IDENTIFICATION OF *Salmonella enterica* serovar Typhi PUTATIVE VIRULENCE FACTORS USING A YEAST MORPHOLOGY ASSAY

PONG SZE YEN

UNIVERSITI SAINS MALAYSIA

2018

IDENTIFICATION OF *Salmonella enterica* serovar Typhi PUTATIVE VIRULENCE FACTORS USING A YEAST MORPHOLOGY ASSAY

by

PONG SZE YEN

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

September 2018

ACKNOWLEDGEMENT

I would like to express my gratitude to everyone throughout my master study that had led to the accomplishment of this research project and thesis.

Lecturers,

Supervisor, Dr. Eugene Ong Boon Beng — for his valuable advices and guidance.

Co-supervisor, Prof Dr. Mohd Nazalan Bin Mohd Najimudin— for the sharing of ideas.

INFORMM lecturers — for their ideas and advice given during project presentations.

Typhoid lab members — for their assistance, technical help and support.

INFORMM students and staffs — for their friendship and help.

Universiti Sains Malaysia and INFORMM — for use of facilities and candidature matters.

FRGS grant — for the funding to support this research project.

MyBrains — for the scholarship that had supported my study for two semesters.

Lastly, I would like to express my sincere gratitude to my beloved family and friends for their support, encouragement and love throughout the days of conducting this project.

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ABBREVIATIONS

%	Percentage
µl	Microliter
µm	Micrometer
6×His-tag	Polyhistidine-tag
bp	Base pair
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
E. coli	*Escherichia coli*
ER	Endoplasmic reticulum
h	Hour
HR	Homologous region
HT	High-throughput
g	Gram
GFP	Green fluorescent protein
kbp	Kilo-base pair
LB	Luria broth
LPS	Lipopolysaccharides
M	Molar
MB	Megabyte
MCS	Multiple cloning site
min	Minutes
ml	Milliliter
mM	Millimolar
OD	Optical density

ORF	Open reading frame
PCR	Polymerase chain reaction
s	Seconds
SC	Synthetic complete medium
S. cerevisiae	*Saccharomyces cerevisiae*
SCV	*Salmonella*-containing vacuole
Sif	*Salmonella*-induced filaments
SPI	*Salmonella* pathogenicity Island
S. Typhi	*Salmonella enterica* serovar Typhi
T3SS	Type III secretion system
VF	Virulence factor
YPD	Yeast extract-Peptone-Dextrose medium

PENGENALPASTIAN FAKTOR VIRULENS PUTATIF *Salmonella enterica serovar* Typhi MENGGUNAKAN ASAI MORFOLOGI YIS

ABSTRAK

Demam kepialu adalah jangkitan bakteria bersimtomatik yang disebabkan *Salmonella enterica serovar* Typhi (*S*. Typhi) yang biasanya berlaku akibat system sanitasi yang tidak baik. Pengenalpastian faktor virulens (VF) novel *S*. Typhi akan membolehkan pemahaman yang lebih baik mengenai patogenesis jangkitan demam kepialu dan membantu pengurusan penyakit ini. Terdapat VF yang masih tidak dikenalpasti dan untuk VF yang telah diketahui, peranan VF tersebut dalam patogenesis masih belum difahami sepenuhnya. Dalam kajian ini, yis *Saccharomyces cerevisiae* (*S. cerevisiae*) digunakan sebagai organisma model untuk mengenal pasti VF *S*. Typhi putatif. Sebanyak 192 VF putatif dipilih daripada proteom *S*. Typhi CT18 melalui semakan kajian literatur dan carian pangkalan data atas talian. Kemudian, gen VF putatif terpilih diklon dan diekspres dalam yis menggunakan format plat 96-telaga. Sebanyak 190 VF putatif diklonkan ke dalam plasmid yang mempunyai protein fluoresen hijau dan diekspres dalam yis. Dengan menggunakan asai pertumbuhan mikroplat yang mengukur kepadatan optik kultur cair, sembilan VF putatif didapati merencat pertumbuhan yis. Lokasi protein VF putatif dalam yis diperhatikan dan dikategorikan, satu protein efektor berlokasi di membran sel dan empat protein efektor masing-masing berlokasi di retikulum endoplasma dan sitoplasma. Asai morfologi yis menunjukkan bahawa lima VF putatif menjalani pertumbuhan filamen dan 98 VF putatif menunjukkan perubahan saiz sel yang ketara. Kesimpulannya, pemerhatian morfologi yis adalah penting untuk mengenal pasti VF putative yang boleh

menyebabkan gangguan pada proses intrasel tetapi tidak semestinya menghalang pertumbuhan yis.

IDENTIFICATION OF *Salmonella enterica* serovar Typhi PUTATIVE VIRULENCE FACTORS USING A YEAST MORPHOLOGY ASSAY

ABSTRACT

Typhoid fever is a symptomatic bacterial infection caused by *Salmonella enterica* serovar Typhi that usually occurs due to poor sanitation and hygiene. The identification of novel virulence factors (VF) of *S.* Typhi will allow for better understanding of the pathogenesis of typhoid infection and potentially help disease management. There are still unidentified VFs and for those already reported, their roles in pathogenesis are still not completely understood. In this study, *Saccharomyces cerevisiae* was used as a model organism to identify putative *S.* Typhi. A total of 192 putative VFs were selected from the proteome of *S.* Typhi CT18 through cross-reference with published literature and database searches. Then, the selected VF genes were cloned and expressed in yeast in a 96-well plate format. A total of 190 putative VFs were cloned into plasmid with green fluorescent protein and expressed in yeast. By using a microplate growth assay that measures optical density of liquid culture, nine putative VFs were found to inhibit the yeast growth. Then the putative VF proteins' localization in yeast were observed and categorised, one effector protein localized at cell membrane and four effector proteins each at endoplasmic reticulum and cytoplasm. Result of the morphology assay showed that five of the putative VFs undergo filamentous growth and 98 putative VFs showed significant cell size changes. In conclusion, yeast morphology observation is essential to identify putative VFs that can interfere with yeast intracellular process but not necessarily inhibit yeast growth.

CHAPTER 1
INTRODUCTION

1.1 General Introduction

Salmonella enterica serovars Typhi (*S.* Typhi) is one of the oldest human diseases and believed to cause the "plague of Athens" during the Peloponnesian war of 430 B.C (Galan, 2016). *S.* Typhi is a human host-restricted organism that causes typhoid fever. It remains a major global health problem with the disease burden of 9.9 to 24.2 million cases and 75,000 - 208,000 deaths per year (Das *et al.*, 2017). High areas of typhoid endemicity include South-central Asia, South-east Asia, and Southern Africa (Mogasale *et al.*, 2014).

Virulence factors (VF) are molecules produced by pathogens that act individually or together to facilitate the replication and dissemination of the bacterium within a host. These factors are either secretory, membrane associated or cytosolic in nature (Sharma *et al.*, 2017). Research efforts have been focused to understand the pathogenicity of *S.* Typhi and the identification of *S.* Typhi VFs can clarify *S.* Typhi pathogenesis mechanisms and thus provide a roadmap for future treatment strategies and prevention of typhoid fever (Hurley *et al.*, 2014).

