The University Of Sheffield.

The Contribution of Mobile Genetic Elements to the Molecular Biology of *Streptococcus pyogenes* Genotype *emm*4

A thesis submitted for the degree of

Doctor of Philosophy

University of Sheffield

By

Alex Remmington BSc (Hons) MSc

Department of Molecular Biology and Biotechnology

University of Sheffield

Department of Molecular Biology and Biotechnology, University of Sheffield,

Firth Court, Western Bank,

Sheffield, S10 2TN

2020

Abstract

Genotype *emm*4 *Streptococcus pyogenes* constitute a major disease-causing *emm*-type worldwide, and have been associated with scarlet fever. Genotype *emm*4 are lysogenised by three toxigenic prophage, and an atoxigenic SpyCI, thought to regulate the DNA mismatch repair operon by growth-phase dependent excision. Phylogenetic analyses determined that *emm*4 in our collection and internationally, clustered with either reference genomes MGAS10750 or MEW427, the latter of which were found to have undergone substantial gene loss within prophage-encoding regions. Gene loss affected modules pertaining to lysogeny, genetic regulation, replication, and encapsidation, but not cognate toxins. Corresponding elements of $M4_{complete}$ *emm*4 were full-length and undegraded. The lineages associated with either MGAS10750 or MEW427 were thus denoted $M4_{complete}$ and $M4_{degraded}$, respectively. Prophage of $M4_{degraded}$ isolates were shown to be cryptic, and incapable of excision, whereas prophage of $M4_{complete}$ isolates were inducible. Unexpectedly, this had little effect on cognate toxin expression, which was independent of induction and replication. Further, the *emm*4 SpyCI was found to be immobile, irrespective of gene content, and appears not to regulate the DNA MMR operon, as had been reported. A number of assays did not demonstrate lineage-specific phenotypes, however, $M4_{complete}$ isolates were significantly more virulent than $M4_{degraded}$ isolates in a *Galleria mellonella* infection model. Intriguingly, efforts to explore this response in more infection-relevant conditions revealed that co-culture with human tonsil cells greatly enhanced expression of the prophage-associated superantigen *ssa*, among both $M4_{complete}$ and $M4_{degraded}$ isolates, and this was also observed among *emm*3 and *emm*12. There may also be a hitherto unappreciated role for major bacterial regulators in the control and elicitation of this response. These data suggest that the relationship between *S. pyogenes*, their viral parasites and the human host may be much more complex than had been anticipated, and present a number of exciting avenues for further study.

Collaborations

I am indebted to a number of students that I had the privilege to co-supervise over the course of the project, in large part for helping me push the research forward and expand its scope by assisting in generating valuable data, but also for friendship and support. These are credited in the present work where applicable.

- **Samuel Haywood (2018-2019) – MSc student, University of Sheffield**
  Thank you for your work with SpyCI induction and *mutL* and *mutS* transcripts PCRs. Congratulations on the paper.

- **Julia Edgar (2019) – International Intern Programme, Princeton University**
  Thank you for your indomitable perseverance with SpyCI investigations, the hundred or so DNA extractions, your work with *mutL* and *mutS* transcription and all the sequence analysis. You were such a great help. Congratulations on the paper.

- **Jordan Finlay (2019-2020), MSc student, University of Sheffield**
  Thank you for your valuable adhesion work with the *emm*4's and for tolerating the very fraught, near-end-of-PhD version of me that you were unfortunately lumbered with, and for producing such a great dissertation despite the pandemic. You did great.

Acknowledgements

I would like firstly to thank the Florey Institute for Host-Pathogen Interactions for funding my PhD project. I would also like to thank Professor Simon Foster for permitting the parasitism of your laboratory in the early days, while ours was still under construction. I do hope what began as an obligate parasitism, is now much more altruistic, with only occasional theft.

Thank you also to all members of the Ewok Defence League for your ongoing comradery and support, what a bunch. I would like to extend special thanks to my long-suffering friends and colleagues Ruairidh Wilson, Henna Khalid and Callum Mynett; I could not have done this without you.

This project would not have been possible either in its conception, execution, direction and conclusion without the infinite wisdom, boundless support and endless patience of my sole supervisor, Dr. Claire E. Turner. It is challenging to express how truly grateful I am for all that you have taught me (and tried to teach me), or should I say, everything that I have learned over these past 4 years, since your arguable lapse in judgement in taking me on as your first PhD student. It seems apposite to draw a comparison between the construction of F19 and my journey as a PhD student: what began a shambolic and crumbling mess, I think is now rather functional, though perhaps more so the lab than I. I cannot thank you enough for the opportunities you have given me, nor can I put into words how much I have come to respect and admire you. Thank you.

