MOLECULAR ANALYSIS OF *Salmonella enterica* serovar Typhi ISOLATES USING PULSED-FIELD GEL ELECTROPHORESIS; COMPARISON BETWEEN ACUTE AND CARRIERS OF TYPHOID FEVER

by

NOR FADHILAH KAMARUZZAMAN

Thesis submitted in fulfillment of the requirements
for the degree of
Master of Science

May 2011

ACKNOWLEDGEMENTS

First and foremost I offer my sincerest gratitude to my supervisor, Assoc. Prof. Dr. Phua Kia Kien, who has guided and supported me thoughout my study with his patience and knowledge, whilst allowing me to work in my own flexible way. I attribute the level of my Masters degree to his encouragement and effort and without him this thesis, too, would not have been completed or written. I am also grateful to my co-supervisor Dr. Zaidah Abd. Rahman and Prof. Thong Kwai Lin for their guidance throughout my whole study. I am also thankful to the head of the Typhoid Research Cluster, Prof. Asma Ismail, for her invaluable guidance and advice. I wish to thank Dr. Chan Yean Yean from the Department of Clinical Microbiology and Parasitology, School of Medical Sciences, USM, for giving me permission to use the PFGE equipment.

To all my dear fellow seniors and colleagues in the laboratory, Dr. Aziah Ismail, Izzzati, Rasyada, Faezah, Mariam, Ezzy, Jack, Zafri, Siti Norazura, Haslizai, Hafiza and all members of the typhoid group at INFORMM, I would like to express my heartfelt appreciation for their encouragement and support that have helped me through numerous obstacles during my study. I would like to extend my warm and sincere thanks to my other labmates and friends in the Department of Clinical Microbiology and Parasitology; Balqis, Nik, Donn, Elina, Ang Lim, Geik Yong, Choo Yee, Low and Izana who have helped me with my work and provided a warm environment for research during my PFGE work in the department. I am grateful to all lecturers and admininstrative staff of INFORMM for helping me in every way they could and continously giving moral support during my candidature. Special

thanks goes to the USM for providing USM fellowship to support my M.Sc studies and also providing a Research University grant (Grant No: CIPPM/1001/812038) and a Postgraduate Research grant (Grant No: CIPPM/1001/8133003) to fund this study. Finally, I thank my parents and sisters for their love, understanding and endless support for me to complete this study.

TABLE OF CONTENTS

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF SYMBOLS, ABBREVIATIONS AND ACRONYMS

Symbols / Abbreviations / Acronyms	Definitions
+	Positive
-	Negative or Minus
>	Greater than
≥	Greater than or equal to
<	Less than
≤	Less than or equal to
%	Percentage
°	Degree
°C	Degree Celsius
0.5×	0.5 times
10×	10 times
cm	Centimeter
dH_20	Distilled water
DNA	Deoxyribonucleic acid
g	Gram
H_2S	Hydrogen sulphide
IgG	Immunoglobulin G
IgM	Immunoglobulin M
I	Intermediate
kb	Kilo base
kDa	Kilo Dalton
l	Liter
mg	Milligram
ml	Milliliter
mm	Millimeter
MW	Molecular weight
nm	Nanometer
OD	Optical density
pH	Potential hydrogeni
R	Resistant
RE	Restriction enzyme
rpm	Revolutions per minute
sec	Second
sp.	Species
S	Sensitive
TIFF	Tagged Image File Format
U	Unit
UV	Ultraviolet
USA	United States of America
V	Voltage
w/v	Weight per Volume
v/v	Volume per Volume
µM	Micromolar
µl	Microliter
µg	Microgram

KAJIAN ANALISA MOLEKUL TERHADAP ISOLAT *Salmonella enterica* serovar Typhi DENGAN MENGGUNAKAN KAEDAH 'PULSED-FIELD GEL ELECTROPHORESIS'; PERBANDINGAN DI ANTARA PENYAKIT TIFOID AKUT DAN PEMBAWA

ABSTRAK

Demam tifoid adalah sejenis penyakit sistemik yang disebabkan oleh bakteria *S*.Typhi, sejenis kuman Gram negatif yang hanya menjangkiti manusia. Penyakit ini adalah endemik di Kelantan, di mana kadar kejadian adalah paling tinggi di Malaysia. Walaupun penyakit ini endemik di Kelantan, namun demikian sehingga kini tiada maklumat tentang jenis-jenis strain *S*.Typhi yang wujud dan taburannya di negeri ini. Justeru itu, penyelidikan ini dijalankan untuk mengetahui kepelbagaian genetik dan taburan strain *S*.Typhi yang dipencilkan daripada pesakit dengan jangkitan tifoid akut, pembawa tanpa gejala dan sampel dari persekitaran yang diperoleh di Kelantan, pada tahun 2002 hingga 2009. Kajian molekular epidemiologi telah dijalankan dengan menggunakan kaedah Pulsed-Field Gel Electrophoresis (PFGE) untuk mencirikan 264 isolat *S*.Typhi individu yang tersedia untuk kajian ini. Sejumlah 260 isolat *S*.Typhi telah berjaya dicirikan menggunakan PFGE melalui enzim penyekat XbaI, dan 38 jenis strain telah dijumpai yang seterusnya dikenali sebagai X001 sehingga X038.

Lima jenis strain yang paling kerap dijumpai adalah X001, X002, X009, X011, dan X037 yang telah membawa kepada 76.5% daripada jumlah keseluruhan isolat yang dikaji. Strain utama adalah X001 (44%) telah dijumpai secara konsistennya pada setiap tahun dan di setiap daerah di negeri ini. Strain ini juga telah menyebabkan wabak tifoid pada tahun 2005. Dengan sedemikian, strain ini dianggap sebagai strain

‘endemik’ di atas keupayaannya untuk kekal di negeri ini. Analisis PFGE ke atas isolat *S*.Typhi yang dipencilkan daripada pembawa menunjukkan jenis strain yang sama ditemui di kalangan isolat yang dipencilkan daripada pesakit yang mengalami tifoid akut, menandakan tiada perbezaan di antara kedua-dua genom yang dapat dikenali melalui kaedah PFGE.