In this study, the yeast *Saccharomyces cerevisiae* (*S. cerevisiae*) was used as a model organism to study *S.* Typhi VFs. *S. cerevisiae* has been shown to be a good model for characterizing bacterial VFs although it cannot be used as a physiological model for human infections (Aleman *et al.*, 2009). The use of yeast cells as a model of eukaryotic cells can reduce time and cost needed for this study (Aleman *et al.*, 2009; Slagowski *et al.*, 2008). When VFs are expressed in yeast cells, it can interfere with normal cellular processes of yeast cells and lead to growth inhibition. Thus, inhibition

in yeast growth can serve as an indicator to screen for the presence of VFs (Slagowski *et al.*, 2008). Additionally, previous studies also showed that VFs when expressed in yeast can cause morphological changes to their cells (Aleman *et al.*, 2009; Lesser & Miller, 2001).

The putative VFs genes were selected through bioinformatic analysis and literature review. Then the selected genes were cloned into plasmid that will allow for green fluorescent protein (GFP) tagging. Protein expression was carried out in a high-throughput format. This was followed by screening of putative VFs using a yeast growth inhibition assay and yeast morphology observation. The subcellular localization of the selected VF proteins were also determined. The overall approach used is outlined in Figure 1.1.

1.2 Hypothesis and Aim of Thesis

Not all *S.* Typhi VFs have been identified, and for those identified not all their functions are known. Thus, the aim of this study is to identify putative *S.* Typhi VFs by identifying proteins that can cause morphological change in *S. cerevisiae* cells.

Specific aims are:

1. To curate a list of putative *S.* Typhi VFs through bioinformatics analysis and literature review.
2. To develop a high-throughput cloning and protein expression platform.
3. To determine the growth inhibitory activity of the putative VFs.
4. To identify morphological changes caused by VFs in yeast cells through microscopy.

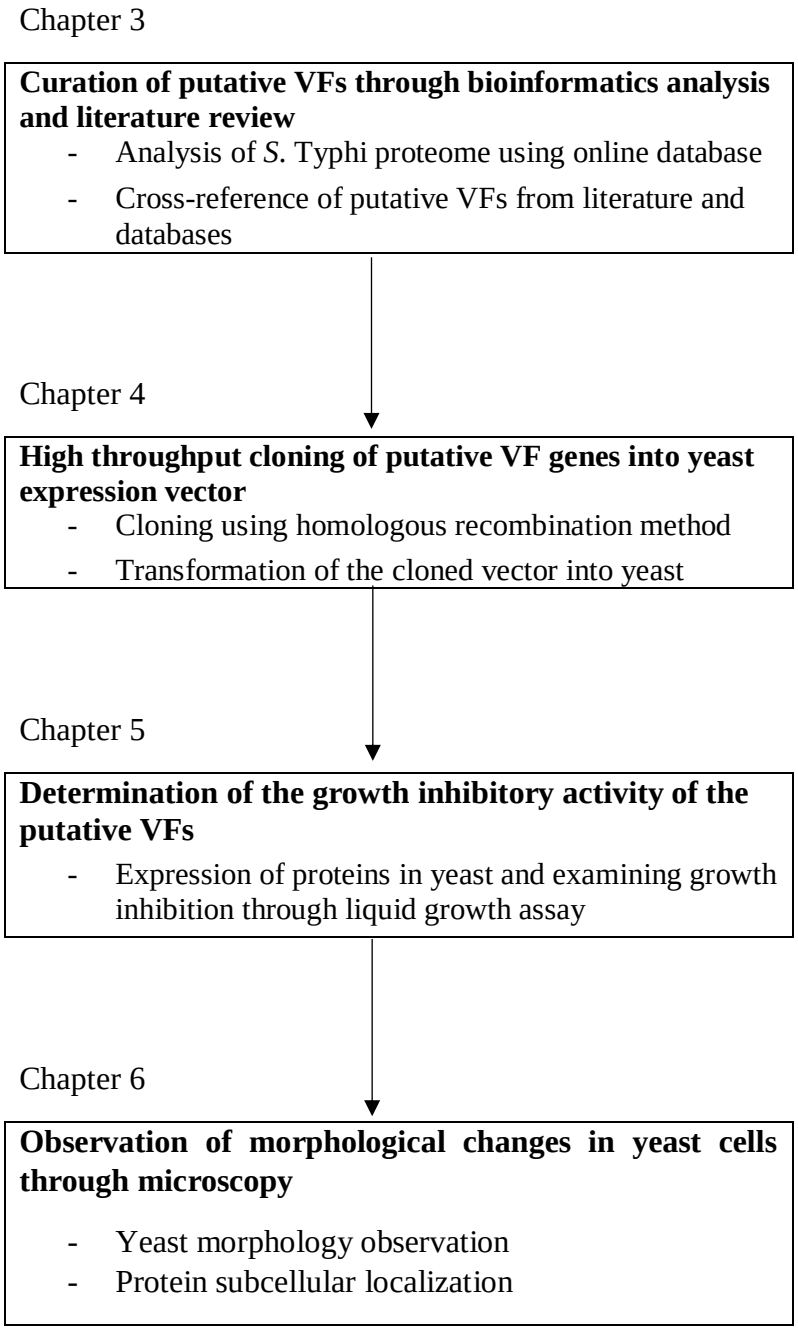

Figure 1.1 Outline and flow of thesis.
Experimental approach taken to identify *S.* Typhi VFs and study of its effect on cellular morphology.

CHAPTER 2

LITERATURE REVIEW

2.1 *Salmonella enterica* serovar Typhi

Salmonella enterica serovar Typhi (*S*. Typhi) is a Gram-negative, facultative anaerobic and flagellated bacilli from the family of Enterobacteriaceae (Lamas *et al.*, 2018). The genus of *Salmonella* contains 2600 serotypes (termed "serovar") and classified into 2 species, which is *Salmonella enterica* and *Salmonella bongori* (Ryan *et al.*, 2017).

Based on the Kauffman-White classification scheme, *S. enterica* is further divided into six distinct subspecies which are *enterica*, *salamae*, *arizonae*, *diarizonae*, *houtenae* and *indica*. Of these six subspecies, only subspecies *enterica* is associated with disease in warm-blooded animals (Baddam *et al.*, 2014).

Human-adapted *S*. Typhi is a worldwide foodborne pathogen and the causative agent of the typhoid fever (Lamas *et al.*, 2018). Typhoid fever remains a common infection in regions with poor economic development and limited public health infrastructure (Dougan & Baker, 2014). It is transmitted through unhygienic water sources, poor hygiene practices and unsanitary living conditions (Mogasale *et al.*, 2014; Wain *et al.*, 2015).

