NON-PROTEIN CODING RNA GENES AS THE POTENTIAL MARKERS FOR THE DETECTION And DIFFERENTIATION OF *Salmonella enterica* serovar Typhi and *Salmonella enterica* serovar Paratyphi A USING PCR

NITHYA RAVICHANTAR

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

September 2014

I dedicate this thesis to my family for all their sacrifices,

unconditional love and support

ACKNOWLEDGEMENTS

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Assoc. Prof Dr. Tang Thean Hock, for providing me the opportunity to pursue MSc. degree under his supervision. His guidance, patience and constructive criticism have always been supportive at all times during my work. I am grateful for the chance given to work with his enthusiastic group. I would like to acknowledge my Co-supervisor, Dr. Doblin Sandai for his advice and support. I owe my deepest appreciation to my field-supervisor Dr. Timofey Rozhdestvensky, Institute of Experimental Pathology, University of Muenster, Germany, for his guidance and valuable advices throughout this work. I am thankful to Dr. Venkata Suresh Chinni, Aimst University, for useful discussions and suggestions.

I thank the Department of Microbiology & Parasitology, USM (Kelantan), Advanced Diagnostic Laboratory (AMDI) and Veterinarian Research institute, Ipoh for providing the bacterial cultures, which has contributed significantly to this work. I am indebted to AMDI, USM for the sanction of AMDI Student Research Grant and personal financial assistance via Graduate Assistance Scheme. I am thankful to the Ministry of Higher Education (MOHE) for MyBrain 15 scholarship. I am grateful to the management staffs of AMDI, USM who rendered me with valuable experiences outside of the laboratory. I would also like to thank the family of Infectious Disease Cluster (now Infectomic Cluster), AMDI, in particular Mdm Siti Aminah, Dr. Hoe Chee Hock,

Dr. Citartan Marimuthu, Ms. Lee Li Pin, Ms. Thiviyaa Othaya Kumar, Mrs. Priya Kaniappan and Ms. Toh Saw Yee for their assistance and support.

Last but not least, I am grateful to the Almighty God for all His guidance, blessings, lessons and 'accompanionship' when I was away from my dearest family to pursue my MSc. Words alone can't explain the degree of gratitude towards my family for showering me with love, care and motivations for me to strive and complete this work successfully. This thesis would not have been possible without their supports. Lastly, I offer my sincerest thanks to my friends and to those who supported me in any respect during the completion of the project.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND SYMBOLS

A	Adenine
Bp	Base pair(s)
BLAST	Basic Local Alignment Search Tool.
Megablast	Mega Basic Local Alignment Search Tool
C	Cytosine
°C	Degrees Celsius
CFU	Colony Forming Unit
dATP	Deoxyadenosine triphosphate
dCTP	Deoxycytidine triphosphate
ddH_2O	Double-distilled water
dGTP	Deoxyguanosine triphosphate
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
dsDNA	Double-stranded DNA
dTTP	Deoxythymidine triphosphate
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic Acid
ELISA	Enzyme-Linked Immunosorbent Assay
et al.	and others
EtBr	3, 8-diamino-5-Ethyl-6-phenyl phenanthridinium Bromide

G	Gram
G	Guanine
GAPDH	Glyceraldehyde-3-phospate-dehydrogenate
HCl	Hydrochloric acid
H	Hour(s)
IAC	Internal Amplification Control
KH_2PO_4	Potassium phosphate
KCl	Potassium chloride
kDa	Kilodalton
L	Litre
LB	Luria Bertani
M	Mol/Liter, molar
Mg	Milligram
$MgCl_2$	Magnesuim Chloride
Min	Minute(s)
mL	Milliliter
mM	Millimolar
mRNA	Messenger RNA
mPCR	Multiplex Polymerase Chain Reaction
NaCl	Sodium chloride
NaOAc. $3H_2O$	Sodium acetate trihydrate
NaOH	Sodium hydroxide

Ng	Nanogram
NH_2HPO_4	Sodium hydrogen phosphate
Nt	Nucleotide(s)
OD	Optical Density
npcRNA	Non-protein-coding gene
PBS	Phosphate buffered saline
PCR	Polymerase Chain Reaction
Pg	Piccogram
RBS	Ribosomal Binding Site
RNA	Ribonucleic acid
RNase	Ribonuclease
Rpm	Rotations per minute
rRNA	Ribosomal RNA
RT	Room temperature
S	Second(s)
SDS	Sodium dodecyl sulfate
siRNA	Small interfering RNA
sRNA	Small RNA
SPI	Salmonella Patogenecity Island
ssDNA	Single-stranded DNA
T	Thymine
TAE	Tris–Acetic Acid–EDTA

Tris	Tris-(Hydroxymethyl)-Aminomethane
Tris-HCl	Tris-(Hydroxymethyl)-Aminomethane Hydrocloride
tRNA	Transfer RNA
UV	Ultraviolet
V	Volt (s)
v/v	Volume per volume
XLD	Xylose Lysine Deoxycholate
µL	Micro liter
µM	Micro molar
%	Percentage
B	Beta

GEN 'RNA BUKAN PENGKOD PROTIN' SEBAGAI SASARAN POTENSI UNTUK TUJUAN PENGESANAN DAN PEMBEZAAN *Salmonella enterica* serovar Typhi dan *Salmonella enterica* Serovar Paratyphi A DENGAN MENGGUNAKAN PCR

