

**Investigation in to the functional role of  
Rgg Quorum Sensing Systems in  
*Streptococcus pneumoniae***

Thesis submitted for the degree of  
Doctor of Philosophy at the  
University of Leicester

Xiangyun Zhi

Department of Infection, Immunity and Inflammation

University of Leicester

September 2017



### **Statement of originality**

The accompanying thesis submitted for the degree of PhD entitled “Investigation in to the functional role of Rgg Quorum Sensing Systems in *Streptococcus pneumoniae*” is based on work conducted by the author in the Department of Infection, Immunity and Inflammation at the University of Leicester mainly during the period between October 2013 and September 2016. All the work recorded in this thesis is original unless otherwise acknowledged in the text or by references.

None of the work has been submitted for another degree in this or any other University.

Signed: \_\_\_\_\_ . Date: \_\_\_\_\_ .

## **Investigation in to the functional role of Rgg Quorum Sensing Systems in *Streptococcus pneumoniae***

**Xiangyun Zhi**

### Abstract

The members of microbial community communicate with each other by using quorum sensing (QS) systems, and modulate their collective ‘behavior’ for in host colonization and virulence, biofilm formation, interspecies competition, and environmental adaptation. Recent influx in genome data availability reveals the presence of several putative QS sensing circuits in microbial pathogens, but many of these have not been functionally characterized despite their utility as drug targets. To increase the repertoire of functionally characterized QS systems in bacteria, we studied Rgg144/Shp144 and Rgg939/Shp939, two putative QS systems in the important human pathogen *Streptococcus pneumoniae*. I find that both of these QS circuits are induced by short hydrophobic peptides (Shp) upon sensing sugars found in the respiratory tract, such as galactose and mannose. Microarray analysis using cultures grown on mannose and galactose revealed that the expression of large number of genes is controlled by these QS systems, especially those encoding for essential physiological functions and virulence related genes such as the capsular locus. Moreover, the array data revealed evidence for cross talk between these systems. Finally, these Rgg systems play a key role in colonisation and virulence, as deletion mutants of these QS systems are attenuated in the mouse models of colonisation and pneumonia.

## **Acknowledgement**

Firstly I would like to express my appreciation to my supervisor Dr Hasan Yesilkaya, He always inspire me and motivate me when I felt frustrated. Thank you for always making the time for me, and be patient, especially when I am new here, I don't know that much about microbiology. Thank you for believing in me and for always pushing to get the best out of me.

I would like thank Professor Peter Andrew. Thanks for his constant support, fantastic suggestions on experimental design trouble shooting. Thanks for being patient and understanding.

I would like to thank Prof Russell Wallis and Dr Christopher Furze for their help with protein purifications.

I would like to thank my friends Yan, Vicky, Shiqiu and the members in lab 125 for their supporting and friendship.

I need to say thank you to my family. My parents, they are always caring and supporting. Thanks for their love and everything they have done for me.

---

**Abbreviations**

<b>ABC</b>	ATP-binding cassette	<b>NAD</b>	Nicotinamide adenine dinucleotide
<b>ACL</b>	acyl-homoserine lactone	<b>g</b>	Gram
<b>AIPs</b>	auto-inducing peptides	<b>ng</b>	Nanogram
<b>APS</b>	Ammonium persulphate	<b>OD</b>	Optical density
<b>BAB</b>	Blood agar base	<b>ONPG</b>	2-Nitrophenyl $\beta$ -Dgalactopyranoside
<b>BgaA</b>	$\beta$ -galactosidase A	<b>PAGE</b>	polyacrylamade gel electrophoresis
<b>BHI</b>	Brain heart infusion	<b>PavA</b>	pneumococcal adhesion and virulence A protein
<b>bp</b>	Base pair	<b>PBS</b>	Phosphate buffered saline
<b>BSA</b>	Bovine serum albumin	<b>PCR</b>	Polymerase chain reaction
<b>Cbp</b>	choline-binding proteins	<b>PCV</b>	Pneumococcal conjugate vaccine
<b>CcpA</b>	Carbon catabolite protein A	<b>PiaA)</b>	pneumococcal iron acquisition A
<b>CCR</b>	Carbon catabolite repression	<b>PiuA</b>	pneumococcal iron uptake A protein
<b>CDM</b>	Chemically defined medium	<b>pmol</b>	picomole
<b>CFU</b>	Colony forming unit	<b>pNPNANA</b>	2-O-(p-nitrophenyl)- $\alpha$ -D-N-acetylneuraminic acid

<b>COG</b>	Clusters of orthologous groups	<b>Ply</b>	pneumolysin
<b>CPS</b>	capsule polysaccharide synthesis	<b>PsaA</b>	pneumococcal surface antigen A
<b>CRP</b>	C-reactive protein	<b>QQ</b>	Quorum quenching
<b>CSP</b>	Competence stimulating peptide	<b>QS</b>	Quorum sensing
<b>dH<sub>2</sub>O</b>	Distilled water	<b>Rgg</b>	regulator genes of glucosyl-transferase
<b>DNA</b>	Deoxyribonucleic acid	<b>RNA</b>	Ribonucleic acid
<b>dNTP</b>	Deoxynucleotide triphosphate	<b>RNAP</b>	RNA polymerase
<b>DTT</b>	Dithiothreitol	<b>ROS</b>	Reactive oxygen species
<b>EDTA</b>	Ethylenediaminetetraacetic acid	<b>rpm</b>	revolutions per minute
<b>EMSA</b>	Electrophoretic mobility shift assay	<b>RR</b>	response regulator
<b>Eng</b>	endo- $\alpha$ -N-acetylgalactosaminidase	<b>SDS</b>	Sodium dodecyl sulphate
<b>Eno</b>	Enolase	<b>SHP</b>	Short hydrophobic peptide
<b>FDRs</b>	false discovery rates	<b>SodA</b>	Superoxide dismutase
<b>Gal</b>	Galactose	<b>SpuA</b>	Pullulanase
<b>GlcNAc</b>	<i>N</i> -acetyl glucosamine	<b>SpxB</b>	Pyruvate oxidase
<b>Glu</b>	Glucose	<b>StrH</b>	$\beta$ -Nacetylglucosaminidase

<b>His</b>	Histidine	<b>TAE</b>	Tris acetic acid EDTA
<b>Hyl</b>	Hyaluronidase	<b>TB</b>	Tris boric acid
<b>IgG</b>	Immunoglobulin G	<b>TCSs</b>	Two-Component Signal Transduction Systems
<b>IPTG</b>	Isopropyl $\beta$ -D-1-thiogalactopyranoside	<b>TEMED</b>	Tetramethylethylenediamine
<b>kb</b>	Kilobase	<b>TFs</b>	Transcription factors
<b>kDa</b>	Kilodalton	<b>TpxD</b>	Thiol peroxidase
<b>HK</b>	histidine kinase	<b><math>\mu</math>g</b>	Microgram
<b>IPD</b>	Invasive pneumococcal disease	<b><math>\mu</math>l</b>	Microliter
<b>LA</b>	Luria-Bertani agar	<b>UPD-Glc</b>	Uridine diphosphate-glucose
<b>LB</b>	Luria-Bertani broth	<b>UV</b>	Ultraviolet
<b>LDH</b>	Lactate dehydrogenase	<b><math>\mu</math>M</b>	Micromolar
<b>LytA</b>	Autolysin A	<b><math>\mu</math>M</b>	Micromolar
<b>Man</b>	Mannose	<b>w/v</b>	Weight per volume
<b>mg</b>	Milligram	<b>x g</b>	Gravity force
<b>MIC</b>	Minimum inhibition concentration	<b>XIP</b>	SigX inducing peptide
<b>ml</b>	Millilitre assay		

# Table of Contents

Chapter I: Introduction .....	1
1.1 <i>Streptococcus pneumoniae</i> biology .....	2
1.2 Epidemiology .....	4
1.3 Antibiotic therapy, resistance, and vaccine .....	5
1.4 Virulence determinants.....	7
1.5 Pneumococcal adaptation in host niches .....	12
15.1 Metal homeostasis .....	13
15.2 Resistance to oxidative stress .....	14
15.3 Carbohydrate acquisition and metabolism .....	15
1.6 Bacterial transcriptional regulation .....	17
1.6.1 Transcription factors (TFs).....	19
1.6.2 Two-Component Signal Transduction Systems .....	20
1.6.3 Quorum sensing systems in bacteria .....	22
1.6.4 Transcriptional regulators can be targeted to develop novel antiinfectives.....	27
1.7 Pneumococcal transcriptional regulation.....	30
1.7.1 Pneumococcal standalone regulators.....	30
1.7.2 Pneumococcal two component systems .....	31
1.7.3 Pneumococcal Quorum sensing systems .....	34
1.8 Rgg regulators .....	35
1.9 Hypothesis .....	40
1.9.1 Hypothesis on Rgg’s involvement in non-glucose sugar metabolism.....	40
1.9.2 Hypothesis on Rggs role in oxidative stress resistance .....	41
1.9.3 Hypothesis on Rggs role in biofilm formation .....	43
1.9.4 Hypothesis on Rggs role of pneumococcal virulence.....	43
1.9.5 Hypothesis on Rgg-SHP pairs being part of a quorum sensing system .....	44
1.10 Aims .....	45
Chapter II : Materials and Methods.....	46
2.1 Bacterial strains and plasmids .....	47
2.2 Chemicals .....	49

---

2.2.1 Antibiotics .....	50
2.3 Media Preparation .....	50
2.3.1 Brain Heart Infusion Medium .....	50
2.3.2 Blood agar base .....	51
2.3.3 Lauria Bertani medium.....	51
2.3.4 Todd-Hewitt yeast extract medium (THY) .....	51
2.3.5 Chemically defined medium.....	51
2.4 Preparation of cultures.....	53
2.4.1 Growth of <i>Escherichia coli</i> .....	53
2.4.2 Growth of <i>Streptococcus pneumoniae</i> .....	53
2.4.3 Stock preparation of bacterial strains .....	54
2.5 Assessment of colony forming unit counts .....	54
2.6 Plasmid extraction from <i>E. coli</i> .....	55
2.7 Extraction of pneumococcal chromosomal DNA.....	55
2.8 Polymerase Chain Reaction (PCR) .....	56
2.9 Agarose gel electrophoresis.....	59
2.10 Restriction digestion.....	60
2.11 DNA Ligation.....	60
2.12 Method for the production of chemically competent <i>E. coli</i> .....	61
2.13 Genetic transformation of <i>Escherichia coli</i> .....	62
2.14 Transformation of <i>Streptococcus pneumoniae</i> .....	62
2.15 RNA extraction from <i>S. pneumoniae</i> .....	63
2.16 Treatment of RNA with DNase I.....	64
2.17 Synthesis of complementary DNA .....	64
2.18 Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) .....	65
2.19 Microarray experiments.....	67
2.20 Gene Splicing by Overlap Extension PCR (SOEing PCR) .....	67
2.21 H <sub>2</sub> O <sub>2</sub> survival assay .....	70
2.22 Paraquat susceptibility assay .....	70
2.23 Bacterial growth assay.....	70
2.24 Preparation of <i>Streptococcus pneumoniae</i> cell lysate by sonication.....	71
2.25 Neuraminidase activity assay .....	71

---

2.26 Hemolytic assay .....	72
2.27 Glucuronic acid assay.....	72
2.28 Quantification of protein concentration using Bradford Assay .....	73
2.29 Denaturing polyacrylamide gel electrophoresis (PAGE) .....	73
2.30 Protein purification.....	75
2.30.1 Purification of Rgg939 .....	75
2.30.2 Purification of Rgg144 .....	76
2.30.3 Purification of inclusion bodies .....	76
2.30.4 Solubilization and Refolding.....	77
2.30.5 Dialysis.....	78
2.30.6 Gel Filtration .....	78
2.31 Electrophoretic mobility shift assay (EMSA) .....	78
2.31.1 <i>In silico</i> analysis of putative promoters, and DNA probe preparation .....	78
2.32 Construction of <i>lacZ</i> -fusions .....	80
2.33 $\beta$ -galactosidase activity assay.....	81
2.34 <i>In vivo</i> virulence studies .....	83
2.34.1 Preparation of bacterial inoculum.....	83
2.34.2 Pneumonia model .....	83
2.34.3 Colonisation experiments .....	84
2.35 Statistical analysis .....	85
Chapter III: Results .....	86
Section A. Construction of isogenic mutants and genetically complemented strains .....	87
3.1 Mutagenesis by gene splicing overlap extension PCR (SOEing PCR) .....	87
3.1.1 Construction of $\Delta rgg144$ , $\Delta rgg144/shp144$ , and $\Delta shp144$ .....	88
3.1.2 Construction of $\Delta rgg939$ , $\Delta rgg939/shp939$ , and $\Delta shp939$ .....	91
3.1.3 Construction of $\Delta rgg144/939$ and $\Delta rgg/shp144/939$ .....	94
3.1.4 Confirmation of mutagenesis by PCR.....	94
3.2 Complementation of mutant strains.....	100
3.2.1 Amplification of <i>rgg</i> genes for complementation .....	101
Section B. Phenotypic characterisation of the Rgg mutants.....	105
3.3 Growth of pneumococcal strains in BHI.....	105
3.4 Growth of pneumococcal strains in CDM with different sugars.....	107

---

3.5 Rggs confer protection against H <sub>2</sub> O <sub>2</sub> .....	111
3.6 Rggs involve in paraquat resistance .....	113
3.7 Phenotypic characterisation of Rgg/Shp circuits.....	115
3.7.1 Neuraminidase activity .....	115
3.7.2 Hemolytic activity .....	115
3.7.3 Capsule content determination by glucuronic acid assay.....	116
Section C. Construction of <i>lacZ</i> reporter systems and $\beta$ -galactosidase assays.....	118
3.8 Construction of <i>lacZ</i> -fusions .....	118
3.8.1 Identification and amplification of the putative promoters of selected genes .....	119
3.8.2 Preparation and digestion of plasmid pPP2.....	121
3.9 Assessing the inducibility of Rgg-Shp circuits.....	124
3.9.1 Promoter induction by Shp .....	124
3.9.2 Shp144 is a secreted peptide.....	126
3.9.3 SHP144-C12 is the most active variant of SHP144 peptides .....	128
3.9.4 SHP939-C8 is the most active variant of SHP939 peptides .....	130
3.9.5 Induction of $\beta$ -galactosidase activity depends on the concentration of peptide ...	132
3.9.6 Rgg144 is required for C12 induction of P <sub><i>shp144</i></sub> .....	134
3.9.7 Rgg939 is required for C8 induced activation of P <sub><i>shp939</i></sub> .....	135
3.9.8 The role of carbon source on the expression of <i>rggs</i> .....	137
Section D. Determination of Rgg144 and Rgg939 by Microarray analysis .....	140
3.10 Microarray analysis for $\Delta$ <i>rgg144</i> .....	140
3.11 Microarray analysis for $\Delta$ <i>rgg939</i> .....	141
3.12 Confirmation of microarray results by qRT-PCR .....	144
Section E. Expression and purification of recombinant proteins .....	150
3.13 Expression and purification of Rgg144.....	150
3.13.1 Amplification of target genes and cloning .....	150
3.13.2 Cloning, transformation and DNA sequencing .....	151
3.13.3 Small-scale protein expression .....	151
3.13.4 Large-scale protein expression and purification.....	151
3.14 Expression and Purification of recombinant Rgg939.....	153
3.14.1 Expression and purification.....	154
Section F. Electrophoretic mobility shift assay (EMSA) .....	156

---

3.15 <i>In silico</i> promoter analysis.....	157
3.16 Amplification of putative promoter regions .....	159
3.17 Electrophoretic mobility shift assay EMSA .....	160
3.17.1 Recombinant Rgg144 interacts with P <sub>shp144</sub> , P <sub>SPD_0315</sub> , P <sub>SPD_1127</sub> , P <sub>SPD_1370</sub> and P <sub>SPD_2030</sub> .....	160
Section G. <i>In vivo</i> virulence studies .....	162
3.18 Contribution of Rggs and Shps in pneumococcal virulence.....	162
3.19 Contribution of Rgg144 and Rgg939 in pneumococcal colonisation.....	166
Chapter IV: Discussion .....	168
Future plan.....	183
Appendix .....	186
Reference.....	194

# **Chapter I: Introduction**

## **1.1 *Streptococcus pneumoniae* biology**

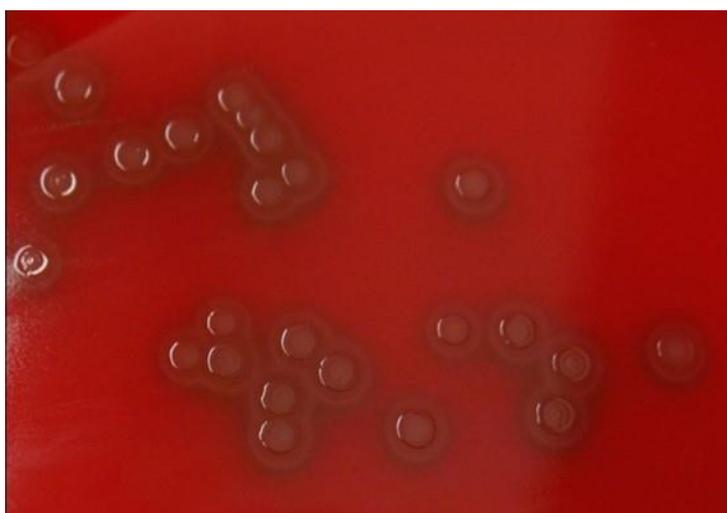
*Streptococcus pneumoniae* is a member of genus *Streptococcus*, it is an encapsulated Gram-positive, alpha-hemolytic, aerotolerant coccoid bacterium (Ryan and Ray, 2004) (Figure 1.1 and 1.2). It was previously called *Diplococcus pneumoniae*. In 1974, it was renamed as *Streptococcus pneumoniae* since in liquid media it appears in chains (Deibel and Seeley Jr, 1974). The diameters of individual cells are between 0.5 and 1.5  $\mu\text{m}$ . Under electron microscope, it presents as a "lancet-shaped" diplococci but it is also seen as short chains or single cells (Zhang et al., 2006).



**Figure 1.1 Gram staining of *Streptococcus pneumoniae* (X100 magnification).** *Streptococcus pneumoniae* was grown in BHI broth, gram stain has been down and the slide was visualised under a 100X objective. *Streptococcus pneumoniae* is Gram-positive which grows in pairs and chains.

Pneumococci do not have spores or a flagellum. They are non-motile, facultative anaerobic and prefer to grow in 5% carbon dioxide. Being similar to other Streptococci, it lacks catalase and are frequently incubated on blood agar plates. *S. pneumoniae* can create  $\alpha$ -haemolysis, this type of haemolysis is characterised by a green colour around the colonies due to incomplete breakdown of hemoglobin (Manco et al., 2006) (Figure 1.2). *S. pneumoniae* is ethylhydrocupreine hydrochloride (optochin) sensitive and normally they can form a zone of inhibition around an optochin disc. *S. pneumoniae* can also lyse in bile, which makes *S. pneumoniae*

different from other alpha hemolytic Streptococci. On blood agar base it is cultured at 37 °C for 24 hours, and forms small gray transparent or translucence shinny and flat colonies. *S. pneumoniae* depends on choline when grown in defined media. Although it can grow in a pH range of 6.5-8.3, the optimum being 7.8. The temperature range of growth is 25-42 °C. It is a fermentative organism, and can ferment upto 32 sugars, some of which include glucose, galactose, fructose, lactose, maltose, raffinose, glycogen and inulin (Bidossi et al., 2012, Buckwalter and King, 2012). It is a facultative anaerobe with marked tendency to accumulate hydrogen peroxide when cultivated aerobically. The polysaccharide capsule envelopes the microbe completely for protection against phagocytosis. Based on the difference in capsular antigenicity, over 90 different pneumococcal serotypes have been identified (Mavroidi et al., 2004).



**Figure 1.2** *Streptococcus pneumoniae* colonies on blood agar plate. *Streptococcus pneumoniae* was streaked out on blood agar plate with 5% horse blood and incubated overnight at 37 °C with 5% CO<sub>2</sub>. Colonies are surrounded by a zone of  $\alpha$ -hemolysis.

## **1.2 Epidemiology**

*Streptococcus pneumoniae* is a part of the human microflora which frequently asymptotically colonises the mucosal surfaces of the human nasopharynx and upper airway. Around 20-30% of healthy individuals, particularly infants, carry *S. pneumoniae* (Ghaffar et al., 1999). However, asymptomatic carriage is a crucial step of infection and will lead to spread of this microbe. In addition to being commensal, *S. pneumoniae* is also a crucial pathogenic bacterium and displays a strong virulence activity when it encounters an immune compromised host (Gratz et al., 2015). The microbe can disseminate from the upper to the lower respiratory tract and pneumococcal disease develops as the pathogen spreads to the sinuses, the middle ear or by invasion into the blood stream, thus giving rise to a number of diseases including acute sinusitis, otitis media, meningitis, bacteremia, endocarditis, peritonitis, pericarditis and brain abscess which are responsible for millions of deaths each year (Siemieniuk et al., 2011).

Pneumococcal diseases pose a huge threat to public health, and is one of the crucial causative agents responsible for large amount of human morbidity and mortality (Obaro et al., 2000). It has been estimated that each year approximately 1.9 million individuals die from invasive pneumococcal disease (IPD) all around the world, a large proportion of these cases happens unfortunately in resource restricted countries (Hoskins et al., 2001), due to lack of efficient detection technologies and treatment options. Even in the United States and Europe, the annual incidence of invasive pneumococcal disease ranges from 10 to 100 cases per 100,000 populations. It has been estimated that 15 to 30 per 100,000 people in developed countries suffer from a serious pneumococcal infection, which makes *S. pneumoniae* a very dangerous pathogen (Zahlten et al., 2014).

Pneumococcal pneumonia is the most common case of pneumococcal infection, and is frequently linked to young children from 6 months to 4 years old whose immune

system is not fully functional to ward off infection, and those aged over 60 and individuals with immune-deficiencies (Ryan and Ray, 2004). In recent decades, although the number of child death decreased from 9.9 million in 2000 to 6.3 million in 2013, epidemiological surveys demonstrated that pneumococcal pneumonia is still responsible for an estimated 935, 000 death (15% of child death) under the age of five in 2013. This is more than AIDS (2%), malaria (7%) and measles (2%), which is a huge burden for human health (Liu et al., 2015). It has been reported that approximately 60-87% of bacteremia cases are attributable to pneumococcal pneumonia in the United States which contributes to respiratory failure and meningitis (Carbon et al., 2006). Moreover, it has been estimated that *S. pneumoniae* is one of the most prevalent agents of meningitis, as 47% of bacterial meningitis cases are caused by the pneumococcus in the United States (Swartz, 2004). According to Black *et al.* (2004), each year around 10.6 million children suffer from invasive pneumococcal disease (IPD). The incidence of pneumococcal disease in developing countries is much higher than the incidence in developed countries. Other than the young children and aged, individuals at high risk also include those suffering from HIV. It has been reported that the incidence of IPD is higher among individuals with HIV, 97 cases in every 100,000 (Fedson et al., 1999).

### **1.3 Antibiotic therapy, resistance, and vaccine**

Antibiotics are always considered to be the popular therapeutic strategy for pneumococcal infections. The  $\beta$ -lactam antibiotics including benzyl-penicillin (Penicillin G), ampicillin, cephalosporin C, ceftriaxone and aztreonam compounds are usually used for pneumococcal infections (Normark and Normark, 2002). However, antibiotic resistant *S. pneumoniae* are continuously isolated in many different countries and regions in more recent decades (Zhou et al., 2012). The clonal expansion and spread of multi-resistant *S. pneumoniae* due to misuse of antibiotics, result in the increase in antimicrobial resistance. *S. pneumoniae* isolates are resistant

to various antibiotics, such as penicillin, macrolides and tetracycline. Antibiotic resistance poses a huge challenge for the treatment of pneumococcal infections especially in developing countries which lack efficient medicine and detection methods (Zhou et al., 2012). Previous studies have shown that there is a high level of pneumococcal resistance to penicillin throughout the world with a MIC (Minimum inhibitory concentration) of  $\geq 2$  mg/L (Song, 2013). In Hong Kong, Taiwan and South Korea the frequency of *S. pneumoniae* resistance is more than 50% (Reinert et al., 2005). Efforts to treat pneumococcal infections have been compromised by this continuous increase in spread of antibiotic resistant clones.

More recently, the introduction of the pneumococcal vaccines has markedly reduced the pneumococcal infection. Given the huge threat of IPD on individuals and community, widespread vaccination against this serious situation appears to be an essential and highly recommendable preventive approach (de Paz et al., 2015). Currently, there are two types of vaccine to prevent pneumococcal diseases, the 23-valent polysaccharide vaccine (PS) and a 7-valent conjugate vaccine (PCV-7). The PS vaccine consists of 23 polysaccharides from the most commonly isolated serotypes (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F, 18C, 19A, 19F, 20, 22F, 23F and 33F) (Bogaert et al., 2004), and it is protective for individuals with a fully developed and functioning immune system. However, it has several limitations when applied to young children under the age of 2 years old and individuals with a compromised immune system (Bogaert et al., 2007). This is because capsule polysaccharide (CPS) is a T-cell-independent antigen, and has decreased immunogenicity in young children. Thus PS is relatively poor in triggering immune responses to prevent pneumococcal disease (Bogaert et al., 2004).

PCV-7 is designed to provide protection against seven of the most frequent serotypes (4, 6B, 9V, 14, 18C, 19F, and 23F) associated with invasive disease. It is very effective in children under five years old (Black et al., 2000) and immune-deficient individuals, among which it has greatly reduced the number of pneumococcal

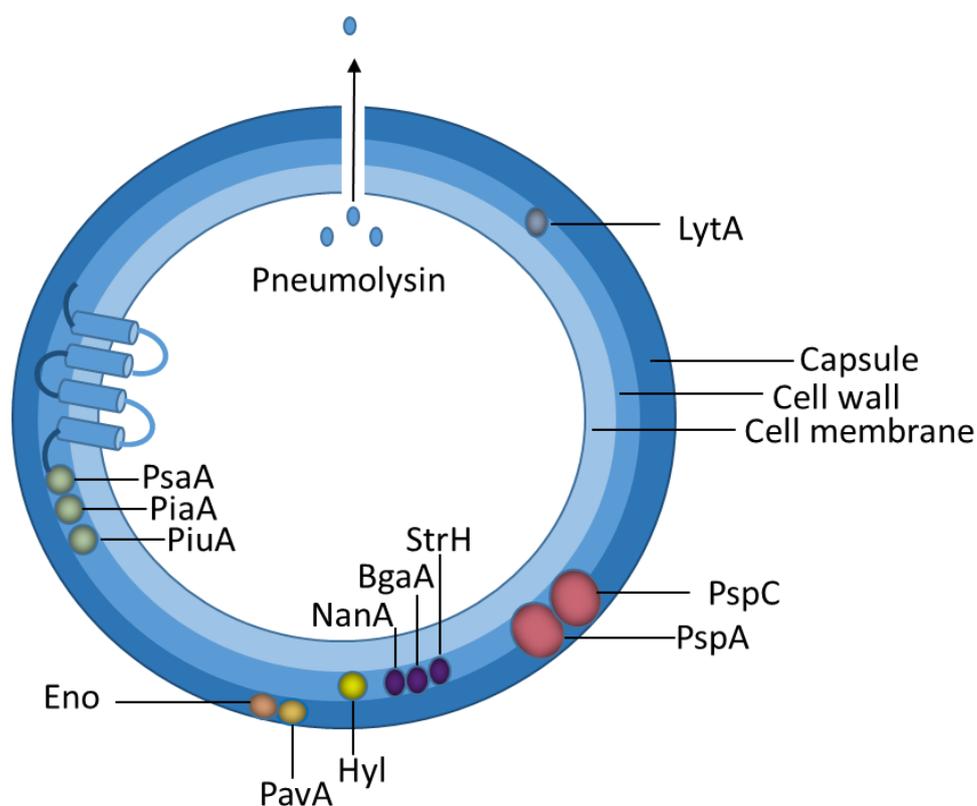
infections (de Paz et al., 2015). PCV-7 contains polysaccharides conjugated to protein carriers, which can induce a T-cell dependent immune response, producing immunoglobulin G (IgG) and memory B-cells. However, it is very expensive to produce so it is not always available in resource-limited countries such as Africa and South East Asia. Moreover, the prevalence of pneumococcal serotypes varies geographically. In addition, PCV-7 is serotype specific and it also suffers from troubles in serotype replacement (Muñoz-Almagro et al., 2008), which refers to shift in the prevalence of circulating serotypes due to mass vaccination. In other words, serotypes that are not represented in vaccine formulation increase in circulation while those that are included decrease. In 2010, PCV13 has been introduced and replaced the PCV-7. PCV13 contains 13 capsule polysaccharide (1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F, 23F) conjugated to protein carrier CRM197 (de Paz et al., 2015). As diphtheria toxoid is conjugated to the 13-valent polysaccharide, PCV13 is competent to induce memory and activate T cell immune responses in both infants and the elderly (Namkoong et al., 2016). It was reported that since routine PCV13 vaccination, IPD decreased 42% overall, and 53% among children (Kaplan et al., 2013). However, PCV13 vaccine cannot prevent pneumococcal carriage (Dagan et al., 2013, Kataoka et al., 2017).

The therapy of pneumococcal diseases is hampered by the increasing prevalence of antibiotic-resistant strains and the limited efficacy of the existing vaccines. Thus, greater understanding of pneumococcal biology, metabolism, and its genetic character is essential so that we can develop new vaccines and antimicrobials to reduce the burden of pneumococcal diseases.

## **1.4 Virulence determinants**

Many pneumococcal virulence determinants are involved in bacterial invasion and modulating host defenses. These virulence factors have various effects on protecting the bacteria against the damage from the host immune system, mediating adhesion,

enabling invasion and improving acquisition of nutrients. Some well-known virulence determinants of *S. pneumoniae* include: the pneumococcal capsule, pneumococcal surface proteins A and C (PspA and PspC) (López and García, 2004), pneumolysin (Ply) (Tilley et al., 2005), pneumococcal surface antigen A (PsaA) (Tseng et al., 2002), pneumococcal adhesion and virulence A protein (PavA), autolysin A (LytA), choline-binding proteins (Cbp), the neuraminidase proteins, hyaluronidase (Hyl), enolase (Eno), pneumococcal iron acquisition A (PiaA) and pneumococcal iron uptake A protein (PiuA) (Kadioglu et al., 2008) (Figure 1.3).



**Figure 1.3. Schematic representation of pneumococcal virulence factors.** NanA, neuraminidase A; StrH,  $\beta$ -N-acetylglucosaminidase; BgaA,  $\beta$ -galactosidase; Hyl, hyaluronate lyase; Eno, enolase; PavA, pneumococcal adherence and virulence factor A; LytA, autolysin; PspC, pneumococcal surface protein C; PspA, pneumococcal surface protein A; PiaA, pneumococcal iron acquisition A, PiuA, pneumococcal iron uptake A, PsaA, pneumococcal surface antigen A, (adapted from Kadioglu et al., 2008 ).

The pneumococcal capsule consists of sugar polymers, and is located in the most outer layer of *S. pneumoniae*, and is approximately 200-400 nm thick, which protects bacteria from phagocytosis by inhibiting complement deposition (Hyams et al., 2010, Kadioglu et al., 2008). It is reported that all the clinical *S. pneumoniae* isolates have a capsule and mutants without a capsule are found to be avirulent (Cohen et al., 2012). Moreover, pneumococcal capsules are usually highly charged at physiological pH, and this may influence the function of phagocytes (Lee et al., 1991). The pneumococcal capsule can produce a shield that protects against the Fc region of Immunoglobulin G (IgG) and prevents the recognition by the receptors on phagocytic cells (Winkelstein, 1981). The capsule is capable of weakening the deposition of complement on the bacterial surface (Abeyta et al., 2003). CPS can inhibit classical complement pathway by preventing C-reactive protein (CRP) and natural antibody access to cell-wall phosphocholine, and inhibits acquired antibody binding to sub-capsular protein antigens (Hyams et al., 2010).

To date, more than 90 distinct CPS types have been identified in the pneumococcus (Henrichsen, 1995). Each CPS serotype is encoded by a unique genetic locus, composed of distinct polysaccharide structure, and serological character. Biosynthesis of most CPS follows a Wzy-dependent mechanism. Only serotypes 3 and 37 utilise a synthase-dependent mechanism. Wzy-dependent CPS are composed of repeating units of three or four monosaccharides, the order and the glycosidic linkages between the monosaccharide is different in each serotype (Standish and Morona, 2015). In contrast, synthase-dependent capsules are composed of one or two sugars. At the beginning of colonisation, CPS need to be up-regulated to escape from the mucus. However, thick CPS is disadvantageous for attachment to epithelial cells. Because the pneumococci need to expose adhesins, such as PspC, to be able to adhere to and colonize the nasopharynx. An array of regulators and proteins have been reported to involve in CPS regulation. For instance, RegM (catabolite control protein A, also known as CcpA) has been shown to affect transcription of the *cps* locus, which indicates that different carbon source can affect capsule production (Giammarinaro and Paton,

2002). In addition, Pgm, phosphoglucomutase, converts the glucose-6-phosphate to glucose-1-phosphate, and GalU, uridylyltransferase, which catalyses glucose-1-phosphate to form uridine diphosphate-glucose (UPD-Glc), the precursor for synthesis of pneumococcal capsule, play important roles in capsule synthesis (Cieslewicz et al., 2001, Mollerach et al., 1998, Kadioglu et al., 2008).

Pneumococcal surface protein A (PspA) is a protective antigen found in the cell wall of pneumococci (López et al., 1992) and reported to be present in every strain of the pneumococcus (Crain et al., 1990). The molecular size of PspA varies from 67 to 99 kDa. Its high charge is supposed to block the binding of complement proteins (Jedrzejewski et al., 2001), so that it prevents *S. pneumoniae* from clearance and phagocytosis mediated by complement (Briles et al., 1997). The PspA mutant was shown to be attenuated in virulence, and it had increased C3 deposition in normal human serum, and increased complement-dependent phagocytosis by mouse macrophages (Ren et al., 2012). In addition, PspA is found to be a lactoferrin binding protein, which can protect *S. pneumoniae* against the bactericidal activity from apolactoferrin (Shaper et al., 2004, Baril et al., 2006).

Pneumolysin A (PlyA) is a 53 kDa protein, also known as cholesterol-dependent cytolysin, and expressed during the late log phase of growth (McAllister et al., 2004, Mitchell and Mitchell, 2010). When pneumolysin is released from *S. pneumoniae*, it binds to membrane cholesterol and produces large pores (Tilley et al., 2005). PlyA has ability to block the ciliary motion of the epithelium in the respiratory tract, so that it promotes pneumococcal adhesion to epithelial cells in the host (Steinfors et al., 1989). PlyA also has a role in breaking the barrier of the alveolar capillary bed which enables the pneumococcus to spread to the blood stream resulting in bacteraemia (Rubins et al., 1996). It was reported that PlyA blocks the respiratory burst of polymorphonuclear cells (PMN) (Paton et al., 1983) thus inhibiting the production of hydrogen peroxide (Paton et al., 1983), deletion of *ply* will lead to attenuated virulence of pneumococcus (Schachern et al., 2013).

Pneumococcal surface antigen A (PsaA) is a 37 kDa protein located on the cell membrane of *S. pneumoniae* (Lawrence et al., 1998), belonging to the cluster A-I family of bacterial transporters of the essential metal ions. PsaA acts as a metal binding protein that binds manganese ions, and it is composed of an ATP-binding cassette (ABC) transporter. The growth of a *psaA* mutant need additional  $Mn^{2+}$  (Ogunniyi and Paton, 2015). PsaA is shown to protect *S. pneumoniae* from toxic reactive oxygen species (Tseng et al., 2002), because  $Mn^{2+}$  acts as a co-factor for superoxide dismutase (SodA), which protects the cell from superoxide. The virulence is dramatically reduced in *psaA* mutants in experimental pneumococcal infections (McAllister et al., 2004). PsaA is highly conserved in all the pneumococcal serotypes and its immunogenicity enables it to be a good candidate for the production of vaccines (Baril et al., 2006).

Neuraminidases, also known as sialidase, cleave terminal sialic acid from complex host glycans (King et al., 2006). *S. pneumoniae* possess three types of neuraminidases, NanA, NanB and NanC. All pneumococcal strains have *nanA* and most also has *nanB*, but only half of the strains, mainly involved in meningitis, can encode *nanC* (Pettigrew et al., 2006). Neuraminidase structure varies (Tong et al., 2002), and only NanA contains a C-terminal sequence LPXTG anchorage motif (LPXTGX), which can bind to peptidoglycan of the cell wall (Camara et al., 1994). Pneumococcal neuraminidases can cleave the terminal N-acetylneuraminic acid residues from cell surface glycolipids and glycoproteins, in order to expose the host cell-surface receptors for pneumococcal adhesins (Manco et al., 2006, Camara et al., 1994). NanA is reported to boost pneumococcal colonization and plays a major role in otitis media pathogenesis by degrading the sialic acid in the host surface as demonstrated in experimental otitis media model in chinchilla (Tong et al., 2000). Comparing with NanA, NanB has a much lower acidic pH optimum (Berry et al., 1996), but both NanA and NanB are crucial for pneumococcal survival in the respiratory tract and bloodstream, NanC mainly associated with meningitis (Pettigrew et al., 2006). However, a report has shown that NanA has no effect on *S. pneumoniae* D39

colonisation in the nasopharynx in an infant rat model or attachment in human upper airway epithelium (King et al., 2006). The discrepancy could be caused by the diversity of receptors and glycosylation patterns between cell types.

In addition to the virulence determinants mentioned above, pneumococcus also possess several other LPxTG-anchored proteins, including a variety of glycosidases ( $\beta$ -galactosidase (BgaA),  $\beta$ -*N*-acetylglucosaminidase (StrH) (Burnaugh et al., 2008, King et al., 2006), pullulanase (SpuA) (Abbott et al., 2010), endo- $\alpha$ -*N*-acetylglucosaminidase (Eng) (Marion et al., 2009), and surface-exposed proteases (e.g. PrtA, ZmpABC) (Bethe et al., 2001, Kilian et al., 1979). These proteins either play crucial roles in cleaving and liberating carbohydrate in host niches, or directly involved in host pathogen interactions, which are vital for pneumococcal virulence. Above all, the pneumococcus possess an abundant repertoire of virulence determinates, which contribute to bacterial fitness, nutrition acquirement, host-pathogen interaction, consequently protecting bacteria against host and competing with other microbes. Intensive studies focusing on virulence determinants in the past 30 years have significantly contributed our understanding of pneumococcal pathogenesis (Ogunniyi and Paton, 2015). A systematic assessment of the precise role of each virulence determinant and in depth understanding of how pneumococcus use these virulence factors to liberate carbohydrates from its environment, and to escape from host immune response may provide new clues for developing highly effective vaccines and future treatments for pneumococcal diseases.

## **1.5 Pneumococcal adaptation in host niches**

Investigations on *S. pneumoniae* virulence attributes enhanced our understanding of pathogenesis of pneumococcal infections. However, these studies tended to concentrate on few virulence determinants that have a role in host-pathogen interaction, such as the capsule and cell-wall proteins (Wizemann et al., 2001, Durmort and Brown, 2015). However, it is becoming clear that pneumococcal

acquisition of nutrients and adaptation in varying host ecological niches during infection are also the crucial aspect of pneumococcal virulence and fitness. Although the efficient acquisition and catabolism of host nutrients and adaptation *in vivo* crucial facets of virulence, the knowledge on these pneumococcal traits is relatively limited. *S. pneumoniae* can successfully survive in host and cause an array of diseases in various anatomical sites, which may be restricted in nutrients required for bacterial growth, and contain different oxygen concentration and temperature range (Charpentier et al., 2000). The exact concentrations of specific nutrients will vary with anatomical site, and are also affected by inflammation (Durmort and Brown, 2015). Due to the dynamic host environment, pneumococcus has developed strict control processes that involve sensing of environmental changes and adaptive responses to regulate its gene expression. Some of the *in vivo* environmental factors encountered by the pneumococcus that are relevant for this study have been covered in detail below.

### **15.1 Metal homeostasis**

Metals are very important for bacterial physiology and pathology as they act as cofactors or crucial structural components of various proteins that are important for cellular metabolism and virulence (Honsa et al., 2013, Andreini et al., 2008). However, over accumulation of metals is detrimental for bacterial cells on multiple cellular pathways. During infection, host pumps out necessary metals such as manganese and iron and pumps in toxic metals such as copper and zinc to eliminate the pathogen (Jabado et al., 2000, Forbes and Gros, 2003, White et al., 2009, Botella et al., 2011). The concentrations of metals vary dramatically in the dynamic host niches, for instance, free iron concentration is approximately 138  $\mu\text{M}$  in the blood but is surged to 257  $\mu\text{M}$  during infection. The zinc concentration in blood is 16  $\mu\text{M}$  but is elevated to 641  $\mu\text{M}$  during infection. The concentration of copper in blood is 11  $\mu\text{M}$ , whereas it is 6  $\mu\text{M}$  in nasopharynx (Honsa et al., 2013, McDevitt et al., 2011). Therefore pneumococcus should have the ability to maintain the metal ion

homeostasis to provide sufficient nutrients for cellular metabolism, and at the same time restrict the excess level in the cell to avoid the cellular damages. Manganese, iron and zinc are imported into pneumococcal cells by transporters PsaBCA, PitABCD and AdcCBA, respectively, whereas excessive amount of metal ions is eliminated by specific efflux systems, including the zinc exporter CzcD (Kloosterman et al., 2007), the manganese exporter MntE (Rosch et al., 2009) and the copper exporter P1-type ATPase transporters CopA (Veldhuis et al., 2009). Metal ion haemostasis has been shown to be important as elimination of, CopA, PsaBCA, or PitABCD abrogate pneumococcal virulence in various experimental infection models (Johnston et al., 2004, Veldhuis et al., 2009, Brown et al., 2002).

## **15.2 Resistance to oxidative stress**

The pneumococcus confronts various oxidative stress conditions at different stages of the infection process (Yesilkaya et al., 2013). Pneumococcus is found on top of the nasopharyngeal mucus layer which has around 20% of oxygen, then disseminate to lower respiratory tract where it is exposed lower concentration of oxygen, around 5%. In blood and brain the pneumococcus has to adjust its metabolism to anaerobic setting (Yesilkaya et al., 2013). Reactive oxygen species (ROS) are generated as the by-product of oxygen metabolism, and they include hydrogen peroxide ( $H_2O_2$ ), hydroxyl radical ( $\bullet OH$ ) and superoxide anions ( $\bullet O_2^-$ ) (Saleh et al., 2013). High levels of ROS could cause deleterious effects on proteins, DNA and lipids, consequently lead to metabolic dysfunction and impair cellular structure and even lead to cell lysis (Niki, 2009, Henry et al., 2012). In addition, pneumococci are also able to produce large quantities of  $H_2O_2$  by the activity of pyruvate oxidase, SpxB, as a by-product of aerobic pyruvate metabolism (Pericone et al., 2002). Hydrogen peroxide production can be detrimental to host tissues and contribute to defeat other pneumococcal competitors including *Haemophilus influenzae*, *Neisseria meningitidis* (Margolis, 2009). Therefore, pneumococci need to have oxidative stress resistance mechanisms.

Interestingly, pneumococcus lacks the major peroxide scavenging enzyme catalase. Unlike other streptococci, *S. pneumoniae* does not have an alkyl hydroperoxide reductase C (AhpC) (Lechardeur et al., 2010) nor PerR, which has important roles in the oxidative stress response and *in vivo* fitness (Gryllos et al., 2008, Yesilkaya et al., 2013). Pneumococci possess variety of enzymes implicated in the oxidative stress response, including superoxide dismutase (SOD) (Yesilkaya et al., 2000), NADH oxidase (NoxA) (Auzat et al., 1999), alkyl hydroperoxidase (AhpD) and thiol peroxidase (TpxD) (Hajaj et al., 2012). SOD has the capacity to convert superoxide radicals to H<sub>2</sub>O<sub>2</sub> and O<sub>2</sub>. In the absence of MnSOD (SodA), the pneumococcus shows significant growth defect under aerobic conditions (Yesilkaya et al., 2000). Alkyl hydroperoxidase (Ahp) catalyzes the conversion of toxic peroxide compounds to alcohol and water, and consequently contributes to oxidative stress resistance (Mishra and Imlay, 2012). Deletion of *ahpD* renders pneumococcus less virulent in both pneumonia and bacteremia mouse models (Paterson et al., 2006). NADH oxidase (Nox) can detoxify oxidative stress by converting O<sub>2</sub> to H<sub>2</sub>O (Auzat et al., 1999). Studies have shown that inactivation of *nox* leads to attenuated virulence of *S. pneumoniae* (Yu et al., 2001). In addition, the thiol peroxidase (TpxD) scavenges H<sub>2</sub>O<sub>2</sub> (Hajaj et al, 2012; Yesilkaya et al, 2013). The mutation of *tpxD* affects pneumococcal growth and survival under aerobic conditions *in vitro*, and renders pneumococcus less virulent in an *in vivo* murine model (Hajaj et al, 2012). Hence, the ability to sense and response to oxidative stress is important for pneumococcal *in vivo* fitness. However, our understanding of how the pneumococcus modulates its response to the agents of ROS is limited. One of the aims of this study is to identify transcriptional regulators involved in pneumococcal oxidative stress resistance.

### **15.3 Carbohydrate acquisition and metabolism**

*Streptococcus pneumoniae* relies exclusively on its host for its nutrients. The microbe is a strictly fermentative bacterium that relies on glycolytic metabolism to obtain

energy for survival, proliferation and causing infection in various sites within the host (Hoskins et al., 2001). Pneumococcus has been reported to have 21 phosphotransferase systems (PTS) and 8 ATP-binding cassette (ABC) transporters that import at least 32 distinct carbohydrates (Bidossi et al., 2012). Adaptation to host carbon source is closely linked to pneumococcal *in vivo* fitness. The PTS and ABC transporters are important for pneumococcal colonization, transmission, and virulence, since inactivation of transporters renders pneumococci less virulent *in vivo* (Bidossi et al., 2012). Despite the capacity to ferment large number of sugars *in vitro*, most of these sugars are not found *in vivo*. The concentration and composition of carbon source varies in different host niches, for instance, the concentration of glucose in human respiratory airway is less than 1 mM, whereas it is higher in blood, around 4-6 mM (Shelburne et al., 2008, Philips et al., 2003). Galactose is abundant in the airway glycoconjugates (Terra et al., 2010b) mannose is known to be rich in N linked glycans found in respiratory tract and blood (Sharma et al., 2014). As the concentration of free carbohydrates is limited in respiratory tract, such as glucose, the host glycoproteins including O- and N-linked glycans, and glycosaminoglycans are reported to provide carbon sources for pneumococcal growth (Marion et al., 2012, Burnaugh et al., 2008). *S. pneumoniae* possesses various glycosidases including BgaC ( $\beta$ -galactosidase C), Eng (endo- $\alpha$ -N-acetylgalactosaminidase), NanA (Neuraminidase A), StrH ( $\beta$ -N-acetylglucosaminidase), BgaA( $\beta$ -galactosidase A), NanB (Neuraminidase B), Hyl (hyaluronate lyase), which could deglycosylate glycans and liberate free sugars for growth and expose the host receptors for pneumococcal attachment (Terra et al., 2010a, Jeong et al., 2009, Marion et al., 2009). It has been shown that pneumococcal glycosidases work in concert to liberate sugars found in complex host glycans, such as sialic acid mainly by NanA and NanB, galactose by BgaA and BgaC, and N-acetyl glucosamine by StrH (Paixão et al., 2015). It has been shown by our research group that the pneumococcus is able to utilise these liberated sugars. The pneumococcus is shown to have the required catabolic machinery for utilisation of galactose, mannose, sialic acid and N-acetyl glucosamine. However, among these catabolic pathways, only

elimination of galactose catabolic pathways, Leloir and tagatose, have shown to attenuate pneumococcal colonisation and virulence in experimental infection model, indicating the importance of galactose catabolism for pneumococcal *in vivo* survival (Paixão et al., 2015). Currently, it is not known why this happens, and how the pneumococcus modulates its galactose catabolism.

In addition to their nutritional value, some of the sugars liberated by pneumococcus can act as signal molecule, such as sialic acid and galactose. It has been shown that free sialic acid acts as a signal that promotes *S. pneumoniae* invasion of nasal tissue and nonhematogenous invasion of the central nervous system (Hatcher et al, 2016). Trappetti et al., (2009) showed that sialic acid enhances pneumococcal biofilm formation, colonisation and virulence. A recent study by Blanchette et al., (2016) also reported that galactose promotes biofilm formation during colonization.

As summarised above, *S. pneumoniae* is exposed to dynamic and fluctuating conditions in host tissues. This suggests the presence of efficient regulatory mechanisms to tightly control the expression of genes for efficient survival, replication and virulence. As the main focus of this study is to understand the pneumococcal transcriptional regulation, next I would like to introduce transcriptional regulation mechanisms in bacteria with a particular focus on *S. pneumoniae*.

## **1.6 Bacterial transcriptional regulation**

Prokaryotes can use their genetic material with great effectiveness to regulate the production of proteins in the appropriate amounts at the required time in response to fluctuating environmental conditions. This regulation typically includes genome-wide alternations in the level of transcription, as well as posttranscriptional and posttranslational changes (Dmitriev et al., 2006). In contrast with post-transcriptional regulations, which mainly depend on small-RNAs, transcriptional regulation is efficient and rapid for bacteria to change gene expression in response to different

growth conditions (López-Maury et al., 2008). The most important element in bacterial transcriptional regulation is the multi-subunit DNA-dependent RNA polymerase (RNAP) (Browning and Busby, 2004). The active site of the polymerase is composed of the large  $\beta$  and  $\beta'$  subunits, which has the components for the binding of both template DNA and the RNA product during transcription. The two  $\alpha$  subunits contain two separate folded domains that are joined by a linker. The larger amino-terminal domain ( $\alpha$ NTD) has effect on the assembly of the  $\beta$  and  $\beta'$  subunits and the smaller carboxy-terminal domain ( $\alpha$ CTD) is a DNA-binding module which is crucial for certain promoters. For initiation of transcription, RNAP should be recruited to a specific locus in the upstream region of the gene, referred to as promoter. The  $\sigma$  subunit is required for the formation of holoenzyme when RNA polymerase interacts with a specific promoter. The  $\sigma$ -factors contain four different domains joined by linkers, which can recognize specific promoter sequences, locate the RNA polymerase holoenzyme at a target promoter, and also have a role in unwinding of the DNA duplex close to the transcript start site (Gross et al., 1998, Browning and Busby, 2004).

In Prokaryotes, the promoter contains two principal elements which are known as the -10 hexamer and the -35 hexamer, which are located 10 and 35 nucleotides upstream from the transcript start site. The -10 region can be specifically recognized by  $\sigma_2$  subunit and -35 region can be specifically recognized by  $\sigma_4$  subunit. The other important element for docking the RNAP within the promoter is the UP element, which is located upstream of the -35 hexamer. The UP element can be recognized by the C-terminal domain of the RNAP  $\alpha$  subunits. These promoter elements have separate roles in docking the RNAP, and initiation of transcription. Deficiencies in one element can be compensated by another (Browning and Busby, 2004). For transcription initiation, RNAP binds to the promoter DNA and forms an open complex. The duplex DNA is unwound in the open complex close to the transcript start point (Zupancic and Record, 1998).

### **1.6.1 Transcription factors (TFs)**

TFs are a number of proteins that regulate the expression of the target gene. TFs have DNA-binding domains. Some TFs bind to a DNA promoter region near the transcription start site, and help form the transcription initiation complex. Other TFs bind to regulatory sequences, such as enhancer sequence, and can either activate or repress transcription of wide repertoire of downstream target genes. TFs with capacity to regulate genes belonging to diverse functional categories with potential to respond to a wide range of environmental conditions are called global TFs, and others targeting specific gene(s) and/or operon(s) are called local TFs.

Transcription factors can be grouped in two categories: activators and repressors. While activators can enhance the interaction between the promoter and RNA-polymerase, repressors bind to the region close to promoter and stop RNA-polymerase progress along the DNA strand, resulting in repression. Moreover, in certain cases, a repressor also may suffer from allosteric competition against a determined activator to repress gene expression. The overlapping DNA-binding motifs could be recognized by both activators and repressors, which will induce a physical competition to occupy the site of binding. If the affinity of repressor is higher for its motif than the activator, transcription would be effectively blocked in the presence of the repressor. In the bacterial cell, RNA polymerase is faced with an array of nearly 2,000 promoter sequences. Since the RNAP is limited, tight regulatory control is achieved by TFs, ensuring that RNAP is correctly distributed among the competing promoters (Browning and Busby, 2004).

Generally bacterial transcriptional regulation systems relayings on extracellular signals can be classified broadly as two-component signal transduction systems (TCR) and Quorum sensing systems as discussed in detail below (Chang et al., 2011).

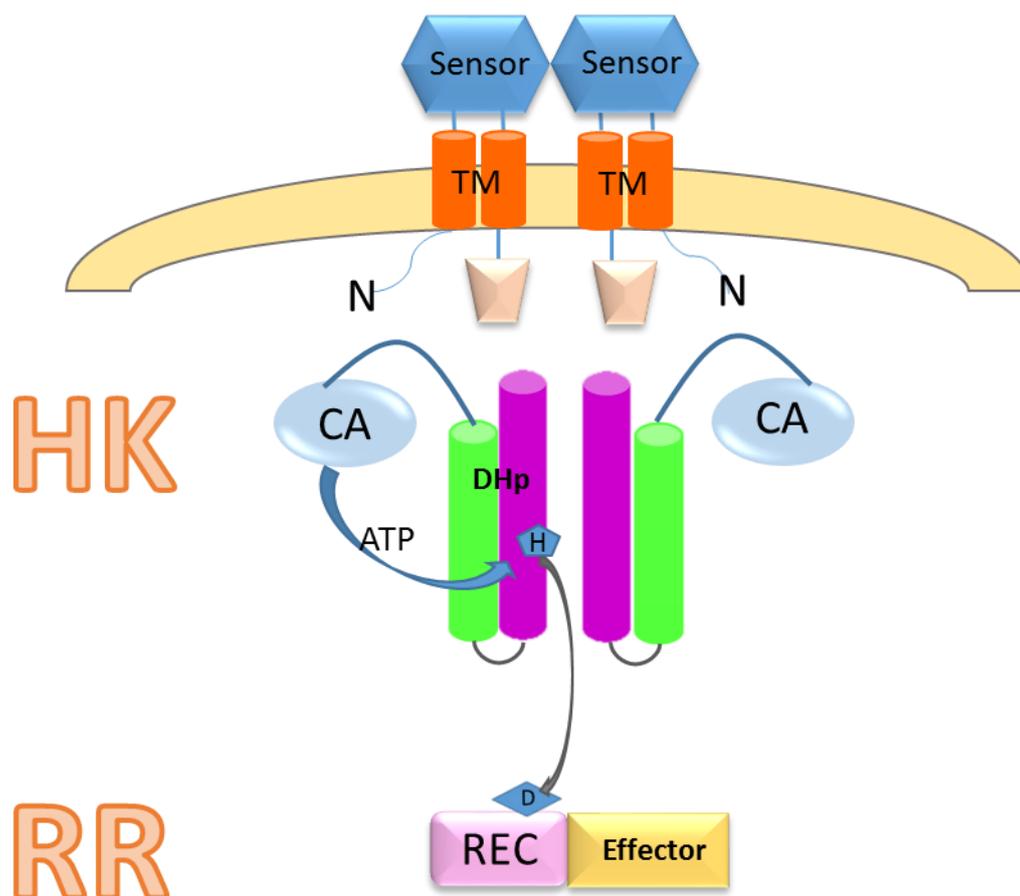
## **1.6.2 Two-Component Signal Transduction Systems**

Two-Component Signal Transduction Systems (TCSs) act as a stimuli-response system, function in detecting environmental signals, transducing signals to effectors and coordinating gene expression. These systems are able to respond to an extremely broad spectrum of diverse stimuli, and control a vast array of cellular function, such as basic metabolism, virulence, cell division and differentiation. TCSs are widespread in a variety of microorganisms, also present in some plants, but not in mammals (Tatsuno et al., 2014). Basically, TCSs consist of a sensor histidine kinase and a response regulator (RR) (Tatsuno et al., 2014). The sensor histidine kinase (HK) senses stimulus or signaling molecules and autophosphorylates at a histidine residue, transfers the phosphoryl group to an aspartate residue in the response regulator protein (RR), which activates effector domain in RR. RRs are mostly DNA-binding transcription factors whose affinities for their target promoters are modulated by phosphorylation (Szurmant, 2012). Consequently, alternation of the phosphorylated state of RRs results in gene expression profiles (Gao and Stock, 2009).

HKs are roughly divided in to two regions, an N-terminal sensing and a C-terminal catalytic kinase core. The N-terminal domain is composed of dimeric receptors framed by two transmembrane (TM) helices (Casino et al., 2010). The C-terminal kinase core is localised in the cytoplasm. The kinase catalytic core of the HKs is composed of two well-differentiated domains: a dimerization of a histidine phosphotransfer domain (DHP) and a catalytic ATP binding domain (CA). The DHP domain allocates the phosphorylatable His residue and mediates HK dimerisation. The CA domain binds the ATP molecule for the autophosphorylation reaction (Casino et al., 2012). RRs contain an N-terminal receiver domain (REC) and a C-terminal effector domain. The REC is well-conserved, generally it consists of five-stranded parallel  $\beta$ -sheet surrounded by five amphipathic  $\alpha$ -helices. It comprises the phosphorylatable Asp residues, which are involved in the phosphotransfer reaction and

in the propagation of long-range conformational changes accompanying receiver phosphorylation (Beier, 2012). However, the effector domains are diverse and mediate specific response to environmental stimuli, such as oxidative stress, and osmotic stress (Stock et al., 2000) (Figure 1.4).

In order to cope with a variety of environmental condition, some TCSs were found exist as a network, consisting of cross-talk and the cascade of signal transductions between different TCSs (Utsumi, 2008). Cross-talk is a signaling system in which distinct TCSs are connected through phosphotransfer from the HK to a non- cognate RR (Laub and Goulian, 2007). On the other hand, in cross-regulation, heterodimers or connector proteins and peptides are involved in signal transduction between TCSs.



**Figure 1.4. The mechanisms of two component regulatory system.** The system is formed by a dimeric histidine kinase (HK) and a response regulator (RR). The HK contains an extracellular sensor domain connected by cytoplasmic domains containing two transmembrane helices (TM). The cytoplasmic domains are composed of a DHp domain containing the phosphorylatable His residue in the H box and a CA domain. The His (H) phosphorylation is mediated by ATP followed by subsequent phosphoryl transfer to the phosphorylatable Asp (D) in the N-terminal receiver domain (REC) of the RR. The effector in C-terminal is in charge of initiating the final response. Figure is constructed based on (Casino et al., 2010).

### 1.6.3 Quorum sensing systems in bacteria

Unlike the TCSs, quorum sensing pathways depend on bacterial population density and pheromones (Jimenez and Federle, 2014). Bacterial adaptation to environmental stimuli is a complex process. Bacteria coordinate their behavior via intercellular chemical signaling, also known as pheromone, produced by community members in a

cell density dependent manner. This process is known as quorum sensing (QS) (Cook and Federle, 2014), which was shown to be important for a variety of functions, for instance, biofilm formation and dispersal (Chang et al., 2011), regulation of virulence factors (Zhu and Mekalanos, 2003), competence development for bacterial natural transformation (Fontaine et al., 2010) and sporulation (Steiner et al., 2012).

The core of bacterial QS communication systems are the intercellular chemical signals (Cook and Federle, 2014). When there are a few cell present, the production of the signal molecule is low, so signal molecule levels remain low. However, as the bacterial population increases in a confined environment, so does the concentration of the signal molecules. In this way, a bacterium can conduct a chemical census of the surrounding population. Bacterial cells sense when signal molecule levels exceed a certain threshold concentration via a specific sensor (or receptor), and therefore can adjust gene expression accordingly.

The most intensively investigated intercellular signal molecules are the N-acylhomoserine lactones (AHLs) produced by Gram-negative bacteria. AHLs are synthesized by LuxI synthases from S-adenosyl-methionine (SAM) and an acyl chain carried by an acyl carrier protein (Grandclément et al., 2015, Pereira et al., 2013). Gram negative bacteria are capable of producing a variety of different AHL molecules. All of the natural AHLs consist of a HSL ring covalently linked to an acyl chain via an amide bond. Structural variation of these molecules is due to differences in the length and degree of saturation of the N-acyl chain (4-18 carbons, presence or absence of one or more double bonds) and oxidation state at the 3-position of the N-acyl side-chains. Delicate differences in the structures of AHL molecules can dramatically change their ability to bind to LuxR family proteins and activate target gene expression.

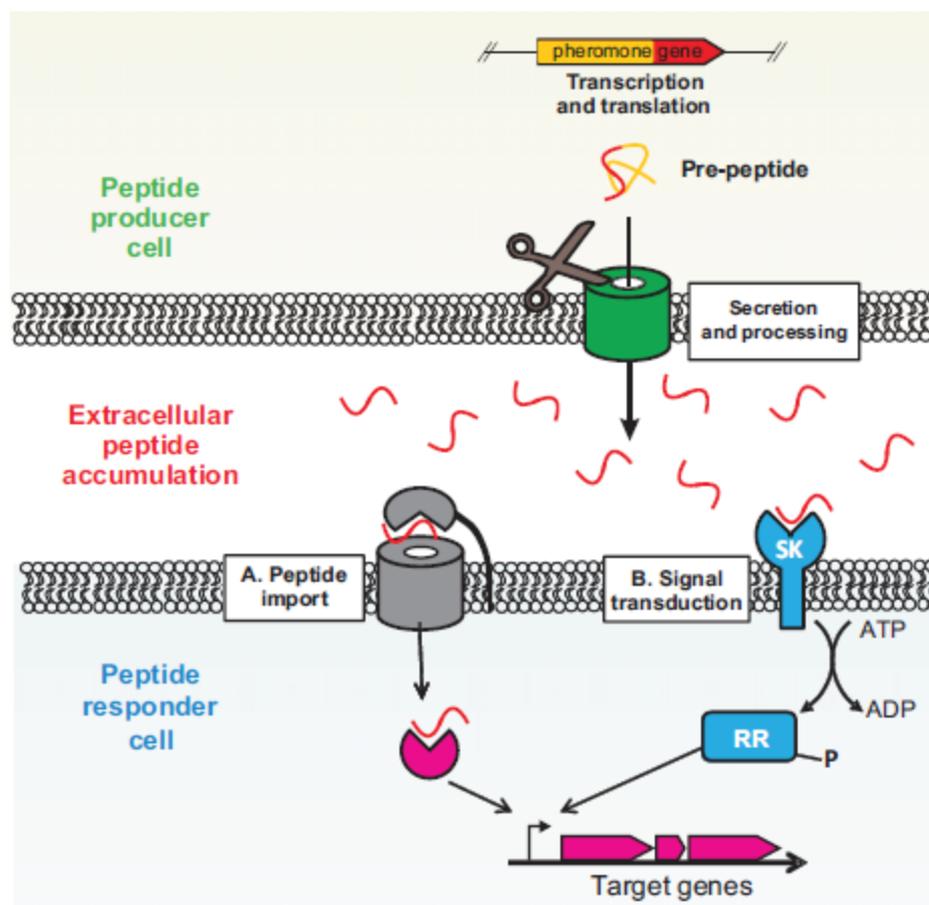
AHLs are not only involved in regulation of bioluminescence in a few marine species of bacteria, but also they control the production of  $\beta$ -lactam antibiotic and multiple

extracellular enzymes associated with plant tissue destruction in *Erwinia. carotovora* (Holden et al., 2007). AHLs also can regulate virulence gene expression including genes coding for extracellular enzymes, toxins and motility, including in *Pseudomonas aeruginosa*, *Burkholderia spp.* and *Chromobacterium violaceum* (Holden et al., 2007). The first identified AHL-mediated regulatory system is related to bioluminescence in *Vibrio. fischeri*. Generally, when *V. fischeri* population live freely in the ocean, they are dark as the population density is low, with less than  $10^2$  cells/ml. *V. fischeri* utilizes a small diffusible AHL molecule, *N*-(3-oxohexanoyl)-L-homoserine lactone (3-oxo-C6-HSL) which is generated by *LuxI*, to gauge the cell density (Milton, 2006). LuxR is a response protein that can sense and detect AHL, and acts as the transcriptional activator of the *lux* bioluminescence gene cluster. LuxR contains an N-terminal AHL-binding domain and a C-terminal DNA-binding domain with a helix-turn-helix motif. When the bacterial cell density increases, the concentration of AHL surpasses a threshold concentration and activates LuxR, resulting in the induction of the *lux* operon. This results in the production of more LuxR and AHL, creating a positive feedback loop, initiating the production of light. However, not all of the Gram-negative bacteria produce AHLs, some produce more than one class of quorum sensing signal molecule. For example, *Pseudomonas aeruginosa* is able to produce AHL, 2-alkyl-4-quinolone (AQ) and cyclic dipeptide signal molecules. *E. coli* and *Salmonella* cannot produce AHL, instead they utilize AI-2 (a group of inter convertible furanones derived from dihydroxypentanedione) and AI-3 as quorum sensing signal molecules (Holden et al., 2007).

The main processes of QS systems in both Gram positive and negative bacteria are similar. However, certain differences exist. In contrast with Gram-negative bacteria, Gram-positive bacteria do not produce *N*-acyl-homoserine lactone (AHLs). Gram-positives use small peptides as signal molecule to mediate QS (Fleuchot et al., 2013). Signal peptides either can be sensed outside the cell by a sensor kinase (SK) at the surface of the cells, or interacts directly by its response regulator after internalization (Fleuchot et al., 2013; Cook & Federle, 2014). QS in Gram positives is initiated by

transcription and translation of peptide precursor, which are processed post-translationally via pheromone-specific peptidases (Eep), followed by secretion of peptides into extracellular environment. Unlike most AHLs, which can freely diffuse across cell membranes, peptides usually require specialized transport mechanisms such as ATP binding cassette transporters, also known as ABC transporters. Mature signal molecules gradually accumulate in the extracellular milieu, and once the concentration reaches to the threshold level, the matured peptide is re-imported into cells by oligopeptide permease transporters (opp) or Ami peptide uptake systems. Proteolytical cleavage occurs to release the active peptide, and then peptide will interact with its cognate cytoplasmic regulatory protein (Fleuchot et al., 2013).

In addition, the peptides also can interact with a SK of two component system, this process leads to the autophosphorylations of the SK, and the signal is transmitted intracellularly by phosphorylation of its response regulator (RR) (Cook & Federle, 2014) (Figure 1.5). The activation of the response regulator by signal peptide leads to conformational changes in its C-terminal domain and alters DNA binding capacity of response regulator, modulating its ability to activate or repress the transcription of its target gene(s).



