Chapter 1 General introduction

1.1 The Genus *Salmonella* : Taxonomical and disease classification

The taxonomy of *Salmonella* is exceedingly complex. At present, the genus *Salmonella* is recognized to consist of only one species. This single *Salmonella* species has been named *Salmonella enterica* (LeMinor, 1984). *Salmonella* enterica is comprised of 7 subspecies. The subdivision into this sub-species is dependent on biochemical reactions of the strains with dulcitol, lactose, o-nitrophenol-β-P-galactopyranoside (ONPG), salicin, D-tartarate, mucate, gelatinase, sorbitol and potassium cyanide (KCN). This biochemical sub-division has been supported by DNA/DNA hybridization and serological methods. The formal sub-species epithets are *Salmonella enterica* subsp. Enterica (1), subsp. salamae (11), subsp. arizonae (111a), subsp. diarizonae (111b), subsp. houtenae (1V), subsp. bongori (V) and subsp. indica (V1) (LeMinor, 1988). Each subspecies in turn is divided into serovars/serotypes according to the O-antigen (somatic) and H-antigen (flagellar) based on the Kauffmann-White antigenic scheme (Kauffmann, 1966). Based on this scheme more than 2,375 *Salmonella* serovars have been identified (Popoff and Leminor, 1992; Popoff et al., 1994). *Salmonella* are widely distributed in nature with humans and animals being the primary reservoirs. Majority of the *Salmonella* strains (99.7%) isolated from humans belong to subsp. 1.

1.1.1 Salmonellosis

Salmonellae are the etiologic agents of different diseases collectively referred to as salmonellosis. Human salmonellosis can be divided into 4 syndromes : enteric fever

(typhoid like disease), gastroenteritis (food poisoning), bacteremia with or without gastroenteritis and the asymptomatic carrier state. The first two being the most important.

(a) Enteric fever

Typhoid fever is an acute systemic illness caused by the infection with *S. typhi*. Paratyphoid fever is a pathologically and clinically similar but generally milder illness that is caused by *S. paratyphi* A, *S. paratyphi* B and *S. paratyphi C* (Hoffman, 1991).

(b) Gastroenteritis

The commonest recognized clinical syndrome caused by Salmonellae is gastroenteritis which can be caused by any of the serotypes. The most common serotypes associated with the non-typhoidal Salmonellosis are *S. typhimurium* and *S. enteritidis* (Rodriguez et al., 1990; Philips and George, 1994; Pang et al., 1995).

1.2 Overview of *Salmonella typhi*

1.2.1 Characteristics

Salmonella typhi is classified under *Salmonella enterica* subsp. enterica (1). It is a gram-negative, facultatively anaerobic and non-spore forming bacilli. It has 3 major antigens : H-antigen (flagellar), O-antigen (somatic) and Vi-antigen (capsular). Most strains are motile by means of peritrichous flagella but there exists non-motile strains. *S. typhi* produces acid from glucose, maltose and sorbitol without the production of gas, does not ferment lactose, sucrose, rhamnose and few other sugars. Nitrite is produced from nitrate and hydrogen sulfide production is observed. Optimum growth temperature is 37°C and they are able to grow on a wide range of relatively simple media that distinguishes it from other members of the Enterobacteriaceae family (Holt, 1984).

1.2.2 Epidemiology

S. typhi is a human pathogen that causes typhoid fever. Humans are the only host for *S. typhi*. Typhoid fever is acquired by the ingestion of food or water that has been directly or indirectly contaminated by excreta (primary feces) of infected persons including carriers (Weissfeld et al., 1994).

Typhoid fever prevails mostly in developing or third-world countries where it normally causes a challenge to public health. In general the incidence of this disease is decreasing worldwide (e.g. in Thailand, Chile and Peru) but it must be noted that certain regions continue to have high incidences of this disease (e.g. Indonesia, Vietnam and Papua New Guinea). There are approximately 17 million cases of typhoid fever annually with nearly 600,000 deaths mainly in Asia and Africa (Pang et al., 1998a). Countries which have reported the highest annual incidences are Papua New Guinea and Indonesia where there are more than 1,000 cases per 100, 000 population. Another trend is that typhoid fever usually affects the young. This pattern can be seen in Indonesia where 91% of the cases occur in those 13-19 years of age (Ivanoff, 1995). High infection rates in children less than 5 years of age have also been reported in Bangladesh, India, Jordan and Iran (Ivanoff, 1998).

1.2.3 Pathogenesis and clinical manifestations

The incubation period for *S. typhi* ranges from 1 week to a month after initial ingestion of the bacteria. It is presumed to enter the body through M-cells. Diarrhoea seldom occurs at this point but ulceration of the intestine can occur during the invasion phase. *S. typhi* multiplies in the submucosal area, after which it enters the bloodstream and spreads throughout the body. Multiplication of the bacteria occurs again in the spleen and

liver. After a certain period of multiplication the bacteria will be released once more into the blood stream in large quantities. This stage of the disease, which can last for 2 to 3 weeks, is characterized by a dry cough, high fever, a flushed appearance and anorexia. Chills, convulsions and delirium can also occur. Diarrhoea is not a common feature but many patients will complain of constipation. These symptoms which are usually felt by typhoid patients are probably due to LPS-mediated release of cytokines. *S. typhi* has no known exotoxins. The bacteria eventually moves from the liver to the gall bladder and are shed in bile into the intestine for the second time. At this point, severe ulceration of the intestinal mucosa can occur and this stage is usually fatal. In some people, *S. typhi* persists in the gall bladder and can be shed in their feces for years. These individuals are known as chronic carriers (Salyers and Whitt, 1994).

1.2.4 Diagnosis, treatment and prevention

Diagnosis of *Salmonella typhi* infection is made by isolation of the organism from blood, feces, urine, bone marrow aspirate and bile. Biochemical and serological tests, including the classical Widal test are done to confirm the diagnosis (Sonnenwirth, 1980). Immunological tests such as enzyme-linked-immunosorbent assay (ELISA) have also been used to complement culture and Widal tests (Rodriguez et al., 1993).

Molecular methods such as rDNA hybridization with genes coding for *S. typhi* capsular Vi antigen has been used in complementation with serological tests to detect *S. typhi* in blood samples (Rubin et al., 1989). Hashimoto et al. (1995) developed a nested PCR based on the *via*B sequence for the detection of *S. typhi* from clinical specimens.

More recently, Lim et al. (1998) developed a simple and rapid new test (TUBEX) which aids routine diagnosis of typhoid fever. This test detects anti- *Salmonella* O9 (both

1.3.1 Phenotypic methods

1.3.1.1 Biotyping (Biochemical typing)

Biotyping refers to establishing the pattern of activity of cellular metabolic enzymes/substrate for species identification (Maslow et al., 1993).

Biotyping can be used in the epidemiological investigation of *Salmonella* as a sole typing method for serovars that cannot be typed by other available typing schemes. It can also be used in conjunction with an already established typing method such as phage typing (Barker and Old, 1989). An efficient biotyping scheme was devised by Duguid et al. (1975) for the study *S. typhimurium*. This typing scheme was performed easily and could be practised without the need for expensive equipment. Many biotyping studies on *S. typhimurium* were based on this particular typing scheme. For *S. typhimurium* biotyping has always been preferred in conjunction with phage typing to characterize/subdivide strains from this serovar (Barker and Old, 1979; Barker et al., 1980). In another example, biotyping in conjunction with ribotyping and IS200 typing showed that *Salmonella salinatis* was closely related to *Salmonella sandiego* (Old et al., 1999).

Hickman and Farmer (1978) showed that 100 *S. typhi* strains tested had uniform biochemical reactions thus limiting the usefulness of this typing method for this particular organism. A study carried out by Fica et al.(1996) showed that analysis performed by biotyping and molecular methods appears to indicate that *S. typhi* strains in Chile belongs to a single clone with a worldwide distribution described previously by Selander et al. (1990).

