Investigating galactose metabolism in *Streptococcus pneumoniae*

Kimberley Taylor McLean
B. Sc. (Hons)

A thesis submitted in fulfilment of the
Degree of Doctor of Philosophy

Research Centre for Infectious Diseases
School of Biological Sciences
The University of Adelaide
Adelaide, South Australia

August 2022

This thesis is dedicated to

Michelle Carmel (Taylor) McLean

My Best Friend. My Mum.

15.5.1967 – 10.6.2020

Table of Contents

Abbreviations

α	Alpha
A^o	Absorbance
ABC	ATP Binding cassette
ACRF	Australian Cancer Research Foundation
AdcA	Zinc transport system substrate-binding protein
Aga	Alphagalactosidase
AGRF	Australian Genomics Research Facility
AI-2	Autoinducer 2
AIP	Autoinducer signalling peptide
AI-2P	Phosphorylated AI-2
Ala	Alanine
APC	Antigen Presenting Cell
ATP	Adenosine tri-phosphate
β	Beta
BA	Blood agar
BAM	Binary Alignment Map
BBB	Blood Brain Barrier
BgaA	Beta-galactosidase
Bp	Basepair/s
BSA	Bovine serum albumin
^{o}C	Degrees Celsius
CaCl$_2$	Calcium chloride
Cbp	Choline binding protein
CCL2	C-C motif chemokine ligand 2
CcpA	Catabolite control protein A
CCR	Carbon catabolite repression
CDM	Chemically defined medium
CFU	Colony forming units
ChoP	Phosphorylcholine
CO$_2$	Carbon dioxide
com	Competence locus
CpG	5'-C-phosphate-G-3'
CPS	Capsular polysaccharide

cre	Catabolite-responsive elements
CRM197	Non-toxic variant of diphtheria toxin used in pneumococcal conjugate vaccines
CSF	Cerebrospinal fluid
CSP	Competence-stimulating peptide
C+Y	Casein-based, semisynthetic liquid medium
Δ	Delta or deletion of gene
DNA	Deoxyribonucleic acid
DMEM	Dulbecco's Modified Eagle's Medium
DOC	Deoxycholate
ECM	Extracellular matrix
EDTA	Ethylenediaminetetraacetic acid
EndoD	Endo-β-1,4-N-acetylglucosaminidase
Eno	Enolase
Ery	Erythromycin
Etrx	Surface-exposed thioredoxin family lipoprotein
FACS	Fluorescence-associated cell sorting
fba	Fructose-bisphosphate aldolase
FCS	Fetal calf serum
Fe	Iron
Fru	Fructose
FruA	Fructose specific transport system, EIIBC or EIIC component
F6P	Fructose-6-phosphate
Gal	Galactose
Gal-1-P	Galactose-1-phosphate
GAP	Glyceraldehyde-3-phosphate
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase
Gent	Gentamicin
Gki	Glucokinase, house-keeping gene
Glc	Glucose
Glc-1-P	Glucose-1-phosphate
GlpO	Alpha-glycerophosphate oxidase
GM	Geometric mean
GM-CSF	Granulocyte macrophage colony-stimulating factor
h	Hour/s

HCl	Hydrochloric acid
HEPES	N-2-hydroxyethylpipeazine-N'-2-ethanesulfonic acid
His	Histidine
HPr	Histidine-containing phosphocarrier protein
HyL	Hyaluronate lyase
H_2O_2	Hydrogen peroxide
iC3b	Inactivated C3b
IFN	Interferon
IgA	Immunoglobulin A
IgG	Immunoglobulin G
IL	Interleukin
IgM	Immunoglobulin M
IMDM	Iscove's Modified Dulbecco's Medium
I.N.	Intranasal
IPD	Invasive pneumococcal disease
IRF7	Interferon regulatory factor 7
Kan	Kanamycin
kDa	Kilodalton
l	Litre/s
LB	Luria-Bertani
LC-MS	Liquid chromatography mass spectrometry
LRT	Lower respiratory tract
LTA	Lipotechoic Acid
luxS	S-ribosylhomocysteine lyase
LytA	Autolysin
LPXTG	Sortase-anchored surface protein motif
mal	Maltose regulon
Mn	Manganese
MAPK	Mitogen-activated protein kinase
MCP-1	Monocyte chemoattractant protein 1
MHC	Major histocompatibility complex
AMΦ	Alveolar macrophages
iMΦ	Interstitial macrophages
μ	Micro
μg	microgram/s

μl	microlitre/s
μM	micromolar
mg	Milligram/s
ml	Millilitre/s
mM	Millimolar
min	Minutes
iMono	Inflammatory monocytes
rMono	Resident monocytes
Myd88	Myeloid differentiation primary response gene 88
n	Nano
NA	Not available
NaCl	Sodium chloride
NADH	Reduced form of nicotinamide adenine dinucleotide molecules
NanA	Neuraminidase
NaOH	Sodium hydroxide
NCBI	National Centre for Biotechnology Information
NCSP	Non-classical surface protein lacking secretion signals or anchorage motifs
NF-κB	Nuclear factor-κB
NH_4Cl	Ammonium chloride
NK	Natural killer
NLR	Nucleotide-binding oligomerisation domain-like receptor
NLRP3	NLR protein 3
nM	Nanomolar
NOD	Nucleotide domain-containing protein
OD	Optical density
OM	Otitis media
PAF	Platelet-activating factor
PAFR	Platelet-activating factor receptor
PAMP	Pathogen-Associated Molecular Pattern
Pav	Adherence and virulence protein
PBP	Penicillin binding protein
PBS	Phosphate-buffered solution
PCR	Polymerase chain reaction
PCV7	7-valent CPS conjugate vaccine

PCV13	13-valent CPS conjugate vaccine
PECAM1	Platelet endothelial cell adhesion molecule 1
PepO	Putative endopeptidase
PFA	Paraformaldehyde
pgm	Phosphoglucomutase
PiaA	Pneumococcal iron acquisition protein A
Pht	Polyhistidine triad
PIGR	Polymeric immunoglobulin receptor
PitA	Pneumococcal iron transporter protein A
PiuA	Pneumococcal iron uptake protein A
Ply	Pneumolysin
PMA	Phorobol Myristate Acetate
PPSV23	23-valent CPS vaccine
PpmA	Foldase protein
PRR	Pattern Recognition Receptor
PrtA	Cell wall-associated serine protease
PsaA	Pneumococcal surface antigen A
psaBCA	Manganese permease complex
PspA	Pneumococcal surface protein A
PspC	Pneumococcal surface protein C
PsrP	Pneumococcal serine-rich repeat protein, pathogenicity island
PTS	Phosphoenolpyruvate phosphotransferase system
PurR	Purine operon repressor
PyMOL	Molecular visualisation software
QC	Quality control
qRT-PCR	One-step relative quantitative real-time reverse transcription PCR
QS	Quorum sensing
Raf	Raffinose
RBC	Red blood cell
Rrg	Structural subunit proteins of the pilus
RIN	RNA Integrity Number
rlrA	Pathogenicity islet encoding Rrg pilus proteins
RMSD	Root mean square deviation
RNA	Ribonucleic acid
rRNA	Ribosomal RNA

RPMI	Roswell Park Memorial Institute
rpsL1	Small ribosomal subunit with an S12L mutation
SAM	Sequence Alignment Map
SB	Serum broth
SD	Standard deviation
sec	Second/s
SlrA	Streptococcal lipoprotein rotamase A
SNP	Single nucleotide polymorphism
SodA	Mn-cofactored superoxide dismutase
SPD	*Streptococcus pneumoniae* database
Spec	Spectinomycin
SpsA	Pneumococcal surface protein C
SpMsrAB2	*Streptococcus pneumoniae* methionine sulfoxide reductases
SpxB	Pyruvate oxidase
ST	Sequence type
StkP	Serine-threonine kinase
Strep	Streptomycin
StrA	Pneumococcal sortase transpeptidase
StrH	*N*-acetylglucosaminidase
T6P	Tagatose-6-phosphate
TAE	Tris, acetate, EDTA
TCSTS	Two-component signal transduction system
Th	T helper cell
TIR	MyD88/Toll-IL-1 receptor
TLR	Toll-like receptor
TMD	Transmembrane domain
TNF	Tumour necrosis factor
TRAM	TRIF-related adaptor molecule
TRIF	TIR domain-containing adaptor inducing beta interferon
UDP	Uridine diphosphate
URT	Upper respiratory tract
WCV	Whole cell vaccine
Zn	Zinc

Declaration

I certify that this work contains no material which has been accepted for the award of any other degree or diploma in my name in any University or other tertiary institution and, to the best of my knowledge and belief, contains no material previously published or written by another person, except where due reference has been made in the text. In addition, I certify that no part of this work will, in the future, be used in a submission in my name for any other degree or diploma in any University or other tertiary institution without the prior approval of the University of Adelaide and where applicable, any partner institution responsible for the joint award of this degree.

The author acknowledges that copyright of published works contained within this thesis resides with the copyright holder(s) of those works.

I give permission for the digital version of my thesis to be made available on the web, via the University's digital research repository, the Library Search and also through web search engines, unless permission has been granted by the University to restrict access for a period of time.

_____

Kimberley McLean

August 2022

Acknowledgements

It is hard to know how to sum up your appreciation for the monumental amount of people who support you through something like a PhD – but I'm going to give it my best shot!

Firstly, thank you to everyone in the Research Centre for Infectious Diseases and the MLS for your friendship, support, and guidance over the last few years. I'm very grateful to have been able to complete my studies surrounded by such an incredible bunch of people. In particular, thank you to the incredible members of the Paton, Wilson, Morona, Alsharifi Labs and GPN. I'm so grateful to have been able to call Level 4 home for the past five years.

To Issy and Marina, thank you for also deciding to pursue PhDs – I would have been a bit lost without my Honours Buddies! I'm so proud of both of you, you're going to do such incredible things in this life.

To Cathy, I've said it before, and I'll say it again – we would all be lost without you! You are the glue that holds the Paton Lab together and we are all so grateful for you. Thank you for always being such a happy and positive and caring presence, I feel very grateful to have been able to work with you.

To Jacqui, Eve, Carla, Chloe, and Miriam, thank you for always being up for a coffee! The caffeine was always needed.

To Alex, thank you for making me a 'special colleague, with Twitter!'. It has been a pleasure working with you over the last few years.

To Vik and Hannah, the best PhD Buddies a person could ask for. I'm so grateful for your friendship, guidance, and support over the years. You were both so integral to this whole experience and I'm so glad to have been able to share it with you.

To Erin, thank you for being such an incredible friend to me over the years. We've bonded over things that people our age shouldn't have to be able to bond over, but

it has been so comforting to have someone who I can speak so openly with about all the weird and wonderful things in life. I'm grateful to have had the chance to learn from you, you're an exceptional scientist!

To Claudia, grazie mille! I can't even begin to thank you enough for everything – and I mean everything – over the last four and a half years. Your support has been nothing short of extraordinary. Thank you for always being a listening ear and for passively teaching me Italian. I can kind of count to 20 now! Grazie per tutto, significa molto. I can't think of anyone better to have had as a supervisor.

To James, you're simply the best supervisor I could have hoped for. Thank you for taking a chance on the student from down the hall. I have grown in so many ways as a scientist and as a person under your guidance, thank you doesn't even begin to cut it. I'm so grateful to have had the opportunity to learn from you over the last few years. Your support, both personally and professionally, is appreciated in a way I can't properly describe. Mum always said that your Lab is where I needed to be, and I couldn't agree with her more. Thank you.

To my family and friends who have all been along for this crazy ride, thank you. I am so grateful to have had your love and support on this journey. There are a few who warrant a special mention!

To Helen, Tammi, Tess, Maci, Bree, Ava, Tash and Ross, thank you for becoming a second family to us in the last two years. You have no idea how much our coffees and catch ups mean to me. You are a constantly reminding me of how proud Mum would be, I can't express how much I have needed those words. Thank you.

To Nana and Pa, if it weren't for me opening your home to me, I wouldn't be here today. I'm so grateful to have had you both there to support me during my Undergraduate years and beyond. Those years of living with you both will always hold a very special place in my heart. Thank you.

To Kenz, you have heard me talk about study for nearly a decade now – please consider this your honorary PhD! Thank you for always being there, I am so grateful to have had a best friend like you along for this ride.

To Aidan, I am so grateful to have had you by my side in the tail end of this journey. Thank you for making me laugh until I can barely breathe and smile until my cheeks hurt. Thank you for making me always feel incredibly loved, appreciated, and supported. I'm so excited for what the future holds and can't wait for more adventures. I appreciate you so very much and can't believe how lucky I am to share life with you.

To Dad, I'm so proud of how you have stepped into this new role in the last two years. Thank you for instilling in me the importance of hard work and persistence. You've always shown me that nothing is impossible, it just takes some grit and determination. I'm so proud to have you as my Dad.

To my darling siblings, Jarrod, and Danielle. I honestly struggle to find the words to describe how much I love and appreciate you both. You two are my rock, and I know you've got my back just as I have yours. You make me so incredibly proud; I think I am so lucky to have a brother and sister who I adore as much as you two. You have been such a huge part of this journey and I'm so glad I've been able to share it with you both. Let me firstly say, thank you for dealing with a stressed-out sister who sometimes didn't know left from right! You always somehow knew what to say, a skill I know you both learned from Mum.

Jarrod let's not beat around the bush, you're my favourite brother! Thank you for always being the logical voice of reason in my life. You have grown so much in the last few years; you've matured into this strong and level-headed young man. I know Mum is so proud of you! The world is your oyster bud, and I can't wait to see what you do with it. You are so witty, kind, and loving, I admire you so much.

As for you Daniellina, you are absolutely my favourite sister! You continue to amaze me every day. Your ability to persevere amazes me, you remind me so much of Mum it's not even funny! You too, have the world at your fingertips and I can't see what you do with it. You are good and kind and caring; people are all the better for knowing you. I would be nothing without you both. Thank you for believing in me and for loving me unconditionally. I love you both more than you know.

Finally, Mum.

My absolute best friend and my reason. You raised me to believe that I can be anything – anything at all so long as it makes me happy. You have always been my greatest cheerleader in life, backing me at every single step. I have never once doubted your support on every journey I've embarked on. You knew that this journey was right for me long before I did, you always did have a knack for that sort of thing. Thank you for always believing in me, Mum. What I would give to have you here to see this finally finished, you might have even had a celebratory drink with me! You were there to calm every anxious thought, celebrate every successful experiment and achievement and catch every tear. Thank you for raising me to be strong and passionate, because if not for that, I would not have made it through the last two years. These are just two of the tendencies I've learned from having you by my side for the first twenty-three years of my life. I could not be prouder to have had such a courageous and resilient woman by my side to help guide and shape me into who I am today. Even now, you continue to amaze me with just how strong you were. I simply don't have the right words to express my love and gratitude for you, Mum. All I can say, is thank you. I love you. I miss you more than I knew was humanly possible. I know you would be so incredibly proud. You were, and always will be, my sunshine.

This was only ever possible, because of you.

Presentations, Conferences and Funding

Publications	
Thesis Chapter	**Title**
Chapters 3 and 4	**McLean KT**, Tikhomirova A, Brazel EB, Legendre S, Hassbroek G, Minhas V, Paton JC, Trappetti C. (2020) Site-specific mutations of GalR affects galactose metabolism in *Streptococcus pneumoniae.* J Bacteriol. 203:e00180-20

Publications and Authorship Statements are included after the Appendices

Presentations		
Presentation Type	**Title**	**Conference**
Poster Presentation	'A novel link between sugar metabolism and cell-to-cell signalling in *S. pneumoniae*'	International Symposium on Pneumococci and Pneumococcal Disease (ISPPD 11), Melbourne, Victoria, April 2018
Oral Presentation	'Autoinducer 2 and Galactose: How Quorum Sensing Dictates Sugar Metabolism in *Streptococcus pneumoniae*'	European Meeting on the Molecular Biology of the Pneumococcus (EuroPneumo 14), Griefswald, Germany, June 2019
Oral Presentation	'Autoinducer 2 and Galactose: How Quorum Sensing Dictates Sugar Metabolism in *Streptococcus pneumoniae*'	Australian Society for Microbiology Meeting, Adelaide, South Australia, July 2019
Abstract Accepted	'Phosphorylation of GalR is critical for galactose metabolism in *Streptococcus pneumoniae*'	International Symposium on Pneumococci and Pneumococcal Disease (ISPPD 12), Online – Meeting postponed due to COVID-19, 2020

Oral Presentation	'Investigating how sugar metabolism contributes to invasive pneumococcal disease	Australian Society for Microbiology (ASM) Zoom-In Session, September 2021
Oral Presentation	'Investigating galactose metabolism in *Streptococcus pneumoniae*'	Adelaide Protein Group (APG) AwardsFest, September 2021
Oral Presentation	'Investigating galactose metabolism in *Streptococcus pneumoniae*'	Australian Society for Microbiology (ASM) Virtual Bacterial Pathogenesis Conference (vBacPath), September 2021
Poster Presentation	'A novel link between sugar metabolism and cell-to-cell signalling in *S. pneumoniae*'	Ingenuity, October 2021

Funding		
Project Title	**Funding Body**	**Amount**
Understanding how sugar contributes to invasive pneumococcal disease	Channel 7 Children's Research Foundation	$40,000.00 (Commencing 2022)

Abstract

Streptococcus pneumoniae is a formidable human pathogen. Responsible for between 1 and 2 million deaths annually, the pneumococcus makes a major contribution to global morbidity and mortality. In order to cause disease, the pneumococcus must first colonise the human nasopharynx. This colonisation is typically asymptomatic and provides the ideal niche from which the pneumococcus can transmit itself to new hosts. However, in some cases, the pneumococcus will undergo a 'switch' from harmless coloniser to invasive pathogen, transiting to deeper, usually sterile niches in the body and causing invasive disease. A key determinant of successful colonisation of the nasopharynx is the ability to metabolise the different carbon sources that are available. While the pneumococcus typically prefers to metabolise glucose, this carbon source is actively eliminated from the human nasopharynx in an attempt to maintain airway sterility. In the absence of glucose, galactose is the predominant sugar in this niche. Galactose can be metabolised by two pathways in the pneumococcus, the Leloir pathway and the tagatose-6-phosphate pathway.

A study by Trappetti *et al.,* in 2017 was the first to show a link between carbohydrate metabolism and cell-to-cell signalling in the pneumococcus, demonstrating that the quorum sensing molecule Autoinducer 2 (AI-2) is likely phosphorylated during import into the cell. Phosphorylated AI-2 is then proposed to either directly or indirectly phosphorylate GalR, the regulator of the Leloir pathway, driving an increase in galactose metabolism and a subsequent hypervirulent phenotype. They propose that this phosphorylation occurs at the putative phosphorylation sites identified by Sun *et al.,* in 2010: Serine 317, Threonine 319 and Threonine 323.

To better understand the role of these putative phosphorylation sites, a series of amino acid substitution mutants were generated in which each of the sites were replaced, either singly or in combination, with either the non-phosphorylatable residue alanine (A) or the phosphomimetic aspartic acid (D) or glutamic acid (E). While the use of phosphomimetic residues proved somewhat challenging, the use of non-phosphorylatable alanine residues revealed that each of the three putative phosphorylation sites are required for growth in galactose, successful activation of the Leloir pathway and disease progression in a murine model of infection.

What became clear during this study was that despite having two functional pathways encoded for galactose metabolism, there was an inability for one pathway to rescue the other during times of metabolic distress. This indicated that these pathways may not be as discreet as once thought. To further investigate this potential interplay, a series of mutants were generated, deleting key genes from either the Leloir or the T6P pathways. This approach revealed that deleting genes from either pathway resulted in an inability to metabolise galactose, as well as transcriptional changes indicating that there is indeed interplay between these two pathways, with GalR possibly playing a key role as the central regulator of both pathways. Additionally, we found that there is differential accumulation of metabolites intracellularly as a result of these mutations, which may hold the key to deciphering exactly how these two pathways are linked.

Finally, using dual *in vivo* RNA sequencing, we have revealed that GalR and its putative phosphorylation sites play an important role in virulence, leading to skewing of the immune response during infection.

Collectively, the findings of this thesis have significantly advanced our understanding of pneumococcal galactose metabolism, particularly in terms of its regulation via GalR. Additionally, we have shed light on the interplay between the Leloir and T6P pathways, showing for the first time that there is a definitive link that requires both pathways to be present and functional in order to survive in the presence of galactose, much like what is found in the human nasopharynx. We have also shown that the putative GalR phosphorylation sites play a key role in pneumococcal galactose metabolism, pneumococcal virulence and the host response to infection. This project provides the foundation for further investigation into the regulation of pneumococcal galactose metabolism, and the wide-reaching impacts this pathway has on pneumococcal virulence and disease.

Chapter 1: Introduction

1.1. *Streptococcus pneumoniae*

Streptococcus pneumoniae (the pneumococcus) is a Gram-positive bacterium that continues to be associated with significant global morbidity and mortality. *S. pneumoniae* is able to asymptomatically colonise the human nasopharynx, from where it can transit to new, uninfected hosts, to ensure its survival as a species. From this niche, however, the pneumococcus can also undergo a switch from harmless coloniser to invasive pathogen, by translocating to deeper, usually sterile sites in the body. When this occurs, localised and invasive disease can arise, ranging from chronic rhinosinusitis and otitis media through to pneumonia, meningitis, bacteraemia, and sepsis. Despite decades of research and development of effective vaccination and antibiotic treatment regimes, the pneumococcus continues to have significant impacts on healthcare.

1.2. History of *Streptococcus pneumoniae*

The pneumococcus was first described in 1881 independently by Louis Pasteur in France (Pasteur, 1881) and George M. Sternberg in the United States of America (Sternberg, 1881). They observed lancet shaped coccoid bacteria present in human saliva. While this bacterium had previously been reported in the literature (Klebs, 1875, Eberth, 1880), they were the first to demonstrate the pathogenic potential of the pneumococcus upon injection into rabbits. Pasteur used the saliva of a child who had succumbed to rabies while Sternberg used his own, after which they independently identified diplococci present within the blood of the infected rabbits. Pasteur termed the bacterium *Microbe speticemique du salive* (Pasteur, 1881) while Sternberg named it *Microbe pasteuri* (Sternberg, 1885). Also noted at this time by Pasteur was the presence of what is now known as the polysaccharide capsule in the form of material surrounding each cell (Austrian, 1981). By 1886, the pathogen

had a strong association with cases of pulmonary pneumonia and as such was then referred to as the *Pneumococcus*. In 1920, the name was again changed to *Diplococcus pneumoniae*, paying homage to the diplococci that were commonly seen as the causative agent of pulmonary pneumonia (Winslow, 1920). Finally, in 1974 the name was changed to the now known *Streptococcus pneumoniae*, reflecting the presence of cocci in chains when grown in liquid media (Deibel, 1974). The discovery of *S. pneumoniae* has had significant impacts on the broader scientific community, being the pathogen in which many key discoveries have been made. Gram staining was amongst the first of these discoveries, showing for the first time that there were indeed differences in the structural composition of the outermost layers of bacteria (Gram, 1884). They demonstrated that while some bacteria were able to retain Crystal Violet stain, others were not, underpinning the finding that all bacteria are either Gram-positive or Gram-negative (Gram, 1884).

The Quellung Reaction was another such discovery, allowing for the recognition of different pneumococcal capsular serotypes. In the early 20th Century, Neufeld showed that the addition of antiserum to a suspension of *S. pneumoniae* would result in serotype-specific macroscopic agglutination and swelling of the pneumococcal capsule (Neufeld, 1902). This technique, termed the Quellung Reaction, allowed for the identification of specific pneumococcal serotypes, and was able to be applied to other encapsulated bacteria such as *Klebsiella, Neisseria* and *Haemophilus* spp. In the century that followed, extensive research has resulted in the recognition of at least 100 structurally distinct capsular polysaccharide serotypes of *S. pneumoniae* (Ganaie et al., 2020).

Arguably amongst the most important scientific discoveries of the 20th century were the phenomenon of capsular transformation and the identification of DNA as the molecule that conveys genetic information. Stryker first reported in 1916 that virulent

pneumococci became less virulent when cultured in broth containing homologous immune sera (Stryker, 1916). They also saw that these strains produced less capsule, were more readily ingested by phagocytes, and displayed different antigenic properties to their virulent counterparts. Griffith expanded on these findings (Griffith, 1923), adapting the terminology of Arkwright (Arkwright, 1921) determining 'Smooth' (S) colonies as those with a 'lustrous and mucoid' appearance attributed to the presence of a polysaccharide capsule, much like what was observed in virulent pneumococci. Unlike the 'S' colonies, 'Rough' (R) colonies didn't possess capsule, were avirulent in mice and didn't lead to the production of antisera (Griffith, 1928). Griffith was able to show that some 'R' forms could revert to their original 'S' phenotype *in vivo*, while others could not. Even some 'R' colonies that didn't spontaneously revert could be transformed back to their 'S' phenotype through a procedure in which mice were injected with heat-killed 'S' pneumococci along with the non-revertible 'R' strain. He found that the 'R' form could not only revert to its original capsular type, but could also acquire the capsular type of the heat-killed 'S' pneumococci. These findings were quickly verified in following years by other groups (Dawson, 1928, Neufeld, 1928, Dawson, 1931a, Dawson, 1931b, Sia, 1931), thus describing the first events of 'natural transformation' occurring within the pneumococcus. Alloway further expanded on these findings, showing that the events of transformation could be replicated *in vitro*, using extracts of lysed pneumococci (Alloway, 1932, Alloway, 1933). A landmark study by Oswald Avery and colleagues showed that DNA was in fact the molecule that carried hereditary genetic information (Avery, 1944). They did so by treating lysates from a 'S' pneumococcal strain with DNase prior to completing a capsular transformation reaction. They saw that this DNase treatment negated the ability for the reaction to proceed, thereby confirming that DNA was indeed genetic material. These

discoveries in particular were deemed critically important to both the scientific and medical communities, and 'initiated the era of molecular biology' (Watson et al., 1993).

1.3. Global Disease Burden

Globally, lower respiratory tract infections (LRTIs) are a significant cause of both morbidity and mortality (Feldman and Anderson, 2020). Previous studies have shown that LRTIs accounted for greater than 2.3 million deaths in 2016. Of these fatalities, more than 650,000 occurred in children under the age of 5, while nearly 1.1 million were in people over the age of 70 (G. B. D. Lower Respiratory Infections Collaborators, 2018). *S. pneumoniae* was found to be the primary cause of morbidity and mortality among these LRTIs, resulting in more deaths than *Haemophilus influenzae* type b, Influenza and Respiratory Syncytial Virus combined (G. B. D. Lower Respiratory Infections Collaborators, 2018).

As evidenced above, pneumococcal infections most commonly occur in children under the age of five, the immunocompromised and the elderly. Once established in the nasopharynx, *S. pneumoniae* must make a career choice, either remaining asymptomatically within the nasopharynx or invading deeper, typically sterile tissues such as the lungs, blood, or cerebrospinal fluid (CSF) to cause both localised and invasive disease (**Fig. 1.1.**) (Weiser et al., 2018). Localised spread typically results in conditions such as sinusitis and otitis media (OM), of which the pneumococcus is a primary causative agent (Brook, 2013, Danishyar, 2019). It is estimated that up to 80% of children will experience at least one bout of OM in their lifetime (Danishyar, 2019). Rates of OM are particularly high in Australian Aboriginal children, up to 10 times higher than the world average (World Health Organization, 1998). Invasive
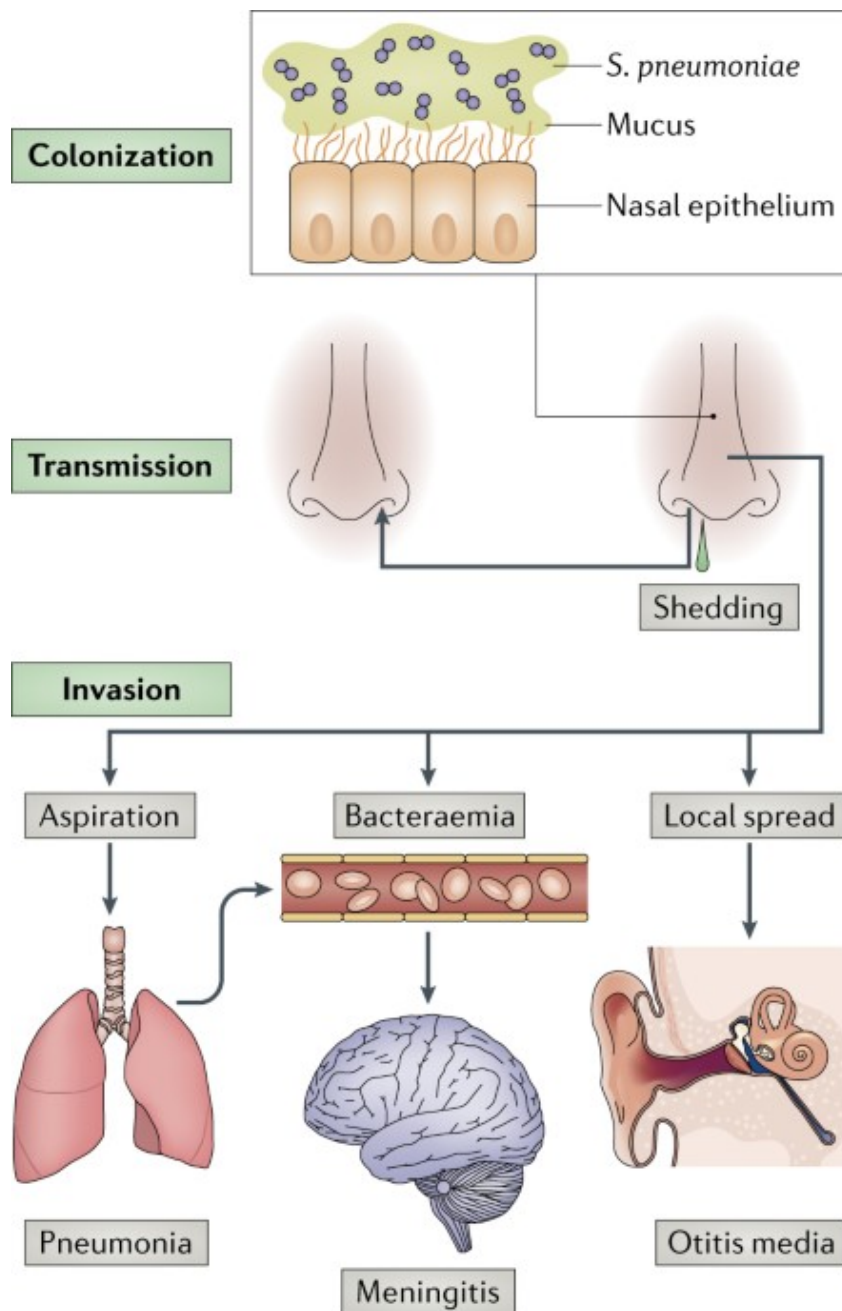
Figure 1.1. Overview of pneumococcal colonisation and invasive disease

Streptococcus pneumoniae must firstly colonise the epithelial mucosa of the nasopharynx. Colonisation is a prerequisite for both transmission to new hosts, along with invasive disease in the index host. *S. pneumoniae* can be shed in the nasal secretions of a carrier, thereby facilitating transmission. Dissemination to other niches, whether by aspiration, bacteraemia, or local spread, can result in invasive disease such as pneumonia, bacteraemia and meningitis, or otitis media. Reproduced from Weiser *et al.,* (2018).

pneumococcal diseases (IPD) take the form of pneumonia, meningitis, bacteraemia, and sepsis. Of the 120 million reported cases of pneumonia in 2010 – 2011, 14 million progressed to severe disease (Walker et al., 2013). Similar to what is observed in the case of OM, rates of IPD are also reported to be higher in Indigenous populations than non-Indigenous populations (Forrest et al., 2000, Australian Institute of Health and Welfare, 2018).

As for many infectious diseases, there are certain risk factors that increase the likelihood of disease. In the case of pneumococcal carriage, risk factors include malnutrition, poverty, poor access to healthcare, and time spent in areas with high levels of crowding such as childcare centres (Lynch and Zhanel, 2010). In the adult population, lifestyle factors such as smoking, along with asthma, immunosuppression, age, and prior history of respiratory infection can all increase the risk of carriage (Lynch and Zhanel, 2010). Alongside these factors, living in developed vs. developing countries can have great impact on the risk of pneumococcal infection. In Australia, the fatality rate as a result of IPD remains at less than 5% in both adults over the age of 65 and children under the age of 5 (Australian Institute of Health and Welfare, 2018). Conversely, the fatality rate for IPD in children less than 5 years of age in developing countries is 10 – 40%. This increase may be attributed to the lack of access to quality health care, malnutrition, and limited availability of treatment options (Bogaert et al., 2004, Harboe et al., 2009).

Co-infection with other upper respiratory tract pathogens has also been shown to increase the risk of pneumococcal infection. Influenza in particular has been shown to enhance pneumococcal infection, leading to increased morbidity and mortality (Siegel et al., 2014, Sender et al., 2021). This phenomenon was observed most dramatically during the 1918-19 Spanish Influenza Pandemic, during which a

significant proportion of fatalities were attributed to secondary bacterial infection, most commonly by *S. pneumoniae* (Morens et al., 2008). Alongside these environmental and socioeconomic predisposing factors, the pneumococcus also possesses a range of virulence factors that contribute to its success as a pathogen.

1.4. Pneumococcal Virulence Factors

1.4.1. Capsular Polysaccharide

The Capsular Polysaccharide (CPS) is the predominant pneumococcal surface structure and acts as a major virulence determinant (Paton and Trappetti, 2019). There are now 100 known pneumococcal serotypes, each with their own propensity to cause localised or invasive disease (Ganaie et al., 2020). First described by Louis Pasteur in the 1880s, the CPS has long since been a subject of investigation. It was first isolated by Dochez and Avery in 1917 (Dochez and Avery, 1917) and determined to be composed of polysaccharide in 1925 (Avery and Heidelberger, 1925). The CPS forms the outermost layer of encapsulated pneumococci as shown in **Figure 1.2.**, and ranges from 200 – 400 nm in thickness (Skov Sorensen et al., 1988). Of the 100 identified serologically and structurally distinct pneumococcal serotypes (Geno et al., 2015, Croucher et al., 2018), the majority have their CPS covalently bound to the outer surface of the cell wall peptidoglycan (Sørenson, 1990). The simplest form of CPS are linear polymers composed of repeats of two or more monosaccharides, while more complicated structures involve branched polysaccharides with repeat units composed of up to six monosaccharides, plus additional side chains (Geno et al., 2015). CPS contributes to pneumococcal pathogenesis in a number of ways. It can inhibit both the classical and alternative complement pathways, acting as a physical barrier that prevents the binding of immunoglobulins, C-reactive protein and complement components to deeper
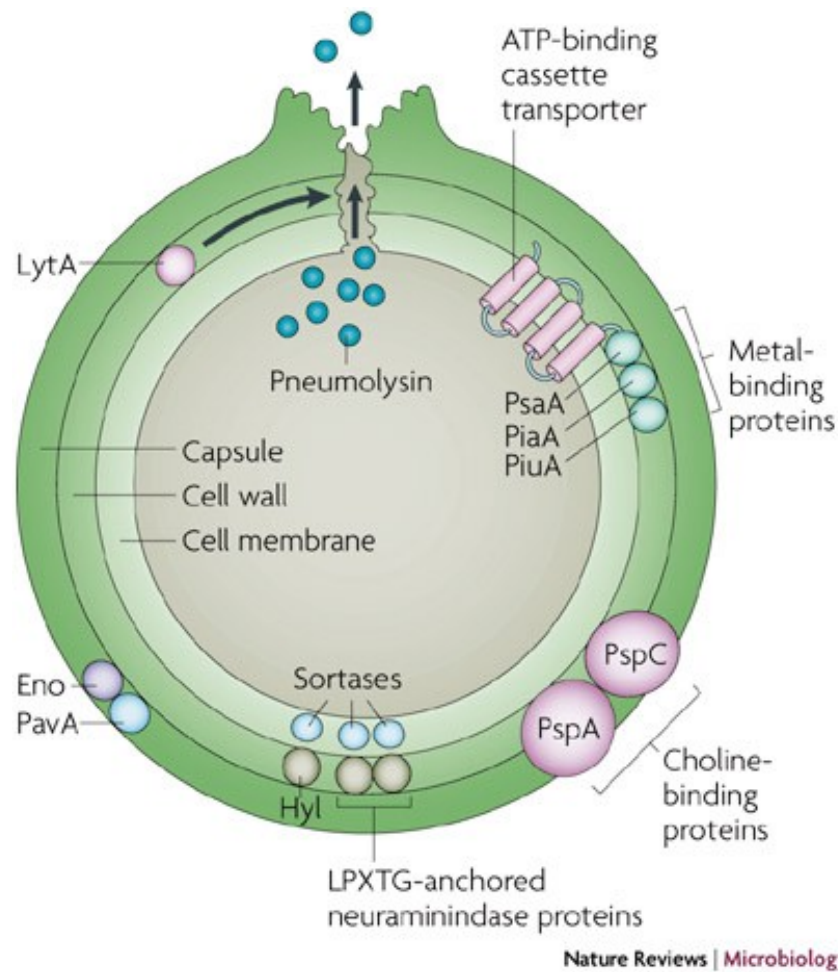
Figure 1.2. Pneumococcal structure and virulence factors
The pneumococcus has multiple virulence mechanisms that allow it to achieve both colonisation and invasive disease. The most important of these include the cell wall, the capsule, choline-binding proteins (PspA, PspC), metal binding proteins (PsaA [manganese import], PiaA, PiuA [iron import and uptake]), LPXTG-anchored neuraminidase proteins, pneumolysin, autolysin A (LytA), enolase (Eno) and pneumococcal adhesion and virulence protein A (PavA). Reproduced from Kadioglu et al., (2008).

surface structures. The CPS is responsible for reducing opsonisation with C3b/iC3b components of the complement pathway, blocking the interaction between bound C3b/iC3b or Fc regions of immunoglobulins with their cognate receptors on phagocytic cells (Abeyta et al., 2003, Hyams et al., 2010). Most CPS serotypes are highly charged at physiological pH, which is thought to aid in colonisation of the nasopharynx (Nelson et al., 2007). This charge aids in avoiding entrapment in nasal mucous and facilitates attachment of pneumococci to the nasopharyngeal epithelium by repelling sialic acid-rich mucopolysaccharides present in this niche. CPS may also aid in inhibiting entrapment of pneumococci within neutrophils (Wartha et al., 2007) and preventing the recognition of TLR ligands present on the surface of bacteria, thereby aiding in the progression from harmless coloniser to invasive pathogen (de Vos et al., 2015).

1.4.2. Virulence Proteins

The pneumococcus possesses a vast array of factors that aid in its virulence. This has been an important area of research over the past 30 years, as the roles of key proteins in virulence have become progressively clearer. While there are dozens of proteins that have been implicated in virulence (**Table 1.1**), there are a few that have also been identified as promising vaccine candidates and therefore have been selected for further discussion below.

1.4.2.1. Pneumolysin

There are many studies showing that pneumolysin (Ply) is a wide-ranging, potent virulence factor. The amino acid sequence of Ply is highly conserved and present in almost all pneumococcal isolates, with only few variants described (Lock et al., 1996, Kirkham et al., 2006). Ply is a 52 kDa protein that oligomerises in the membrane of cells, forming a large ring-shaped pore that traverses the membrane. The pore is composed of approximately 40 monomers and is 260 Å in diameter

Table 1.1. Overview of key pneumococcal virulence factors, adapted from Weiser *et al* 2018.

Virulence Factor	Description	Function
CPS	Capsular polysaccharide, major surface antigen divided into 100 structurally distinct serotypes	Prevents entrapment by mucous, inhibits opsonophagocytosis
ChoP on teichoic acid	PAFR ligand	Binds PAFR on surface of endothelial and epithelial cells, facilitate adherence and invasion
Lipopeptides, lipoteichoic acid and peptidoglycan fragments	Pathogen-Associated Molecular Patterns (PAMPs)	Promote inflammation
Ply	Pneumolysin, pore-forming toxin, TLR4 ligand	Cytotoxic and pro-apoptotic for many host cells, activates classical complement pathway, pro-inflammatory
PspA	Choline Binding Protein	Limits C3 deposition on pneumococcal cell surface, protects against lactoferrin
CbpA/PspC/SpsA	Choline Binding Protein	Binds C3, factor H and sIgA, limits C3b deposition, facilitates adherence and invasion of respiratory epithelium and blood-brain barrier
LytA	Choline Binding Protein, Autolysin	Digests cell wall, releases pro-inflammatory cell wall fragments and Ply, mediates capsule shedding during invasion

CbpD	Choline Binding Protein, Murein hydrolase	Mediates fratricide and release of extracellular DNA, promotes biofilm formation
CbpE/Pce	Choline Binding Protein, Phosphorylcholine esterase	Decreases neutrophil activity by inactivation of host PAF, binds plasminogen
CbpG	Choline Binding Protein serine protease	Implicated in mucosal and invasive disease, cell-attached form promotes adherence, extracellular form degrades fibronectin
CbpL	Choline Binding Protein	Binds collagen, elastin and C-reactive protein, inhibits phagocytosis promoting dissemination from nasopharynx to lungs and blood
NanA	Neuraminidase A, LPXTG	Unmasks receptors for adhesins, cleaves terminal sialic acid from host glycoconjugates, triggers TGF-β signalling to facilitate endothelial invasion
BgaA	β-galactosidase, LPXTG	Cleaves sugars from host glycoconjugates in sequential manner
StrH	β-*N*-acetylglucosaminidase, LPXTG	Cleaves sugars from host glycoconjugates in sequential manner
EndoD	Endo-*N*-acetylglucosaminidase	Cleaves sugars from host glycoconjugates in sequential manner
Hyl	Hyaluronate lyase, LPXTG	Degrades ECM and facilitates tissue penetration

PrtA	Cell wall-associated serine protease, LPXTG	Cleaves lactoferrin, possible adhesin
ZmpA/IgA1	Zinc metalloprotease, LPXTG	Cleaves human IgA$_1$
ZmpB	Zinc metalloprotease, LPXTG	Possible adhesin
ZmpC	Zinc metalloprotease, LPXTG	Cleaves human matrix metalloprotease 9
PepO	Endopeptidase	Facilitates adherence and invasion, binds fibronectin and plasminogen, binds C1q to inhibit classical complement pathway
PsrP	Large *O*-glycosylated serine-rich repeat protein, LPXTG	Adhesion, binds to lung cells via keratin 10, mediates biofilm formation and bacterial aggregation in the lung
RgrA, RgrB and RgrC	LPXTG proteins, structural components of pilus 1, encoded by *rlrA* pathogenicity islet	Adhesins, bind to range of glycans, facilitate biofilm formation and colonisation, RgrA binds PIGR and PECAM1 on endothelium of the blood-brain barrier promoting invasion
PsaA	Lipoprotein, solute-binding component of Mn-specific ABC transporter	Mn uptake within host, essential for resistance to oxidative stress *in vivo*
AcdA and AdcAII	Lipoproteins, solute-binding components of Zn-specific ABC transporter	Zn acquisition *in vivo*
PiuA, PiaA and PitA	Lipoproteins, solute-binding components of iron-specific ABC transporters	Fe acquisition *in vivo*

SlrA and PpmA	Lipoproteins, peptidyl-prolyl isomerases	Contribute to nasopharyngeal colonisation
PhtA, PhtB, PhtB and PhtE	Family of surface proteins with unusual His-triad motifs	Putative adhesins, facilitate Zn acquisition *in vivo* along with AdcAII, may reduce C3 deposition on pneumococcal surface by binding factor H
PavA and PavB	Fibronectin-binding proteins, NCSP	Adhere to host surfaces, important during sepsis and meningitis
Eno	Enolase, NCSP	Binds and activates plasminogen, facilitates tissue invasion
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase, NCSP	Binds and activates plasminogen, facilitates tissue invasion
SpxB	Pyruvate oxidase	Generates H_2O_2
GlpO	α-Glycerophosphate oxidase	Generates H_2O_2
SodA	Mn-dependent superoxide dismutase	Resistance to oxidative stress
Etrx1 and Etrx2	Surface-exposed thioredoxin-family lipoproteins	Resistance to oxidative stress
SpMsrAB2	Methionine sulfoxide reductase	Redox partner of Etrx1 and Etrx2

(Tilley et al., 2005). Pore-formation leads to lysis of a wide range of eukaryotic cell types, but at sub-lytic concentrations, Ply is associated with a plethora of cell-modulatory activities. These include inhibition of ciliary beating on the respiratory epithelium and brain ependyma, inhibition of phagocyte respiratory burst and induction of cytokine synthesis and subsequent T-cell activation and chemotaxis (Paton and Ferrante, 1983, Hirst et al., 2004, Kadioglu et al., 2004). Ply can also activate the classical complement pathway in the absence of antibody, reducing serum opsonic activity (Paton et al., 1984) a property that is independent of its pore-forming activity (Mitchell et al., 1991).

Mutagenesis of the *ply* gene has been shown to significantly reduce the virulence of *S. pneumoniae* in mice, providing direct *in vivo* evidence for a role in pathogenesis (Berry et al., 1989b). Both the cell-modulatory and complement-activation attributes of Ply have been shown to contribute to virulence *in vivo* (Rubins et al., 1996, Alexander et al., 1998, Jounblat et al., 2003). Further indirect evidence for a role in pathogenesis stems from the fact that immunisation of mice with native pneumolysin (Paton et al., 1983) or non-toxic recombinant Ply derivatives (Paton et al., 1991, Alexander et al., 1994) confers significant protection against multiple serotypes of *S. pneumoniae*. A Ply derivative has also been shown to be an effective carrier for type 19F CPS in an experimental conjugate vaccine (Paton et al., 1991).

1.4.2.2. Cell Surface Proteins

Cell-surface proteins are strongly considered vaccine targets as they have the potential to stimulate the production of opsonic antibodies if they are sufficiently exposed on the cell surface, contributing to clearance of pneumococcal infection. There are three major classes of pneumococcal cell-surface proteins that have been identified: choline-binding proteins, lipoproteins and proteins covalently anchored to the cell wall via a LPXTG sortase motif (Kadioglu et al., 2008).

1.4.2.2.1. Choline-Binding Proteins

The pneumococcus expresses phosphorylcholine (ChoP) moeties as a component of cell-wall teichoic acids and membrane-bound lipoteichoic acids which anchor choline-binding proteins to the cell wall. The majority of these proteins have 20 amino acid repeat sequences that facilitate attachment of the proteins to the cell surface via ChoP, with the N-terminal sequences varying widely (Jedrzejas, 2001). The pneumococcus encodes between 10 and 15 choline-binding proteins (Bergmann and Hammerschmidt, 2006), including pneumococcal surface proteins A and C (PspA and PspC) and autolysin (LytA).

PspA contains three structural domains, with an N-terminal region composed of repeated α-helices that protrude from the surface of pneumococci (Jedrzejas et al., 2001). Between the N- and C- termini is a proline-rich region of between 60 and 8 amino acids that likely contributes to flexibility of the protein. The C-terminus typically houses the choline-binding region, anchoring the protein to the cell, although the choline-binding region can be found at the N-termini in some proteins (Jedrzejas et al., 2001). PspA interferes with the fixation of complement on the cell surface, thereby inhibiting opsonization (Kadioglu et al., 2008). It is also a lactoferrin binding protein, which is thought to protect pneumococci from the bactericidal activity of apolactoferrin (Shaper et al., 2004). While the contribution of PspA to these processes has been clearly demonstrated, the role it plays *in vivo* has been harder to elucidate. Studies of both serotype 3 (McDaniel et al., 1987, Ren et al., 2004) and serotype 4 (Hava and Camilli, 2002) pneumococci have shown that PspA is required for growth *in vivo*, but the same was not found in a serotype 2 strain (Berry and Paton, 2000, Abeyta et al., 2003, Orihuela et al., 2004).

PspC is a multifunctional cell surface protein that is known by several names (also choline binding protein A [CbpA] and SpsA), each reflecting different aspects of

protein function (Kadioglu et al., 2008). It has been shown to play a role in adherence *in vitro*, with a *pspC* knockout mutant having a decreased ability to bind to both epithelial cells and sialic acid, and showing reduced nasopharyngeal colonisation compared to wildtype (Rosenow et al., 1997). PspC is also able to bind to the polymeric immunoglobulin receptor that transports secretory IgA, hence the alternative name SpsA (secretory pneumococcal surface protein A) (Hammerschmidt et al., 2000, Zhang et al., 2000). This binding may be one of the first stages of translocation across the respiratory epithelium, and is consistent with a *pspC* knockout mutant being less virulent in a murine pneumonia model of infection (Jounblat et al., 2003). PspC can also bind to Factor H (Janulczyk et al., 2000, Dave et al., 2004), which has been shown to prevent formation of C3b via the alternative complement pathway, thereby inhibiting opsonization and complement-mediated clearance (Quin et al., 2005).

LytA is an amidase that cleaves the *N*-acetylmuramoyl-$_L$-alanine bond in pneumococcal peptidoglycan (Howard and Gooder, 1974). The action of this enzyme results in the cell lysis observed when growing pneumococci in batch culture, as well as being implicated in cell-wall biosynthesis and turnover. There is a strong association between LytA and virulence, with LytA mutants showing decreased virulence in both pneumonia and bacteraemia murine models of infection (Berry et al., 1989a, Canvin et al., 1995, Berry and Paton, 2000, Orihuela et al., 2004). It is thought that this contribution to virulence may be mediated by the LytA-dependent release of pneumolysin, along with release of inflammatory peptidoglycan and teichoic acid fragments from lysed cells.

1.4.2.2.2. Cell-Surface Lipoproteins

There are 45 pneumococcal cell-surface lipoproteins that have been reported in the literature (Bergmann and Hammerschmidt, 2006). Some of the most important of

these in terms of novel vaccine development are those required for the acquisition of divalent metal ions from the host environment. Of particular importance is PsaA, a component of the pneumococcal manganese uptake machinery, along with PiaA and PiuA, both of which are required for the acquisition and uptake of iron (Kadioglu et al., 2008).

PsaA was originally proposed to be a pneumococcal adhesin due to its degree of sequence similarity to putative adhesins from other streptococci species (Sampson et al., 1994), a theory that was supported in the literature with reports of *psaA* mutants exhibiting decreased binding to mammalian cells *in vitro* (Berry and Paton, 1996, Briles et al., 2000). Further studies however confirmed that PsaA is the divalent metal-ion-binding lipoprotein component of the ATP-binding cassette (ABC) transport system used for the uptake of manganese (Dintilhac et al., 1997, McAllister et al., 2004). The effect that PsaA has on adherence is likely a result of the pleiotropic impact of a deficiency in manganese transport on expression of other genes, including those for actual adhesins. This is consistent with previous structural studies indicating that PsaA is unlikely to protrude beyond the cell wall (Lawrence et al., 1998, Jedrzejas, 2001). It has also been shown that when the other genes within the *psaBCA* operon are mutated, there are comparable impacts on adherence observed, even though they encode components of the ABC transporter that are embedded in the cell membrane or are on the cytoplasmic face (Johnston et al., 2004). Importantly, mutagenesis studies of *psaA* have shown a role for manganese uptake in resistance of *S. pneumoniae* to oxidative stress. Hydrogen peroxide is a by-product of pneumococcal aerobic metabolism and the host also generates reactive oxygen species during the innate immune response, both of which expose the bacterium to oxidative stress. This role in managing oxidative

stress is a likely explanation for the avirulence of *psaA* mutants in murine models of both colonisation and invasive disease (Tseng et al., 2002, McAllister et al., 2004). There are three operons that are implicated in iron uptake in the pneumococcus, *pia, piu* and *pit* (Brown et al., 2001a, Brown et al., 2002). Each operon encodes a metal binding protein, a membrane permease and an ATPase; however, the *pia* and *piu* systems have been shown to be particularly important (Brown et al., 2002). PiaA and PiaU comprise the lipoprotein metal-binding components of these systems. The role of these proteins in virulence has been established, with immunisation with these proteins deemed protective in past studies (Brown et al., 2001b). While there is some redundancy in the three identified iron uptake systems, only a *piu-pia* double mutant strain showed significantly attenuated growth in iron-deficient media. Nevertheless, single *pia* and *piu* mutants exhibited decreased virulence in murine models of pneumonia and bacteraemia, although a double *piu-pia* mutant was attenuated to a significantly greater extent (Brown et al., 2002).

1.4.2.2.3. LPXTG Motif Proteins

The covalent anchorage of certain surface proteins to the peptidoglycan of Gram-positive bacteria is mediated by sortase transpeptidases that recognise the amino-acid sequence LPXTG within surface proteins, where the 'X' can be any amino acid. While some pneumococci encode a single sortase, others encode multiple predicted sortase genes (Bergmann and Hammerschmidt, 2006). In strains with multiple sortase genes, sortase A (StrA) is believed to be responsible for anchoring most LPXTG-containing proteins, while other sortase proteins mediate anchorage of a subset of surface proteins, possibly in response to specific environmental cues (Hava and Camilli, 2002). There are as many as 20 pneumococcal proteins anchored via an LPXTG motif, including neuraminidases (Bergmann and Hammerschmidt, 2006).

Neuraminidases, also referred to as sialidases, cleave the terminal sialic acid residues present on glycoproteins, glycolipids and cell-surface oligosaccharides. Studies have shown that neuraminidases can remove the sialic acid residues from key soluble proteins such as lactoferrin and IgA_2 (King et al., 2004). There are at least three neuraminidase genes present in available pneumococcal genomes: *nanA, nanB* and *nanC.* All pneumococcal genomes have *nanA*, almost all have *nanB*, and approximately 50% have *nanC* (Pettigrew et al., 2006). Neuraminidases can be secreted from the pneumococcal cell, but only NanA contains an LPXTG motif, and NanA and NanB have distinct pH optima (Berry et al., 1996), suggesting these enzymes may have distinct functions *in vivo*. Experiments utilising loss-of-function mutants in murine models of acute pneumonia have shown that both NanA and NanB are important for survival of the pneumococcus in both the respiratory tract and bloodstream (Manco et al., 2006). This study also showed that the NanA mutant was quickly cleared from the respiratory tract, while the NanB mutant was able to persist without any increase in bacterial load, indicating these enzymes have differing roles (Manco et al., 2006). While there is no experimental evidence for NanC playing a significant biological role, analysis of *nanC* sequences across pneumococcal isolates indicated that any role it does play may be tissue-specific, with *nanC* being more commonly present in CSF isolates compared to carriage isolates (Pettigrew et al., 2006). Collectively, the pneumococcus is able to utilise these virulence factors to facilitate its pathogenesis.

1.5. Pathogenesis

1.5.1. Transmission and Carriage

Transmission is an essential factor for the survival of the pneumococcus as a species. It has long been known that for the pneumococcus to spread, there needed

to be close contact with a carrier, typically young children, and that these transmission events were more common during the colder months (Numminen et al., 2015). This would typically coincide with viral infections of the upper respiratory tract when shedding of respiratory secretions becomes more prevalent, creating the ideal environment to facilitate transmission (Gwaltney et al., 1975, Musher, 2003). The intricacies involved in this transmission, however, have only been realised in recent years with the development of suitable animal and human-to-human transmission models.

The pneumococcus is typically found in the mucous layer that overlays the epithelium of the upper respiratory tract. In this environment the release of Ply results in increased inflammation at the site, allowing for increased shedding of pneumococci and transmission to new hosts (Zafar et al., 2017). Similarly, co-infection with upper respiratory viruses such as influenza virus can lead to increased shredding of pneumococci, as a result of increased stimulation of secretions (Weiser et al., 2018). When pneumococci are ingested into the phagosome upon detection by the immune system, they are killed by lysozyme (Davis et al., 2011). This results in the release of Ply and subsequent pore formation, triggering a pro-inflammatory immune response (Parker et al., 2011, Karmakar et al., 2015). The increase in inflammation and respiratory secretions creates an unhospitable environment for the pneumococcus, triggering transmission to a new, more suitable host (Weiser et al., 2018). Once shed from the index host, the pneumococcus must survive for long enough in the external environment to ensure its transmission to a new carrier.

The pneumococcus can survive for days under ambient conditions with sufficient nutrients, and has even been cultured from soft toys recently handled by colonised children in previous studies (Marks et al., 2014). In conditions which are considered nutrient poor, the expression of pneumolysin can increase the survival of

pneumococci *ex vivo* as a result of there being increased nutrients present within respiratory secretions (Zafar et al., 2017). The ability for the pneumococcus to form a biofilm can also aid in increasing survival times in nutrient-poor environments, as they enter a 'dormant' state, but are capable of reverting to their original phenotype when conditions become more favourable (Walsh and Camilli, 2011, Marks et al., 2014).

Nasopharyngeal colonisation, also referred to as 'carriage' is often asymptomatic, and a naturally occurring event within the first few months of life (Bidossi et al., 2012). Studies have indicated that pneumococcal carriage can occur in >50% of healthy children and between 4 – 10% of healthy adults (Regev-Yochay et al., 2004, Hosseini et al., 2015, Wang et al., 2017). There are several factors that are required to ensure the establishment of colonisation. The first line of defence that the pneumococcus must overcome in the upper respiratory tract is mucous entrapment (Feldman et al., 1992) as shown in **Figure 1.3**. A gel-like mucin glycoprotein overlays the epithelium in the upper respiratory tract containing immunoglobulins and antimicrobial peptides (Rose and Voynow, 2006). While this mucin overlay is designed to prevent resident pathogens of the upper respiratory tract from interacting with the epithelium, it provides an ideal environment for bacteria to reside with ample nutrients (Weiser et al., 2018). As discussed in Section 1.4, the pneumococcus possesses a range of adhesins that aid in its ability to bind to the host epithelium, even increasing their expression in response to inflammation to further promote adherence. Some of these factors include surface-located adhesins such as pneumococcal adherence and virulence protein A (PavA) and PavB, which bind to fibronectin and plasminogen, components of the extracellular matrix (Holmes et al., 2001, Jensch et al., 2010). ChoP present on cell wall teichoic acid are able to bind to platelet-activating factor receptor (PAFR), while
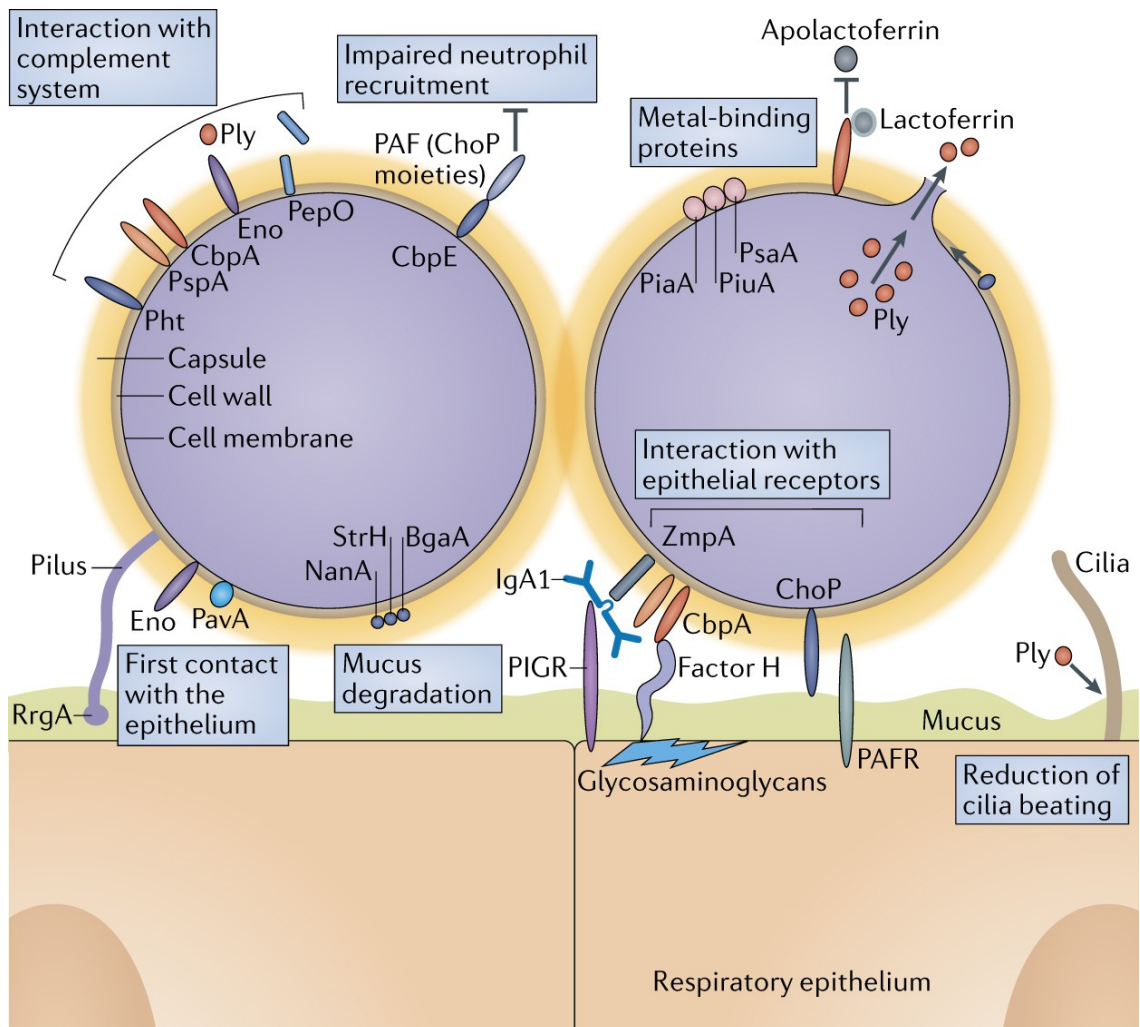
Figure 1.3. Molecular mechanisms involved in pneumococcal colonisation

Streptococcus pneumoniae must first establish contact with the epithelium and epithelial receptors to facilitate colonisation. During colonisation, the pneumococcus must interact and overcome the complement system to prevent clearance by the immune system. It must also degrade the mucous present on the epithelial layer to gain access to receptors present on these cells to promote adherence and colonisation. During this process, the pneumococcus must also act to impair neutrophil recruitment and activation at the site to again prevent clearance by the innate arm of the immune system. The pneumococcus will also utilise the pro-inflammatory effects of pneumolysin during colonisation, to gain access to the respiratory epithelium.
Reproduced from Weiser *et al.,* (2018).

PspC is able to bind to the polymeric IgG receptors (PIGR), along with host factor H and vitronectin (Cundell et al., 1995, Zhang et al., 2000). Additionally, the pneumococcus encodes at least 10 extracellular glycosidases that can enhance adherence to the host epithelium by modifying glycoconjugates to reveal glycan receptors (King, 2010). These glycosidases can release carbohydrates which can be imported and metabolised. This will be discussed in more detail in Section 1.9. Two of these glycosidases, Neuraminidase A (NanA) and β-galactosidase (BgaA) have additional lectin domains that appear to function as adhesins independently from their enzymatic activity (Uchiyama et al., 2009, Limoli et al., 2011).

Another factor that must be considered during colonisation is interactions between the pneumococcus and other commensal nasopharyngeal microflora, which can be either competitive or cooperative (Shak et al., 2013). Human colonisation studies have shown that acquisition of *S. pneumoniae* after intranasal challenge leads to increased diversity within the microbiota (Cremers et al., 2014). A stable state of inflammation in the upper respiratory tract (URT) is the most favourable for the presence of *S. pneumoniae*, such as that which occurs during coinfection with upper respiratory viruses (Nakamura et al., 2011). Viruses such as influenza result in expression of pro-inflammatory cytokines that upregulate expression of the epithelial receptors used by the pneumococcus for adherence (McCullers and Rehg, 2002). This compromises the nasopharyngeal epithelium, providing more nutrients in the immediate environment and promoting colonisation (Avadhanula et al., 2006).

1.5.2. Invasive Pneumococcal Disease

Establishing stable colonisation in the nasopharynx provides the optimum lifestyle for *S. pneumoniae*. Here it is able to survive for long periods, while also having the opportunity to readily transmit to new hosts. As mentioned above, viral infection can result in a pro-inflammatory immune response, causing damage to the respiratory

epithelium. These conditions can cause increases in bacterial loads in the URT. This not only facilitates transmission, but can also increase the likelihood of invasion of other host tissues, resulting in both localised and invasive disease, as outlined in **Figure 1.4**. Progression to invasive disease in particular is most likely when the pneumococcus transits to the lungs, or crosses the epithelial and/or endothelial layers housing the blood or cerebrospinal fluid (CSF) (Weiser et al., 2018).

The process of invasion is a complex one. Firstly, the pneumococcus must avoid entrapment in the mucous overlaying the epithelial layer, as explored in Section 1.5.1. The negatively charged capsule and degradation of IgA1 by the zinc metalloprotease ZmpA aid in preventing this entrapment. From here, glycosidases such as NanA, BgaA and StrH are able to unmask glycan targets present on the epithelium, revealing locations for pneumococcal adhesins to bind, while pneumolysin prevents ciliary beating (Kadioglu et al., 2008). There are five cell surface structures in particular that mediate adherence to the apical surface of epithelial cells; ChoP, CbpA, PavA, the large surface-exposed glycoprotein PsrP and finally the ancillary pilus subunit RgrA which is located at the tip of pili. When *S. pneumoniae* binds to PAFR via ChoP, or to PIGR via CbpA (Zhang et al., 2000), it subverts the host receptor recycling pathways, triggering endocytosis into epithelial cells and release at the basolateral surface upon exocytosis (Cundell et al., 1995). The actions of pneumolysin and hydrogen peroxide can provide an alternative method of transit to the interstitium as a result of disrupting the integrity of the epithelium, while hyaluronate lyase (Hyl) and plasmin can degrade the extracellular matrix, enabling paracellular invasion. Binding of ChoP to PAFR and CbpA to PIGR also promote traversing of the endothelium, allowing access to the bloodstream and subsequent haematogenous spread (Cundell et al., 1995, Zhang et al., 2000).
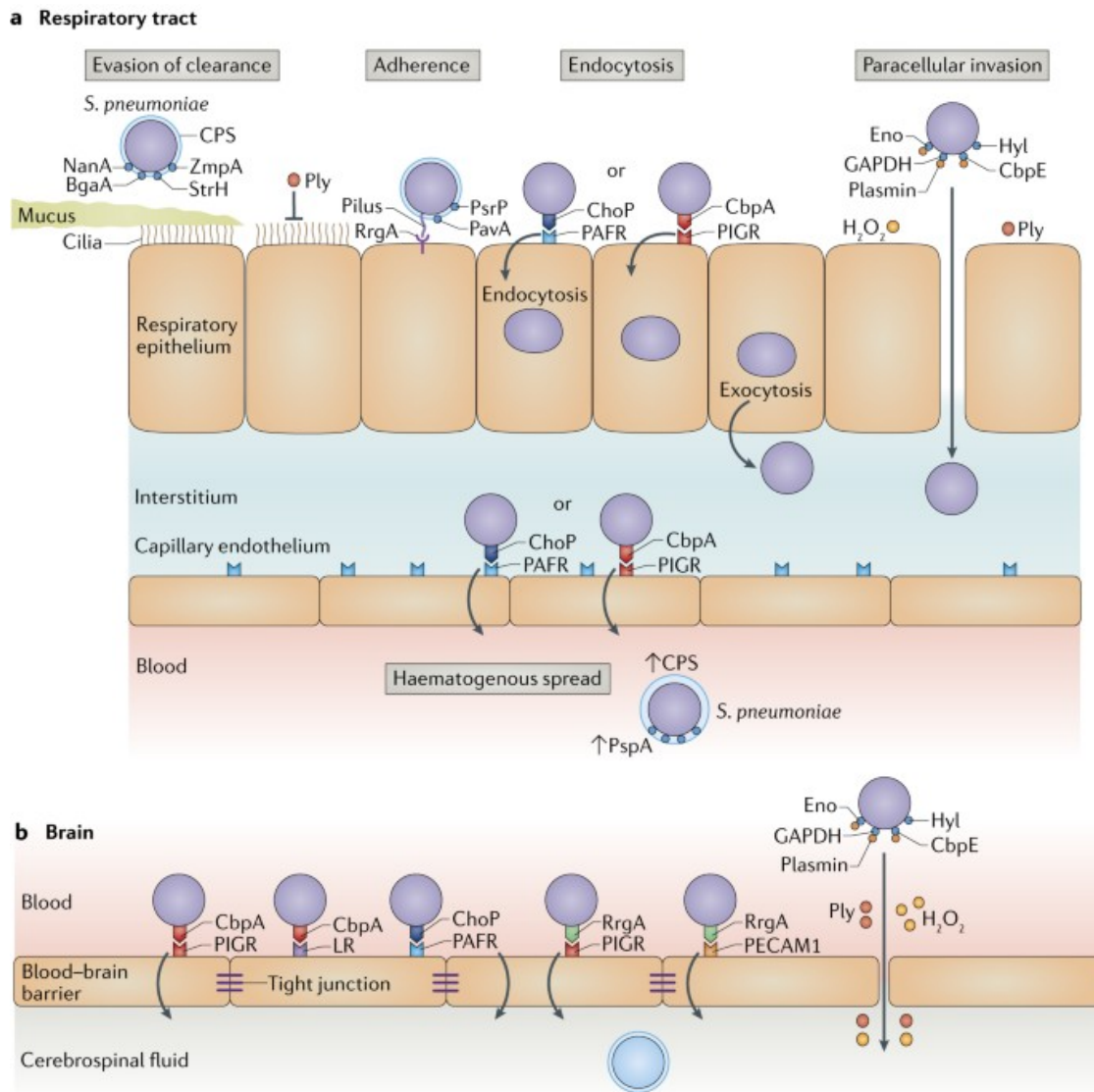
Figure 1.4. Stages of pneumococcal adherence and invasion

To cause invasive disease, *S. pneumoniae* must first adhere to the epithelium of the upper respiratory tract **(a)** after evading entrapment in mucous and mucociliary clearance. Binding via ChoP to PAFR and CbpA to PIGR triggers endocytosis of the pneumococcus into the cell, allowing transit to the basolateral surface. Alternatively, pneumolysin and the release of hydrogen peroxide damages the integrity of the epithelium, allowing paracellular invasion by pneumococci. This permits haematogenous spread from the interstitium into the blood stream **(b)**, allowing subsequent penetration of the blood-brain barrier and access to the cerebrospinal fluid. Reproduced from Weiser *et al.,* (2018).