**Figure 1.5. Peptide signaling in Gram-positive bacteria (from Cook & Federle, 2014).** Production of Gram-positive peptide pheromones involves transcription and translation of precursors followed by processing and secretion. Once in the extracellular environment, peptides are often further processed before interacting with surrounding cells. To exert effects on neighboring cells, peptides are either directly imported into the cell where they interact with their cognate receptor (A) or interact with a surface exposed sensor kinase (SK) (B). Following peptide interaction with an SK, a signal is transmitted intracellularly in the form of phosphorylation of a response regulator (RR). The phosphorylated RR or peptide/receptor combination can then alter gene expression by either directly binding DNA or interacting with transcriptional regulators such as sigma factors and RNA polymerase.

There are four classes of QS systems in Gram positive bacteria, which are distinguished by their characteristics of the autoinducers and their receptors, including RNPP (Rap, NprR, PlcR, and PrgX) family of regulators, Agr-type cyclical pheromones, peptides with double-glycine (Gly-Gly) processing motifs, and Rgg

family regulators (Cook and Federle, 2014). The best-studied regulatory protein family, RNPP includes aspartyl phosphate phosphatases (Rap proteins), the neutral protease regulator (NprR), the phospholipase C regulator (PlcR) and sex pheromone receptor (PrgX). Rap has ability to bind to response regulators and inactivate it, which represses the initiation of sporulation and competence development in *Bacillus subtilis* (Rocha-Estrada et al., 2010). The phospholipase C (PlcR) regulator is a transcriptional activator controlling the expressions of many virulence-related genes, especially in *Bacillus cereus* (Rocha-Estrada et al., 2010). PlcR exerts its activity by binding to the cognate peptide PapR, which alters its helix-turn-helix (HTH) positioning to bind specific target promoters to initiate transcription of the PlcR virulence regulon (Declerck et al., 2007). On the other hand, PrgX can either interact with the chromosomally encoded activator peptide (cCF10) or the plasmid-encoded inhibitor peptide (iCF10) (LaSarre et al., 2013), regulating the expression of the conjugative transfer genes of the *Enterococcus faecalis* plasmid pCF10 via conformational changes that reduce the oligomerization state (Shi et al., 2005). In addition, NprR interacts with the NprX peptide, promoting the NprR HTH domain binding to DNA, thus activating transcription (Zouhir et al., 2013, Rocha-Estrada et al., 2010).

#### **1.6.4 Transcriptional regulators can be targeted to develop novel antiinfectives**

Research on QS in molecular level is beneficial to elucidate the mechanisms of intra- and intercellular communication and phenotypic adaptation of pathogens. In addition to understanding of microbial pathogenesis, this knowledge can be used to develop novel antiinfectives. Antibiotic resistance is one of the greatest challenges of the twenty-first century. The problems associated with multiple-antibiotic resistance have been compounded over the last few decades by the failure of developing new antibacterial agents with truly novel modes of action (Holden et al., 2007). Traditional

approaches to combat bacterial infection were primarily identified in and derived from other microorganisms. These approaches rely on the disruption of the growth cycle by preventing the synthesis and assembly of key components of bacterial processes such as cell wall synthesis, DNA replication and protein synthesis. Although these strategies and compounds were highly effective, they resulted in substantial stress on the target bacterium, which rapidly selected for resistant subpopulations. The increased understanding of bacterial pathogenesis and cellular communication systems has revealed many potential strategies to develop novel drugs to treat bacteria-mediated diseases. Quorum quenching (QQ) emerged as a new term referring to all processes involved in the disturbance of QS (Dong et al., 2001). QQ molecular actors such as enzymes and chemical compounds could affect the QS pathways including QS-signal cleavage and competitive inhibition (Byers et al., 2002, Yates et al., 2002, Delalande et al., 2005).

Usually, the enzymes that inactivate QS signals are named QQ enzymes. QQ enzymes are reported to clear QS signal, and involved in recycling of signals and detoxication. For instance, AiiA, the prototypic AHL lactonase, has been identified in several strains of members of the *Bacillus* genus (Dong et al., 2000). AiiA hydrolyzes a large number of substrates with acyl chain length ranging from 4 to 12 carbon atoms, with or without oxo substitution at carbon 3 (Dong et al., 2000). On the other hand, the chemicals disrupting QS pathways are called QS inhibitors (QSIs). QSIs may interfere with the synthesis of the autoinducers, the cell-to-cell exchange of autoinducers, and the perception and transduction of autoinducer signal through its interaction with sensor/transcriptional factors (Grandclément et al., 2015). Although to date, only a few clinical experiments have been conducted, with promising results such as those using azithromycin as a QSI in the treatment of ventilator associated pneumonia (Van Delden et al., 2012), it is clear that characterization of QS systems in bacteria can help device new antiinfectives in the future.

Transcriptional regulators have been considered as targets for anti-virulence drug discovery strategies due to their role in bacterial virulence. Anti-virulence signaling strategies may specifically interfere with the ability of the bacteria to recognize host signals that alert the bacteria that they are at the locus of infection and activate specific virulence traits that are needed to establish infection. By preventing the expression or activity of virulence traits, the bacteria are less able to colonize the host. In addition, as this strategy does not directly kill the bacteria, there is presumably less evolutionary pressure for the development of resistant clones than with traditional antibiotics (Rasko and Sperandio, 2010). Moreover, these anti-virulence drugs could potentially be used in combination with established or novel antimicrobials in a synergistic manner to extend the lifespan of these drugs. Some of cell-to-cell signaling pathways are conserved in many bacterial species, and targeting of these systems for therapeutics may yield broad-spectrum drugs. For instance, the membrane-bound QseC histidine sensor kinase QseC homologues are present in at least 25 important human and plant pathogens, and *qseC* mutants of enterohaemorrhagic *E. coli* (EHEC), *S. typhimurium* and *F. tularensis* have been shown to be attenuated in infection (Rasko and Sperandio, 2010). Histidine sensor kinases are ideal therapeutic targets as they are absent in mammalian cells. High throughput screening (HTS) of a library of 150,000 small organic compounds led to the identification of a lead structure, LED209 (*N*-phenyl-4-[[[(phenylamino) thioxomethyl] amino] - benzenesulphonamide), which blocks the binding of signaling molecules (AI-3, adrenaline and noradrenaline) to QseC. This prevents the autophosphorylation of QseC and consequently inhibits QseC-mediated activation of the expression of virulence genes. LED209 decreased virulence of EHEC in *in vitro* systems, whereas *in vivo* studies did not show a statistically significant decrease in infection. It is thought this failure of LED209 in the EHEC animal model is due to the rapid systemic absorption of the compound into the peripheral circulatory system. Despite the drawbacks, bacterial adaptation mechanisms are viable anti-infective

targets. Therefore, it is important to study pneumococcal adaptation mechanisms to combat drug resistant and tolerant pneumococci.

## **1.7 Pneumococcal transcriptional regulation**

The factors enabling *S. pneumoniae* to sense and respond to the environmental changes, to compete with other bacteria for nutrients, and to protect from the bactericidal activity of the immune system are highly regulated (Kadioglu et al., 2008). However, studies on how the pneumococcus deals with the changing environmental factors are relatively limited. Although a few regulatory systems from *S. pneumoniae* have been studied, the function and mode of action of most transcriptional regulators are still unknown.

### **1.7.1 Pneumococcal standalone regulators**

In Gram negative bacteria, sigma factors are used to coordinate gene expression, for instance, *E. coli* use an alternative sigma factor RpoS to regulate transcription in stationary phase (Dmitriev et al., 2006). In contrast, *S. pneumoniae* genome does not have typical stationary-phase sigma factors, instead, only an alternate sigma factor, SigX, has been detected, which possesses limited functions, such as ClpP-dependent degradation (Luo and Morrison, 2003). The term ‘stand-alone’ regulator was originally used to describe transcriptional regulators and their paralogs that control global virulence regulons in response to changing environmental conditions, yet where the exact signals and cognate sensory elements involved remain undefined (McIver, 2009).

The examples of stand-alone regulators in *S. pneumoniae* include Mga-like repressor A (MgrA), Catabolite Control Protein A (CcpA) and CodY (Carvalho et al., 2011, Giammarinaro and Paton, 2002). CcpA is a member of the LacI family of transcriptional repressors/activators, and is the main transcriptional regulator

functioning at the core of carbon catabolite control (Iyer et al., 2005). CcpA binds to catabolite-responsive elements (*cre*), and repress  $\beta$ -galactosidase and  $\beta$ -glucosidase. Furthermore, CcpA also has been implicated in the regulation of many virulence factors (Carvalho et al., 2011). MgrA coordinate the expression of numerous virulence-associated genes, which encode M family proteins, C5a peptidase, and a secreted inhibitor of complement, which is closely related to pneumococcal pathogenicity. The MgrA deficient mutant was shown to have a decreased ability for pneumococcal nasopharyngeal carriage and lung infection (Hemsley et al., 2003). Pneumococcal CodY is a transcriptional repressor, which mediates nutritional regulation and virulence. CodY effects the expression of many genes that are involved in amino acid metabolism, biosynthesis, and uptake. Moreover, previous work has demonstrated that CodY is necessary for bacterial adherence and colonization of the nasopharynx in murine model (Hendriksen et al., 2008). GlnR is a transcriptional regulator involved in the regulation of glutamine and glutamate metabolism, the absence of pneumococcal GlnR leads to attenuation in virulence and colonisation in a murine model (Al-Bayati et al., 2017). Although lots of studies work on pneumococcal gene regulation, there are still many more yet to be functionally characterised.

### **1.7.2 Pneumococcal two component systems**

In addition to the stand-alone regulators, the pneumococcal genome contains 13 putative complete TCSs plus one orphan response regulator. Pneumococcal TCSs could mediate various cellular responses including stress response, bacteriocin (pneumocin) production, competence and pathogenicity (Paterson et al., 2006). Ten of the pneumococcal TCSs have been shown to be important for pneumococcal virulence (Throup et al., 2000, Hava and Camilli, 2002, Lau et al., 2001), only TSC03, TCS10 and TCS11 have not been found associated with virulence. For instance, CiaR/CiaH (TCS05), the first TCS identified in pneumococcus, has been

demonstrated to contribute to virulence, probably in part through control of expression of *htrA*, which encodes a serine protease that is a major virulence factor (Guenzi et al., 1994). Recently, TCS06 has been found to regulate the expression of *cbpA*, which codes for a major adhesin that is also a protective antigen. TCS09 also has been reported to involve in pneumococcal virulence, a *rr09* mutant in strain D39 was essentially avirulent in mouse models of pneumonia and bacteremia (Blue and Mitchell, 2003). In contrast to wild-type-infected mice, all the mice infected with *rr09*-mutant survived during infection, with bacteria being rapidly cleared. RR09 (and presumably TCS09) therefore has the potential to contribute significantly to pneumococcal virulence. TCS04 was also identified as a pneumococcal virulence factor, as the TIGR4 *rr04* mutant was attenuated relative to its parental wild-type during murine pneumonia. However, considerable variation appears in TCS04 role among different pneumococcal strains, reflecting the diversity in natural pneumococcal populations (McCluskey et al., 2004). A TCS13 mutant was found to dramatically attenuated in respiratory tract infection model with bacterial lung counts reduced by about  $10^4$ -fold compared with wild-type strain 0100993. This system was found to control a 16-gene regulon, regulating the synthesis and export of bacteriocin-like peptides (BlpC) and immunity proteins by microarray analyzes (De Saizieu et al., 2000). Furthermore, RR489/RitR is an orphan response regulator. Unlike the other pneumococcal TCSs, *rr489* is not located in the genome next to a cognate HK. The *rr489* has significant impact on pneumococcal virulence, as *rr489* deficient mutant displayed a more than  $10^4$  fold reduction in pulmonary bacterial counts compared with its wild-type parent in a murine pneumonia model (Throup et al., 2000). RitR is a regulator of iron transport, contributes to the Piu iron uptake system (Ulijasz et al., 2004). Despite being essential for the growth of most bacteria, iron also can be detrimental via the Fenton reaction, where it catalyses the synthesis of reactive oxygen intermediates from hydrogen peroxide. Thus RitR is important for regulation of iron level in pneumococcus. Absence of *ritR* leads to increased iron uptake, hence a *ritR* mutant is more vulnerable to iron-dependent killing by streptonigrin (Paterson et

al., 2006). To date, how RitR works in the absence of a cognate HK remains unclear.

TCS12 has been demonstrated to involve in pneumococcal competence and virulence (Paterson et al., 2006). *Pneumococcus* is naturally competent and the competence is activated through TCS12. TCS12 consists of the HK encoded by *comD* and the RR encoded by *comE*, which together respond to competence stimulating peptide, CSP, which is derived from a ribosomally synthesized precursor (ComC). The CSP secreted and processed by the ComA/B ABC transporter to the extracellular environment (Paterson, et al., 2006). When the concentration of CSP reach to a certain threshold, the ComD HK is activated. The activated HK transfer the phosphoryl to its cognate RR. The phosphorylated RR could mediate the expression of *comC/D/E* and *comA/B* operons, and activates the transcription of *comX*, which encodes an alternative sigma factor ComX, activating genes involved in competence. Phosphorylated ComE could mediate the expression of *comW*, which contribute to ComX stabilization (Luo and Morrison, 2003, Sung and Morrison, 2005). In addition, previous research has demonstrated a *comD* (*hk12*) mutant in D39 was attenuated in models of both pneumonia and bacteremia (Bartilson et al., 2001). On the other hand, *comD* was identified as a virulence factor in pneumonia model (Hava and Camilli, 2002).

TCS13 is involved in pneumococcal bacteriocin (pneumocin) production, which is controlled by the *blp* locus. The *blp* locus is stimulated by the accumulation of peptide BlpC. Pre-BlpC is processed at a double glycine motif and secreted out of the cell via its cognate transporter BlpAB. BlpC binds to and activates the histidine kinase receptor, BlpH, resulting in phosphorylation of the response regulator, BlpR (De Saizieu et al., 2000). Phosphorylated BlpR upregulates the 4-6 operons in the *blp* locus including a variety of genes found in the Bacteriocin-Immunity Region (BIR) of the locus, in turn regulate the bacteriocin production (Wholey et al., 2016).

Additionally, TCS02 is essential for pneumococcal viability (Lange et al., 1999, Throup et al., 2000). A recent study showed that TCS02 contributes to the regulation

of cell wall and fatty acid biosynthesis as well as expression of the virulence factor *pspA*, codes for a surface protein (Mohedano et al., 2005, Ng et al., 2003, Ng et al., 2004). TCS10 has been implicated in tolerance to the antibiotic vancomycin (Haas et al., 2005). The remaining TCSs are relatively poorly characterized although most have been shown to contribute to virulence.

### **1.7.3 Pneumococcal Quorum sensing systems**

*Pneumococcus* also adapts its metabolism to different environmental niches by using its regulatory mechanisms, including QS systems. Similar to QS systems in other Gram positive bacteria, QS systems use the oligopeptide as chemical signal to sense the environmental cues, and interact with cytoplasmic receptor to coordinate gene expression. The competence regulon is the best studied QS system as discussed above. In addition, recently TprA/PhrA regulatory system which is similar to PlcR/PapR in *Bacillus* has been studied by Hoover et al. (2015). The TprA/PhrA system in *S. pneumoniae* D39 was reported to regulate a lantibiotic biosynthesis gene cluster, which is linked to produce antimicrobial peptides that may be important for competitive fitness during nasopharynx colonization (Hoover et al., 2015). The PhrA-signaling peptide is derived from a precursor protein that is predicted to be exported through the Sec pathway and processed outside the cell by proteases to release the mature peptide. Once the PhrA peptide concentration has reached a threshold in the extracellular environment, it interacts with an oligopeptide permease (Opp) whereby it is brought into the cytoplasm and antagonize the inhibitory activity of TprA to induce expression of the peptide-encoding gene, *tprA*, and the lantibiotic biosynthesis gene cluster (Hoover et al., 2015).

*S. pneumoniae* colonize the nasopharynx, where complex, polymicrobial communities organized as biofilms. Thus, colonization involves interspecies competition, which drives temporal changes in the bacterial composition of the nasopharyngeal microbiome. LuxS/autoinducer-2 is a Quorum sensing pathway involved in

pneumococcal biofilm formation. LuxS (S-ribosylhomocysteinase) is a protein required for the biosynthesis of a type 2 autoinducer (AI-2) by converting S-ribosylhomocysteine to homocysteine and 4,5-dihydroxy-2,3-pentanedione (DPD). DPD then spontaneously cyclises to active AI-2 that is involved in quorum sensing systems throughout various bacterial genera (Stroeber et al., 2003). The accumulation of secreted AI-2 in the external milieu stimulates planktonic bacteria to initiate early formation of the biofilm structure (Vidal et al., 2011). LuxS/AI-2 was found to be responsible for the regulation of *S. pneumoniae* biofilms on both abiotic surfaces and human respiratory epithelial cells (HREC) (Galante et al., 2015). Although its involvement in the biofilm development and virulence in several bacterial species is well studied, the regulation and mechanism of LuxS conferred phenotypes are not clear to date (Kyd et al., 2016).

While characterization of pneumococcal regulatory networks has been studied in some detail, there is more research need to be done with respect to their contribution in pneumococcal adaptation in different tissue sites with varying environmental conditions. One of the transcriptional regulatory classes that have not been studied in detail in *S. pneumoniae* is Rgg type regulators. Hence, the objective of this study is to identify a role for Rgg type regulators in pneumococcal biology.

## **1.8 Rgg regulators**

Rgg, 'regulator gene of glucosyl-transferase', family proteins are also known as GadR, MutR, RopB and LasX (Federle, 2012). Rgg proteins are widespread in *Firmicute* specie, they function as stand-alone transcriptional regulators in low-G+C Gram-positive microbes (Chang et al., 2011). Structural prediction algorithms revealed potentially similar secondary and tertiary structure to PlcR and PrgX, members of the RNPP protein family. Both Rgg and RNPP family regulators contain a HTH in N terminal domain and a tetratricopeptide repeat (TPR) in C terminal region (Lasarre et al., 2013). N-terminal end HTH region function as a region to bind to the target

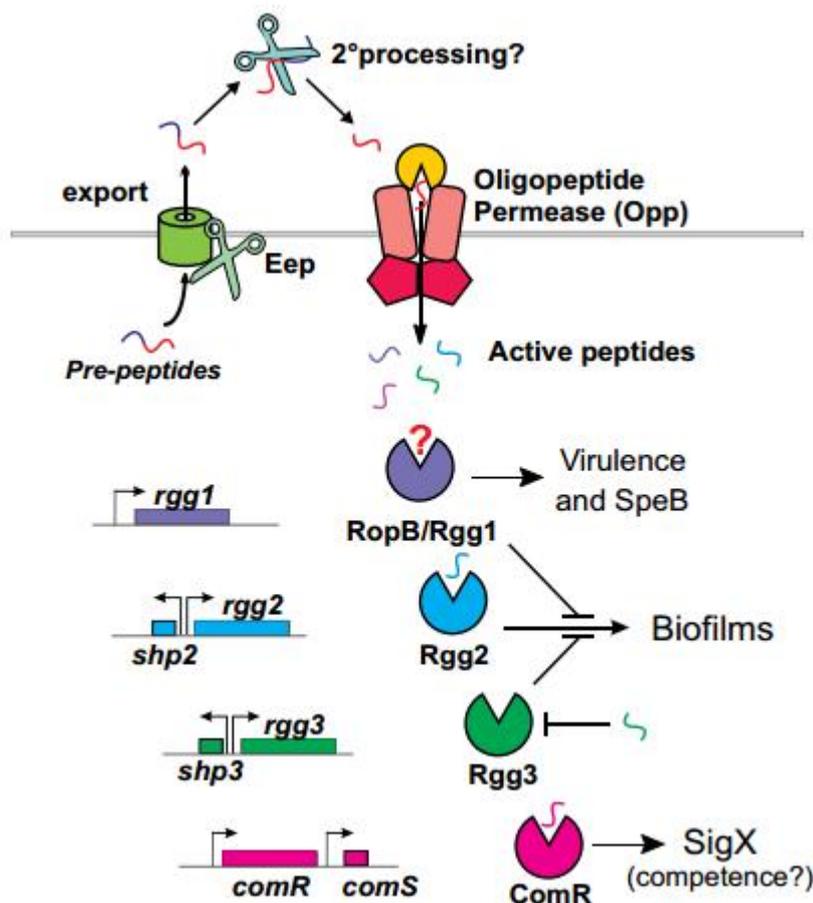
promoter regions, thus HTH motif plays a significant role in transcriptional regulation (Anbalagan et al., 2011). Rgg proteins are predicted to contain a tetratricopeptide repeat (TPR) at C-terminal ends, which may have an effect on binding to peptide pheromones (Anbalagan et al., 2011).

Rgg members, once considered to be ‘standalone regulators’ (Kreikemeyer et al., 2003), shown to respond to short hydrophobic peptide (SHP) that serve as pheromones. *In silico* analysis has showed that in streptococcal genomes there are some small open reading frames that encode short hydrophobic peptides, which are flanked by Rgg-regulated genes (Lasarre et al., 2013). These SHPs have been shown to be signal peptide for Rggs. However, it should be noted that not all *rggs* have a cognate *shp*. A new family of short coding sequences (CDSs) encoding SHPs was firstly identified by Ibrahim (Ibrahim et al., 2007), which are located upstream of Rggs. The size of SHPs range between 22 to 23 aa, containing at least one positively charged amino acid (lysine) in the N-terminal and one glycine in the C-terminal. The major component of SHPs is hydrophobic amino acids, which usually flanked by a negatively charged amino acid such as glutamate or aspartate (Ibrahim et al., 2007). The genome-wide survey of the short CDSs for SHPs revealed abundance of unknown Rgg-Shp QS pathways in *Streptococcus*. A previous report has analyzed 90 different genomes to identify *rgg*-like genes, including in various species of *Lactobacillale* and *Listeriaceae* genus (Fleuchot et al., 2011). A total of 494 *rgg* genes were identified along with 61 adjacent *shp* genes (Cook & Federal, 2013). Rgg regulators were found in all the analysed genomes except *Lactobacillus salivarius*. SHPs have been classified into three groups via phylogenetic study of Rgg amino acid sequences. In group I, the SHPs have a conserved glutamate (Glu) residue, while in group II; the SHPs have a conserved aspartate (Asp) residue at N terminal. In both groups, the *shp* and *rgg* genes are transcribed divergently. Contrastingly, in group III, the *shp* genes located at downstream of *rgg* genes, in a convergent orientation and the SHPs have a glutamate or an aspartate residue. A previous study has shown that the ‘activity’ of SHP in *S. pyogenes* has attenuated by changing its Asp amino acid with

an amide-bearing residue; however, the activity is maintained by changing an Asp residue with a Glu. Therefore, all mature SHPs have an Asp or Glu at their N terminals. These conserved residues are believed to be important for the recognition of the precursor by the protease involved in their maturation and the activity of the mature SHP (Fleuchot et al., 2013).

It has been reported that Rgg family members have a variety of physiological functions. They are involved in regulation of a diverse set of genes, thus enabling bacteria adapting to different environmental stimuli (Zheng et al., 2011). For instance, Rgg regulators have effect on thermal adaption in *Streptococcus thermophilus* (Henry et al., 2011), on the expression of the glucosyltransferase gene (*gtfR*) in *S. oralis* (Fujiwara et al., 2000), on the pathogenicity in *S. pyogenes* and *S. mutans*, and on control of biofilm formation in *S. pyogenes* (Chang et al., 2011). They are implicated in H<sub>2</sub>O<sub>2</sub> and paraquat resistance in *S. pyogenes*, and bacteriocin production in *S. mutans* (Qi et al., 1999) and *Lactobacillus sakei* (Rawlinson et al., 2005, Skaugen et al., 2002). They also affect the non-glucose carbohydrate metabolism, and are necessary for the virulence of *S. suis* (Zheng et al., 2011). GadR regulates the expression of genes involved in glutamate-dependent acid resistance in *Lactococcus lactis* (Sanders et al., 1998).

It has been demonstrated in other streptococci that Rgg-SHP pairs act as quorum sensing system. For example, in *S. pyogenes*, three out of four Rggs, Rgg2 and Rgg3, are associated with SHPs. These SHP peptides are synthesized as inactive precursors, processed by pheromone-specific peptidase (Eep), and released into extracellular medium by the general secretory (Sec) system or the ABC-type transporters. Mature peptides are imported into cytoplasm by an oligopeptide permease transport system called Opp or Ami, member of the ubiquitous ATP-binding cassette superfamily (ABC transporters) (Linton and Higgins, 2007). Once inside the cells, active peptides activate or repress the transcription of target genes to affect bacterial behaviors such as biofilm formation (Figure 1.6) (Cook and Federle, 2014).

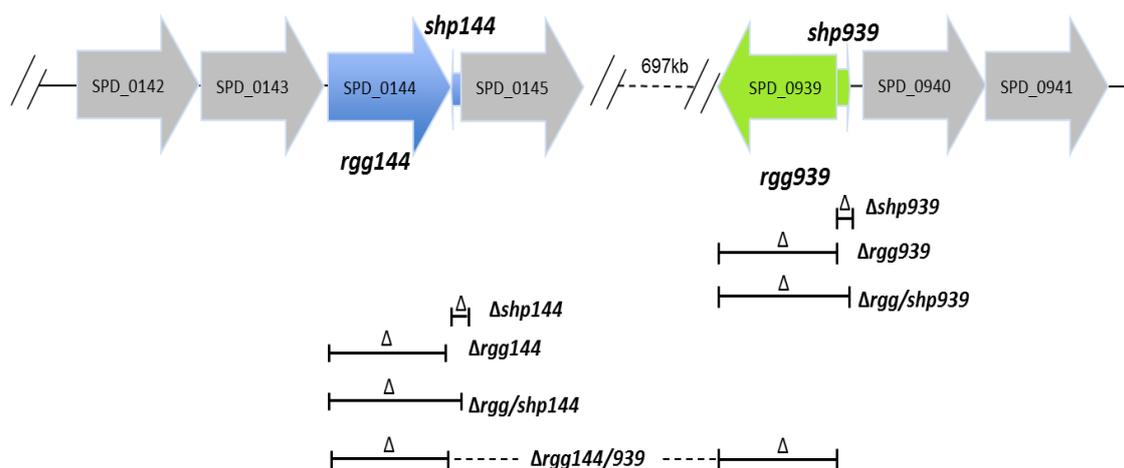


**Figure 1.6. Proposed model of Rgg-dependent quorum sensing in *Streptococcus pyogenes* (Cook & Federle, 2013).**

Nearly all streptococci genomes contain one copy of *rgg/shp*, but some streptococci have multiple copies. This suggests that Rgg QS systems are widespread in the streptococci (Fleuchot et al., 2011). Genome analysis has shown that there are 6 Rgg homologs in *Streptococcus pneumoniae* type 2 D39 strain, which are coded by SPD\_0112, SPD\_0144, SPD\_0939, SPD\_0999, SPD\_1518, and SPD\_1952, whose deduced amino acid sequence exhibits 23 to 28% identity and 42 to 48% similarity to Rgg transcriptional regulators in *S. gordonii*, which is regarded as a Rgg prototype. Only SPD\_0144 and SPD\_0939 are associated with *shp* genes, coding for group III and I peptides, respectively. This is in contrast to *S. pyogenes*, which has only group I

peptides, which may indicate the mechanism of action of Rggs in *S. pneumoniae* and *S. pyogenes* are different.

The functional characterization of pneumococcal Rggs have not been done with the exception of SPD\_1952, which has been shown to be important for pneumococcal virulence, oxidative stress resistance, and biofilm formation (Bortoni et al., 2009). Therefore the objective of this study is to determine Rggs involvement in pneumococcal biology. Particularly, I will study the function of two Rggs, SPD\_0144 and SPD\_0939, which are associated with putative *shp* genes (Figure 1.7). The study will gather evidence to support the following hypotheses:



**Figure 1.7. Rgg regulators in the *S. pneumoniae* D39 genome.** The location of *rgg144* and *rgg939* is indicated by the blue and green box, respectively, and their neighboring genes is indicated by the gray box. Gene deletions are indicated below the box and correspond to genotypes listed in Table 2.1.

## **1.9 Hypothesis**

### **1.9.1 Hypothesis on Rgg's involvement in non-glucose sugar metabolism**

In order to proliferate and colonise, *S. pneumoniae* needs to obtain energy from its host. As a preferred carbon source, glucose represses the expression of systems involved in the use of secondary carbon sources through a process known as carbon catabolite repression (CCR). Although the pneumococcus has specific phosphotransferase systems (PTS) for fructose, glucose, lactose, mannitol, and trehalose (Bidossi et al., 2012), most of these sugars are either not found or very rare in the respiratory tract. Galactose is the prevailing sugar in the human respiratory tract, while human blood is rich in glucose. Previous studies suggested that *S. pneumoniae* can break down complex carbohydrates for their demand of nutrient (Yesilkaya et al., 2008). In the nasopharynx, *S. pneumoniae* most likely relies on mucin glycoproteins present in the mucus layer coating the epithelial cell surfaces of airway structures (Rose and Voynow, 2006, Yesilkaya et al., 2008b). These large macromolecular multifunctional glycoproteins are highly glycosylated with a molecular weight of 2-20 x 10<sup>5</sup> Da, containing high carbohydrate content, up to 90% (Rose and Voynow, 2006). Mucin contains N-acetyl-glucosamine, galactose, fucose and sialic acids in its structure, which can be utilised by the microbe (Terra et al., 2010). Galactose and N-acetyl-galactosamines were demonstrated to be the most abundant sugars, and forms approximately 45% of sugar content in mucin (Terra et al., 2010a). The pneumococcus is able to cleave sugars from host glycans with a large array of glycosidases including neuraminidases, galactosidases, glycanase and hexosaminidases (King, 2010), and provide free sugars that can potentially be used for pneumococcal growth (Burnaugh et al., 2008, Marion et al., 2012).

The availability of carbon source varies in different host niches (King, 2010). Thus, nutrient quantity and quality can be a stimuli that trigger adaptive changes to promote survival of microbes in a changing environment. According to Dmitriev et al., (2006), *ropB*, an ortolog of pneumococcal Rgg in *S. pyogenes*, is important for fructose, mannose and sucrose utilization as the primary carbon sources. *In vitro* growth assay shows that *rgg* deficient mutant has difficulties when grown in Chemically Defined Medium (CDM) supplemented with fructose, mannose and sucrose (Dmitriev et al., 2006). In addition, mannose acts as the environmental indicator, which can significantly induces the Rgg-SHP system in *S. pyogenes* (Chang et al., 2015). Alternation of carbon source requires Rgg2 and SHPs to modulate gene expression, which is dependent upon the expression of the inducible mannose phosphotransferase system, *ptsABCD* (Chang et al., 2015). Therefore, it is reasonable to hypothesis that pneumococcal Rggs are also important for regulation of sugar metabolism.

### **1.9.2 Hypothesis on Rggs role in oxidative stress resistance**

The pneumococcus has to sense and respond to the changing concentrations of oxygen during infection (Bortoni et al., 2009). Hence, effective strategies are essential pneumococcal adaptation to the varying concentration of oxygen. *S. pneumoniae* is an anaerobic fermentative organism but it can also survive in oxygenated environments, like the nasopharynx of humans. The pneumococcus lacks many proteins that have been shown to protect against oxidative stress in other bacterial species such as NADP peroxidase, catalase, and the regulators of oxidative stress resistance such as SoxRS and OxyR (Tettelin et al., 2015), although *S. pneumoniae* still appears to deal effectively with the high levels of reactive oxygen species (ROS). In addition, during aerobic growth, *S. pneumoniae* generates exceptionally high concentrations of H<sub>2</sub>O<sub>2</sub>, which mainly results from the activity of pyruvate oxidase, which converts pyruvate to acetyl phosphate, CO<sub>2</sub> and H<sub>2</sub>O<sub>2</sub> (Spellerberg et al., 1996). H<sub>2</sub>O<sub>2</sub> poses a constant threat to the survival of bacteria in the aerobic environment since it results in severe

oxidative damage to DNA, protein, and lipids (Martindale and Holbrook, 2002), and finally mediate cell death at high concentration. A number of key enzymes have been implicated in the oxidative stress response, which include superoxide dismutase (SOD), NADH oxidase, alkyl hydroperoxidase and thiol peroxidase. It has been reported that some pneumococcal regulators are involved in modulation of oxidative stress resistance, such as “TCS04, PsaR, CiaRH, Rgg, MerR/NlmR, RitR, and SpxR” (Yesilkaya et al., 2013). Although effector mechanisms responsible for ROS resistance has been studied to certain extent, a detailed understanding of pneumococcal regulatory response to oxidative stress is still unclear. Some recent studies have shown that Rgg have a role in oxidative stress response. It has been reported that Rgg is involved in resistance to paraquat, which is a compound that can produce superoxide radical, in *S. pyogenes* (Chaussee et al., 2004). In contrast to the wild type strain, the *rgg* mutant was more susceptible to killing by paraquat (Chaussee et al., 2004). Moreover, in *S. pyogenes*, Rgg is involved in the repression of H<sub>2</sub>O<sub>2</sub> resistance and decomposition (Pulliainen et al., 2008). Compared to wild type strain, *rgg* deficient mutant could grow much better in presence of 3 mM H<sub>2</sub>O<sub>2</sub>. On the other hand, H<sub>2</sub>O<sub>2</sub> decomposition was significantly faster by stationary *rgg* deficient mutant. The mRNA levels for the general stress response components GroEL and DnaK, as well as the CovR response regulator were down-regulated in the *rgg* deficient mutant strain (Pulliainen et al., 2008). Furthermore, according to Bortoni et al. (2009), in absence of pneumococcal *rgg* (SPD1952), pneumococci were more susceptible to oxidative stress. These studies show that *rgg* mutation presents significant growth impairment under aerobic conditions, and leads to sensitivity to paraquat. Taken together, it is reasonable to hypothesis that pneumococcal Rggs play a role in oxidative stress resistance.

### **1.9.3 Hypothesis on Rggs role in biofilm formation**

Biofilm is formed by bacterial aggregation and adherence to inert surfaces (Talekar et al., 2014). A high percentage of bacterial infections are caused by biofilm forming cells. In contrast with planktonic cells, biofilm-associated cells are much more resistant to antibiotics, which will contribute to the upsurge of antibiotic resistance (Thomas et al., 2006).

Pneumococci colonise on the surface of human nasopharynx, closely attached to the epithelial surface to form biofilms (Hakansson et al., 2015, Mayanskiy et al., 2015). The biofilm-associated cells cannot be completely eliminated by using antibiotics (Marks et al., 2014). Moreover, bacteria can be released from the biofilm and disseminate to lower respiratory tract, the blood stream and brain to cause life-threatening diseases. However, how pneumococci regulate biofilm formation is still unclear. According to Chang et al., (2011), Rgg regulators associated with biofilm formation in *S. pyogenes*. Therefore, I hypothesize that Rggs are able to control the biofilm formation in pneumococci.

### **1.9.4 Hypothesis on Rggs role of pneumococcal virulence**

The pneumococcus resides in the mucosal surfaces of the host nasopharynx and upper airway. Through a combination of virulence-factor activity and an ability to evade the early components of the host immune response, this organism can spread from the upper respiratory tract to the sterile regions of the lower respiratory tract, which leads to pneumonia. To date, various studies demonstrated that certain transcriptional regulators are able to control the expression of virulence factors, effect the nutrition acquisition *in vivo*, and mediate environmental adaptation. For instance, it has been reported that in *S. pyogenes* RopB which is a Rgg class of regulator, is linked to bacterial virulence by coordinating the expression of various virulence related factors, including the M protein which binds to the CD4 receptor on mammalian cells and

inhibits phagocytosis, streptococcal pyrogenic exotoxin B (SpeB), which affects host-pathogen interactions by destroying the extracellular matrix and degradation of complement factors in host (Kappeler et al., 2009), and streptolysin O (SLO), which is a toxin and mediates the microbe's haemolytic activity (Feil et al., 2014). In addition, RopB modulates the expression of regulatory systems such as Mga, CovRS, and FasBCAX, effecting the expression of genes encoding proteins involved in adherence and resistance to phagocytosis (Dmitriev et al., 2006). Due to the similarity in the sequence of Rggs, it is reasonable to hypothesise that pneumococcal Rgg regulators might affect pneumococcal pathogenicity by coordinating virulence-associated gene expression or via attenuated nutrient acquisition or weakened fitness to environmental stress.

### **1.9.5 Hypothesis on Rgg-SHP pairs being part of a quorum sensing system**

Intercellular communication is of great significance for bacterial survival, which is crucial for gene expression in response to extracellular signals (Lasarre et al., 2013). Recent studies suggest that several Rgg proteins involve in quorum sensing together with SHP signaling peptides (Chang et al., 2011). To date, there are 484 Rgg-like proteins, and 68 *rgg/shp* copies have been identified, although all these SHP pheromones are generally similar (Fleuchot et al., 2013). For instance, in *S. pyogenes*, Rgg2 and Rgg3 can respond to SHP2 and SHP3, respectively, to control biofilm formation (Lasarre et al., 2013). In addition, Rgg1358 can respond to SHP1358 in *S. thermophilus* LMD-9 (Fleuchot et al., 2011). Moreover, SHP1520 and RovS, which form a QS circuit in *S. agalactiae*, has been identified as a cross-species signalling system that is able to activate SHP/Rgg circuit of *S. pyogenes* (Cook and Federle, 2014). Both pneumococcal *rgg144* and *rgg939* lie adjacent to small open reading frames encoding SHP144 and SHP939. These findings arouse interest to investigate whether the pneumococcal Rggs and SHPs are components of a novel QS regulatory

system, although the genetic context of Rgg/SHPs and signalling in *S. pneumoniae* might be different from other streptococci.

## **1.10 Aims**

Hence the aims of this work are to evaluate a role for Rgg regulators in the pneumococcus, elucidate how Rgg mediated gene regulation occurs, and to determine if selected SHPs interact with their cognate Rggs. To investigate these questions, several isogenic Rgg and SHP mutants were constructed, and characterised *in vitro* by growth studies, and microarray analysis. The Rgg mutants and wild type strains were challenged with oxidative stress and different carbon source to check if the Rgg regulators had any roles in oxidative stress resistance and sugar metabolism. The Rgg regulators were overexpressed and purified, and the direct protein-DNA interactions were determined using the purified proteins and several putative promoter probe targets. In order to determine the regulatory roles for Rgg regulators and their response to host-derived sugars, transcriptional *lacZ*-fusions to the promoters of the *rgg* and *shp* in different strain background were constructed and analysed. To determine the impact of Rgg regulators on pneumococcal virulence and nasopharyngeal colonisation, Rgg mutants and wild type strains were characterised in a mouse model of pneumococcal infection.

## **Chapter II : Materials and Methods**

## **2.1 Bacterial strains and plasmids**

The strains and plasmids used in this study, and their characteristics have been provided in Table 2.1.

**Table 2.1. Bacterial strains and plasmids**

Strain or plasmid*	Description	Source
D39	<i>Streptococcus pneumoniae</i> type 2 virulent strain	Laboratory collection of Dr Hasan Yesilkaya
$\Delta rgg144$	Same as D39 but SPD_0144 was deleted	Constructed in this study
$\Delta rgg144/shp144$	Same as D39 but SPD_0144 and SHP_0144 were deleted	Constructed in this study
$\Delta rgg939$	Same as D39 but SPD_0939 gene was deleted	Constructed in this study
$\Delta rgg939/shp939$	Same as D39 but SPD_0939 and SHP_0939 gene were deleted	Constructed in this study
$\Delta rgg144/939$	SPD_0144 and SPD_0939 gene were deleted	Constructed in this study
$\Delta rgg144/shp144/rgg939/shp939$	SPD_0144, SHP_0144, SPD_0939, and SHP_0939 gene were deleted	Constructed in this study
$\Delta shp144$	Same as D39 but SHP_0144 was deleted	Constructed in this study
$\Delta shp939$	Same as D39 but SHP_0939 was deleted	Constructed in this study
$\Delta rgg144Comp$	Complementation of D39 $\Delta rgg144$	Constructed in this study

<i>Δrgg144/shp144</i> Comp	Complementation of <i>D39Δrgg144/shp144</i>	Constructed in this study
<i>Δ rgg939</i> Comp	Complementation of <i>D39Δrgg939</i>	Constructed in this study
<i>Δrgg939/shp939</i> Comp	Complementation of <i>Δrgg939/shp939</i>	Constructed in this study
XZ1	D39 with integrated $P_{rgg144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ2	D39 with integrated $P_{shp144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ3	D39 with integrated $P_{rgg939}$ - <i>lacZ</i> reporter	Constructed in this study
XZ4	D39 with integrated $P_{shp939}$ - <i>lacZ</i> reporter	Constructed in this study
XZ5	<i>D39Δrgg144</i> with integrated $P_{rgg144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ6	<i>D39Δrgg144</i> with integrated $P_{shp144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ8	<i>D39Δrgg144</i> with integrated $P_{shp939}$ - <i>lacZ</i> reporter	Constructed in this study
XZ10	<i>D39Δrgg939</i> with integrated $P_{shp144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ11	<i>D39Δrgg939</i> with integrated $P_{rgg939}$ - <i>lacZ</i> reporter	Constructed in this study
XZ12	<i>D39Δrgg939</i> with integrated $P_{shp939}$ - <i>lacZ</i> reporter	Constructed in this study
XZ18	<i>D39Δrgg144/939</i> with integrated $P_{shp144}$ - <i>lacZ</i> reporter	Constructed in this study
XZ20	<i>D39Δrgg144/939</i> with integrated $P_{shp939}$ - <i>lacZ</i> reporter	Constructed in this study

XZ26	D39 $\Delta$ <i>shp144</i> with integrated P <sub><i>shp144</i></sub> - <i>lacZ</i> reporter	Constructed in this study
XZ36	D39 $\Delta$ <i>shp939</i> with integrated P <sub><i>shp939</i></sub> - <i>lacZ</i> reporter	Constructed in this study
<i>E. coli</i> BL21 DE3	<i>F-ompT hsdSB (rb- mb-) gal dcm</i> (DE3)	Novagen
<i>E. coli</i> DH5 $\alpha$	Antibiotic sensitive strain used for plasmid propagation	Laboratory collection
pPP2*	Contains promoterless LacZ gene, tetracycline and ampicillin resistant	Halfmann <i>et al.</i> , 2007
pCEP*	Source of kanamycin resistance cassette, used for genetic complementation	Guiral <i>et al.</i> , 2006
pDL278*	source of spectinomycin resistance cassette	Yesilkaya <i>et al.</i> , 2000
pE-SUMO*	T7, SUMO-6His-Tag, ampicillin resistant, used for Rgg939 protein expression	LifeSensors
pLEICS-01*	6His-Tag, ampicillin resistant, used for Rgg144 protein expression	PROTEX, UK

## **2.2 Chemicals**

All the chemicals used in this work were purchased from Qiagen, (UK), Sigma-Aldrich Ltd (UK), Invitrogen (UK) or Fisher Scientific (UK) unless otherwise stated. Primers were ordered from MWG (UK). DNA polymerase and restriction enzymes were purchased from New England Biolabs (UK).

### **2.2.1 Antibiotics**

The list of antibiotics used in this study is shown in Table 2.2.

**Table 2.2. Antibiotic solutions used**

<b>Antibiotic</b>	<b>Stock solution</b>	<b>Working concentration*</b>
Kanamycin	100 mg/ml in H <sub>2</sub> O	250 µg/ml
Spectinomycin	100 mg/ml in H <sub>2</sub> O	100 µg/ml
Tetracycline	15 mg/ml in 50% (v/v) ethanol	3 µg/ml
Ampicillin	100 mg/ml in H <sub>2</sub> O	100 µg/ml

Antibiotics prepared in distilled water (dH<sub>2</sub>O) or 50% (v/v) ethanol and filtered by using a 0.2 µm syringe filter, divided into stocks of 500 µl, and stored at -20 °C. \* Working concentration for *S. pneumoniae*, 100 µg/ml of kanamycin was used for *E. coli*.

## **2.3 Media Preparation**

All media used was prepared according to manufactures instructions and autoclaved at 121°C for 15 minutes at 15 psi pressure.

### **2.3.1 Brain Heart Infusion Medium**

BHI broth was prepared by dissolving 3.7 g in 100 ml of distilled water and autoclaved. The media was stored at room temperature.

### **2.3.2 Blood agar base**

Blood agar base (BAB) was prepared by dissolving 4.0 g in 100 ml distilled water autoclaved and then cooled to about 50 °C. Appropriate antibiotics and 5% of defibrinated horse blood were added to the medium, mixed well and poured into plates at about 20 ml per plate.

### **2.3.3 Lauria Bertani medium**

Lauria Bertani (LB) broth was prepared by dissolving 1 g of NaCl, 0.5 g of yeast extract and 1 g of tryptone in 100 ml of distilled water, and autoclaved. Lauria Bertani agar was prepared by addition of 1 g agar into LB medium.

### **2.3.4 Todd-Hewitt yeast extract medium (THY)**

THY broth was prepared by 30 g of the Todd-Hewitt in one liter of distilled water with 0.5% (w/v) yeast extract. Then, broth was autoclaved at 121 °C for 15 minutes.

### **2.3.5 Chemically defined medium**

The composition of the defined medium is described in Table 2.3

**Table 2.3.** Composition of CDM used for growth of pneumococcal strains.

<b>Ingredient</b>	<b>g/L</b>	<b>Ingredient</b>	<b>g/L</b>
<b>Basal Solution (pH 6.5)</b>		<b>Amino acids (pH 6.5)</b>	
Na <sub>2</sub> -β-glycerophosphate	26.0	Alanine	0.24
KH <sub>2</sub> PO <sub>4</sub>	1.0	Arginine	0.124

(NH <sub>4</sub> ) <sub>3</sub> citrate	0.6	Asparagine	0.352
Na-Acetate	1.0	Aspartate	0.4
Cysteine-HCl	0.4	Glutamate	0.5
<b>Vitamins (pH 6.5)</b>		Glutamine	0.392
Na-p-Aminobenzoate	0.005	Glycine	0.176
D-Biotin <sup>2</sup>	0.0025	Histidine	0.152
Folic acid <sup>2</sup>	0.001	Isoleucine	0.212
Nicotinic acid	0.001	Leucine	0.456
Ca(D+)Pantothenate	0.001	Lysine	0.44
Pyridoxamine-HCL	0.0025	Methionine	0.124
Pyridoxine-HCl	0.002	Phenylalanine	0.276
Riboflavin1	0.001	Proline	0.676
Thiamine-HCl	0.001	Serine	0.34
DL-6,8-Thioctic acid	0.0015	Threonine	0.224
Vitamin B12	0.001	Tryptophane	0.052
Nitrogenous bases		Valine	0.324
Adenine	0.01	Micronutrients	
Uracil	0.01	MgCl <sub>2</sub>	0.20
Xanthine	0.01	CaCl <sub>2</sub>	0.038
Guanine	0.01	ZnSO <sub>4</sub>	0.005
Other Chemicals			
Pyruvate	0.1	Choline-HCl	0.01

The nitrogen base was dissolved in 0.1 M NaOH, d-biotin and folic acid was dissolved in 2 M NaOH. Basal solutions, vitamins and amino acids was adjust to pH6.5. Glucose/ Galactose/ Mannose/ *N*-actyl glucosamine were added to the appropriate amount.

## **2.4 Preparation of cultures**

### **2.4.1 Growth of *Escherichia coli***

*Escherichia coli* was inoculated into LB medium. The culture was incubated at 37 °C on a shaking platform at 200 revolution/minute (rpm). When grown on agar, LA plates were used, and they were incubated at 37 °C overnight.

### **2.4.2 Growth of *Streptococcus pneumoniae***

*Streptococcus pneumoniae* was grown in BHI broth, or on BA plates supplemented with 5% (v/v) defibrinated horse blood at 37 °C in a candle jar, and incubated at 37 °C overnight. Pneumococcal strains were also grown in chemical defined medium (CDM) microaerobically and anaerobically supplemented either with glucose, galactose, mannose, or *N*-acetyl glucosamine (Table 2.3). Pneumococcal strains were grown initially on BAB plates, one single colony was picked up and transferred to 10 ml BHI, incubated at 37 °C until mid-exponential phase. The cultures were centrifuged at 3500 rpm (Sorvall legend T, Thermo Scientific) for 10 min at room temperature and the supernatant was discarded. The pellets were resuspended with 1 ml of CDM, and 200 µl was transferred into 10 ml of CDM supplemented with selected carbon source. The cultures were then incubated at 37 °C until mid- or late-exponential phase as required.

### **2.4.3 Stock preparation of bacterial strains**

The frozen stocks of the bacterial strains were prepared initially by streaking on a blood agar plate with or without appropriate antibiotics, and were incubated at 37 °C in a candle jar overnight. A sweep of colonies was inoculated in 10 ml BHI or LB broth and the tubes were incubated at 37 °C until optical density at 600 nm (OD<sub>600nm</sub>) had reached 0.4-0.5. Then the cultures were centrifuged at 4500 rpm for 10 min in Allegra™ X-22 centrifuge (Beckman Coulter, USA). The supernatant was discarded and the pellet was suspended in 2 ml of BHI or LB containing 15% (v/v) glycerol. Then the suspension was divided into 10 aliquots in 1.5 micro-centrifuge tubes and stored at -80 °C.

### **2.5 Assessment of colony forming unit counts**

Bacterial colony forming unit (CFU) counts in bacterial suspensions were determined as described previously (Miles et al., 1938). 20 µl of *S. pneumoniae* bacterial suspension was mixed with 180 µl of sterile PBS in a 96-well microtitre plate, followed by serial dilution. 60 µl of the dilutions were plated onto BA plates supplemented with antibiotic when necessary. The plates were allowed to dry and incubated overnight in a candle jar. The next day the plates were counted. The viable colonies were counted in the sections where between 30 to 300 colonies were observed. The number of colony forming units per ml was calculated using the following formula:

$$\text{CFU/ml} = (\text{numbers of colonies} \times \text{dilution factor}) \times (1000/60)$$

## **2.6 Plasmid extraction from *E. coli***

Plasmid DNA was extracted using the QIAprep spin Miniprep kit (Qiagen, UK). Bacterial cultures were incubated for 12-16 h at 37 °C in LB broth with appropriate antibiotics with vigorous shaking. 1.5 ml of an overnight bacterial culture was placed in a micro-centrifuge tube, and centrifuged at 13, 000 rpm for 10 minutes at room temperature. The supernatant was discarded and the pellet was re-suspended in 250 µL of buffer P1 containing RNase A. 250 µL of buffer P2 was added and mixed by inverting the tube several times. After this, 350 µL of buffer N3 was added and the solution was mixed immediately by inverting the tube several times. The reaction was centrifuged for 10 minutes at 13, 000 rpm and the supernatant was transferred to a QIAprep spin column with collection tube. Then the tube was centrifuged for 30 seconds at 13, 000 rpm. The liquid was discarded and the column placed back onto the collection tube. 750 µL of buffer PE was added and centrifuged for 30 seconds at 13, 000 rpm. The liquid was discarded and a further centrifugation of 1 minute at 13, 000 rpm was performed to remove any remaining buffer PE. To elute the DNA, the column was placed on a 1.5 ml micro-centrifuge tube, 50 µL of DNase/RNase free water were added to the centre of the column, left to stand for 1 min, and centrifuged at 13, 000 rpm for 1 min. The eluted DNA was stored at -20 °C until further use.

## **2.7 Extraction of pneumococcal chromosomal DNA**

Pneumococcal chromosomal DNA was prepared as described previously (Saito and Miura, 1963). *S. pneumoniae* was incubated on blood agar (BA) plates in a candle jar overnight at 37 °C. Then, one single colony was transferred in 10 ml of BHI broth and incubated overnight at 37°C. 200 µl of overnight culture of bacteria (OD<sub>600nm</sub>: 0.9-1.1) was inoculated to 10 ml BHI, grown at 37 °C for 6-8 h. The pneumococcal culture was centrifuged at 3500 rpm for 15 min, the pellet was re-suspend in 400 µl TE buffer (1 M Tris-HCl and 500 mM EDTA, pH 8.0) containing 25% (w/v) of

sucrose, 60  $\mu$ l of 500 mM EDTA and 40  $\mu$ l of 10% (w/v) of SDS and 2  $\mu$ l of Proteinase K (12.5 mg/ml) were added. Samples were incubated at 37 °C for 1-2 h to obtain clear lysate. After incubation, samples were centrifuged at 13, 000 rpm for 5 min. The upper aqueous phase was transferred to a fresh tube, and an equal volume of phenol:chloroform:isoamyl alcohol (25:24:1, v/v, respectively) (Invitrogen, UK) was added. The samples were centrifuge at 13, 000 rpm for 10 min, and the upper aqueous phase was transferred to a fresh tube, without disturbing the white protein layer. This process was repeated once more, the upper aqueous phase was transferred to a fresh tube, and mixed with 5 volumes of 100% (v/v) ethanol and 0.1(v/v) volume of sodium acetate (3M pH 5.2). Nucleic acids were pelleted by centrifugation, and ethanol was discarded. Finally, the pellet was washed with 500  $\mu$ l of 70% (v/v) of ethanol, the pellet was dried at room temperature and re-suspended in 250  $\mu$ l TE buffer. DNA samples were kept at 4 °C until needed. Concentration of DNA was determined by using a NanoDrop<sup>TM</sup> spectrophotometer at OD<sub>260nm</sub> (Thermo Scientific, UK).

## **2.8 Polymerase Chain Reaction (PCR)**

In this study polymerase chain reaction (PCR) was performed (Saiki et al., 1988) for mutation detection, genetic modification, and amplification of desired fragments. The list of primers used in this study has been provided in Table 2.4. A typical PCR reaction mixture was prepared in a final volume of 20  $\mu$ l as indicated in Table 2.5.

**Table 2.4. Primers used in this study.**

<b>Primers</b>	<b>SEQUENCE</b>	<b>TM* ( °C)</b>
SPD0144LF	TGGAGGATATTGTCTTCCCAAAG	58.9
SPD0144RR	CTGGGTACGAGGAAAGATATAGGTTG	63.2
SPD0144LR	TATTCACGAACGAAAATCGATTTC AATC ATCCTACCACCTCC	70.4

SPD0144RF	AACAATAAACCCCTTGCATAACATAGATT AGTTAAAAATGGAA	64.6
SHP0144RF	AACAATAAACCCCTTGCATAACCCAGGG AAATAATCAAATCTATCA	68.5
SPD0144LF	TGGAGGATATTGCTTCCCAAAG	58.9
specF	ATCGATTTTCGTTTCGTGAATACATGTTAT	59.6
specR	GTTATGCAAGGGTTTATTGTTTTCTA	56.9
SPD0939RF	TTACTGGATGAATTGTTTTAGCATAGTTT TCCCATTTCCCA	67.5
SPD0939LR	TGCGCGCCATGGTAGCACCTCTAAATGA AATTATTCTCTATC	71.4
SPD0939RR	TCCAATGTACCAATAACTCGAGG	61.0
SPD0939LF	ATCCACTGAAGAGATTGAGGAAGC	61.0
Shp0939RF	TTACTGGATGAATTGTTTTAGGCTTAGG TGTACACCTAAGTC	71.4
Shp0939RR	CCACTTAGTTGAAGCAGCATC	57.9
KanF	GAGGTGCTACCATGGCGCGCATGC	69.5
KanR	CTAAAACAATTCATCCAGTAAAATATAAT A	55.4
Spd0939NcoI	<b>CGCCATGG</b> ATTTACGAGAATGCCTTAC	66.6
Spd0939BamHI	<b>ACGGGATCC</b> CTTCATACTCTGACTTCCTT	66.7
SHP0939BamHI	<b>ACGGGATCC</b> GATACTCTAAAGACTTAGG T	65.3
Spd0144BamHI	<b>ACGGGATCCC</b> GGAGTGGGTATTGTCA	68.0
SHP0144BamHI	<b>ACGGGATCC</b> CAAGTACAGTATAACACG	66.0
Spd0144NcoI	<b>CGCCATGG</b> CAATCACCAAAGTGTCAG	66.4
Prgg939SphI	<b>GACGCATGCT</b> CTTCATACTCTGACTTCC	66.6
Prgg144SphI	<b>GACGCATGC</b> ACGAGGATTGAGTTCTG	66.4
Pshp939SphI	<b>GACGCATG</b> CAATTTGTTTGCCCTTACG	65

<i>Pshp144Bam</i> HI	<b>ACGGGATCC</b> AGTAGGAGTATCAGTTGG	66.5
<i>Prgg144Bam</i> HI	<b>ACGGGATCCCCAA</b> ACGTAATTCCTTG	64.8
<i>Prgg939Bam</i> HI	<b>ACGGGATCCG</b> TTTGCCCTTACGAATTT	65
<i>Pshp939R</i>	<b>ACGGGATCCT</b> CTTCATACTCTGACTTCC	66.6
Shp0144LR	TATTCACGAACGAAAATCGATGTGGGTTA TTGTCATTCCATTT	68.5
Shp0939LR	TGCGCGCCATGGTAGCACCTCTCTTCATA CTCTGACTTCCTTC	75
Shp144LF	TGCTGTCTGCGGATAAGTTGA	57.9
Shp0939LR	CCACTGGTTGATTGGCAAC	56.7
SPD0939 <i>Bsa</i> I	<b>TAGGTCTCGAGGT</b> ATGAAATCAAACCTT GGTGTAC	68.3
SPD0939 <i>Xba</i> I	<b>GCTCTAGATT</b> ATTTTTTGTTTACTGGT TTATTAATGG	63.1
SPD0144N	<b>TACTTCCAATCCATG</b> ATGATTGAAAAAA TGGAACCTGGGGG	69.4
SPD0144C	<b>TATCCACCTTTACTGTC</b> ATAAGTTCTTTA TATTGCTGAAAACGC	68.5
malF	GCTTGAAAAGGAGTATACTT	51.1
pCEPR	AGGAGACATTCCTTCCGTATC	57.9
Fusion-seqR	AGGCGATTAAGTTGGGTAAC	55.2
Fusion-seqF	CTACTTGGAGCCACTATCGA	57.3
T7 forward	TAATACGACTCACTATAGG	50.2
T7 terminator	GCTAGTTATTGCTCAGCGG	56.7

Bold typeface nucleotides are homologous to cloning sites. \*TM, melting temperature.

**Table 2.5. The ingredients used in a typical PCR reaction.**

Ingredients	Volume	Stock
Forward primer	1 $\mu$ l	10 pmol
Reverse primer	1 $\mu$ l	10 pmol
Template DNA	2 $\mu$ l	15 ng/ $\mu$ l
Enzyme mix	10 $\mu$ l	2 X
DNase/RNases free water	6 $\mu$ l	

PCR settings: 95 °C for 10 minutes, followed by 30 cycles of 95 °C for 45 seconds of denaturation, 55 °C for 45 seconds of annealing (varied depending on the  $T_m$  of primers) and 72 °C of 1 min/kb of extension, the final extension for 10 minutes at 72 °C and hold at 4 °C.

## **2.9 Agarose gel electrophoresis**

Agarose gel electrophoresis was used to visualise DNA fragments as described in (Sambrook et al., 1989). Routinely, agarose (Biolone, UK) at a concentration of 1% (w/v) was prepared in 100 ml 1X TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0). Ethidium bromide (from a stock solution of 10 mg/ml in water) was added to a concentration of 0.5  $\mu$ g/ml. The DNA samples were mixed with 3  $\mu$ l of 6X loading dye (New England Biolabs, UK) and carefully transferred to wells of the agarose gel using a disposable micropipette. 1 Kb or 100 bp DNA ladder (New England Biolabs, UK) were loaded into the wells of the agarose gel to measure the approximate concentration and/or size of the DNA samples. Electrophoresis was achieved by applying a constant voltage of around 90 V, and the DNA fragments were visualized under UV light using a long-wave UV transilluminator.

## **2.10 Restriction digestion**

For a restriction digestion, 2.5 µg of DNA was first mixed with the required amount of DNase/RNase free water and the appropriate buffer, and then the restriction endonuclease was added. The reaction incubated for a minimum of one and a half hours at the temperature required for optimum enzyme activity (usually 37 °C). The volume of the digestion reaction varied according to the amount of DNA to digest. A typical reaction mixture has been illustrated in Table 2.6.

**Table 2.6. A composition of typical restriction endonuclease reaction**

<b>Components</b>	<b>Quantities</b>
DNA	25 µl (1 µg)
Buffer (10X)	5 µl
Restriction enzyme 1	1 µl (10 unit)
Restriction enzyme 2	1 µl (10 unit)
DNase/RNase free water	18 µl
Total	50 µl

After incubation of the restriction digestion for three hours, a sample of the reaction mixture was analysed on a DNA electrophoresis gel. DNA was purified using the QIAquick PCR purification kit, and the restricted DNA was stored at -20 °C.

## **2.11 DNA Ligation**

Ligation reaction was used to join two separate DNA molecules. The following reaction was set up in a micro-centrifuge tube on ice (Table 2.7).

**Table 2.7. A typical ligation reaction used in this study.**

<b>Component</b>	<b>Amount</b>
10X T4 DNA Ligase Buffer*	2 $\mu$ l
purified Vector DNA	50 ng (0.025 pmol)
purified Insert DNA	50 ng (0.076 pmol)
Nuclease-free water	to 20 $\mu$ l
T4 DNA Ligase (400 U/ $\mu$ l)	1 $\mu$ l (5 U)

Ligation reaction was incubated at 16 °C for 16 h, and then heat inactivated at 65 °C for 10 min.

## **2.12 Method for the production of chemically competent *E. coli***

Chemically competent *E. coli* was prepared according to the protocol described by (Sambrook et al., 1989). One single colony of *E. coli* DH5 $\alpha$  was picked up and transferred into 5 mL of LB broth containing 20 mM MgSO<sub>4</sub> and incubated overnight at 37 °C. 1 mL of overnight culture was transferred into 100 mL LB broth in a conical flask and incubated until OD<sub>550 nm</sub> had reached 0.8. The bacteria were centrifuged at 3000 rpm (Sorvall legend T, Thermo Scientific) for 10 minutes, and the pellet was re-suspended in 30 mL of ice cold Tfb I (Table 2.8) buffer and kept on ice for 30 min. The bacterial suspension was centrifuged at 3000 rpm for 10 min at 4 °C, and the pellet was re-suspended in 4 mL of Tbf II buffer (Table 2.8). 60  $\mu$ L aliquots were distributed into ice-cold micro-centrifuge tubes, and stored at -80 °C.

**Table 2.8. Buffers for preparation of competent *E. coli* cells**

Tfb I buffer		Tfb II buffer	
Ingredients	Quantities	Ingredients	Quantities
K-acetate	3 mM	Na-MOPS	10 mM
MnCl <sub>2</sub>	50 mM	CaCl <sub>2</sub>	75 mM
KCl	100 mM	KCl	10 mM
CaCl <sub>2</sub>	10 mM	Glycerol	15 %
Glycerol	15 % (v/v)		

### **2.13 Genetic transformation of *Escherichia coli***

*Escherichia coli* competent cells were thawed on ice before use. 5 ng of DNA was added to a 50 µL aliquot of competent cells and left on ice for 30 minutes. After that the aliquot was quickly placed into the water bath at 42 °C for exactly 45 seconds, and then incubated on ice for 2 minutes. 0.5 mL of pre-warmed S.O.C medium (2% tryptone, 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl<sub>2</sub>, 10 mM MgSO<sub>4</sub>, 20 mM glucose) were added into the reaction and incubated at 37 °C for 1 h on a shaking platform (200 rpm) (New Brunswick Scientific, USA). After this, an appropriate amount of culture was plated on LA plates containing the appropriate antibiotic, and incubated overnight for the selection of transformant colonies.

### **2.14 Transformation of *Streptococcus pneumoniae***

Transformation was done as described previously (Bricker and Camilli, 1999). *S. pneumoniae* D39 was incubated on blood agar (BA) plates in a candle jar overnight at 37 °C. Then one single colony was transferred in 10 ml of BHI broth and incubated overnight at 37 °C. 1: 50 dilutions of overnight grown D39 was done until OD<sub>600nm</sub>

had reached between 0.05-0.08. 860  $\mu$ l of culture and the other ingredients (Table 2.9) were added into 1.5 ml micro-centrifuge tubes. The reaction was incubated at 37 °C for 3 hours, and then every hour 330  $\mu$ l of culture was plated out on BA plates with appropriate antibiotics. The plates were incubated overnight at 37 °C.

**Table 2.9. Chemicals used in transformation of *S. pneumoniae***

Ingredient (stock)	Volume	Final concentration
100 mM NaOH	100 $\mu$ l	10 mM/ml
20% BSA (w/v)	10 $\mu$ l	0.2%
100 mM CaCl <sub>2</sub>	10 $\mu$ l	0.2%
50 ng/ml CSP (w/v)	2 $\mu$ l	1 mM/ml

## **2.15 RNA extraction from *S. pneumoniae***

The method used for extracting RNA from *S. pneumoniae* was described by (Stewart et al., 2002). Bacteria were grown in BHI broth or CDM under micro-aerobic conditions until mid-log phase (approximately OD<sub>600nm</sub> of 0.3-04). 3 ml of bacterial culture was harvest and resuspended in 500  $\mu$ l Trizol reagent (Invitrogen, UK). Samples were vortexed for 15 sec, and then 100  $\mu$ l of chloroform was added. The bacterial suspension was quickly transferred to a RiboLyser blue matrix tube (Haybaid) and processed in a PowerLyzer™ 24 homogeniser (MO BIO, USA). After this, the tube was left for 5 min at room temperature, and was centrifuged at 12,000 g for 15 min at 4 °C. The upper phase was carefully transferred to a fresh micro-centrifuge tube. 250  $\mu$ L of isopropanol was added to precipitate the RNA and the mixture was left to stand at room temperature for 15 min. Then, the tube was centrifuged at 12,000 g for 10 min at 4 °C. The supernatant was removed, and the pellet was washed with 75% (v/v) ethanol to eliminate remaining isopropanol,

followed by centrifugation at 12,000 g for 5 min at 4 °C. The supernatant was carefully removed and the pellet was re-suspended in 100 µL of RNase free water.

## **2.16 Treatment of RNA with DNase I**

RNA was treated with the Amplification Grade DNase from Invitrogen following the manufacturer's instructions to remove any trace of contaminating DNA. 5 µg of RNA was mixed with 5 µl of 10x DNase buffer (Invitrogen) and 5 µl of amplification grade DNase I (1 unit/µl) in a final reaction volume of 50 µl. The reaction mixture was incubated at room temperature for 15 min. After this time, 5 µl of 25 mM EDTA, pH 8.0, was added, and the reaction was heated for 10 min at 65 °C to inactivate the enzyme. The reaction mixture was briefly placed on ice for 1 min and was then kept in 20 °C for future use.