ABSTRAK

RNA-bukan-pengkod-protein (npcRNA) merupakan RNA yang tidak di terjemahkan kepada penghasilan protein. Sebaliknya, ia memainkan peranan dalam mengawal-atur pelbagai process dalam sel. Kebelakangan ini, ekspresi npcRNA tertentu telah dikaitkan dengan pelbagai penyakit manusia dan barah. Salmonellosis merupakan penyakit berjangkit yang disebabkan oleh bakteria, *Salmonella spp* melalui pengambilan makanan dan minuman yang tercemar. Pengesanan dan pembezaan antara Salmonella spp, termasuk *S.* Typhi dan *S.* Paratyphi A, adalah sangat penting untuk mengawal kadar jangkitan dan kematian. Oleh itu, keberkesanan dalam penyasaran gen-gen npcRNA untuk pengesanan dan pembezaan Salmonellosis telah dibincang dalam tesis ini, khususnya ujian multipleks PCR (mPCR) yang mengesan *S.* Typhi dan *S.* Paratyphi A secara spesifik telah dibangunkan. Had pengesanan ujian mPCR tersebut adalah 10pg. Selanjutnya, dalam eksperimen pencemaran *S.* Typhi dan *S.* Paratyphi A dalam najis, had pengesanan yang dicapai adalah 1.5×10^1 dan 1.5×10^0 CFU/mL selepas 4 jam inkubasi dalam media kultur "selenite cystine". Secara lansung, ujian mPCR yang berasaskan npcRNA yang di bangunkan adalah sensitif dan spesifik untuk digunakan

dalam diagnostik klinikal dan pengawasan penyakit tersebut. Oleh itu, npcRNA mewakili sasaran molekul yang sangat baik untuk pengesanan penyakit salmonellosis.

NON-PROTEIN CODING RNA GENES AS THE POTENTIAL MARKERS FOR THE DETECTION And DIFFERENTIATION OF *Salmonella enterica* serovar Typhi and *Salmonella enterica* serovar Paratyphi A USING PCR

ABSTRACT

Non-protein-coding RNAs (npcRNAs) are not translated into protein but are key regulatory molecules in the cell. Recently, the expression of specific npcRNAs has been linked to various human diseases and cancer. Salmonellosis is a communicable disease caused by members of the *Salmonella* species, which infect humans through contaminated food or water. It is of paramount importance that accurate detection methods are available for differentiating the *Salmonella* serovars that produce severe infection in humans, including *S.* Typhi and *S.* Paratyphi A. The efficacy of using npcRNA genes from *S.* Typhi as molecular targets in detection and differentiation of S. Typhi and S.Paratyphi A using a a multiplex polymerase chain reaction (mPCR) assay was reported in this study. The detection limit for this mPCR assay was 10pg of gDNA. Moreover, in a stool-seeding experiment with *S.* Typhi and *S.* Paratyphi A, a respective detection limit of 1.5×10^1 and 1.5×10^0 CFU/mL following 4h of enrichment in selenite cystine broth offers an early clinical diagnostics. Collectively, this study has successfully developed npcRNA genes based mPCR that is able to detect and differentiate *S.* Typhi, *S.* Paratyphi A in a typhoid infection besides discriminating them from a general salmonellosis. With its comparable specificity and sensitivity level to other studies/ products, this mPCR could prove to be an excellent alternative for clinical diagnosis and disease monitoring of salmonellosis.

CHAPTER 1

INTRODUCTION

1.1 Non-protein coding RNA

Cells consist of two types of RNAs: 1) RNAs that are translated into proteins which are the messenger RNA (mRNA) and 2) RNA that are not translated into proteins but involved in many regulatory functions in the cell. They are often referred as small RNAs (sRNAs) as for short bacterial npcRNAs while non-coding RNAs (ncRNAs) have been the most common term for eukaryotic ncRNAs along with other synonyms such as functional RNAs (fRNAs) and to a lesser extent, non-messenger RNA (nmRNAs) (Eddy, 2001, Huttenhofer *et al*., 2002, Huttenhofer *et al*., 2005, Mattick, 2004). To have a common term for all such RNAs, we have opted to apply the term non-protein coding RNA (npcRNA) to all these functional RNAs, irrespective of the realm of life in which they might appear.

Recent advances in molecular biology have revealed vast range of npcRNAs in all three domains of life. For examples, Huttenhofer *et al* 2001 reported the identification of 201 npcRNAs from mouse brain (Huttenhofer *et al*., 2001, Mattick, 2001), Tang et al 2002 identified 86 candidates for small non-messenger RNAs from the archaeon *Archaeoglobus fulgidus* and further reveals a the link between splicing of archaeal introns and rRNA processing in one of the candidate (Tang *et al*., 2002a, Tang *et al*., 2002b). npcRNA was also been reported from bacteria, for example, Vogel et al

disclosed the interaction of 2 npcRNAs in an antisense mechanism in *Escherichia coli* (Vogel *et al*., 2003). npcRNAs are reported to regulate the genes involved in the pathogenecity of the *Vibrio cholerae* and *Mycobacterium tuberculosis* which, contribute to their ability to cause disease and to survive, respectively (Arnvig and Young, 2012, Bardill and Hammer, 2012).

The npcRNA participate in myriad of functions in a genome, *i.e* transcriptional regulation, post-transcriptional gene silencing, protein stability and mRNA stability (Flintoft, 2013, Kornienko *et al*., 2013, Mellor, 2010, Szymanski *et al*., 2003). Most npcRNAs seemed to fine-tune cellular responses during environmental changes by integrating environmental signals into global regulations. They act via mechanisms such as RNA-RNA base pairing, RNA-DNA base pairing, RNA-protein interaction and intrinsic RNA activity (Castello *et al*., 2013, Leontis and Westhof, 2001, Nacher, 2013, Rios, 2006).

1.1.1 npcRNA as Biomarkers in Human Diseases and Cancer

The differential expressions and tissue specificity has prompted npcRNA to be an ideal diagnostic markers in various human diseases and cancers (Lu *et al*., 2005, Ren *et al*., 2013, Rosenfeld *et al*., 2008, Walley, 2013). For instances, the circulating *miR-195* in blood has been found to be significantly elevated in breast cancer patients compare to healthy controls. Differential expression of miRNAs were also reported in patients with other malignancies such as colon, prostate, renal cancer (Heneghan *et al*., 2010a,

Heneghan *et al*., 2010b). Besides, circulating npcRNAs such as *miR-1*, *miR-133*, *miR-208* and *miR-499* were observed to be significantly increased in plasma of acute myocardial infarction patients at early stage of the disease, notably, at 4^{th} hours of symptoms onset, thus they have been proposed to be an excellent markers for diagnostics (Ha, 2011, Kinet *et al*., 2013).