Inflammatory cytokines present at the site of invasion will cause upregulation of PAFR expression further facilitating invasion via this pathway (Weiser et al., 2018). In order to traverse the blood-brain barrier (BBB), the pneumococcus again utilises ChoP-PAFR and CbpA-PIGR binding, along with CbpA-laminin receptor binding on the brain microvascular endothelium (Orihuela et al., 2009, Brown et al., 2014). This allows *S. pneumoniae* to cross the BBB, leading to meningitis. Similar to invasion of the respiratory epithelium, pneumolysin and hydrogen peroxide generated by α-glycerophosphate oxidase (GlpO), as well as binding of pneumococcal surface proteins enolase (Eno), glyceraldehyde 3-phosphate dehydrogenase (GAPDH) and choline binding protein E (CbpE) to plasmin can all lead to decreased integrity of the BBB (Rayner et al., 1995, Attali et al., 2008, Mahdi et al., 2012), promoting invasion of these tissues and subsequent invasive disease

1.6. Host Immune Response to Pneumococcal Infection

1.6.1. Innate Immunity

The innate immune response is essential for controlling colonisation of the URT and for providing protection during invasive disease (Koppe et al., 2012). Innate immunity provides a non-specific, yet rapid response initiated by the recognition of anything considered 'non-self'. Recognition of invading pathogens is mediated by Pattern Recognition Receptors (PRRs). They can detect Pathogen-Associated Molecular Patterns (PAMPs), resulting in the initiation of a signalling cascade leading to what is known as the innate immune response (Janeway et al., 2001, Paterson and Mitchell, 2006).

Toll-like Receptors (TLRs) are one such PRR that reside throughout human immune cells (Koppe et al., 2012). TLR2, TLR4 and TLR9, specifically, play roles in recognition and response to pneumococcal infection (Srivastava et al., 2005,

Richard et al., 2014, Tomlinson et al., 2014). TLR2 is located within the plasma membrane of cells and is able to recognise PAMPs in the form of pneumococcal cell wall components such as lipoteichoic acid and lipoproteins (Yoshimura et al., 1999, Schroder et al., 2003, Schmeck et al., 2006). TLR4 also resides in the plasma membrane and recognises Ply as a PAMP (Malley et al., 2003, Koppe et al., 2012). TLR9 is located within endosomes and can recognise pneumococcal DNA containing unmethylated CpG motifs (Koppe et al., 2012, Albiger et al., 2005). There is increasing evidence to suggest that TLR2 and TLR9 can enhance pneumococcal phagocytosis and intracellular killing in leucocytes (Albiger et al., 2005, Letiembre et al., 2005). TLRs can differentially engage with four major adaptor molecules, myeloid differentiation primary response 88 gene (MyD88), TIR-domain-containing adaptor-inducing-interferon β (TRIF) protein, MyD88-adaptor-like (MAL) protein, and TRIF-related adaptor molecule (TRAM). Interaction with these adaptors triggers signalling pathways resulting in activation of transcription factors such as NF-κB, IRF3 and IRF7 (O'Neill and Bowie, 2007). MyD88 specifically is also able to mediate downstream signalling of interleukin (IL) – 1 and IL-18 receptors (Koppe et al., 2012), resulting in inflammation.

In addition to TLRs, Nucleotide-binding and Oligomerisation Domain- (NOD-) like Receptors (NLRs) play an important role in the innate immune response, primarily activating NF-κB-dependent pro-inflammatory gene expression (Koppe et al., 2012). NOD1 is activated through recognition of Gram-negative bacteria, while NOD2 is activated by the peptidoglycans of essentially 'all' bacteria (Sorbara and Philpott, 2011). *In vitro* studies have previously shown that upon detection of internalised pneumococci, there is activation of NF-κB (O'Neill and Bowie, 2007), while NOD2-mediated recognition of *S. pneumoniae* mediated MCP-1 (CCL2) production, leading to recruitment of macrophages in the upper respiratory tract.

NOD2, together with TLR2 contribute to the clearance of pneumococcal colonisation within the host (Davis et al., 2011). Other NLRs, such as NLRP3, can form protein complexes known as inflammasomes (Schroder and Tschopp, 2010), which regulate IL-1β and IL-18 post-translationally by cleaving their zygomatic pro-forms to form functional cytokines (Koppe et al., 2012).

The actions of TLRs and NLRs lead to the recruitment of immune cells to the site of pathogen recognition. *S. pneumoniae* activates phagocytic cells and is then destroyed through different mechanisms involving TLRs, resulting in the induction of B cells to produce cytokines such as tumour necrosis factor α (TNFα), IL-6 and pro-IL-1β (Albiger et al., 2007, Ku et al., 2007, Wang et al., 2007, Dessing et al., 2008, Dogan et al., 2011, Koppe et al., 2012). Complement is activated through a C3-dependent cascade in response to infection (Kerr et al., 2005), leading to B-cell activation through the complement receptors CD21 and CD35 (Carroll, 2004). After antigen presentation, naïve B cells differentiate into IgM$^+$ memory cells, producing pneumococcus-specific IgM without the aid of T cells. Later, during hypermutation and class switching, some of these specific IgM$^+$ B cells will differentiate into pneumococcus-specific IgG$^+$ and IgA$^+$ memory B cells or plasma cells, thereby branching from innate immunity into adaptive immunity (Richards et al., 2010).

1.6.2. Adaptive Immunity

Adaptive, or acquired, immunity takes longer to develop than innate immunity, but it is pathogen-specific and highly effective. The adaptive immune response can be divided into two main branches: cell-mediated immunity and humoral immunity.

Cell-mediated immunity is dependent on the production of T-cells, which play a key role in the clearance of pneumococcal infection. T-cells can be further divided into two major classes, CD4$^+$ helper T-cells and CD8$^+$ cytotoxic T-cells (Mook-Kanamori et al., 2011). T-cell activation occurs upon recognition of antigens presented on the

major histocompatibility complex (MHC) of antigen presenting cells (APCs) such as macrophages, dendritic cells and B-cells (Brooks and Mias, 2018). Once activated, CD4$^+$ helper T-cells further differentiate into either Th1, Th2 or Th17 subsets which each play unique roles in the progression of the immune response. Th1 cells typically produce pro-inflammatory cytokines, driving an inflammatory immune response and recruiting additional immune cells to the site of infection. Th2 cells will interact with B-cells to trigger class switching and the generation of antibodies (Romagnani, 1999). Th17 cells are IL-17 producing T-helper cells that have been shown to play a role in anti-pneumococcal immunity. The release of IL-17 from these cells promotes recruitment of monocytes, neutrophils and macrophages to the site of infection, resulting in enhanced bacterial clearance (Hoe et al., 2017). Conversely, CD8$^+$ T-cells will directly kill infected host cells (van der Poll and Opal, 2009).

Humoral immunity involves the generation of antigen-specific antibodies via B-cells. The activation of B-cells and class switching that follows results in the generation of highly specific and functional antibodies, typically in the form of IgG and IgA, which will promote phagocytosis through opsonisation of the target pathogen (Mook-Kanamori et al., 2011). Post-infection, subsets of both T-cells and B-cells will differentiate into memory cells which will continue to circulate. If re-infection were to occur, these memory cells would be capable of undergoing rapid activation, allowing for a faster, pathogen-specific adaptive immune response and enhanced pathogen clearance (Sprent, 1994). This principle of 'memory' within the immune response is what makes prevention methods like vaccination such a robust and reliable method for preventing infection.

1.7. Prevention and Treatment of Pneumococcal Disease

1.7.1. Vaccines and Serotype Replacement

The first trials of pneumococcal vaccines were reported in the early 1910s, with a crude whole-cell vaccine (WCV) used to immunise South African gold mine workers who showed a high incidence of severe pneumococcal infections (Wright, 1914). This led to a raft of clinical trials to determine the safety and efficacy of pneumococcal vaccination against particular serotypes. The validity of these studies, however, was soon questioned due to a lack of randomisation and methodological flaws (Lister, 1917, Ekwurzel, 1938, Felton, 1938). In the 1940s, trials commenced on bivalent, trivalent and quadrivalent pneumococcal capsular polysaccharide vaccines which proved more successful and showed that there could indeed be immunity generated as a result of vaccination (MacLeod et al., 1945, Kaufman, 1947). It was around this time that antibiotics effective in the treatment of pneumococcal infections were introduced into clinical use. In particular, the efficacy of penicillin against pneumococcal disease was deemed 'miraculous' and led to the widespread conclusion that these infections were completely curable and didn't need effective vaccines to control the spread and severity of infection. Because of this, all pneumococcal vaccines were effectively withdrawn from the market by the late 1950s (Butler et al., 1999). This changed however in 1964, when Robert Austrian and Jerome Gold presented data on more than 2,000 pneumococcal pneumonia cases that presented at Kings County Hospital in Brooklyn between 1952 and 1962 (Austrian, 1964). They reported that despite the efficacy of macrolides in the treatment of infections, mortality rates within these patients were as high as 1 in 4, with this mortality being more pronounced in the elderly. This prompted Austrian and others to reassess the development of an efficacious polyvalent pneumococcal polysaccharide vaccine (U.S. Congress,

1979). With the use of a double-blind randomised trial (Austrian, 1984), Austrian and others were able to show rates of protective efficacy against invasive pneumococcal infection and pneumonia between 76 – 92% post-vaccination (Austrian et al., 1976, Smit et al., 1977). This resulted in the licensure of the first 14-valent pneumococcal polysaccharide vaccine in 1977. The implementation of this vaccine saw a reduction in both morbidity and mortality in the healthy population, but overall rates of pneumonia remained mostly unaffected (Riley et al., 1977). In 1983, the introduction of an expanded 23-valent polysaccharide vaccine (PPSV-23) provided coverage against 80-90% of disease-causing pneumococcal serotypes and this vaccine is still in use today. While PPSV-23 reduced the incidence of invasive pneumococcal disease within the target population, the overall rates of pneumococcal carriage and pneumonia have remained unchanged (Fine et al., 1994, Stanek et al., 2016). One additional factor complicating the overall efficacy of PPSV-23 is the low levels of immunogenicity it elicits in the most vulnerable target populations, these being young children, the elderly and immunocompromised individuals. The low levels of immunogenicity here is a result of the CPS being a T-cell independent antigen. The majority of antibody responses occurring in response to this antigen are a direct result of expansion of cognate B cells (Daniels et al., 2016). This means that, while able to induce an immune response, it is not an adequate response in these age groups, and thus the effectiveness of PPSV-23 is compromised (Heilmann, 1990, Shapiro et al., 1991). The continued high risk of invasive pneumococcal disease within these groups led to the development of polysaccharide-protein conjugate vaccine (PCVs), in which each of the component polysaccharides is conjugated to a protein carrier, thereby promoting a T-cell dependent immune response, antibody maturation and generation of memory cells. A 7-valent PCV (PCV7) covering serotypes 4, 6B, 9V, 14, 18C, 19F and 23F, in

which the capsular polysaccharide was conjugated to the protein carrier CRM197, a non-toxic variant of diphtheria toxin, was licensed in 1999. PCV7 elicited a strong immune response, even in children under the age of two. Widespread use in paediatric populations resulted in a reduction in both carriage and IPD in this age group, as well as in the adult population as a consequence of herd immunity (Whitney et al., 2003, Hammitt et al., 2006, von Gottberg et al., 2014). In 2010, PCV7 was expanded to include six additional serotypes (1, 3, 5, 6A, 7F and 19A), to cover serotypes particularly prevalent in developing countries (1 and 5), as well as others which had remained significant causes of disease globally. Types 6A and 19A were included to address the poorer than expected degree of cross-protection elicited by the 6B and 19F antigens included in the original formulation (Bryant et al., 2010).

Notwithstanding their high immunogenicity, PCVs elicit largely serotype-dependent protection, and whilst their widespread use has proven highly successful in the reduction of invasive disease caused by included serotypes, rates of disease due to several non-vaccine serotypes has increased, partially off-setting the overall reductions in morbidity and mortality. This phenomenon is termed 'serotype replacement' and has been widely reported in the literature since the introduction of PCVs (Dagan, 2009, Gladstone et al., 2017). This underpins the need for non-serotype specific vaccines that will target non-capsular virulence factors common to the majority of pneumococcal serotypes. There are a range of target protein antigens that have been investigated for their potential use as vaccine candidates, including neuraminidase, LytA, Ply, CbpA, PspA, pneumococcal surface antigen A (PsaA) and polyhistidine triad protein D (PhtD) to name a few (Lock et al., 1988, McDaniel et al., 1991, Paton et al., 1993, Sampson et al., 1997, Chen et al., 2015, Pichichero, 2017, Afshar et al., 2020). Protein antigens such as these are expected

to elicit a T-cell-dependent response that results in long term immunological memory. There is also now increasing evidence to support the use of WCVs in the prevention of pneumococcal disease, much like the principles employed in the first documented vaccine from 1911 (Wright et al., 1914). The newer WCVs are based on killed, non-encapsulated pneumococci, that display a wide array of protein antigens on their outer surface. This elicits a broad anti-surface protein response that mimics the natural human immunity that develops in late childhood-early adulthood through previous exposure to pneumococci. Indeed, anti-protein antibodies rather than anti-capsular antibodies are largely responsible for natural human immunity (Wilson et al., 2015). WCVs based on mutated derivatives of the un-encapsulated *S. pneumoniae* Rx1 strain have previously been shown to provide protection against both carriage and sepsis in mice, indicating that the ability to generate antibodies against a range of pneumococcal antigens can play a role in preventing invasive disease and carriage (Malley et al., 2001, Lu et al., 2010, Babb et al., 2016). These WCVs were inactivated either by chemical killing (Moffitt and Malley, 2016) or gamma-irradiation (Babb et al., 2016, Babb et al., 2017, David et al., 2019), and are at the late pre-clinical or early clinical trial stages of development. WCVs offer the prospect of cheap broadly-effective protection against all pneumococci, regardless of serotype (Morais et al., 2019).

1.7.2. Antibiotics and Emerging Resistance

The introduction of beta-lactams such as penicillin in the 1940s led to a significant reliance on antibiotics for the treatment of pneumococcal infections. This was a suitable approach until the emergence of resistant pneumococcal strains, the first of which was reported in a paediatric patient in Australia in 1967 (Hansman and Bullen, 1967). In the years that followed, resistant isolates were continually isolated from both patients who had previously undergone beta-lactam treatment, and those

who hadn't (Hansman et al., 1971). The misuse and overuse of beta-lactams played a significant role in the development of this resistance and continues to be an issue to this day (Cherazard et al., 2017). In 2017, penicillin-non-susceptible *S. pneumoniae* was listed as one of 12 'priority pathogens' that posed the greatest threat to human health (World Health Organization, 2017). The mechanisms underpinning this resistance are now understood to arise from structural modification of one or more penicillin binding proteins (PBPs), such that penicillin can no longer bind and reduce peptidoglycan synthesis (Cornick and Bentley, 2012). Rates of penicillin resistance vary globally, with non-susceptibility rates surpassing 30% in Spain and 50% in the Middle East and East Asia, while Finland, Sweden and Germany maintain resistance rates as low as 5% (Bruinsma et al., 2004, Lynch and Zhanel, 2009, Mamishi et al., 2014, Cherazard et al., 2017, El Moujaber et al., 2017).

In addition to beta-lactams, macrolides such as erythromycin, azithromycin and clarithromycin have become standard in antibiotic treatment regimes (Cherazard et al., 2017). Macrolides are characterised by the presence of a large lactone ring varying from 12 to 16 atoms in size (Zhanel et al., 2001). Macrolides act by binding to the 23S rRNA, thereby inhibiting RNA-dependent protein biosynthesis and causing cell death (Cherazard et al., 2017). Current resistance rates to macrolides range from 10% to 70% across the globe, with these rates continuing to rise (Zhanel et al., 2014). Continued overuse of antibiotics is applying strong selective pressure, which is driving the increasing rates of resistance that are seen, leaving us dependent on new or combined treatment regimens, which are much more expensive. It is also increasing the need for improved prevention strategies in the fight against pneumococcal disease. Gaining a fundamental understanding of the

processes undertaken by the pneumococcus that facilitate invasive disease can provide new avenues for potential treatments.

1.8. Quorum Sensing (QS)

At different stages of infection, the pneumococcus is capable of forming biofilms. Biofilms are highly organised and dense structures, with more than 99% of microbial life thought to exist in this state (Watnick and Kolter, 2000). Within these complex structures, individual cells can interact with neighbouring cells to coordinate a vast array of cellular processes. This process is known as Quorum Sensing (QS) and is governed by the coordinated sensing and release of small signalling molecules known as autoinducers (Miller and Bassler, 2001, von Bodman et al., 2008). This process is linked to a number of critical features in the pneumococcus, as detailed below:

1.8.1. Oligopeptide two-component type QS

The oligopeptide two – component type QS system, found only in Gram-positive bacteria, utilises an autoinducer signalling peptide (AIP) and a two-component signal transduction system (TCSTS) (Novick, 2003, Claverys et al., 2006). TCSTSs consist of a histidine kinase and a response regulator protein. When an AIP is sensed by the histidine kinase, it transfers a phosphate group to the response regulator protein, resulting in differential gene expression (Stock et al., 2000). The CSP/Com system controlling genetic competence in pneumococci is an example of this type of QS system. Indeed, it was the first QS system to be described in the literature (Tomasz, 1965).

1.8.2. Competence and Natural Transformation

Competence is the feature governing the natural transformability of the pneumococcus (Vidal et al., 2013). This refers to a transient state developed by a

range of bacterial species, during which 'competent' cells upregulate genes responsible for the uptake of foreign DNA, enabling natural transformation and the rapid acquisition of favourable traits from neighbouring cells (Moreno-Gamez et al., 2017). Pneumococcal competence is regulated by the *comABCD* locus. In short, the signalling peptide required by this system is Competence Stimulating Peptide (CSP), encoded by the gene *comC*. The CSP precursor, pre-CSP is transported from the cytoplasm across the cell membrane to the extracellular environment. During this process, pre-CSP is cleaved to its active form CSP and begins to accumulate in the extracellular environment, where it is able to be sensed by both cells producing CSP and neighbouring cells via the histidine kinase ComD. This results in phosphorylation of the cognate response regulator ComE, which then recognises and binds specific DNA binding motifs, resulting in differential gene expression, including late competence genes that encode proteins required for uptake and integration of extracellular DNA (Galante et al., 2015). The release and sensing of CSP by pneumococci during development of competence is the earliest example of bacterial quorum sensing (Tomasz, 1965).

While competence has always been critical for the acquisition of new virulence genes (Zhu and Lau, 2011), studies have also shown that the competence regulon can cross-regulate virulence (Lau et al., 2001, Ibrahim et al., 2004, Guiral et al., 2005, Claverys et al., 2007, Kowalko and Sebert, 2008). Studies assessing the role of ComB, a component of the ABC transporter required for CSP export, have shown that loss of function of this protein results in attenuation of *S. pneumoniae* in a murine model of infection (Lau et al., 2001). Similar attenuation has also been observed in *S. pneumoniae* mutants lacking the histidine kinase ComD (Lau et al., 2001).

1.8.3. S-ribosyl-homocysteine lyase (LuxS)-mediated Autoinducer 2 QS

A QS system common to both Gram-positive and Gram-negative bacteria that allows them to communicate with their neighbours is the LuxS-mediated autoinducer 2 (AI-2) QS system (Wang et al., 2012). This system is dependent on the synthesis of the keto-pentose sugar-like molecule Autoinducer-2 by the metabolic enzyme LuxS (S-ribosyl-homocysteine lyase). AI-2 synthesis by LuxS occurs as a by-product of the activated methyl cycle in the pneumococcus, arising during the conversion of *S*-ribosyl-homocysteine to homocysteine (Yadav et al., 2018). Homologues of LuxS have been found across almost all bacterial species, indicating this is a highly conserved method of cell-to-cell communication (Kaur et al., 2018). Despite this, the mechanism whereby AI-2 accumulation mediates transcriptional responses is poorly understood, particularly in Gram-positive bacteria.

1.9. Pneumococcal Carbohydrate Metabolism

S. pneumoniae is a strictly fermentative organism, relying solely on glycolysis for energy production. The pneumococcus lacks the genetic material required to encode the tricarboxylic acid cycle and electron transport chain, leaving it reliant on the availability of carbon sources present within the human host for survival (Hoskins et al., 2001). The importance of carbohydrate metabolism in the pneumococcus is exemplified by the sheer number of sugar transporters encoded within the genome. Of all transporters present, 30% of them are dedicated to the uptake of carbohydrates (Hoskins et al., 2001, Bidossi et al., 2012). To date, there have been 32 carbohydrates identified as potential energy sources in *S. pneumoniae* that can be readily imported by 21 phosphoenolpyruvate phosphotransferase system transporters (PTSs) and 8 ATP-binding cassette (ABC) transporters (Bidossi et al.,

2012). These two transporter families differ primarily in their ability to phosphorylate incoming cargo. ABC transporters are composed of an extracellular solute binding domain, capable of binding solute and delivering it to the transporter. The solute then traverses the membrane through the transmembrane domain (TMD). Directly beneath the TMD sits the nucleotide binding domain, which is able to hydrolyse ATP, providing energy to facilitate transport (Marion et al., 2011, Buckwalter and King, 2012). The PTS on the other hand is composed of two cytoplasmic proteins, Enzyme I and Histidine Protein (HPr) which transfer phosphate groups from phosphoenolpyruvate to the multi-domain Enzyme II, a sugar-specific transporter facilitating import of the desired carbon source (Buckwalter and King, 2012, Postma et al., 1993).

During the course of colonisation and infection, the pneumococcus encounters a range of niches, each with distinct nutrient availability. This means that niche adaptation by the pneumococcus is critical. Glucose (Glc) is the preferred carbon source for the pneumococcus, followed closely by sucrose. The presence of these sugars results in suppression of the metabolic pathways required for the utilisation of other, non-preferred carbon sources. This suppression occurs via a process known as 'Carbon Catabolite Repression' (CCR), preventing the simultaneous utilisation of sugars to maintain metabolic efficiency. During CCR, the transcription factor carbon catabolic protein A (CcpA) will bind to catabolite repression elements (*cre*) present within the promotor region of genes responsible for non-preferred carbon source utilisation, resulting in repression. This process has been shown to play critical roles in pneumococcal disease progression, with deletion of *ccpA* resulting in decreased colonisation of the nasopharynx and reduced transit of pneumococci to the lungs and blood in a murine model of infection (Iyer et al., 2005). Other upper respiratory tract co-colonisers such as *Haemophilus influenzae,*

Neisseria meningitidis and *Moraxella catarrhalis* have far greater limitations on their ability to metabolise carbohydrates, but they are not solely reliant on them for survival (Shakhnovich et al., 2002).

1.9.1. Glucose as a Carbon Source

As mentioned in Section 1.9, Glc is the preferred carbon source for the pneumococcus (Carvalho et al., 2011). Glc is typically found in the blood of the human host ranging in concentration from 3.5 mM to 5.5 mM in healthy adults, along with inflamed tissues, however its availability differs between niches (Guemes et al., 2016). Glc is predominantly imported by the PTS MalN-MalLM system in the pneumococcus (Bidossi et al., 2012), resulting in phosphorylation of the incoming cargo to Glc-1-P, permitting entry into the glycolytic pathway. In addition, there is an additional ABC transporter that can import Glc, which is then phosphorylated to Glc-6-P by the glucokinase Gki before entering the glycolytic pathway (Bidossi et al., 2012). In niches such as the nasopharynx, where Glc concentrations are below 1mM (Philips et al., 2003), the pneumococcus employs a range of glycosidases such as NanA, BgaA and StrH that are able to cleave sugars from host *N*-linked glycoconjugates (King, 2010). This allows these liberated sugars to then be taken up by their cognate PTS or ABC transporters to support energy production.

1.9.2. Galactose as a Carbon Source

Galactose (Gal) is a monosaccharide sugar that is the primary carbon source found in the upper respiratory tract (Afzal et al., 2015). This is a tactic employed by the human host to maintain airway sterility by actively eliminating Glc from this niche. This then limits the bacteria that can persist there to those that can metabolise the carbon sources available in the absence of Glc. Gal can be metabolised by two pathways in *S. pneumoniae*, the Leloir pathway and the Tagatose-6-Phoshate (T6P) pathway.

1.9.2.1. Leloir Pathway

The Leloir Pathway is the primary pathway for metabolism of Gal in the pneumococcus. It is not a discrete pathway found only in the pneumococcus, rather it is found in both Gram-positive and Gram-negative bacteria. Gal is able to enter the cell via an as yet unidentified ABC transporter. A study by Bidossi and collaborators identified a putative ABC transporter that may be involved in Gal import, SPD_0088-0090, but the role it plays as a transporter is yet to be confirmed (Bidossi et al., 2012). Imported Gal is firstly converted to α-Gal by the aldolase 1-epimerase GalM. This α-Gal is then phosphorylated intracellularly at the C1 position by the kinase GalK, yielding α-Gal-1-phosphate (αGal1P), as shown in **Figure 1.5**. αGal1P is then converted to αGlc1P by GalT, a hexose-1-phosphate uridyltransferase, while the UDP-Glc epimerase GalE simultaneously convert UDP-Gal to UDP-Glc. The resultant αGlc1P generated by GalT can then feed into two subsequent pathways. The first of these is glycolysis, when αGlc1P is converted to Glc-6-phosphate (G6P) by a phosphoglucomutase (*pgm)*. The second is when αGlc1P is converted to UDP-Glc by the UDP-Glc pyrophosphorylase GalU, resulting in the generation of a key precursor for the synthesis of the various nucleotide sugars required for capsular polysaccharide biosynthesis (Carvahlo et al., 2011, Paixão et al., 2015)

There is little known about the regulation of the Leloir pathway in the pneumococcus. A study by Afzal *et al.,* was the first to show that GalR acts a transcriptional activator of the Leloir pathway by activating *galKT* (Afzal et al., 2015) (see **Figure 1.6**). Using gene expression analyses, they were able to show that the expression of both *galK* and *galT* was significantly higher in the presence of Gal compared to Glc in a serotype 2 wildtype strain (D39). They further concluded that GalR acts as the regulator of these genes by again performing transcriptional analyses in a *galR-*

deletion mutant (D39Δ*galR*), and assessing expression of both *galK* and *galT* in the presence of Gal. GalR belongs to the LacI-family of transcriptional regulators, a group that is primarily involved in the regulation of catabolic pathways utilising sugars such as lactose, ribose, fructose and maltose, to name a few (Nguyen and Saier, 1995). In these pathways the sugars themselves, or their phosphorylated counterparts, act as the effector molecule for these transcriptional regulators. Members of the LacI-family are mostly repressors; however, some have been reported to act as both activators and repressors, with Swint-Kruse and colleagues reporting that LacI-family transcriptional regulators typically have altered DNA-binding affinity after binding with an effector ligand (Swint-Kruse and Matthews, 2009). GalR has previously been shown to be the transcriptional repressor of the *gal* operon in both *S. mutans* and *Escherichia coli* in the absence of Gal (Weickert and Adhya, 1993, Ajdic and Ferretti, 1998), so the same action was proposed to occur in *S. pneumoniae*. Afzal and colleagues reported that upon deletion of *galR*, *galKT* became inactivated, indicating that GalR does indeed act as a transcriptional repressor in *S. pneumoniae*. They additionally performed growth analyses, comparing the growth of both D39 and a *galK* deletion mutant (D39Δ*galK*) in the presence of Gal. They reported that GalK is required for the proper growth of *S. pneumoniae* in the presence of Gal. These findings were among the first to elucidate the importance of both GalR and GalK in pneumococcal Gal metabolism (Afzal et al., 2015).

1.9.2.2. Tagatose-6-Phosphate Pathway

The Tagatose-6-Phosphate (T6P) Pathway is primarily utilised for the metabolism of lactose in the pneumococcus, but it is also able to breakdown Gal. The T6P is encoded by two separate operons, *lacABCD* under the control of the repressor LacR2 and *lacTFEG* under the control of the activator LacT (Afzal et al., 2014) (see

Figure 1.6). Uptake of Gal by this system is believed to be mediated by several proposed PTSs, encoded by SPD_0262 - 264, SPD_0559 – 0561 and more recently, SPD_0066 – 0069 (Bidossi et al., 2012). LacE2 is a component of the PTS transporter responsible for importing Gal, resulting in the conversion of Gal to Gal6P. This Gal6P is converted to tagatose-6-phosphate by the Gal-6-phosphate isomerase LacAB, as shown in **Figure 1.5**. Tagatose-6-phosphate is then converted to tagatose-1,6-bisphosphate by the tagatose-6-phosphate kinase LacC. Finally, tagatose-1,6-bisphosphate is converted to dihydroxyacetone-phosphate and D-glyceraldehyde-3-phosphate by the tagatose-1,6-bisphosphate aldolase LacD. The intermediates from this conversion then feed into pyruvate metabolism (Afzal et al., 2014). Gal taken up via the T6P pathway can enter the glycolytic pathway but does not generate Glc1P and so does not impact capsule biosynthesis. The existence of two distinct pathways indicates that Gal utilisation is complex, perhaps with a failsafe should one pathway become compromised. Until recently, however, the degree of interplay between alternative Gal utilisation pathways and pneumococcal QS, along with the potential impacts on both virulence and survival has remained unclear.

1.10. The Interplay Between Quorum Sensing and Galactose Metabolism

A study by Trappetti *et al.,* was the first to show a link between pneumococcal quorum sensing and sugar metabolism (Trappetti et al., 2017). Using a serotype 2 strain (D39) that had the gene *luxS* deleted ($\Delta luxS$), they were able to demonstrate a decreased ability for this mutant to grow in the presence of Gal compared to the wild-type (D39). They were also able to show that the growth of the $\Delta luxS$ mutant in Gal was able to be 'rescued' by the addition of the QS molecule AI-2 to the media in finite amounts. This gave the first indication that QS is linked to carbohydrate metabolism, as this relationship between sugar and signalling was confined to Gal
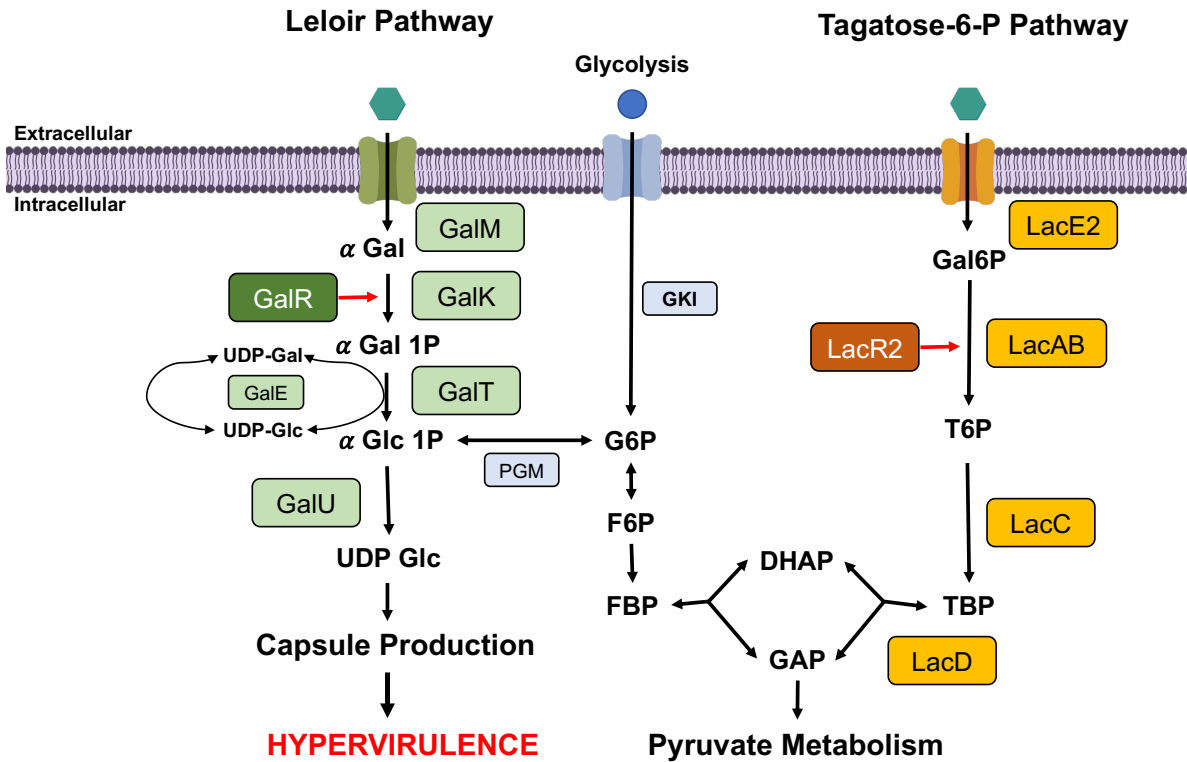
Figure 1.5. Overview of Gal metabolism in *Streptococcus pneumoniae*
There are two reported pathways for the metabolism of Gal in the pneumococcus. The first of these is the Leloir pathway (green). Here, Gal enters the cell via an ABC transporter where it is subsequently metabolised by the proteins encoded by *galM, galK, galT, galE* and *galU*. *galK* and *galT* are under the control of the repressor GalR. The intermediates generated by this pathway can feed in to either pyruvate metabolism/glycolysis (blue) or capsule production. Conversely, the tagatose-6-phosphate pathway (orange) imports Gal into the cell via a phosphoenolpyruvate phosphotransferase (PTS) transporter encoded by the *lacTFEG* operon and phosphorylates it. The resultant Gal-6-P is subsequently metabolised by proteins encoded by *lacA, lacB, lacC* and *lacD.* The intermediates from this pathway can then feed into pyruvate metabolism, resulting in energy production for the cell.

Leloir Pathway
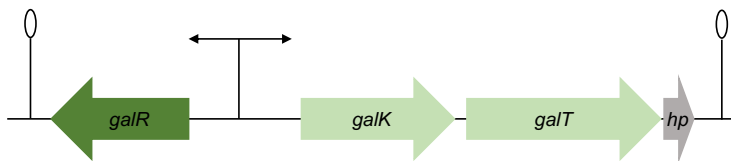
Tagatose-6-Phosphate Pathway
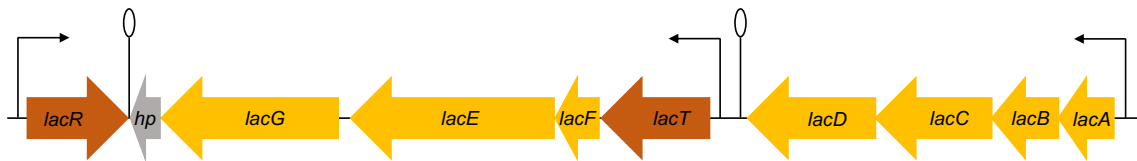
Figure 1.6. Organisation of genes responsible for galactose metabolism in *Streptococcus pneumoniae*

There are two reported pathways for the metabolism of Gal in the pneumococcus. The first of these is the Leloir pathway (green). The regulator GalR (dark green) is located upstream of the genes *galK* and *galT* (light green), acting as a transcriptional activator. The tagatose-6-phosphate pathway (orange) consists of two operons, *lacABCD* (light orange) under control of the repressor *lacR* (dark orange) and *lacTFEG* (light orange) under control of the activator *lacT* (dark orange). Black arrows represent transcriptional activator sites. Black ovals represent transcriptional terminator sites. Adapted from Afzal et al., 2014 and Afzal et al., 2015.

(i.e. the Δ*luxS* mutant grew as well as the wild type D39 in Glc). They further confirmed this link by quantifying total capsule present on both the wild-type and Δ*luxS* strains in the presence of Gal. There was a significant decrease in the amount of total capsule present on the Δ*luxS* mutant compared to the wild-type. Interestingly, upon addition of AI-2, they reported an increase in total capsule on the Δ*luxS* strain to a level comparable to that of D39. As explored in Section 1.9.2.1, while capsule is primarily generated from Glc metabolism, it can also be generated through the metabolism of Gal. The 'rescue' effect that is evident between both growth and capsule production upon addition of AI-2 in the presence of Gal further confirms the link between QS and sugar metabolism proposed in this study. Additional RNA sequencing in the presence of AI-2 revealed candidate genes that may be involved in this relationship. This included *fruA*, which exhibited a 1.7-fold change in expression between the wild-type and the Δ*luxS* mutant. Further capsule studies revealed that there was no change in the total amount of capsule in a Δ*fruA* mutant strain in the presence or absence of AI-2, while gene expression studies showed no change in the expression of the Leloir pathway genes *galR, galK,* and *galT* in the Δ*fruA* mutant in the presence of AI-2. These changes led them to conclude that FruA may be involved in the sensing and uptake of AI-2 into the cell (Trappetti et al., 2017). FruA is a PTS transporter which will phosphorylate incoming cargo during uptake. This led to the proposal that FruA-mediated import of AI-2 results in phosphorylation of the signalling molecule, meaning it can no longer freely diffuse out of the cell, thus becoming trapped intracellularly. From there, they hypothesised that this phosphorylated AI-2 (AI-2-P) may be able to directly or indirectly facilitate phosphorylation of GalR, the activator of the Leloir pathway (e.g. by acting as a phosphate donor), resulting in the upregulation of this pathway and a subsequent increase in galactose metabolism. (**Figure 1.7.**). This left the question
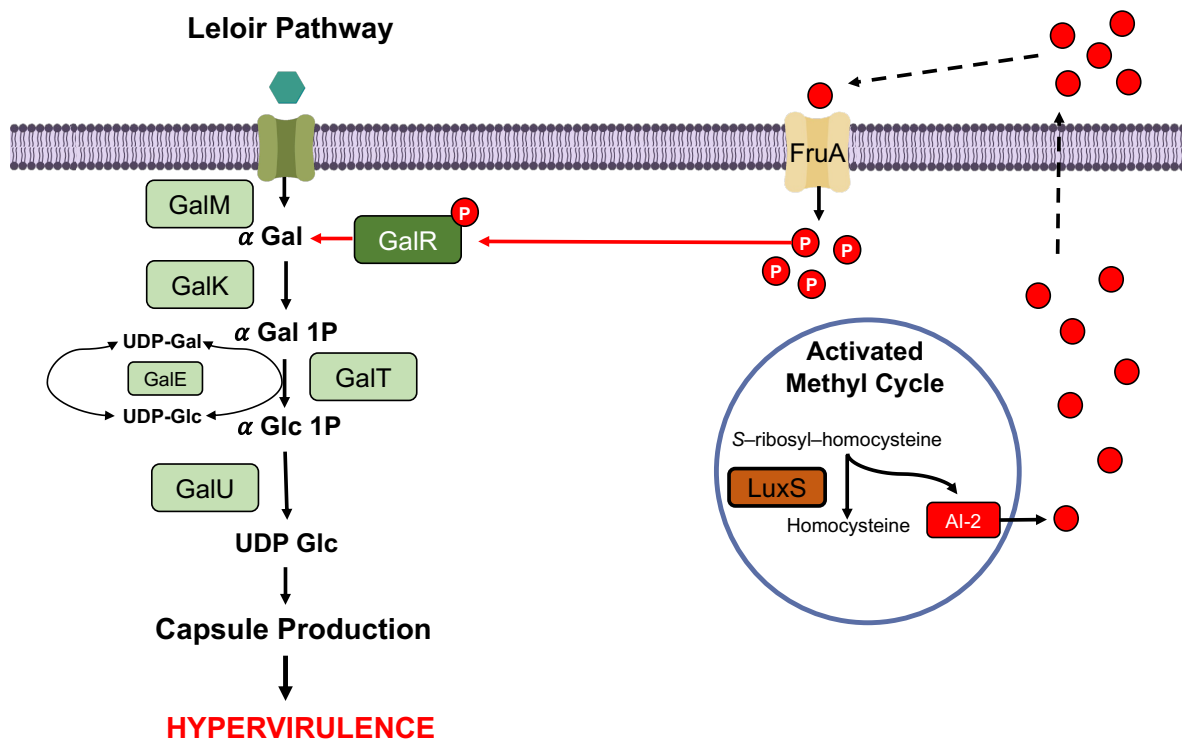
Figure 1.7. Visualising the interplay between Quorum Sensing and galactose metabolism in the pneumococcus

As proposed by Trappetti *et al.,* (2017), FruA may be the cognate transporter for the signalling molecule Autoinducer 2 (AI-2). When pneumococci undergo cell-to-cell communication (QS), AI-2 is produced as a by-product of the activated methyl cycle, where it can then diffuse to the extracellular environment and then be sensed by neighbouring cells. They propose that this AI-2 is able to be sensed and imported via FruA, a PTS family transporter involved in fructose uptake. Being a PTS, FruA can phosphorylate this AI-2 during import, thereby trapping the AI-2P in the intracellular environment. They propose that AI-2P is then able to directly or indirectly mediate phosphorylation of GalR, the regulator of the Leloir pathway. This phosphorylation is proposed to occur at three putative phosphorylation sites identified by Sun *et al.,* (2010): S317, T319 and T323. Phosphorylation of GalR is proposed to upregulate expression of the Leloir pathway in the presence of Gal, resulting in increased generation of the capsule precursor UDP-Glc. Increased UDP-Glc production results in increased total capsule, resulting in a hypervirulent phenotype.

of exactly where this phosphorylation may be occurring within GalR to permit this activation of the Leloir pathway and the subsequent hypervirulent phenotype resulting from increased capsule production. A previous study by Sun *et al* utilised phosphoproteomics to identify residues that may be subject to phosphorylation within the serotype 2 (D39) pneumococcal genome (Sun et al., 2010). There were three such residues identified within GalR: Serine 317 (S317), Threonine 319 (T319) and Threonine 323 (T323). Combined with these findings, Trappetti *et al* proposed that phosphorylation of GalR via AI-2-P is likely occurring at the putative phosphorylation sites S317, T319 and T323, however the exact mechanisms allowing for this to occur was yet to be established.

1.11. Research Project

1.11.1. Rationale

The findings of Trappetti *et al.,* (2017) provide the foundation for this study. They concluded that mutational analysis of the putative GalR phosphorylation sites may prove informative in elucidating the link between QS and Gal metabolism. This led to the development of the following hypothesis and aims:

1.11.2. Hypothesis, Aims and Strategy

Metabolism of Gal is critical for the survival of *S. pneumoniae* in the human upper respiratory tract. This metabolism has been shown to occur primarily through the Leloir pathway in the pneumococcus. Previous studies have shown that GalR is the regulator of this pathway and may be essential for the ability to metabolise Gal. Here, I hypothesise that GalR is essential for the ability to grow in the presence of Gal. Further to this, I propose that the putative GalR phosphorylation sites S317, T319 and T323 play a key role in this regulation of Gal metabolism, allowing for the

upregulation of genes within the Leloir pathway to permit growth when Gal is the principal carbon source. To address this hypothesis, I have developed the following aims:

Aim 1: Elucidate the role that the putative GalR phosphorylation sites play in the ability to metabolise Gal.

Aim 2: Investigate the interplay between the Leloir and Tagatose-6-Phosphate pathways.

Aim 3: Determine the broader impacts of Gal metabolism in a mouse model of pneumococcal infection.

Strategy: Firstly, a series of GalR amino acid substitution mutants will be constructed to determine which of the putative phosphorylation sites play a role in the ability to grow in Gal. These strains will then be further characterised by assessing expression of related genes and their impact on invasive disease. Moving on from here, additional gene deletion mutants will be constructed, deleting key genes from either the Leloir or Tagatose-6-Phosphate pathways. These strains will then be characterised in terms of growth and gene expression, before performing metabolomics to identify the level of interplay, if any, between these two pathways. Finally, RNA sequencing will be utilised to examine the greater transcriptomic responses occurring in both the murine and pneumococcal genomes in response to infection with GalR mutant vs wild type pneumococci.

Chapter 2: Materials and Methods

2.1. Bacterial Strains and Growth Conditions

The bacterial strains used in this study are listed in **Table 2.1**. *S. pneumoniae* strains were routinely cultured in semi-synthetic casein-based liquid media (C+Y) and Chemically Defined Media (CDM) comprising RPMI 1640 medium (Sigma-Aldrich, St. Louis, USA) supplemented with amino acids (**Table 2.2a**), vitamins (**Table 2.2b**), choline and catalase as previously described (Kloosterman et al., 2006). Carbon sources were then added to CDM at a final concentration of 0.5%, as required. Bacteria were routinely plated on Columbia blood agar supplemented with 5% (vol/vol) defibrinated horse blood (BA) with or without antibiotics as required (**Table 2.3**) and incubated at 37 ^{o}C with 5% CO_2 overnight. For murine studies, *S. pneumoniae* strains were grown in Serum Broth (10g/l Peptone [Oxoid], 10 g/l Lab Lemco Powder [Oxoid], 5 g/l NaCl and 10% (v/v) heat-inactivated horse serum [Gibco, Auckland, New Zealand]) to an approximate concentration of 1 x 10^8 CFU/ml before being aliquoted and stored. All *S. pneumoniae* strains were routinely stored at -80 ^{o}C. The viable count of Serum Broth stocks was retrospectively confirmed by serial dilution and spread plating on BA.

2.2. Polymerase Chain Reaction

Polymerase Chain Reactions (PCR) were performed in a Mastercycler Flexilid Thermal Cycler (Eppendorf, NSW, Australia). Reactions were performed using 2 × Phusion Flash PCR Master Mix (Thermo Fisher Scientific, Victoria, Australia), as per the manufacturer's instructions. Primers used in these reactions are listed in **Table 2.4**. Reaction conditions included an initial denaturation at 98 ^{o}C for 5 minutes (min), followed by 35 cycles of denaturation at 98 ^{o}C for 5 seconds (sec), annealing for 30 sec and extension at 72 ^{o}C for 15 – 30 sec per kilobase of the expected PCR

Table 2.1. Bacterial strains used in this study

Strain	Description (Resistance Phenotype)	Source
D39	*S. pneumoniae* serotype 2, wild type	Avery et al., 1944
ΔgalR	*S. pneumoniae* serotype 2 with *galR* deletion (Ery^R)	Dr. Claudia Trappetti
ΔgalR Janus	*S. pneumoniae* serotype 2 with *rpsL1* allele swap and *galR* replacement with Janus cassette (Kan^R)	Dr. Claudia Trappetti
GalR$_{AAA}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317A, T319A and T323A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{ATT}$	*S. pneumoniae* serotype 2, with amino acid substitution S317A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{SAT}$	*S. pneumoniae* serotype 2, with amino acid substitution T319A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{STA}$	*S. pneumoniae* serotype 2, with amino acid substitution T323A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{AAT}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317A and T319A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{ATA}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317A and T323A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{SAA}$	*S. pneumoniae* serotype 2, with amino acid substitutions T319A and T323A in *galR* (Strep^R)	McLean et al., 2020
GalR$_{DDD}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317D, T319D and T323D in *galR* (Strep^R)	This Study
GalR$_{DTT}$	*S. pneumoniae* serotype 2, with amino acid substitution S317D in *galR* (Strep^R)	This Study
GalR$_{SDT}$	*S. pneumoniae* serotype 2, with amino acid substitution T319D in *galR* (Strep^R)	This Study
GalR$_{STD}$	*S. pneumoniae* serotype 2, with amino acid substitution T323D in *galR* (Strep^R)	This Study
GalR$_{DDT}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317D and T319D in *galR* (Strep^R)	This Study
GalR$_{DTD}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317D and T323D in *galR* (Strep^R)	This Study

GalR$_{SDD}$	*S. pneumoniae* serotype 2, with amino acid substitutions T319D and T323D in *galR* (Strep^R)	This Study
GalR$_{EEE}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317E, T319E and T323E in *galR* (Strep^R)	This Study
GalR$_{ETT}$	*S. pneumoniae* serotype 2, with amino acid substitution S317E in *galR* (Strep^R)	This Study
GalR$_{SET}$	*S. pneumoniae* serotype 2, with amino acid substitution T319E in *galR* (Strep^R)	This Study
GalR$_{STE}$	*S. pneumoniae* serotype 2, with amino acid substitution T323E in *galR* (Strep^R)	This Study
GalR$_{EET}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317E and T319E in *galR* (Strep^R)	This Study
GalR$_{ETE}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317E and T323E in *galR* (Strep^R)	This Study
GalR$_{SEE}$	*S. pneumoniae* serotype 2, with amino acid substitutions T319E and T323E in *galR* (Strep^R)	This Study
GalR$_{DEE}$	*S. pneumoniae* serotype 2, with amino acid substitution S317D, T317E and T323E in *galR* (Strep^R)	This Study
GalR$_{DAA}$	*S. pneumoniae* serotype 2, with amino acid substitution S317D, T317A and T323A in *galR* (Strep^R)	This Study
GalR$_{ADA}$	*S. pneumoniae* serotype 2, with amino acid substitution S317A, T319D and T323A in *galR* (Strep^R)	This Study
GalR$_{AAD}$	*S. pneumoniae* serotype 2, with amino acid substitution S317A, T319A, T323D in *galR* (Strep^R)	This Study
GalR$_{DDA}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317D and T319D and T323A in *galR* (Strep^R)	This Study
GalR$_{DAD}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317D, T319A and T323D in *galR* (Strep^R)	This Study
GalR$_{ADD}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317A, T319D and T323D in *galR* (Strep^R)	This Study
GalR$_{EAA}$	*S. pneumoniae* serotype 2, with amino acid substitution S317E, T317A and T323A in *galR* (Strep^R)	This Study
GalR$_{AEA}$	*S. pneumoniae* serotype 2, with amino acid substitution S317A, T319E and T323A in *galR* (Strep^R)	This Study

GalR$_{AAE}$	*S. pneumoniae* serotype 2, with amino acid substitution S317A, T319A, T323E in *galR* (Strep^R)	This Study
GalR$_{EEA}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317E and T319E and T323A in *galR* (Strep^R)	This Study
GalR$_{EAE}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317E, T319A and T323E in *galR* (Strep^R)	This Study
GalR$_{AEE}$	*S. pneumoniae* serotype 2, with amino acid substitutions S317A, T319E and T323E in *galR* (Strep^R)	This Study
Δ*galK*	*S. pneumoniae* serotype 2 with *galK* deletion (Spec^R)	McLean et al., 2020
Δ*lacAB*	*S. pneumoniae* serotype 2 with *lacAB* deletion (Spec^R)	McLean et al., 2020
Δ*lacD*	*S. pneumoniae* serotype 2 with *lacD* deletion (Spec^R)	McLean et al., 2020
Δ*lacD*	*S. pneumoniae* serotype 2 with *lacD* deletion (Ery^R)	McLean et al., 2020
Δ*galR*Δ*lacD*	*S. pneumoniae* serotype 2 with *galR* and *lacD* deletion (Ery^R, Spec^R)	McLean et al., 2020
Δ0088	*S. pneumoniae* serotype 2 with SPD_0088 deletion (Ery^R)	Dr. Claudia Trappetti
Δ*lacR2*	*S. pneumoniae* serotype 2 with *lacR2* deletion (Spec^R)	This Study

Table 2.2a. Amino acids added to Chemically Defined Media

Amino Acids	Concentration
Adenine	1.0 mg/ml
Alanine	4.8 mg/ml
Arginine	2.5 mg/ml
Asparagine	7.0 mg/ml
Aspartic Acid	6.0 mg/ml
Cysteine Hydrochloride	6.0 mg/ml
Glutamine	7.8 mg/ml
Glycine	3.5 mg/ml
Histidine	3.0 mg/ml
Isoleucine	4.3 mg/ml
Leucine	9.5 mg/ml
Lysine	8.8 mg/ml
Methionine	2.5 mg/ml
Phenylalanine	5.5 mg/ml
Proline	13.5 mg/ml
Serine	6.8 mg/ml
Threonine	4.5 mg/ml
Tryptophan	1.0 mg/ml
Uracil	1.0 mg/ml
Valine	6.5 mg/ml

Table 2.2b. Vitamins added to Chemically Defined Media

Vitamins	Concentration
Biotin	0.01 mg/ml
Niacinamide	0.1 mg/ml
Pantothenic Acid	0.1 mg/ml
Pyridoxal Hydrochloride	0.2 mg/ml
Thiamine	0.1 mg/ml

Table 2.3. Antibiotics used in this study

Antibiotic	Abbreviation	Working Concentration
Erythromycin	Ery	0.2 µg/ml
Gentamicin	Gent	40 µg/ml
Kanamycin	Kan	200 µg/ml
Spectinomycin	Spec	200 µg/ml
Streptomycin	Strep	200 µg/ml

Table 2.4. Mutagenesis oligonucleotides used in this study

Name	Sequence	Reference
GalR$_{AAA}$ F	ACTCCACGGTCGCAAAATTCCTGCGCTGGCCATGCTGGGAGCCAGACTGACATTAAGA	McLean et al., 2020
GalR$_{AAA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCATGGCCAGCGCAGGAATTTTGCGACCGTGGAGT	McLean et al., 2020
GalR$_{ATT}$ F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGACCATGCTGGGAACCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020
GalR$_{ATT}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGTCCCAGCATGGTCAGGGCAGGAATTTTGCGACCGGGAGTACT	McLean et al., 2020
GalR$_{SAT}$ F	AGTACTCCACGGTCGCAAAATTCCTAGCCTGGCCATGCTGGGAACCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020
GalR$_{SAT}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGTTCCCAGCATGGCCAGGCTAGGAATTTTGCGACCGTGGAGTACT	McLean et al., 2020
GalR$_{STA}$ F	AGTACTCCACGGTCGCAAAATTCCTAGCCTGACCATGCTGGGAGCCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020
GalR$_{STA}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGTCAGGCTAGGAATTTTGCGACCGTGGAGTACT	McLean et al., 2020
GalR$_{AAT}$ F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGGCCATGCTGGGAACCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020
GalR$_{AAT}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGTTCCCAGCATGGCCAGGGCAGGAATTTTGCGACCGTGGAGTACT	McLean et al., 2020
GalR$_{ATA}$ F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGACCATGCTGGGAGCCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020
GalR$_{ATA}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGTCAGGGCAGGAATTTTGCGACCGTGGAGTACT	McLean et al., 2020
GalR$_{SAA}$ F	AGTACTCCACGGTCGCAAAATTCCTAGCCTGGCCATGCTGGGAGCCAGACTGACATTAAGAGAAAGTACCC	McLean et al., 2020

GalR$_{SAA}$ R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGCCAGGCTAGGAATTTTGCGACCGTGGAGTACT	McLean et al., 2020
GalR$_{DDD}$ F	ACTCCACGGTCGCAAAATTCCTGACCTGGACATGCTGGGAGACAGACTGACATTAAGAG	This Study
GalR$_{DDD}$ R	CTCTTAATGTCAGTCTGTCTCCCAGCATGTCCAGGTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{DTT}$ F	ACTCCACGGTCGCAAAATTCCTGACCTGACCATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR$_{DTT}$ R	TCTTAATGTCAGTCTGGTTCCCAGCATGGTCAGGTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{SDT}$ F	ACTCCACGGTCGCAAAATTCCTAGCCTGGACATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR$_{SDT}$ R	TCTTAATGTCAGTCTGGTTCCCAGCATGTCCAGGCTAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{STD}$ F	ACTCCACGGTCGCAAAATTCCTAGCCTGACCATGCTGGGAGACAGACTGACATTAAGA	This Study
GalR$_{STD}$ R	TCTTAATGTCAGTCTGTCTCCCAGCATGGTCAGGCTAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{DDT}$ F	ACTCCACGGTCGCAAAATTCCTGACCTGGACATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR$_{DDT}$ R	TCTTAATGTCAGTCTGGTTCCCAGCATGTCCAGGTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{DTD}$ F	ACTCCACGGTCGCAAAATTCCTGACCTGACCATGCTGGGAGACAGACTGACATTAAGA	This Study
GalR$_{DTD}$ R	TCTTAATGTCAGTCTGTCTCCCAGCATGGTCAGGTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR$_{SDD}$ F	ACTCCACGGTCGCAAAATTCCTAGCCTGGACATGCTGGGAGACAGACTGACATTAAGA	This Study

GalR_{SDD} R	TCTTAATGTCAGTCTGTCTCCCAGCATGTCCAGGCTAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{EEE} F	ACTCCACGGTCGCAAAATTCCTGAACTGGAAATGCTGGGAGAAAGACTGACATTAAGA	This Study
GalR_{EEE} R	TCTTAATGTCAGTCTTTCTCCCAGCATTTCCAGTTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{ETT} F	ACTCCACGGTCGCAAAATTCCTGAACTGACCATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR_{ETT} R	TCTTAATGTCAGTCTGGTTCCCAGCATGGTCAGTTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{SET} F	ACTCCACGGTCGCAAAATTCCTAGCCTGGAAATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR_{SET} R	TCTTAATGTCAGTCTGGTTCCCAGCATTTCCAGGCTAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{STE} F	ACTCCACGGTCGCAAAATTCCTAGCCTGACCATGCTGGGAGAAAGACTGACATTAAGA	This Study
GalR_{STE} R	TCTTAATGTCAGTCTTTCTCCCAGCATGGTCAGGCTAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{EET} F	ACTCCACGGTCGCAAAATTCCTGAACTGGAAATGCTGGGAACCAGACTGACATTAAGA	This Study
GalR_{EET} R	TCTTAATGTCAGTCTGGTTCCCAGCATTTCCAGTTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{ETE} F	ACTCCACGGTCGCAAAATTCCTGAACTGACCATGCTGGGAGAAAGACTGACATTAAGA	This Study
GalR_{ETE} R	TCTTAATGTCAGTCTTTCTCCCAGCATGGTCAGTTCAGGAATTTTGCGACCGTGGAGT	This Study
GalR_{SEE} F	ACTCCACGGTCGCAAAATTCCTAGCCTGGAAATGCTGGGAGAAAGACTGACATTAAGA	This Study

GalR$_{SEE}$ R	TCTTAATGTCAGTCTTTCTCCCAGCATT TCCAGGCTAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{DEE}$ F	ACTCCACGGTCGCAAAATTCCTGACCT GGAAATGCTGGGAGAAAGACTGACATT AAGA	This Study
GalR$_{DEE}$ R	TCTTAATGTCAGTCTTTCTCCCAGCATT TCCAGGTCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{DAA}$ F	ACTCCACGGTCGCAAAATTCCTGACCT GGCCATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{DAA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCAT GGCCAGGTCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{ADA}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGACATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{ADA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCAT GTCCAGGGCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{AAD}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGCCATGCTGGGAGACAGACTGACATT AAGA	This Study
GalR$_{AAD}$ R	TCTTAATGTCAGTCTGTCTCCCAGCATG GCCAGGGCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{DDA}$ F	ACTCCACGGTCGCAAAATTCCTGACCT GGACATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{DDA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCAT GTCCAGGTCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{DAD}$ F	ACTCCACGGTCGCAAAATTCCTGACCT GGCCATGCTGGGAGACAGACTGACATT AAGA	This Study
GalR$_{DAD}$ R	TCTTAATGTCAGTCTGTCTCCCAGCATG GCCAGGTCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{ADD}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGACATGCTGGGAGACAGACTGACATT AAGA	This Study

GalR$_{ADD}$ R	TCTTAATGTCAGTCTGTCTCCCAGCATG TCCAGGGCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{EAA}$ F	ACTCCACGGTCGCAAAATTCCTGAACT GGCCATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{EAA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCAT GGCCAGTTCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{AEA}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGAAATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{AEA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCATT TCCAGGGCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{AAE}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGCCATGCTGGGAGAAAGACTGACATT AAGA	This Study
GalR$_{AAE}$ R	TCTTAATGTCAGTCTTTCTCCCAGCATG GCCAGGGCAGGAATTTTGCGACCGTG GAGT	This Study
GalR$_{EEA}$ F	ACTCCACGGTCGCAAAATTCCTGAACT GGAAATGCTGGGAGCCAGACTGACATT AAGA	This Study
GalR$_{EEA}$ R	TCTTAATGTCAGTCTGGCTCCCAGCATT TCCAGTTCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{EAE}$ F	ACTCCACGGTCGCAAAATTCCTGAACT GGCCATGCTGGGAGAAAGACTGACATT AAGA	This Study
GalR$_{EAE}$ R	TCTTAATGTCAGTCTTTCTCCCAGCATG GCCAGTTCAGGAATTTTGCGACCGTGG AGT	This Study
GalR$_{AEE}$ F	ACTCCACGGTCGCAAAATTCCTGCCCT GGAAATGCTGGGAGAAAGACTGACATT AAGA	This Study
GalR$_{AEE}$ R	TCTTAATGTCAGTCTTTCTCCCAGCATT TCCAGGGCAGGAATTTTGCGACCGTGG AGT	This Study
Janus F	CCGTTTGATTTTTAATGGATAATG	Sung et al., 2001
Janus R	AGAGACCTGGGCCCCTTTCC	Sung et al., 2001

Δ*galR* F	AAGACAAGCCAGAACCATTTGGG	McLean et al., 2020
Δ*galR* Ery F	CGGGAGGAAATAATTCTATGAGGAAAG TACCCTAAATCAAGAATAG	McLean et al., 2020
Δ*galR* Ery R	TTGTTCATGTAATCACTCCTTCTGTGCA ATGTCTTTTAAGGTAGCC	McLean et al., 2020
Δ*galR* R	GACAAGGTTGTTGTTATCGGTGAT	McLean et al., 2020
Δ*galK* F	CCTTCATTAAGTCATAGCCAGA	McLean et al., 2020
Δ*galK* Spec F	AAATAACAGATTGAAGAAGGTATAAGAA GTAGTTGGATACGCTCC	McLean et al., 2020
Δ*galK* Spec R	TATGTATTCATATATATCCTCCTCCAAAT TGATACGACCTGGTG	McLean et al., 2020
Δ*galK* R	GTAGGACAGACATTGGCCA	McLean et al., 2020
Δ*lacAB* F	AATATCGGACAAGCTGGT	McLean et al., 2020
Δ*lacAB* Spec F	AAATAACAGATTGAAGAAGGTATAAGCT CAACAAACAGACGC	McLean et al., 2020
Δ*lacAB* Spec R	TATGTATTCATATATATCCTCCTCATCTC AAACCTGCAGCATC	McLean et al., 2020
Δ*lacAB* R	TTCGTCTGAGCTATCTACATC	McLean et al., 2020
Δ*lacD* F	CTGTATCGCTATTCTCCACG	McLean et al., 2020
Δ*lacD* Ery F	CGGGAGGAAATAATTCTATGAGGGTAT CATCTCAGCTCTTGC	McLean et al., 2020
Δ*lacD* Ery R	TTGTTCATGTAATCACTCCTTCGCAAGA GCTGAGATGATACC	McLean et al., 2020
Δ*lacD* Spec F	AAATAACAGATTGAAGAAGGTATAATGA AGCAGCAGCTCGCGAAT	McLean et al., 2020
Δ*lacD* Spec R	TATGTATTCATATATATCCTCCTCGCAA GAGCTGAGATGATACC	McLean et al., 2020

Δ*lacD* R	TGATCTGCTAGCTTCTGAC	McLean et al., 2020
Δ*0088* F	AAGATGCCATTGCTTCCCTA	Dr. Claudia Trappetti
Δ*0088* Ery F	CGGGAGGAAATAATTCTATGAGTTCGT GAAGAAACAGTAGGAC	Dr. Claudia Trappetti
Δ*0088* Ery R	TTGTTCATGTAATCACTCCTTCGAGTCA ACTGAGGAATGGTTA	Dr. Claudia Trappetti
Δ0088 R	GTAGCGAACAAACATCTTCTT	Dr. Claudia Trappetti
Δ*lacR2* F	GTCATCGATGTTTACGCTGC	This Study
Δ*lacR2* Spec F	AAATAACAGATTGAAGAAGGTATAAGCT ATCTAAACACATTA	This Study
Δ*lacR2* Spec R	TATGTATTCATATATATCCTCCTCCGTC CATAATCTGTTTAACTG	This Study
Δ*lacR2* R	CAATCGATATCCAGTTGACCTC	This Study
Spectinomycin F (J253)	GAGGAGGATATATATGAATACATACG	Trappetti et al., 2017
Spectinomycin R (J254)	TTATACCTTCTTCAATCTGTTATTTAAAT AGTTTATAGTTA	Trappetti et al., 2017
Erythromycin F	GAAGGAGTGATTACATGAACAA	Trappetti et al., 2017
Erythromycin R	ATCATAGAATTATTTCCTCCCG	Trappetti et al., 2017
galR **Seq F**	AATCTATCATGATGAACTGGTC	McLean et al., 2020
galR **Seq R**	CATAATGGAGGGCGTATGG	McLean et al., 2020

product. The annealing temperature was routinely modified depending on × the Tm^o of the primer set to be used. PCR products were analysed by gel electrophoresis using an 0.8% (w/v) agarose gel (Agarose Low EEO; AppliChem, Germany) in 1 × TAE buffer (40 mM Tris, 20 mM Acetic Acid, 1 mM ethylenediaminetetraacetic acid [EDTA]) supplemented with 0.5 µl/10 ml RedSafe Nucleic Acid Solution (Intron Biotechnology) To prepare samples for loading, PCR products were mixed with 1/10th volume of 10 × Loading Buffer (15% [w/v] Ficoll, 0.1% [w/v] bromophenol dye, 100 ng/ml RNAse A). Gels were electrophoresed in 1 × TAE buffer at 180 V for 45 min. DNA bands were visualised via transillumination with short wavelength ultraviolet light using a Gel/Chemi Doc XR system (Bio-Rad, New South Wales, Australia) and analysed using Quantity One v 4.6.9 software (Bio-Rad, New South Wales, Australia).

2.3. Transformation and Mutagenesis of *S. pneumoniae*

To generate *S. pneumoniae* mutant strains, cells were made competent and transformed with linear DNA as previously described (Iannelli and Pozzi, 2004, Minhas et al., 2019). Briefly, PCR was firstly utilised to generate linear DNA containing the desired mutation – typically replacement of the gene of interest with an antibiotic resistance cassette - with flanking regions homologous to the 5' and 3' regions of the gene of interest. Strains targeted for mutagenesis were then grown overnight on BA at 37 °C. The following morning, pneumococci were taken from these plates and grown in C+Y to an OD_{600} of 0.25. Strains were then diluted 1:10 in fresh C+Y media supplemented with 10 µg/ml Competence Stimulating Peptide 1 (CSP-1) and incubated at 37 °C for 15 min. Following this incubation, the PCR product described above was added to the competent cells and left to incubate at 37 °C for 3 hours (h). Cells were then pelleted via centrifugation at 13,000 × *g* for 5

min before removing the supernatant and resuspending the cells in 50 µl fresh C+Y media. The resuspended cells were then spread plated on BA supplemented with the required antibiotic for selection of mutants and incubated overnight at 37 °C. Successfully transformed colonies were subsequently selected and grown again on BA supplemented with required antibiotic overnight at 37 °C to generate sufficient biomass for long-term storage at -80 °C. Mutants were then confirmed using PCR.

2.4. Mutagenesis Utilising the Janus Cassette System

In addition to gene deletion mutants (Section 2.3), there was a need for point-amino acid mutants within the gene *galR* in this study. The substitutions were achieved via allelic exchange mutagenesis using the Janus cassette system (Harvey et al., 2014, Sung et al., 2001). This is a three-step process, in which the endogenous *rspL* gene conferring streptomycin sensitivity, was first replaced with the streptomycin-resistant *rpsL1* allele. The Janus cassette, containing a kanamycin resistance marker and a dominant counter-selectable $rpsL^+$ marker, was then used to replace the gene of interest by direct transformation with a linear PCR product comprising the Janus cassette flanked by 5' and 3' sequences homologous to the regions directly 5' and 3' of *galR*. This strain was denoted as Δ*galR* Janus (Kan^R, $Strep^S$). The final step in this process was to perform direct transformation of the Δ*galR* Janus strain with a linear PCR product comprising the *galR* gene containing the desired amino acid substitutions, flanked by 5' and 3' regions homologous to the regions directly 5' and 3' of *galR*. Successful mutants were counter selected on Strep, confirming loss of the Janus cassette and reinstates the $Strep^R$ phenotype conferred by the initial mutation of the *rpsL* gene. Mutants were confirmed via PCR and Sanger Sequencing.

