescent dye (VIC, PET, FAM and NED) were pooled together as one sample to run on an automated DNA sequencer. The use of capillary electrophoresis has made the fragment size determination far more accurate than any agarose gel based system used previously (20, 27). Having fluorescent dyes on the universal M13 primer only, has significantly reduced the cost for our initial screening for variable VNTRs as it required only four dye-labeled universal primers rather than 38 dye-labeled VNTR specific primers for the 38 VNTRs screened.

**Level of polymorphism and discriminatory power of the nine VNTR loci**

Fragment sizes were converted to number of repeats for each VNTR. For the genome sequenced strains, CT18 and Ty2, there were discrepancies between theoretical and observed number of repeats for three VNTRs. TR4699, SAL02 and SAL16 of CT18 consist of 7, 10 and 14 repeats respectively according to the genome sequence but 13, 9 and 8 repeats respectively according to our typing. Similarly for Ty2, TR4699, SAL02 and SAL16 showed 23, 15 and 16 copies in genome sequence but 20, 16 and 10 copies in ours respectively. For both strains, we confirmed the copy numbers by sequencing. SAL16 was previously found to be inconsistent for serovar Typhimurium strain LT2 (27). Possible explanations for this observation are either genome sequencing errors or strain divergence between our strains and the sequenced strains or a combination of the two.

All except TR4500 can be converted into copy numbers without ambiguity. TR4500 had a case that the conversion is not perfect and was found to have an insertion of 55 bp between the second and third repeat. The different number of repeats was assigned as alleles. We sequenced represent isolates for all alleles observed and confirmed the number of repeats. Overall, two to 28 alleles were observed and the repeat units ranged from a single repeat, to as many as 40 repeats. TR4699 had the highest number of alleles followed by TR2, SAL02, TR1, SAL20 and SAL16. Continuous range of repeat units was observed in TR1, from 5 to 16 repeats. For all VNTRs except TR2, a particular number of repeats are predominant with TR4500, TR4699, SAL02, SAL16 and TR1 having mostly 3, 6, 11, 7 and 16 repeats respectively. In SAL20, the most frequent repeats are 15 and 16, equally shared by 17 isolates each, while in TR2, the most...
frequent repeats are 11, 12 and 17, equally shared by six isolates each. The allelic variability of the nine VNTR loci is summarised in Table 2. The distributions of number of repeats for most VNTR loci were different between our data and the other two studies. The predominant repeat number for SAL02 and SAL16 was 14 and 14 repeats in Ramisse et al. (27) but 11 and 7 repeats in ours. In TR1, the predominant number of repeats is 12, 17, 16 in Liu et al. (20), Ramisse et al. (27), and ours respectively.

The MLVA data of the 73 Typhi isolates are presented as allelic numbers in Supplementary Table 2, which were distinguished into 70 MLVA profiles. Only three pairs of isolates (SARB63 with IP.E88 374; CC6 with CC7; and 444Ty with 702Ty) were identical for all VNTR loci. The discriminatory power of each VNTR and all nine VNTRs combined was measured by Simpson’s Index of diversity (D) (10). The D value could also be used to determine the usefulness of these VNTRs loci for typing. The closer the value is to one, the better the marker is to differentiate the isolates for epidemiological purpose. The discriminatory power of each VNTR differed; the D value ranged from 0.042 to 0.964 (Table 2). The three most variable VNTR loci, TR2, TR4699 and SAL02 exhibited D values of 0.964, 0.947 and 0.927 respectively. For all nine VNTR loci, the D value is 0.999 suggesting that these VNTRs are highly discriminating and they are expected to be highly valuable for local epidemiology. The D value for MLVA was much higher than SNP typing (23) and ribotyping (15) where the discriminatory power for both was only 0.87. The same level of discrimination can also be achieved by typing only five of the nine VNTR loci, including TR4699, SAL02, SAL20, TR1 and TR2.