1.3.1.2 Serotyping

It has been recognized that microorganisms of the same species can differ in terms of the expression of antigenic determinants on the cell surface. Serotyping which involves the detection of such diffrences by the means of immunological techniques has been one of the classic basis for epidemiological studies of gram-negative bacteria including bacteria from the genus Salmonella (Maslow et al., 1993).

Classification of strains of *Salmonella* spp into serovars/serotypes are based on the diversity of the lipopolysaccharide (O) antigens and the flagellar proteins (H) antigen, a scheme which was instituted by White (1926). Kauffmann (1972) has extended this typing scheme and it currently consists of 2,375 antigenically distinct serovars (Popoff and Leminor, 1992; Popoff et al., 1994).

Serotyping is used not only for the diagnosis and routine taxonomic grouping of species, but also for a more analytical differentiation of epidemiological investigations as seen with *E.coli*, *P. aeruginosa* and *Klebsiella* spp., where serotyping has been an important typing method used for differentiating strains obtained in different epidemiological situation. On the other hand, serotyping in combination with phage typing has been an important tool for routine identification of *Salmonella* serovars which include *S. paratyphi* A and *S. typhimurium* (Meitert and Meitert, 1978).

Serotyping helps in the study of the prevalence of particular serovars (Palmer and Rowe, 1986) and identification of the serovars alone may be sufficient in outbreaks caused by uncommon serovars. For example, an outbreak of *S. typhi* infection in England and Wales in 1982 was detected and epidemiological investigations quickly identified 2 types of chocolate covered bars imported from Italy as the vehicle of infection. The prompt recognition of this outbreak prevented thousands of other infections (Gill et al., 1983).

Serotyping studies demonstrated that *S. enteritidis* displaced *S. typhimurium* in 1988 as the most common serovar causing food poisoning in humans in England and Wales (Anon, 1989).

Investigation of an outbreak of water-borne typhoid fever in Catalonia by serotyping in complementation with other phenotypic and genotypic typing methods showed that the source of outbreaks was a drinking fountain (Usera et al., 1995).

However the development of serotyping as an epidemiological test has been limited by its lack of discriminatory power and the need for high quality antisera production. Moreover, serotyping is available only in reference laboratories.

1.3.1.3 Phage typing

In 1938, the first phage typing scheme based on phage adaptation was developed and used on an international scale to differentiate *S. typhi* (Craigie and Yen, 1938 a,b). This typing scheme uses the progressive adaptation of Vi phage 11, which is phage specific for the Vi (capsular) antigen of *S. typhi* (Felix and Pit, 1934) and based on this, eleven phage types were initially identified. *S. typhi* Vi phage typing was standardized in 1947 (Craigie and Felix, 1947). By 1986, with further adaptation of Vi-phage 11, a further 95 types had been defined and internationally recognized thus bringing the total number of Vi-phage types to 106 (Edelman and Levine, 1986).

Phage typing has been the method of choice for differentiation within *Salmonella* serovars in reference laboratories since the 1950s. This technique has proven to be a valuable adjunct to serological identification and has played an important role in epidemiological investigations of *Salmonella* outbreaks.

A typhoid outbreak in Aberdeen in 1964 was suspected to be the result of contaminated canned corned beef and phage typing showed that the epidemic strain of *S. typhi* was ivolved. This typing method also revealed that a number of other typhoid outbreaks were due to *S. typhi* Vi-type 34 strains (Anderson and Hobbes, 1973). Conn et al. (1972) carried out an investigation of a small series of typhoid fever cases caused by river water in Britain between 1963 and 1970 and found that all were caused by a single source which was a rare Vi-phage type K1 of *Salmonella typhi*. The contamination of the river was due to incorrect sewage connection with a surface water drain outfall into the river.

Use of the phage typing approach for epidemiological study of outbreaks of typhoid fever is well documented (Hickman-Brenner and Farmer, 1983). More recently, phage typing and genome fingerprinting by pulsed-field gel electrophoresis was able to show that an unusually prolonged outbreak of typhoid fever (from 1988 to 1994) in Terrassa, Spain was caused by a casual food handler who was a carrier (Xercavins, et al., 1997)

Phage typing has been particularly important in support of epidemiological investigations of *S. hadar* (Rowe et al., 1980), *S. typhimurium* (Martini et al., 1992), *S. virchow* (Torre et al., 1993) and *S. enteritidis* (Laconcha et al., 1998).

1.3.1.4 Antibiotic resistance typing (antibiogram)

Antibiotic resistance typing is routinely performed by clinical microbiology laboratories in the evaluation of most bacterial isolates (Maslow et al., 1993). Even though the antibiotic pattern (antibiogram) is normally easy to determine, resistance to a range of antibiotic is often associated with the unstable extrachromosomal R plasmid, which can undergo frequent and sudden changes. The potential to produce different results are high and this reduces its discrimination power. This typing method is useful to identify recurring

resistance patterns in outbreak situations but such patterns have little value when comparative studies involving sporadic cases are considered (Towner and Cockayne, 1993).

Antibiotic resistance in *Salmonella* strains is generally encoded by plasmids which have been acquired as a consequence of selective pressure by antibiotics used in human and veterinary medicine (Anderson 1968; Threlfall et al., 1978). The majority of resistance plasmids have acquired their resistance genes by transposition, either from other plasmids in the strain, from plasmids carried by other bacterial species in the host, or from the chromosome or spontaneous mutation of chromosomal genes which again can result from sudden selective pressures. Therefore, antibiotic resistance typing cannot be satisfactorily used as a primary method for discrimination within serovars. This typing method can be used in conjunction with other methods such as serotyping, biotyping and phage typing for epidemiological purposes (Threlfall and Frost, 1990).

Chloramphenicol resistant *S. typhi* strains belonging to different phage types which caused major outbreaks in Mexico, India and South-East Asia in the 1970s were succesfully characterized by antibiogram typing (Anderson, 1975; Rowe and Threlfall, 1984).

Antibiogram typing in conjunction with phage typing showed the existence of 2 multidrug resistant (MDR) strains of *S. typhi* which caused a typhoid epidemic in Tajikistan. It is the first report of epidemic ciprofloxacin-resistant S. typhi (Murdoch et al., 1998).

Wain et al.(1998) reported the emergence of MDR *S. typhi* strains during the 1990s in the Mekong delta of Vietnam using resistance typing and plasmid analysis. There was a rapid increase in quinolone resistant *S. typhi* strains.

Hitchcock and Brown (1983) showed that LPS profiles within a serovar were very homogeneous while slight variations were detected between serovars. Another study by Jimenez-Lucho and Foulds (1990) suggested that LPS heterogeneity may exist among *S. typhi* strains. However, LPS analysis has not been widely used for studying the epidemiology of *Salmonella* spp.

1.3.1.6 Multilocus enzyme electrophoresis (MLEE)

Multilocus enzyme electrophoresis (MLEE) has long been a standard method in eukaryotic population genetic and systematic studies. Recently it has been used in large scale studies to estimate the genetic diversity and structure in natural populations of a variety of bacteria. Selander et al. (1986) established basic population genetic frameworks for the analysis of variation in serotypes and other phenotypic characters and has provided extensive data for the systematic study and useful marker systems for epidemiology.

In MLEE, genetically controlled variants of enzymes are detected by shifts in electrophoretic mobility as a consequence of one or more amino acid substitutions which results in a change in net electrophoretic charge. The alloenzymes are identified by comparing the relative mobilities of the enzyme variants on the gel after histochemical staining with specific substrate (Gargallo-Viola, 1989).

The first study of the genetics of natural populations of a bacterium using enzyme electrophoresis was performed in 1973 by Milkman. Allelic variations of the five enzyme loci were measured to test the neutral theory of molecular polymorphism and evolution. This pioneering work in evolutionary genetics was then extended by Selander and Levine (1980). The primary objective of most of these investigations was to estimate the overall genomic relationships among strains within species. For this purpose, large numbers of

enzyme loci were assayed to minimize the errors in estimates of relatedness resulting from interlocus variance in allelic diversity and evolutionary rate of change.