Analisis PFGE ke atas isolat yang dipencilkan dari sampel persekitaran telah menemui jenis strain yang turut ditemui dalam 2 isolat yang dipencilkan daripada pesakit yang mengalami tifoid akut, justeru membuktikan keupayaan *S*.Typhi untuk bertahan dalam persekitaran dan menyebabkan infeksi kepada manusia. Seterusnya, pencirian terhadap isolat *S*.Typhi yang dipencilkan semasa wabak tifoid telah berjaya menemui strain yang utama bagi setiap wabak yang berjumlah di antara 70 sehingga 88% daripada keseluruhan isolat untuk setiap wabak. Penemuan ini menunjukkan keupayaan kaedah PFGE untuk membezakan di antara kes tifoid yang terlibat dengan wabak dan yang tidak terlibat dengan wabak. Kesemua jenis corak PFGE yang ditemui telah berjaya dikatalogkan di dalam pengkalan data *Salmonella* Typhi Bacilli Bank, untuk tujuan pemantauan transmisi *S*.Typhi di negeri ini.

Secara kesimpulannya, kajian ini telah menunjukkan bahawa kaedah PFGE mempunyai peranan penting dalam penentuan molekular epidemiologi demam tifoid (contohnya: mengesan punca infeksi dan mengenalpasti transmisi yang baru atau yang berulangan). Oleh itu, pemusatan kajian penjenisan kaedah PFGE untuk semua isolat *S*.Typhi yang dipencilkan daripada kes tifoid di negeri ini adalah dicadangkan bagi memberi tinjauan molekular epidemiologi penyakit ini, yang akan membantu dalam pengenalpastian punca kepada infeksi dan seterusnya menghapuskan penyakit ini di Kelantan.

MOLECULAR ANALYSIS OF *Salmonella enterica* serovar Typhi ISOLATES USING PULSED-FIELD GEL ELECTROPHORESIS; COMPARISON BETWEEN ACUTE AND CARRIERS OF TYPHOID FEVER

ABSTRACT

Typhoid fever is a systemic disease caused by *S*.Typhi, a Gram negative bacterium that infects exclusively humans. The disease is endemic in Kelantan, which has the highest incidence in Malaysia. Despite the endemicity, not much information is known regarding the type of *S*.Typhi strains and its distribution in this region. Thus, this study was conducted to ascertain retrospectively the genetic diversity and distribution of *S*.Typhi strains isolated from patients with acute typhoid, asymptomatic carriers and environmental samples in Kelantan, between the year 2002 and 2009. Molecular epidemiology study was carried out using Pulsed-Field Gel Electrophoresis (PFGE) method to characterize 264 individual *S*.Typhi isolates available in this study. A total of 260 *S*.Typhi isolates were successfully typed by PFGE using XbaI restriction enzyme, and 38 strains were found, which were designated as X001 to X038.

The 5 most common strains were X001, X002, X009, X011 and X037 that accounted for 76.5% of the total isolates. The predominant strain was X001 (44%), which was consistently found in every year and in every district, and was the same strain that caused a major outbreak in 2005. Thus, this strain was considered as the endemic strain due to its ability to persist in this region. PFGE analysis of *S*.Typhi isolated from carriers showed the same strains distributed amongst the isolates from patients with acute infection, indicating that there was no genomic difference between them that could be identified by the PFGE method. PFGE analysis performed on isolates

from environmental samples uncovered a strain that was also found in 2 isolates from patients with acute typhoid infection, proved the ability of *S*.Typhi to survive in the environment and cause infection in humans. Characterization of *S*.Typhi isolated from 4 defined typhoid outbreaks successfully identified a predominant strain for each outbreak which accounted for between 70 to 88% of the total isolates in each outbreak. These findings showed the ability of the PFGE method to successfully discriminate between outbreak-related and non-outbreak-related typhoid cases. All PFGE patterns identified in this study were successfully catalogued in the *Salmonella* Typhi Bacilli Bank database, thus laying the foundation for a database for a development of a national database for monitoring transmission of *S*.Typhi strains in this state.

In conclusion, this study has shown that the PFGE method has a role to play in the molecular epidemiology of typhoid fever (i.e. tracing the source of infection, recognizing new or recurrent typhoid transmission). Therefore, centralization of PFGE typing for all isolates recovered from typhoid cases in the state is proposed in order to provide molecular epidemiological surveillance of the disease, which will help in the identification of the sources of infection towards eradication of the disease in Kelantan.

CHAPTER 1

INTRODUCTION

1.1 Typhoid Fever

Typhoid fever is a systemic infection caused by the bacteria *Salmonella enterica* subspecies *enterica* serovar Typhi (*S*.Typhi) (WHO, 2003). A person is confirmed to have typhoid fever when he has fever (38°C and above) for at least 3 days, and a laboratory-confirmed positive culture (blood, bone marrow or stool) for *S*.Typhi (WHO, 2003). *S*.Typhi is transmitted by faecal-oral route, acquired by ingestion of food or water contaminated by human wastes that contain *S*.Typhi (Crum, 2003). The disease is restricted to humans as its host and reservoir (Uzzau *et al*., 2002). This acute systemic illness is characterized by prolonged fever, abdominal pain, and persistent bacteremia with seeding of the reticuloendothelial system. Between 10 to 20% of patients develop acute diarrhoea after ingestion of *S*.Typhi, which may last for several days. Fevers develop within 7 to 14 days after the infection (Crum, 2003).

Physical findings in typhoid include prolonged fever, brown coated tongue, hepatomegaly, splenomegaly, confusion, decreased auditory acuity and nuchal rigidity (Crum, 2003). Rose spots may occur in some patients on their abdomen and chest (Levine, 2009). On the second week of infection, the fever persists and the patient may appear severely ill. Gastrointestinal complications, such as bleeding or perforation may occur at any time but usually at the third week of infection (Levine, 2009). The disease is treatable if diagnosed early and appropriately treated with the right choice of antibiotics.

1.2 Typhoid Carriers

Generally, a typhoid carrier is a person who carries the bacteria without showing any clinical signs or symptoms. Chronic carriers are recognized as those who still excrete *S.*Typhi in stools or urine (or are repeatedly positive for duodenal string cultures) for longer than 1 year after the onset of acute typhoid fever (WHO, 2003; Kementerian-Kesihatan-Malaysia, 2006).

