

## Research Article

# Identification of Five Novel *Salmonella* Typhi-Specific Genes as Markers for Diagnosis of Typhoid Fever Using Single-Gene Target PCR Assays

Yuan Xin Goay,<sup>1</sup> Kai Ling Chin,<sup>1</sup> Clarissa Ling Ling Tan,<sup>1</sup>  
Chiann Ying Yeoh,<sup>1</sup> Ja'afar Nuhu Ja'afar,<sup>1</sup> Abdul Rahman Zaidah,<sup>2</sup>  
Suresh Venkata Chinni,<sup>3</sup> and Kia Kien Phua<sup>1</sup>

<sup>1</sup>Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia (USM), Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia

<sup>2</sup>Department of Medical Microbiology and Parasitology, Universiti Sains Malaysia (USM), Health Campus, 16150 Kubang Kerian, Kelantan, Malaysia

<sup>3</sup>Faculty of Applied Sciences, AIMST University, Jalan Bedong-Semeling, 08100 Bedong, Kedah, Malaysia

Correspondence should be addressed to Kia Kien Phua; [kkphua7@gmail.com](mailto:kkphua7@gmail.com)

Received 22 June 2016; Revised 27 September 2016; Accepted 18 October 2016

Academic Editor: Hao-Teng Chang

Copyright © 2016 Yuan Xin Goay et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Salmonella* Typhi (*S. Typhi*) causes typhoid fever which is a disease characterised by high mortality and morbidity worldwide. In order to curtail the transmission of this highly infectious disease, identification of new markers that can detect the pathogen is needed for development of sensitive and specific diagnostic tests. In this study, genomic comparison of *S. Typhi* with other enteric pathogens was performed, and 6 *S. Typhi* genes, that is, STY0201, STY0307, STY0322, STY0326, STY2020, and STY2021, were found to be specific *in silico*. Six PCR assays each targeting a unique gene were developed to test the specificity of these genes *in vitro*. The diagnostic sensitivities and specificities of each assay were determined using 39 *S. Typhi*, 62 non-Typhi *Salmonella*, and 10 non-*Salmonella* clinical isolates. The results showed that 5 of these genes, that is, STY0307, STY0322, STY0326, STY2020, and STY2021, demonstrated 100% sensitivity (39/39) and 100% specificity (0/72). The detection limit of the 5 PCR assays was 32 pg for STY0322, 6.4 pg for STY0326, STY2020, and STY2021, and 1.28 pg for STY0307. In conclusion, 5 PCR assays using STY0307, STY0322, STY0326, STY2020, and STY2021 were developed and found to be highly specific at single-gene target resolution for diagnosis of typhoid fever.

## 1. Introduction

To date, there are more than 2,500 serotypes identified within the *Salmonella enterica* species [1]. Most are harmless to humans but one serotype, *Salmonella enterica* subspecies *enterica* serovar Typhi (*S. Typhi*), causes typhoid fever, a severe and life-threatening systemic infection in humans. Worldwide, typhoid fever causes 269,000 deaths from 26.9 million new cases each year [2]. Travellers, children, the elderly, and immune-compromised individuals are especially at risk [3, 4]. The clinical manifestations of typhoid fever are similar to other febrile illnesses. Therefore, diagnosis based on clinical signs and symptoms alone is difficult [5]. The

emergence of multidrug-resistant *S. Typhi* strains and development of the typhoid carrier state have further complicated the management of typhoid fever [6, 7]. Delay in diagnosis and initiation of antibiotic treatment can cause serious clinical complications and fatality [8]. Thus, early and correct laboratory diagnosis of typhoid fever is critical to reduce the morbidity and mortality, as well as curtail transmission of the disease.

DNA-based detection methods, such as polymerase chain reaction (PCR), have proven to be sensitive, specific, and rapid compared to conventional culture-based methods for the diagnosis of many infectious diseases [9–11]. Several target genes have been used for *S. Typhi* identification using

PCR, such as the O antigen somatic genes (*tyv* and *prt*) [12], H antigen flagellar gene (*fliC-d*) [13], and Vi capsular antigen gene (*viaB*) [14]. However, these genes cannot stand alone as single *S. Typhi*-specific diagnostic marker since they are not specific to *S. Typhi* and are also found in other *Salmonella* serotypes. Thus, these markers provide provisional rather than differential diagnosis of typhoid fever. For example, the *fliC-d* gene of *S. Typhi* shares the same nucleic acid sequence as *S. Muenchen* [15]; the *prt* gene is present in *S. Typhi*, *S. Paratyphi A*, and *S. Enteritidis* [12]; and the *viaB* gene is found not only in *S. Typhi* but also in *S. Dublin*, a few strains of *S. Paratyphi C* [16] and *Citrobacter freundii* [17]. Due to the lack of specificity of these target genes, a combination of different pairs of primers using multiplex PCR [18] or nested PCR [19] are needed to increase the sensitivity and specificity of the PCR diagnostic test. This, however, will increase the cost, time, and complexity of the laboratory diagnosis.

Diagnostic markers which can detect pathogens at single-gene target resolution could lead to a simpler, cost-effective, and more functional DNA-based detection method since less primers are needed for target detection. Many approaches, such as subtractive hybridization [20], next generation sequencing [21], and microarray [22] techniques, have been used to identify genes that are specific or unique to a pathogen. However, these high-end technologies are cumbersome and expensive and sometimes yield false negative or false positive results [23]. Since bacterial genome databases have expanded tremendously over the past decade and advancement in computing technologies has made nucleic acid sequence alignment services readily accessible at NCBI, *in silico* comparative hybridization approach coupled with *in vitro* PCR (wet-lab) validation is sufficient to facilitate the translation of genomic data into diagnostic marker discoveries. In this study, a low-cost and simple attempt was made to identify new DNA diagnostic markers specific for *S. Typhi* by utilizing genome data (stored in NCBI databases) and nucleic acid sequence alignment tools (BLASTn) that are readily available in the public domain. The diagnostic sensitivities and specificities of the primers designed for amplifying whole gene sequences can be validated using a panel of confirmed bacteria isolates selected from *S. Typhi*, non-*Typhi Salmonella*, and non-*Salmonella* clinical isolates. To serve as a control for the PCR reaction, 16S rRNA gene, that is ubiquitous among bacteria species, can be used as a PCR amplification control [24].