This was demonstrated by Reeves et al. (1989) whose study on genetic relatedness of *Salmonella* serotypes was based on 24 metabolic enzymes. It was shown that 32 distinct allelic profiles or electrophoretic types (ETs) were found in the 48 strains of *Salmonella* serotypes studied, other than *S. typhi*. The data showed the potential application of MLEE in epidemiologic and taxonomic studies of *Salmonella*, although it is not differential for *S. typhi*, which belonged to a single clone.

Beltran et al. (1988) used 23 metabolic enzymes in their MLEE studies to demonstrate a clonal population structure which existed within the salmonella serotypes studied. In their studies of MLEE to evaluate evolutionary genetic relationship among *Salmonella* serotypes. Selander et al. (1990) used 24 different metabolic enzymes. The analysis revealed no close phylogenetic relationships among clones of different human adapted serovars other than *S. paratyphi* A and *S. sendai*.

Another conclusion drawn from the studies was that two clones of *S. typhi* existed, one globally distributed and another apparently confined to Africa, but both clones were distantly related to those of all other serovars studied.

A MLEE analysis of 96 strains of Salmonellae distinguished 80 electrophoretic types (ETs) and placed them in 8 groups, seven of which corresponds precisely to the 7 taxonomic groups (1, 11, 111a, 111b, 1V, V, V1). The eighth group was a variation of group 1V (Boyd et al.1996).

1.3.2 Genotypic methods

1.3.2.1 Plasmid profile analysis and plasmid fingerprinting

Plasmids are autonomous replicating extrachromosomal DNA of bacteria, coding for a number of genetic determinants that aid the bacteria in survival. The medical importance of plasmids is that many code for antibiotic resistance (R-factors) that are transferable between strains, species and genera. In the clinical laboratory, plasmid DNA is important in the epidemiological investigation of outbreaks caused by pathogenic organisms.

The development of rapid and inexpensive techniques for extracting plasmid DNA (Birboim and Doly, 1979; Kado and Liu, 1981) and separating plasmids on the basis of their size by agarose gel electrophoresis led to the widespread use of plasmids in epidemiological investigations.

Technically, plasmid profile analysis and plasmid restriction endonuclease analysis (plasmid fingerprinting) are the simplest of the DNA based typing methods. It was first used in Britain in 1979 to study *Salmonella* epidemiology when the plasmid content of antibiotic resistant strains belonging to *S. typhimurium* phage types 204 to 193 was analysed (Willshaw et al., 1980).

In 1982, plasmid profile typing and restriction endonuclease fingerprinting were used in the United States to confirm that contaminated marijuana was responsible for apparently unrelated outbreaks of *S. muenchen* amongst young adults (Schmidt et al., 1982). Weide-Botjes et al. (1996) showed that plasmid and macrorestriction analysis proved to be the most suitable for the molecular typing of *S. cholerasuis* and the differentiation of both live vaccine strains from field isolates of this serovar.

Plasmid characterization and pulsed field electrophoresis analysis demonstrated that ampicillin-resistant strains of *Salmonella enteritidis* phage type 6a are derived from *S. enteritidis* phage type 4 (Ridley et al., 1996). Multi- drug resistant *S. paratyphi* C isolates from Zaire were shown to be different from sensitive strains isolated in India, Greece, East Africa and Iraq by plasmid profiling, plasmid fingerprinting and pulsed field gel electrophoresis (Kariuki et al., 1999).

Mhand et al. (1999) described the spread of a closely related beta-lactamase producing *S. typhimurium* strain in children and nurses by plasmid analysis in conjunction with other phenotypic and genotypic methods.

Although plasmid profiles are valuable in epidemiological investigations of certain *Salmonella* serovars they are not universally applicable. Many *Salmonella* serovars are plasmid free, harbour plasmids infrequently or have serovar specific plasmid. Plasmid analysis may be of limited value in serovars such *S. dublin* where the majority of strains carry only the serovar specific plasmid (Threlfall et al., 1990),

Murray et al. (1985) showed that the majority of the *S. typhi* strains studied were plasmid-free and those that had plasmids produced digest patterns that were similar. Most of the *S. typhi* isolates from Chile and Peru studied by Maher et al. (1980) were also plasmid-free and similarities among plasmid digests were also noticed.

The importance of plasmid analysis for *S. typhi* has increased recently due to the emergence of multi-drug resistant strains. Hampton et al. (1998) were able to subdivide phage types of *S. typhi* MDR strains by plasmid profile typing and pulsed field gel electrophoresis. The rapid emergence of *S. typhi* MDR strains during the 1990s in the Mekong Delta of Vietnam was shown by plasmid analysis, antibiotic typing and other genotypic typing methods (Wain et al., 1998; 1999).

1.3.2.2 Pulsed- field gel electrophoresis

A solution to the problem of analysis caused by complex chromosomal fingerprint patterns is to use low frequency clevage restriction endonucleases to generate only a limited number of large DNA fragments (McClelland et al., 1987). Large DNA fragments which are bigger than 50 Kb cannot be separated readily by conventional agarose gel electophoresis. The only technique available currently that enables the separation of larger molecular weight DNA by physical means was introduced by Schwartz et al. (1982). The concept behind this technique was that DNA molecules larger than 50Kb could be separated by using two alternating electrical fields and was subsequently termed as pulsed-field gel electrophoresis (PFGE).

It applies the principle that after the first electric field is removed, the DNA coil must change conformation and re-orient itself before it can migrate in the direction of the second electric field. Fractionation of DNA thus occurs because the speed at which DNA molecules change their direction of movement through the gel is size-dependent (Lai et al., 1989).

Schwartz and Cantor (1984), developed the first PFGE apparatus to separate *Saccharomyces cerevisiae* chromosomes (200 to 30 000 Kb). It employed 2 alternating electric fields, one homogeneous and the other non-homogeneous. Carle and Olson (1984) developed a similar apparatus but using 2 non-homogeneous electric fields which they named orthogonal field alternating gel electrophoresis (OFAGE). The drawback of these 2 apparatus was that the fragments were not separated in straight tracks due to non-uniform electric fields. This led to the modification of Schwartz's pulse field system in 1986 by Gardiner et al. who re-orientated the electric field transversely to the gel. This system was named transversed alternating field electrophoresis (TAFE). The characteristic bent lanes

were eliminated, but the drawback was that molecules still did not move at a constant velocity over the length of gel. A process referred to as field inversion gel electrophoresis (FIGE) was described by Carle et al (1986) which helped overcome this problem.

The major advancement in PFGE took place when Chu et al. (1986) arranged 24 electrodes in a hexagonal contour which offers re-orientation angles of 60° or 120°, generating homogeneous electric fields. This system was developed into contour-clamped homogenous electric field (CHEF) electrophoresis and was used to separate yeast chromosomes.

DNA molecules as large as 12 Mb have been separated using this technique (Orbach, 1988). Due to the capability of analyzing such large fragments, physical maps have been constructed. Examples of the use of PFGE in mapping are the mapping of the *Salmonella typhimurium* LT2 genome (Liu et al., 1992), *S. paratyphi* B genome (Liu et al., 1994) and *S. typhi* Ty2 genome (Liu and Sanderson, 1995).

PFGE has been used widely in the epidemiological studies of *S. typhi*. Thong et al. (1994) reported that PFGE, following *Xba*I and *Spe*I restriction digest of *S. typhi* chromosomal DNA, is a useful method for differentiating isolates of *S. typhi* and in the molecular analysis of isolates involved in outbreaks of typhoid fever in Malaysia. PFGE analysis served as a useful tool for delineating common *S. typhi* phage types of diverse origins from Southeast Asia (Nair et al., 1994; Thong et al., 1995).

A waterborne typhoid fever outbreak in Catalonia in 1994 was shown to originate from a drinking fountain by PFGE in conjunction with other genotypic and phenotypic methods (Usera et al., 1995).

