Review

Overview of Molecular Typing Tools For The Characterization of Salmonella enterica in Malaysia

Soo Tein Ngoi¹, Cindy Shuan Ju Teh², Lay Ching Chai¹, and Kwai Lin Thong¹, ⁶

Salmonella is a member of the family Enterobacteriaceae. This genus comprises two species, namely Salmonella enterica and Salmonella bongori. Salmonella enterica is further divided into six subspecies, namely enterica, salamae, arizonae, diarizonae, houtenae, and indica. To date, over 2500 serovars of Salmonella enterica have been described[1] on the basis of their antigenic properties defined by the somatic (O) and flagellar (H) antigens displayed on the bacterial cell surface, according to the White-Kaufmann-Le Minor Scheme[2]. Salmonella serovars can be categorized as typhoidal or non-typhoidal on the basis of their host-specificity and disease manifestations in humans. Typhoidal Salmonella serovars are host-specific, causing invasive bacteraemia only in humans. On the other hand, non-typhoidal Salmonella serovars have a wider host range, are ubiquitously found in the environment, and are capable of causing various cross-species infections[3]. Salmonella infection causes the majority of foodborne diseases worldwide, both in developed and developing countries[4-8]. According to laboratory surveillance data obtained in 2006, an estimated 93.8 million gastroenteritis cases caused by Salmonella occur worldwide on a yearly basis, resulting in 155,000 deaths[9]. Moreover, 86% of cases were related to foodborne Salmonella infections[9]. Approximately 40,000 cases of salmonellosis are reported every year in the United States (CDC, Atlanta, USA)[10-13]. In developed countries, large-scale salmonellosis outbreaks are often caused by contamination in commercially prepared food or food ingredients[14-17].

In Malaysia, the incidence of food and waterborne diseases is 48.51 cases per 100,000 people[18]. Typhoid fever (1.5%) is one of the most prevalent diseases in Malaysia[18]. The majority of other cases are generally categorized as food poisoning (92.6%), without stating the etiologic agents in reports published by the Ministry of Health (MOH) Malaysia[18]. Although not clearly stated in the MOH reports, Salmonella is frequently isolated from various sources in Malaysia, including humans, food, and animals[5-6,19-20], with Enteritidis and Typhimurium being the two most commonly isolated serovars[1-6]. These sources include clinical specimens such as stools and blood, beef, pork, chicken meat, indigenous vegetables, ready-to-eat food, buffalo, poultry and others. In the past 10 years, an overall increase in resistance to amikacin, chloramphenicol, and ciprofloxacin and a consistently high rate of resistance to tetracycline, ampicillin, and trimethoprim sulfamethoxazole have been observed among clinical Salmonella strains in Malaysia[21]. These antimicrobial agents are often used in the treatment of salmonellosis, and therefore, the increasing resistance to antibiotics among Salmonella strains is a rising public health concern.

As a means of active monitoring and surveillance of foodborne diseases, including that of Salmonella infections, the Ministry of Health Malaysia has established the Food Safety Information System of Malaysia (FoSIM), which plays an important role in monitoring the safety of imported food in Malaysia to protect consumers[22]. This monitoring system is responsible for managing and overseeing the proper notification and testing, as well as the analysis of all food consignments, imported into Malaysia prior to distribution[22]. Moreover, FoSIM also allows for online and up-to-date distribution of information regarding food safety issues (http://fsis2.moh.gov.my/fosimv2/HOM/frmHOMPage.aspx). To achieve effective surveillance of foodborne diseases and warrant successful epidemiological investigation of Salmonella outbreaks, accurate strain subtyping is of the utmost importance.

Various subtyping tools that detect genotypic and phenotypic variations among Salmonella strains have been used in investigations of the outbreaks.


1. Institute of Biological Sciences, Faculty of Science, University of Malaya, 50603 Kuala Lumpur, Malaysia; 2. Department of Medical Microbiology, Faculty of Medicine, University of Malaya, 50603 Kuala Lumpur, Malaysia
Phenotyping methods, such as serotyping and phage typing, and genotyping methods, such as pulsed-field gel electrophoresis (PFGE) and multilocus sequence typing (MLST), are some of the commonly used methods for identifying and characterizing Salmonella strains. Phenotyping methods have often faced problems, such as suboptimal discriminatory power, low throughput, technical expertise required, and time inefficiency.

Currently, mainly molecular technologies are used for detailed strain subtyping of Salmonella. Since the 1990’s, the development and application of various molecular techniques have greatly improved the resolution and speed of strain subtyping in epidemiological studies and investigations of the outbreaks of bacterial pathogens. These molecular approaches examine the genetic composition of the organisms instead of their phenotypic characteristics. Hence, molecular techniques could overcome the limitations of phenotyping methods and therefore provide the higher resolution essential for epidemiological studies. For several globally applied genotyping methods, such as PFGE, MLST, and MLVA, standardized protocols have been developed to allow reproducible data for inter-center communication.

Salmonella serovars are commonly isolated from food, humans, and animals in Malaysia. Several research groups in Malaysia have been actively working on isolation, detection, and characterization of Salmonella (Table 1). The majority of these studies were more focused on the isolation and identification of Salmonella strains rather than detailed characterization of strains using molecular techniques. Nevertheless, the use of several molecular tools in the characterization of Salmonella strains, mostly based on polymerase chain reaction (PCR), have also been reported in some Malaysian studies (Table 1). Here, we provide an overview of several low-cost and less technically demanding molecular subtyping methods for characterization of Salmonella strains with a focus on the applicability of these techniques in Malaysia. The performance, advantages, and drawbacks for each subtyping method are assessed based on the reported studies.