## 2. Materials and Methods

**2.1. Bacterial Strains.** A total of 111 bacteria isolates including 39 *S. Typhi*, 62 non-*Typhi Salmonella* serotypes, and 10 non-*Salmonella* strains were used in this study. *S. Typhi* strains consisted of 1 *S. Typhi* reference strains (ATCC 7251) and 38 different pulsed-field types (PFTs) representing all strains in the state of Kelantan in Malaysia. These 38 PFTs were the result of screening 279 *S. Typhi* clinical isolates using pulsed-field gel electrophoresis (PFGE) [25]. Non-*Typhi Salmonella* serotypes were closely related *Salmonella* species made up of 26 different serotypes (Table 2) and 10 ATCC strains including *S. Paratyphi A* (ATCC

9150), *S. Paratyphi B* (ATCC BAA 1250), *S. Paratyphi C* (ATCC 9068), *S. Enteritidis* (ATCC 13076), *S. Typhimurium* (ATCC 14028), *S. Weltevreden* (NCTC 6534), *S. Agona* (ATCC 51957), *S. Heidelberg* (ATCC 8326), *S. Poona* (ATCC 04840), and *S. Braenderup* (ATCC BAA-664). In addition, 10 other non-*Salmonella* strains such as *Shigella dysenteriae*, *Shigella flexneri*, *Shigella boydii*, *Shigella sonnei*, *Vibrio cholera*, *Enterohemorrhagic E. coli*, *Enteropathogenic E. coli*, *Aeromonas hydrophila*, *Yersinia enterocolitica*, and *Klebsiella pneumoniae* were also included. All clinical strains were procured from the Department of Clinical Microbiology and Parasitology, Hospital Universiti Sains Malaysia (HUSM), Kelantan, Malaysia, and the Biobank of the Institute for Research in Molecular Medicine (INFORMM), Kelantan, Malaysia. All bacteria strains were stored in glycerol stocks at  $-80^{\circ}\text{C}$  until being ready for use. Ethical clearance for this project was obtained from the Human Research Ethics Committee, Universiti Sains Malaysia (reference number USM/KK/PPP/JEPeM [235.3.(16)]).

**2.2. Culture Conditions and Confirmation Tests.** All bacteria isolates used in this study were confirmed by traditional culture, biochemical, and serotyping methods as described in ISO6579 with some modifications. Bacteria isolates were revived from frozen glycerol stocks by pipetting  $100\ \mu\text{L}$  thawed cells into 10 mL nutrient broth and incubated at  $37^{\circ}\text{C}$  for 18 hours in an orbital shaker at 200 rpm. The bacteria were streaked on Xylose Lysine Deoxycholate (XLD) selective agar and incubated at  $37^{\circ}\text{C}$  for 18 hours. Colonies grown on the agar were tested with a panel of biochemical tests, including Triple Sugar Iron (TSI), urease, Methyl Red Voges Proskauer (MRVP), citrate, and indole tests. Suspected *Salmonella* isolates were then sent to the *Salmonella* Reference Centre, Institute for Medical Research (IMR), Malaysia, to confirm their serotypes using specific antisera and latex agglutination method.

**2.3. Identification of *S. Typhi*-Specific Genes Using Bioinformatics (In Silico).** Full genome sequence of *S. Typhi* CT18 (GenBank accession number AL513382) was downloaded from the National Center for Biotechnology Information database (NCBI) and used as the reference genome. The 2 plasmids, namely, pHCM1 and pHCM2, which resided in *S. Typhi* CT18 were excluded since plasmids are genetically unstable. The 6 complete *S. Typhi* whole-genome sequences available in NCBI were used for data mining. They comprised CT18 (Genbank accession number AL513382) [27], Ty2 (Genbank accession number AE014613) [28], P-stx-12 (Genbank accession number CP003278) [29], Ty21a (Genbank accession number CP002099) [30], B/SF/13/03/195 (Genbank accession number CP012151) [31], and PM016/13 (Genbank accession number CP012091) [32]. In order to ascertain whether the genomic regions were conserved and specific to *S. Typhi*, the nucleotide Basic Local Alignment Search Tool (BLASTn), a free online software for nucleic acid analysis, was used to compare the whole-genome sequence of *S. Typhi* CT18 with the other 5 complete *S. Typhi* genomes and other bacteria genomes in the NCBI database (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>). Genes found in

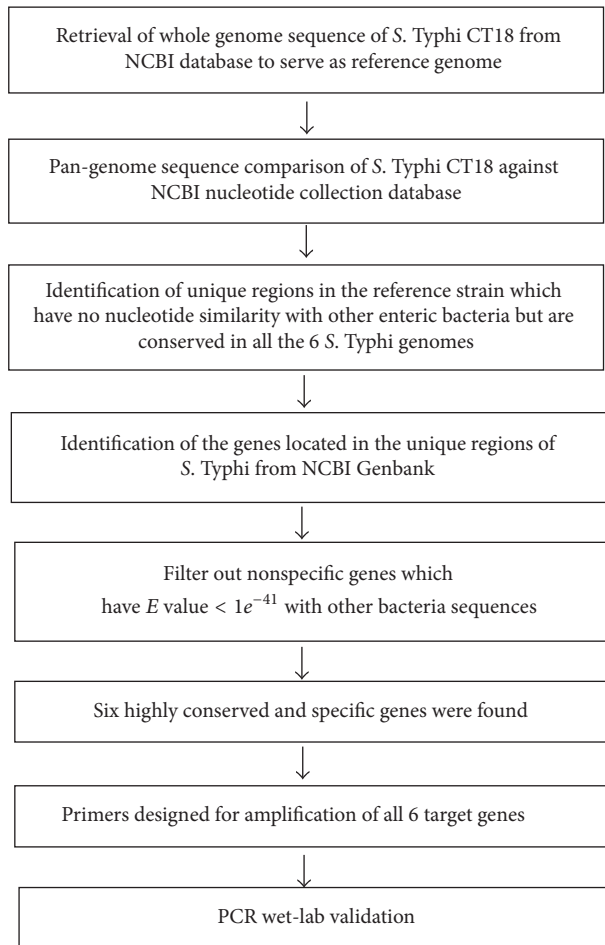


FIGURE 1: Experimental workflow describing the comparative genomic and wet-lab approaches used to identify and validate *S. Typhi*-specific DNA diagnostic markers.