1.1.2 npcRNA in Bacteria

A plethora array of npcRNA candidates has been identified in bacteria by various experimental strategies (Brosius and Tiedge, 2004, Chinni *et al*., 2010, Heidrich *et al*., 2006, Huttenhofer *et al*., 2002, Storz, 2002, Vogel *et al*., 2003, Washietl *et al*., 2005). Discoveries of growing numbers of npcRNAs have changed the previous impression that proteins are the only relevant players in bacterial gene expressions. Bacterial small npcRNAs usually constitute a structurally diverse class of molecules that are typically 50-500 nucleotides in length. Besides lacking of an sizeable open reading frame (ORF) (Altuvia, 2007, Vogel *et al*., 2003) the majority of the known npcRNAs are found to have regulatory roles by base pairing with target mRNAs which theoretically, involve sequence-specific binding to a target RNA and therefore are referred to as antisense RNAs. The antisense RNAs in bacteria are divided into *cis*-encoded antisense RNAs and *trans*-encoded antisense RNAs.

Cis-encoded antisense RNAs are npcRNAs that perfectly complementary to their respective target and transcribed from the same locus, in an opposite orientation. By

contrast, *trans*-encoded npcRNAs are encoded by separate genetic loci. *Trans*-encoded npcRNAs are generally short and would exist in non-contiguous complimentarity to the target mRNA, therefore requires a chaperon protein such as Hfq to facilitate the binding in general (Jousselin *et al*., 2009, Vogel *et al*., 2003).

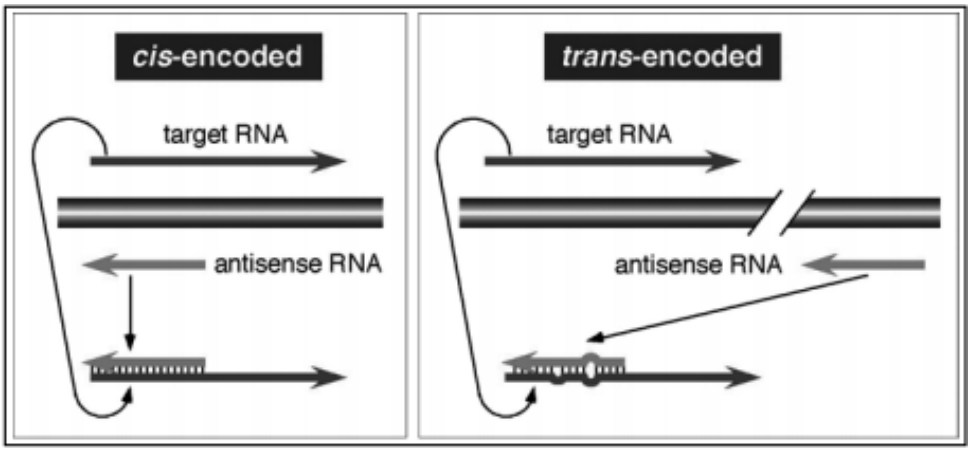

Figure 1.1: The schematic diagram of both *cis*-encoded and *trans*-encoded antisense RNAs. Complexes formed by interaction between antisense and target RNAs are simplified (often, *cis*-encoded antisense/target RNA complexes are not fully paired). Non-contiguous helices are indicated in the right hand panels (adapted from Wagner *et a*l., 2003).

1.1.2.1 Identification of Novel npcRNAs in *S*. Typhi

Generally, the identification and characterization of npcRNA can be achieved via biocomputational or experimental. Notebly, the experimental RNomics has paid greater attention to the identification, classification, functional characterization and gene regulation of npcRNA in bacteria (Bartel, 2004, Huttenhofer *et al*., 2005, Papenfort and Vogel, 2010). By an experimental RNomic approach, Chinni *et al*., 2010 have discovered 97 novel npcRNAs candidates from the human pathogen *Salmonella enterica* serovar Typhi (*S.* Typhi), the causative agent of typhoid. Based on the location in the genome (Salmonella Pathogenicity Island-derived npcRNAs) and the differential expression of these npcRNA, the authors have proposed that some of the detected npcRNA candidate to be involved in regulatory mechanisms of virulence, antibiotic resistance and pathogenic specificity of *S*. Typhi. Interestingly, some candidates are thought to be specific to *S.* Typhi only and/or in *Salmonella* species.

However, further characterization, especially in the diagnostic potential of these npcRNA candidates, was not carried out. Therefore, this work looks into the potential of these npcRNA as molecular marker for the detection and differentiation of salmonellosis infection.

1.2 The Genus *Salmonella* and Its Characteristics

In the year 1900, Lignieres coined the name *Salmonella* after Daniel Elmer Salmon, the bacteriologist who identified *Salmonella Choleraesuis* (*S.* Choleraesuis) in 1885 from pigs (Threlfall and Frost, 1990). The genus *Salmonella* is phylogenetically clustered in the family of *Enterobacteriaceae* (Chart, 2003). The *salmonella* is divided into 2 species; *S. enteric*a and *S. bongori* and has been assigned into more than 2500 different serovars (CDC,2011), which has been classified according to the Kauffman-white scheme based on the antigenic variation at lipopolysaccharide moieties – O antigen (somatic antigen), H antigen (flagellar antigen), and Vi antigen (capsular polysaccharides) (Brenner *et al*., 2000, Popoff *et al*., 2000, Schrader *et al*., 2008).