**SEROTYPING AND BIOTYPING**

PCR targeting a specific serovar has often been used for the rapid detection of Salmonella. The increasingly available bacterial genomic sequences have allowed the development of PCR to conduct more specific functions, such as serotyping and biotyping. Salmonella strains, based on their surface antigens, are classified into different serovars (somatic and flagellar antigens), according to the White-Kauffmann-Le Minor Scheme. Conventionally, Salmonella serotyping has been conducted using the slide agglutination method. However, this method sometimes provides ambiguous results, which may lead to misidentification of Salmonella serovars. Therefore, conventional serotyping of Salmonella is only conducted in the Reference Laboratories of the Institute of Medical Research and the Veterinary Research Institute in Malaysia. A PCR-based approach for the serotyping of Salmonella strains has been developed to overcome the limitations presented with the use of conventional methods. The presence of genes encoding surface antigens is detected by PCR. In Malaysia, the potential of PCR serotyping was explored by Lim et al., followed by Nori & Thong. In both studies, a 100% concordance of PCR and conventional serotyping results was reported. The PCR serotyping scheme developed by Nori & Thong successfully identified 14 Salmonella serovars from five serogroups. On screening a panel of 122 Salmonella strains, it was found that 77% were completely serotyped by PCR. The successfully serotyped Salmonella strains included the most commonly encountered and clinically important Salmonella serovars in Malaysia, namely Enteritidis, Typhimurium, Weltevreden, Hadar, Typhi, and Paratyphi A and Paratyphi B. The remaining strains could not be serotyped because of the limitations of the PCR targets, i.e., the strains comprised antigens not included in the PCR scheme. A previous study in Spain also showed that PCR serotyping presented with a strong correlation (99.2%) with the conventional method, using a triple multiplex PCR scheme aiming detection of 32 serovars containing the most common somatic and flagellar antigens. In the same study, there was a 0.8% discrepancy between conventional and PCR serotyping; however, it was eventually confirmed that the PCR method was accurate, showing that the conventional method had caused misidentification. Misidentification by
### Table 1. Summary of Published Studies in Malaysia on Isolation and Characterization of *Salmonella* (1983-2013)

<table>
<thead>
<tr>
<th>Year</th>
<th>Sources</th>
<th>Identification and Characterization Methods</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1983</td>
<td>Human</td>
<td>Phage typing</td>
<td>[32]</td>
</tr>
<tr>
<td>1984</td>
<td>Human</td>
<td>Conventional serotyping</td>
<td>[33]</td>
</tr>
<tr>
<td>1994</td>
<td>Human</td>
<td>Pulsed-field gel electrophoresis; ribotyping</td>
<td>[34]</td>
</tr>
<tr>
<td>1995</td>
<td>Food</td>
<td>Conventional isolation and biochemical identification; conventional serotyping</td>
<td>[35]</td>
</tr>
<tr>
<td>1995</td>
<td>Animal</td>
<td>Antimicrobial susceptibility test; plasmid profiling</td>
<td>[36]</td>
</tr>
<tr>
<td>1995</td>
<td>Human</td>
<td>Pulsed-field gel electrophoresis; ribotyping</td>
<td>[37]</td>
</tr>
<tr>
<td>1995</td>
<td>Human</td>
<td>Pulsed-field gel electrophoresis</td>
<td>[38]</td>
</tr>
<tr>
<td>1995</td>
<td>Human</td>
<td>Conventional serotyping</td>
<td>[39]</td>
</tr>
<tr>
<td>1996</td>
<td>Environment; human</td>
<td>Pulsed-field gel electrophoresis</td>
<td>[40]</td>
</tr>
<tr>
<td>1997</td>
<td>Human</td>
<td>Conventional serotyping</td>
<td>[41]</td>
</tr>
<tr>
<td>1998</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification</td>
<td>[42]</td>
</tr>
<tr>
<td>1998</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; pulsed-field gel electrophoresis</td>
<td>[43]</td>
</tr>
<tr>
<td>2002</td>
<td>Environment; human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping; pulsed-field gel electrophoresis</td>
<td>[44]</td>
</tr>
<tr>
<td>2003</td>
<td>Food</td>
<td>Conventional isolation and biochemical identification; conventional serotyping</td>
<td>[45]</td>
</tr>
<tr>
<td>2003</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping</td>
<td>[46]</td>
</tr>
<tr>
<td>2003</td>
<td>Laboratory culture collection</td>
<td>Pulsed-field gel electrophoresis</td>
<td>[47]</td>
</tr>
<tr>
<td>2004</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping</td>