2.5. Sanger Sequencing

All Sanger Sequencing was performed through the Australian Genome Research Facility (AGRF). To prepare samples for sequencing, PCR reactions were performed using the GalR F and GalR R primers listed in **Table 2.4**. Successfully amplified PCR reactions were then purified using the QIAquick PCR Purification Kit (Qiagen, Hilden, Germany) as per the manufacturer's instructions. Purified PCR products were assessed for total DNA content and quality using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Victoria, Australia). Samples were then diluted in RNase/DNase-free water to a final concentration of 25 ng/μl. In an Eppendorf tube, 2 μl of PCR template was added to 9 μl of water with 1 μl of either the GalR Seq F or GalR Seq R primer, which was then sent for sequencing.

2.6. Structural Modelling of GalR

The GalR amino acid sequence (SPD_1635) was obtained from the NCBI database and input into SWISS-MODEL (Biasini et al., 2014). A homology model was generated based on the 2.4-Å structure (PDB: 1JFS) of the *Escherichia coli* purine repressor (PurR) W147F mutant (Huffman et al., 2002). The cartoon representation of the GalR homology model and the aligned PurR template were generated in PyMOL version 2.3.3 (Schrödinger). The root mean square deviation (RMSD) between the GalR model and the PurR template was determined by alignment in PyMOL. The DNA binding domain and putative sugar binding residues were identified by the NCBI conserved domain search (Lu et al., 2020), and the locations of the putative phosphorylated residues were determined based on the phosphor-proteomic findings (Sun et al., 2010).

2.7. Growth Assay

For growth assays, *S. pneumoniae* strains were firstly grown overnight on plain BA at 37 °C. The following morning, the strains were used to inoculate CDM supplemented with either 0.5% Glucose (Glc) (CDM + Glc), 0.5% Galactose (Gal) (CDM + Gal) or no sugar (CDM) to an OD_{600} of 0.05. A final volume of 200 µl of inoculated culture was added to a Costar 96-well flat-bottom cell culture plate in technical triplicate (Corning Incorporated, New York, USA) and incubated at 37 °C with 5% CO_2 for 18 – 24 h. OD_{600} readings were taken at 30 min intervals in either a SPECTROstar Omega Spectrophotometer or FLUOstar Omega Spectrophotometer (BMG Labtech). Data were visualised using GraphPad Prism v 8.3.0 (GraphPad Software, San Diego, California).

2.8. RNA Extraction

For RNA extraction, *S. pneumoniae* were initially plated on plain BA, incubated overnight and then resuspended to a final OD_{600} of 0.25 in either CDM + Glc or CDM + Gal. Strains were then incubated for 30 min at 37 °C with 5% CO_2 before adding 500 µl Bacterial RNA Protect (Qiagen, Hilden, Germany). Cells were then incubated at room temperature for 5 min prior to pelleting via centrifugation at 13,000 × *g*. The supernatant was then removed, and the pellets then used to perform an RNA extraction using the RNeasy Extraction Kit (Qiagen, Hilden, Germany) with on-column DNase treatment as per the manufacturer's instruction.

2.9. Quantitative Real Rime RT-PCR

Real-time PCR was performed using a SYBR Green One Step qRT-PCR kit (Thermo Fisher Scientific, Victoria, Australia) according to the manufacturer's instructions, with the oligonucleotides listed in **Table 2.5.** Each reaction was

Table 2.5. qRT-PCR oligonucleotides used in this study

Name	Sequence	Reference
gyrA RT F	ACTGGTATCGCGGTTGGGAT	Minhas et al., 2019
gyrA RT R	ACCTGATTTCCCCATGCAA	Minhas et al., 2019
galK RT F	CACGTTTCTCTGGAGCATGA	Trappetti et al., 2017
galK RT R	ATGGCACAGCCACTAAAACC	Trappetti et al., 2017
galR RT F	TCTCTATCGCCGACCGTATCC	Trappetti et al., 2017
galR RT R	GGTGTAGCCCAGCTCTTCAG	Trappetti et al., 2017
galT RT F	GTGGGAGAAGGTGTTTTGGA	Trappetti et al., 2017
galT RT R	ACGCGCAGTCTGACTATCCT	Trappetti et al., 2017
lacAB RT F	CGTGATTGATGCTTATGGAG	McLean et al., 2020
lacAB RT R	AGCCAATTCATCACCAACAAG	McLean et al., 2020
lacD RT F	CATCGGTTCTGAGTGTGTGG	Trappetti et al., 2017
lacD RT R	AAAGCGTGGGTCTGAAAAGA	Trappetti et al., 2017
adhA RT F	TGTCGCACCTGACTCCATAG	Dr. Erin Brazel
adhA RT R	TGTTCAAAAAGGGGACAAGG	Dr. Erin Brazel
phtE RT F	AGCACCTCAAGGAAATGGTG	Dr. Erin Brazel
phtE RT R	TAGGGTCACTCCCCACATTC	Dr. Erin Brazel
SPD_1774 RT F	GTGCATTCGACAGAAAGC	This Study
SPD_1774 RT R	CCATAGCCCAAGTGTCTG	This Study
strH RT F	CAACCGACCATCCATACG	This Study
strH RT R	GTTGAACTAACTTCCTCTTGTGG	This Study
blpB RT F	TGCGTTTATTCTCATTATTTCAAC	This Study
blpB RT R	TCCAGCATTTACCATATCTCC	This Study
lacR2 RT F	GCTCAAACAAGAAAAATTAGCC	This Study
lacR2 RT R	GCAAATCTTTATCAGCTAATTCCTG	This Study

performed in a final volume of 10 µl containing PCR-grade water, SYBR Green Master Mix, forward and reverse primers, Taq polymerase and template RNA. All primers were used at a final concentration of 200 nM per reaction. Primers specific for the housekeeping gene Gyrase A (*gyrA*) were used as an internal control. Amplification was performed using a LightCycler 480 II cycler (Roche, Switzerland), including an initial denaturation at 95 °C for 2 min, followed by 35 cycles of denaturation at 95 °C for 15 sec, annealing at 60 °C for 30 sec and extension at 72 °C for 30 sec (Minhas et al., 2020). Data were analysed using the comparative critical threshold cycle ($2^{-\Delta\Delta CT}$) method (Livak and Schmittgen, 2001). Assays were performed in triplicate over a minimum of two independent experiments. Statistical analyses were performed using two-tailed Student's t test; P values of <0.05 were deemed statistically significant.

2.10. Murine Infection Model

All animal experiments were approved by the University of Adelaide Animal Ethics Committee. Female outbred 4- to 6-week-old CD-1 (Swiss) mice were anaesthetized by intraperitoneal injection with ketamine (8 mg/ml) and xylazine (0.8 mg/ml) at a dose of 10 µl/g body weight. Anaesthetised mice were then intranasally inoculated with 1×10^7 CFU of the required *S. pneumoniae* strain in a total of 50 µl, as previously described (McLean et al., 2020). The challenge dose was retrospectively confirmed by serial dilution and plating on BA. At 24 h post-infection, mice were euthanized by CO_2 asphyxiation before harvesting the blood, lungs, nasal tissue, brain and ear tissue. Pneumococci were enumerated from homogenised tissue as described previously by serial dilution and plating on BA supplemented with 40 µg/ml gentamicin (McLean et al., 2020). Statistical analyses

of log-transformed CFU data were performed using two-tailed Student's t test; P values of <0.05 were deemed statistically significant.

2.11. Adherence Assays

To assess adherence of *S. pneumoniae*, Detroit 562 human nasopharyngeal carcinoma cells were used. Cells were initially cultured in Roswell Park Memorial Institute (RPMI) media supplemented with 10 % Fetal Calf Serum (FCS), 0.2 U/ml Penicillin and 100 μg/ml Streptomycin to confluency. Once confluent, cells were detached by treatment with 0.05% Trypsin, following which 2 x 10^5 cells per well were seeded into a 24 well micro plate and again grown to confluency. The day prior to performing the adherence assay, the required *S. pneumoniae* strains were streaked on plain BA and grown overnight at 37 °C with 5% CO_2. The next morning, the Detroit cells were washed twice with PBS to remove any remaining antibiotic, then seeded with 500 μl of *S. pneumoniae* at OD_{600} of 0.25 in CDM + 0.5% Gal. Detroit cells and *S. pneumoniae* were then incubated at 37°C with 5% CO_2 for 2 h. Following this incubation, the medium was removed, and the cells then washed three times in PBS. To each well, 100 μl of 0.25% Trypsin – 0.02% EDTA and 400 μl of 0.025% Triton X-100 was added and vigorously pipetted up and down to detach cells from the plate. Cells were then serially diluted and spot plated on BA and incubated overnight at 37 °C with 5% CO_2, with single colonies then enumerated the following day (Trappetti et al., 2011). Assays were performed in triplicate over a minimum of two independent experiments. Statistical analyses were performed using two-tailed Student's t test; P values of <0.05 were deemed statistically significant.

2.12. Uronic Acid Assay

To assess total capsule, *S. pneumoniae* strains were first grown overnight on plain BA at 37 °C with 5% CO_2. The following morning, strains were grown to mid-log phase in Serum Broth before being concentrated to a final OD_{600} of 0.5. Cells were then pelleted via centrifugation at 3,273 × *g* for 20 min and the supernatant removed. Pelleted cells were resuspended in 500 µl of 150 mM Tris and 1 mM $MgSO_4$ at pH 7.0. Cells were then lysed by addition of 0.1% DOC and incubated at 37 °C for 30 min. DNase I and RNase A were then added to the samples at a final concentration of 100 µg/ml, along with 100 U of mutanolysin and subsequently incubated overnight at 37 °C. The following morning, samples were treated with 100 µg/ml of Proteinase K and incubated at 56 °C for 4 h. Samples were diluted two-fold in duplicate to a final volume of 100 µl in ultrapure water, including 100 µl of water as a blank. To each sample, 600 µl of sodium tetraborate solution in concentrated sulphuric acid (12.5 mM $Na_2B_4O_4$ in concentrated H_2SO_4) was added prior to vortexing and then heated at 100 °C for 5 min. The samples were then cooled on ice prior to adding 10 µl of either 0.5% NaOH or a solution of 0.15% 3-phenylphenol in 0.5% NaOH to each of the duplicate tubes. Tubes were then immediately shaken before transferring 200 µl to a 96-well microtitre tray and reading Absorbance at 520 nm (A_{520}) using a Spectramax M2 spectrometer (Molecular Devices, California, USA) (Morona et al., 2006, Trappetti et al., 2017).

2.13. Metabolomics Sample Preparation

For metabolomics, *S. pneumoniae* strains were firstly grown overnight on plain BA at 37 °C. The following morning, these strains were used to inoculate CDM + 0.5% Glu or CDM + 0.5% Gal to an OD_{600} of 0.25 and incubated at 37 °C for 30 min. Following incubation, 10 ml of *S. pneumoniae* culture at OD_{600} 0.25 was infused into

30 ml of chilled PBS and incubated on ice for 5 min. Cell suspensions were then centrifuged at 8,000 × g for 10 min at 0 °C. The supernatant was removed and the remaining cells resuspended in 1 ml chilled PBS and transferred to a 1.5 ml Eppendorf tube. Cells were then washed with 1 ml chilled PBS twice more before flash freezing the pellet in liquid nitrogen and storing at -80 °C. Pellets were then sent to Dr. David De Souza at Metabolomics Australia (University of Melbourne) for further analysis.

2.14. RNA Extraction of Infected Murine Lung Tissue

Groups of 12 outbred 4 – 6 week old female Swiss (CD-1) mice were anaesthetised and challenged as described in Section 2.10. The challenge dose was retrospectively confirmed by serial dilution and spread plating on BA. Mice were euthanised by CO_2 asphyxiation at 24 h post-infection and the lungs perfused with 1 × PBS. Once removed, the lungs were placed in 1 ml of TRIzol reagent (Thermo Fisher Scientific, Victoria, Australia) in screw-cap Eppendorf tubes. Tissue was then homogenised using the Precellys® 24 tissue homogeniser (Bertin Technologies, France). The homogenised tissues were then transferred to fresh Eppendorf tubes before pelleting the cell material, removing the supernatant and resuspending the pellet in 200 μl of TRIzol. This was then transferred to a tube containing 40 - 50 mg of acid-washed 0.1 μM glass beads (Sigma-Aldrich, St. Louis, USA). The tissue was then homogenised for an additional three cycles of 30 sec at 5,000 × g with 30 sec breaks in between. After homogenising, the tubes were left at room temperature for 5 min to allow for dissociation of nucleoprotein complexes before centrifuging for 1 min at 10,000 × g to pellet the glass beads. The supernatant was then transferred to a fresh Safe-Lock Eppendorf tube containing 300 μl of chloroform and shaken vigorously by hand for 15 sec. These tubes were then left at room temperature for

2 - 3 min before centrifuging at 12,000 × g for 15 min at 4 ℃. Following this step, approximately 400 µl of the aqueous phase was transferred to a fresh Safe-Lock Eppendorf before adding an equal volume of 70% ethanol. These tubes were then vortexed for 10 sec before transferring 700 µl of the sample to a Qiagen RNeasy Mini Kit Spin Column (Minhas et al., 2020). From here, RNA extraction proceeded as per the Qiagen RNeasy Mini Kit directions. Prior to sending samples for RNA sequencing, extracted RNA was then pooled (4 mice lungs per replicate) and assessed for total RNA quantity using a Nanodrop spectrophotometer (Thermo Fisher Scientific, Victoria, Australia).

2.15. RNA Library Preparation and Sequencing

RNA Library Preparation and Sequencing was performed at the Australian Cancer Research Foundation (ACRF) Cancer Genomics Facility. RNA submitted for sequencing was checked for quality using a Bioanalyser (Agilent Technologies, Santa Clara, USA) as an initial step in the QC process. rRNA depletion was performed using the Qiagen QIAseq FastSelect rRNA HMR Removal Kit (Qiagen, Hilden, Germany) to remove murine rRNA and the NEBNext rRNA Depletion Kit (Bacteria) (New England Biolabs, Massachusetts, USA) for bacterial rRNA depletion, as per the manufacturer's instruction. Stranded cDNA synthesis and library preparation was performed using the KAPA RNA HyperPrep Kit (Roche, Switzerland) as per the manufacturer's instructions. Sequencing was performed on an Illumina NextSeq High Output Flowcell in 150 bp paired end mode. Libraries were de-multiplexed and subject to further in-house analyses.

2.16. Bioinformatic Analysis

The quality of raw libraries was first checked using FastQC (v0.11.8, Babraham Bioinformatics, UK). Read quality was further enhanced be trimming adaptor sequences using Trimmomatic v0.38 (Bolger et al., 2014). The quality of trimmed reads was again checked using FastQC before proceeding with downstream analyses. The published *Streptococcus pneumoniae* D39 (Assembly: GCA_000014365.2) and *Mus musculus* genomes (Assembly: GRCm38.p6) were used as reference genomes in this project. Trimmed reads were aligned to these genomes to determine the number of reads that successfully mapped back to the reference genomes using Bowtie2 (Langmead and Salzberg, 2012). SAMtools (Li et al., 2009) was used to generate sorted BAM files that could be used for downstream differential gene expression analyses. To determine which genes were differentially expressed in the pneumococcal genome, the Galaxy Australia platform provided by the Research Computing Centre at the University of Queensland was used. Specifically, Degust (Powell, 2019) was used to quantify differentially expressed genes. To determine which genes were differentially expressed in the murine genome, R (R Core Team, 2021) was used. Specifically, Rsubread (Liao et al., 2019) was used to import sorted BAM files into R and determine which genes were differentially expressed. Limma (Ritchie et al., 2015) was used to identify differentially expressed genes.

2.17. Flow Cytometry Analysis of Infected Murine Lung Tissue

Flow cytometric analysis of infected murine tissue was performed essentially as previously described (David et al., 2019). At 24 h post-pneumococcal challenge with 1×10^7 CFU of either the D39, $\Delta galR$ or D39$_{AAA}$ strain, groups of 8 outbred 4-6 week old female Swiss (CD-1) mice were euthanised by CO_2 asphyxiation and the lungs

perfused with PBS. Lungs were then removed and finely macerated in 1 ml of pre-warmed digestion media (Dulbecco's Modified Eagle Medium (DMEM) [Gibco], 5% FCS, 10 mM HEPES, 2.5 mM $CaCl_2$, 0.2 U/ml Penicillin, 50 µg/ml Gentamicin, 1 mg/ml Collagenase 1A [Sigma Aldrich] and 30 U/ml DNase [Sigma Aldrich]). Lungs were incubated for a total of 1 h with mixing every 20 min. Following incubation, macerated lung tissue was filtered through a 70 µM filter (Becton, Dickson and Company, New Jersey, USA) to ensure a single cell suspension. Tissue samples then underwent red cell lysis by incubation in red cell lysis buffer (155 mM NH_4Cl and 170 mM Tris-HCl) for 5 min at 37 °C. Cells were then washed in PBS and stored on ice while awaiting cell counting. Single cell suspensions were added to 96-well U-bottom trays at 1 x 10^6 cells per well and pelleted by centrifugation at 400 × g for 2 min. Cells were then washed once more in PBS before proceeding with staining.

2.17.1. Intracellular Cytokine Staining

For intracellular cytokine staining, cells were resuspended in 50 µl pre-warmed Iscove's Modified Dulbecco's Medium (IMDM) [Gibco, Auckland, New Zealand] supplemented with 50 ng/ml Phorbol Myristate Acetate (PMA), 1 µg/ml ionomycin and GolgiStop (Becton, Dickinson and Company, New Jersey, USA) diluted 1:1500 and incubated at 4 °C for 4 h. Following this incubation, cells were pelleted and resuspended in 85 µl of Cytofix/Cytoperm solution (Becton, Dickinson and Company, New Jersey, USA) before incubating for a further 20 min at 4 °C. Cells were then washed twice in Permwash buffer (Becton, Dickinson and Company, New Jersey, USA) before adding antibodies as listed in **Table 2.6** at pre-determined concentrations. Cells were incubated in the presence of antibody for 30 min at 4 °C before washing again in Permwash buffer, followed by PBS. Cells were then fixed in 200 µl PBS + 1% PFA.

2.17.2 Extracellular Surface Marker Staining

For extracellular surface marker staining, cells were washed once in FACS buffer (PBS + 1% BSA + 0.04% sodium azide) before pelleting and resuspending in 50 µl per well Mouse δ-Globulin with 1:1000 Live/Dead Stain (Becton, Dickinson and Company, New Jersey, USA). Cells were then incubated for 15 min in the dark at room temperature. Following this incubation, cells were washed once in FACS buffer before being resuspended in FACS buffer containing the antibodies of interest as listed in **Table 2.6** at pre-determined concentrations. Cells were incubated in the presence of antibody for 15 min at room temperature in the dark before washing twice in PBS + 0.04% sodium azide. Cells were then fixed in 200 µl PBS + 1% PFA. Data acquisition for all fixed cells was performed on a BD LSRFortessa X-20 flow cytometer and analysed using FlowJo™ v10.7.2 Software (Becton, Dickinson and Company, New Jersey, USA). Statistical analyses were performed using two-tailed Student's t test; P values of <0.05 were deemed statistically significant.

2.18. Phenotypic Microarray

Caron source microarray was performed using the Phenotypic Microarray (PM) plates PM1 and PM2 (Biolog, Inc., Hayward, California). These microplates test for the catabolism of 190 different carbon sources (Appendix A), with each well of the microarray containing a different carbon source. In short, cells were resuspended in the provided buffer according to manufacturer's instructions to an Absorbance 590 nm (A_{590}) of 0.37. Following this, 100 µl of the cell suspension was added to each well and the A_{590} measured following 18 h incubation at 37 °C in a SpectraMax M2 Microplate Reader (Molecular Devices, California, USA). Catabolism was measured as a result of the reduction of tetrazolium dye by NADH, with blank corrected A_{590} values greater than 0.65 being considered positive.

Table 2.6. Antibodies used in this study

Antibody	Company	Identifier
BUV395 Rat α-Mouse CD4 (Clone GK1.5)	BD Biosciences	Cat# 563790
BUV737 Rat α-Mouse CD8a (Clone 52-6.7)	BD Biosciences	Cat# 612759
BV421 Rat α-Mouse GM-CSF (Clone MP1-22E9)	BD Biosciences	Cat# 564747
Alexa Fluor® 488 Rat α-Mouse IGN-γ (Clone XMG1.2)	BD Biosciences	Cat# 557724
PE Rat α-Mouse IL-17A (Clone TC11-18H10)	BD Biosciences	Cat# 559502
PE-Cy™7 Rat α-Mouse CD45R/B220 (Clone RA3-6B2)	BD Biosciences	Cat# 552772
APC Hamster α-Mouse CD3e (Clone 145-2C11)	BD Biosciences	Cat# 553066
BV786 Hamster α-Mouse CD11c (Clone HL3)	BD Biosciences	Cat# 563735
BV711 Rat α-Mouse CD24 (Clone M1/69)	BD Biosciences	Cat# 563450
BUV396 Rat α-Mouse Ly-6G (Clone 1A8)	BD Biosciences	Cat# 563978
BV650 Rat α-Mouse I-A/I-E (Clone M5/114.15.2)	BD Biosciences	Cat# 563415
PE α-Mouse/Human CD11b (Clone M1/70)	BioLegend	Cat# 101208
FITC α-Mouse CD45 (Clone 30-F11)	BioLegend	Cat# 103107
BV421 α-Mouse CD64 (Clone X54-5/7.1)	BioLegend	Cat# 139309
PerCP/Cyanine5.5 α-Mouse Ly-6C (Clone Hk1.4)	BioLegend	Cat# 128011

Chapter 3: Site-specific mutations of GalR affect galactose metabolism in *Streptococcus pneumoniae*

3.1. Introduction

This chapter includes data published in the Journal of Bacteriology (McLean et al., 2020), which focuses on examination of the regulation of Gal metabolism in *S. pneumoniae* by mutation of the Leloir pathway regulator, GalR. As reported in Chapter 1, the pneumococcus is a human-adapted bacterium often carried asymptomatically in the nasopharynx. Colonisation of the URT is an essential prerequisite for invasive disease; however, the availability of carbon sources is scarce in this environment (Pezzulo et al., 2011). Gal is the most abundant sugar in the URT (Paixão et al., 2015), so it follows that the ability to metabolise it may be a crucial factor for pneumococcal survival within this host niche.

The pneumococcus possesses two pathways for Gal metabolism, the Leloir pathway and the tagatose 6-phosphate (T6P) pathway. The Leloir pathway will be the main focus of this chapter. In the Leloir pathway, as shown in **Figure 1.5**, Gal is proposed to enter the cell via the ABC transporter SPD_0088 (Bidossi et al., 2012) and is converted to αGal by GalM. αGal is then phosphorylated intracellularly at the C1 position by the kinase GalK to yield αGal1P. This is then converted into αGlc1P by the combined action of the hexose 1-phosphate uridyltransferase GalT and the UDP-Glc epimerase GalE, with simultaneous conversion of UDP-Glc to UDP-Gal. From here, αGlc1P can be converted to Glc6P by phosphoglucomutase (Pgm) and enter the glycolytic pathway, or be converted by GalU to UDP-Glc, the precursor of activated nucleotide sugars required for CPS biosynthesis (Paixão et al., 2015). Regulation of this pathway, specifically *galK* and *galT*, is controlled by GalR, a member of the GalR/LacI family of transcriptional regulators. A previous study (Afzal et al., 2015) showed that GalR is the transcriptional activator of the Leloir pathway. An additiona study has indicated that GalR possesses three putative phosphorylation sites: S317, T319 and T323 (Sun et al., 2010).

As explored in Section 1.10, a previous study from our group showed that the QS signalling molecule AI-2 promotes Gal metabolism by upregulation of *galKT* via the regulator *galR* (Trappetti et al., 2017). This study also showed that AI-2 promotes transition of the pneumococcus from harmless coloniser to invasive pathogen. AI-2 is generated as a by-product of the activated methyl cycle in *S. pneumoniae* and is dependent on the metabolic enzyme LuxS, a S-ribosylhomocysteine lyase.

We therefore propose that the quorum sensing molecule AI-2 is able to (directly or indirectly) promote phosphorylation of the three putative phosphorylation sites GalR, resulting in upregulation of the Leloir pathway. The mechanisms behind GalR function, the role of the putative phosphorylation sites reported by Sun *et al* (Sun et al., 2010) and the extent of its role in regulation of Gal metabolism remains poorly understood. In this chapter, site-specific mutagenesis techniques have been used to selectively mutate these putative phosphorylation sites, either singly or in combination, to better elucidate their role in pneumococcal Gal metabolism.

3.2. Results

3.2.1. The location of the putative GalR phosphorylation sites

There is currently no published crystal structure available for GalR, but structures of closely related proteins have been solved. In preparation for future site-directed mutagenesis studies, and in the absence of structural information for GalR, we sought to examine the location of the putative phosphorylated residues (S317, T319, and T323) by generating a structural homology model. A homology model of GalR was constructed using SWISS-MODEL based on the homodimeric 2.4-Å structure (PDB: 1JFS) of the *Escherichia coli* PurR W147F mutant (35% sequence similarity, 93% sequence coverage) (Huffman et al., 2002). Alignment of the GalR model (green) with the template (cyan) revealed a moderate level of variation (root mean square deviation [RMSD], 2.868 Å), with an additional loop present in the GalR model that was absent in PurR, corresponding to residues 183 to 191 (**Fig. 3.1A**). To complement these studies, we performed a conserved domain search to investigate whether any of the putative phosphorylated residues were located within regions of possible functional importance (**Fig. 3.1B**). The putative Gal binding residues (magenta) are situated at the proposed dimer interface of GalR, suggesting a role in protein dimer stabilisation upon sugar binding, while the N-terminal region of GalR harbors the helix-turn-helix domain (blue) responsible for the interaction with DNA. All of the putative phosphorylation sites (orange spheres) were situated in a region distinct from the residues proposed to be involved in Gal binding and DNA binding (**Fig. 3.1B**). As any functional impact of S317, T319, or T323 phosphorylation is more likely a consequence of allosteric changes rather than a direct impact on sugar or DNA binding, we investigated the contribution of each putative phosphorylated residue to GalR function, but not before gaining a greater understanding of the role of GalR itself in Gal metabolism.
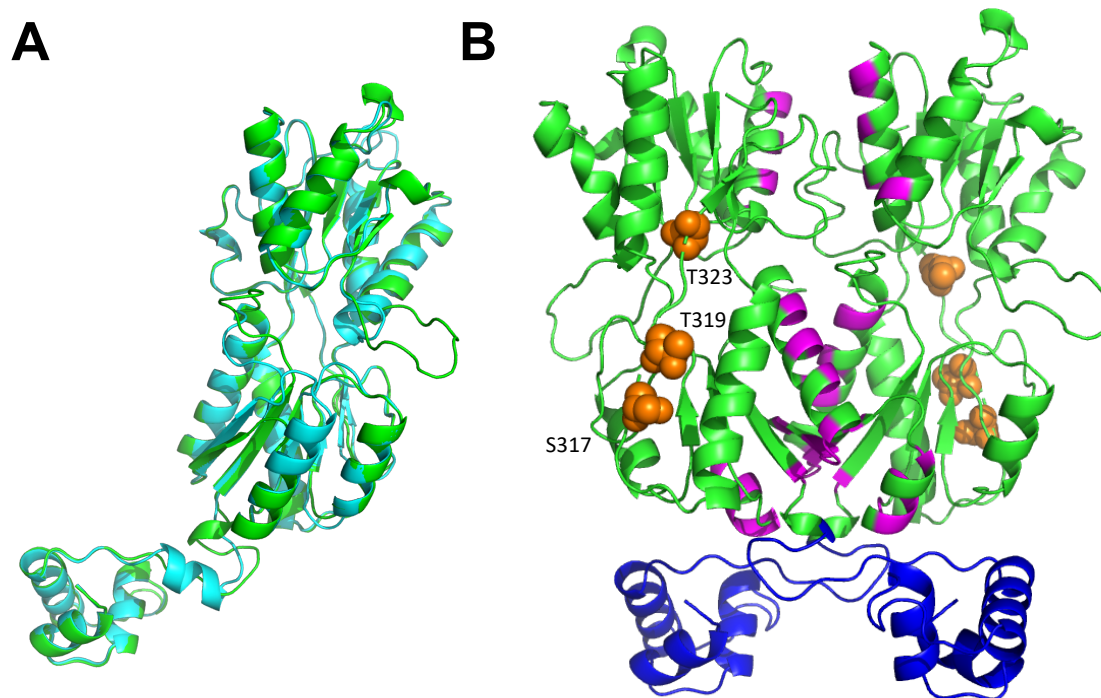
FIGURE 3.1 **Structural homology model of *S. pneumoniae* GalR**
(A) Cartoon representation of the protomeric homology model of *S. pneumoniae* GalR (green) based on the 2.9 Å structure of *Escherichia coli* PurR W147F mutant (cyan; RMSD: 2.868 Å). (B) Cartoon representation of the dimeric homology model of GalR. DNA binding helix-turn-helix domain is shown in blue and putative sugar binding regions are highlighted in magenta. The serine (S317) and threonine (T319 and T323) residues hypothesised to be phosphorylated are depicted as orange spheres.

3.2.2. Impact of GalR on galactose metabolism

While investigating the regulation of the Leloir pathway, Afzal *et al.,* (Afzal et al., 2015) determined that deletion of *galK* resulted in a decreased ability to grow in M17 medium with Gal as the primary carbon source. Given our interest in the function of GalR, we wanted to determine the impact that deletion of *galR* had on the ability to grow when Gal is the only carbon source present. To do this, we used a minimal chemically defined media (CDM) that will not permit growth of *S. pneumoniae* without the addition of a carbon source. First, we utilised a strain in which the *galR* gene was deleted and replaced with an erythromycin (Ery) resistance cassette, named Δ*galR* that had been previously generated (**Table 2.1**). Growth in chemically defined medium with Glc as the sole carbon source (CDM + Glc) (**Fig. 3.2A**) showed that the Δ*galR* mutant grew similarly to the D39 wild-type strain. When grown in chemically defined medium with Gal as the sole carbon source (CDM + Gal) **(Fig. 3.2B**), the Δ*galR* strain was unable to grow, thus indicating that GalR plays a key role in the ability to utilise Gal as a carbon source.

3.2.3. Impact of SPD_0088 deletion on import of galactose and galactose metabolism

Now knowing the importance of GalR in the metabolism of Gal, we wanted to better understand how Gal enters the cell in the first instance. The import of Gal remains poorly understood in the pneumococcus; however, there is one known importer, a PTS transporter belonging to the Tagatose-6-Phosphate pathway. Import via this pathway, however, is unable to generate intermediates for use in the Leloir pathway, meaning additional transporters must be present. As mentioned in Section 1, a previous study (Bidossi et al., 2012) investigated the array of carbohydrate transporters in the pneumococcus and identified the ABC transporter SPD_0088 as

A

B

FIGURE 3.2 **Impact of *galR* deletion on bacterial growth**
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth
was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points
are mean OD_{600} of triplicate assays.

a potential importer of Gal. To determine if this was indeed the case, a previously generated strain (**Table 2.1**) in which the SPD_0088 gene had been deleted (ΔSPD_0088) was assessed.

In CDM + Glc the growth of both the wild-type and ΔSPD_0088 was comparable (**Fig. 3.3A**), while in CDM + Gal (**Fig. 3.3B**) ΔSPD_0088 showed only slightly decreased growth compared to the wild-type. This indicates that SPD_0088 plays only a minor role in the import of Gal into the cell under these conditions, something that might be expected given the array of known transporters in the pneumococcus.

3.2.4. Generation of putative GalR phosphorylation site amino acid substitution mutants

Previous work in our laboratory (Trappetti et al., 2017) was the first to show that AI-2 signalling in the presence of Gal results in increased capsule production and hypervirulence. This appeared to be dependent on import of AI-2 via the FruA PTS system, leading to upregulation of the Leloir pathway. PTS systems phosphorylate their cargo during import, suggesting the possibility that phosphorylated AI-2 might be capable of mediating phosphorylation of GalR (directly or indirectly). As mentioned in Section 1.10, a previous study by Sun *et al.,* (Sun et al., 2010) utilised phosphoproteomics to determine residues that may be subject to phosphorylation throughout the pneumococcal genome. This study revealed three putative phosphorylation sites within GalR: Serine (Ser) 317 (S317), Threonine (Thr) 319 (T319) and Thr 323 (T323). To determine whether these proposed phosphorylation sites are involved in GalR-mediated upregulation of the Leloir pathway, we performed site-directed mutagenesis to selectively replace these residues with select amino acids that would either mimic or prevent phosphorylation at these positions.

A

CDM + Glc

B

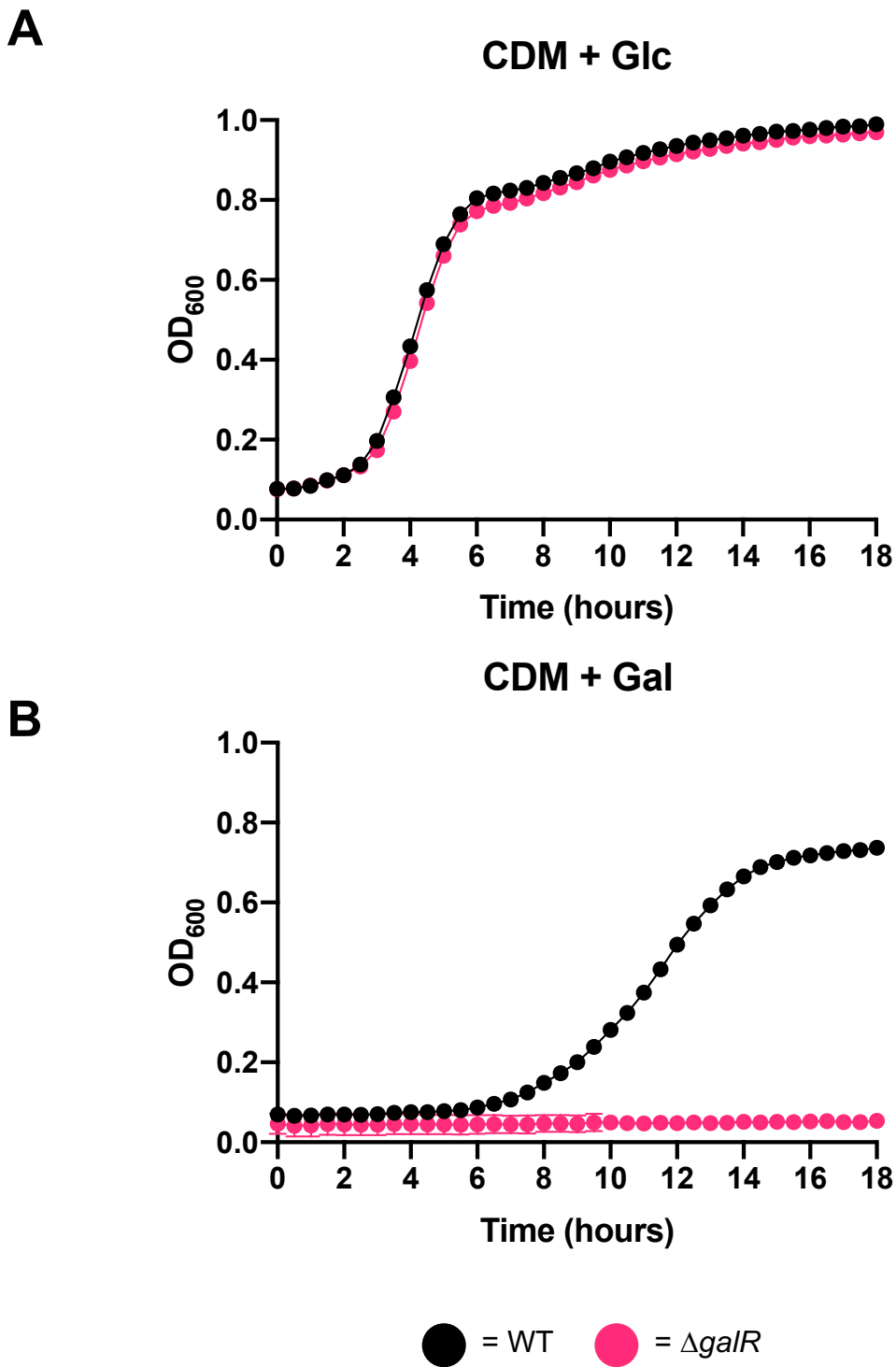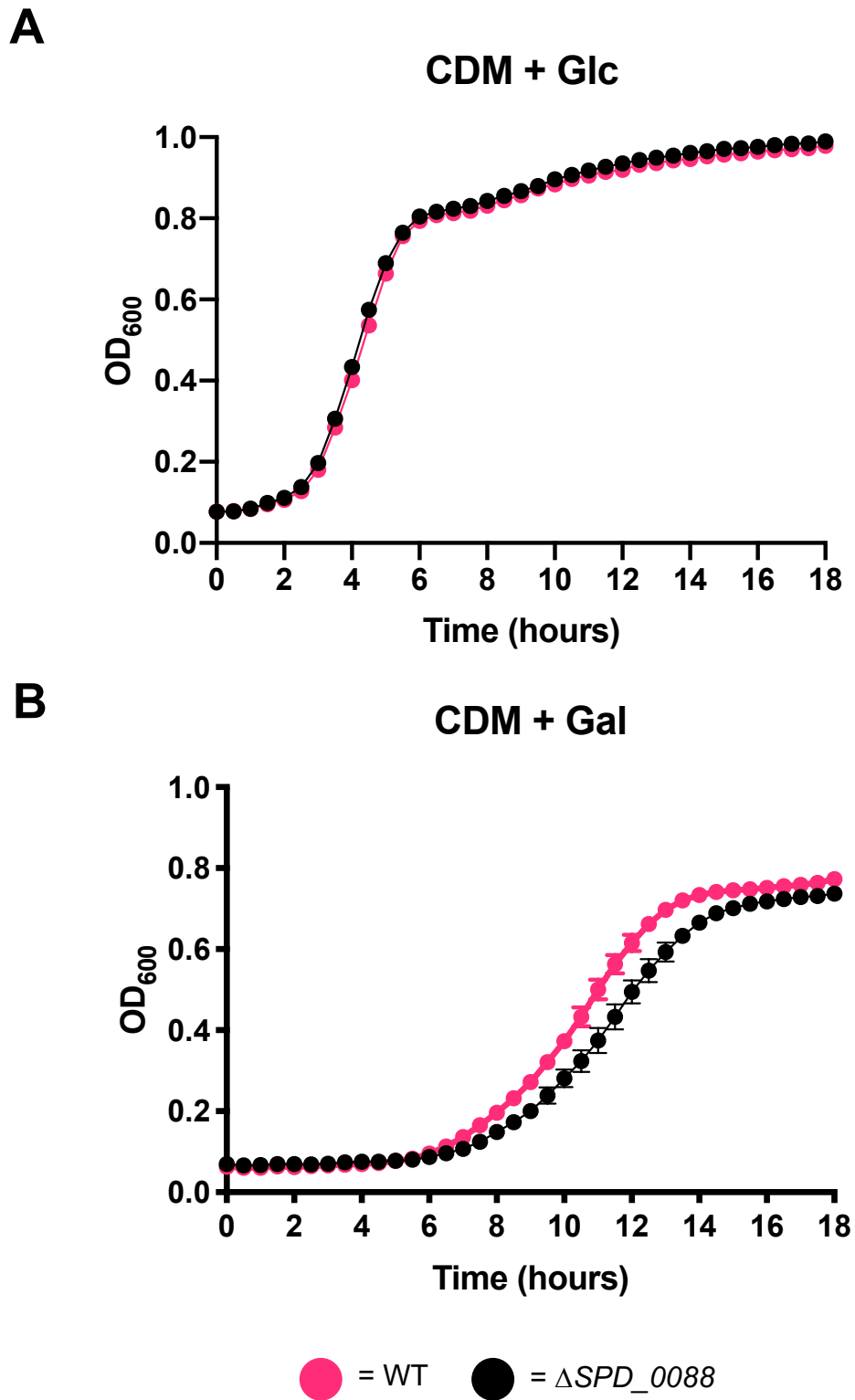### CDM + Gal

● = WT   ● = ΔSPD_0088

FIGURE 3.3   Impact of SPD_0088 deletion on bacterial growth
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

To do this, the Janus cassette was used (Sung et al., 2001). As mentioned in Section 2.4, this system firstly requires the introduction of an *rpsL* mutation within the parent strain. Replacement of the streptomycin-sensitive *rpsL* with the streptomycin-resistant *rpsL1* allele allows for counter-selection at later stages. From here, the gene of interest, in this case *galR,* is replaced with the Janus cassette which renders the strain kanamycin resistant and streptomycin sensitive, as shown in **Figure 3.4**, thus generating the Δ*galR* Janus strain into which the amino acid substitutions could then be introduced. Following this, PCR was used to generate linear DNA fragments incorporating the desired *galR* amino acid substitutions. This required the use of primer sets that would amplify from outside of the gene to the site of the desired amino acid mutations, as shown in **Figure 3.5.** This approach generates a donor PCR product which can then be transformed into the Δ*galR* Janus strain as described in Section 2.3. Transformation results in replacement of the Janus cassette with this PCR product, restoring a modified version of *galR* to the genome with the required amino acid substitutions included. These mutants can then be selected for based on their resistance profiles, firstly by resistance to streptomycin, then by concurrent resistance to kanamycin. Using this system, we are able to replace wild-type residues with either alanine (A), which lacks the hydroxyl group required for phosphorylation, or aspartic acid (D) or glutamic acid (E), which places a positive charge at the respective position, thereby constitutively mimicking the phosphorylated state. The various single and combination substitution mutants (in the D39 background) and the abbreviation code used to designate each of them are detailed in **Table 2.1.** This code comprises three subscript letters indicating whether the three putative GalR phosphorylation sites at amino acids 317, 319 and 323 are the wild type (S or T), or the substitution mutants (A, D or E). The phenotypes of the various mutants were then determined.
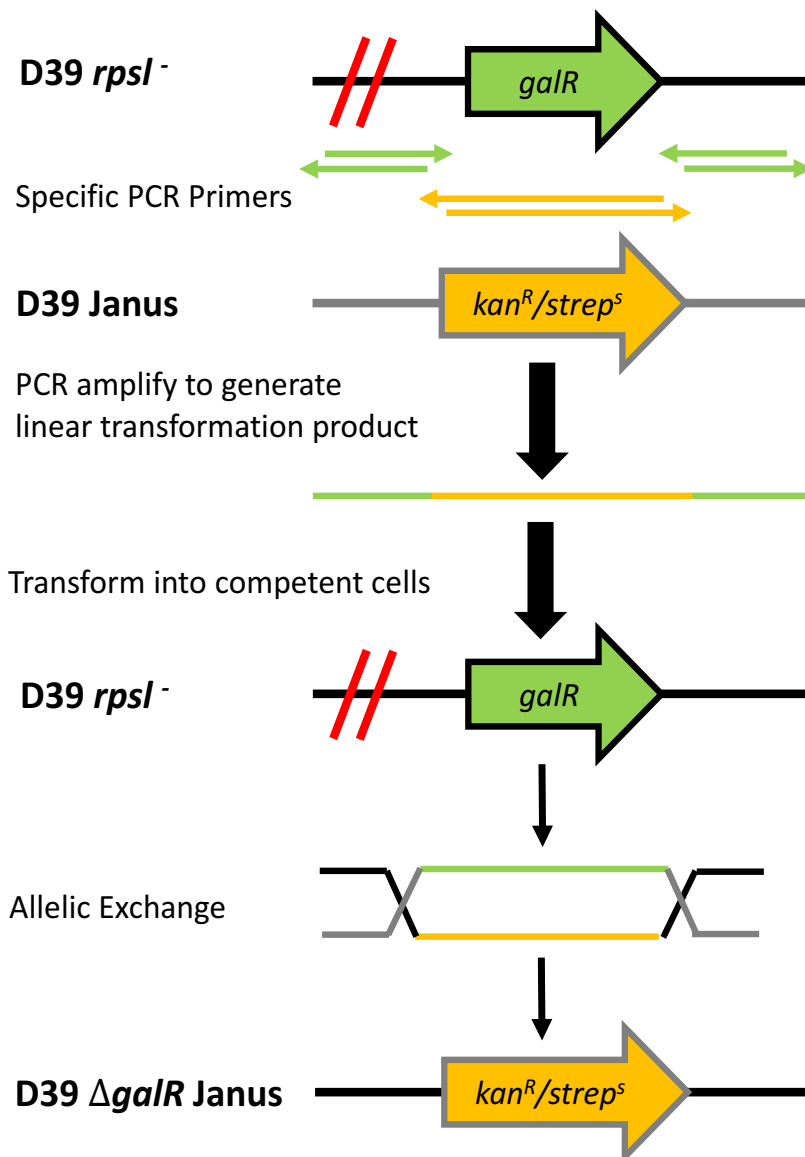
FIGURE 3.4. **Generation of the Δ*galR* Janus strain**

An overview of the process undertaken to generate the Δ*galR* Janus strain. First, the streptomycin-sensitive *rpsl* allele is replaced with the streptomycin-resistant *rpsL1* allele. The strain is then transformed with linear DNA housing the Janus cassette flanked by DNA that is homologous to the regions directly up and downstream of the gene of interest, in this case *galR*, yielding the Δ*galR* Janus strain.

A

D39 WT

Specific PCR Primers

PCR amplify to generate mutations

Donor PCR Product

Transform into competent cells

D39 ΔgalR Janus   kan^R/strep^s

Allelic Exchange

GalR_XXX Mutant   galR

B

3 kb
2 kb
1.6 kb
1 kb
800 bp
700 bp
600 bp

1.6 kb Reverse Flank

700 bp Forward Flank

C

3 kb
2 kb
1.6 kb
1 kb
800 bp
700 bp
600 bp

2.3 kb Overlap Donor PCR Product
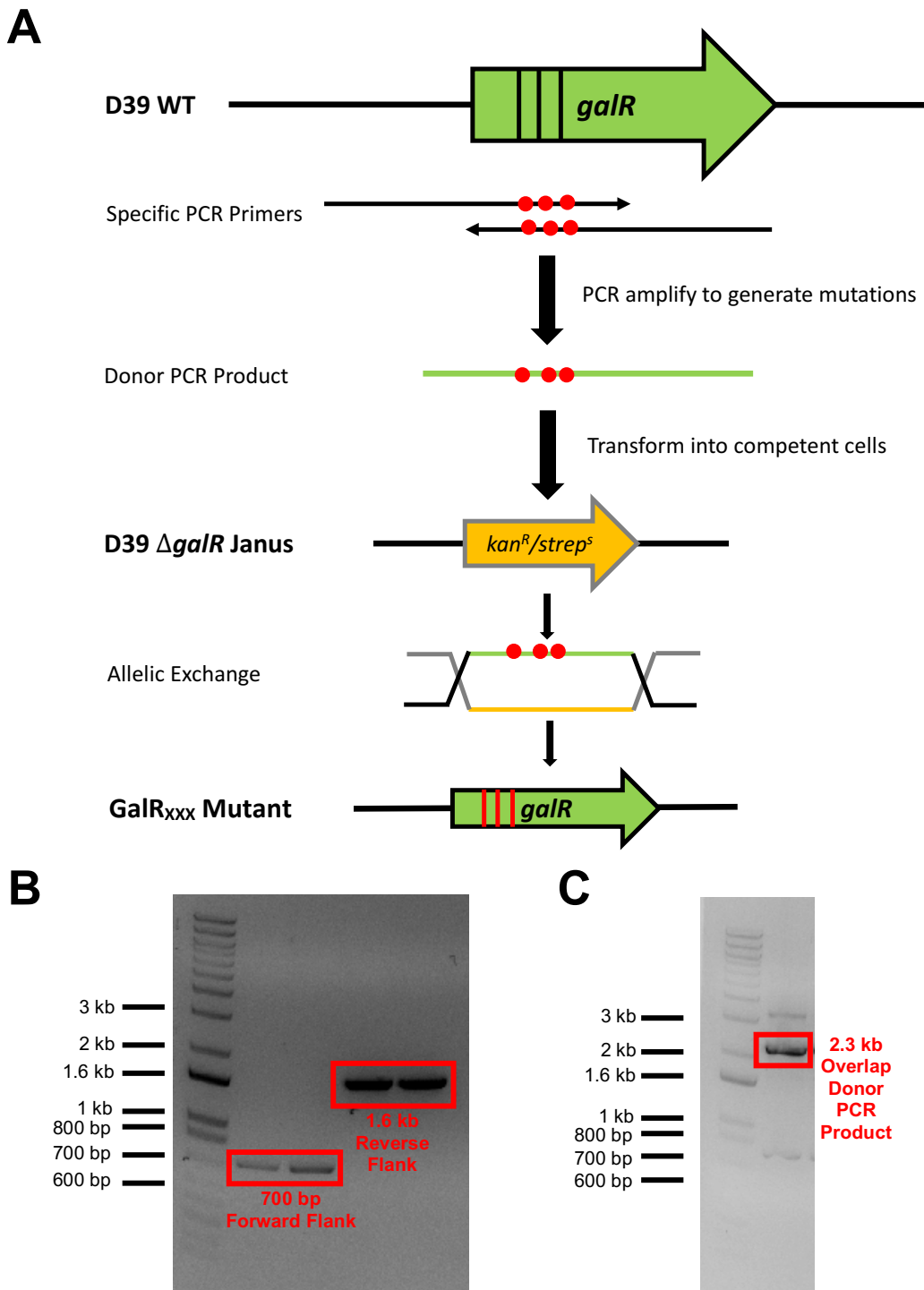
FIGURE 3.5   Generation of GalR amino acid substitution mutants

Overview of the process used to generate the GalR amino acid substitution mutants (A). Using the primers listed in **Table 2.4.**, PCR was utilised to firstly generate forward and reverse products housing the desired amino acid substitutions (B). These products were then used as the template in an Overlap Extension PCR to yield the final donor PCR product (C). This product was then transformed into the ΔgalR Janus strain. Successful mutants were selected for based their resistance phenotypes and confirmed by Sanger sequencing.

3.2.5. Impact of GalR putative phosphorylation sites on galactose metabolism

3.2.5.1. Alanine Substitution Mutants

By substituting each of the putative phosphorylation sites with Ala residues, phosphorylation is prevented from occurring, allowing us to determine which residues, or combination of residues, must be phosphorylated to enable growth in Gal. There were seven Ala substitution mutants generated, including a control strain in which all three putative phosphorylation sites were replaced with Ala residues (GalR$_{AAA}$), three strains in which each putative site was substituted individually (GalR$_{ATT}$, GalR$_{SAT}$ and GalR$_{STA}$) and three strains in which two putative sites were substituted (GalR$_{AAT}$, GalR$_{ATA}$ and GalR$_{SAA}$). Growth in CDM + Glc revealed that each mutant grew comparably to the wild-type strain (**Fig. 3.6A**). Conversely, when grown in CDM + Gal, significant growth differences between strains became apparent (**Fig. 3.6B**). The GalR$_{AAA}$ strain was almost completely unable to grow in Gal, while GalR$_{SAA}$ showed a delay in growth, with a slower generation time and a decrease in final culture density compared to wild type D39. The remaining GalR substitution mutants, GalR$_{ATT}$, GalR$_{SAT}$, GalR$_{STA}$, GalR$_{AAT}$, and GalR$_{ATA}$, exhibited a capacity to grow in CDM + Gal similar to that of D39. This indicates that mutation of any one of the three GalR phosphorylation sites alone does not significantly impact the capacity to utilise Gal. However, mutation of both T319 and T323, as occurs in GalR$_{SAA}$, reduced the capacity of the strain to grow in this medium. Thus, the first GalR phosphorylation site (S317) on its own is insufficient to fully sustain growth in Gal. Substitution of all three sites, however, essentially abolished growth, indicating that these sites collectively play an essential role during growth in Gal.
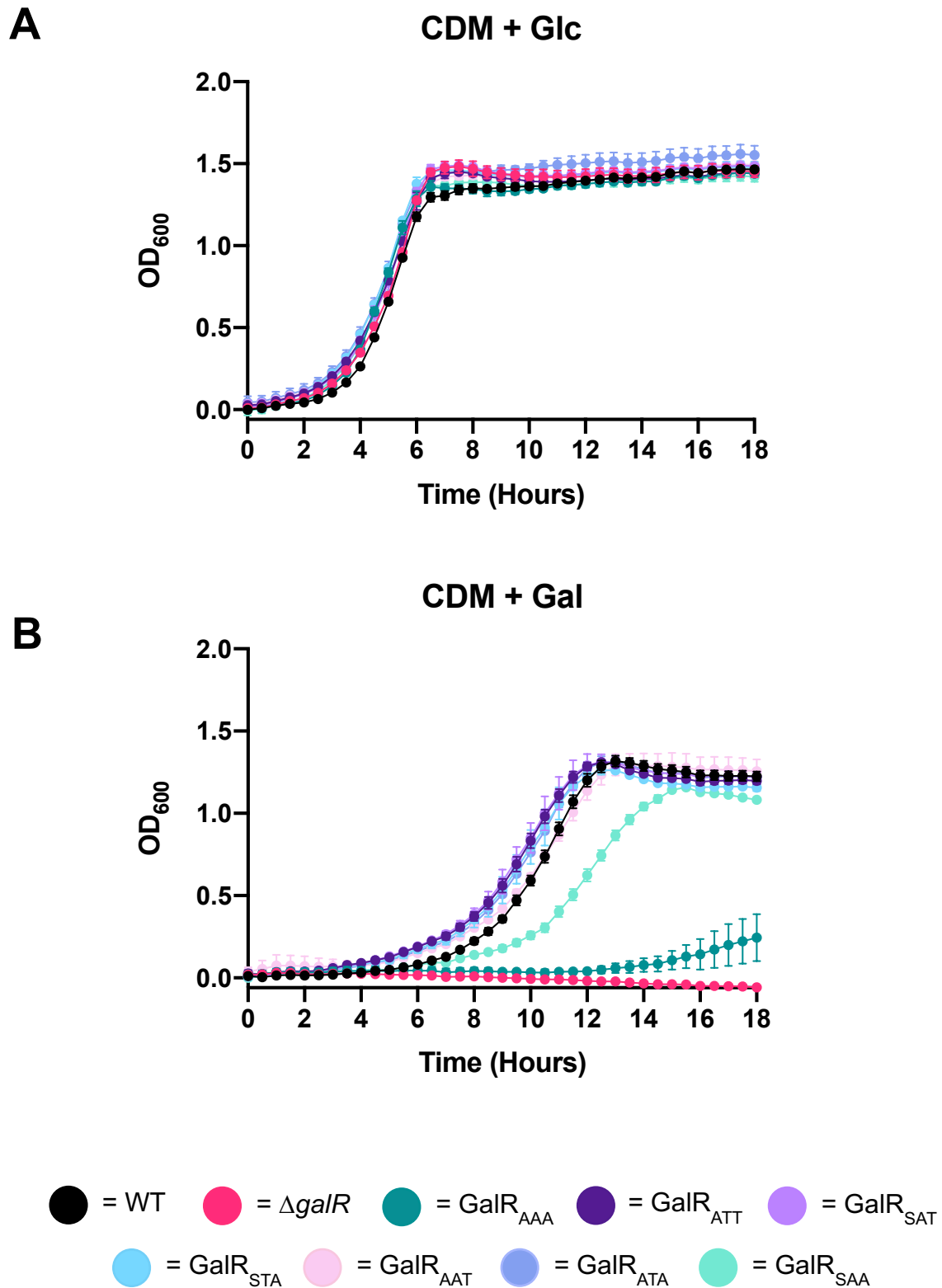
FIGURE 3.6 Impact of alanine substitution on bacterial growth
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

3.2.5.2. Aspartic Acid Substitution Mutants

Preventing phosphorylation at the putative GalR phosphorylation sites by mutating to Ala gave insight as to whether phosphorylation is essential for the utilisation of Gal as a carbon source. We subsequently wanted to investigate the effect that mimicking constitutive phosphorylation had on Gal metabolism. This can be achieved through the use of an amino acid with a negatively charged carboxyl group. Asp is one such amino acid predicated to mimic a state of constitutive phosphorylation at the site of interest (Chen and Cole, 2015). Given this, seven Asp substitution mutants were generated: a strain in which all three putative phosphorylation sites were replaced with Asp residues (GalR$_{DDD}$), three strains in which each putative site was substituted individually (GalR$_{DTT}$, GalR$_{SDT}$ and GalR$_{STD}$) and three strains in which two of the putative sites were substituted (GalR$_{DDT}$, GalR$_{DTD}$ and GalR$_{SDD}$). Growth in CDM + Glc showed that each mutant again grew comparably to the D39 wild-type strain (**Fig. 3.7A**). When grown in CDM + Gal (**Fig. 3.7B**), all substitution mutants with the exception of GalR$_{STD}$, where only the third site has been substituted with Asp, displayed an inability to grow in the presence of Gal. Interestingly, the GalR$_{DDD}$ strain, in which all of the putative phosphorylation sites are substituted, fails to grow in the presence of Gal. This suggests that introduction of a negative charge at all three of the putative sites does not adequately mimic the effect of natural phosphorylation or causes a conformational change in GalR that interferes with its function. The fact that the GalR$_{STD}$ strain has similar growth to wild type in CDM + Gal implies that the third putative phosphorylation site may play a relatively minor role in regulation of Gal metabolism, consistent with the near wild type growth of the Ala substitution mutant, GalR$_{STA}$ in CDM + Gal. These phenotypes may also be largely attributable to natural phosphorylation of the wild type residues at the first and second sites.

A

B

FIGURE 3.7 Impact of aspartic acid substitution on bacterial growth
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

3.2.5.3. Glutamic Acid Substitution Mutants

One possible reason for the apparent inability of Asp residues to mimic phosphorylation at the various sites, at least as judged by impact on growth in CDM + Gal, could be due to selection of sub-optimal phosphomimetic residues. Given that phosphomimetics require the presence of a negatively charged carboxyl group, Glu is another amino acid that can be used to mimic phosphorylation. Based on size similarity, Ser residues should be substituted with Asp, while Thr residues should preferably be substituted with Glu (Thorsness and Koshland, 1987). A set of seven Glc substitution mutants were therefore constructed, including a strain in which all three putative phosphorylation sites were replaced with Glu residues (GalR$_{EEE}$), three strains in which each putative site was substituted individually (GalR$_{ETT}$, GalR$_{SET}$ and GalR$_{STE}$) and three strains in which two of the putative sites were substituted (GalR$_{EET}$, GalR$_{ETE}$ and GalR$_{SEE}$). As with the previously generated mutants, growth in CDM + Glc showed that the mutants grew comparably to the D39 wild-type strain (**Fig. 3.8A**). Interestingly, growth in CDM + Gal (**Fig. 3.8B**) revealed that only two amino acid substitution mutants could grow under these conditions, GalR$_{SET}$ and GalR$_{STE}$, where only the second or third site was substituted with Glu. All remaining substitution mutants failed to grow when Gal was the sole carbon source. While it is possible that the second and third sites may play minor roles in Gal metabolism, the capacity of GalR$_{SET}$ and GalR$_{STE}$ to grow in CDM + Gal may also be attributable to natural phosphorylation of the wild type residues at the other two sites in each case.

3.2.5.4. Alanine and Aspartic Acid Substitution Mutants

Analysis of the Ala substitution mutants provided insights into the requirement for phosphorylation at one or more of the three putative sites for Gal metabolism. However, the inability of either GalR$_{DDD}$ or GalR$_{EEE}$ to grow in CDM + Gal could have
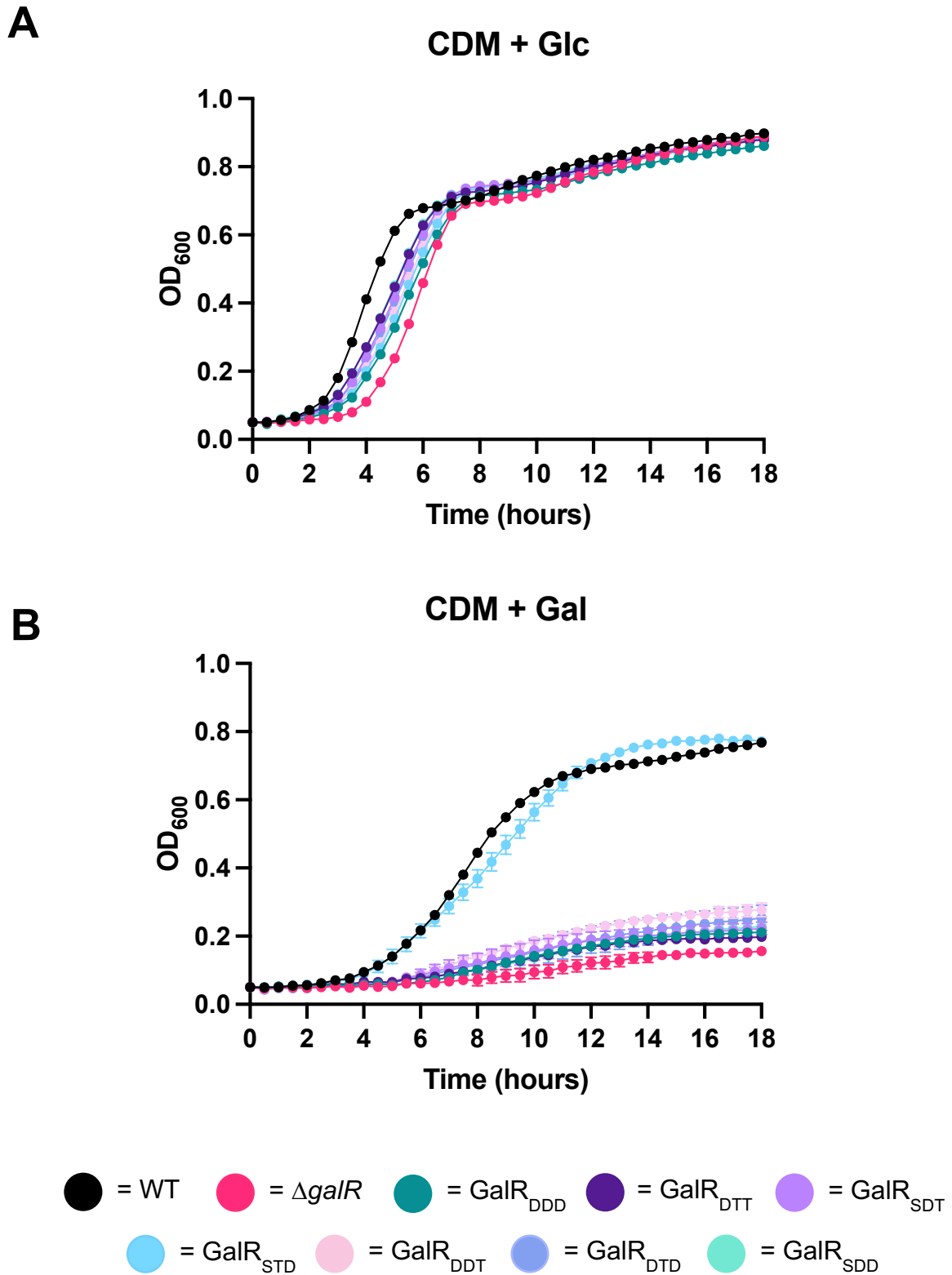
FIGURE 3.8 Impact of glutamic acid substitution on bacterial growth
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth
was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points
are mean OD_{600} of triplicate assays.

been due to the inability of either residue to effectively mimic constitutive phosphorylation, or conformational changes induced by the density of negative charges that impacted GalR function. Interpretation of the phenotypes of the various single and double D and E substitution mutants was also complicated by the presence of wild type S or T at the other site(s). To address this, a set of six additional mutants were generated; in three of these individual phosphorylation sites were predicted to mimic the 'ON' state by substitution with aspartate and the remaining two 'OFF' by substitution with Ala (GalR$_{DAA}$, GalR$_{ADA}$ and GalR$_{AAD}$). The other three strains comprised combinations of two phosphorylation sites were 'ON' and the remaining site 'OFF' (GalR$_{DDA}$, GalR$_{DAD}$ and GalR$_{ADD}$). Growth in CDM + Glc (**Fig. 3.9A**) showed that the mutants grew comparably to the D39 wild-type strain. However, none of the mutant strains were able to grow in CDM + Gal (**Fig. 3.9B**). This indicates that the presence of the phosphomimetic residues at any of the putative GalR phosphorylation sites is not sufficient to enable metabolism of Gal when the other sites are mutated to Ala.

3.2.5.5. Alanine and Glutamic Acid Substitution Mutants

A further six combined Ala and Glu mutants were then also generated; in three strains only individual phosphorylation sites were rendered 'ON' with the remaining two 'OFF' (GalR$_{EAA}$, GalR$_{AEA}$ and GalR$_{AAE}$); in the other three strains a combination of two phosphorylation sites were 'ON' and the remaining site 'OFF' (GalR$_{EEA}$, GalR$_{EAE}$ and GalR$_{AEE}$). Growth in CDM + Glc again showed that all mutant strains could grow comparably to the wild-type (**Fig. 3.10A**). However, growth in CDM + Gal (**Fig. 3.10B**) showed that the mutants were again unable to grow when Gal is the sole carbon source. These findings indicate that substitution of any or all of the sites with either of the phosphomimetic residues (Asp or Glu) is insufficient for Gal metabolism.

A

B

FIGURE 3.9 Impact of combined alanine and aspartic acid substitution on bacterial growth
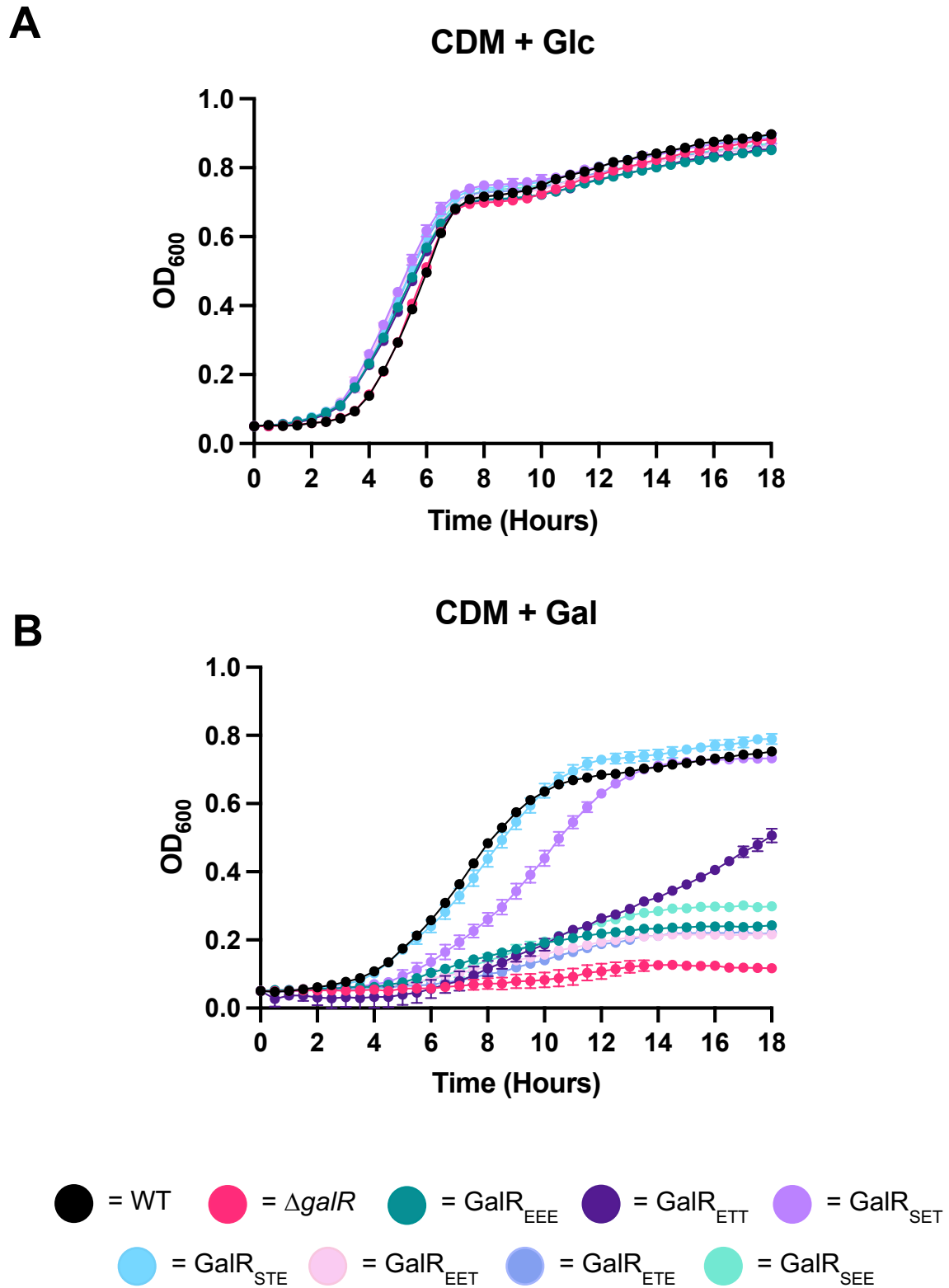
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

A

B

FIGURE 3.10 **Impact of combined alanine and glutamic acid substitution on bacterial growth**

The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

<u>3.2.5.6. Combined Aspartic and Glutamic Acid Residue Substitution</u>

The results so far support the finding that Asp is the most appropriate phosphor-Ser mimic, while Glu is the most appropriate phospho-Thr mimic (Thorsness and Koshland, 1987). Given this, a final mutant strain incorporating the optimum phosphomimetic residues at each site (GalR$_{DEE}$) was constructed. When grown in CDM + Glc (**Fig. 3.11A**), GalR$_{DEE}$ grows comparably to the wild-type. However, when grown in CDM + Gal (**Fig. 3.11B**) there was no detectable growth of the GalR$_{DEE}$ strain, similar to what was observed for both the GalR$_{DDD}$ and GalR$_{EEE}$ strains. Collectively, these findings cast doubt on whether substitution of putative phosphorylatable residues with acidic amino acids adequately mimics the wild type phosphorylated state. Alternatively, constitutive activation of GalR due to incorporation of phosphomimetic residues may lead to some form of feed-back inhibition either via GalR or some other regulatory pathway or result in conformational changes within the protein which significantly impact its function.