It is characterized as a gram-negative, ubiquitous, non-encapsulated, straight-rod shaped, facultative, non-spore forming, and generally motile with peritrichous flagella (Fig. 1.2) (McCarter, 2005), with the exception of the poultry-specific serovars, *Salmonella enterica* serovar Gallinarum (*S.* Gallinarum) and *Salmonella enterica* serovar Pullorum (*S.* Pullorum) (Ribeiro et al., 2009). The bacterium has a width of 0.7 to 1.5 μm and a length of 2.0 to 5.0 μm (Holt *et al*., 1994). It can grow within a wide range of temperatures from 8 to 45°C, but the optimal growth temperature is within the range of 35 to 40°C (D'Aoust, 1991). Typically, *Salmonella's* pH growth range lies within 4.5 to 9.0. However, the most favorable pH for growth is above 5.5 (Foster, 1995).

Salmonella are usually aerogenic - producing gas from glucose, and can utilize citrate as their sole carbon source. The failure to ferment lactose and the ability to produce Hydrogen sulphide (H_2S) gas from Sulphur-containing amino acids are features used for the identification of *Salmonella* colonies on primary culture media. Nevertheless, *S. Typhi* is a facultative anaerobe, incapable of producing gas from fermentation of glucose, and does not utilize citrate (WHO 2000). The *Salmonella* bacterium is non-tolerant to oxidase and can convert nitrate to nitrite (Gilberthorpe and Poole, 2008, Vazquez-Torres *et al*., 2000) .

In addition, lysine and ornithine are decarboxylated and the bacterium do not hydrolyse urea nor deaminate tryptophan or phenylalanine (Viala *et al*., 2011,Wray 2000). These unique characteristics provide a form of demarcation when identifying *Salmonella* from other closely related organisms.

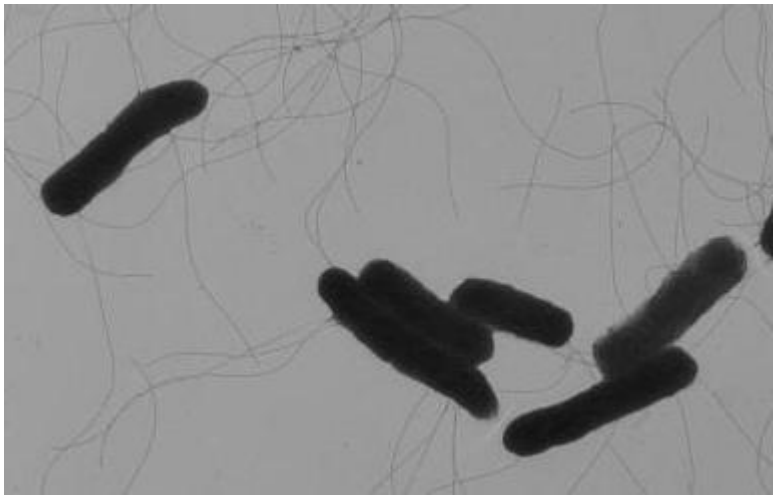

Figure 1.2: The peritrechous flagellated *Salmonella* bacteria under the electron microscope (adapted from Eureka.com 2011).

1.3 *Salmonella enterica*

Salmonellosis is an infection caused by *Salmonella* spp which can be divided into two species, enteric and bongori, *S. enterica* can be subdivided into **I)** *S. enterica* subspecies enterica; **II)** *S. enterica* subspecies salamae; **IIIa**, *S. enterica* subspecies arizonae; **IIIb)** *S. enterica* subspecies diarizonae**; IV)** *S. enterica* subspecies houtenae and **V)** *S. enterica* subspecies indica These organisms are commonly, recovered from nearly all vertebrates, insects and environment (Choo *et al*., 2011, Gerlach and Hensel, 2007). However, the majority (99.5%) of isolated serovars from animal and human belongs to *S. enterica* subspecies I, thus, responsible for human salmonellosis (typhoidal and non-typhoidal disease in human). Currently, there are over 2,300 serovars identified within subspecies with different prevalency and disease that they caused in different host. The clinical severity is largely based on the serovar affecting specific host (De Lappe *et al*., 2009).

The serovars have to overcome both the specific and non-specific host immune system during the course of adaptation. Thus, the pathogenecity of host adapted serovars results from the development of ways for the survival in host which is influenced by the genetic factor and environment in respective host (Uzzau *et al*., 2000). Among the *S. enterica* subspecies I that invades immune system of warm-blooded animals (Porwollik *et al*., 2004), *S.* Typhi is a well known serovar with its ability to modify the physiological function (such as intracellular engulfment, apoptosis, transfers of antigens by M cells and others) of the host for their conveniences (Gordon, 2008, Holt *et al*., 2009, Spano *et al*., 2011).

Although, most of the serovars from subspecies I causes an acute, short duration localized gastroenteritis (or non-typhoidal salmonellosis in humans) rather than a systemic disease, such as *Salmonella enterica* serovar Typhimuruim and *Salmonella enterica* serovar Enteritidis (Asseva *et al*., 2012, Uzzau *et al*., 2000). However the human-restricted pathogen, *S.* Typhi and *Salmonella enterica* serovar Paratyphi A (*S.* Paratyphi A) causes severe systemic disease (typhoidal disease in humans) presented as septicemia, fever, severe diarrhea and could be life-threatening (Gal-Mor *et al*., 2012).

Therefore, *Salmonella enterica* subspecies enterica especially *S.* Typhi and *S.* Paratyphi A, are the most pathogenic subspecies of *Salmonella genus* with highest

medical significance. Therefore, these subspecies deserve for greater clinical attention compare to other subspecies.