<td>[48]</td>
</tr>
<tr>
<td>2005</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping</td>
<td>[49]</td>
</tr>
<tr>
<td>2007</td>
<td>Food; human</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping; ERIC-PCR</td>
<td>[50]</td>
</tr>
<tr>
<td>2008</td>
<td>Animal; food</td>
<td>Antimicrobial susceptibility test; conventional serotyping</td>
<td>[19]</td>
</tr>
<tr>
<td>2008</td>
<td>Food</td>
<td>ERIC-PCR; PCR-restriction fragment length polymorphism; random amplified polymorphic DNA fingerprinting</td>
<td>[51]</td>
</tr>
<tr>
<td>2008</td>
<td>Food</td>
<td>Antimicrobial susceptibility test; conventional serotyping; ERIC-PCR; random amplified polymorphic DNA fingerprinting</td>
<td>[52]</td>
</tr>
<tr>
<td>2009</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional serotyping</td>
<td>[53]</td>
</tr>
<tr>
<td>2009</td>
<td>Food</td>
<td>PCR virulotyping (20 virulence genes)</td>
<td>[54]</td>
</tr>
<tr>
<td>2009</td>
<td>Laboratory culture collection</td>
<td>PCR serotyping</td>
<td>[55]</td>
</tr>
<tr>
<td>2010</td>
<td>Food</td>
<td>Most probable number (MPN) analysis; PCR identification</td>
<td>[56]</td>
</tr>
<tr>
<td>2010</td>
<td>Animal; human</td>
<td>Antimicrobial susceptibility test; PCR resistance genes profiling; plasmid profiling; pulsed-field gel electrophoresis</td>
<td>[57]</td>
</tr>
<tr>
<td>2010</td>
<td>Food</td>
<td>Conventional isolation and biochemical identification; commercial identification kit; PCR identification; pulsed-field gel electrophoresis</td>
<td>[5]</td>
</tr>
<tr>
<td>2010</td>
<td>Food; human</td>
<td>PCR serotyping</td>
<td>[58]</td>
</tr>
<tr>
<td>2010</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional serotyping; pulsed-field gel electrophoresis; REP-PCR</td>
<td>[59]</td>
</tr>
<tr>
<td>2011</td>
<td>Animal</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; commercial identification kit; heavy metal tolerance test</td>
<td>[60]</td>
</tr>
<tr>
<td>2011</td>
<td>Laboratory culture collection</td>
<td>PCR identification</td>
<td>[61]</td>
</tr>
<tr>
<td>2011</td>
<td>Food</td>
<td>Conventional isolation; MPN-multiplex PCR</td>
<td>[62]</td>
</tr>
<tr>
<td>2011</td>
<td>Animal; environment; food; human</td>
<td>Conventional biotyping; PCR biotyping; pulsed-field gel electrophoresis; REP-PCR</td>
<td>[63]</td>
</tr>
<tr>
<td>2011</td>
<td>Human</td>
<td>Antimicrobial susceptibility test; conventional serotyping; pulsed-field gel electrophoresis</td>
<td>[20]</td>
</tr>
<tr>
<td>2011</td>
<td>Food</td>
<td>Antimicrobial susceptibility test; conjugation; PCR resistance genes profiling</td>
<td>[64]</td>
</tr>
<tr>
<td>2012</td>
<td>Food</td>
<td>Plate count</td>
<td>[65]</td>
</tr>
<tr>
<td>2013</td>
<td>Animal; environment</td>
<td>Antimicrobial susceptibility test; conventional isolation and biochemical identification; conventional serotyping; plasmid extraction</td>
<td>[66]</td>
</tr>
<tr>
<td>2013</td>
<td>Animal; human</td>
<td>PCR virulotyping (22 virulence genes), REP-PCR</td>
<td>[67]</td>
</tr>
<tr>
<td>2013</td>
<td>Animal</td>
<td>Conventional isolation and biochemical identification; conventional serotyping</td>
<td>[68]</td>
</tr>
<tr>
<td>2013</td>
<td>Animal; food; human</td>
<td>Antimicrobial susceptibility test; multi-locus variable number tandem repeat analysis; PCR identification; pulsed-field gel electrophoresis</td>
<td>[69]</td>
</tr>
<tr>
<td>2013</td>
<td>Animal; human</td>
<td>Antimicrobial susceptibility test; multi-locus variable number tandem repeat analysis; PCR identification; pulsed-field gel electrophoresis</td>
<td>[70]</td>
</tr>
</tbody>
</table>
conventional serotyping is not uncommon, particularly when serotyping rough, monophasic and non-motile strains\[^{69,75}\]. PCR detection of the flagellar antigens successfully identified a monophasic variant that was conventionally serotyped as a *Salmonella* Typhimurium strain\[^{69}\]. However, the main drawback of PCR serotyping has been that if the genes encoding the H-antigen are allelically diverse or unrecognized, it will return partial or incorrect serotyping results\[^{23}\]. Nevertheless, PCR serotyping is a useful method when conventional serotyping facilities are not available. In Malaysia, the commonly isolated *Salmonella* serovars (e.g., Typhimurium, Enteritidis, etc.) generally consist of the most common second-phase flagellar antigens\[^{76}\]. Hence, the PCR method is a suitable alternative to conventional serotyping.