Since being a child, I have been exceedingly fortunate to encounter a number of people who I can truly say, through their friendship, tutelage, guidance, assurance, encouragement and inspiration, are pivotal in my having gotten to this point. Having the opportunity to finally thank you all has motivated me throughout my life, even if some of you may likely forgotten me, I shall never forget you. So, thank you, in order of appearance, to: my mother, Jacqui Remmington, my nan, Josie Paddock, Jamie Wilson, Nasrin Choudhry, Charlotte Gates, Dr. Alyson Pavitt, Dr. Sarah J. Pitt, Bethan Thomas, Dr. Sarah Maxwell, and Dr. Carol Yates.

Thank you finally to Stevie Nicks, for the endless hours of sage advice and divine wisdom.

Thank you.

Table of Contents

Abstract 2

Collaborations 3

Acknowledgements 4

Table of Contents 5

List of figures 11

List of tables 13

Abbreviations 14

1. Introduction 16

1.1 *Streptococcus pyogenes* 16

1.2 Clinical manifestations of *S. pyogenes* 16

1.2.1 Tonsillopharyngitis 16

1.2.2 Scarlet fever 17

1.2.3 Impetigo 18

1.2.4 Erysipelas 18

1.2.5 Cellulitis 18

1.2.6 Necrotising fasciitis 19

1.2.7 Puerperal sepsis 20

1.2.8 Streptococcal toxic shock syndrome 21

1.2.9 Post-infectious immunological sequelae 21

1.2.10 Asymptomatic carriage 22

1.3 Control and prevention of *S. pyogenes* infection 23

1.3.1 Treatment 23

1.3.2 Antimicrobial resistance 24

1.3.3 Vaccines 24

1.4 The *S. pyogenes* genome and genotypes 25

1.5 Mobile genetic elements and *S. pyogenes* 26

1.5.1 Lysogenic bacteriophage 26

1.5.2 General genomic features of streptococcal phage 27

1.5.3 Satellite and cryptic prophage 29

1.5.4 Integrative and conjugative elements of *S. pyogenes* 35

1.5.5 *S. pyogenes* chromosomal islands....35
1.6 *S. pyogenes* and host immunity....36
1.6.1 Innate immunity....36
1.6.2 Complement....40
1.6.3 Antibodies and adaptive immunity....41
1.7 *S. pyogenes* virulence factors....42
1.7.1 M protein and M-like proteins....42
1.7.2 Hyaluronic acid capsule....44
1.7.3 Streptolysin S....46
1.7.4 Streptolysin O and NADase....46
1.7.5 Chromosomally-encoded hyaluronidases....47
1.7.6 Serum opacity factor....47
1.7.7 Streptococcal pyrogenic exotoxin B....48
1.7.8 Immunoglobulin degrading enzymes....49
1.7.9 Immunoglobulin binding proteins....49
1.7.10 SpyCEP....51
1.7.11 C5a peptidase....52
1.7.12 Streptococcal inhibitor of complement....52
1.7.13 Secreted phospholipase....53
1.7.14 DNases of *S. pyogenes*....53
1.7.15 Streptococcal superantigens....54
1.8 Transcriptional regulation in *S. pyogenes*....55
1.8.1 Two-component regulatory systems....55
1.8.2 Standalone Transcriptional Regulators....58
1.9 Mobile genetic elements & *S. pyogenes*: impact on host phenotype....59
1.9.1 Enhanced toxin expression at the host-pathogen interface....59
1.9.2 Enhanced nasopharyngeal shedding....60
1.9.3 Clonal expansion and extinction....60
1.9.4 Remodelling of the bacterial chromosome....61
1.9.5 Regulation of gene expression at attachment sites....61
1.9.6 Enhanced adhesion and colonisation....62
1.9.7 Enhanced virulence....63
1.10 Conclusions and research justifications....64
1.10.1 Hypotheses....64
1.10.2 Project aims....65