unique regions which have no nucleotide similarity with other enteric organisms were identified and retrieved from the Genebank of NCBI. These genes were further screened individually using similarity searches against the NCBI non-redundant nucleotide (nr/nt) database to reconfirm their specificities. The program was set for “somewhat similar sequences search,” which allowed nucleotide sequence matching down to 7 bases (the smaller the nucleotide size, the more sensitive the result). Realizing the high genome similarity among the enteric pathogens and the possibility that different geographical areas may result in different bacterial genotypes, only genes which have 100% sequence conservation (an  $E$ -value threshold = 0.0) in all 6 complete *S. Typhi* genomes and had little or no similarity ( $E$ -value threshold  $\geq 1e^{-41}$ ) to other bacterial sequences in the NCBI database were considered as potential targets and were subjected to wet-lab analysis. The experimental pipeline is as shown in Figure 1.

**2.4. Design of Oligonucleotide Primers for PCR Amplification.** Primers were designed manually to amplify the *S. Typhi*-specific genes identified previously, including the start and

the stop codons. A pair of primers specific for 16S rRNA gene amplification as described by Marchesi and colleagues [24] were also incorporated into each PCR assay to serve as an internal amplification control (IAC). This is a universal gene target which is highly conserved in bacteria [24]. All primers were synthesized by Integrated DNA Technologies (IDT) Pte. Ltd., Malaysia.

**2.5. Template DNA Extraction.** DNA from all bacteria isolates were extracted using DNeasy Blood & Tissue kit® (Qiagen, USA) according to the manufacturer’s instructions. The purity and concentration of the extracted DNA were determined using Nanodrop Spectrophotometer ND-1000 (Thermo Fisher Scientific, USA). DNA concentration was measured from the absorbance at 260 nm. Ratio of the absorbance at 260 and 280 nm ( $A_{260/280}$ ) and ratio of the absorbance at 230 and 260 nm ( $A_{230/260}$ ) were used to evaluate the DNA quality. The extracted DNAs were diluted to a final stock concentration of 50 ng/ $\mu$ L using ultrapure water and stored at  $-20^{\circ}\text{C}$  until ready for PCR amplification.

**2.6. Optimization of PCR.** Each PCR assay was optimized using a modified Taguchi method as described by Cobb and Clarkson [33]. The effects and interactions of the 4 main PCR components (IAC primers, *S. Typhi*-specific gene primers,  $\text{MgCl}_2$ , and annealing temperatures) each at 3 different levels (IAC primers: 0.05, 0.10, and 0.15  $\mu\text{M}$ ; *S. Typhi* primers: 1.00, 1.50, and 2.00  $\mu\text{M}$ ;  $\text{MgCl}_2$ : 2.00, 2.50, and 3.00 mM, and annealing temperatures: 50, 55, and  $60^{\circ}\text{C}$ ) were investigated in a balanced orthogonal array of 9 experimental combinations. The PCR amplifications were carried out in a total reaction volume of 20  $\mu\text{L}$ , and the PCR products were analysed on a 1.2% (w/v) agarose gel containing SYBR® Safe DNA Gel Stain (Invitrogen, USA), visualized using a blue-light transilluminator (Syngene, UK).

**2.7. Analytical Specificities of Genes Unique to *S. Typhi*.** Analytical specificities of the PCR assays were assessed by running each PCR assay on a panel of bacteria strains consisting of 39 *S. Typhi*, 62 non-*Typhi Salmonella*, and 10 non-*Salmonella* clinical isolates.

**2.8. Detection Limit of the PCR Assays.** Detection limit of the PCR assays was defined as the minimum amount of *S. Typhi* DNA (ng/ $\mu\text{L}$ ) that yielded positive PCR amplicons. The assay sensitivities were determined by amplification of a 5-fold serial dilution of *S. Typhi* ATCC 7251 DNA, ranging from 50 ng to 25.6 fg. Two microliters of the DNA was subjected to PCR amplification. The analytical sensitivity was indicated by the presence of visible PCR product bands on the agarose gel using the transilluminator as described above.

**2.9. DNA Sequencing.** To confirm the PCR products were indeed derived from the *S. Typhi* strains, PCR amplicons from all assays produced using Phusion® High-Fidelity DNA Polymerase (New England Biolabs, USA) were purified and sent to First BASE Laboratories Pte. Ltd., Malaysia, for sequencing. The resultant nucleotide sequences were compared with the reference *S. Typhi* CT18 gene sequences in NCBI using BioEdit software.

TABLE 1: List of primers targeting *S. Typhi*-specific genes for the development of 6 PCR assays.

Target genes	Primer labels	Primer sequences (5'-3')	Target lengths (bp)
STY0201	0201F	ATGCTTTTAAAAAACACAACATGG	1176
	0201R	TTACGGATAGGTGATTGAAAATTG	
STY0307	0307F	ATGAAACCTTTATTCTCAGTGC	495
	0307R	TTAGCGTAATTCCCAGAACC	
STY0322	0322F	ATGAAATATAAAAAATAAGAG	678
	0322R	CTATGGATTCATTTCCATTTC	
STY0326	0326F	ATGAATACGAATAATTCACC	261
	0326R	TTACCCTCCCCATGTCAC	
STY2020	2020F	ATGCCTGTTATGCATAATTG	429
	2020R	TTATGCTGTTAACGAGTCGTC	
STY2021	2021F	ATGAGTTTAGCGCAGCCTAAATCC	732
	2021R	TTAGAAGTCTCCTGCCTGGAAAC	
16S rRNA <sup>a</sup>	16SF	CAGGCCTAACACATGCAAGTC	1362
	16SR	GGGCGGTGTGTACAAGGC	

<sup>a</sup> 16S ribosomal RNA gene served as internal amplification control (IAC) [24].

F represents forward primer.

R represents reverse primer.

### 3. Results

Using the bioinformatic method for whole-genome comparison (Figure 1), 6 potential diagnostic markers with NCBI locus tags, STY0201, STY0307, STY0322, STY0326, STY2020, and STY2021, were found. They exhibit 100% query coverage and identity ( $E$ -value = 0) with all 6 *S. Typhi* gene sequences but had low or no significant similarity ( $E$ -value  $\geq 1e^{-41}$ ) with other enteric bacteria nucleotide sequences as of 11 March 2016. These genes were found to be (bioinformatically) highly conserved and specific and thus were selected for further wet-lab validation using PCR method. The primers designed to amplify these selected genes are shown in Table 1.