WHO estimated that only about 95% of typhoid patients completely recover from the disease after treatment while the remaining 5% becoming asymptomatic carriers of the disease (WHO, 2003). The rate of carriage was reported to be slightly higher among female patients older than 50 years old, and patients with cholelithiasis or schistomiosis (Levine *et al*., 1982). The gallbladder and the liver have being suggested as the main niche for *S.*Typhi in chronic typhoid carriers (Nath *et al*., 2010). The fact that these carriers are asymptomatic made them unaware that they could transmit the disease to other people. Treatment of typhoid carriers include administration of the antibiotic Amoxicillin or Ampicillin (100 mg per kg per day) plus Probenecid (Benemid®) (1 g orally or 23 mg per kg for children) or Trimethoprim-sulfamethoxazole (1600 to 800 mg twice daily) for 6 weeks (WHO, 2003). However, in the case of individuals with gallstones, antibiotic therapy as well as cholecystectomy are recommended (Crum, 2003).

1.3 *Salmonella enterica* subspecies *enterica* serovar Typhi (*S*.Typhi)

S.Typhi is a rod-shaped, Gram negative, facultative anaerobic, non-encapsulated, flagellated bacilli belonging to the family of *Enterobacteriaceae* (WHO, 2003). It is motile with peritrichous flagella. Serologically it is positive for LPS antigen 09 and 012, protein flagella antigen H-d, and polysaccharide capsular antigen Vi (WHO, 2003). *S*.Typhi can be distinguished from other *Salmonella* serotypes by its biochemical properties that can be tested as described in Table 1.1. The genus *Salmonella* was named after the pathologist Salmon, who first isolated the organisms from animal intestines. There are currently 2,463 serotypes of *Salmonella* (Popoff, 2000). The nomenclature system used for the genus *Salmonella* is based on recommendations from the Centers for Disease Control and Prevention (CDC), USA (Brenner *et al*., 2000). Figure 1.1 shows the nomenclature of *Salmonella.*

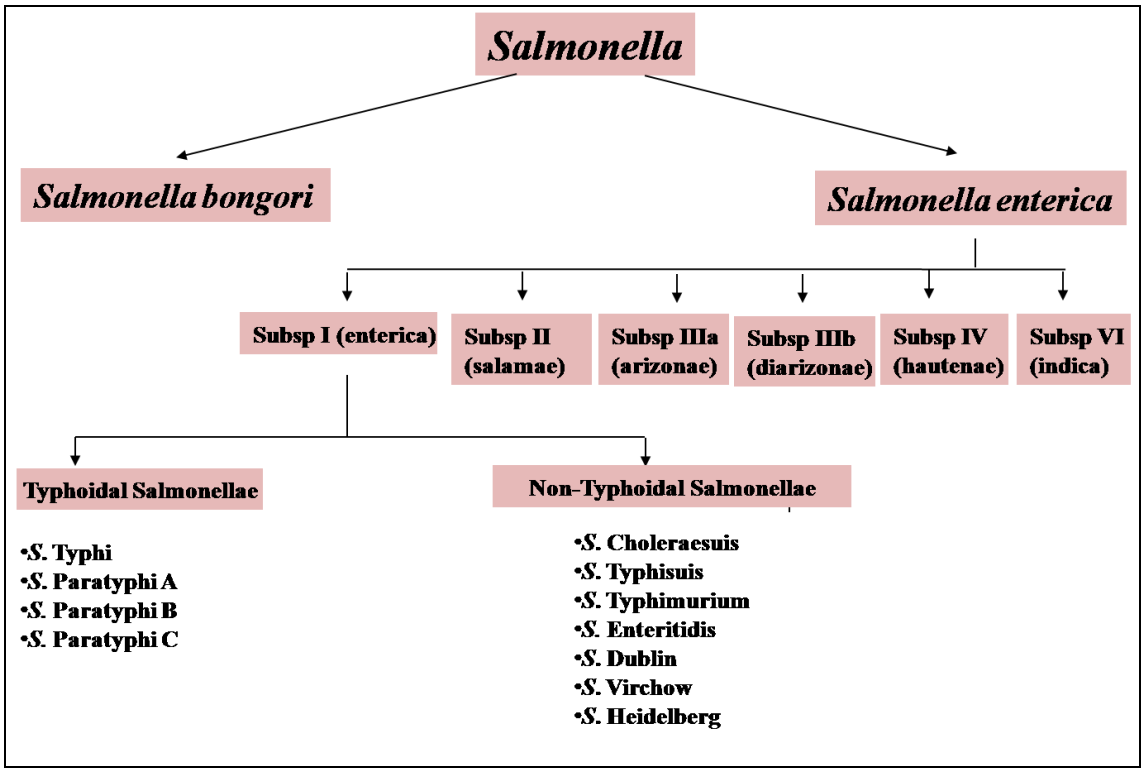

Figure 1.1: Nomenclature of *Salmonella* according to the Kauffman-White scheme (Adapted from CDC, USA, 2000)

Table 1.1: Biochemical tests for differentiation of *S*.Typhi from related subspecies (Adapted from WHO, 2003)

Organism	Kliger's Iron Agar Test				Motility, Indole, Urea Tests			Citrate Test
	Slant	Butt	H_2S	Gas	Motility	Indole	Urea	-
S.Typhi	Alkaline	Acid	Wk+	-	+	-	-	-
S.Paratyphi A	Alkaline	Acid	-	+	+	-	-	-
Other *Salmonella* sp.	Alkaline	Acid	V	V	+	-	-	V

'+' = Positive
'-' = Negative
Wk + = Weak positive
V = Variable result
H_2S = Hydrogen sulphide

1.4 Laboratory Diagnosis of Typhoid Fever

1.4.1 Isolation and Identification of *S*.Typhi (Culture Method)

Diagnosis of typhoid fever is dependent on the isolation of *S*.Typhi from blood, bone marrow, or specific anatomical lesions of the patient. Bone marrow aspirate culture is the gold standard for the diagnosis of typhoid fever, and is particularly valuable for patients who have been treated, who have a long history of illness, and for those with repetitive negative blood cultures (WHO, 2003). However, bone marrow aspirate from patient is difficult to obtain and relatively invasive (Bhutta, 2006). For these reasons, blood samples are widely used for diagnosis of the acute stage of the disease. There are a few selective media which are frequently used in the isolation of *S*.Typhi. On MacConkey agar, *S*.Typhi produces lactose non-fermenting smooth colorless colonies. On Deoxycholate Citrate agar, *S*.Typhi produces lactose non-fermenting colonies with black centers. On Xylose Lysine Deoxycholate agar, *S*.Typhi produces transparent red colonies with black centers. The non-lactose fermenting colonies from agar plates need to be identified by biochemical tests, as described in Table 1.1, followed by serotyping tests using specific antisera for *S*.Typhi (WHO, 2003)

1.4.2 Serological Diagnostic Method

Unlike bacterial culture methods that take 3 to 5 days to perform, serological diagnostic tests are used for rapid diagnosis of typhoid fever. There are a few serological methods currently being used to diagnose typhoid fever.