## **2.17 Synthesis of complementary DNA**

First-strand cDNA synthesis was performed according to the manufacturer's instructions using SuperScript III reverse transcriptase (Invitrogen). DNase-treated total RNA (~1 µg) was mixed with 1 µl of 300 ng random primers (Invitrogen) and 1 µl of 10 mM of dNTP mixture (Promega, UK) in a total volume of 20 µl with nuclease-free water. The mixture was heated to 65 °C for 5 min. This followed by addition of 4 µl of 5X first-strand buffer and 2 µl of 0.1 M dithiothreitol (DTT). Contents were mixed gently and incubated at 25 °C for 2 min. 200 U SuperScript III reverse transcriptase in 1 µl was added, and mixed gently by pipetting up and down. Reverse transcription reaction was incubated at 42 °C for 50 min and then heated to 70 °C for 15 min to stop the reaction. The cDNA sample was diluted 5 fold with DNase-RNase free water, and was stored at -20 °C until needed.

## **2.18 Quantitative reverse transcriptase polymerase chain reaction (qRT-PCR)**

This assay was used to determine the relative gene expression of selected targets. The SensiMix™ SYBR® Hi-ROX kit (Bioline, UK) was used for qRT-PCR. 2 µl of diluted cDNA (~15 ng), 10 µl 2X SensiMix™ SYBR Hi-ROX, 2 µl of gene specific primers (3 pmol of each) (Table 2.10) and 6 µl of DNase/RNase free water were mixed together and placed in the Corbett RG-6000 PCR system (Qiagen, UK). Real time PCR was started at 95 °C for 10 min, followed by 40 cycles of 98 °C for 20 s, 55 °C for 30 s, 72 °C for 20 s and 74 °C for 20 s. The transcription levels of the genes were normalised to transcription of *gyrB*. Negative controls which were not treated with reverse transcriptase were run to check any DNA contamination. All samples were set up in triplicate. The real time PCR data was analysed by the comparative C<sub>T</sub> method as described by Livak and Schmittgen (Schmittgen and Livak, 2008). Twofold or greater differences in gene expression were considered as significant (Yesilkaya et al., 2009).

**Table 2.10. Gene specific primers used in RT-PCR**

<b>Primers</b>	<b>Sequence</b>	<b>T<sub>m</sub> (°C)</b>
gyrBRTF	TCGTGTGGCTGCCAAGCGTG	63.5
gyrBRTR	GGCTGATCCACCAGCTGAGTC	63.7
SPD145RTF	TTGCAGAAACAATGGAATTGA	52.0
SPD145RTR	ACCACTTCTTCCACCACCTC	59.4
SPD0146RTF	TGCAGTATGGTTGGATAACGA	55.9
SPD0146RTR	ATTCCATGAAGCATGCAGAG	55.3
SPD0147RTF	TTTACAGGAGCTACGGCATT	55.9

SPD0147RTR	AAACAGCGGAACCAAGTACA	55.3
SPD0187RTF	AAACGCCGTGTAGAAGAGTG	57.3
SPD0187RTR	TATCTAGTCGGCAGAAACGG	57.3
SPD0189RTF	AATCAGGAACAGGAAATGGG	55.3
SPD0189RTR	AAAGTTACTGAGGCGCAACC	57.3
SPD0219RTF	TCGTGACAACCTGAAGCTCGT	57.3
SPD0219RTR	TTCAGATGCGATTCATTACG	54.0
SPD0264RTF	ACGTCTACTTCACGGTCAGG	59.4
SPD0264RTR	ACATTACCTGGAGCTGCTTG	57.3
SPD0327RTF	TGAAGAGCTGGAAGCATCTG	57.3
SPD0327RTR	TTTCCTCATCAAGCACTTCG	55.3
SPD0445RTF	TGCGAAAGCTCTTCTTGAAA	53.2
SPD0445RTR	AGACCAAGGAAGCCTTCAGA	57.3
SPD0561RTF	TGGGTGTTACTCCAGCTCTC	59.4
SPD0561RTR	GGCACCTACACCTTTAGCAA	57.3
SPD0915RTF	GCAAATAAGATAAATGATGCTGA	53.5
SPD0915RTR	TGAGGCTGCTAACGGTGTAT	57.3
SPD0942RTF	GGAGGGATATGAGGTTCCAA	57.3
SPD0942RTR	AAGGGCCGTGCATAATTA	53.2
SPD0945RTF	TGGTAATGGCGATATCTGGA	55.3
SPD0945RTR	CTTTCACGGCCTGTTACAAA	55.3
SPD0949RTF	AGGTGAAGTTGAGGGATTGG	57.3
SPD0949RTR	CAAATGCAACAATTGCTCT	53.2
SPD1050RTF	GGTTCTGAGTGTGTGGCTGA	59.4

SPD1050RTR	AAAGCGTGGGTCTGAAAAGA	55.2
SPD1133RTF	TCATGATGTTTCTCCGTGTG	55.3
SPD1133RTR	GCATGAGGATTGCTGTTTCT	55.3
SPD1634RTF	GGGCTGATCAACGTGCTATT	57.3
SPD1634RTR	CTCAGCACGACGTTCATTGT	57.3
SPD1932RTF	CTTGCATGACCTTGCTGACT	57.3
SPD1932RTR	CTTCGTCAAGCTCGATACCA	57.3

## **2.19 Microarray experiments**

These experiments were done to determine regulons for Rggs in collaboration with Prof Oscar Kuipers, Groningen University. *S. pneumoniae* D39 and its isogenic mutant strains were grown anaerobically in CDM supplemented with 55 mM galactose or mannose as the main carbon source. RNA was extracted as described in section 2.15, and the samples were sent on ice to the Netherlands for cDNA synthesis, labeling, and data analysis as described before (Shafeeq et al., 2015).

## **2.20 Gene Splicing by Overlap Extension PCR (SOEing PCR)**

SOEing (Gene Splicing by Overlap Extension) PCR mutagenesis was used for the construction of mutated DNA sequences (Figure 2.1). During the first step, the genetic locus surrounding the region to be mutated was individually amplified using the proof reading DNA polymerase (PrimerSTAR HS, Takara). The flanking fragments were amplified by using XX-LF and XX-LR, and XX-RF and XX-RR primer sets, respectively (where XX indicates the gene designations) (Table 2.11). The XX-LR and XX-RF primers were designed to contain a homologous region for spectinomycin resistance gene (SpecR). The SpecR gene was amplified from the

plasmid pDL278 using the primers SpecF/R that are listed in Table 2.4. In the second step, all the products were mixed in equal ratios and fused with the outermost primer pair (XX-LF and XX-RR). PCR conditions were the same as the first step. Then, the DNA fragments were separated on 1% (w/v) agarose, the band of interest was excised and purified using the Wizard SV Gel and PCR Clean-Up System (Promega).

**Table 2.11. Ingredients used in SOEing PCR**

**First step:**

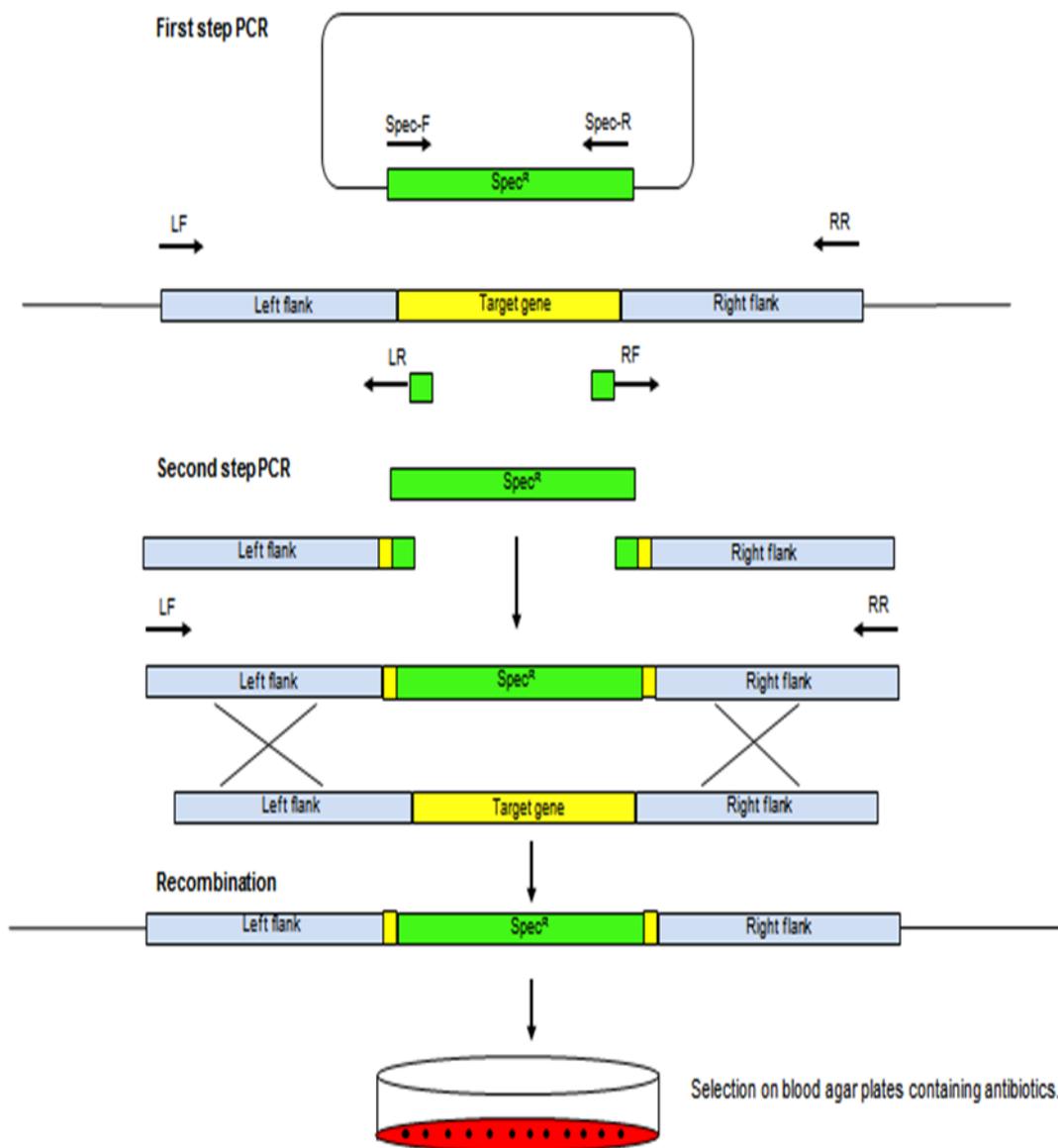
Ingredient		Volume
Template DNA	D39 genome DNA	3 $\mu$ l (15 ng/ $\mu$ l)
Primers	Reaction 1: XX-LF+XX-LR	3 $\mu$ l (10 pmol/ $\mu$ l)
	Reaction 2: XX-RF+XX-RR	3 $\mu$ l (10 pmol/ $\mu$ l)
Enzyme	PrimerSTAR HS (premix 2X)	25 $\mu$ l
Sterilized distilled water		19 $\mu$ l

PCR conditions: 30 cycles of 98 °C for 10 seconds, 55 °C for 15 seconds and 72 °C for 1 minute.

**Second step**

Ingredient		Volume
Template DNA	mixed, purified PCR products from first step	3 $\mu$ l (15 ng/ $\mu$ l)
Primers	XX-LF+XX-RR	3 $\mu$ l (10 pmol/ $\mu$ l)
Enzyme	PrimerSTAR HS (premix 2X)	25 $\mu$ l
Sterilized distilled water		19 $\mu$ l

The products were stored at -20 °C for future use.



**Figure 2.1. Schematic diagram of mutagenesis by SOEing PCR.** The bold crosses represent the homologous recombination during transformation. The LR and RF primers were designed to contain a homologous region for spectinomycin resistance gene (Spec<sup>R</sup>). Figure constructed based on Song et al (2005).

## **2.21 H<sub>2</sub>O<sub>2</sub> survival assay**

Wild type D39 and its isogenic mutants were grown in CDM under aerobic conditions to an OD<sub>600nm</sub> of 0.3. Then 1 ml culture was incubated either with 20 mM or 40 mM H<sub>2</sub>O<sub>2</sub> at 37 °C for 15 min. 20 µl of bacterial culture was then serially diluted in 180 µl of PBS. 40 µl of each dilution was spotted onto blood agar plate, and incubated overnight at 37 °C. The number of colonies was counted on the appropriate dilution. The percent survival was calculated by dividing the CFU of cultures after exposure to H<sub>2</sub>O<sub>2</sub> by the CFU in the control culture without H<sub>2</sub>O<sub>2</sub>.

## **2.22 Paraquat susceptibility assay**

The pneumococcal inoculum was prepared as described above (Bortoni et al., 2009) and exposed to either 0.05 or 0.1 mM paraquat for 1 h. The survival was determined by serial dilution and plating on blood agar plates. The results were expressed as percentage survival relative to the control, which had not been treated with paraquat.

## **2.23 Bacterial growth assay**

Growth studies were done by using a MultiskanTM GO Microplate Spectrophotometer (Thermo Scientific, UK). BHI or CDM (198 µl) supplemented with 55 mM of desired carbon source was added into the wells of a flat bottom microtitre plate along with 2 µl bacterial suspension containing ~5 x 10<sup>7</sup> cfu/ml in triplicate. The absorbance was determined automatically every hour by a spectrophotometre with an adjustable temperature unit. Bacterial growth rates ( $\mu$ ) were calculated using the following equation:

$$(\mu)h^{-1} = \ln OD_1 - \ln OD_2 / t_2 - t_1$$

Where  $\ln$  is the natural logarithm of a number,  $t_2$  and  $t_1$  are different time points on the growth curve,  $OD_2$  and  $OD_1$  are the cell densities at  $t_2$  and  $t_1$ , respectively (Widdel, 2007).

## **2.24 Preparation of *Streptococcus pneumoniae* cell lysate by sonication**

*Streptococcus pneumoniae* was grown in 10 mL of BHI broth to mid log phase and centrifuged for 10 min at 3000 g. The supernatant was discarded, and the pellet re-suspended in 1.5 ml of PBS. The bacterial suspension was transferred into a microcentrifuge tube and left on ice for 10 min. The sonicator (Soniprep 150) was set to an amplitude of 8 microns. Sonication pulses of 15 sec, every 30 sec, were given to the bacterial suspension. The cells were kept on ice at all times and the sonication pulses were repeated 6 to 8 times. The cell lysate was centrifuged at 18,000 g for 1 minute, and the supernatant filtered through a 25  $\mu$ m Acrodisc Syringe Filter (Pall Corporation) into a fresh microcentrifuge tubes. The filtered lysate was stored at -80°C until further use.

## **2.25 Neuraminidase activity assay**

The total level of neuraminidase activity in *S. pneumoniae* wild type and isogenic mutants were determined by using a quantitative assay utilising the substrate 2-O-(*p*-nitrophenyl)- $\alpha$ -D-N-acetylneuraminic acid (pNPNANA, Sigma). The pNP-NANA is cleaved by neuraminidase with the release of free *p*-nitrophenol (pNP), which can be assayed by its absorbance at 405 nm. Overnight pneumococcal culture was harvested and resuspended in PBS, and the cell suspension was sonicated (amplitude 8 microns) for 15 sec with 45 sec intervals. The samples were then centrifuged at 9460 g for 1 minute at 4 °C and the supernatant was stored on ice. 25  $\mu$ l of clear cell lysate was added in triplicate to each well of a 96 well plate along with 25  $\mu$ l 0.3 mM pNP-

NANA dissolved in a dilution buffer (38.4 ml 100 mM citric-acid phosphate, pH 6.6, 3.12 ml 25 mg/ml BSA, 35.5 ml dH<sub>2</sub>O, 160 µl 10% (w/v) sodium azide). The reaction mixture was incubated statically for 2 h at 37 °C. The reaction was stopped by adding 100 µl of ice cold 0.5 M Na<sub>2</sub>CO<sub>3</sub> (pH 9.6). Absorbance at 405 nm was measured in a BioRad Plate Reader (BioRad, UK). To determine the activity of the neuraminidase enzyme, a standard curve was prepared using known concentrations of *p*-NP.

## **2.26 Hemolytic assay**

Hemolytic activity of pneumococcal strains was tested to determine pneumolysin activity as previously described with some modifications (Pan et al., 2009). Cell lysate was prepared from wild type D39 and its isogenic mutants as described in section 2.24. Serial two fold dilutions (150 µl) of samples were prepared in microtiter plates with PBS (pH7.4) as the diluent. 150 µl 4% sheep red blood cells were added. After incubation for 30 min at 37 °C, the plates were centrifuged at 1000 g for 10 min, and 150 µl of supernatant was transferred to a new plate for spectrophotometric measurements at OD<sub>540nm</sub>. Hemolytic activity units (HU) were defined as the reciprocal of the highest dilution of supernatant inducing at least 50% lysis of the red blood cells.

## **2.27 Glucuronic acid assay**

Capsular polysaccharide (CPS) production was quantified by the method described previously (Lai et al., 2003). Five hundred microliters of bacterial culture was mixed with 100 µl of 1% (v/v) Zwittergent 3-14 detergent (Sigma-Aldrich) in 100 mM citric acid (pH 2.0), and then the mixture was incubated at 50 °C for 20 min. The CPS was precipitated with 1 ml of absolute ethanol. The pellet was dissolved in 200 µl of distilled water, and a 1200 µl of 12.5 mM borax (Sigma-Aldrich) in H<sub>2</sub>SO<sub>4</sub> was added. The mixture was vigorously vortexed, boiled for 5 min, and cooled, and then 20 µl of

0.15% 3-hydroxydiphenol (Sigma-Aldrich) was added. The absorbance of the mixture at 520 nm was measured. The uronic acid content was determined from a standard curve of glucuronic acid (Sigma-Aldrich).

## **2.28 Quantification of protein concentration using Bradford Assay**

The Bradford assay was performed following the protocol provided with the Bradford reagent from Bio-Rad, UK (Bradford, 1976). The procedure started with the preparation of the BSA (Bovine Serum Albumin) standards in nH<sub>2</sub>O. Standards containing 0 to 0.8 mg/ml BSA were prepared and assayed. For this, 10 µl of each standard was pipetted into a 96 well microtitre plate, also 10 µl of each of the unknown proteins were pipetted into the same microtitre plate followed by addition of 200 µl of Bradford reagent in working concentration (diluted 1:5). The reaction was incubated for 5 min at room temperature. Finally the OD was measured at 595 nm in a Bio-Rad microplate reader Model 680. The standards were plotted and the concentration of protein determined by using the equation from the trend line adjusted to the points obtained by plotting the standards.

## **2.29 Denaturing polyacrylamide gel electrophoresis (PAGE)**

A Sodium dodecyl sulphate gel (SDS PAGE) was necessary for the analysis of expression of recombinant proteins. SDS-PAGE was performed using the Mini-PROTEAN 3 cell (Bio-rad). SDS PAGE consisted of two gels, the resolving and stacking gel. The gel was made as described in Table 2.12. The resolving gel was added into the assembled cassettes, up to 1 cm below the level of the comb teeth. After polymerisation of the resolving gel, the stacking gel was added on the top of the resolving gel. Comb was placed in the stacking gel to create the sample wells, and the

gels were left to polymerise before loading the samples.

**Table 2.12. Composition of SDS PAGE**

<b>Components</b>	<b>15% Resolving gel (10mL)</b>	<b>Components</b>	<b>15% stacking gel (10mL)</b>
H <sub>2</sub> O	2.3 mL	H <sub>2</sub> O	2.7 mL
30% Acrylamide	5 mL	30% Acrylamide	670 µl
1.5 M Tris pH8.8	2.5 mL	1.0 M Tris pH 6.8	0.5 mL
10% SDS	100 µl	10% SDS	40 µl
TEMED	4 µl	TEMED	4 µl
10% ammonium persulphate	150 µl	10% ammonium persulphate	50 µl

The samples (20 µl) were mixed with 10 µl of 10X SDS loading buffer (6.25 ml 0.5 M Tris-HCl pH 6.8, 1 g SDS, 10 mM DTT, 50 mg bromophenol blue, 10 ml glycerol and 6.25 ml dH<sub>2</sub>O). The samples were boiled for 5 minute before loading. The gel mould was placed in the electrophoresis tank (Bio-Rad, UK) and 1X SDS running buffer (25 mM Tris-HCl pH 8.3, 192 mM Glycine and 0.1 % (w/v) SDS) was added. The comb was removed carefully, and 10 µl of each sample with 4X NuPAGE LDS (lithium dodecyl sulphate) sample buffer (Invitrogen) and protein marker Precision Plus (Bio-Rad, UK) were loaded into the wells. The gels were set to run at 150 V for 60 min. Once the electrophoresis was completed the gel cassette was gently separated and the gel carefully removed into a large Petri dish. The gels were stained with Coomassie Brilliant Blue staining solution (0.4 g of 0.1% (w/v) coomassie brilliant blue, 40% (v/v) methanol, 10% (v/v) acetic acid and dH<sub>2</sub>O up to a final volume of 400 ml) overnight with gentle shaking. The staining solution was discarded, and the gels

were rinsed with tap water, then placed in the destaining solution (25% (v/v) isopropanol, 10% (v/v) acetic acid and dH<sub>2</sub>O up to 400 ml) for 1-2 hours with gentle shaking. Gels were then photographed.

## **2.30 Protein purification**

### **2.30.1 Purification of Rgg939**

The SPD\_0939 gene was amplified using SPD0939*Xba*I and SPD0939*Bsa*I primers (Table 2.4) and cloned into pE-SUMO expression vector bearing N-terminal His-6 and SUMO tags (LifeSensors, USA). The newly constructed plasmid was transformed into *E. coli* DH5 $\alpha$  for propagation and sequenced using T7 forward primer (TAATACGACTCACTATAGGG). The sequenced recombinant plasmids were transformed into *E. coli* BL21 (DE3) for protein expression. A single colony of *E. coli* BL21 (DE3) carrying the desired construct was inoculated into 10 ml of LB supplemented with 100  $\mu$ g/ml ampicillin, incubated overnight at 37 °C in a shaking incubator at 220 rpm. The overnight culture was diluted into power prime broth (AthenaES, USA) supplemented with 100  $\mu$ g/ml ampicillin to OD<sub>600nm</sub> 0.05, and the culture was incubated at 37 °C in a shaking incubator set at 220 rpm. Expression of Rgg939 was induced at an OD<sub>600nm</sub> of 1.4 with 0.5 mM IPTG (Isopropyl  $\beta$ -D-1-thiogalactopyranoside) for 6 h at 30 °C. The culture was centrifuged at 10,000 x g for 20 min at 4 °C, and the supernatant was discarded. Cell pellets were resuspended in Buffer A (20 mM sodium chloride, pH7.5; 1.5 mM Tris, 5 mM MgCl<sub>2</sub>, with 0.5 mg/ml lysozyme, 0.01 mg/ml DNaseI and protease inhibitor). Cells were disrupted by sonication on ice, and cellular debris was removed by centrifugation at 10,000 xg for 20 min at 4 °C in an Avanti J-E refrigerated centrifuge (Beckman Coulter). The PD-10 columns (GE Healthcare Life Sciences, UK) were used for protein purification. 2 ml of TALON Metal Affinity Resin (Clontech, USA) was added on the polyethylene filter in the PD-10 columns. Then, the column was calibrated with 15 ml of the

binding buffer (20 mM Tris, 150 mM NaCl, pH 7.5). Protein samples were pass through the column. His6-SUMO-Rgg939 was adsorbed to Talon column. The column was washed with 100 ml binding buffer and 15 ml 20 mM imidazole buffer. Afterwards, protein was eluted with 5 ml 500 mM imidazole buffer, concentration was estimated by measuring the OD at 280 nm. The tagged protein was treated with SUMO protease for two hours at room temperature, and re-applied to a talon column to separate His6-SUMO from Rgg939. Pure Rgg939 was obtained in the flow-through and was concentrated using an Amicon Ultracel 10K centrifuge concentrator. Finally, aliquots were frozen in liquid nitrogen and stored at -80 °C.

### **2.30.2 Purification of Rgg144**

The SPD\_0144 was amplified using primers SPD0144N and SPD0144C and cloned into pLEICS-01 expression vector bearing N-terminal His-6 (PROTEX, UK). The recombinant plasmid was transformed into *E. coli* DH5 $\alpha$  for propagation, and sequenced by using T7 forward primer. After this, it was transformed into *E. coli* BL21 (DE3) for protein expression. A single colony of *E. coli* BL21 (DE3) carrying the desired construct was inoculated into 10 ml of LB supplemented with 100  $\mu$ g/ml ampicillin, incubated overnight at 37 °C in a shaking incubator at 220 rpm. The overnight culture was diluted into 1 L Power prime broth (AthenaES, USA) supplemented with 100  $\mu$ g/ml ampicillin to OD<sub>600nm</sub> 0.05. Expression of His6-Rgg144 was induced at an OD<sub>600nm</sub> of 1.4 with 1 mM IPTG overnight at 37 °C in a shaking incubator at 220 rpm.

### **2.30.3 Purification of inclusion bodies**

Overnight bacteria growth culture carrying the recombinant plasmid was centrifuged at 20, 000 rpm at 4 °C for 20 minutes in an Avanti J-E refrigerated centrifuge (Beckman Coulter). Pellet was re-suspended in PBS and spun down at 20, 000 rpm at

4 °C for 20 minutes. Pellet was re-suspended in 50 mL of Lysis buffer 1 (302.85 mg 25 mM Tris pH8.0, 876.6 mg 150 mM NaCl, 500 µl 0.5% (v/v) Triton-x-100, 200 µl 1 mM EDTA, 50 mg 0.5 mg/mL Lysozyme and use nH<sub>2</sub>O top-up to 100 ml) with 2 tablets of protease inhibitor cocktail (Roche), and incubated at room temperature for 15 min with shaking. 50 µl DNase and 250 µl MgCl<sub>2</sub> were added to degrade the chromosomal DNA, and incubated at room temperature for 15-20 min. The sample was topped up to 70 ml with Lysis buffer 1, and was sonicated (section 2.24) followed by centrifugation at 20000 rpm for 20 min at 4 °C. Afterwards the pellet was re-suspend using Buffer 2 (25 mM Tris pH8.0, 0.5 M NaCl, 500 µl 0.5% Triton-x-100, 200 µl 1 mM EDTA and use nH<sub>2</sub>O top-up to 100 ml), briefly sonicated, and centrifuged at 20000 rpm at 20 min at 4 °C. The pellet was then re-suspended in Buffer 3 (25 mM Tris pH8.0, 0.5 M NaCl, 1 mM EDTA and 1 M Urea), briefly sonicated, and centrifuged at 20000 rpm for 20 min at 4 °C. The pellet was re-suspend using 10 ml BugBuster® Master Mix containing Benzonase® Nuclease and rLysozyme Solution (Merck, UK) and spun down for 10 minutes at 13, 000 rpm. Finally, the pellet was re-suspend using 25 mM Tris (pH8.0), and centrifuged for 10 minutes at 13, 000 rpm. The supernatant was discarded, and the inclusion body was kept in -80 °C.

#### **2.30.4 Solubilization and Refolding**

Inclusion body was re-suspend in 10 ml of buffer 5 (25 mM Tris pH8.0, 6M Guanidine Hydrochloride and 5 mM D.T.T) and centrifuged for 5 minutes at 13, 000 rpm. The supernatant was collected, and protein concentration was estimated by Bradford assay (section 2.28). The sample was diluted to approximately to 2 mg/ml using buffer 5, and then diluted into refolding buffer (50 mM MES, 9.6 mM NaCl, 0.4 mM KCl, 2 mM MgCl<sub>2</sub>, 2 mM CaCl<sub>2</sub>, 0.5 M arginine, 0.05% Polyethylene glycol 3550, and 1 mM DTT, pH6.0) at 1:20 by administering the protein solution slowly into the refolding buffer at 4 °C.

### **2.30.5 Dialysis**

To increase the purity of protein, refolded mixture was added into dialysis membrane (Thermo Fisher, UK) and kept in dialysis buffer (25 mM Tris pH7.4 and 150 mM NaCl) after diluting 1 in 10 overnight at 4 °C with constant stirring. The dialysis buffer changed twice, and the dialysed sample was passed through the PD-10 column with TALON Metal Affinity Resin. Finally, the protein was eluted with different concentrations of imidazole elution buffer (20 and 500 mM) and 10 fractions were collected.

### **2.30.6 Gel Filtration**

A Superdex 200 16/600 HiLoad column from GE Healthcare was equilibrated with solution containing 50 mM Tris-HCl, pH 7.5, and 150 mM NaCl. Once equilibrated, the 5 ml sample was injected into the loading loop of the AKTA purifier (GE Healthcare Life Sciences, UK). The protein was run over the column at a flow rate of 1 ml/minute according to the manufacturer's specifications, and the fractions were collected. The fractions were analysed using SDS-PAGE to determine the successful recovery of eluted protein. Finally, the protein fractions were concentrated using an Amicon Ultracel 10K centrifuge concentrator. Aliquots were frozen in liquid nitrogen, and stored at -80 °C.

## **2.31 Electrophoretic mobility shift assay (EMSA)**

### **2.31.1 *In silico* analysis of putative promoters, and DNA probe preparation**

*In silico* analysis was used to determine the putative promoter regions of target genes using the bacterial promoter prediction tool (BPROM, Softberry) (Solovyev and

Salamov, 2011) and the Motif-based sequence analysis tools (The MEME Suite) (Bailey and Elkan, 1994). Based on *in silico* analysis, primers were designed to amplify 100-300 bp upstream regions of each gene containing their putative promoters. One of the primers fluorescently labeled at 5' ends. The PCR was carried out using D39 genomic DNA, PrimeSTAR HS premix, and the primers listed in Table 2.13. The amplified PCR products were analysed by agarose gel electrophoresis and purified by using Wizard SV Gel and PCR Clean-Up System (Promega).

**Table 2.13. Primers used for DNA probes in EMSA**

Primers	Sequence	T <sub>m</sub> (°C)
Shp144EMSAF	CAAGTACAGTATAACACGA	50.2
Shp144EMSAR	AGTAGGAGTATCAGTTGG	51.4
149EMSAF	AATTTTACTTTCCCAATCG	48.0
149EMSAR	CAATAGCTGCCACAGGAGC	58.8
315EMSAF	TTGGTTCGCGGGAAGTCTAC	59.4
315EMSAR	CTTCGCTTCACTTTCTGTG	54.5
1041EMSAF	CAAAACACAGGATATAGTTTC	52.0
1041EMSAR	GAAACGCTTGGTCATTT	47.9
1370EMSAF	GTATCAAACCATAAGAACAGG	54.0
1370EMSAR	CTTCAATGTTTGGACGAATG	53.2
1127EMSAF	CTGCATAATTCTCCTATTC	50.2
1127EMSAR	CATGCGTGTGCCAGTTC CA	58.8
1517EMSAF	GGGGATAAAGAAATTAGAGTC	54.0
1517EMSAR	CCAACAGCACTTATCATTA	52.0
2030EMSAF	ATTGGTTTGTAGAGGTTAAAG	52.8
2030EMSAR	CAAACGCAATCAATCCTACTAG	56.5

Shp939EMSAF	GACTAATTTGTTTGCCCTTACG	56.5
Shp939EMSAR	TTTCTTCATACTCTGACTTCC	54

EMSA was performed as described by (Lasarre et al., 2013) with some modifications. 0-0.8  $\mu$ M recombinant protein was incubated in binding buffer (20 mM HEPES pH7.9, 20 mM KCl, 5 mM MgCl<sub>2</sub>, 0.2 mM EDTA, pH8.0, 0.5 mM dTT, 0.5 mM CaCl<sub>2</sub>, 12% glycerol) with 10 nM DNA probe for 30 minutes at room temperature. Synthetic peptide was added into the mixture at final concentration 2  $\mu$ M for 20 min when required.

Non-denaturing polyacrylamide gels (8%, v/v) were used to separate DNA-protein complexes. Each gel was composed of 4.1 ml dH<sub>2</sub>O, 0.3 ml 10X TB buffer (Tris base 89 mM + Boric acid 89 mM), 1.6 ml 30% (v/v) acrylamide, 75  $\mu$ l 10% (w/v) ammonium persulfate (APS) and 5  $\mu$ l TEMED. The gels were pre-run (without samples) for 20 min at 200 V in 0.5X TB buffer to remove all traces of APS. The gels were kept cool using iced 0.5X TB buffer. Following pre-run, the samples (20  $\mu$ l) were loaded into the wells and run for additional 45 min at 200 V. After electrophoresis, the gels were carefully removed, and the gel shifts were detected by fluorescence imaging using a TYPHOON Trio+scanner (GE Healthcare Life Sciences, UK) with a 526 nm short-pass wavelength filter.

## **2.32 Construction of *lacZ*-fusions**

Transcriptional *lacZ*-fusions were constructed according to the method described previously (Halfmann et al., 2007). An integrative plasmid pPP2, which derived from pBR322, was selected for constructing the *lacZ* reporter system (Halfmann et al., 2007). *In silico* analysis was used to determine the putative promoter regions of target genes using the bacterial promoter prediction tool (BPROM, Softberry) (Solovyev and Salamov, 2011). The putative promoter regions were amplified using primers

incorporated with the *SphI* and *BamHI* recognition sites (Table 2.4). The PCR was carried out using PrimeSTAR HS premix and D39 chromosome DNA. The amplified PCR products were analysed on agarose gel.

The plasmid and inserts were digested using the *SphI* and *BamHI* as previously described in section 2.10. T4 DNA Ligase was used to ligate the each putative promoter region into the pPP2 plasmid. Afterwards, the ligation reaction was transformed into *E. coli* DH5. The successful cloning was confirmed by PCR using seqF and seqR, and by sequencing (Table 2.4).

The recombinant reporter plasmids were transformed into different pneumococcal strains. The incorporation of reporter construct in pneumococcal genome is mediated through homologous recombination. The transformants were selected on blood agar plates containing 3 µg/ml of tetracycline. Colony PCR was used for the incorporation of reporter construct using the same primers as above.

### **2.33 β-galactosidase activity assay**

For β-galactosidase activity assay, a previously published protocol was used (Miller, 1972). The activity in the cell lysates was measured by monitoring the development of yellow colour, which occurs due to the hydrolysis of the chromogenic substrate *O*-Nitrophenyl-β-galactopyranoside (ONPG) by β-galactosidase. For this assay 10 ml CDM supplemented with specific sugar with or without additions of synthetic peptides was inoculated with *S. pneumoniae* reporter strains until late-exponential phase. The synthetic peptides used in this study were shown in table 2.14. At this stage, 3 ml of bacterial culture was taken, and centrifuged at 3500 rpm (Hettich MIKRO 22R, Germany) for 15 minutes at 4 °C. The supernatant was removed and the pellet was re-suspend using 3 ml of chilled Z buffer (0.80 g Na<sub>2</sub>HPO<sub>4</sub>•7H<sub>2</sub>O, 0.28 g NaH<sub>2</sub>PO<sub>4</sub>•H<sub>2</sub>O, 0.5 ml 1 M KCl, 0.05 ml 1 M MgSO<sub>4</sub>, 0.175 ml β-mercaptoethanol (BME), 40 ml dH<sub>2</sub>O, pH 7.0), and the OD<sub>600</sub> was measured. Z buffer was used as a

blank. The remaining 2 ml culture was diluted into 1:1 ratio with Z buffer. For higher activity 1:9 ratio Z buffer: cell suspension was also used. Subsequently, 100 µl of chloroform and 50 µl freshly prepared 0.1% (w/v) SDS was added. The samples were mixed and incubated for 10 minutes at 28°C water bath.

After the cells lysis, 200 µl of ONPG (ortho-Nitrophenyl-β-galactoside) (4 mg/ml stock) was added into the samples, vortexed and incubated at 30 °C. The reaction was stopped after sufficient yellow colour was observed by adding 500 µl of 1 M Na<sub>2</sub>CO<sub>3</sub>. All the reactions were stopped after 90 minutes if sufficient yellow colour was not observed. Samples were then centrifuged at 14, 000 rpm (Microfuge, Sigma) for 15 minutes. 1 ml of the supernatant was used to measure the OD<sub>420nm</sub>. Finally the specific β-galactosidase activity was determined by the following equation:

Miller Units=1000\*(OD<sub>420</sub>)/ (T\*V\*OD<sub>600</sub>) where T= time of the reaction in minutes (duration of the enzymatic reaction took place), V= Volume of culture used in the assay in mls, Miller unit represents nmol of *o*-nitrophenol/min/ml of cells/OD<sub>600</sub>.

**Table 2.14. Synthetic peptides used in this study**

Peptide	Sequence
SHP144-C8	VIPFLTNL
SHP144-C10	VIVIPFLTNL
SHP144-C11	WVIVIPFLTNL
SHP144-C12	EWVIVIPFLTNL
SHP144-C13	SEWVIVIPFLTNL
SHP144-C13REV	LNTLFPVIVWES
SHP144-C8REV	LNTLFPVIV
SHP144-C14	ISEWVIVIPFLTNL
SHP144-C15	LISEWVIVIPFLTNL

SHP939-C8	DIIIVGG
SHP939C8REV	GGVIIIID
SHP939-C9	MDIIIVGG

## **2.34 *In vivo* virulence studies**

Virulence studies were conducted using a murine model as described previously (Terra et al., 2016). Virulence studies aimed at determining survival time after intranasal infection, bacterial numbers in blood, and the ability of pneumococcal strains to colonise the respiratory tract.

### **2.34.1 Preparation of bacterial inoculum**

To prepare the standard inoculum to be used for *in vivo* virulence studies, pneumococcal strains were grown overnight at 37 °C until OD<sub>500nm</sub> had reached 1.4-1.6. The cultures were then harvested at 3000 rpm for 15 min and the supernatant was discarded. The pellet was resuspended in 1 ml of BHI serum broth (80% (v/v) BHI and 20% (v/v) filter sterilised fetal calf serum). 700 µl of resuspended culture was added to fresh BHI serum broth to bring the OD<sub>500nm</sub> to 0.7, and incubated until OD<sub>500nm</sub> reached 1.6. Then, 500 µl aliquots were made, and stored at -80 °C until needed.

### **2.34.2 Pneumonia model**

Female CD1 mice (Charles River, UK) at 8-10 weeks of age were used in this study. Mice were lightly anaesthetised with 2.5% (v/v) Isoflurane (Isocare, UK) over oxygen (1.4 to 1.6 litres/min), in an anaesthetic box. 50 µl of PBS containing approximately 2 x 10<sup>6</sup> CFU of *S. pneumoniae* was administered into the nostrils of the mouse. The mouse was then laid on its back inside the cage to allow for the inoculum to reach the

lungs. The inoculum dose was confirmed by colony counting on blood agar after administration of the dose. Mice were monitored for signs of illness for 7 days (Progressively starry coat, hunched appearance and lethargy) (Morton and Griffiths, 1985). Mice were culled when they reached a very lethargic state, the time to this point was defined as ‘survival time’.

To determine the development of bacteremia in each mouse, approximately 20  $\mu$ l of venous blood was taken by tail bleeding from mouse at 24 and 48 h post infection. Viable counts in blood were determined by serial dilution in sterile PBS and plating on the blood agar plates. Data were analyzed by analysis of variance followed by the Bonferroni posttest. *P* values of <0.05 were considered statistically significant.

### **2.34.3 Colonisation experiments**

CD1 mice were administered with  $5 \times 10^5$  CFU/mouse of *S. pneumoniae* in 20  $\mu$ l PBS. Decreasing the amount of inoculum was required to avoid pneumococci disseminate into the lower respiratory tract. The pneumococcal counts in inoculum were determined by plating on blood agar plates after infection. The colonisation of pneumococci in the nasopharynx was determined as described previously (Richards et al., 2010). Mice were deeply anaesthetised with 5% (v/v) isoflurane over oxygen, and the mice were subsequently killed by cervical dislocation at 0 and 7 days post-infection. Nasopharyngeal tissue was collected and transferred into 5 ml of sterile PBS. Tissue samples were weighed, and then homogenised in an ultra-turrax homogeniser (IKA-Werk, Germany). Viable counts in homogenates were determined by serial dilution in sterile PBS, and plating on the blood agar plates.

## **2.35 Statistical analysis**

GraphPad prism version 7 (Graphpad, California, USA) was used to analyse all data. The experimental results were expressed as mean  $\pm$  standard error of the mean (SEM). One- and two-way analysis of variance (ANOVA) followed by Tukey multiple comparison test were used to compare the groups for growth studies, H<sub>2</sub>O<sub>2</sub> and paraquat killing assay, and  $\beta$ -galactosidase assay. The Mann Whitney test was used for *in vivo* survival assay whereas one-way ANOVA followed by Tukey's multiple comparisons test was used to compare the groups for bacteraemia development and colonisation experiment. Less than 0.05 of  $p$  value was regarded statistically significant. Significance was defined as \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$  and \*\*\*\*  $p < 0.0001$ .

## **Chapter III: Results**

## **Section A. Construction of isogenic mutants and genetically complemented strains**

The *Streptococcus pneumoniae* D39 serotype 2 have 5 *rgg*-like genes, *rgg144* (SPD\_0144), *rgg939* (SPD\_0939), *rgg999* (SPD\_0999), *rgg1518* (SPD\_1518), and *rgg1952* (SPD\_1952), which have significant homology (between 23 to 28 % identity and 42 to 48 % similarity at the amino acid sequence level) to what is considered as the Rgg prototype the Rgg from *Streptococcus gordonii* (SGO0496). Off these five *rggs*, *rgg144* and *rgg939* are associated with two unannotated ORFs, predicted to code for short hydrophobic peptides (Shp). These *shp* genes are proximal to *rgg144* and *rgg939*, and are designated as *shp144* and *shp939*, respectively, in this study. To determine the biological roles of *rgg144* and *rgg939*, as well as to test the hypothesis that *rgg144/shp144* and *rgg939/shp939* act as quorum-sensing mediators to regulate gene expression in response to pheromones, single and double mutations were made to create  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta shp144$ ,  $\Delta shp939$ ,  $\Delta rgg144/shp144$ ,  $\Delta rgg939/shp939$ ,  $\Delta rgg144/939$  and  $\Delta rgg144/939/shp144/939$  by allelic replacement mutagenesis. The mutant strains were used to determine the phenotypic contributions of Rgg/Shp systems in pneumococcal biology by growth studies, microarray analysis, paraquat killing assay, and *in vivo* tests. Transcriptional reporter systems were also constructed in order to determine the regulatory interactions between individual components of Rgg/Shp systems, and to assess the environmental conditions important for the induction of each Rgg/Shp.

### **3.1 Mutagenesis by gene splicing overlap extension PCR (SOEing PCR)**

Previously scientists described a novel approach to introduce targeted mutagenesis in bacteria using a method called DNA splicing by overlap extension, which is based on

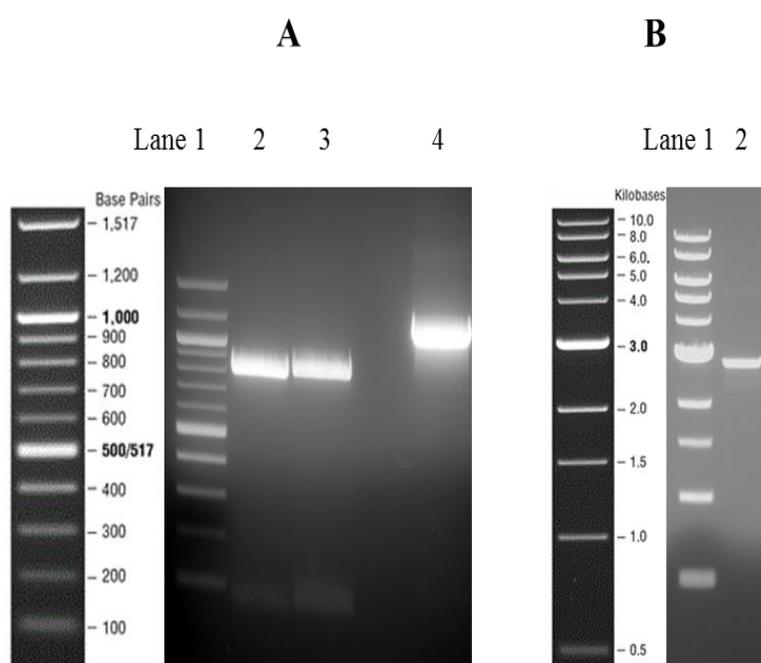
the polymerase chain reaction (SOEing PCR) (Horton et al., 1989). In contrast to plasmid based mutagenesis systems, SOEing PCR is a rapid and efficient strategy to generate desired modification at any position of a DNA molecule *in vitro* without a need for cloning and plasmid propagation (Akerley et al., 2002). *S. pneumoniae* is naturally competent, which refers to a physiological state that makes it possible to take up and incorporate extracellular DNA into pneumococcal genome, makes the SOEing method ideal to introduce mutations in the pneumococcus (Seitz and Blokesch, 2014).

To introduce the mutation, initially two separate PCR reactions were set up. While the first reaction used LF/LR, RF/RR and SpecF/SpecR or KanF/KanR primers, the second one utilised LF/RR primers. These primer sets amplified the left and right flanks of the target genes, omitting the amplification of the target gene, by virtue of primers XX-LR and XX-RF spanning at the either side of target. Left and right flank regions were amplified using D39 genomic DNA as a template, while pDL278 or pCEP were used to amplify *aadA*, for spectinomycin resistance, or *aph*, for kanamycin resistance, respectively. As part of the primers XX-LR and XX-RF overlapped with the amplicons representing *aadA* or *aph*, these fragments could be joined together in a second round of PCR by using the mixture of three fragments as template from the first round PCR. Finally, the fused PCR products were transformed into *S. pneumoniae* D39. The homologous recombination between the flanks leads to the deletion of the target gene and insertion of antibiotic cassette.

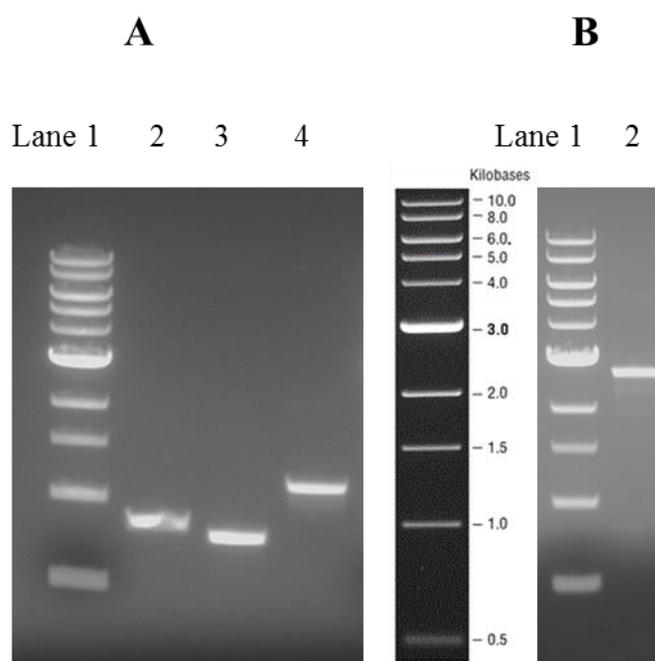
### **3.1.1 Construction of $\Delta rgg144$ , $\Delta rgg144/shp144$ , and $\Delta shp144$**

*S. pneumoniae* D39, a virulent type 2 laboratory strain, was used for the construction of  $\Delta rgg144$ ,  $\Delta rgg144/shp144$ , and  $\Delta shp144$ . To mutate *rgg144*, *shp144*, or both the genetic regions surrounding *rgg144*, and *shp144* were amplified from D39 DNA, and the gene conferring resistance to spectinomycin resistance cassette was amplified

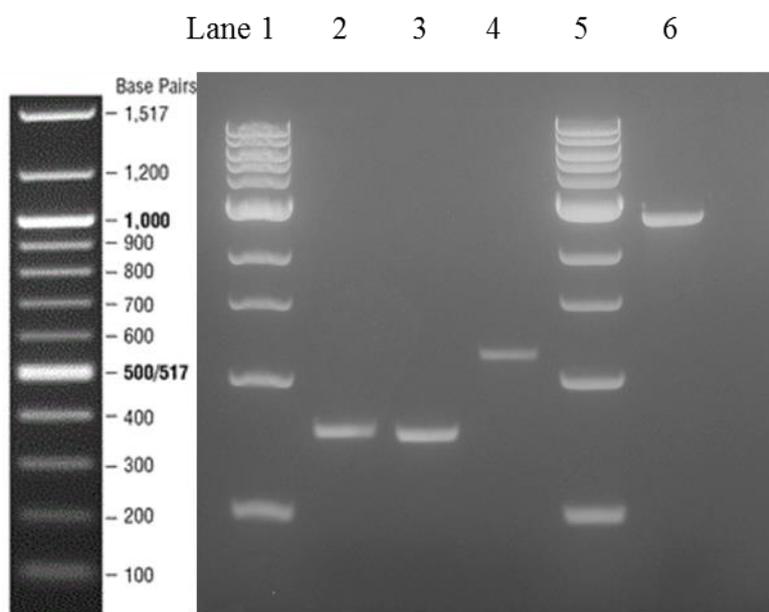
from pDL278 (Yesilkaya, 1999), which are shown in Figure 3.1A, 3.2A and 3.3. The amplicons were subjected to gel electrophoresis and the corresponding bands were extracted from the gel in order to obtain the right sized products. Subsequently, the purified products were mixed and joined together with the outer primers by PCR (Figure 3.1B, 3.2B and 3.3). After purification from the agarose gel, the *in vitro* mutagenized constructs were transformed individually into D39. Spectinomycin resistant colonies were selected for mutation confirmation by PCR and sequencing.



**Figure 3.1. Agarose gel electrophoresis analysis showing the amplified antibiotic resistance gene, the flanking regions of *rgg144*, and the fused SOEing fragment.** Lane 1 in both panels is 100 bp DNA ladder (New England Biolab, UK) **A.**, lane 2 is left flank of the *rgg144*, which is approximately 812 bp, lane 3 is the right flank of the *rgg144* gene, which is approximately 857 bp, lane 4 is spectinomycin resistance cassette which is approximately 1158 bp. **B.** lane 2 is the fused SOEing fragment, which is approximately 2827 bp.



**Figure 3.2. Agarose gel electrophoresis analysis showing the amplified antibiotic resistance gene, the flanking regions of *rgg144/shp144*, and the fused SOEing fragment.** Lane 1 is 100 bp DNA ladder (New England Biolab, UK) in both panels. **A.** lane 2 is left flank of the *rgg/shp144*, which is approximately 812 bp, lane 3 is the right flank of the *rgg144/shp144* gene, which is approximately 749 bp, lane 4 is spectinomycin cassette. **B.** lane 2 is the fused SOEing fragment, which is approximately 2724 bp.

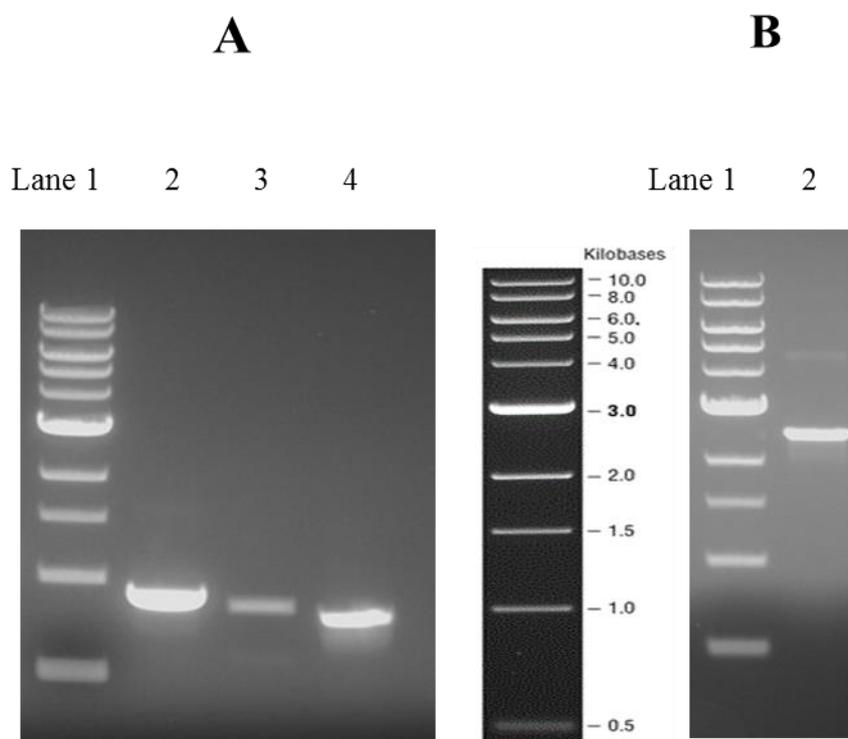


**Figure 3.3. Agarose gel electrophoresis analysis showing the amplified antibiotic resistance genes, the flanking regions of *shp144*, and the fused SOEing fragment.** Lane 1 and 5 is 1 kb DNA ladder (New England Biolab, UK), lane 2 is left flank of the *shp144*, which is approximately 760 bp, lane 3 is the right flank of the *shp144* gene, which is approximately 749 bp, lane 4 is spectinomycin cassette, which is approximately 1158 bp. Lane 6 is the fused SOEing fragment, which is approximately 2667 bp.

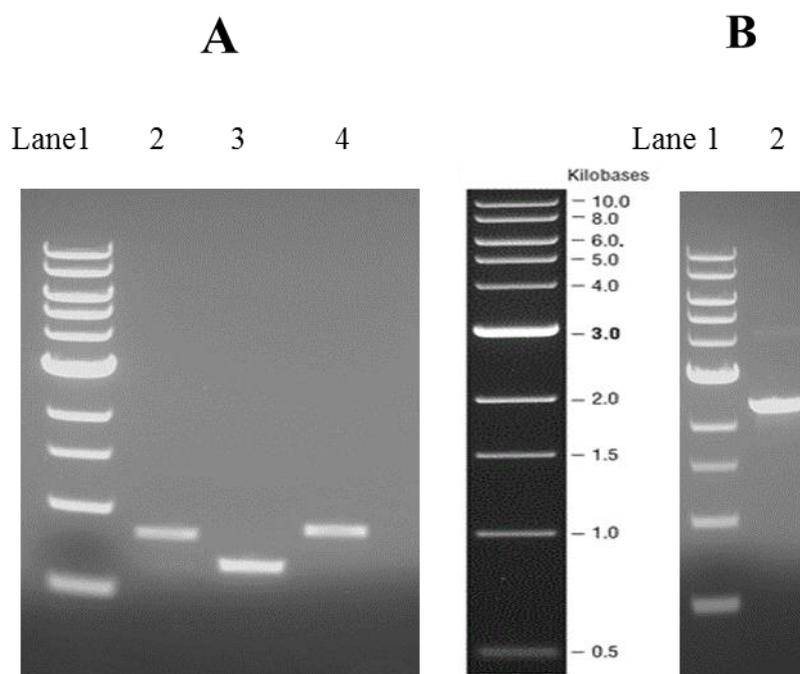
### **3.1.2 Construction of $\Delta rgg939$ , $\Delta rgg939/shp939$ , and $\Delta shp939$**

To construct  $\Delta rgg939$ ,  $\Delta rgg939/shp939$ , and  $\Delta shp939$ , the genetic regions surrounding *rgg939*, *rgg939/shp939* and *shp939* genes, as well as the gene conferring resistance to kanamycin resistance cassette were amplified from the D39 DNA and pCEP, respectively. The amplicons were subjected to gel electrophoresis and the corresponding bands were extracted from the gel in order to obtain the right products (Figure 3.4A, 3.5A and 3.6). Subsequently, the purified products were mixed and joined together with the outer primers (Figure 3.4B, 3.5B and 3.6). After purification from the agarose gel, the *in vitro* mutagenised constructs were transformed individually into D39. Kanamycin resistant colonies were selected for mutation

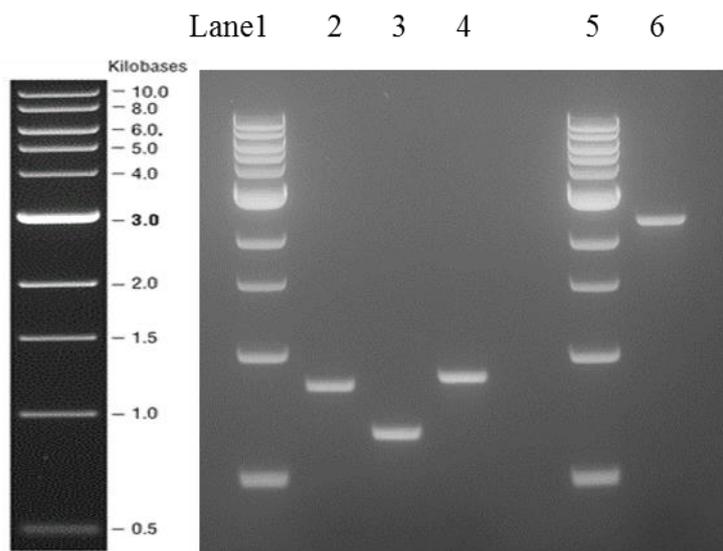
confirmation by PCR and sequencing.



**Figure 3.4. Agarose gel electrophoresis analysis showing the amplified antibiotic resistance gene, the flanking regions of *rgg939*, and the fused SOEing fragment. .** Lane1 is 1kb DNA ladder (New England Biolab, UK) in both panels. **A**, lane 2 is kanamycin cassette which is approximately 895 bp, lane 3 is the left flank of *rgg939*, which is approximately 848 bp. lane 4 contains the amplicons for the right flank of *rgg939*, which is approximately 806 bp. **B**. lane 2 represents the amplicons for fused SOEing fragment, which is approximately 2549 bp.



**Figure 3.5. Agarose gel electrophoresis analysis showing the amplified antibiotic resistance gene, the flanking regions *rgg/shp939*, and the fused SOEing fragment.** **A.** Lane 1 contains 1 kb DNA ladder (New England Biolab, UK), lane 2 represents the amplicons kanamycin cassette, which is approximately 895 bp, lane 3 is for the PCR products for the left flank of *rgg939/shp939*, which is approximately 848 bp, lane 4 is right flank *rgg939/shp939*, which is approximately 626 bp. **B.** Lane 1 is 1 kb DNA ladder (New England Biolab, UK), lane 2 is the fused SOEing fragment, which is approximately 2369 bp.



**Figure 3.6.** Agarose gel electrophoresis analysis showing the amplified antibiotic resistance gene, the flanking regions of *shp939*, and the fused SOEing fragment. Lane 1 and 5 is 1 kb DNA ladder (New England Biolab, UK), lane 2 is the left flank of *shp939*, which is approximately 836 bp, lane 3 is right flank *shp939*, approximately 626 bp. lane 4 is kanamycin cassette, which is approximately 895 bp. Lane 6 contains PCR products for the fused SOEing fragment, which is approximately 2357 bp.

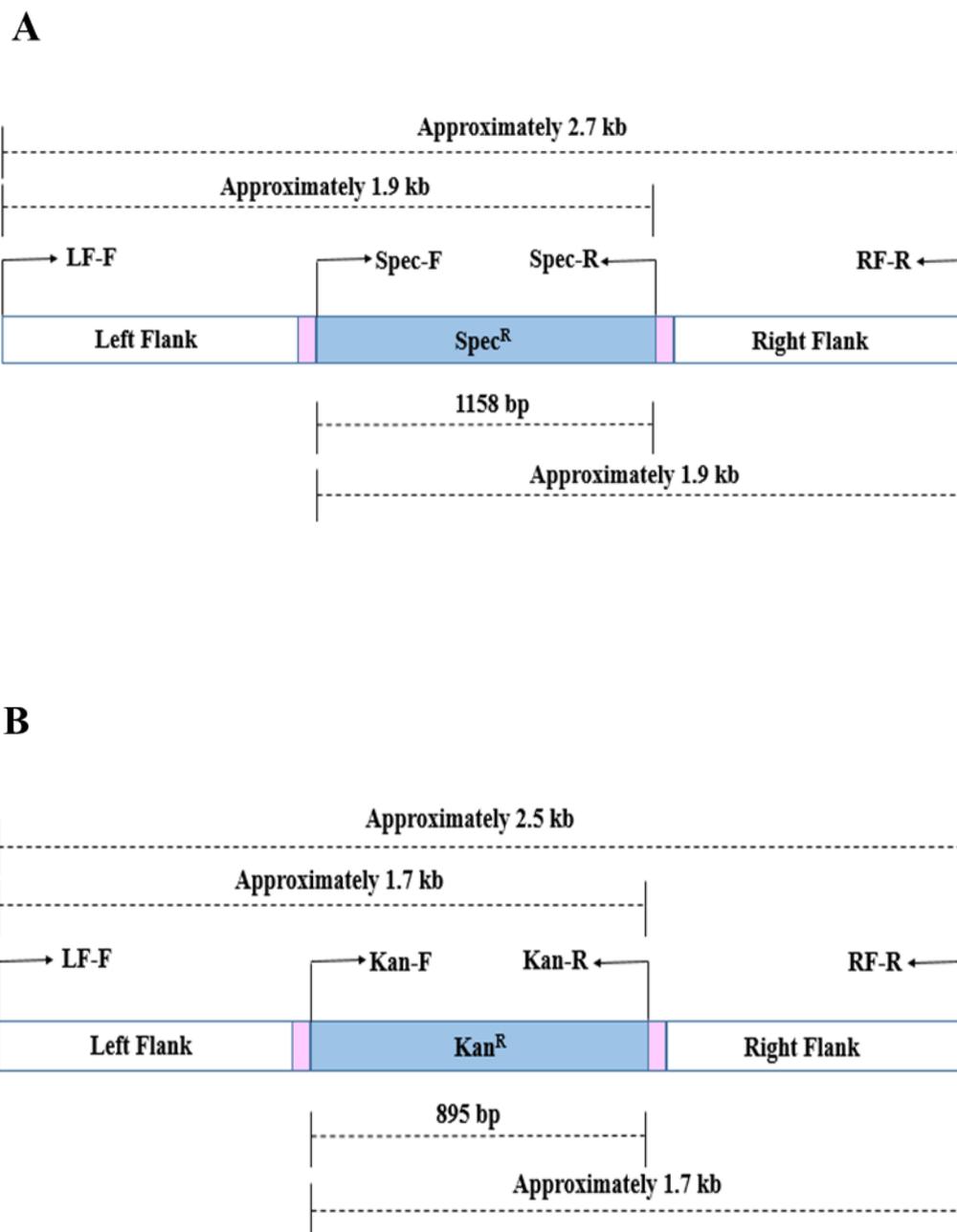
### **3.1.3 Construction of $\Delta rgg144/939$ and $\Delta rgg/shp144/939$**

*Streptococcus pneumoniae*  $\Delta rgg144$  and  $\Delta rgg144/shp144$ , were used for the construction of *S. pneumoniae*  $\Delta rgg144/939$  and  $\Delta rgg/shp144/939$  double mutants, respectively. To construct  $\Delta rgg144/939$ , the genetic region containing *rgg939* gene and kanamycin resistant gene were successfully amplified, and the joined product was transformed into  $\Delta rgg144$ . Similarly, for  $\Delta rgg/shp144/939$  mutation, the mutated locus from  $\Delta rgg939/shp939$  was amplified and transformed into the  $\Delta rgg144/shp144$  mutant. The kanamycin resistance gene replaces the chromosomal copy of the *rgg939* or *rgg939/shp939* gene through homologous recombination, thereby generate a double mutant. For each mutation event, a selected spectinomycin and kanamycin resistant transformant was tested by PCR to show the successful mutagenesis of both targets in the same strain background.