The results showed that all 6 designed primer pairs successfully amplified their target genes with amplicon sizes of 1176, 495, 678, 261, 429, and 732 bps, respectively. DNA sequencing results of the amplicons showed 100% identity with their corresponding *S. Typhi* genes, confirming the fidelity and sensitivity of the primers.

The 6 single-gene target PCR assays were then optimized using Taguchi method with the incorporation of IAC which targeted the 16S rRNA gene. The optimized master mix for the PCR assays targeting STY0201, STY0307, and STY2020 genes consisted of 1x Green GoTaq Flexi Buffer, 2.0 mM MgCl<sub>2</sub>, 0.2 mM dNTPs, 1.5  $\mu$ M *S. Typhi*-specific gene primers, 0.10  $\mu$ M IAC primers, 0.75 U GoTaq Flexi DNA Polymerase (Promega, USA), and 5% glycerol in a total volume of 20  $\mu$ L. Two microliters of test DNA (50 ng/ $\mu$ L) was added to the master mix and amplified using the following optimized thermal-cycling parameters: initial denaturation at 95°C for 1 min, followed by 30 cycles elongation at 95°C for 30 s, 55°C for 30 s, 72°C for 1 min and a final extension at 72°C for 5 min. Similar PCR conditions were used for amplification of STY0322, STY0326, and STY2021 genes except for the concentration of MgCl<sub>2</sub> and IAC primers which were set at 3.0 mM and 0.15  $\mu$ M, respectively. The optimal annealing

temperature was set at 50°C. Under these conditions, the IAC primer pair produced an amplicon of 1,362 bp for all bacteria isolates tested (111/111).

The optimized PCR assays for STY0307, STY0322, STY0326, STY2020, and STY2021 correctly identified all *S. Typhi* (39/39) isolates, whereas none of the non-Typhi *Salmonella* (0/62) and none of the non-*Salmonella* (0/10) isolates were detected. This showed a 100% sensitivity and 100% specificity for the PCR assays (Table 2) and indicate that the 5 genes were unique to *S. Typhi* (Figures 2, 3, and 4).

The results of serial dilution of *S. Typhi* genomic DNA showed that the detection limit of the optimized PCR assays was 32 pg for gene STY0322, 6.4 pg for genes STY0326, STY2020, and STY2021, and 1.28 pg for gene STY0307.

Although gene STY0201 exhibited 100% sensitivity (detection of 39/39 *S. Typhi* isolates), it showed cross-reactivity with *S. Oslo* and *S. Kissi* (Table 2), resulting in a specificity of only 97.2% (detection of 2/72 of non-Typhi isolates). Sequencing of their PCR products showed a substitution of nucleotide C  $\rightarrow$  T at position 89 and T  $\rightarrow$  C at positions 354 and 1,026 for both *S. Kissi* and *S. Oslo*. The sequence variation between *S. Kissi* and *S. Oslo* with the *S. Typhi* CT18 reference genome was very small (only 3 nucleotide differences), indicating that the false positive results were due to sequence similarity among themselves.

### 4. Discussion

The diagnosis of typhoid fever based on clinical signs and symptoms is often ambiguous, while phenotypic detection of *S. Typhi* bacteria based on biochemical and serotyping methods is laborious and time-consuming. Thus, rapid molecular detection methods, such as nucleic acid-based amplification, such as PCR assay, is critically needed to help diagnose this contagious disease. Development of this test requires diagnostic markers that are sensitive and specific.

TABLE 2: Evaluation of the specificities of the 6 target genes for identification of *S. Typhi* using PCR (total of 111 clinical isolates).

Test bacteria strains	Positive PCR amplification for each target gene					
	STY0201	STY0307	STY0322	STY0326	STY2020	STY2021
<i>S. Typhi</i> ( $n = 39$ )	39/39	39/39	39/39	39/39	39/39	39/39
<i>S. Paratyphi A</i> ( $n = 10$ )	0/10	0/10	0/10	0/10	0/10	0/10
<i>S. Paratyphi B</i> ( $n = 10$ )	0/10	0/10	0/10	0/10	0/10	0/10
<i>S. Paratyphi C</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Enteritidis</i> ( $n = 10$ )	0/10	0/10	0/10	0/10	0/10	0/10
<i>S. Typhimurium</i> ( $n = 10$ )	0/10	0/10	0/10	0/10	0/10	0/10
<i>S. Weltevreden</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Agona</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Hadar</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Heidelberg</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Poona</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Braenderup</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Albany</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Oslo</i> ( $n = 1$ )	1/1	0/1	0/1	0/1	0/1	0/1
<i>S. Kibi</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Newport</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Tshiongwe</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Uppsala</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Richmond</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Bardo</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Emek</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Kissi</i> ( $n = 1$ )	1/1	0/1	0/1	0/1	0/1	0/1
<i>S. Virchow</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Bordeaux</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Regent</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Java</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>S. Farsta</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Shigella dysenteriae</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Shigella flexneri</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Shigella sonnei</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Shigella boydii</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Vibrio cholerae</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Enterohemorrhagic E. coli</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Enteropathogenic E. coli</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Aeromonas hydrophila</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Yersinia enterocolitica</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1
<i>Klebsiella pneumoniae</i> ( $n = 1$ )	0/1	0/1	0/1	0/1	0/1	0/1

This is the first report on the use of genes STY0307, STY0322, STY0326, STY2020, and STY2021 as *S. Typhi*-specific diagnostic markers. Unlike other *S. Typhi* PCR targets that were selected based on immunological properties, these genes are individually highly specific for *S. Typhi* and therefore can be used as single-gene target PCR assays without the need for nested or multiplex PCR. Also, these targets are whole gene sequences (from start to stop codon for the purpose of whole gene amplification) unlike other diagnostic markers which are only partial gene sequences. The idea of using this strategy is that if the whole gene sequence is specific to the bacteria then primers can be designed at any location

of the gene. Thus, these gene sequences not only serve as specific targets for PCR assay, but also are suitable for more advanced diagnostic tests that require multiple DNA sites, such as loop-mediated isothermal amplification (LAMP) and strand displacement amplification (SDA) which requires multiple primer annealing sites [34]. These genes could be utilized for the development of innovative Point-of-Care (POC) diagnostics to address the need for low-cost, simple, rapid, and accurate diagnostics for low resource settings.