1.4.2 (a) Felix-Widal Test

This test works by measuring agglutinating antibodies against O and H antigens of *S*.Typhi, in sera of typhoid patients. However, this test only provides moderate sensitivity. It has been reported that as much as 30% of patients, who tested positive by bacterial culture, were negative by this test (WHO, 2003). Also, the target antigen used in this method, which are the O and H antigens, are shared with other *Salmonella* serotypes, can lead to false positive results (Clegg *et al*., 1994). As such, this test is also low in specificity. However, the Felix-Widal test is still in use today in many areas that cannot afford more expensive diagnostic methods (Benedikt *et al*., 2010).

1.4.2 (b) TYPHIDOT® Test

This diagnostic test kit is developed under license from USM based on the principle of the dot-blot enzyme immune-assay (Dot-EIA), for detecting IgM and IgG antibodies that are specific for a 50 kDa antigen in *S*.Typhi (Ismail *et al*., 1991b). The 50 kDa protein has been found to be a specific protein on the outer membrane of *S*.Typhi (Ismail, 1991a). The test showed sensitivity greater than 90% and specificity of 75% (Choo *et al*., 1994). A subsequently test kit referred as TYPHIDOT-M was later developed to detect specific IgM anti-50 kDa antibody by first removing IgG from the test sera. In convalescence or possibly acute relapse cases of typhoid fever, IgG is boosted by the secondary immune response and could mask IgM detection (Ismail *et al*., 1998). This test was reported to be useful in typhoid endemic areas, since it can differentiate between convalescence and new cases. TYPHIDOT-M has since been shown to have a sensitivity of 92% and a specificity of 100% in independent evaluation studies for all age groups of typhoid fever (Bhutta, 2006).

1.5 Epidemiology of Typhoid Fever

The most recent exercise to estimate the global morbidity and mortality burden of typhoid fever was reported in the year 2000, where it was estimated that there were 21.6 million cases of typhoid fever with 216,510 deaths (Crump *et al*., 2004). This estimation was based on blood-culture positive cases in 22 population-based studies (Crump *et al*., 2004). However, the true incidence of typhoid fever may be underestimated because the clinical picture is often confused with other febrile infections, and the facility and expertise needed for performing bacteriological confirmation tests is often lacking in less developed countries (Ochiai *et al*., 2008).

The incidence of typhoid and other enteric fevers is high in the less developed countries where fecal contamination of water and food is common and hygiene practice is primitive that leads to outbreaks of the disease (Bhutta, 2006). Typhoid remains a major problem in South- and South-east Asia, the Middle-east, North-east Africa, Sub-Sahara Africa, Central America, and some parts of the Caribbean Islands (Crump *et al*., 2004). South-east Asia was recognized as a high endemic region by WHO in the year 2000 (Figure 1.2) (Crump *et al*., 2004). In Malaysia, the incidence rate of typhoid fever was estimated at 1.9 per 100,000 population in the year 2004 (Kementerian-Kesihatan-Malaysia, 2006). The state with the highest number of cases is Kelantan, located in the east of Peninsular Malaysia, where in the year 2005, the incidence rate recorded in this endemic state was 56.7 per 100,000 population (Kementerian-Kesihatan-Malaysia, 2006). However, in the year 2009, the incidence of typhoid in Kelantan decreased to 5.2 per 100,000 population (Figure 1.3). This may be due to the continuous intervention programs by the Kelantan Public Health Department of the Ministry of Health Malaysia, and other health agencies such as

Hospital Universiti Sains Malaysia and the Institute for Research in Molecular Medicine (INFORMM), Universiti Sains Malaysia, in their concerted efforts to eradicate the disease.

Figure 1.2: Global distribution of typhoid fever in the year 2000 (Adapted from Crump *et al*., 2004)

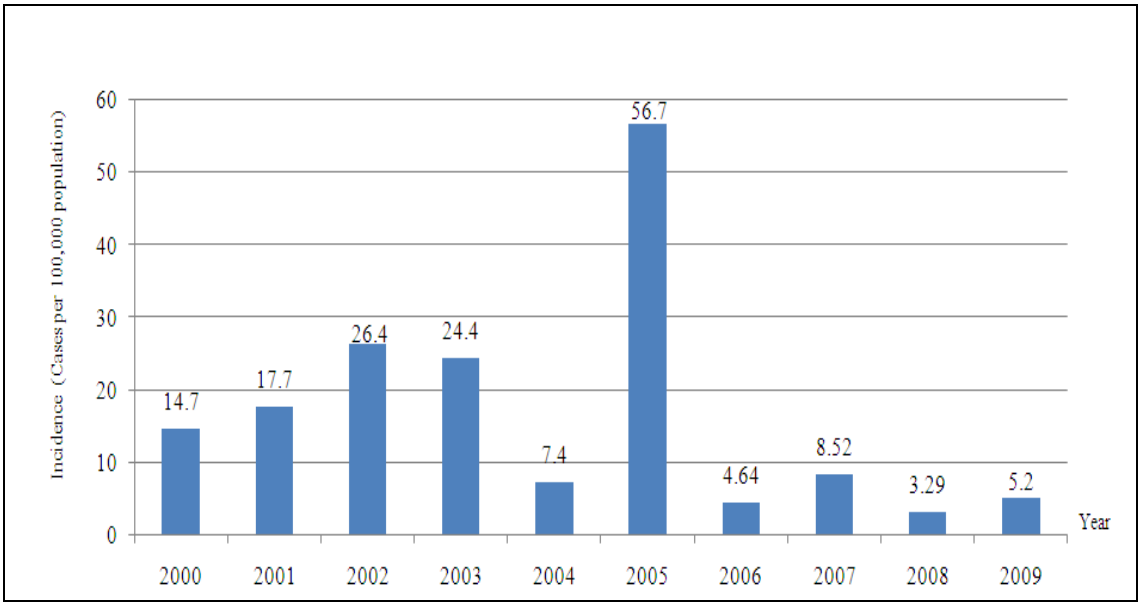

Figure 1.3: Incidence of typhoid fever in Kelantan per 100,000 population between the year 2000 and 2009 (Data from the Public Health Department of Kelantan, 2010)