Seven of the nine VNTRs were from previous studies (20, 27) allowing a direct comparison for diversity. For the six loci (SAL02, SAL06, SAL10, SAL16, SAL20, TR1) common with Ramisse et al. (27), the combined D values were similar (0.945 in our study versus 0.936 in Ramisse et al.). The discriminatory power of two loci, SAL10 and SAL06, were lowest in both studies but the D values were much lower in this study (0.04 and 0.24) than those (0.14 and 0.43) in Ramisse et al. (27). The D values of the remaining four loci were comparable between the two studies with slightly higher values in this study. Two loci, TR1 and TR2, are also in the Liu et al. (20) study with D values of 0.91 and 0.97, respectively which is slightly higher than those
observed in this study (0.89 and 0.96). The differences between the two studies are likely due to differences in sampling.

Fast evolving makers such as VNTRs give high discriminatory power for outbreak investigations. However, a fast evolving marker may vary during an outbreak and such marker may be considered unstable. SAL16 (STTR5) which was also used in the Typhimurium MLVA scheme has been shown to vary in three of the eight different Typhimurium outbreaks studied (9). To accommodate such variation for epidemiological typing, it has been suggested that a single locus variant may be classified as belonging to the same outbreak (9, 22). The stability of TR1 and TR2 used in this study were tested in laboratory culture using four isolates by Liu et al. (20). After passaging 40 rounds, no changes were observed in these two VNTR loci (20). It remains to be tested whether any of the markers used in this study varies during outbreaks.

**Phylogenetic relationships based on the MLVA data**

Initially, a dendrogram was generated using UPGMA to illustrate the genetic relationships among the MLVA profiles. The dendrogram showed that there were seven major clusters (Figure 1). These clusters were supported by certain VNTR alleles. For TR4500, the majority of isolates had allele 2 in cluster 1, allele 4 in cluster 2, allele 3 in clusters 3 to 6, and allele 1 in cluster 7. In addition, cluster 2 was supported by allele 4 of TR4699, allele 6 of SAL02, allele 4 of SAL16 and allele 12 of TR1; cluster 3 by allele 3 of SAL16 and allele 8 of SAL20; cluster 4 by allele 5 of TR1, cluster 5 by allele 9 of TR1, cluster 6 by allele 4 of SAL06 and allele 8 of TR1; and cluster 7 by allele 8 of SAL16 and allele 1 of TR1. However, these cluster supporting alleles were not unique for that particular cluster. Only allele 1 for TR4500 was found within cluster 7 while the other alleles were also present in other clusters. Although different clusters could be distinguished on the UPGMA dendrogram, none of them was supported by high bootstrap values. Only four branches were supported by bootstrap values greater than 50%. This suggests that the VNTR data were too diverse to infer a robust relationship.

A minimum-spanning tree (MST) was often used to display relationships based on MLVA data (5, 27, 30). An MST was also constructed to visualise the relationships of MLVA profiles
(Supplementary Figure 1). The majority of the MLVA profiles differed by four VNTR loci, with the minimum difference of one and the maximum of five. Only two pairs of MLVA profiles, 50 and 58, and 56 and 58, differed by one VNTR, SAL20 and TR2 respectively. There was no apparent major clustering in the MST. The seven major clusters found by UPGMA were not observed in the MST. This further suggests that the MLVA data were too divergent to give a meaningful MST as the VNTRs used, evolved rapidly.

The phylogenetic analysis allowed us to determine whether there was any epidemiological link for isolates obtained from the same region. Nineteen of the 73 isolates were from Canada and the VNTR data showed that these isolates were not clustered together, suggesting that none of these isolates were epidemiologically related. This conclusion is also supported by our previous study using SNP data which did not cluster them together (23). This observation was not surprising since in developed countries Typhi infections are often related to travel to endemic countries as sporadic cases (2). On the other hand, for the three pairs of isolates with identical MLVA profiles, all were likely to be epidemiologically related. One pair was isolated from Thailand in the same year, although no other details were available.