1.4 Pathogenesis of *S*. Typhi

S. Typhi (and *S.* Paratyphi A) causes enteric fever, acquired through fecal-oral ingestion leading to systemic infection. However, least observation has been made on the pathogenesis of typhoid as the host-range of *S.* Typhi and *S.* Paratyphi A restricted to human. Experimentation using *S.* Typhimuruim and surrogate host (mice) convey that invasion of *S.* Typhi occurs through M cells on peyer's pathches (Figure 1.3). *S.* Typhi will finally localize to the bone marrow and ended-up at gall bladder where the internal transmission cycle is completed where the organisms are shed in bile, probably in high numbers. *S.* Typhi (but not *S.* Paratyphi A) expresses the Vi capsular polysaccharides, which possesses immuno-modulatory properties that capable to overcome host immune defenses profoundly (Gordon, 2008, Holt *et al*., 2009, Spano *et al*., 2011, Wain *et al*., 1998, Wain *et al*., 2001).

Thus, *S.* Typhi probably appreciates a simple default pathway by limiting the activation of inflammatory response which results in the lacking of gastroenteritis symptoms at early stage of invasion. Moreover, human do not react clinically to the initial invasion step since there is an incubation period before the emergence of the symptoms, which usually occur during the phase of systemic infection. Therefore, an early diagnostics as soon as symptoms emerge could help to control the rate of

morbidity and mortality of the disease. A disease symptom that is indistinguishable from typhoid infection clinically is an additional caveat for an effective and appropriate early diagnostics (Raffatellu *et al*., 2005a, Raffatellu *et al*., 2005b).

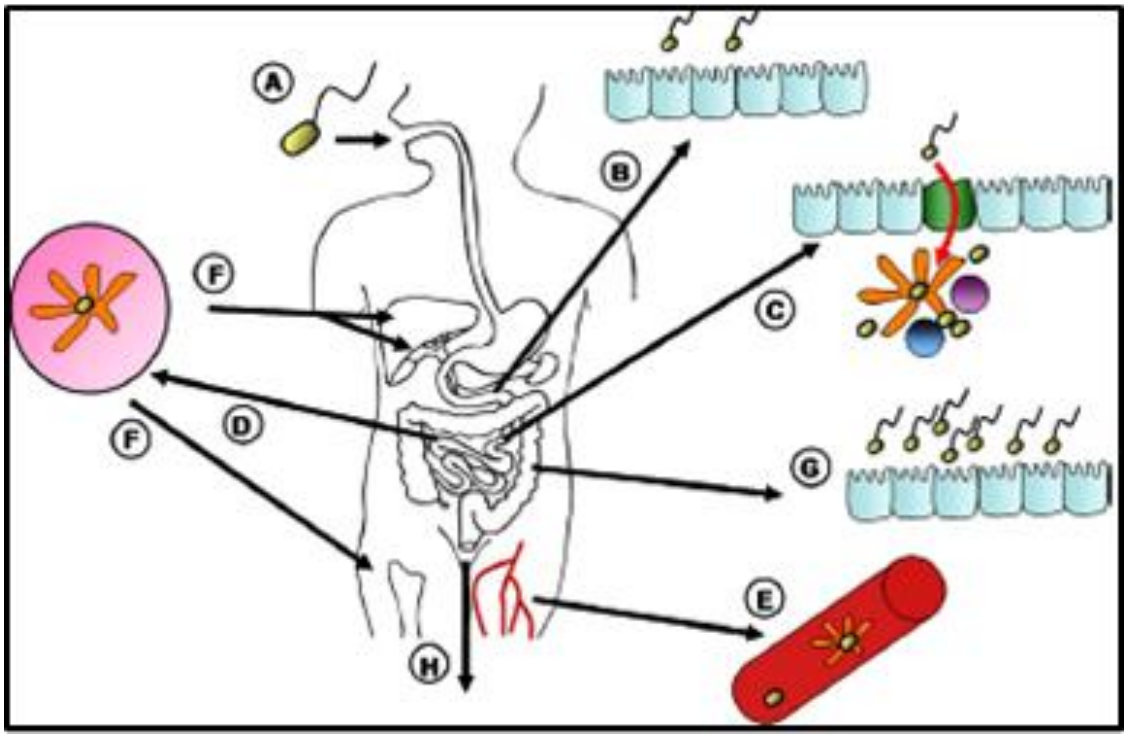

Figure 1.3: **The life-cycle of *S.* Typhi in human host**. **A)** *S.* Typhi usually enters the human through fecal-oral ingestion of an infectious dose. **B)** *S.* Typhi enters large intestine, at this stage host may shed the organism in the stool in sporadic or limited amount as *S.* Typhi do not replicate in large intestines. **C)** Invasion of *S.* Typhi through the terminal ileum probably via the M cells, this perhaps takes place a short time after ingestion. **D)** *S.* Typhi will be trafficked to the reticulo-endothelial system via monocytic cells, potentially in semi-dormant state. **E)** Upon the activation of acquired immune response, *S.* Typhi re-emerges from the reticulo-endothelial system and re-enters the blood stream in low numbers. **F)** *S.* Typhi seeds into liver, the gall bladder, and the bone marrow where it can reside and may be detected for months or years. **G)** *S.* Typhi capable of entering into the bile duct and shedded in high numbers into the environment via intestines. **H)** Excretion of *S.* Typhi in stool (Baker *et al*., 2010).

1.5 Typhoid and Diagnostics

Typhoid fever is the global problem that claims thousands of lives yearly. Infants, childrens, elderly people and immuno-compromised patients are at high risk (WHO, 2005). However, estimation of typhoid is difficult as only severe cases are ever reported. Awareness in the rising incidence of the disease, such as paratyphoid infections is important for the improved control of the disease. Therefore, an accurate diagnostics is not only important for the treatment but also plays an important role in disease vigilance (Teh *et al*., 2008).

1.5.1 Global Burden of Typhoid

Globally *S.* Typhi and *S.* Paratyphi A account for more than 21.7 million and 5.4 million infections per year, respectively (Crump *et al*., 2004), with highest incidence being reported in developing and under-developing countries in South–East Asia, representing 394.2 cases per 100,000 populations (Buckle *et al*., 2012). Typhoid strike Malaysia with sporadic outbreak from time to time with an incidence rate of 0.84 per 100,000 (MOH, 2012). However, this rate is predicted to raise due to the emergence of foreign workers from endemic neighbouring countries such as Indonasia and India (Baker *et al*., 2008, Holt *et al*., 2012).