Biotyping discriminates *Salmonella* strains on the basis of their ability to ferment certain substrates. Most of the *Salmonella* serovars of clinical importance can be biotyped, such as those of Typhi, Paratyphi A, Paratyphi B, Typhimurium. *Salmonella* Paratyphi B strains are separated into biotype Java (d-tartrate fermenting) and biotype Paratyphi B (d-tartrate non-fermenting). Traditionally, biotyping was carried out via the phenotypic lead acetate test. PCR biotyping has been used in several studies to differentiate the *Salmonella* Paratyphi B strains isolated in Malaysia\[^{63,77}\]. These studies have augmented the PCR approach with the conventional biochemical methods. In both studies, the concordance between the two subtyping methods was high (97.7%-100%). Therefore, PCR biotyping is a promising method for the rapid and reliable identification of *Salmonella* Paratyphi B biotypes. However, the discriminatory capacity of PCR biotyping is insufficient for detailed strain subtyping for outbreak investigations and epidemiological studies.

**VIRULENCE AND RESISTANCE GENE PROFILING**

The pathogenesis of *Salmonella* depends on a wide array of virulence and antimicrobial resistance genes contained in its genome. Virulence determinants in *Salmonella* genome are the genes involved in host cell invasion (*bapA, siiE, sopB*), motility (*flic*), intracellular survival (*sseF, sseG*), plasmids (*spvB, spvC*), and ion acquisition (*corA, mgtA, mgtB*)\[^{78}\]. A previous study used virulence genes (*invA* and *spvC*) as genetic markers for the rapid identification of *Salmonella* serovars\[^{79}\]. Since then, PCR virulotyping has been used to detect and characterize pathogenic *Salmonella* strains. In this method, the strains are screened for a panel of selected virulence factors using the PCR method. Although not generally applied, this method has also been used to characterize *Salmonella* strains isolated in Malaysia. Khoo et al.\[^{54}\] typed 114 *Salmonella* strains from 38 different serovars via multiplex PCR detection of 20 virulence genes associated with *Salmonella* pathogenicity islands and quorum sensing. In the same study, all *Salmonella* strains were found to possess up to 70% of the virulence factors examined. Moreover, Khoo and colleagues reported a 100% reproducibility of the PCR virulotyping method and found the method to be rapid and effective as a molecular tool for the monitoring of pathogenic *Salmonella* strains\[^{54}\]. However, the virulence genes present were not specifically associated with any *Salmonella* serovar\[^{54}\]. Elemfareji and Thong\[^{67}\] used 22 virulence genes to subtype 181 strains from Typhi and Enteritidis serovars. The targeted genes were mostly associated with *Salmonella* pathogenicity islands and were involved in adhesion, invasion, intracellular survival, colonization, and systemic infection of *Salmonella*\[^{67}\]. These virulence genes were found to be widely distributed among the two *Salmonella* serovars examined. Several plasmid and fimbrial genes are missing from the genome of all *Salmonella* Typhi strains; while only the *cdtB* gene is completely absent in all *Salmonella* Enteritidis strains\[^{67}\]. Such diversity in virulence genes among different serovars may either be the reason or the result of host adaptation. PCR virulotyping is a rapid and highly reproducible method for effectively monitoring the pathogenic *Salmonella* strains. Virulotyping is serovar-specific to an extent, when a wide range of virulence factors are examined\[^{80}\]. Some virulence genes are specifically present exclusively in certain serovars. Therefore, if a large panel of virulence factors is examined, PCR virulotyping can differentiate *Salmonella* serovars to an extent, based on the presence of these serovar-specific virulence factors. However, this method does not further clarify the genetic relationships among the various *Salmonella* strains and whether or not they are epidemiologically related. Hence, the utility of virulotyping is limited in investigations of *Salmonella* outbreaks.

Besides virulence genes, resistance gene profiling has also been used to characterize
Salmonella serovars\textsuperscript{81}, in which PCR is used to detect the presence of genes associated with antimicrobial resistance [e.g., aac, aad, aph, strA/B, bla\textsubscript{TEM}, bla\textsubscript{CMY}, sul\textit{I}, tet(A,B,C,D), dfrA, etc.]. This method is useful in characterizing multidrug-resistant organisms and has been used by researchers to characterize antimicrobial-resistant Salmonella isolated from food (e.g., poultry and beef products, ready to eat food, etc.), animals (e.g., cattle, chicken, swine, fish, frog, etc.) and human samples (stools and blood) in Malaysia\textsuperscript{157,64}. The antimicrobials tested include tetracycline, sulfonamide, streptomycin, nalidixic acid, trimethoprim-sulfamethoxazole, ampicillin, chloramphenicol, cephalothin, kanamycin, ciprofloxacin, gentamicin, cefoxitin, amoxicillin-clavulanate, and amikacin. Resistance gene profiling, similar to virulotyping, does not provide information on genetic relationships among Salmonella strains.

**PLASMID PROFILING**

Plasmids are often associated with antimicrobial resistance and virulence in Salmonella and play an important role in the intra-species and interspecies dissemination of resistance and virulence factors\textsuperscript{82-83}. The presence of plasmids and the variability of their sizes allow for the characterization of Salmonella via plasmid profiling. In this method, the extracted plasmid DNA molecules are resolved into specific banding patterns using agarose gel electrophoresis. Plasmid profiling has frequently been used to characterize Salmonella strains isolated from humans and food animals in Malaysia\textsuperscript{36,57,66,84-85}. In all the studies mentioned, plasmids were detected in the majority of the isolated Salmonella strains, which comprised both multidrug-resistant and non-multidrug-resistant strains. Multidrug-resistant strains generally harbored more than one plasmid, but plasmids were sometimes absent. The presence of plasmids is often linked to the type and level of antimicrobial resistance among the Salmonella strains that are isolated from different sources to understand the dissemination of antimicrobial resistance determinants among Salmonella in Malaysia\textsuperscript{36,57,66,84-85}. Radu et al.\textsuperscript{36} showed, via a conjugation experiment, that the presence of plasmid is associated with tetracycline resistance in Salmonella Enteritidis. In Budiati's study\textsuperscript{66}, antimicrobial-resistant Salmonella strains were simultaneously detected in both fish and pond water.

The presence of similarly sized plasmids in the samples isolated from these two sources suggested the possible transmission of antimicrobial resistance phenotypes from the environment (or feed) to the host organisms. Similar observations were documented on duck farms by Adzitey et al.\textsuperscript{84}. The studies performed by Budiati et al.\textsuperscript{66} and Adzitey et al.\textsuperscript{84} showed that plasmid profiling aided in the identification of potential health hazards for humans working in animal farming or on the production line of meat or poultry products. Generally, plasmid profiling does not serve the purpose of measuring genetic diversity among bacterial strains. Plasmid profiling has a limited application in source identification and investigation of the outbreaks. This method can only be used to characterize Salmonella strains that harbor plasmids. Furthermore, the instability and low level of diversity of plasmids also contributes to the less-than-desirable discriminatory power of plasmid profiling for Salmonella strains\textsuperscript{86}. Therefore, plasmid profiling may help researchers in understanding the mechanisms and modes of transmission of the virulence or antimicrobial resistance determinants among Salmonella (or even other bacterial) strains; however, it is not recommended for the fine-resolution subtyping of Salmonella strains.