1.6 Multi-drug Resistance of *S*.Typhi

Failure of treatment due to multi-drug resistant strains of *S*.Typhi to the first line of antibiotics have been reported in Pakistan, India, Vietnam, Indonesia, Korea, Chile, and Bangladesh (Paniker & Vimala, 1972; Anderson, 1975; Rowe *et al*., 1997). Beginning in the year 1991, the use of third generation Cephalosporin group of antibiotics (such as Ceftriaxone) and Fluoroquinolones group of antibiotics (such as Ciprofloxacin) for the treatment of typhoid was recommended (Trujillo *et al*., 1991). Today, they are extensively used in both developing and developed countries. Recently it has been reported that the emergence of *S*.Typhi strains that have additional resistance towards Ciprofloxacin, made the treatment for typhoid more challenging (Kownhar *et al*., 2007). In Malaysia, studies by the National Antibiotic Resistance Surveillance group conducted in the year 2000 found that 10.6% of *S*.Typhi isolated in government hospitals showed resistance to Ampicillin and 8.5% resistance to Chloramphenicol (Kementerian-Kesihatan-Malaysia, 2006).

1.7 Molecular Epidemiology

Molecular epidemiology is a discipline that combines ‘molecular’, i.e. the use of techniques in molecular biology to characterize the genotype and phenotype, with ‘epidemiology’, i.e. the study of determinants of disease occurrence and its distribution in human populations (Foxman & Riley, 2001). Previously, the traditional way of monitoring the spread of infectious diseases was by performing structured questionnaire interviews to the patients involved, either by public health workers or by primary health care providers involved (Bruisten & Schouls, 2010). Studies on pathogen dissemination that rely solely on these subjective data can yield bias to the information, especially if the data required involve sensitive information from the patients (Bruisten & Schouls, 2010). The field of ‘molecular epidemiology’ arose when the application of molecular typing tools e.g. strain typing, was found to be useful to elucidate epidemiological features of infectious diseases which were previously achieved only by using statistical method (Riley, 2009). Tibayrence (2009) has broaden the definition of molecular epidemiology as i) the definition, identification and tracking of relevant pathogenic species, subspecies, strains, clones and genes by means of molecular detection technology and evolutionary biology, and ii) the evaluation of the impact of a pathogen’s genetic diversity on its relevant medical effects, such as multi-drug resistance trait.

The basic concept of molecular typing for epidemiological purposes is the understanding that isolates of an infectious agent that are part of the same chain of transmission are clonally related; i.e. the progeny of the same ancestor cell are identical or closely related to each other (Struelens, 1998; Olive & Bean, 1999). Thus, clonally related isolates exhibit significantly more similar characteristics than

epidemiological unrelated isolates. Within a population of microbial pathogens of the same species, extensive genomic and phenotypic diversity, which reflects the evolutionary divergence that arise from mutation and gene efflux, make the discrimination of the species beyond the strain level possible using molecular typing methods (Struelens, 1998). The markers that distinguish the pathogen at the strain level are called epidemiological markers and these markers can be resolved using typing techniques at the DNA level. Bacterial epidemiological typing generates isolate-specific genotypic or phenotypic characters that can be used to elucidate the sources and routes of spread of the bacteria (Van Belkum *et al*., 2007). In the infectious disease field, the ability to discriminate the pathogen beyond the strain level has provided opportunity or possibility for intervention and prevention of the disease (Tenover *et al*., 1997; Van Belkum *et al*., 2007).

Application of strain typing in the epidemiological setting is performed for several reasons; i) aiding in source tracing to elucidate whether there is transmission from patient-to-patient or from environmental-to-patient (Branderburg *et al*., 1996; Gutierrez, 1998; Grundman *et al*., 2005), ii) to reveal phenotypical properties of the pathogen, whether certain types of pathogen possesses virulence factors that needs extra attention from the clinician and the public health workers to ensure this type of pathogen is not further spread, e.g. multi-drug resistant strains (Al-Sanouri *et al*., 2008), iii) assessing treatment activity whether patients experience relapse of the pathogen or possibility being re-infected by another strain. This application has been found to be useful to differentiate between relapse and new cases in typhoid, tuberculosis and meliodosis (Wain *et al*., 1999; Maharjan *et al*., 2005; Narayanan, 2004), and iv) monitoring the geographical spread of endemic important virulent

strains, as well as recognizing emergence of new strains of pathogen within a region in prospective surveillance of certain diseases, e.g. spread of methillin resistant *Staphylococcus aureus* (Sitchenko *et al.,* 2009).

Molecular epidemiology discipline has been applied successfully in public health services to monitor the transmission pattern of diseases and the possibility of outbreaks that might otherwise be undetected because of the misleading nature of the disease (Foxman & Riley, 2001). This approach has been extensively used to study and solve problems in infectious diseases (Maslow *et al*., 1993), including food-borne associated diseases (Barret *et al*., 1994; Chiou *et al*., 2000; Schalch *et al*., 2003; Gebreyes *et al*., 2006; Oliviera *et al*., 2007; Mullner *et al*., 2010), hospital acquired associated diseases (Lelievre *et al*., 1999; Kalkanci *et al*., 2007; Loomba *et al*., 2010), and sexually transmitted diseases (Spaargaren *et al*., 2001; Starnino *et al*., 2008; Morris *et al*., 2009).