3.2.6. Impact of GalR amino acid substitution on gene expression

Notwithstanding the uncertainties emanating from the growth studies with phosphomimetic mutants above, analysis of the Ala substitution mutants (Section 3.2.5.1) suggested a clear requirement for phosphorylatable residues at one or other of the putative phosphorylation sites for growth in Gal. While growth is informative, it still leaves questions regarding the effects of these substitutions on the regulation of both Leloir pathway specific genes and also those in other related pathways. Therefore, to complement the above growth data, gene expression analyses were conducted on all the GalR Ala substitution mutants to assess the impact on expression of both Leloir and T6P pathway genes. Strains were grown overnight on blood agar, washed and resuspended in CDM + Gal, and then

A

CDM + Glucose

B

CDM + Galactose

● = WT   ● = GalR$_{DEE}$

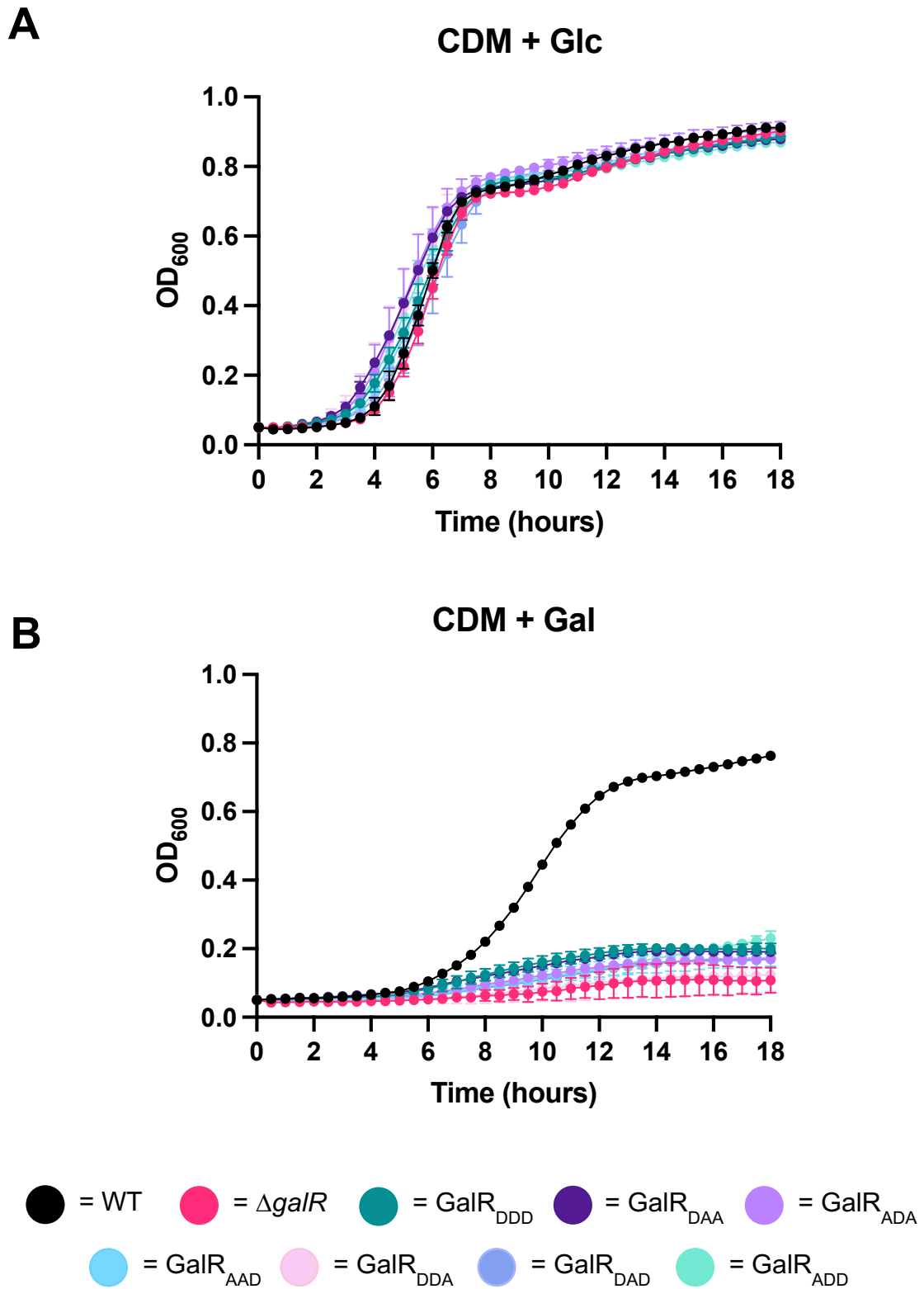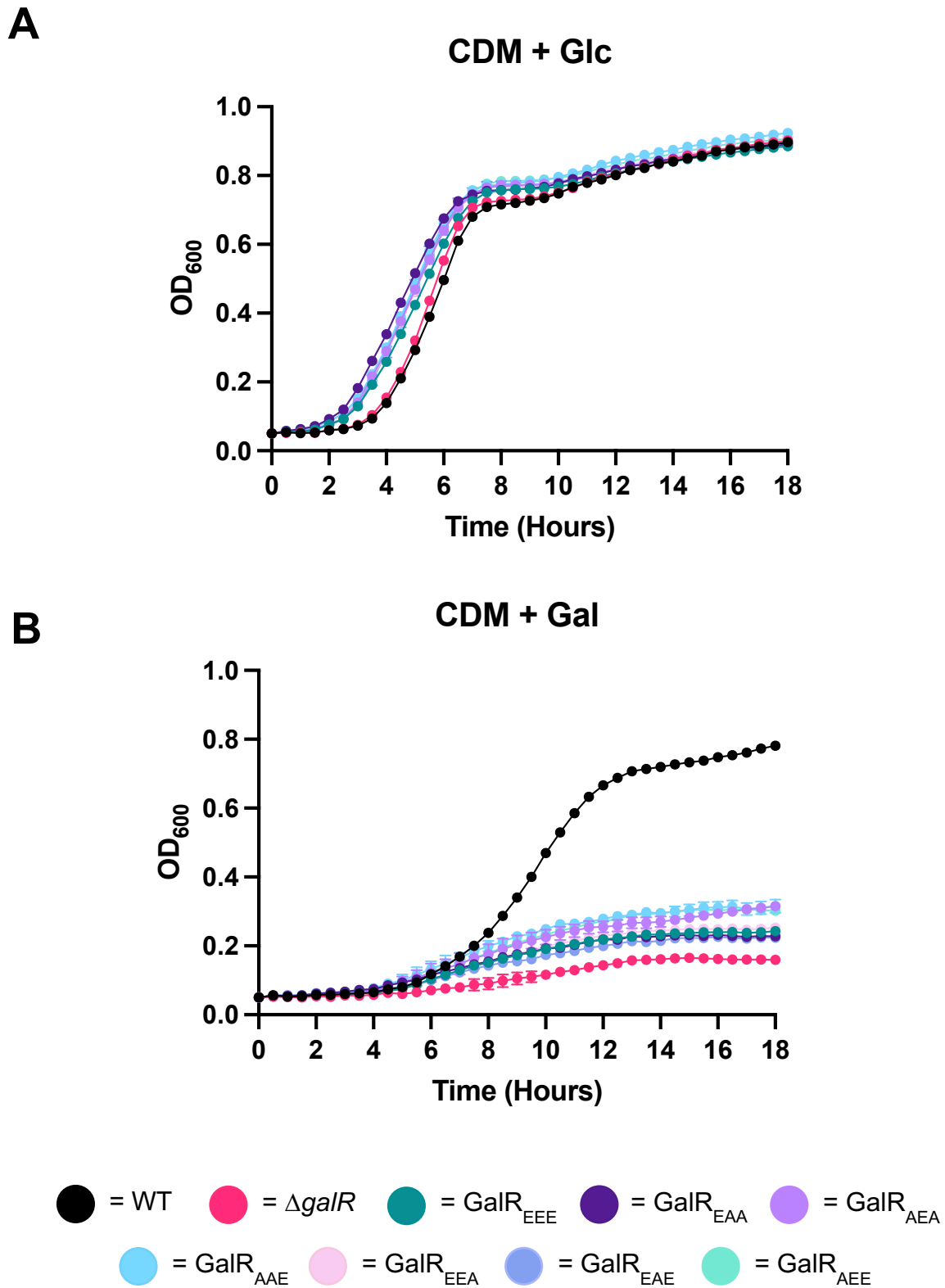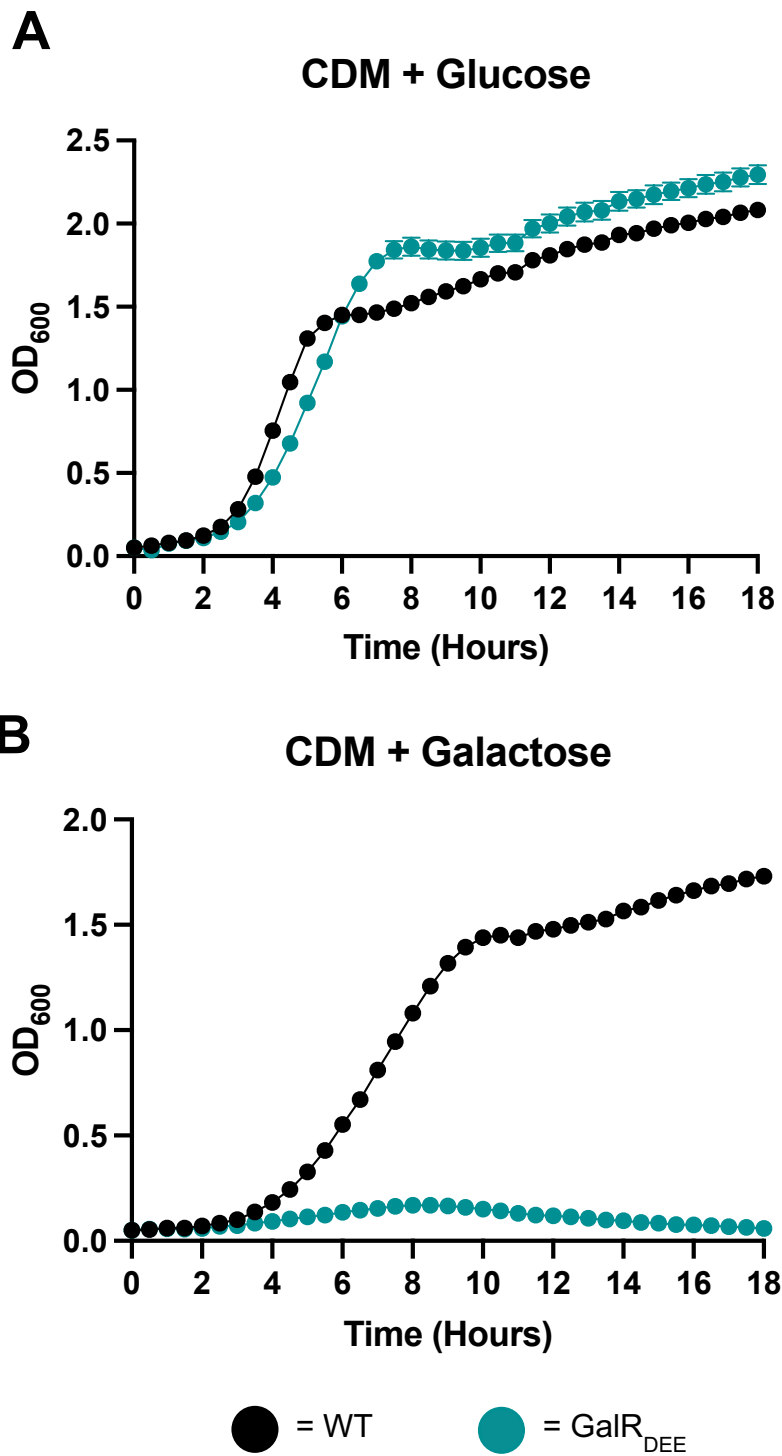**FIGURE 3.11** **Impact of combined aspartic and glutamic acid substitution on bacterial growth**
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD$_{600}$ every 30 min for a total of 18 h. Data points are mean OD$_{600}$ of triplicate assays.

incubated for 30 min. RNA was then extracted, and expression of *galR*, *galK*, and *lacD* was assessed by quantitative reverse transcription-qualitative PCR (qRT-PCR) (see Section 2.9) (**Fig. 3.12**).

Expression of *galR* itself was, as expected, undetectable in Δ*galR and* was also significantly downregulated in all the mutants tested. GalR$_{AAA}$ was the most affected, showing an 88% reduction in *galR* expression (**Fig. 3.12A**). As *galR* has been previously shown to *regulate* the *galKT* operon (Afzal et al., 2015), we expect to see similarly decreased *galK* expression in these strains. There was significantly decreased expression of *galK* in all GalR mutants compared to wild-type D39 (**Fig. 3.12B**). In particular, the Δ*galR* and GalR$_{AAA}$ strains showed similarly low levels of *galK* expression (≥98% reduction). These findings are largely consistent with the effects of the mutations on the expression of *galR* itself (**Fig. 3.12A**). It is worth noting that only those mutants with >98% reduction in *galK* expression (Δ*galR* and GalR$_{AAA}$) exhibited severe growth defects in CDM + Gal (**Fig. 3.12B**). We also assessed whether the absence of functional Leloir pathway expression had any effect on the expression of the T6P pathway by examining *lacD* expression. *lacD* encodes the last enzyme of the T6P pathway and is responsible for the conversion of tagatose 1,6-bP to dihydroxyacetone-P and D-glyceraldehyde-3-P, which can then feed into the glycolytic pathway. Interestingly, *lacD* expression was significantly (30 to 60%) lower in all GalR mutants compared to D39 (**Fig. 3.12C**), indicating a direct or indirect role for GalR phosphorylation in the expression of T6P pathway genes. However, there was no apparent association between reduced *lacD* expression and the relative ability of the various strains to grow in Gal.

A *galR*

B *galK*

C *lacD*
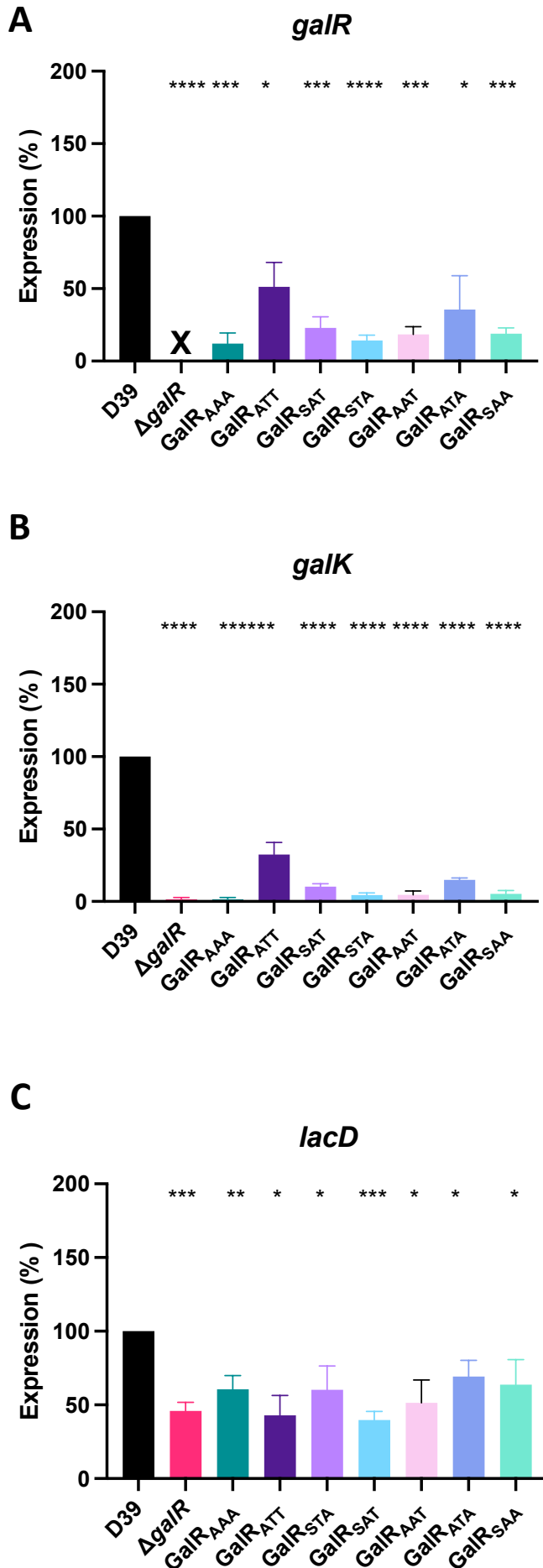
FIGURE 3.12. Differential gene expression in GalR alanine substitution mutants

D39, $\Delta galR$, $galR_{AAA}$, $galR_{ATT}$, $galR_{SAT}$, $galR_{STA}$, $galR_{AAT}$, $galR_{ATA}$ and $galR_{SAA}$ were cultured overnight on blood agar plates, washed and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. RNA was then extracted and levels of *galR* (A), *galK* (B) and *lacD* (C) mRNA were quantitated by qRT-PCR using *gyrA* as an internal control (see Materials and Methods). Data presented are the mean ± standard deviation from three independent experiments, expressed as a percentage of that for D39. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$, unpaired *t*-test (relative to D39); ns, not significant; X, transcript absent due to gene deletion.

3.2.7. Substitution of key GalR residues results in a decreased ability to adhere to nasopharyngeal cells

The human host environment provides an array of carbon sources that can potentially be utilised by bacteria. The human nasopharynx and upper respiratory tract is particularly rich in Gal, either free of as a component of cell surface glycoconjugates (Pezzulo et al., 2011). Thus, the ability to acquire and metabolise Gal may be an important factor in successful host colonisation. Following from our growth and transcriptional studies, we wanted to determine if the putative phosphorylation sites play a role in the ability to adhere to the nasopharyngeal epithelium. To assess this, Detroit 562 nasopharyngeal cells were first grown to confluency before being incubated with *S. pneumoniae* in the presence of CDM + Gal and examined for adherent pneumococci after two hours (Section 2.11) (**Fig. 3.13**). With the exception of GalR$_{STA}$ and GalR$_{ATA}$, there was an observed decrease in the ability to adhere to Detroit cells when compared to D39. Interestingly, the difference in adherence between $\Delta galR$ and each of the amino acid substitution mutants appears to be negligible. Collectively, these analyses show that GalR may play an important role in the ability to adhere to the nasopharyngeal epithelium in an environment where Gal is the predominant carbon source, and that the phosphorylation sites themselves may aid in this.

3.2.8. Substitution of key GalR residues has a minimal impact on total capsule production

The polysaccharide capsule is one of the primary virulence factors of the pneumococcus, and is known to impact the capacity of pneumococci to adhere to epithelial cells *in vitro* (Hammerschmidt et al., 2005). We therefore wanted to determine whether there were any differences in the total amount of capsular
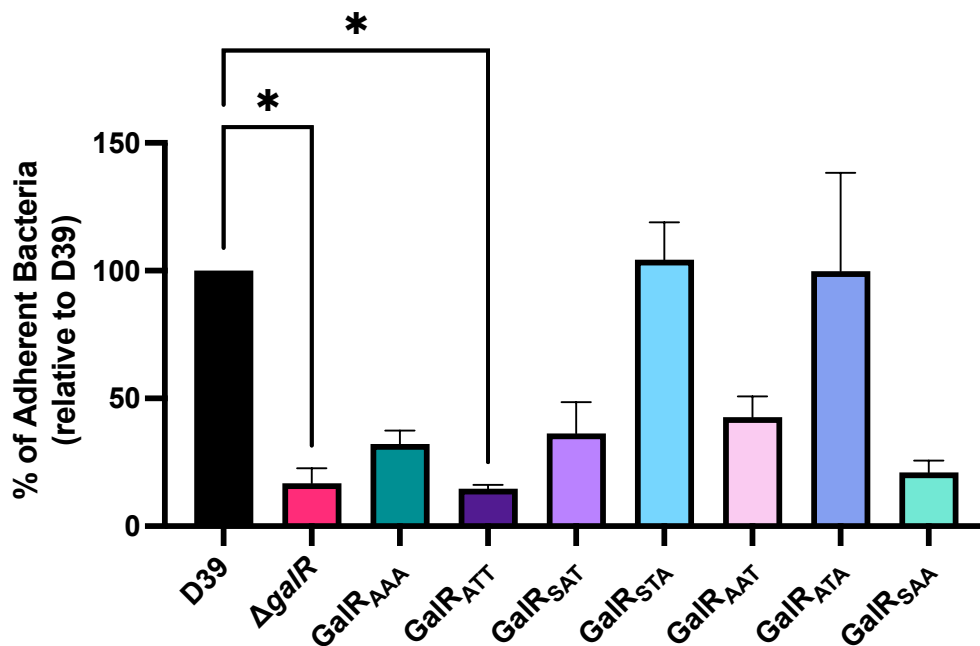
FIGURE 3.13. Impact of amino acid substitution on adherence to Detroit 562 cells

D39, ΔgalR, GalR$_{AAA}$, GalR$_{ATT}$, GalR$_{SAT}$, GalR$_{STA}$, GalR$_{AAT}$, GalR$_{ATA}$ and GalR$_{SAA}$ and galR$_{ATA}$ were inoculated at a final OD$_{600}$ of 0.2 onto Detroit 562 monolayers in CDM + 10% FCS + 0.5% Gal. Cells were then incubated for 2 hours at 37°C before being assessed for adherence. Each strain was assayed in technical triplicate over a minimum of two independent experiments. Data presented are the mean ± standard error of the mean, expressed as a percentage of that for D39. *, $P < 0.05$; one-way ANOVA (relative to D39).

polysaccharide present on Δ*galR* and GalR$_{AAA}$ before moving on to *in vivo* studies. When compared to D39, there was no significant difference in capsule production in the Δ*galR* and GalR$_{AAA}$ strains when grown in Serum Broth (data not shown). Serum Broth was used as opposed to CDM + Gal as this is the media used to grow murine challenge strains. This also circumvented the issue of these strains being unable to grow in CDM + Gal, as shown earlier in this chapter. The results indicate that despite these strains showing differences in growth when Gal is the sole carbon source, they are able to produce similar amounts of capsular polysaccharide, presumably exploiting an alternative carbon source under these conditions.

3.2.9. Impact of GalR and GalR phosphorylation in a mouse model of pneumococcal infection

Having now characterised the growth and transcriptional phenotypes of the Ala substitution mutants, along with the influence of the substitutions on adherence to nasopharyngeal cells and total capsule production, we next wanted determine if these substitutions, along with GalR, also had any significant impact on the progression of invasive pneumococcal disease *in vivo*. Groups of mice were challenged intranasally with 1×10^7 CFU of wild type D39, Δ*galR* or GalR$_{AAA}$, and the numbers of pneumococci present in various niches (nasal tissue, ear, lungs, and blood) were determined at 24 h post-challenge (**Fig. 3.14**). The GalR$_{AAA}$ strain exhibited a significantly attenuated virulence phenotype, with reduced bacterial loads relative to D39 in nasal tissue (**Fig. 3.14A**), ears (**Fig. 3.14B**), and lungs (**Fig. 3.14C**). A similar trend was observed for bacterial loads in the blood, although this did not reach statistical significance (**Fig. 3.14D**). The Δ*galR* strain showed a significantly decreased bacterial load in the ear, but not in any other niches.
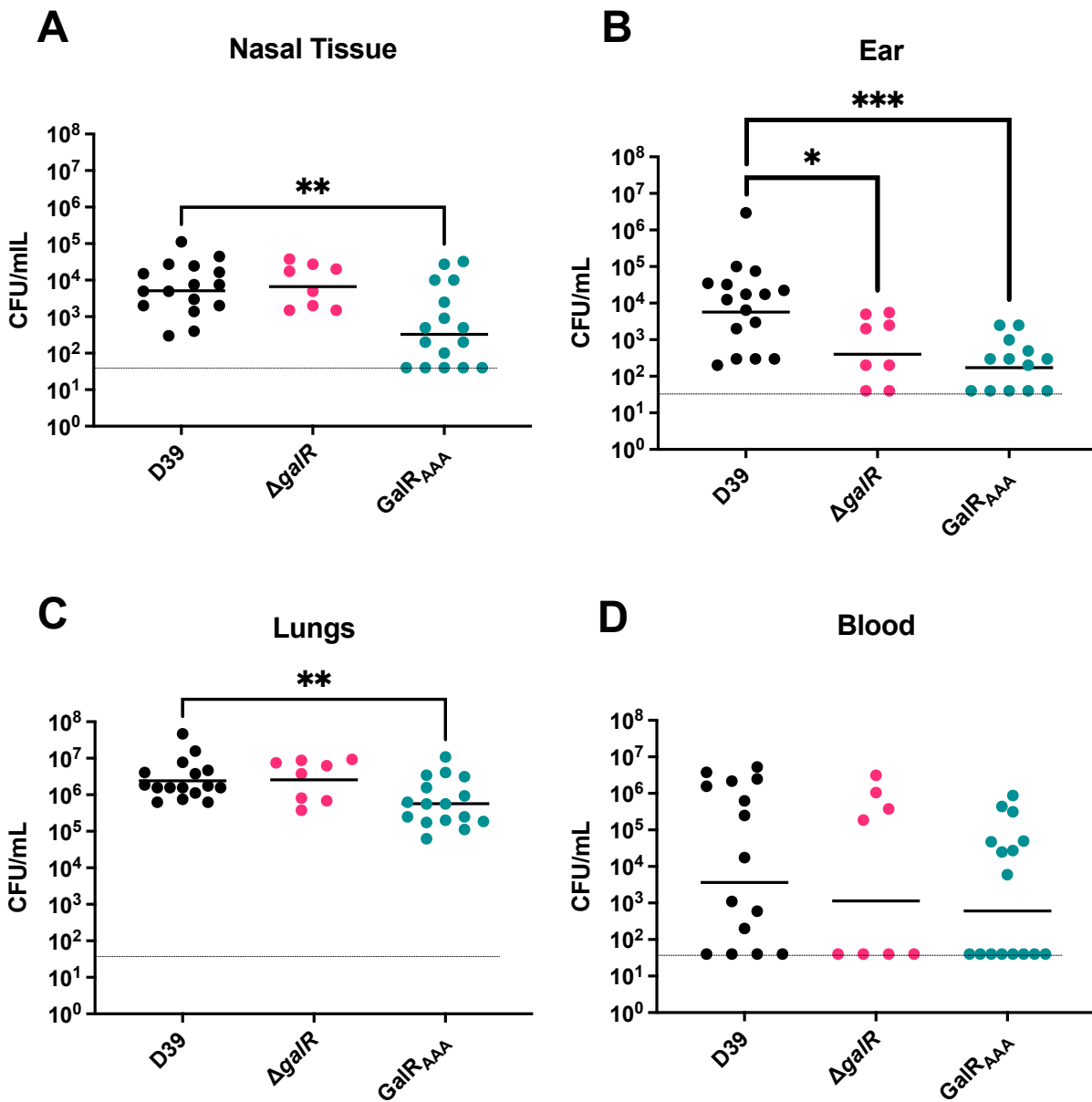
FIGURE 3.14. Virulence phenotypes of D39, Δ*galR* and GalR$_{AAA}$

Groups of mice were infected intranasally with 10^7 CFU of the indicated strain. At 24 h, mice were euthanised and the number of pneumococci isolated from the nasal tissue (A), lungs (B), ears (C) and blood (D) were enumerated. Viable counts (total CFU per tissue, or CFU per ml blood) are shown for each mouse in each niche; the horizontal bars indicate the geometric mean (GM) CFU for each group; the dotted line indicates the threshold of detection. Significance of differences in GM bacterial load between groups were analysed using unpaired *t* tests: *, $P < 0.05$; ** $P < 0.01$; ***, $P < 0.001$.

TABLE 3.1. Summary of GalR substitution mutant growth phenotypes.

Strain	Growth in Glucose	Growth in Galactose
Alanine Substitution Mutants		
GalR$_{AAA}$	Y	N
GalR$_{ATT}$	Y	Y
GalR$_{SAT}$	Y	Y
GalR$_{STA}$	Y	Y
GalR$_{AAT}$	Y	Y
GalR$_{ATA}$	Y	Y
GalR$_{SAA}$	Y	Y
Aspartic Acid Substitution Mutants		
GalR$_{DDD}$	Y	N
GalR$_{DTT}$	Y	N
GalR$_{SDT}$	Y	N
GalR$_{STD}$	Y	Y
GalR$_{DTT}$	Y	N
GalR$_{DTD}$	Y	N
GalR$_{SDD}$	Y	N
Glutamic Acid Substitution Mutants		
GalR$_{EEE}$	Y	N
GalR$_{ETT}$	Y	N
GalR$_{SET}$	Y	Y
GalR$_{STE}$	Y	Y
GalR$_{EET}$	Y	N
GalR$_{ETE}$	Y	N
GalR$_{SEE}$	Y	N
Combined Alanine and Aspartic Acid Mutants		
GalR$_{ADD}$	Y	N
GalR$_{ADA}$	Y	N
GalR$_{AAD}$	Y	N
GalR$_{DDA}$	Y	N
GalR$_{DAD}$	Y	N
GalR$_{ADD}$	Y	N
Combined Alanine and Glutamic Acid Mutants		
GalR$_{AEE}$	Y	N
GalR$_{AEA}$	Y	N
GalR$_{AAE}$	Y	N
GalR$_{EEA}$	Y	N
GalR$_{EAE}$	Y	N
GalR$_{AEE}$	Y	N

This indicates that the putative GalR phosphorylation sites do indeed play a role in the pathogenesis of invasive disease, while GalR may play a role in the ability to persist in the ear during infection.

3.3. Discussion

The results of this chapter have expanded on previous findings showing that GalR is important for Gal metabolism in *S. pneumoniae* (Afzal et al., 2015), revealing that the putative GalR phosphorylation sites (S317, T319, and T323) are important for growth in a defined medium with Gal as the sole carbon source. Mutation of all putative phosphorylation sites to Ala (GalR$_{AAA}$) significantly abrogated growth in Gal and reduced expression of *galR* and *galK* to levels comparable to those of the Δ*galR* strain (**Fig. 3.6B** and **Fig. 3.12**). Growth of all substitution mutants is summarised in **Table 3.1.** These mutations also had a profound impact on the ability to adhere to nasopharyngeal epithelial cells *in vitro* (**Fig. 3.13**). Moreover, the substitution of these amino acids with Ala appears to alter the interaction of GalR with the *galR* operator sequence; as a result, the defects observed in *galK* expression may be at least partially attributable to a reduction in GalR abundance, rather than a direct alteration in binding to the *galK* operator. The precise mechanism behind these defects remains unknown but may be due to effects on folding, dimerization, or binding of effector molecules rather than by directly preventing phosphorylation. However, structural modelling of GalR shows that these residues are located in a distinct region to the putative Gal and DNA binding domains (**Fig. 3.1B**). Additional studies exploring the interaction of purified GalR and GalR$_{AAA}$ with each operator DNA sequence, e.g. using electrophoretic mobility shift assays, may provide greater insight into the regulation of the Leloir pathway genes.

In *S. pneumoniae*, the only reported kinase responsible for phosphorylation of Ser and Thr residues is StkP (Echenique et al., 2004, Novakova et al., 2005). A previous proteomic study performed in medium containing Glc failed to identify GalR as a target of StkP, but LacA, the Gal6P isomerase responsible for converting Gal6P to T6P, was identified as a possible target for StkP-mediated phosphorylation (Hirschfeld et al., 2019). Given that this study was performed in Glc means the failure to detect GalR as a potential StkP target is not entirely unsurprising. However, it would be worth investigating this further in the future. Preliminary data from our laboratory have revealed that a D39Δ*stkP* strain was unable to grow in Gal (E. Brazel, 2018, Unpublished), but this may be a consequence of the defect in LacA phosphorylation rather than failure to phosphorylate GalR. The inability of an *stkP* mutant to grow in Gal, however, indicates that this may indeed be the kinase required to phosphorylate GalR. As a result, additional studies to directly examine the role of StkP in GalR phosphorylation are required, along with further investigation into what may be acting as the phosphor-donor in this case.

Gene expression studies demonstrated that of the various GalR mutants constructed in this study, those with the greatest defects in growth in CDM + Gal, namely, Δ*galR* and GalR$_{AAA}$, exhibited virtually undetectable levels of *galK* expression (>98% reduction) (**Fig. 3.12B**), showing a direct link between Leloir pathway gene expression and growth in Gal. The single and double-point mutants also showed significantly reduced expression of *galR* and *galK*, but this level of expression still enabled sufficient Leloir pathway activity to sustain growth in CDM + Gal (**Fig. 3.12A** and **Fig. 3.12B**). Interestingly, of the single or double Ala substitution mutants, the GalR$_{SAA}$ mutant was the only one to show a defect in growth in Gal compared to D39 (**Fig. 3.6B**), but gene expression was similar to that in the other mutants that grew comparably to the wild type (**Fig. 3.6B** and **Fig. 3.12**).

Therefore, growth in Gal can occur even at low levels of *galK* expression. Thus, the effects on Leloir pathway gene expression and growth in CDM + Gal could be attributable to inadequate levels of GalR in the cell, reduced capacity of the respective GalR phosphorylation site mutants to activate Leloir pathway genes such as *galK*, or a combination of both.

The present study has shown that the GalR putative phosphorylation sites play a significant role in adherence to the nasopharyngeal epithelial cells (**Fig. 3.13**) and pneumococcal infection (**Fig. 3.14**). Mice infected with GalR$_{AAA}$ displayed significantly reduced bacterial loads in the nasopharynx, middle ear, and lungs relative to those infected with wild-type D39 (**Fig. 3.14A-D**), while mice infected with Δ*galR* showed decreased bacterial loads in the ear. These findings are compatible with previous studies showing reduced nasopharyngeal colonization and reduced systemic virulence of D39 *galK* and *lacD* deletion mutants after intranasal, but not intravenous, challenge of mice (Paixão et al., 2015). However, the impact of the putative GalR phosphorylation sites has not previously been investigated. Clearly, the capacity to metabolize Gal is important for survival and proliferation in the upper respiratory tract and the middle ear, where it is an important carbon source (Paixão et al., 2015). Moreover, metabolism of Gal by pneumococci *in vitro* is known to lead to increased production of CPS relative to cells growing on Glc, which may be the basis for the altered virulence profiles (Trappetti et al., 2017). We previously showed that treatment with the quorum-sensing molecule AI-2 upregulates Leloir pathway gene expression and CPS production in the presence of Gal *in vitro*, as well as virulence in an intranasal challenge model (Trappetti et al., 2017). This upregulation was dependent on the PTS component FruA, which is presumed to be the bacterial surface receptor for AI-2. This signalling molecule is a di-ketopentose and may structurally mimic the natural cargo of FruA, namely, fructose, and if AI-2 is capable

of internalization via the FruA PTS system, then it would be expected to be phosphorylated during import. It is tempting to speculate that such phosphorylated AI-2 may play a direct or indirect role in GalR phosphorylation, perhaps acting as a phosphate donor, thereby mediating upregulation of the Leloir pathway. Further to this, we can speculate that it may be StkP that is responsible for transferring the phosphate group from AI-2P to GalR. We also know that deletion of *luxS* results in an inability to produce AI-2, as the conversion of S-ribosyl-homocysteine to homocysteine in the Activated Methyl Cycle cannot occur. An initial aim of this project was to determine the impact of exogenous AI-2 on growth and gene expression in the GalR substitution mutants, however we encountered some significant issues with both the biological activity and acquisition of the product itself. Future experiments should include the opportunity to further investigate the relationship between *luxS, stkP* and *galR* with biologically active AI-2 to further elucidate the relationship between carbohydrate metabolism and cell-to-cell signalling in the pneumococcus. An additional step to be taken in future experiments would be to utilise techniques such as phosphor-proteomics to determine the level of phosphorylation that is required for activation of GalR and subsequently, the Leloir pathway. Performing this in strains such as D39, $\Delta galR$ and GalR$_{AAA}$ would allow for the determination of the precise amount needed for function. We know that the phosphorylation sites themselves are important, however the amount of phosphorylation required for activation remains a mystery.

The overall inability of these amino acid substitution mutants to grow is made all the more perplexing by the presence of the tagatose-6-phosphate pathway in these cells. The T6P and Leloir pathways have been previously reported to be discreet pathways, notwithstanding functional similarities in terms of capacity to metabolise Gal (Afzal et al., 2014, Afzal et al., 2015). However, it could be theorised that in the

event of the Leloir pathway being compromised in some manner, for example by amino acid substitution at critical GalR phosphorylation sites, the T6P pathway should be able to operate in place of the Leloir pathway to ensure survival under these nutritional conditions. Interestingly, this wasn't the case, opening the doorway for reasoning that the role of GalR in pneumococcal Gal metabolism may be more widespread than initially thought.

The impact that these amino acid substitutions may have on the overall structure and function of GalR cannot be overlooked. While these substitutions should theoretically act as intended, it would be remiss to argue that there is potentially no change to the folding of GalR, and thus the ability to bind DNA and successfully regulate the repression of both *galK* and *galT*. Further studies should aim to assess whether this is indeed occurring in these mutants, whether this be by use of electrophoretic mobility shift assays (EMSAs) or more intricate DNA-binding assays that utilised recombinantly expressed and purified GalR protein variants.

This study has shown that, collectively, the GalR putative phosphorylation sites play a key role in virulence and the ability to metabolise Gal and has revealed a complex interplay between the Gal metabolic pathways in *S. pneumoniae*.

Chapter 4: The interplay between the Leloir and Tagatose – 6 – Phosphate pathways

4.1. Introduction

The findings of Chapter 3 have shown the importance of GalR in Gal metabolism, and further, a role for the putative phosphorylation sites therein. What remains unclear is how the Leloir and T6P pathways may be able to compensate for each other during times of metabolic stress. This chapter includes the remaining data published in the *Journal of Bacteriology* (McLean *et al.,* 2020).

S. pneumoniae can metabolise upwards of 30 different carbon sources, each through unique pathways and via a range of different transporters. As discussed in Chapter 1, the human host acts to maintain airway sterility by actively eliminating the Glc present in this niche (Pezzulo et al., 2011). In the absence of Glc, Gal is the predominant sugar, meaning that uptake and metabolism of this sugar is critical for bacterial survival. In *S. pneumoniae*, Gal metabolism has traditionally been thought to occur via two discreet pathways, the Leloir pathway, and the tagatose-6-phosphate (T6P) pathway.

As shown in Chapter 3, in cases where GalR is shut down, there is a persistent inability to grow in the presence of Gal, despite the presence of a theoretically functional secondary pathway in the form of the T6P pathway, suggesting a potential link between the two. In this chapter, mutagenesis techniques were employed to generate targeted knockouts of genes within either the Leloir or T6P pathways before characterising their growth, transcriptional and adherence phenotypes in the presence of different carbon sources. Finally, metabolomics was utilised to determine if there were differences in the abundance of key metabolites/intermediates present in either the Leloir or T6P pathways in Gal media.

4.2. Results

4.2.1. Both the Leloir and Tagatose-6-Phosphate pathways are required for growth in Gal

In order to assess the contribution of the Leloir and T6P pathways to Gal metabolism, growth of a range of both Leloir and T6P pathway mutants were analysed in both CDM + Glc and CDM + Gal. Focussing firstly on the Leloir pathway (**Fig. 4.1**), growth in CDM + Glc appears consistent for the wild-type and ΔgalR, as shown in Chapter 3, and Δ*galK* strains (**Fig. 4.1A**). When grown in CDM + Gal, both the Δ*galR* (lacking the regulator of *galKT*) and Δ*galK* strains (lacking the kinase responsible for converting αGal to αGal1P) are unable to grow (**Fig. 4.1B**), supporting the hypothesis that the Leloir pathway is indeed essential for the utilisation of Gal as a carbon source. When considering the T6P pathway mutants (Δ*lacAB* and Δ*lacD*) (**Fig. 4.2**), it was initially thought that these mutations wouldn't have a significant impact on the ability to grow when Gal was the sole carbon source, but this was not the case, even though the Leloir pathway remains intact in these strains. When grown in CDM + Glc, the Δ*lacAB* (lacking the isomerase that converts Gal6P to T6P) and Δ*lacD* mutants (lacking the tagatose-1,6-bisphosphate aldolase) grow comparably to the WT (**Fig. 4.2A**). Interestingly, both the Δ*lacAB* and Δ*lacD* mutants fail to grow in CDM + Gal (**Fig. 4.2B**), showing that the T6P pathway is also necessary for the metabolism of Gal when it is the sole carbon source. This also supports the possibility of a link between the Leloir and T6P pathways as these strains both have an intact Leloir pathways that might have rescued growth in these conditions. Finally, we wanted to determine whether creating mutations in both the Leloir and T6P pathways (Δ*galR*Δ*lacD*) would reveal any alternative pathways for Gal metabolism in the pneumococcus (**Fig. 4.3**). When grown in CDM + Glc (**Fig.**
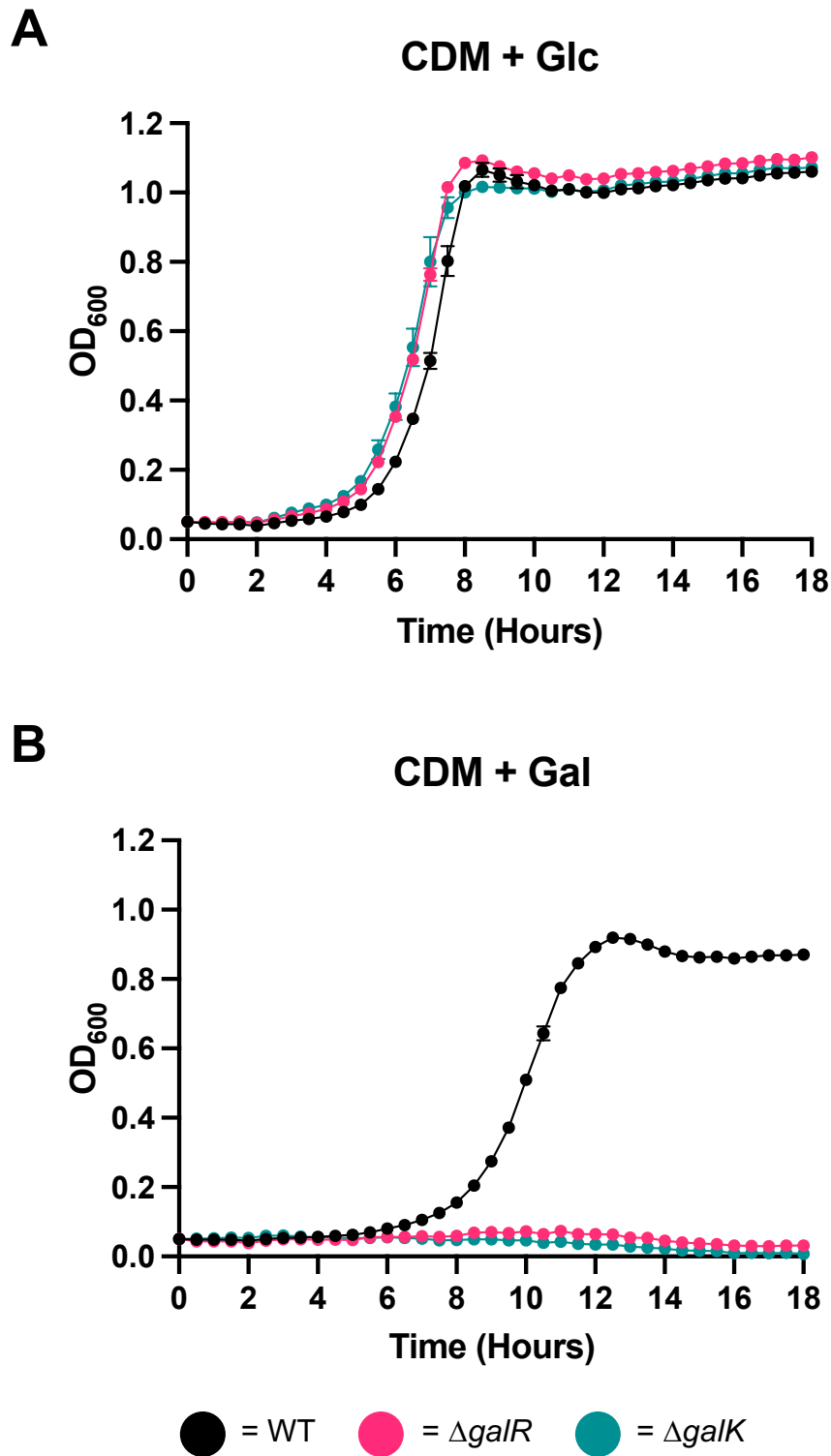
Figure 4.1. Impact of Leloir pathway mutation on growth in galactose
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

A

B
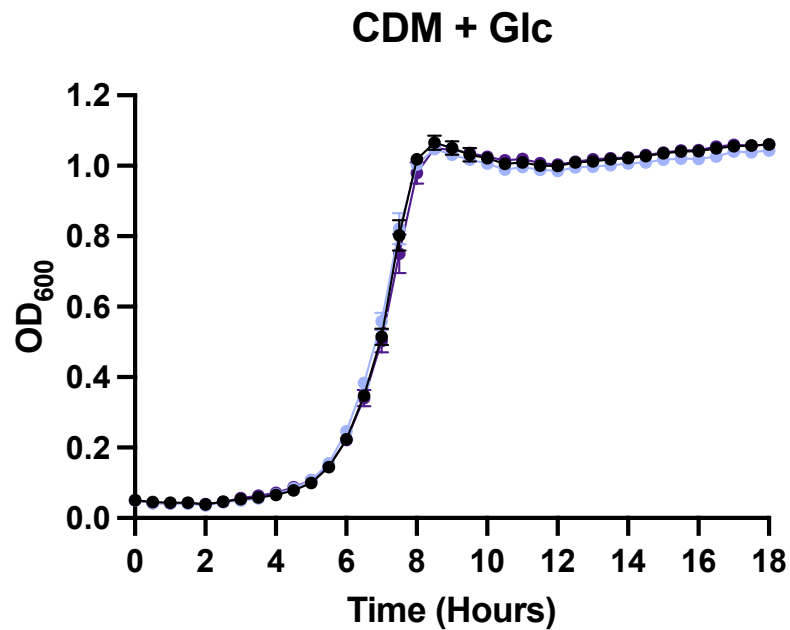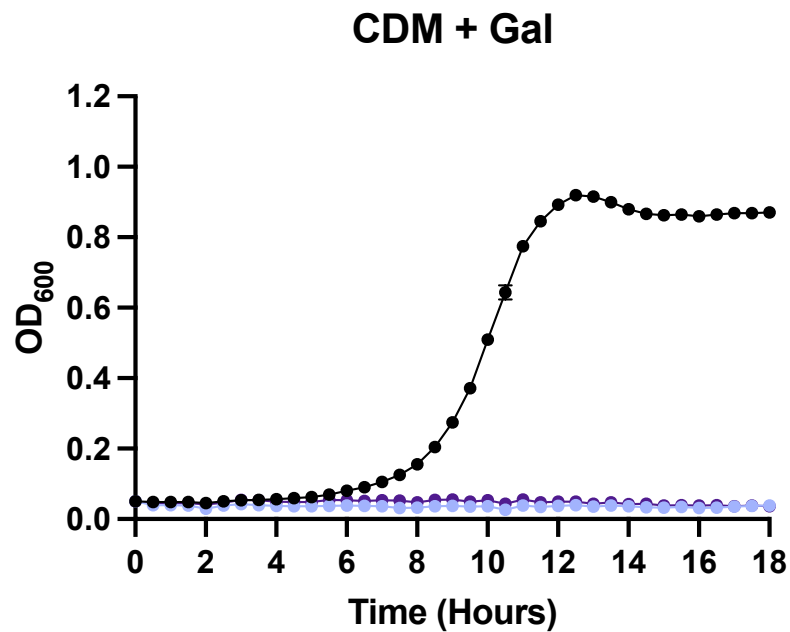
Figure 4.2. Impact of tagatose-6-phosohate pathway mutation on growth in galactose

The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

A

B

● = WT ● = $\Delta galR\Delta lacD$
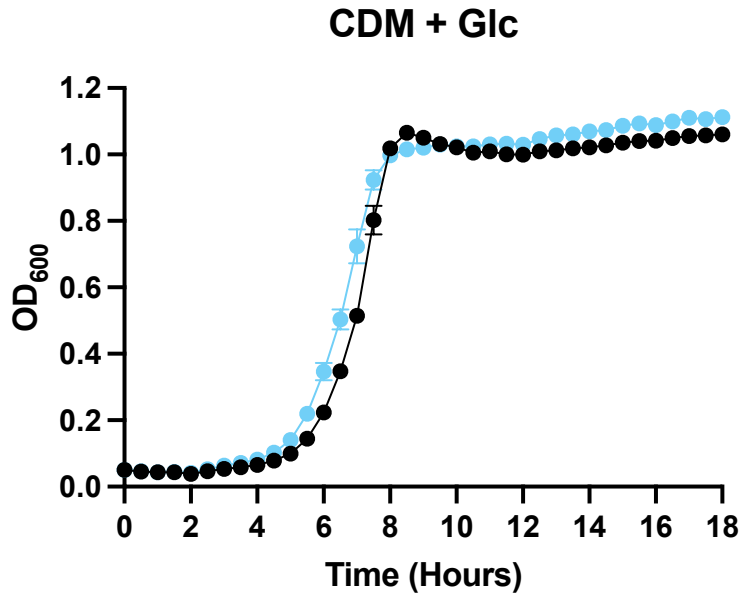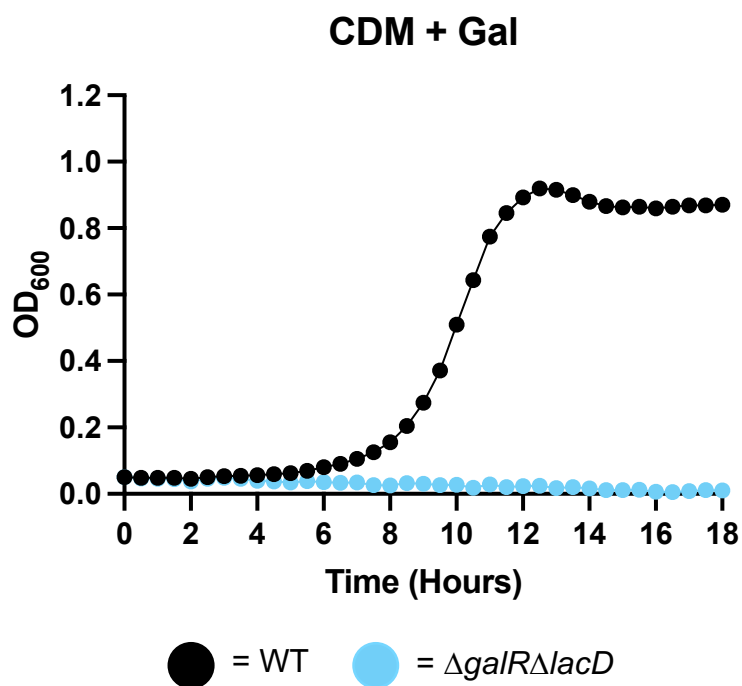
Figure 4.3. Impact of combined Leloir and tagatose-6-phosphate pathway mutation on growth in galactose

The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.

4.3A), the ∆*galR*∆*lacD* mutant grows comparably to the wild-type; however, it is unable to grow in CDM + Gal (**Fig. 4.3B**). Collectively, these findings indicate essential roles for both the Leloir and T6P pathways in the successful metabolism of Gal, given that they must both be functional in order to permit growth when Gal is the sole carbon source, indicating a degree of interplay between the two.

4.2.2. Contribution of GalR to the regulation of pneumococcal Gal metabolism

To further understand the apparent cross talk between the Leloir and T6P pathways, we analysed the expression of *galR*, *galK*, *lacA*, and *lacD* in D39, in the Leloir pathway mutants (∆*galR*, ∆*galK* and GalR$_{AAA}$), the tagatose pathway mutants (∆*lacAB*, ∆*lacD)* and the combined Leloir and T6P pathway double mutant (∆*galR*∆*lacD*) (**Fig. 4.4**). The expression of *galR* was significantly upregulated in both ∆*galK* and ∆*lacD* relative to D39, whereas it was unaffected in ∆*lacAB* (**Fig. 4.4A**). *galK* expression was unaffected in ∆*lacD*, but was significantly downregulated relative to D39 in ∆*lacAB* (**Fig. 4.4B**). Both *galR* and *galK* expression were also largely abrogated in ∆*galR* and GalR$_{AAA}$, confirming the requirement for functional GalR for activation of the Leloir pathway, as previously shown in Chapter 3. Unsurprisingly, the expression of these two genes was also abrogated in ∆*galR*∆*lacD*. *lacA* expression (**Fig. 4.4C**) was significantly upregulated in ∆*galK* relative to D39. In the remaining mutants, the expression of both *lacA* and *lacD* was downregulated relative to D39 (**Fig. 4.4C and 4.4D**). Collectively, these findings underscore the requirement for GalR (and the putative phosphorylation sites therein) in the activation and upregulation of the T6P pathway when the Leloir pathway is blocked. This could be due to feedback inhibition from accumulation of either Leloir or T6P intermediates, as has been shown for other species of
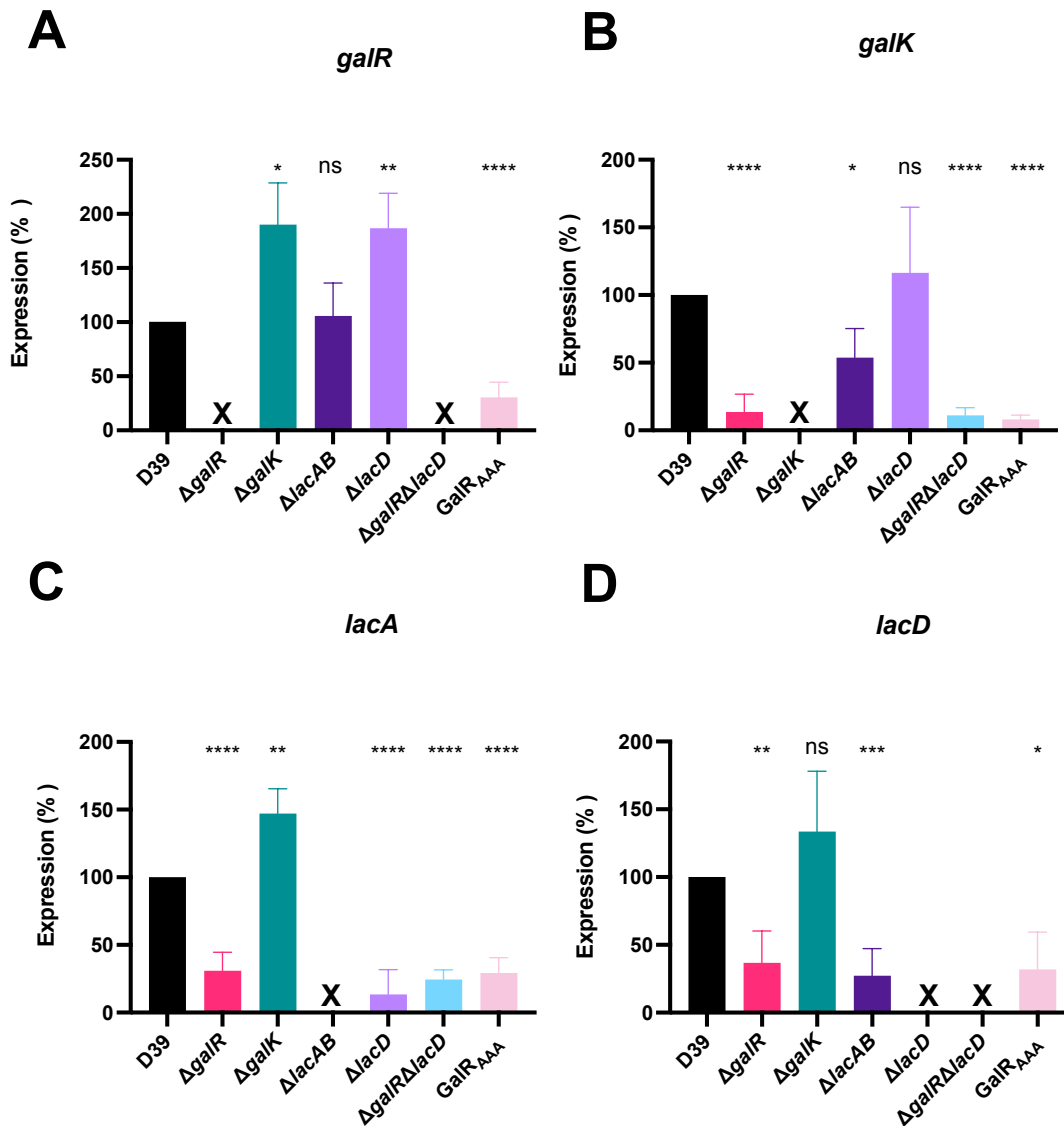
Figure 4.4. Impact of Leloir and tagatose-6-phosphate pathway deletion on regulation of galactose metabolism

Differential gene expression in Leloir and T6P pathway mutants. D39, Δ*galR*, Δ*galK*, Δ*lacAB*, Δ*lacD*, Δ*galR*Δ*lacD*, and GalR$_{AAA}$ were cultured overnight on blood agar, washed, and resuspended to a final OD$_{600}$ of 0.25 in CDM + Gal and incubated for 30 min. RNA was then extracted, and qRT-PCR was used to assess levels of *galR* (A), *galK* (B), *lacAB* (C), and *lacD* (D) mRNA, using *gyrA* as an internal control. Data presented are the mean ± standard deviation from three independent experiments, expressed as a percentage of that for D39. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$, by unpaired *t* test (relative to D39); ns, not significant; X, transcript absent due to gene deletion.
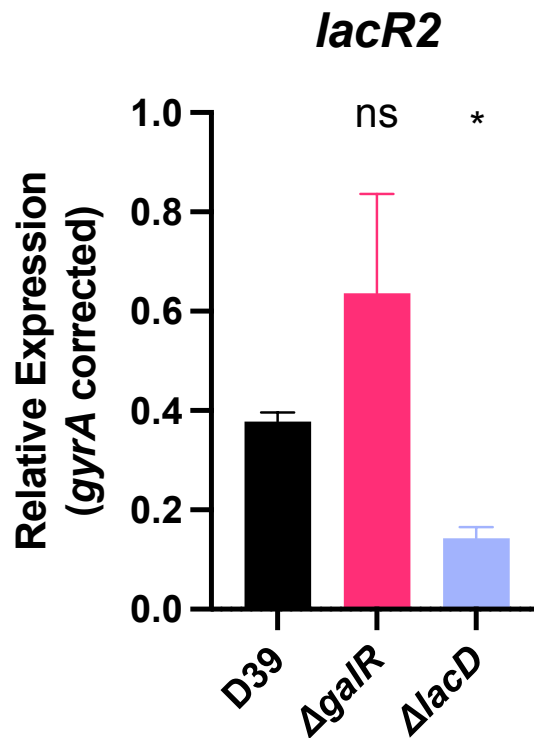
Figure 4.5. Assessing regulation of tagatose-6-phosphate pathway expression in response to galactose

D39, Δ*galR* and Δ*lacD* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. RNA was then extracted, and qRT-PCR was used to assess levels of *lacR2* mRNA, using *gyrA* as an internal control. Data presented are the mean ± standard error of the mean from three independent experiments, expressed as relative expression. *, $P < 0.05$; by unpaired *t* test (relative to D39); ns, not significant.

Streptococci (Zeng et al., 2010). We also assessed expression of *lacR2*, the repressor of the *lacI* operon (Afzal et al., 2014), in which we expected to see decreased expression in response to Gal. In the presence of Gal, there was a decrease in expression of *lacR2* in D39Δ*lacD* relative to D39, implying de-repression of the operon to promote upregulation of the T6P pathway genes. However, there was no significant difference in *lacR2* expression between D39 and Δ*galR*, indicating that effects of GalR on T6P pathway gene expression are not mediated via *lacR2* (**Fig. 4.5**)

4.2.3. Both the Leloir and Tagatose-6-Thosphate pathways are required for adherence to the nasopharyngeal epithelium

The ability to metabolise galactose in the human upper respiratory tract seems an essential feature for pneumococcal survival. To test this hypothesis, I performed an *in vitro* adherence assay using Detroit 562 nasopharyngeal epithelial cells. These cells were grown to confluency as explained in Section 2.11 and then incubated in the presence of either D39, Δ*galR,* Δ*lacD* or Δ*galR*Δ*lacD* pneumococci in CDM + Gal for two hours. Following this incubation, the number of adhered pneumococci were determined. There was a significant decrease in the adherence observed for the Δ*galR,* Δ*lacD* and Δ*galR*Δ*lacD* strains compared to D39, as shown in **Figure 4.6**. This potentially indicates that there is also a role for the T6P pathway in colonisation of the URT, not just the Leloir pathway and the putative GalR phosphorylation sites, as shown in Section 3.2.7.

4.2.4. Differentiating between intracellular and extracellular galactose

To investigate the metabolism of intracellular Gal in *S. pneumoniae* we examined the in *vitro* growth kinetics of Δ*galR*, Δ*lacD*, and the double mutant Δ*galR*Δ*lacD*
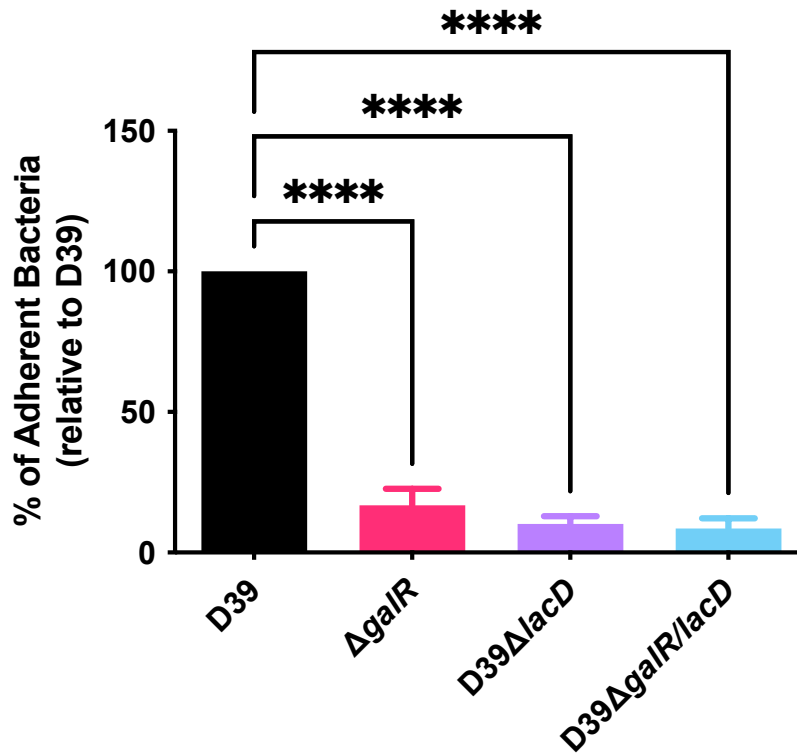
Figure 4.6. Assessing adherence of Leloir and tagatose-6-phosphate pathway mutants to Detroit 562 nasopharyngeal cells

D39, ΔgalR, ΔlacD and ΔgalRΔlacD were inoculated at a final OD_{600} of 0.2 onto Detroit 562 monolayers in CDM + 10% FCS + 0.5% Gal. Cells were then incubated for 2 hours at 37°C before being assessed for adherence. Each strain was assayed in technical triplicate over a minimum of two independent experiments. Data presented are the mean ± standard deviation, expressed as a percentage of adherent bacteria relevant to D39. ****, $P < 0.0001$; by One Way ANOVA (relative to D39).

in CDM with raffinose (Raf) as the sole carbon source. Raf is a trisaccharide composed of Gal, Glc, and fructose (Fru), which is internalised via an ABC transporter, after which the terminal Gal is released by an α-galactosidase encoded by the Raf operon *aga* gene (Hobbs et al., 2019). It has been previously reported that Gal is the intracellular inducer of the Leloir pathway in *S. mutans* (Zeng et al., 2010), although this is yet to be shown for *S. pneumoniae*. The Δ*lacD* strain exhibited growth kinetics in CDM + Raf nearly identical to that of the wild-type under the same conditions (Fig. 4.7). However, when analysing the Δ*galR* mutant, we observed that this strain exhibited a markedly slower growth phenotype than D39 and was not able to reach an equivalent culture density, confirming the role of the Leloir pathway in the metabolism of intracellular Gal. Interestingly, deletion of genes from both the Leloir and T6P pathways (Δ*galR*Δ*lacD* mutant) resulted in a strain with severe growth perturbation in CDM + Raf, indicating a role for these pathways in the utilisation of intracellular Gal.