Following infection through ingestion, symptoms begin after an incubation period that usually lasts 7-14 days. The onset of diseases is marked by fever as high as 39° to 40°C. Followed by abdominal distension, constipation, headache, rash, malaise, loss of appetite, nausea, vomiting, hepatosplenomegaly and leukopenia (Akinyemi *et al.*, 2005). Some serious complications such as gastrointestinal bleeding,

intestinal perforation and typhoid encephalopathy may occur in 10-15% of typhoid patients (Non *et al.*, 2015).

It is estimated that 26.9 million cases of disease per year reported worldwide, resulting in 216 000 deaths (Buckle *et al.*, 2012). High incidence areas of typhoid endemicity include South-central Asia, Southeast Asia, and Southern Africa (Figure 2.1) (Mogasale *et al.*, 2014). Although the global burden of typhoid fever has reduced, the emergence of multi-drugresistant *S.* Typhi (MDRST) is still a serious health problem worldwide (Chiu *et al.*, 2002).

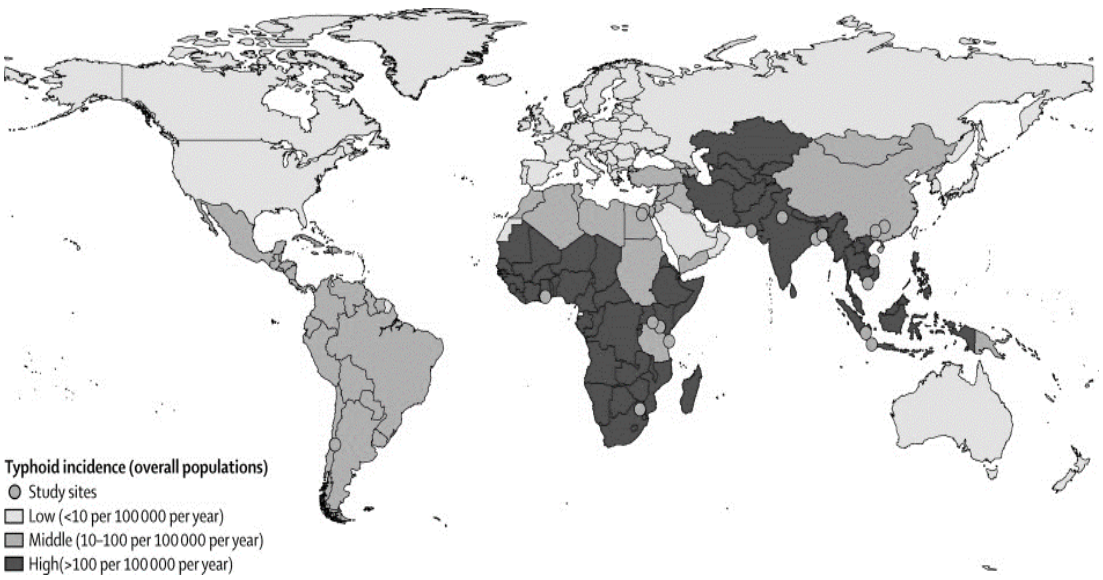

Figure 2.1 Geographical distribution of typhoid incidence.
Typhoid have high incidence at South-central Asia, Southeast Asia, and Southern Africa (adapted from Mogasale *et al*., 2014).

2.2 *Salmonella* Infection

S. Typhi enter human digestive system through contaminated food or water. After surviving the low pH of stomach acid，*S.* Typhi come in contact with the epithelial cells of intestines. *Salmonella* preferentially invades the intestinal barrier through microfold (M) cells and later disseminated into lymph nodes, spleen and liver (Haraga *et al.*, 2008; Suarez & Russmann, 1998) (Figure 2.2A).

Internalization of *Salmonella* into host cells occur depending on the invading cells type, such as phagocytic and non-phagocytic cells using a type III secretion system (T3SS1) (Figure 2.2B). Phagocytosis in *S.* Typhi involves multiple receptors to activate different signalling pathways in the phagocyte (Ibarra & Steele-Mortimer, 2009). Whereas, T3SS1 mediated invasion is a highly specific process that is regulated by the expression of a number of effector proteins (SipA, SipC, SopB/SigD, SopD, SopE2 and SptP) (Kage *et al.*, 2008).

When contact with the intestinal epithelial cells, *S.* Typhi inject T3SS1, which is a needle like complex into the host cells membranes and assemble a channel between the *S.* Typhi and cell membranes. Translocated proteins will enter the host cell cytosol through the T3SS1. These protein will trigger a complex set of signalling events in the host cell which cause cytoskeletal rearrangements (Schmidt & Hensel, 2004). The rearrangement of the host cells actin cytoskeleton will result in membrane ruffling that uptake the *S.* Typhi into the host epithelial cells (Boumart *et al.*, 2014).

Besides phagocytosis and T3SS1 mediated invasion, fimbriae and adhesins on the surface of *S.* Typhi also facilitate the attachment and internalization through T3SS-1 independent process (Ibarra & Steele-Mortimer, 2009).

Once the epithelial barrier has been breached, *S.* Typhi will activate various virulence mechanisms for survival in the host cells environment. This promotes bacterial replication and subsequently invade into the reticuloendothelial system (RES) (Haraga *et al.*, 2008).