The gene STY0201 has been used as a PCR target, and the PCR assays that were developed based on this gene were reported to be 100% sensitivity and specificity [35, 36].

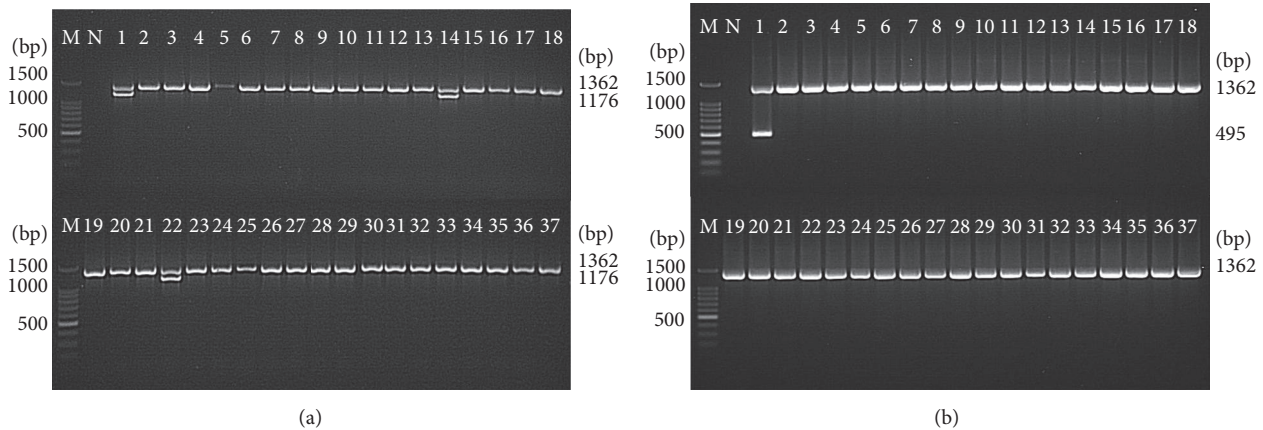


FIGURE 2: Analytical specificity of the PCR assay for detection of (a) STY0201 and (b) STY0307 genes (representative figures). Lane: M = 100 bp DNA ladder (Promega); N = negative control; 1 = *S. Typhi*; 2 = *S. Paratyphi A*; 3 = *S. Paratyphi B*; 4 = *S. Paratyphi C*; 5 = *S. Enteritidis*; 6 = *S. Typhimurium*; 7 = *S. Weltevreden*; 8 = *S. Agona*; 9 = *S. Heidelberg*; 10 = *S. Poona*; 11 = *S. Hadar*; 12 = *S. Braenderup*; 13 = *S. Albany*; 14 = *S. Oslo*; 15 = *S. Kibi*; 16 = *S. Newport*; 17 = *S. Tshiongwe*; 18 = *S. Uppsala*; 19 = *S. Richmond*; 20 = *S. Bardo*; 21 = *S. Emek*; 22 = *S. Kissi*; 23 = *S. Virchow*; 24 = *S. Bordeaux*; 25 = *S. Regent*; 26 = *S. Java*; 27 = *S. Farsta*; 28 = *Shigella dysenteriae*; 29 = *Shigella flexneri*; 30 = *Shigella sonnei*; 31 = *Shigella boydii*; 32 = *Vibrio cholerae*; 33 = *Enterohemorrhagic Escherichia coli*; 34 = *Enteropathogenic Escherichia coli*; 35 = *Aeromonas hydrophila*; 36 = *Yersinia enterocolitica*; and 37 = *Klebsiella pneumonia*. The PCR amplicon sizes for genes 16S rRNA, STY0201, and STY0307 were 1326 bp, 1176 bp, and 732 bp, respectively.