Previously, we have typed the 73 global Typhi isolates using 38 SNPs (23). These SNPs were able to differentiate the isolates into 23 SNP profiles and 4 clusters. It can be seen from Figure 1 that most of the UPGMA clusters established by VNTR data did not correspond to the SNP clustering. For example, SNP cluster I was split and SNP profiles fell within two VNTR clusters, clusters 2 and 3. There were phylogenetic signals supporting SNP cluster I, including allele 4, 4, 6, 4, 12 for TR4500, TR4699, SAL02, SAL16 and TR1 respectively. A subset of SNP cluster 3 profiles were also grouped together as VNTR cluster 3. However, the other SNP clusters were virtually disappeared and were distributed into different VNTR clusters. SNP profiles that contain multiple isolates were further differentiated by MLVA. However, MLVA profiles with these same SNP profiles were not always grouped together.

A composite tree based on both SNP and VNTR data
SNPs were able to establish the phylogenetic relationships of these isolates but lack discriminatory power as shown previously (23). In contrast, this study showed that VNTRs were highly discriminating but gave only meaningful phylogenetic relationships of closely related isolates. Therefore, a composite tree was constructed (Figure 2). Firstly, an MST was constructed using SNP data to resolve the 73 global Typhi isolates into their corresponding SNP profiles and clusters which is the same tree as previously described (23). For SNP profiles that contained more than two MLVA profiles (SNP profiles 2, 3, 5, 10, 13 and 18), an MST of the MLVA profiles was generated for each SNP profile. As can be seen in Figure 2, MLVA profiles within SNP profiles 10, 13 and 18 differed by at least four VNTR loci. The high levels of VNTR diversity within these three SNP profiles were due to poor differentiation by the SNPs used as discussed previously (23) which were obtained from comparison of CT18 and Ty2 genomes and were outside the evolutionary path of these SNP profiles due to phylogenetic discovery bias. On the other hand, MLVA profiles belonging to SNP profiles 2, 3 and 5, all of which belong to SNP cluster I, had the least difference and the majority differed by one or two VNTR loci. As noted above SNP cluster I is largely consistent with MLVA clustering as the isolates within this cluster are more closely related.

**Concluding comments**

In this study we applied MLVA of nine VNTRs to type global Typhi isolates using fluorescent universal primer hemi-nested PCR and capillary electrophoresis. We differentiated 73 Typhi isolates into 70 MLVA profiles. We have further shown that MLVA offers a higher discriminatory power than SNP typing previously applied to these isolates (23). The nine VNTR markers have different levels of variability with seven (TR4500, TR4699, TR1, TR2, SAL02, SAL16 and SAL20) being highly variable. Although typing the five most variable VNTRs (TR2, TR4699, SAL02, TR1 and SAL20) is sufficient to achieve maximum differentiation of isolates used in this study, applying all seven VNTRs will offer higher discriminatory power for outbreak investigations. We suggest these seven VNTRs to be used as a standardized typing scheme for Typhi.
Capillary electrophoresis was used to accurately size the VNTR PCR products in this study in contrast to the two previous studies (20, 27) using agarose gel electrophoresis which has far lower resolution than capillary electrophoresis. The VNTR data expressed as alleles based on number of repeats will allow comparison of data between laboratories. We have set up a database and deposited our data at http://minisatellites.u-psud.fr/MLVAnet. This resource will be useful for rapid computerised identification and classification of any Typhi isolates for epidemiological study.

The isolates used in this study were previously typed using 38 SNPs (23), allowing a direct comparison of SNPs and VNTRs as markers for typing Typhi. The highly polymorphic VNTRs appeared to evolve too fast to have retained sufficient phylogenetic information for determination of the evolutionary relationships of the 73 global Typhi isolates studied. The VNTRs did not give any support for SNP cluster II to IV previously established. Only some phylogenetic signals were evident in the isolates belonging to SNP cluster I corresponding to VNTR cluster 2. This suggests that VNTRs could only be used to determine the relationships of closely related isolates. Therefore a combination of SNP typing and VNTR typing will be more appropriate for global surveillance of Typhi.

Acknowledgements

This work was supported by a grant from the National Health and Medical Research Council of Australia and UNSW faculty research program. We thank Drs Ken Sanderson (the University of Calgary) and Gordon Dougan (Imperial College London) for generously providing us the strains. We also thank the referees for helpful suggestions.