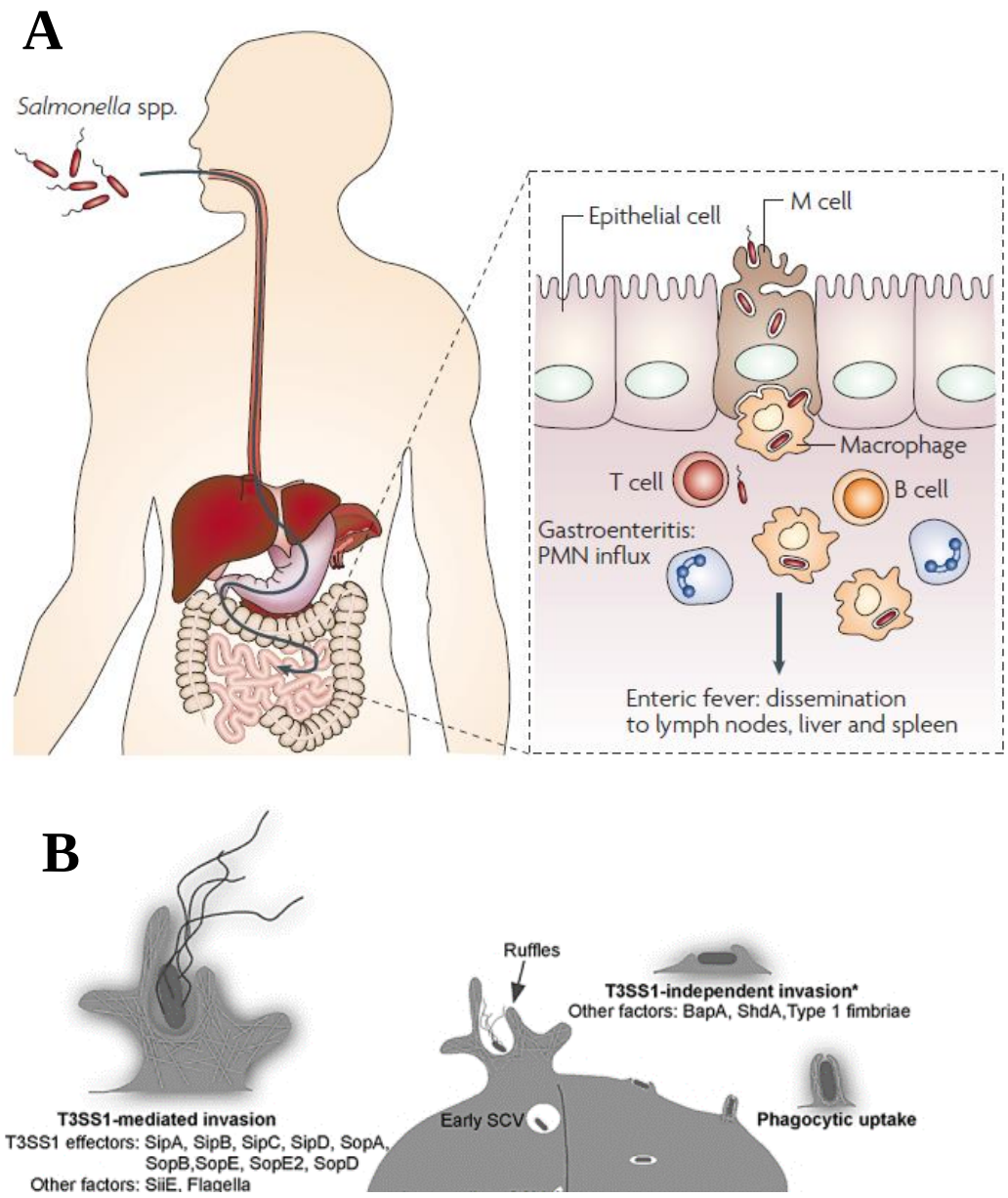

Figure 2.2 Overview of *Salmonella* infection.
A, Ingested *Salmonella* invade host intestinal epithelial cells through M cell and cross the intestinal barrier before disseminated to other parts of body (adapted from Haraga *et al.*, 2008). B, Different methods of internalization of *Salmonella* into host cells (adapted and modified from Ibarra & Steele-Mortimer, 2009).

2.3 Host Specificity of *S*. Typhi

S. Typhi is human restricted pathogens which only develop and caused diseases in human host. Studies showed that chimpanzees can be infected with *S.* Typhi but do not develop typhoid fever. Indeed, the *S.* Typhi-infected animals developed symptoms which are more "nontyphoidal Salmonella" infection alike than those of typhoid fever (Song *et al.*, 2013). It is hypothesized that typhoid toxin which only encoded by typhoidal *Salmonella* serovars *S.* Typhi and *S.* Paratyphi play important role in the development of typhoid fever (Galan, 2016).

Glycoarray analysis showed typhoid toxin prefer to bind to termini with the consensus sequence Neu5Ac2-3Galβ1-3 and does not bind to termini with N-glycolylneuraminic acid (Neu5Gc) (Galan, 2016). This results provided perception for *S.* Typhi human-host specificity because sialoglycans on human cells were terminated in N-acetylneuraminic acid (Neu5Ac) (L. Deng *et al.*, 2014). Whereas for other primates and mammals, it is terminated in Neu5Gc. This difference was caused by a mutation in human gene which emerged after hominids separated from other primates. This mutation encodes enzyme CMP-N-acetylneuraminic acid hydroxylase (CMAH), which converts Neu5Ac to Neu5Gc. This insight is strengthed when typhoid toxin is capable to induce typhoid fever symptoms in mice. The expression pattern of CMAH in mice is variables, displaying sialoglycans terminated in both Neu5Ac and Neu5Gc (Hedlund *et al.*, 2007). Thus, these explaining reasons *S.* Typhi only cause typhoid fever in human host.

2.4 Virulence Factors

Virulence factors (VF) are molecules produced by pathogens to promote the invasion and survival of pathogens in the host cells environment. The ability of bacteria to infect a host determined by multiple virulence factors acting individually or together at different stages of infection (Wu *et al.*, 2008). Some VFs regulate each other or antagonistically to promote infections. VFs secreted during late stage of infection may modulate the functions of those secreted during early stage, changing the bacterial virulence strategy to promote infection. Also, a single VF can have multiple functions and host targets, such as *S*. Typhi SopB which involve in rearrangement of actin cytoskeleton and also stimulating fluid secretion which cause diarrhoea (Hurley *et al.*, 2014; Ong *et al.*, 2010; Sharma *et al.*, 2017).

Identification of *S*. Typhi VFs are important to understand the mechanism of *S*. Typhi pathogenesis and their interactions with the host, thus provide a roadmap for future treatment strategies and prevention of typhoid fever. Here are some example of *Salmonella* VFs: Pathogenicity islands, fimbriae, flagellar, Vi antigen, lipopolysaccharides and many more (Figure 2.3) (Wu *et al.*, 2008).

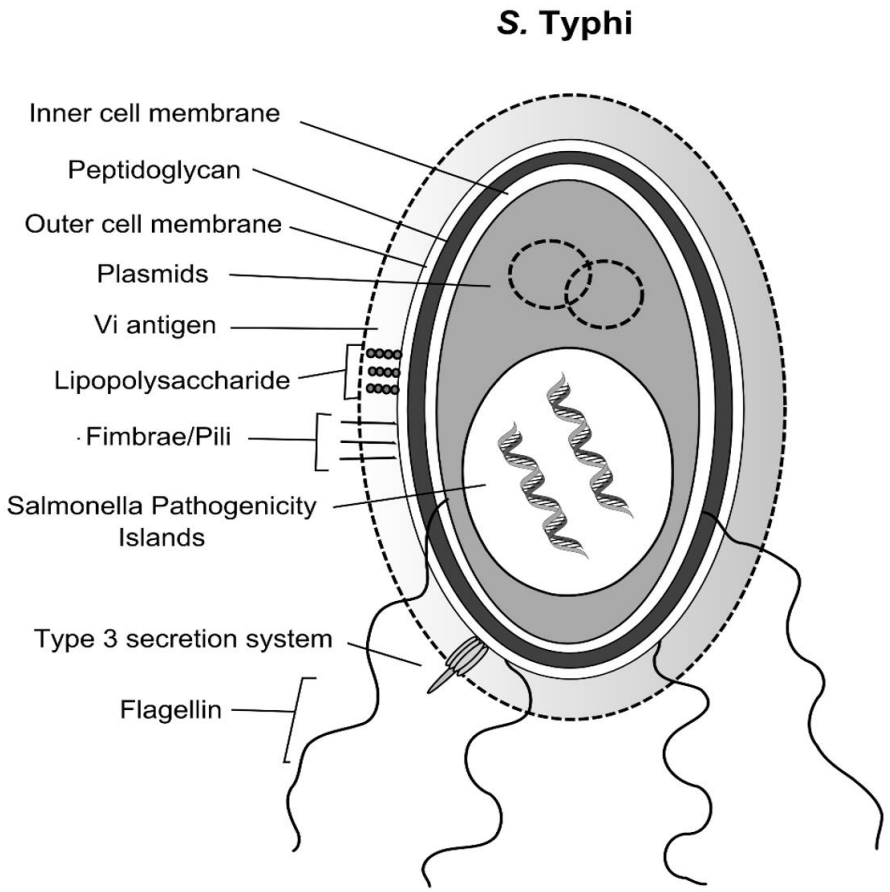

Figure 2.3 Overview of *S*. Typhi virulence factors.
Different types *S*. Typhi virulence factors (adapted and modified from de Jong *et al.*, 2012).