1.8 Types of Typing Methods

Typing methods are divided into 2 methods; *phenotypic* methods and *genotypic* methods (Maslow *et al*., 1993). Phenotypic methods are those that characterize the products of gene expression in order to differentiate strains. These include biotyping, antimicrobial susceptibility patterns and serotyping (Maslow *et al*., 1993; Tenover *et al*., 1997; Singh *et al*., 2006; Van Belkum *et al*., 2007; Goering, 2010). Typing methods based on phenotypes have been found to be useful for discrimination of bacteria at the species level, e.g. serotyping for identification *Salmonella*'s serotypes (Van Belkum *et al*., 2007). Conversely, genotypic methods are those that are based on direct analysis of the genetic structure of an organism using molecular-based typing methods that allow differentiation of organism at the strain level (Maslow *et*

al., 1993; Tenover *et al.*, 1997; Van Belkum *et al*., 2007; Singh *et al*., 2006; Goering, 2010). These methods assess variation in the genomes of bacterial isolates with respect to composition (e.g. presence or absence of plasmids), overall structure (e.g. restriction endonuclease profiles, number and positions of repetitive elements), or precise nucleotides sequence (of 1 or more genes or intergenic regions) (Van Belkum *et al*., 2007). These methods can be divided into 3 general categories; i) those based on restriction enzymes analysis of the bacterial DNA, ii) those based on polymerase chain reaction (PCR) amplification of particular genetic targets, and (iii) those based on the identification of DNA sequence polymorphisms (Singh *et al*., 2006; Goering, 2010) .

Within restriction enzyme-based methods, few methods are available today including plasmid analysis, Restriction Fragment Length Polymorphism (RFLP) analysis, Ribosomal typing (Ribotyping), Insertional Sequences detection (IS)-RFLP, and Pulsed-Field Gel Electrophoresis (PFGE). Within amplification-based methods, the methods available include Amplified Fragment Length Polymorphisms (AFLP), Random Amplified Polymorphic DNA PCR (RAPD-PCR), Repetitive element PCR (Rep-PCR), and Multiple Locus Variable Number of Tandem Repeat Analysis (MLVA). Within sequence-based methods, the methods available are Multilocus-Sequencing typing (MLST), and Single Nucleotide Polymorphisms (SNP) analysis (Van Belkum *et al*., 2007; Singh *et al*., 2006; Li *et al*., 2009; Goering, 2010).

The method chosen for molecular epidemiological purposes must be based on the questions that need to be answered. For example, to understand the global distribution and evolution rate of a certain organism, methods that focus on a slowly evolved molecular marker (e.g. housekeeping genes), such as Multilocus Enzyme

Electrophoresis (MLEE) and MLST are more suitable. MLST involve amplification of the housekeeping genes and then sequencing to detect the changes in the bases (Maiden *et al*., 1998). This method has become the method of choice because of its automated procedure and centralization of all data generated in a database, which is available for all MLST users (Aanansen & Spratt, 2005). The centralization of data has allowed tracing and monitoring of the movement of pathogens worldwide (Maiden, 2006). However, for analyzing and understanding organisms in small geographic places and within a restricted time period (such as during an outbreak), methods which are more sensitive that can detect recent evolution of organism is needed. These methods include AFLP, RAPD, PFGE, MLVA and SNPs typing (Van Belkum *et al*., 2007; Singh *et al*., 2006; Li *et al*., 2009; Goering, 2010).

All typing methods must meet several criteria in order to be broadly useful. The criteria include typeability; all organisms within a species must be typeable by the methods used (Maslow & Mulligan, 1996; Van Belkum *et al*., 2007). The method used also must have high discriminatory power; which is being able to clearly differentiate unrelated strains, such as those that are geographically distinct from the source organism and at the same time be able to demonstrate the relationship of all organisms isolated from individuals infected from the same source (Maslow & Mulligan, 1996; Van Belkum *et al*., 2007). Another important criteria for typing methodologies is reproducibility. This refers to the ability of a technique to yield the same result when a particular strain is repeatedly tested. The typing method must also satisfy epidemiological concordance, which refers to its capacity to correctly classify isolates into the same cluster of isolates from a well described outbreak (Maslow & Mulligan, 1996; Van Belkum *et al*., 2007).

1.9 Pulsed-Field Gel Electrophoresis

Pulsed-Field Gel Electrophoresis (PFGE) was first introduced in 1984 by Schwartz and Cantor as a method to separate large DNA molecules (Schwartz & Cantor, 1984). Conventional electrophoresis, which relies on separation of DNA in a constant electric field, was only able to separate DNA up to 50 kb in size. The separation is inversely proportional to the size of the genome. The larger the molecule, the slower it migrates through the agarose gel (Maule, 1998). However, when the molecular size exceeds the threshold of 50 kb, all fragments exhibit size dependent mobility (Carle *et al*., 1986). Due to that, the larger molecules will undergo reptation (snaking movement), and co-migrate with the smaller molecules.

The introduction of PFGE has overcome this problem as this method is able to separate molecules as large as 12 MB. The technique relies on a direct current electric field that periodically changes direction relative to the agarose gel. The time interval, during which the field is in any 1 direction, is called the pulse time (Maule, 1998). The duration of pulse time is the single most important factor in determining the molecular size range over which separation is possible. In response to changes in the electric field, large DNA molecules are able to migrate through the agarose matrix by zig-zag orientation (Maule, 1998). This technique has allowed further manipulation of larger DNA fragments including physical mapping of eukaryotic genome.

Taking advantages on this method, it has been applied to complement the chromosomal restriction fragment analysis performed on the genome of bacteria. The conventional preparation of DNA samples in the form of a solution is not suitable for

PFGE purposes, because large DNA molecules are susceptible to shearing forces, leading to strand breakage (Maule, 1998). To overcome the mechanical shearing, DNA is prepared and protected by encapsulation in ultrapure low-melting temperature agarose. Briefly, the procedure involves embedding organisms in agarose, lysing the organisms *in situ* and digesting the chromosomal DNA with restriction endonuclease that cleaves the DNA at specific sites (Tenover *et al*., 1995; Herschleb *et al*., 2007). The resultant fragments are then separated using the PFGE apparatus. The separated fragments are visualized and comparison of each pattern from different sets of bacterial isolates can be made to determine their relatedness.