Molecular characterization of *S. typhi* isolates from various geographic locales and different epidemiological settings by PFGE has gained importance over the years. For

example, typing of MDR *S. typhi* strains from Vietnam and the Indian sub-continent (Hampton et al., 1998; Wain et al., 1999; Connerton et al., 2000; Shanahan et al., 2000) and analysis of *S. typhi* obtained from patients with fatal and non-fatal typhoid fever in Papua New Guinea (Thong et al., 1996b) The study of sporadic cases of typhoid fever in Taiwan (Tsen et al., 1999) and the diffrentiation of MDR *S. typhi* strains and *S. typhi* sensitive strains in India (Shanahan et al., 1998) are other examples.

Chromosomes from other *Salmonella* serovars have also been analyzed for epidemiological purposes. They include *S. brandenburg* (Baquar et al., 1994), *S. agona* (Threlfall et al., 1996), *S. panama* (Stanley et al., 1995), *S. cholerasuis* (Weide-Botjes et al., 1996), *S. stanley* and *S. bovismorbificans* (Puohiniemi et al., 1997), *S. enteriditis* (Laconcha et al., 1998), *S. paratyphi* C (Kariuki et al., 1999) and *S. typhimurium* (Mhand et al., 1999).

1.3.2.3 Genome analysis by specific DNA probe

Chromosomal DNA can be analyzed by examination of restriction endonuclease cleavage fragment patterns. DNA restriction patterns generated by enzymes that have frequent restriction sites is not easy as these enzymes tend to generate too many DNA fragments which do not generally separate well by conventional agarose gel electrophoresis. These complex patterns of restriction fragments can be easily analyzed by hybridization techniques which involves the use of labelled DNA probes which decreases the number of fragments to be compared (Towner and Cockayne, 1993).

Several different types of probes can be used :

a) Ribosomal DNA (ribotyping)

b) Probes derived from insertion sequences (IS200 typing)

c) Probes that consists of randomly cloned DNA fragments

d) Probes derived from various genes

1.3.2.3.1 Ribotyping

Ribotyping involves the fingerprinting of genomic DNA restriction fragments that contain all or part of the rRNA genes and/or their adjacent regions. The 2 major advantages of ribotyping is that genes coding for rRNA are highly conserved, enabling a single probe to subtype all eubacteria and secondly, most bacteria contain multiple ribosomal operons which produce a reasonable number of fragments after probing to allow interspecies discrimination (Stull et al., 1988 ; Grimont and Grimont, 1991).

Grimont and Grimont (1986) used isotopically labelled ribosomal 16S and 23S rRNA which was extracted from *Escherichia coli* to probe preparations of restriction endonuclease cleaved chromosomal DNA. The variation in copy number and size of fragments carrying rRNA genes (rDNA) was detected and used as a method for strain discrimination. The name ribotyping was coined by Stull et al (1988) who demonstrated that this method is a widely applicable system in determining the molecular epidemiology of genetically diverse gram negative organisms.

Altwegg et al. (1989) showed that analysis of ribosomal RNA gene restriction patterns allowed separation of most independently isolated *S. typhi* strains of identical phage types. The increased sensitivity over phage typing may render it useful for analysis of epidemics caused by common phage types. Similar results were confirmed by Pang et al. (1992a) who used this method to study the genetic variation among Malaysian isolates of *S. typhi*. They found considerable genetic variation among *S. typhi* strains and suggested that

rRNA gene restriction patterns are of value in differentiating individual strains and may complement phage typing for epidemiological purposes.

rRNA gene restriction pattern analysis was used by Nastasi et al. (1991) to characterize 169 strains of *S. typhi* phage types C1, C4, D1 and D4. Once again, it was suggested that this typing method appears to be a reliable tool to complement phage typing in characterizing isolates of *S. typhi*, and may effectively contribute to investigation of outbreaks and trace epidemiologically related from unrelated isolates of *S. typhi*.

Ribotyping in conjunction with other genotypic and phenotypic method were used to assess typhoid epidemics in Chile. The result of the study suggested that the Chilean epidemic was probably produced by multiple sources of infection because of deficient sanitary conditions (Fica et al., 1996). Gruner et al (1997) identified 2 clonally related *S.typhi* isolates from an outbreak of typhoid fever in Zürich by using ribotyping and PFGE. A similar study by Echeita and Usera (1998) showed that two *S. typhi* strains were involved in 3 typhoid outbreaks in Spain between 1989 to 1994.

When *Salmonella* serotypes Reading, Senftenberg and Typhimurium were characterized by ribotyping, it was found that this method was able to establish different ribotypes within a given serotype regardless of host or geographic origin of the isolates. Ribotyping could be used as a tool for the classification of *Salmonella* isolates and is complementary to phage typing, antibiotic sensitivity testing or plasmid analysis (Esteban et al., 1993). Ribotyping has also been established to be useful in epidemiological studies of other *Salmonella* serovars such as *S. enteritidis* (Powell et al., 1995), *S. brandenburg* (Baquar et al., 1994), *Salmonella* serogroup D1 (Stanley et al., 1995), *S. salinitis*, *S. duisburg*, *S. sandiego* (Old et al., 1991) and *S. typhimurium* (Guerra et al., 1997).

1.3.2.3.2 IS200 typing

The mobile genetic element IS200, which caused a mutation in the histidine operon of *Salmonella typhimurium*, first described by Lam and Roth (1983), is known to be the smallest bacterial insertion sequence. This 708 bp nucleotide sequence has been shown to be distributed on conserved loci on the chromosome of many *Salmonella* serotypes (Gilbert et al., 1990), but some later studies have shown that other bacteria like *E. coli* or *Shigella* spp. can harbour IS200 (Biseraic and Ochman, 1993). It has been recently reported by Stanley et al. (1991) that additional IS200 copies can be detected on plasmid DNA.

IS200 transposes only rarely under laboratory conditions (Casedesus et al., 1989) and its stable nature has favored its use as a probe for discrimination within *Salmonella* serotypes and for phylogenetic purposes (Threlfall et al., 1994). The resolution of chromosomal genotypes by IS200 profiling in *Salmonella* is precise and discriminating, especially when a serovar contains moderately high copy numbers of the element, as does *S. heidelberg* and *S. typhimurium* (Stanley et al., 1992b; Baquar et al., 1993).

Threlfall et al. (1993) examined *S. typhi* strains of Vi phage types A, E1 and M1 by IS200 fingerprinting. IS200 typing was shown not to be as highly discriminatory as phage typing but it was able to distinguish between drug-sensitive and drug-resistant strains of Vi-phage types E1 and M1. They suggested that IS200 fingerprinting might be useful in epidemiological investigations, particularly in investigating the origins of strains with multiple drug-restriction. However a study done by Hermans et al. (1996) showed that IS200 typing was able to distinguish sensitive and multi-drug resistant *S. typhi* strains from Bangladesh.

Threlfall et al. (1994) probed 49 Vi-phage type *S. typhi* strains with a PCR amplified IS200 gene probe and identified 14 IS200 fingerprints, in which two fingerprints

dominated. It was concluded that IS200 cannot be used as an alternative typing method and it should only be used in conjunction with more traditional discrimination methods such as phage typing which still remains the method of choice for the primary differentiation of *S. typhi*.

Stanley et al. (1993) managed to identify eight IS200 profile patterns with strains belonging to 9 of the 10 common phage types of *S.typhimurium* and it has been suggested that IS200 typing may be suitable for the genotypic subdivision of this serotype. This finding was also supported by Baquar et al. (1993) who managed to differentiate a single phage type of *S. typhimurium* DT 204C into several chromosomal subtypes using IS200 probe generated by PCR. In another study, Baquar et al. (1994) showed that IS200 profiling was able to subtype one of the most common phage types of *S. typhimurium*, phage type 193 and it was concluded that this method is appropriate either for primary strain discrimination or for the subdivision of certain phage types.

Olsen et al. (1994) discovered that *S. enteritidis* strains from 33 phage types could be grouped into two clonal lines and have suggested that IS200 typing be used in parallel with ribotyping when performing phylogenetic or epidemiological studies on this serovar. IS200 typing was also found to be useful for the strains belonging to *Salmonella* serogroup D especially those lacking in plasmids (Stanley et al., 1994).