2.4.1 *Salmonella* Pathogenicity Island

Majority of the *Salmonella* VFs are clustered on *Salmonella* Pathogenicity Islands (SPI) on the chromosome. To date, there are 21 SPI reported. However, not all SPI are present in all *Salmonella* serotypes, only SPI-1, SPI-2, SPI-3, SPI-4 and SPI-5 were present in all serotypes. Based on the complete genome sequence of *S.* Typhi CT18, there are at least 5 more regions can be designated as SPIs (Parkhill *et al.*, 2001).

Salmonella pathogenicity island 1 (SPI-1) is 40 kb DNA region, bearing the genetic information encoding for type III secretion system (T3SS1) at centisome 63 at *Samonella* chromosome (Galan, 2001). T3SS1 play important roles during interaction between the *Salmonella* and host intestinal epithelial cells helping the entry into host cells and initiation of the diseases. Whereas, SPI-2 located at the centisome 30.7 of the bacterial chromosome and encoding for two-component regulatory system and T3SS2 (Schmidt & Hensel, 2004)

Other than these two major SPIs, other SPIs (SPI-3, SPI-4, SPI-5, SPI-7, SPI-8, SPI-9 and SPI-10) have also been identified to play important roles in virulence and survival of the bacteria (Table 2.1)(Ong *et al.*, 2010). However, most of the SPIs function are not well studied yet.

Table 2.1 SPIs and their functions in virulence (adapted from Edwards *et al.*, 2000; Ong *et al.*, 2010; Valdez *et al.*, 2009).

SPI	Encode	Function in virulence
SPI-1	- T3SS1	Invasion of epithelial cells
SPI-2	-Two-component regulatory system -T3SS2	Survival of *S.* Typhi inside *Salmonella* Containing Vacuole (SCV)
SPI-3	- Mg2+ uptake system	Salmonella survival in the SCV
SPI-4	- Nonfimbrial adhesin	Bacterial adherence to the apical surface of polarized cells
SPI-5	- Effector proteins that are secreted by SPI-1 and SPI-2 T3SS	*Salmonella* enteropathogenesis
SPI-6	- *saf* and *tcf* fimbrial operons	Bacterial adherence
SPI-7	- Vi biosynthetic genes - Type IV fimbrial operon	Host cell cytoskeletal rearrangement
SPI-8	- Colicin/pyocin	Kill other bacteria to compete for nutrients
SPI-9	- Type I secretion system - Repeats-in-toxin (RTX) proteins	Adhesion in intestine
SPI-10	- *sef* fimbrial operon	Colonization and penetration of the intestinal barrier.

2.4.2 Type III Secretion System

T3SSs are supramolecular complexes that play a major role in the invasion of host cells, it serves as a channel for the effector proteins to enter into host cytosol (Ibarra & Steele-Mortimer, 2009). The T3SS proteins can be categorized into export apparatus, needle complex, translocons, regulators, effectors, and chaperones (Figure 2.4).

SPI-1 encoded T3SS (T3SS1) genes are expressed when *S*. Typhi contact with host epithelial cells under regulation of T3SS1 regulatory proteins (InvF, HilA, HilC and HilD) (Srikanth *et al.*, 2011). The translocated effector proteins will enter into to T3SS1 needle complex (PrgI and InvG) through the export apparatus (SpaO, PrgH and PrgK) which assembled at the *S*. Typhi inner membrane (burkinsaw). SpaS and InvA function as gate for controlling the access into the needle complex. On the other hand, translocon proteins (SipB, SipC and SipD) which located at the other end of needle complex and responsible for forming pores on host cell membranes (Haraga *et al.*, 2008). Once the effector proteins enter into host cytosol, it will induce host actin cytoskeleton rearrangement to form a 'pocket' to engulf and uptake the *S*. Typhi into the host cells (Ramos-Morales, 2012).

The SPI-2 T3SS genes are only expressed inside host SCV and regulated by three important two-component regulatory systems including SpiR/SsrB, PhoP/PhoQ and EnvZ/OmpR in response to acidic pH and nutrient limitation in the SCV(Ramos-Morales, 2012). The activated SPI-2 T3SS facilitated the translocation of the effector proteins across SCV membrane into host cells through the needle complex (SsaG) (Kaur & Jain, 2012). After entering into host cells, these VFs will interfere host cellular processes for replication and survival in the host cells (Kaur & Jain, 2012). For instance, SpiC, a protein translocated into host macrophages cytosol will interfere normal secretory pathway of the host cell which protects the *Salmonella* for bactericidal compounds (Haraga *et al.*, 2008).