**TRACKING SALMONELLA STRAIN CLONALITY AND GENETIC RELATEDNESS**

In the epidemiological studies of Salmonella, molecular subtyping methods play an important role in determining genetic relationships among the isolates. Genotyping tools are essential in investigations of the outbreaks to pinpoint the possible sources of Salmonella infections. Generally, genotyping methods for Salmonella could be classified into gel- and sequence-based typing methods. According to the applicability of these methods in the Salmonella studies in Malaysia, gel-based genotyping methods, such as pulsed-field gel electrophoresis (PFGE) and PCR-based DNA fingerprinting, and sequence-based genotyping methods, such as multi-locus variable-number tandem repeat analysis (MLVA) and multi-locus sequence typing (MLST), are discussed in this review. To evaluate the use of genotyping tools for determining bacterial relatedness, multiple factors, such as reproducibility, stability, discriminatory power, and typeability, of the method need to be
considered\cite{87}. A summary of the discriminatory power of different genotyping methods used in subtyping Malaysian \textit{Salmonella} strains is shown in Table 2.

**PCR-BASED DNA FINGERPRINTING**

In Malaysia, three PCR-based DNA fingerprinting methods are used in the subtyping of isolated \textit{Salmonella} strains: random amplified polymorphic DNA (RAPD) fingerprinting, enterobacterial repetitive intergenic consensus (ERIC)-PCR, and repetitive extragenic palindromic sequences (REP)-PCR\cite{50,52,63,67}. These methods examine the genomic contents of the organisms, and are thus useful for differentiation of the strains and identification of the source for investigation of the outbreaks. ERIC-PCR and REP-PCR utilize the occurrence of short interspersed repetitive DNA sequences in the bacterial genome for DNA fingerprinting\cite{88}. Meanwhile, RAPD fingerprinting uses a single arbitrary primer to amplify random sites throughout the bacterial genome to detect polymorphisms\cite{89}.

ERIC-PCR was applied for the subtyping of \textit{Salmonella} strains originating from poultry, food, and clinical samples in Malaysia\cite{50,52}. In Tunung et al.’s study\cite{50}, ERIC-PCR could successfully discriminate \textit{Salmonella} strains from different serovars and resolve the strains from different sources (street food and clinical samples) into different clusters. Lee et al.\cite{52} applied and compared the discriminatory capacity of ERIC-PCR and RAPD fingerprinting in subtyping \textit{Salmonella} strains isolated from poultry and food samples and reported that ERIC-PCR (Simpson’s index of diversity, D=0.78) was less discriminative as compared to RAPD fingerprinting (D=0.92). Nonetheless, both ERIC-PCR and RAPD fingerprinting were able to produce serovar-specific clusters, whereby strains from Weltevreden and Agona serovars were grouped into different clusters\cite{52}.

Thong and colleagues have used REP-PCR for the genotypic characterization of \textit{Salmonella} strains isolated from humans and food animals\cite{59,63,67}. Elemfareji & Thong\cite{67} reported a concordance between the results of REP-PCR and PCR virulotyping in the grouping of \textit{Salmonella} Typhi isolated from three countries (Malaysia, Papua New Guinea, and Indonesia). When comparing serovars, the discriminatory power of REP-PCR was lower (D=0.57) in subtyping \textit{Salmonella} Typhi (D=0.57) than \textit{Salmonella} Enteritidis (D=0.81) (Table 2). Moreover, REP-PCR could not distinguish \textit{Salmonella} Enteritidis strains based on the source and year of isolation when the strains were isolated from a single locality (Malaysia)\cite{67}. The authors inferred that the lower discriminatory capacity of REP-PCR was due to the clonal

<table>
<thead>
<tr>
<th>Reference</th>
<th>REP-PCR</th>
<th>ERIC-PCR</th>
<th>RAPD</th>
<th>PFGE</th>
<th>MLVA</th>
<th>Combined Analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>[5]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[20]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.91</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[34]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.86</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[37]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.15</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[38]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.96-0.99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[40]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[44]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.93</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[50]</td>
<td>-</td>
<td>0.96-0.99</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[52]</td>
<td>-</td>
<td>0.78</td>
<td>0.92</td>
<td>-</td>
<td>-</td>
<td>0.92</td>
</tr>
<tr>
<td>[57]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[59]</td>
<td>0.96</td>
<td>-</td>
<td>-</td>
<td>0.98</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[63]</td>
<td>0.93</td>
<td>-</td>
<td>-</td>
<td>0.99</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[67]</td>
<td>0.81 (S. Enteritidis)</td>
<td>0.57 (S. Typhi)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>[69]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.99</td>
<td>0.76</td>
<td>-</td>
</tr>
<tr>
<td>[70]</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.96</td>
<td>0.82</td>
<td>-</td>
</tr>
</tbody>
</table>
nature of the endemic Salmonella serovars (Typhi and Enteritidis). However, REP-PCR generally provided good discriminatory power in subtyping other Salmonella serovars, comparable to that of pulsed-field gel electrophoresis (PFGE), although marginally lower. The discriminatory power of REP-PCR when subtyping Salmonella strains consisting of multiple serovars (D=0.96) was comparable to that of a single serovar (Paratyphi B, D=0.93). However, REP-PCR yielded less discernible bands, causing difficulties in gel analysis. In addition, the low reproducibility of REP-PCR has been a major problem. RAPD fingerprinting and ERIC-PCR are also known to present the same problem of low reproducibility. This is probably caused by the low annealing temperatures of these PCRs which allow for mismatches and lead to inconsistent DNA banding patterns. In addition, other factors, such as the DNA extraction method, model of thermocycler, differences in primer synthesis, different supplier for PCR reagents, and even technical skills of the operator at times, may affect the reproducibility of PCR fingerprinting methods.

Nevertheless, when the ease of operation, cost, and speed of analysis are considered, these PCR-based methods are suitable for Salmonella subtyping. Strict adherence to an optimized protocol is essential to achieve reproducible results. The discriminatory power can be increased by combining the analysis of two or more DNA fingerprinting methods.