Epidemiologically, typhoid fever has been superseded by paratyphoid infection (Ochiai *et al*., 2005), which is now responsible for 30–50% of enteric fevers in South–East Asia, particularly in the countries with largest population, China and India (Maskey

et al., 2008). Unfortunately, there is no current paratyphoid vaccine, and administration of *S.* Typhi vaccine provides little to no protection against *S.* Paratyphi A (Pakkanen *et al*., 2014, Pokharel *et al*., 2006, Sahastrabuddhe *et al*., 2013, Tankhiwale *et al.*, 2003). Moreover, paratyphoid has become increasingly prevalent among vaccinated travelers, and these has pose a threat to developed countries as well (Connor and Schwartz, 2005). Therefore, there is a requisite rapid diagnostic test for clinical management, contact tracing, and identification of convalescent/chronic fecal carriers.

1.5.2 Tradisional Diagnostics of Typhoid

Current diagnostic of typhoid infection depend on isolation of organisms from patient, most commonly blood culture and the serological based test (Wain and Hosoglu, 2008). An obvious obstacle in blood culturing during the early stage of infection will be the low number of bacteria in the blood that might lead to false negativity. There are situation where high number of bacteria present in 'unculturable' form, for example, the rapid transfer of bacteria in dormant form from within an intracellular vacuole to laboratory media may kill the fastidious and delicate organisms (Baker *et al*., 2010). Besides that, it takes about 3-5 days to obtain the results.

Widal test is a crude assay that monitors agglutinating antibodies that react with *S.* Typhi. Since *S.* Typhi is a member of the *Enterobactericae*, where many of the surface antigens demonstrates significant conservation and induce antibodies that are cross-reactive. Besides, individuals in endemic areas have cross-reactive antibodies even

though they have no clinical record of typhoid, thus any such serological based test would likely to yield a significant false-positive result (Baker *et al*., 2010, House *et al*., 2005, House *et al*., 2001).

1.5.3 PCR-based Diagnostics of Typhoid

The advancement in molecular techniques adopted Polymerase Chain Reaction (PCR) during the urgent raise for the rapid, sensitive and specific diagnostic of typhoid. PCR is capable of amplifying small fragment of DNA into thousands of copies, ensuring the rapid detection of the organism at early stage of an infection without the expense of sensitivity and specificity. Food industry has used PCR for several decades and published a standardized guidelines for the detection of *Salmonella* in food samples (Malorny *et al*., 2008), while there are several studies that are looking into specific detection of *S.* Typhi from clinical sample .

Current PCR-based assays for diagnosing typhoid or salmonellosis have been largely based on the detection of protein-coding genes and 16S rRNA. For example, multiplex PCR was designed based on O, H and Vi antigen genes for the rapid detection of *S.* Typhi and *S.* Paratyphi A (Hirose *et al*., 2002). Elsewhere, Aziah et al have reported the usage of ~50kDa channel protein gene as the target for the specific detection of *S.* Typhi (Aziah *et al*., 2007). Additionally, Ngan and coworkers have developed a multiplex PCR based assay for the precise detection of Salmonella spp, *S.* Typhi and *S.* Paratyphi A by targeting outer membrane protein gene, putative protein

gene and fimbrial protein gene (Ngan *et al.*, 2010). Recently, Thong and Chua have patented a specific detection of Salmonella spp., *S.* Typhi and *S.* Paratyphi A based on protein-coding genes (Thong & Chua 2013). Besides protein-coding genes, 16S rRNA which is highly conserved in all bacteria, have also been used for the detection of *S.* Typhi (Woo *et al.*, 2001, Zhu *et al*., 1996).

Recently, the specificity and stability of npcRNAs as biomarkers in various human diseases and cancer have suggests that these molecules may serve as an excellent molecular target for the detection of infectious diseases such as salmonellosis.

1.6 Research Objectives

Feasibility of npcRNA genes as target in multiplex PCR could provide an alternate choice for the specific and sensitive detection of Salmonella spp., *S.* Typhi and *S.* Paratyphi A. Therefore, the objectives of this project were;

1. To identify npcRNA genes that is specific tor *S.* Typhi, *S.* Paratyphi A and *Salmonella* species for the PCR primers design.
2. To optimize multiplex PCR assay for the detection and differentiation of *S.* Typhi, *S.* Paratyphi A and *Salmonella* species
3. To estimate the sensitivity and specificity of multiplex PCR.
4. To evaluate the sensitivity and specificity of the developed multiplex PCR in seeded stool sample.

CHAPTER 2

MATERIALS AND METHODS

2.1 Materials

2.1.1 Chemicals and Reagents

All the chemicals were of analytical grade. Chemicals and reagents used in this study are listed below in alphabetical order according to the manufacturer's name.

Manufacturer	Chemicals
Amresco	Sodium acetate; Sodium hydroxide; Potassium chloride anhydrous.
Bio basic inc (Canada)	EDTA; PCR primers.
Bio-Rad (Hercules,USA)	Chelax; Ethidium bromide (10mg/mL); Sodium deodecyl sulfate (SDS)
Biotools (Canada)	10X PCR buffer; Magnesium chloride ; DNTP mix; DNA polymerase.
Laborotarious Conda (Madrid,Spain)	LB agar (lennox); LB broth (lennox); Mac Conkey
Invitrogen ® (Carlsbad,CA)	Ultra-pure Tris
Merck KGaA (Darmstadt, Germany	Ethanol; Acetic acid (glacial) 100%; Sodium hydroxide; Glycerol
Promega (Madison,USA)	100bp DNA ladder; Agarose
Sigma	XLD agar; selenite cystine broth.