2. Materials and Methods 66

2.1 Materials 66

2.1.1 *S. pyogenes* isolates 66

2.1.2 Bacteriological growth media 67

2.1.3 Antibiotics 68

2.1.4 Human cell lines 68

2.1.5 Human cell culture media and supplements 68

2.1.6 Enzymes 69

2.1.7 Standard solutions and buffers 69

2.1.8 Kits 70

2.1.9 Miscellaneous reagents 71

2.1.10 Primers 72

2.2 Methods 73

2.2.1 Bacterial strains and media 73

2.2.2 *S. pyogenes* DNA extraction 74

2.2.3 *S. pyogenes* RNA Extraction 74

2.2.4 PCR and agarose gel electrophoresis 74

2.2.5 Detection of prophage excision by PCR 75

2.2.6 Purification of PCR products 76

2.2.7 Purification of DNA following electrophoresis 76

2.2.8 Plasmid purification 76

2.2.9 Construction of plasmids for quantitative real-time PCR 77

2.2.10 qPCR primer matrices 78

2.2.11 Reverse transcription and quantitative real-time PCR 78

2.2.12 Maintenance of human cell lines 80

3. Bioinformatic characterisation of *emm*4 isolates 81

3.1 Introduction 81

3.2 Chapter aims and hypotheses 82

3.3 Methods 83

3.3.1 Genomic comparison of reference genomes 83

3.3.2 Bioinformatic comparison of prophage-encoding regions 83

3.3.3 Variable factor typing of BSAC *emm*4 isolates 83

3.3.4 Lineage specific SNPs 83

3.3.5 Pangenome analysis 84

3.4 Results 84
3.4.1 Gene loss within prophage-encoding regions of reference genome MEW427 84
3.4.2 Bioinformatic features of MGAS10750 and MEW427 prophage 85
3.4.3 Superantigen and DNase typing of BSAC *emm*4 isolates 86
3.4.4 Bioinformatic features of BSAC *emm*4 prophage 87
3.4.5 Bioinformatic features of MGAS10750 and MEW427 SpyCI 91
3.4.6 $M4_{degraded}$ lineage-specific SNPs 92
3.4.7 Additional genetic characteristics of $M4_{complete}$ and $M4_{degraded}$ 97
3.4.8 Non-canonical prophage among *emm*4 isolates resolved by long-read sequencing 98
3.5 Discussion 99
3.5.1 Chapter 3: Key findings 104
4. Prophage induction and gene expression 105
4.1 Introduction 105
4.2 Chapter hypotheses and aims 106
4.3 Methods 106
4.3.1 Prophage and SpyCI induction using mitomycin C 106
4.3.2 Spontaneous prophage and SpyCI induction experiments 107
4.3.3 Detection of gene expression and copy number by RT-qPCR and qPCR 107
4.3.4 Semi-quantitative detection of *mutS* and *mutL* transcripts 107
4.4 Results 108
4.4.1 Prophage of $M4_{complete}$ are inducible but prophage of $M4_{degraded}$ are not 108
4.4.2 Prophage induction during *S. pyogenes* growth in THB 114
4.4.3 Prophage toxin gene expression during bacterial growth 119
4.4.4 SpyCIM4 of $M4_{complete}$ and $M4_{degraded}$ are not inducible 122
4.5 Discussion 124
4.5.1 Chapter 4: Key findings 127
5. Phenotypic characterisation of *emm*4 isolates 128
5.1 Introduction 128
5.2 Chapter hypotheses and aims 128
5.3 Methods 129
5.3.1 Lancefield assay 129
5.3.2 Superantigenicity assay 129
5.3.3 *Galleria mellonella* virulence assay 130

5.3.4 Gram's crystal violet adhesion assay....130
5.4 Results....130
5.4.1 Growth in THB....130
5.4.2 Lancefield assay: survival of *emm*4 in whole human blood....131
5.4.3 Superantigenicity of *emm*4 culture supernatants....132
5.4.4 *Galleria mellonella* virulence assay with *emm*4 isolates....133
5.4.5 Biofilm formation and adhesion of *emm*4 isolates....135
5.5 Discussion....136
5.5.1 Chapter 5: Key findings....139
6. Modulation of superantigen expression by HTE co-culture....140
6.1 Introduction....140
6.2 Chapter hypotheses and aims....141
6.3 Methods....141
6.3.1 Storage and recovery of HTE cells and A549s....141
6.3.2 *S. pyogenes* co-culture with HTE cells and A549s....142
6.4 Results....142
6.4.1 Growth of *S. pyogenes* in RPMI +10% (v/v) FCS and following HTE co-culture....142
6.4.2 Human tonsillar epithelium does not enhance prophage induction....143
6.4.3 Modulation of superantigen expression following HTE co-culture....145
6.4.4 Expression of *ssa* was also enhanced by A549 cells....148
6.4.5 Enhanced *ssa* expression was not associated with an increase in gene copies....149
6.4.6 Expression of *ssa* is enhanced by human tonsillar epithelium among other genotypes....151
6.5 Discussion....152
6.5.1 Chapter 6: Key findings....154
7. Discussion....156
7.1 Genotype *emm*4 prophage....156
7.1.1 Two clades within clinical *emm*4 isolates....156
7.1.2 Prophage degradation within MEW427 clade....156
7.1.3 $M4_{degraded}$ prophage are cryptic....157
7.1.4 A novel international *emm*4 lineage characterised by cryptic prophage....157
7.2 Redundancy of prophage mobility regulation among *emm*4 *S. pyogenes*....158
7.2.1 Prophage induction is not associated with enhanced toxin transcription....158
7.2.2 Genotype *emm*4 DNA MMR operon is not regulated by $SpyCI^{M4}$....158