**PULSED-FIELD GEL ELECTROPHORESIS (PFGE)**

PFGE remains the gold standard for Salmonella subtyping. This method provides sufficiently high discriminative power for successful source identification and investigation of outbreaks of Salmonella strains from various serovars. The high discriminatory capacity of PFGE in the subtyping of Salmonella strains is also observed among different Salmonella serovars in Malaysia (Table 2). Generally, PFGE is able to discriminate Salmonella strains from different sources with a discriminatory index of 0.90 (Simpson’s Index of Diversity, D) and above. The only exception occurred in a study that investigated the Salmonella Enteritidis strains obtained from a single hospital, wherein the D-index was only 0.15. However, in another study, PFGE was used to subtype Salmonella Enteritidis strains from different sources; the discriminatory index was relatively high (D=0.96). A previous study performed by Kerouanton et al. showed that PFGE may be less sensitive in subtyping some Salmonella serovars, including that of Salmonella Enteritidis, the most commonly encountered serovar. Nonetheless, the aforementioned molecular subtyping methods, such as MLST, phage typing, RAPD fingerprinting, and plasmid profiling, yielded suboptimal discriminatory power compared to that of PFGE. To date, PFGE is the reference method for the evaluation of new molecular typing methods.

PFGE involves enzymatic restriction of the bacterial chromosomal DNA using a rare cutter, followed by the electrophoretic separation of DNA fragments using an alternating electric current (pulsed field). The choice of restriction enzymes is based on the GC content of the bacterial species. For example, the restriction enzymes XbaI, NotI, SpeI, and SfiI are for Gram-negative bacteria; and Smal, CspI, and SgrA1 are for Gram-positive bacteria. The primary restriction enzyme for Salmonella is XbaI. However, PFGE is technically demanding, parameters, such as the amount of bacterial cells, temperature used for lysis and washing, concentration of buffers and enzymes used, pulse time, and run time, can affect the outcome of PFGE analysis. Therefore, protocol standardization is essential to obtain reproducible results. Strict adherence to a standardized protocol should be practiced by researchers and technicians to allow reproducible results for inter-laboratory comparison.

Before the standardization of PFGE protocols for subtyping Salmonella, different protocols had been used by different research groups. Since the 1990’s, Thong and colleagues had made an effort to optimize the PFGE protocol and have applied it to subtyping Salmonella strains in Malaysia. Several technical improvements were made to the PFGE protocol by Thong et al. in 2003 in an attempt to reduce the number of steps and time required (Table 3). In that study, the DNA banding pattern of the marker strain (Salmonella Braenderup, H9812) prepared according to the modified protocol was indistinguishable from that of the standardized protocol by the Centers for Disease Control and Prevention (CDC, Atlanta, USA); well-separated DNA fragments were also produced from all quality control strains provided by the CDC. Therefore, the results of the protocol modified by Thong et al. were comparable to those of the CDC PulseNet protocol. Moreover, the modified protocol is reproducible and applicable to bacterial species.
other than Salmonella (Vibrio cholerae, Pseudomonas aeruginosa, Shigella spp., and pathogenic Escherichia coli). Similar technical modifications were made to the rapid PFGE protocol published in 2006 for standardization among PulseNet participating laboratories\(^{[31]}\). To date, the modified PFGE protocol is being used in Malaysia to characterize the isolated Salmonella strains\(^{[57,63,69-70]}\). In these studies, PFGE proved to be more discriminative than plasmid profiling, REP-PCR, biotyping, antimicrobial susceptibility profiling, and MLVA in subtyping Salmonella strains in Malaysia.

**MULTI-LOCUS VARIABLE NUMBER TANDEM REPEAT ANALYSIS (MLVA)**

MLVA is a molecular subtyping technique that utilizes repetitive DNA sequences in the bacterial genome. This method assesses the variability of the genetic entity called variable number tandem repeat (VNTR). VNTRs are present at multiple loci and may vary in terms of nucleotide sequence and unit size. Such variations are often the strain-defining parameters\(^{[104]}\). MLVA has gained increasing popularity as an alternative to PFGE for subtyping Salmonella strains since 2003\(^{[105-109]}\). Most of these studies have documented the high discriminatory power of MLVA in subtyping various Salmonella serovars. MLVA subtyping of Salmonella was first described for the differentiation of genetically homogeneous Salmonella serovars, such as Typhimurium (DT104) and Enteritidis (PT4)\(^{[106-107]}\). Furthermore, MLVA has proven to be a rapid, high throughput, and relatively cheap method that could be easily be standardized for inter-laboratory comparisons\(^{[29,110-113]}\). MLVA is easier to perform than PFGE because the protocol involves only a simple PCR step followed by capillary electrophoresis. In addition, MLVA has a high reproducibility.

In Malaysia, Thong and colleagues were the first to apply MLVA to bacterial strain subtyping, including Salmonella\(^{[69-70,112-113]}\). MLVA has comparable discriminatory capacity to PFGE in subtyping Vibrio cholerae\(^{[113]}\) and Shigella sonnei\(^{[112]}\) strains in Malaysia. However, MLVA was less discriminative than PFGE in subtyping Salmonella serovars of Typhimurium and Enteritidis strains in Malaysia\(^{[69-70]}\). The discriminatory power of MLVA (D=0.76 for Salmonella Typhimurium; D=0.82 for Salmonella Enteritidis) was lower than that of PFGE (D=0.99 for Salmonella Typhimurium; D=0.96 for Salmonella Enteritidis). This finding is contrary to those of European studies, which mostly documented a higher discriminatory power of MLVA over PFGE\(^{[106,114]}\). The low discriminatory capacity of MLVA was mainly caused by the invariability and sometimes absence of certain VNTR loci used in the MLVA scheme. For instance, three VNTR loci (SENTR6, SENTR7, and SE7) in Salmonella Enteritidis strains showed extremely low polymorphism\(^{[70]}\); and two VNTR loci (STR6 and STR10pl) were largely absent in the Salmonella Typhimurium strains in Malaysia\(^{[69]}\). Nonetheless, MLVA analyses in both studies showed concordance with PFGE analyses of Salmonella Typhimurium and Salmonella Enteritidis strains, confirming the previous notion on the circulation of genetically homogeneous strains of these two serovars in Malaysia. Although less discriminative, MLVA was found to complement PFGE in further distinguishing Salmonella Enteritidis strains that could not be differentiated by PFGE (Wallace coefficient=0.76)\(^{[70]}\). This observation is in agreement with Best et al.\(^{[114]}\), who suggested that MLVA could enhance the discriminatory ability of PFGE, based on their study on Salmonella Typhimurium strains isolated from animals and humans. Moreover, the discriminatory power of MLVA can be improved by selecting more polymorphic VNTRs and increasing the number of VNTRs.