PFGE is 1 of the most reproducible and highly discriminatory method available, and it is generally the method of choice for many epidemiologic evaluations. Thus it is considered the 'gold standard' for bacteria subtyping (Tenover *et al*., 1997; Singh *et al*., 2006; Goering, 2010). It has been applied to at least 40 pathogens or pathogens group (Singh *et al*., 2006) which made it the most widely used method for typing bacteria. The PFGE method has been successfully applied in CDC, USA during the food-borne outbreak caused by *Escherichia coli* O157:H7 (*E.coli* O157:H7) which contaminated hamburgers served in a fast food-restaurant chain in western United States in the year 1993. The method was found suitable to be used during outbreak investigations due to its ability to characterize individual strains, both from clinical and food isolates of *E.coli* O157:H7. Since then, the method has been implemented by the reference public health laboratories in the USA for food-borne bacterial disease surveillance. In 1996, CDC established the national molecular subtyping network for food-borne disease surveillance, known as PulseNet (Swaminathan *et al*., 2001). This network works by standardizing all PFGE protocols as well as

gathering PFGE patterns in the national database submitted by reference public health laboratories in USA, thus providing an early warning system for food-borne disease outbreaks (Swaminathan *et al*., 2001). In 2006, PulseNet International was established, that gather reference laboratories in different continents including Canada, Europe, the Asia Pacific region and Latin America. This establishment has allowed rapid recognition multi-national food-borne disease outbreaks for early intervention (Swaminathan *et al*., 2006).

Standardization of protocols to ensure reproducibility of the method has been achieved through developments of PFGE protocols for organisms involve in food-borne disease including *Salmonella* by the PulseNet (Ribot *et al*., 2006). The HARMONY project, funded by the European Union (EU) was dedicated for harmonization of the PFGE protocols for Methicillin-resistant *Staphylococcus aureus* in the European countries, to ensure reproducibility of the method in order to track global spread of the organism (Murchan *et al*., 2003). These 2 important projects by the CDC and EU groups showed that the PFGE method has greatly been accepted and implemented everywhere. The method is almost standard for all organisms varying only in a few chemicals and restriction enzymes for Gram-negative and Gram-positive bacteria, making the method easier to be implemented in reference laboratories for routine analysis for different types of organisms.

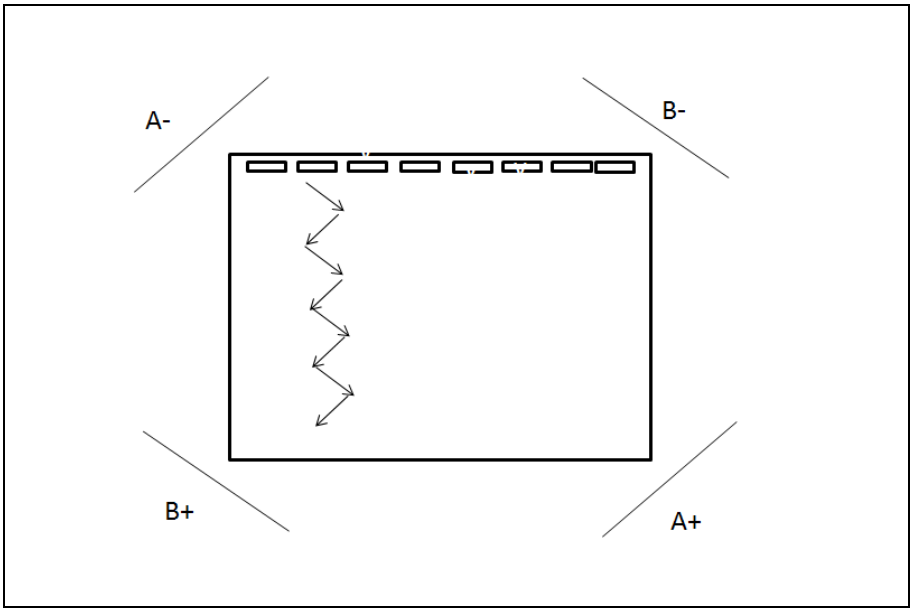

Figure 1.4: General principle of PFGE

An agarose gel is represented by the box; the series of short horizontal rectangles indicate the wells into which the DNA plugs are loaded. A and B represent 2 sets of electrodes. When the A electrodes are activated, the DNA is driven anodically down ward and to the right as indicated by the first arrow. When the A electrodes are turned off, the B electrodes are immediately activated. The DNA now moves downward and to the left. The path of the DNA molecules is subjected to continued alternation of field directions as shown by the arrows (Sambrook & Russell, 2000).

CHAPTER 2

AIMS AND RATIONALE OF STUDY

Typhoid fever remains a public health problem in the state of Kelantan, Malaysia. Typhoid cases are not only reported during outbreaks, but also continually throughout the year as sporadic cases. Since typhoid is a notifiable disease in Malaysia, every confirmed case of typhoid must be notified to the Public Health Department. However, despite the burden of typhoid in Kelantan, molecular epidemiology research has never been conducted to understand the extent of genomic diversity of *S.*Typhi in this region; i.e. whether all typhoid cases were due to a single strain or diverse strains.

Molecular epidemiology involves the study of microorganisms beyond subspecies level using typing methods such as PFGE. PFGE method has been successfully applied to study many microorganisms that caused important diseases in humans, such as *Salmonella* species, *Staphylococcus* species, *Campylobacter* species and *Shigella* species, to understand their epidemiological distribution as well as their strain diversity. Understanding the epidemiological distribution as well as the diversity of *S.*Typhi is important as this will determine whether typhoid cases in this region are due to a single or multiple sources that act as reservoir for the bacteria, or whether they are foreign acquired due to migration. To gain insight into *S.*Typhi strains that circulate in this region, specific emphasis was directed to the analysis of *S.*Typhi isolates recovered during 4 defined outbreaks using PFGE. Outbreaks occur when large numbers of people are infected simultaneously by bacteria from the same

source, whereas sporadic cases occur as isolated individual cases that are not apparently linked with each other.

Typhoid outbreaks are recognized when the number of cases exceeds the number of normal sporadic cases. Outbreaks usually occur within a defined period of time and in a specified district or location. It is not known whether strains that caused sporadic cases are also involved in triggering outbreaks of typhoid fever, or whether there are distinct strains that cause sporadic and outbreak infections. Thus, the PFGE method was applied to study isolates recovered from outbreaks and sporadic cases, in order to ascertain whether there are any genomic differences between them. A laboratory method that has the ability to differentiate between strains that cause outbreaks and strains that cause sporadic cases and thus to define its socio-economic burden to the community could be very helpful to the Public Health Department, as this kind of information can assist them prioritize their efforts to find the source of the transmission.