References


Figure 1. Unweighted pair group method with arithmetic means (UPGMA) dendrogram of the 70 MLVA profiles. The assigned clusters are indicated by roman numerals. Bootstrap values, if greater than 50%, are presented at the nodes of the tree. Table on the right corresponds to the allelic pattern for each MLVA profile. The shaded cells correspond to the alleles that supported the clustering of the MLVA profiles.

Figure 2. A composite tree combining a minimum spanning tree (MST) generated using SNP and an MST using VNTR data. A secondary MST was generated only for SNP profiles containing more than two MLVA profiles and the MST for the corresponding SNP profile is located adjacent to the given SNP profile. The numbers on the MLVA MST trees indicate the number of differences in VNTR loci. The SNP clusters are indicated with roman numerals on the node. Members of SNP cluster II were omitted for clarity as there was no further MLVA MST for any of the SNP profiles in this cluster.
Table 1. Primers used for VNTR typing

<table>
<thead>
<tr>
<th>VNTR Name</th>
<th>Positions</th>
<th>Primer Sequence 5' -&gt; 3'</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR4500</td>
<td>4,500,182</td>
<td>U cacgacgttaaaagacCGTGTGCTGCTCCGAAAT</td>
</tr>
<tr>
<td></td>
<td>4,500,203</td>
<td>L GCGTGAAAGTGAAAAAG</td>
</tr>
<tr>
<td>TR4699</td>
<td>4,699,223</td>
<td>U cacgacgttaaaagacTATTCTACTTCAGTCCCC</td>
</tr>
<tr>
<td></td>
<td>4,699,322</td>
<td>L AACCTCCCCGCTATCTACAA</td>
</tr>
<tr>
<td>TR1(^1)</td>
<td>2,017,168</td>
<td>U cacgacgttaaaagacAGAACCAGCAATGCGCAACGA</td>
</tr>
<tr>
<td></td>
<td>2,017,251</td>
<td>L CAAGAAGTGCGCATACTACCC</td>
</tr>
<tr>
<td>TR2(^1)</td>
<td>2,557,000</td>
<td>U cacgacgttaaaagacCCTGTGTTCTTGCTGATACG</td>
</tr>
<tr>
<td></td>
<td>2,557,222</td>
<td>L CAGAGGATATCGCAACAATCGG</td>
</tr>
<tr>
<td>SAL02(^1)</td>
<td>666,037</td>
<td>U cacgacgttaaaagacGGAAGACTGGCGCAACAAAT</td>
</tr>
<tr>
<td></td>
<td>666,093</td>
<td>L TCGCAATACCATGAGTACG</td>
</tr>
<tr>
<td>SAL06(^1)</td>
<td>764,489</td>
<td>U cacgacgttaaaagacTTGTCGCGGAACCTATAACTG</td>
</tr>
<tr>
<td></td>
<td>764,518</td>
<td>L CTTCGCTGATGCGGACTCC</td>
</tr>
<tr>
<td>SAL10(^1)</td>
<td>2,016,742</td>
<td>U cacgacgttaaaagacAAGCGACGTTCCTCTGCAAC</td>
</tr>
<tr>
<td></td>
<td>2,016,765</td>
<td>L TGGATATGATGGCATGACG</td>
</tr>
<tr>
<td>SAL16(^1)</td>
<td>3,041,486</td>
<td>U cacgacgttaaaagacCCATGGCTGAGTTAATTTCT</td>
</tr>
<tr>
<td></td>
<td>3,041,568</td>
<td>L TGATACGCTTTTACAGCTGC</td>
</tr>
<tr>
<td>SAL20(^1)</td>
<td>3,643,794</td>
<td>U cacgacgttaaaagacCACGCGACACAACTAACGA</td>
</tr>
<tr>
<td></td>
<td>3,643,841</td>
<td>L ACTGTACCGTGCCGTTT</td>
</tr>
</tbody>
</table>