<table>
<thead>
<tr>
<th>Table 3. Technical Modifications on PFGE Protocol for Salmonella Subtyping by Thong et al.(^{[47]})</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step</strong></td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>Cell lysis</td>
</tr>
<tr>
<td>Agarose plug washing</td>
</tr>
<tr>
<td>Pre-lysis</td>
</tr>
</tbody>
</table>
loci examined. In a nutshell, MLVA subtyping alone is not suitable for the assessment of genetic diversity of Malaysian *Salmonella* strains due to their clonal nature. However, joint analyses of MLVA and PFGE may provide further insights into the genetic relationships among the strains.

One major disadvantage of MLVA subtyping of *Salmonella* strains is the serovar-specificity of this method. A total of 58 MLVA markers have been developed to subtype *Salmonella* strains, whereby different sets of VNTR loci were used to subtype different serovars\(^{[110]}\). Some \(n=15\) VNTR loci are universal among different *Salmonella* serovars\(^{[110]}\). However, other VNTR loci are unique to a certain serovar. On the other hand, PFGE is universal and requires only two or less restriction enzymes (XbaI and AvaII) for the analysis of all *Salmonella* serovars. However, efforts have been made to develop an MLVA subtyping scheme that can differentiate multiple serovars of *Salmonella*, with promising results\(^{[115]}\). Van Cuyck and colleagues selected 11 VNTR loci from the *Salmonella* Typhimurium LT2 genome and developed an MLVA scheme to differentiate *Salmonella enterica* strains isolated from humans, animals, and food. The MLVA scheme used has successfully identified 31 serovars, with strains from the same serovar tightly clustered in the phylogeny tree generated, except for serovars Derby, Schwarzengrund, Stanley, and Weltevreden\(^{[115]}\). Hence, MLVA can be a promising tool for rapidly identifying and genotyping *Salmonella* strains from a sample pool consisting of multiple serovars.

**MULTI-LOCUS SEQUENCE TYPING (MLST)**

In the bacterial genome, MLST characterization of species is based on polymorphisms in selected housekeeping genes. The PCR amplicons of the target genes are sequenced, and detailed nucleotide variations are used to determine the sequence types of the organism. Online MLST databases have been established to allow sharing of information and sequence type analysis\(^{[116]}\). The current MLST scheme that examines seven housekeeping genes, namely *aroC*, *dnaN*, *hemD*, *hisD*, *purE*, *sucA*, and *thrA* has been developed for *Salmonella*\(^{[117]}\), enabling the interlaboratory comparison of subtyping data. MLST subtyping is able to group *Salmonella* strains of multiple serovars into clusters of genetically closely related strains, which generally correspond to a serovar\(^{[117]}\). However, MLST is generally less successful in discriminating genetically similar strains from a single serovar\(^{[95-96]}\). High sequence identity and slow mutation in the housekeeping genes of strains from the same serovar may be the reason behind such limitations\(^{[96]}\). To differentiate strains of the same serovar, PFGE, which examines genetic variations at the whole genome level, is more useful\(^{[118]}\). Nevertheless, MLST remains effective as a routine subtyping tool because of the robustness of its data and support for epidemiological studies and evolutionary analyses of *Salmonella* to be carried out. In Malaysia, *Salmonella enterica* is endemic, with multiple serovars persisting in the population. Hence, MLST is potentially an effective molecular tool for subtyping *Salmonella* strains in this region. Indeed, the advent of the high throughput sequencer has made DNA sequencing more affordable, thus promoting the use of MLST for subtyping *Salmonella* Typhi strains in Malaysia (unpublished study).

**CURRENT ADVANCES IN MOLECULAR SUBTYPING AND THE POTENTIAL FOR FUTURE APPLICATION IN MALAYSIA**

Due to the limited genetic diversity of *Salmonella* strains isolated in Malaysia, especially that of the most commonly isolated serovars (Typhimurium and Enteritidis), we have limited options for molecular subtyping tools that provide satisfactory discriminatory power and reproducibility. However, current advances in genome technologies, especially the advent of next generation sequencing, have contributed greatly to developments in molecular subtyping of *Salmonella*. Next, generation sequencing (NGS) allows for high throughput genome sequencing, thus drastically reducing the time and cost required. Moreover, a continually expanding database of genomic sequences of *Salmonella* is available on publicly accessible domains such as NCBI GenBank. Such genomic database enables the sharing of information and subsequently, the improvement of molecular subtyping methods. For instance, comparative genomic analysis was used to identify serogroup-specific genes in *Salmonella* in order to develop a PCR-serogrouping scheme\(^{[119]}\). Besides that, the whole genome subtyping of *Salmonella enterica* could be achieved by comparing the core genomes in order to investigate molecular evolution, and also to identify potential gene markers for genotyping\(^{[120]}\). Whole genome sequencing has been applied to subtype *Salmonella* strains, and is proven
superior to PFGE, the classical gold standard for subtyping *Salmonella*\textsuperscript{[121]}. However, similar to any other molecular subtyping tool, whole genome sequencing needs to be supported with epidemiological data in order to determine the relatedness of the *Salmonella* strains in a particular outbreak\textsuperscript{[121]}. Due to the relatively high cost of operation, whole genome sequencing has only been used in the phylogenomic study of *Salmonella* in Malaysia\textsuperscript{[122]}, but not for routine subtyping of the organism. The vast amount of data obtained from whole genome sequencing has been used to elucidate the genomic evolution of *Salmonella* Typhi in Malaysia, via a comparison of outbreak, sporadic, and carrier strains\textsuperscript{[122]}. However, as the cost of new technology tends to decrease gradually over time, the routine application of this high-resolution genotyping tool in Malaysia is possible in the future. For a more sustainable approach using NGS, there is a need for high initial costs for infrastructure, genome data storage and bioinformatics support.