Having large numbers of people being infected within a short period of time in a particular location made the transmission easily recognizable so that specific efforts to locate and break the chain of transmission could be enforced. However, when there is no outbreak occurring during a period of time, connection between these sporadic cases are hard to tell. For this reason, *S.*Typhi isolates recovered as sporadic cases should be analyzed using PFGE to study if there are any genomic similarity between them, or if there are any clusters of infection that are dispersed in different places in this region.

WHO has estimated that approximately 5% of typhoid patients develop into carrier state 1 year after their acute infection. These typhoid carriers harbour the bacteria in their gallbladder, but remain asymptomatic and therefore ideal reservoir for the bacteria. Occasionally, the bacterium is secreted together with bile into the gastrointestinal system. Unhygienic typhoid carriers involved in preparation of food will have the risk of spreading the bacteria to other people, either as sporadic cases or as typhoid outbreak. Therefore, it is very important to identify typhoid carriers, as this will help break the cycle of typhoid transmission in this region. Diagnosis of typhoid carriers is not only important to prevent transmission of the disease, but early detection of carriers can prevent the development of gallbladder cancer, as the bacteria residing in the gallbladder can act as a mutagen and inducing carcinoma in the host (Shukla *et al*., 2001).

The mechanism behind the development of typhoid carriers remain unknown, as not much is known regarding factors involved in the development of the carrier state; whether the manifestation of the disease is due to a lack of immunity of the host, or due to virulent strain(s) of *S*.Typhi, or probably both factors may be involved. No study has been carried out to correlate between the typhoid carrier state and specific strains of *S*.Typhi that infect them, probably due to the difficulty of isolating the organism from carriers. Taking the advantage of having *S*.Typhi isolates recovered from 14 typhoid carriers in Kelantan, this study was initiated to find any genomic correlation between the bacteria and the disease state of the subject in which the bacteria was isolated from. PFGE, the gold standard method for fingerprinting of organism, was used to study genomic differences between *S*.Typhi strains isolated from acute and carrier states of the disease. Comparison of isolates, both from

patients with acute typhoid as well as typhoid carriers, was made to see if there were any strain(s) that could be specific for typhoid carriers, thus demonstrating its special ability to induce typhoid carrier state in humans. If this is found to be true, patients infected by these specific strains of *S.*Typhi should be given special attention in terms of treatment before they are discharged, and follow-up laboratory investigations can be thorough, because they have a higher chance of becoming carriers, compared to other typhoid patients. As typhoid endemicity has also been associated with contamination of water systems, this study was also conducted to ascertain whether water has been the reservoir for typhoid infection in this region, and the correlation was carried out by comparing *S.*Typhi isolates recovered from water as well as acute typhoid patients in the locality. In summary, the objectives for this study are:

Using Pulsed-Field Gel Electrophoresis;

1. To ascertain the diversity of *S.*Typhi isolated from patients with acute typhoid, typhoid carriers, asymptomatic food handlers, and the environment in Kelantan between the year 2002 and 2009.
2. To identify specific strains of *S.*Typhi from typhoid carriers by comparing isolates from acute typhoid patients and typhoid carriers.
3. To profile *S.*Typhi isolates associated with outbreaks of typhoid fever between the year 2002 and 2009.
4. To assess any possible transmission among sporadic cases of typhoid fever by comparing PFGE profile of *S.*Typhi isolates from different districts in Kelantan between the year 2002 and 2009.
5. To identify correlation between outbreak and sporadic cases of typhoid fever.

6. To identify possible transmission of *S.*Typhi within household members during typhoid infection.
7. To ascertain the correlation between drinking water sources and acute typhoid infections in Kelantan.

CHAPTER 3

MATERIALS AND METHODS

3.1 Materials

3.1.1 Bacterial Culture

A total of 264 *S*.Typhi isolated between the year 2002 and 2009 were included in this study. This include *S*.Typhi isolates recovered from patients with acute typhoid (n=249), typhoid carriers (n=4), asymptomatic food handlers (n=10) and also environmental sample (n=1). *S*.Typhi isolates from patients with acute typhoid were obtained from the Microbiology and Parasitology Department, School of Medical Sciences, Universiti Sains Malaysia and Kuala Krai Hospital. All isolates included in this study must satisfy at least 2 epidemiological information, the date of isolation and also the origin (district) of patients which isolates were recovered from.

Salmonella Braenderup from the American Type Culture Collection (ATCC) was used as the standard marker for all PFGE analysis, as recommended by the PulseNet network, CDC, USA (Hunter *et al*., 2005).

All isolates were stored in -70°C in INFORMM Bacterial Bank.

All bacterial isolates were maintained in the Nutrient broth-glycerol (15%) and kept in -70°C freezer. For experimental purposes, the organisms were grown in Nutrient broth at 37°C for 18 hours.

3.1.2 Media and Chemicals

All media and chemical used are listed in Table 3.1.

All the media prepared for experimental purposes were autoclaved at 121°C, 15 psi for 15 minutes and tested for sterility by incubating at 37°C overnight to check for contamination. Then, the media were stored at 4°C until used.

Table 3.1: List of chemicals, reagents and media used in this study

Name	Supplier
Nutrient broth	Oxoid Ltd. UK
Nutrient agar	Oxoid Ltd. UK
MacConkey agar	Oxoid Ltd. UK
Xylose lysine agar	Oxoid Ltd. UK
Triple sugar iron agar	Oxoid Ltd. UK
Mueller-Hinton agar	Oxoid Ltd. UK
Urease agar	Oxoid Ltd. UK
Sulphate Indole Motility medium	Oxoid Ltd. UK
Simmons Citrate agar	Oxoid Ltd. UK
Methyl red agar	Oxoid Ltd. UK
Antisera poly O, poly H, H-d, Vi and 09	Remel, UK
Methyl red solution and Kovac reagents	Remel, UK
Tris	Bio-Rad USA
EDTA	Sigma, USA
Boric Acid	Sigma, USA
Sarkosyl	Sigma, USA
Proteinase K	Invitrogen, USA
Antibiotics; Ampicillin (10µg), Ciprofloxacin (5µg), Ceftriaxone (30µg), Nalidixic Acid (30µg), Trimethoprim-Sulfamethoxazole (25µg) and Chloramphenicol (30µg).	Becton, Dickinson, USA
Seakem Gold Agarose	Lonza, USA
Low Melting Agarose	Lonza, USA
Restriction endonucleae XbaI	Fermentas, USA