We wanted to gain a better understanding of the transcriptional changes occurring in response to discrimination between intracellular and extracellular Gal. To do this, we examined the expression of both *galR* and *lacD* in key Leloir pathway mutants (Δ*galR* and Δ*galK*), T6P pathway mutants (Δ*lacAB* and Δ*lacD*) and combined deletion (Δ*galR*Δ*lacD*) mutants in the presence of either CDM + Gal or CDM + Raf (**Fig. 4.8**). Looking at *galR* expression (**Fig. 4.8A**), these levels remained consistent in both media, with the exception of the Δ*lacD* strain where there was a decrease in expression observed in the presence of Raf. Given that there is a 'functional' Leloir pathway present in this strain and there was only minor growth perturbation observed during growth in CDM + Raf, this was not unexpected. *lacD* expression was significantly decreased when strains were exposed to Raf, indicating that the Leloir pathway may play a more significant role in the utilisation of intracellular Gal.
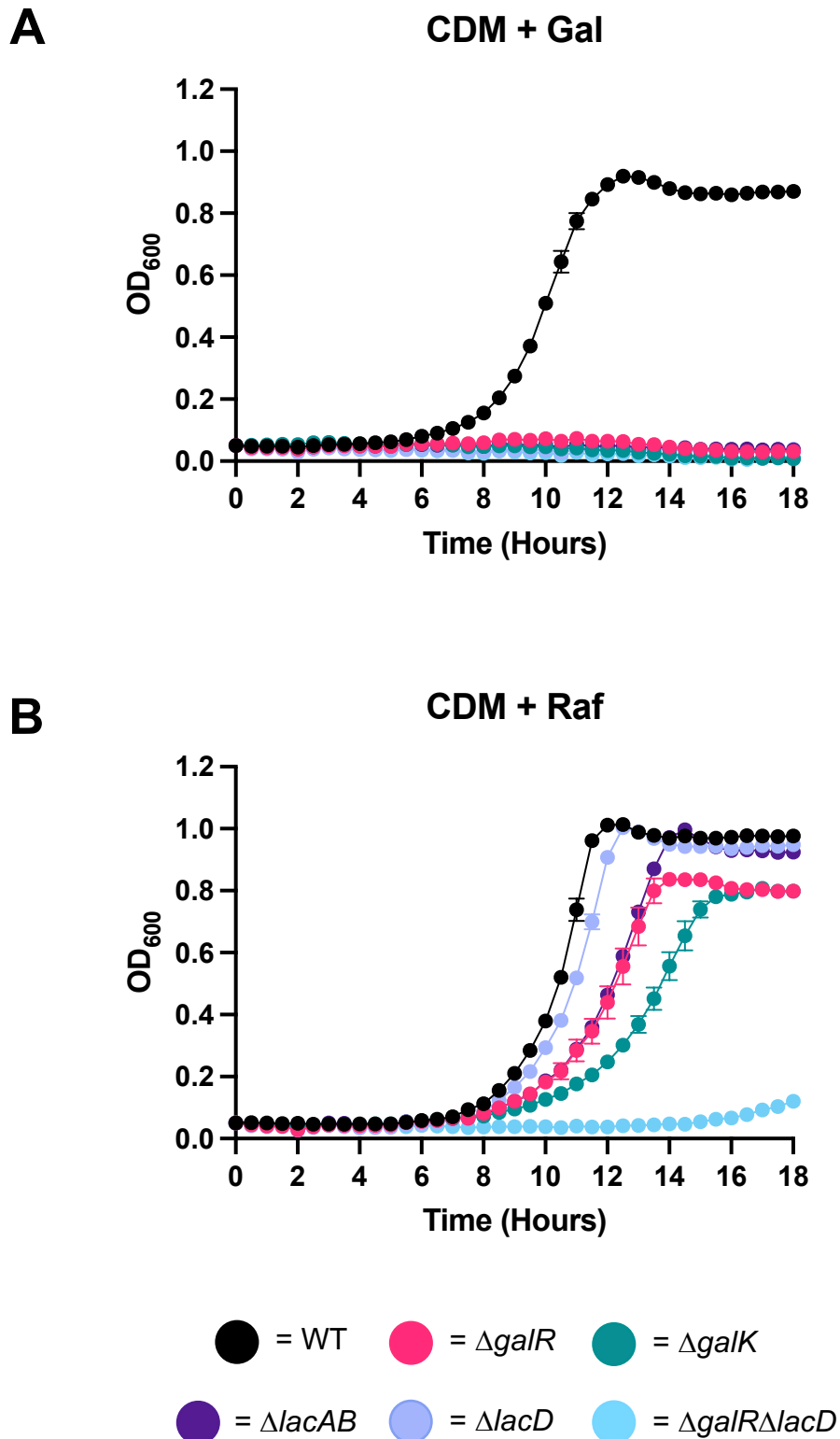
Figure 4.7. Impact of Leloir and tagatose-6-phosphate mutation on utilisation of intracellular galactose
The indicated strains were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring OD_{600} every 30 min for a total of 18 h. Data points are mean OD_{600} of triplicate assays.
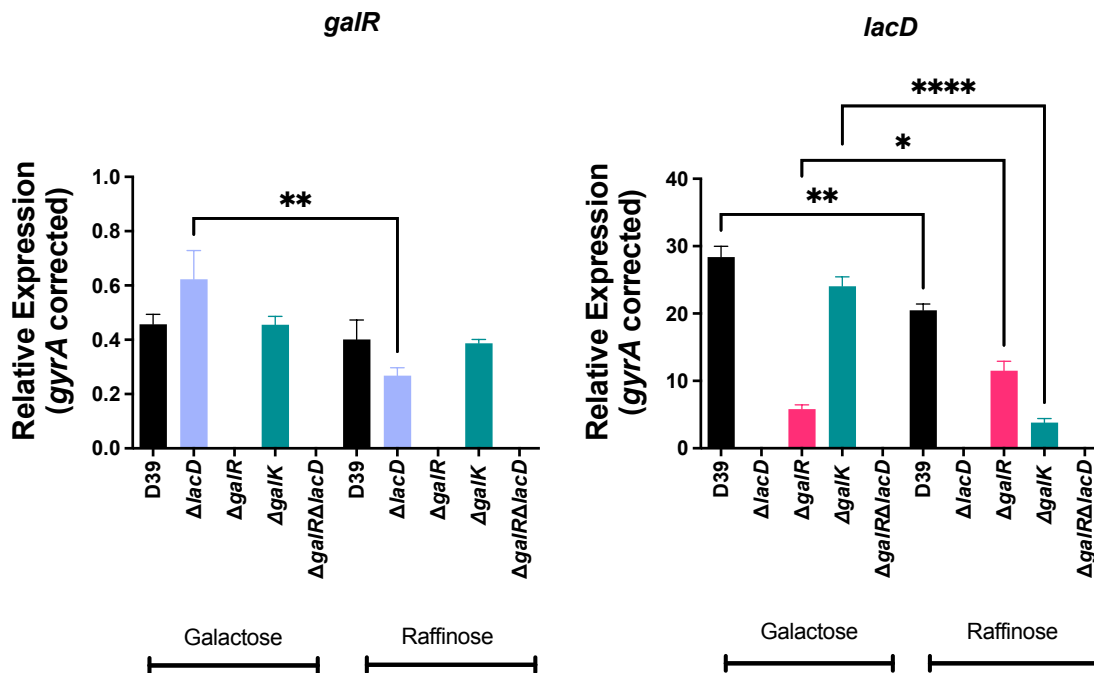
Figure 4.8. Regulation of the Leloir and tagatose-6-phosohate pathways in response to intracellular galactose

D39, D39Δ*galR*, D39Δ*galK*, D39Δ*lacD* and D39Δ*galR*Δ*lacD* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal or CDM + Raf and incubated for 30 min. RNA was then extracted, and qRT-PCR was used to assess levels of *galR* and *lacD* mRNA, using *gyrA* as an internal control. Data presented are the mean ± standard error of the mean from three independent experiments, expressed as relative expression. *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$, by One Way ANOVA (Multiple Comparisons).

This is further confirmed by the growth phenotypes observed in **Fig. 4.7B**. Further, the decreased *lacD* expression (**Fig. 4.8B**) in the ΔgalK strain compared to ΔgalR aids in supporting this theory, as the presence of *galR* would initially signal that the Leloir pathway is present and functional. With time, the decreased ability of this strain to metabolise Raf to Gal and its other intermediates would become apparent, as is seen in **Fig. 4.7B**. Collectively, these findings indicate that while the Leloir pathway may play a key role in the utilisation of intracellular Gal, it is actually both the Leloir and T6P pathways that are responsible for growth in these conditions.

4.2.5. Utilisation of carbon sources differs between Leloir and Tagatose-6-Phosphate pathway mutants

Given the apparent link between the Leloir and T6P pathways that has been identified through growth and transcriptional analyses, we wanted to investigate whether there was differential metabolism of any other carbon sources between the wild-type and the mutants ΔgalR, ΔlacAB and ΔlacD. To do this, we conducted phenotypic microarrays as described in Section 2.18, which allows for visualisation of the ability to metabolise 96 different carbon sources. As shown in **Table 4.1**, there were distinct differences observed.

There was a subset of carbon sources that could be metabolised by all four test strains, these being L-arabinose, N-acetyl-G-glucosamine, D-trehalose, D-mannose, D-ribose, D-fructose, α-D-glucose, maltose, α-D-lactose, sucrose, maltotriose, D-cellobiose, N-acetyl-β-D-mannosamine and L-lyxose. There was a small subset of carbon sources that could exclusively be metabolised by the wild-type, including D-Gal, glycerol, L-fucose, D-glucuronic acid and L-rhamnose. There were an additional five carbon sources that were differentially metabolised between

Table 4.1. Overview of differences in carbon utilisation in *S. pneumoniae*

Metabolite	D39	ΔgalR	ΔlacAB	ΔlacD
L-Arabinose	+	+	+	+
N-Acetyl-G-Glucosamine	+	+	+	+
D-Galactose	+	-	-	-
D-Trehalose	+	+	+	+
D-Mannose	+	+	+	+
Glycerol	+	-	-	-
L-Fucose	+	-	-	-
D-Glucuronic Acid	+	-	-	-
D-Xylose	+	+	-	+
D-Ribose	+	+	+	+
L-Rhamnose	+	-	-	-
D-Fructose	+	+	+	+
α-D-Glucose	+	+	+	+
Maltose	+	+	+	+
α-D-Lactose	+	+	+	+
Lactulose	+	+	-	-
Sucrose	+	+	+	+
β-Methyl-D-Glucoside	+	-	-	+
Maltotriose	+	+	+	+
D-Cellobiose	+	+	+	+
Acetoacetic Acid	+	-	-	+
N-Acetyl-β-D-Mannosamine	+	+	+	+
D-Psicose	-	-	+	-
L-Lyxose	+	+	+	+

the mutant strains. D-xylose could be metabolised by D39, $\Delta galR$ and $\Delta lacD$ but not

$\Delta lacAB$. Lactulose could be metabolised by D39 and $\Delta galR$ but not $\Delta lacAB$ and

$\Delta lacD$. β-D-methyl-glucoside and acetoacetic acid could only be metabolised by

D39 and $\Delta lacD$. Finally, $\Delta lacAB$ was the only strain able to metabolise D-psicose.

The remaining carbon sources present in the microarray plate were unable to be

metabolised by any of the strains tested. The use of phenotypic microarray revealed

differences in the ability of different Leloir and T6P pathway mutants to metabolise

varied carbon sources. Given that the intracellular accumulation of metabolic

intermediates has been shown to be toxic in other *Streptococci* (Zeng et al., 2010),

we wanted to further investigate these differences using a more comprehensive

metabolomic approach.

4.2.6. Deletion of key tagatose-6-phosphate pathway genes alters the metabolome in the presence of Gal

To gain insight into whether any metabolites/intermediates were toxically

accumulating intracellularly in the various mutants, cells were analysed using Liquid

Chromatography Mass Spectrometry and Gas Chromatography Mass Spectrometry

at Metabolomics Australia by Dr. David De Souza and Dr. Brunda Nijagal. In this

pilot study, two strains were sent for analysis: D39 and $\Delta lacAB$. Strains were firstly

grown on BA overnight and resuspended to a final OD_{600} of 0.25 in CDM + Gal

before being incubated at 37°C for 30 min. Bacterial pellets were then harvested as

explained in Section 2.13 and sent for analysis.

As can be seen in **Figures 4.9.1** and **4.9.2**, there were 3 metabolites reported to be

significantly more abundant in $\Delta lacAB$ relative to D39. These were Glc-6-P, D-

Glyceraldehyde-3-P and Sedoheptulose-7-P. In addition, there were 7 metabolites
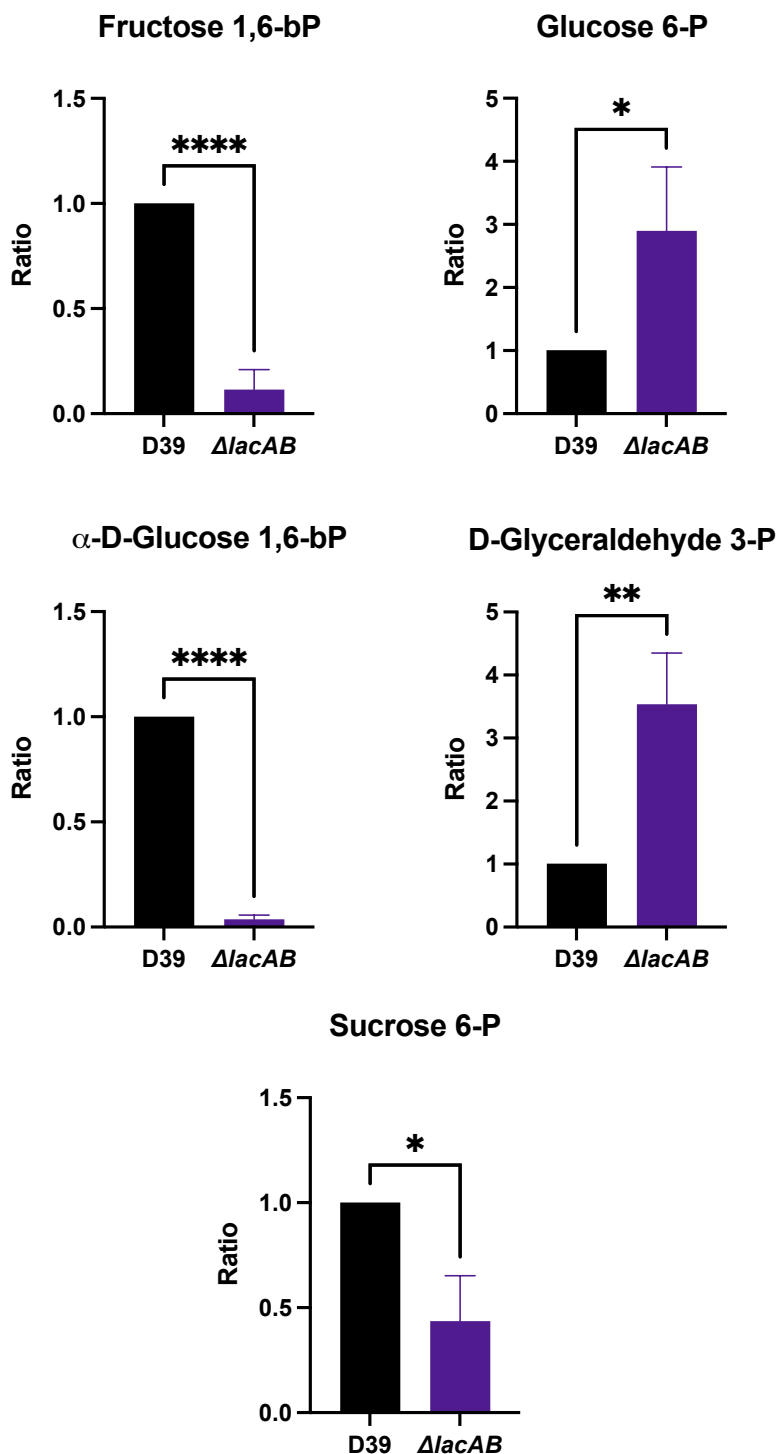
Figure 4.9.1. Assessing significant changes in the metabolome in response to tagatose-6-phosphate pathway mutation

D39 and Δ*lacAB* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. Cells were then infused in cold PBS and pelleted prior to snap freezing. Pellets were then shipped to Metabolomics Australia for analysis. Metabolite level in the mutant strain is expressed as a ratio relative to that in D39. *, $P < 0.05$; **, $P < 0.01$; ****, $P < 0.0001$, by unpaired *t* test.
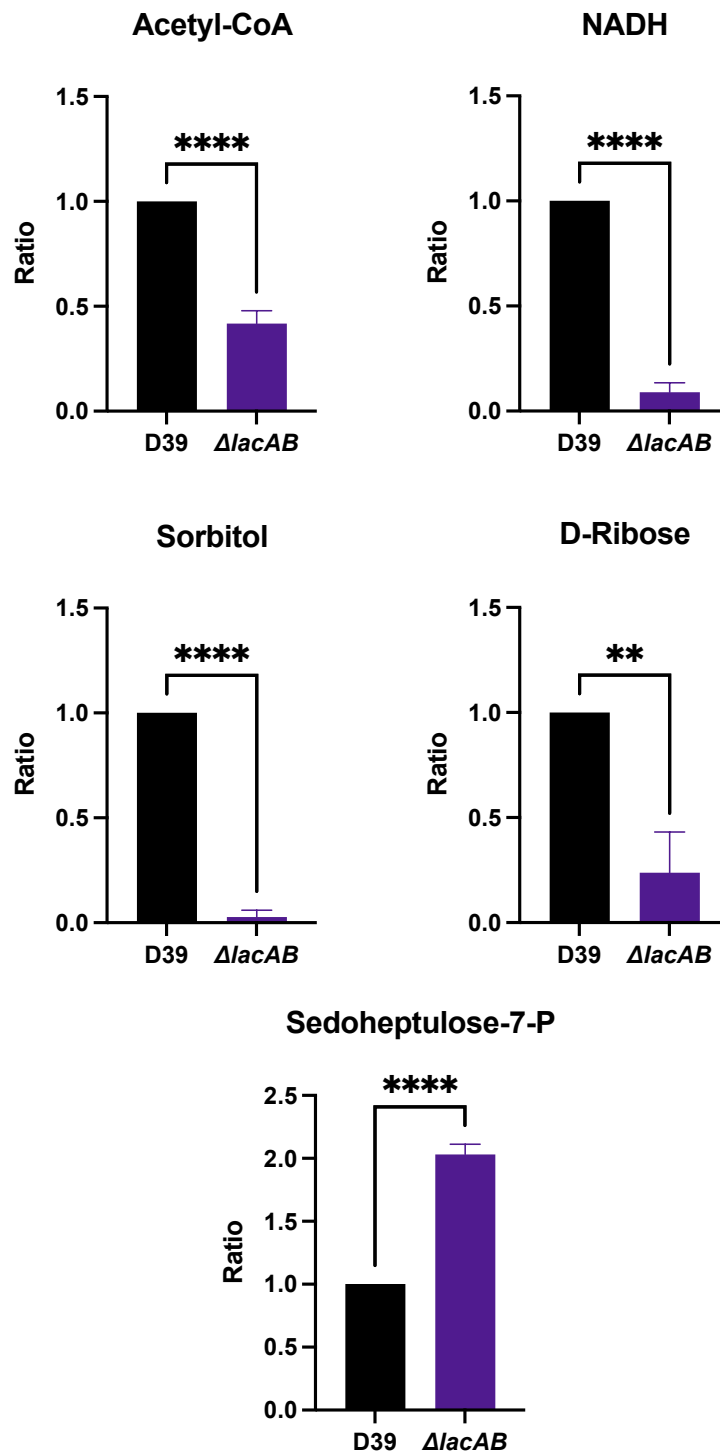
Figure 4.9.2. Assessing significant changes in the metabolome in response to tagatose-6-phosphate pathway mutation

D39 and Δ*lacAB* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. Cells were then infused in cold PBS and pelleted prior to snap freezing. Pellets were then shipped to Metabolomics Australia for analysis. **, $P < 0.01$; ****, $P < 0.0001$, by unpaired t test.

reported to be significantly less abundant in Δ*lacAB* compared to D39, these being Fru-1,6-bP, α-D-Glc-1,6-bP, Sucrose-6-P, Acetyl CoA, NADH, Sorbitol and D-Ribose. While some of these metabolites are found within the T6P pathway, they can also be associated with other metabolic pathways. Non-significant metabolomics results can be found in **Appendix A**.

Given that there were these changes observed within the metabolome of the Δ*lacAB* strain, we wanted to better understand what this may mean for intracellular metabolite accumulation within the strain. To do this, we utilised a schematic diagram of key metabolic pathways within *S. pneumoniae* that had been previously generated by Dr. Erin Brazel (**Figure 4.10**). In doing so, we can highlight exactly where the mutation of Δ*lacAB* impacts metabolism and how these differences in metabolite abundance may be linked. Of the ten metabolites identified as being significantly different in their abundance between D39 and Δ*lacAB*, five were able to be pinpointed. These were Sedoheptulose-7-P, Glc6P, D-Glyceraldehyde-3-P, Fru-1,6-bP and Acetyl-CoA. Sedpheptulose-7-P was increased in the Δ*lacAB* strain compared to D39, and is an intermediate of the Pentose Phosphate pathway, along with nucleotide and aromatic amino acid biosynthesis. This increase may be expected in response to the lack of conversion of Gal6P to T6P in this strain. Glc6P, D-Glyceraldehyde-3-P and Fri-1,6-bP are intermediates of the glycolytic pathway in the pneumococcus; however, they are also central to many other essential metabolic pathways as shown in **Figure 4.10**. Finally, Acetyl-CoA was significantly less abundant in the Δ*lacAB* strain compared to D39. This is an essential intermediate for both pyruvate metabolism and cell membrane biosynthesis. This decrease could again be explained by the lack of conversion of Gal6P to T6P. However, given the distance of these two intermediates from each other in their
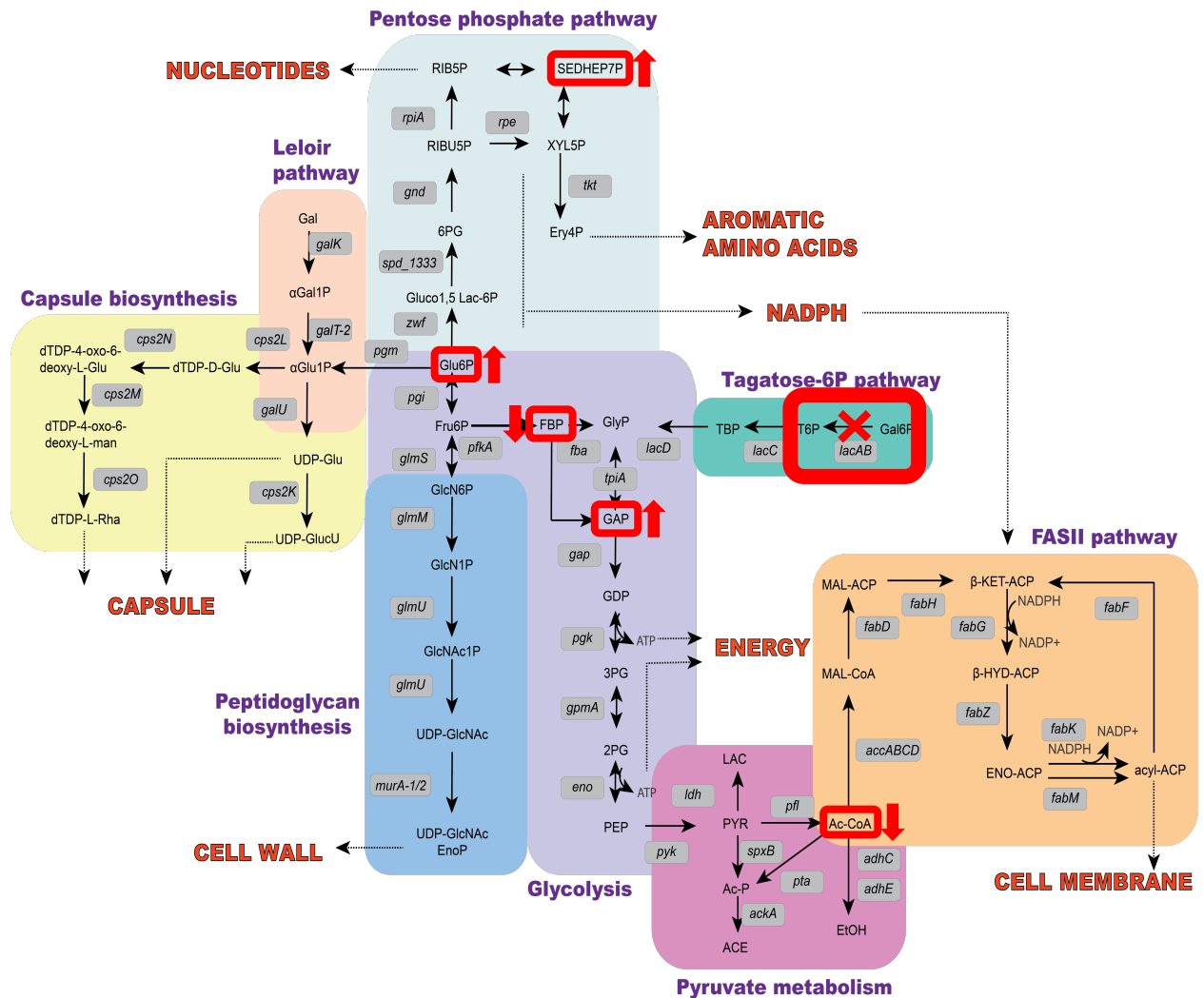
Figure 4.10. Overview of metabolic changes occurring in ΔlacAB

Schematic diagram previously generated by Dr. Erin Brazel. Metabolites that were significantly up- or down-regulated in abundance in comparison to D39 are indicated by a red rectangle surrounding the name and an arrow indicating whether it was up- or down-regulated. Additionally, the location of the mutation within the ΔlacAB strain relative to the pathways is shown as a red rectangle with an 'X' above the location of LacAB.

respective pathways, much like all the metabolites identified in this approach, there would need to be additional studies conducted to confirm these changes.

4.3. Discussion

This chapter has focussed on delving deeper into the interaction between the Leloir and T6P pathways through use of mutagenesis, gene expression and metabolism studies. An unexpected finding of the current study was that all the investigated GalR mutants exhibited significantly reduced expression of the T6P pathway gene *lacD* as shown in Section 3.2.6, indicating that GalR has some effect on the T6P pathway, either directly or indirectly. This may provide a mechanistic basis for a previously proposed subtle regulatory link between the Leloir and T6P pathways (Paixão et al., 2015). This was further examined by comparing the growth and gene expression phenotypes of D39 mutants with deletions in *galR*, *galK*, *lacAB*, *lacD*, and *galR* plus *lacD*. Unlike growth in CDM + Glc, none of these mutants were capable of growth in CDM + Gal (**Fig.4.2 – 4.4**) indicating that both the Leloir and T6P pathways are essential for growth when Gal is the sole carbon source.

Other groups have previously attempted to determine the roles and responsibilities of these pathways in Gal metabolism; however, the area still remains poorly elucidated. A previous study of a similar D39 *galK* deletion mutant (Afzal et al., 2015) reported that growth in the presence of Gal was reduced relative to wild-type D39, but not completely abrogated as shown in the present study. This discrepancy may be attributable to the use of a more nutrient-rich M17 medium in that study (Afzal et al., 2015), as discussed in Chapter 3. A separate study reported complete abrogation of growth of a D39 *galK* deletion mutant in Gal, whereas a *lacD* deletion mutant was able to grow logarithmically in Gal, albeit after a long lag phase (Paixão et al., 2015).

In the present study, gene expression analyses have shown that maximal expression of the T6P pathway genes *lacA* and *lacD* require functional GalR and that all three putative phosphorylation sites of GalR are necessary to achieve this function (**Fig. 4.4**). Additionally, there appears to be a link between *galK* and *lacAB* gene expression. Deletion of *galK* upregulated *lacA* expression, perhaps as a consequence of the upregulation of *galR* in this mutant. On the other hand, deletion of *lacAB* significantly reduced *galK* expression, but did not impact expression of *galR*. Furthermore, *galR* expression was significantly elevated in the *lacD* mutant. Thus, there is a complex interplay between the Leloir and T6P pathways in the various mutants, potentially mediated by intracellular concentrations of intermediates or end products of either pathway.

Given that the regulator of the T6P pathway in the pneumococcus is known, we wanted to assess the changes in expression that occur in this regulator in the presence of Gal. LacR2 acts as repressor of the *lacABCD* operon (Afzal et al., 2014), indicating that we should observe decreased expression in the presence of Gal. As expected, in the case of the ΔlacD strain, there was decreased *lacR2* expression relative to D39, indicating de-repression of the operon to utilise Gal as a carbon source. Interestingly, in the case of the ΔgalR strain, there was an increase in *lacR2* expression relative to D39. Given that the ΔgalR strain is unable to grow in Gal, it might be expected that there would be de-repression of the *lacABCD* operon in response to this, promoting utilisation of Gal, but this is not the case. While these findings don't reveal a mechanism whereby these two pathways may be linked, it certainly adds to the evidence for these two pathways being linked.

We next wanted to assess the ability for the pneumococcus to utilise intracellular vs. extracellular Gal. To do this, Raf was used as the sole carbon source in both growth and gene expression assays. Raf is a trisaccharide sugar composed of Glc,

Gal and Fru. Raf is a plant-derived sugar that is not found in the human host, but it can be obtained from dietary sources such as starchy fruits and vegetables. Raf is typically imported via an ABC transporter, in which the substrate binding protein RafE delivers Raf to the transporter comprising RafEFG (Hobbs et al., 2019). Once inside the cell, the alpha-galactosidase encoded by *aga* cleaves Raf into sucrose and Gal. The sucrose is then further broken down by GtfA to Fru and Glc-1-P. The generation of these metabolites provides options for the pneumococcus in terms of which sugars to metabolise first. Given that Gal is the first sugar liberated by this pathway, it provides a good option for determination of intracellular vs. extracellular Gal metabolism. Utilising the same Leloir and T6P pathway mutants, we saw interesting growth results in the presence of Raf. Looking firstly at the T6P pathway mutants, the $\Delta lacD$ strain was able to grow comparably to the wild-type, whereas the $\Delta lacAB$ strain showed a slight growth delay, but was nevertheless able to reach the same final OD_{600} as both the $\Delta lacD$ and wild-type strains. Interestingly, the Leloir pathway mutants ($\Delta galR$ and $\Delta galK$) both exhibited delayed growth and an inability to reach the same OD_{600} as the wild-type and T6P pathway mutants. This indicates that the Leloir pathway may play a greater role in the metabolism of intracellular Gal. This is also supported by the overall decrease in total *lacD* expression in all assessed strains in the presence of Raf. If we look at this in terms of the import of Gal, the T6P pathway utilises a PTS transporter for import, resulting in phosphorylated Gal accumulating intracellularly, whereas the Leloir pathway is proposed to use an ABC transporter, leaving the galactokinase GalK to phosphorylate Gal in the early stages of metabolism. Given that Raf is also imported via an ABC transporter and the cleaving of Raf via Aga doesn't result in phosphorylation, it is likely that this pathway would be used in preference to the T6P pathway. It is important to note that while this assay was designed to assess growth

with intracellular Gal, there are other 'more preferred' carbon sources liberated in this pathway such as Fru and Glc1P, meaning growth may be a result of CCR enabling growth on these sugars as opposed to Gal. While this may be the case for the single mutants, the inability for the ΔgalRΔlacD double mutant to grow in Raf indicates that both the Leloir and T6P pathways are required to metabolise Gal that is liberated from the metabolism of Raf.

The question remains as to whether there is indeed interplay between the Leloir and T6P pathways. It was clear by this stage that we would not gain answers from mutagenesis studies and that a more in-depth intracellular approach would be required to elucidate this link. In the first instance, we performed a phenotypic microarray that would allow for the visualisation of metabolism of 95 different carbon sources. This revealed differences in the ability for D39, ΔgalR, ΔlacAB and ΔlacD to metabolise different carbon sources. Unsurprisingly, D39 was the only strain capable of metabolising galactose, further confirming the findings of our growth studies. While there was a large proportion of carbon sources that were unable to be metabolised by any of the strains, there was a subset that was differentially metabolised between the mutants, including lactulose, D-xylose and acetoacetic acid. While none of these carbon sources directly link to Gal metabolism in the pneumococcus, these results do confirm that there are some key differences occurring in the metabolism of these mutants. However, there are still questions regarding the link between the Leloir and T6P pathways. Given the complexity of these two pathways, metabolomics was suggested as a way to gain a broad snapshot of the changes occurring in metabolite abundance in the presence of Gal in response to Leloir and/or T6P pathway mutation. While time was a limiting factor, we were able to gain some insight into the differences occurring between both the wild-type and the ΔlacAB mutant in the presence of Gal. One of the key metabolites

that we wanted to assess here was Gal-6-phosphate (Gal6P), given that this should accumulate in the Δ*lacAB* strain in the absence of the Gal6P isomerase encoded by this gene. If Gal6P and other metabolites were found to be detectable, we planned to then send additional samples for testing to better elucidate the network of metabolites that may be accumulating in different mutants. One of the first issues encountered was the ability to accurately resolve Gal6P from other sugars, particularly T6P and Fru-6-phosphate (F6P). Unfortunately, these three sugars would co-elute with each other on the LC-MS, meaning that while we could gain understanding of crude differences in abundance between strains, these differences could not be accurately attributed to specific sugars. Taking this into account, there were other metabolites that could be assessed, which yielded some rather interesting results. Specifically, we saw significant differences in the abundance of both sorbitol and D-glyceraldehyde 3-P, both of which are metabolites involved in pneumococcal galactose metabolism. While the link to sorbitol is difficult to decipher, the link to D-glyceraldehyde 3-P is much clearer. The deletion of Δ*lacAB* means the cell can no longer convert Gal-6-P to T6P through the T6P pathway. Theoretically, this would result in decreased T6P by-products, one of which is D-glyceraldehyde 3-P that is generated as a result of the tagatose 1,6-diphosphate aldolase LacD. In this study, however, D-glyceraldehyde 3-P was found to be present in 3.5 times greater abundance within the Δ*lacAB* mutant compared to D39. If this were the only way for the cell to generate the tagatose 1,6-diphosphate required to generate D-glyceraldehyde 3-P, then this result would make little sense, but it is not the only way. It would be logical to theorise that these cells would be under an element of stress when in the presence of CDM + Gal at the time of their harvesting, meaning the cell would be using any means necessary to attempt to metabolise this sugar, even bypassing the need for LacA and LacB to get the job

done. In addition to these metabolites, there were significant differences observed in other metabolites associated more broadly with either carbon metabolism or the Pentose Phosphate pathway, as shown in **Figure 4.10**. Upon further investigation, it became apparent that there were even more Pentose Phosphate pathway metabolites that were being differentially metabolised in the ∆*lacAB* mutant, specifically, Fru-1,6-bP, Glc6P, D-Ribose and Sedoheptulose-7-P. This led to the discovery of these metabolites often overlapping with other significant metabolic pathways such as glycolysis, particularly in the case of Fru-1,6-bP, Glc6P and D-Glyceraldehyde-3-P. Metabolism is incredibly complex, and these results certainly highlight this. While it is possible to think of the relationship between metabolite abundance and *lacAB* deletion in a closed system, it does not reflect what is truly occurring intracellularly. As such, any significant conclusions drawn from this data would need to go through additional validation prior to reporting. This, however, does not preclude us from commenting on interesting differences that were observed. Perhaps most interestingly of these was the significant decrease in abundance of NADH in the ∆*lacAB* strain. Reduced Nicotinamide Adenine Dinucleotide molecules (NADH) play important roles in various biological processes across all levels of life. A previous study by (Afzal et al., 2018) investigated the transcriptional responses that occurred in the *S. pneumoniae* D39 strain in the presence of NADH. They found that there was increased expression in a range of genes, however there were two identified that is of interest in this study. They observed increased expression of both *fba, gapN* and *gap* genes in response to NADH. *fba* encodes a fructose-bisphosphate aldolase which converts D-Frc-1,6-bP to D-glyceraldehyde-3-P, while the *gap* genes encode glyceraldehyde-3-phosphate dehydrogenases which convert D-glyceraldehyde-3-P to the energy rich intermediate glyceraldyhyde-1,3-bisphosphate. This study proves the link between

NADH and levels of Frc-1,6-bP and D-glyceraldehyde-3-P that may be present intracellularly. We also observed decreased levels of Acetyl-CoA in the $\Delta lacAB$ mutant compared to D39. Acetyl-CoA is another important metabolite within the pneumococcus that is associated with a range of cellular processes. Studies have linked the availability of Acetyl-CoA intracellularly to capsule biosynthesis (Echlin et al., 2016) and also to the altered flux of pyruvate metabolism depending on the availability of carbon sources (Echlin et al., 2020). These studies again underpin the conclusion that metabolism is an incredibly complex process that, in the context of this study, requires additional studies to elucidate.

While only a pilot study, these results show that it would be worthwhile to investigate the differences in the metabolome present in all Leloir and T6P pathway mutants discussed in this chapter. While we still don't have a clear answer as to how these two pathways may be linked, there is now evidence to suggest that there may intracellular accumulation of metabolites occurring within these strains. If this accumulation were toxic to the cell, it may result in an inability for the mutants to grow in the presence of Gal, leading to cell death and the subsequent growth phenotypes we observed in **Figures 4.1** to **4.3**.

Unfortunately, due to the SARS-CoV-2 pandemic across 2020 and 2021, we were unable to continue with additional metabolomics experiments as planned. Had we been able to, the plan was to send the additional Leloir and T6P pathway mutants, including GalR$_{AAA}$ for analysis. While it was likely that we would have continued to face issues with attempting to resolve Gal6P from other sugar species, we still would have been able to gain insight into the differences in metabolite abundance occurring between strains.

We can conclude at this point that the Leloir and T6P pathways are likely not discreet as previously reported in the literature (Afzal et al., 2014, Afzal et al., 2015,

Paixão et al., 2015). The combination of growth and transcriptional analyses conducted in this chapter, along with assessing the difference in metabolism of D39 and the Δ*lacAB* mutant reveals that there is a definitive link between these two pathways. In future, it would be ideal to repeat metabolomics analyses of D39 and the Δ*lacAB* strain, along with other Leloir and T6P pathway mutants, to allow for comprehensive comparison of the abundance of key metabolites between mutants to shed light on exactly how these two pathways are able to influence each other.

Chapter 5: Assessing the role of GalR and the putative phosphorylation sites *in vivo*

5.1. Introduction

The findings of this thesis so far have shown an important role for GalR not only in the regulation and utilisation of Gal, but also in virulence. The findings from Chapter 3 revealed that while deletion of *galR* itself may not have significant impact on bacterial loads in the lungs, substitution of the putative GalR phosphorylation sites does. It also revealed a critical role for these sites in Gal metabolism. Chapter 4 built on these findings, showing that there is a likely link between the Leloir and T6P pathways, with both being required for the successful metabolism of Gal. These findings also indicated that GalR may be responsible for this regulation.

In this chapter, we wanted to explore the impact that GalR and its putative phosphorylation sites may have during infection and determine the transcriptional changes occurring in both the host and the pathogen. A previous study from our group was among the first to utilise the technique 'dual *in vivo* RNAseq' in a murine model of infection. This cutting-edge approach allows for simultaneous analysis of the transcriptomes from both the pneumococcus and the host in a particular tissue, which in this case was the lung. This led to the conclusion that single nucleotide polymorphisms within the regulator of the *raf* operon, RafR, resulted in transcriptomic differences in both pathogen and host that led to differential neutrophil recruitment during infection and distinct tissue tropism (Minhas et al., 2020). Here, we employed the same approach with the strains D39, Δ*galR* and GalR$_{AAA}$ in a murine model of infection, harvesting RNA from the lungs and performing RNA sequencing. By using this broad-spectrum approach, the aim was to identify any genes that were differentially expressed in both genomes during infection, possibly identifying key differences in the immune responses to each of the three different strains.

5.2. Results

5.2.1. Comparative Host/Pathogen Transcriptomics

As shown in Chapter 3, there are differences in the abundance of D39 and GalR$_{AAA}$ in the lungs at 24 h post-infection. Before commencing this study, we needed to determine an appropriate time point to harvest RNA for dual *in vivo* RNAseq. Having previously assessed bacterial loads at 24 h for the three test strains (D39, $\Delta galR$ and GalR$_{AAA}$), we reassessed this data to see if there were any major differences in abundance at this time point. **Figure 3.14** shows that while there was a difference in the abundance of these strains at 24 hours, they were small enough that any changes observed in downstream experiments would likely not be due to a dose effect. Additionally, these bacterial loads should be sufficient to yield ample RNA for sequencing, in theory.

Accordingly, mice were challenged as described in the Section 2.14. In brief, groups of 12 mice were anaesthetised and challenged with 1 x 10^7 CFU of either D39, $\Delta galR$ or GalR$_{AAA}$. At 24 hours post-infection, the mice were humanely euthanised and the lungs harvested, washed and placed in Trizol before proceeding with downstream processing (see Section 2.15). Once the RNA was extracted from each lung sample, RNA quality was assessed using a Nanodrop Spectrophotometer (Thermofisher Scientific, Waltham, Massachusetts) and subsequently pooled into three replicates, such that each replicate was representative of 4 mice. The RNA was then sent to the Australian Cancer Research Foundation (ACRF) Cancer Genomics Facility for sequencing and analysis. A more detailed explanation of the methods used by ACRF can be found in Section 2.15.

Once sequencing was completed, the reads generated were firstly subject to an initial 'clean up' pipeline to remove adaptor sequences and trim reads, as explained in the Section 2.16. Following this, reads were then mapped back to the

pneumococcal genome as shown in **Table 5.1**. For D39, there was an average of 6,345 reads that mapped back to the pneumococcal genome. For Δ*galR*, there were on average 19,515 successfully mapped reads and 16,792 reads on average for GalR$_{AAA}$. While there was no threshold set for the number of reads that needed to be successfully mapped back to the genome to proceed with analyses, this was substantially lower than the approximately 1 – 2 million reads mapping back to the pneumococcal genome in the dataset used by Minhas et al., (2020). This was not improved by adjusting the programs used for adaptor removal and read trimming. Notwithstanding this deficiency, we decided to proceed with downstream analyses, as detailed in the Section 2.16, to determine if there were any major differences in the *in vivo* transcriptomes of the three strains.

5.2.2. Pneumococcal transcriptional changes occurring in response to infection

There were five genes identified as being differentially expressed in the pneumococcal genome – *adhA*, *phtE*, *pflA* (SPD_1774)*, strH* (SPD_0063) and *blpB*. Given that these genes reached significance within the RNAseq dataset, we immediately sought to confirm these results using qRT-PCR. Using the same RNA that was sent for RNA sequencing, we performed qRT-PCR with primers specific for each of these genes, as listed in **Table 2.4**. This approach revealed that there were no statistically significant differences in the expression of *adhA, phtE* or *pflA*. However, the expression of *strH* (SPD_0063) was significantly upregulated in GalR$_{AAA}$ comparted to D39 and Δ*galR*, as shown in **Figure 5.1**. As mentioned in Chapter 1, *strH* is an N-acetylglucosaminidase responsible for cleaving host glycoconjugates in the upper respiratory tract. Additionally, expression of *blpB* was significantly downregulated in

Table 5.1. Number of *S. pneumoniae*-specific reads from RNAseq

Sample ID	Number of Reads	% of Total Reads
D39 – 1	14,367	0.22%
D39 – 2	4,020	0.11%
D39 – 3	648	0.01%
Δ*galR* – 1	18,372	0.25%
Δ*galR* – 2	12,144	0.16%
Δ*galR* – 3	28,031	0.43%
GalR$_{AAA}$ – 1	30,906	0.45%
GalR$_{AAA}$ – 2	6,903	0.10%
GalR$_{AAA}$ – 3	12,567	0.18%
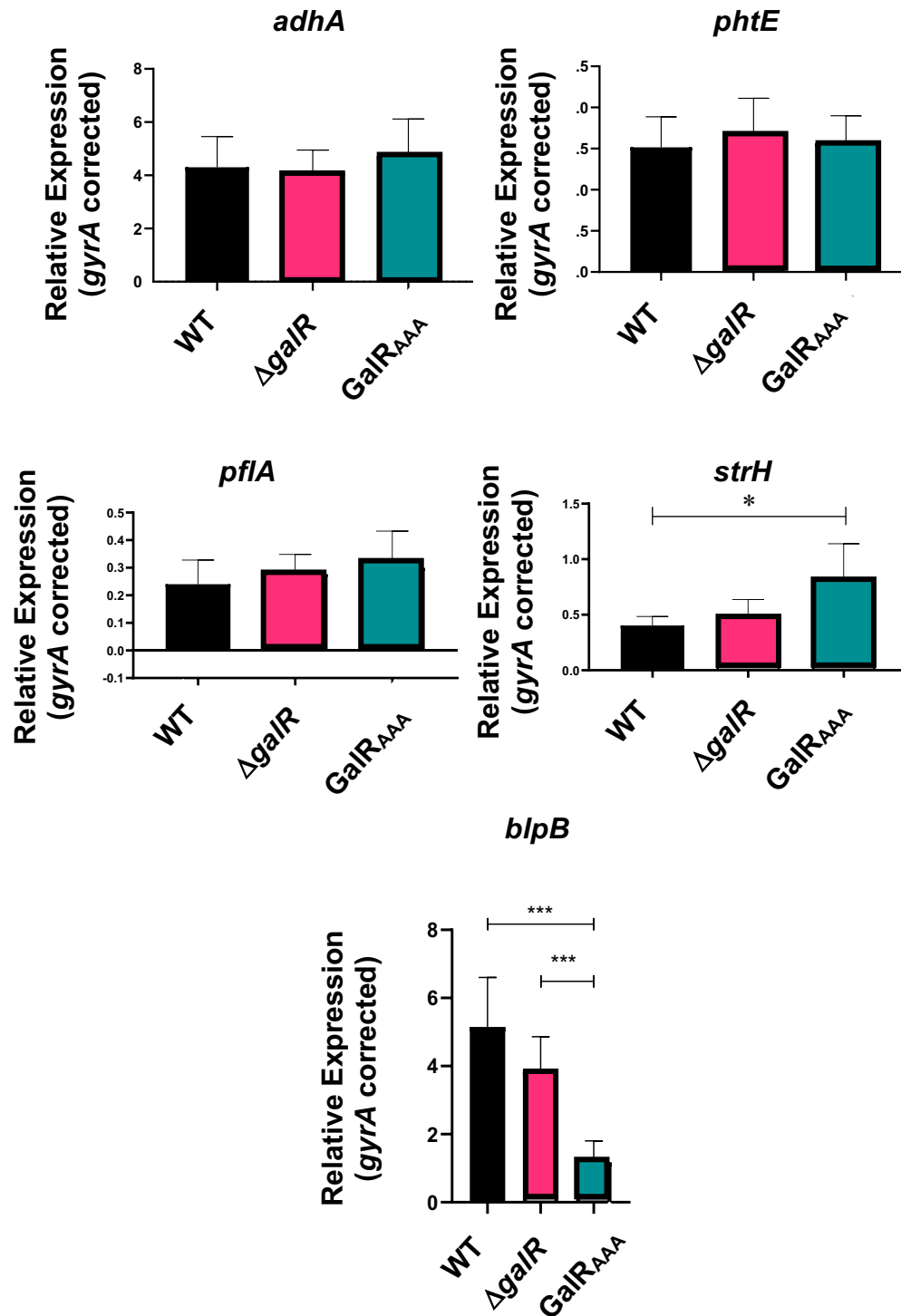
Figure 5.1. Validation of differentially expressed pneumococcal genes *in vivo*
The RNA from the lungs of female 4-6 week old Swiss mice infected with either D39, Δ*galR* or GalR$_{AAA}$ was firstly extracted and then pooled. qRT-PCR was then used to assess levels of *adhA, phtE, pflA*, *strH* and *blpB* mRNA with *gyrA* as an internal control. Data presented are the mean ± standard error of the mean expressed as relative expression. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; by a Student's *t*-test.

$\Delta galR$ compared to D39, and in GalR$_{AAA}$ compared to $\Delta galR$. BlpB is a bacteriocin-like protein that is described as a pseudogene (Kjos et al., 2016). Due to this, the role of this gene was not further explored in any detail, with focus shifting to differentially-expressed murine genes identified within the dataset.

5.2.3. Murine transcriptional changes occurring in response to pneumococcal infection

Dual *in vivo* RNAseq provides a unique advantage in that you can assess the transcriptomes of both the host and pathogen simultaneously during infection. As might be expected given the greater mass of host tissue relative to bacterial cells, the number of reads mapping to the murine genome was much higher than for the pneumococcal genome (**Table 5.2**). There was an average of 4,713,108 murine specific reads across the D39 samples. For the $\Delta galR$ samples, there was an average of 7,015,534 reads mapping to the murine genome and an average of 6,709,8773 murine reads across the GalR$_{AAA}$ samples. As with the pneumococcal reads, the murine reads were firstly subjected to adaptor removal and trimming prior to mapping back to the murine genome. We immediately moved through our analysis pipeline, as described in the Section 2.16 before conducting statistical analyses to determine which genes, if any, were differentially expressed in the lungs of these mice during infection.

Interestingly, there were no significantly differentially expressed genes identified when comparing $\Delta galR$ infected mice to GalR$_{AAA}$ infected mice, indicating these strains are quite similar in the context of host response to infection. However, there were genes identified when comparing both D39 vs. $\Delta galR$ infection and D39 vs. GalR$_{AAA}$ infection as shown in **Table 5.3** and **Table 5.4** that were differentially expressed. All genes listed have an adjusted *p*-value of less than 0.05 and a greater

Table 5.2. Number of *Mus musculus*-specific reads from RNAseq

Sample ID	Number of Reads	% of Total Reads
D39 – 1	6,370,678	99.78%
D39 – 2	3,590,923	99.89%
D39 – 3	4,177,723	99.99%
$\Delta galR$ – 1	7,221,437	99.75%
$\Delta galR$ – 2	7,257,886	99.84%
$\Delta galR$ – 3	6,567,280	99.57%
GalR$_{AAA}$ – 1	6,726,994	99.55%
GalR$_{AAA}$ – 2	6,701,552	99.90%
GalR$_{AAA}$ – 3	6,701,074	99.82%

than 1 $\log_2$-fold change. **Table 5.3** lists genes that were significantly upregulated in mice infected with D39 compared to both Δ*galR* and GalR$_{AAA}$. There were 15 genes identified in this approach with some of these having roles implicated in the immune response ranging from cell recruitment to inflammation. These include the *ccl12* gene encoding the chemokine CCL12, a ligand of CCR2 which recruits mononuclear phagocytes to the site of infection, *s100a8*, a gene with a role in inflammation and neutrophil activation, and the gene encoding the T cell activation marker CD69. These findings indicate that in the case of infection with D39, there may be an increase in both T cell infiltration to the infection site and also the recruitment of neutrophils along with other immune cells.

Table 5.4 list genes that were significantly downregulated in D39-infected mice compared to both Δ*galR* and GalR$_{AAA}$. While there were no obvious pathways in the immune response that appeared to be exclusively down-regulated in comparison to the mutants, there were still differences observed in genes linked to cell mobility (*Actg2*), cell signalling (*Ptp4a3*) and macrophage scavenger receptors (*Scara3*). This shows that there are defined differences within the immune response when comparing D39 to the GalR mutants, however further analyses are required to fully understand the impact that GalR and the putative phosphorylation sites play in the murine immune response. The remaining differentially expressed murine genes yielded from this approach are listed in **Appendix C**.

Table 5.3. Consistently up-regulated murine genes in D39 infection

Gene	Log FC	Adjusted *p*-value
Gnb1l	3.4443	0.0389
LOC115489778	2.5595	0.0446
Ccl20	2.0133	0.0352
Kcnip1	1.6763	0.0361
S100a8	1.5690	0.0289
Gm15564	1.5233	0.0383
Gm6377	1.4625	0.0389
Magt1	1.3692	0.0312
Ccl12	1.3059	0.0268
Gm15056	1.2816	0.0299
Ubd	1.2471	0.0204
Cd69	1.1455	0.0381
Vnn1	1.1186	0.0289
Gphn	1.0687	0.0358
Camk1d	1.01224	0.0358

Table 5.4. Consistently down-regulated murine genes in D39 infection

Gene	Log FC	Adj. *p*-value	Gene	Log FC	Adj. *p*-value
Nr1d1	-2.4536	0.0204	Hr	-1.4103	0.0475
Slc2a4	-2.0522	0.0268	Chst3	-1.3956	0.0381
Mhrt	-1.919	0.0389	Fbln1	-1.3610	0.0327
Grip2	-1.8703	0.0358	Speg	-1.3589	0.0204
Wdr83os	-1.8663	0.0478	Gli2	-1.3441	0.0204
Nptx1	-1.7879	0.0427	Adcy5	-1.3245	0.0287
Megf6	-1.6444	0.0252	Lgr6	-1.3239	0.0204
Olfml2a	-1.6277	0.0389	Asb2	-1.3218	0.0358
Atp1b2	-1.6238	0.0204	Acta2	-1.3120	0.0268
Fbln2	-1.5989	0.0499	Tafa5	-1.3011	0.0358
Stum	-1.5964	0.0289	Kcnk3	-1.2983	0.0204
Ltbp4	-1.5809	0.0204	Dnm1	-1.2933	0.0327
Tnxb	-1.5495	0.0204	Smad9	-1.2835	0.0372
Hmcn2	-1.5433	0.0426	Sod3	-1.2797	0.0249
D430041D05Rik	-1.5139	0.0381	Col27a1	-1.2683	0.0204
Col16a1	-1.4757	0.0358	Eln	-1.2654	0.0317
Cd248	-1.4649	0.0478	Zfp469	-1.2630	0.0204
Slc4a3	-1.4607	0.0389	Ntrk3	-1.2615	0.0227
Mn1	-1.4602	0.0496	Emc10	-1.2532	0.0496
Mrgprf	-1.4594	0.0358	Inmt	-1.2252	0.0395
Fgfr4	-1.4495	0.0344	Dmpk	-1.2247	0.0252
Stac	-1.4461	0.0204	Aoc3	-1.2175	0.0358
Synpo2l	-1.4214	0.0204	Slc8a2	-1.2080	0.0379

Table 5.4 (Contd.) Consistently down-regulated murine genes in D39 infection.

Gene	Log FC	Adj. *p*-value	Gene	Log FC	Adj. *p*-value
Ptgis	-1.1993	0.0204	Ptprs	-1.1196	0.0204
Ppp1r14a	-1.1992	0.0496	Gstm1	-1.1147	0.0204
Des	-1.1937	0.0268	Lgi4	-1.1077	0.0443
Ldb3	-1.1923	0.0327	Notch3	-1.1024	0.0229
Emid1	-1.1919	0.0227	Smtn	-1.0960	0.0268
2610027K06Rik	-1.1902	0.0360	Cspg4	-1.0958	0.0204
9330158H04Rik	-1.1753	0.0268	Col14a1	-1.0957	0.0438
Ntf3	-1.1747	0.0268	Heyl	-1.0884	0.0358
Sftpc	-1.1738	0.0268	Scara3	-1.0759	0.0465
Dgkb	-1.1730	0.0389	Kcnq4	-1.0624	0.0358
Clip3	-1.1683	0.0370	Actg2	-1.0563	0.0360
Wscd2	-1.1649	0.0327	Ptp4a3	-1.0558	0.0327
Jph2	-1.1630	0.0204	Meg3	-1.0550	0.0357
Gpc6	-1.1615	0.0358	Aqp1	-1.0539	0.0381
Acacb	-1.1610	0.0496	Cnn1	-1.0486	0.0360
Adcyap1r1	-1.1557	0.0426	Adamtsl5	-1.0465	0.0357
Hlf	-1.1551	0.0268	Cand2	-1.0438	0.0449
Ncam1	-1.1363	0.0443	Zfyve28	-1.0403	0.0227
Mtss2	-1.1351	0.0227	Slc45a4	-1.0390	0.0204
Podn	-1.1326	0.0360	Smoc1	-1.0368	0.0358
Per3	-1.1254	0.0268	Pgm5	-1.0191	0.0400
Igsf9b	-1.1226	0.0204			

5.2.4. Interrogation of the murine immune response to pneumococcal infection

When considering pneumococcal infection, it is important to factor in the range of different cell subsets present within a particular niche that the pneumococcus must interact with. In the context of the murine lung, there are endothelial and epithelial cells, parenchymal, and immune cells, both resident and recruited, just to name a few. The innate immune response with alter the trafficking and recruitment of specific immune cells subsets in response to infection (Luster et al., 2005). Therefore, to complement the transcriptional responses examined in Section 5.2.3, we wanted to investigate the differences in immune cell recruitment occurring during infection by flow cytometry.

Typically, qRT-PCR would be used to validate the results obtained from RNA sequencing, as was performed by Minhas *et al* 2020. In their study, there was a consistent correlation between the results obtained from both RNA sequencing and qRT-PCR. Given that we utilised the same methods as that of Minhas *et al*, we were considering whether our results needed to be validated using the same methods. Additionally, as a result of the COVID-19 pandemic, there were significant delays in obtaining standard DNA oligos for use in qRT-PCR. Nevertheless, potentially complementary information could be gained using Flow Cytometry to analyse immune cell subsets present during infection and determine whether the observed transcriptomic differences impacted cell trafficking within the infected lung. An antibody panel was designed to allow for visualisation of differences in cytokine production, along with the panel used by Minhas *et al* 2020, to assess differences in monocyte, macrophage and dendritic cell subsets. These antibodies, as well as the methods used, are detailed in the Section 2.17 and **Table 2.6**. The gating strategies used can be found in **Appendix D**. This approach would use murine lung tissue, infected in the same manner as was done to harvest RNA for sequencing,

from which single cells would be isolated, stained with required antibodies and fixed before acquiring data via Flow Cytometry.

Figure 5.2. shows the breakdown of broad cell subsets present in murine lungs in response to infection with either D39, $\Delta galR$ or GalR$_{AAA}$. While an uninfected lung control may have assisted interpretation, we were principally interested in differences between lungs infected with the two GalR mutants vs D39. Thus, in addition to attempting to minimise the number of mice used, an uninfected control was not included. There were no significant differences in the abundance of B cells, T cells or CD4+ T cells, between any of the infected tissues. The only significant difference detected was in the number of CD8+ T cells present in response to infection with GalR$_{AAA}$ when compared to $\Delta galR$.

Figures 5.3 and 5.4 show the proportion of CD4+ and CD8+ T cells, respectively, expressing the cytokines GM-CSF (Granulocyte-Macrophage Colony-Stimulating Factor), IL (Interleukin) -17A and Interferon (IFN) γ. To determine this, single cells were stimulated with Phorbol Myristate Acetate (PMA) and Ionomycin prior to staining with antibodies specific to the cytokines themselves as explained in Section 2.17. The process of stimulation can be quite harsh and often results in a high degree of cell death occurring within the sample, meaning that the results for each group of mice were concatenated to gain one singular representative figure indicative of 8 mice in total. As such, there are no statistical analyses that can be conducted for these data. The most striking finding across both of these figures is the increase in recruitment of GM-CSF producing CD4+ and CD8+ positive T cells to the site of GalR$_{AAA}$ infection relative to the other two strains. GM-CSF is generated by a range of cells in response to inflammation, particularly T cells, mast cells and macrophages (Shi et al., 2006). The increase in these cell subsets may provide an

explanation as to why bacterial loads in the lungs were lower for GalR$_{AAA}$ than for D39 and $\Delta galR$ at 24 h post pneumococcal challenge (**Fig. 3.14**).

Figures 5.4 and 5.6 assess the abundance of monocyte, macrophage and dendritic cell subsets as performed by Minhas *et al* 2020. Using this antibody panel, we would be able to assess the abundance of neutrophils, eosinophils, dendritic cells, NK cells, macrophages (alveolar and interstitial) and monocytes (resident and inflammatory). There were no significant differences observed in any of the reported immune cell subsets using this panel.

Overall, the most striking findings from this study was the increased expression of *strH* in the GalR$_{AAA}$ strain during infection, indicating some rewiring of pneumococcal carbohydrate metabolism. Given the issues encountered with numbers of pneumococcal reads obtained during RNA extraction, it would be worthwhile to reoptimize these protocols to either enrich for pneumococcal-specific reads to give us more confidence in these results. Excitingly, we also observed rewiring of the murine immune response during infection, with an increase in the abundance of CD4+ and CD8+ GM-CSF expressing T cells during to GalR$_{AAA}$ infection. Collectively, these results underpin the role of galactose metabolism in pneumococcal infection, however further work must be done to fully elucidate this relationship.
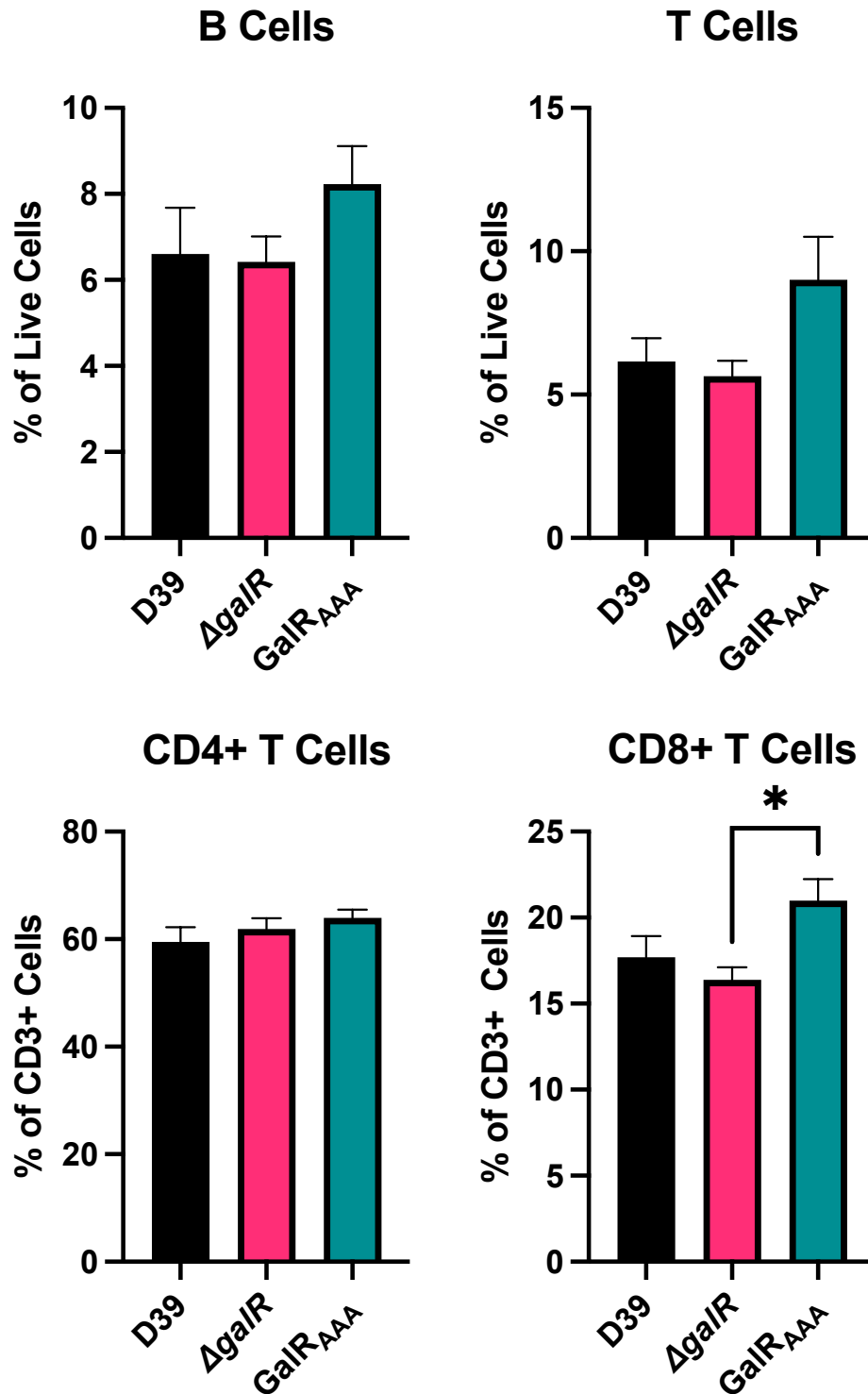
Figure 5.2. Quantification of immune cell subsets in murine lungs at 24 hours post-infection.