The clonal nature of *Salmonella* Enteritidis has been a major problem for investigations of the outbreak

\begin{table}[!h]
\centering
\begin{tabular}{|l|l|l|}
\hline
\textbf{Molecular Characterization Method} & \textbf{Advantages} & \textbf{Disadvantages} \\
\hline
PCR serotyping & Rapid; reliable (high concordance with conventional serotyping); easy to perform & Genetically diverse or not previously recognized H-antigens may cause misinterpretation; does not elucidate genetic relationship among strains \\
PCR biotyping & Rapid; reliable (high concordance with conventional biotyping), easy to perform & Does not elucidate genetic relationship among strains \\
Virulence genes profiling & Rapid; reproducible; serovar-specific to some extent & Does not elucidate genetic relationship among strains \\
Resistance genes profiling & Rapid; reproducible; useful in characterizing antimicrobial-resistant strains & Does not elucidate genetic relationship among strains \\
Plasmid profiling & Useful in characterizing virulent and/or antimicrobial resistant strains (only if plasmids are present) & Plasmids are unstable and often have low level of diversity; not suitable for the study of genetic diversity; not all strains contain plasmids \\
ERIC-PCR & Rapid; easy to perform; low cost & Moderate discriminatory power; low reproducibility; difficult to analyse banding patterns due to inconsistency in band intensity \\
REP PCR & Good discriminatory power (comparable to PFGE); rapid; easy to perform; low cost & Less discernible bands adversely affecting gel analysis; low reproducibility; difficult to analyse banding patterns due to inconsistency in band intensity \\
RAPD fingerprinting & Good discriminatory power (higher than ERIC-PCR); rapid; easy to perform; low cost & Low reproducibility; difficult to analyse banding patterns due to inconsistency in band intensity \\
PFGE & Good discriminatory power; high reproducibility; standardized protocol & Technically demanding; time-consuming \\
MLVA & High reproducibility; rapid; high throughput; easy to perform; low cost; robust data for inter-laboratory comparison & Moderate discriminatory power (due to invariability or absence of specific VNTR loci); serovar-specific assay \\
MLST & High reproducibility; high throughput; easy to perform; robust data for interlaboratory comparison & Relatively high cost of operation \\
\hline
\end{tabular}
\caption{Summary of Advantages and Disadvantages of the Molecular Characterization Tools for *Salmonella* Strains in Malaysia}
\end{table}

The pros and cons of each molecular tool discussed in this review are summarized in Table 4.

**CONCLUDING REMARKS**

The pros and cons of each molecular tool discussed in this review are summarized in Table 4.
and epidemiological investigations of \textit{Salmonella} infections. PFGE is the gold standard recommended for investigations of \textit{Salmonella} outbreaks, since this method is highly discriminative and 100\% reproducible with a standardized protocol, but requires at least two days to obtain results. Alternatively, REP-PCR is a second choice in such investigations as it is faster and easier to perform than PFGE, with a comparable discriminatory power. Although MLVA may also be suitable for subtyping \textit{Salmonella} strains in Malaysia, it is not commonly used in Malaysia due to the lack of suitable instrumentation in most laboratories. Conversely, PCR-based serotyping, biotyping, virulotyping, resistotyping, and plasmid profiling are recommended for general molecular characterization of \textit{Salmonella} strains, irrespective of whether or not they are epidemiologically linked. MLST is a good option to study the evolutionary relationships between \textit{Salmonella} strains. In most molecular epidemiological studies of \textit{Salmonella}, a combined use of two or more subtyping methods frequently provides higher resolution than that provided by a single subtyping method. Moreover, the decreasing cost of next generation sequencing has made the routine application of this high-resolution genotyping method in the near future more feasible developing countries with limited resources, such as Malaysia. In conclusion, currently, there is no single ideal method for subtyping bacterial strains. The choice of subtyping methods depends on various parameters, such as cost effectiveness, discriminatory capacity, reproducibility, stability, and sensitivity of the molecular tools, as well as the ease of interpretation and analysis of data. Furthermore, the interpretation of the data obtained from molecular subtyping must be linked to the background and other epidemiological data of the bacterial strains.

There is no one ideal method as each has its own limitation. Molecular subtyping methods are generally more sensitive, discriminative, reproducible, time-saving, and have higher throughput compared with the conventional phenotypic characterization methods. Hence, molecular tools are suitable for effective monitoring

\textbf{ACKNOWLEDGMENTS}

We thank the University of Malaya for the facilities provided. This work was financially supported by the University of Malaya Research Grant (UMRG) (RP003A-13BIO) and UM Postgraduate Research Fund (PPP) (PS319/2010B). NST was supported by a fellowship from University of Malaya.

\textbf{DECLARATION OF INTEREST}

All authors declare that there is no conflict of interests.

*Correspondence should be addressed to Kwai Lin Thong, Tel: 603-79674437, Fax: 603-79675908, E-mail: thongkl@um.edu.my

Biographical note of the first author: Ngoi Soo Tein, female, born in 1986, major in molecular microbiology. Received: January 20, 2015; Accepted: June 10, 2015

\textbf{REFERENCES}