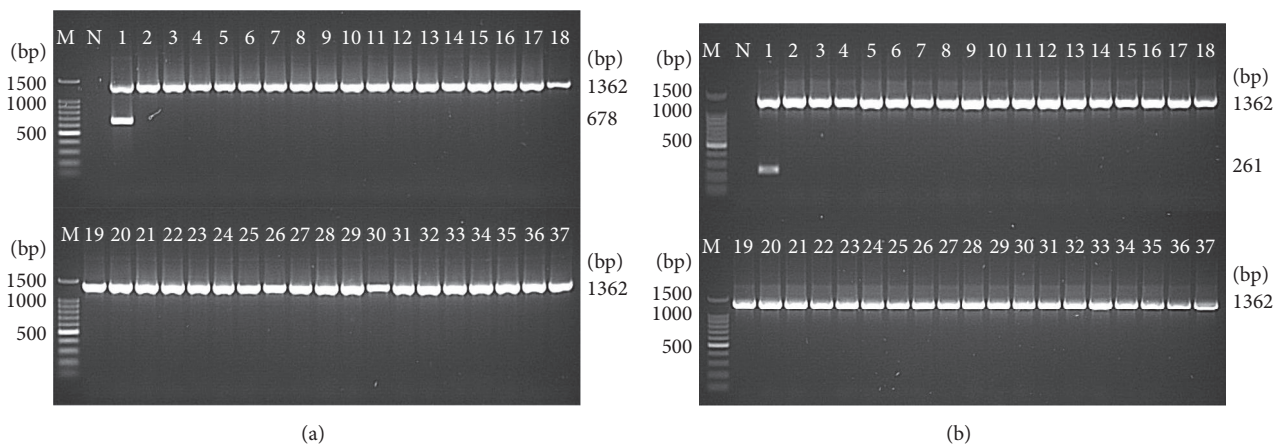


FIGURE 3: Analytical specificity of the PCR assay for detection of (a) STY0322 and (b) STY0326 genes, respectively (representative gels). Lane: M = 100 bp DNA ladder (Promega); N = negative control; 1 = *S. Typhi*; 2 = *S. Paratyphi A*; 3 = *S. Paratyphi B*; 4 = *S. Paratyphi C*; 5 = *S. Enteritidis*; 6 = *S. Typhimurium*; 7 = *S. Weltevreden*; 8 = *S. Agona*; 9 = *S. Heidelberg*; 10 = *S. Poona*; 11 = *S. Hadar*; 12 = *S. Braenderup*; 13 = *S. Albany*; 14 = *S. Oslo*; 15 = *S. Kibi*; 16 = *S. Newport*; 17 = *S. Tshiongwe*; 18 = *S. Uppsala*; 19 = *S. Richmond*; 20 = *S. Bardo*; 21 = *S. Emek*; 22 = *S. Kissi*; 23 = *S. Virchow*; 24 = *S. Bordeaux*; 25 = *S. Regent*; 26 = *S. Java*; 27 = *S. Farsta*; 28 = *Shigella dysenteriae*; 29 = *Shigella flexneri*; 30 = *Shigella sonnei*; 31 = *Shigella boydii*; 32 = *Vibrio cholerae*; 33 = *Enterohemorrhagic Escherichia coli*; 34 = *Enteropathogenic Escherichia coli*; 35 = *Aeromonas hydrophila*; 36 = *Yersinia enterocolitica*; and 37 = *Klebsiella pneumonia*. The PCR amplicon sizes for genes 16S rRNA, STY0322, and STY0326 were 1,326 bp, 678 bp, and 261 bp, respectively.

However, this study found that this gene was only 97.2% specific and cross-reacted with *S. Oslo* and *S. Kissi*. The incorrect bioinformatic prediction of the specificity of gene STY0201 may be due to the incomplete genome sequence available for the 2 bacteria in the NCBI database that limit the matching accuracy of the BLASTn search. This is a limitation of the alignment-based marker identification method, as it relies on the availability of a complete genome sequence. Thus, whenever new sequence data becomes available for the target organism, the bioinformatic analysis should be

repeated to align the current diagnostic markers with the new gene sequence to ensure the specificity.

The other 5 genes identified in this study showed no sequence homology to proteins of known function using protein BLAST (BLASTp) programs. Genes STY0307, STY0322, and STY0326 encode for hypothetical proteins, while genes STY2020 and STY2021 encode for putative bacteriophage proteins. Interestingly, genes STY0307, STY0322, and STY0326 are located in the *Salmonella* Pathogenicity Island 6 (SPI-6). Yet, their role in bacteria virulence and

TABLE 3: Details of the 5 target genes and their description, antigenicity prediction, protein coverage, and identity with *S. Paratyphi A*.

Number	Target genes (NCBI locus tag)	Gene description	GC content (%)	Antigenicity prediction*	Protein coverage with <i>S. Paratyphi A</i> (%)	Protein identity with <i>S. Paratyphi A</i> (%)
1	STY0307	Hypothetical protein	43	0.66	0	0
2	STY0322	Hypothetical protein	29	0.37	21	33
3	STY0326	Conserved hypothetical protein	37	0.79	0	0
4	STY2020	Putative bacteriophage protein	42	0.66	0	0
5	STY2021	Putative bacteriophage protein	42	0.27	0	0

\*Antigenicity of the *S. Typhi* proteins predicted using SCRATCH Protein Prediction software [26].

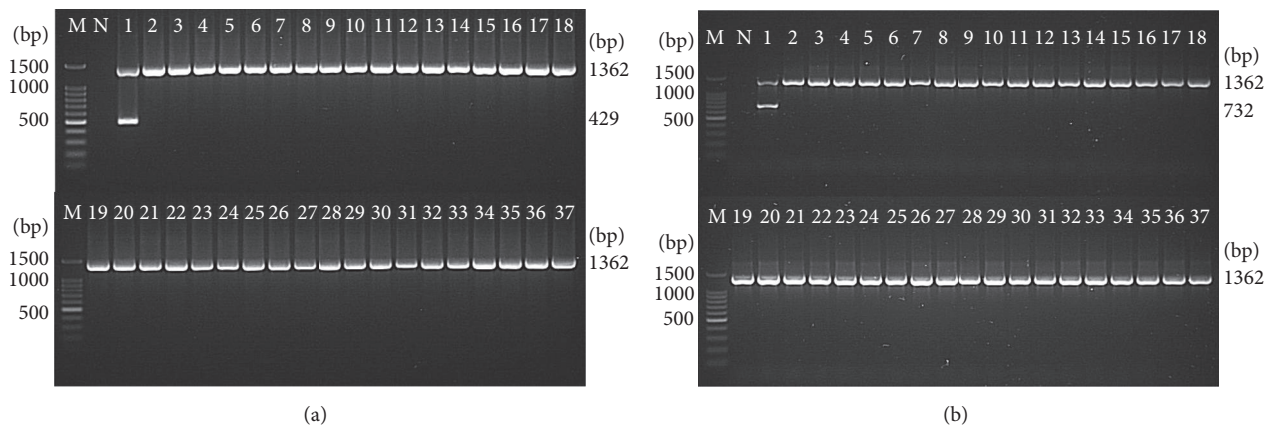


FIGURE 4: Analytical specificity of the PCR assay for detection of (a) STY2020 and (b) STY2021 genes, respectively (representative figures). Lane: M = 100 bp DNA ladder (Promega); N = negative control; 1 = *S. Typhi*; 2 = *S. Paratyphi A*; 3 = *S. Paratyphi B*; 4 = *S. Paratyphi C*; 5 = *S. Enteritidis*; 6 = *S. Typhimurium*; 7 = *S. Weltevreden*; 8 = *S. Agona*; 9 = *S. Heidelberg*; 10 = *S. Poona*; 11 = *S. Hadar*; 12 = *S. Braenderup*; 13 = *S. Albany*; 14 = *S. Oslo*; 15 = *S. Kibi*; 16 = *S. Newport*; 17 = *S. Tshiongwe*; 18 = *S. Uppsala*; 19 = *S. Richmond*; 20 = *S. Bardo*; 21 = *S. Emek*; 22 = *S. Kissi*; 23 = *S. Virchow*; 24 = *S. Bordeaux*; 25 = *S. Regent*; 26 = *S. Java*; 27 = *S. Farsta*; 28 = *Shigella dysenteriae*; 29 = *Shigella flexneri*; 30 = *Shigella sonnei*; 31 = *Shigella boydii*; 32 = *Vibrio cholerae*; 33 = *Enterohemorrhagic Escherichia coli*; 34 = *Enteropathogenic Escherichia coli*; 35 = *Aeromonas hydrophila*; 36 = *Yersinia enterocolitica*; and 37 = *Klebsiella pneumoniae*. The PCR amplicon sizes for genes 16S rRNA, STY2020, and STY2021 were 1,326 bp, 429 bp, and 732 bp, respectively.