\(^1\) The location of the VNTR on the chromosome of *S. enterica* serovar Typhi strain CT18
\(^2\) cacgacgttaaaagac is the M13 tail attached at the 5' end of the upper primers
\(^3\) Primers designed from previously published literatures (20, 27) with additional of 18-mers M13 tail at 5' end of the upper primer
Table 2. Features of the nine polymorphic VNTRs observed in the 73 Typhi isolates

<table>
<thead>
<tr>
<th>VNTR Name</th>
<th>Gene</th>
<th>Product</th>
<th>Consensus</th>
<th>No. of Repeats</th>
<th>PCR Product Size Range (bp)</th>
<th>No. of alleles</th>
<th>Range of No. of Repeats</th>
<th>D values</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>TR4500</td>
<td>STY4635</td>
<td>hypothetical protein</td>
<td>GGACTC</td>
<td>6 4 5</td>
<td>294-366</td>
<td>5</td>
<td>1, 2, 3, 3, 4</td>
<td>0.691</td>
<td>This study</td>
</tr>
<tr>
<td>TR4699</td>
<td>sefC</td>
<td>outer membrane fimbral usher protein</td>
<td>TGTTGG</td>
<td>6 17 23</td>
<td>163-307</td>
<td>23</td>
<td>3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 20, 21, 22, 25, 26, 27, 30</td>
<td>0.947</td>
<td>This study</td>
</tr>
<tr>
<td>TR1</td>
<td>intergenic region between yedD and yedE</td>
<td>AGAAGAA</td>
<td>7 12 12</td>
<td>227-304</td>
<td>12</td>
<td>5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16</td>
<td>0.893</td>
<td>Liu et al. (20) Ramisse et al. (27)</td>
<td></td>
</tr>
<tr>
<td>TR2</td>
<td>intergenic region between acrD and yffB</td>
<td>CCAGTTCC</td>
<td>8 28 28</td>
<td>284-620</td>
<td>28</td>
<td>1, 3, 4, 5, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 26, 27, 28, 32, 37, 40</td>
<td>0.964</td>
<td>Liu et al. (20)</td>
<td></td>
</tr>
<tr>
<td>SAL02</td>
<td>citT</td>
<td>citrate carrier</td>
<td>TACCAG</td>
<td>6 10 20</td>
<td>146-278</td>
<td>20</td>
<td>6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 23, 24, 26, 28</td>
<td>0.927</td>
<td>Ramisse et al. (27)</td>
</tr>
<tr>
<td>SAL06</td>
<td>STY0765</td>
<td>Pseudogene</td>
<td>CTCAAT</td>
<td>6 5 3</td>
<td>180-198</td>
<td>4</td>
<td>1, 3, 5, 6</td>
<td>0.244</td>
<td>Ramisse et al. (27)</td>
</tr>
<tr>
<td>SAL10</td>
<td>yedD</td>
<td></td>
<td>ACGCCGCTGCCG</td>
<td>12 2 2</td>
<td>197-209</td>
<td>2</td>
<td>1, 2</td>
<td>0.042</td>
<td>Ramisse et al. (27)</td>
</tr>
<tr>
<td>SAL16</td>
<td>Intergenic region between STY3169 (pseudogene) and STY3172</td>
<td>ACCATG</td>
<td>6 14 10</td>
<td>216-276</td>
<td>10</td>
<td>4, 5, 6, 7, 8, 9, 10, 11, 13, 14</td>
<td>0.834</td>
<td>Ramisse et al. (27)</td>
<td></td>
</tr>
<tr>
<td>SAL20</td>
<td>ftsN</td>
<td>cell division protein</td>
<td>CAG</td>
<td>3 16 11</td>
<td>183-219</td>
<td>11</td>
<td>8, 9, 11, 12, 13, 14, 15, 16, 17, 18, 20</td>
<td>0.864</td>
<td>Ramisse et al. (27)</td>
</tr>
</tbody>
</table>

* The isolates had three GGACTC repeats in addition to 55 bp sequence insertions between the second and third repeat, and thus were assigned a new allele number.