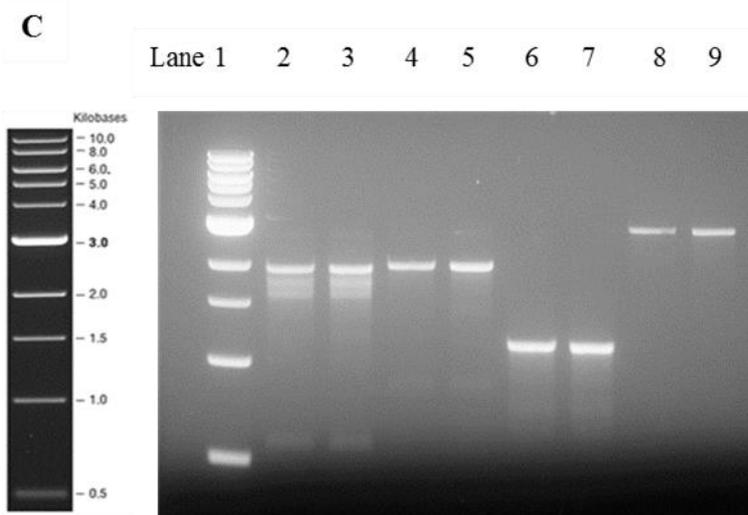
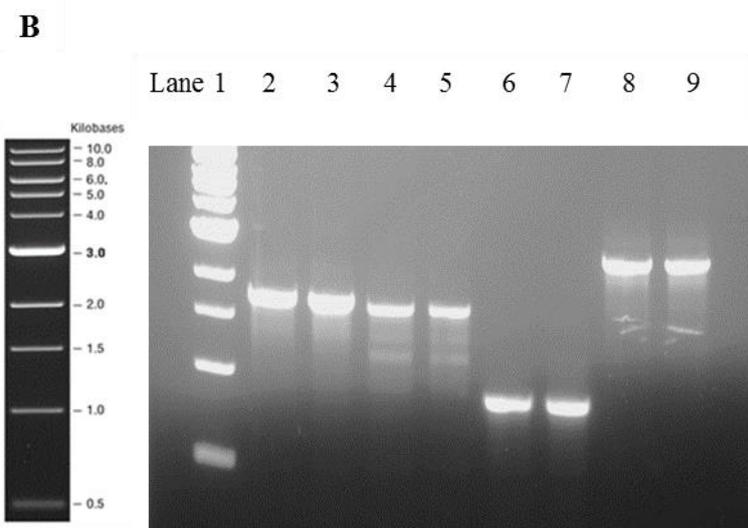
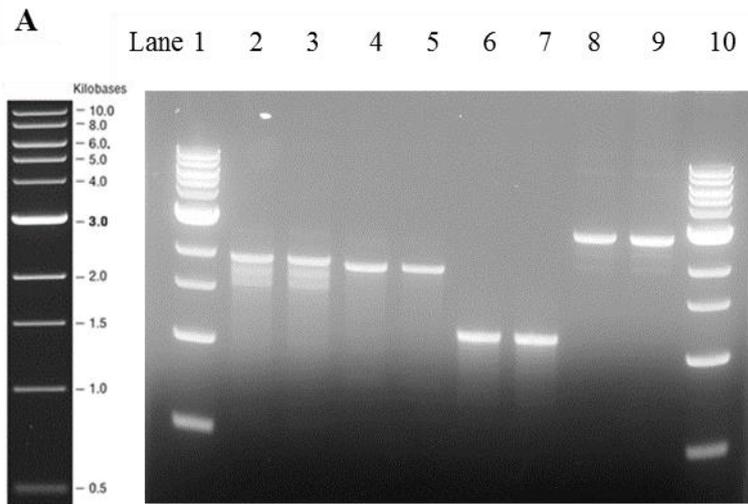
### **3.1.4 Confirmation of mutagenesis by PCR**

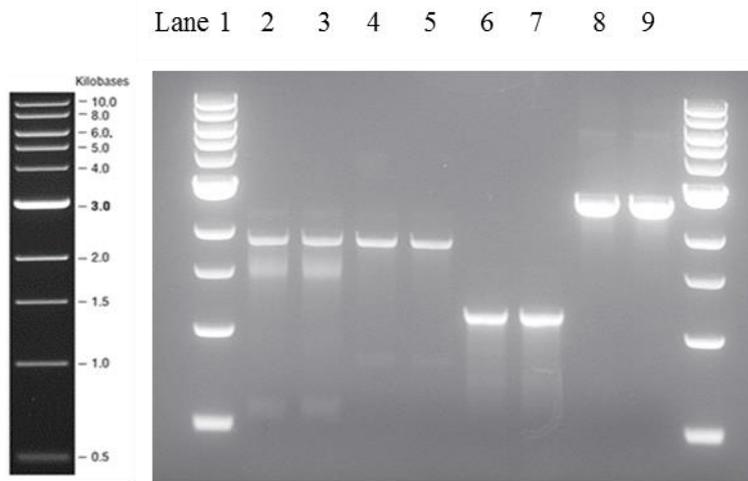
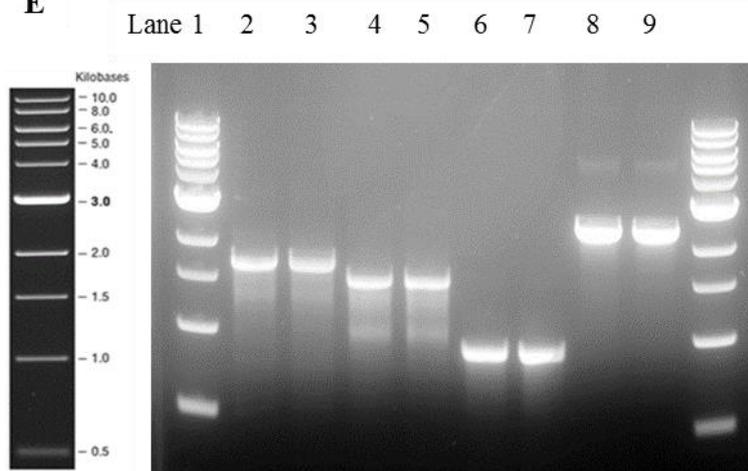
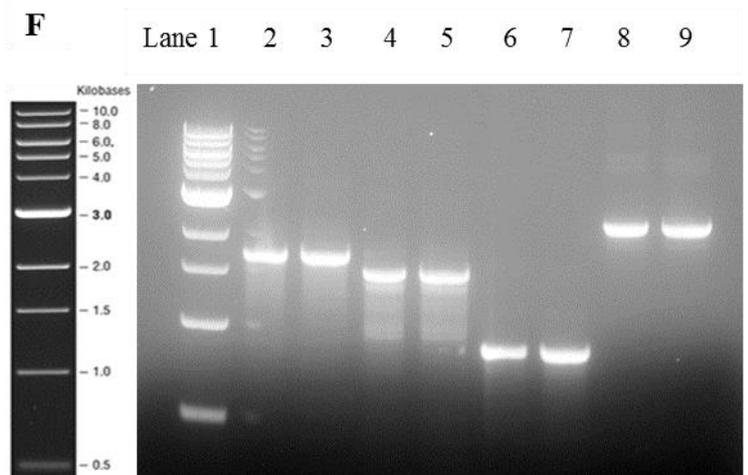
Targeted gene mutagenesis was confirmed by PCR, and the confirmation strategy is illustrated in (Figure 3.7A and B). The Spec<sup>R</sup> or Kan<sup>R</sup> transformants potentially carrying the

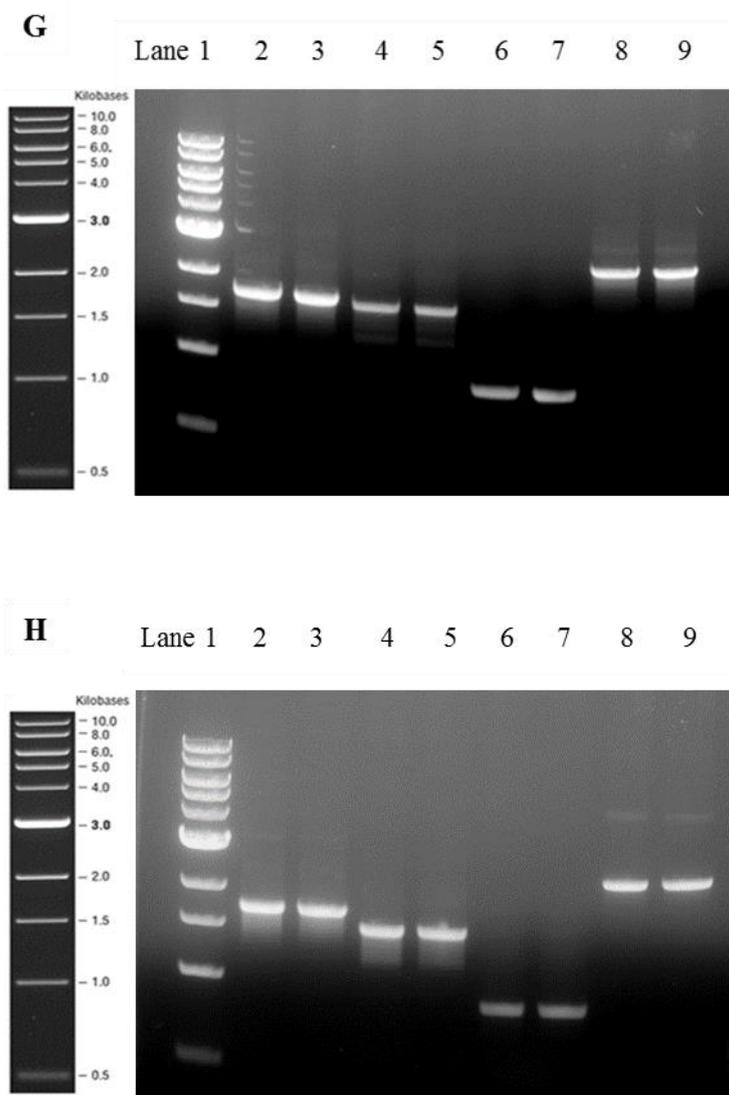
mutated target genes, were grown overnight in BHI at 37 °C. The genomic DNA was extracted from these colonies and used as templates for PCR amplification. Different primer combinations were used to confirm gene replacement for each target. The LF-F and RF-R primers were used to amplify both wild type and mutant DNA. While the DNA template from wild type amplified the genes with their flanking regions, a PCR fragment (left flank of gene 800 bp + Spec<sup>R</sup> gene 1158bp + right flank of gene 800 bp) was amplified from the mutant DNA of each target. In addition, the primers LF- F and Spec-R were used to amplify the left flanking region of each target gene and the Spec<sup>R</sup> cassette whereas the primers Spec-F and RF- R were used to amplify the Spec<sup>R</sup> cassette and the right flanking region of each target gene, resulting PCR products with sizes of 1958 bp. Furthermore, the primers Spec-F and/ Spec-R were used to amplify the Spec<sup>R</sup> cassette (1158 bp) from the relevant mutants (Figure 3.7A). The mutants constructed with kanamycin resistance marker, *rgg939* and *rgg/shp939* and double mutants, were essentially analysed using the same strategy as illustrated in Figure 3.7B. The amplified PCR products were analysed by agarose gel electrophoresis (Figure 3.8).



**Figure 3.7. Illustration showing the PCR strategy to confirm gene replacement in mutant strains.** Confirmation of single mutation using the spectinomycin resistance gene (*aadA*) (**A**) and the kanamycin resistance gene (*aph*) (**B**). LF indicates left flank, RF is for right flank. The generic primers used for confirmation strategy and the expected amplicon sizes have been indicated. The sequence of each specific primer is shown in Table 2.4.



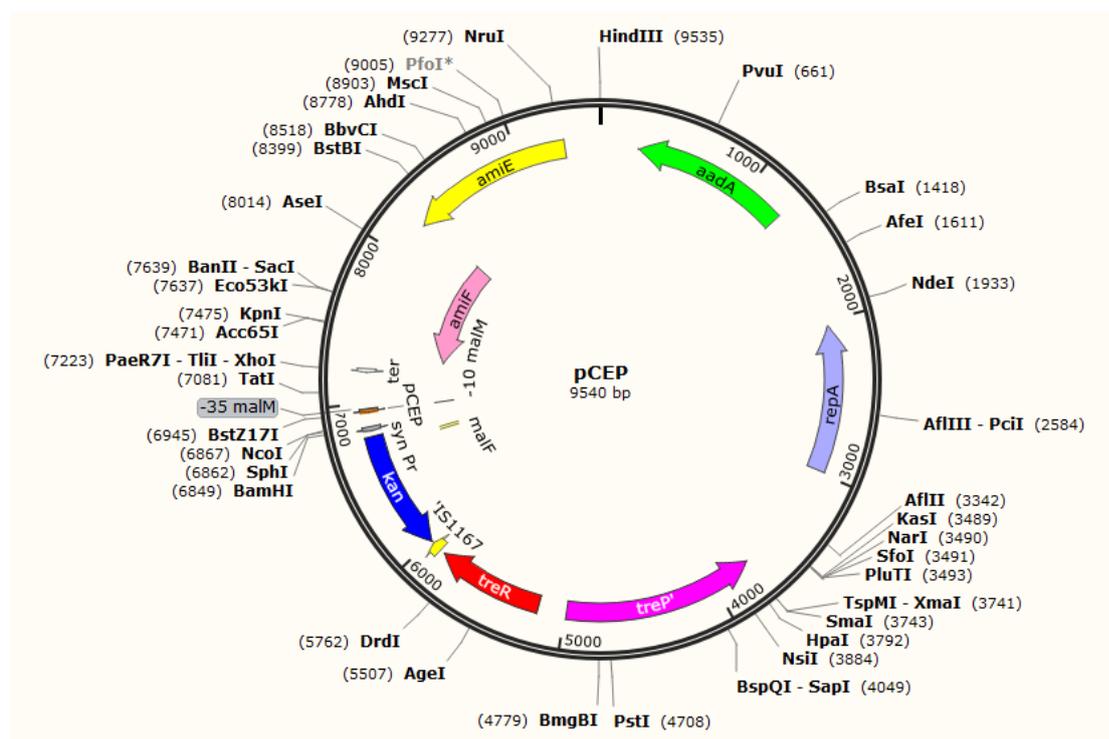
**D****E****F**



**Figure 3.8.** Agarose gel electrophoresis analysis confirming successful construction of (A)  $\Delta$ *rgg144*, (B)  $\Delta$ *rgg144/shp144*, (C)  $\Delta$ *shp144* (D)  $\Delta$ *rgg939*, (E)  $\Delta$ *rgg939/shp939*, (F)  $\Delta$ *shp939*, (G)  $\Delta$ *rgg144/939*, (H)  $\Delta$ *rgg144/939/shp144/939*. In A, B and C, lane 1 is 1 kb DNA ladder (New England Biolab, UK), lane 2 and 3 contain the amplicons for the left flank plus spectinomycin cassette. Lane 4 and 5 are for the right flank plus spectinomycin cassette. Lane 6 and 7 represent the amplicons for the spectinomycin cassette. In D, E, F, G and H, lane 1 is 1 kb DNA ladder (New England Biolab, UK), lane 2 and 3 is the left flank plus kanamycin cassette. Lane 4 and 5 is the right flank plus kanamycin cassette. Lane 6 and 7 have amplicons for the kanamycin cassette. Lane 8 and 9 is the left flank + spectinomycin or kanamycin cassette + right flank.

## **3.2 Complementation of mutant strains**

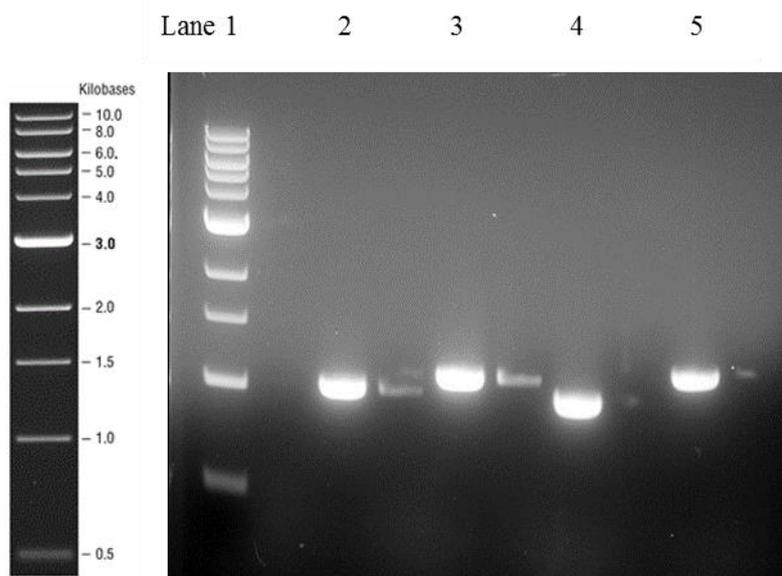
Insertion-deletion mutations may cause polar effect, which means that mutation can affect the expression of downstream genes. Any gene located downstream of the mutation without its own promoter, in an operon organisation, will not be expressed due to termination of transcription (Shapiro, 1969). If the target gene is the first gene of a predicted operon, then the possibility of a polar effect caused by mutation is higher (Guiral et al., 2006). Hence, complementation of  $\Delta rgg144$ ,  $\Delta rgg144/shp144$ ,  $\Delta rgg939$  and  $\Delta rgg939/shp939$  were done to rule out the possibility of polar effect of mutations. Plasmid pCEP is the most popular vector for genetic complementation in *S. pneumoniae*. Plasmid pCEP is derived from pR410, a 9540 bp single copy plasmid, which can replicate in *E. coli*, and illustrated in Figure 3.9. A maltose-inducible promoter and kanamycin resistant gene are separated by multiple cloning sites. Therefore, the chromosomal expression platform (CEP) is flanked by more than 2 kb DNA homologous to pneumococcal genome, which mediates homologous recombination in *S. pneumoniae*. By using pCEP, the intact copy of the gene will be driven by a maltose-inducible promoter and inserted into downstream of the *amiA* operon, which is a transcriptionally silent site without any cellular functions, avoiding disrupting the cell physiology (Alioing et al., 1996, Guiral et al., 2006). The expression of cloned genes can also be driven by their native promoter in this system.



**Figure 3.9. Genetic map of pCEP.** *treR*: Trehalose operon repressor, *amiF* and *amiE*: oligopeptide ABC transporters, *treR* and *treP*: trehalose-utilization system, *kan*: kanamycin resistance cassette. The multiple cloning site MCS consists of *BstZ17I*, *NcoI*, *SphI* and *BamHI* restriction sites. *malR*: maltosaccharide-inducible promoter. *aadA*: the spectinomycin resistance gene, *repA*: required for autonomous plasmid replication.

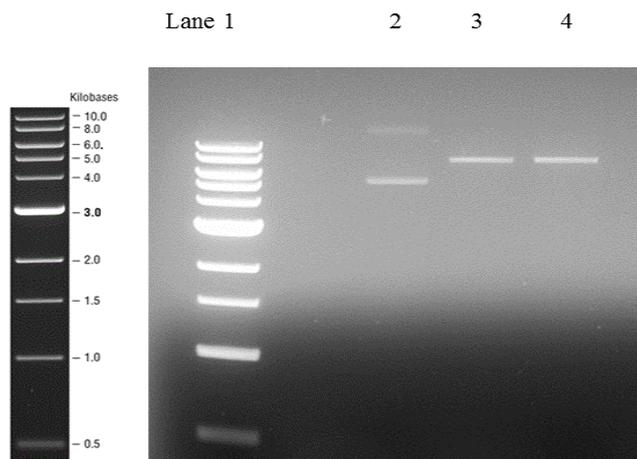
### 3.2.1 Amplification of *rgg* genes for complementation

The entire gene sequence of *rgg144*, *rgg144/shp144*, *rgg939* and *rgg939/shp939*, as well as approximately 100 bp upstream of the respective genes predicted to contain putative promoter regions, was amplified with primers SPD939*BamHI*/SPD939*NcoI*, SHP939*BamHI*/SPD939*NcoI*, SPD144*BamHI*/SPD144*NcoI* and SHP144*BamHI*/SPD144*NcoI*, which introduces *NcoI* and *BamHI* sites into the 5' and 3' ends of the amplicons, respectively. The amplified PCR products were purified using a Wizard SV Gel and PCR Clean-Up System from Promega (UK), and the successful amplification was demonstrated by agarose gel electrophoresis (Figure 3.10).

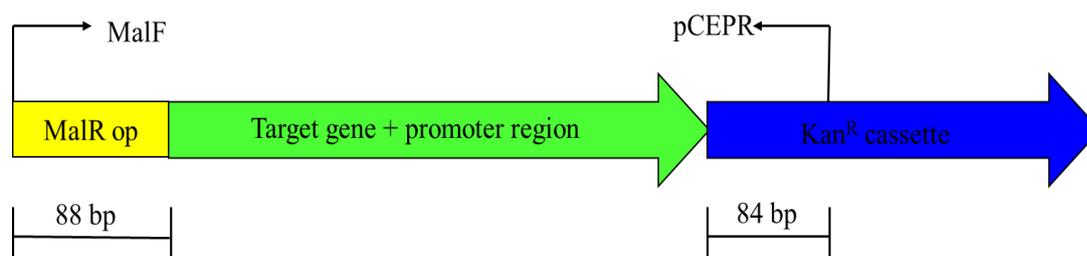


**Figure 3.10. Agarose gel electrophoresis analysis of amplicons for genetic complementation of mutants.** Lane 1 is 1 kb DNA marker (New England Biolabs), Lane 2-5 are amplicons representing the putative promoter and the coding regions of *rgg939* (1012 bp), *rgg939/shp939* (1113 bp), *rgg144* (971 bp) and *rgg144/shp144* (1012 bp).

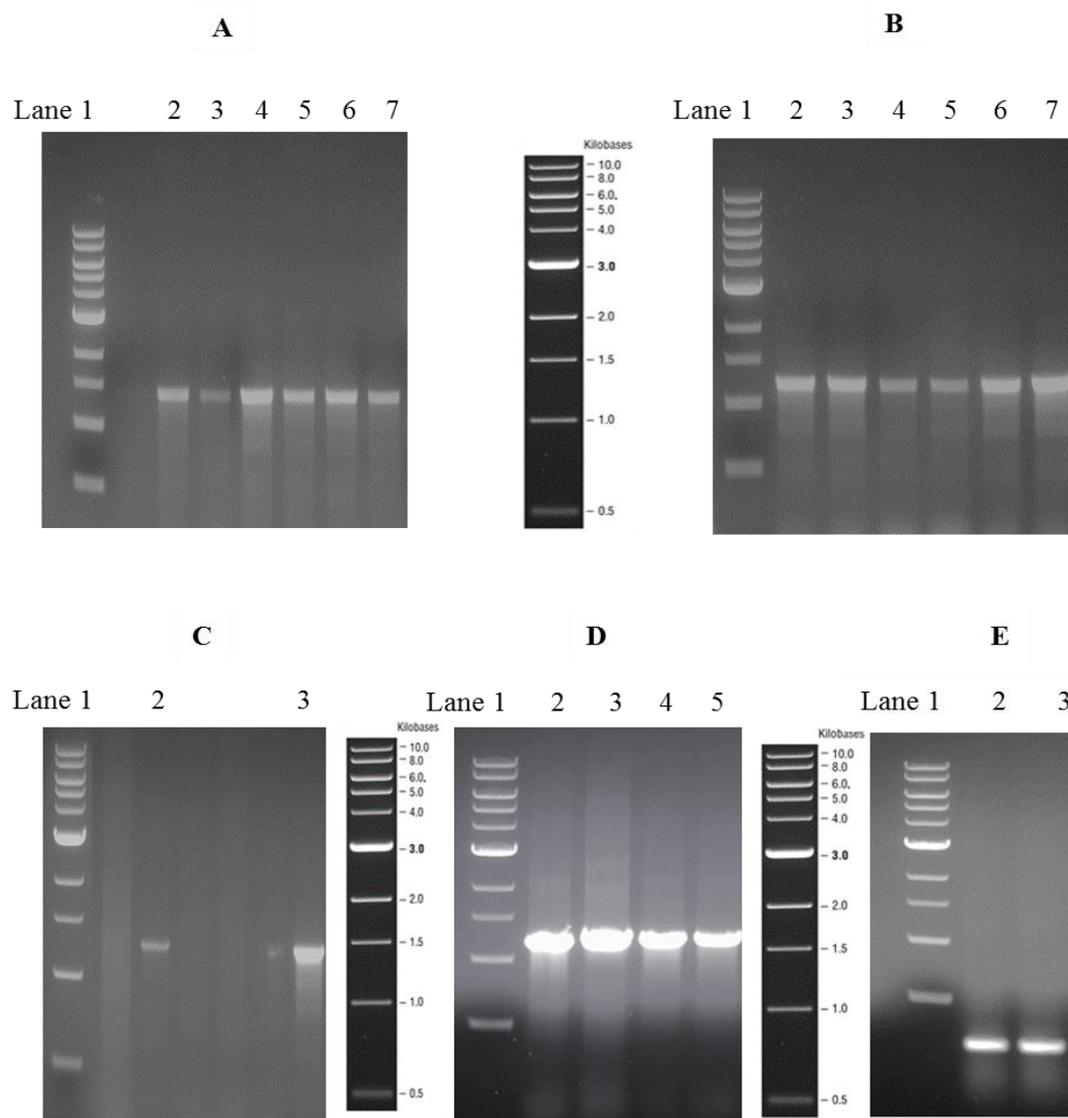
PCR products were digested with *Nco*I and *Bam*HI, and ligated into the *Nco*I and *Bam*HI digested pCEP vector, which has been shown in Figure 3.11. An aliquot of the ligation product was transformed into *E. coli* competent cells *DH5* $\alpha$  for propagation and confirmation purposes. The transformants were selected on Luria-Bertani (LB) agar plates supplemented with kanamycin (150  $\mu$ g/ml). The recombinant plasmids were purified by using the Wizard SV Gel and PCR Clean-Up System from Promega, and tested by PCR using primers MalF and pCEPR, whose recognition sites are localised up- and downstream of the cloning site, respectively. The sequencing result is showing in Appendix 2. 100 ng of respective recombinant plasmid was transformed into  $\Delta$ *rgg144*,  $\Delta$ *rgg144/shp144*,  $\Delta$ *rgg939* and  $\Delta$ *rgg939/shp939*, respectively. The transformants were selected on blood agar plates supplemented with spectinomycin (100  $\mu$ g/ml) and kanamycin (250  $\mu$ g/ml). Successful integration into the pneumococcal genome was confirmed by colony PCR using primers MalF and pCEPR. An illustration of this PCR strategy is shown in Figure 3.12. For the negative control, approximately 264 bp product was amplified in the empty vector. For the positive transformants, approximately 1424 bp, 1235 bp, 1377 bp and 1276 bp product were amplified, respectively (Figure 3.13 A, B, C and D).



**Figure 3.11. Agarose gel electrophoresis analysis showing the digested and undigested pCEP.** Lane 1: 500 ng of 1 kb DNA ladder (NEB). Lane 2 contains undigested pCEP. Lane 3 and 4 contain *NcoI* and *BamHI* digested pCEP, respectively.



**Figure 3.12. Illustration showing the strategy used to confirm successful integration of the inserts carrying the intact copy of each gene with their putative promoters within the pneumococcal genome of  $\Delta rgg144$ ,  $\Delta rgg/shp144$ ,  $\Delta rgg939$  and  $\Delta rgg/shp939$ .** Primers MalF and pCEPR were used to amplify the target genes and their putative promoter regions along with 172 bp of sequence surrounding the cloning site. The resulting PCR products were 1424 bp for *rgg144/shp144*, 1235 bp for *rgg144*, 1377 bp for *rgg939/shp939* and 1276 bp for *rgg939* inserts.



**Figure 3.13. Agarose gel electrophoresis analysis confirming the successful genetic complementation of mutants for  $\Delta rgg/shp144$  (A),  $\Delta rgg144$  (B),  $\Delta rgg/shp939$  (C) and  $\Delta rgg939$  (D), as well as the negative control (E). Lane 1 is 500 ng of 1 kb DNA ladder (NEB). (A) Lane 2-7 are positive transformants, which include the up-stream region and coding sequence of *rgg/shp144*, with an expected amplicon size of approximately 1424bp. (B) Lane 2-7 are positive transformants, which include the up-stream and downstream region of cloning site, promoter regions and *rgg144*, with an expected amplicon size of approximately 1235bp. (C) Lane 2&3 are positive transformants, which include the up-stream and downstream region of cloning site, promoter regions and *rgg/shp939*, with an expected amplicon size of approximately 1377 bp. (D) Lane 2 - 5 are positive transformants, which include the up-stream and downstream region of cloning site, promoter regions and *rgg939*, with an expected amplicon size of approximately 1276bp. (E) Lane 2 & 3 show PCR product approximately 264 bp obtained with pCEP without any insert using the MalF and pCEPR primers.**

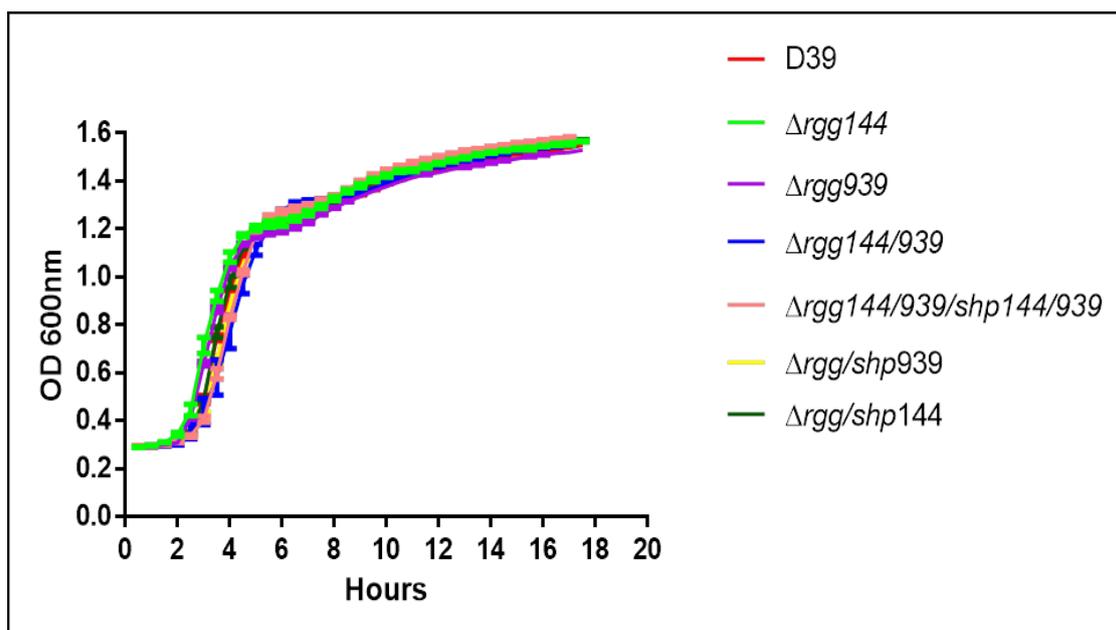
## **Section B. Phenotypic characterisation of the Rgg mutants.**

It was reported that Rgg regulators have a variety of physiological functions. They are involved in thermal adaption in *S. thermophilus* (Henry et al., 2011), pathogenicity in *S. pyogenes* and *S. mutans*, and have an effect on control of biofilm formation in *S. pyogenes* (Chang et al., 2011). They are implicated in H<sub>2</sub>O<sub>2</sub> and paraquat resistance in *S. pyogenes*, and bacteriocin production in *Streptococcus mutans* (Qi et al., 1999) and *Lactobacillus sakei* (Rawlinson et al., 2005). They also affect the non-glucose carbohydrate metabolism, and are necessary for the virulence of *S. suis* (Zheng et al., 2011). However, the biological roles of pneumococcal Rgg have not been investigated. Hence, the effects of Rgg deletion were investigated by growth assay, H<sub>2</sub>O<sub>2</sub> and paraquat killing assay, and *in vitro* virulence assays.

The possible effect of the Rgg deficiency on *S. pneumoniae* growth was determined by growing the wild-type D39 and the isogenic mutants in BHI broth or CDM supplemented with 55 mM glucose, galactose, mannose, *N*-actyl glucosamine (GlcNAc), fructose or sucrose as the sole carbon source as previously described in section 2.23. Any attenuation in growth on selected sugar would imply the direct or indirect involvement of the *rgg* regulation.

### **3.3 Growth of pneumococcal strains in BHI**

In BHI broth, similar growth rates and yields were measured for D39 wild type and Rgg deficient mutants under micro-anaerobic growth environment (Figure 3.14), ( $p > 0.05$ ). Calculated growth rates and yield are shown in Table 3.1. The growth rates ( $\mu$ ) ranged between 0.567 to 0.738 h<sup>-1</sup>. Bacterial yield (maximal OD<sub>600nm</sub>) for different strains ranged between 1.525 and 1.566. The results suggest that the mutants are not deficient in growth in rich medium.



**Figure 3.14. Pneumococcal growth curves performed micro-aerobically in BHI.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using oneway ANOVA and Dunnett's multiple comparisons test. ( $p>0.05$ ).

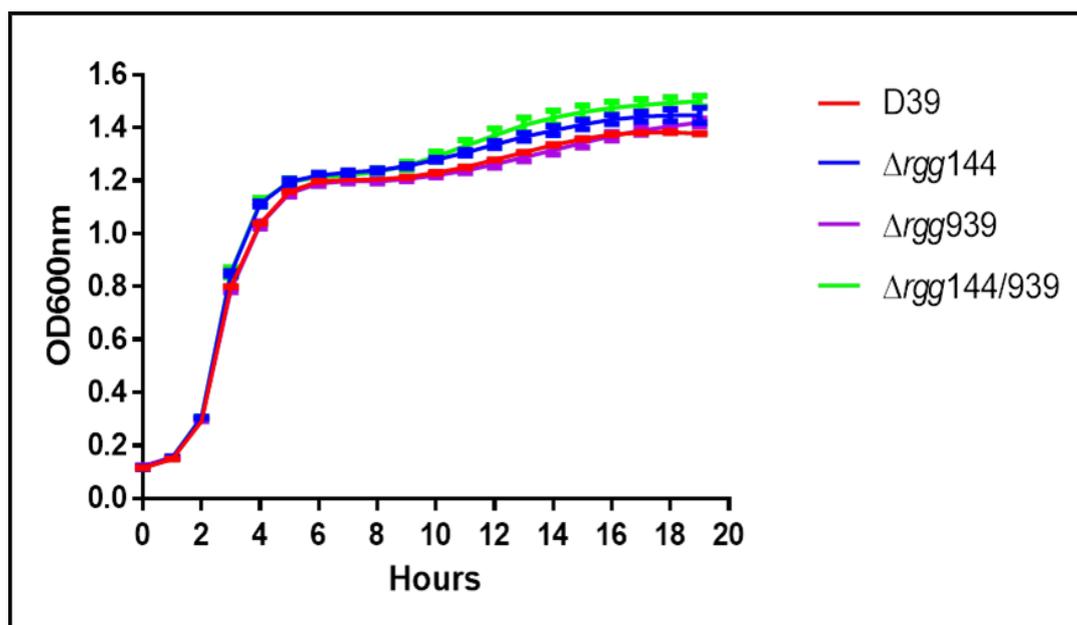
**Table 3.1. Growth rate ( $\mu$ ) and yield when pneumococcal strains grow in BHI.** Values are average of at least three independent experiments each with three replicates. ' $\pm$ ' indicates standard error of means (SEM).

Strains	Growth rate	Yield
D39	0.66 $\pm$ 0.036	1.57 $\pm$ 0.003
<i>Δrgg144</i>	0.73 $\pm$ 0.031	1.53 $\pm$ 0.006
<i>Δrgg144/shp144</i>	0.72 $\pm$ 0.039	1.58 $\pm$ 0.006
<i>Δrgg939</i>	0.74 $\pm$ 0.002	1.56 $\pm$ 0.010
<i>Δrgg939/shp939</i>	0.73 $\pm$ 0.013	1.57 $\pm$ 0.003
<i>Δrgg144/939</i>	0.67 $\pm$ 0.034	1.56 $\pm$ 0.013
<i>Δrgg144/939/shp144/939</i>	0.68 $\pm$ 0.005	1.56 $\pm$ 0.006

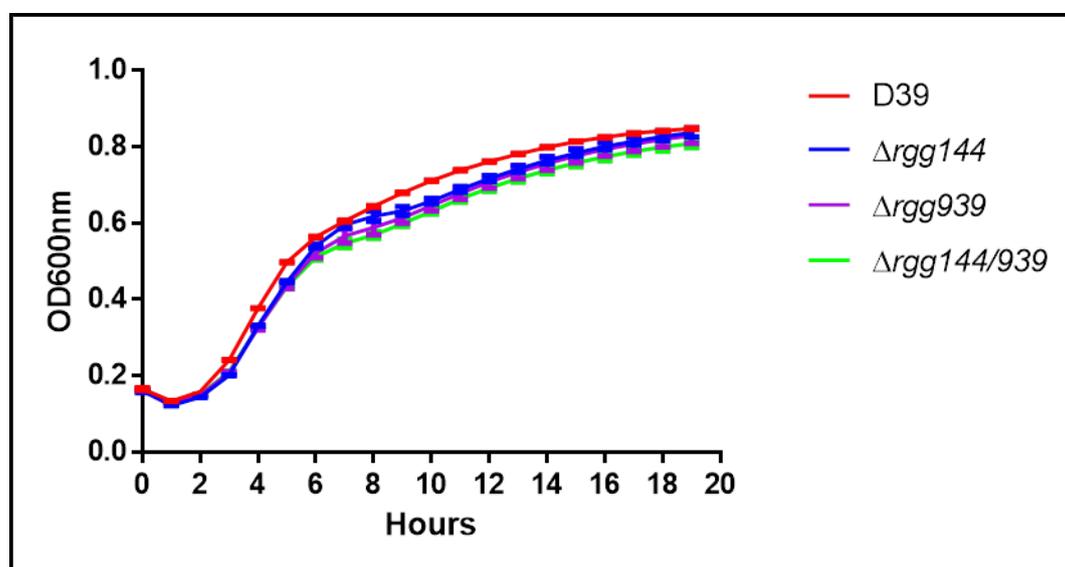
### **3.4 Growth of pneumococcal strains in CDM with different sugars**

In order to test the role of Rggs in carbohydrate utilisation, D39 and the isogenic *rgg/shp* mutants were incubated in chemically defined medium supplied with 1% (w/v) glucose, galactose, mannose, fructose, sucrose or *N*-acetyl glucosamine (GlcNAc) as the primary carbon source. There was no significant difference in growth between the wild type and the Rgg mutants when glucose, galactose, fructose, sucrose or *N*-acetyl glucosamine (GlcNAc) were used as primary carbon source ( $p > 0.05$ ) (Figure 3.15, 3.16, 3.17, 3.18 and 3.19). The result implied that *rgg* inactivation did not decrease the glucose, galactose, fructose, sucrose and GlcNAc utilization of the organism under *in vitro* conditions. Complemented mutants had the similar phenotype as the wild type. However, when mannose was used as the sole carbon source (Figure 3.20), Rgg mutants displayed attenuated growth; the yields and growth rates were shown in Table 3.2. The *rgg144* and *rgg939* displayed a lower growth yield ( $1.03 \pm 0.024$  and  $0.95 \pm 0.048$ , respectively) compare to the wild type D39 ( $1.21 \pm 0.007$ ), ( $p < 0.0001$ ). In addition, *rgg144* and *rgg939* display slower growth rates ( $0.35 \pm 0.006$  and  $0.33 \pm 0.039$ , respectively) compare to the wild type ( $0.40 \pm 0.009$ ) ( $p < 0.05$ ). Moreover, the double mutants  $\Delta rgg144/939$  had the lowest growth rate ( $0.30 \pm 0.014$ ) ( $p < 0.001$ ) and yield ( $0.90 \pm 0.016$ ) compared to the wild type ( $p < 0.0001$ ). The growth rate and yield of  $\Delta rgg144/939$  is significant lower than  $\Delta rgg144$  and  $\Delta rgg939$  ( $p < 0.001$  and  $p < 0.05$ , respectively). This result indicates that *rgg144* and *rgg939* are required for pneumococcal growth on mannose, implying their involvement in mannose metabolism.

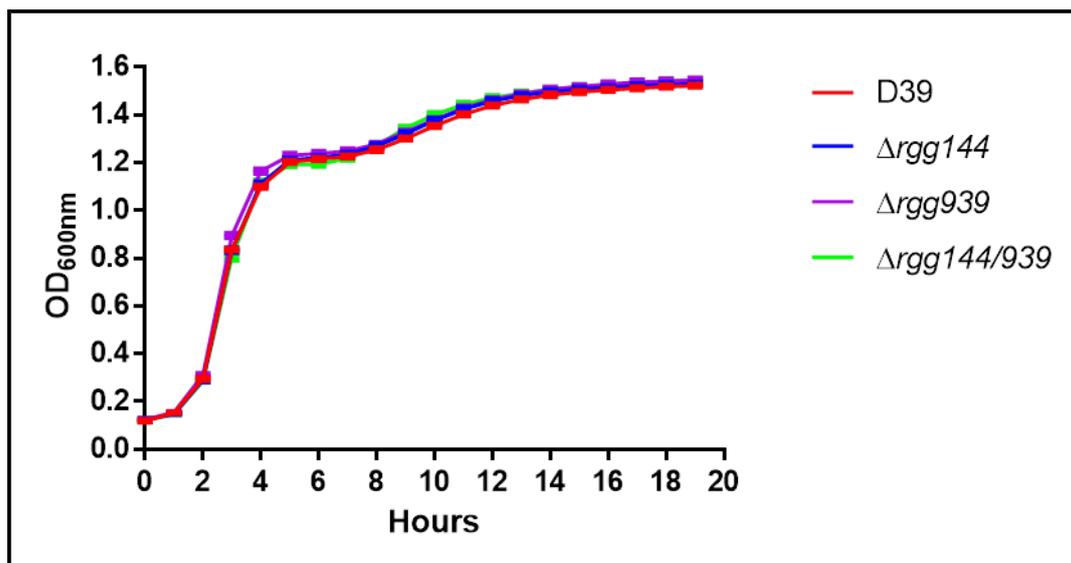
The genetically complemented strains  $\Delta rgg144$ Comp and  $\Delta rgg939$ Comp were also grown micro-aerobically in CDM supplemented with 55 mM mannose as the sole carbon source (Tables 3.2). Introduction of an intact copy of the *rggs* restored the growth rates of the complemented strains, indicating that the observed phenotype in the mutants were not due to polar effect of mutation.



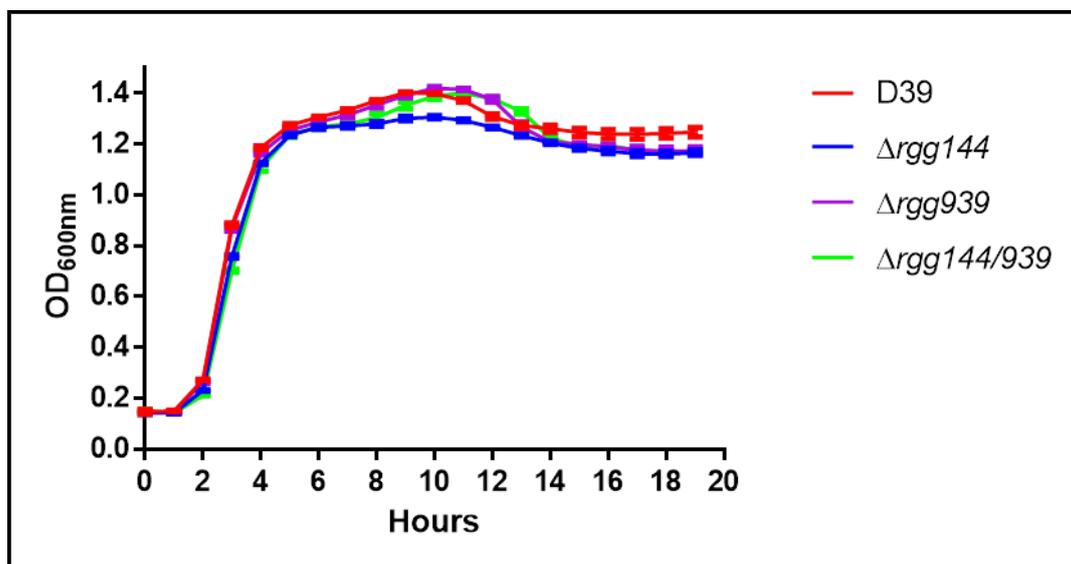
**Figure 3.15. Pneumococcal growth curves performed micro-aerobically in CDM supplemented with 1% (w/v) glucose.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using one way ANOVA and Dunnett's multiple comparisons test ( $p>0.05$ ).



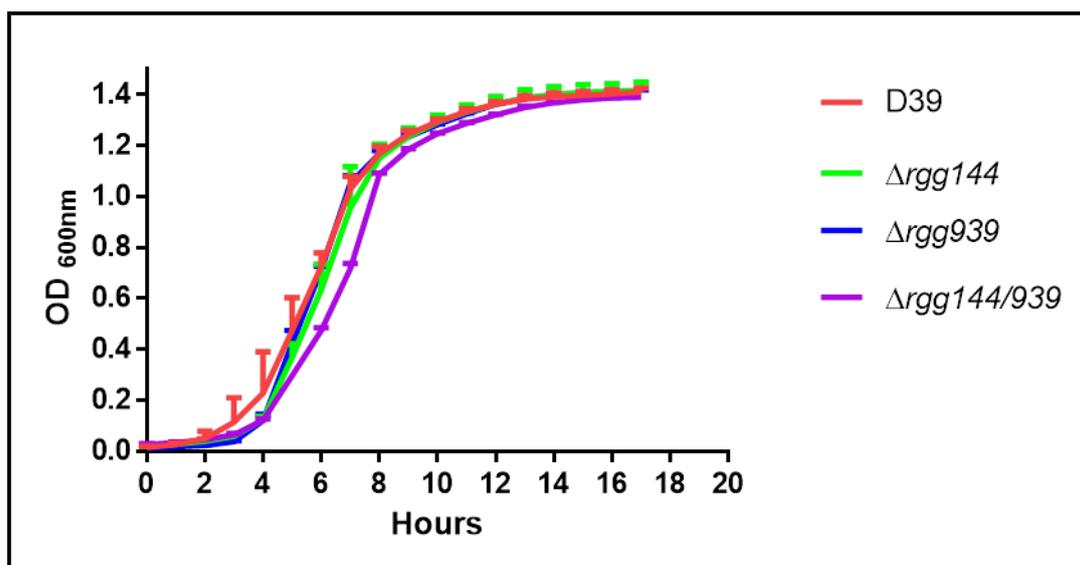
**Figure 3.16. Pneumococcal growth profiles performed micro-aerobically in CDM supplemented with 1% (w/v) galactose.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using oneway ANOVA and Dunnett's multiple comparisons test ( $p>0.05$ ).



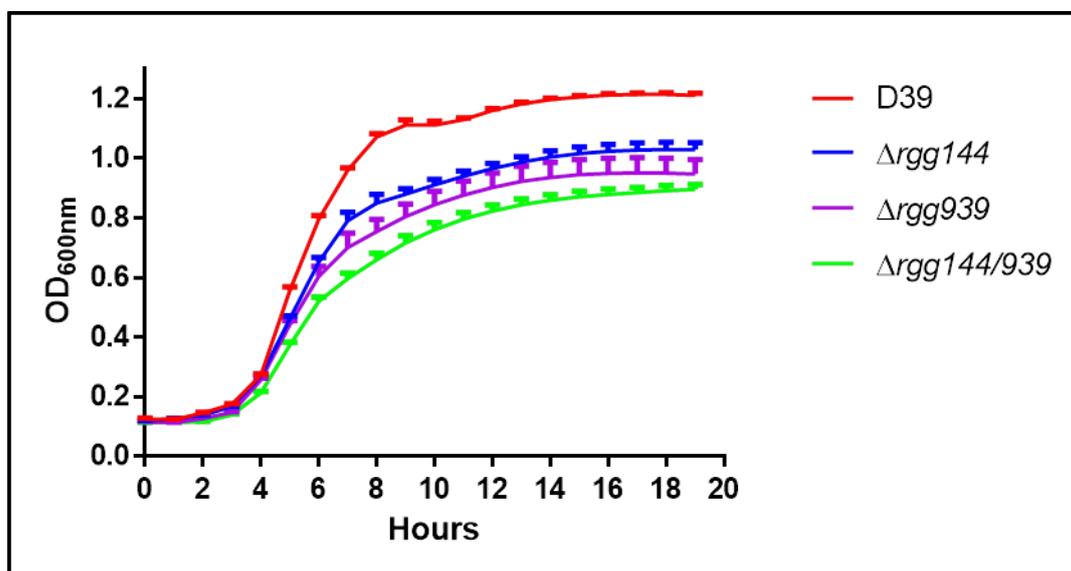
**Figure 3.17. Pneumococcal growth profiles performed micro-aerobically in CDM supplemented with 1% (w/v) sucrose.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using oneway ANOVA and Dunnett's multiple comparisons test ( $p>0.05$ ).



**Figure 3.18. Pneumococcal growth curves performed micro-aerobically in CDM supplemented with 1% (w/v) fructose.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using oneway ANOVA and Dunnett's multiple comparisons test ( $p>0.05$ ).



**Figure 3.19. Pneumococcal growth curves performed micro-aerobically in CDM supplemented with 1% (w/v) GlcNAc.** Error bars show the standard error of the mean for three individual measurements each with three replicates. No significant differences were seen when comparing the growth rates of mutant strains to the wild type D39 using oneway ANOVA and Dunnett's multiple comparisons test ( $p > 0.05$ ).



**Figure 3.20. Pneumococcal growth profiles in CDM supplemented with 1% (w/v) mannose.** Error bars show the standard error of the mean for three individual measurements each with three replicates. Rgg deficient mutants have decreased growth rate and lower yield relative to the wild type strain ( $n=9$ ). (Growth rate:  $\Delta rgg144$  & D39:  $*p < 0.05$ ,  $\Delta rgg939$  & D39:  $*p < 0.05$ ,  $\Delta rgg144/939$  & D39:  $*** p < 0.001$ ,  $\Delta rgg144$  &  $\Delta rgg144/939$ :  $*** p < 0.001$ ,  $\Delta rgg939$  &  $\Delta rgg144/939$ :  $*p < 0.05$ ; yield:  $\Delta rgg144$  & D39,  $****p < 0.0001$ ,  $\Delta rgg939$  & D39,  $****p < 0.0001$ ,  $\Delta rgg144/939$  & D39,  $****p < 0.0001$ ,  $\Delta rgg144$  &  $\Delta rgg144/939$ :  $*** p < 0.001$ ,  $\Delta rgg939$  &  $\Delta rgg144/939$ :  $*p < 0.05$ ).

**Table 3.2. Growth rate ( $\mu$ ) and yield of pneumococcal strains grow in CDM with 1% (w/v) mannose.** Values are average of at least three independent experiments. ‘ $\pm$ ’ indicates standard error of means (SEM).

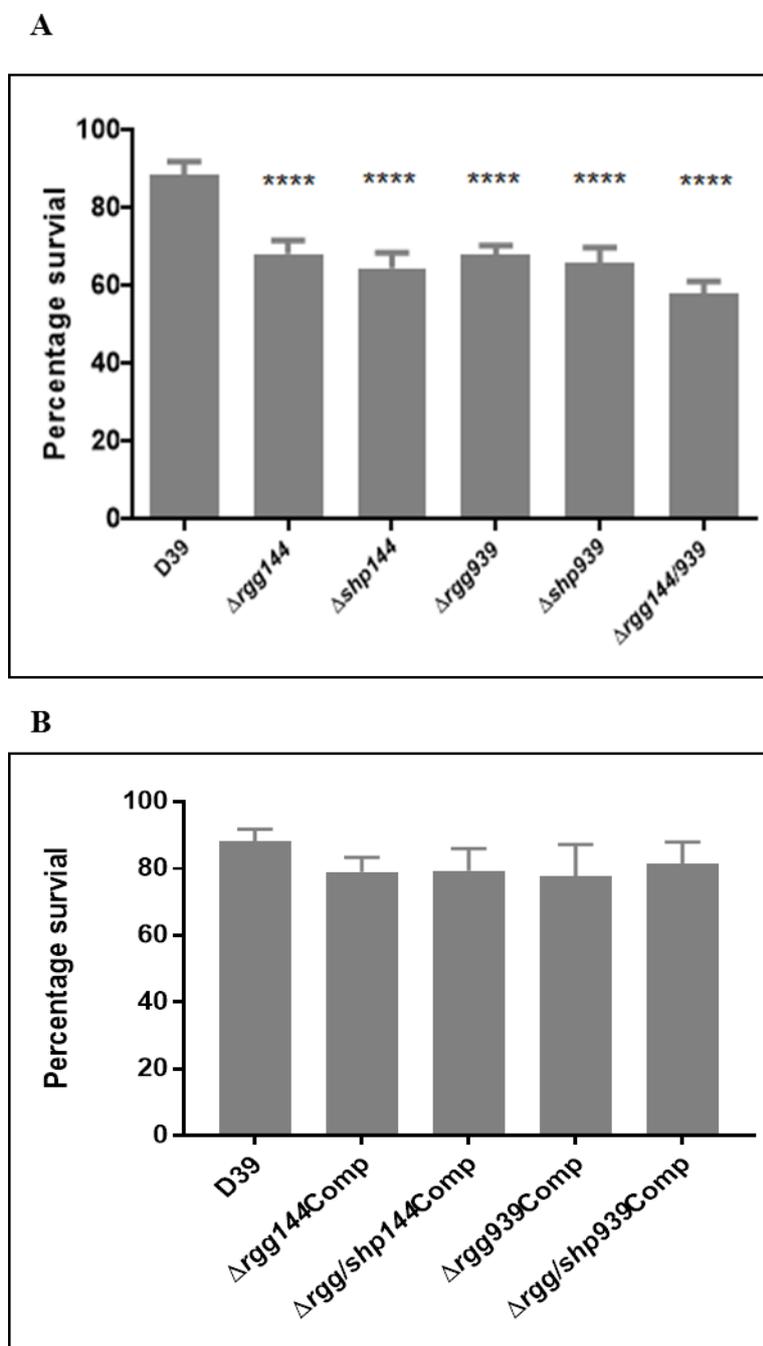
Strains	Growth rate	Yield
D39	0.40 $\pm$ 0.009	1.21 $\pm$ 0.007
$\Delta rgg144$	0.35 $\pm$ 0.006	1.03 $\pm$ 0.024
$\Delta rgg939$	0.33 $\pm$ 0.039	0.95 $\pm$ 0.048
$\Delta rgg144/939$	0.30 $\pm$ 0.014	0.90 $\pm$ 0.016
$\Delta rgg144$ Comp	0.39 $\pm$ 0.0012	1.22 $\pm$ 0.005
$\Delta rgg939$ Comp	0.41 $\pm$ 0.008	1.20 $\pm$ 0.009

### 3.5 Rggs confer protection against H<sub>2</sub>O<sub>2</sub>

*Streptococcus pneumoniae* produces large amounts of H<sub>2</sub>O<sub>2</sub> by the activity of the enzyme pyruvate oxidase during its aerobic sugar metabolism but the organism itself can tolerate high concentrations of H<sub>2</sub>O<sub>2</sub> that would normally lead to oxidative stress and result in peroxidation of lipids, DNA lesions, protein oxidation and even cell death in other bacterial species (Birben et al., 2012). It was reported that Rgg family regulators are involved in reactive oxygen species (ROS) resistance. The Rgg mutant in *S. pyogenes* was less tolerant to paraquat (Chaussee et al., 2004). In addition, pneumococcal *rgg1952* mutant strain showed susceptibility to oxidative stress (Bortoni et al., 2009).

To test the contribution of *rggs* and *shps* in peroxide resistance, pneumococci were grown in CDM supplemented with 1% (w/v) glucose, and when OD<sub>600 nm</sub> reached 0.3, 20 mM H<sub>2</sub>O<sub>2</sub> was added into bacterial suspension and incubated for 30 min at 37 °C. The percentage of surviving bacteria was determined by serial dilution relative to the control that had not been treated by H<sub>2</sub>O<sub>2</sub> (Figure 3.21A). The challenge of pneumococci with 20 mM of H<sub>2</sub>O<sub>2</sub> led to 88.3  $\pm$  2.8%, 68.1  $\pm$  8.3%, 64.4  $\pm$  5.4% 67.7  $\pm$  3.0%, 65.7  $\pm$  2.1%, and 57.8  $\pm$  2.5%, survival of D39,  $\Delta rgg144$ ,  $\Delta shp144$ ,  $\Delta rgg939$ ,  $\Delta shp939$ ,  $\Delta rgg144/939$ , respectively. The percentage survival of  $\Delta rgg144$ ,  $\Delta shp144$ ,  $\Delta rgg939$ ,  $\Delta shp939$ ,  $\Delta rgg144/939$ , were statistically lower than the wild type ( $p < 0.0001$ ). The percentage survival of  $\Delta rgg144/939$  was significant lower

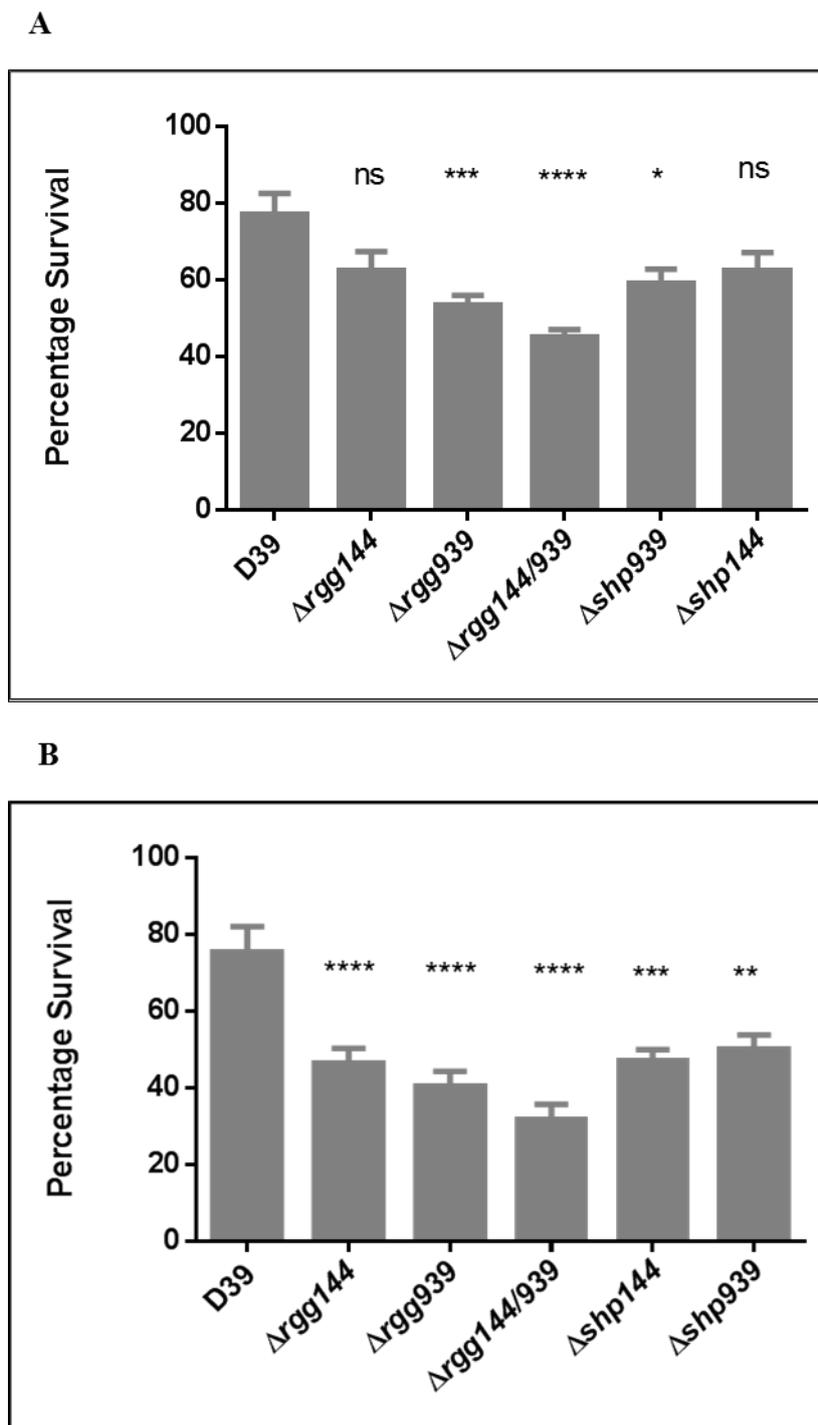
than  $\Delta rgg144$  and  $\Delta rgg939$  ( $p < 0.05$ ). This result suggests that both Rgg/Shp systems play a role in protection against  $H_2O_2$ . Complemented mutants had the similar phenotype as the wild type ( $79.1 \pm 1.9\%$ ,  $73.4 \pm 2.7\%$ ,  $75.8 \pm 2.9\%$  and  $75.3 \pm 4.2\%$ ) ( $p > 0.05$ ), (Figure 3.21B), which demonstrated the role of Rgg on ROS resistance is not due to the polar effect.



**Figure 3.21.** The susceptibility of pneumococcal strains to 20 mM  $H_2O_2$ . The percentage of surviving pneumococci was determined by serial dilution relative to the control that had not been treated by  $H_2O_2$ . Error bars indicate the SEM ( $n=8$ , \*\*\*\* $p < 0.0001$ ).

### **3.6 Rgg involve in paraquat resistance**

To determine whether Rgg/Shp systems are involved in protection against superoxide, pneumococci were grown in CDM supplemented with 1% (w/v) glucose, and when OD<sub>600</sub> nm reached 0.3, 0.5 mM or 1 mM paraquat (a redox active compound that generates intracellular superoxide in the presence of oxygen) were added into bacterial suspension and incubated for 30 min at 37 °C. The percentage of surviving bacteria was determined by serial dilution relative to the control that had not been treated with paraquat. The challenge of pneumococci with 0.5 mM of paraquat (Figure 3.22A) led to a 76.9 ± 5.6 %, 62.4 ± 4.9%, 53.4 ± 2.5, 45.1 ± 1.9%, 59.0 ± 3.6 % and 62.5 ± 4.5 % survival of D39, *Δrgg144*, *Δrgg939*, *Δrgg144/939*, *Δshp939* and *Δshp144*, respectively, which is illustrated in Figure 3.22A. In comparison to D39, the percentage of survival significantly decreased for *Δrgg939* ( $p < 0.001$ ) *Δrgg144/939* ( $p < 0.0001$ ) and *Δshp939* ( $p < 0.05$ ). There was no significant difference between D39 and *Δrgg144* ( $p > 0.05$ ), D39 and *Δshp144* ( $p > 0.05$ ). However, the challenge of pneumococci with 1mM of paraquat led to a 75.3 ± 6.8%, 46.3 ± 3.9, 40.5 ± 3.7%, 31.7 ± 3.9%, 49.9 ± 1.3% and 47.1 ± 2.73% survival of D39, *Δrgg144*, *Δrgg939*, *Δrgg144/939*, *Δshp939* and *Δshp144*, respectively, which is illustrated in Figure 3.22B. The percentage survival of *Δrgg144*, *Δrgg939* and *Δrgg144/939*, *Δshp939* and *Δshp144* was statistically lower than D39 ( $p < 0.0001$ ,  $p < 0.0001$ ,  $p < 0.0001$ ,  $p < 0.001$ ,  $p < 0.01$ , respectively). The percentage survival of *Δrgg144/939* was significant lower than *Δrgg144* and *Δrgg939* ( $p < 0.05$ ). These results suggest that *rgg144*, *rgg939*, *shp939* and *shp144* are all involved in protection against superoxide.



**Figure 3.22. The susceptibility of pneumococcal strains to 0.5 mM (A) and 1 mM (B) paraquat.** The percentage of surviving bacteria was determined by serial dilution relative to the control that had not been treated by paraquat. Error bars indicate the SEM, (n=8, \* $p$ <0.05, \*\* $p$ <0.01, \*\*\*  $p$ <0.001, \*\*\*\*  $p$ <0.0001, ns:  $p$ >0.05).

## **3.7 Phenotypic characterisation of Rgg/Shp circuits**

Rggs were found to be involved in the pathogenicity of various streptococcal species, for instance, in *S. suis* Rgg mutant shows increased adhesion to Hep-2 cells and haemolytic activity *in vitro* (Zheng et al., 2010), thus several enzymatic activities were tested to identify the effect of pneumococcal Rgg on virulence. In this section, Rggs' contributions to pneumococcal neuraminidase activity, haemolytic potential and capsule synthesis were assessed individually.

### **3.7.1 Neuraminidase activity**

*Streptococcus pneumoniae* has three neuraminidases, including NanA, NanB and NanC, which could directly cleave terminal sialic residues (Kadioglu et al., 2008). To investigate the effect of Rgg on neuraminidase activity, wild type D39 and Rgg-Shp mutants were assayed at pH 6.6 using the chromogenic pNP-NANA substrate as described in section 2.25. The results are shown in Table 3.3. Neuraminidase activity for wild type pneumococci was similar as its isogenic mutants, it ranged between  $53.8 \pm 3.1$  to  $57.2 \pm 2.6$ . The difference was not statistically significant ( $p > 0.05$ ), which shows that the deletion of Rggs do not have effect on neuraminidase activity.

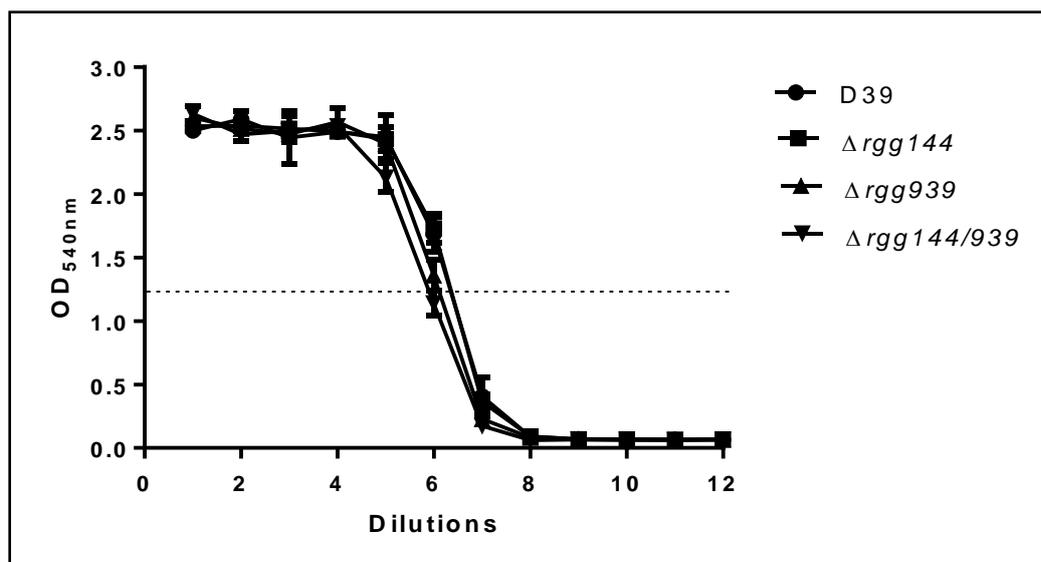
**Table 3.3: Neuraminidase activity in *S. pneumoniae*.** Neuraminidase activity was assessed at pH 6.6. Average values are expressed as nmole p-NP released per minute per  $\mu\text{g}$  total protein, and represent the average of three separate experiments, using triplicates in each experiment. '±' indicates standard error of means (SEM).

Strains	Neuraminidase Unit
D39	$54.0 \pm 3.6$
$\Delta\text{rgg}144$	$57.2 \pm 2.6$
$\Delta\text{rgg}939$	$54.9 \pm 1.9$
$\Delta\text{rgg}144/939$	$53.8 \pm 3.1$

### **3.7.2 Hemolytic activity**

Pneumolysin is an important virulence factor in *S. pneumoniae*, the absence of pneumolysin

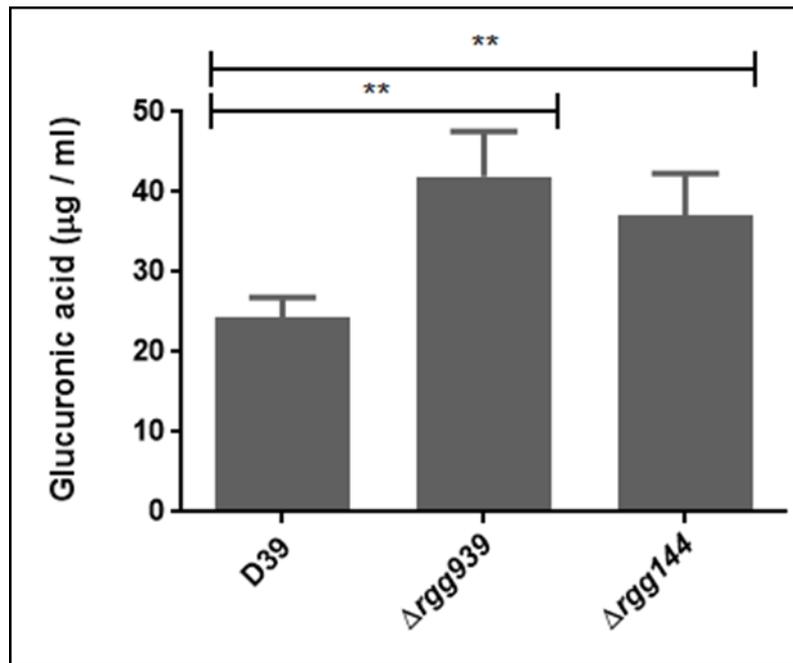
protein abolishes hemolytic activity (Alexander et al., 1998). The effect of Rgg on hemolytic activity was tested in the cell lysates of wild type D39 and Rgg mutants using 4% of sheep red blood cells as described in section 2.26. The results are shown in Figure 3.23. Hemolytic activity for wild type pneumococci was similar to its isogenic mutants ( $p>0.05$ ). They all produced approximately 64 HU in 1 mg total protein, which indicated that the deletion of Rggs do not affect hemolytic activity.