Application of IS200 typing and ribotyping, combined with classical microbiological methods, was able to distinguish the five *Salmonella* serotypes 06, 7:C:1,5 groups : Cholerasuis sensu stricto, Cholerasuis var Kunzeddorf, Cholerasuis var Decatur, Paratyphi C and Typhisuis (Uzzau et al., 1999).

Clonal lineage studies based on IS200 profiling have also been carried out with *S. berta* (Stanley et al., 1992a), *S. heidelberg* (Stanley et al., 1992b), *S. livingstone* and *S.*

eimsbeuttel (Crichton et al., 1998). IS200 as a genotypic subtyping scheme has been used successfully for S. *paratyphi* B and *S. java* (Ezuerra et al., 1993) and *S. infantis* (Pelkonan et al., 1994).

Weide-Botjes et al.(1998) and Chrisholm et al.(1999) showed that IS200 typing had limited value in the epidemiological studies of *S. hadar* and *S. thompson* as these serovars lacked the IS200 element or had none at all.

1.3.2.3.3 Randomly cloned probes

Chromosomal fingerprinting was first demonstrated by Tompkins et al. (1986) who used a randomly cloned DNA fragment of unknown function to differentiate between strains of Salmonella species. These probes highlighted restriction site heterogeneity and the typing scheme managed to differentiate *S. typhimurium* and *S. dublin* strains but was unsuccessful in the case of *S. enteritidis*.

Waterhouse and Glover (1993) developed a method which enables the cloning and identification of procaryotic repetitive DNA suitable for use as DNA fingerprinting probes of gram positive and gram negative organisms. Fragments generated by RAPD-PCR were used succesfully by Hermans et al. (1996) to detect multi-drug resistant *S. typhi* strains from Bangladesh.

1.3.2.3.4 Various gene probes

The use of specific gene probes as molecular markers for epidemiological purposes have increased over the years. Hybridization of restricted chromosomal DNA with various random gene probes was able to differentiate virulent and non-virulent *Toxoplasma gondii* strains (Sibley and Boothroyd, 1992). The existence of 2 distinct genetic subgroups among

eight *Streptococcus pyogenes* strains was shown by hybridization assays with a M protein gene (Kaufhold et al., 1992).

Smeltzer et al. (1996) showed that genomic fingerprinting with the *cna*, *fnb*A, *fnb*B and *hlb* gene probes can provide an important epidemiological tool for the identification of clinical isolates of *S. aureus*. Southern blot hybridization analysis using the *vip*R gene showed that the *via*B regions of *Salmonella typhi*, *S. paratyphi* C and *S. dublin* were identical as compared to *Citrobacter freundii* (Hashimoto and Khan, 1997). Shanahan et al. (1998) used specific gene probes in hybridization experiments to identify multi-drug resistant *S. typhi* strains harbouring plasmid mediated antibiotic resistant genes.

1.3.2.4 Genome fingerprinting by polymerase chain reaction

The polymerase chain reaction (PCR) for amplification of specific nucleic acid sequences was introduced by Saiki et al. (1988). The basic principle of PCR is that it uses a thermostable polymerase to amplify specific nucleic acid sequences quickly and exponentially (Towner and Cockayne, 1993).

The potential of PCR was recognized in the early stages as this method offered a sensitive approach to the detection and identification of specific microorganism even in trace amounts from a variety of samples such as clinical and environmental specimens (Towner and Cockayne, 1993). Some of the applications of PCR to the specific detection of the genus *Salmonella* are as follows. A PCR based test was developed by Jae-Hoon et al. (1993) for the detection of *Salmonella typhi* in the blood specimens of patients with typhoid fever using 2 pairs of oligonucleotide primers that were able to amplify a 343bp fragment of the flagellin gene of *S. typhi*. The same studies were carried out by Chaudry et

al. (1994) for the detection of *S. typhi* in the blood of patients using primers to amplify a 486bp fragment from the specific region of the dH flagellin gene of *S. typhi*.

Zhu Qing and Lim (1994) designed unique PCR primers based on the 5S and 23S RNA gene sequences of *S. typhi* Rawlings (type strain) to detect *S. typhi* from food samples. They were able to detect *S. typhi* genomic DNA as low as 0.1 pg. Hashimoto et al. (1995) devised a combination of PCR primers based on the *via*B region (Vi antigen) and, using a nested PCR strategy, they were able to use it to specifically detect *Salmonella typhi* from clinical specimens including blood samples.

An enzyme-linked immunosorbent assay (DIG-ELISA) to detect PCR products of the *rfb*S gene from serogroup D Salmonellae offers a fast and accurate way of detecting this infectious agent in stool specimens (Luk et al., 1997). A polymerase chain reaction test based on the amplification of a 389 bp *inv*A gene fragment to detect *S. typhi* in blood specimens was developed by Cocolin et al. (1998).

A number of variations of the basic PCR technique has led to various strategies for the epidemiological typing of microorganisms. These strategies can be divided into a few major groups :

(a) Randomly amplified polymorphic DNA fingerprinting

(b) PCR-ribotyping

(c) Rep-PCR

(d) Analysis of RFLPs generated by PCR with specific primers (PCR-RFLP)

(e) Amplified fragment length polymorphism (AFLP)

(f) PCR-single strand conformation polymorphism (PCR-SSCP).

1.3.2.4.1 Random amplified polymorphic DNA

The PCR technique has recently been used to detect genomic polymorphism at the strain level by using a single arbitrary primer in a reaction referred to as arbitrarily primed-PCR (AP-PCR) (Williams et al., 1990) or random amplified polymorphic DNA-PCR (RAPD-PCR) (Welsh and McLelland, 1990). This technique requires no previous genetic knowledge about the target organism and is based on the use of simple arbitrary primers combined with 2 cycles of PCR at low stringency and multiple cycles at high stringency to generate a discrete and reproducible set of amplification products that are characteristic of a particular genome (Towner and Cockayne, 1993).

Hermans et al. (1996) manage to differentiate between MDR *S. typhi* strains (plasmid encoded) and sensitive strains from Dhaka by RAPD analysis. RAPD was not able to distinguish outbreaks of *S. typhi* in Zürich (Gruner et al., 1997).

Sixty three *S. typhi* strains isolated from Taiwan and various other regions divided into 21 RAPD types by RAPD fingerprinting. The interesting point of this study was the differentiation of a vaccine *S. typhi* Ty 21a strain from the rest of the strains (Shangkuan and Lin, 1998).

RAPD in conjunction with PFGE and ribotyping showed the spread of a closely related strain of *S. typhimurium* in children and nurses in a hospital (Mhand et al., 1999). Analysis of other *Salmonella* serotypes by RAPD has been carried out as well, with *Salmonella dublin* (Kerouanton et al., 1996) and *S. enteritidis* (Landeras and Mendoza, 1998).

1.3.2.4.2 PCR- ribotyping

PCR ribotyping involves the use of specific primers to amplify particular regions of the genome where rRNA are localized. The polymorphism in rRNA genes or intergenic spacer regions can be detected by PCR banding patterns that are produced by the PCR amplification of these RNA regions which are reproducible and stable in a given isolate and thus can be used for typing purposes (Towner and Cockayne, 1993).

Kostman et al. (1992) showed that PCR mediated amplification of the intergenic region of isolates of *Pseudomonas cepaciae* yielded a variable number of amplified fragments and further demonstrated that epidemiologically unrelated and related strains could be differentiated based in the patterns of fragments generated by PCR- ribotyping. Since then PCR- ribotyping has been used in the epidemiological studies of various microorganisms such as *Listeria, Staphylococcus Salmonella* (Jensen et al., 1993), *Burkholderia cepacia* (Dansen et al., 1999) and *Staphylococcus* species (Maes et al., 1997).

PCR- ribotyping has been successfully used by Nastasi and Mammina (1995) to characterize sporadic and outbreak *S. typhimurium* isolates. In a study done by Lagatolla et al. (1996) *Salmonella* isolates belonging to different serotypes were correctly differentiated by PCR-ribotyping. In contrast,Kerouanton et al. (1996) and Landeras et al. (1998) found this technique to be of limited value while studying *S. dublin* and *S. enteritidis* strains respectively.