2.1.2 Buffers/Solutions

Buffers/solutions together with their components used in this study are listed below in alphabetical order.

Manufacturer	Chemicals
5X loading dye (Orange Dye)	6.75 g of Ficoll; 0.25 g orange G; dissolve in ddH_2O by heating.
10X PBS buffer	80 g NaCl; 2.0 g KCL; 14.4 g of Na_2HPO_4; 2.4 g KH_2PO_4 in 1L of ddH_2O pH 7.4.
10X TAE buffer	242 Tris base; 57.1mL Acetic acid; 100 mL 0.5 M EDTA, adjust pH to 8.5, add ddH_2O to final volume of 1 L.
Buffer for Genomic DNA extraction	
1) Solution I	Dissolve 20 % sucrose, 50 mM Tris-Hcl, pH 8.0, 1% SDS, 0.2M Sodium hydroxide, 25 mM EDTA, pH 8.0 and 0.1 M NaCl.
2) Solution II	3M Sodium acetate, pH 6.4 (pH adjusted using glacial Acetic acid).
3) Solution III	95% Ethanol
4) Solution IV	3 M sodium acetate, pH 5.2 (pH adjusted using glacial Acetic acid).

2.1.3 Bacteria Strains

Strains	**Origin/Reference**
Salmonella enterica Serovar	
Typhi Paratyphi A Paratyphi C	Department of Microbiology & Parasitology, School of Medical Sciences, Universiti Sains Malaysia

Pullorom Jawa Kedougou Mikawashima Give Hadar Corvalis Bareily Newport Agona Tennessee Typhimuruim Weltevreden Enteritidis Albany Paratyphi B Braenderup Infantis	Veterinary Research Institute Ipoh,Malaysia
Non-Salmonella strains	
Klebsiella pneumonia *Pseudomonas aeruginosa* *Esherichia coli* *Shigella flexnari* *Vibrio cholera* *Acenitobacter baumanii* *Aeromonas hydrophiila* *Neisseria meningitides* *Streptococcus spp.* *Staphylococcus epidermis* *Providencia spp.*	Clinical Trial Centre, Advanced Medical and Dental Institute, Universiti Sains Malaysia

2.1.4 Plasmid

Plasmid	Origin/References
pL50 (Seegene, USA)	Forever 100bp ladder personalizer Gene information: c-myc gene isolated from mouse Imprinting gene region tissue.

2.1.5 Culture Media

Culture media	Ingredients	Manufacturer
LB agar (lennox)	Trypton 10 g; yeast extract5g;NaCl 5g; bacteriological agar 15 g; ddH_2O to 1 L.	Laborotorious conda (Madrid,Spain)
LB broth (lennox)	Trypton 10g;Yeastextract 5g;NaCl5g;Bacteriological agar 15g; ddH_2O to 1L.	Laborotorious conda (Madrid,Spain)
Mac conkey	Peptone 17g; Proteose Peptone 3g; NaCl 5g; Cystal violet 1mg; Neutral red 30 mg; Bile salts 1.5 g; Bacteriological agar 13.5 g; ddH_2O to1 L	Laborotorious conda (Madrid,Spain)
Selenite cystine broth	Sodium phosphate 10g; Peptone mixture 5 g; Lactose 4 g; Sodium selenite 4 g; L-cystine 0.01 g; ddH_2O to 1 L.	Sigma Aldrich (Missouri. USA)
XLD agar	Yeast extract 3 g; Lactose 7. 5g; Sucrose 7.5 g; Xylose 3.5 g; L-lysine 5 g; Ferric Ammonium Citrate 0.8 g; Phenol red 0.08 g;	Laborotorious conda (Madrid,Spain)

	Sodium chloride 5 g; Sodium deoxycholate 2.5 g; Sodium thiosulfate 6.8 g; Bacteriological agar 13.5 g; ddH_2O to 1 L	

2.2 Methods

2.2.1 Bacteria growth

Single colony of bacteria aforementioned in section 2.1.3 was inoculated into 10 mL of LB broth (Lennox) and incubated at 37ºC at 200 revolutions per minute (rpm) for 16 hours (h) to produce fresh overnight culture.

2.2.2 Bacterial Culture Storage Conditions

For short-term storage, fresh culture was streaked onto LB agar/Mac Conkey incubated for 16 h at 37°C to obtain single colonies and the culture was sub-cultured every 2 weeks. For longer storage, glycerol was aseptically added to reach final concentration of 20% v/v to 200 µL of overnight bacterial culture stored at -80°C. In case of revival the 100 µL of glycerol stock will be added into 10 mL of LB broth.

2.2.3 Bacterial Isolates

Twenty-two *Salmonella* species and 15 other bacterial strains (Gram-negative and Gram-positive) were obtained from the Veterinary Research Institute, Clinical trial centre, Advanced Medical and Dental Institute (AMDI) and Department of Medical Microbiology and Parasitology, School of Medical Science, University Sains Malaysia, Kelantan (Section 2.3.1).

2.2.4 Bacterial Genomic DNA Extraction

Genomic DNA extraction was performed using an in-house protocol. In brief, the bacterial strains were cultured in Luria–Bertani (LB) broth for 16 h (37°C, 200 rpm). Subsequently, 1.5 mL of each culture was centrifuged (Kubota, Japan) at 13,000 rpm for 1 minute (min), and the supernatants were discarded. The pellets were resuspended in 200 μL of Solution I, which was followed by addition of 200 μL of Solution II. The tubes were gently inverted, incubated for 5 min at room temperature, and centrifuged (13,000 rpm, 5 min). The supernatants were transferred to fresh tubes. One mL of solution III was added to each tube, followed by 40 μL of solution IV (pH 5.2). The tubes were then incubated at –80°C for 20 min and centrifuged (13,000 rpm at 4°C, 13 min). The supernatants were discarded. The pellets were washed with 1 mL of 70% Ethanol, air-dried, and dissolved in 50 μL autoclaved distilled water. The concentration of bacterial DNA was measured using UV spectrophotometer (Eppendorf, Germany).