pathogenicity remains unknown. More importantly, antigenicity prediction scores using SCRATCH protein prediction software [26] showed that genes STY0201, STY0207, STY0307, STY0326, and STY2020 were highly antigenic and may have potential to serve as antigens for serodiagnosis of typhoid fever (Table 3). When compared with the deduced amino acid sequence of *S. Paratyphi A*, which is the closest relative of *S. Typhi* [37], the putative proteins showed weak or no similarity to *S. Typhi* (Table 3). These findings provide an opportunity for gene cloning and protein expression to investigate their serodiagnostic value for development of low-cost antibody-based diagnostic tests or vaccines for typhoid fever.

In conclusion, 5 *S. Typhi*-specific genes, namely, STY0307, STY0322, STY0326, STY2020, and STY2021, were found

to be highly conserved among *S. Typhi* strains. Wet-lab experiments found no false positive reaction with non-Typhi serotypes or non-*Salmonella* enteric pathogens. These genes could serve as useful diagnostic markers for development of DNA-based diagnostics for sensitive and specific detection of typhoid fever.

### Competing Interests

The authors declare there is no conflict of interests regarding the publication of this paper.

### Acknowledgments

This work was supported by Universiti Sains Malaysia (USM) Research University Individual (RUI) Grant

(1001/CIPPM/812096) and the Ministry of Higher Education Prototype Research Grant Scheme (PRGS) (203/CIPPM/6740030). Yuan Xin Goay was a recipient of the National Science Fellowship (NSF) (M/0071/03/2010/S&T) and USM Fellowship [P-NFD0004/12(R)]. The authors would like to thank the Department of Clinical Microbiology and Parasitology, Hospital Universiti Sains Malaysia (HUSM), and INFORMM Biobank, Kelantan, Malaysia, for providing the bacterial strains.

## References

- [1] M. Akiba, M. Kusumoto, and T. Iwata, "Rapid identification of *Salmonella enterica* serovars, typhimurium, choleraesuis, infantis, hadar, enteritidis, dublin and gallinarum, by multiplex PCR," *Journal of Microbiological Methods*, vol. 85, no. 1, pp. 9–15, 2011.
- [2] G. C. Buckle, C. L. Walker, and R. E. Black, "Typhoid fever and paratyphoid fever: systematic review to estimate global morbidity and mortality for 2010," *Journal of Global Health*, vol. 2, article 10401, 2012.
- [3] T. C. Darton, C. J. Blohmke, and A. J. Pollard, "Typhoid epidemiology, diagnostics and the human challenge model," *Current Opinion in Gastroenterology*, vol. 30, no. 1, pp. 7–17, 2014.
- [4] J. N. Ja'afar, Y. X. Goay, N. F. Mohammed Zaidi et al., "Epidemiological analysis of typhoid fever in Kelantan from a retrieved registry," *Malaysian Journal of Microbiology*, vol. 9, no. 2, pp. 147–151, 2013.
- [5] T. V. T. Nga, A. Karkey, S. Dongol et al., "The sensitivity of real-time PCR amplification targeting invasive *Salmonella* serovars in biological specimens," *BMC Infectious Diseases*, vol. 10, article 125, 2010.
- [6] M. D. Ganjali, P. Abdesahian, K. Sudesh, and K. K. Phua, "Optimization of *Salmonella* Typhi biofilm assay on polypropylene microtiter plates using response surface methodology," *Biofouling*, vol. 32, no. 4, pp. 477–487, 2016.
- [7] S. A. Zaki and S. Karande, "Multidrug-resistant typhoid fever: a review," *Journal of Infection in Developing Countries*, vol. 5, no. 5, pp. 324–337, 2011.
- [8] A. I. Ugochukwu, O. C. Amu, and M. A. Nzegwu, "Ileal perforation due to typhoid fever—review of operative management and outcome in an urban centre in Nigeria," *International Journal of Surgery*, vol. 11, no. 3, pp. 218–222, 2013.
- [9] M. Hayashi, T. Natori, S. Kubota-Hayashi et al., "A new protocol to detect multiple foodborne pathogens with PCR dipstick DNA chromatography after a six-hour enrichment culture in a broad-range food pathogen enrichment broth," *BioMed Research International*, vol. 2013, Article ID 295050, 10 pages, 2013.
- [10] G. Kumar, C. B. Pratap, O. P. Mishra, K. Kumar, and G. Nath, "Use of urine with nested PCR targeting the flagellin gene (*fliC*) for diagnosis of typhoid fever," *Journal of Clinical Microbiology*, vol. 50, no. 6, pp. 1964–1967, 2012.
- [11] J. Wang, Z. Q. Xu, P. H. Niu et al., "A two-tube multiplex reverse transcription PCR assay for simultaneous detection of viral and bacterial pathogens of infectious diarrhea," *BioMed Research International*, vol. 2014, Article ID 648520, 9 pages, 2014.
- [12] K. Hirose, K.-I. Itoh, H. Nakajima et al., "Selective amplification of *tyv* (*rfbE*), *prt* (*rfbS*), *viaB*, and *fliC* genes by multiplex PCR for identification of *Salmonella enterica* serovars Typhi and Paratyphi A," *Journal of Clinical Microbiology*, vol. 40, no. 2, pp. 633–636, 2002.
- [13] J.-H. Song, H. Cho, M. Y. Park, D. S. Na, H. B. Moon, and C. H. Pai, "Detection of *Salmonella typhi* in the blood of patients with typhoid fever by polymerase chain reaction," *Journal of Clinical Microbiology*, vol. 31, no. 6, pp. 1439–1443, 1993.
- [14] S. Kolyva, H. Waxin, and M. Y. Popoff, "The Vi antigen of *Salmonella typhi*: molecular analysis of the *viaB* locus," *Journal of General Microbiology*, vol. 138, no. 2, pp. 297–304, 1992.
- [15] G. Frankel, S. M. Newton, G. K. Schoolnik, and B. A. Stocker, "Intragenic recombination in a flagellin gene: characterization of the *HI-j* gene of *Salmonella typhi*," *The EMBO Journal*, vol. 8, pp. 3149–3152, 1989.
- [16] H. M. Seth-Smith, "SPI-7: salmonella's Vi-encoding Pathogenicity Island," *Journal of Infection in Developing Countries*, vol. 2, no. 4, pp. 267–271, 2008.
- [17] E. M. Daniels, R. Schneerson, W. M. Egan, S. C. Szu, and J. B. Robbins, "Characterization of the *Salmonella paratyphi C* Vi polysaccharide," *Infection and Immunity*, vol. 57, pp. 3159–3164, 1989.
- [18] A. Kumar, Y. Balachandran, S. Gupta, S. Khare, and Suman, "Quick PCR based diagnosis of typhoid using specific genetic markers," *Biotechnology Letters*, vol. 32, no. 5, pp. 707–712, 2010.
- [19] S. Khan, B. N. Harish, G. A. Menezes, N. S. Acharya, and S. C. Parija, "Early diagnosis of typhoid fever by nested PCR for flagellin gene of salmonella enterica serotype typhi," *Indian Journal of Medical Research*, vol. 136, no. 5, pp. 850–854, 2012.
- [20] P. G. Agron, R. L. Walker, H. Kinde et al., "Identification by subtractive hybridization of sequences specific for salmonella enterica serovar enteritidis," *Applied and Environmental Microbiology*, vol. 67, no. 3–12, pp. 4984–4991, 2001.
- [21] P. Leekitcharoenphon, E. M. Nielsen, R. S. Kaas, O. Lund, and F. M. Aarestrup, "Evaluation of whole genome sequencing for outbreak detection of *Salmonella enterica*," *PLoS ONE*, vol. 9, no. 2, Article ID e87991, 2014.
- [22] S. Porwollik, R. M.-Y. Wong, and M. McClelland, "Evolutionary genomics of *Salmonella*: gene acquisitions revealed by microarray analysis," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 13, pp. 8956–8961, 2002.
- [23] F. Poly, D. Threadgill, and A. Stintzi, "Identification of *Campylobacter jejuni* ATCC 43431-specific genes by whole microbial genome comparisons," *Journal of Bacteriology*, vol. 186, no. 14, pp. 4781–4795, 2004.
- [24] J. R. Marchesi, T. Sato, A. J. Weightman et al., "Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA," *Applied and Environmental Microbiology*, vol. 64, no. 2, pp. 795–799, 1998.
- [25] N. F. Kamaruzzaman, N. J. Jaafar, M. H. Hani et al., "Pulsed-field gel electrophoresis analysis of *Salmonella enterica* serovar typhi isolates in the north-east region of Peninsular Malaysia between 2002 and 2009," *Journal of Applied Life Sciences International*, vol. 5, no. 2, pp. 2394–1103, 2016.
- [26] C. N. Magnan, M. Zeller, M. A. Kayala et al., "High-throughput prediction of protein antigenicity using protein microarray data," *Bioinformatics*, vol. 26, no. 23, pp. 2936–2943, 2010.
- [27] J. Parkhill, G. Dougan, K. D. James et al., "Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi CT18," *Nature*, vol. 413, no. 6858, pp. 848–852, 2001.
- [28] W. Deng, S.-R. Liou, G. Plunkett III et al., "Comparative genomics of *Salmonella enterica* serovar Typhi strains Ty2 and CT18," *Journal of Bacteriology*, vol. 185, no. 7, pp. 2330–2337, 2003.