1.3.2.4.3 rep-PCR

rep-PCR involves the use of oligonucleotide primers based on families of short, extragenic repetitive sequences which are represented by the 38bp repetitive extragenic palindromic (REP) element and the 126bp enterobacterial repetitive intergenic consensus

(ERIC) sequence (Woods et al., 1992). These repetitive DNA sequences appear to be conserved among many members of the family Enterobacteriaceae as well as other bacterial species (Hulton et al., 1991).

DNA band patterns which are of different sizes represent polymorphism in the distances between repetitive sequence elements. DNA fingerprints that are generated allows distinctions to be made between species and strains (Versalovic et al., 1991). By using PCR analysis, Versalovic et al. (1991) used consensus primers such as ERIC sequences to amplify DNA sequences located between successive repetitive elements in gram negative bacteria and suggested the potential of this method for subtyping gram-negative enteric bacteria.

rep-PCR also appears to be a useful technique for the molecular fingerprinting of *Legionella* species as it was able to discriminate between strains within a serogroup of *L. pneumophila* and between different *Legionella* species (Georghiou et al., 1994).

ERIC-PCR showed limited discrimination of 70 *Salmonella* strains previously characterized by ribotyping (Millemann et al., 1996). Since then progress has been made in this typing method. Burr et al. 1998 managed to obtain different ERIC-PCR profiles from the 89 *Salmonella* isolates from 22 serotypes studied. The drawback was that the isolates could not be separated into different serotypes.

rep-PCR fingerprinting offered an attractive choice as a primary method for the discrimination of strains within *Salmonella* serotype Saintpaul. Epidemic strains were adequately discriminated from cases apparently not related to the epidemic (Beyer et al., 1998)

1.3.2.4.4 PCR-RFLP

PCR-RFLP refers to the restriction endonuclease digestion of PCR amplified DNA to produce RFLPs for analysis. The target sequence which is amplified usually falls within the range of 1-2 kb and the resulting product cut with a restriction endonuclease. Resolution of the fragments by agarose or polyacrylamide gel electrophoresis permits comparisons of different isolates (Kerr, 1994). Kilger and Grimont (1993) used this method for differentiation between non-motile variants and flagellar type *Salmonella* serovars. This was achieved by cleaving the amplified phase 1 flagellin *fli*C gene of major samples by using 2 endonucleases, *Taq*I and *Sac*I.

Restriction fragment length polymorphism of the *flj*B gene (phase 2 flagellar gene) coupled with IS200 analysis showed that *S. panama*, *S. kapemba*, *S. goettingen*, *S. zaiman* and *S. mendoza* could have been possible ancestors of serogroup D1 serovar *Salmonella* 9,12: 1,V (Burnens et al., 1996). Phase-1 and phase-2 flagellin gene RFLP allowed identification of all diphasic *Salmonella* serovars studied although the diversity uncovered did not precisely match that evidenced by flagellar agglutination (Dauga et al., 1998).

PCR-RFLP has been useful in the epidemiological studies of *Chlamydia* spp (Herrman et al., 1996) and *Veittonella* spp (Sato et al. 1997). This technique has also showed importance in taxonomical studies as seen with the genus *Acinetobacter* (Vaneechoutte et al., 1995) and *Listeria monocytogenes* (Wiedman et al., 1997).

1.3.2.4.5 Amplified fragment length polymorphism (AFLP)

This method was originally developed for plant genetic studies, but it is now considered an universal method for fingerprinting DNA from any source including human, animal, plant and microorganismal DNAs (Zabeau and Vos, 1993; Vos et al., 1995).

The AFLP technique is based on the selective PCR amplification of restriction fragments from total digested genomic DNA. It basically involves 3 steps :

(i) restriction endonuclease digestion of genomic DNA with 2 enzymes. (an average cutting frequency enzyme like *Eco*RI, *Pst*I, *Hind*III or *Apa*I and the second enzyme with a higher cutting frequency like *Mse*I or *Taq*I). Restriction is followed by ligation to specifically designed double stranded oligonucleotide adaptors.

(ii) The restricted-ligated DNA is then subjected to 2 PCR amplification cycles under stringent conditions. The first amplification is achieved by using the adaptor and restriction site sequence on target sites for primer annealing. The second amplification (selective amplification) is achieved by adaptor- specific primers that have at their 3' end an extension of 1 to 3 nucleotides running into the unknown chromosomal restriction fragment. The number of fragments amplified can vary depending on the length of nucleotide extension.

(iii) The amplified fragments are electrophoresed either on a polyacrylamide gel or an agarose gel. Detection of fragments are done by radioactively or fluorescently labeling the PCR primer which spans the average frequence restriction site (e.g. *Eco*RI site). The complex banding patterns can be analyzed by computer software for standardization. (Vos et al., 1995; Janssen et al., 1996 ; Savelkoul et al., 1999).

AFLP has been used in the taxonomical studies of *Aeromonas* (Huys et al., 1996) and *Acinetobacter*, where Koeleman et al. (1998) confirmed the grouping of several clinical isolates of *A. baumanni* within the *A. baumanni* species. This technique has also been successfully used in epidemiological studies of *Staphylococcus epidermidis* where relatedness of strains isolated from a sequential blood culture obtained from a single patient

was confirmed. These strains were distinguishable from epidemiologically unrelated strains (Sloos et al., 1998).

In another study, AFLP led to the identification of a cluster of French *Chlamydia psittaci* strains and permitted differentiation among strains in relation to host origin and clinical syndrome (Boumedine and Rodolakis, 1998). The AFLP profile derived from *Hind*III fragments differentiated strains of *H. pylori* from unrelated individuals and confirmed the common origin of the strain in some family members (Gibson et al.,1998).

Recently AFLP analysis of *Salmonella dublin* strains with known PFGE types identified unique AFLP patterns suggesting that both these techniques have about the same discriminating typing power for Salmonellae. However, limited discrimination was seen between *S. enteritidis* strains (Duim et al., 1997). Aarts et al. (1998) analyzed *Salmonella* strains from 62 different serotypes by AFLP and showed that each serotype and in some cases even strains had their own specific AFLP patterns.

1.3.2.4.6 PCR-SSCP

PCR-SSCP analysis is one of the simplest and perhaps one of the most sensitive methods for detection of mutations based on PCR technology. It involves PCR amplifications of a target sequence in genomic DNA and subsequent denaturation of the amplified product to two single stranded DNA (ssDNA). The ssDNA is subjected to non-denaturing polyacrylamide gel electrophoresis. Under non-denaturing conditions the ssDNA has a secondary structure that is determined by the nucleotide sequence.

The sensitivity of PCR-SSCP (the ability to detect mutations) depend on how mutations affect the secondary structure of the ssDNA and, in turn, how the electrophoretic mobility of the ssDNA depends on the secondary structure. Band shifts on the gel indicates

the presence of a mutation which can be detected non-radioactively (by silver staining the gel) or radioactively (by using radioactively labeled PCR primers) (Hayashi, 1992).

This technique was first designed to detect mutations in oncogenes and allelic variants in the human genome (Orita et al, 1989). Since then this technique has been widely used for the detection of mutations in genes responsible for various hereditary diseases (Ainsworth et al., 1991) and somatic mutations of oncogenes (Murakami et al., 1991).

One of the first studies to locate mutated sequences in bacterial genes was carried out by Widjojoatmodjo et al. (1994) where PCR-SSCP (based on the 16s rRNA gene) was used for the detection and identification of a wide spectrum of bacteria. Species specific SSCP profiles were observed for *Clostridium* spp, *Listeria* spp, *Pseudomonas* spp and *Enterobacter* spp.

Variable regions in the *tox* and *dtxr* genes from *Corynebacterium diphtheriae* outbreak strains isolated in the Russian Federation was detected by PCR-SSCP and showed the epidemiological usefulness of this technique (Nakao et al., 1997). A further PCR-SSCP study based on the 16S-23S intergenic spacer region by Nakao and Popovic (1998) provided better differentiation of *Corynebacterium diphtheriae* strains as compared to the traditional ribotyping.