2.2.5 Multiplex PCR (mPCR) Amplification

PCR parameters were optimized and amplification was performed in 20 μL reaction volume, which contained the following: 5 pmol of each forward and reverse primer for *StyR-3, StyR-36, StyR-143,* and *pL50*; 200 μM dNTPs; 3 mM $MgCl_2$; 2 U Taq DNA polymerase in 1X PCR buffer. For template DNA, 1 μL (100 ng) of genomic DNA extracted from bacterial cultures or 2 μL of DNA extracted from spiked stool samples. PCR was performed in a Bio-Rad thermocycler (Bio-Rad, USA) with an initial denaturation at 95°C for 1 min, followed by 30 amplification cycles (30 seconds (s) denaturation at 95°C, 30 s annealing at 66°C, 30 s extension at 72°C), and a final elongation of 2 min at 72°C.

2.2.6 PCR Product Analysis

The PCR products were analyzed by electrophoresis using a 2% agarose gel in TAE buffer containing 0.5 μg/mL of Ethidium bromide. The PCR reaction products (20 μL) were electrophoresed at 60 V for 60 min and visualized using a gel-imaging system (Bio-Rad, USA).

2.2.7 Specificity of the mPCR Assay

Thirty-seven (37) bacterial strains genomic DNA from different sources (Section 2.1.3) were used to assess the analytical specificity of the optimized mPCR assay. One hundred

(100) ng of genomic DNA of each *Salmonella* and other bacteria strains was subjected to PCR. Specificity were confirmed through the amplification of target organisms.

2.2.8 Sensitivity of the mPCR Assay

The analytical sensitivity represents the lowest concentration of template DNA that can be amplified to produce visible bands upon gel electrophoresis.

2.2.8.1 Sensitivity of the mPCR using Genomic DNA

To determine sensitivity of the assay, genomic DNA from *S.* Typhi, *S.* Paratyphi A, and *S.* Paratyphi B (*i.e.* non-*S.* Typhi and non-*S.* Paratyphi A) was extracted (Section 2.2.4) and serially diluted 10-fold from 100 ng to 1 pg. The lowest concentration that produces amplicon will be regarded to be 'detection limit' of the assay.

2.2.8.2 Sensitivity of the mPCR using Spiked Human Stool Sample

An artificial stool contamination experiment (stool spiking) was carried out based on a method described by Kongmoung, with slight modification (Kongmuang *et al*., 1994). Single colonies of *S.* Typhi and *S.* Paratyphi A were inoculated and grown overnight in LB broth. The bacterial cultures were adjusted to 0.5 MacFarland turbidity standard which is used to standardize the approximate number of bacteria in a liquid suspension by visually comparing the turbidity of suspension with MacFarland standard value. Turbidity of bacteria cell suspension were adjusted to 0.132 at OD_{600} which is equivalent to 1.5 x 10^8 CFU/mL, using UV spectrophotometer,. A ten-fold serial dilution was

performed (1.5 x 10^8 CFU/mL to 1.5 x 10^0 CFU/mL), and 1 mL of each dilution was spiked into 0.2 g of healthy human stool (confirmed to be *Salmonella* negative through culturing). The infected feces were then mixed with 9 mL of selenite cystine broth. The spiked stool samples (before and after enrichment) were processed for PCR amplification. In brief, 1.5 mL of each spiked stool sample was centrifuged (13,000 rpm, 1 min). The supernatants were then discarded, and the pellets were washed with 500 μL of 0.01 M phosphate-buffered saline (PBS). After resuspension with PBS, the samples were centrifuged (13,000 rpm, 10 min), and the supernatants were again discarded. The pellets were then resuspended in 50 μL of 10% chelex and boiled for 10 min. Following centrifugation (13,000 rpm, 15 min), 2 μL of each of the supernatants was used as template in the mPCR assay. The lowest CFU/mL that produces amplicon will be regarded to be 'detection limit' of the assay.

CHAPTER 3

RESULTS

3.1 Selection of the npcRNA Gene Candidates as Diagnostic Markers in mPCR Assay.

StyR-3, *StyR-36* and *StyR-143* are the specific npcRNAs that was experimentally identified in the genome of *S.* Typhi (Table 3.1). *StyR-3* was reported to be present in Salmonella spp., *StyR-36* was specific to *S*. Typhi and *StyR-143* was found to be specific for both *S*. Typhi and *S*. Paratyphi A. Therefore, *m*PCR amplification was carried out targeting these npcRNA genes which were predicted to participate in the pathogenecity of Salmonella (Chinni *et al*., 2010).

Table 3.1: The sequence of npcRNA candidates

npcRNA candidates	**Sequence (5 - 3')**
StyR-3	TTACTCACTCATAATCAAGGGCTGCCGCATGAAGTGGTAGAAAAGCAT ATTGCAGGCCATGCGATAAGCCGTCTCACAATTTGTGTGGTTATTACTA TGCTTATTGCTGTTGCCGTAAATGTGCGGTGCGGGAGCCGCTGACGA
StyR-36	CCATGCGCTTGCGCTAAGAGACGTCAGGTATCTATGGAGGAACAAGTT ATGGATACAAACGAACTTGGCTTAGTTAAGGCGCGTGTTGAACTGATC ACCGCTATGCTCAAATGCGCAACCGCGTTTGTTGGCTTAGTTGGTGCGG TTTACGCCGTTCTTAACATGGCCTTCAACT
StyR-143	ATTCTTACTCAAAAAGACAAGGGAGGAATGCCGCAAGAAACCAAAGA AAATCATGGGTTTCATTAAACTTCATTATTGAAGAGGTTTAATAAAGCT GGTTCTATAGGTGCGCGCCTGCTCGTCTTTCATTGTGCCAGCTTTTCT