Groups of 8 mice were anaesthetised and challenged with 1×10^7 CFU of the respective strains (see Section 2.17) and humanely euthanised at 24 h post-infection. Lungs were removed and single cell suspensions were prepared, stained with appropriate antibodies (**Table 2.6**), fixed and examined by flow cytometry, as described in Section 2.17. Populations represented include B cells, T cells, CD4+ T cells and CD8+ T cells. All quantitative data are presented as mean ± S.D., analysed by One-Way ANOVA (*$p < 0.05$).
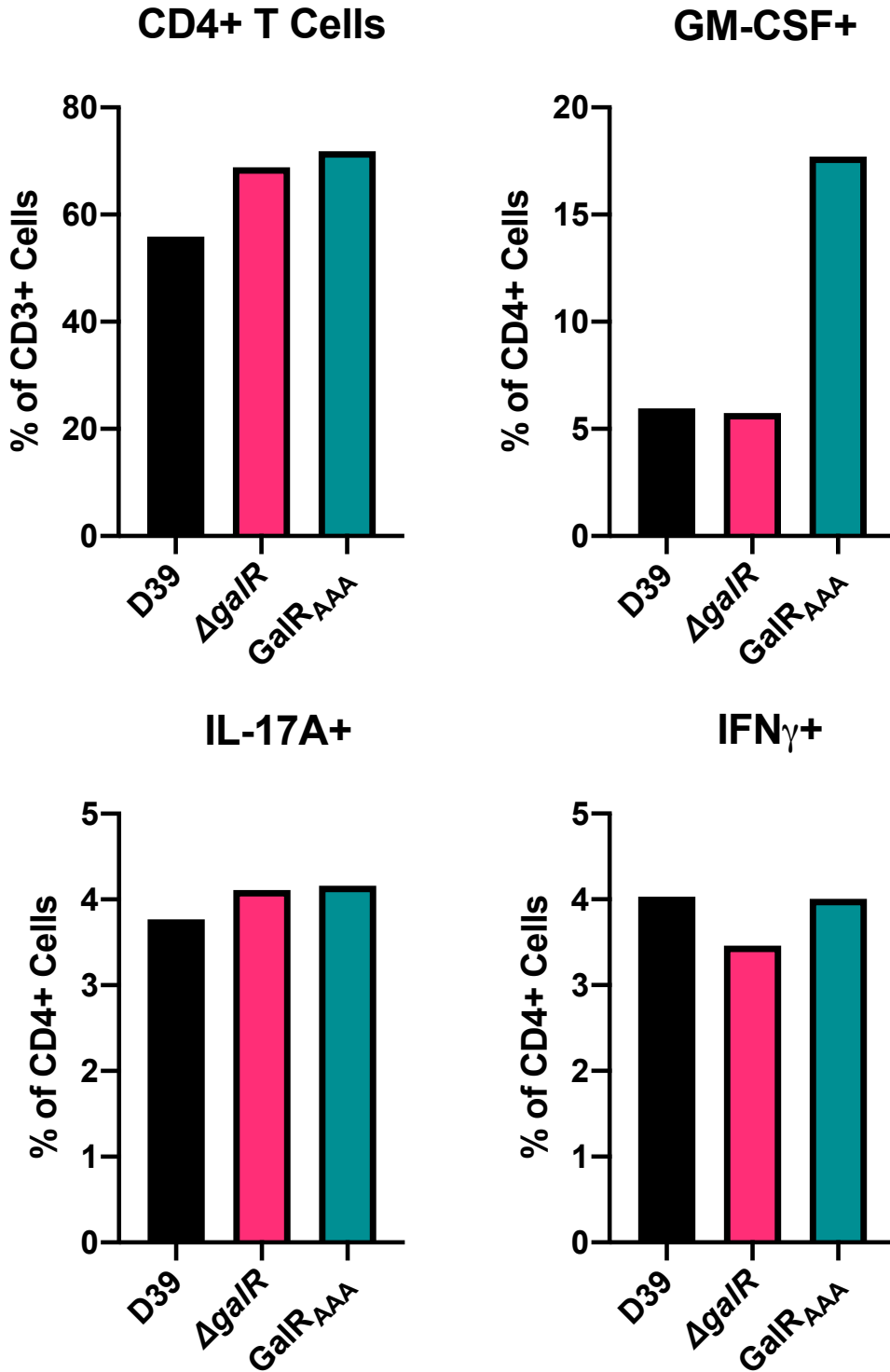
Figure 5.3. Cytokine expression by CD4+ T cells in infected lung tissue.
Groups of 8 mice were anaesthetised and challenged with 1 x 10^7 CFU of the
respective strains (see Section 2.17) and humanely euthanised at 24 h post-
infection. Lungs were removed and single cell suspensions were prepared,
stimulated with PMA and ionomycin, stained with appropriate antibodies (**Table
2.6**), fixed and examined by flow cytometry, as described in Section 2.17.
Populations represented include CD4+ T cells secreting GM-CSF, IL-17A and
IFNγ.
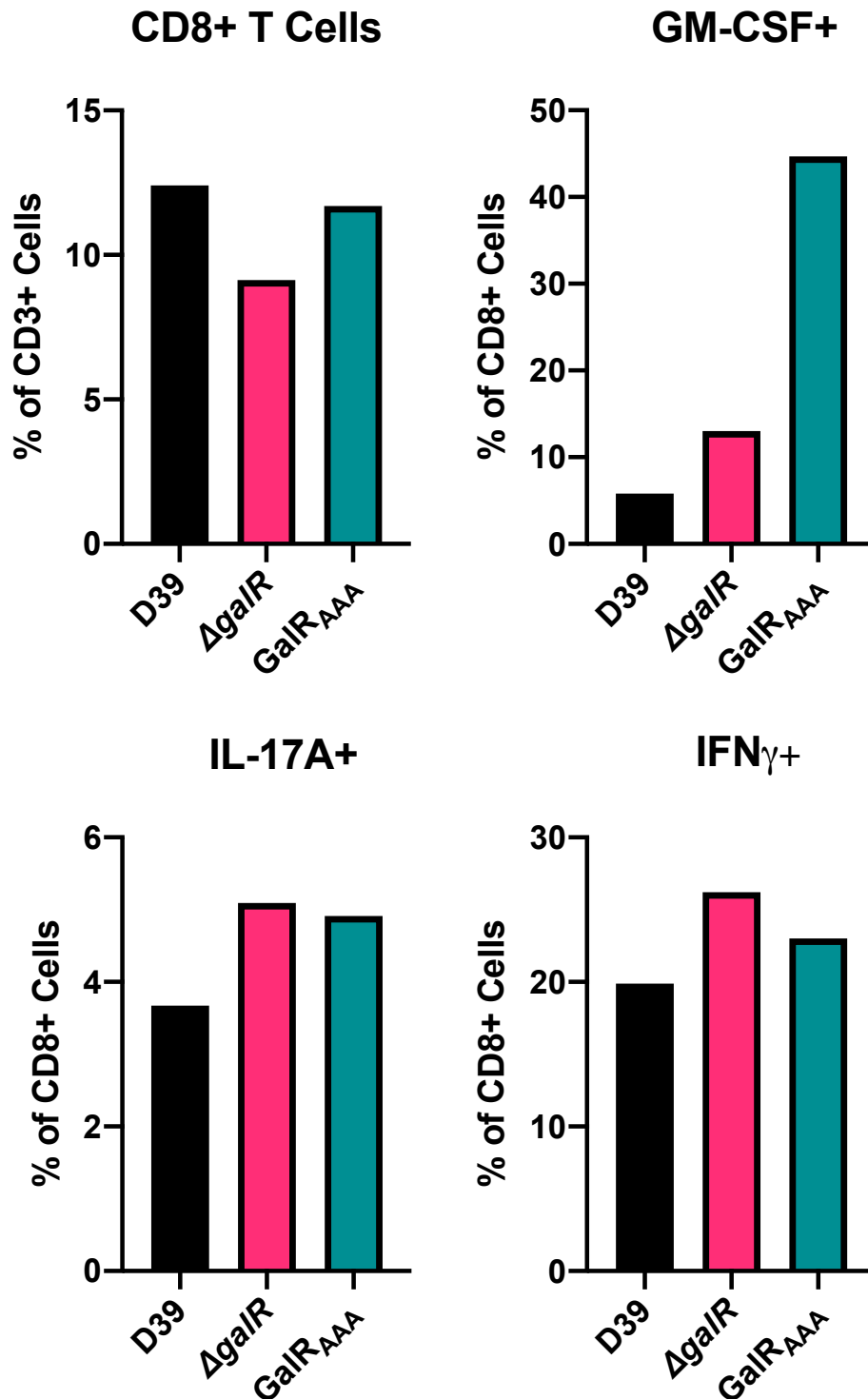
Figure 5.4. Cytokine expression by CD8+ T cells in infected lung tissue.
Groups of 8 mice were anaesthetised and challenged with 1 x 10^7 CFU of the respective strains (see Section 2.17) and humanely euthanised at 24 h post-infection. Lungs were removed and single cell suspensions were prepared, stimulated with PMA and ionomycin, stained with appropriate antibodies (Table 2.6), fixed and examined by flow cytometry, as described in Section 2.17. Populations represented include CD4+ T cells secreting GM-CSF, IL-17A and IFNγ.
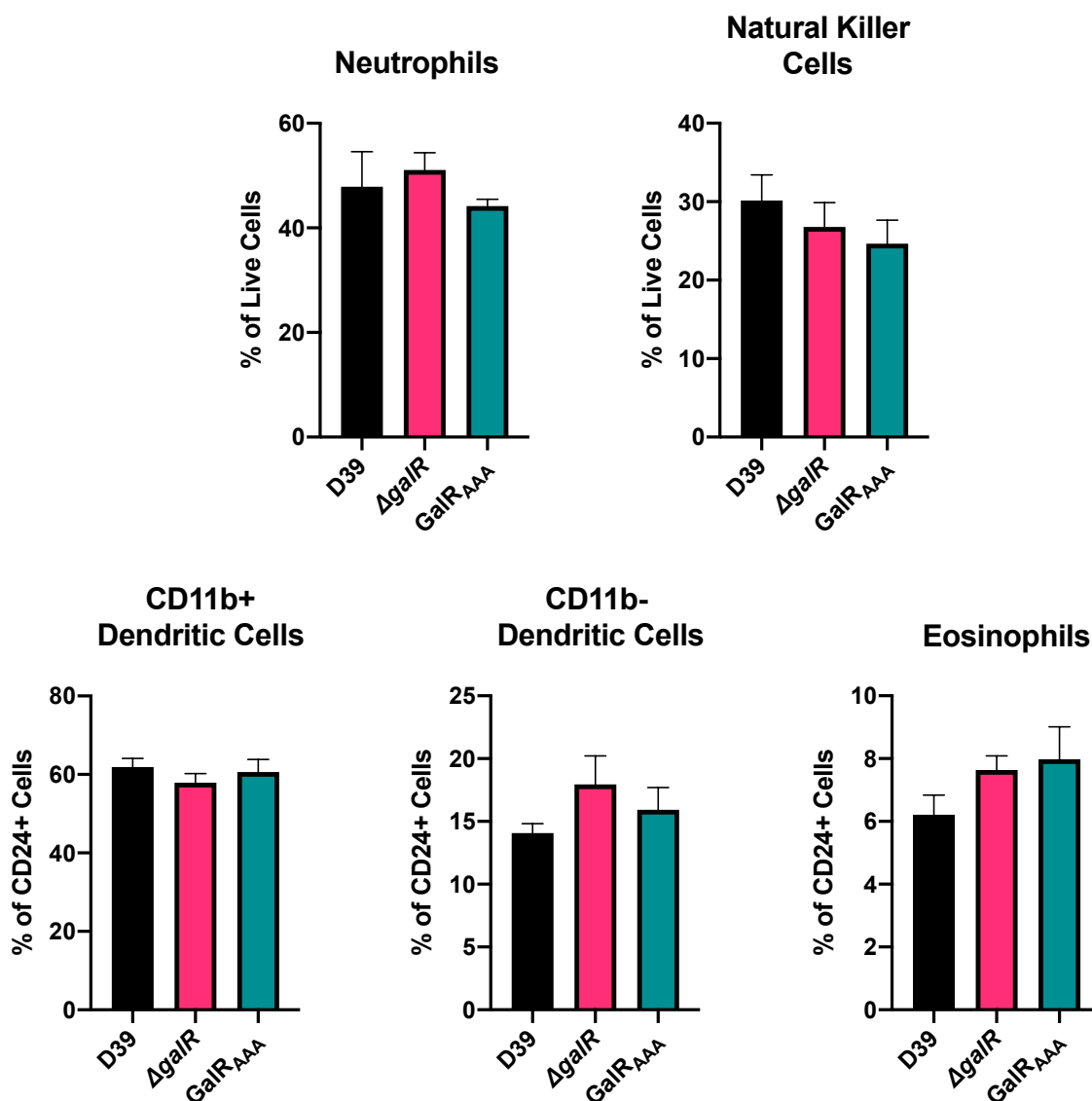
Figure 5.5. Quantification of immune cell subsets in murine lungs at 24 hours post-infection.

Groups of 8 mice were anaesthetised and challenged with 1×10^7 CFU of the respective strains (see Section 2.17) and humanely euthanised at 24 h post-infection. Lungs were removed and single cell suspensions were prepared, stained with appropriate antibodies (Table 2.6), fixed and examined by flow cytometry, as described in Section 2.17. Populations represented neutrophils, eosinophils, natural killer cells (NK cells), CD11b+ dendritic cells (CD11b+ DCs) and CD11b- dendritic cells (CD11b- DCs). All quantitative data are presented as mean ± S.D., analysed by One-Way ANOVA.
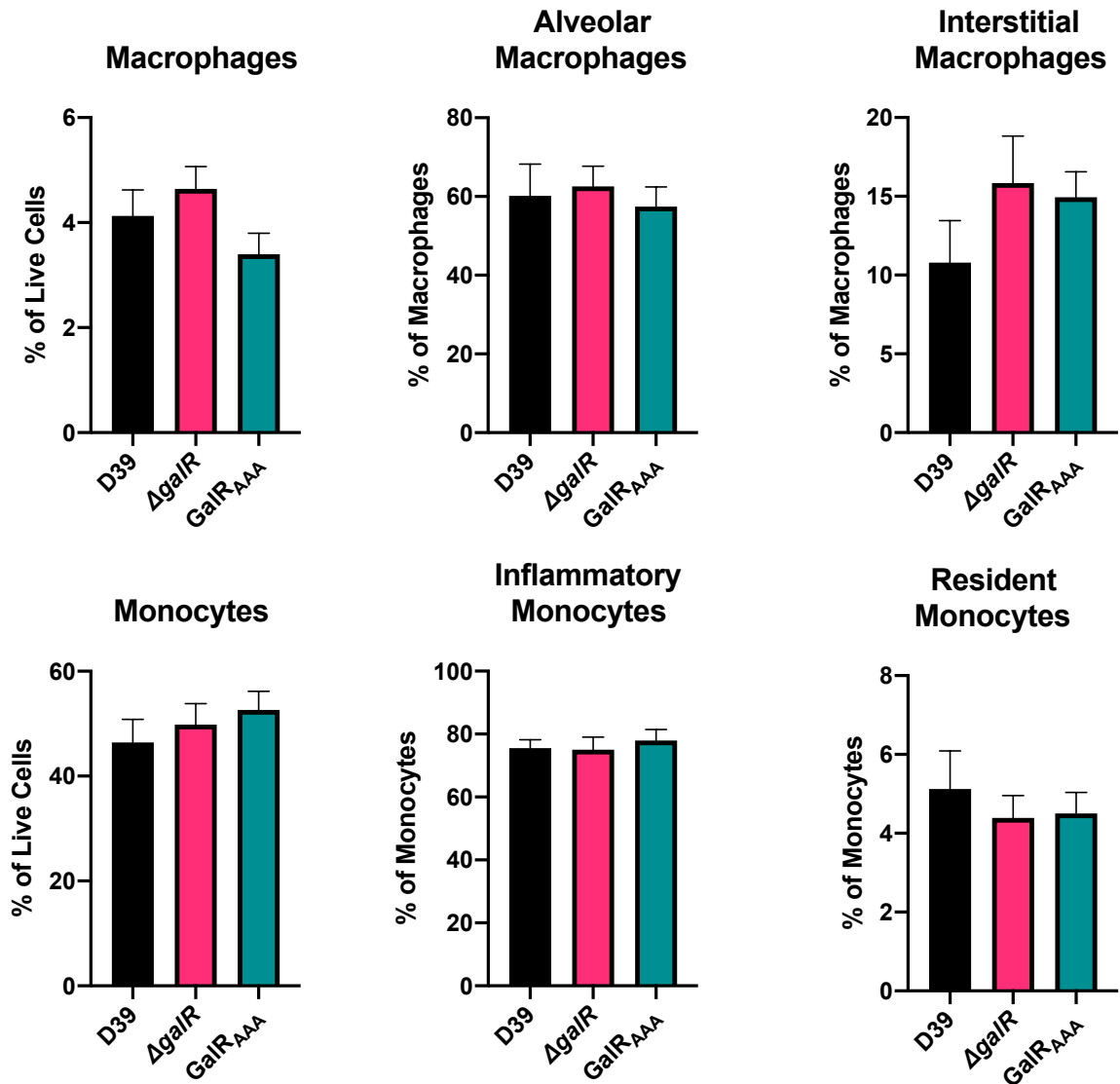
Figure 5.6. Quantification of immune cell subsets in murine lungs at 24 hours post-infection.

Groups of 8 mice were anaesthetised and challenged with 1×10^7 CFU of the respective strains (see Section 2.17) and humanely euthanised at 24 h post-infection. Lungs were removed and single cell suspensions were prepared, stained with appropriate antibodies (Table 2.6), fixed and examined by flow cytometry, as described in Section 2.17. Populations represented include macrophages (MΦ), alveolar macrophages (AMΦ), interstitial macrophages (IMΦ), monocytes, inflammatory monocytes (iMono) and resident monocytes (rMono). All quantitative data are presented as mean ± S.D., analysed by One-Way ANOVA.

5.3. Discussion

In this chapter, a combination of microbiological, immunological and bioinformatic techniques were employed to shed light on the role of GalR and its putative phosphorylation sites in pneumococcal infection.

The process of dual RNAseq is a relatively new technique, with our group being amongst the first to show that alterations in pneumococcal sugar metabolism can rewire the host immune response to infection. The study by Minhas *et al* in 2020 built on their previous paper which showed that a single nucleotide polymorphism within the regulator of the raffinose operon *rafR* resulted in specific niche tropism during infection (Minhas et al., 2020). Using dual *in vivo* RNAseq, the RNA from lung tissue of mice infected with *S. pneumoniae* was able to be isolated and sequenced. Using this technique, they demonstrated that the SNP resulted in differential expression of multiple sugar transporters which effectively fine-tuned carbohydrate metabolism in those strains. Additionally, they showed that this SNP also resulted in extensive rewiring of the host transcriptome, including differential expression of genes encoding cytokine and chemokine ligands and receptors. Given that this process allows visualisation of the transcriptome from both the host and pathogen perspective, it was an ideal approach to employ in this project.

Using the same methods and sequencing approach as Minhas *et al*, we planned to perform dual *in vivo* RNA sequencing that would allow examination of the impact of *galR* mutations on the host and pathogen transcriptomic response during infection. Unfortunately, the number of reads that mapped back to the pneumococcal genome was suboptimal for all samples, suggesting poor yield of bacterial RNA, potentially compromising the sensitivity of the bacterial transcriptomic analyses. AGRF subsequently analysed the quality of the RNA samples to determine the RNA Integrity Number (RIN) (a value between 1 and 10 that indicates the quality of the

sampled RNA; the lower the RIN, the more degraded the RNA) (Schroeder et al., 2006) (See **Appendix B**). The RIN values for the sequenced samples ranged between 5.9 and 7.1, whereas values between 7 and 8 would typically be considered necessary for reasonable quality analyses (Puchta et al., 2020). Poorer than expected RNA integrity may have contributed to the low read depth for bacterial transcripts, which clearly impacts sensitivity. There is also a possibility of distortion of the findings if a given class of RNA was more sensitive to degradation than another.

Despite the poor RNA quality and low read count, there were still 5 genes that were identified as being significantly differentially expressed within the pneumococcal genome. Two of these genes had differential expression confirmed via qRT-PCR, namely *strH* and *blpB*. As mentioned in Section 1.4.2, *strH* is a surface glycosidase that aids in pneumococcal invasion by revealing glycan targets on host epithelial cells which allow adhesins to bind. In both the RNA seq data and qRT-PCR validation, expression of *strH* was significantly upregulated in GalR$_{AAA}$ compared to $\Delta galR$. Given that galactose metabolism is so important for survival in the upper respiratory tract, it would make some sense that expression of *strH* would be upregulated in the phosphorylation site mutant. We know from our functional studies in Chapter 3 that the phosphorylation site mutations render the GalR$_{AAA}$ strain unable to grow in or metabolise galactose. It is possible that as a result of the amino acid substitutions in this strain, the cell is attempting to upregulate enzymes that will liberate alternative carbon sources, such as GlcNac, which will enable metabolism. While interesting, this hypothesis requires further investigation that was beyond the scope of this thesis. In the case of *blpB*, this is a bacteriocin-like protein that was significantly downregulated in the GalR$_{AAA}$ strain, as shown in **Figure 5.1**. BlpB is a component of the ABC transporter responsible for exporting the bacteriocin BlpC in

the pneumococcus (Dawid et al., 2007). It has been previously shown that BlpB is responsible for inhibiting sensitive pneumococcal strains during colonisation and also provides a fitness advantage in both *in vitro* biofilms and competitive colonisation models. Additionally, the production of *blpC* and its exporter *blpAB* are under the control of *comE,* a gene also required for competence and natural transformation (Kjos et al., 2016, Wholey et al., 2016). This provides a link between competence, Gal metabolism and invasion capacity in the lungs, similar to that reported by Trappetti et al., (2017), who demonstrated a link between carbohydrate metabolism and cell-to-cell communication. Despite this, *blpB* is classified as a pseudogene in the serotype 2 strain D39 according to the Kyoto Encyclopaedia of Genes and Genomes (KEGG), meaning that while it resembles a gene, it has likely lost the ability to encode a functional protein over time as a result of accumulated mutations. Pseudogenes will typically be subject to degradation and eventual removal as a result of these continued mutations (Kuo and Ochman, 2010). Consequently, no additional investigations were performed into this gene.

While it was disappointing that there was not more information that could be gleaned from the pneumococcal-specific RNA sequencing analysis, the murine transcriptomic analysis yielded more promising results. There was a much greater number of reads that successfully mapped back to the murine genome, giving greater confidence in the data that were generated. Genes were identified that were differentially expressed and associated with the murine immune response and immune cell recruitment. While there were no differentially expressed genes identified when comparing $\Delta galR$ and GalR$_{AAA}$ infected mice, there were differences between both WT and $\Delta galR$ and WT and GalR$_{AAA}$ infected mice. In particular, there was a subset of genes that were upregulated in both the WT vs. $\Delta galR$ and WT vs. GalR$_{AAA}$ comparisons of infection. These included *ccl12*, a gene encoding a CCR2

receptor ligand involved in monocyte recruitment and *ccl20*, a gene involved in Th17 cell recruitment. There was also differential expression of *cxcl9* and *cxcl13,* genes involved in Th1 and T follicular helper recruitment and B cell recruitment, respectively. Additionally, there were genes that are associated with inflammation that were upregulated including *s100a8, s100a9* and *ms4a4a*. The RNA sequencing results were indicative of potentially increased T cell infiltration in D39 infection compared to both Δ*galR* and GalR$_{AAA}$, along with a potential increase in the number of neutrophils, granulocytes and eosinophils present during infection. The immune cell subsets that we proposed may be impacted were similar to those assessed by Minhas *et al.,* 2020. As such, we decided that our next steps would be to perform flow cytometry analysis to determine if there were indeed any differences in the abundance of different immune cell subsets in response to infection, replicating the antibody panel used in their study. Additionally, we wanted to further probe whether there were differences in the cytokines being expressed by immune cells during infection, which led to the development of a secondary antibody panel which allowed us to visualise expression of IL-17A, IFN-γ and GM-CSF.

Through use of FACS, we were able to determine that there was a significant difference in the number of CD8+ T cells present during infection in response to the GalR$_{AAA}$ strain when compared to Δ*galR* infection. Further, there was an increase in the number of CD4+ and CD8+ T cells expressing GM-CSF in response to GalR$_{AAA}$ infection compared to infection with the other two strains. GM-CSF is a pro-inflammatory cytokine that is associated with a range of inflammatory autoimmune diseases in humans such as rheumatoid arthritis and inflammatory lung disorders (Hamilton et al., 1980, Shi et al., 2006). It has also been identified as a mediator in lung inflammatory models, controlling the number of responding neutrophils and macrophages, as well as limiting expression of Toll-like Receptor 4, a plasma-bound

receptor that is able to recognise pneumolysin during infection (Malley et al., 2003, Shi et al., 2006, Koppe et al., 2012). Interestingly, GM-CSF has also been shown to enhance monocyte recruitment and activation to sites of infection. Resting T cells typically do not produce GM-CSF, but they can do so transiently upon activation (Shi et al., 2006). It has been shown on multiple occasions that T cells can produce GM-CSF in response to activation with anti-CD3, although its direct effect on T cells remains poorly defined (Quill et al., 1989, Levitt et al., 1991, Himes et al., 2000). The receptor for GM-CSF comprises both an α-chain and a β-chain. The β-chain of the receptor is not expressed on most resting cells, meaning that there must be elevated levels of GM-CSF present to trigger production of the α-chain on these cells (Shi et al., 2006). Experimental models in mice deficient in GM-CSF have revealed that while GM-CSF is essentially dispensable in terms of maintaining normal levels of hematopoietic cells and their precursors, it is essential for both resistance to local infection and normal pulmonary physiology (Wada et al., 1997). T cells isolated from GM-CSF deficient mice have diminished Th1 and Th2 responses, indicating that this cytokine is critical for regulation of T cell mediated immune responses (Wada et al., 1997, Shi et al., 2006). Based on these findings we can propose that the response to the GalR$_{AAA}$ mutant results in skewing of the immune response away from a traditional Th1 mediated response toward GM-CSF from both CD4+ and CD8+ cells. Interestingly, this does correlate with the findings of our RNA sequencing data, showing that there may be increases in both inflammation and Th1 cell recruitment.

Unfortunately, there were no significant differences observed in the additional immune cell subsets shown across **Figures 5.2** to **5.6.** However, the differences observed in production of GM-CSF in response to infection warrant further investigation. As mentioned in the Section 5.2.4, the way in which the data were

presented for cytokine production in CD4+ and CD8+ T cells required the concatenation of the results of all 8 mice per group into a single data point. In future, a repeat of this experiment with either greater mouse numbers or a means of enrichment for these cytokines would be beneficial to validate these results.

Nevertheless, the finding that GalR$_{AAA}$ infection results in remodelling of the immune response relative to the wild type shows that galactose metabolism does indeed impact the immune response to pneumococcal infection and warrants further investigation to fully characterise this relationship.

Chapter 6: Final Discussion

This thesis has focussed on attempting to decipher the intricacies of pneumococcal Gal metabolism. Despite carbohydrate metabolism being critical to survival, it remains poorly characterised in *S. pneumoniae*.

Chapter 3 focussed on gaining greater insight into regulation of Gal metabolism in the pneumococcus. We observed an essential role for GalR in the use of Gal as a carbon source, as well as altered expression of the Leloir pathway due to deletion of the gene. There was also a marked loss in the ability of the GalR mutant to adhere to the nasopharyngeal epithelium. While the role of GalR as the regulator of the Leloir pathway has been known for some time (Afzal *et al.,* 2015), its importance in metabolism, regulation and adherence had not been shown. While essential for these processes, we observed that there was minimal impact of GalR deletion on virulence *in vivo*. Given that a pneumococcal pneumonia model was used here, this is not entirely unsurprising. We expect that use of a nasopharyngeal colonisation model may yield more significant results, but this was beyond the scope of the project. Having uncovered more of the functions of GalR within the pneumococcus, we wanted to further investigate the role of the putative phosphorylation sites identified by Sun *et al.,* While these sites have not been further characterised in the literature, much less confirmed as phosphorylation sites, we wanted to determine whether they contributed in any way to Gal metabolism in the first instance. Selective mutation of each of these sites to non-phosphorylatable Ala residues showed that the presence of even a single wild-type putative phosphorylation site was sufficient to permit growth in Gal. Combined mutation of all three sites to Ala, however, showed significant growth perturbation, indicating that these sites do play an essential role in the ability to metabolise Gal. We know from our structural modelling studies that the putative phosphorylation sites are not located near the DNA or Gal binding domains of the GalR dimer, so while these substitutions may

result in structural changes within the protein, they should not directly impact the ability to bind galactose or the promoter region of the *galKT* operon. By assessing transcriptional changes occurring in response to these amino acid substitutions, we discovered that amino acid substitution results in decreased expression of *galK*, but that only a minimal amount of *galK* expression is required to successfully metabolise Gal. Interestingly, we again observed that there was no significant upregulation of the T6P pathway gene *lacD* in strains that were unable to metabolise Gal, namely GalR$_{AAA}$. This was amongst the first evidence suggesting that there may be a link between these two pathways which was further investigated in Chapter 4. We also revealed a potential role for the putative phosphorylation sites in adherence to the nasopharyngeal epithelium, with the GalR$_{AAA}$ strain showing a decreased ability to adhere to Detroit cells in the presence of Gal. Finally, we revealed a role for these sites in pneumococcal virulence, with bacterial loads of the GalR$_{AAA}$ strain being significantly lower in the lungs, nose and ears at 24 hours post-challenge. What remained unclear at this stage, was whether the phosphomimetic amino acid substitutions were functioning sufficiently to show true growth phenotypes as a result of particular sites being 'phosphorylated'.

The findings of Trappetti *et al* suggested that in the presence of Gal, phosphorylated AI-2 may be able to phosphorylate GalR, either directly or indirectly, resulting in upregulation of the Leloir pathway and a subsequent hypervirulent phenotype as a result of increased UDP-Glc generation. While the exact mechanism of GalR phosphorylation is yet to be determined, the wanted to investigate impact of exogenous AI-2 on the putative phosphorylation sites as a function of growth in Gal. We know from the findings of Trappetti *et al.,* that AI-2 must be present in a 'goldilocks' concentration intracellularly to elicit positive effects on growth in Gal. That is, too much exogenous AI-2 results in growth inhibition in Gal media (Trappetti

et bal., 2017). Thus, it is possible that the phosphomimetic amino acid substitutions might actually mimic a level of phosphorylation that interferes with GalR function, thereby accounting for the failure of most of the mutants to grow in CDM + Gal. We also know that the presence of *luxS* means a pneumococcal cell is capable of constitutively generating its own AI-2 as a by-product of the Activated Methyl Cycle, potentially adding to the extent of phosphorylation at the wild type sites. Deciphering the exact mechanism of GalR phosphorylation is complex, and was unfortunately beyond the scope of this PhD. This, however, does not mean that is should not be investigated further. The inability to source biologically active AI-2 significantly derailed attempts to fully elucidate the mechanisms behind phosphorylation of GalR, however gaining access to this in the future, coupled with techniques such as phosphoproteomics, could provide strong insight. Nonetheless, the findings gained from the alanine substitution mutants alone was sufficient to identify the importance of the putative phosphorylation sites in Gal metabolism, regulation and virulence.

What became evident during investigation of the putative phosphorylation sites was that there is a degree of interplay occurring between the Leloir and the T6P pathways. Given that these two pathways have been reported as being independent of each other in the literature, the seeming interplay between the two was interesting. As there is a fitness cost involved in maintaining complete pathways within the genome, there must be a reason for both pathways being present. Either they are discreet pathways with different roles, or they are linked, like what has previously been shown for *S. mutans* (Zeng *et al.,* 2010). From the growth analyses conducted with both Leloir and T6P pathway mutants in Chapter 4 it is evident that these pathways are incapable of rescuing each other in times of metabolic distress. Similarly, transcriptional studies show that deletion of either *galK* or *lacD* results in similar upregulation of GalR. Previous studies by Afzal *et al.,* have shown that LacD

is under the control of the repressor LacR (Afzal *et al.,* 2014). However, this result indicates that there may be some cross talk between these two pathways, with GalR potentially acting as a master regulator. Interestingly, additional transcriptional studies looking at expression of *lacR* showed decreased expression in the Δ*lacD* mutant, as expected, indicating de-repression of the *lacABCD* operon to promote growth in Gal. Conversely, *lacR* expression was upregulated in the Δ*galR* mutant, indicating once again that GalR may be required for correct regulation of these two pathways.

Given that there is a defined link between these two pathways in *S. mutans*, our next step was to determine if there may be toxic accumulation of metabolic intermediates such as Gal6P that may be resulting in an inability to grow when Gal is the sole carbon source. Utilising both phenotypic microarray analyses and metabolomics, we did observe some differential metabolism occurring between the mutants, albeit none that could definitively be linked to toxicity when accumulating intracellularly. This is an area of the project that warrants further investigation in order to fully characterise the metabolome of the each of the Leloir and T6P pathway mutants, so definitive differences in metabolism can be identified. While this particular question remains unanswered, Chapter 4 has definitively shown that there is indeed a link between the Leloir and T6P pathways in the pneumococcus, a stark contrast to what has previously reported in the literature (Afzal et al., 2014, Afzal et al., 2015, Paixão et al 2015).

Having now gained a better understanding of the interplay between the Leloir and T6P pathways and the regulation of Gal metabolism via GalR and its putative phosphorylation sites, the next question to answer was the role of GalR in virulence. While its impact in a pneumonia model of infection was answered in Chapter 3, what

remained unclear is what transcriptome changes occur in both the pathogen and the host during infection. From the findings of Minhas *et al.,* this question could be answered using dual *in vivo* RNA sequencing. There were issues with the number of reads recovered that were specific to the pneumococcal genome, as detailed in Chapter 5. Unfortunately, these low read numbers meant there were no substantial conclusions that could be drawn from the bacterial transcriptomic dataset, aside from the consistent upregulation of the surface exoglycosidase *strH* in the $GalR_{AAA}$ strain. As discussed in Section 5.3, this would not be unexpected as the strain is unable to metabolise Gal and is therefore attempting to upregulate enzymes that would result in the liberation of other carbon sources, such as the GlcNac that is released by StrH. Fortunately, there was a much greater recovery of murine-specific reads that could be interpreted with increased confidence. The results of the murine-specific RNA sequencing indicated that there were potential changes occurring in the murine immune response during infection with each of the three strains that were assessed: D39, $\Delta galR$ and $GalR_{AAA}$. Flow cytometry analysis confirmed that there were indeed changes occurring in the immune response during infection, particularly with $GalR_{AAA}$. There was a significant increase in the number of CD8+ T cells recruited to the infected lung tissue. CD8+ T cells are a critical component of the adaptive immune response and play a key role in the clearance of intracellular pathogens. The increase in this cell subset may explain why we observed decreased bacterial loads of $GalR_{AAA}$ in the lungs at 24 hours post infection, as shown in Chapter 3. Additionally, we observed an increase in the amount of GM-CSF being expressed from both CD4+ and CD8+ T cells in response to $GalR_{AAA}$ infection. Collectively, this indicates a skewing of the immune response away for the standard Th1 response toward a GM-CSF response. This underpins a critical role

for the putative GalR phosphorylation sites in the development of an adequate immune response, a finding that has not yet been reported in the literature.

The findings presented in this thesis have detailed significant advancement in our understanding of Gal metabolism in the pneumococcus from a range of perspectives. Starting at the most basic level, we have shown that GalR is critical for the successful utilisation of Gal as a carbon source *in vitro*. Further, we have revealed that the putative phosphorylation sites S317, T319 and T323 also play critical roles in the ability to utilise Gal, while also resulting in altered transcription of key Leloir and T6P genes. For the first time, we have shown that the Leloir and T6P pathways are linked in the pneumococcus, as shown by the inability for one pathway to rescue the other in the event of metabolic distress. Much like *S. mutans*, these pathways are likely linked in terms of their intracellular metabolites and may be under the control of a universal regulator in the form of GalR. However, further studies are required to confirm this. Finally, we have identified a role for the putative GalR phosphorylation sites in pneumococcal virulence, showing that substitution of these residues to non-phosphorylatable Ala results in rewiring of the murine immune response towards a GM-CSF-specific response. We now have a greater understanding of exactly how Gal is utilised and the internal checks that must be achieved in order to permit this. The true importance of GalR is now being realised, with further studies likely to show that the role of GalR extends beyond the ability to metabolise Gal. We have also gained a greater appreciation for the role of carbohydrate metabolism in other, wider reaching intracellular processes within the pneumococcus such as virulence.

References

References

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.

Afshar, D., Rafiee, F., Kheirandish, M., Ohadian Moghadam, S. & Azarsa, M. (2020). Autolysin (LytA) recombinant protein: a potential target for developing vaccines against pneumococcal infections. *Clin Exp Vaccine Res* 9**,** 76-80.

Afzal, M., Shafeeq, S. & Kuipers, O. P. (2014). LacR is a repressor of *lacABCD* and LacT is an activator of *lacTFEG*, constituting the *lac* gene cluster in *Streptococcus pneumoniae*. *Appl Environ Microbiol* 80**,** 5349-5358.

Afzal, M., Shafeeq, S. & Kuipers, O. P. (2018). NADH-mediated gene expression in *Streptococcus pneumoniae* and role of Rex as a transcriptional repressor of the Rex-regulon. *Front Microbiol* 9**,** 1300.

Afzal, M., Shafeeq, S., Manzoor, I. & Kuipers, O. P. (2015). GalR acts as a transcriptional activator of *galKT* in the presence of galactose in *Streptococcus pneumoniae*. *J Mol Microbiol Biotechnol* 25**,** 363-371.

Ajdic, D. & Ferretti, J. J. (1998). Transcriptional regulation of the *Streptococcus mutans gal* operon by the GalR repressor. *J Bacteriol* 180**,** 5727-5732.

Albiger, B., Dahlberg, S., Sandgren, A., Wartha, F., Beiter, K., Katsuragi, H., Akira, S., Normark, S. & Henriques-Normark, B. (2007). Toll-like receptor 9 acts at an early stage in host defence against pneumococcal infection. *Cell Microbiol* 9**,** 633-644.

Albiger, B., Sandgren, A., Katsuragi, H., Meyer-Hoffert, U., Beiter, K., Wartha, F., Hornef, M., Normark, S. & Normark, B. H. (2005). Myeloid differentiation factor 88-dependent signalling controls bacterial growth during colonization and systemic pneumococcal disease in mice. *Cell Microbiol* 7**,** 1603-1615.

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.

Alexander, J. E., Lock, R. A., Peeters, C. C., Poolman, J. T., Andrew, P. W., Mitchell, T. J., Hansman, D. & Paton, J. C. (1994). Immunization of mice with pneumolysin

toxoid confers a significant degree of protection against at least nine serotypes of *Streptococcus pneumoniae. Infect Immun* 62**,** 5683-5688.

Alloway, J. L. (1932). The transformation *in vitro* of R pneumococci into S forms of different specific types by the use of filtered pneumococcus extracts. *Journal of Experimental Medicine* 55(1)**,** 91 - 99.

Alloway, J. L. (1933). Further observations on the use of pneumococcus extracts in effecting transformation of type *in vitro. Journal of Experimental Medicine***,** 265 - 278.

Arkwright, J. A. (1921). Variation in bacteria in relation to agglutination both by salts and by specific serum. *Journal of Pathology and Bacteriology***,** 36 - 60.

Attali, C., Durmort, C., Vernet, T. & Di Guilmi, A. M. (2008). The interaction of *Streptococcus pneumoniae* with plasmin mediates transmigration across endothelial and epithelial monolayers by intercellular junction cleavage. *Infect Immun* 76**,** 5350-5356.

Australian Institute of Health and Welfare. 2018. *Pneumococcal Disease in Australia* [Online]. Australian Institute of Health and Welfare. Available: https://www.aihw.gov.au/reports/immunisation/vaccine-preventable-diseases/contents [Accessed].

Austrian, R. (1981). Pneumococcus: the first one hundred years. *Rev Infect Dis* 3**,** 183-189.

Austrian, R. (1984). A reassessment of pneumococcal vaccine. *New England Journal of Medicine* 310**,** 651 - 653.

Austrian, R., Douglas, R. M. & Schiffman, G. (1976). Prevention of pneumococcal pneumoniae by vaccination. *Transactions of the Association of Americam Physicians* 89**,** 184 - 189.

Austrian, R. & Gold, J. (1964). Pneumococcal bacteremia with especial reference to bacteremic pneumococcal pneumonia. *Annals of Internal Medicine* 60**,** 759 - 776.

Avadhanula, V., Rodriguez, C. A., Devincenzo, J. P., Wang, Y., Webby, R. J., Ulett, G. C. & Adderson, E. E. (2006). Respiratory viruses augment the adhesion of

References

bacterial pathogens to respiratory epithelium in a viral species- and cell type-dependent manner. *J Virol* 80, 1629-1636.

Avery, O. T. & Heidelberger, M. (1925). Immunological relationships of cell constituents of pneumococcus : II. *J Exp Med* 42, 367-376.

Avery, O. T., MacLeod, C. M. & McCarty, M. (1944). Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a desoxyribonucleic acid fraction isolated from pneumococcus type III. *Journal of Experimental Medicine* 79, 137 - 157.

Babb, R., Chen, A., Hirst, T. R., Kara, E. E., Mccoll, S. R., Ogunniyi, A. D., Paton, J. C. & Alsharifi, M. (2016). Intranasal vaccination with gamma-irradiated *Streptococcus pneumoniae* whole-cell vaccine provides serotype-independent protection mediated by B-cells and innate IL-17 responses. *Clin Sci (Lond)* 130, 697-710.

Babb, R., Chen, A., Ogunniyi, A. D., Hirst, T. R., Kara, E. E., Mccoll, S. R., Alsharifi, M. & Paton, J. C. (2017). Enhanced protective responses to a serotype-independent pneumococcal vaccine when combined with an inactivated influenza vaccine. *Clin Sci (Lond)* 131, 169-180.

Bergmann, S. & Hammerschmidt, S. (2006). Versatility of pneumococcal surface proteins. *Microbiology (Reading)* 152, 295-303.

Berry, A. M., Lock, R. A., Hansman, D. & Paton, J. C. (1989a). Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infect Immun* 57, 2324-2330.

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.

Berry, A. M. & Paton, J. C. (1996). Sequence heterogeneity of PsaA, a 37-kilodalton putative adhesin essential for virulence of *Streptococcus pneumoniae*. *Infect Immun* 64, 5255-5262.

Berry, A. M. & Paton, J. C. (2000). Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infect Immun* 68, 133-140.

Berry, A. M., Yother, J., Briles, D. E., Hansman, D. & Paton, J. C. (1989b). Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. *Infect Immun* 57, 2037-2042.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L. & Schwede, T. (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42, W252-258.

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.

Bogaert, D., De Groot, R. & Hermans, P. W. (2004). *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* 4, 144-154.

Bolger, A. M., Lohse, M. & Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114-2120.

Briles, D. E., Ades, E., Paton, J. C., Sampson, J. S., Carlone, G. M., Huebner, R. C., Virolainen, A., Swiatlo, E. & Hollingshead, S. K. (2000). Intranasal immunization of mice with a mixture of the pneumococcal proteins PsaA and PspA is highly protective against nasopharyngeal carriage of *Streptococcus pneumoniae*. *Infect Immun* 68, 796-800.

Brook, I. (2013). Acute sinusitis in children. *Pediatr Clin North Am* 60, 409-424.

Brooks, L. R. K. & Mias, G. I. (2018). *Streptococcus pneumoniae*'s virulence and host immunity: Aging, diagnostics, and prevention. *Front Immunol* 9, 1366.

Brown, A. O., Mann, B., Gao, G., Hankins, J. S., Humann, J., Giardina, J., Faverio, P., Restrepo, M. I., Halade, G. V., Mortensen, E. M., Lindsey, M. L., Hanes, M., Happel, K. I., Nelson, S., Bagby, G. J., Lorent, J. A., Cardinal, P., Granados, R.,

Esteban, A., Lesaux, C. J., Tuomanen, E. I. & Orihuela, C. J. (2014). *Streptococcus pneumoniae* translocates into the myocardium and forms unique microlesions that disrupt cardiac function. *PLoS Pathog* 10**,** e1004383.

Brown, J. S., Gilliland, S. M. & Holden, D. W. (2001a). A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Mol Microbiol* 40**,** 572-585.

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, J. S., Ogunniyi, A. D., Woodrow, M. C., Holden, D. W. & Paton, J. C. (2001b). Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection. *Infect Immun* 69**,** 6702-6706.

Bruinsma, N., Kristinsson, K. G., Bronzwaer, S., Schrijnemakers, P., Degener, J., Tiemersma, E., Hryniewicz, W., Monen, J., Grundmann, H. & European Antimicrobial Resistance Surveillance, S. (2004). Trends of penicillin and erythromycin resistance among invasive *Streptococcus pneumoniae* in Europe. *J Antimicrob Chemother* 54**,** 1045-1050.

Bryant, K. A., Block, S. L., Baker, S. A., Gruber, W. C., Scott, D. A. & Group, P. C. V. I. S. (2010). Safety and immunogenicity of a 13-valent pneumococcal conjugate vaccine. *Pediatrics* 125**,** 866-875.

Buckwalter, C. M. & King, S. J. (2012). Pneumococcal carbohydrate transport: food for thought. *Trends Microbiol* 20**,** 517-522.

Butler, J. C., Shapiro, E. D. & Carlone, G. M. (1999). Pneumococcal vaccines: history, current status and future directions. *The Americal Journal of Medicine* 107**,** 69 - 76.

Canvin, J. R., Marvin, A. P., Sivakumaran, M., Paton, J. C., Boulnois, G. J., Andrew, P. W. & Mitchell, T. J. (1995). The role of pneumolysin and autolysin in the pathology of pneumonia and septicemia in mice infected with a type 2 pneumococcus. *J Infect Dis* 172**,** 119-123.

Carroll, M. C. (2004). The complement system in regulation of adaptive immunity. *Nat Immunol* 5**,** 981-986.

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.

Chen, A., Mann, B., Gao, G., Heath, R., King, J., Maissoneuve, J., Alderson, M., Tate, A., Hollingshead, S. K., Tweten, R. K., Briles, D. E., Tuomanen, E. I. & Paton, J. C. (2015). Multivalent pneumococcal protein vaccines comprising pneumolysoid with epitopes/fragments of CbpA and/or PspA elicit strong and broad protection. *Clin Vaccine Immunol* 22**,** 1079-1089.

Chen, Z. & Cole, P. A. (2015). Synthetic approaches to protein phosphorylation. *Curr Opin Chem Biol* 28**,** 115-122.

Cherazard, R., Epstein, M., Doan, T. L., Salim, T., Bharti, S. & Smith, M. A. (2017). Antimicrobial resistant *Streptococcus pneumoniae*: prevalence, mechanisms, and clinical implications. *Am J Ther* 24**,** e361-e369.

Claverys, J. P., Martin, B. & Havarstein, L. S. (2007). Competence-induced fratricide in *Streptococci*. *Mol Microbiol* 64**,** 1423-1433.

Claverys, J. P., Prudhomme, M. & Martin, B. (2006). Induction of competence regulons as a general response to stress in Gram-positive bacteria. *Annu Rev Microbiol* 60**,** 451-475.

Cornick, J. E. & Bentley, S. D. (2012). *Streptococcus pneumoniae*: the evolution of antimicrobial resistance to beta-lactams, fluoroquinolones and macrolides. *Microbes Infect* 14**,** 573-583.

Cremers, A. J., Zomer, A. L., Gritzfeld, J. F., Ferwerda, G., Van Hijum, S. A., Ferreira, D. M., Shak, J. R., Klugman, K. P., Boekhorst, J., Timmerman, H. M., De Jonge, M. I., Gordon, S. B. & Hermans, P. W. (2014). The adult nasopharyngeal microbiome as a determinant of pneumococcal acquisition. *Microbiome* 2**,** 44.

References

Croucher, N. J., Lochen, A. & Bentley, S. D. (2018). Pneumococcal vaccines: host interactions, population dynamics, and design principles. *Annu Rev Microbiol* 72**,** 521-549.

Cundell, D. R., Gerard, N. P., Gerard, C., Idanpaan-Heikkila, I. & Tuomanen, E. I. (1995). *Streptococcus pneumoniae* anchor to activated human cells by the receptor for platelet-activating factor. *Nature* 377**,** 435-438.

Dagan, R. (2009). Impact of pneumococcal conjugate vaccine on infections caused by antibiotic-resistant *Streptococcus pneumoniae*. *Clin Microbiol Infect* 15 Suppl 3**,** 16-20.

Daniels, C. C., Rogers, P. D. & Shelton, C. M. (2016). A review of pneumococcal vaccines: current polysaccharide vaccine recommendations and future protein antigens. *J Pediatr Pharmacol Ther* 21**,** 27-35.

Danishyar, A. & Ashurst, J. V. 2019. *Otitis, Media, Acute. StatPearls* [Online]. Available: https://www.ncbi.nlm.nih.gov/books/NBK470332/ [Accessed].

Dave, S., Carmicle, S., Hammerschmidt, S., Pangburn, M. K. & Mcdaniel, L. S. (2004). Dual roles of PspC, a surface protein of *Streptococcus pneumoniae*, in binding human secretory IgA and factor H. *J Immunol* 173**,** 471-477.

David, S. C., Laan, Z., Minhas, V., Chen, A. Y., Davies, J., Hirst, T. R., Mccoll, S. R., Alsharifi, M. & Paton, J. C. (2019). Enhanced safety and immunogenicity of a pneumococcal surface antigen A mutant whole-cell inactivated pneumococcal vaccine. *Immunol Cell Biol* 97**,** 726-739.

Davis, K. M., Nakamura, S. & Weiser, J. N. (2011). Nod2 sensing of lysozyme-digested peptidoglycan promotes macrophage recruitment and clearance of *S. pneumoniae* colonization in mice. *J Clin Invest* 121**,** 3666-3676.

Dawid, S., Roche, A. M. & Weiser, J. N. (2007). The *blp* bacteriocins of *Streptococcus pneumoniae* mediate intraspecies competition both *in vitro* and *in vivo*. *Infect Immun* 75**,** 443-451.

Dawson, M. H. (1928). The interconvertability of "R" and "S" forms of pneumococcus. *Journal of Experimental Medicine* 47**,** 577 - 591.

Dawson, M. H. & Sia, R. H. P. (1931a). *In vitro* transformation of pneumococcal types: I. A technique for inducing transformation of pneumococcal types *in vitro. Journal of Experimental Medicine* 54**,** 681 - 699.

Dawson, M. H. & Warbasse, A. (1931b). Further observations on the transformation of type-specific pneumococci by *in vitro* procedures. *Proceedings of the Society for Experimental Biology and Medicine* 29**,** 149 - 151.

De Vos, A. F., Dessing, M. C., Lammers, A. J., De Porto, A. P., Florquin, S., De Boer, O. J., De Beer, R., Terpstra, S., Bootsma, H. J., Hermans, P. W., Van 'T Veer, C. & Van Der Poll, T. (2015). The polysaccharide capsule of *Streptococcus pneumoniae* partially impedes MyD88-mediated immunity during pneumonia in mice. *PLoS One* 10**,** e0118181.

Deibel, R. H., Seeley, H. W. Jr., 1974. *Family II: Streptococcaceae. Fam. nov. In: Buchanan RE, Gibbons NE eds Bergy's manual of determinative bacteriology,* Baltimore, Williams & Wilkins.

Dessing, M. C., Schouten, M., Draing, C., Levi, M., Von Aulock, S. & Van Der Poll, T. (2008). Role played by Toll-like receptors 2 and 4 in lipoteichoic acid-induced lung inflammation and coagulation. *J Infect Dis* 197**,** 245-252.

Dintilhac, A., Alloing, G., Granadel, C. & Claverys, J. P. (1997). Competence and virulence of *Streptococcus pneumoniae*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. *Mol Microbiol* 25**,** 727-739.

Dochez, A. R. & Avery, O. T. (1917). The elaboration of specific soluble substance by pneumococcus during growth. *J Exp Med* 26**,** 477-493.

Dogan, S., Zhang, Q., Pridmore, A. C., Mitchell, T. J., Finn, A. & Murdoch, C. (2011). Pneumolysin-induced CXCL8 production by nasopharyngeal epithelial cells is dependent on calcium flux and MAPK activation via Toll-like receptor 4. *Microbes Infect* 13**,** 65-75.

Eberth, C. J. (1880). Zur Kenntniss der mykotischen Processe. *Deutsches Archiv für Klinische Medizin (Leipzig)* 3**,** 1 - 42.

References

Echenique, J., Kadioglu, A., Romao, S., Andrew, P. W. & Trombe, M. C. (2004). Protein serine/threonine kinase StkP positively controls virulence and competence in *Streptococcus pneumoniae*. *Infect Immun* 72**,** 2434-2437.

Echlin, H., Frank, M., Rock, C. & Rosch, J. W. (2020). Role of the pyruvate metabolic network on carbohydrate metabolism and virulence in *Streptococcus pneumoniae*. *Mol Microbiol* 114**,** 536-552.

Echlin, H., Frank, M. W., Iverson, A., Chang, T. C., Johnson, M. D., Rock, C. O. & Rosch, J. W. (2016). Pyruvate oxidase as a critical link between metabolism and capsule biosynthesis in *Streptococcus pneumoniae*. *PLoS Pathog* 12**,** e1005951.

Ekwurzel, G. D. (1938). Studies of immunizing substances in pneumococci. VIII. Report on field test to determine the prophylactic value of a pneumococcus antigen. *Public Health Reports* 53**,** 1877 - 1893.

El Moujaber, G., Osman, M., Rafei, R., Dabboussi, F. & Hamze, M. (2017). Molecular mechanisms and epidemiology of resistance in *Streptococcus pneumoniae* in the Middle East region. *J Med Microbiol* 66**,** 847-858.

Feldman, C. & Anderson, R. (2020). Recent advances in the epidemiology and prevention of *Streptococcus pneumoniae* infections. *F1000Res* 9. F1000 FacultyRev-338

Feldman, C., Read, R., Rutman, A., Jeffery, P. K., Brain, A., Lund, V., Mitchell, T. J., Andrew, P. W., Boulnois, G. J., Todd, H. C. et al., (1992). The interaction of *Streptococcus pneumoniae* with intact human respiratory mucosa *in vitro*. *Eur Respir J* 5**,** 576-583.

Felton, L. D. (1938). Studies on immunizing substances in pneumococci. VII. Response in human beings to antigenic pneumococcus polysaccharides, type I and II. *Public Health Reports***,** 2855 - 2877.

Fine, M. J., Smith, M. A., Carson, C. A., Meffe, F., Sankey, S. S., Weissfeld, L. A., Detsky, A. S. & Kapoor, W. N. (1994). Efficacy of pneumococcal vaccination in adults. A meta-analysis of randomized controlled trials. *Arch Intern Med* 154**,** 2666-2677.

Forrest, J. M., Mcintyre, P. B. & Burgess, M. A. (2000). Pneumococcal disease in Australia. *Commun Dis Intell* 24**,** 89-92.

Galante, J., Ho, A. C., Tingey, S. & Charalambous, B. M. (2015). Quorum sensing and biofilms in the pathogen, *Streptococcus pneumoniae*. *Curr Pharm Des* 21**,** 25-30.

Ganaie, F., Saad, J. S., Mcgee, L., Van Tonder, A. J., Bentley, S. D., Lo, S. W., Gladstone, R. A., Turner, P., Keenan, J. D., Breiman, R. F. & Nahm, M. H. (2020). A new pneumococcal capsule type, 10D, is the 100th serotype and has a large *cps* fragment from an oral *Streptococcus*. *mBio* 11. e00973-20.

G. B. D. Collaborators (2018). Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory infections in 195 countries, 1990-2016: a systematic analysis for the Global Burden of Disease Study 2016. *Lancet Infect Dis* 18**,** 1191-1210.

Geno, K. A., Gilbert, G. L., Song, J. Y., Skovsted, I. C., Klugman, K. P., Jones, C., Konradsen, H. B. & Nahm, M. H. (2015). Pneumococcal capsules and their types: past, present, and future. *Clin Microbiol Rev* 28**,** 871-899.

Gladstone, R. A., Devine, V., Jones, J., Cleary, D., Jefferies, J. M., Bentley, S. D., Faust, S. N. & Clarke, S. C. (2017). Pre-vaccine serotype composition within a lineage signposts its serotype replacement - a carriage study over 7 years following pneumococcal conjugate vaccine use in the UK. *Microb Genom* 3**,** e000119.

Gram, C. (1884). Ueber die isolierte Färbung der Schnizomyceten in Scnittund Trockenpräparten. *Fortschr Med* 12**,** 185 - 189.

Griffith, F. 1923. The influence of immune serum on the biological properties of pneumococci. *Reports on public health and medical subjects.* London: His Majesty's Stationery Office.

Griffith, F. (1928). The significance of pneumococcal types. *Journal of Hygeine* 27**,** 113 - 159.

Guemes, M., Rahman, S. A. & Hussain, K. (2016). What is a normal blood glucose? *Arch Dis Child* 101**,** 569-574.

References

Guiral, S., Mitchell, T. J., Martin, B. & Claverys, J. P. (2005). Competence-programmed predation of noncompetent cells in the human pathogen *Streptococcus pneumoniae*: genetic requirements. *Proc Natl Acad Sci U S A* 102**,** 8710-8715.

Gwaltney, J. M., Jr., Sande, M. A., Austrian, R. & Hendley, J. O. (1975). Spread of *Streptococcus pneumoniae* in families. II. Relation of transfer of S*. pneumoniae* to incidence of colds and serum antibody. *J Infect Dis* 132**,** 62-68.

Hamilton, J. A., Stanley, E. R., Burgess, A. W. & Shadduck, R. K. (1980). Stimulation of macrophage plasminogen activator activity by colony-stimulating factors. *J Cell Physiol* 103**,** 435-445.

Hammerschmidt, S., Tillig, M. P., Wolff, S., Vaerman, J. P. & Chhatwal, G. S. (2000). Species-specific binding of human secretory component to SpsA protein of *Streptococcus pneumoniae* via a hexapeptide motif. *Mol Microbiol* 36**,** 726-736.

Hammerschmidt, S., Wolff, S., Hocke, A., Rosseau, S., Muller, E. & Rohde, M. (2005). Illustration of pneumococcal polysaccharide capsule during adherence and invasion of epithelial cells. *Infect Immun* 73**,** 4653-4667.

Hammitt, L. L., Bruden, D. L., Butler, J. C., Baggett, H. C., Hurlburt, D. A., Reasonover, A. & Hennessy, T. W. (2006). Indirect effect of conjugate vaccine on adult carriage of *Streptococcus pneumoniae*: an explanation of trends in invasive pneumococcal disease. *J Infect Dis* 193**,** 1487-1494.

Hansman, D. & Bullen, M. M. (1967). A resistant pneumococcus. *Lancet* 2**,** 264 - 265.

Hansman, D., Glasgow, H., Sturt, J., Devitt, L. & Douglas, R. (1971). Increased resistance to penicillin of pneumococci isolated from man. *N Engl J Med* 284**,** 175-177*.*

Harboe, Z. B., Thomsen, R. W., Riis, A., Valentiner-Branth, P., Christensen, J. J., Lambertsen, L., Krogfelt, K. A., Konradsen, H. B. & Benfield, T. L. (2009). Pneumococcal serotypes and mortality following invasive pneumococcal disease: a population-based cohort study. *PLoS Med* 6**,** e1000081.

Harvey, R. M., Hughes, C. E., Paton, A. W., Trappetti, C., Tweten, R. K. & Paton, J. C. (2014). The impact of pneumolysin on the macrophage response to *Streptococcus pneumoniae* is strain-dependent. *PLoS One* 9, e103625.

Hava, D. L. & Camilli, A. (2002). Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol* 45, 1389-1406.

Heilmann, C. (1990). Human B and T lymphocyte responses to vaccination with pneumococcal polysaccharides. *APMIS Suppl* 15, 1-23.

Himes, S. R., Reeves, R., Attema, J., Nissen, M., Li, Y. & Shannon, M. F. (2000). The role of high-mobility group I(Y) proteins in expression of IL-2 and T cell proliferation. *J Immunol* 164, 3157-3168.

Hirschfeld, C., Gomez-Mejia, A., Bartel, J., Hentschker, C., Rohde, M., Maass, S., Hammerschmidt, S. & Becher, D. (2019). Proteomic investigation uncovers potential targets and target sites of pneumococcal Serine-Threonine Kinase StkP and phosphatase PhpP. *Front Microbiol* 10, 3101.

Hirst, R. A., Kadioglu, A., O'callaghan, C. & Andrew, P. W. (2004). The role of pneumolysin in pneumococcal pneumonia and meningitis. *Clin Exp Immunol* 138, 195-201.

Hobbs, J. K., Meier, E. P. W., Pluvinage, B., Mey, M. A. & Boraston, A. B. (2019). Molecular analysis of an enigmatic *Streptococcus pneumoniae* virulence factor: The raffinose-family oligosaccharide utilization system. *J Biol Chem* 294, 17197-17208.

Hoe, E., Anderson, J., Nathanielsz, J., Toh, Z. Q., Marimla, R., Balloch, A. & Licciardi, P. V. (2017). The contrasting roles of Th17 immunity in human health and disease. *Microbiol Immunol* 61, 49-56.

Holmes, A. R., Mcnab, R., Millsap, K. W., Rohde, M., Hammerschmidt, S., Mawdsley, J. L. & Jenkinson, H. F. (2001). The *pavA* gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. *Mol Microbiol* 41, 1395-1408.

Hoskins, J., Alborn, W. E., Jr., Arnold, J., Blaszczak, L. C., Burgett, S., Dehoff, B. S., Estrem, S. T., Fritz, L., Fu, D. J., Fuller, W., Geringer, C., Gilmour, R., Glass, J.

References

S., Khoja, H., Kraft, A. R., Lagace, R. E., Leblanc, D. J., Lee, L. N., Lefkowitz, E. J., Lu, J., Matsushima, P., Mcahren, S. M., Mchenney, M., Mcleaster, K., Mundy, C. W., Nicas, T. I., Norris, F. H., O'gara, M., Peery, R. B., Robertson, G. T., Rockey, P., Sun, P. M., Winkler, M. E., Yang, Y., Young-Bellido, M., Zhao, G., Zook, C. A., Baltz, R. H., Jaskunas, S. R., Rosteck, P. R., Jr., Skatrud, P. L. & Glass, J. I. (2001). Genome of the bacterium *Streptococcus pneumoniae* strain R6. *J Bacteriol* 183**,** 5709-5717.

Hosseini, S. M., Poorolajal, J., Karami, M. & Ameri, P. (2015). Prevalence of nasopharyngeal carriage of *Streptococcus pneumoniae* in Iran: a meta-analysis. *J Res Health Sci* 15**,** 141-146.

Howard, L. V. & Gooder, H. (1974). Specificity of the autolysin of *Streptococcus (Diplococcus) pneumoniae*. *J Bacteriol* 117**,** 796-804.

Huffman, J. L., Lu, F., Zalkin, H. & Brennan, R. G. (2002). Role of residue 147 in the gene regulatory function of the *Escherichia coli* purine repressor. *Biochemistry* 41**,** 511-520.

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.

Iannelli, F. & Pozzi, G. (2004). Method for introducing specific and unmarked mutations into the chromosome of *Streptococcus pneumoniae*. *Mol Biotechnol* 26**,** 81-86.

Ibrahim, Y. M., Kerr, A. R., Mccluskey, J. & Mitchell, T. J. (2004). Control of virulence by the two-component system CiaR/H is mediated via HtrA, a major virulence factor of *Streptococcus pneumoniae*. *J Bacteriol* 186**,** 5258-5266.

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.

Janeway, C. a. J., Travers, P., Walport, M. & Shlomchik, M. J. 2001. *Immunobiology, 5th edition,* New York, Garland Science.

Janulczyk, R., Iannelli, F., Sjoholm, A. G., Pozzi, G. & Bjorck, L. (2000). Hic, a novel surface protein of *Streptococcus pneumoniae* that interferes with complement function. *J Biol Chem* 275**,** 37257-37263.

Jedrzejas, M. J. (2001). Pneumococcal virulence factors: structure and function. *Microbiol Mol Biol Rev* 65**,** 187-207.

Jedrzejas, M. J., Lamani, E. & Becker, R. S. (2001). Characterization of selected strains of pneumococcal surface protein A. *J Biol Chem* 276**,** 33121-33128.

Jensch, I., Gamez, G., Rothe, M., Ebert, S., Fulde, M., Somplatzki, D., Bergmann, S., Petruschka, L., Rohde, M., Nau, R. & Hammerschmidt, S. (2010). PavB is a surface-exposed adhesin of *Streptococcus pneumoniae* contributing to nasopharyngeal colonization and airways infections. *Mol Microbiol* 77**,** 22-43.

Johnston, J. W., Myers, L. E., Ochs, M. M., Benjamin, W. H., Jr., 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.

Jounblat, R., Kadioglu, A., Mitchell, T. J. & Andrew, P. W. (2003). Pneumococcal behavior and host responses during bronchopneumonia are affected differently by the cytolytic and complement-activating activities of pneumolysin. *Infect Immun* 71**,** 1813-1819.

Kadioglu, A., Coward, W., Colston, M. J., Hewitt, C. R. & Andrew, P. W. (2004). CD4-T-lymphocyte interactions with pneumolysin and pneumococci suggest a crucial protective role in the host response to pneumococcal infection. *Infect Immun* 72**,** 2689-2697.

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.

Karmakar, M., Katsnelson, M., Malak, H. A., Greene, N. G., Howell, S. J., Hise, A. G., Camilli, A., Kadioglu, A., Dubyak, G. R. & Pearlman, E. (2015). Neutrophil IL-1beta processing induced by pneumolysin is mediated by the NLRP3/ASC

inflammasome and caspase-1 activation and is dependent on K+ efflux. *J Immunol* 194**,** 1763-1775.

Kaufman, P. (1947). Pneumonia in old age, active immunization against pneumonia with pneumococcus polysaccharides; results of a six year study. *Archives of Internal Medicine***,** 518 - 531.

Kaur, A., Capalash, N. & Sharma, P. (2018). Quorum sensing in thermophiles: prevalence of autoinducer-2 system. *BMC Microbiol* 18**,** 62.

Kerr, A. R., Paterson, G. K., Riboldi-Tunnicliffe, A. & Mitchell, T. J. (2005). Innate immune defense against pneumococcal pneumonia requires pulmonary complement component C3. *Infect Immun* 73**,** 4245-4252.

King, S. J. (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., Gould, J. M., Bae, D., Peterson, S., Cline, R. T., Fasching, C., Janoff, E. N. & Weiser, J. N. (2004). Phase variable desialylation of host proteins that bind to *Streptococcus pneumoniae in vivo* and protect the airway. *Mol Microbiol* 54**,** 159-171.

Kirkham, L. A., Jefferies, J. M., Kerr, A. R., Jing, Y., Clarke, S. C., Smith, A. & Mitchell, T. J. (2006). Identification of invasive serotype 1 pneumococcal isolates that express nonhemolytic pneumolysin. *J Clin Microbiol* 44**,** 151-159.

Kjos, M., Miller, E., Slager, J., Lake, F. B., Gericke, O., Roberts, I. S., Rozen, D. E. & Veening, J. W. (2016). Expression of *Streptococcus pneumoniae* bacteriocins is induced by antibiotics via regulatory interplay with the competence system. *PLoS Pathog* 12**,** e1005422.

Klebs, E. (1875). Beiträge zur Kenntniss der Schistomyceten. VII. Die Monadinen. *Archiv für Experimentelle Pathologie und Pharmakologie (Leipzig)* 4**,** 409 - 488.

Kloosterman, T. G., Bijlsma, J. J. E., Kok, J. & Kuipers, O. P. (2006). To have neighbour's fare: extending the molecular toolbox for *Streptococcus pneumoniae*. *Microbiology (Reading)* 152**,** 351-359.

Koppe, U., Suttorp, N. & Opitz, B. (2012). Recognition of *Streptococcus pneumoniae* by the innate immune system. *Cell Microbiol* 14**,** 460-466.

Kowalko, J. E. & Sebert, M. E. (2008). The *Streptococcus pneumoniae* competence regulatory system influences respiratory tract colonization. *Infect Immun* 76**,** 3131-3140.

Ku, C. L., Von Bernuth, H., Picard, C., Zhang, S. Y., Chang, H. H., Yang, K., Chrabieh, M., Issekutz, A. C., Cunningham, C. K., Gallin, J., Holland, S. M., Roifman, C., Ehl, S., Smart, J., Tang, M., Barrat, F. J., Levy, O., Mcdonald, D., Day-Good, N. K., Miller, R., Takada, H., Hara, T., Al-Hajjar, S., Al-Ghonaium, A., Speert, D., Sanlaville, D., Li, X., Geissmann, F., Vivier, E., Marodi, L., Garty, B. Z., Chapel, H., Rodriguez-Gallego, C., Bossuyt, X., Abel, L., Puel, A. & Casanova, J. L. (2007). Selective predisposition to bacterial infections in IRAK-4-deficient children: IRAK-4-dependent TLRs are otherwise redundant in protective immunity. *J Exp Med* 204**,** 2407-2422.

Kuo, C. H. & Ochman, H. (2010). The extinction dynamics of bacterial pseudogenes. *PLoS Genet* 6.

Langmead, B. & Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9**,** 357-359.

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.

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.

Letiembre, M., Echchannaoui, H., Bachmann, P., Ferracin, F., Nieto, C., Espinosa, M. & Landmann, R. (2005). Toll-like receptor 2 deficiency delays pneumococcal phagocytosis and impairs oxidative killing by granulocytes. *Infect Immun* 73**,** 8397-8401.

References

Levitt, L. J., Nagler, A., Lee, F., Abrams, J., Shatsky, M. & Thompson, D. (1991). Production of granulocyte/macrophage-colony-stimulating factor by human natural killer cells. Modulation by the p75 subunit of the interleukin 2 receptor and by the CD2 receptor. *J Clin Invest* 88, 67-75.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. & Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Liao, Y., Smyth, G. K. & Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res* 47, e47.

Limoli, D. H., Sladek, J. A., Fuller, L. A., Singh, A. K. & King, S. J. (2011). BgaA acts as an adhesin to mediate attachment of some pneumococcal strains to human epithelial cells. *Microbiology (Reading)* 157, 2369-2381.

Lister, F. S. (1917). Prophylactic inoculation of man against pneumococcal infections, and more particularly against lobar pneumonia. *Publications from the South African Institute for Medical Research* 10, 304 - 322.

Livak, K. J. & Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402-408.

Lock, R. A., Paton, J. C. & Hansman, D. (1988). Comparative efficacy of pneumococcal neuraminidase and pneumolysin as immunogens protective against *Streptococcus pneumoniae. Microb Pathog* 5, 461-467.

Lock, R. A., Zhang, Q. Y., Berry, A. M. & Paton, J. C. (1996). Sequence variation in the *Streptococcus pneumoniae* pneumolysin gene affecting haemolytic activity and electrophoretic mobility of the toxin. *Microb Pathog* 21, 71-83.

Lu, S., Wang, J., Chitsaz, F., Derbyshire, M. K., Geer, R. C., Gonzales, N. R., Gwadz, M., Hurwitz, D. I., Marchler, G. H., Song, J. S., Thanki, N., Yamashita, R. A., Yang, M., Zhang, D., Zheng, C., Lanczycki, C. J. & Marchler-Bauer, A. (2020). CDD/SPARCLE: the conserved domain database in 2020. *Nucleic Acids Res* 48, D265-D268.

Lu, Y. J., Leite, L., Goncalves, V. M., Dias Wde, O., Liberman, C., Fratelli, F., Alderson, M., Tate, A., Maisonneuve, J. F., Robertson, G., Graca, R., Sayeed, S., Thompson, C. M., Anderson, P. & Malley, R. (2010). GMP-grade pneumococcal whole-cell vaccine injected subcutaneously protects mice from nasopharyngeal colonization and fatal aspiration-sepsis. *Vaccine* 28**,** 7468-7475.

Luster, A. D., Alon, R. & Von Andrian, U. H. (2005). Immune cell migration in inflammation: present and future therapeutic targets. *Nat Immunol* 6**,** 1182-1190.

Lynch, J. P., 3rd & Zhanel, G. G. (2009). *Streptococcus pneumoniae*: does antimicrobial resistance matter? *Semin Respir Crit Care Med* 30**,** 210-238.

Lynch, J. P., 3rd & Zhanel, G. G. (2010). *Streptococcus pneumoniae*: epidemiology and risk factors, evolution of antimicrobial resistance, and impact of vaccines. *Curr Opin Pulm Med* 16**,** 217-225.

MacLeod, C. M., Hodges, R. G., Heidelberger, M. & Bernhardm W. G. (1945). Prevention of pneumococcal pneumonia by immunization with specific capsular polysaccharides. *Journal of Experimental Medicine* 82**,** 445 - 465.

Mahdi, L. K., Wang, H., Van Der Hoek, M. B., Paton, J. C. & Ogunniyi, A. D. (2012). Identification of a novel pneumococcal vaccine antigen preferentially expressed during meningitis in mice. *J Clin Invest* 122**,** 2208-2220.

Malley, R., Henneke, P., Morse, S. C., Cieslewicz, M. J., Lipsitch, M., Thompson, C. M., Kurt-Jones, E., Paton, J. C., Wessels, M. R. & Golenbock, D. T. (2003). Recognition of pneumolysin by Toll-like receptor 4 confers resistance to pneumococcal infection. *Proc Natl Acad Sci U S A* 100**,** 1966-1971.

Malley, R., Lipsitch, M., Stack, A., Saladino, R., Fleisher, G., Pelton, S., Thompson, C., Briles, D. & Anderson, P. (2001). Intranasal immunization with killed unencapsulated whole cells prevents colonization and invasive disease by capsulated pneumococci. *Infect Immun* 69**,** 4870-4873.

Mamishi, S., Moradkhani, S., Mahmoudi, S., Hosseinpour-Sadeghi, R. & Pourakbari, B. (2014). Penicillin-Resistant trend of *Streptococcus pneumonia*e in Asia: A systematic review. *Iran J Microbiol* 6**,** 198-210.

References

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.

Marion, C., Aten, A. E., Woodiga, S. A. & King, S. J. (2011). Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in *Streptococcus pneumoniae. Infect Immun* 79**,** 4193-4200.

Marks, L. R., Reddinger, R. M. & Hakansson, A. P. (2014). Biofilm formation enhances fomite survival of *Streptococcus pneumoniae* and *Streptococcus pyogenes. Infect Immun* 82**,** 1141-1146.

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.

McCullers, J. A. & Rehg, J. E. (2002). Lethal synergism between influenza virus and *Streptococcus pneumoniae*: characterization of a mouse model and the role of platelet-activating factor receptor. *J Infect Dis* 186**,** 341-350.

McDaniel, L. S., Sheffield, J. S., Delucchi, P. & Briles, D. E. (1991). PspA, a surface protein of *Streptococcus pneumoniae*, is capable of eliciting protection against pneumococci of more than one capsular type. *Infect Immun* 59**,** 222-228.