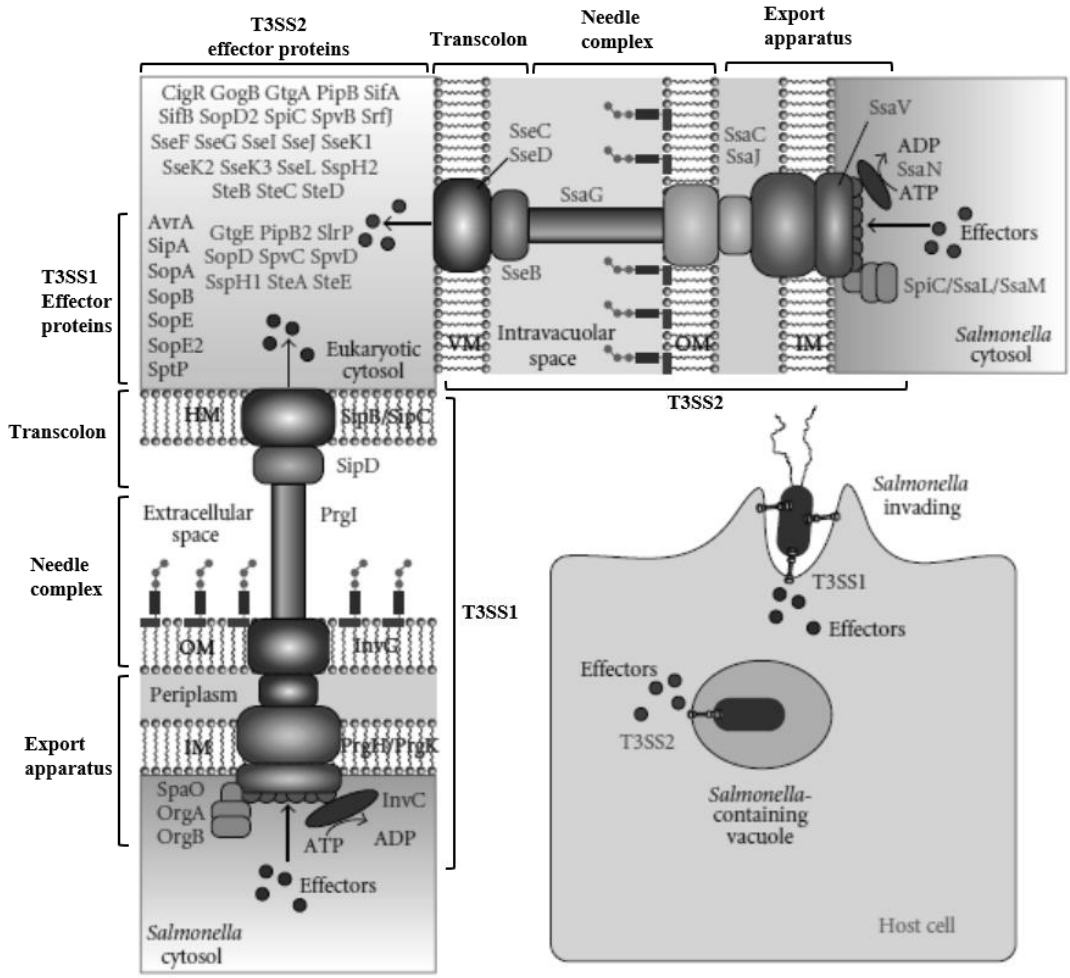

Figure 2.4 Structural overview of T3SS and location of the proteins.
The export apparatus located within the inner membrane of *Salmonella*. Needle components are assemble into the needle complex passing through *Salmonella* outer membrane into extracellular environment. The transcolon components form pores on host cell membrane and transferring *Salmonella* effector proteins into the host cell cytoplasm. SPI-1 encoded T3SS (T3SS1) translocate VFs from bacterial cytoplasm to host cells, across the host plasma membrane whereas SPI-2 T3SS (T3SS2) is expressed inside SCV and translocate VFs to host cytoplasm, across the vacuolar membrane (adapted and modified from Ramos-Morales, 2012).

2.4.3 Others Virulence Factors

Besides SPI, other VFs of *Salmonella* also play important roles in bacterial pathogenesis. Some example of the VFs are fimbriae and flagellar, Vi antigen, lipopolysaccharides (Wu *et al.*, 2008).

Flagella are complex motility structures which have been associated with virulence in many pathogens, including *Salmonella*. Reduces adherence ability to human intestinal epithelial cell lines are show in flagellar mutants strains, suggesting that flagella are important in adherence of *Salmonella* to intestinal epithelial cells during invasion (Dibb-Fuller *et al.*, 1999).

Fimbriae are hair-like appendages that protrude outside the bacterial surface membrane which play important role in biofilm formation, adhesion and colonization of bacteria in the host cells such as epithelial, lymphoid and endothelial cells (Proft & Baker, 2009).

Vi antigen found to maintain bacterial survival during host immune responses by preventing recognition of host pattern recognition receptors (PRRs) (de Jong *et al.*, 2012). Studies have shown loss of Vi antigen reduced salmonella virulence in mice thus suggested the importance of Vi antigen in bacteria pathogenesis (Raffatellu *et al.*, 2006).

Lipopolysaccharides (LPS) is a major component of the outer membrane of Salmonella. It consists a hydrophobic region (lipid A) that anchors it to the bacterial outer membrane, a non-repeated core oligosaccharide, and a repeated polysaccharide (O antigen) (Raetz & Whitfield, 2002). LPS helps to protect *Salmonella* from the acidic environment of the gastrointestinal tract and important in host-pathogen interactions with host immune system.

2.5 Universal Protein Resource

Universal Protein Resource (UniProt) is a freely accessible database, combining the view of protein sequence and functional annotations through integration of data from different resources (UniProt Consortium, 2018). UniProt can be reached at https://www.uniprot.org. UniProt is composed of several important component parts and each designed for different uses, the UniProt Knowledgebase (UniProtKB), the UniProt Reference Clusters and the UniProt Metagenomic and Environmental Sequence Database (UniProt, 2013).

2.6 Yeast as Model System

Saccharomyces cerevisiae or usually known as baker's yeast has been widely used as a model organisms for more than 50 years (Duina *et al.*, 2014) (Figure 2.5). Up to 30% of genes involve in human disease have orthologs in the yeast sequences and thus conserving most eukaryotic cellular process (Franssens *et al.*, 2013; Karathia *et al.*, 2011). With these criteria, the use of yeast has been shown to be a powerful approach for understanding the interactions of bacterium and its host, providing further pieces of knowledge of their enzymatic functions (Curak *et al.*, 2009; Siggers & Lesser, 2008; Valdivia, 2004).

Previous studies showed yeast was used as a surrogate system to characterize the activities of the Zika viral genome (Li *et al.*, 2017). The studies showed that the Zika viral protein subcellular localization was overall parallel to its predicted protein structure. Some of the Zika viral proteins showed various levels of cytopathic effects such as regulate cell growth, induced hypertrophy, or cellular oxidative stress which lead to cell death. So, yeast can be used as a model to study VFs either with yeast cell growth, cell morphology and also the protein subcellular localization.