**Figure 3.23. The hemolytic activity of the wild type and *rgg* mutants.** The dotted horizontal line means the highest dilution of the supernatant inducing at least 50% lysis of red blood cells. The error bars indicates the SEM. ( $p>0.05$ ).

### **3.7.3 Capsule content determination by glucuronic acid assay.**

Capsular polysaccharide (CPS) is the most important pneumococcal virulence factor, which can protect bacteria from phagocytosis, and plays a crucial role in the pneumococcal survival in different environment. In order to determine whether Rggs have any role in capsule production, polysaccharide capsule isolation followed by glucuronic acid assay were carried out as described in section 2.27. The results are shown in Figure 3.24. Compared with the wild type strain D39 ( $23.4 \pm 2.4 \mu\text{g}/\text{ml}$ ), the *rgg939* mutant strain produced the most glucuronic acid ( $41.9 \pm 2.3 \mu\text{g}/\text{ml}$ ) ( $p<0.01$ ), followed by *rgg144* mutant, which produced  $37.16 \pm 2.10 \mu\text{g}/\text{ml}$  of glucuronic acid ( $p<0.01$ ). The result shows that Rgg may inhibit the capsule synthesis.



**Figure 3.24. The glucuronic acid concentration from capsules of pneumococcal strains.** Wild-type D39 and Rgg mutants were grown in CDM supplemented with 55 mM mannose. Glucuronic acid quantification assay was done by using prepared pneumococcal capsule. The error bars indicate the SEM, (n=9, \*\* $p < 0.01$ ).

## **Section C. Construction of *lacZ* reporter systems and $\beta$ -galactosidase assays**

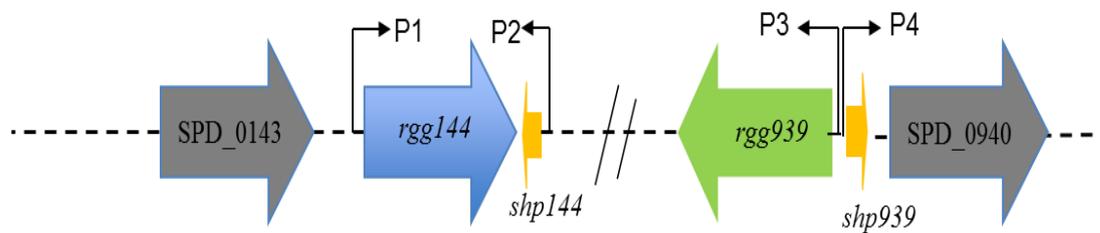
### **3.8 Construction of *lacZ*-fusions**

Transcriptional fusions are frequently used to analyze gene expression patterns, monitor the regulatory systems, and investigate their response to different environmental niches, such as pheromones, carbon source and reactive oxygen species (ROS) in both prokaryotic and eukaryotic organisms. Transcriptional fusions between promoter regions of interest and reporter gene encoding an easily assayable product makes possible to monitor target gene under defined environmental conditions. The available reporter genes include the *E. coli*  $\beta$ -galactosidase *lacZ* gene (Shapira et al., 1983), the green fluorescent protein gene *gfp* from the jelly fish *Aequorea victoria* (Prasher et al., 1992), the luciferase gene, *luxAB*, from the firefly *Photinus pyralis* (Bronstein et al., 1994), the *E. coli* chloramphenicol acetyl transferase gene, *cat*, (Gorman et al., 1982) and the *E. coli*  $\beta$ -glucuronidase *gus* gene (Jefferson et al., 1987).

The *E. coli lacZ* gene has been used in this study as it is well established in monitoring the regulation of gene expression, and the *lacZ* activity can be easily assayed by a  $\beta$ -galactosidase assay (Miller, 1972). The integrative plasmid pPP2, derived from pBR322, was selected for constructing the *lacZ* reporter systems, which is shown in Figure 3.25. This plasmid was designed to have a promoterless  $\beta$ -galactosidase gene, *lacZ* gene, from *E. coli* (Halfmann et al., 2007). Translation start signal is from the protease gene, *htrA*, of *S. pneumoniae*. The ampicillin resistance gene  $\beta$ -lactamase gene (*bla*) and tetracycline resistance gene *tetM* can be used to select the transformants in *E. coli* and *S. pneumoniae*, respectively. The other crucial feature of this plasmid is that it is designed to integrate into the genetic region of *S. pneumoniae* containing endogenous  $\beta$ -galactosidase gene, *bgaA* (SPD\_0562), via homologous recombination. This abolishes the background  $\beta$ -galactosidase activity. Once transcriptional fusions are constructed, the promoter of target gene will drive the expression of promoterless *lacZ* gene. Therefore, the regulatory scheme can be investigated by *lacZ* activity.



A.



B.

ATTAGAACGAGGATTGAGTTCTGTATTGTACAAGGCTCGGTCCTTTTAGAGTCAG  
 CTTAAGGCTGGCTTTTTCAATCACCAAAGTGTGAGAATGTT**TTGACA**AATGAACA  
 CAAATAATGA**TATAAT**ATGCAAAGCTAGGAGGTGGTAGGATGATTGAAAA**ATG**  
 GAACTGGGGGAATTT

C.

CTTTTCTT**CAT**CTATTATATTCCTCCTTGTTAGTAATAACTTTATTATAACCAGGGAAAT  
 AATCAAATCTATCAAAATCGCAAATAAGAAATTTCTATAAGAAAAATATCAAATAT  
 GCGATTT**TTAAA**ATAAGCCAATTTTCGTGT**TATACT**GTACTTGTAAGCACTTGA  
 AGCAAATCCTAGGTCGCAGAAGTGGTTTACAAATGAAGATAATTGAAGGAGTGTA  
 AAAGATGTTAAATTTACAATTTGCAGAAACAATGGAATTGACAGAAGCTGAGTTG  
 CAAGATGTTAGAGGAGGC

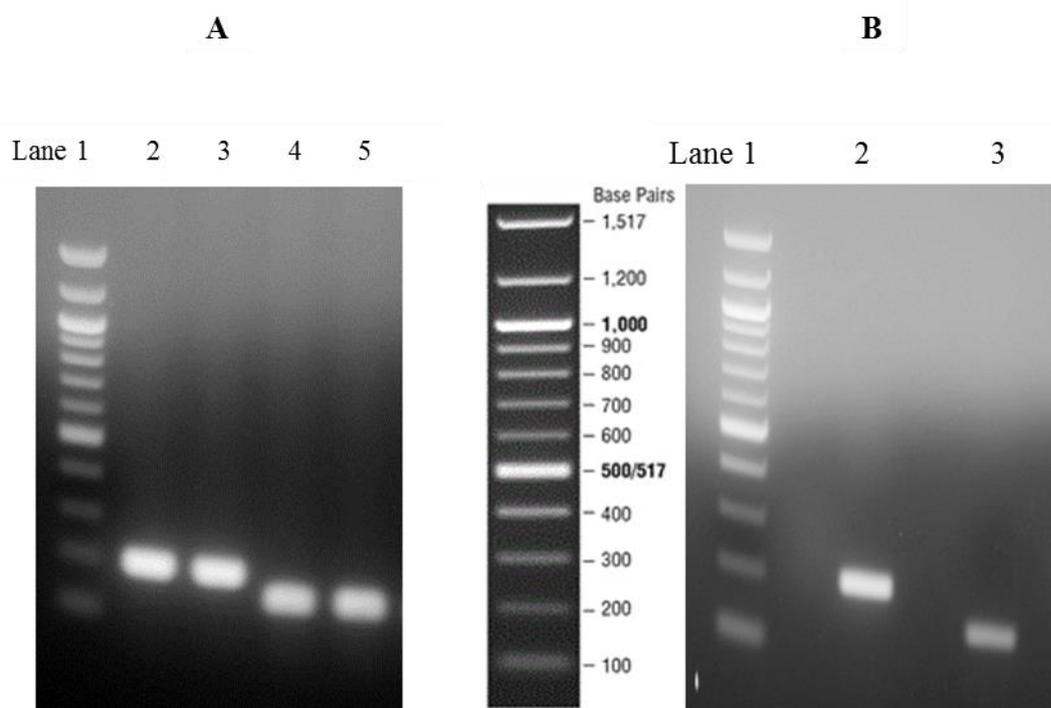
D.

AATTTTTCTAAGTGTGACACCAAGTTTTGATTT**CAT**AGTTTTCCCATTTTCCC**AACA**  
**AA**AATTA ACTTATTTCTTATT**TATTGT**ACAATAAATAAAAAATGAAAGGAAGTCAG  
 AGT**ATG**AAGAAAATTC

**Figure 3.26. Rgg/Shp pairs in the *S. pneumoniae* D39 genome (A) and intergenic region maps for *rgg144* (B) *shp144* (C), *rgg939* and *shp939* (D) gene pairs. A.** The bent arrows indicate the putative promoter regions (B, C and D). Sense DNA strand for the promoter region of each gene. The start codon of each gene is in green, the putative -10 regions are in purple, the -35 regions are in red. The *rgg939* and *shp939* genes may share the same promoter region as the upstream regions of these two genes overlap.

After identification, these four putative promoter regions were amplified by using primers listed in Table 2.4, which are modified to incorporate *SphI* and *BamHI* sites. After the amplification of the putative promoter regions by PCR, the samples were cleaned up and all the samples were analysed by agarose gel electrophoresis to illustrate their successful

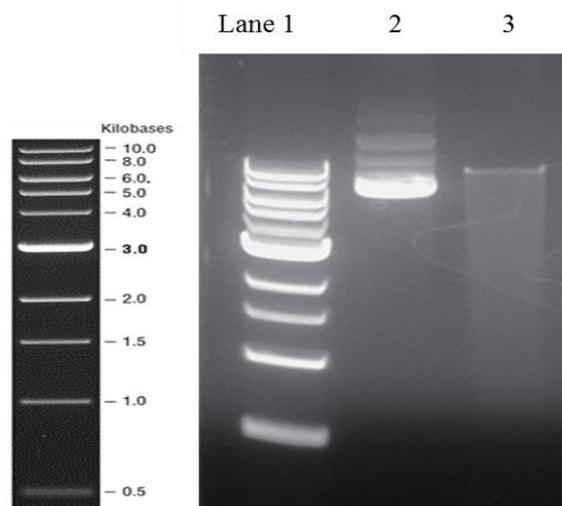
amplification. Figure 3.27A and 3.27B illustrate the amplification of all the promoter regions of *rgg144*, *rgg939*, *shp144* and *shp939*.



**Figure 3.27. Agarose gel electrophoresis showing the amplified putative promoter regions of *shp* (A) and *rgg* (B) genes.** Lane 1: 500 ng of 100 bp DNA ladder (NEB). (A). Lane 2 & 3: amplified promoter region of *shp144*, which is approximately 184 bp. Lane 4 & 5: amplified promoter region of *shp939*, which is approximately 136 bp. (B). Lane 2: amplified upstream region of *rgg144* (193 bp). Lane 3: amplified putative promoter region of *rgg939* (136 bp).

### **3.8.2 Preparation and digestion of plasmid pPP2.**

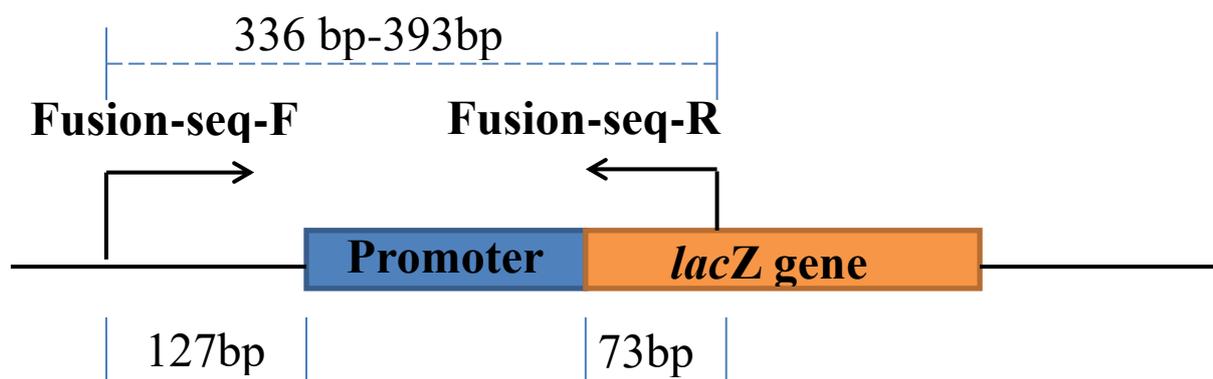
Plasmid pPP2 was double digested using the restriction enzymes *SphI* and *BamHI*. Successful digestion of pPP2 was illustrated by agarose gel electrophoresis (Figure 3.28), and was compared to the undigested pPP2.



**Figure 3.28. Agarose gel electrophoresis of the digested and undigested pPP2.** Lane1: 500 ng of 1 kb DNA ladder (NEB); Lane 2: undigested pPP2. Lane 3 pPP2 double digested with *SphI* and *BamHI*. 3.8.3 Cloning of putative promoters to pPP2

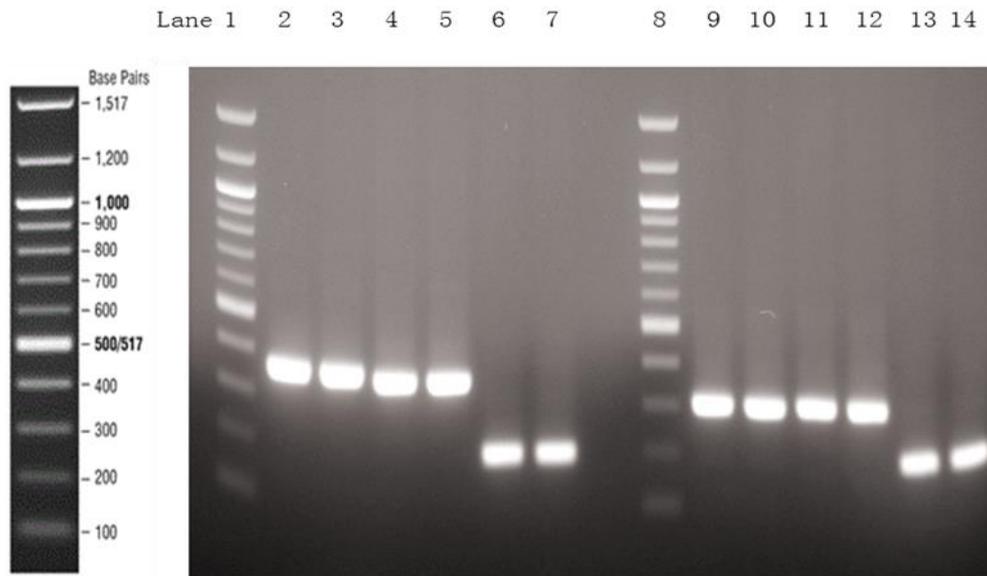
T4 DNA Ligase was used to ligate the each putative promoter region into the pPP2 plasmid. Afterwards, the ligation reaction was transformed into *E. coli* DH5 $\alpha$ . Colony PCR was done by using primers Fusion-Seq-F and Fusion-Seq-R for the confirmation of successful construction of the recombinant plasmids. PCR amplification using pPP2 plasmid as template produces approximately 200 bp products, which served as a negative control. Additionally, after the recombinant plasmids were extracted, they were sequenced.

Sequencing results indicated that all transcriptional fusions had been constructed successfully. The new constructed plasmids were transformed into the wild type or mutant pneumococci, and were integrated through homologous recombination. The transformants were selected on blood agar plates containing 3  $\mu$ g/ml of tetracycline. Again, colony PCR was used for the identification and confirmation of the positive colonies using the Fusion-Seq- F and Fusion-Seq-R primers. The colony PCR in both cases was performed as it illustrated in Figure 3.29. The resulting reporter strains were designated as XZ1 ( $P_{rgg144}$ -*lacZ*-wt), XZ2 ( $P_{shp144}$ -*lacZ*-wt), XZ3 ( $P_{rgg939}$ -*lacZ* -wt), XZ4 ( $P_{shp939}$ -*lacZ*-wt), XZ5 ( $P_{rgg144}$ -*lacZ*- $\Delta$ *rgg144*), XZ6 ( $P_{shp144}$ -*lacZ*- $\Delta$ *rgg144*), XZ8 ( $P_{shp939}$ -*lacZ* - $\Delta$ *rgg144*), XZ10 ( $P_{shp144}$ -*lacZ*- $\Delta$ *rgg939*), XZ11 ( $P_{rgg939}$ -*lacZ*- $\Delta$ *rgg939*), XZ12 ( $P_{shp939}$ -*lacZ*- $\Delta$ *rgg939*), XZ18 ( $P_{shp144}$ -*lacZ*- $\Delta$ *rgg144/939*), XZ20 ( $P_{shp939}$ -*lacZ*- $\Delta$ *rgg144/939*), XZ26 ( $P_{shp144}$ -*lacZ* - $\Delta$ *shp144*), XZ36 ( $P_{shp939}$ -*lacZ* - $\Delta$ *shp939*).



**Figure 3.29. Illustration showing the strategy used to confirm successful integration of the transcriptional *lacZ* fusions within the pneumococcal genome.** Fusion-Seq-F and Fusion-Seq-R primers were used to amplify the inserts containing putative promoter regions. The resulting PCR products ranged from 336 to 393 bp, for *rgg144* and *rgg939*, respectively, in size while the empty pPP2 plasmid yielded PCR products of 200 bp in size using the same set of primers.

The amplified PCR products were analysed by agarose gel electrophoresis (Figure 3.30). As can be seen in Figure 3.30, lanes 2-5 and 9-12 show successful amplification of inserts representing promoter regions of each gene, and 200 bp upstream and downstream of the cloning site in pPP2. The PCR products in lanes 2-5 and 9-12 had the expected approximate sizes for  $P_{rgg144}$  (393 bp),  $P_{shp144}$  (384 bp),  $P_{rgg939}$  (336 bp) and  $P_{shp939}$  (336 bp), respectively. In addition, lane 6, 7, 13 and 14 show a promoterless fragment of the pPP2 plasmid (200 bp) using the same set of primers, Fusion-Seq-F and Fusion-Seq-R.



**Figure 3.30. Agarose gel electrophoresis confirming the successful integration of the recombinant pPP2 into wild type D39 genome.** Lanes 1 and 8 are 100 bp DNA ladder (NEB). Lanes 6, 7, 13 and 14 are negative control, which use empty pPP2 plasmid as template (200 bp), Lanes 2 and 3 are the amplicons from  $P_{rgg144}$  (393 bp), lanes 4 and 5 are amplicons from  $P_{shp144}$  (384 bp). Lanes 9, and 10 are amplicons from the  $P_{rgg939}$  (336 bp). Lanes 11, and 12 are amplicons from the  $P_{shp939}$  (336 bp). All of the PCR reactions used the primers Fusion-Seq-F and Fusion-Seq-R.

### **3.9 Assessing the inducibility of Rgg-Shp circuits**

The protocol used here is derived from previously published reports (Miller, 1972, Zhang and Bremer, 1995). The promoters that were investigated in this study were cloned in front of the promoterless *lacZ* gene as described in section 2.32. The recombinant reporter systems were transformed into *S. pneumoniae*. If the promoter is responsive to the selected stimuli, it would drive the expression of promoterless *lacZ* gene, leading to the production of  $\beta$ -galactosidase. Afterwards, the level of  $\beta$ -galactosidase activity as a measure of promoter activity could be determined in the cell extracts by measuring the appearance of yellow colour over time due to hydrolysis of *O*-nitrophenyl-D-galactoside (ONPG), a chromogenic lactose mimic, by  $\beta$ -galactosidase.

#### **3.9.1 Promoter induction by Shp**

In *S. pneumoniae* D39 genome, there are two small, unannotated open reading frames (ORF)

which were designated as the genes for short hydrophobic peptides (*shp*), which are proximal to *rgg* genes. The Rgg-Shp pairs are proposed to act as quorum sensing systems (Chang et al., 2011). These SHP peptides are usually synthesized as inactive precursors, processed by pheromone-specific peptidase (Eep), and released to extracellular medium by the general secretory (Sec) system or the ABC-type transporters. Processed peptides are re-imported into cytoplasm by an oligopeptide permease transport (Linton and Higgins, 2007). Once inside the cells, active peptides activate or repress the transcription of target genes to affect bacterial phenotype such as biofilm formation (Figure 3.31).

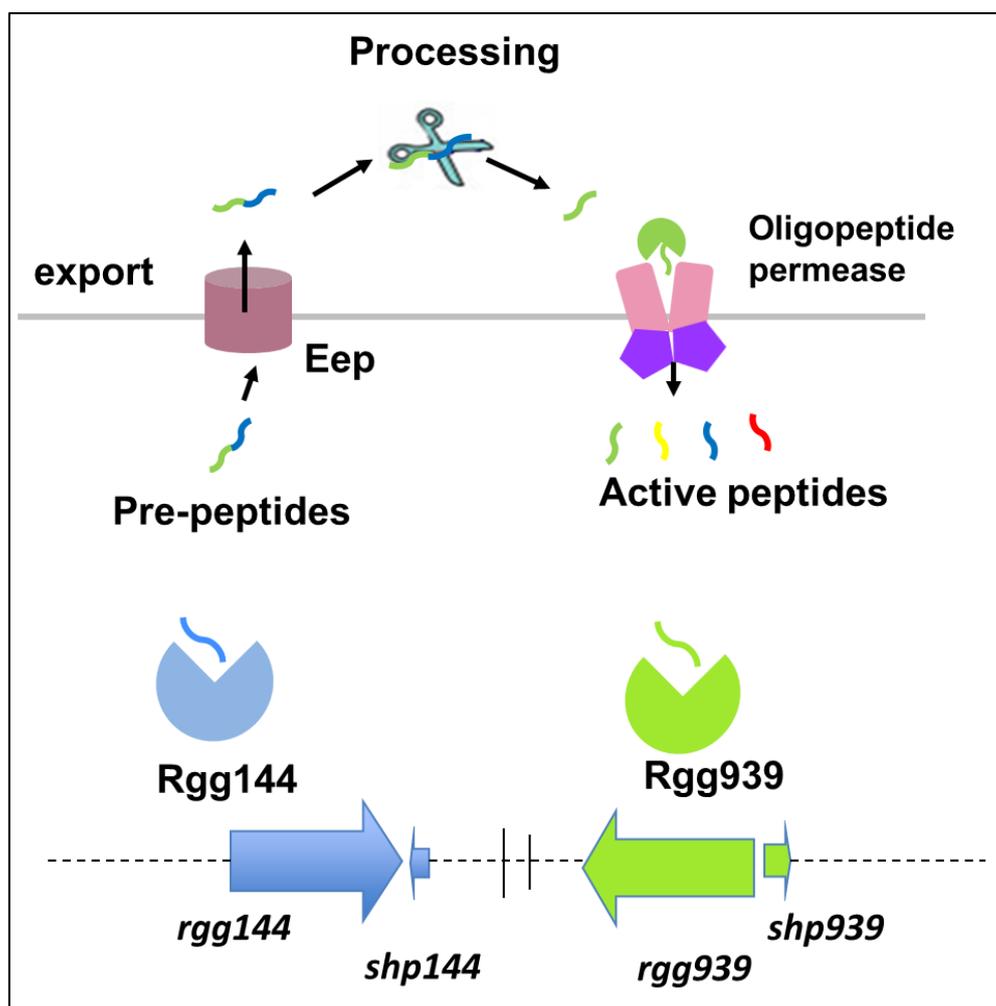


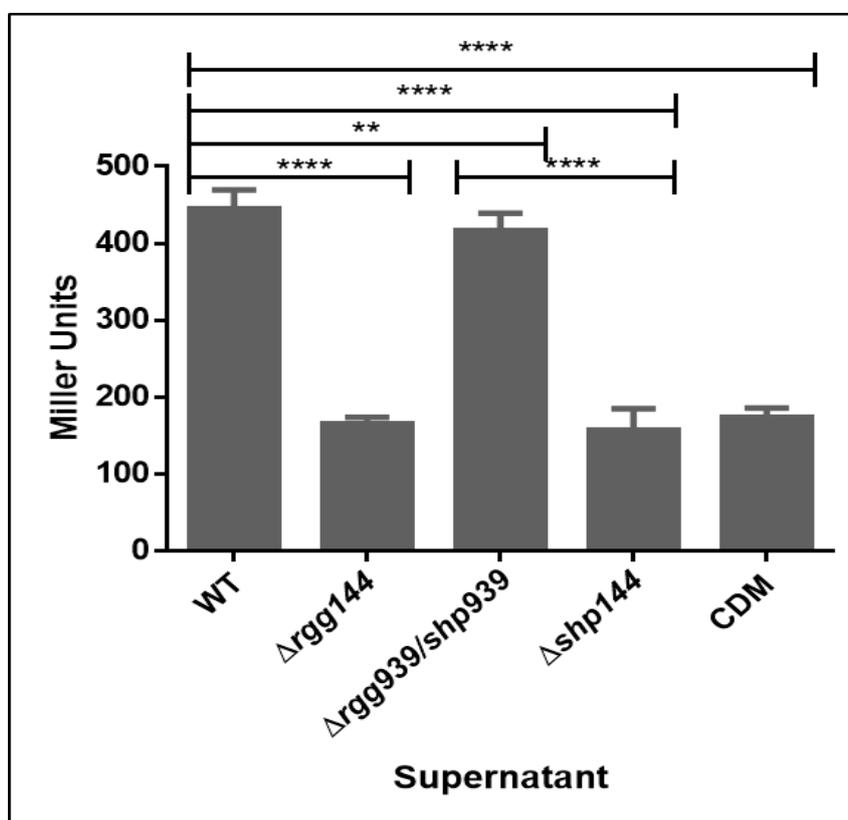
Figure 3.31. Proposed model of Rgg-dependent quorum sensing in *Streptococcus pneumoniae*.

To test the regulatory interaction between Rgg and Shp for each Rgg-Shp circuit, DNA segments likely to encompass the promoters of the *shp144* and *shp939* genes ( $P_{shp144}$  and  $P_{shp939}$ ) were selected for fusion to the bacterial  $\beta$ -galactosidase gene, *lacZ*. Thus,  $P_{shp144}$ -*lacZ* and  $P_{shp939}$ -*lacZ* transcriptional fusions were constructed and integrated into wild type and their respective isogenic mutants. The reporter strains were grown in CDM supplemented with glucose, microaerobically, and  $\beta$ -galactosidase activity was determined to evaluate the responsiveness of respective promoters to the synthetic peptides. Specifically, it was aimed to determine if (1) Shp peptides are secreted, (2) the type active variant of Shp144 and Shp939, (3) if the induction of  $\beta$ -galactosidase activity depends on the concentration of peptide, (4) if Rgg144 or Rgg939 are required for regulation of *shp144* or *shp939*, respectively.

### **3.9.2 Shp144 is a secreted peptide**

A previous study (Chang et al., 2011) in Group A *Streptococcus* reported that *shp* genes encode short nascent peptides, approximately 26 amino acid, which function as signaling molecules involved in quorum sensing pathway to mediate gene expression. Since microbial cell-to-cell communication systems typically rely on secreted signaling molecules in Gram positive bacteria, the Shp144 in pneumococcal genome was proposed to be the precursor of the pheromone. To test if the pheromone is secreted, the late exponential phase cell-free culture supernatants from wild type strain (which contains intact *rgg* and *shp* copies),  $\Delta$ *rgg144*,  $\Delta$ *rgg939/shp939* and  $\Delta$ *shp144* were mixed with  $P_{shp144}$ -*lacZ*- $\Delta$ *shp144* mutant background. The mutant strain background was used for the reporter construct to eliminate induction by endogenously produced Shp144. The expression of the *shp144* promoter was then tested (Figure 3.32). The activity is expressed in nmol *p*-nitrophenol/min/ml. Fresh uninoculated CDM was used as negative control. The  $\beta$ -galactosidase activity of the reporter strain was shown to be  $445.2 \pm 7.0$  MU,  $165.4 \pm 2.3$  MU,  $416.5 \pm 6.5$  MU,  $157.3 \pm 8.7$  MU and  $173.5 \pm 3.8$  when either wild type,  $\Delta$ *rgg144*,  $\Delta$ *rgg939/shp939*,  $\Delta$ *shp144* or CDM, respectively, were used. The  $\beta$ -galactosidase activity of the reporter strain incubated with the supernatant from  $\Delta$ *shp144* was significantly lower than that treated with wild type supernatant ( $p < 0.001$ ). This was consistent with  $\Delta$ *shp144* being devoid of pheromone to promote the induction of *shp144* promoter, and induction of *shp* by their own peptide has been shown in other streptococci (Lasarre et al., 2013, Fleuchot et al., 2011). This demonstrates the functional complementation of the cells of the reporter strain by the late

exponential phase supernatant of the wild type strain, very likely through secreted product of the *shp144*. Moreover, the  $\beta$ -galactosidase activity of the reporter strain  $P_{shp144}\text{-lacZ-}\Delta shp144$  incubated with the late exponential phase supernatant from  $\Delta rgg939/shp939$  was significantly higher than that treated with  $\Delta shp144$  supernatant ( $p < 0.0001$ ), though the activity induced by  $\Delta rgg939/shp939$  strain supernatant was lower than that induced by the wild type spent culture supernatant ( $p < 0.01$ ), indicating that deletion of *rgg939/shp939* affects the production or secretion of Shp144. As expected the  $\beta$ -galactosidase activity in  $P_{shp144}\text{-lacZ-}\Delta shp144$  was similar when the supernatant from  $\Delta rgg144$  or fresh CDM were used ( $p > 0.05$ ), indicating that Rgg144 is required for the induction of *shp144*. In addition, the Rgg144/Shp144 circuit could not respond to the supernatants prepared from early exponential phase.



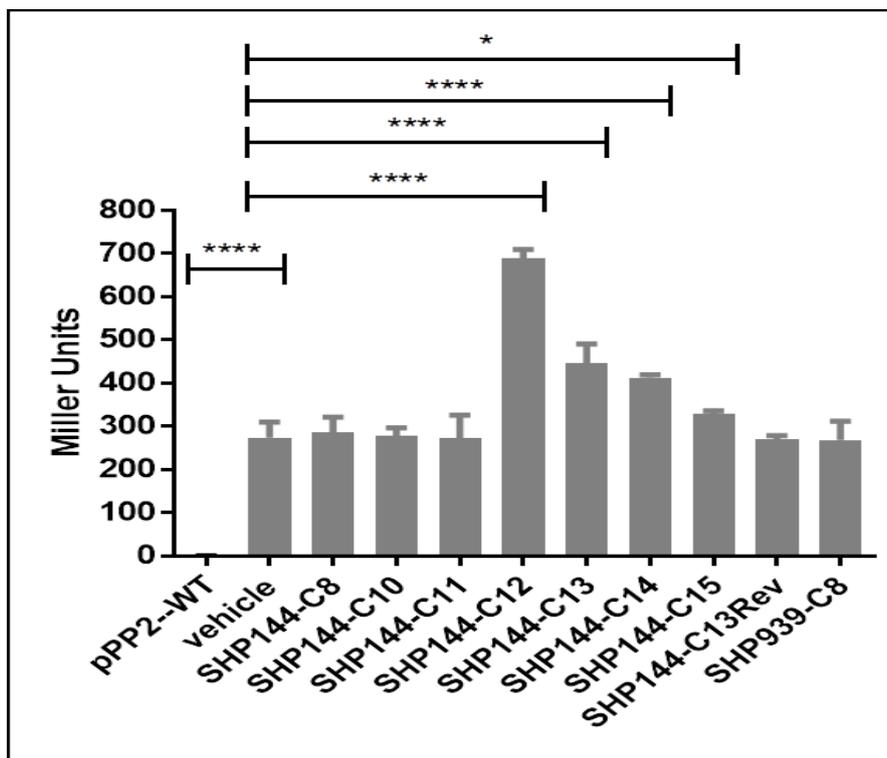
**Figure 3.32.**  $\beta$ -galactosidase activity in  $P_{shp144}\text{-lacZ-}\Delta shp144$  incubated with the late exponential phase culture supernatants from wild type,  $\Delta rgg144$ ,  $\Delta rgg939/shp939$ ,  $\Delta shp144$  and CDM (uninoculated). The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. The error bars indicated the SEM. (\*\* $p < 0.01$ , \*\*\*\*  $p < 0.0001$ ).

Similar experiments were done to test the expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoter of *shp939* in *shp939* mutant background (strain XZ36) mixed with cell-free culture supernatant prepared from late exponential phase wild type,  $\Delta rgg939$ ,  $\Delta rgg144/shp144$ ,  $\Delta shp939$  and CDM (uninoculated) supplemented with 55 mM of glucose. None of the added culture supernatants had the ability to increase the induction of  $\beta$ -galactosidase activity of *lacZ*-fusions to the promoters of *shp939* in *shp939* mutant background. This was thought to be due low level secreted peptide to reach the threshold to trigger the expression in the assay condition used.

### **3.9.3 SHP144-C12 is the most active variant of SHP144 peptides**

Previous studies (Chang et al., 2011, Aggarwal et al., 2014) demonstrated that 8 amino acid residue peptide representing C-terminal portion (15 to 23) of SHP3 in *S. pyogenes*, which also have Rgg-Shp QS systems, is necessary for  $P_{shp3-luxAB}$  induction. To determine the amino acid sequence of the active form of the secreted SHP144 pheromone, different length synthetic peptides, 8 to 15 amino acid residues, representing C-terminal end of SHP144, SHP-C8, -C10, -C11, -C12, -C13, -C14, -C15, - were added into the cultures of XZ3 ( $P_{shp144-lacZ}$ - wt) which contains intact copies of *rgg/shp*. In addition, C13Rev, reverse C13 peptide, and SHP939-C8, representing 8 residue long SHP939 peptide, were included in the assay as control peptides to determine the specificity of induction. The strain XZ0 containing promoterless pPP2 (pPP2-WT) was used as a control to demonstrate if an inducible promoter is present in the reporter strain. In addition, XZ3 ( $P_{shp144-lacZ}$ - wt) without added peptide was also included as a control (vehicle). The  $\beta$ -galactosidase activity of  $P_{shp144}$  reporter in response to different size synthetic SHP variants was illustrated in Figure 3.33. The Expression levels (in Miller units) were shown in Table 3.4. While SHP144-C13, -C14, and -C15 could also induce the expression of the *shp* promoter significantly compared to the vehicle ( $447.2 \pm 12.6$  MU,  $412.9 \pm 3.2$  MU, and  $328.6 \pm 3.8$  MU, respectively), SHP144-C12 generated the highest  $\beta$ -galactosidase activity ( $689.4 \pm 7.1$  MU), compared to vehicle ( $274.2 \pm 10.3$  MU) ( $p < 0.0001$ ). These data suggest that SHP144-C12 is the most active variant of SHP144 peptides, and this form is possibly the putative product of a C-terminal cleavage of the 23 amino acid peptide precursor. However, SHP144-C8, -C10 and -C11 could not induce  $P_{shp144}$  driven  $\beta$ -galactosidase activity above background level ( $286.3 \pm 12.4$  MU,  $278.4 \pm 6.5$  MU,  $272.9 \pm 19.0$  MU, respectively), compared to the vehicle ( $p > 0.05$ ). Moreover, SHP939-C8 does not induce the  $P_{shp144}$  expression ( $270.5 \pm 4.3$  MU) compared to the vehicle,

( $p > 0.05$ ).



**Figure 3.33.** Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoters of *shp144* in a wild-type background strain (D39) grown microaerobically in CDM supplemented with different synthetic SHP144 variant or vehicle (no added synthetic peptide). The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. (\*  $p < 0.05$ , \*\*\*\*  $p < 0.0001$ ). Error bars indicate the SEM.

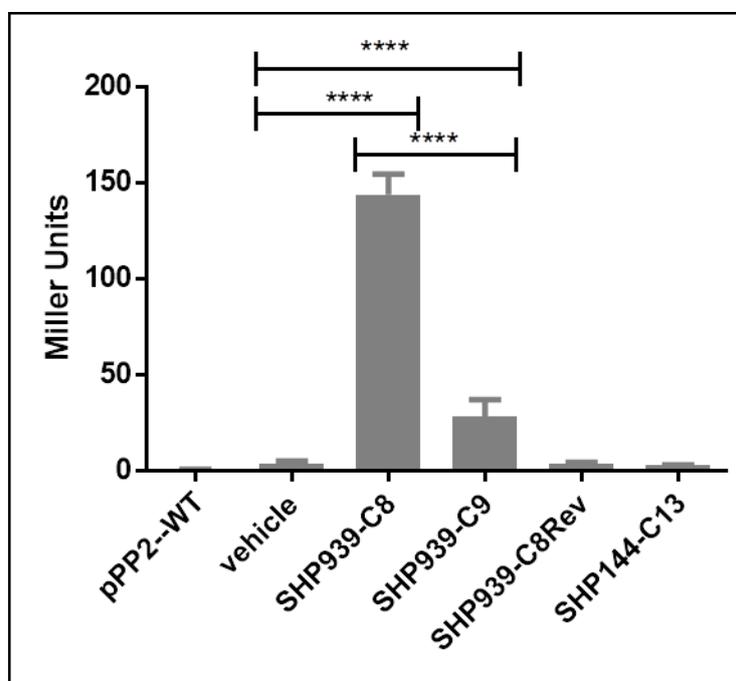
**Table 3.4. Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoters of *shp144* in a wild-type background strain (D39) in CDM supplemented with different synthetic SHP144 variant or vehicle (without adding any synthetic peptide). The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. ‘±’ indicates standard error of means (SEM).**

Peptide	Sequence	β-galactosidase activity (Miller units)
vehicle	-----	274.2 ± 10.3
SHP144-C8	VIPFLTNL	286.3 ± 12.4
SHP144-C10	VIVIPFLTNL	278.4 ± 6.5
SHP144-C11	WVIVIPFLTNL	272.9 ± 19.0
SHP144-C12	EWVIVIPFLTNL	689.4 ± 7.1
SHP144-C13	SEWVIVIPFLTNL	447.2 ± 12.6
SHP144-C14	ISEWVIVIPFLTNL	412.9 ± 3.2
SHP144-C15	LISEWVIVIPFLTNL	328.6 ± 3.8
SHP144-C13Rev	LNTLFPVIVWES	268.7 ± 15.5
SHP939-C8	DIIIVGG	270.5 ± 4.3

#### **3.9.4 SHP939-C8 is the most active variant of SHP939 peptides**

To determine the most active variant of SHP939 pheromone, different length of synthetic peptides, SHP-C8, -C9, -C8Rev (reverse peptide) and SHP144-C13 variants, were added into the cultures of  $P_{shp939}$  reporter in a wild type background strain (XZ4) As SHP939 is 90% identity to SHP3 of *S. pyogenes*, Chang et al., (2011) already showed SHP3-C8 is the most active variant of SHP3, so in this study, we selected only SHP939-C8 and C9 to test the induction of  $P_{shp939}$ . As before, the strain XZ0 containing promoterless pPP2 plasmid (pPP2-WT) was used as a control to demonstrate if an inducible promoter is present in the reporter strain, and XZ4 ( $P_{shp939}$ -*lacZ*-wt) without added peptide was included to show that the

induction is due to added peptide. The  $\beta$ -galactosidase activity of  $P_{shp939}$  reporter in response to synthetic SHP variants was illustrated in Figure 3.34. The Expression levels (in Miller units) were shown in Table 3.5. Without adding synthetic peptide (vehicle), the *shp939* promoter was not induced. The  $\beta$ -galactosidase activity of XZ4 containing  $P_{shp939}$ -*lacZ* construct was  $3.7 \pm 0.4$  MU. This was because wild type strain (XZ4) could not produce sufficient quantities of pheromone to induce Rgg939 signaling autonomously, and therefore it was concluded that  $P_{shp939}$  required an exogenous supply of pheromone to induce observable  $\beta$ -galactosidase activity in reporter assay. Similar result was obtained when the reporter strain treated with SHP939-C8 reverse peptide ( $3.5 \pm 0.3$  MU) compared to the vehicle ( $p>0.05$ ). On the other hand, as expected both SHP-C8 and C9 were capable of inducing  $\beta$ -galactosidase activity in XZ4 strain ( $P_{shp939}$  - *lacZ*- wt) significantly at the concentrations tested (250 nM) compared to the activity in vehicle ( $p<0.0001$ ). In contrast to SHP939-C9 ( $28.2 \pm 4.4$  MU), C8 induction resulted in much higher activity ( $143.7 \pm 3.1$  MU) in XZ4 ( $p<0.0001$ ). This result demonstrated that SHP939-C8 is the most active variant of SHP939 peptides. Moreover, the induction by SHP939-C8 was specific as SHP144-C13 peptide could not induce the promoter of *shp939* ( $3.0 \pm 0.1$  MU), ( $p>0.05$ ).



**Figure 3.34.  $\beta$ -Galactosidase activity of the XZ4 in response to synthetic SHP variants.** Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoters of *shp939* in a wild-type background strain (D39) grown microaerobically in CDM supplemented with different synthetic SHP939 variants, or control peptides. The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent

experiments each with three replicates. Error bars indicated the SEM. (\*\*\*\* $p < 0.0001$ ). Vehicle control is XZ4 ( $P_{shp939-lacZ-wt}$ ) without adding any peptide.

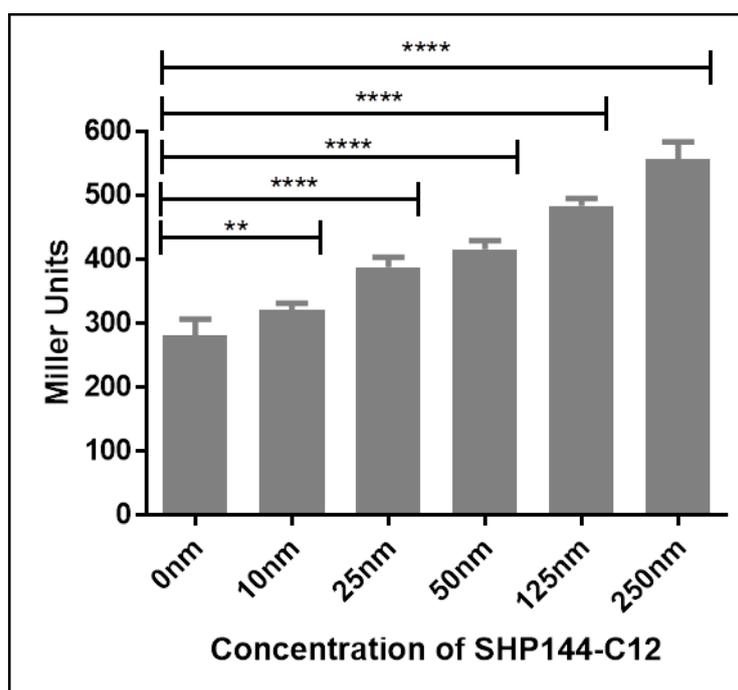
**Table 3.5. Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoters of *shp939* in a wild-type background strain (D39) in CDM supplemented with different synthetic SHP939 variant or vehicle (without adding any synthetic peptide).** The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. ‘±’ indicates standard error of means (SEM).

Peptide	Sequence	β-galactosidase activity (Miller units)
vehicle	-----	3.7 ± 0.4
SHP939-C8	DIIIIVGG	143.7 ± 3.1
SHP939-C9	MDIIIIVGG	28.2 ± 4.4
SHP939C8REV	GGVIIIID	3.5 ± 0.3
SHP144-C13	SEWVIVIPFLTNL	3.0 ± 0.1

### **3.9.5 Induction of β-galactosidase activity depends on the concentration of peptide**

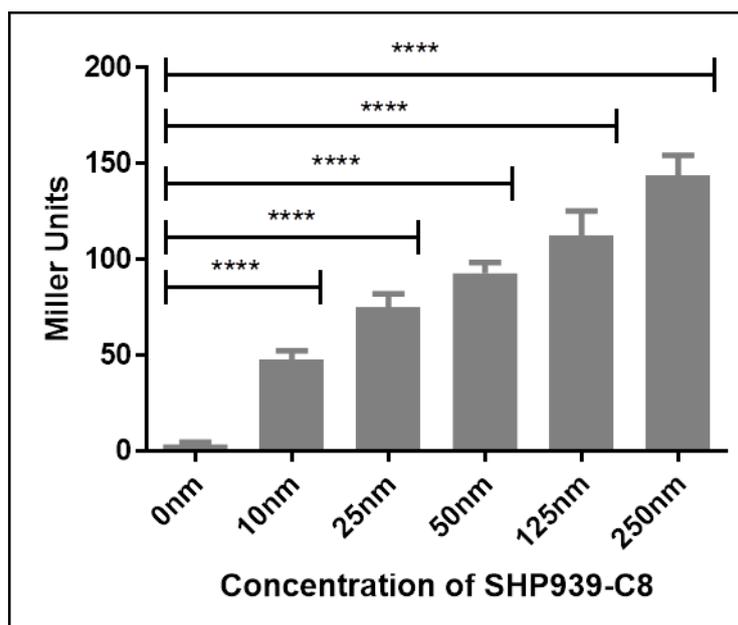
The experiments described above demonstrated that various SHP peptides are produced by *S. pneumoniae* and that each variant has a distinct potential to induce Rgg-dependent transcription in reporter strains. Quorum-sensing mechanisms involve signalling molecules reaching a threshold concentration to trigger the expression of their target gene(s) (Winzer et al., 2002, Podbielski and Kreikemeyer, 2004, Siehnel et al., 2010). To check whether this applies to the Rgg/SHP144 and Rgg/SHP939 systems, the expression of the *shp144* and *shp939* promoters were measured in XZ3 ( $P_{shp144-lacZ-wt}$ ) and XZ4 ( $P_{shp939-lacZ-wt}$ ) supplemented with individual specific synthetic peptides at different concentrations (Fig. 3.35 and 3.36). The strain XZ0 (pPP2-WT) was used as a control to demonstrate if an active promoter was present in the reporter strains. The XZ3 ( $P_{shp144-lacZ-wt}$ ) without synthetic peptide addition was used as negative control. The β-galactosidase activity of the XZ3 strain

was shown to be  $280.6 \pm 9.2$  MU,  $320.6 \pm 4.0$  MU,  $388.5 \pm 5.4$  MU,  $416.0 \pm 4.9$  MU,  $484.1 \pm 4.1$  MU and  $557.0 \pm 9.7$  MU when the culture medium was added with 0 nM, 10 nM, 25 nM, 50 nM, 125 nM, and 250 nM of synthetic SHP144-C12, respectively, ( $p < 0.01$ ,  $p < 0.0001$ ,  $p < 0.0001$ ,  $p < 0.0001$ ,  $p < 0.0001$ ). Increased induction of  $\beta$ -galactosidase activity by the SHP144-C12 was observed, and was correlated with its concentration in the medium.



**Figure 3.35.** Expression levels (in Miller units) of XZ3 ( $P_{shp144-lacZ-wt}$ ) in response to different concentrations of synthetic SHP144-C12. The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. Error bar indicate the SEM. (\*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ , relative to the XZ3 ( $P_{shp144-lacZ-wt}$ ) without adding any synthetic peptide.)

Similar results were obtained with  $P_{shp144}$  induction. The  $\beta$ -galactosidase activity of the XZ4 ( $P_{shp939-lacZ-wt}$ ) was shown to be  $3.7 \pm 0.4$  MU,  $48.1 \pm 1.3$  MU,  $75.2 \pm 2.0$  MU,  $93.0 \pm 1.6$  MU,  $112.5 \pm 3.7$  MU and  $143.7 \pm 3.1$  MU when treated with 0 nM, 10 nM, 25 nM, 50 nM, 125 nM, and 250 nM of synthetic SHP939-C8, respectively ( $p < 0.0001$ ), (Figure 3.36).

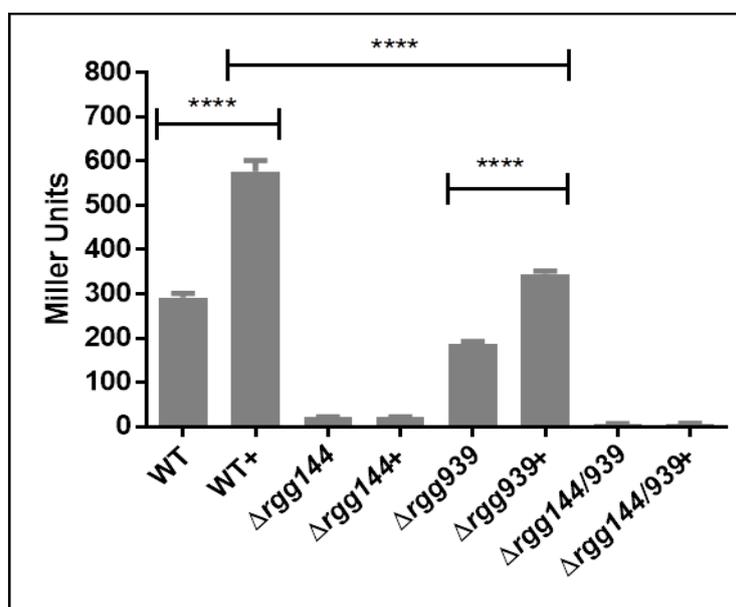


**Figure 3.36. Expression levels (in Miller units) of XZ4 ( $P_{shp939}$ -*lacZ*-wt) in response to different concentrations of synthetic SHP939-C8.** The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of three independent experiments each with three replicates. Error bars indicate the SEM. (\*\*\*\* $p < 0.0001$ , relative to the XZ4 ( $P_{shp939}$ -*lacZ*-wt) without adding any synthetic peptide).

### **3.9.6 Rgg144 is required for C12 induction of $P_{shp144}$**

To evaluate the role of Rggs in *shp144* regulation,  $P_{shp144}$ -*lacZ* fusion was transformed into wild type,  $\Delta rgg939$ ,  $\Delta rgg144$  and  $\Delta rgg144/939$  strains. It was hypothesised that if Rggs had any role in *shp144* regulation,  $\beta$ -galactosidase activity in strain containing  $P_{shp144}::lacZ$  construct would be different between the wild type and mutants. The  $\beta$ -galactosidase activity of these strains were determined in CDM with our without addition of SHP144-C12. The  $\beta$ -galactosidase activity conferred by  $P_{shp144}$ -*lacZ* construct in the wild type background was  $291.5 \pm 3.3$  MU without peptide, but when SHP144-C12 was added, the  $\beta$ -galactosidase activity increased significantly to  $577.1 \pm 8.8$  MU, ( $p < 0.001$ ). This result shows that wild type strain with intact copies of *rgg/shp* could respond to the synthetic SHP144-C12 peptide. On the other hand, the activity level in  $P_{shp144}$ -*lacZ*- $\Delta rgg939$  strains was  $344.8 \pm 2.5$  MU with the addition of SHP939-C8, but only  $186.3 \pm 2.4$  MU without peptide ( $p < 0.0001$ ). Thus in the absence of *rgg939*, SHP144-C12 still could induce the expression of  $P_{shp144}$ , but the  $\beta$ -

galactosidase activity was lower than that induced by  $P_{shp144-lacZ}$  in wild type background, indicating that Rgg939 is required for full induction of  $P_{shp144}$ . However, when  $P_{shp144-lacZ}$  construct was in the  $\Delta rgg144$  background, there was no induction above background level regardless of addition of SHP144-C12. Indeed, the activity level was  $21.5 \pm 0.2$  MU and  $21.5 \pm 0.3$  MU, with or without the addition of SHP144-C12, respectively, ( $p > 0.05$ ). Similarly, in  $\Delta rgg144/939$ , the activity level of  $P_{shp144-lacZ}$  was  $6.1 \pm 0.3$  MU without peptide, and it was  $7.1 \pm 0.2$  MU with the peptide as expected ( $p > 0.05$ ).  $P_{shp144-lacZ}$  could not respond to SHP144-C12 in  $\Delta rgg144$  and  $\Delta rgg144/939$  backgrounds (Figure 3.37). These data demonstrated that *rgg144* is necessary for  $P_{shp144}$  responsiveness to SHP144-C12, and the presence of Rgg939 is required for full induction of  $P_{shp144}$ .

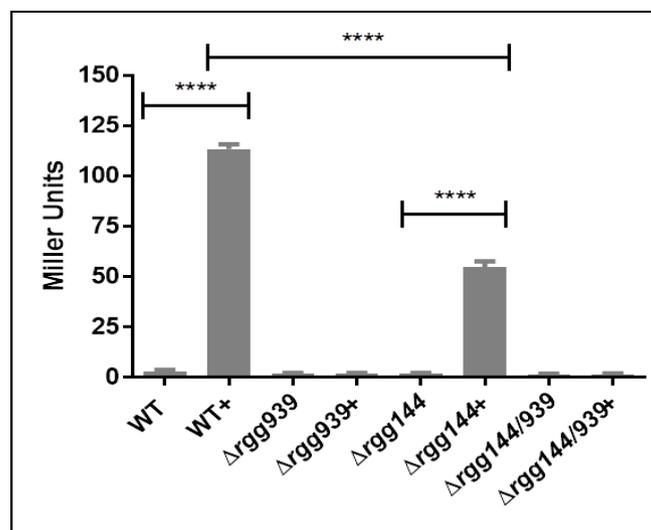


**Figure 3.37.** Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoter of *shp144* in wild type or  $\Delta rgg144$ ,  $\Delta rgg939$  and  $\Delta rgg144/939$  with (+) or without SHP144-C12 synthetic peptide. The strains were grown micro-aerobically in CDM supplemented with 55 mM of glucose. The activity is expressed in nmol *p*-nitrophenol/min/ml. ( $n=9$ , \*\*\*\*  $p < 0.0001$ ). Error bars indicate the SEM.

### **3.9.7 Rgg939 is required for C8 induced activation of $P_{shp939}$**

To determine the function of Rggs in *shp939* regulation,  $P_{shp939-lacZ}$  fusion was transformed into the wild type, and the mutant backgrounds:  $\Delta rgg939$ ,  $\Delta rgg144$  and  $\Delta rgg144/939$ . As

before, the  $\beta$ -galactosidase activity was determined in CDM with or without addition of inducing SHP939-C8 peptide. The results are shown in Figure 3.38. The  $\beta$ -galactosidase activity of the  $P_{shp939}$ -*lacZ* fusion in the wild type background was  $3.1 \pm 0.7$  MU without peptide, but when SHP939-C8 was added, the  $\beta$ -galactosidase activity increased significantly to  $113.6 \pm 0.8$  MU ( $p < 0.001$ ). This result shows that  $P_{shp939}$ -*lacZ* strain in wild type background with intact copies of *rgg/shp* could respond to the synthetic SHP939-C8 peptide. However, the activity level driven by  $P_{shp939}$ -*lacZ* in the  $\Delta$ *rgg939* background was  $1.9 \pm 0.2$  MU and  $1.9 \pm 0.1$  MU with and without the addition of SHP939-C8, respectively, ( $p > 0.05$ ). This shows that without Rgg939,  $P_{shp939}$ -*lacZ* cannot respond to SHP939-C8, indicating that Rgg939 is required for  $P_{shp939}$  responsiveness to SHP939-C8. As expected, in  $\Delta$ *rgg144/939*, the activity level was  $1.6 \pm 0.2$  MU and  $1.7 \pm 0.1$  MU with and without the addition of SHP939-C8, respectively. There was no significant difference between presence and absence of synthetic peptide in  $\Delta$ *rgg144/939* background ( $p > 0.05$ ). On the other hand, the activity level driven by  $P_{shp939}$ -*lacZ* was  $55.0 \pm 0.9$  MU in the  $\Delta$ *rgg144* with the addition of SHP939-C8, but only  $1.9 \pm 0.11$  MU without peptide ( $p < 0.0001$ ). This result shows that Rgg144 is required for full induction of the promoter of *rgg939*, because the induction of  $P_{shp939}$ -*lacZ* by the peptide in wild type background was significantly higher than that of  $\Delta$ *rgg144* ( $p < 0.0001$ ).



**Figure 3.38.** Expression levels (in Miller units) of pneumococcal transcriptional *lacZ*-fusions to the promoter of *shp939* in wild type or  $\Delta$ *rgg144*,  $\Delta$ *rgg939* and  $\Delta$ *rgg144/939* with (+) or without SHP939-C8 synthetic peptide. The strains were grown microaerobically in CDM supplemented with 55 mM of glucose. The activity is expressed in nmol *p*-nitrophenol/min/ml ( $n=9$ , \*\*\*\*  $p < 0.0001$ ). Error bars indicate the SEM.

### **3.9.8 The role of carbon source on the expression of *rggs***

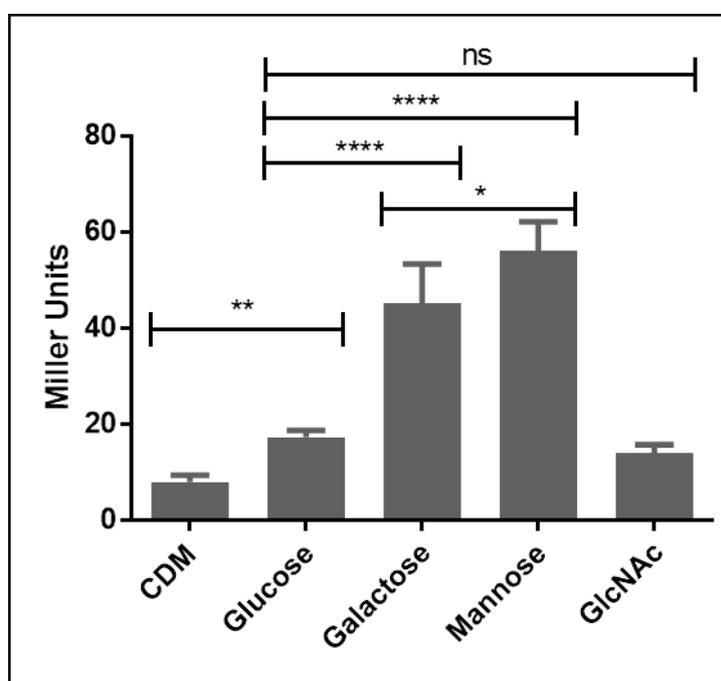
Microbial nutrient acquisition and metabolic pathways are very important for pneumococcal adaptation and *in vivo* fitness. Therefore, I investigated Rgg/Shp circuits role in nutrient metabolism. Initially, their induction by different sugars was determined. In this study, glucose, galactose, mannose and *N*-acetyl glucosamine (GlcNAc) were used to determine the role of the Rgg regulators in host derived sugar metabolism, as these sugars are abundant in the human respiratory tract or in blood (Rose and Voynow, 2006).

In order to evaluate the responsiveness of *rgg* promoters to different carbon source, the reporter strains XZ1 ( $P_{rgg144}$ -*lacZ*-wt) and XZ3 ( $P_{rgg939}$ -*lacZ*-wt) were grown in CDM supplemented with glucose, galactose, mannose or *N*-acetyl glucosamine microaerobically, and  $\beta$ -galactosidase activity was determined. The CDM without adding any sugar was used as negative control ( $8.0 \pm 0.6$  MU). The results indicated that the highest induction of *lacZ* was obtained when XZ1 was grown on mannose ( $56.1 \pm 2.5$ ,  $n=9$ , \*\*\*\* $p<0.0001$  compare to glucose), then by galactose ( $45.2 \pm 3.3$ ,  $n=9$ , \*\*\*\* $p<0.0001$  compare to glucose) and glucose ( $17.3 \pm 0.6$ ,  $n=9$ ), while the presence of *N*-acetyl glucosamine led to the lowest  $\beta$ -galactosidase activity ( $14.0 \pm 0.7$ ,  $n=9$ ) (Figure 3.39) (Table 3.6). The growth on mannose resulted in higher induction than galactose ( $p<0.05$ ). Similarly, the growth in glucose led to a significantly higher induction than CDM control that did not contain any sugar ( $p<0.01$ ).

The XZ3 ( $P_{rgg939}$ -*lacZ* -wt) displayed a similar expression profile as XZ1 ( $P_{rgg144}$ -*lacZ*-wt) in mannose, and the highest activity was obtained on mannose ( $18.6$  MU  $\pm$  1.7,  $n=9$ ,  $p<0.001$  relative to glucose), then by galactose ( $12.1 \pm 1.2$ ,  $p>0.05$  relative to glucose) and the lowest was on *N*-acetyl glucosamine ( $7.5$  MU  $\pm$  0.9,  $n=9$ ,  $p>0.05$  relative to glucose) (Table 3.7). The induction by mannose was higher than that by galactose ( $12.1 \pm 1.2$  MU), ( $p<0.01$ ). These results suggest that the induction of *rgg* promoters is influence by the type of sugars used, and it is very likely that *rgg144* and *rgg939* play an important role in mannose and galactose metabolism.

**Table 3.6.** Expression levels of strain XZ1 grown microaerobically in CDM supplemented with 55 mM of glucose, galactose, mannose or *N*-acetyl glucosamine. '±' indicates standard error of means (SEM).

Carbon source	Miller Units
CDM	8.0 ± 0.6
Glucose	17.3 ± 0.6
Galactose	45.2 ± 3.3
Mannose	56.1 ± 2.5
GlcNAc	14.0 ± 0.7



**Figure 3.39.** Expression levels of strain XZ1 grown microaerobically in CDM supplemented with 55 mM of glucose, galactose, mannose or *N*-acetyl glucosamine (GlcNAc). The activity is expressed in nmol *p*-nitrophenol/min/ml. Values are average of at least three independent experiments each with three replicates. Error bars indicate the SEM. (\* $p < 0.05$ , \*\* $p < 0.01$ , \*\*\*\* $p < 0.0001$ , ns: no significant).

**Table 3.7. Expression levels of strain XZ3 grown microaerobically in CDM supplemented with 55 mM of glucose, galactose, mannose or *N*-acetyl glucosamine. '±' indicates standard error of means (SEM).**

<b>Carbon source</b>	<b>Miller Units</b>
CDM	4.7 ± 0.3
Glucose	9.8 ± 0.5
Galactose	12.1 ± 1.2
Mannose	18.6 ± 1.7
GlcNAc	7.5 ± 0.9

## **Section D. Determination of Rgg144 and Rgg939 by Microarray analysis**

DNA microarrays are small, solid supports onto which the sequences from thousands of different genes are immobilised, or attached at fixed locations. They are used to obtain a comprehensive genome-wide survey of gene expression patterns by using standard statistical algorithms to arrange genes according to similarity in pattern of gene expression (Brown and Botstein, 1999). In addition, they are utilised for various other applications, such as for detection of different microbial species, cell types, for evaluation of effect of different stimuli on cells, and for assessment of gene function. When microarrays are used for gene expression analysis, fluorescently labeled cDNA is hybridized to microarray chip, and subsequently the bound cDNA is detected using laser technology. Then, the data is analysed using various computational methods (Harrington et al., 2000).

To reveal the wider influence of Rggs on pneumococcal biology and the genes regulated directly or indirectly by Rggs, D39,  $\Delta rgg144$  and  $\Delta rgg939$  were analysed by using DNA microarrays in collaboration with Prof Oscar Kuipers and Dr Sulman Shafeeq, University of Groningen, the Netherlands. D39,  $\Delta rgg144$  and  $\Delta rgg939$  strains were inoculated microaerobically in CDM supplemented with 55 mM, galactose or mannose as the main source of carbon, due to inducibility of *rgg* genes by these sugars, and then total RNA was extracted in Leicester. The samples were shipped to Groningen for microarray analysis.

### **3.10 Microarray analysis for $\Delta rgg144$**

The largest number of differentially expressed genes in *rgg* mutants was seen on mannose. On mannose, 154 genes were differentially expressed in  $\Delta rgg144$  relative to the wild type; of these 131 increased in expression and 23 decreased (Appendix). These findings confirm the role of Rgg144 as a global regulator with a prevalent repressor function. All of the significantly differentially transcribed ORFs have been classified into COG category (Figure 3.40A and B). The notable genes upregulated by Rgg144 included those putatively involved in i.) replication, recombination and repair, ii.) translation, ribosomal structure and biogenesis, iii.) capsule biosynthesis, iv.) nucleotide, transport and metabolism, and v.) those coding for hypothetical proteins. Furthermore, the locus adjacent to Rgg SPD\_1518, encoding

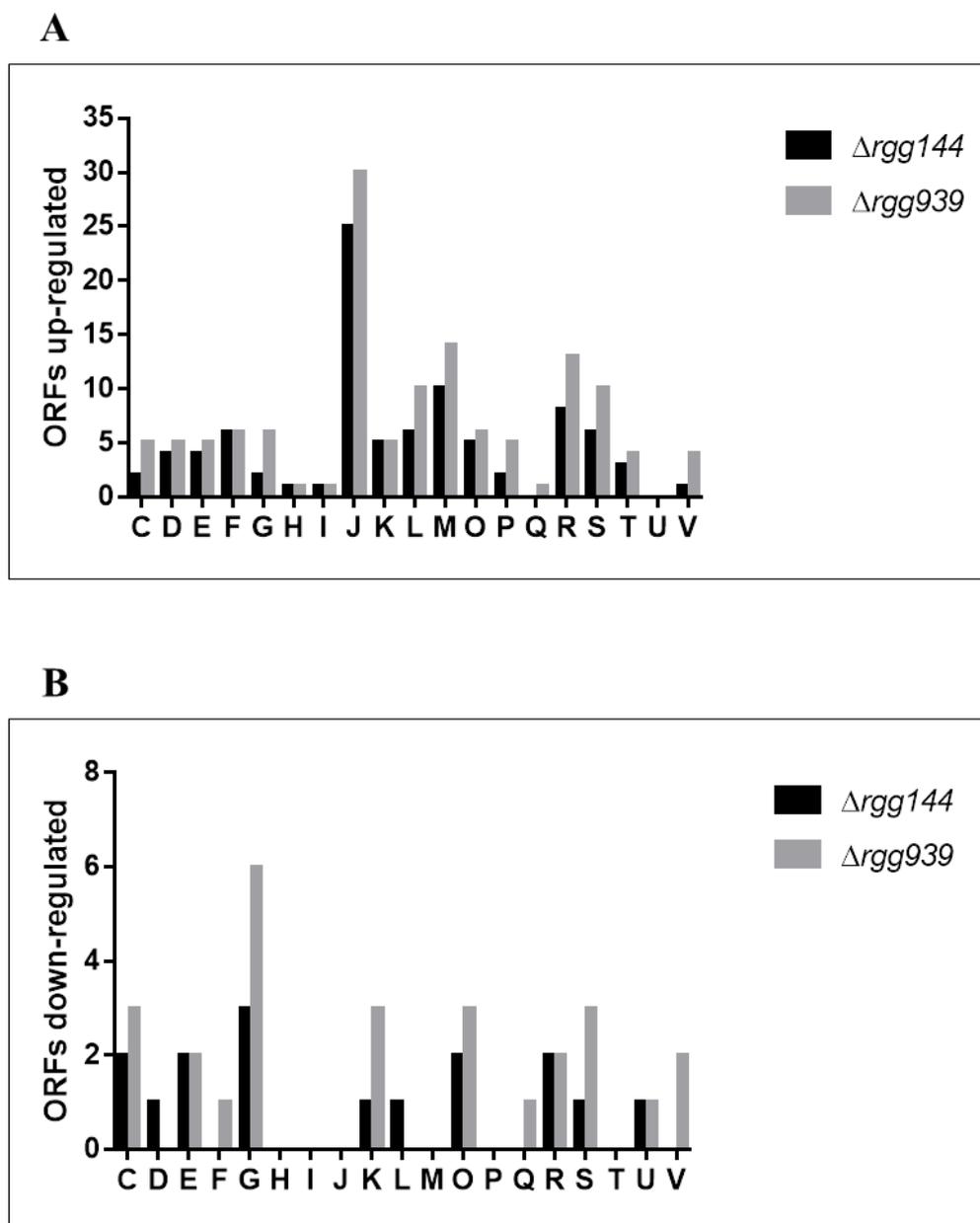
SPD\_1513-SPD\_1517, is also negatively regulated by Rgg144. This reveals a potential regulatory interaction between Rgg144 and Rgg1518. The notable genes upregulated by Rgg144 included those putatively involved in oxidative stress response (*gor* oxidoreductase (SPD\_0685), sugar metabolism and amino acid transport and metabolism (*pepA*, *pepC* and *pepQ*). The genes positively regulated by Rgg144 included the adjacent VP1 peptide and downstream genes (SPD\_145-147), which have been shown to have a role in biofilm formation and virulence, and to be regulated by Rgg144 (Cuevas et al., 2017).