McDaniel, L. S., Yother, J., Vijayakumar, M., Mcgarry, L., Guild, W. R. & Briles, D. E. (1987). Use of insertional inactivation to facilitate studies of biological properties of pneumococcal surface protein A (PspA). *J Exp Med* 165**,** 381-394.

McLean, K. T., Tikhomirova, A., Brazel, E. B., Legendre, S., Haasbroek, G., Minhas, V., Paton, J. C. & Trappetti, C. (2020). Site-specific mutations of GalR affect galactose metabolism in *Streptococcus pneumoniae. J Bacteriol* 203. e00180-20.

Miller, M. B. & Bassler, B. L. (2001). Quorum sensing in bacteria. *Annu Rev Microbiol* 55**,** 165-199.

Minhas, V., Aprianto, R., McAllister, L. J., Wang, H., David, S. C., McLean, K. T., Comerford, I., McColl, S. R., Paton, J. C., Veening, J. W. & Trappetti, C. (2020). *In*

vivo dual RNA-seq reveals that neutrophil recruitment underlies differential tissue tropism of *Streptococcus pneumoniae*. *Commun Biol* 3**,** 293.

Minhas, V., Harvey, R. M., McAllister, L. J., Seemann, T., Syme, A. E., Baines, S. L., Paton, J. C. & Trappetti, C. (2019). Capacity to utilize raffinose dictates pneumococcal disease phenotype. *mBio* 10. e02956-18.

Mitchell, T. J., Andrew, P. W., Saunders, F. K., Smith, A. N. & Boulnois, G. J. (1991). Complement activation and antibody binding by pneumolysin via a region of the toxin homologous to a human acute-phase protein. *Mol Microbiol* 5**,** 1883-1888.

Moffitt, K. & Malley, R. (2016). Rationale and prospects for novel pneumococcal vaccines. *Hum Vaccin Immunother* 12**,** 383-392.

Mook-Kanamori, B. B., Geldhoff, M., Van Der Poll, T. & Van De Beek, D. (2011). Pathogenesis and pathophysiology of pneumococcal meningitis. *Clin Microbiol Rev* 24**,** 557-591.

Morais, V., Texeira, E. & Suarez, N. (2019). Next-generation whole-cell pneumococcal vaccine. *Vaccines (Basel)* 7. 151.

Moreno-Gamez, S., Sorg, R. A., Domenech, A., Kjos, M., Weissing, F. J., Van Doorn, G. S. & Veening, J. W. (2017). Quorum sensing integrates environmental cues, cell density and cell history to control bacterial competence. *Nat Commun* 8**,** 854.

Morens, D. M., Taubenberger, J. K. & Fauci, A. S. (2008). Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness. *J Infect Dis* 198**,** 962-970.

Morona, J. K., Morona, R. & Paton, J. C. (2006). Attachment of capsular polysaccharide to the cell wall of *Streptococcus pneumoniae* type 2 is required for invasive disease. *Proc Natl Acad Sci U S A* 103**,** 8505-8510.

Musher, D. M. (2003). How contagious are common respiratory tract infections? *N Engl J Med* 348**,** 1256-1266.

References

Nakamura, S., Davis, K. M. & Weiser, J. N. (2011). Synergistic stimulation of type I interferons during influenza virus coinfection promotes *Streptococcus pneumoniae* colonization in mice. *J Clin Invest* 121**,** 3657-3665.

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.

Neufeld, F. 1902. Ueber die Agglutination der Pneumokokken und über die Theorien der Agglutination. *Zeitschrift für Hygiene und Infektionskrankheiten* 40, 54-72.

Neufeld, F. & Levinthal, W. (1928). Beiträge zur Variabilität der Pneumokokken. *Zeitschrift für Immunitätforschung und Experimentelle Therapie* 55**,** 324 - 340.

Nguyen, C. C. & Saier, M. H., Jr. (1995). Phylogenetic, structural and functional analyses of the LacI-GalR family of bacterial transcription factors. *FEBS Lett* 377**,** 98-102.

Novakova, L., Saskova, L., Pallova, P., Janecek, J., Novotna, J., Ulrych, A., Echenique, J., Trombe, M. C. & Branny, P. (2005). Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of *Streptococcus pneumoniae* and identification of kinase substrates. *FEBS J* 272**,** 1243-1254.

Novick, R. P. (2003). Autoinduction and signal transduction in the regulation of staphylococcal virulence. *Mol Microbiol* 48**,** 1429-1449.

Numminen, E., Chewapreecha, C., Turner, C., Goldblatt, D., Nosten, F., Bentley, S. D., Turner, P. & Corander, J. (2015). Climate induces seasonality in pneumococcal transmission. *Sci Rep* 5**,** 11344.

Office of Technology Assessments, U.S. Congress. (1979) *A review of selected federal vaccine and immunization policies.* Washington, DC: U.S. Government Printing Office.

O'Neill, L. A. & Bowie, A. G. (2007). The family of five: TIR-domain-containing adaptors in Toll-like receptor signalling. *Nat Rev Immunol* 7**,** 353-364.

Orihuela, C. J., Gao, G., Francis, K. P., Yu, J. & Tuomanen, E. I. (2004). Tissue-specific contributions of pneumococcal virulence factors to pathogenesis. *J Infect Dis* 190, 1661-1669.

Orihuela, C. J., Mahdavi, J., Thornton, J., Mann, B., Wooldridge, K. G., Abouseada, N., Oldfield, N. J., Self, T., Ala'aldeen, D. A. & Tuomanen, E. I. (2009). Laminin receptor initiates bacterial contact with the blood brain barrier in experimental meningitis models. *J Clin Invest* 119, 1638-1646.

Paixão, L., Oliveira, J., Verissimo, A., Vinga, S., Lourenco, E. C., Ventura, M. R., Kjos, M., Veening, J. W., Fernandes, V. E., Andrew, P. W., Yesilkaya, H. & Neves, A. R. (2015). Host glycan sugar-specific pathways in *Streptococcus pneumoniae*: galactose as a key sugar in colonisation and infection [corrected]. *PLoS One* 10, e0121042.

Parker, D., Martin, F. J., Soong, G., Harfenist, B. S., Aguilar, J. L., Ratner, A. J., Fitzgerald, K. A., Schindler, C. & Prince, A. (2011). *Streptococcus pneumoniae* DNA initiates type I interferon signaling in the respiratory tract. *mBio* 2, e00016-00011.

Pasteur, N. D. L. (1881). Maison de Louis Pasteur à Arbois - Publications de Louis Pasteur dans les Comptes rendus. *Comptes Rendus de l'Académie Des Sciences de Paris* 92.

Paterson, G. K. & Mitchell, T. J. (2006). The role of *Streptococcus pneumoniae* sortase A in colonisation and pathogenesis. *Microbes Infect* 8, 145-153.

Paton, J. C., Andrew, P. W., Boulnois, G. J. & Mitchell, T. J. (1993). Molecular analysis of the pathogenicity of *Streptococcus pneumoniae*: the role of pneumococcal proteins. *Annu Rev Microbiol* 47, 89-115.

Paton, J. C. & Ferrante, A. (1983). Inhibition of human polymorphonuclear leukocyte respiratory burst, bactericidal activity, and migration by pneumolysin. *Infect Immun* 41, 1212-1216.

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.

References

Paton, J. C., Lock, R. A., Lee, C. J., Li, J. P., Berry, A. M., Mitchell, T. J., Andrew, P. W., Hansman, D. & Boulnois, G. J. (1991). Purification and immunogenicity of genetically obtained pneumolysin toxoids and their conjugation to *Streptococcus pneumoniae* type 19F polysaccharide. *Infect Immun* 59**,** 2297-2304.

Paton, J. C., Rowan-Kelly, B. & Ferrante, A. (1984). Activation of human complement by the pneumococcal toxin pneumolysin. *Infect Immun* 43**,** 1085-1087.

Paton, J. C. & Trappetti, C. (2019). *Streptococcus pneumoniae* capsular polysaccharide. *Microbiol Spectr* 7(2).

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.

Pezzulo, A. A., Gutierrez, J., Duschner, K. S., Mcconnell, K. S., Taft, P. J., Ernst, S. E., Yahr, T. L., Rahmouni, K., Klesney-Tait, J., Stoltz, D. A. & Zabner, J. (2011). Glucose depletion in the airway surface liquid is essential for sterility of the airways. *PLoS One* 6**,** e16166.

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.

Pichichero, M. E. (2017). Pneumococcal whole-cell and protein-based vaccines: changing the paradigm. *Expert Rev Vaccines* 16**,** 1181-1190.

Postma, P. W., Lengeler, J. W. & Jacobson, G. R. (1993). Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. *Microbiol Rev* 57**,** 543-594.

Powell, D. R. (2019). Degust: interactive RNA-seq analysis. *Zenodo*.

Puchta, M., Boczkowska, M. & Groszyk, J. (2020). Low RIN value for RNA-Seq library construction from long-term stored seeds: a case study of barley seeds. *Genes (Basel)* 11. 1190.

Quill, H., Gaur, A. & Phipps, R. P. (1989). Prostaglandin E2-dependent induction of granulocyte-macrophage colony-stimulating factor secretion by cloned murine helper T cells. *J Immunol* 142**,** 813-818.

Quin, L. R., Carmicle, S., Dave, S., Pangburn, M. K., Evenhuis, J. P. & Mcdaniel, L. S. (2005). *In vivo* binding of complement regulator factor H by *Streptococcus pneumoniae*. *J Infect Dis* 192**,** 1996-2003.

R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

Rayner, C. F., Jackson, A. D., Rutman, A., Dewar, A., Mitchell, T. J., Andrew, P. W., Cole, P. J. & Wilson, R. (1995). Interaction of pneumolysin-sufficient and -deficient isogenic variants of *Streptococcus pneumoniae* with human respiratory mucosa. *Infect Immun* 63**,** 442-447.

Regev-Yochay, G., Raz, M., Dagan, R., Porat, N., Shainberg, B., Pinco, E., Keller, N. & Rubinstein, E. (2004). Nasopharyngeal carriage of *Streptococcus pneumoniae* by adults and children in community and family settings. *Clin Infect Dis* 38**,** 632-639.

Ren, B., Szalai, A. J., Hollingshead, S. K. & Briles, D. E. (2004). Effects of PspA and antibodies to PspA on activation and deposition of complement on the pneumococcal surface. *Infect Immun* 72**,** 114-122.

Richard, A. L., Siegel, S. J., Erikson, J. & Weiser, J. N. (2014). TLR2 signaling decreases transmission of *Streptococcus pneumoniae* by limiting bacterial shedding in an infant mouse Influenza A co-infection model. *PLoS Pathog* 10**,** e1004339.

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.

Riley, I. D., Tarr, P. I., Andrews, M., Pfeiffer, M., Howard, R., Challands, P. & Jennison, G. (1977). Immunisation with a polyvalent pneumococcal vaccine. Reduction of adult respiratory mortality in a New Guinea Highlands community. *Lancet* 1**,** 1338-1341.

References

Ritchie, M. E., Phipson, B., Wu, D., Hu, Y., Law, C. W., Shi, W. & Smyth, G. K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 43**,** e47.

Romagnani, S. (1999). Th1/Th2 cells. *Inflamm Bowel Dis* 5**,** 285-294.

Rose, M. C. & Voynow, J. A. (2006). Respiratory tract mucin genes and mucin glycoproteins in health and disease. *Physiol Rev* 86**,** 245-278.

Rosenow, C., Ryan, P., Weiser, J. N., Johnson, S., Fontan, P., Ortqvist, A. & Masure, H. R. (1997). Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae. Mol Microbiol* 25**,** 819-829.

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.

Sampson, J. S., Furlow, Z., Whitney, A. M., Williams, D., Facklam, R. & Carlone, G. M. (1997). Limited diversity of *Streptococcus pneumoniae psaA* among pneumococcal vaccine serotypes. *Infect Immun* 65**,** 1967-1971.

Sampson, J. S., O'connor, S. P., Stinson, A. R., Tharpe, J. A. & Russell, H. (1994). Cloning and nucleotide sequence analysis of *psaA*, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein homologous to previously reported *Streptococcus sp*. adhesins. *Infect Immun* 62**,** 319-324.

Schmeck, B., Huber, S., Moog, K., Zahlten, J., Hocke, A. C., Opitz, B., Hammerschmidt, S., Mitchell, T. J., Kracht, M., Rosseau, S., Suttorp, N. & Hippenstiel, S. (2006). Pneumococci induced TLR- and Rac1-dependent NF-kappaB-recruitment to the IL-8 promoter in lung epithelial cells. *Am J Physiol Lung Cell Mol Physiol* 290**,** L730-L737.

Schroder, K. & Tschopp, J. (2010). The inflammasomes. *Cell* 140**,** 821-832.

Schroder, N. W., Morath, S., Alexander, C., Hamann, L., Hartung, T., Zahringer, U., Gobel, U. B., Weber, J. R. & Schumann, R. R. (2003). Lipoteichoic acid (LTA) of

Streptococcus pneumoniae and *Staphylococcus aureus* activates immune cells via Toll-like receptor (TLR)-2, lipopolysaccharide-binding protein (LBP), and CD14, whereas TLR-4 and MD-2 are not involved. *J Biol Chem* 278**,** 15587-15594.

Schroeder, A., Mueller, O., Stocker, S., Salowsky, R., Leiber, M., Gassmann, M., Lightfoot, S., Menzel, W., Granzow, M. & Ragg, T. (2006). The RIN: an RNA integrity number for assigning integrity values to RNA measurements. *BMC Mol Biol* 7**,** 3.

Sender, V., Hentrich, K. & Henriques-Normark, B. (2021). Virus-induced changes of the respiratory tract environment promote secondary infections with *Streptococcus pneumoniae*. *Front Cell Infect Microbiol* 11**,** 643326.

Shak, J. R., Vidal, J. E. & Klugman, K. P. (2013). Influence of bacterial interactions on pneumococcal colonization of the nasopharynx. *Trends Microbiol* 21**,** 129-135.

Shakhnovich, E. A., King, S. J. & Weiser, J. N. (2002). Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of *Neisseria meningitidis* and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. *Infect Immun* 70**,** 7161-7164.

Shaper, M., Hollingshead, S. K., Benjamin, W. H., Jr. & Briles, D. E. (2004). PspA protects *Streptococcus pneumoniae* from killing by apolactoferrin, and antibody to PspA enhances killing of pneumococci by apolactoferrin [corrected]. *Infect Immun* 72**,** 5031-5040.

Shapiro, E. D., Berg, A. T., Austrian, R., Schroeder, D., Parcells, V., Margolis, A., Adair, R. K. & Clemens, J. D. (1991). The protective efficacy of polyvalent pneumococcal polysaccharide vaccine. *N Engl J Med* 325**,** 1453-1460.

Shi, Y., Liu, C. H., Roberts, A. I., Das, J., Xu, G., Ren, G., Zhang, Y., Zhang, L., Yuan, Z. R., Tan, H. S., Das, G. & Devadas, S. (2006). Granulocyte-macrophage colony-stimulating factor (GM-CSF) and T-cell responses: what we do and don't know. *Cell Res* 16**,** 126-133.

Sia, R. H. P. & Dawson, M. H. (1931). *In vitro* transformation of pneumococcal types: II. The nature of the factor resposible for the transformation of pneumococcal types. *Journal of Experimental Medicine* 54**,** 701 - 710.

References

Siegel, S. J., Roche, A. M. & Weiser, J. N. (2014). Influenza promotes pneumococcal growth during coinfection by providing host sialylated substrates as a nutrient source. *Cell Host Microbe* 16**,** 55-67.

Skov Sorensen, U. B., Blom, J., Birch-Andersen, A. & Henrichsen, J. (1988). Ultrastructural localization of capsules, cell wall polysaccharide, cell wall proteins, and F antigen in pneumococci. *Infect Immun* 56**,** 1890-1896.

Smit, P., Oberholzer, D., Hayden-Smith, S., Koornhof, H. J. & Hilleman, M. R. (1977). Protective efficacy of pneumococcal polysaccharide vaccines. *Journal of the American Medical Association* 238**,** 2613 - 2616.

Sorbara, M. T. & Philpott, D. J. (2011). Peptidoglycan: a critical activator of the mammalian immune system during infection and homeostasis. *Immunol Rev* 243**,** 40-60.

Sørenson, U. B. S., Henrichson, J., Chen, H. C. & Szu, S. C. (1990). Covalent linkage between the capsular polysaccharide and the cell wall peptidoglycan of *Streptococcus pneumoniae* revealed by immunochemical methods. *Microb Pathog* 8**,** 325 - 334.

Sprent, J. (1994). T and B memory cells. *Cell* 76**,** 315-322.

Srivastava, A., Henneke, P., Visintin, A., Morse, S. C., Martin, V., Watkins, C., Paton, J. C., Wessels, M. R., Golenbock, D. T. & Malley, R. (2005). The apoptotic response to pneumolysin is Toll-like receptor 4 dependent and protects against pneumococcal disease. *Infect Immun* 73**,** 6479-6487.

Stanek, R. J., Norton, N. B. & Mufson, M. A. (2016). A 32-year study of the effect of pneumococcal vaccines on invasive *Streptococcus pneumoniae* Disease. *Am J Med Sci* 352**,** 563-573.

Sternberg, G. M. (1881). A fatal form of septicaemia in the rabbit, produced by the subcutaneous injection of human saliva. *Annual Reports for the National Board of Health* 2**,** 87 - 108.

Sternberg, G. M. (1885). The pneumonia-coccus of Friedlander (*Micrococcus Pasteuri*, Sternberg). *American Journal of Medical Science* 90**,** 106 - 123.

Stock, A. M., Robinson, V. L. & Goudreau, P. N. (2000). Two-component signal transduction. *Annu Rev Biochem* 69**,** 183-215.

Stryker, L. M. (1916). Variations in the pneumococcus induced by growth in immune serum. *Journal of Experimental Medicine* 24**,** 49 - 68.

Sun, X., Ge, F., Xiao, C. L., Yin, X. F., Ge, R., Zhang, L. H. & He, Q. Y. (2010). Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae. J Proteome Res* 9**,** 275-282.

Sung, C. K., Li, H., Claverys, J. P. & Morrison, D. A. (2001). An *rpsL* cassette, janus, for gene replacement through negative selection in *Streptococcus pneumoniae. Appl Environ Microbiol* 67**,** 5190-5196.

Swint-Kruse, L. & Matthews, K. S. (2009). Allostery in the LacI/GalR family: variations on a theme. *Curr Opin Microbiol* 12**,** 129-137.

Thorsness, P. E. & Koshland, D. E., Jr. (1987). Inactivation of isocitrate dehydrogenase by phosphorylation is mediated by the negative charge of the phosphate. *J Biol Chem* 262**,** 10422-10425.

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.

Tomasz, A. (1965). Control of the competent state in Pneumococcus by a hormone-like cell product: an example for a new type of regulatory mechanism in bacteria. *Nature* 208**,** 155-159.

Tomlinson, G., Chimalapati, S., Pollard, T., Lapp, T., Cohen, J., Camberlein, E., Stafford, S., Periselneris, J., Aldridge, C., Vollmer, W., Picard, C., Casanova, J. L., Noursadeghi, M. & Brown, J. (2014). TLR-mediated inflammatory responses to *Streptococcus pneumonia*e are highly dependent on surface expression of bacterial lipoproteins. *J Immunol* 193**,** 3736-3745.

Trappetti, C., Gualdi, L., Di Meola, L., Jain, P., Korir, C. C., Edmonds, P., Iannelli, F., Ricci, S., Pozzi, G. & Oggioni, M. R. (2011). The impact of the competence

quorum sensing system on *Streptococcus pneumoniae* biofilms varies depending on the experimental model. *BMC Microbiol* 11**,** 75.

Trappetti, C., McAllister, L. J., Chen, A., Wang, H., Paton, A. W., Oggioni, M. R., McDevitt, C. A. & Paton, J. C. (2017). Autoinducer 2 signaling via the phosphotransferase FruA drives galactose utilization by *Streptococcus pneumoniae*, resulting in hypervirulence. *mBio* 8. e02269-16.

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.

Uchiyama, S., Carlin, A. F., Khosravi, A., Weiman, S., Banerjee, A., Quach, D., Hightower, G., Mitchell, T. J., Doran, K. S. & Nizet, V. (2009). The surface-anchored NanA protein promotes pneumococcal brain endothelial cell invasion. *J Exp Med* 206**,** 1845-1852.

Van der Poll, T. & Opal, S. M. (2009). Pathogenesis, treatment, and prevention of pneumococcal pneumonia. *Lancet* 374**,** 1543-1556.

Vidal, J. E., Howery, K. E., Ludewick, H. P., Nava, P. & Klugman, K. P. (2013). Quorum-sensing systems LuxS/autoinducer 2 and Com regulate *Streptococcus pneumoniae* biofilms in a bioreactor with living cultures of human respiratory cells. *Infect Immun* 81**,** 1341-1353.

Von Bodman, S. B., Willey, J. M. & Diggle, S. P. (2008). Cell-cell communication in bacteria: united we stand. *J Bacteriol* 190**,** 4377-4391.

Von Gottberg, A., De Gouveia, L., Tempia, S., Quan, V., Meiring, S., Von Mollendorf, C., Madhi, S. A., Zell, E. R., Verani, J. R., O'brien, K. L., Whitney, C. G., Klugman, K. P., Cohen, C. & Investigators, G.-S. (2014). Effects of vaccination on invasive pneumococcal disease in South Africa. *N Engl J Med* 371**,** 1889-1899.

Wada, H., Noguchi, Y., Marino, M. W., Dunn, A. R. & Old, L. J. (1997). T cell functions in granulocyte/macrophage colony-stimulating factor deficient mice. *Proc Natl Acad Sci U S A* 94**,** 12557-12561.

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. *Lancet* 381**,** 1405-1416.

Walsh, R. L. & Camilli, A. (2011). *Streptococcus pneumoniae* is desiccation tolerant and infectious upon rehydration. *mBio* 2**,** e00092-00011.

Wang, L., Fu, J., Liang, Z. & Chen, J. (2017). Prevalence and serotype distribution of nasopharyngeal carriage of *Streptococcus pneumoniae* in China: a meta-analysis. *BMC Infect Dis* 17**,** 765.

Wang, Q., He, Z., Hu, Y., Jiang, Y., Ma, R., Tang, Z., Liang, J., Liu, Z. & Huang, Z. (2012). *luxS* mutant regulation: quorum sensing impairment or methylation disorder? *Sensors (Basel)* 12**,** 6155-6175.

Wang, R., Han, G., Wang, J., Song, L., Chen, G., Xu, R., Yu, M., Qian, J., Shen, B. & Li, Y. (2007). The role of STAT3 in antigen-IgG inducing regulatory CD4(+)Foxp3(+)T cells. *Cell Immunol* 246**,** 103-109.

Wartha, F., Beiter, K., Albiger, B., Fernebro, J., Zychlinsky, A., Normark, S. & Henriques-Normark, B. (2007). Capsule and D-alanylated lipoteichoic acids protect *Streptococcus pneumoniae* against neutrophil extracellular traps. *Cell Microbiol* 9**,** 1162-1171.

Watnick, P. & Kolter, R. (2000). Biofilm, city of microbes. *J Bacteriol* 182**,** 2675-2679.

Watson, D. A., Musher, D. M., Jacobson, J. W. & Verhoef, J. (1993). A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. *Clin Infect Dis* 17**,** 913-924.

Weickert, M. J. & Adhya, S. (1993). Control of transcription of *gal* repressor and isorepressor genes in *Escherichia coli*. *J Bacteriol* 175**,** 251-258.

Weiser, J. N., Ferreira, D. M. & Paton, J. C. (2018). *Streptococcus pneumoniae*: transmission, colonization and invasion. *Nat Rev Microbiol* 16**,** 355-367.

Whitney, C. G., Farley, M. M., Hadler, J., Harrison, L. H., Bennett, N. M., Lynfield, R., Reingold, A., Cieslak, P. R., Pilishvili, T., Jackson, D., Facklam, R. R.,

References

Jorgensen, J. H., Schuchat, A. & Active Bacterial Core Surveillance of the Emerging Infections Program, N. (2003). Decline in invasive pneumococcal disease after the introduction of protein-polysaccharide conjugate vaccine. *N Engl J Med* 348**,** 1737-1746.

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.

Wilson, R., Cohen, J. M., Jose, R. J., De Vogel, C., Baxendale, H. & Brown, J. S. (2015). Protection against *Streptococcus pneumoniae* lung infection after nasopharyngeal colonization requires both humoral and cellular immune responses. *Mucosal Immunol* 8**,** 627-639.

Winslow, C. E. A., Broadhurst, J., Buchanan, R. E., Krumweide, C. Jr., Rogers, L. A. & Smith, G. H. (1920). The families and genera of the bacteria: final report of the committee of the Society of American Bacteriologists on characterization and classification of bacterial types. *Journal of Bacteriology* 5**,** 191 - 229.

World Health Organization. 1998. Prevention of hearing impairment from chronic otitis media. Report of a WHO/CIBA foundation workshop. London.

World Health Organization. 2017. *WHO Publishes List of Bacteria for Which New Antibiotics are Urgently Needed* [Online]. Available: https://www.who.int/news-room/detail/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed [Accessed February 22, 2022].

Wright, A. E., Morgan, W. P., Colebrook, L. & Dodgson, R. W. (1914). Observations on prophylactic inoculation against pneumococcus infections, and on the results which have been achieved by it. *The Lancet* 183**,** 87 - 95.

Yadav, M. K., Vidal, J. E., Go, Y. Y., Kim, S. H., Chae, S. W. & Song, J. J. (2018). The LuxS/AI-2 quorum-sensing system of *Streptococcus pneumoniae* is required to cause disease, and to regulate virulence- and metabolism-related genes in a rat model of middle ear infection. *Front Cell Infect Microbiol* 8**,** 138.

Yoshimura, A., Lien, E., Ingalls, R. R., Tuomanen, E., Dziarski, R. & Golenbock, D. (1999). Cutting edge: recognition of Gram-positive bacterial cell wall components by the innate immune system occurs via Toll-like receptor 2. *J Immunol* 163**,** 1-5.

Zafar, M. A., Wang, Y., Hamaguchi, S. & Weiser, J. N. (2017). Host-to-host transmission of *Streptococcus pneumoniae* Is driven by its inflammatory toxin, pneumolysin. *Cell Host Microbe* 21**,** 73-83.

Zeng, L., Das, S. & Burne, R. A. (2010). Utilization of lactose and galactose by *Streptococcus mutans*: transport, toxicity, and carbon catabolite repression. *J Bacteriol* 192**,** 2434-2444.

Zhanel, G. G., Dueck, M., Hoban, D. J., Vercaigne, L. M., Embil, J. M., Gin, A. S. & Karlowsky, J. A. (2001). Review of macrolides and ketolides: focus on respiratory tract infections. *Drugs* 61**,** 443-498.

Zhanel, G. G., Wolter, K. D., Calciu, C., Hogan, P., Low, D. E., Weiss, K. & Karlowsky, J. A. (2014). Clinical cure rates in subjects treated with azithromycin for community-acquired respiratory tract infections caused by azithromycin-susceptible or azithromycin-resistant *Streptococcus pneumoniae:* analysis of Phase 3 clinical trial data. *J Antimicrob Chemother* 69**,** 2835-2840.

Zhang, J. R., Mostov, K. E., Lamm, M. E., Nanno, M., Shimida, S., Ohwaki, M. & Tuomanen, E. (2000). The polymeric immunoglobulin receptor translocates pneumococci across human nasopharyngeal epithelial cells. *Cell* 102**,** 827-837.

Zhu, L. & Lau, G. W. (2011). Inhibition of competence development, horizontal gene transfer and virulence in *Streptococcus pneumoniae* by a modified competence stimulating peptide. *PLoS Pathog* 7**,** e1002241.

Appendices

Appendix A: Non-Significant Metabolomics Results

Assessing changes in the metabolome in response to tagatose-6-phosphate pathway mutation

D39 and D39Δ*lacAB* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. Cells were then infused in cold PBS and pelleted prior to snap freezing. Pellets were then shipped to Metabolomics Australia for analysis.

D-Ribose 5-P

D-Ribulose 5-P

Ribitol

Guanosine monophosphate

myo-Inositol
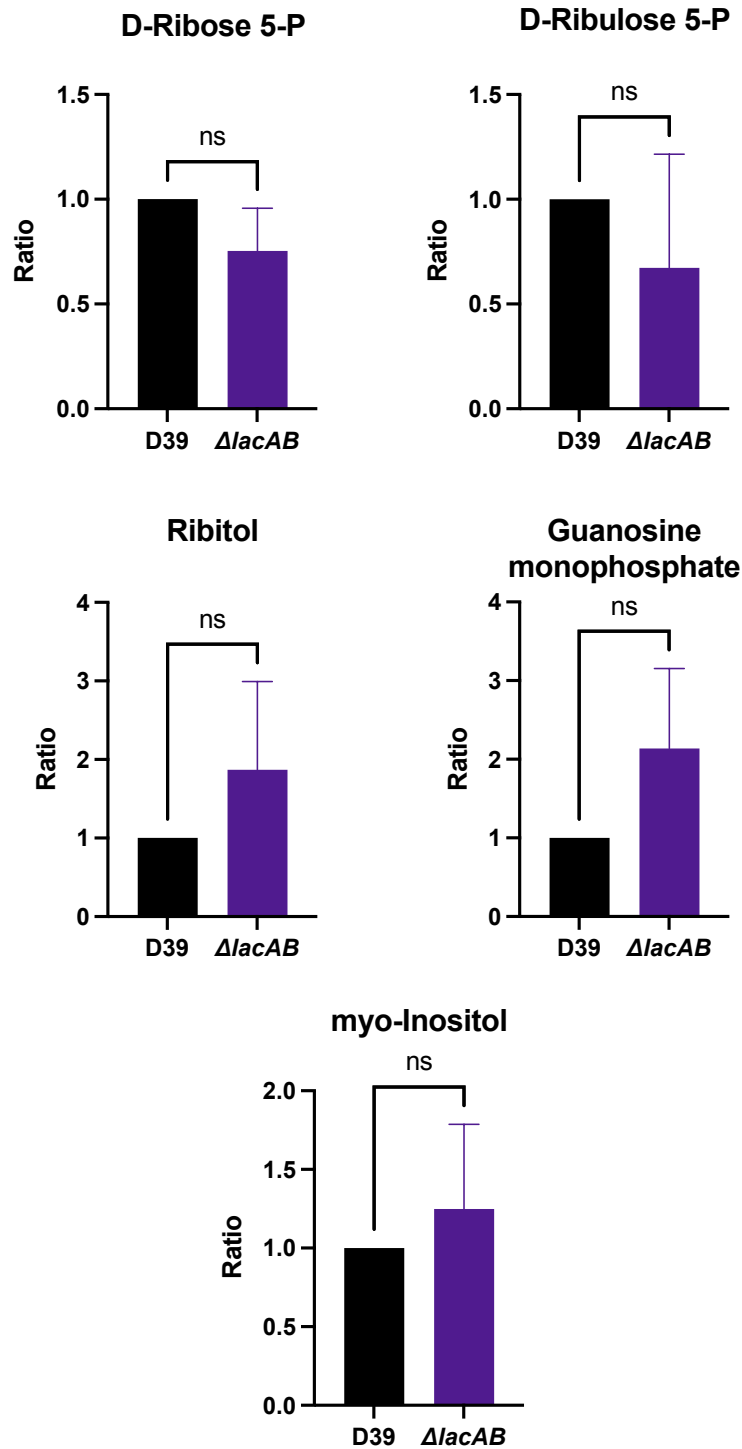
Assessing changes in the metabolome in response to tagatose-6-phosphate pathway mutation

D39 and D39Δ*lacAB* were cultured overnight on blood agar, washed, and resuspended to a final OD_{600} of 0.25 in CDM + Gal and incubated for 30 min. Cells were then infused in cold PBS and pelleted prior to snap freezing. Pellets were then shipped to Metabolomics Australia for analysis.

Appendix B: RNA Sequencing Bioanalyser Results

Electrophoresis File Run Summary
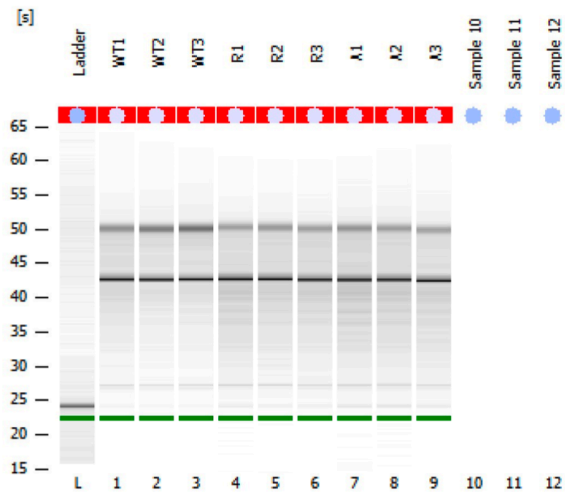
Instrument Information:

Instrument Name:	DE54704589	Firmware:	C.01.069
Serial#:	DE54704589	Type:	G2938C

Assay Information:

Assay Origin Path: C:\Program Files\Agilent\2100 bioanalyzer\2100 expert\assays\RNA\Eukaryote Total RNA Nano Series II.xsy

Assay Class: Eukaryote Total RNA Nano

Version: 2.6

Assay Comments: Total RNA Analysis ng sensitivity (Eukaryote)

© Copyright 2003 - 2009 Agilent Technologies, Inc.

Chip Information:

Chip Lot #:

Reagent Kit Lot #:

Chip Comments:
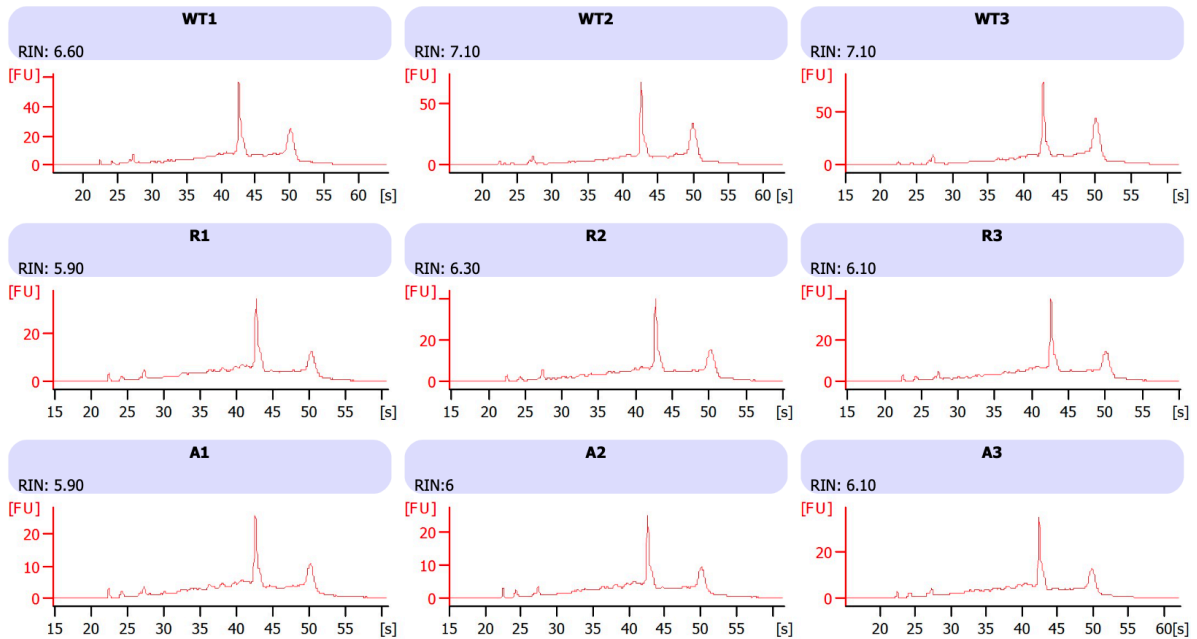
Bioanalyser results of combined *Mus musculus* and *Streptococcus pneumoniae* RNA

RNA sent for sequencing was subjected to Bioanalyser Analysis before proceeding with sequencing on the Illumina platform. A summary of these results are shown above.

Appendix C: Differentially Expressed Murine Genes

Table C1 – Differentially expressed *M. musculus* genes: D39 vs. Δ*galR*

Gene	Fold Change	Adjusted p-value
Ubd	1.2471	0.0204
Alk	1.0523	0.0204
Clec12a	1.0023	0.0204
Ms4a4a	1.0760	0.0213
Ccl12	1.3059	0.0268
S100a8	1.5690	0.0289
Vnn1	1.1186	0.0289
Gm15056	1.2816	0.0299
Magt1	1.3692	0.0312
Cxcl9	1.3420	0.0312
Ccl20	2.0133	0.0352
Gphn	1.0687	0.0358
Gm34790	1.2393	0.0358
Cxcl13	1.1564	0.0358
Camk1d	1.0122	0.0358
Kcnip1	1.6763	0.0361
S100a9	1.1309	0.0361
Cd69	1.1455	0.0382
Gm15564	1.5233	0.0383
Gnb1l	3.4443	0.0389
Gm6377	1.4625	0.0389
Gm30339	1.0313	0.0389
LOC115489778	2.5595	0.0446

Table C2 - Differentially expressed *M. musculus* genes: D39 vs. Δ*galR*

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Nr1d1	-2.4536	0.0204	Esr2	-1.4334	0.0268
Atp1b2	-1.6239	0.0204	Acta2	-1.3121	0.0268
Ltbp4	-1.5810	0.0204	Des	-1.1937	0.0268
Tnxb	-1.5495	0.0204	9330158H04Rik	-1.1753	0.0268
Stac	-1.4462	0.0204	Ntf3	-1.1748	0.0268
Synpo2l	-1.4214	0.0204	Sftpc	-1.1739	0.0268
Speg	-1.3589	0.0204	Hlf	-1.1551	0.0268
Gli2	-1.3442	0.0204	Mamdc2	-1.1385	0.0268
Lgr6	-1.3239	0.0204	Per3	-1.1254	0.0268
Kcnk3	-1.2983	0.0204	Smtn	-1.0961	0.0268
Col27a1	-1.2683	0.0204	Adcy5	-1.3246	0.0288
Zfp469	-1.2631	0.0204	Stum	-1.5965	0.0289
Ptgis	-1.1994	0.0204	Map2k6	-1.2302	0.0289
Jph2	-1.1630	0.0204	Ptpru	-1.0046	0.0302
Igsf9b	-1.1226	0.0204	Eln	-1.2655	0.0317
Ptprs	-1.1197	0.0204	Fbln1	-1.3610	0.0327
Gstm1	-1.1148	0.0204	Dnm1	-1.2933	0.0327
Cspg4	-1.0959	0.0204	Ldb3	-1.1924	0.0327
Slc45a4	-1.0390	0.0204	Wscd2	-1.1650	0.0327
Ntrk3	-1.2615	0.0228	Ptp4a3	-1.0558	0.0327
Emid1	-1.1920	0.0228	Fgfr4	-1.4495	0.0344
Mtss2	-1.1352	0.0228	Meg3	-1.0551	0.0357
Zfyve28	-1.0404	0.0228	Adamtsl5	-1.0466	0.0357
Notch3	-1.1024	0.0229	Col16a1	-1.4758	0.0358
Sod3	-1.2798	0.0250	Kcnq4	-1.0625	0.0358
Megf6	-1.6445	0.0253	Grin2c	-2.8305	0.0358
Dmpk	-1.2247	0.0253	Grip2	-1.8704	0.0358
Slc2a4	-2.0522	0.0268	Mrgprf	-1.4594	0.0358

Table C2 (Contd.)

Gene	Log FC	Adjusted p-value	Gene	Log FC	Adjusted p-value
Asb2	-1.3219	0.0358	Inmt	-1.2252	0.0395
Tafa5	-1.3011	0.0358	Pgm5	-1.0191	0.0400
Aoc3	-1.2175	0.0358	Ephx1	-1.0175	0.0424
Gpc6	-1.1616	0.0358	Hmcn2	-1.5434	0.0427
Heyl	-1.0884	0.0358	Adcyap1r1	-1.1558	0.0427
Smoc1	-1.0369	0.0358	Nptx1	-1.7880	0.0427
2610027K06Rik	-1.1903	0.0361	Col14a1	-1.0957	0.0438
Podn	-1.1326	0.0361	Ncam1	-1.1363	0.0444
Actg2	-1.0563	0.0361	Lgi4	-1.1077	0.0444
Cnn1	-1.0486	0.0361	Cand2	-1.0438	0.0450
Clip3	-1.1684	0.0370	Scara3	-1.0759	0.0465
Smad9	-1.2836	0.0373	Hr	-1.4104	0.0475
Slc8a2	-1.2081	0.0379	Wdr83os	-1.8663	0.0479
D430041D05Rik	-1.5140	0.0382	Cd248	-1.4650	0.0479
Chst3	-1.3956	0.0382	Ppp1r14a	-1.1992	0.0496
Aqp1	-1.0540	0.0382	Mn1	-1.4603	0.0497
Mhrt	-1.9193	0.0389	Emc10	-1.2532	0.0497
Olfml2a	-1.6278	0.0389	Acacb	-1.1611	0.0497
Slc4a3	-1.4607	0.0389	Fbln2	-1.5990	0.0500
Dgkb	-1.1730	0.0389	Grin2a	-1.1879	0.0500

Table C3 - Differentially expressed *M. musculus* genes: D39 vs. GalR$_{AAA}$

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Hexb	5.9386	0.0013	Slc6a14	1.4700	0.0034
Lars2	5.6611	0.0013	Hsp90aa1	1.2585	0.0034
Gm36876	5.6535	0.0013	Gm46595	4.2743	0.0035
Jarid2	4.4706	0.0013	Epcam	1.1406	0.0035
Arhgap21	3.0548	0.0013	Peak1	1.6814	0.0036
Hsp90b1	1.2180	0.0013	Zfp791	1.1511	0.0036
Ubd	1.5795	0.0023	LOC115488073	3.4751	0.0036
Gphn	1.6425	0.0025	LOC115488074	3.4541	0.0036
Sspn	6.1008	0.0025	LOC115488065	3.3337	0.0036
Gm46178	4.4274	0.0027	Ssb	1.2144	0.0036
Cntnap5b	1.9865	0.0027	Qrfprl	1.2442	0.0036
Gm15564	1.8440	0.0027	LOC115488075	3.5562	0.0037
Dnah7b	1.2846	0.0027	Pof1b	1.1514	0.0037
Gnb1l	5.0678	0.0027	LOC115489767	3.4369	0.0038
Il31ra	5.6407	0.0028	Adad1	1.1896	0.0038
Il1rapl1	1.6723	0.0028	Defb41	1.1942	0.0038
Gm30821	6.3549	0.0029	Plet1	1.4442	0.0039
Camk1d	1.3163	0.0029	Gm15056	1.1888	0.0039
Vmn2r53	1.1438	0.0030	Atf2	1.1394	0.0039
Kcnip1	2.0681	0.0032	Rgs1	1.3729	0.0042
LOC105245440	5.9545	0.0033	Zfp990	1.1468	0.0043
D130009I18Rik	5.5965	0.0033	LOC115489778	2.9756	0.0044
4933404O12Rik	5.2941	0.0033	LOC115488077	3.0969	0.0047
Ccl20	2.3448	0.0033	Vnn1	1.1351	0.0048
Tceal9	1.4908	0.0033	Hmgn5	1.5488	0.0066
Mal2	1.1467	0.0033	Gm6377	1.3072	0.0067
Akr1c14	1.3163	0.0033	Tmem56	1.0791	0.0069
LOC115488060	3.4070	0.0034	Ttpa	1.1007	0.0072

Table C3 (Contd.)

Gene	Log FC	Adjusted p-value	Gene	Log FC	Adjusted p-value
mt-Rnr1	3.2089	0.0074	Magt1	1.1036	0.0159
Lrriq1	1.1322	0.0079	Gm34259	1.5663	0.0179
Mns1	1.1231	0.0081	A530046M15Rik	1.2330	0.0183
Tfap2d	1.0450	0.0082	LOC115486438	1.2850	0.0186
Mis18bp1	1.0342	0.0085	Ccdc113	1.0126	0.0191
4931440J10Rik	1.1508	0.0085	Spink5	1.2523	0.0193
Cd69	1.0006	0.0086	Enkur	1.0261	0.0214
Sfr1	1.2094	0.0087	Hypk	1.0324	0.0216
Rnu2-10	1.6531	0.0089	Gm41810	1.5446	0.0227
Cd300ld	1.0731	0.0089	Slco1a5	1.2142	0.0227
Zc3h7a	1.5488	0.0090	Ccdc160	1.0307	0.0236
Zfp820	1.3849	0.0095	mt-Rnr2	1.9969	0.0254
Mageb16	1.0273	0.0114	LOC115487917	1.2770	0.0259
Gm19935	1.0592	0.0119	S100a8	1.0337	0.0262
Ccl12	1.1051	0.0121	Zfp804a	1.3102	0.0323
Cxcl5	1.2404	0.0125	Syt10	1.0960	0.0334
Ctla2a	1.3090	0.0126	Gm41042	1.0579	0.0376
D130079A08Rik	1.6308	0.0126	Col6a6	1.0286	0.0408
LOC115487918	1.7882	0.0129	Serpinb3a	1.3010	0.0420
LOC115487919	2.1634	0.0155	Wfdc17	1.0036	0.0482
Serpinb2	1.9968	0.0158			

Table C4 - Differentially expressed *M. musculus* genes: D39 vs. GalR$_{AAA}$

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Nr1d1	-3.1378	0.0013	Zmiz2	-1.3866	0.0031
Sod3	-2.0027	0.0013	Mcf2l	-1.0697	0.0031
Ltbp4	-1.9786	0.0013	Bcar1	-1.1291	0.0031
Csrnp1	-1.0074	0.0013	Psd	-1.4074	0.0032
Gse1	-1.2225	0.0021	Ptgis	-1.3738	0.0032
Tns2	-1.4503	0.0024	Atp1b2	-1.9549	0.0033
Tgfb1i1	-1.1230	0.0025	Des	-1.6226	0.0033
Cspg4	-1.5081	0.0025	Ptprs	-1.3011	0.0033
Jak3	-1.5143	0.0025	Efhd2	-1.1090	0.0033
Shank3	-1.1071	0.0025	Fblim1	-1.0600	0.0033
Sdc3	-1.0845	0.0025	Ripor1	-1.0495	0.0033
Gm42002	-1.1425	0.0026	Fhl3	-1.3895	0.0033
Socs1	-1.8030	0.0028	Prrc2a	-1.3314	0.0033
Rnasek	-1.6582	0.0028	Nr4a1	-1.2723	0.0033
Gli2	-1.5309	0.0028	Pdlim7	-1.1554	0.0033
Nbl1	-1.3411	0.0028	Glis2	-1.1123	0.0033
Gipc3	-1.6982	0.0029	Crtc2	-1.0705	0.0033
Zfp469	-1.6353	0.0029	Slc2a4	-2.4536	0.0033
Sphk2	-1.5780	0.0029	Grip2	-2.1291	0.0033
Col27a1	-1.4303	0.0029	Xirp1	-1.9247	0.0033
Hspg2	-1.3186	0.0029	Mrgprf	-1.8748	0.0033
Mfap4	-1.1244	0.0029	Speg	-1.7707	0.0033
Hspb7	-2.0079	0.0029	Dmpk	-1.7502	0.0033
Ldb3	-1.6930	0.0030	Snai1	-1.6972	0.0033
Midn	-1.4351	0.0030	Slc43a1	-1.6390	0.0033
Hspb1	-1.4259	0.0030	Btbd2	-1.5356	0.0033
Syvn1	-1.1676	0.0030	Acta2	-1.5113	0.0033
Atp2a1	-2.1703	0.0030	Adamtsl5	-1.5054	0.0033

Table C4 (Contd.)

Gene	Log FC	Adjusted p-value	Gene	Log FC	Adjusted p-value
Kcnk3	-1.4803	0.0033	Sox7	-1.0481	0.0033
Hlf	-1.4454	0.0033	Ndrg2	-1.0445	0.0033
Ncln	-1.4411	0.0033	Trp53	-1.0389	0.0033
Slc25a39	-1.4406	0.0033	Tspan4	-1.0262	0.0033
Tmem259	-1.4272	0.0033	Leng8	-1.0150	0.0033
Jph2	-1.3541	0.0033	Pip5k1c	-1.0042	0.0033
Acacb	-1.3528	0.0033	Cand2	-1.3281	0.0034
Per1	-1.3413	0.0033	Hdac5	-1.2361	0.0034
Kcnq4	-1.3120	0.0033	Pde4a	-1.0865	0.0034
Dvl3	-1.2935	0.0033	Mlip	-1.3625	0.0035
Gpr153	-1.2852	0.0033	Igsf9b	-1.1832	0.0035
Ftl1	-1.2672	0.0033	Egr1	-1.1476	0.0035
Lgr6	-1.2667	0.0033	Ubald1	-1.5012	0.0036
Armc5	-1.2611	0.0033	Ucp2	-1.4687	0.0036
Steap3	-1.2578	0.0033	Carmn	-1.1705	0.0036
Pear1	-1.2295	0.0033	Asb2	-1.6402	0.0036
Id3	-1.2246	0.0033	Tnxb	-1.5880	0.0036
Actg2	-1.2170	0.0033	Cx3cl1	-1.5064	0.0036
Cdc42ep1	-1.2016	0.0033	Cyp27a1	-1.2027	0.0036
Ttyh3	-1.1758	0.0033	Notch3	-1.1699	0.0036
Loxl1	-1.1623	0.0033	Coro6	-2.2430	0.0036
Selenow	-1.1365	0.0033	Flnc	-1.8921	0.0036
Pim3	-1.1187	0.0033	Actc1	-1.6858	0.0036
Wbp2	-1.1115	0.0033	Ccdc97	-1.2348	0.0036
Zyx	-1.0680	0.0033	Ntrk3	-1.1478	0.0036
Mlf2	-1.0600	0.0033	Ipo4	-1.0278	0.0036
Nlgn2	-1.0561	0.0033	Tafa5	-1.4406	0.0036
Col6a2	-1.0538	0.0033	Pnpla2	-1.4134	0.0036

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Plbd2	-1.3948	0.0036	Adcy5	-1.5231	0.0039
Zfp414	-1.3838	0.0036	Osr1	-1.3862	0.0039
Adamtsl2	-1.2053	0.0036	Hmcn2	-1.8662	0.0039
Pom121	-1.1730	0.0036	Srl	-1.3373	0.0039
Agpat1	-1.0713	0.0036	Coq8a	-1.6855	0.0039
Mknk2	-1.1861	0.0036	Zfp703	-1.6032	0.0039
Mrvi1	-1.0506	0.0036	Kcnd3	-1.3206	0.0039
Slc25a23	-1.3718	0.0037	Stac	-1.1821	0.0039
Il3ra	-1.5587	0.0037	Adgrl1	-1.1089	0.0039
Gm29943	-1.8660	0.0037	Per3	-1.1373	0.0040
Zcchc3	-1.6034	0.0037	Sgca	-2.1788	0.0041
Smtn	-1.4792	0.0037	Rexo1	-1.4761	0.0041
Rgma	-1.2177	0.0037	Lmf2	-1.3227	0.0041
Plekhg2	-1.2085	0.0037	Slc25a22	-1.3164	0.0041
Adamtsl4	-1.1402	0.0037	Zfp651	-1.2339	0.0041
Eef2k	-1.1271	0.0037	Grik3	-1.2269	0.0041
Tcf3	-1.1830	0.0038	Gga1	-1.1843	0.0041
Nacad	-1.4518	0.0038	Myo18b	-1.5494	0.0041
Emc10	-1.9698	0.0038	Synpo2l	-2.1457	0.0041
Fbln1	-1.6131	0.0038	Mbd6	-1.5141	0.0041
Lrrn2	-1.3165	0.0038	Mylk4	-1.3957	0.0041
Dtx3	-1.2569	0.0038	Klf15	-1.3038	0.0041
Ntf3	-1.2339	0.0038	Rbm38	-1.3734	0.0041
Plcd1	-1.2041	0.0038	Tpm2	-1.2883	0.0041
Bcl9l	-1.1746	0.0038	Cic	-1.1192	0.0042
Plekha4	-1.0126	0.0038	Polr2a	-1.0489	0.0042
Igsf8	-1.1509	0.0039	Notch1	-1.0470	0.0042
Dock6	-1.0537	0.0039	Fbln2	-1.7475	0.0042

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Cntnap1	-1.1771	0.0042	Scube3	-1.0613	0.0045
Fbxo31	-1.1888	0.0042	Pde6a	-1.3574	0.0045
Stum	-1.6929	0.0043	Ryr1	-1.8752	0.0045
Copz2	-1.2431	0.0043	Bgn	-1.1862	0.0045
Rnf126	-1.3285	0.0043	Vwa1	-1.8109	0.0046
Dok7	-1.7320	0.0043	Hr	-1.4913	0.0046
Cygb	-1.3599	0.0043	Kctd15	-1.3198	0.0046
Mir1a-1hg	-2.4703	0.0044	Heyl	-1.2083	0.0046
Mn1	-1.6842	0.0044	Syngap1	-1.0574	0.0046
Zfp579	-1.4840	0.0044	Rbck1	-1.1984	0.0046
Pfkm	-1.3960	0.0044	Pdk2	-1.6879	0.0046
Rfng	-1.3121	0.0044	Tob2	-1.1361	0.0046
Ptp4a3	-1.1943	0.0044	Phf1	-1.0988	0.0046
Pde4dip	-1.0727	0.0044	Ephb2	-1.0243	0.0046
Usp21	-1.0637	0.0044	Nlgn3	-1.3438	0.0046
Npdc1	-1.0580	0.0044	Dact3	-1.5532	0.0046
Rasl11b	-1.0535	0.0044	Nfic	-1.0464	0.0046
Pcnx3	-1.0445	0.0044	Mlycd	-1.0558	0.0047
Banf1	-1.1219	0.0044	BC034090	-1.1529	0.0047
Snta1	-1.0895	0.0044	Mfsd10	-1.3906	0.0048
Pfn1	-1.0671	0.0044	Cd3e	-1.2660	0.0048
Gm16576	-1.0485	0.0044	Grina	-1.1602	0.0048
Numbl	-1.0434	0.0044	Acbd4	-1.0027	0.0048
Obscn	-1.6159	0.0044	Tle5	-1.1804	0.0048
Abat	-1.0909	0.0044	Pgm5	-1.0722	0.0048
Robo4	-1.0792	0.0044	Lrrc32	-1.0879	0.0048
Sf3a2	-1.2154	0.0044	Tmem182	-1.2862	0.0048
Abca2	-1.2530	0.0044	Wdr83os	-2.4740	0.0049

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
G6pc3	-1.6669	0.0049	Pdlim3	-1.1237	0.0060
Nr1h2	-1.0388	0.0049	Tubb4a	-1.4985	0.0060
Zdhhc8	-1.0979	0.0049	Clip3	-1.4811	0.0060
Cpsf4l	-1.0853	0.0049	Lncppara	-1.1445	0.0061
Rara	-1.0092	0.0049	AI837181	-1.5986	0.0061
Emid1	-1.2463	0.0050	Ky	-1.4296	0.0061
Atn1	-1.2364	0.0050	Bcat1	-1.0918	0.0061
Rtn4rl2	-1.0329	0.0050	9330158H04Rik	-1.0912	0.0061
Eln	-1.3702	0.0050	Tecr	-1.3014	0.0061
Gm42339	-1.0237	0.0050	H2ac19	-1.7871	0.0062
Phactr3	-1.8051	0.0051	Kcnj12	-1.3485	0.0062
Wnk2	-1.5973	0.0051	Pdzd7	-1.1829	0.0062
Ap5b1	-1.4733	0.0051	Miga2	-1.0267	0.0062
Josd2	-1.0585	0.0051	N4bp3	-1.3928	0.0062
Ncdn	-1.2188	0.0051	Osbp2	-1.3758	0.0062
Sgsh	-1.3307	0.0051	Pkd1	-1.0243	0.0062
Stc2	-1.1614	0.0053	Gas6	-1.3056	0.0063
Popdc2	-1.1385	0.0054	Thra	-1.0229	0.0063
Cwc22	-3.4081	0.0054	Cdr2l	-1.0019	0.0063
Aoc3	-1.5406	0.0054	Asb18	-2.0928	0.0063
Irf2bp1	-1.0864	0.0054	Cd248	-2.0067	0.0063
Gas2l1	-1.1442	0.0055	Lamb2	-1.0664	0.0063
Lrg1	-1.5849	0.0055	Gm39606	-1.4985	0.0064
Erf	-1.0952	0.0056	Mgst3	-1.3859	0.0065
Gpsm3	-1.3667	0.0057	Alpk3	-1.0425	0.0065
Cacnb3	-1.0105	0.0058	Col7a1	-1.0984	0.0065
Cdh13	-1.1046	0.0059	Cilp	-1.8500	0.0065
Sipa1	-1.1240	0.0059	Med29	-1.2262	0.0065

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Slc35b2	-1.1169	0.0065	Slc7a5	-1.0651	0.0070
Fasn	-1.0673	0.0065	Cacna1s	-1.5620	0.0070
Htra3	-1.8334	0.0065	Eno2	-1.2823	0.0072
Col9a3	-1.0245	0.0065	Fcgrt	-1.0542	0.0072
Ebf2	-1.4682	0.0065	Repin1	-1.3015	0.0072
H4c4	-1.1865	0.0065	Prrt2	-1.7811	0.0072
Xirp2	-1.1426	0.0065	Pcbp4	-1.2386	0.0072
Sbk1	-1.0247	0.0065	Mbd3	-1.1676	0.0072
H19	-1.6469	0.0066	Mybpc2	-1.1298	0.0072
Mt2	-1.0561	0.0066	Tead3	-1.0278	0.0072
Mmp17	-1.2841	0.0066	Mvk	-1.0398	0.0072
Ints5	-1.2201	0.0066	Cdk18	-1.1737	0.0072
Six5	-1.4460	0.0066	Csf2ra	-1.1695	0.0072
Plekhh3	-1.3650	0.0067	Ace	-1.1493	0.0072
Nfatc4	-1.1911	0.0068	Gpr182	-1.1312	0.0072
Mxra7	-1.0929	0.0068	Nr2f6	-1.1286	0.0072
Slc4a3	-1.6707	0.0068	Map3k11	-1.0044	0.0072
Meg3	-1.1772	0.0069	Ncs1	-1.3533	0.0073
Igfbp6	-1.6471	0.0069	Shmt2	-1.1238	0.0073
Sox18	-1.5533	0.0069	Fastk	-1.0502	0.0073
Myh2	-1.4775	0.0069	Crip2	-1.5116	0.0074
Smyd1	-1.4069	0.0069	Vamp2	-1.3280	0.0074
Slc29a1	-1.3535	0.0069	Atcayos	-2.0184	0.0074
Adamts7	-1.2199	0.0069	Tpra1	-1.2966	0.0074
Gstm1	-1.0932	0.0069	Emp3	-1.0645	0.0074
Fam110a	-1.0406	0.0069	Col5a3	-1.3696	0.0074
Pkdcc	-1.0262	0.0069	Bves	-1.1044	0.0074
Scara3	-1.2246	0.0069	Espn	-1.1053	0.0074

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Pth1r	-1.1565	0.0076	Prr12	-1.0302	0.0082
Tmem38a	-1.3356	0.0077	Mast3	-1.1222	0.0082
Lamc3	-1.0443	0.0077	Gata5	-1.3153	0.0082
Mdga1	-1.2323	0.0077	Gm10914	-1.1875	0.0082
Slc39a3	-1.0142	0.0077	Gsn	-1.0143	0.0082
Nr1d2	-1.0121	0.0077	Ryr3	-1.0053	0.0082
Scn4b	-1.7900	0.0078	Flt3l	-1.1959	0.0082
Fcgbp	-1.5323	0.0079	Gm33449	-2.1390	0.0083
Ap1m1	-1.0224	0.0079	Podn	-1.1820	0.0083
Gpc1	-1.1515	0.0080	Npepl1	-1.1259	0.0084
Asmt	-1.6560	0.0080	Ticam1	-1.0784	0.0084
Lrp3	-1.3865	0.0080	Vegfb	-1.2143	0.0084
Wscd2	-1.1717	0.0080	Ckb	-2.0230	0.0085
Pdlim4	-1.2907	0.0080	Myot	-1.9415	0.0085
Trim47	-1.2253	0.0080	Slc24a2	-1.1996	0.0085
Kdm6b	-1.1380	0.0080	Tbx15	-1.4709	0.0085
Adrm1	-1.0956	0.0080	Inka1	-1.3204	0.0085
Myg1	-1.0436	0.0080	Slc22a17	-1.1677	0.0085
Mfsd12	-1.0278	0.0080	Cbx7	-1.0044	0.0086
Ltbp3	-1.0187	0.0081	Tysnd1	-1.0454	0.0086
Aatk	-1.1691	0.0081	Eef1a2	-1.2618	0.0086
Tagln	-1.3140	0.0081	Efhd1	-1.7534	0.0086
Sh3glb2	-1.2248	0.0081	Adgrb1	-1.2787	0.0086
Rnf26	-1.1907	0.0081	Mfsd5	-1.2992	0.0086
Itga7	-1.0009	0.0081	Pygm	-1.5220	0.0086
Dqx1	-1.2930	0.0081	Gm30974	-1.7032	0.0086
Hamp	-2.2191	0.0082	Carns1	-1.4431	0.0087
Actn2	-1.2066	0.0082	Mapk8ip1	-1.3829	0.0087

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Pias3	-1.0630	0.0087	Hs3st5	-1.5337	0.0097
Cacna1h	-1.7482	0.0087	Cyc1	-1.0202	0.0097
D630003M21Rik	-1.5901	0.0087	Tldc2	-1.0892	0.0097
Celsr2	-1.0174	0.0087	Mlxipl	-1.6294	0.0097
Aldoart2	-1.3143	0.0087	C1qtnf6	-1.3200	0.0099
3425401B19Rik	-1.1315	0.0087	Dll4	-1.0847	0.0099
Slc6a17	-1.1172	0.0088	Cox6a2	-1.9978	0.0099
Dnm1	-1.3323	0.0088	Nrbp2	-1.1132	0.0099
Plscr3	-1.1425	0.0088	Scube1	-1.0937	0.0099
Eno3	-1.3912	0.0089	Klhl30	-1.5083	0.0099
Islr	-1.4083	0.0089	AY036118	-1.3136	0.0100
Traf4	-1.1618	0.0089	Gaa	-1.1002	0.0100
Myl9	-1.0046	0.0089	A330074K22Rik	-1.3246	0.0100
Zfpl1	-1.0858	0.0089	Acta1	-1.9157	0.0101
Ndrg4	-1.6491	0.0090	Mypn	-1.3453	0.0102
Kctd17	-1.3636	0.0090	Cdh22	-1.0242	0.0102
Gm36583	-1.1749	0.0090	Hspb6	-1.7773	0.0102
Chac1	-1.3158	0.0091	Cnn1	-1.1745	0.0102
Dbp	-2.1919	0.0091	Mb	-1.4472	0.0102
Mt1	-1.0037	0.0091	LOC115489726	-1.1008	0.0102
Tlr9	-1.0291	0.0092	Smoc1	-1.0561	0.0102
Klhl17	-1.0654	0.0092	H4c14	-1.8228	0.0102
Tmem201	-1.1311	0.0092	Adcyap1r1	-1.0745	0.0103
Galnt16	-1.2885	0.0093	Fstl3	-1.3163	0.0104
Sptb	-1.2018	0.0093	Col16a1	-1.4284	0.0105
Isyna1	-1.0656	0.0093	Scn5a	-1.5528	0.0105
Cdh24	-1.5642	0.0094	Shank1	-1.3282	0.0105
Gm34554	-1.8497	0.0096	Fam171b	-1.5042	0.0106

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Smarcd3	-1.1931	0.0106	Rnpepl1	-1.0240	0.0119
Ablim2	-1.0487	0.0107	Actr3b	-1.0228	0.0119
Apc2	-1.0225	0.0107	Fam89b	-1.3605	0.0119
2610027K06Rik	-1.0876	0.0107	Nsun5	-1.1184	0.0119
Igf2	-1.4663	0.0107	Ptov1	-1.0836	0.0119
1700017B05Rik	-1.0208	0.0108	2310011J03Rik	-1.1658	0.0120
Pias4	-1.0028	0.0108	Plpp2	-1.0997	0.0121
Hmg20b	-1.0227	0.0108	Fkbp10	-1.0116	0.0121
Armh4	-1.3702	0.0108	Myom2	-1.0324	0.0121
Klhdc8b	-1.1584	0.0109	Mybphl	-1.3651	0.0122
Klf16	-1.0671	0.0109	Nctc1	-2.3592	0.0123
St6galnac4	-1.0420	0.0109	Nptx1	-1.5370	0.0123
Hgfac	-1.0685	0.0109	Cyp26b1	-1.3796	0.0123
Abhd17a	-1.1347	0.0109	Tmem150a	-1.1063	0.0123
Atp13a2	-1.0384	0.0111	5430431A17Rik	-1.5492	0.0123
Hdac10	-1.1373	0.0112	Trim72	-1.2789	0.0123
Gm34213	-1.0945	0.0113	Slc16a11	-1.2690	0.0123
Srpx	-1.0155	0.0113	Gm41530	-1.1659	0.0123
Rogdi	-1.4230	0.0113	Fxyd6	-1.1051	0.0123
Pdzrn4	-1.0893	0.0113	Rex1bd	-1.0757	0.0124
Gm30807	-1.5421	0.0114	Kcnk18	-2.5162	0.0124
Slc25a34	-1.7879	0.0114	Tssc4	-1.2420	0.0124
Krtcap2	-1.0522	0.0115	Prr5	-1.1050	0.0124
Mospd3	-1.3391	0.0115	Gm34189	-1.1725	0.0126
Chst3	-1.3186	0.0115	Kcna1	-1.7872	0.0126
Adssl1	-1.0439	0.0118	Rnf165	-1.1905	0.0126
Phlda3	-1.3482	0.0118	Mtss2	-1.0833	0.0126
Ckmt2	-1.2297	0.0119	Scn4a	-1.3273	0.0129

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Ccdc8	-1.1690	0.0129	Klhl31	-1.1936	0.0147
Ddah2	-1.4958	0.0129	Gsg1l	-1.1832	0.0149
Slc8a2	-1.2413	0.0130	Gm32908	-1.0687	0.0150
Inmt	-1.0404	0.0130	Ppp1r13l	-1.0369	0.0151
Ppp1r14a	-1.4909	0.0130	n-TGgcc13	-2.1718	0.0151
Atp5d	-1.1939	0.0130	Efs	-1.2246	0.0151
Gm46655	-1.2647	0.0130	Prss22	-1.6149	0.0151
Tmem8b	-1.2370	0.0130	Nkg7	-1.3572	0.0152
Smad9	-1.1257	0.0130	Myh1	-1.2528	0.0153
Ano5	-1.7587	0.0131	Vars2	-1.1112	0.0153
Ifrd2	-1.2924	0.0131	Samd1	-1.3467	0.0154
Rbm20	-1.2091	0.0132	Spib	-1.0837	0.0154
H4c2	-1.5281	0.0132	Sln	-1.3106	0.0154
Gapdh	-1.0079	0.0136	Shisal1	-1.1298	0.0155
Cyp8b1	-1.0804	0.0137	Npy1r	-1.0738	0.0155
Trarg1	-1.2933	0.0138	Erfe	-1.1613	0.0157
Fgf12	-1.0311	0.0138	Gpaa1	-1.1140	0.0157
Gm11266	-1.0433	0.0138	Cavin3	-1.1749	0.0158
Tepsin	-1.1042	0.0139	Gm46652	-1.3326	0.0159
Acd	-1.2702	0.0140	Hcfc1r1	-1.1269	0.0159
Capn3	-1.3585	0.0141	Tead2	-1.0224	0.0159
Wfdc1	-1.1805	0.0141	Adcy1	-1.2694	0.0160
Grhpr	-1.1017	0.0144	Faap100	-1.1665	0.0160
Mustn1	-1.1292	0.0144	Lynx1	-1.1350	0.0161
Dhtkd1	-1.7735	0.0145	Sema6b	-1.0972	0.0161
Gmpr	-1.5465	0.0145	Sypl2	-1.9074	0.0162
Rnf207	-1.5992	0.0146	Thap7	-1.4859	0.0162
Trim54	-1.3215	0.0146	Fgfr4	-1.0595	0.0162

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Ebf3	-1.1608	0.0162	Dnlz	-1.0395	0.0197
Cacng7	-1.2294	0.0164	Plin4	-1.3185	0.0200
Shisa4	-1.2453	0.0167	Hba-a2	-1.4901	0.0203
Atp6v0e2	-1.1462	0.0173	Prima1	-1.2548	0.0204
Tnnt3	-1.0455	0.0173	Myoz3	-1.2937	0.0205
Rtl8b	-1.0078	0.0174	Myl3	-1.7354	0.0205
Med16	-1.3464	0.0174	Lzts3	-1.1232	0.0208
Apobec2	-1.7212	0.0174	Eva1b	-1.0008	0.0209
Dcst2	-1.1758	0.0174	Gm31251	-1.3008	0.0209
Zfp358	-1.1814	0.0175	1110002E22Rik	-1.8892	0.0210
Mamstr	-1.0831	0.0175	Gm33496	-1.6130	0.0214
H2ac18	-1.5840	0.0176	Nkain1	-1.0912	0.0214
Ccdc3	-1.2428	0.0180	Opcml	-1.0676	0.0215
Acan	-1.9048	0.0182	Pmm1	-1.3749	0.0215
Clcn1	-1.2360	0.0182	Tesc	-1.1716	0.0215
Dnajc30	-1.3216	0.0182	Egf	-1.1749	0.0217
Bdh1	-1.0092	0.0183	1110018N20Rik	-1.0240	0.0217
Dusp27	-1.1775	0.0183	Pusl1	-1.2112	0.0217
Myl7	-1.1098	0.0188	Hjv	-1.0064	0.0219
Gm11725	-1.1575	0.0188	Nrap	-1.1929	0.0219
Myl4	-1.2376	0.0188	Atxn7l2	-1.1736	0.0219
Cacna1i	-1.1860	0.0192	Lgi4	-1.3233	0.0221
Mybpc1	-1.0590	0.0193	Sox10	-2.8947	0.0223
Musk	-1.1034	0.0194	Serpinf1	-1.1058	0.0224
Kcnt1	-1.2536	0.0194	Polrmt	-1.0310	0.0224
Prob1	-1.0894	0.0195	Adgrb2	-1.0687	0.0224
Rufy4	-1.0165	0.0195	C130080G10Rik	-1.0598	0.0225
Gm9899	-1.1355	0.0196	Ncam1	-1.1253	0.0226

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Sbk2	-1.6193	0.0229	Cstad	-1.0363	0.0262
Myo7b	-1.1836	0.0229	Ptprcap	-1.6167	0.0262
Abhd11	-1.0837	0.0230	Syp	-1.0596	0.0262
Gm35438	-1.0232	0.0230	Arhgap33	-1.1119	0.0268
Pianp	-1.9248	0.0231	Rpph1	-1.0471	0.0271
Fndc5	-1.1823	0.0231	Myoz1	-1.8308	0.0274
Crlf2	-1.2381	0.0233	Zfp385c	-1.0596	0.0275
Tcap	-1.4531	0.0234	Ankrd13b	-1.1375	0.0277
Rflna	-1.1657	0.0236	Pla2g4e	-1.4593	0.0277
Gm33470	-1.1440	0.0237	Rab3a	-1.1773	0.0277
Atp1a2	-2.7748	0.0238	Clec3b	-1.0019	0.0282
LOC102636795	-1.8666	0.0239	Acr	-1.0872	0.0282
Myh6	-1.2012	0.0241	Ost4	-1.2418	0.0286
Sec16b	-1.2522	0.0241	Adprhl1	-1.5805	0.0288
Csdc2	-2.3329	0.0242	Faim2	-1.2259	0.0288
Cdt1	-1.2450	0.0244	Atf5	-1.0043	0.0289
Fjx1	-1.0742	0.0246	Plin1	-2.0479	0.0296
LOC115488495	-1.0300	0.0246	Mfap5	-1.2802	0.0300
Pold1	-1.1444	0.0247	Fmod	-1.3397	0.0301
Agbl1	-1.1218	0.0248	Ldhd	-1.0136	0.0301
Srm	-1.0955	0.0249	Ndufa11	-1.1883	0.0305
Ppan	-1.1557	0.0251	Rtl1	-1.3492	0.0306
Hba-a1	-1.3784	0.0251	Mhrt	-1.8419	0.0307
Dph1	-1.0729	0.0253	Zdhhc12	-1.6895	0.0307
Nos1	-2.0681	0.0254	Abcc8	-1.0195	0.0309
Ap5s1	-1.2138	0.0256	Jrk	-1.2606	0.0312
Megf6	-1.1495	0.0260	D430041D05Rik	-1.0727	0.0314
Dhh	-1.0755	0.0261	Camk2a	-1.8670	0.0314

Table C4 (Contd.)