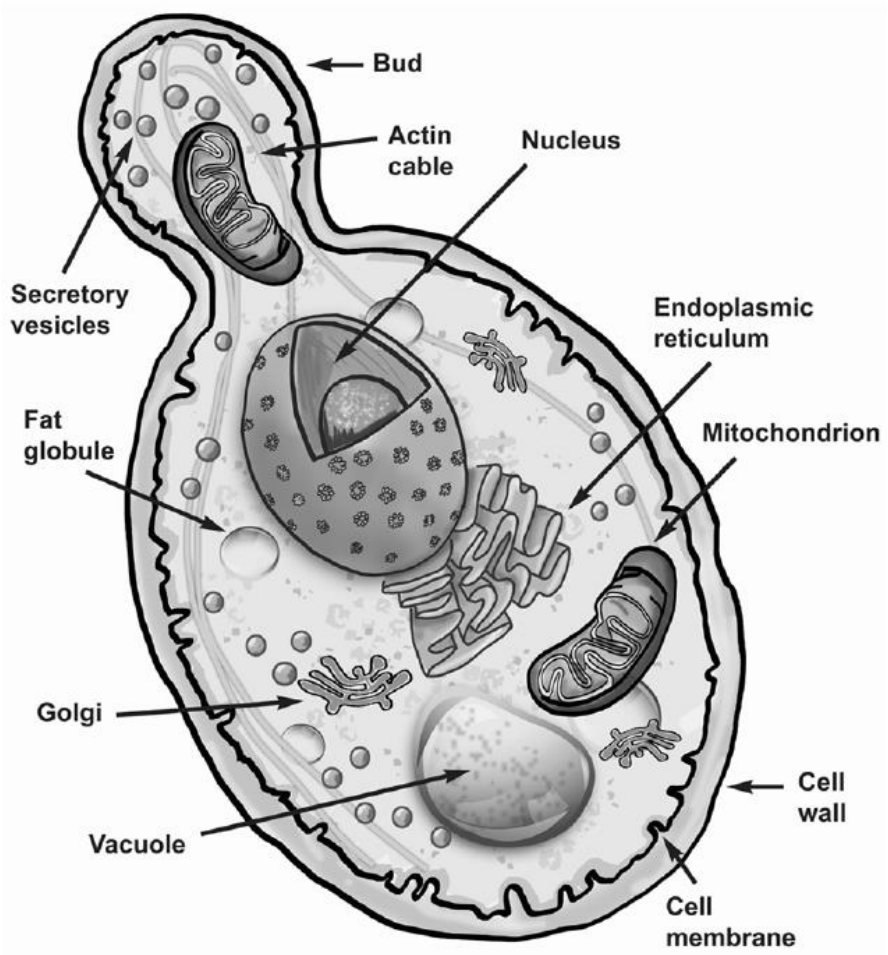

Figure 2.5 Overview of *Saccharomyces cerevisiae* subcellular compartments. Yeast cells contain various subcellular (adapted from Walker, 2016).

2.6.1 Yeast growth Inhibition as a Reporter for *S*. Typhi Virulence Factors

There are increasing studies revealed that heterologous expression of bacterial effector proteins in yeast will lead to yeast growth inhibition (Siggers & Lesser, 2008). For example, effector proteins from *Pseudomonas* aeruginosa, *Shigella* flexneri, *Salmonella* Typhimurium, *Legionella* pneumophila and *Yersinia* species have been observed to inhibit growth when expressed in yeast (Siggers & Lesser, 2008). These growth inhibition is the indirect result of effector proteins interfere with the yeast cytoskeleton, organelle membranes, or specific signalling pathways (Popa *et al.*, 2016). Thus proving that yeast growth inhibition is a sensitive and specific reporter to screen for bacteria effector proteins or virulence factors which can perturb host cellular processes (Aleman *et al.*, 2009).

2.6.2 Yeast to study *S*. Typhi Effector Proteins Subcellular Localization

Previous studies suggested that bacterial effector heterologous expressed in yeast could accurately reflect their localization when injected into host cells during invasion (Siggers & Lesser, 2008). This shows that the molecular mechanisms of effector targeting are likely conserved from yeast to human. For example, *Salmonella* effector proteins SopE2 are localized at yeast plasma membrane and PipB targeted to endoplasmic reticulum, precisely matching the localization which previously reported in mammalian cells (Weigele *et al.*, 2017).

Few studies visualized the protein subcellular localization by fusing the protein of interested with fluorescent proteins. Green fluorescent protein (GFP) is one of the fluorescent proteins which is widely used to tagged proteins of interest. The subcellular localization of the effector proteins can be categorised by the GFP localization pattern or with the help of other organelles fluorescent stains as guides (Weigele *et al.*, 2017)

(Figure 2.6). This localization patterns of proteins in yeast providing useful tool to identify the subcellular localization and molecular mechanisms of the proteins (Lippincott-Schwartz *et al.*, 2003).

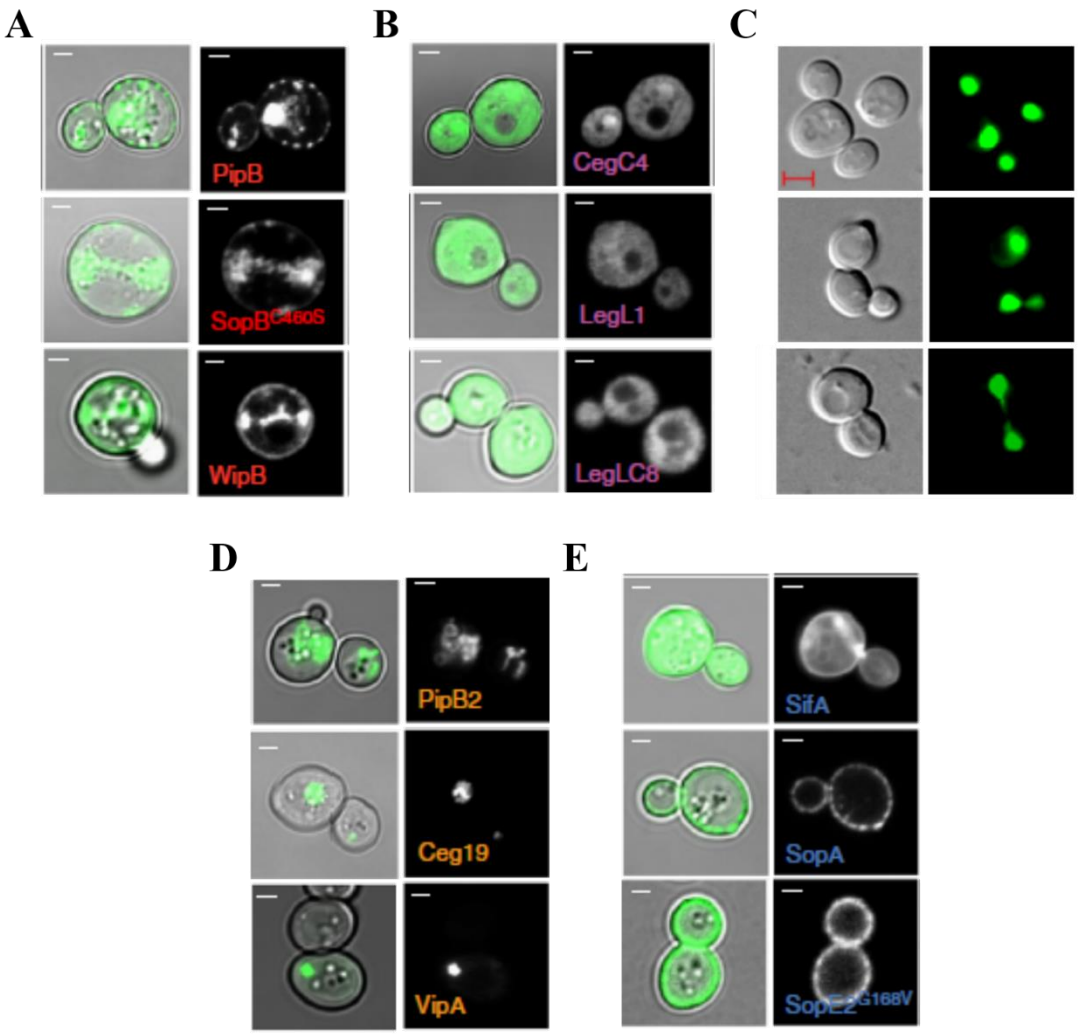

Figure 2.6 Bacterial effector protein localization in yeast.
Subcellular localization of GFP tagged proteins in yeast cells. A, endoplasmic reticulum; B, cytoplasm; C, nucleus D, vacuoles E, cell membrane (adapted and modified from Henke *et al.*, 2011; Weigele *et al.*, 2017).