### **3.11 Microarray analysis for $\Delta$ rgg939**

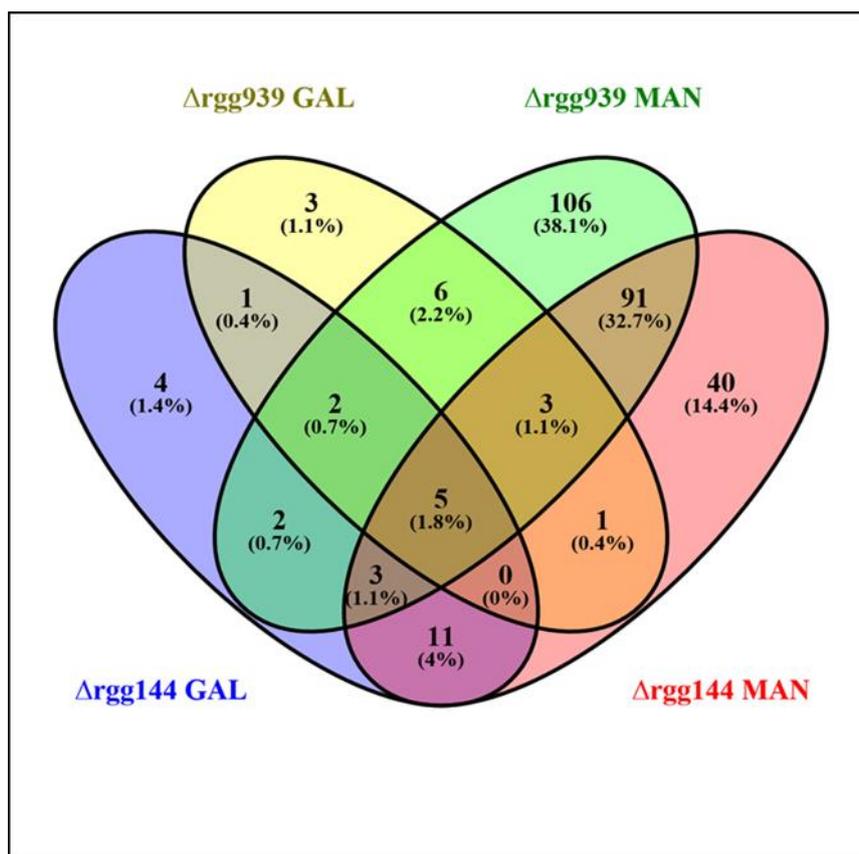
218 genes were differentially regulated in the wildtype versus the Rgg939 deletion mutant. Of these 177 are negatively regulated and 41 positively regulated by Rgg939 (Appendix). All of the significantly differentially transcribed ORFs have been classified into COG category (Figure 3.40A and B). Interestingly, the genes upregulated in  $\Delta$ rgg939 had a substantial overlap to those upregulated in  $\Delta$ rgg144. The Venn diagram (Figure 3.41) clearly shows the relative transcript levels in  $\Delta$ rgg144 and  $\Delta$ rgg939 in mannose were compared to  $\Delta$ rgg144 and  $\Delta$ rgg939 in galactose as generated by VENNY (<http://bioinfogp.cnb.csic.es/tools/venny/index.html>). There are 102 genes that are significantly differentially transcribed in both  $\Delta$ rgg144 Man and  $\Delta$ rgg939 Man, showing that Rgg proteins have a core regulon. These included genes putatively encoding for bacteriocin (BlpU) (SPD\_0046), amino acid transport and metabolism (SPD\_0150), inorganic ion transport and metabolism (SPD\_0151), replication, recombination and repair (SPD\_0179-SPD\_0181), translation, ribosomal structure and biogenesis (SPD\_191-SPD\_219), capsule biosynthesis locus (SPD\_0315-SPD\_0327), cell wall/membrane/envelope biogenesis (SPD\_1200), and lipid transport and metabolism (SPD\_1127). In addition to this overlap, a number of loci were found to be differentially regulated only by Rgg939. These included genes encoding for putative cell division proteins (SPD\_0007-SPD\_0011), iron transport (SPD\_0915-SPD\_0920), cell membrane biogenesis (SPD\_0940-SPD\_0950), ATP synthase (SPD\_1338-SPD\_1340), and choline transport (SPD\_1642-SPD\_1644). Moreover, similar to Rgg144, it was found that Rgg939 also influences the expression of genes regulated by other *rgg* genes. Specifically, Rgg939 upregulates the Rgg144-regulated VP1 locus (SPD\_0145-SPD\_0147). These interactions suggest cooperative behaviors across these Rggs. Moreover, the regulon overlap suggests that Rgg proteins have a core regulon that may be related to

generalized functions of this protein family, and finally the differences between the regulons demonstrates that each Rgg also has specific roles under the same environmental condition.

When the mutants were grown on galactose, the size of putative regulon for both *rgg* genes were smaller than that on mannose (Appendix). For Rgg144, the affected loci included SPD\_0144-SPD\_0149 and SPD\_1514-SPD\_1516, which were down regulated in the mutant relative to the wild type. On the other hand, in  $\Delta rgg939$ , SPD\_0146-SPD\_0147 and SPD\_0940-SPD\_0950 expression went up, while the expression of SPD\_1513-SPD\_1516 went down. Downregulation of SPD\_1513-SPD\_1516 was noteworthy because this locus is upregulated in  $\Delta rgg939$  on mannose. This shows that under different environmental conditions, the same Rgg can act either as repressor or the activator for the same target gene. We also found that the regulation exerted by different Rggs on the same target is influenced by the carbon source. For example, for SPD\_145-SPD\_147 while Rgg939 is a repressor on galactose, Rgg144 acts as an activator for the same locus on mannose.



**Figure 3.40. Numbers of genes significantly differentially transcribed in  $\Delta rgg144$  (black bars) or  $\Delta rgg939$  (grey bars) on mannose relative to the wild type. (A.) The up-regulated ORFs. (B.) The down-regulated ORFs. Gene classes: [C] Energy production and conversion; [D] Cell cycle control, cell division; chromosome partitioning; [E] Amino acid transport and metabolism; [F] Nucleotide transport and metabolism; [G] Carbohydrate transport and metabolism; [H] Coenzyme transport and metabolism; [I] Lipid transport and metabolism; [J] Translation, ribosomal structure and biogenesis; [K] Transcription; [L] Replication, recombination and repair; [M] Cell wall/membrane/envelope biogenesis; [O] Posttranslational modification, protein turnover, chaperones; [P] Inorganic ion transport and metabolism; [Q] Secondary metabolites biosynthesis, transport and catabolism; [R] General function prediction only; [S] Function unknown; [T] Signal transduction mechanisms; [U] Intracellular trafficking, secretion, and vesicular transport; [V] Defense mechanisms;**

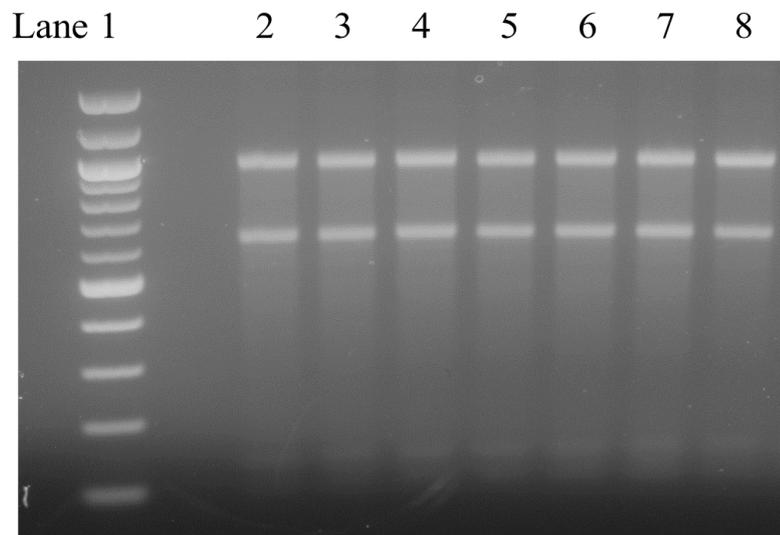


**Figure 3.41. Venn diagram of the genes significantly differentially transcribed due to loss of *rgg144* and *rgg939*.** The relative transcript levels in  $\Delta rgg144$  and  $\Delta rgg939$  in mannose were compared to  $\Delta rgg144$  and  $\Delta rgg939$  in galactose. For all intersections, which are not drawn to scale, the numbers of genes are indicated. The total number of genes influenced in each condition tested was:  $\Delta rgg144$  Man: 154 genes,  $\Delta rgg939$  Man: 218,  $\Delta rgg144$  Gal: 28 genes and  $\Delta rgg939$  Gal: 21 genes.

### **3.12 Confirmation of microarray results by qRT-PCR**

Microarray results were confirmed by quantitative reverse transcriptase PCR (qRT-PCR). The pneumococcal strains wild type D39,  $\Delta rgg144$  and  $\Delta rgg939$  were inoculated microaerobically in CDM supplemented with 55 mM glucose, galactose or mannose as the carbon source. Total RNA was prepared from mid-exponential phase, and was treated with DNase I to eliminate the DNA contamination. The total RNA was confirmed by agarose gel electrophoresis (Figure 3.42). Lanes 2-8 show the total extracted RNA from wild type D39,

$\Delta rgg144$  and  $\Delta rgg939$ , respectively. The quality and concentration of each RNA samples was determined by Nano-drop spectrophotometer.



**Figure 3.42. Agarose gel electrophoresis showing the pneumococcal total RNA.** L- 500 ng of 100 bp DNA ladder (New England Biolabs (NEB), UK). Lane 2: D39 RNA prepared in galactose, Lane 3:  $\Delta rgg144$  RNA prepared in galactose. Lane 4:  $\Delta rgg939$  RNA prepared in galactose. Lane 5: D39 RNA prepared in glucose. Lane 6: D39 RNA prepared in mannose. Lane 7:  $\Delta rgg144$  RNA prepared in mannose. Lane 8:  $\Delta rgg939$  RNA prepared in mannose.

First strand cDNA was synthesised using SuperScript III reverse transcriptase and random primers as described in section 2.17. The transcription levels of the genes were normalised to transcription of DNA gyrase B gene (*gyrB*) as its expression is known to remain constant in different environmental conditions (Goerke et al., 2001). The transcript level of each gene was analysed using the gene specific primers (Table x as described in section x.) The data was data was analysed by the comparative  $C_T$  method ( $2^{-\Delta\Delta C_T}$ ) as described by Livak and Schmittgen (Schmittgen & Livak, 2008). Twofold or greater differences in gene expression were considered as significant (Yesilkaya et al., 2009). The results showed that qRT-PCR analysis was in accordance with microarray data (Table 3.8-3.13). Although there was a difference in absolute fold changes of each gene between qRT-PCR and microarray analysis, the trend of expression was, however, consistent in both methods.

**Table 3.8. Fold difference in expression of selected genes in *Δrgg939* relative to wild type D39 strain grown micro-anaerobically in CDM supplemented with galactose. \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.**

Gene tag	Function	Fold difference *	
		qRT-PCR	Microarray
SPD_0145	CAAX amino terminal protease family protein	-10.61	-6.39
SPD_0146	CAAX amino terminal protease family protein	-6.81	-5.49
SPD_0147	transporter, major facilitator family protein	-5.63	-3.92

**Table 3.9. Fold difference in expression of selected genes in *Δrgg144* relative to wild type D39 strain grown micro-anaerobically in CDM supplemented with galactose. \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.**

Gene tag	Function	Fold difference *	
		qRT-PCR	Microarray
SPD_0942	hypothetical protein	8.07	4.72
SPD_0945	AMP-binding enzyme, putative	2.66	3.4
SPD_0949	bacterial transferase hexapeptide (three repeats), putative	5.83	4.14
SPD_0451	type I restriction-modification system, S subunit, putative	3.95	2.26

**Table 3.10. Fold difference in expression of selected genes in *Δrgg144* relative to wild type D39 strain grown micro-anaerobically in CDM supplemented with mannose. \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.**

Gene tag	Function	Fold difference *	
		qRT-PCR	Microarray
SPD_0146	CAAX amino terminal protease family protein	-6.82	-2.67
SPD_0187	anaerobic ribonucleoside-triphosphate reductase	2.95	4.08
SPD_0189	acetyltransferase, GNAT family protein	2.20	3.95
SPD_0445	phosphoglycerate kinase	-2.46	-3.92

**Table 3.11. Fold difference in expression of selected genes in *Argg939* relative to wild type D39 strain grown micro-anaerobically in CDM supplemented with mannose. \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.**

Gene tag	Function	Fold difference *	
		qRT-PCR	Microarray
SPD_0187	anaerobic ribonucleoside-triphosphate reductase	2.40	4.72
SPD_0189	acetyltransferase, GNAT family protein	2.53	4.85
SPD_0942	hypothetical protein, Carbohydrate transport and metabolism	280.14	4.56
SPD_1932	maltodextrin phosphorylase	-4.55	-3

**Table 3.12. Fold difference in expression of selected genes in wild type D39 grown micro-anaerobically in CDM supplemented with galactose relative to glucose. \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.**

Gene tag	Function	Fold difference *	
		qRT-PCR	Microarray
SPD_0187	anaerobic ribonucleoside-triphosphate reductase	-2.67	-5.87
SPD_0561	PTS system, IIC component, putative	48.17	5.59
SPD_1133	aspartate carbamoyl-transferase	1.21	-14.92
SPD_1050	tagatose 1,6-diphosphate aldolase	129.79	11.93
SPD_1633	galactose-1-phosphate uridylyltransferase	21.67	11.23

**Table 3.13. Fold difference in expression of selected genes in D39 grown micro-anaerobically in CDM supplemented with mannose relative to glucose.** \* Fold difference  $\geq 2$  were considered to be significant, - indicates down regulation of genes,  $\pm$  represents the standard deviation for three individual measurements.

Gene tag	Function	Fold difference*	
		qRT-PCR	Microarray
SPD_0219	ribosomal protein L17	-7.69	-3.58
SPD_0264	PTS system, mannose-specific IAB components	-6.25	-2.07
SPD_1133	aspartate carbamoyl transferase	-2.27	-3.75
SPD_0915	iron-compound ABC transporter, iron compound-binding protein	-3.33	-2.4
SPD_0327	UDP-galactopyranose mutase	-2.28	-4.56

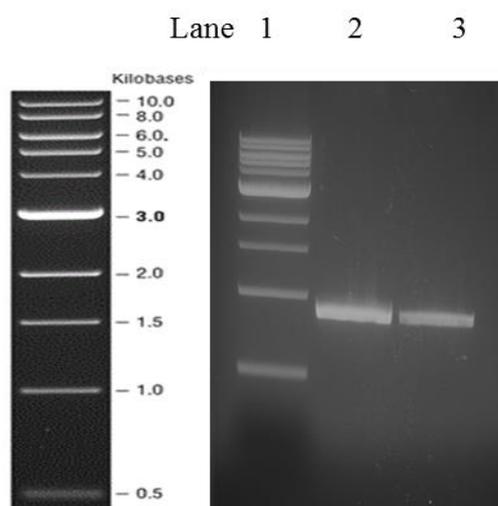
## **Section E. Expression and purification of recombinant proteins**

Recombinant proteins for Rgg144 and Rgg939 were produced to investigate their direct interaction with *shp* genes, SPD\_0315, SPD\_1127, SPD\_1370 and SPD\_2030 by electrical mobility shift assay (EMSA).

### **3.13 Expression and purification of Rgg144**

#### **3.13.1 Amplification of target genes and cloning**

The *rgg144* was amplified with PrimeSTAR HS premix using the primers listed in Table 2.4. The amplified PCR products were analysed by agarose gel electrophoresis (Figure 3.43). The results demonstrated the successful amplification of the target gene. Figure 3.43 show the expected amplicon sizes for *rgg144* (864 bp). The amplicons were then purified from the agarose gel slabs using the Wizard SV Gel and PCR Clean-Up System to remove the remaining enzyme and salt from the PCR products.



**Figure 3.43. Agarose gel electrophoresis showing the amplicons for *rgg144*.** L1- 500 ng of 1 kb DNA ladder (NEB, UK); Lane 2-3 *rgg144* which is approximately 864 bp.

### **3.13.2 Cloning, transformation and DNA sequencing**

The gel-purified amplicons for the transcriptional regulators were cloned into pLEICS-01 (PROTEX, University of Leicester), and a portion of ligation mixture was then transformed into *E. coli* DH5 $\alpha$  competent cells by heat shock. The extracted recombinant plasmids were sequenced using T7 promoter-F and pLIECS-01-Seq-R primers whose recognition sites are localised in immediately up and downstream of the cloning site, respectively. The sequencing results confirmed successful cloning of the genes and the absence of mutations. The recombinant plasmids were then transformed into *E. coli* BL21 (DE3) pLysS competent cells for protein expression.

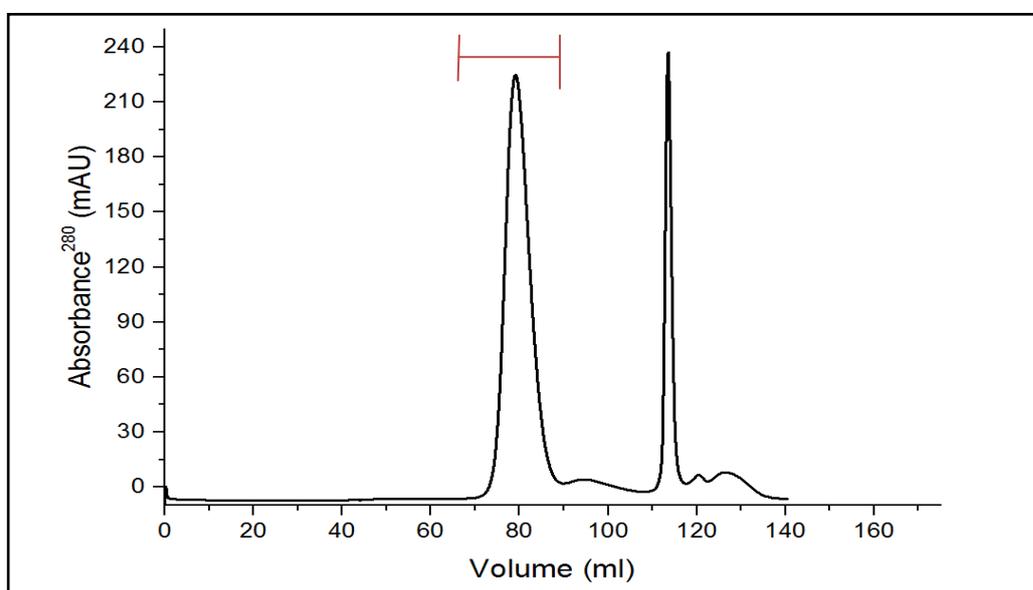
### **3.13.3 Small-scale protein expression**

Small-scale protein expression was done to detect the expressed proteins before moving to large scale expression and protein purification. Different conditions were used to optimize the expression of each recombinant protein. *E. coli* BL21 (DE3) pLysS carrying the desired construct were induced at different growth phases (when OD<sub>600nm</sub> was 1.2-1.6) in power broth medium with different IPTG concentrations (0.1, 0.5 and 1 mM) at different temperatures (18, 24, 30 and 37 °C). The optimal conditions for Rgg144 protein expression was found to be induction with 1 M IPTG when the OD<sub>600</sub> reached between 1.4, and further incubation overnight at 37 °C in a shaking incubator at 220 rpm. SDS-PAGE results showed that recombinant proteins Rgg144 was not soluble, as large amount of protein Rgg144 appears in the pellets rather than the supernatants (Figure 3.44A).

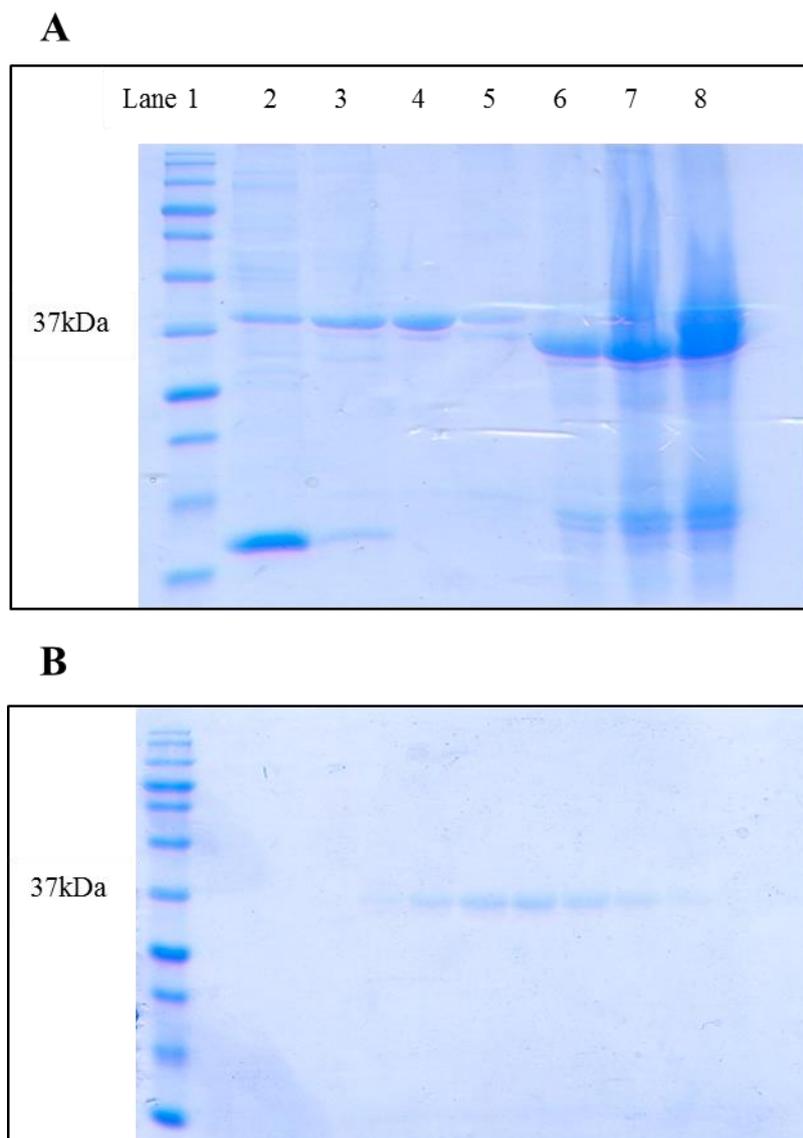
### **3.13.4 Large-scale protein expression and purification**

The optimal conditions determined in the small-scale expression were selected and used in large-scale protein expression and purification. Large-scale samples (before induction, after induction, after sonication and clear cell lysate) were prepared as described previously. Inclusion bodies were purified and solubilized as described in section 2.30. The Rgg144 was refolded and dialyzed. The dialyzed sample was passed through the purification column, washed with washing buffer and the samples were collected. Finally, the protein was eluted with different concentrations of imidazole elution buffer (20, and 500 mM) and 10 of 1 ml

fractions were collected. The histidine-tagged recombinant proteins were bound to the TALON Metal affinity Resin inside the purification column, which has excellent affinity and specificity for his-tagged proteins. Rgg144 was purified by gel filtration using a Superdex 200 16/600 HiLoad column (Figure 3.45). After purification, the eluted protein fractions were analysed by SDS-PAGE along with control samples. Purified fractions of each protein, encoded by *rgg144* are seen in Figure 3.44B. The computed molecular weight of Rgg144 including the histidine tag (1.7 kDa) is as follows: 35.7 kDa. This was defined visually by seeing the highest yields, and later by the Bradford assay.



**Figure 3.44. Purification of Rgg144 on a Superdex 200 16/600 HiLoad Column.** Rgg144 was purified by affinity Chromatography on TALON column and eluted by imidazole, and then loaded onto the Superdex 200 HiLoad 16/600 column and purified by gel filtration. Fractions were collected and those between the red lines were run on polyacrylamide gels to confirm the presence and purity of the sample.



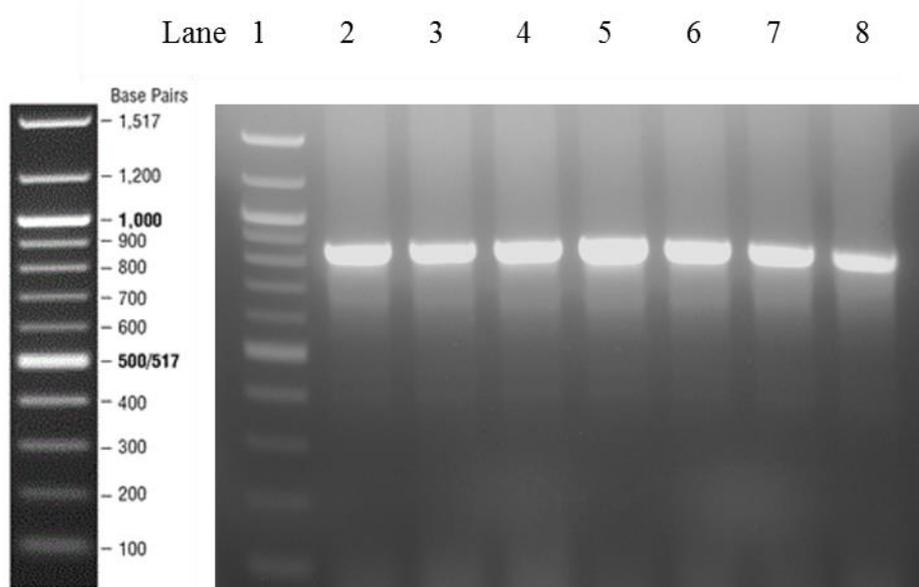
**Figure 3.45. SDS-polyacrylamide gel electrophoresis showing purification of Rgg144.** (A): Lane 1 is protein ladder. A: lane 2, 3, 4 and 5 is the supernatant after treated with lysis buffer 1, 2, 3 and 4 respectively. Lane 6, 7, 8 are 2  $\mu$ l, 4  $\mu$ l and 6  $\mu$ l of the inclusion body. (B): is the different fraction collected by gel filtration.

□

### **3.14 Expression and Purification of recombinant Rgg939**

The pE-SUMO expression vector (LifeSensor) was selected to construct a new recombinant plasmid for *rgg939* expression as it can dramatically enhance the solubility of recombinant proteins in *E. coli*. The pE-SUMO contains N-terminal His-6, SUMO tags and T7 promoters.

The *S. pneumoniae* D39 *rgg939* gene (SPD\_0939) was amplified using gene-specific primers SPS939XbaI /SPD939BsaI. The amplified PCR products were analysed by agarose gel electrophoresis (Figure 3.46). The results demonstrated the successful amplification of the target gene *rgg939*. The PCR product was digested with *BsaI* and *XbaI*, while vector was digested with *BsaI*. Insert and vector was ligated. The newly constructed recombinant plasmid pXZ1, was sequenced and transformed into *E. coli* BL21 (DE3) before each purification.

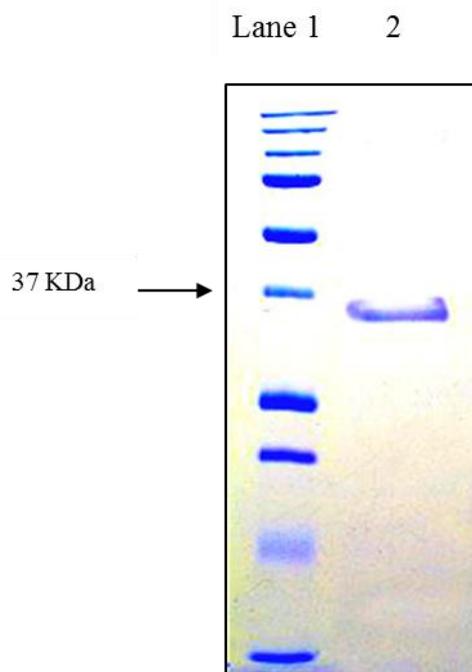


**Figure 3.46. Agarose gel electrophoresis showing the amplicons for *rgg939*.** L1- 500 ng of 100 bp DNA ladder (NEB, UK); Lane 2-8 *rgg939* which is approximately 855 bp.

### **3.14.1 Expression and purification**

The optimal conditions determined in the small-scale expression were selected and used in large-scale protein expression and purification (as described in section 2.30). Bacteria were grown in 2X YT medium until  $OD_{600nm}$  was 0.8-1.2. At this point 0.5 mM IPTG was added to induce the protein expression. Induced bacteria were grown at 30 °C for further 6 hrs. Cell pellet was lysed and the clear cell lysate was passed through the purification column. Finally, the protein was eluted with different concentrations of imidazole elution buffer (20, 40,100,

300, 500 mM) and 1-10 ml fractions were collected. The recombinant proteins were bound to the TALON metal affinity resin inside the purification column, which has excellent affinity and specificity for his-tagged proteins. After purification, the eluted protein fractions were analysed by SDS-PAGE along with control samples. Purified fractions of Rgg939 encoded by the genes *rgg939* is seen in lanes 2 (Figure 3.47). The computed molecular weight of each protein including the histidine tag (1.7 kDa) is as follows: 34.9 kDa.



**Figure 3.47. SDS-polyacrylamide gel electrophoresis showing purification of Rgg939.** Lane 1 is protein ladder. Lane 2 is Rgg939 collected by gel filtration.

## **Section F. Electrophoretic mobility shift assay (EMSA)**

The electrophoretic mobility shift assay (EMSA) also called gel shift assay is the most common affinity electrophoresis technique used quantitative analysis of interacting systems, such as protein and nucleic acid interactions. The most common methods was described by (Garner and Revzin, 1986). This method depends on electrophoretic separation of a protein and DNA, DNA fragments and proteins are mixed in a suitable buffer and binding is allowed to occur, since the probe binds to the protein, The mixture is then separated by nondenaturing gel electrophoresis, the complex of DNA and protein will stay up on the gel as its moves more slowly than the free probe (Molloy, 2000). For visualization of the DNA-protein interaction, the DNA probe is commonly labelled with a radioactive, fluorescent or biotin label (Rodgers et al., 2000). Although ethidium bromide staining method is used in some experiments, this method is less sensitive to detect small amount of interactions. The isotopic label with  $^{32}\text{P}$ -phosphate may suffer from the short half-life of the label and safety problem of the operator (Jing et al., 2003). When using a biotin label, it needs additional step with blotting the gel to a positively charged nylon membrane, which are time consuming. The EMSA assays using FAM labeled DNA probe is sensitive, fast, and does not need addition of detective step for visualization.

To date, there are several different approach have been used for detection of protein-nucleic acid interactions. The most popular approaches are nitrocellulose filter binding and Chromatin Immunoprecipitation (ChIP) (Carey et al., 2009). However, filter binding can not be applied in assays which need competitor proteins. In addition, CHIP is mainly used to detect the location of DNA binding sites on the genome for a target protein in living cells or tissues but it is prone to false positive results (Reid et al., 2000, Ren et al., 2000). On the other hand, EMSA assay is sensitive, relatively easy to perform and fast to get results. Therefore, EMSA assay was selected to detect the interaction of Rgg144 and selected promoters of *shp144*, SPD\_0315, SPD\_1041, SPD\_1127, SPD\_1370, SPD\_1517, and SPD\_2030. These targets were selected for EMSA because the whole operon of SPD\_0315 (The capsule locus), SPD\_1041(the *nrd* operon, which involved in nucleic acid transport and metabolism), SPD\_1127 (*lipD* operon involved in lipid transport and metabolism), SPD\_1370 (the *rps* operon, the ribosomal protein), SPD\_1517 (the operon involved in defense mechanism), SPD\_2030 (The operon involved in ribosomal structure and

biogenesis.) were all differentially regulated in *rgg144* and *rgg939* mutants in microarray analysis.

### **3.15 In silico promoter analysis**

The upstream regions of *shp144*, SPD\_0315, SPD\_1041, SPD\_1127, SPD\_1370, SPD\_1517 and SPD\_2030 were screened using the bacterial promoter prediction tool (BPROM) and Motif-based sequence analysis tool (MEME) (<http://meme.nbcr.net/meme/tools/meme>) to detect the putative promoter regions and the putative binding motifs for the target genes, respectively. An illustration of each predicted promoter region and binding motifs are shown in Figure 3.48. The promoter prediction tool detected the core promoter elements -10 and -35 in the upstream of all screened genes: *shp144*, SPD\_0315, SPD\_1041, SPD\_1127, SPD\_1370, SPD\_1517, and SPD\_2030. Moreover, the motif-based sequence analysis tool MEME identified the putative binding site within the promoter regions of the screened genes (Figure 3.49). Therefore, these genetic regions were selected for EMSA analysis.

**A.**

AGTAGGAGTATCAGTTGGATTTTTCTTTTCTTCATCTATTATATTCCTCCTTGTTAGTAATAAC  
TTTATTATACCAGGGAATAATCAAATCTATCAAATCGCAAATAAGAAATTTCTATAAGAA  
AAAATATCAAATATGCGATTTTATAATAAGCCAATTTTCGTGTTATACTGTACTTG

**B.**

TTGGTTCGCGGGAAGTCTACTAAGATACTTAAAGATGCAGATAGTGAAAAAAGGTGTAGACA  
TTACCGTAAAAAAGTGATATAATCGTAAGATGTTCAATGTATAGGTGTTAATCATGAGTAGA  
CGTTTTAAAAAATCACGTTACAGAAAGTGAAGCGAAG

**C.**

CAAAACACAGGATATAGTTCTTTAAAAACAATTTATAGGATAATCATTGCTATTTACAATACA  
AACACTATAIATTTGTGTTTATATATAAAAAACAACAAAAGAACCACCTATATTTAGTTCAAATCA  
TTGACAAAATTTTATTTTCTGATATACTAAGATAGTATTAATTTTGAAGAGGAGTTACACAATG  
GTAACCGTTTATTCTAAAAACAATTGTGTCCAATGTAATGACCAAGCGTTTCTTG

**D.**

CTGCATAATTCTCCTATTC TAGAAGGGGAGGACCAGTATTTCTTTATGATAGGACTAGATTGTG  
GTATAATAGAGAGAATAAGTTTTTTTAGTAAGACAAAGGAGAAAATAGATGATTTATGCAGGA  
ATTCTTGCCGGTGGAACTGGCACACGCATG

**E.**

GTATCAAACCATAAGAACAGGAAAACGCCCATGTGGCGTTTTTTTCGAATTCTAGTGTTTTACT  
 TCGTAAAAAATTTTTTCAGTATAATAGTTTATTGTGAGCGAACCTCACTTACCCCTTGCAA  
 AGTCTTGGGGTCATTAGACCAAAGGAGGAACATATCAATGGCTAAATACGAAATCTTTATAT  
 CATTGTCCAAACATTGAAG

F.

GGGATAAAGAAATTAGAGTCTATACATAATAAGTAAAGGAGGAATTTGTATGATATAACCATCG  
 TTTAGAATAAGTCTATATTTAAAATAGAA0GTTATAAGGAATCTTCTTGAATAAAAAATGAAAG  
 GAAACAGAAAAAATGAAAATTAAGAAATTATTGAAAATGGTTATTCCTGTTTAATGATAAGTG  
 CTGTTG

G.

ATTGGTTTGTAGAGGTTAAAGAATCTTAATTTCAATATGTGTAAAGGTAGGTTACTGAATT  
 GTAAACTACCTTTTCTTATAATTGATAGAATTACTGATTACACTTTTAAAAAGTCGCTTTTTG  
 AGGGGTATTATGGTATAATGGGTGCCAAGAGGTTTTGAATGAAAAAATTTTATGTAAGTCCA  
 ATTTTTCTATTCTAGTAGGATTGATTGCGTTTG

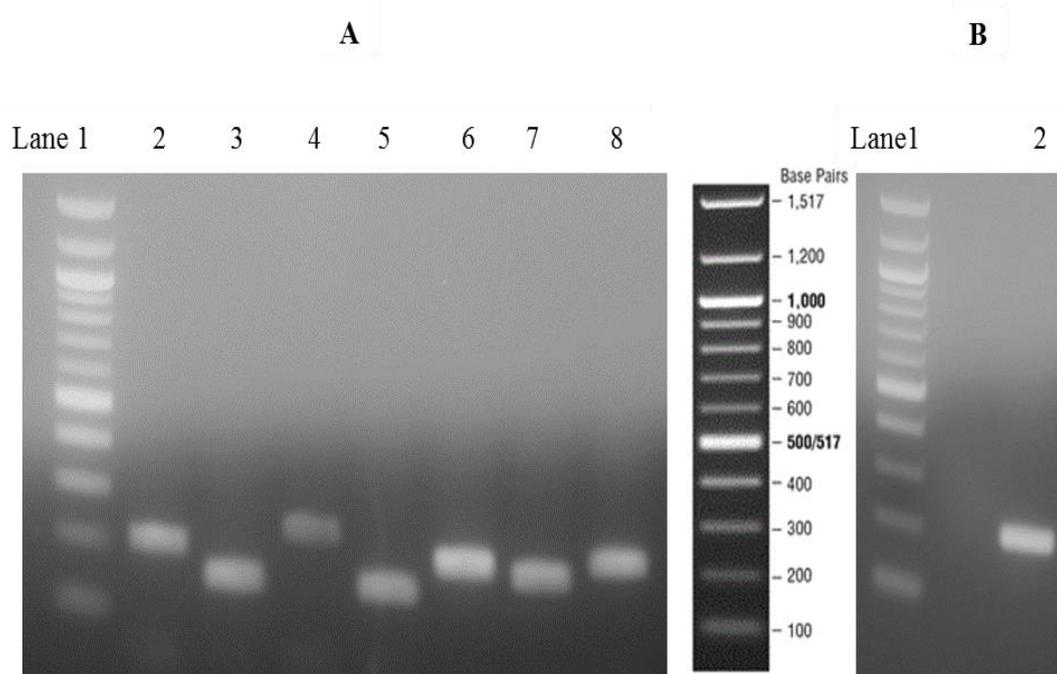
**Figure 3.48. Illustration showing the analysis of the predicted promoter region of *shp144* (A), SPD\_0315 (B), SPD\_1041(C), SPD\_1127 (D), SPD\_1370 (E), SPD\_1517 (F), SPD\_2030 (G). Sense DNA strand for the promoter region of each genes. The start codon of each gene is in green, the putative -10 regions are in purple, the -35 regions are in red. The primer sequences were highlighted in yellow.**



**Figure 3.49. The putative Rgg binding motif, generated *de novo* from the upstream regions of SPD\_0315, SPD\_1127, SPD\_1370, SPD\_2030 and *shp144*.**

### **3.16 Amplification of putative promoter regions**

The putative promoter regions for SPD\_0315, SPD\_1041, SPD\_1127, SPD\_1370, SPD\_1517, SPD\_2030 and *pflB* (which was used as the negative control, because *pflB* is not differentially expressed in *rgg144* mutant) with their putative binding sites were amplified using the FAM labeled primers listed in Table 2.13. The amplified PCR products were purified using a Wizard SV Gel and PCR Clean-Up System from Promega (UK). The amplified promoter regions were analysed on agarose gel electrophoresis (Figure 3.50).

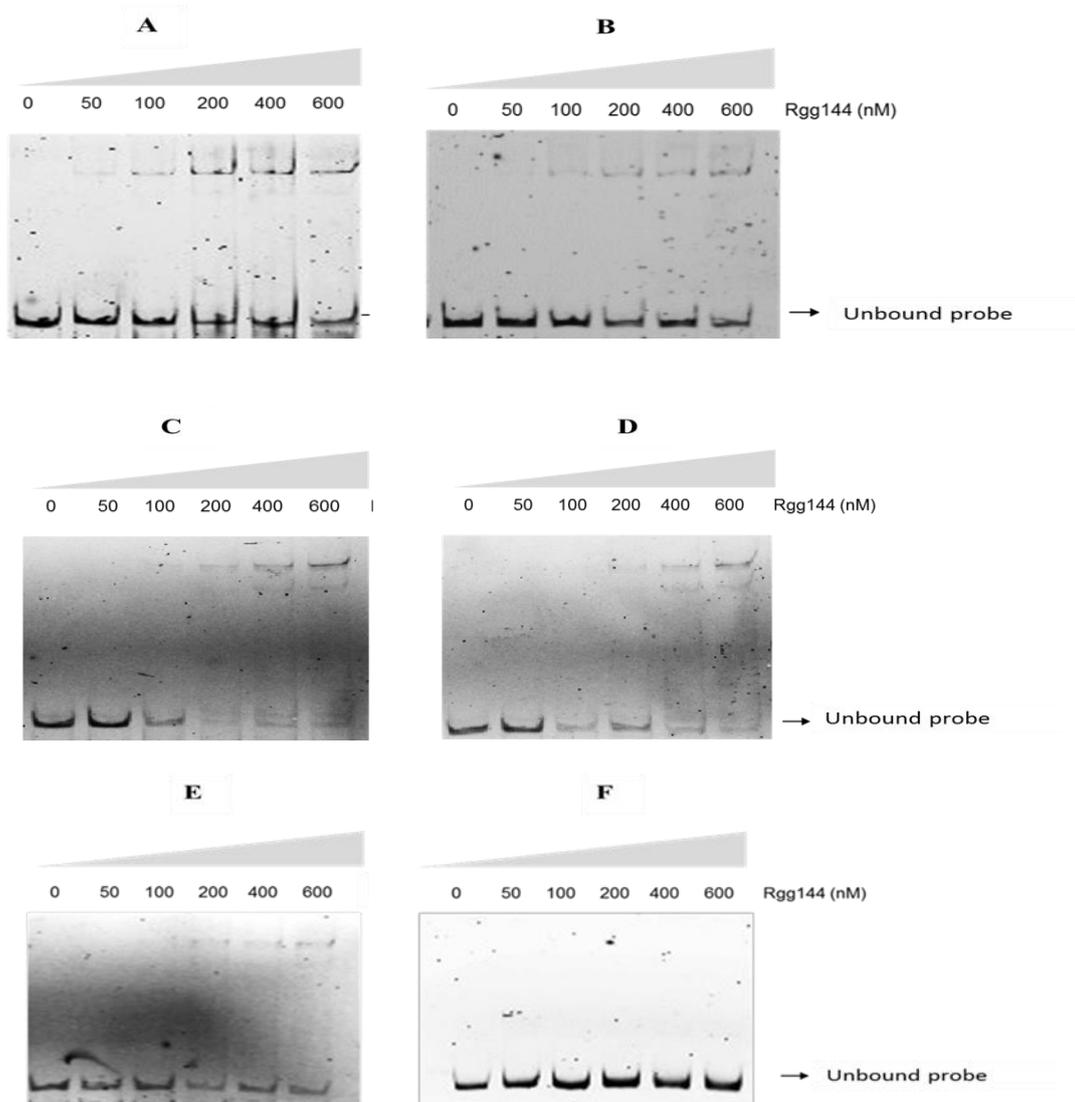


**Figure 3.50. Agarose gel electrophoresis showing the amplicons for each probes.** L1- 500 ng of 100 bp DNA ladder (NEB, UK); (A) Lane 2, P<sub>*pflB*</sub> (210 bp). Lane 3; P<sub>SPD\_0315</sub> (162 bp), Lane 4: P<sub>SPD\_1041</sub> (249 bp), Lane 5: P<sub>SPD\_1127</sub> (158 bp), Lane 6: P<sub>SPD\_1370</sub> (212 bp). Lane 7: P<sub>SPD\_1517</sub> (198bp), Lane 8: P<sub>SPD\_2030</sub> (223bp). (B) Lane 2: P<sub>shp144</sub> (184 bp). ‘P’ refers to promoter.

## **3.17 Electrophoretic mobility shift assay EMSA**

### **3.17.1 Recombinant Rgg144 interacts with $P_{shp144}$ , $P_{SPD\ 0315}$ , $P_{SPD\ 1127}$ , $P_{SPD\ 1370}$ and $P_{SPD\ 2030}$ .**

EMSA was used to analyse the interaction of the Rgg144 with  $P_{shp144}$ ,  $P_{SPD\ 0315}$ ,  $P_{SPD\ 1127}$ ,  $P_{SPD\ 1370}$  and  $P_{SPD\ 2030}$  according to the protocol described in section 2.31 ('P' means promoter). Increasing amounts (0.05 -0.6  $\mu$ M) of the purified His-tagged recombinant Rgg144 was mixed with constant amount of labeled  $P_{shp144}$ ,  $P_{SPD\ 0315}$ ,  $P_{SPD\ 1127}$ ,  $P_{SPD\ 1370}$  and  $P_{SPD\ 2030}$  DNA probes individually, and DNA-protein mixtures were analysed in non-denaturing condition by PAGE. The results showed that Rgg144 was able to bind to  $P_{shp144}$ ,  $P_{SPD\ 0315}$ ,  $P_{SPD\ 1127}$ ,  $P_{SPD\ 1370}$  and  $P_{SPD\ 2030}$ , as the mobility of DNA probe in the presence of protein was different than that of DNA alone (Figures 3.51). As shown in Figures 3.51, lanes 2-6, Rgg144 bound to the each probe in a dose-dependent manner and the band shift increased with increasing protein concentration. To demonstrate the specificity of Rgg144 binding to  $P_{shp144}$ ,  $P_{SPD\ 0315}$ ,  $P_{SPD\ 1127}$ ,  $P_{SPD\ 1370}$  and  $P_{SPD\ 2030}$ ,  $P_{pflB}$  was included in the assay as a negative control. As can be seen in Figure 3.51, there was no change in mobility of  $P_{pflB}$  in the presence of Rgg144, indicating the specific binding of Rgg144 to the selected DNA probes.



**Figure 3.51. Electrophoretic mobility shift assays indicating the Rgg144 binding to different DNA probes.** DNA binding of the His-tagged Rgg144 protein to promoter regions of the target genes of the *shp144* (A), SPD\_0315(B), SPD\_1127(C), SPD\_1370 (D), SPD\_2030 (E) and  $P_{pflB}$  (F), the negative control. Lane 2- 6: contains 10 nM probes and 0.05 -0.6  $\mu$ M Rgg144.

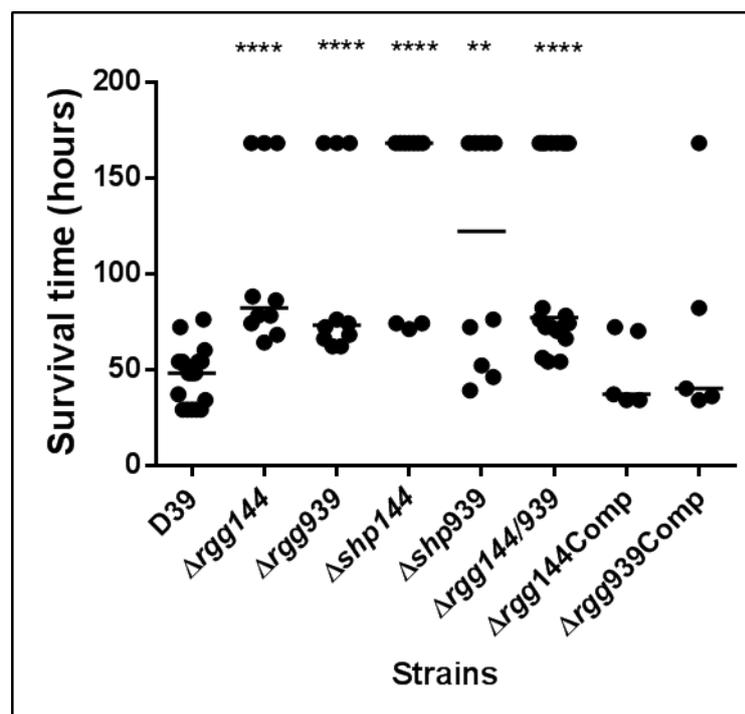
## **Section G. In vivo virulence studies**

It was reported that Rgg family regulators have crucial roles in virulence, for instance, *rgg* mutant in *S. suis* displayed attenuated virulence in an experimental piglet infection model (Zheng et al., 2011). In addition, *S. pyogenes* exotoxin B (SpeB), which affects host-pathogen interactions by destroying the extracellular matrix and degradation of complement factors in host (Kappeler et al., 2009), is regulated by Rgg. Thus, it was hypothesised that Rgg is involved in pneumococcal survival and virulence. To support this hypothesis we tested  $\Delta rgg144$ ,  $\Delta shp144$ ,  $\Delta rgg939$  and  $\Delta shp939$  in a mouse model of pneumococcal pneumonia and colonisation that develops after intranasal infection. Animal models are crucial tools to investigate the infectious disease (Chiavolini et al., 2008). For pneumococcal diseases, the mouse represents the most commonly used animal model, as mouse cost is low, easy to handle, and researchers can use sufficient amount of mouse to obtain statistically significant results for analysis (Mohawk and O'Brien, 2011). The inbred strains show more-uniform responses to experimental treatments including vaccines, and antibiotics, as they have tightly controlled immune system. Therefore, the inbred strains are widely used to test the efficiency of vaccination, and drugs. Outbred mouse strains can display maximum phenotypic diversity, although the genetic diversity in mice is lower than that of humans, they could be used to mimic the natural variation in response to infection, thus the outbred mouse are popular in analysing pathogenicity mechanisms (Chiavolini et al., 2008). For pneumococcal pneumonia model, the infection through intranasal aspiration is the most popular approach, as it is fast and easy to operate, do not need invasive surgical procedures. Moreover, it mimics the natural infection process in humans (Chiavolini et al., 2008)

### **3.18 Contribution of Rggs and Shps in pneumococcal virulence**

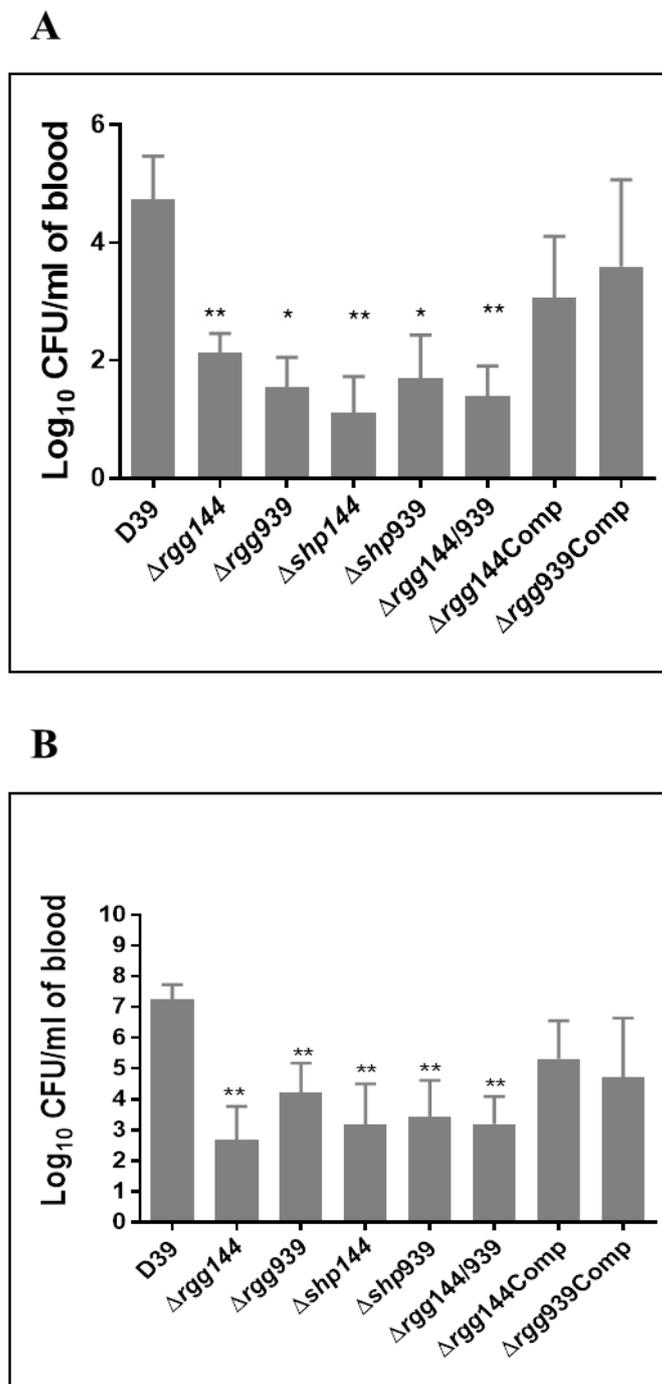
The contribution of Rggs and Shps to pneumococcal virulence was determined in a mouse model of pneumococcal infection as previously described in section 2.34.2. The mutant strains  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta shp144$ ,  $\Delta shp939$ , the double mutant  $\Delta rgg144/939$  (double mutants was used to check if *rggs* have synergistic effect on pneumococcal virulence), the genetically complemented mutants,  $\Delta rgg144Comp$ ,  $\Delta rgg939Comp$ , and the wild type strain D39 were included in the assay.

The survival times of mice infected with the strains are shown in Figure 3.52. The results showed that the median survival time of mice infected intranasally with  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta shp144$ ,  $\Delta shp939$ , and  $\Delta rgg144+939$  ( $104 \pm 14.2$ ,  $98 \text{ h} \pm 15.3$ ,  $139 \text{ h} \pm 14.5$ ,  $112 \text{ h} \pm 18.8$  and  $109 \text{ h} \pm 11.2$ , respectively,  $n=10$ ) were significantly longer than the wild type infected group ( $46 \text{ h} \pm 3.5$ ,  $n=10$ ) ( $p<0.01$ ), hence there is no synergistic effect among Rggs *in vivo*. The introduction of intact copies of *rgg144* and *rgg939* into the mutants,  $\Delta rgg144$  and  $\Delta 939$ , respectively, reconstituted the virulence of these strains as the median survival times of mice infected with  $\Delta rgg144\text{Comp}$  ( $49 \text{ h} \pm 8.8$ ,  $n=5$ ) and  $\Delta rgg939\text{Comp}$  ( $72 \text{ h} \pm 25.5$ ,  $n=5$ ) were not significantly different from the wild type infected cohort ( $p>0.05$ ). The results obtained with the complemented strains rule out the possibility of polar effect of mutations, and relate to the significant attenuation of virulence due to the loss of Rggs and Shps.



**Figure 3.52. Survival time of mice infected intranasally with approximately  $2 \times 10^6$  pneumococci.** Symbols show the times mice became severely lethargic. The horizontal bars mark the median times to the severely lethargic state. Each point is the mean of data from ten mice. Significant differences in survival times are seen comparing with the wild type strain using Mann Whitney test. (\*\* $p<0.01$ , \*\*\*\*  $p<0.001$ ).

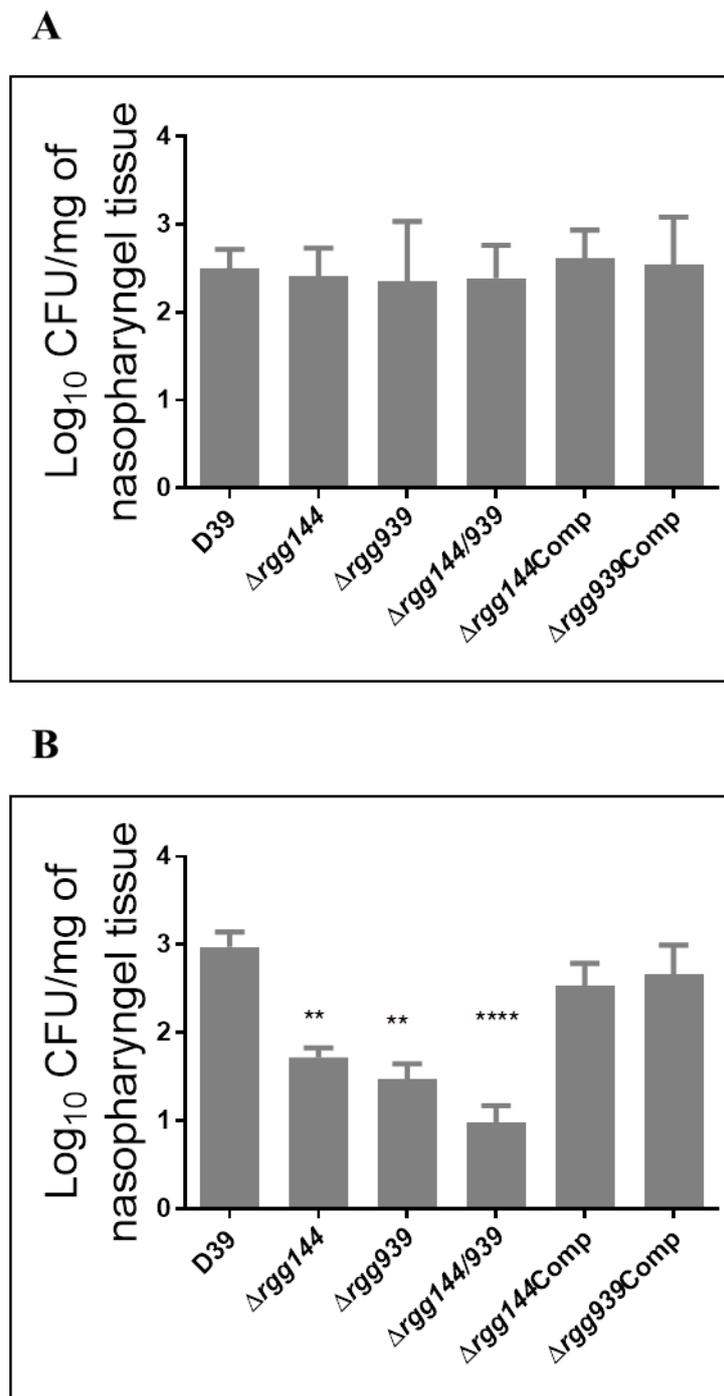
The progression of bacteraemia in animals infected with the pneumococcal strains was also determined. The  $\log_{10}$  CFU/ml of pneumococci retrieved from the blood of mice infected with pneumococcal strains is shown in Figure 3.53. The bacterial load in the  $\Delta rgg144$  infected group was significantly lower at 24 and 48 h post-infection ( $\log_{10} 2.13 \pm 0.33$  CFU/ml and  $\log_{10} 2.68 \pm 1.01$  CFU/ml respectively, n=10) compared to the wild type strain at 24 h ( $\log_{10} 4.72 \pm 0.74$  CFU/ml, n=10) and 48 h post-infection ( $\log_{10} 7.25 \pm 0.48$  CFU/ml, n=10) ( $p < 0.0001$ ). Moreover, numbers of  $\Delta rgg939$  in the blood of mice was significantly lower at 24- and 48 h post-infection ( $\log_{10} 1.54 \pm 0.51$  CFU/ml and  $\log_{10} 4.20 \pm 0.97$  CFU/ml respectively, n=10) compared to the wild type strain at 24 h ( $\log_{10} 4.98 \pm 0.32$  CFU/ml, n=10) and 48 h post-infection ( $\log_{10} 8.43 \pm 0.18$  CFU/ml, n=10) ( $p < 0.01$ ). Moreover the progression of bacteraemia in cohort infected with complemented strains  $\Delta rgg144$ Comp at 24 h ( $\log_{10} 3.07 \pm 1.03$  CFU/ml, n=5) and 48 h post-infection ( $\log_{10} 5.32 \pm 1.24$  CFU/ml, n=5) was similar to that of wild type infected cohort ( $p > 0.05$ ). Moreover, both complemented strains  $\Delta rgg144$ Comp and  $\Delta rgg939$ Comp had significantly higher bacterial load in the blood of infected mice at 24 and 48 h of post infection compared to the respective mutant strains  $\Delta rgg144$  and  $\Delta rgg939$  ( $p < 0.01$ ).



**Figure 3.53. Progression of bacteraemia in mice infected intranasally with D39,  $\Delta$ rgg144,  $\Delta$ rgg939,  $\Delta$ shp144,  $\Delta$ shp939, and  $\Delta$ rgg144/939 and their genetically complemented strains at 24 h (A) and 48 h (B) post-infection.** Each point is the mean of data from ten mice. Error bars show the SEM. Significant differences in bacterial counts are seen comparing with the wild type strain using oneway ANOVA and Tukey's multiple comparisons test. (\*  $p < 0.05$  and \*\*  $p < 0.01$ ).

### **3.19 Contribution of Rgg144 and Rgg939 in pneumococcal colonisation**

The role of Rgg144 and Rgg939 in nasopharyngeal colonisation was investigated as previously described in section 2.34.3. Pneumococci were administered intranasally and the numbers of bacteria were determined by serial dilution of homogenised nasopharyngeal tissue of infected mice at the time of infection and 7 days after infection. The number of colony forming unit was assessed for per mg of homogenised tissue. One hour after infection, the bacterial load in the nasopharyngeal tissue for all strains ( $\log_{10} 2.41 \pm 0.16$  CFU/mg,  $\log_{10} 2.24 \pm 0.34$  CFU/mg,  $\log_{10} 2.39 \pm 0.19$  CFU/mg,  $\log_{10} 2.61 \pm 0.16$  CFU/mg and  $\log_{10} 2.54 \pm 0.27$  CFU/mg, for  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta rgg144/939$ ,  $\Delta rgg144Comp$  and  $\Delta rgg939Comp$ , respectively, n=5) was similar to that of wild type ( $\log_{10} 2.49 \pm 0.11$  CFU/mg, n=5) ( $p > 0.05$ ) (Figure 3.54A). On the other hand, at 7 days post-infection the colony counts for  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta rgg144/939$  ( $\log_{10} 1.72 \pm 0.11$  CFU/mg,  $\log_{10} 1.47 \pm 0.18$  CFU/mg and  $\log_{10} 0.98 \pm 0.19$  CFU/mg respectively, n=5) were significantly lower than the counts of wild type strain ( $\log_{10} 2.98 \pm 0.17$  CFU/mg, n=5) ( $p < 0.01$ ,  $p < 0.01$  and  $p < 0.0001$  for  $\Delta rgg144$ ,  $\Delta rgg939$ ,  $\Delta rgg144/939$ , respectively) (Figure 3.54B). No significant differences were seen in the bacterial load of the complemented strains,  $\Delta rgg144Comp$  and  $\Delta rgg939$  ( $\log_{10} 2.53 \pm 0.26$  CFU/mg and  $\log_{10} 2.67 \pm 0.33$  CFU/mg respectively, n=5) compared to the wild type ( $p > 0.05$ ). These results strongly suggest that  $\Delta rgg144$  and  $\Delta rgg939$  contribute to pneumococcal colonisation.



**Figure 3.54. Pneumococcal strains defective in *rgg144*, *rgg939* and *rgg144/939* were less able to colonise nasopharynx.** Mice were infected approximately with  $5 \times 10^5$  CFU pneumococci. At day 0 (A) and day 7 (B), five mice were culled, and CFU/mg of bacteria were determined by serial dilutions of nasopharyngeal homogenates. Each column represents the mean of data from five mice. Error bars show the standard error of the mean. Significant differences in bacterial counts are seen comparing with the wild type strain using one-way ANOVA and Tukey's multiple comparisons test. \*\*  $p < 0.01$  and \*\*\*\*  $p < 0.0001$ .

## **Chapter IV: Discussion**

*Streptococcus pneumoniae* can cause a wide range of disease including sinusitis, otitis media, meningitis, pneumonia and sepsis, which usually associated with high mortality and morbidity (O'Brien et al., 2009). Worldwide, pneumonia is still the leading cause of death among children less than 5 years old, causing 1.6 million deaths annually (Kim et al., 2016). Persons at higher risk for invasive pneumococcal disease (IPD) include infants, adults  $\geq$  65 years old, people suffering chronic diseases, and those with compromised immune system (Walker et al., 2013, Robinson et al., 2001). Antibiotics have been a first choice of therapeutic approach for IPD. However, after overuse of antibiotic, multi-resistant pneumococcal clones emerged and disseminated worldwide. Penicillin, erythromycin, and trimethoprim-sulfamethoxazole resistant pneumococci are wide spread all over the world, tetracycline, chloramphenicol, fluoroquinolone resistances were also identified with various levels among different countries (Klugman, 1990, Jones et al., 2013, Kim et al., 2016). The prevalence of antibiotic resistance creates a difficulty for treating pneumococcal infections. Therefore, there is an urgent need to develop new antiinfectives. One of the ways to design efficient antiinfectives relies on the identification of the pathways that are critical for *in vivo* survival of the microbe so that they can be exploited to develop efficient antimicrobials.

I hypothesised that the environmental adaption is critical for pneumococcal survival *in vivo*. The pneumococcus is exposed to different environmental parameters in different host tissues during colonisation and invasive disease, including different composition of carbon source, oxidative stress and osmotic pressure. The microbe has a high level of adaptive capacity. These adaptive mechanisms are vital for the *in vivo* survival of the pneumococcus and therefore represent a viable route for the treatment of bacterial infections. However, our knowledge about how the pneumococcus adjusts to the dynamic host niches is limited.