- [29] S. Y. Ong, C. B. Pratap, X. Wan et al., "Complete genome sequence of *Salmonella enterica* subsp. *Enterica* serovar typhi P-stx-12," *Journal of Bacteriology*, vol. 194, no. 8, pp. 2115–2116, 2012.
- [30] D. Xu, J. O. Cisar, F. Poly et al., "Genome sequence of *Salmonella enterica* serovar Typhi oral vaccine strain Ty21a," *Genome Announcements*, vol. 1, no. 4, Article ID e00650-13, 2013.
- [31] S. M. Harish, K. S. Sim, F. Mohd Nor et al., "Complete genome sequence of *Salmonella enterica* subsp. *enterica* serovar Typhi isolate B/SF/13/03/195 associated with a Typhoid carrier in pasir mas, Kelantan, Malaysia," *Genome Announcements*, vol. 3, no. 6, Article ID e01285-15, 2015.
- [32] S. Muhamad Harish, K.-S. Sim, N. Najimudin, and I. Aziah, "Genome sequence of *Salmonella enterica* subsp. *enterica* serovar Typhi isolate PM016/13 from untreated well water associated with a typhoid outbreak in pasir mas, Kelantan, Malaysia," *Genome Announcements*, vol. 3, no. 6, Article ID e01261-15, 2015.
- [33] B. D. Cobb and J. M. Clarkson, "A simple procedure for optimising the polymerase chain reaction (PCR) using modified Taguchi methods," *Nucleic Acids Research*, vol. 22, no. 18, pp. 3801–3805, 1994.
- [34] C.-C. Chang, C.-C. Chen, S.-C. Wei, H.-H. Lu, Y.-H. Liang, and C.-W. Lin, "Diagnostic devices for isothermal nucleic acid amplification," *Sensors (Switzerland)*, vol. 12, no. 6, pp. 8319–8337, 2012.
- [35] C. B. Pratap, G. Kumar, S. K. Patel et al., "Targeting of putative fimbrial gene for detection of *S. typhi* in typhoid fever and chronic typhoid carriers by nested PCR," *Journal of Infection in Developing Countries*, vol. 7, no. 7, pp. 520–527, 2013.
- [36] G. J. Yin Ngan, L. M. Ng, R. T. P. Lin, and J. W. P. Teo, "Development of a novel multiplex PCR for the detection and differentiation of *Salmonella enterica* serovars Typhi and Paratyphi A," *Research in Microbiology*, vol. 161, no. 4, pp. 243–248, 2010.
- [37] K. E. Holt, N. R. Thomson, J. Wain et al., "Pseudogene accumulation in the evolutionary histories of *Salmonella enterica* serovars Paratyphi A and Typhi," *BMC Genomics*, vol. 10, article 36, 2009.