2.6.3 Yeast to Study the Expressed *S.* Typhi Proteins mediated Morphology Changes

Besides using subcellular localization to study eukaryotic cellular processes, another approach is to screen for effector proteins that can mediate the yeast morphology changes. For example, many pathogens interfere with the host actin cytoskeleton during the invasion process. Such as *Shigella* and *Salmonella* which mediate their uptake into normally nonphagocytic cells by delivering effector proteins that induce the formation of membrane ruffles in the cells (Cossart & Sansonetti, 2004; Ly & Casanova, 2007). Both yeast and mammals conserved the molecular switches Rho GTPases that regulate the actin cytoskeleton (Etienne-Manneville & Hall, 2002).

Some effector proteins will cause different responses when introduced into different cell types, such as expression of *Salmonella* SopE2 that activates Cdc42, results in formation of membrane ruffles in mammalian cells (Rho GTPases pathway). While in yeast, it is associated with activation of the filamentous growth pathway (MAPK pathway), causing the yeast to filament (Figure 2.7) (Rodriguez-Pachon *et al.*, 2002).

Besides the filamentous shape of yeast cells, expression of bacterial proteins also caused changes in yeast cell size, such as enlargement or reduce in yeast cell sizes, due to the disruption of the yeast intracellular processes by bacterial effector proteins. The changes in cell size will indirectly changing the cell volume or surface area which will interfere with cell normal metabolic flux, biosynthetic capacity, nutrient exchange and most importantly cell division process (Marshall *et al.*, 2012).

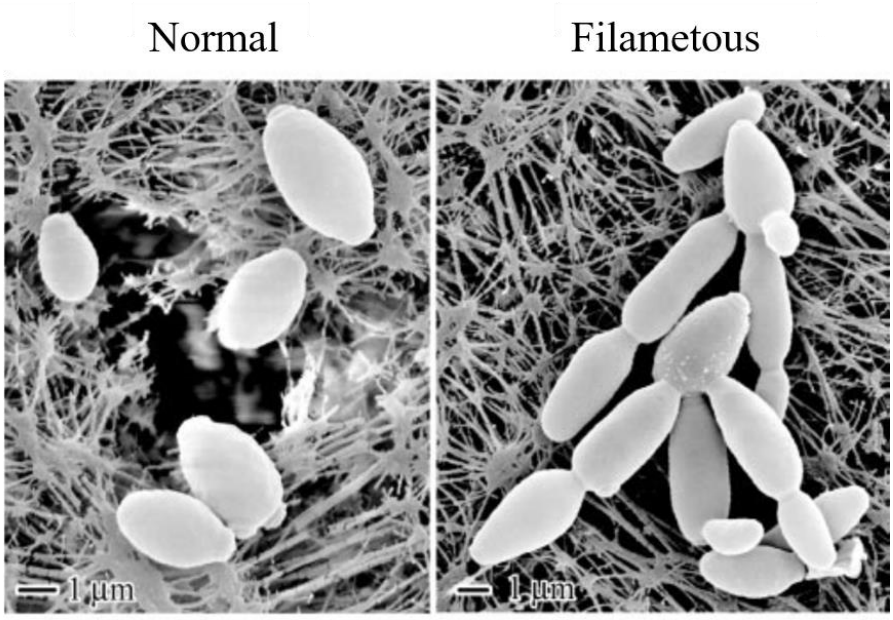

Figure 2.7 Normal and filamentous yeast
Scanning electron micrographs showing normal and filamentous yeast cells which expressing *Salmonella* SopE2 (adapted and modified from Rodriguez-Pachon *et al.*, 2002).

CHAPTER 3

CURATION OF PUTATIVE VIRULENCE FACTORS THROUGH BIOINFORMATICS ANALYSIS AND LITERATURE REVIEW

3.1 Introduction

VFs are molecules produced by pathogens to assist the colonization and survival of pathogens in the host cells environment. These factors are either secretory, membrane associated or cytosolic (Sharma *et al.*, 2017). All these factors are working together or individually during different stages of infection to ensure the establishment of the diseases (Wu et al., 2008). Secretory factors play important roles in helping the bacterium to survive the host immune response and also host cell–bacteria interactions. Whereas the membrane associated VFs are involved in adhesion and invasion of host cells. The cytosolic factors facilitate the bacterium metabolic and physiological processes (Sharma *et al.*, 2017).

UniProtKB/Swiss-Prot contain manually curated records with information extracted from literature and curator-evaluated computational analysis. UniProtKB/Swiss-Prot currently contains about half a million sequences and continues to grow as new proteins are experimentally reported (UniProt, 2014). Whereas for the sequences which are unreviewed collected in UniProtKB/TrEMBL. The entries of UniProtKB/TrEMBL are automatically generated annotation supplemented by computationally analyzed records (Poux *et al.*, 2017). Both (UniProtKB/Swiss Prot) and (UniProtKB/TrEMBL) provide protein information including biological process, molecular function and cellular component (UniProt, 2013). Therefore, UniProtKB is a useful online tool to facilitate the process of acquiring of protein sequences and functional information of *S*. Typhi proteomes in this study.

Besides UniProt, VF databases such as Virulence Factor Database (VFDB) and PathoSystems Resource Integration Center (PATRIC) were also referred when curating a list of possible VFs for testing. VFDB (http://www.mgc.ac.cn/VFs/) is a free online database that serves to provide a source for scientists to access to current knowledge of VFs from various bacterial pathogens (Chen *et al.*, 2013). On the other hand, the PATRIC database provide complete bioinformatics resource for pathogens, including genomics, proteome and metabolic pathway data to facilitate fundamental biomedical research on bacterial caused diseases (Wattam *et al.*, 2017).

3.2 Experimental Approach

The UniProtKB database was used to search and download *S.* Typhi complete proteome. Possible VFs were selected by analysing the proteins based on protein general annotation, gene ontology (GO) provided in the database and further cross-referenced to putative VFs reported in literature and virulence databases (VFDB, PATRIC).

3.3 Methods

3.3.1 Selection of VFs through Bioinformatic Analysis and Literature Review

S. Typhi proteins with general annotation including protein names, gene names, biological processes, functions, subcellular location and length were downloaded from UniProtKB (http://www.uniprot.org/uniprot/?query= proteome:UP000000541+AND+proteomecomponent:%22Chromosome%22) in Microsoft Excel file format. A list of possible VFs were selected through the analysis of the proteins based on the general annotation and filtered with keywords related to VFs. The keywords used were pathogenesis, virulence, Pathogenicity Island, Type III secretion, secreted, flagella, fimbria, toxin, Vi polysaccharide, lipopolysaccharide (LPS), adhesion and invasion. The selected VFs were then cross-referenced with