The pneumococcus has several regulatory proteins, and they modulate the microbe's phenotype under different environmental conditions. Examples include 13 TCS and multiple stand-alone regulators (Paterson et al., 2006). In addition, the pneumococcal genomes contain several regulatory systems resembling QS systems (Greenberg, 2003). Microbes can talk to each other by using QS circuits, and QS systems enable microbes to modulate their phenotype in a cell density dependent manner. At the start of this project, the knowledge on pneumococcal QS systems was limited mainly to ComABCDE, BlpABC SRH and the LuxS/Autoinducer 2 systems (Galante et al., 2015). Therefore, my objective was to determine whether these *in silico* identified putative circuits would operate as QS systems, and what

roles they would play in pneumococcal biology. Research on QS pathways is beneficial for identification of new therapeutic targets for pneumococcal diseases. Unlike traditional antibiotics, QS system inhibitors do not interfere with the essential metabolic processes, hence there will be less selective pressure for emergence of resistance (Rasko and Sperandio, 2010).

The landmark study of Feluchot et al (2011) reported a comprehensive list of *rgg* genes associated with short hydrophobic peptides (Shp) in streptococci. A number of Rgg proteins have been investigated due to their central roles in the control of important physiological processes in streptococci other than pneumococcus (Sulavik et al., 1992, Fernandez et al., 2006, Pulliainen et al., 2008). In addition to stand-alone *rgg* that do not have an associated pheromone coding gene, genome analysis has revealed that some *rgg* homologs are located in the genome adjacent to small open reading frames that encode putative short hydrophobic peptides, SHP (Ibrahim et al., 2007). Possibly due to the small size (20–23 aa), the genes encoding SHPs are mostly unannotated in sequenced genomes, but they have been researched along with *rgg*, and the available evidence has revealed SHP-induced modulation of Rgg quorum sensing pathways (Chang et al., 2011, Fleuchot et al., 2013, Ibrahim et al., 2007). Accumulating evidence suggests that Rgg regulators, in combination with cognate peptides, provide a common mechanism for genetic regulation among the streptococci. Functional Rgg-peptide examples now available from *S. thermophilus* (Rgg1358-Shp1358c) (Fleuchot et al., 2011) and ComR-ComS (Fontaine et al., 2010), *S. mutans* (ComR-ComS) (Mashburn-Warren et al., 2010), and *S. pyogenes* Rgg2/3 (Chang et al., 2015), as also presented in my study in pneumococcus, substantiate the suggestion that Rgg proteins are cytoplasmic regulatory proteins that directly bind to the cognate peptides.

The knowledge on the functional roles of pneumococcal Rggs has been largely unknown, except a recent work detailing the involvement of a peptide in the Rgg144 regulon in biofilm formation and pneumococcal pathogenesis (Cuevas et al., 2017). Despite the structural similarities among different Rggs in streptococci, the study of Rggs in *S. pneumoniae* is worthy not because the genetic complexity of Rggs in *S. pneumoniae* is higher than the other streptococci, for example, different pneumococcal strains contain between 2 to 7 Rgg homologs (personal communication with Dr Luisa Hiller) whereas *S. pyogenes* only has 4 Rgg homologs. In addition, the signalling in *S. pneumoniae* is also different than in other Streptococci. For example in certain pneumococcal strains such as Taiwan 19F-14 there is no

SHP associated-Rgg, and in other pneumococcal strains the signalling peptides belong to either Group I, III or both, whereas in *S. pyogenes*, the signalling peptide can be either SHP Group I or XIP. Hence, due these attributes of pneumococcal Rggs, it was thought at the outset of this project that Rggs would play unique roles in pneumococcal biology.

In this study I characterised two new peptide-dependent signaling systems, Rgg144/Shp144 and Rgg939/Shp939, in *S. pneumoniae*. QS circuits in Gram positive bacteria known to have the following features: 1) they use processed peptides as pheromones, 2) the activity of QS circuits is growth phase dependent, and 3) the cognate Rgg regulators act as cytoplasmic receptors for intracellular signalling peptides. The newly characterised Rgg/Shp144 and Rgg/shp939 circuits fulfilled all the criteria of quorum sensing system, my data support the hypothesised model in which SHP144 and SHP 939 is derived from a precursor peptide, produced and exported to extracellular environment, and processed to functionally active state. Once the peptide had reached a sufficient extracellular concentration, it would be re-imported into cell via the oligopeptide permease and bind to the Rgg regulators. The Rgg regulators will have conformational changes, and bind to the promoter regions of Rgg regulons, switch on or switch off the transcription. This mechanism resembles quorum-sensing pathways of other Gram-positive bacteria.

QS systems are active when high cell density is present. The Rgg144/Shp144 circuit could respond to the spent late exponential culture supernatants from wild type strain rather than the early exponential phase supernatants. At late exponential phase of growth, the system would be active due the presence of higher concentration of pheromone in the extracellular environment, stimulating the expression of *shp*. This result was in line with the findings in other streptococcal Rgg/Shp systems (Chang et al., 2011, Fleuchot et al., 2011). However, *P<sub>shp939</sub>::lacZ* driven  $\beta$ -galactosidase activity could not be induced irrespective of the growth phase of culture supernate. This suggests that in the culture supernates used the concentration of inducing peptide was low probably due not to use optimal condition that would induce the synthesis of active pheromone.

These QS circuits were responsive to the synthetic peptides representing the C-terminal end of each Shp. SHP144-C12 and SHP939-C8 were found to be the most active variants for activation of the respective circuits. Difference in size of the activating peptides for each QS system suggests that the processing cascades for these circuits could differ. In addition, it also

implies that the binding sites of these peptides may display some structural differences. For different Rgg/Shp circuits, the size of the most active peptide has been shown to display a wide variation. For example, SHP1358-C9 is the most active peptide for Rgg1358 in *S. thermophilus* (Fleuchot et al., 2011), C-terminal eight amino acids is the minimum length of SHP required for the expression of *shp3* in *S. pyogenes* (Chang et al., 2011). This may be caused by the diversity in the structure of the cognate Rgg. The concave surface of C-terminal repeat domain of each Rgg could be different, which was identified as the binding site for their cognate peptide pheromones by X-ray crystal structure analysis of Rgg2 in *S. pyogenes* (Parashar et al., 2015). Aggarwal and his colleagues demonstrated that multiple variants of SHP peptide can be produced by GAS streptococcus. Their results showed that different SHP peptide variants has different affinities for binding to Rgg proteins, and the binding affinities are related to transcriptional activity (Aggarwal et al., 2014). They also reported that unprocessed pre-peptide do not have the ability to induce the quorum sensing circuit.

It is intriguing to point out that nearly identical orthologs of *rgg939/shp939* are present in *S. pyogenes* and *S. thermophilus* with 90% and 87% identity, respectively. The orthologs of *rgg144* are found in *S. mitis*, *S. pseudopneumoniae* and *S. intermedius*, with 93%, 93%, and 63% identity, respectively. It is speculated that these signaling pathways provide useful quorum-sensing functions for each species, and that, in the case of *S. pneumoniae* which share the nasopharynx with *S. pyogenes*, interspecies signalling may occur (Cook et al., 2013). However, it should be noted that the genetic location surrounding *rgg* orthologs in these related species have not been conserved. For example, although pneumococcal *rgg939* and *spy49\_0449* are nearly identical, the genes surrounding them are different. The *rgg939* is surrounded by genes coding for hypothetical proteins and for a UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase (SPD\_0940), whereas the upstream of *spy49\_0449* has a putative operon coding for *vicR/vicK* two-component response regulator, while the downstream region contains *aroE*, codes shikimate 5-dehydrogenase, and genes for protein of unknown function. Hence, the different genomic backgrounds may lead to the different functional roles of Rggs.

Very little is known regarding the processing of SHP peptides. It will be interesting to determine how SHP peptides are exported from the cell. Some of the specific questions include if export is coupled to peptide maturation or if these are temporally and/or spatially

distinct steps in the circuit. Chang et al demonstrated that a metalloprotease (Eep, coded by spy49\_1620c) has a role in processing of SHP3, but not absolutely required for induction of the  $P_{shp3}$  promoter when SHP3 is abundant. Pneumococcus has an *eep* ortholog (SPD\_0245) which is 62.9% identical to spy49\_1620c at nucleotide sequence level, which may have similar role in SHP peptide processing and should be tested in future. Moreover, recently, an ABC transporter called PptAB was identified, which contribute to Rgg-SHP signalling pathway in *S. agalactiae* (Perez-Pascual et al., 2015) and *S. pyogenes* (Chang and Federle, 2016). The PptAB has been demonstrated as the primary transporter for SHP pheromones in GAS, as *pptAB* mutant lost the ability for robust  $P_{shp3-lux}$  induction, whereas the addition of synthetic SHP-C8 pheromone complemented the functional deficiency, which demonstrated a role for pheromone production rather than signal detection (Chang and Federle, 2016). Interestingly, the ortholog of gene encoding PptAB has also been found in *S. pneumoniae* D39 by *in silico* analysis. SPD\_0464 encoding an ABC transporter shares 76.8% identity with PptAB. Hence I hypothesis that the ABC transporter encoded by SPD\_0464 could be the transporter for SHP pheromones in pneumococcus. Hence, further work is required to determine the role of SPD\_0464 in pheromone export in future as detailed in Future plan. In addition, how the SHP is re-imported into the cell is still not clear. It was reported that *ami* operon, which encodes a multicomponent oligopeptide transporter, is involved in the re-importation of the SHP1358-C9 pheromone into the cell in *S. thermophilus* (Fleuchot 2012). The expression of the *shp1358* gene is abolished in an *amiCDE::spec* mutant. Pneumococcus also has an *ami* operon (SPD\_1667-SPD\_1671). Hence it is reasonable to assume that pneumococcal Ami transporter is responsible for SHP peptide importation. The role of pneumococcal Ami transporter in SHP peptide importation could be characterised in detail in future.

My data support the hypothesis that there is a potential cross-talk between Rgg144/Shp144 and Rgg939/Shp939. There are several lines of evidence to support this conclusion. Firstly, the full induction of one Rgg/Shp system required the presence of other system. In addition, the microarray data showed that there is a large overlap of regulon for each Rgg/Shp systems on mannose. Furthermore, a detailed analysis of microarray data also revealed that one Rgg/Shp circuit controlled the expression of loci potentially controlled by other Rgg. For example, Rgg939 upregulates the Rgg144-regulated VP1 locus (SPD\_0145-SPD\_0147). Currently, wider implications of this potential cross-talk are not known. It has been well-

documented in *S. pyogenes* that Rgg2 and Rgg3 together tightly regulate two linked genes encoding new peptide pheromones. Rgg2 was reported to activate transcription of and is required for full induction of the pheromone genes, while Rgg3 was shown to play an antagonistic role and represses pheromone expression. The authors reported that the regulatory interaction between Rgg2/Rgg3 systems determined production of biofilms (Chang et al., 2011).

The exploitation of QS systems for development of potential therapeutic targets rely largely on their detailed genetic characterisation. This involves how microbes process and transport QS signals. Hence, future studies should concentrate on how microbes process and transport pheromones as well as on structural characterisation and activation mechanism of Rgg/Shp circuits.

## **4.1 Rggs have effect on pneumococcal virulence and survival**

Rgg regulators have been reported to modulate gene expression in GAS, and implicated in regulation of microbial stress response and pathogenicity (Aggarwal et al., 2014, Chaussee, 2002, Chaussee et al., 2003, Chaussee et al., 2004). Rggs were reported to enhance the bacterial fitness in various host niches, and provide protection from host immune response. In recent decades, various researchers attempted to unravel the role of Rggs in virulence in different species of streptococci. However, only RopB and Rgg2/3 have been characterised in detail in pathogenesis (Chaussee, 2002, Chaussee et al., 2003, Chaussee et al., 2004, Chang et al., 2011, Lasarre et al., 2013). Little is known about the contribution of Rggs in pneumococcal pathogenesis.

The hypothesis on the role of Rgg quorum sensing systems in pneumococcal virulence and colonisation was tested in an *in vivo* murine model. The results demonstrated that mutation of *rgg144*, *rgg939*, *shp144* and *shp939* rendered the pneumococcus less virulent (Section 3.15), and less able to colonise the nasopharynx in a mouse model of pneumonia and colonisation (Section 3.16). The reduction in colonisation and virulence in the mutants is very likely due to the inability of mutants to utilize mannose efficiently, increased susceptibility to oxidative stress, and Rggs' regulatory role over several genes potentially involved in several important

biological pathways as discussed below.

### **A. Rggs are involved in mannose utilisation**

*Streptococcus pneumoniae* relies exclusively on carbohydrates as a carbon source and devotes 30% of all transport mechanisms to carbohydrate import. It is predicted that there are 21 phosphotransferase systems (PTS) and up to eight ATP-binding cassette (ABC) transporters that import at least 32 distinct carbohydrates (Bidossi et al., 2012). My data showed that Rggs play an important role in mannose utilisation. Host glycans found in the respiratory tract, such as O- and N-linked glycans as well as those in blood, such as transferrin, contain mannose. The pneumococcus is known to have mannosidase activity, and has the pathways to utilise mannose (Kahya et al., 2017, King, 2010). The assertion that Rggs play an important role in mannose utilisation is supported by the fact that the expression of both *rggs* was stimulated by galactose and mannose, and the absence of Rggs led to the reduced utilisation of mannose. Moreover, the transcriptional *lacZ* reporter assay results further confirmed the importance of Rgg growth in presence of mannose, as the reporter assays can directly evaluate the responsiveness of target promoter to selected conditions (Pessi et al., 2001). The result demonstrated that the highest induction of *lacZ* was obtained when XZ1 (P<sub>rgg144</sub>-*lacZ*-wt) was grown on mannose, then by galactose.

Despite pneumococcal Rggs' responsiveness to mannose and galactose, I did not detect any differentially expressed genes, as far as known, involved directly in mannose or galactose catabolism in putative Rgg144 and Rgg939 regulons. I hypothesise that galactose and mannose act as signals to alter pneumococcal phenotype *in vivo*. Pneumococcus usually occupies the upper respiratory system, but free carbohydrates are scarce in the airway, especially for glucose, which is the preferred carbon source for several microbes (Carvalho et al., 2011, Bidossi et al., 2012). The concentration of glucose is reported to be less than 1 mM, in contrast to its content in blood (~4-6 mM) (Philips et al., 2003, Shelburne et al., 2008). Therefore, *in vivo* growth in the nasopharynx requires alternative carbon sources. The pneumococcus acquires carbon through modification and import of complex glycans (Burnaugh et al., 2008, Yesilkaya et al., 2000, Marion et al., 2012, King, 2010, Carvalho et al., 2011). On the surface of the human respiratory tract there is a constant interaction between the pneumococcus and high molecular weight glycoproteins covering the apical epithelial surfaces of respiratory tract, such as mucin, which is rich in galactose and mannose.

Recent findings indicate that sugars derived from deglycosylation provide suitable carbon and energy sources for nasopharyngeal growth (Yesilkaya et al., 2008a, Terra et al., 2010a, Burnaugh et al., 2008). *S. pneumoniae* is equipped with at least 10 extracellular (exo- or endo) glycosidases with a broad range of specificities, including galactosidase, heaxosaminidase, and neuraminidases (Kahya et al., 2017, King, 2010). These enzymes can break down O-linked glycans (Marion et al., 2009, Terra et al., 2010a, Jeong et al., 2009, King et al., 2006), N-linked glycans (e.g. NanA, StrH, BgaA, NanB) (Burnaugh et al., 2008, King et al., 2006), and glycosaminoglycans (hyaluronic acid) (e.g. Hyl) (Marion et al., 2012), providing free sugars that can potentially be used by the pneumococcus to grow (Marion et al., 2012, Burnaugh et al., 2008).

The initial breach of glycan component of mucin is prevented due to the presence of terminal sialic acid (Manco et al., 2006). Lack of access to mucin sugars ensures that Rggs down regulate large number of genes involved in protein and capsule synthesis as well as those required for iron uptake and cell division. Such an expression profile ensures a lower growth rate and promotes a stable commensal existence on mucosal surface. However, once the sialic acid removed in parallel to gradual increase in pneumococcal numbers, the microbe will eventually have access to the sugars located 'below' sialic acid, such as galactose and mannose. Access to these sugars will subsequently increase the expression of cognate *shp* genes, hence the synthesis of Shp peptides, which then interact with their Rgg proteins, activating Rgg/Shp circuits to modulate target gene expression.

A similar phenotype for Rgg in mannose utilisation was reported in *Streptococcus pyogenes* (Dmitriev et al., 2006, Chang et al., 2015). Inactivation of *rgg* (*ropB*) disrupted coordinate expression of genes associated with the metabolism of nonglucose carbon sources, such as fructose, mannose, and sucrose. The changes were associated with an inability of the mutant strain to grow using these compounds as the primary carbon source (Dmitriev et al., 2006). Furthermore, Chang and his colleague demonstrated mannose could activate the Rgg-SHP signalling in *Streptococcus pyogenes*. The detailed mechanism of how mannose regulates the Rgg-SHP pathway is still poorly understood. The inactivation of Rgg has effect on mannose utilization might be due to the changes in regulatory activities from an intermediary transcriptional factor of PTS systems rather than the direct influence from mannose uptake (Chang et al., 2015), or due to inefficient cleavage and catabolism of mannose.

In the past few decades, the study of *S. pneumoniae* has been heavily focused on factors that directly impinge on host-pathogen interactions, such as toxins, cell wall components, adhesins and capsule (Kadioglu et al., 2008). In contrast, investigation of pneumococcal physiology has only recently been addressed, in spite of it being a fundamental aspect of pneumococcal survival *in vivo*. Historically, the study of bacterial catabolism of complex carbohydrates has contributed to understanding basic bacterial physiology. The limiting access of primary nutrients to potential pathogenic microbes is a crucial approach to cut down the habitability of pathogens on susceptible host tissues (Chang et al., 2015). Successful colonisation of the human host and invasion of normally sterile sites, such as the blood or subcutaneous tissues, is likely to require adaptation to growth with mannose which are found in N- and O-linked glycans. Therefore, increased understanding of these Rgg-mediated regulatory pathways is going to be extremely beneficial for rational interruption strategies and thereby weaken the ability of pneumococci to invade or colonise host tissue.

### **B. Rggs have role in capsule synthesis**

The success of *S. pneumoniae* as a commensal depends on its ability to colonise and be transmitted between humans (Wyllie et al., 2016, Hammerschmidt et al., 2005). The primary ecological niche for *S. pneumoniae* is the epithelium of the upper airways, presumably within the nasopharynx. For successful colonisation of the upper airways, the pneumococcus must overcome host-mediated trapping, killing and clearing, and adapt to a new environment and the amount of available nutrients (Wyllie et al., 2016). The optimal strategy for successful colonisation may be a compromise between requirements for a thick capsule to escape early innate immunity and for a thinner capsule when *S. pneumoniae* adheres to the host cells. My data shows that Rggs play a role in pneumococcal colonisation through their role in control of capsule expression.

The pneumococcal capsule is an important virulence factor. The expression of capsule was reported to reduce pneumococcal entrapment in the mucus, thereby allowing the pneumococcus to access the epithelial surfaces. This allows the microbe to gain a foothold *in vivo* (Vernatter and Pirofski, 2013, Nelson et al., 2007, Kadioglu et al., 2008). The pneumococcus can regulate capsule expression in response to environmental signals via a number of mechanisms. Down-regulation of capsule expression during the transparent phase is assumed to promote successful colonisation (Hammerschmidt et al., 2005), whereas up-

regulation in the opaque phase contributes to evasion of host immune responses after tissue invasion (Wyllie et al., 2016, Kim et al., 1999). Hence, it is very likely that the mutant strains, due to Rggs role in downregulation of capsule locus, produced more capsule in the nasopharynx, which led to the decrease in colonisation ability of the mutants. The extent of colonisation could have been also influenced by changes in biofilm formation, as Rgg144 positively regulates virulence peptide1 (VP1), which increases biofilm development (Cuevas et al., 2017). Combined, Rggs may increase adherence via downregulation of the capsule synthesis and increase in biofilm development via upregulation of VP1. This explanation is consistent with the *in vitro* results, which showed elevated level of capsule synthesis in the mutants compared to the wild type on mannose.

The positive association between Rgg expression and virulence is less intuitive, given that capsule production has been shown to enhance pneumococcal virulence. This contradiction, the reduced virulence in the mutants despite increased expression of *cps* locus, can be explained by different scenarios. Firstly, despite the increased expression of capsule in less-virulent Rgg deletion mutants, Rggs influence other genes that may play a role in virulence. For example, we have shown that the Rgg144-regulated VP1 is a potent virulence factor, thus lower levels of VP1 in the mutant may contribute to the decrease in virulence. Secondly, as our array data showed the regulation exerted by Rggs is condition specific. Therefore, as the pneumococcus migrates into the deeper tissue sites, its encounter with mannose may be limited than the concentration of mannose used in our *in vitro* regulon determination. The microarray data reveals further other possible mechanisms for reduction in colonisation and virulence. For example, we have seen reduction in expression of iron transport locus, changes in the expression of genes responsible for choline binding proteins, ATPase synthase, and cell division, which are known to be important for pneumococcal attachment, proliferation and energetics (Stenz et al., 2011).

### **C. Pneumococcal Rggs are involved in ROS resistance**

In addition, the reduced virulence in Rgg mutants also could be due to their susceptibility to oxidative stress, and consequently a decreased ability to colonise and invade. In this study, I determined the susceptibility of pneumococcus to ROS by exposing the pathogen to 0.5-1 mM paraquat (a superoxide-generating agent) or 20 and 40 mM hydrogen peroxide. It was found that deletion of *rgg144* and *rgg939* rendered pneumococci more susceptible to paraquat

and H<sub>2</sub>O<sub>2</sub>.

Although *S. pneumoniae* is a part of the commensal flora in the upper respiratory tract of humans, the microbe can migrate to the lower respiratory tract or can cause systemic disease when individuals have a weakened immune system (Potter et al., 2012). However, it is assumed that the pneumococci are exposed to almost normal air (20% O<sub>2</sub>) on top of the nasopharyngeal mucus layer, to micro-aerobic conditions (around 5% O<sub>2</sub>) in the lower respiratory tract, to anaerobic conditions when in blood and the cerebral spinal fluid (CSF) (Yesilkaya et al., 2013). During growth aerobically, *S. pneumoniae* produces high levels of H<sub>2</sub>O<sub>2</sub> (about 0.5–2 mM) due to pyruvate oxidase activity which can be further reduced to yield the damaging hydroxyl radicals via the Fenton reaction (Pesakhov et al., 2007). During pneumococcal inflammation, neutrophils in particular are recruited to the site of infection and release diverse ROS such as H<sub>2</sub>O<sub>2</sub>, OH●, and O<sub>2</sub>● through the oxidative burst, which can have adverse effects on cell viability, DNA, and other cellular compounds. Apart from the damaging effect of oxygen radicals, the variation in oxygen concentration may affect pneumococcal gene expression, which may have effect on capsular polysaccharide synthesis, carbon metabolism, competence development and membrane composition (Bortoni et al., 2009, Weiser et al., 2001, Pesakhov et al., 2007). Hence, efficient survival mechanisms are required to remove the adverse effects of reactive oxygen species (ROS).

Pneumococci have the ability to survive in various concentrations of oxygen and previous studies have demonstrated that the ability to resist the toxicity of ROS is vital for survival in oxygenated environments of host tissues (Hajaj et al., 2012). It has been reported that superoxide dismutase, thiol peroxidase, NADH oxidase and alkyl hydroperoxidase are involved in reducing levels of ROS in *S. pneumoniae* (Yesilkaya et al., 2013). The pneumococcus has neither catalase nor oxidative stress response regulators that are present in other bacteria, such as OxyR, SoxRS and PerR (Tettelin et al., 2002). Although several reports demonstrated that the pneumococcus contains various proteins and regulators to protect themselves from oxidative stress (Hajaj et al., 2017, Yesilkaya et al., 2000, Yesilkaya et al., 2013, Ulijasz et al., 2004), the mechanism of detoxification and the regulatory network of oxidative stress are still poorly understood.

Since pneumococcus does not express catalase, it is expected to rely on some other ways to decompose H<sub>2</sub>O<sub>2</sub>. The inactivation of Rggs led to pneumococci more susceptible to oxidative

stress could be explained as (i) Rgg might directly regulate the oxidative response through regulatory impact on oxygen detoxifying enzymes, or (ii) Rgg regulates the oxidative stress by mediating the expression of other genes involved in cellular metabolism. For example, capsule thickness and efficient production of ATP is known to be important for microbial oxidative stress resistance (Carvalho et al., 2013). The microarray data shows that inactivation of *rgg144* lead to 2.6 fold down-regulation of *gor*, codes for glutathione reductase. It is reasonable to suggest that *S. pneumoniae* can utilise and import extracellular glutathione by ABC transporters to detoxify the oxidative stress due to its disulfide reductase activity. Potter and his colleagues showed that the mutation of *gor* renders pneumococci susceptible to superoxide, and decreases its ability to colonise and invade in a murine model (Potter et al., 2012, Yesilkaya et al., 2013, Masip et al., 2006). Hence, the inactivation of *rgg* may affect the expression of *gor* and consequently renders pneumococcus more susceptibility to superoxide. In addition, Rgg regulates the oxidative stress by controlling the expression of the gene coding ABC transporters, which are mainly implicated in microbial nutrient uptake and removal of toxins and antibiotics (Ulijasz et al., 2004, Yesilkaya et al., 2013). The microarray data show several ABC transporter genes up-regulated in both  $\Delta$ *rgg144* and  $\Delta$ *rgg939*, including SPD\_0151, SPD\_0915, SPD\_0916, SPD\_0917, and SPD\_1463, which are involved in inorganic ion transport and metabolism. Rggs' involvement in oxidative stress response could be explained by the fact that Rggs act as the repressor of these iron transporters, preventing pneumococcus from oxidative stress by reducing levels of intracellular iron, hence controlling the hydroxyl ion producing from the Fenton reaction (Touati, 2000). Currently, it is not known whether the regulatory impact of Rgg on *gor* or the transporter genes is direct or indirect. This can be studied in future by EMSA and reporter assays as widely used in this study.

Similar phenotypic effect has also been reported for an *rgg* mutant of *S. pyogenes*, which was more sensitive to paraquat (Chaussee et al., 2004). In addition, the deletion of one of the pneumococcal *rggs* (SDP\_1952) rendered microbe more susceptible to paraquat rather than peroxide, which might be due to up regulation of thiol peroxidase in the mutants (Bortoni et al., 2009). However, in *S. pyogenes* M49 serotype strain deletion of *rgg* mutation rendered the mutant more resistant to hydrogen peroxide than the wild type strain and mutants also could decompose H<sub>2</sub>O<sub>2</sub> more efficiently (Pulliainen et al., 2008). Difference in Rggs role in oxidative stress resistance can be due to these microbes capacity to deal with reactive oxygen

species. More research on the relation between Rgg and Gor, and the iron transporters in oxidative stress will unravel the mechanism of Rgg mediated oxidative stress resistance. A comprehensive understanding of the mechanism of pneumococcal resistance to oxidative stress may contribute to our understanding of how microbe successfully colonise and develop invasive infection.

The role of Rgg on bacterial virulence has been reported in other streptococci as well. However, it was demonstrated that Rgg mutant was more virulent than its wild type strain in *S. pyogenes* in a murine intraperitoneal infection model (Chaussee et al., 2003, Chaussee et al., 2004, Pulliainen et al., 2008). The discrepancy between my study and that of previous one could be due to several reasons (i) Genome polymorphism: Although *S. pneumoniae* and *S. pyogenes* are members of *Firmicutes*, each species has a distinct biology, which is evident from the range of diseases each cause. Hence, in different species of streptococcus Rggs may play unique roles. (ii) The Rgg sequence diversity means that each Rgg may be controlling different set of genes. Indeed Rggs have various roles in different microorganisms. Such as RopB could regulate the expression of a secreted cysteine proteinase (SpeB) in GAS (Ajdic and Ferretti, 1999, Pulliainen et al., 2008, Lyon et al., 1998), also act as a global transcriptional regulator, mediate the expression of several extracellular proteins including M protein, C5a peptidase, the SLO cytolysin, and streptokinase (Chaussee, 2002, Watson et al., 2001), and control the expression of genes coding virulence factors including *mga*, *csrRS/covRS*, *fasBCA*, and *saga* (Chaussee et al., 2003, Chaussee et al., 2004).

Pneumococcal Rggs mainly regulate the expression of capsule locus, iron transporters, oxidative stress response, sugar metabolism and amino acid metabolism, which are supported by our microarray data. Strikingly, despite low similarity between Rgg144 and Rgg939, we observe substantial overlap among the regulons of these Rggs. It should be also noted that though there is an overlap, there are also unique targets regulated by each Rgg.

Rgg regulators are part of RRNPP (Rgg, regulator gene of glucosyltransferase; Rap, response regulator aspartate phosphatases; NprR, neutral protease regulator; PlcR, phosphatidylinositol-specific phospholipase C gene regulator; PrgX, pheromone-responsive transcription factor) family proteins (Lasarre et al., 2013, Rocha-Estrada et al., 2010, Cook and Federle, 2014). Structure–function studies showed that Rap, NprR, PlcR, and PrgX employ a structurally similar C-terminal tetratricopeptide (TPR)-like repeat domain to bind

their cognate peptide pheromones (Parashar et al., 2015). It may be possible that under different environmental conditions the conserved structural properties in different Rggs respond to same stimuli, which leads to the regulation of same genes, while differences in folding pattern or its affinity for the target DNA regulatory elements may provide target specificity to each Rgg, resulting in differences in regulon composition. Currently, there is no established paradigm for the action mechanism for Rggs, and the future structure-function studies similar to those done for PlcR and PrgX can test these hypotheses (Parashar et al., 2015).

Antibiotics have been instrumental in the fight against infectious diseases (Parashar et al., 2015), but the effectiveness of existing antibiotics is in danger due to the rapid rise of resistant bacteria. Soon after mainstream antibiotic usage, multi-resistant pneumococcal clones emerged and disseminated worldwide. During the 1970s and 1980s, pneumococci resistant to penicillin (MIC of 0.1 g/ml), erythromycin, and trimethoprim-sulfamethoxazole (TMP-SMX) spread rapidly globally, including to Australia, Papua New Guinea, Israel, Spain, Poland, South Africa, and the United States (Kim et al., 2016, Henderson et al., 1988, Klugman et al., 1986). There is an urgent need to identify new microbial targets for anti-infectives, which will allow the development of new classes of antibiotics. Most antibiotics act by directly inhibiting key central cell functions, namely DNA, protein or cell wall synthesis (Kohanski et al., 2010). A different approach is to target virulence factors, metabolic functions or environment responsive elements (Alksne and Projan, 2000). Our data clearly show that Rgg144 and Rgg939 can be potential targets for next-generation drugs. Hence, the future studies need to focus on the methods to interfere with the interaction between signal peptide and Rgg proteins to modulate pneumococcal virulence.

## Future plan

The pneumococcal Rgg-Shp quorum sensing systems are complex circuits carrying important cellular functions. Their importance in microbial physiology and virulence requires more detailed analysis in future. One area of research could be how peptide signals of Rggs are exported and imported back into the cell. It was reported that the Ami transporters are involved in the re-importation of the SHP1358 (15-23) pheromone into the cell in *S. thermophilus* (Fleuchot et al., 2011). The pneumococcal *amiA* gene is 85.6% identical to *amiA* of *S. thermophilus*. Hence, it is reasonable to hypothesise that pneumococcal Ami transporter is responsible for SHP peptide importation. The role of pneumococcal Ami transporter in SHP peptide importation can be characterised by *lacZ* reporter assay. The *amiA* mutant will be constructed by SOEing PCR. Then  $P_{shp144}$ -*lacZ* reporter can be transformed into  $\Delta$ *amiA* $\Delta$ *shp144*. The  $\beta$ -galactosidase activity will be tested for the new constructed strains:  $P_{shp144}$ -*lacZ*- $\Delta$ *amiA* $\Delta$ *shp144* with or without in addition of synthetic peptide Shp144-C12. The XZ26 ( $P_{shp144}$ -*lacZ*) will be use as a positive control. It is expected that the  $P_{shp144}$ -*lacZ*- $\Delta$ *amiA* $\Delta$ *shp144* strain can not response to the synthetic peptide. Hence, the sythetic peptide will not be sensed or imported by the cells lacing the oligopeptide transporter AmiA, but it could be sensed and imported into cells by the reporter strain with *ami* operons (XZ26). Similar experiment will be done for  $P_{shp939}$ -*lacZ* reporter.

How pneumococcus exports and processes SHP peptide is still unknown. The Eep protease plays a key role in the maturation of various pheromones and anti-pheromones of *E. faecalis* (An et al., 1999) and *S. gordonii* (Vickerman et al., 2010). Chang et al (2011) demonstrated that Eep encoded by spy49\_1620c has a role in processing of SHP3. Pneumococcus has an *eep* ortholog (SPD\_0245) which is 62.9% identical to spy49\_1620c; hence it is reasonable to hypothesis that pneumococcal Eep has a similar role in SHP peptide processing. Moreover, recently, an ABC transporter called PptAB was identified contributing to Rgg-SHP signalling pathway in *S. agalactiae* (Perez-Pascual et al., 2015) and *S. pyogenes* (Chang and Federle, 2016). The PptAB has been demonstrated as the primary transporter for SHP pheromones in GAS (Chang and Federle, 2016). The ortholog of gene encoding PptAB has been found in *S. pneumoniae* D39 by *in silico* analysis, the SPD\_0464, which shares 76.8% identity with PptAB. Hence I hypothesis that the ABC transporter encode by SPD\_0464 may be the

transporter for SHP pheromones in pneumococcus. To test this hypothesis, I would utilise a strain mutated in PptAB and compare the mutant strain against the wild type by mass spectrometry. This will be done in order to determine any changes in the composition and quantity of secreted SHP compared to the wild type in the spent culture supernate and in cell homogenates. Any change in composition and quantity of SHP peptide in PptAB mutant relative to the wild type would imply PptAB's involvement in export and processing of SHP in *S. pneumoniae*.

SHP/Rgg QS mechanisms are widespread among species of Streptococci, it would be interesting to investigate how Rgg-dependent quorum sensing mediate inter- and intra-species communication and if other pneumococcal Rggs could respond to Shp144 and Shp939. In order to do this, the  $P_{shp144-lacZ}$  reporter will be fused into *rgg999*, *rgg1518* and *rgg1952* mutants. The  $\beta$ -galactosidase activity will be tested with or without addition of sythtic peptide Shp144-C12. The XZ2 ( $P_{shp144-lacZ-wt}$ ) Strain will be used as a positive control. The new constructed reporter strains  $P_{shp144-lacZ-rgg999}$ ,  $P_{shp144-lacZ-rgg1518}$  and  $P_{shp144-lacZ-rgg1952}$  are expected have lower  $\beta$ -galactosidase activity compare with XZ2 ( $P_{shp144-lacZ-wt}$ ) with addition of Shp144-C12. These results will demonstrate that pneumococcal *rgg999*, *rgg1518* and *rgg1952* are necessary for  $P_{shp144}$  responsiveness to SHP144-C12. To test of the hypothesis of Rgg-dependent quorum sensing mediate intra-species communication, the XZ26 ( $P_{shp144-lacZ-\Delta shp144}$ ) will be grown to log phase and re-suspend in cell-free spent culture supernatants from various strains *S. pneumoniae* wild type (positive control), *S. pneumoniae*  $\Delta shp144$  (negative control), *S. pyogenes*, *S. agalactiae*, *S. thermophilus* and *S. dysgalactiae* subsp. *Equisimilis*. If any of the supernatants have the ability to induce the reporter strain, it will indicate this stain could produced a secreted peptide capable of stimulating the pneumococcal Rgg/Shp144 system, and the pneumococcal Rgg/Shp144 QS system can respond to SHPs produced by other streptococcal species.

EMSA assay will be done to detect the interaction of Rgg939 and selected promoters of *shp939*, SPD\_0315, SPD\_1041, SPD\_1127, SPD\_1370, SPD\_1517, and SPD\_2030. These targets were selected for EMSA because the whole operon of SPD\_0315 (The capsule locus), SPD\_1041 (the *nrd* operon, which involved in nucleic acid transport and metabolism), SPD\_1127 (*lipD* operon involved in lipid transport and metabolism), SPD\_1370 (the *rps* operon, the ribosomal protein), SPD\_1517 (the operon involved in defense mechanism), SPD\_2030 (The operon involved in ribosomal structure and biogenesis.) were all

differentially regulated in *rgg144* and *rgg939* mutants in microarray analysis. In addition, it is also worthy to test if SHP peptide have any effect on the interaction of Rgg and *shp*. The DNA binding reactions, coupling probes to the Rgg939 protein with or without the synthetic peptide Shp939-C8. One of the hypotheses is Shp939 increase the binding of Rgg939 to the *shp939* promoter, the synthetic peptide Shp939-C8 will be added after the mixture of Rgg939 and the *shp939* probe. The other hypothesis is Shp939 decrease the binding of Rgg939 to the *shp939* promoter, then Shp939-C8 will be mixed with Rgg939 and incubated for 10min, then continue to add the *shp939* probe. The mixture will be loaded on a native PAGE, and visualize by a TYPHOON Trio+scanner (GE Healthcare Life Sciences, UK) with a 526 nm short-pass wavelength filter.

It is also noteworthy to investigate whether pneumococcal Rgg-Shp quorum sensing pathways can be used as a potential target to develop anti-infective compounds against pneumococcal infections. Using Rgg-SHP as a model receptor-ligand target, it is interesting to identify chemical compounds that could specifically inhibit Rgg quorum-sensing circuits. A previous described method (Aggarwal et al., 2015) will be used to study interactions between Rgg144 proteins and their cognate peptide ligands Shp144, an *in vitro* high-throughput, compound library screen was utilized to identify compounds that specifically interfere with Rgg144-SHP144 interactions. Briefly, the mixture of Rgg144 and Shp144-C12 and each compound will be incubated and tested by fluorescence polarization assay. Then these compounds will be test by a *lacZ* reporter assay. The XZ26 ( $P_{shp144}$ -*lacZ*- $\Delta shp144$ ) will be supplemented with Shp144-C12 and different compounds, the  $\beta$ -galactosidase activity will be tested.

## Appendix

**Table 1. Microarray analysis of gene expression in  $\Delta$ rgg144 or  $\Delta$ rgg939 relative to wild type D39 grown micro-anaerobically in CDM supplemented with mannose or galactose\*.**

Locus	Name	Fold change of Rgg144		Fold change of Rgg939		Function
		Mannose	Galactose	Mannose	Galactose	
SPD_0001	<i>dnaA</i>	2.1		3.09		chromosomal replication initiator protein DnaA
SPD_0007				2.28		S4 domain protein
SPD_0008		2.43		5.96		septum formation initiator, putative
SPD_0009				2.62		hypothetical protein
SPD_0010				2.9		hypothetical protein
SPD_0011	<i>tilS</i>			2.17		tRNA(Ile)-lysidine synthetase
SPD_0015		3.62	2.87			
SPD_0017		3.08				
SPD_0020		6.75	2.17			
SPD_0046	<i>blpU</i>	2.91		2.09		bacteriocin BlpU
SPD_0047		6.79		5.46		hypothetical protein
SPD_0089					2.17	ABC transporter, permease protein
SPD_0093		2.19		2.61		hypothetical protein
SPD_0094		2.18				hypothetical protein
SPD_0116			2.05		2.4	hypothetical protein
SPD_0120				2.1		hypothetical protein
SPD_0121				1.92	2.64	hypothetical protein
SPD_0123				2.07		hypothetical protein
SPD_0126	<i>pspA</i>			2.66		pneumococcal surface protein A
SPD_0150		4.5		2.22		ABC transporter, substrate-binding protein
SPD_0151		4.04		2.92		lipoprotein
SPD_0179		2.71		2.8		lipoprotein, putative
SPD_0180		2.83		2.61		hypothetical protein
SPD_0181		2.78		2.45		conserved hypothetical protein TIGR00250
SPD_0187	<i>nrdD</i>	4.08		4.72		anaerobic ribonucleoside-triphosphate reductase
SPD_0189		3.95		4.85		acetyltransferase, GNAT family protein
SPD_0191		2.88		2.61		hypothetical protein
SPD_0192	<i>rpsJ</i>	2.72		2.53		ribosomal protein S10
SPD_0193	<i>rplC</i>	2.47		3.19		ribosomal protein L3
SPD_0194	<i>rplD</i>	2.16		2.74		ribosomal protein L4
SPD_0195	<i>rplW</i>			2.19		ribosomal protein L23
SPD_0197	<i>rpsS</i>			2.15		ribosomal protein S19
SPD_0198	<i>rplV</i>	2.7		3.06		ribosomal protein L22
SPD_0199	<i>rpsC</i>	2.28		2.64		ribosomal protein S3

SPD_0200	<i>rplP</i>	2.08		2.06		ribosomal protein L16
SPD_0201	<i>rpmC</i>	3.24		3.6		ribosomal protein L29
SPD_0202	<i>rpsQ</i>	2.84		3.55		ribosomal protein S17
SPD_0203	<i>rplN</i>	2.48		2.74		ribosomal protein L14
SPD_0214	<i>adk</i>	2.04				adenylate kinase
SPD_0215	<i>infA</i>	3.17		2.78		translation initiation factor IF-1
SPD_0216	<i>rpsM</i>	2.82		2.59		ribosomal protein S13
SPD_0217	<i>rpsK</i>	3.11		2.96		ribosomal protein S11
SPD_0218	<i>rpoA</i>	3.2		2.71		DNA-directed RNA polymerase, alpha subunit
SPD_0219	<i>rplQ</i>	4.54		5.03		ribosomal protein L17
SPD_0251	<i>rpsL</i>			2.2		ribosomal protein S12
SPD_0256				2.37		conserved hypothetical protein TIGR00053
SPD_0257				2.97		hypothetical protein
SPD_0308	<i>clpL</i>	5.3		2.18		ATP-dependent Clp protease, ATP-binding subunit
SPD_0313		2.44		3.35		hypothetical protein
SPD_0315	<i>cps2A</i>	2.98		3.15		integral membrane regulatory protein <i>Cps2A</i>
SPD_0316	<i>cps2B</i>	2.96		268		tyrosine-protein phosphatase <i>CpsB</i>
SPD_0317	<i>cps2C</i>	3.07		3.42		chain length determinant protein/polysaccharide export protein, MPA1 family protein
SPD_0318	<i>cps2D</i>	3.23		2.9		tyrosine-protein kinase <i>Cps2D</i> cytosolic ATPase domain
SPD_0319	<i>cps2E</i>	2.34		2.2		undecaprenylphosphate glucosephosphotransferase <i>Cps2E</i>
SPD_0320	<i>cps2T</i>			2.32		glycosyl transferase, group 1 family protein, putative
SPD_0322	<i>cps2G</i>	2.32		2.6		glycosyl transferase, group 1 family protein
SPD_0323	<i>csp2H</i>	3.58		4.44		Polysaccharide polymerase
SPD_0325		2.62		3.32		hypothetical protein
SPD_0326	<i>cps2K</i>	3.13		3.18		UDP-glucose 6-dehydrogenase, putative
SPD_0327	<i>cps2P</i>	4.48		4.9		UDP-galactopyranose mutase
SPD_0334	<i>aliA</i>	2.12		2.15		oligopeptide ABC transporter, oligopeptide-binding protein <i>AliA</i>
SPD_0337	<i>recU</i>	2.07		2.46		recombination protein U
SPD_0342				2.03		hypothetical protein
SPD_0365	<i>tig</i>			2.15		trigger factor
SPD_0373		4.6		2.19		hypothetical protein
SPD_0400				2.04		Glycosyl transferase family protein 8, putative
SPD_0420	<i>pflB</i>	2.39				formate acetyltransferase
SPD_0422				2.11		hypothetical protein
SPD_0441		2.04		2.04		DNA-directed RNA polymerase, delta subunit, putative
SPD_0451			2.27	2.47	2.26	type I restriction-modification system, S subunit, putative
SPD_0460	<i>dnaK</i>			2.14		chaperone protein DnaK
SPD_0473	<i>blpY</i>	2.61		2.61	2.13	immunity protein BlpY
SPD_0474		2.3				hypothetical protein

SPD_0493				2.25		hypothetical protein
SPD_0548				2.09		HIT family protein
SPD_0550	<i>rplK</i>	2.67		2.59		ribosomal protein L11
SPD_0551	<i>rplA</i>	3.21		3.58		ribosomal protein L1
SPD_0558	<i>prtA</i>			2.83		cell wall-associated serine protease PrtA
SPD_0577	<i>zmpB</i>	2.44		2.16		zinc metalloprotease ZmpB
SPD_0692				2.36		hypothetical protein
SPD_0702	<i>ciaH</i>	2.35				sensor histidine kinase CiaH
SPD_0710	<i>ezrA</i>	3.59		4.12		septation ring formation regulator EzrA
SPD_0712		3.19		2.02		transposase family protein
SPD_0726				2.35		purine nucleoside phosphorylase, family protein 2
SPD_0728				2.23		hypothetical protein
SPD_0750				2.48		hypothetical protein
SPD_0756		3.23	2.35			
SPD_0757	<i>rpsA</i>	2.34		2.35		ribosomal protein S1
SPD_0784				2.87		type I restriction-modification system, R subunit, putative
SPD_0793				2.43		hypothetical protein
SPD_0794				2.12		
SPD_0835	<i>frr</i>	2.29		3.24		ribosome recycling factor
SPD_0847	<i>infC</i>			2.25		translation initiation factor IF-3
SPD_0863	<i>smpB</i>			2.22		SsrA-binding protein
SPD_0876				2.02		hypothetical protein
SPD_0878				2.22		hypothetical protein
SPD_0899		2.48	2.9			
SPD_0905		2.09		3.72		acetyltransferase, GNAT family protein
SPD_0913					2	hypothetical protein
SPD_0915		3.32		9.31		iron-compound ABC transporter, iron compound-binding protein
SPD_0916				2.14		iron-compound ABC transporter, permease protein
SPD_0917				2.27		iron-compound ABC transporter, permease protein
SPD_0919				2.24		hypothetical protein
SPD_0920				3		hypothetical protein
SPD_0932				3.1		hypothetical protein
SPD_0940				2.46	1.84	UDP-N-acetyl-D-mannosaminuronic acid dehydrogenase, putative
SPD_0941				5.92	4.4	hypothetical protein
SPD_0942				4.56	4.72	hypothetical protein
SPD_0945				3.69	3.4	AMP-binding enzyme, putative
SPD_0945					3.4	AMP-binding enzyme, putative
SPD_0949				2.44	4.14	bacterial transferase hexapeptide (three repeats), putative

SPD_0950				2.06	2.61	transporter, major facilitator family protein
SPD_0961		2.25		2.82		glycosyl transferase, group 1
SPD_0963				2.49		hypothetical protein
SPD_0968				2.7		acetyltransferase, GNAT family protein
SPD_0969				2.4		hypothetical protein
SPD_0972				2.45		IS1381, transposase OrfB
SPD_0989	<i>rplU</i>	2.62				ribosomal protein L21
SPD_1041	<i>nrdH</i>	2.7		3.75		glutaredoxin-like protein NrdH
SPD_1042	<i>nrdE</i>	2.56		2.56		ribonucleoside-diphosphate reductase, alpha subunit
SPD_1043	<i>nrdF</i>	3.13		2.31		ribonucleoside-diphosphate reductase, beta subunit
SPD_1055				2.43		
SPD_1064				2.6		hemolysin A, putative
SPD_1076	<i>srtA</i>	2.24				sortase
SPD_1077	<i>gyrA</i>			2.09		DNA gyrase, A subunit
SPD_1080				2.52		type II restriction endonuclease, putative
SPD_1083	<i>vicX</i>			2.16		vicX protein
SPD_1084				2.64		sensory box sensor histidine kinase
SPD_1106		5.83				
SPD_1115	<i>leuB</i>			2.01		3-isopropylmalate dehydrogenase
SPD_1125	<i>pck</i>	2.12		2.75		choline kinase
SPD_1126				2.07		alcohol dehydrogenase, zinc-containing
SPD_1127	<i>ispD</i>	2.16		2.72		2-C-methyl-D-erythritol 4-phosphate cytidyltransferase
SPD_1129	<i>licD1</i>			2.59		phosphotransferase LicD1
SPD_1167				2.31		ABC transporter, ATP-binding protein
SPD_1175		2.49		7.3		hypothetical protein
SPD_1177				2.92		drug efflux ABC transporter, ATP-binding/permease protein
SPD_1199				2.18		glycosyl transferase, group 2 family protein
SPD_1200		2.11		2.66		glycosyl transferase, group 1 family protein
SPD_1297		2.48		2.42		pyridoxine biosynthesis protein
SPD_1338	<i>atpH</i>			3.47		ATP synthase F1, delta subunit
SPD_1339	<i>atpF</i>	2.43		3.77		ATP synthase F0, B subunit
SPD_1340	<i>atpB</i>			2.43		ATP synthase F0, A subunit
SPD_1346				3.25		Uncharacterized BCR, putative
SPD_1350				2.14		hypothetical protein
SPD_1365		2.03				hypothetical protein
SPD_1367		2.25		2.42		Cof family protein/peptidyl-prolyl cis-trans isomerase, cyclophilin type
SPD_1368	<i>rpsR</i>	2.71		2.25		ribosomal protein S18
SPD_1369	<i>ssb</i>	2.76		2.17		single-strand binding protein
SPD_1370	<i>rpsF</i>	2.65		2.78		ribosomal protein S6
SPD_1404	<i>tpiA</i>			2.23		triosephosphate isomerase
SPD_1410		5.12				

SPD_1422				2.21		hypothetical protein
SPD_1463				2.63		ABC transporter, substrate binding lipoprotein
SPD_1474	<i>divIVA</i>	2.5		2.8		cell division protein DivIVA
SPD_1475	<i>ylmH</i>	2.19		2.06		YlmH protein
SPD_1495				2.3		sugar ABC transporter, sugar-binding protein
SPD_1504	<i>nanA</i>			3.29		sialidase A precursor
SPD_1513		2.08		3.43		
SPD_1514		2.91		3.53		ABC transporter, ATP-binding protein
SPD_1515		3.23		3.95		hypothetical protein
SPD_1516		4.51		4.97		hypothetical protein
SPD_1517		2.17		2.19		hypothetical protein
SPD_1522		2.15		3.37		hypothetical protein
SPD_1529		2.12		2.14		
SPD_1558				2.08		conserved hypothetical protein TIGR00253
SPD_1566		2.24				hypothetical protein
SPD_1567	<i>trx</i>	2.28				thioredoxin
SPD_1581		8.85	2.27	2.13		
SPD_1588		4.88		4.54	2.34	hypothetical protein
SPD_1589		4.04		4.53		lipoprotein, putative
SPD_1590		3.24		2.34		general stress protein 24, putative
SPD_1591		4.23		3.4		hypothetical protein
SPD_1595				4.34		hypothetical protein
SPD_1642	<i>proWX</i>			2.02		choline transporter (glycine betaine transport system permease protein)
SPD_1643	<i>proV</i>			2.22		choline transporter
SPD_1644				2.27		hypothetical protein
SPD_1671	<i>amiA</i>	3.13		3.62		oligopeptide ABC transporter, oligopeptide-binding protein AmiA
SPD_1682			4.73			
SPD_1685		4.06				
SPD_1686		10.79				
SPD_1687			2			
SPD_1688		7.85	1.99	2.41		
SPD_1689		4.12				
SPD_1690		2.6				
SPD_1691		8.45	2.03			
SPD_1692		4.18	2.22			
SPD_1693		3.53				
SPD_1694		5.49	2.12			
SPD_1695		2.5	2.31			
SPD_1696		3.48	3.13			
SPD_1697		2.88				
SPD_1698		6.26				
SPD_1707				2.06		hypothetical protein

SPD_1730		2.11		3.11		
SPD_1739	<i>recA</i>			2.84		recA protein
SPD_1873				2.27		hypothetical protein
SPD_1879		4.76	4.56			
SPD_1881		2.14	2.33			
SPD_1882		4.07				
SPD_1896	<i>glxX</i>	2.45		4.03		glutamyl-tRNA synthetase
SPD_1898		6.73		6.78		hypothetical protein
SPD_1899		3.33		2.19		glutamine amidotransferase, class 1
SPD_1906				2.1		IS1381, transposase OrfB
SPD_1963	<i>rpmF</i>	2.43		2.43		ribosomal protein L32
SPD_1965	<i>pcpA</i>	4.56		4.49		choline binding protein PcpA
SPD_1967				2.11		IS1381, transposase OrfB
SPD_1981				2.6		hypothetical protein
SPD_2017	<i>cbpA</i>			2.42		choline binding protein A
SPD_2019				2.06		sensor histidine kinase
SPD_2020		2.71		4.35		DNA-binding response regulator
SPD_2030	<i>dnaB</i>	2.17				replicative DNA helicase
SPD_2031	<i>rplI</i>	2.01		2.02		ribosomal protein L9
SPD_2032		2.45		2.19		DHH subfamily 1 protein
SPD_2033	<i>yfiA</i>	2.84				ribosomal subunit interface protein
SPD_2037	<i>cysK</i>	3.85				cysteine synthase A
SPD_2043		3.59		3.98		secreted 45 kDa protein precursor
SPD_2059				5.41		
SPD_2060				2.89		transcriptional regulator, TetR family protein
SPD_2066		3.62				
SPD_2068		7.35		4.11		serine protease
SPD_2069		5.71		3.52	2.56	SpoJ protein
	<i>adhA</i>			-2.17		
SPD_0039				-2.24		hypothetical protein
SPD_0116				-2.85		hypothetical protein
SPD_0144			-2.47			hypothetical protein
SPD_0145		-6.75	-6.39	-2.45		hypothetical protein
SPD_0146		-2.67	-5.49	-2.64	2.6	CAAX amino terminal protease family protein
SPD_0147		-2.4	-3.92	-2.18	2.36	CAAX amino terminal protease family protein
SPD_0148			-3.65			hypothetical protein
SPD_0149			-1.83			ABC transporter, ATP-binding protein
SPD_0213	<i>secY</i>			-2.01		preprotein translocase, SecY subunit
SPD_0261	<i>pepC</i>	-2.16				aminopeptidase C
SPD_0276				-2.05		hypothetical protein
SPD_0339		-2.4				hypothetical protein
SPD_0425				-2.01		hypothetical protein
SPD_0442	<i>pyrG</i>			-2.12		CTP synthase

SPD_0445	<i>pgk</i>	-3.92				phosphoglycerate kinase
SPD_0458	<i>hrcA</i>			-2.98		heat-inducible transcription repressor HrcA
SPD_0459	<i>grpE</i>			-2.09		heat shock protein GrpE
SPD_0463				-2.06		Hit-like protein involved in cell-cycle regulation, putative
SPD_0465		-2.16				ABC transporter, permease protein, putative
SPD_0485	<i>hrcA</i>	-2.63		-2.03		hypothetical protein
SPD_0511	<i>metF</i>			-2.15		5,10-methylenetetrahydrofolate reductase
SPD_0582		-3.56		-2.38		hypothetical protein
SPD_0588				-2.09		transcriptional regulator, putative
SPD_0646				-2.15		hypothetical protein
SPD_0685	<i>gor</i>	-2.56				glutathione-disulfide reductase
	<i>gpmA</i>	-2.43				
SPD_0770				-2.62		
SPD_0772				-2.09		1-phosphofructokinase, putative
SPD_0900	<i>asd</i>			-2.32		transcriptional regulator
SPD_0997	<i>hup</i>	-2.06				DNA-binding protein HU
SPD_1004	<i>gapN</i>			-2.12		glyceraldehyde-3-phosphate dehydrogenase, NADP-dependent
SPD_1138	<i>htpX</i>	-2.34		-2.26		heat shock protein HtpX
SPD_1139	<i>lemA</i>	-2.77		-2.19		LemA protein
SPD_1255				-2.11		ABC transporter, ATP-binding protein
SPD_1295		-2.35				hemolysin
SPD_1320		-2.25				glycerol uptake facilitator protein, putative
SPD_1354				-2.04		hypothetical protein
SPD_1415		-2.07		-2.01		oxidoreductase, pyridine nucleotide-disulfide, class I
SPD_1418	<i>pepQ</i>	-2.22				proline dipeptidase PepQ
SPD_1506				-2.07		acetyl xylan esterase, putative
SPD_1513					-2.59	
SPD_1514			-2.04		-2.92	hypothetical protein
SPD_1515			-1.99		-2.2	hypothetical protein
SPD_1516			-2.41		-2.78	
SPD_1646		-2.1				hypothetical protein
SPD_1647	<i>pepA</i>	-2.33				glutamyl aminopeptidase PepA
SPD_1663	<i>treC</i>			-2.06		alpha,alpha-phosphotrehalase
SPD_1682				-3.83		
SPD_1683				-2.13		
SPD_1687				-2.38		
SPD_1735				-2.31		hypothetical protein
SPD_1790	<i>rpmH</i>	-2.01				ribosomal protein L34
SPD_1797	<i>ccpA</i>			-2.19		catabolite control protein A
SPD_1834				-2.41		alcohol dehydrogenase, iron-containing
SPD_1843				-2.05		hexulose-6-phosphate isomerase, putative

---

SPD_1917				-2.14		hypothetical protein
SPD_1932	<i>malP</i>			-3		maltodextrin phosphorylase
SPD_1933	<i>malQ</i>			-2.43		4-alpha-glucanotransferase
SPD_1956	<i>ilvD</i>			-2.08		dihydroxy-acid dehydratase
SPD_1984		-2.4		-2.56		hypersensitive-induced reaction protein 4
SPD_2006				-2.31		hypothetical protein
SPD_2009				-2.02		hypothetical protein

\*Fold changes  $\geq 2$  or  $\leq -2$ . All P-values are  $< 0.001$ .

## Reference

- ABBOTT, D. W., HIGGINS, M. A., HYRNUIK, S., PLUVINAGE, B., LAMMERTS VAN BUEREN, A. & BORASTON, A. B. 2010. The molecular basis of glycogen breakdown and transport in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 77, 183-199.
- ABEYTA, M., HARDY, G. G. & YOTHER, J. 2003. Genetic alteration of capsule type but not PspA type affects accessibility of surface-bound complement and surface antigens of *Streptococcus pneumoniae*. *Infect. Immun.*, 71, 218-225.
- AGGARWAL, C., JIMENEZ, J. C., LEE, H., CHLIPALA, G. E., RATIA, K. & FEDERLE, M. J. 2015. Identification of Quorum-Sensing Inhibitors Disrupting Signaling between Rgg and Short Hydrophobic Peptides in Streptococci. *MBio*, 6, e00393-15.
- AGGARWAL, C., JIMENEZ, J. C., NANAVATI, D. & FEDERLE, M. J. 2014. Multiple length peptide-pheromone variants produced by *Streptococcus pyogenes* directly bind Rgg proteins to confer transcriptional regulation. *J Biol Chem*, 289, 22427-36.
- AJDIC, D. & FERRETTI, J. J. 1999. The rgg gene of *Streptococcus pyogenes* NZ131 positively influences extracellular SPE B production. *Infect. Immun.*, 67, 1715-1722.
- AKERLEY, B. J., RUBIN, E. J., NOVICK, V. L., AMAYA, K., JUDSON, N. & MEKALANOS, J. J. 2002. A genome-scale analysis for identification of genes required for growth or survival of *Haemophilus influenzae*. *PNAS.*, 99, 966-971.
- AL-BAYATI, F. A., KAHYA, H. F., DAMIANOU, A., SHAFEEQ, S., KUIPERS, O. P., ANDREW, P. W. & YESILKAYA, H. 2017. Pneumococcal galactose catabolism is controlled by multiple regulators acting on pyruvate formate lyase. *Sci. Rep.*, 7.
- ALEXANDER, J. E., BERRY, A. M., PATON, J. C., RUBINS, J. B., ANDREW, P. W. & MITCHELL, T. J. 1998. Amino acid changes affecting the activity of pneumolysin alter the behaviour of pneumococci in pneumonia. *Microb. Pathog.*, 24, 167-174.
- ALIOING, G., GRANADEL, C., MORRISON, D. A. & CLAVERYS, J. P. 1996. Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 21, 471-478.
- ALKSNE, L. E. & PROJAN, S. J. 2000. Bacterial virulence as a target for antimicrobial chemotherapy. *Curr. Opin. Biotechnol.*, 11, 625-636.
- AN, F. Y., SULAVIK, M. C. & CLEWELL, D. B. 1999. Identification and characterization of a determinant (eep) on the *Enterococcus faecalis* chromosome that is involved in production of the peptide sex pheromone cAD1. *J. Bacteriol.*, 181, 5915-5921.
- ANBALAGAN, S., MCSHAN, W. M., DUNMAN, P. M. & CHAUSSEE, M. S. 2011. Identification of Rgg binding sites in the *Streptococcus pyogenes* chromosome. *J Bacteriol*, 193, 4933-42.
- ANDREINI, C., BERTINI, I., CAVALLARO, G., HOLLIDAY, G. L. & THORNTON, J. M. 2008. Metal ions in biological catalysis: from enzyme databases to general principles. *JBIC J. Biol. Inorg. Chem.*, 13, 1205-1218.
- AUZAT, I., CHAPUY - REGAUD, S., LE BRAS, G., DOS SANTOS, D., OGUNNIYI, A. D., LE THOMAS, I., GAREL, J. R., PATON, J. C. & TROMBE, M. C. 1999. The NADH oxidase of *Streptococcus pneumoniae*: its involvement in competence and virulence. *Mol. Microbiol.*, 34, 1018-1028.
- BAILEY, T. L. & ELKAN, C. 1994. Fitting a mixture model by expectation maximization to discover motifs in bipolymers.
- BARIL, L., DIETEMANN, J., ESSEVAZ - ROULET, M., BENIGUEL, L., COAN, P.,

- BRILES, D., GUY, B. & COZON, G. 2006. Pneumococcal surface protein A (PspA) is effective at eliciting T cell - mediated responses during invasive pneumococcal disease in adults. *Clin. Exp. Immunol.*, 145, 277-286.
- BARTILSON, M., MARRA, A., CHRISTINE, J., ASUNDI, J. S., SCHNEIDER, W. P. & HROMOCKYJ, A. E. 2001. Differential fluorescence induction reveals *Streptococcus pneumoniae* loci regulated by competence stimulatory peptide. *Mol. Microbiol.*, 39, 126-135.
- BEIER, D. 2012. Deviations From the Rule: Orphan and Atypical Response Regulators. *Two-component Systems in Bacteria*, 109.
- BERRY, A. M., LOCK, R. A. & PATON, J. C. 1996. Cloning and characterization of nanB, a second *Streptococcus pneumoniae* neuraminidase gene, and purification of the NanB enzyme from recombinant *Escherichia coli*. *J. Bacteriol.*, 178, 4854-4860.
- BETHE, G., NAU, R., WELLMER, A., HAKENBECK, R., REINERT, R. R., HEINZ, H.-P. & ZYSK, G. 2001. The cell wall-associated serine protease PrtA: a highly conserved virulence factor of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.*, 205, 99-104.
- BIDOSSI, A., MULAS, L., DECOROSI, F., COLOMBA, L., RICCI, S., POZZI, G., DEUTSCHER, J., VITI, C. & OGGIONI, M. R. 2012. A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. *PLoS ONE*, 7, e33320.
- BIRBEN, E., SAHINER, U. M., SACKESSEN, C., ERZURUM, S. & KALAYCI, O. 2012. Oxidative stress and antioxidant defense. *World Allergy Organ J.*, 5, 9.
- BLACK, S., SHINEFIELD, H., FIREMAN, B., LEWIS, E., RAY, P., HANSEN, J. R., ELVIN, L., ENSOR, K. M., HACKELL, J. & SIBER, G. 2000. Efficacy, safety and immunogenicity of heptavalent pneumococcal conjugate vaccine in children. *Pediatr. Infect. Dis. J.*, 19, 187-195.
- BLUE, C. E. & MITCHELL, T. J. 2003. Contribution of a response regulator to the virulence of *Streptococcus pneumoniae* is strain dependent. *Infect. Immun.*, 71, 4405-4413.
- BOGAERT, D., DE GROOT, R. & HERMANS, P. W. 2007. *Streptococcus pneumoniae*: From Colonization and Infection Towards Prevention Strategies. *Molecular Biology of Streptococci*, 247.
- BOGAERT, D., DE GROOT, R. & HERMANS, P. W. M. 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis*, 4, 144-154.
- BORTONI, M. E., TERRA, V. S., HINDS, J., ANDREW, P. W. & YESILKAYA, H. 2009. The pneumococcal response to oxidative stress includes a role for Rgg. *Microbiology*, 155, 4123-34.
- BOTELLA, H., PEYRON, P., LEVILLAIN, F., POINCLOUX, R., POQUET, Y., BRANDLI, I., WANG, C., TAILLEUX, L., TILLEUL, S. & CHARRIÈRE, G. M. 2011. Mycobacterial P 1-type ATPases mediate resistance to zinc poisoning in human macrophages. *Cell host & microbe*, 10, 248-259.
- BRADFORD, M. M. 1976. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.*, 72, 248-254.
- BRICKER, A. L. & CAMILLI, A. 1999. Transformation of a type 4 encapsulated strain of *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.*, 172, 131-135.
- BRILES, D. E., HOLLINGSHEAD, S. K., SWIATLO, E., BROOKS-WALTER, A., SZALAI, A., VIROLAINEN, A., MCDANIEL, L. S., BENTON, K. A., WHITE, P. & PRELLNER, K. 1997. PspA and PspC: their potential for use as pneumococcal vaccines. *Microb. Drug Resist.*, 3, 401-408.
- BRONSTEIN, I., FORTIN, J., STANLEY, P. E., STEWART, G. S. & KRICKA, L. J. 1994. Chemiluminescent and bioluminescent reporter gene assays. *Anal. Biochem.*, 219,

- 169-181.
- BROWN, J. S., GILLILAND, S. M., RUIZ-ALBERT, J. & HOLDEN, D. W. 2002. Characterization of pit, a *Streptococcus pneumoniae* iron uptake ABC transporter. *Infect. Immun.*, 70, 4389-4398.
- BROWN, P. O. & BOTSTEIN, D. 1999. Exploring the new world of the genome with DNA microarrays. *Nat. Genet.*, 21.
- BROWNING, D. F. & BUSBY, S. J. 2004. The regulation of bacterial transcription initiation. *Nat. Rev. Microbiol.*, 2, 57-65.
- BUCKWALTER, C. M. & KING, S. J. 2012. Pneumococcal carbohydrate transport: food for thought. *Trends Microbiol.*, 20, 517-522.
- BURNAUGH, A. M., FRANTZ, L. J. & KING, S. J. 2008. Growth of *Streptococcus pneumoniae* on human glycoconjugates is dependent upon the sequential activity of bacterial exoglycosidases. *J. Bacteriol.*, 190, 221-230.
- BYERS, J. T., LUCAS, C., SALMOND, G. P. & WELCH, M. 2002. Nonenzymatic turnover of an *Erwinia carotovora* quorum-sensing signaling molecule. *J. Bacteriol.*, 184, 1163-1171.
- CAMARA, M., BOULNOIS, G., ANDREW, P. & MITCHELL, T. 1994. A neuraminidase from *Streptococcus pneumoniae* has the features of a surface protein. *Infect. Immun.*, 62, 3688-3695.
- CARBON, C., VAN RENSBURG, D., HAGBERG, L., FOGARTY, C., TELLIER, G., RANGARAJU, M. & NUSRAT, R. 2006. Clinical and bacteriologic efficacy of telithromycin in patients with bacteremic community-acquired pneumonia. *Lancet Respir Med*, 100, 577-585.
- CAREY, M. F., PETERSON, C. L. & SMALE, S. T. 2009. Chromatin immunoprecipitation (chip). *Cold Spring Harbor Protocols*, 2009, pdb. prot5279.
- CARVALHO, S. M., ANDISI, V. F., GRADSTEDT, H., NEEF, J., KUIPERS, O. P., NEVES, A. R. & BIJLSMA, J. J. 2013. Pyruvate oxidase influences the sugar utilization pattern and capsule production in *Streptococcus pneumoniae*. *PLoS ONE*, 8, e68277.
- CARVALHO, S. M., KLOOSTERMAN, T. G., KUIPERS, O. P. & NEVES, A. R. 2011. CcpA ensures optimal metabolic fitness of *Streptococcus pneumoniae*. *PLoS ONE*, 6, e26707.
- CASINO, P., LÓPEZ-REDONDO, M. & MARINA, A. 2012. Structural Basis of Signal Transduction and Specificity in Two components Systems. *Two-component Systems in Bacteria*, 21.
- CASINO, P., RUBIO, V. & MARINA, A. 2010. The mechanism of signal transduction by two-component systems. *Curr. Opin. Struct. Biol.*, 20, 763-771.
- CHANG, J. C. & FEDERLE, M. J. 2016. PptAB Exports Rgg Quorum-Sensing Peptides in *Streptococcus*. *PLoS ONE*, 11, e0168461.
- CHANG, J. C., JIMENEZ, J. C. & FEDERLE, M. J. 2015. Induction of a quorum sensing pathway by environmental signals enhances group A streptococcal resistance to lysozyme. *Mol Microbiol.*
- CHANG, J. C., LASARRE, B., JIMENEZ, J. C., AGGARWAL, C. & FEDERLE, M. J. 2011. Two group A streptococcal peptide pheromones act through opposing Rgg regulators to control biofilm development. *PLoS Pathog*, 7, e1002190.
- CHARPENTIER, E., NOVAK, R. & TUOMANEN, E. 2000. Regulation of growth inhibition at high temperature, autolysis, transformation and adherence in *Streptococcus pneumoniae* by clpC. *Mol. Microbiol.*, 37, 717-726.
- CHAUSSEE, M. A., CALLEGARI, E. A. & CHAUSSEE, M. S. 2004. Rgg regulates growth phase-dependent expression of proteins associated with secondary metabolism and stress in *Streptococcus pyogenes*. *J Bacteriol*, 186, 7091-9.