Gene	Log FC	Adjusted *p*-value	Gene	Log FC	Adjusted *p*-value
Gm34147	-1.0381	0.0314	Abcb9	-1.0299	0.0385
Ache	-2.3098	0.0321	Fbxl22	-1.1334	0.0387
Mylk2	-1.2619	0.0322	Htr1b	-1.0236	0.0391
gene888	-2.4552	0.0331	Npas4	-1.7374	0.0392
Myl1	-1.2167	0.0332	Myh13	-1.1414	0.0392
Cacng1	-2.2953	0.0332	Cidec	-1.4093	0.0392
Drp2	-1.0891	0.0334	Gm12963	-1.1823	0.0394
Cadm3	-1.1969	0.0342	Dohh	-1.1089	0.0396
Frem2	-1.2575	0.0345	Crhr2	-1.3804	0.0399
Map1lc3a	-1.6207	0.0345	Gm39793	-2.0682	0.0407
Hectd2os	-1.0926	0.0350	Gm35657	-1.1331	0.0413
Gm39335	-1.1458	0.0350	Gm33289	-1.0262	0.0420
Aspa	-1.0083	0.0352	Gm33320	-2.5942	0.0421
Nid2	-1.0388	0.0355	Pkmyt1	-1.0382	0.0428
Lrrc4b	-1.0676	0.0357	Car3	-1.2436	0.0431
Itgb4	-1.0050	0.0361	Slc8a3	-1.0662	0.0433
Hhipl1	-1.0919	0.0367	Myoz2	-1.0879	0.0441
Dxo	-1.0033	0.0367	Myh4	-1.2683	0.0449
Igdcc4	-1.9571	0.0368	Gm39938	-1.1046	0.0456
Atp1a4	-1.1165	0.0369	Lrrn1	-1.3483	0.0463
Fbn2	-1.2988	0.0371	Ckm	-1.7686	0.0464
Lrrc27	-1.0643	0.0376	Gm10863	-1.1774	0.0464
Klhl41	-1.2757	0.0379	gene768	-2.0514	0.0489
Olfml2a	-1.3079	0.0382	Slc25a47	-1.2754	0.0496
Vat1l	-1.0736	0.0385	Cdh19	-1.1463	0.0496

Appendix D: FACS Gating Strategies
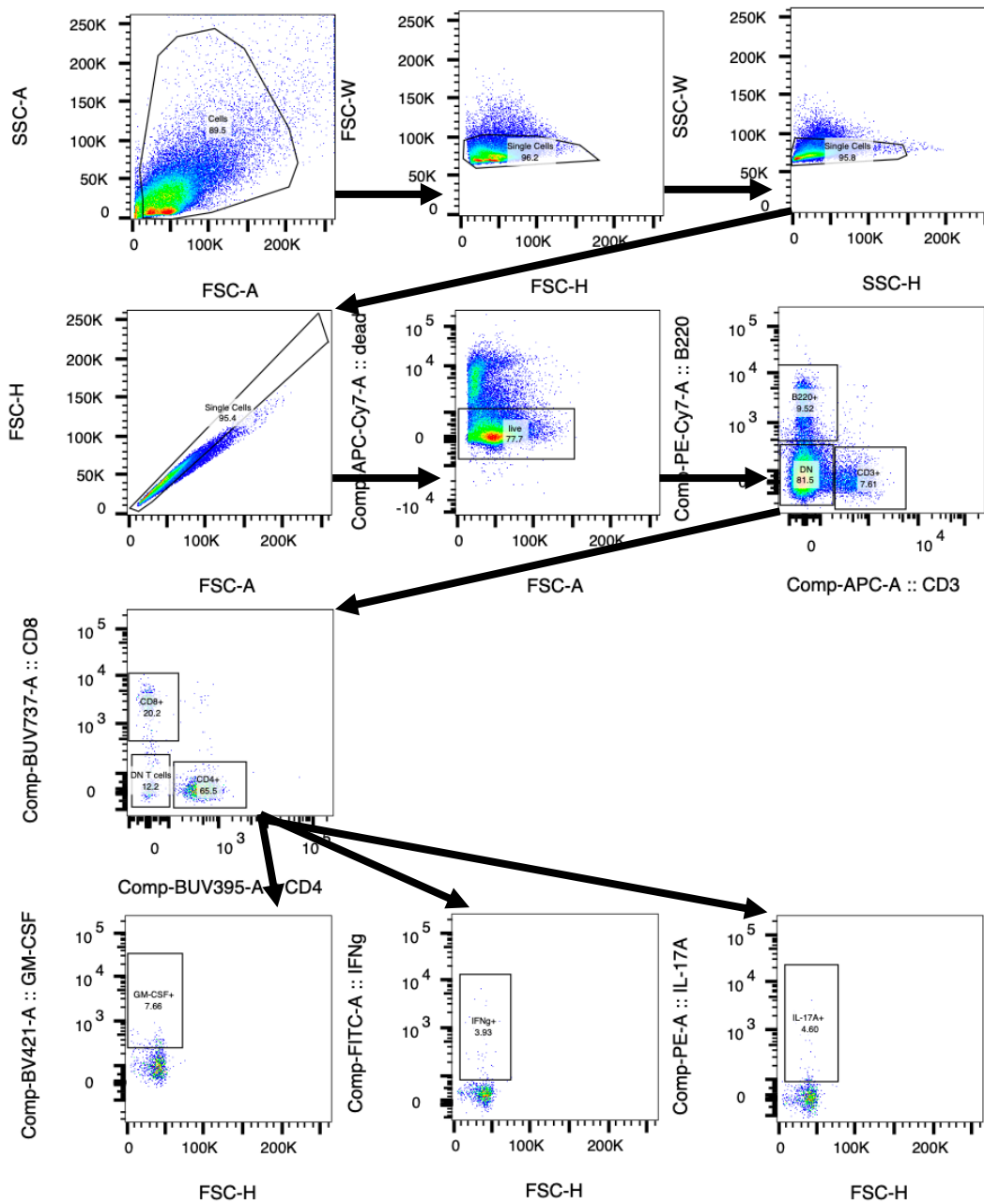
Gating strategy for identification of cytokine-expressing immune cells in pneumococcal infected murine lung tissue

Cells were first gated for size and singularity, before excluding dead cells. Remaining live cells were then gated to select for B cells, T cells and double negative (DN) cells. T cells were selected and gated to identify CD4+ and CD8+ cells. Both of these subsets were then further gated to identify cells expressing either GM-CSF, IFN-γ or IL-17A.

Statement of Authorship

Title of Paper	Site-Specific Mutations of GalR Affect Galactose Metabolism in *Streptococcus pneumoniae*
Publication Status	☑ Published ☐ Accepted for Publication ☐ Submitted for Publication ☐ Unpublished and Unsubmitted work written in manuscript style
Publication Details	McLean, K. T., Tikhomirova, A., Brazel, E. B., Legendre, S., Haasbroek, G., Minhas, V., Paton, J. C., & Trappetti, C. (2020). Site-Specific Mutations of GalR Affect Galactose Metabolism in *Streptococcus pneumoniae*. Journal of bacteriology, 203(1), e00180-20. https://doi.org/10.1128/JB.00180-20

Principal Author

Name of Principal Author (Candidate)	Kimberley Taylor McLean
Contribution to the Paper	Conceptualization, data curation, formal analysis, investigation, methodology, writing - original draft, writing - review & editing
Overall percentage (%)	70%
Certification:	This paper reports on original research I conducted during the period of my Higher Degree by Research candidature and is not subject to any obligations or contractual agreements with a third party that would constrain its inclusion in this thesis. I am the primary author of this paper.
Signature	Date 15th June 2022

Co-Author Contributions

By signing the Statement of Authorship, each author certifies that:

    i.   the candidate's stated contribution to the publication is accurate (as detailed above);

    ii.   permission is granted for the candidate in include the publication in the thesis; and

    iii.   the sum of all co-author contributions is equal to 100% less the candidate's stated contribution.

Name of Co-Author	Alexandra Tikhomirova
Contribution to the Paper	Investigation, writing - original draft, writing - review & editing
Signature	Date 15/06/2022

Name of Co-Author	Erin Brazel
Contribution to the Paper	Formal analysis, investigation, writing - original draft, writing - review & editing
Signature	Date 15/06/2022

Name of Co-Author	Salome Legendre		
Contribution to the Paper	Investigation		
Signature		Date	

Name of Co-Author	Gian Haasbroek		
Contribution to the Paper	Investigation		
Signature		Date	20/06/2022

Name of Co-Author	Vikrant Minhas		
Contribution to the Paper	Investigation, methodology, writing - review & editing		
Signature		Date	**15/06/22**

Name of Co-Author	James Paton		
Contribution to the Paper	Conceptualization, formal analysis, funding acquisition, project administration, supervision, writing - original draft, writing - review & editing		
Signature		Date	22/06/22

Name of Co-Author	Claudia Trappetti		
Contribution to the Paper	Conceptualization, formal analysis, funding acquisition, investigation, methodology, project administration, supervision, writing - original draft, writing - review & editing		
Signature		Date	15/06/22

Site-Specific Mutations of GalR Affect Galactose Metabolism in *Streptococcus pneumoniae*

Kimberley T. McLean,ᵃ Alexandra Tikhomirova,ᵃ Erin B. Brazel,ᵃ Salomé Legendre,ᵃ Gian Haasbroek,ᵃ Vikrant Minhas,ᵃ James C. Paton,ᵃ Claudia Trappettiᵃ

ᵃResearch Centre for Infectious Diseases, Department of Molecular and Biomedical Science, University of Adelaide, Adelaide, Australia

ABSTRACT *Streptococcus pneumoniae* (the pneumococcus) is a formidable human pathogen that is capable of asymptomatically colonizing the nasopharynx. Progression from colonization to invasive disease involves adaptation to distinct host niches, which vary markedly in the availability of key nutrients such as sugars. We previously reported that cell-cell signaling via the autoinducer 2 (AI-2)/LuxS quorum-sensing system boosts the capacity of *S. pneumoniae* to utilize galactose as a carbon source by upregulation of the Leloir pathway. This resulted in increased capsular polysaccharide production and a hypervirulent phenotype. We hypothesized that this effect was mediated by phosphorylation of GalR, the transcriptional activator of the Leloir pathway. GalR is known to possess three putative phosphorylation sites, S317, T319, and T323. In the present study, derivatives of *S. pneumoniae* D39 with putative phosphorylation-blocking alanine substitution mutations at each of these GalR sites (singly or in combination) were constructed. Growth assays and transcriptional analyses revealed complex phenotypes for these GalR mutants, with impacts on the regulation of both the Leloir and tagatose 6-phosphate pathways. The alanine substitution mutations significantly reduced the capacity of pneumococci to colonize the nasopharynx, middle ear, and lungs in a murine intranasal challenge model.

IMPORTANCE Pneumococcal survival in the host and capacity to transition from a commensal to a pathogenic lifestyle are closely linked to the organism's ability to utilize specific nutrients in distinct niches. Galactose is a major carbon source for pneumococci in the upper respiratory tract. We have shown that both the Leloir and tagatose 6-phosphate pathways are necessary for pneumococcal growth in galactose and demonstrated GalR-mediated interplay between the two pathways. Moreover, the three putative phosphorylation sites in the transcriptional regulator GalR play a critical role in galactose metabolism and are important for pneumococcal colonization of the nasopharynx, middle ear, and lungs.

KEYWORDS GalR, *Streptococcus pneumoniae*, carbon metabolism, galactose, pneumococcus, protein phosphorylation, virulence

Streptococcus pneumoniae is a human-adapted bacterium often carried asymptomatically in the nasopharynx. However, in a proportion of carriers, it can spread to other sites of the body and cause a wide range of illnesses, including otitis media and sinusitis, as well as severe diseases such as bacteremia, pneumonia, and meningitis (1, 2). Globally, *S. pneumoniae* infections account for 1 million to 2 million deaths every year, making it one of the world's foremost bacterial pathogens (3, 4). Colonization of the upper respiratory tract (URT) is an essential prerequisite for invasive disease. However, in this environment, carbon sources are scarce and the host actively eliminates glucose (Glc) to help maintain airway sterility (5). Galactose (Gal) is the most

abundant sugar in the URT (6), so it follows that the ability to metabolize it may be a crucial factor for pneumococcal survival within host niches.

In a previous study, we discovered a direct link between carbohydrate utilization and virulence (7). In particular, we found that the quorum-sensing signaling molecule autoinducer 2 (AI-2) promotes transition of the pneumococcus from colonizer to pathogen. Importantly, AI-2 signaling via the fructose-specific phosphoenolpyruvate phosphotransferase system (PTS) component FruA enables the bacterium to utilize Gal as a carbon source and upregulates the Leloir pathway, thereby leading to increased production of capsular polysaccharide (CPS) and a hypervirulent phenotype (7). CPS precursors are synthesized from Glc 1-phosphate, which can be produced via either the Glc 6-phosphate pathway in the presence of Glc (as occurs in the blood) or the Leloir pathway when Gal is the predominant sugar (as occurs in the URT).

S. pneumoniae possesses two pathways for galactose metabolism, the Leloir pathway and the tagatose 6-phosphate (T6P) pathway. In the T6P pathway, extracellular Gal is transported into the cell and phosphorylated through a PTS (unrelated to FruA) (8). The resultant Gal 6-phosphate is then converted into T6P by the enzyme Gal 6-phosphate isomerase (encoded by *lacAB*). The T6P kinase (encoded by *lacC*) then converts T6P into tagatose 1,6-bP. Finally, *lacD* codes for the tagatose 1,6-bP aldolase, which converts tagatose 1,6-bP into dihydroxyacetone-P and D-glyceraldehyde-3-P (9, 10). In the Leloir pathway, Gal enters the cell via a proposed ABC transporter (8). It is then phosphorylated intracellularly at the C1 position by a specific kinase (encoded by *galK*) to yield Gal 1-phosphate, which is then converted into Glc 1-phosphate by hexose 1-phosphate uridyltransferase (encoded by *galT*) and UDP-glucose epimerase (encoded by *galE*). The transcriptional regulator of this pathway is GalR, which is believed to possess three putative phosphorylation sites, S317, T319, and T323 (11). We previously proposed that phosphorylated AI-2 imported via FruA facilitates phosphorylation of GalR at these sites, thereby activating (or relieving repression of) the *gal* operon (7). In the present study, we conducted mutational analysis of the putative GalR phosphorylation sites and examined the impact on expression of Leloir and T6P pathway genes and Gal metabolism.

RESULTS

The location of the GalR putative phosphorylation sites. In the absence of structural information for GalR, we sought to examine the location of the putative phosphorylated residues (S317, T319, and T323) by generating a structural homology model. A homology model of GalR was constructed using SWISS-MODEL based on the homodimeric 2.4-Å structure (PDB: 1JFS) of the *Escherichia coli* PurR W147F mutant (35% sequence similarity, 93% sequence coverage) (12). Alignment of the GalR model (green) with the template (cyan) revealed a moderate level of variation (root mean square deviation [RMSD], 2.868 Å), with an additional loop present in the GalR model that was absent in PurR, corresponding to residues 183 to 191 (Fig. 1A). To complement these studies, we performed a conserved domain search to investigate whether any of the putative phosphorylated residues were located within regions of possible functional importance (Fig. 1B). The putative galactose binding residues (magenta) are situated at the putative dimer interface of GalR, suggesting a role in protein dimer stabilization upon sugar binding, while the N-terminal region of GalR harbors the helix-turn-helix domain (blue) responsible for the interaction with DNA. All of the putative phosphorylated residues (orange spheres) were situated in a region distinct from the residues proposed to be involved in galactose binding and DNA binding (Fig. 1B). As any functional impact of S317, T319, or T323 phosphorylation is more likely a consequence of allosteric changes rather than a direct impact on sugar or DNA binding, we investigated the contribution of each putative phosphorylated residue to GalR function. We constructed a series of GalR amino acid substitution mutants in *S. pneumoniae* D39 in which S317, T319, and T323 were replaced, either singly or in combination, with the nonphosphorylatable residue alanine (A), using the Janus cassette system (see Materials and Methods). A total of 7 substitution mutants were
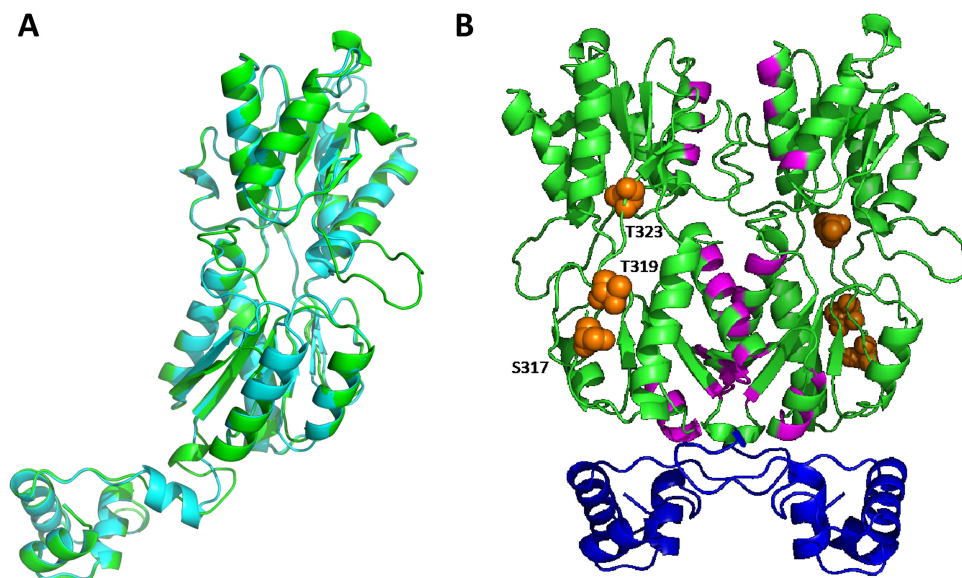
FIG 1 Structural homology model of *S. pneumoniae* GalR. (A) Cartoon representation of the protomeric homology model of *S. pneumoniae* GalR (green) based on the 2.9-Å structure of the *Escherichia coli* PurR W147F mutant (cyan; RMSD, 2.868 Å). (B) Cartoon representation of the dimeric homology model of GalR. The DNA binding helix-turn-helix domain is shown in blue, and the putative sugar binding regions are highlighted in magenta. The serine (S317) and threonine (T319 and T323) residues hypothesized to be phosphorylated are depicted as orange spheres.

generated (designated D39$_{AAA}$, D39$_{ATT}$, D39$_{SAT}$, D39$_{STA}$, D39$_{AAT}$, D39$_{ATA}$, and D39$_{SAA}$), as well as a *galR* deletion mutant (D39Δ*galR*) (see Table 1).

Impact of GalR putative phosphorylation sites on galactose metabolism. Growth in chemically defined medium with Glc as the sole carbon source (CDM + Glc) revealed that each mutant grew comparably to the D39 wild-type strain (Fig. 2A). Conversely, when grown in chemically defined medium with Gal as the sole carbon source (CDM + Gal), significant growth differences between strains became apparent (Fig. 2B). First, in comparison to D39, the D39Δ*galR* strain displayed complete abrogation of growth in the presence of Gal, indicating an essential role for *galR* in the ability to utilize galactose, as previously shown (13). The D39$_{AAA}$ strain, where the three putative phosphorylation sites are nonphosphorylatable, showed almost a complete inability to grow in galactose. D39$_{SAA}$ showed a delay in growth, with a slower generation time and a decrease in final cell density compared to D39. The remaining GalR substitution mutants, D39$_{ATT}$, D39$_{SAT}$, D39$_{STA}$, D39$_{AAT}$, and D39$_{ATA}$, had a capacity to utilize Gal similar to that of D39. This indicates that mutation of any one of the three

TABLE 1 Strains used in this study

Strain	Description[a]	Source
D39	Capsular serotype 2	NCTC 7466
D39Δ*galR*	*galR* deletion-replacement mutant, erythromycin[r]	This study
D39Δ*galK*	*galK* deletion-replacement mutant, spectinomycin[r]	This study
D39Δ*lacAB*	*lacAB* deletion-replacement mutant, spectinomycin[r]	This study
D39Δ*lacD*	*lacD* deletion-replacement mutant, erythromycin[r]	This study
D39Δ*galR*Δ*lacD*	*galR*, *lacD* deletion-replacement mutant, erythromycin[r] and spectinomycin[r]	This study
D39Δ*galR* Janus	*galR* deletion-replacement mutant utilizing Janus cassette, kanamycin, streptomycin[s]	This study
D39$_{AAA}$	*galR* S317A, T319A, T323A amino acid substitution mutant, kanamycin[s], streptomycin[s]	This study
D39$_{ATT}$	*galR* S317A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study
D39$_{SAT}$	*galR* T319A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study
D39$_{STA}$	*galR* T323A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study
D39$_{AAT}$	*galR* S317A, T319A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study
D39$_{ATA}$	*galR* S317A, T323A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study
D39$_{SAA}$	*galR* T319A, T323A amino acid substitution mutant, kanamycin[s], streptomycin[r]	This study

[a]Superscript "r" and "s" following antibiotic names indicate resistance and sensitivity, respectively.
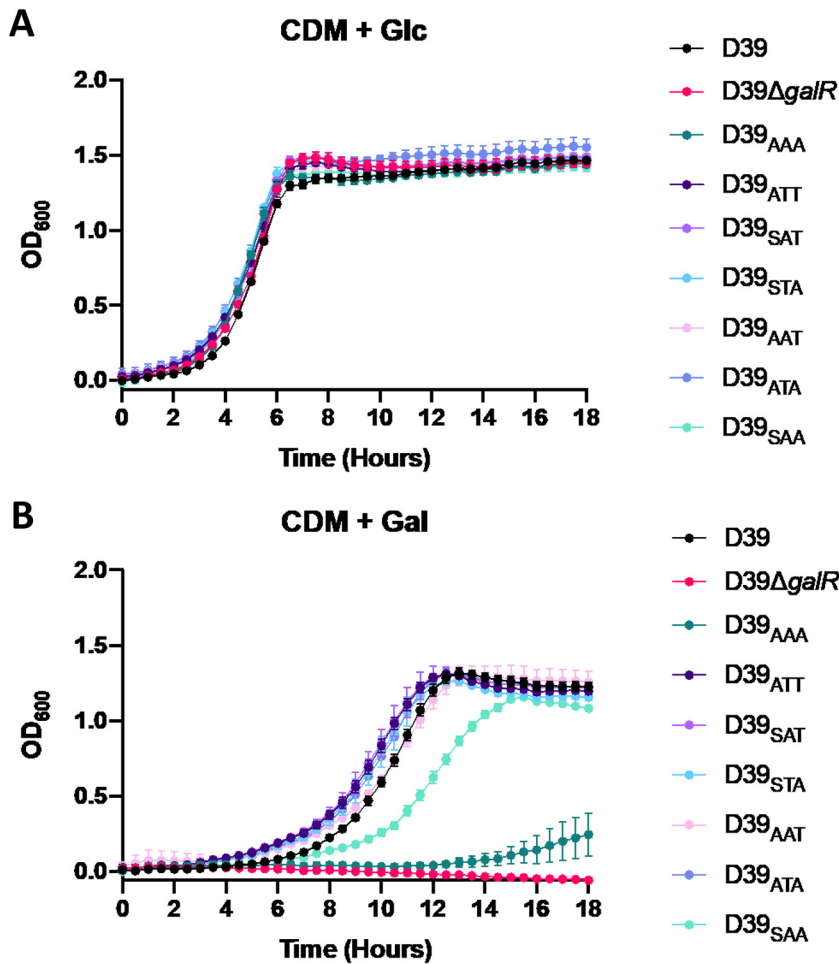
FIG 2 Impact of GalR mutations on bacterial growth. D39, D39ΔgalR, D39_AAA, D39_ATT, D39_SAT, D39_STA, D39_AAT, D39_ATA, and D39_SAA were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring the OD_{600} every 30 min for a total of 18 h. Data points are the mean OD_{600} from triplicate assays.

GalR phosphorylation sites alone does not significantly impact the capacity to grow in CDM + Gal. However, mutation of both T319 and T323, as occurs in D39_SAA, reduced the capacity of the strain to grow in this medium. Thus, the first GalR phosphorylation site (S317) on its own is insufficient to fully sustain growth in Gal.

To complement these data, gene expression analyses were conducted on all the GalR mutants to assess the impact on expression of both Leloir and T6P pathway genes. Strains were grown overnight on blood agar, washed and resuspended in CDM + Gal, and then incubated for 30 min. RNA was then extracted, and expression of galR, galK, and lacD was quantitated by reverse transcription-qualitative PCR (qRT-PCR) (see Materials and Methods) (Fig. 3). Expression of galR itself was, as expected, undetectable in D39ΔgalR. galR expression was also significantly downregulated in all the mutants tested, with D39_AAA being the most affected, showing an 88% reduction in expression (Fig. 3A).

It was previously shown that galR regulates the galKT operon (14). Here, there was significantly decreased expression of galK in all GalR mutants compared to wild-type D39 (Fig. 3B). In particular, the D39ΔgalR and D39_AAA strains showed similarly low levels of galK expression (≥98% reduction). These findings are largely consistent with the effects of the mutations on the expression of galR itself (Fig. 3A). It is worth noting that only those mutants with >98% reduction in galK expression (D39ΔgalR and D39_AAA) exhibited severe growth defects in CDM + Gal (Fig. 2B).
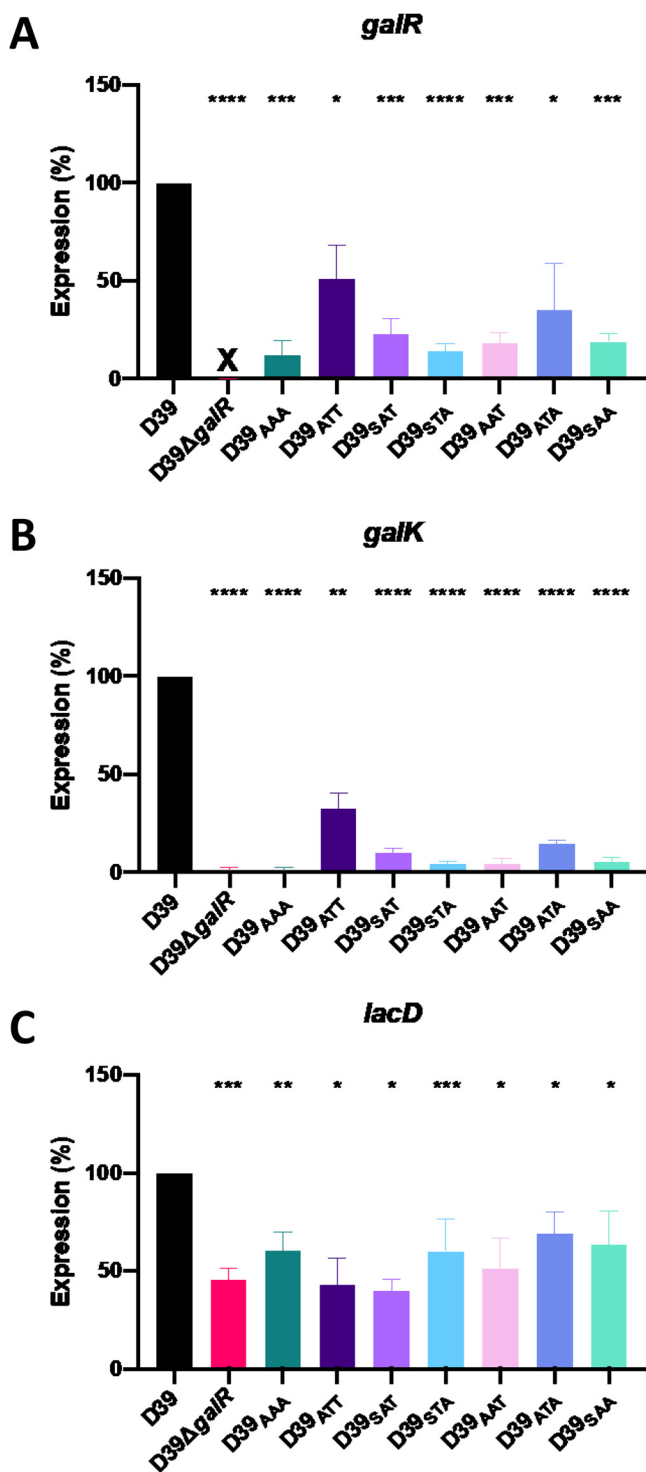
FIG 3 Differential gene expression in GalR mutants. D39, D39Δ*galR*, D39$_{AAA}$, D39$_{ATT}$, D39$_{SAT}$, D39$_{STA}$, D39$_{AAT}$, D39$_{ATA}$, and D39$_{SAA}$ were cultured overnight on blood agar plates, washed, and resuspended to a final OD$_{600}$ of 0.25 in CDM + Gal and incubated for 30 min. RNA was then extracted and the levels of *galR* (A), *galK* (B), and *lacD* (C) mRNA were quantitated by qRT-PCR using *gyrA* as an internal control. Data presented are the mean ± standard deviation from three independent experiments, expressed as a percentage of the result for D39. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$, unpaired t test (relative to D39); ns, not significant; X, transcript absent due to gene deletion.
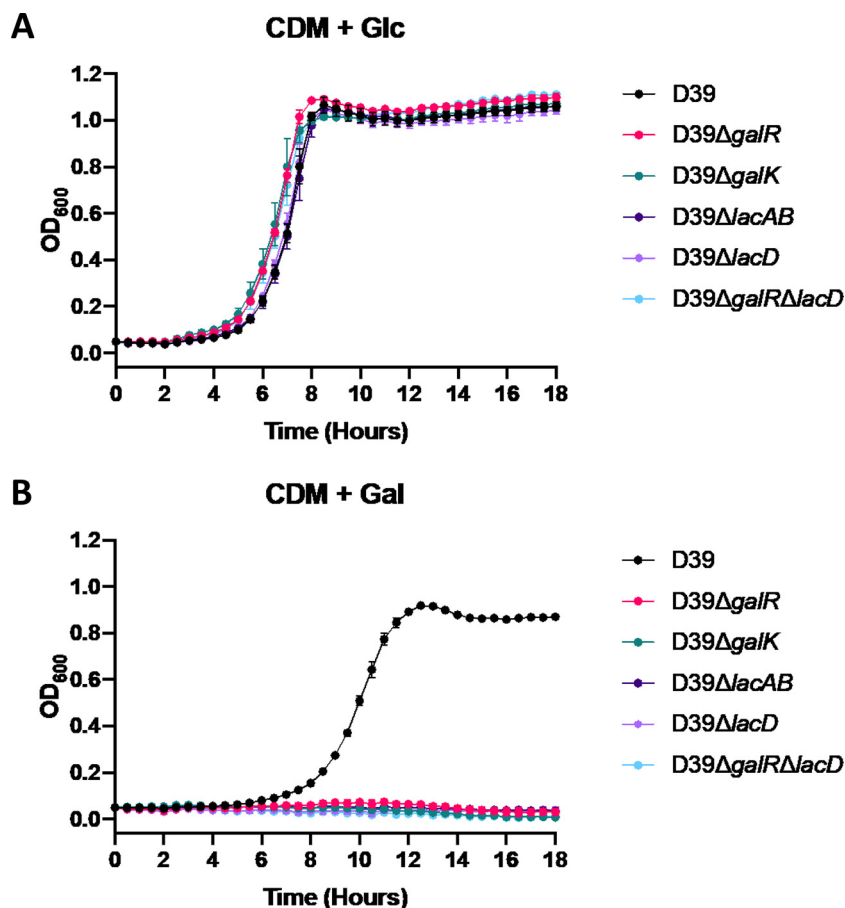
FIG 4 Differential growth of Leloir and T6P pathway mutants. D39, D39Δ*galR*, D39Δ*galK*, D39Δ*lacAB*, D39Δ*lacD*, and D39Δ*galR*Δ*lacD* were grown in CDM + Glc (A) or CDM + Gal (B). Growth was monitored by measuring the OD_{600} every 30 min for 18 h. Data points are the mean OD_{600} from triplicate assays.

We also assessed whether the absence of functional Leloir pathway expression had any effect on the expression of the T6P pathway by examining *lacD* expression. *lacD* encodes the last enzyme of the T6P pathway and is responsible for the conversion of tagatose 1,6-bP to dihydroxyacetone-P and D-glyceraldehyde-3-P, which can then feed into the glycolytic pathway. Interestingly, *lacD* expression was significantly (30 to 60%) lower in all GalR mutants than in D39 (Fig. 3C), indicating a direct or indirect role for GalR phosphorylation in the expression of T6P pathway genes. However, there was no apparent association between reduced *lacD* expression and the relative ability of the various strains to grow in CDM + Gal (Fig. 2B). Collectively, these analyses indicate that all three putative GalR phosphorylation sites, S317, T319, and T323, are essential for galactose metabolism in *S. pneumoniae* and are required for the activation of both the Leloir and T6P pathways.

Both the Leloir and T6P pathways are required for growth in galactose. In order to assess the contribution of the Leloir and T6P pathways to Gal metabolism, growth of D39, D39Δ*galR*, D39Δ*galK*, D39Δ*lacAB*, D39Δ*lacD*, and D39Δ*galR*Δ*lacD* was analyzed in both CDM + Glc and CDM + Gal (Fig. 4). All mutant strains grew as well as D39 when Glc was the only carbon source (Fig. 4A). In contrast, in CDM + Gal, all mutant strains displayed complete abrogation of growth (Fig. 4B). Thus, the presence of both the functional Leloir and T6P pathways is required for growth in Gal, indicating potential interplay between these two pathways.

Contribution of GalR and its putative phosphorylation sites to regulation of Gal metabolism. To further understand the apparent cross talk between the Leloir and T6P pathways, we analyzed the expression of *galR*, *galK*, *lacA*, and *lacD* in D39,
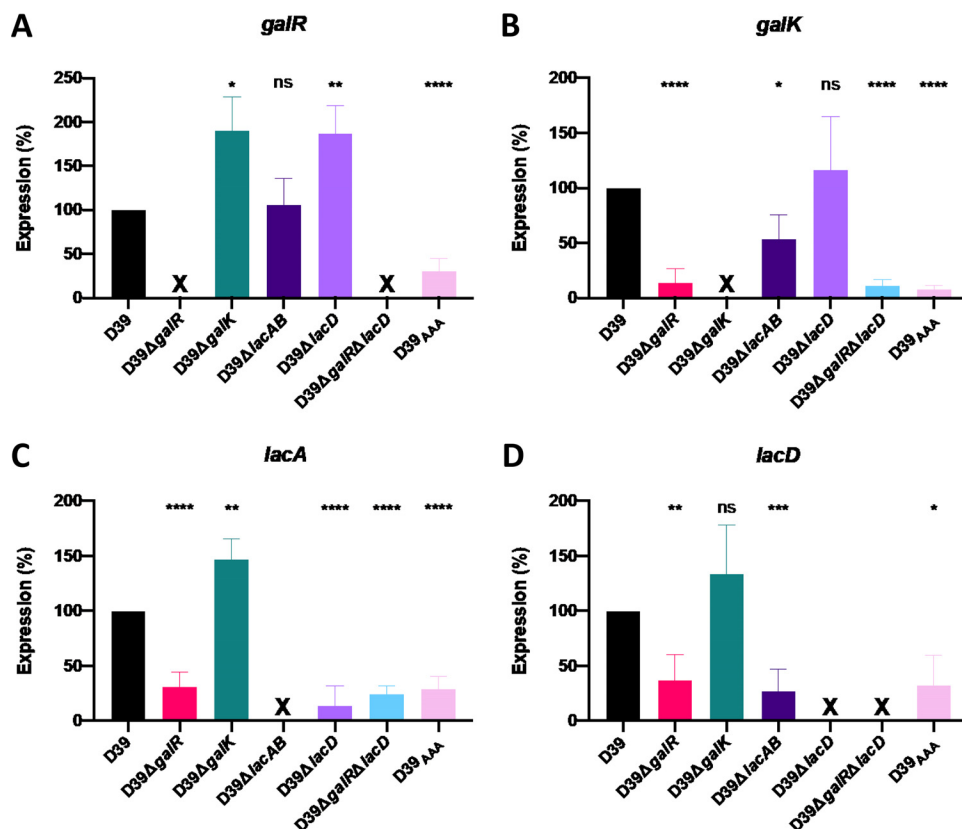
FIG 5 Differential gene expression in Leloir and T6P pathway mutants. D39, D39Δ*galR*, D39Δ*galK*, D39Δ*lacAB*, D39Δ*lacD*, D39Δ*galR*Δ*lacD*, and D39$_{AAA}$ were cultured overnight on blood agar, washed, and resuspended to a final OD$_{600}$ of 0.25 in CDM + Gal and incubated for 30 min. RNA was then extracted, and qRT-PCR was used to assess levels of *galR*, *galK*, *lacAB*, and *lacD* mRNA, using *gyrA* as an internal control. Data presented are the mean ± standard deviation from three independent experiments, expressed as a percentage of that for D39. *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$; ****, $P < 0.0001$, by unpaired *t* test (relative to D39); ns, not significant; X, transcript absent due to gene deletion.

D39Δ*galR*, D39Δ*galK*, D39Δ*lacAB*, D39Δ*lacD*, D39Δ*galR*Δ*lacD*, and D39$_{AAA}$ (Fig. 5). The expression of *galR* was significantly upregulated in both D39Δ*galK* and D39Δ*lacD* relative to D39, whereas it was unaffected in D39Δ*lacAB* (Fig. 5A). *galK* expression was unaffected in D39Δ*lacD* but was significantly downregulated relative to D39 in D39Δ*lacAB* (Fig. 5B). Both *galR* and *galK* expression were also largely abrogated in D39Δ*galR* and D39$_{AAA}$, confirming the requirement for functional GalR for activation of the Leloir pathway, as previously shown in Fig. 3B. Unsurprisingly, the expression of these two genes was also abrogated in D39Δ*galR*Δ*lacD*.

Interestingly, *lacA* expression (Fig. 5C) was significantly upregulated relative to D39 in D39Δ*galK*. In the remaining mutants, the expression of both *lacA* and *lacD* was downregulated relative to D39 (Fig. 5C and D). Collectively, these findings underscore the requirement for GalR (and putative phosphorylation sites therein) for upregulation of the T6P pathway when the Leloir pathway is blocked. This could be due to feedback inhibition from accumulation of either Leloir or T6P intermediates. We also assessed expression of *lacR2*, the repressor of the *lacI* operon. In the presence of Gal, there was a significant decrease in expression of *lacR2* in D39Δ*lacD* relative to D39, implying derepression of the operon to promote upregulation of the T6P pathway genes. However, there was no significant difference in *lacR2* expression between D39 and D39Δ*galR*, indicating that effects of GalR on T6P pathway gene expression are not mediated via *lacR2* (data not shown).

Impact of putative GalR phosphorylation in a mouse model of pneumococcal infection. To determine the role of putative GalR phosphorylation during pneumococ-
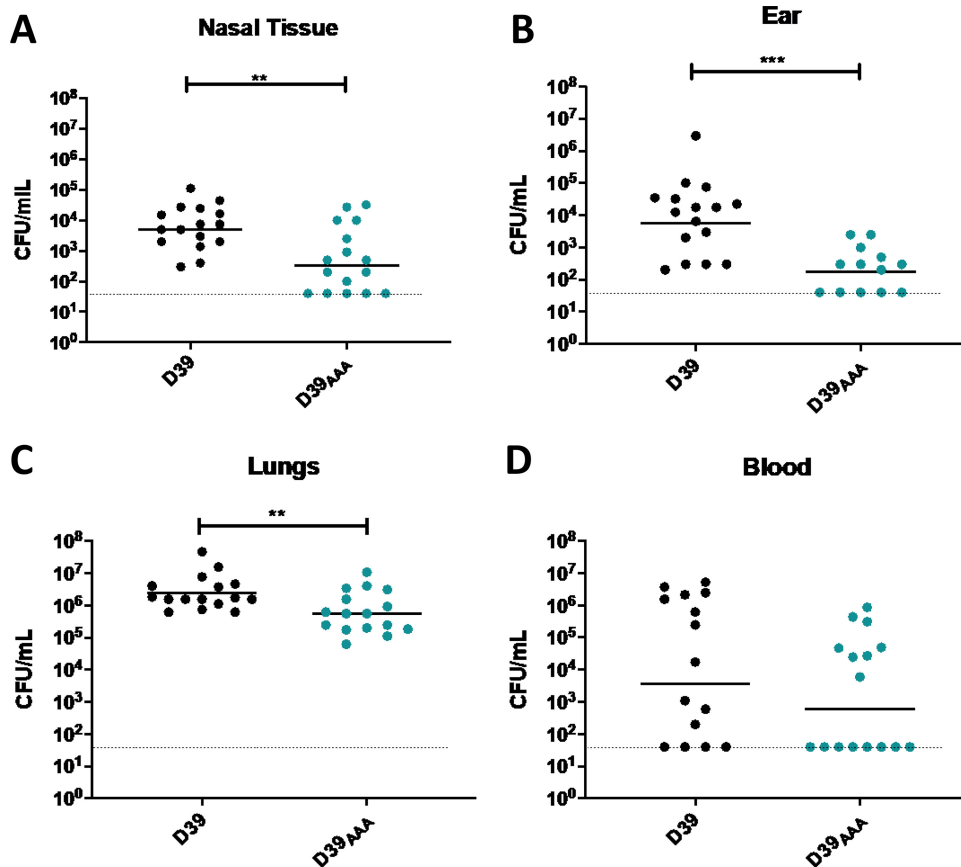
FIG 6 Virulence phenotypes of D39 and D39$_{AAA}$. Groups of 16 mice were infected intranasally with 10^7 CFU of the indicated strain. At 24 h, mice were euthanized, and the numbers of pneumococci isolated from the nasal tissue (A), ears (B), lungs (C), and blood (D) were enumerated. Viable counts (total CFU per tissue or CFU per ml blood) are shown for each mouse in each niche; the horizontal bars indicate the geometric mean (GM) CFU for each group; the dotted line indicates the threshold of detection. The significance of the differences in GM bacterial load between groups was analyzed using unpaired t tests; *, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$.

cal infection, the virulence of the *S. pneumoniae* D39 and D39$_{AAA}$ strains was assessed in a murine model of infection. Groups of mice were challenged intranasally with 1×10^7 CFU of each strain, and the numbers of pneumococci present in various niches (nasal tissue, ear, lungs, and blood) were determined at 24 h postchallenge (Fig. 6). The D39$_{AAA}$ strain exhibited a significantly attenuated virulence phenotype, with reduced bacterial loads relative to D39 in nasal tissue (Fig. 6A), ears (Fig. 6B), and lungs (Fig. 6C). No significant difference in bacterial numbers between D39 and D39$_{AAA}$ was observed in the blood (Fig. 6D).

DISCUSSION

The results of this study support previous findings that GalR is important for galactose metabolism in *S. pneumoniae* (14). Here, we demonstrated that the putative GalR phosphorylation sites (S317, T319, and T323) are required for growth in a defined medium with Gal as the sole carbon source. Mutation of all putative phosphorylation sites to alanine (D39$_{AAA}$) completely abrogated growth in Gal and reduced expression of *galR* and *galK* to levels comparable to those of the D39Δ*galR* strain (Fig. 2B and Fig. 3). Moreover, the substitution of these amino acids with alanine appears to alter the interaction of GalR with the *galR* operator sequence; as a result, the defects observed in *galK* expression may be at least partially attributable to a reduction in GalR abundance, rather than a direct alteration in binding to the *galK* operator. The precise mechanism behind these defects remains unknown but may be due to effects on folding, dimerization, or binding of effector molecules rather than by directly prevent-

ing phosphorylation. However, structural modeling of GalR shows that these residues are positioned in a distinct location to the putative galactose and DNA binding regions (Fig. 1B). Additional studies exploring the interaction of purified GalR and GalR$_{AAA}$ with each operator DNA sequence may provide greater insight into the regulation of the Leloir pathway genes.

In *S. pneumoniae*, the kinase responsible for phosphorylation of serine and threonine residues is StkP (15, 16). A previous proteomic study performed in medium containing Glc failed to identify GalR as a target of StkP, but LacA was identified as a possible target for StkP-mediated phosphorylation (17). Preliminary data have revealed that a D39Δ*stkP* strain was unable to grow in Gal (data not shown), but this may be a consequence of the defect in LacA phosphorylation rather than GalR. As a result, additional studies to directly examine the role of StkP in GalR phosphorylation are required.

Gene expression studies demonstrated that of the various GalR mutants constructed in this study, those with the greatest defects in growth in CDM + Gal, namely, D39Δ*galR* and D39$_{AAA}$, exhibited virtually undetectable levels of *galK* expression (>98% reduction) (Fig. 3B), showing a direct link between Leloir pathway gene expression and growth in Gal. The single and double point mutants also showed significantly reduced expression of *galR* and *galK*, but this level of expression still enabled sufficient Leloir pathway activity to sustain growth in CDM + Gal (Fig. 2B). Interestingly, within the single or double point mutants, the D39$_{SAA}$ mutant was the only one to show a defect in growth in Gal compared to D39 (Fig. 2B), but gene expression was similar to that in the other mutants that grew comparably to the wild type (Fig. 2B and Fig. 3). Therefore, growth in Gal can occur even at low levels of *galK* expression. Thus, the effects on Leloir pathway gene expression and growth in CDM + Gal could be attributable to inadequate levels of GalR in the cell, reduced capacity of the respective GalR phosphorylation site mutants to activate Leloir pathway genes such as *galK*, or a combination of both.

An unexpected finding of the current study was that all the investigated GalR mutants exhibited significantly reduced expression of the T6P pathway gene *lacD* (Fig. 3C), indicating a direct or indirect effect of GalR on the T6P pathway. This may provide a mechanistic basis for a previously proposed subtle regulatory link between the Leloir and T6P pathways (6). This was further examined by comparing the growth and gene expression phenotypes of D39 mutants with deletions in *galR*, *galK*, *lacAB*, *lacD*, and *galR* plus *lacD*. Unlike growth in CDM + Glc, none of these mutants was capable of growth in CDM + Gal (Fig. 4), indicating that both the Leloir and T6P pathways are essential for growth under these conditions. A previous study of a similar D39 *galK* deletion mutant (14) reported that growth in the presence of Gal was reduced relative to wild-type D39 but not completely abrogated as shown in the present study. This discrepancy may be attributable to the use of a more nutrient-rich M17 medium in that study (14). A separate study reported complete abrogation of growth of a D39 *galK* deletion mutant in Gal, whereas a *lacD* deletion mutant was able to grow logarithmically in Gal, albeit after a long lag phase (6).

In the present study, gene expression analyses have shown that maximal expression of the T6P pathway genes *lacA* and *lacD* require functional GalR and that all three putative phosphorylation sites of GalR are necessary to achieve this function (Fig. 5). Additionally, there appears to be a link between *galK* and *lacAB* gene expression. Deletion of *galK* upregulated *lacA* expression, perhaps as a consequence of the upregulation of *galR* in this mutant. On the other hand, deletion of *lacAB* significantly reduced *galK* expression but did not impact expression of *galR*. Furthermore, *galR* expression was significantly elevated in the *lacD* mutant. Thus, there is a complex interplay between the Leloir and T6P pathways in the various mutants, presumably mediated by intracellular concentrations of intermediates or end products of either pathway.

Notwithstanding the above-described complexities, the present study has shown that the GalR putative phosphorylation sites play a significant role in pneumococcal infection. Mice infected with D39$_{AAA}$ displayed significantly reduced bacterial loads in

the nasopharynx, middle ear, and lungs relative to those infected with wild-type D39 (Fig. 6). These findings are compatible with previous studies showing reduced naso-pharyngeal colonization and reduced systemic virulence of D39 *galK* and *lacD* deletion mutants after intranasal, but not intravenous, challenge of mice (6). However, the impact of the putative GalR phosphorylation sites has not previously been investigated. Clearly, the capacity to metabolize Gal is important for survival and proliferation in the upper respiratory tract and the middle ear, where it is an important carbon source (6). Moreover, metabolism of Gal by pneumococci *in vitro* is known to lead to increased production of CPS relative to cells growing on Glc, which may be the basis for the altered virulence profiles (7). We previously showed that treatment with the quorum-sensing molecule AI-2 upregulates Leloir pathway gene expression and CPS production in the presence of Gal *in vitro*, as well as virulence in an intranasal challenge model (7). This upregulation was dependent on the PTS component FruA, which is presumed to be the bacterial surface receptor for AI-2. This signaling molecule is a di-ketopentose and may structurally mimic the natural cargo of FruA, namely, fructose, and if AI-2 is capable of internalization via the FruA PTS system, then it would be expected to be phosphorylated during import. It is tempting to speculate that such phosphorylated AI-2 may play a direct or indirect role in GalR phosphorylation, perhaps acting as a phosphate donor, thereby mediating upregulation of the Leloir pathway. This study has shown that, collectively, the GalR putative phosphorylation sites play a key role in virulence and the ability to metabolize Gal and has revealed a complex interplay between the Gal metabolic pathways in *S. pneumoniae*.

MATERIALS AND METHODS

Structural modeling of GalR. The GalR amino acid sequence (SPD_1635) was obtained from the NCBI database and input into SWISS-MODEL (18). A homology model was generated based on the 2.4-Å structure (PDB: 1JFS) of the *Escherichia coli* purine repressor (PurR) W147F mutant (12). The cartoon representation of the GalR homology model and the aligned PurR template were generated in PyMOL version 2.3.3 (Schrödinger). The root mean square deviation (RMSD) between the GalR model and the PurR template was determined by alignment in PyMOL. The DNA binding domain and putative sugar binding residues were identified by the NCBI conserved domain search (19), and the locations of the putative phosphorylated residues were determined based on the phosphoproteomic findings (11).

Bacterial strains and growth conditions. The *S. pneumoniae* strains used in this study are listed in Table 1. Pneumococci were routinely cultured on Columbia blood agar base supplemented with 5% sterile horse blood overnight at 37°C with 5% CO_2. In order to select for mutant strains, blood agar plates were supplemented with 0.2 μg/ml erythromycin, 200 μg/ml spectinomycin, 200 μg/ml kanamycin, or 200 μg/ml streptomycin, as appropriate. Growth experiments were performed with pneumococci grown in chemically defined medium (CDM) supplemented with vitamins, amino acids, choline, and catalase, as previously described (20), and either 0.5% glucose (CDM + Glc) or galactose (CDM + Gal).

Construction of mutants. Genes were deleted from *S. pneumoniae* using overlap extension PCR using the primers listed in Table 2, followed by transformation, essentially as previously described (21). For strains harboring single amino acid substitutions within *galR*, allelic exchange mutagenesis was performed through use of the Janus cassette system, as described previously (20, 22). Mutants were confirmed by PCR and Sanger sequencing using the primers listed in Table 2 (AGRF, Adelaide, Australia).

Growth assays. Growth assays were performed in flat-bottom 96-well microtiter plates with a final volume of 200 μl. Cells were inoculated at a starting optical density at 600 nm (OD_{600}) of 0.05 in either CDM + Glc or CDM + Gal and then incubated at 37°C with 5% CO_2 (20). The OD_{600} was measured every 30 min for a total of 18 h in a SpectroSTAR Omega spectrophotometer (BMG Labtech). Assays were performed in triplicate with a minimum of two independent experiments.

Bacterial RNA extraction and real-time qRT-PCR. *S. pneumoniae* strains were first cultured over-night on blood agar plates at 37°C with 5% CO_2. Cells were then harvested, washed, and resuspended in 1 ml of CDM + Gal to a final OD_{600} of 0.25. Cells were then incubated for 30 min at 37°C with 5% CO_2, after which RNA was extracted using a Qiagen RNeasy minikit as per the manufacturer's instructions. Gene expression was analyzed using one-step relative real-time qRT-PCR in a Roche LC480 real-time cycler, as described previously (20). The primers used to amplify target genes (listed in Table 2) were used at a final concentration of 200 nM. Amplification data were analyzed using the comparative critical threshold ($2^{\Delta\Delta CT}$) method (23) and are presented as a percentage of total expression relative to that for D39 for each gene. Assays were performed in triplicate with a minimum of two independent experiments. Statistical analyses were performed using two-tailed Student's *t* test; *P* values of <0.05 were deemed statistically significant.

Murine infection model. All animal experiments were approved by the University of Adelaide Animal Ethics Committee. Female outbred 4- to 6-week-old CD-1 (Swiss) mice were anaesthetized by intraperitoneal injection with ketamine and xylazine before being intranasally inoculated with 1×10^7 CFU of D39 or $D39_{AAA}$ in a total of 50 μl, as previously described (24). The challenge dose was

TABLE 2 Primers used in this study

Primer	Sequence (5′–3′)
GalR F	AAGACAAGCCAGAACCATTTGGG
GalR Ery F	CGGGAGGAAATAATTCTATGAGGAAAGTACCCTAAATCAAGAATAG
GalR R	GACAAGGTTGTTGTTATCGGTGAT
GalR Ery R	TTGTTCATGTAATCACTCCTTCTGTGCAATGTCTTTTAAGGTAGCC
LacD F	CTGTATCGCTATTCTCCACG
LacD Ery F	CGGGAGGAAATAATTCTATGAGGGTATCATCTCAGCTCTTGC
LacD R	TGATCTGCTAGCTTCTGAC
LacD Ery R	TTGTTCATGTAATCACTCCTTCGCAAGAGCTGAGATGATACC
Ery F	GAAGGAGTGATTACATGAACAA
Ery R	CTCATAGAATTATTTCCTCCCG
GalK F	CCTTCATTAAGTCATAGCCAGA
GalK Spec F	AAATAACAGATTGAAGAAGGTATAAGAAGTAGTTGGATACGCTCC
GalK R	GTAGGACAGACATTGGCCA
GalK Spec R	TATGTATTCATATATATCCTCCTCCAAATTGATACGACCTGGTG
LacAB F	AATATCGGACAAGCTGGT
LacAB Spec F	AAATAACAGATTGAAGAAGGTATAAGCTCAACAAACAGACGC
LacAB R	TTCGTCTGAGCTATCTACATC
LacAB Spec R	TATGTATTCATATATATCCTCCTCATCTCAAACCTGCAGCATC
LacD F	CTGTATCGCTATTCTCCACG
LacD Spec F	AAATAACAGATTGAAGAAGGTATAATGAAGCAGCAGCTCGCGAAT
LacD R	TATGTATTCATATATATCCTCCTCGCAAGAGCTGAGATGATACC
LacD Spec R	GTTGCGTAATCACATCATAGAT
J253	GAGGAGGATATATATGAATACATACG
J254	TTATACCTTCTTCAATCTGTTATTTAAATAGTTTATAGTTA
AAA F	ACTCCACGGTCGCAAAATTCCTGCGCTGGCCATGCTGGGAGCCAGACTGACATTAAGA
AAA R	TCTTAATGTCAGTCTGGCTCCCAGCATGGCCAGCGCAGGAATTTTGCGACCGTGGAGT
ATT F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGACCATGCTGGGAACCAGACTGACATTAAGAGAGAAAGTACCC
ATT R	GGGTACTTTCTCTTAATGTCAGTCTGGTTCCCAGCATGGTCAGGGCAGGAATTTTGCGACCGTGGAGTACT
SAT F	AGTACTCCACGGTCGCAAAATTCCTAGCCTGGCCATGCTGGGAACCAGACTGACATTAAGAGAGAAAGTACCC
SAT R	GGGTACTTTCTCTTAATGTCAGTCTGGTTCCCAGCATGGCCAGGCTAGGAATTTTGCGACCGTGGAGTACT
STA F	AGTACTCCACGGTCGCAAAATTCCTAGCCTGACCATGCTGGGAGCCAGACTGACATTAAGAGAGAAAGTACCC
STA R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGTCAGGCTAGGAATTTTGCGACCGTGGAGTACT
AAT F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGGCCATGCTGGGAACCAGACTGACATTAAGAGAGAAAGTACCC
AAT R	GGGTACTTTCTCTTAATGTCAGTCTGGTTCCCAGCATGGCCAGGGCAGGAATTTTGCGACCGTGGAGTACT
ATA F	AGTACTCCACGGTCGCAAAATTCCTGCCCTGACCATGCTGGGAGCCAGACTGACATTAAGAGAGAAAGTACCC
ATA R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGTCAGGGCAGGAATTTTGCGACCGTGGAGTACT
SAA R	AGTACTCCACGGTCGCAAAATTCCTAGCCTGGCCATGCTGGGAGCCAGACTGACATTAAGAGAGAAAGTACCC
SAA R	GGGTACTTTCTCTTAATGTCAGTCTGGCTCCCAGCATGGCCAGGCTAGGAATTTTGCGACCGTGGAGTACT
gyr-RT-F	ACTGGTATCGCGGTTGGGAT
gyr-RT-R	ACCTGATTTCCCCATGCAA
galR-RT-F	TCTCTATCGCCGACCGTATCC
galR-RT-R	GGTGTAGCCCAGCTCTTCAG
galK-RT-F	CACGTTTCTCTGGAGCATGA
galK-RT-R	ATGGCACAGCCACTAAAACC
galT-RT-F	GTGGGAGAAGGTGTTTTGGA
galT-RT-R	ACGCGCAGTCTGACTATCCT
lacA-RT-F	CGTGATTGATGCTTATGGAG
lacA-RT-R	AGCCAATTCATCACCAACAAG
lacD-RT-F	CATCGGTTCTGAGTGTGTGG
lacD-RT-R	AAAGCGTGGGTCTGAAAAGA
galR–Seq F	AATCTATCATGATGAACTGGTC
galR–Seq R	CATAATGGAGGGCGTATGG

REFERENCES

1. 2010. Nuorti JP, Whitney CG. Prevention of pneumococcal disease among infants and children: use of 13-valent pneumococcal conjugate vaccine and 23-valent pneumococcal polysaccharide vaccine: recommendations of the Advisory Committee on Immunization Practices (ACIP). MMWR Recomm Rep 59:1–18. https://www.cdc.gov/mmwr/preview/mmwrhtml/rr5911a1.htm.

2. Thigpen MC, Whitney CG, Messonnier NE, Zell ER, Lynfield R, Hadler JL, Harrison LH, Farley MM, Reingold A, Bennett NM, Craig AS, Schaffner W, Thomas A, Lewis MM, Scallan E, Schuchat A, Emerging Infections Programs N. 2011. Bacterial meningitis in the United States, 1998–2007. N Engl J Med 364:2016–2025. https://doi.org/10.1056/NEJMoa1005384.

3. WHO. 2007. Pneumococcal conjugate vaccine for childhood immunization: WHO position paper. Wkly Epidemiol Rec 82:93–104.

4. Gray BM, Converse GM, III, Dillon HC, Jr. 1980. Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. J Infect Dis 142:923–933. https://doi.org/10.1093/infdis/142.6.923.

5. Pezzulo AA, Gutierrez J, Duschner KS, McConnell KS, Taft PJ, Ernst SE, Yahr TL, Rahmouni K, Klesney-Tait J, Stoltz DA, Zabner J. 2011. Glucose depletion in the airway surface liquid is essential for sterility of the airways. PLoS One 6:e16166. https://doi.org/10.1371/journal.pone.0016166.

6. Paixao L, Oliveira J, Verissimo A, Vinga S, Lourenco EC, Ventura MR, Kjos M, Veening JW, Fernandes VE, Andrew PW, Yesilkaya H, Neves AR. 2015. Host glycan sugar-specific pathways in *Streptococcus pneumoniae*: galactose as a key sugar in colonisation and infection [corrected]. PLoS One 10:e0121042. https://doi.org/10.1371/journal.pone.0121042.

7. Trappetti C, McAllister LJ, Chen A, Wang H, Paton AW, Oggioni MR, McDevitt CA, Paton JC. 2017. Autoinducer 2 signaling via the phosphotransferase FruA drives galactose utilization by *Streptococcus pneumoniae*, resulting in hypervirulence. mBio 8:e02269-16. https://doi.org/10.1128/mBio.02269-16.

8. Bidossi A, Mulas L, Decorosi F, Colomba L, Ricci S, Pozzi G, Deutscher J, Viti C, Oggioni MR. 2012. A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. PLoS One 7:e33320. https://doi.org/10.1371/journal.pone.0033320.

9. Afzal M, Shafeeq S, Kuipers OP. 2014. LacR is a repressor of *lacABCD* and LacT is an activator of *lacTFEG*, constituting the *lac* gene cluster in *Streptococcus pneumoniae*. Appl Environ Microbiol 80:5349–5358. https://doi.org/10.1128/AEM.01370-14.

10. Zeng L, Das S, Burne RA. 2010. Utilization of lactose and galactose by *Streptococcus mutans*: transport, toxicity, and carbon catabolite repression. J Bacteriol 192:2434–2444. https://doi.org/10.1128/JB.01624-09.

11. Sun X, Ge F, Xiao CL, Yin XF, Ge R, Zhang LH, He QY. 2010. Phosphoproteomic analysis reveals the multiple roles of phosphorylation in pathogenic bacterium *Streptococcus pneumoniae*. J Proteome Res 9:275–282. https://doi.org/10.1021/pr900612v.

12. Huffman JL, Lu F, Zalkin H, Brennan RG. 2002. Role of residue 147 in the gene regulatory function of the *Escherichia coli* purine repressor. Biochemistry 41:511–520. https://doi.org/10.1021/bi0156660.

13. Fleming E, Lazinski DW, Camilli A. 2015. Carbon catabolite repression by

14. seryl phosphorylated HPr is essential to *Streptococcus pneumoniae* in carbohydrate-rich environments. Mol Microbiol 97:360–380. https://doi.org/10.1111/mmi.13033.

14. Afzal M, Shafeeq S, Manzoor I, Kuipers OP. 2015. GalR Acts as a transcriptional activator of *galKT* in the presence of galactose in *Streptococcus pneumoniae*. J Mol Microbiol Biotechnol 25:363–371. https://doi.org/10.1159/000439429.

15. Echenique J, Kadioglu A, Romao S, Andrew PW, Trombe MC. 2004. Protein serine/threonine kinase StkP positively controls virulence and competence in *Streptococcus pneumoniae*. Infect Immun 72:2434–2437. https://doi.org/10.1128/iai.72.4.2434-2437.2004.

16. Novakova L, Saskova L, Pallova P, Janecek J, Novotna J, Ulrych A, Echenique J, Trombe MC, Branny P. 2005. Characterization of a eukaryotic type serine/threonine protein kinase and protein phosphatase of *Streptococcus pneumoniae* and identification of kinase substrates. FEBS J 272:1243–1254. https://doi.org/10.1111/j.1742-4658.2005.04560.x.

17. Hirschfeld C, Gómez-Mejia A, Bartel J, Hentschker C, Rohde M, Maaß S, Hammerschmidt S, Becher D. 2019. Proteomic investigation uncovers potential targets and target sites of pneumococcal serine-threonine kinase StkP and phosphatase PhpP. Front Microbiol 10:3101. https://doi.org/10.3389/fmicb.2019.03101.

18. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, Kiefer F, Gallo Cassarino T, Bertoni M, Bordoli L, Schwede T. 2014. SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. Nucleic Acids Res 42:W252–W258. https://doi.org/10.1093/nar/gku340.

19. Lu S, Wang J, Chitsaz F, Derbyshire MK, Geer RC, Gonzales NR, Gwadz M, Hurwitz DI, Marchler GH, Song JS, Thanki N, Yamashita RA, Yang M, Zhang D, Zheng C, Lanczycki CJ, Marchler-Bauer A. 2020. CDD/SPARCLE: the conserved domain database in 2020. Nucleic Acids Res 48: D265–D268. https://doi.org/10.1093/nar/gkz991.

20. Minhas V, Harvey RM, McAllister LJ, Seemann T, Syme AE, Baines SL, Paton JC, Trappetti C. 2019. Capacity to utilize raffinose dictates pneumococcal disease phenotype. mBio 10:e02596-18. https://doi.org/10.1128/mBio.02596-18.

21. Iannelli F, Pozzi G. 2004. Method for introducing specific and unmarked mutations into the chromosome of *Streptococcus pneumoniae*. Mol Biotechnol 26:81–86. https://doi.org/10.1385/MB:26:1:81.

22. Sung CK, Li H, Claverys JP, Morrison DA. 2001. An *rpsL* cassette, Janus, for gene replacement through negative selection in *Streptococcus pneumoniae*. Appl Environ Microbiol 67:5190–5196. https://doi.org/10.1128/AEM.67.11.5190-5196.2001.

23. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2(-\Delta\Delta C(T))$ method. Methods 25:402–408. https://doi.org/10.1006/meth.2001.1262.

24. Tikhomirova A, Trappetti C, Standish AJ, Zhou Y, Breen J, Pederson S, Zilm PS, Paton JC, Kidd SP. 2018. Specific growth conditions induce a *Streptococcus pneumoniae* non-mucoidal, small colony variant and determine the outcome of its co-culture with *Haemophilus influenzae*. Pathog Dis 76:fty074. https://doi.org/10.1093/femspd/fty074.