Wain et al. (1998) showed that the *gyr*A gene (involved in quilone resistance) of *S. typhi* multidrug resistant strains from Vietnam had 2 novel point mutations detected by PCR-SSCP. These mutations were not seen in the sensitive strains. Genetic variability was detected in the *rpo*S gene (encodes a sigma factor to overcome exposure to stress) of *Salmonella enterica* enviromental isolates by PCR-SSCP (Jordan et al., 1999).

1.3.2.5 DNA sequencing

Direct comparison of genomic DNA sequences of selected bacterial genes is currently the best means of comparative quantitative analysis of strains (Swaminathan and Mater, 1993). However, DNA sequencing is still too expensive and too complex to be applied in practical situations to type a large number of bacterial isolates (Olive and Bean, 1999). DNA sequencing may be better suited for taxonomic and phylogenetic studies rather than epidemiological investigations. The 16s rRNA genes and 16S-23S intergenic spacer regions of *Salmonella* serovars have been sequenced recently for taxonomical and phylogenetical purposes (Christensen et al., 1998; Luz et al., 1998).

The advent of rapid, automated DNA sequencing techniques (Christensen et al., 1998) may see the more widespread applications of this approach in the future.

1.4 Genes used for molecular analysis of *S. typhi*

There are approximately 3000 genes in the genome of *Salmonella* (Selander et al., 1991) and in the present study a few of the genes associated with virulence have been used in the molecular analysis (PCR-RFLP, PCR-SSCP and hybridization) of *S. typhi*.

1.4.1 *fli*C gene

The phase-1 and phase-2 flagellar proteins are encoded by flagellin genes *fli*C and *fli*B which are found in different locations on the chromosome of *Salmonella* strains. *S. typhi* is monophasic in which it has only the *fli*C gene that expresses the phase 1 flagellar antigen. Most clinically and epidemiologically important Salmonellae are monophasic. The biphasic *Salmonella* isolates produce both types of flagellar proteins (Dauga et al., 1998).

1.4.2 *inv*A gene

Salmonellae causes disease by invading the intestinal epithelium after ingestion. Galan and Curtis (1989) first characterized the *Salmonella* invasin gene *inv*A, the first gene in an operon which is thought to trigger the internalization of salmonellae in cultured epithelial cells.

1.4.3 *via*B gene cluster

The Vi polysaccharide which is an antigen associated with virulence, occurs in only a few organisms such as *S. typhi*, *S. paratyphi* C, a few strains of *Citrobacter freundii* and *Salmonella dublin*. The DNA sequence encoding the Vi antigen is called the *via*B region which contains a set of genes (*vip*A, *vip*B, *vip*C and *vip*R) (Hashimoto et al., 1993).

1.4.4 *rfb*S gene

In *S. enterica*, the enzymes for O antigen biosynthesis are encoded in the *rfb* gene cluster. The O unit is synthesized from nucleotide sugar which are encoded by the following genes : *rfb*B, *rfb*C, *rfb*A and *rfb*D encode the dTDP-rhamnose pathway, *rfb*F, *rfb*G, *rfb*H, *rfb*I and *rfb*J encode the dideoxyhexose pathway, while *rfb*M and *rfb*K encode the GDP-mannose pathway. In *Salmonella* group A and D (which includes *S. typhi*) *rfb*J has been replaced by *rfb*E and *rfb*S which encodes the GDP-tyrelose and GDP-paratose pathways respectively (Xiang et al., 1993).

1.4.5 *omp*C, S1 and S2 genes

Porins are the pore-forming proteins that reside in the outer membrane of gram negative bacteria (e.g. *S. typhi*). They are believed to be involved in the non-selective

passive diffusion of small solutes and even plays a role in the permeation of lactams. The Omp proteins also seems to participate in the invasion of host cells by *Yersinia*, *Shigella* and *S. typhimurium*. *omp* C and *omp*F encodes the synthesis of 2 major porins that are OmpC and OmpF, with OmpA and PhoE also playing an important role. The *omp*S1 and the *omp*S2 gene codes for OMPs that have not been well characterized yet (Martinez-Florez et al., 1998).

1.4.6 *gro*EL gene

When a microbial pathogen (e.g. *S. typhi*) enters a human host, it is confronted by several changes which are highly stressful. These pathogens produce heat shock proteins (HSP) which helps them to survive under such conditions. The *gro*EL gene of *S. typhi* encodes for a GroEL stress protein which is a member of the HSP 60 family of stress protein that serves such purposes of survival (Lindler and Hayes, 1994).

1.4.7 *Salmonella* Pathogenicity Islands (SPI)

Pathogenic islands are chromosomal clusters of virulence genes present in pathogenic organisms but absent from related, non-pathogenic bacteria. Five large pathogenicity islands have been identified in *Salmonella* termed *Salmonella* pathogenicity islands (SPI 1to 5) (Blanc-Potard et al., 1999).

SPI-1 and SPI-2 will only be mentioned as they are relevant in this present study. SPI-1 at 63 minutes on the *S. typhimurium* chromosome, is a 40kb island incorporating more than 28 genes which are involved in the production of type III secretion proteins (*spa, inv, prg*, and *org* genes, effector proteins (*sip, ssp, spt*, genes) and regulatory proteins (*invF*

and *hilA*). These proteins govern the ability of *Salmonella* to invade host and epithelial cells (Groisman and Ochman, 1996; Deiwick et al., 1998; Blanc-Potard et al., 1999).

SPI-2 at 31 minutes mediates survival within macrophages and systemic infections. It is also 40kb and has genes (*ssa, ssr, sse* and *ssc*) that encode for type III secretion proteins, effectors and regulatory proteins and chaperones (Groisman and Ochman, 1996; Deiwick et al., 1998).

1.5 Molecular epidemiology and the purpose of bacterial typing

Molecular epidemiology is the investigation of the epidemiology and pathogenesis of infectious agents by using a wide variety of techniques derived from immunology, biochemistry, molecular biology and genetics (Tackett, 1989; Maslow et al., 1993). It is an independent approach for tracking the transmission of infectious agents by analysing the genetic diversity of field isolates. This approach aims to assess the extend of genetic variability and diversity among population of pathogens in order to implement control measures.

The continuing presence of diseases (such as typhoid fever) in endemic areas and its potential for spread means that improved epidemiological surveillance is essential in epidemiologic investigations. Bacterial typing (subtyping of microorganism) is of great value in the study of epidemiology of infectious disease. Various typing methods have been used to examine sets of bacterial isolates for characteristics that would allow discrimination below species level. Drawbacks of phenotypic typing methods such as the ability of bacteria to alter unpredictably the expression of phenotypic characteristic being assessed has led to the development of genotypic typing. These genotyping methods forms the basis of molecular epidemiology (Swaminathan and Matar, 1993; Olive and Bean, 1999).

Subtyping helps to determine whether previously recognized virulent clones are present in a set of strains, for recognizing outbreaks, and to identify the source of infection in outbreaks and sporadic disease settings (Eisenstein, 1990). These typing techniques can produce data that result in the detection of a cluster of isolates with distinctive features, increased prevalence of particular species, appearance of new species and even the emergence of new multidrug resistant or virulent strains. Information pertaining to these typing data will/can help the epidemiologist to contain or stop a current outbreak by contributing to the knowledge of the patterns of spread of infection. It may influence the design of preventive programmes, especially in relation to the presence of virulent and multidrug resistant strains.

1.6 Variations in the bacterial genome

Genetic variation is a pre-requisite for biological evolution. The genetic diversity of prokaryotes (microorganisms) is advantageous for adaptation in unstable and changing environments and ensures against extinction.

Recent molecular techniques such as pulsed field gel electrophoresis, hybridization and sequencing have been used to assess the variations in bacterial genomes. These variations can be mediated by a number of mechanisms but they generally fall into 3 classes (Ochman and Bergthorsson, 1995; Sanderson and Liu, 1998; Arber, 2000):

(a) point mutations (base pair changes) which include nucleotide substitutions and frameshift mutations.