- CHAUSSEE, M. S. 2002. Rgg Influences the Expression of Multiple Regulatory Loci To Coregulate Virulence Factor Expression in *Streptococcus pyogenes*. *Infect. Immun.*, 70, 762-770.
- CHAUSSEE, M. S., SOMERVILLE, G. A., REITZER, L. & MUSSER, J. M. 2003. Rgg Coordinates Virulence Factor Synthesis and Metabolism in *Streptococcus pyogenes*. *J. Bacteriol.*, 185, 6016-6024.
- CHIAVOLINI, D., POZZI, G. & RICCI, S. 2008. Animal models of *Streptococcus pneumoniae* disease. *Clin Microbiol Rev*, 21, 666-85.
- CIESLEWICZ, M. J., KASPER, D. L., WANG, Y. & WESSELS, M. R. 2001. Functional Analysis in Type Ia Group B *Streptococcus* of a Cluster of Genes Involved in Extracellular Polysaccharide Production by Diverse Species of Streptococci. *Biol. Chem.*, 276, 139-146.
- COHEN, J. M., CHIMALAPATI, S., DE VOGEL, C., VAN BELKUM, A., BAXENDALE, H. E. & BROWN, J. S. 2012. Contributions of capsule, lipoproteins and duration of colonisation towards the protective immunity of prior *Streptococcus pneumoniae* nasopharyngeal colonisation. *Vaccine*, 30, 4453-4459.
- COOK, L. C. & FEDERLE, M. J. 2014. Peptide pheromone signaling in *Streptococcus* and *Enterococcus*. *FEMS Microbiol Rev*, 38, 473-92.
- COOK, L. C., LASARRE, B. & FEDERLE, M. J. 2013. Interspecies communication among commensal and pathogenic streptococci. *MBio*, 4, e00382-13.
- CRAIN, M., WALTMAN, W., TURNER, J., YOTHER, J., TALKINGTON, D., MCDANIEL, L., GRAY, B. M. & BRILES, D. 1990. Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. *Infect. Immun.*, 58, 3293-3299.
- CUEVAS, R. A., EUTSEY, R., KADAM, A., WEST - ROBERTS, J. A., WOOLFORD, C. A., MITCHELL, A. P., MASON, K. M. & HILLER, N. L. 2017. A novel streptococcal cell - cell communication peptide promotes pneumococcal virulence and biofilm formation. *Mol. Microbiol.*
- DAGAN, R., PATTERSON, S., JUERGENS, C., GREENBERG, D., GIVON-LAVI, N., PORAT, N., GURTMAN, A., GRUBER, W. C. & SCOTT, D. A. 2013. Comparative immunogenicity and efficacy of 13-valent and 7-valent pneumococcal conjugate vaccines in reducing nasopharyngeal colonization: a randomized double-blind trial. *Clin. Infect. Dis.*, 57, 952-962.
- DE PAZ, H. D., SELVA, L. & MUÑOZ-ALMAGRO, C. 2015. Pneumococcal Vaccination and Consequences. 41-57.
- DE SAIZIEU, A., GARDÈS, C., FLINT, N., WAGNER, C., KAMBER, M., MITCHELL, T. J., KECK, W., AMREIN, K. E. & LANGE, R. 2000. Microarray-based identification of a novel *Streptococcus pneumoniae* regulon controlled by an autoinduced peptide. *J. Bacteriol.*, 182, 4696-4703.
- DECLERCK, N., BOUILLAUT, L., CHAIX, D., RUGANI, N., SLAMTI, L., HOH, F., LERECLUS, D. & AROLD, S. T. 2007. Structure of PlcR: Insights into virulence regulation and evolution of quorum sensing in Gram-positive bacteria. *PNAS.*, 104, 18490-18495.
- DEIBEL, R. & SEELEY JR, H. 1974. Family II. Streptococcaceae fam. nov. 7. *Streptococcus pneumoniae*. *Bergey's manual of determinative bacteriology*. The Williams & Wilkins Co., Baltimore, 499-500.
- DELALANDE, L., FAURE, D., RAFFOUX, A., UROZ, S., D'ANGELO-PICARD, C., ELASRI, M., CARLIER, A., BERRUYER, R., PETIT, A. & WILLIAMS, P. 2005. N-hexanoyl-L-homoserine lactone, a mediator of bacterial quorum-sensing regulation,

- exhibits plant-dependent stability and may be inactivated by germinating *Lotus corniculatus* seedlings. *FEMS Microbiol. Ecol.*, 52, 13-20.
- DMITRIEV, A. V., MCDOWELL, E. J., KAPPELER, K. V., CHAUSSEE, M. A., RIECK, L. D. & CHAUSSEE, M. S. 2006. The Rgg regulator of *Streptococcus pyogenes* influences utilization of nonglucose carbohydrates, prophage induction, and expression of the NAD-glycohydrolase virulence operon. *J Bacteriol.*, 188, 7230-41.
- DONG, Y.-H., WANG, L.-H., XU, J.-L., ZHANG, H.-B., ZHANG, X.-F. & ZHANG, L.-H. 2001. Quenching quorum-sensing-dependent bacterial infection by an N-acyl homoserine lactonase. *Nature*, 411, 813-817.
- DONG, Y.-H., XU, J.-L., LI, X.-Z. & ZHANG, L.-H. 2000. AiiA, an enzyme that inactivates the acylhomoserine lactone quorum-sensing signal and attenuates the virulence of *Erwinia carotovora*. *PNAS.*, 97, 3526-3531.
- DURMORT, C. & BROWN, J. S. 2015. *Streptococcus pneumoniae* Lipoproteins and ABC Transporters. *Streptococcus Pneumoniae 10*, 181-206.
- FEDERLE, M. 2012. Pathogenic streptococci speak, but what are they saying? *Virulence*, 3, 92-4.
- FEDSON, D. S., ANTHONY, J. & SCOTT, G. 1999. The burden of pneumococcal disease among adults in developed and developing countries: what is and is not known. *Vaccine*, 17, S11-S18.
- FEIL, S. C., ASCHER, D. B., KUIPER, M. J., TWETEN, R. K. & PARKER, M. W. 2014. Structural studies of *Streptococcus pyogenes* streptolysin O provide insights into the early steps of membrane penetration. *J Mol Cell Biol*, 426, 785-792.
- FERNANDEZ, A., BORGES, F., GINTZ, B., DECARIS, B. & LEBLOND-BOURGET, N. 2006. The *rggC* locus, with a frameshift mutation, is involved in oxidative stress response by *Streptococcus thermophilus*. *Arch Microbiol*, 186, 161-9.
- FLEUCHOT, B., GITTON, C., GUILLOT, A., VIDIC, J., NICOLAS, P., BESSET, C., FONTAINE, L., HOLS, P., LEBLOND-BOURGET, N., MONNET, V. & GARDAN, R. 2011. Rgg proteins associated with internalized small hydrophobic peptides: a new quorum-sensing mechanism in streptococci. *Mol Microbiol*, 80, 1102-19.
- FLEUCHOT, B., GUILLOT, A., MEZANGE, C., BESSET, C., CHAMBELLON, E., MONNET, V. & GARDAN, R. 2013. Rgg-associated SHP signaling peptides mediate cross-talk in Streptococci. *PLoS ONE*, 8, e66042.
- FONTAINE, L., BOUTRY, C., DE FRAHAN, M. H., DELPLACE, B., FREMAUX, C., HORVATH, P., BOYAVAL, P. & HOLS, P. 2010. A novel pheromone quorum-sensing system controls the development of natural competence in *Streptococcus thermophilus* and *Streptococcus salivarius*. *J. Bacteriol.*, 192, 1444-1454.
- FORBES, J. R. & GROS, P. 2003. Iron, manganese, and cobalt transport by Nramp1 (Slc11a1) and Nramp2 (Slc11a2) expressed at the plasma membrane. *Blood*, 102, 1884-1892.
- FUJIWARA, T., HOSHINO, T., OOSHIMA, T., SOBUE, S. & HAMADA, S. 2000. Purification, Characterization, and Molecular Analysis of the Gene Encoding Glucosyltransferase from *Streptococcus oralis*. *Infect. Immun.*, 68, 2475-2483.
- GALANTE, J., CY HO, A., TINGEY, S. & M CHARALAMBOUS, B. 2015. Quorum sensing and biofilms in the pathogen, *Streptococcus pneumoniae*. *Curr. Pharm. Des.*, 21, 25-30.
- GAO, R. & STOCK, A. M. 2009. Biological insights from structures of two-component proteins. *Annu. Rev. Microbiol.*, 63, 133.
- GARNER, M. M. & REVZIN, A. 1986. The use of gel electrophoresis to detect and study nucleic acid—protein interactions. *Trends Biochem. Sci.*, 11, 395-396.
- GHAFFAR, F., FRIEDLAND, I. R. & GEORGE H MCCRACKEN, J. 1999. Dynamics of

- nasopharyngeal colonization by *Streptococcus pneumoniae*. *Pediatr. Infect. Dis. J.*, 18, 638-646.
- GIAMMARINARO, P. & PATON, J. C. 2002. Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. *Infect. Immun.*, 70, 5454-5461.
- GORMAN, C. M., MOFFAT, L. F. & HOWARD, B. H. 1982. Recombinant genomes which express chloramphenicol acetyltransferase in mammalian cells. *J Mol Cell Biol*, 2, 1044-1051.
- GRANDCLÉMENT, C., TANNIÈRES, M., MORÉRA, S., DESSAUX, Y. & FAURE, D. D. 2015. Quorum quenching: role in nature and applied developments. *FEMS Microbiol. Rev.*, fuv038.
- GRATZ, N., LOH, L. N. & TUOMANEN, E. 2015. Pneumococcal Invasion. 433-451.
- GREENBERG, E. P. 2003. Bacterial communication and group behavior. *J. Clin. Invest.*, 112, 1288.
- GROSS, C., CHAN, C., DOMBROSKI, A., GRUBER, T., SHARP, M., TUPY, J. & YOUNG, B. The functional and regulatory roles of sigma factors in transcription. Cold Spring Harbor symposia on quantitative biology, 1998. Cold Spring Harbor Laboratory Press, 141-156.
- GRYLLOS, I., GRIFANTINI, R., COLAPRICO, A., CARY, M. E., HAKANSSON, A., CAREY, D. W., SUAREZ-CHAVEZ, M., KALISH, L. A., MITCHELL, P. D. & WHITE, G. L. 2008. PerR confers phagocytic killing resistance and allows pharyngeal colonization by group A *Streptococcus*. *PLoS Pathog.*, 4, e1000145.
- GUENZI, E., GASC, A. M., SICARD, M. A. & HAKENBECK, R. 1994. A two - component signal - transducing system is involved in competence and penicillin susceptibility in laboratory mutants of *Streptococcus pneumoniae*. *Mol. Microbiol.*, 12, 505-515.
- GUIRAL, S., HENARD, V., LAABERKI, M.-H., GRANADEL, C., PRUDHOMME, M., MARTIN, B. & CLAVERYS, J.-P. 2006. Construction and evaluation of a chromosomal expression platform (CEP) for ectopic, maltose-driven gene expression in *Streptococcus pneumoniae*. *Microbiology*, 152, 343-349.
- HAAS, W., KAUSHAL, D., SUBLETT, J., OBERT, C. & TUOMANEN, E. I. 2005. Vancomycin stress response in a sensitive and a tolerant strain of *Streptococcus pneumoniae*. *J. Bacteriol.*, 187, 8205-8210.
- HAJAJ, B., YESILKAYA, H., BENISTY, R., DAVID, M., ANDREW, P. W. & PORAT, N. 2012. Thiol peroxidase is an important component of *Streptococcus pneumoniae* in oxygenated environments. *Infect. Immun.*, 80, 4333-4343.
- HAJAJ, B., YESILKAYA, H., SULMAN SHAFEEQ, X. Z., BENISTY, R., TCHALAH, S., KUIPERS, O. P. & PORAT, N. 2017. CodY Regulates Thiol Peroxidase Expression as Part of the Pneumococcal Defense Mechanism against H<sub>2</sub>O<sub>2</sub> Stress. *Front Cell Infect Microbiol*, 7.
- HAKANSSON, A. P., MARKS, L. R. & ROCHE-HAKANSSON, H. 2015. Pneumococcal Genetic Transformation During Colonization and Biofilm Formation. 129-142.
- HALFMANN, A., HAKENBECK, R. & BRUCKNER, R. 2007. A new integrative reporter plasmid for *Streptococcus pneumoniae*. *FEMS Microbiol Lett*, 268, 217-24.
- HAMMERSCHMIDT, S., WOLFF, S., HOCKE, A., ROSSEAU, S., MÜLLER, E. & ROHDE, M. 2005. Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect. Immun.*, 73, 4653-4667.
- HARRINGTON, C. A., ROSENOW, C. & RETIEF, J. 2000. Monitoring gene expression using DNA microarrays. *Curr. Opin. Microbiol.*, 3, 285-291.
- HAVA, D. L. & CAMILLI, A. 2002. Large - scale identification of serotype 4 *Streptococcus*

- pneumoniae* virulence factors. *Mol. Microbiol.*, 45, 1389-1406.
- HEMSLEY, C., JOYCE, E., HAVA, D. L., KAWALE, A. & CAMILLI, A. 2003. MgrA, an Orthologue of Mga, Acts as a Transcriptional Repressor of the Genes within the rlrA Pathogenicity Islet in *Streptococcus pneumoniae*. *J. Bacteriol.*, 185, 6640-6647.
- HENDERSON, F. W., GILLIGAN, P. H., WAIT, K. & GOFF, D. A. 1988. Nasopharyngeal carriage of antibiotic-resistant pneumococci by children in group day care. *Int. J. Infect. Dis.*, 157, 256-263.
- HENDRIKSEN, W. T., BOOTSMA, H. J., ESTEVAO, S., HOOGENBOEZEM, T., DE JONG, A., DE GROOT, R., KUIPERS, O. P. & HERMANS, P. W. 2008. CodY of *Streptococcus pneumoniae*: link between nutritional gene regulation and colonization. *J. Bacteriol.*, 190, 590-601.
- HENRICHSEN, J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J. Clin. Microbiol.*, 33, 2759.
- HENRY, L. G., MCKENZIE, R. M., ROBLES, A. & FLETCHER, H. M. 2012. Oxidative stress resistance in *Porphyromonas gingivalis*. *Future Microbiol.*, 7, 497-512.
- HENRY, R., BRUNEAU, E., GARDAN, R., BERTIN, S., FLEUCHOT, B., DECARIS, B. & LEBLOND-BOURGET, N. 2011. The *rgg0182* gene encodes a transcriptional regulator required for the full *Streptococcus thermophilus* LMG18311 thermal adaptation. *BMC Microbiol.*, 11, 1.
- HOLDEN, M. T., DIGGLE, S. P. & WILLIAMS, P. 2007. Quorum Sensing. *eLS*.
- HONSA, E. S., JOHNSON, M. D. & ROSCH, J. W. 2013. The roles of transition metals in the physiology and pathogenesis of *Streptococcus pneumoniae*. *Front Cell Infect Microbiol.*, 3.
- HOOVER, S. E., PEREZ, A. J., TSUI, H. C., SINHA, D., SMILEY, D. L., DIMARCHI, R. D., WINKLER, M. E. & LAZAZZERA, B. A. 2015. A new quorum-sensing system (TprA/PhrA) for *Streptococcus pneumoniae* D39 that regulates a lantibiotic biosynthesis gene cluster. *Mol. Microbiol.*, 97, 229-43.
- HORTON, R. M., HUNT, H. D., HO, S. N., PULLEN, J. K. & PEASE, L. R. 1989. Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene*, 77, 61-68.
- HOSKINS, J., ALBORN, W. E., ARNOLD, J., BLASZCZAK, L. C., BURGETT, S., DEHOFF, B. S., ESTREM, S. T., FRITZ, L., FU, D.-J. & FULLER, W. 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J. Bacteriol.*, 183, 5709-5717.
- HYAMS, C., CAMBERLEIN, E., COHEN, J. M., BAX, K. & BROWN, J. S. 2010. The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infect. Immun.*, 78, 704-715.
- IBRAHIM, M., NICOLAS, P., BESSIERES, P., BOLOTIN, A., MONNET, V. & GARDAN, R. 2007. A genome-wide survey of short coding sequences in streptococci. *Microbiology*, 153, 3631-3644.
- IYER, R., BALIGA, N. S. & CAMILLI, A. 2005. Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. *J. Bacteriol.*, 187, 8340-8349.
- JABADO, N., JANKOWSKI, A., DOUGAPARSAD, S., PICARD, V., GRINSTEIN, S. & GROS, P. 2000. Natural resistance to intracellular infections. *J. Exp. Med.*, 192, 1237-1248.
- JEDRZEJAS, M. J., LAMANI, E. & BECKER, R. S. 2001. Characterization of selected strains of pneumococcal surface protein A. *Biol. Chem.*, 276, 33121-33128.
- JEFFERSON, R. A., KAVANAGH, T. A. & BEVAN, M. W. 1987. GUS fusions: beta-glucuronidase as a sensitive and versatile gene fusion marker in higher plants. *The*

- EMBO journal*, 6, 3901.
- JEONG, J. K., KWON, O., LEE, Y. M., OH, D.-B., LEE, J. M., KIM, S., KIM, E.-H., LE, T. N., RHEE, D.-K. & KANG, H. A. 2009. Characterization of the *Streptococcus pneumoniae* BgaC protein as a novel surface  $\beta$ -galactosidase with specific hydrolysis activity for the Gal $\beta$ 1-3GlcNAc moiety of oligosaccharides. *J. Bacteriol.*, 191, 3011-3023.
- JIMENEZ, J. C. & FEDERLE, M. J. 2014. Quorum sensing in group A *Streptococcus*. *Front Cell Infect Microbiol*, 4.
- JING, D., AGNEW, J., PATTON, W. F., HENDRICKSON, J. & BEECHEM, J. M. 2003. A sensitive two - color electrophoretic mobility shift assay for detecting both nucleic acids and protein in gels. *Proteomics*, 3, 1172-1180.
- JOHNSTON, J. W., MYERS, L. E., OCHS, M. M., BENJAMIN, W. H., BRILES, D. E. & HOLLINGSHEAD, S. K. 2004. Lipoprotein PsaA in virulence of *Streptococcus pneumoniae*: surface accessibility and role in protection from superoxide. *Infect. Immun.*, 72, 5858-5867.
- JONES, R. N., SADER, H. S., MENDES, R. E. & FLAMM, R. K. 2013. Update on antimicrobial susceptibility trends among *Streptococcus pneumoniae* in the United States: report of ceftaroline activity from the SENTRY Antimicrobial Surveillance Program (1998–2011). *Diagn. Microbiol. Infect. Dis.*, 75, 107-109.
- KADIOGLU, A., WEISER, J. N., PATON, J. C. & ANDREW, P. W. 2008. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat. Rev. Microbiol.*, 6, 288-301.
- KAHYA, H. F., ANDREW, P. W. & YESILKAYA, H. 2017. Deacetylation of sialic acid by esterases potentiates pneumococcal neuraminidase activity for mucin utilization, colonization and virulence. *PLoS Pathog.*, 13, e1006263.
- KAPLAN, S. L., BARSON, W. J., LIN, P. L., ROMERO, J. R., BRADLEY, J. S., TAN, T. Q., HOFFMAN, J. A., GIVNER, L. B. & MASON JR, E. O. 2013. Early trends for invasive pneumococcal infections in children after the introduction of the 13-valent pneumococcal conjugate vaccine. *Pediatr. Infect. Dis. J.*, 32, 203-207.
- KAPPELER, K. V., ANBALAGAN, S., DMITRIEV, A. V., MCDOWELL, E. J., NEELY, M. N. & CHAUSSEE, M. S. 2009. A naturally occurring Rgg variant in serotype M3 *Streptococcus pyogenes* does not activate speB expression due to altered specificity of DNA binding. *Infect Immun*, 77, 5411-7.
- KATAOKA, K., FUKUYAMA, Y., BRILES, D. E., MIYAKE, T. & FUJIHASHI, K. 2017. Dendritic cell - targeting DNA - based nasal adjuvants for protective mucosal immunity to *Streptococcus pneumoniae*. *Contrib. Microbiol. Immunol.*
- KILIAN, M., MESTECKY, J. & SCHROHENLOHER, R. E. 1979. Pathogenic species of the genus *Haemophilus* and *Streptococcus pneumoniae* produce immunoglobulin A1 protease. *Infect. Immun.*, 26, 143-149.
- KIM, J. O., ROMERO-STEINER, S., SØRENSEN, U. B. S., BLOM, J., CARVALHO, M., BARNARD, S., CARLONE, G. & WEISER, J. N. 1999. Relationship between cell surface carbohydrates and intrastain variation on opsonophagocytosis of *Streptococcus pneumoniae*. *Infect. Immun.*, 67, 2327-2333.
- KIM, L., MCGEE, L., TOMCZYK, S. & BEALL, B. 2016. Biological and epidemiological features of antibiotic-resistant *Streptococcus pneumoniae* in pre-and post-conjugate vaccine eras: a United States perspective. *Clin. Microbiol. Rev.*, 29, 525-552.
- KING, S. 2010. Pneumococcal modification of host sugars: a major contributor to colonization of the human airway ? *Mol Oral Microbiol*, 25, 15-24.
- KING, S. J., HIPPE, K. R. & WEISER, J. N. 2006. Deglycosylation of human

- glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. *Mol. Microbiol.*, 59, 961-974.
- KLOOSTERMAN, T. G., DER KOOI - POL, V., MAGDALENA, M., BIJLSMA, J. J. & KUIPERS, O. P. 2007. The novel transcriptional regulator SczA mediates protection against Zn<sup>2+</sup> stress by activation of the Zn<sup>2+</sup> resistance gene *czcD* in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 65, 1049-1063.
- KLUGMAN, K. P. 1990. Pneumococcal resistance to antibiotics. *Clin. Microbiol. Rev.*, 3, 171-196.
- KLUGMAN, K. P., KOORNHOF, H. J. & KUHNLE, V. 1986. Clinical and nasopharyngeal isolates of unusual multiply resistant pneumococci. *Am. J. Dis. Child.*, 140, 1186-1190.
- KOHANSKI, M. A., DWYER, D. J. & COLLINS, J. J. 2010. How antibiotics kill bacteria: from targets to networks. *Nat. Rev. Microbiol.*, 8, 423-435.
- KREIKEMEYER, B., MCIVER, K. S. & PODBIELSKI, A. 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen–host interactions. *Trends Microbiol.*, 11, 224-232.
- KYD, J. M., KRISHNAMURTHY, A. & KIDD, S. 2016. Interactions and Mechanisms of Respiratory Tract Biofilms Involving *Streptococcus Pneumoniae* and Nontypeable Haemophilus Influenzae. *Microbial Biofilms-Importance and Applications*. InTech.
- LAI, Y.-C., PENG, H.-L. & CHANG, H.-Y. 2003. RmpA2, an activator of capsule biosynthesis in *Klebsiella pneumoniae* CG43, regulates K2 cps gene expression at the transcriptional level. *J. Bacteriol.*, 185, 788-800.
- LANGE, R., WAGNER, C., DE SAIZIEU, A., FLINT, N., MOLNOS, J., STIEGER, M., CASPERS, P., KAMBER, M., KECK, W. & AMREIN, K. E. 1999. Domain organization and molecular characterization of 13 two-component systems identified by genome sequencing of *Streptococcus pneumoniae*. *Gene*, 237, 223-234.
- LASARRE, B., AGGARWAL, C. & FEDERLE, M. J. 2013. Antagonistic Rgg regulators mediate quorum sensing via competitive DNA binding in *Streptococcus pyogenes*. *MBio*, 3.
- LAU, G. W., HAATAJA, S., LONETTO, M., KENSIT, S. E., MARRA, A., BRYANT, A. P., MCDEVITT, D., MORRISON, D. A. & HOLDEN, D. W. 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. *Mol. Microbiol.*, 40, 555-571.
- LAUB, M. T. & GOULIAN, M. 2007. Specificity in two-component signal transduction pathways. *Annu. Rev. Genet.*, 41, 121-145.
- LAWRENCE, M. C., PILLING, P. A., EPA, V. C., BERRY, A. M., OGUNNIYI, A. D. & PATON, J. C. 1998. The crystal structure of pneumococcal surface antigen PsaA reveals a metal-binding site and a novel structure for a putative ABC-type binding protein. *Structure*, 6, 1553-1561.
- LECHARDEUR, D., FERNANDEZ, A., ROBERT, B., GAUDU, P., TRIEU-CUOT, P., LAMBERET, G. & GRUSS, A. 2010. The 2-Cys peroxiredoxin alkyl hydroperoxide reductase c binds heme and participates in its intracellular availability in *Streptococcus agalactiae*. *Biol. Chem.*, 285, 16032-16041.
- LEE, C.-J., BANKS, S. D. & LI, J. P. 1991. Virulence, immunity, and vaccine related to *Streptococcus pneumoniae*. *Crit. Rev. Microbiol.*, 18, 89-114.
- LINTON, K. J. & HIGGINS, C. F. 2007. Structure and function of ABC transporters: the ATP switch provides flexible control. *Pflugers Arch.*, 453, 555-567.
- LIU, L., OZA, S., HOGAN, D., PERIN, J., RUDAN, I., LAWN, J. E., COUSENS, S., MATHERS, C. & BLACK, R. E. 2015. Global, regional, and national causes of child mortality in 2000–13, with projections to inform post-2015 priorities: an updated systematic analysis. *The Lancet*, 385, 430-440.

- LUO, P. & MORRISON, D. A. 2003. Transient association of an alternative sigma factor, ComX, with RNA polymerase during the period of competence for genetic transformation in *Streptococcus pneumoniae*. *J. Bacteriol.*, 185, 349-358.
- LYON, W. R., GIBSON, C. M. & CAPARON, M. G. 1998. A role for Trigger Factor and an Rgg - like regulator in the transcription, secretion and processing of the cysteine proteinase of *Streptococcus pyogenes*. *The EMBO Journal*, 17, 6263-6275.
- LÓPEZ, R. & GARCÍA, E. 2004. Recent trends on the molecular biology of pneumococcal capsules, lytic enzymes, and bacteriophage. *FEMS Microbiol. Rev.*, 28, 553-580.
- LÓPEZ, R., GARCÍA, J., GARCÍA, E., RONDA, C. & GARCÍA, P. 1992. Structural analysis and biological significance of the cell wall lytic enzymes of *Streptococcus pneumoniae* and its bacteriophage. *FEMS Microbiol. Lett.*, 100, 439-447.
- LÓPEZ-MAURY, L., MARGUERAT, S. & BÄHLER, J. 2008. Tuning gene expression to changing environments: from rapid responses to evolutionary adaptation. *Nat. Rev. Genet.*, 9, 583-593.
- MANCO, S., HERNON, F., YESILKAYA, H., PATON, J. C., ANDREW, P. W. & KADIOGLU, A. 2006. Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect. Immun.*, 74, 4014-4020.
- MARGOLIS, E. 2009. Hydrogen peroxide-mediated interference competition by *Streptococcus pneumoniae* has no significant effect on *Staphylococcus aureus* nasal colonization of neonatal rats. *J. Bacteriol.*, 191, 571-575.
- MARION, C., LIMOLI, D. H., BOBULSKY, G. S., ABRAHAM, J. L., BURNAUGH, A. M. & KING, S. J. 2009. Identification of a pneumococcal glycosidase that modifies O-linked glycans. *Infect. Immun.*, 77, 1389-1396.
- MARION, C., STEWART, J. M., TAZI, M. F., BURNAUGH, A. M., LINKE, C. M., WOODIGA, S. A. & KING, S. J. 2012. *Streptococcus pneumoniae* can utilize multiple sources of hyaluronic acid for growth. *Infect. Immun.*, 80, 1390-1398.
- MARKS, L. R., MASHBURN-WARREN, L., FEDERLE, M. J. & HAKANSSON, A. P. 2014. *Streptococcus pyogenes* biofilm growth in vitro and in vivo and its role in colonization, virulence, and genetic exchange. *J Infect Dis*, 210, 25-34.
- MARTINDALE, J. L. & HOLBROOK, N. J. 2002. Cellular response to oxidative stress: signaling for suicide and survival. *J. Cell. Physiol.*, 192, 1-15.
- MASHBURN-WARREN, L., MORRISON, D. A. & FEDERLE, M. J. 2010. A novel double-tryptophan peptide pheromone controls competence in *Streptococcus spp.* via an Rgg regulator. *Mol Microbiol*, 78, 589-606.
- MASIP, L., VEERAVALLI, K. & GEORGIU, G. 2006. The many faces of glutathione in bacteria. *Antioxidants & redox signaling*, 8, 753-762.
- MAVROIDI, A., GODOY, D., AANENSEN, D. M., ROBINSON, D. A., HOLLINGSHEAD, S. K. & SPRATT, B. G. 2004. Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J. Bacteriol.*, 186, 8181-8192.
- MAYANSKIY, A. N., CHEBOTAR, I. V., LAZAREVA, A. V. & MAYANSKIY, N. A. 2015. Biofilm formation by *Streptococcus pneumoniae*. *Molecular Genetics, Mol. Gen. Microbiol. Virol.*, 30, 124-131.
- MCALLISTER, L. J., TSENG, H. J., OGUNNIYI, A. D., JENNINGS, M. P., MCEWAN, A. G. & PATON, J. C. 2004. Molecular analysis of the psa permease complex of *Streptococcus pneumoniae*. *Mol. Microbiol.*, 53, 889-901.
- MCCLUSKEY, J., HINDS, J., HUSAIN, S., WITNEY, A. & MITCHELL, T. 2004. A two - component system that controls the expression of pneumococcal surface antigen A (PsaA) and regulates virulence and resistance to oxidative stress in *Streptococcus*

- pneumoniae*. *Mol. Microbiol.*, 51, 1661-1675.
- MCDEVITT, C. A., OGUNNIYI, A. D., VALKOV, E., LAWRENCE, M. C., KOBE, B., MCEWAN, A. G. & PATON, J. C. 2011. A molecular mechanism for bacterial susceptibility to zinc. *PLoS Pathog.*, 7, e1002357.
- MCIVER, K. S. 2009. Stand-alone response regulators controlling global virulence networks in *Streptococcus pyogenes*. *Bacterial Sensing and Signaling*. Karger Publishers.
- MILES, A. A., MISRA, S. & IRWIN, J. 1938. The estimation of the bactericidal power of the blood. *Epidemiol. Infect.*, 38, 732-749.
- MILLER, J. 1972. Experiments in molecular genetics. Cold Spring Laboratory Press. Cold Spring Harbor, NY.
- MILTON, D. L. 2006. Quorum sensing in vibrios: complexity for diversification. *Int. J. Med. Microbiol.*, 296, 61-71.
- MISHRA, S. & IMLAY, J. 2012. Why do bacteria use so many enzymes to scavenge hydrogen peroxide? *Arch. Biochem. Biophys.*, 525, 145-160.
- MITCHELL, A. & MITCHELL, T. 2010. *Streptococcus pneumoniae*: virulence factors and variation. *Clin. Microbiol. Infect.*, 16, 411-418.
- MOHAWK, K. L. & O'BRIEN, A. D. 2011. Mouse models of Escherichia coli O157: H7 infection and shiga toxin injection. *Biomed Res Int*, 2011.
- MOHEDANO, M. L., OVERWEG, K., DE LA FUENTE, A., REUTER, M., ALTABE, S., MULHOLLAND, F., DE MENDOZA, D., LÓPEZ, P. & WELLS, J. M. 2005. Evidence that the essential response regulator YycF in *Streptococcus pneumoniae* modulates expression of fatty acid biosynthesis genes and alters membrane composition. *J. Bacteriol.*, 187, 2357-2367.
- MOLLERACH, M., LÓPEZ, R. & GARCÍA, E. 1998. Characterization of the galU gene of *Streptococcus pneumoniae* encoding a uridine diphosphoglucose pyrophosphorylase: a gene essential for capsular polysaccharide biosynthesis. *J. Exp. Med.*, 188, 2047-2056.
- MOLLOY, P. L. 2000. Electrophoretic mobility shift assays. *Transcription Factor Protocols*, 235-246.
- MORTON, D. A. & GRIFFITHS, P. 1985. Guidelines on the recognition of pain, distress and discomfort in experimental animals and an hypothesis for assessment. *Vet Rec*, 116, 431-436.
- MUÑOZ-ALMAGRO, C., JORDAN, I., GENE, A., LATORRE, C., GARCIA-GARCIA, J. J. & PALLARES, R. 2008. Emergence of invasive pneumococcal disease caused by nonvaccine serotypes in the era of 7-valent conjugate vaccine. *Clin. Infect. Dis.*, 46, 174-182.
- NAMKOONG, H., ISHII, M., FUNATSU, Y., KIMIZUKA, Y., YAGI, K., ASAMI, T., ASAKURA, T., SUZUKI, S., KAMO, T. & FUJIWARA, H. 2016. Theory and strategy for Pneumococcal vaccines in the elderly. *Hum Vaccin Immunother*, 12, 336-343.
- NELSON, A. L., ROCHE, A. M., GOULD, J. M., CHIM, K., RATNER, A. J. & WEISER, J. N. 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect. Immun.*, 75, 83-90.
- NG, W. L., KAZMIERCZAK, K. M. & WINKLER, M. E. 2004. Defective cell wall synthesis in *Streptococcus pneumoniae* R6 depleted for the essential PcsB putative murein hydrolase or the VicR (YycF) response regulator. *Mol. Microbiol.*, 53, 1161-1175.
- NG, W. L., ROBERTSON, G. T., KAZMIERCZAK, K. M., ZHAO, J., GILMOUR, R. & WINKLER, M. E. 2003. Constitutive expression of PcsB suppresses the requirement for the essential VicR (YycF) response regulator in *Streptococcus pneumoniae* R6.

- Mol. Microbiol.*, 50, 1647-1663.
- NIKI, E. 2009. Lipid peroxidation: physiological levels and dual biological effects. *Free Radic. Biol. Med.*, 47, 469-484.
- NORMARK, B. H. & NORMARK, S. 2002. Evolution and spread of antibiotic resistance. *J. Intern. Med.*, 252, 91-106.
- O'BRIEN, K. L., WOLFSON, L. J., WATT, J. P., HENKLE, E., DELORIA-KNOLL, M., MCCALL, N., LEE, E., MULHOLLAND, K., LEVINE, O. S. & CHERIAN, T. 2009. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *The Lancet*, 374, 893-902.
- OBARO, S. K., ADEGBOLA, R. A., CHANG, I., BANYA, W. A., JAFFAR, S., MCADAM, K. W. & GREENWOOD, B. M. 2000. Safety and immunogenicity of a nonavalent pneumococcal vaccine conjugated to CRM197 administered simultaneously but in a separate syringe with diphtheria, tetanus and pertussis vaccines in Gambian infants. *Pediatr. Infect. Dis. J.*, 19, 463-469.
- OGUNNIYI, A. D. & PATON, J. C. 2015. Vaccine Potential of Pneumococcal Proteins. 59-78.
- PAIXÃO, L., OLIVEIRA, J., VERÍSSIMO, A., VINGA, S., LOURENÇO, E. C., VENTURA, M. R., KJOS, M., VEENING, J.-W., FERNANDES, V. E. & ANDREW, P. W. 2015. Host glycan sugar-specific pathways in *Streptococcus pneumoniae*: Galactose as a key sugar in colonisation and infection. *PLoS ONE*, 10, e0121042.
- PAN, X., GE, J., LI, M., WU, B., WANG, C., WANG, J., FENG, Y., YIN, Z., ZHENG, F., CHENG, G., SUN, W., JI, H., HU, D., SHI, P., FENG, X., HAO, X., DONG, R., HU, F. & TANG, J. 2009. The orphan response regulator CovR: a globally negative modulator of virulence in *Streptococcus suis* serotype 2. *J. Bacteriol.*, 191, 2601-12.
- PARASHAR, V., AGGARWAL, C., FEDERLE, M. J. & NEIDITCH, M. B. 2015. Rgg protein structure–function and inhibition by cyclic peptide compounds. *PNAS.*, 112, 5177-5182.
- PATERSON, G. K., BLUE, C. E. & MITCHELL, T. J. 2006. Role of two-component systems in the virulence of *Streptococcus pneumoniae*. *J. Med. Microbiol.*, 55, 355-63.
- PATON, J. C., LOCK, R. A. & HANSMAN, D. J. 1983. Effect of immunization with pneumolysin on survival time of mice challenged with *Streptococcus pneumoniae*. *Infect. Immun.*, 40, 548-552.
- PEREIRA, C. S., THOMPSON, J. A. & XAVIER, K. B. 2013. AI-2-mediated signalling in bacteria. *FEMS Microbiol. Rev.*, 37, 156-181.
- PEREZ-PASCUAL, D., GAUDU, P., FLEUCHOT, B., BESSET, C., ROSINSKI-CHUPIN, I., GUILLOT, A., MONNET, V. & GARDAN, R. 2015. RovS and its associated signaling peptide form a cell-to-cell communication system required for *Streptococcus agalactiae* pathogenesis. *MBio*, 6.
- PERICONE, C. D., BAE, D., SHCHEPETOV, M., MCCOOL, T. & WEISER, J. N. 2002. Short-sequence tandem and nontandem DNA repeats and endogenous hydrogen peroxide production contribute to genetic instability of *Streptococcus pneumoniae*. *J. Bacteriol.*, 184, 4392-4399.
- PESAKHOV, S., BENISTY, R., SIKRON, N., COHEN, Z., GOMELSKY, P., KHOZIN-GOLDBERG, I., DAGAN, R. & PORAT, N. 2007. Effect of hydrogen peroxide production and the Fenton reaction on membrane composition of *Streptococcus pneumoniae*. *BBA-Biomembranes*, 1768, 590-597.
- PESSI, G., BLUMER, C. & HAAS, D. 2001. lacZ fusions report gene expression, don't they? *Microbiology*, 147, 1993-1995.
- PETTIGREW, M. M., FENNIE, K. P., YORK, M. P., DANIELS, J. & GHAFFAR, F. 2006. Variation in the presence of neuraminidase genes among *Streptococcus pneumoniae*

- isolates with identical sequence types. *Infect. Immun.*, 74, 3360-3365.
- PHILIPS, B. J., MEGUER, J.-X., REDMAN, J. & BAKER, E. H. 2003. Factors determining the appearance of glucose in upper and lower respiratory tract secretions. *Intensive Care Med*, 29, 2204-2210.
- PODBIELSKI, A. & KREIKEMEYER, B. 2004. Cell density-dependent regulation: basic principles and effects on the virulence of Gram-positive cocci. *International Int. J. Infect. Dis.*, 8, 81-95.
- POTTER, A. J., TRAPPETTI, C. & PATON, J. C. 2012. *Streptococcus pneumoniae* uses glutathione to defend against oxidative stress and metal ion toxicity. *J. Bacteriol.*, 194, 6248-6254.
- PRASHER, D. C., ECKENRODE, V. K., WARD, W. W., PRENDERGAST, F. G. & CORMIER, M. J. 1992. Primary structure of the *Aequorea victoria* green-fluorescent protein. *Gene*, 111, 229-233.
- PULLIAINEN, A. T., HYTONEN, J., HAATAJA, S. & FINNE, J. 2008. Deficiency of the Rgg regulator promotes H<sub>2</sub>O<sub>2</sub> resistance, AhpCF-mediated H<sub>2</sub>O<sub>2</sub> decomposition, and virulence in *Streptococcus pyogenes*. *J Bacteriol*, 190, 3225-35.
- QI, F., CHEN, P. & CAUFIELD, P. W. 1999. Functional analyses of the promoters in the lantibiotic mutacin II biosynthetic locus in *Streptococcus mutans*. *Appl. Environ. Microbiol.*, 65, 652-658.
- RASKO, D. A. & SPERANDIO, V. 2010. Anti-virulence strategies to combat bacteria-mediated disease. *Nat Rev Drug Discov*, 9, 117-128.
- RAWLINSON, E. L., NES, I. F. & SKAUGEN, M. 2005. Identification of the DNA-binding site of the Rgg-like regulator LasX within the lactocin S promoter region. *Microbiology*, 151, 813-23.
- REID, J. L., IYER, V. R., BROWN, P. O. & STRUHL, K. 2000. Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase. *Mol. Cell*, 6, 1297-1307.
- REINERT, R. R., REINERT, S., VAN DER LINDEN, M., CIL, M. Y., AL-LAHHAM, A. & APPELBAUM, P. 2005. Antimicrobial susceptibility of *Streptococcus pneumoniae* in eight European countries from 2001 to 2003. *Antimicrob. Agents Chemother.*, 49, 2903-2913.
- REN, B., LI, J., GENSCHMER, K., HOLLINGSHEAD, S. K. & BRILES, D. E. 2012. The absence of PspA or presence of antibody to PspA facilitates the complement-dependent phagocytosis of pneumococci in vitro. *Clin. Vaccine Immunol.*, 19, 1574-1582.
- REN, B., ROBERT, F., WYRICK, J. J., APARICIO, O., JENNINGS, E. G., SIMON, I., ZEITLINGER, J., SCHREIBER, J., HANNETT, N. & KANIN, E. 2000. Genome-wide location and function of DNA binding proteins. *Science*, 290, 2306-2309.
- RICHARDS, L., FERREIRA, D. M., MIYAJI, E. N., ANDREW, P. W. & KADIOGLU, A. 2010. The immunising effect of pneumococcal nasopharyngeal colonisation; protection against future colonisation and fatal invasive disease. *Immunobiology*, 215, 251-263.
- ROBINSON, K. A., BAUGHMAN, W., ROTHROCK, G., BARRETT, N. L., PASS, M., LEXAU, C., DAMASKE, B., STEFONEK, K., BARNES, B. & PATTERSON, J. 2001. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995-1998: Opportunities for prevention in the conjugate vaccine era. *Jama*, 285, 1729-1735.
- ROCHA-ESTRADA, J., ACEVES-DIEZ, A. E., GUARNEROS, G. & DE LA TORRE, M. 2010. The RNPP family of quorum-sensing proteins in Gram-positive bacteria. *Appl Microbiol Biotechnol*, 87, 913-23.

- RODGERS, J. T., PATEL, P., HENNES, J. L., BOLOGNIA, S. L. & MASCOTTI, D. P. 2000. Use of biotin-labeled nucleic acids for protein purification and agarose-based chemiluminescent electromobility shift assays. *Anal. Biochem.*, 277, 254-259.
- ROSCH, J. W., GAO, G., RIDOUT, G., WANG, Y. D. & TUOMANEN, E. I. 2009. Role of the manganese efflux system *mntE* for signalling and pathogenesis in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 72, 12-25.
- ROSE, M. C. & VOYNOW, J. A. 2006. Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol. Rev.*, 86, 245-278.
- RUBINS, J. B., CHARBONEAU, D., FASCHING, C., BERRY, A. M., PATON, J. C., ALEXANDER, J. E., ANDREW, P. W., MITCHELL, T. J. & JANOFF, E. N. 1996. Distinct roles for pneumolysin's cytotoxic and complement activities in the pathogenesis of pneumococcal pneumonia. *Am. J. Respir. Crit. Care Med.*, 153, 1339-1346.
- RYAN, K. & RAY, C. ed.(2004). Sherris Medical Microbiology. McGraw Hill.
- SAIKI, R. K., GELFAND, D. H., STOFFEL, S., SCHARF, S. J., HIGUCHI, R., HORN, G. T., MULLIS, K. B. & ERLICH, H. A. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Science*, 239, 487-491.
- SAITO, H. & MIURA, K.-I. 1963. Preparation of transforming deoxyribonucleic acid by phenol treatment. *BBA-GEN SUBJECTS*, 72, 619-629.
- SALEH, M., BARTUAL, S. G., ABDULLAH, M. R., JENSCH, I., ASMAT, T. M., PETRUSCHKA, L., PRIBYL, T., GELLERT, M., LILLIG, C. H. & ANTELMANN, H. 2013. Molecular architecture of *Streptococcus pneumoniae* surface thioredoxin - fold lipoproteins crucial for extracellular oxidative stress resistance and maintenance of virulence. *EMBO Mol Med*, 5, 1852-1870.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. 1989. *Molecular cloning: a laboratory manual*, Cold spring harbor laboratory press.
- SANDERS, J. W., LEENHOUTS, K., BURGHOORN, J., BRANDS, J. R., VENEMA, G. & KOK, J. 1998. A chloride - inducible acid resistance mechanism in *Lactococcus lactis* and its regulation. *Mol. Microbiol.*, 27, 299-310.
- SCHACHERN, P. A., TSUPRUN, V., GOETZ, S., CUREOGLU, S., JUHN, S. K., BRILES, D. E., PAPARELLA, M. M. & FERRIERI, P. 2013. Viability and virulence of pneumolysin, pneumococcal surface protein A, and pneumolysin/pneumococcal surface protein A mutants in the ear. *JAMA Otolaryngol Head Neck Surg*, 139, 937-943.
- SCHMITTGEN, T. D. & LIVAK, K. J. 2008. Analyzing real-time PCR data by the comparative CT method. *Nat Protoc*, 3, 1101.
- SEITZ, P. & BLOKESCH, M. 2014. DNA transport across the outer and inner membranes of naturally transformable *Vibrio cholerae* is spatially but not temporally coupled. *MBio*, 5, e01409-14.
- SHAFEEQ, S., AFZAL, M., HENRIQUES-NORMARK, B. & KUIPERS, O. P. 2015. Transcriptional profiling of UlaR-regulated genes in *Streptococcus pneumoniae*. *Genom Data*, 4, 57-59.
- SHAPER, M., HOLLINGSHEAD, S. K., BENJAMIN, W. H. & BRILES, D. E. 2004. PspA protects *Streptococcus pneumoniae* from killing by apolactoferrin, and antibody to PspA enhances killing of pneumococci by apolactoferrin. *Infect. Immun.*, 72, 5031-5040.
- SHAPIRA, S. K., CHOU, J., RICHAUD, F. V. & CASADABAN, M. J. 1983. New versatile plasmid vectors for expression of hybrid proteins coded by a cloned gene fused to lacA gene sequences encoding an enzymatically active carboxy-terminal portion of  $\beta$ -

- galactosidase. *Gene*, 25, 71-82.
- SHAPIRO, J. 1969. Mutations caused by the insertion of genetic material into the galactose operon of *Escherichia coli*. *J Mol Cell Biol*, 40, 93-105.
- SHARMA, V., ICHIKAWA, M. & FREEZE, H. H. 2014. Mannose metabolism: More than meets the eye. *Biochem. Biophys. Res. Commun*, 453, 220-228.
- SHELBURNE, S. A., DAVENPORT, M. T., KEITH, D. B. & MUSSER, J. M. 2008. The role of complex carbohydrate catabolism in the pathogenesis of invasive streptococci. *Trends Microbiol*, 16, 318-25.
- SHI, K., BROWN, C. K., GU, Z. Y., KOZLOWICZ, B. K., DUNNY, G. M., OHLENDORF, D. H. & EARHART, C. A. 2005. Structure of peptide sex pheromone receptor PrgX and PrgX pheromone complexes and regulation of conjugation in *Enterococcus faecalis*. *Proc Natl Acad Sci U S A*, 102, 18596-601.
- SIEHNEL, R., TRAXLER, B., AN, D. D., PARSEK, M. R., SCHAEFER, A. L. & SINGH, P. K. 2010. A unique regulator controls the activation threshold of quorum-regulated genes in *Pseudomonas aeruginosa*. *PNAS*, 107, 7916-7921.
- SIEMIENIUK, R. A., GREGSON, D. B. & GILL, M. J. 2011. The persisting burden of invasive pneumococcal disease in HIV patients: an observational cohort study. *BMC Infect. Dis.*, 11, 1.
- SKAUGEN, M., ANDERSEN, E. L., CHRISTIE, V. H. & NES, I. F. 2002. Identification, characterization, and expression of a second, bicistronic, operon involved in the production of lactocin S in *Lactobacillus sakei* L45. *Appl. Environ. Microbiol.*, 68, 720-727.
- SOLOVYEV, V. & SALAMOV, A. 2011. Automatic annotation of microbial genomes and metagenomic sequences. *Metagenomics and its applications in agriculture, biomedicine and environmental studies*, 61-78.
- SONG, J.-H. 2013. Advances in pneumococcal antibiotic resistance. *Expert review of Lancet Respir Med*, 7, 491-498.
- SPELLERBERG, B., CUNDELL, D. R., SANDROS, J., PEARCE, B. J., IDÄNPÄÄN - HEIKKILÄ, I., ROSENOW, C. & MASURE, H. R. 1996. Pyruvate oxidase, as a determinant of virulence in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 19, 803-813.
- STANDISH, A. J. & MORONA, R. 2015. Capsule Structure, Synthesis, and Regulation. 169-179.
- STEINER, E., SCOTT, J., MINTON, N. P. & WINZER, K. 2012. An agr quorum sensing system that regulates granule formation and sporulation in *Clostridium acetobutylicum*. *Appl. Environ. Microbiol.*, 78, 1113-1122.
- STEINFORT, C., WILSON, R., MITCHELL, T., FELDMAN, C., RUTMAN, A., TODD, H., SYKES, D., WALKER, J., SAUNDERS, K. & ANDREW, P. 1989. Effect of *Streptococcus pneumoniae* on human respiratory epithelium in vitro. *Infect. Immun.*, 57, 2006-2013.
- STENZ, L., FRANCOIS, P., WHITESON, K., WOLZ, C., LINDER, P. & SCHRENZEL, J. 2011. The CodY pleiotropic repressor controls virulence in gram-positive pathogens. *FEMS Immunol. Med. Microbiol.*, 62, 123-139.
- STEWART, G. R., WERNISCH, L., STABLER, R., MANGAN, J. A., HINDS, J., LAING, K. G., YOUNG, D. B. & BUTCHER, P. D. 2002. Dissection of the heat-shock response in *Mycobacterium tuberculosis* using mutants and microarrays. *Microbiology*, 148, 3129-3138.
- STOCK, A. M., ROBINSON, V. L. & GOUDREAU, P. N. 2000. Two-component signal transduction. *Annu. Rev. Biochem.*, 69, 183-215.
- STROEHER, U. H., PATON, A. W., OGUNNIYI, A. D. & PATON, J. C. 2003. Mutation of luxS of *Streptococcus pneumoniae* affects virulence in a mouse model. *Infect.*

- Immun.*, 71, 3206-3212.
- SULAVIK, M., TARDIF, G. & CLEWELL, D. 1992. Identification of a gene, *rgg*, which regulates expression of glucosyltransferase and influences the Spp phenotype of *Streptococcus gordonii* Challis. *J. Bacteriol.*, 174, 3577-3586.
- SUNG, C. K. & MORRISON, D. A. 2005. Two distinct functions of ComW in stabilization and activation of the alternative sigma factor ComX in *Streptococcus pneumoniae*. *J. Bacteriol.*, 187, 3052-3061.
- SWARTZ, M. N. 2004. Cellulitis. *N. Engl. J. Med.*, 350, 904-912.
- SZURMANT, H. 2012. Essential two-component systems of Gram-positive bacteria. Norfolk, UK: Caister Academic Press.
- TALEKAR, S. J., CHOCHUA, S., NELSON, K., KLUGMAN, K. P., QUAVE, C. L. & VIDAL, J. E. 2014. 220D-F2 from *Rubus ulmifolius* kills *Streptococcus pneumoniae* planktonic cells and pneumococcal biofilms. *PLoS ONE*, 9, e97314.
- TATSUNO, I., ISAKA, M., OKADA, R., ZHANG, Y. & HASEGAWA, T. 2014. Relevance of the two-component sensor protein CiaH to acid and oxidative stress responses in *Streptococcus pyogenes*. *BMC Res Notes*, 7, 1.
- TERRA, V. S., HOMER, K. A., RAO, S. G., ANDREW, P. W. & YESILKAYA, H. 2010a. Characterization of novel beta-galactosidase activity that contributes to glycoprotein degradation and virulence in *Streptococcus pneumoniae*. *Infect Immun*, 78, 348-57.
- TERRA, V. S., HOMER, K. A., RAO, S. G., ANDREW, P. W. & YESILKAYA, H. 2010b. Characterization of novel  $\beta$ -galactosidase activity that contributes to glycoprotein degradation and virulence in *Streptococcus pneumoniae*. *Infect. Immun.*, 78, 348-357.
- TERRA, V. S., ZHI, X., KAHYA, H. F., ANDREW, P. W. & YESILKAYA, H. 2016. Pneumococcal 6-phospho- $\beta$ -glucosidase (BglA3) is involved in virulence and nutrient metabolism. *Infect. Immun.*, 84, 286-292.
- TETTELIN, H., CHANCEY, S., MITCHELL, T., DENAPAITE, D., SCHÄHLE, Y., RIEGER, M. & HAKENBECK, R. 2015. Genomics, Genetic Variation, and Regions of Differences. 81-107.
- TETTELIN, H., MASIGNANI, V., CIESLEWICZ, M. J., EISEN, J. A., PETERSON, S., WESSELS, M. R., PAULSEN, I. T., NELSON, K. E., MARGARIT, I. & READ, T. D. 2002. Complete genome sequence and comparative genomic analysis of an emerging human pathogen, serotype V *Streptococcus agalactiae*. *PNAS*, 99, 12391-12396.
- THOMAS, J. G., LITTON, I. & RINDE, H. 2006. Economic impact of biofilms on treatment costs. *Biofilms, Infection and Antimicrobial Therapy*, 21-37.
- THROUP, J. P., KORETKE, K. K., BRYANT, A. P., INGRAHAM, K. A., CHALKER, A. F., GE, Y., MARRA, A., WALLIS, N. G., BROWN, J. R. & HOLMES, D. J. 2000. A genomic analysis of two - component signal transduction in *Streptococcus pneumoniae*. *Mol. Microbiol.*, 35, 566-576.
- TILLEY, S. J., ORLOVA, E. V., GILBERT, R. J., ANDREW, P. W. & SAIBIL, H. R. 2005. Structural basis of pore formation by the bacterial toxin pneumolysin. *Cell*, 121, 247-256.
- TONG, H., BLUE, L., JAMES, M. & DEMARIA, T. 2000. Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chinchilla model. *Infect. Immun.*, 68, 921-924.
- TONG, H. H., LIU, X., CHEN, Y., JAMES, M. & DEMARIA, T. 2002. Effect of neuraminidase on receptor-mediated adherence of *Streptococcus pneumoniae* to chinchilla tracheal epithelium. *Acta Otolaryngol.*, 122, 413-419.
- TOUATI, D. 2000. Iron and oxidative stress in bacteria. *Arch. Biochem. Biophys.*, 373, 1-6.

- TSENG, H.-J., MCEWAN, A. G., PATON, J. C. & JENNINGS, M. P. 2002. Virulence of *Streptococcus pneumoniae*: PsaA mutants are hypersensitive to oxidative stress. *Infect. Immun.*, 70, 1635-1639.
- ULIJASZ, A. T., ANDES, D. R., GLASNER, J. D. & WEISBLUM, B. 2004. Regulation of iron transport in *Streptococcus pneumoniae* by RitR, an orphan response regulator. *J. Bacteriol.*, 186, 8123-8136.
- UTSUMI, R. 2008. *Bacterial signal transduction: networks and drug targets*, Springer Science & Business Media.
- VAN DELDEN, C., KÖHLER, T., BRUNNER-FERBER, F., FRANÇOIS, B., CARLET, J. & PECHÈRE, J.-C. 2012. Azithromycin to prevent *Pseudomonas aeruginosa* ventilator-associated pneumonia by inhibition of quorum sensing: a randomized controlled trial. *Intensive Care Med*, 38, 1118-1125.
- VELDHUIS, N. A., GAETH, A. P., PEARSON, R. B., GABRIEL, K. & CAMAKARIS, J. 2009. The multi-layered regulation of copper translocating P-type ATPases. *Biometals*, 22, 177-190.
- VERNATTER, J. & PIROFSKI, L.-A. 2013. Current concepts in host-microbe interaction leading to pneumococcal pneumonia. *Curr. Opin. Infect. Dis.*, 26, 277.
- VICKERMAN, M., FLANNAGAN, S., JESIONOWSKI, A., BROSSARD, K., CLEWELL, D. & SEDGLEY, C. 2010. A genetic determinant in *Streptococcus gordonii* Challis encodes a peptide with activity similar to that of enterococcal sex pheromone cAM373, which facilitates intergeneric DNA transfer. *J. Bacteriol.*, 192, 2535-2545.
- VIDAL, J. E., LUDEWICK, H. P., KUNKEL, R. M., ZÄHNER, D. & KLUGMAN, K. P. 2011. The LuxS-dependent quorum-sensing system regulates early biofilm formation by *Streptococcus pneumoniae* strain D39. *Infect. Immun.*, 79, 4050-4060.
- WALKER, C. L. F., RUDAN, I., LIU, L., NAIR, H., THEODORATOU, E., BHUTTA, Z. A., O'BRIEN, K. L., CAMPBELL, H. & BLACK, R. E. 2013. Global burden of childhood pneumonia and diarrhoea. *The Lancet*, 381, 1405-1416.
- WATSON, R. O., SMOOT, J. C. & MUSSER, J. M. 2001. Identification of Rgg-Regulated Exoproteins of *Streptococcus pyogenes*. *Infect. Immun.*, 69, 822-831.
- WEISER, J. N., BAE, D., EPINO, H., GORDON, S. B., KAPOOR, M., ZENEWICZ, L. A. & SHCHEPETOV, M. 2001. Changes in availability of oxygen accentuate differences in capsular polysaccharide expression by phenotypic variants and clinical isolates of *Streptococcus pneumoniae*. *Infect. Immun.*, 69, 5430-5439.
- WHITE, C., LEE, J., KAMBE, T., FRITSCHKE, K. & PETRIS, M. J. 2009. A role for the ATP7A copper-transporting ATPase in macrophage bactericidal activity. *Biol. Chem.*, 284, 33949-33956.
- WHOLEY, W.-Y., KOCHAN, T. J., STORCK, D. N. & DAWID, S. 2016. Coordinated bacteriocin expression and competence in *Streptococcus pneumoniae* contributes to genetic adaptation through neighbor predation. *PLoS Pathog.*, 12, e1005413.
- WIDDEL, F. 2007. Theory and measurement of bacterial growth. *Di dalam Grundpraktikum Mikrobiologie*, 4.
- WINKELSTEIN, J. A. 1981. The role of complement in the host's defense against *Streptococcus pneumoniae*. *Rev. Infect. Dis.*, 3, 289-298.
- WINZER, K., HARDIE, K. R. & WILLIAMS, P. 2002. Bacterial cell-to-cell communication: sorry, can't talk now gone to lunch! *Curr. Opin. Microbiol.*, 5, 216-222.
- WIZEMANN, T. M., HEINRICHS, J. H., ADAMOU, J. E., ERWIN, A. L., KUNSCH, C., CHOI, G. H., BARASH, S. C., ROSEN, C. A., MASURE, H. R. & TUOMANEN, E. 2001. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect. Immun.*, 69, 1593-1598.

- WYLLIE, A. L., VAN HENTEN, S. C., DUURLAND, C. L., WEINBERGER, D. M., SANDERS, E. A., LIPSITCH, M. & TRZCIŃSKI, K. 2016. Significance of serotype on interactions between *Streptococcus pneumoniae* and the human host. *Molecular surveillance of pneumococcal carriage in all ages*.
- YATES, E. A., PHILIPP, B., BUCKLEY, C., ATKINSON, S., CHHABRA, S. R., SOCKETT, R. E., GOLDNER, M., DESSAUX, Y., CÁMARA, M. & SMITH, H. 2002. N-acylhomoserine lactones undergo lactonolysis in a pH-, temperature-, and acyl chain length-dependent manner during growth of *Yersinia pseudotuberculosis* and *Pseudomonas aeruginosa*. *Infect. Immun.*, 70, 5635-5646.
- YESILKAYA, H. 1999. *Studies on the role of superoxide dismutase (SOD) in the virulence of Streptococcus pneumoniae and the effects of interferon gamma on sensitivity of phagocytes to the toxin pneumolysin*. Microbiology.
- YESILKAYA, H., ANDISI, V. F., ANDREW, P. W. & BIJLSMA, J. J. 2013. *Streptococcus pneumoniae* and reactive oxygen species: an unusual approach to living with radicals. *Trends Microbiol.*, 21, 187-195.
- YESILKAYA, H., KADIOGLU, A., GINGLES, N., ALEXANDER, J. E., MITCHELL, T. J. & ANDREW, P. W. 2000. Role of manganese-containing superoxide dismutase in oxidative stress and virulence of *Streptococcus pneumoniae*. *Infect. Immun.*, 68, 2819-2826.
- YESILKAYA, H., MANCO, S., KADIOGLU, A., TERRA, V. S. & ANDREW, P. W. 2008a. The ability to utilize mucin affects the regulation of virulence gene expression in *Streptococcus pneumoniae*. *FEMS Microbiol Lett*, 278, 231-5.
- YESILKAYA, H., MANCO, S., KADIOGLU, A., TERRA, V. S. & ANDREW, P. W. 2008b. The ability to utilize mucin affects the regulation of virulence gene expression in *Streptococcus pneumoniae*. *FEMS Microbiol. Lett.*, 278, 231-235.
- YESILKAYA, H., SPISSU, F., CARVALHO, S. M., TERRA, V. S., HOMER, K. A., BENISTY, R., PORAT, N., NEVES, A. R. & ANDREW, P. W. 2009. Pyruvate formate lyase is required for pneumococcal fermentative metabolism and virulence. *Infect. Immun.*, 77, 5418-5427.
- YU, J., BRYANT, A. P., MARRA, A., LONETTO, M. A., INGRAHAM, K. A., CHALKER, A. F., HOLMES, D. J., HOLDEN, D., ROSENBERG, M. & MCDEVITT, D. 2001. Characterization of the *Streptococcus pneumoniae* NADH oxidase that is required for infection. *Microbiology*, 147, 431-438.
- ZAHLTEN, J., KIM, Y.-J., DOEHN, J.-M., PRIBYL, T., HOCKE, A. C., GARCÍA, P., HAMMERSCHMIDT, S., SUTTORP, N., HIPPENSTIEL, S. & HÜBNER, R.-H. 2014. *Streptococcus pneumoniae*-induced oxidative stress in lung epithelial cells depends on pneumococcal autolysis and is reversible by resveratrol. *Int. J. Infect. Dis.*, jiu806.
- ZHANG, S., WILSON, F. & PACE, L. 2006. *Streptococcus pneumoniae*-associated cellulitis in a two-month-old Domestic Shorthair kitten. *Journal of veterinary diagnostic investigation*, 18, 221-224.
- ZHANG, X. & BREMER, H. 1995. Control of the *Escherichia coli* *rrnB* P1 promoter strength by ppGpp. *Biol. Chem.*, 270, 11181-11189.
- ZHENG, F., JI, H., CAO, M., WANG, C., FENG, Y., LI, M., PAN, X., WANG, J., QIN, Y., HU, F. & TANG, J. 2011. Contribution of the Rgg transcription regulator to metabolism and virulence of *Streptococcus suis* serotype 2. *Infect Immun*, 79, 1319-28.
- ZHOU, H., HABER, M., RAY, S., FARLEY, M. M., PANOZZO, C. A. & KLUGMAN, K. P. 2012. Invasive pneumococcal pneumonia and respiratory virus co-infections. *Emerg Infect Dis*, 18, 294-297.

- ZHU, J. & MEKALANOS, J. J. 2003. Quorum sensing-dependent biofilms enhance colonization in *Vibrio cholerae*. *Dev. Cell*, 5, 647-656.
- ZOUHIR, S., PERCHAT, S., NICAISE, M., PEREZ, J., GUIMARAES, B., LERECLUS, D. & NESSLER, S. 2013. Peptide-binding dependent conformational changes regulate the transcriptional activity of the quorum-sensor NprR. *Nucleic Acids Res.*, 41, 7920-7933.
- ZUPANCIC, M. L. & RECORD, M. T. 1998. RNA polymerase-promoter interactions: the comings and goings of RNA polymerase. *J. Bacteriol.*, 180, 3019-3025.