(b) chromosomal rearrangement involving deletions, inversions, transpositions and duplications.

Subtyping helps to determine whether previously recognized virulent clones are present in a set of strains, for recognizing outbreaks, and to identify the source of infection in outbreaks and sporadic disease settings (Eisenstein, 1990). These typing techniques can produce data that result in the detection of a cluster of isolates with distinctive features, increased prevalence of particular species, appearance of new species and even the emergence of new multidrug resistant or virulent strains. Information pertaining to these typing data will/can help the epidemiologist to contain or stop a current outbreak by contributing to the knowledge of the patterns of spread of infection. It may influence the design of preventive programmes, especially in relation to the presence of virulent and multidrug resistant strains.

1.6 Variations in the bacterial genome

Genetic variation is a pre-requisite for biological evolution. The genetic diversity of prokaryotes (microorganisms) is advantageous for adaptation in unstable and changing environments and ensures against extinction.

Recent molecular techniques such as pulsed field gel electrophoresis, hybridization and sequencing have been used to assess the variations in bacterial genomes. These variations can be mediated by a number of mechanisms but they generally fall into 3 classes (Ochman and Bergthorsson, 1995; Sanderson and Liu, 1998; Arber, 2000):

(a) point mutations (base pair changes) which include nucleotide substitutions and frameshift mutations.

(b) chromosomal rearrangement involving deletions, inversions, transpositions and duplications.

(c) Horizontal gene transfer due to homologous recombination which results in the acquisition of foreign DNA sequences.

1.6.1 Point Mutation

A mutation is a change in the nucleotide sequence of a short region of a genome. Many mutations are point mutations which results in nucleotide substitution, small deletion/insertion of one or a few nucleotides (Brown, 1999).

A number of studies have demonstrated that genomic DNA restriction fragment patterns differ between different strains of a single bacterial species. Such restriction fragment length polymorphism (RFLP) analysis (e.g. PFGE, ribotyping, AFLP) has been used as a typing method to trace and differentiate strains of bacteria from different sources (Struelens et al., 1992; Pang et al., 1995; Aarts et al., 1998).

Hall (1994) carried out DNA fingerprinting studies on *Enterococcus faecalis* by digesting total DNA with restriction enzyme *Sst*I followed by separation on a conventional agarose gel. It was shown that 25% of the RFLPs were generated by point mutations. The restriction cut sites of *Xba*I and *Bln*I is very different in the genes of *S. paratyphi* B and *S. typhimurium* LT2, most likely due to base pair changes (Liu et al., 1994).

LeClerc et al. (1996) showed that the rate of genetic variation increased in certain mutator phenotypes of *Escherichia coli* and *Salmonella* pathogens which were deficient in methyl-directed mismatch repair (MMR). MMR is a post replicative repair system that corrects mistakes on newly synthesized DNA strands. This is to ensure the precision of chromosome replication. Bacteria defective in MMR show both an increased rate of mutation (a hypermutable phenotype) and increased genetic exchange between diverged DNA sequences (horizontal gene transfer).

I- *CeuI* digestion followed by PFGE was used by Nandi et al. (1997) to study the genomic rearrangements of *Vibrio cholerae* strains belonging to different serovars and biovars. They concluded that major genomic rearrangements in the classical and El Tor biovars were due to homologous recombination between closely linked rRNA genes which resulted in deletion of one of the *rrn* loci. The chromosomes of other genera such as *Streptomyces* are relatively unstable. Nearly 30% of the chromosomal DNA of many *Streptomyces* are due to duplication (amplification) of chromosomal sequences (Birch et al., 1990).

1.6.3 Horizontal gene transfer

The addition of new genes (segments of DNA) through horizontal transfer is now recognised as a major factor driving evolution of the bacterial chromosome, indicating that bacteria can acquire foreign DNA (nucleotide composition that is significantly deviated from that of the host chromosome) despite multiple barriers to chromosomal gene transfer between species (LeClerc and Cebula, 1997). It is estimated that 6- 16% of *E. coli* genes and 20% of *Salmonella* chromosomal DNA sequences were acquired by horizontal transfer (Medigue et al., 1991 ; Lan and Reeves, 1996). Horizontal transfer involves entry of non-homologous genes, most likely by classical transfer methods such as conjugation and transduction, and integration of these genes into the host chromosome (Sanderson and Liu, 1998). The vectors that transport these segments of DNA from donor to recepient organisms involves mobile genetic elements such as phages, plasmids, insertion elements and transposons (Mecsas and Strauss, 1996). These segments of DNA, known as pathogenicity islands intergrate themselves into the host chromosome (forming DNA loops) by an unknown mechanism. However, the existence of insertion elements associated with pathogenicity islands found in *Salmonella*, *Helicobacter* and *Vibrio* ; repeated DNA

motifs at the boundaries of pathogenicity islands in *E. coli*, *Salmonella* and *Yersinia* ; and tRNA genes used as acceptor sites for pathogenicity islands in enteropathogenic and uropathogenic *E. coli* and in *Salmonella*, suggest that homologous recombination events are involved (Mecsas and Strauss, 1996 ; Groisman and Ochman, 1996 ; Hensel et al., 1997). Pathogenicity islands contain discrete segments of DNA (genes) that encode virulence traits. *Salmonella* contains at least 5 pathogenicity islands (Mecsas and Strauss, 1996). The phoN gene (for non- specific acid phosphatase) which has a G+C content of 43% compared to the *Salmonella* average of about 51% is incorporated into the *Salmonella* genome via plasmid conjugation (Groisman et al., 1992).

The EHEC strain *E. coli* 0157 : H7 produces cytotoxins similar to those from *Shigella dysenteriae* 1. *E. coli* 0157 : H7 strains harbour lambdoid prophages that contains the *sit* genes needed to encode the Shiga- like toxin (O'Brien and Holmes, 1987).

Vibrio cholerae 0139 strains may have emerged after acquisition of pathogenicity islands from serotype 01 El Tor. The pathogenicity islands in *V. cholerae* 0139 is flanked by 2 different insertion elements.

1.7 General objectives of the study

The quality of genomic DNA is important for reproducibility of DNA fingerprinting. Hence, the first part of the study was to optimise the conditions for DNA preparation applicable to typing methods such as pulsed-field gel electrophoresis (PFGE), ribotyping, IS200 typing and gene hybridization studies using endonuclease- restricted DNA.

One of the major objectives of this project was to assess the usefulness of PFGE, ribotyping, IS200 typing, amplified fragment length polymorphism (AFLP) analysis, PCR- ribotyping and PCR- restriction fragment length polymorphism (RFLP) as molecular tools for the purpose of delineating epidemiological and evolutionary relationships of sporadic *Salmonella typhi* isolated from Papua New Guinea (PNG) over a period of 3 years, as well as *S. typhi* isolated from various geographic locales. In this study, we used AFLP, which is a recently developed PCR- mediated typing method, PFGE and ribotyping to compare the molecular characteristics of *S. typhi* isolated from PNG patients with fatal and non- fatal typhoid fever and from different body sites in the same patient.

Another important objective of this project was to assess the genetic diversity and genomic relationships of the PNG strains as well as *S. typhi* strains from different regions of the world by Southern blot hybridization of restricted genomic DNA with various gene probes related to virulence. We also tried to assess as well whether IS200 elements played a role in the transposition of genes in the genome of *S. typhi*. Another aspect that was looked into, based on hybridization work was to detect the existence of *Salmonella* pathogenicity islands (SPI-1 and SPI-2) in *S. typhi* environmental and clinical isolates of *S. typhi* from Chile.

The potential utility of a PCR- based method (PCR- SSCP) for detecting genetic diversity in the PNG and geographic *S. typhi* isolates based on base pair mutations within specific genes was assessed. The possiblility of using PCR- SSCP as a molecular tool in the epidemiological studies of *Salmonella* serovars by detecting mutations within the *gro*EL gene (heat shock protein gene) was similarly evaluated.