7.3 Genotype *emm*4 at the host-pathogen interface 159
7.3.1 Phenotypic similarities between $M4_{complete}$ and $M4_{degraded}$ 159
7.3.2 Human tonsillar epithelium modulates superantigen expression 160
7.4 Bacterial regulation of prophage induction 162
7.4.1 Possible role for CovRS in prophage replication and toxin expression 162
7.5 Conclusions 162
7.1 Future work 163
8. Appendix 166
9. References 167

List of figures

Figure 1.1: Oropharyngeal features of streptococcal tonsillopharyngitis....17
Figure 1.2: Cellulitis and lymphangitis....19
Figure 1.3: *S. pyogenes* and cryptogenic necrotising fasciitis: pathogenesis....20
Figure 1.4: Selected transmission and clinical manifestations of *S. pyogenes*....23
Figure 1.5: Diversity within and between *emm* genotypes....25
Figure 1.6: Genomic architecture of *S. pyogenes* lysogenic prophage....27
Figure 1.7: Prophage integration and excision mediated by integrase and excisionase....28
Figure 1.8: The SOS-response....29
Figure 1.9: Genomic architecture of SpyCI....35
Figure 1.10: Complement pathway....41
Figure 1.11: M Protein: selected mechanisms of action....44
Figure 1.12: Variable expression of hyaluronic acid capsule by *S. pyogenes*....46
Figure 1.13: Serum opacification by SOF....48
Figure 1.14: Differential, concentration-dependent binding of IgG by M1....50
Figure 1.15: Neutrophil diapedesis and margination is abrogated by SpyCEP....52
Figure 1.16: Streptococcal DNases promote expedition of bacteria ensnared in NETs....54
Figure 1.17: Streptococcal superantigens: selected mechanisms of action....55
Figure 1.18: Transcriptional regulation by CovRS....57
Figure 1.19: Enhanced nasopharyngeal infection associated with epidemic *emm*3 lineage....60
Figure 1.20: SpyCI mediated regulation of gene expression in *S. pyogenes*....62
Figure 1.21: Enhanced virulence associated with the insertion of a single nucleotide....64
Figure 2.1: Schematic of prophage conformation PCR....75
Figure 2.2: Schematic of the plasmids constructed for qPCR quantification....78
Figure 2.3: Target gene qPCR cycling parameters....79
Figure 3.1. Putative regions of recombination within MEW427-like BSAC isolates....82
Figure 3.2: Comparison of *emm*4 reference genomes MGAS10750 and MEW427....85
Figure 3.3: Comparison of gene content between MGAS10750 and MEW427 prophage....86
Figure 3.4 Variable factors present in BSAC *emm*4 isolates....87
Figure 3.5: Bioinformatic comparison of prophage-encoding regions for *emm*4 isolates....90
Figure 3.6.: Comparison of gene content between SpyCIM4 of MGAS10750 and MEW427....91
Figure 3.7: Bioinformatic comparison of BSAC SpyCIM4....92
Figure 3.8: Additional non-canonical prophage identified by long-read sequencing....99
Figure 3.9: International *emm*4 lineage associated with cryptic prophage....102
Figure 4.1. Φ10750.1/Φ427.1 induction in liquid culture using mitomycin C....109
Figure 4.2.: Φ10750.2/Φ427.2 induction in liquid culture using mitomycin C....111

Figure 4.3. Φ10750.3/Φ427.3 induction in liquid culture using mitomycin C 113
Figure 4.4.: Spontaneous Φ10750.1/Φ427.1 induction during bacterial growth..... 115
Figure 4.5.: Spontaneous Φ10750.2/Φ427.2 induction during bacterial growth..... 117
Figure 4.6.: Spontaneous Φ10750.3/Φ427.3 induction during bacterial growth..... 118
Figure 4.7.: Expression of *speC* and gene copies during growth in THB 120
Figure 4.8: Expression of *spd3* during growth in THB 121
Figure 4.9.: Expression of *ssa* during growth in THB..... 122
Figure 4.10: SpyCI^{M4} induction with mitomycin C and during growth in THB 123
Figure 4.11: *mutS* and *mutL* expression during bacterial growth 124
Figure 5.1. Growth characteristics of *emm*4 isolates..... 131
Figure 5.2. Survival of *emm*4 isolates in whole human blood..... 132
Figure 5.3. Superantigenic activity of *emm*4 isolates 133
Figure 5.4: *Galleria mellonella* virulence assay..... 134
Figure 5.5: Adhesion of *emm*4 isolates to polystyrene..... 135
Figure 5.6: Adhesion of *emm*4 isolates to eukaryotic substrata 136
Figure 6.1: Differential growth of *S. pyogenes* in Green's versus RPMI +10% (v/v) FCS 143
Figure 6.2: Prophage induction following co-culture with human tonsillar epithelium..... 144
Figure 6.3. Expression of *speC* following co-culture with human tonsillar epithelium 146
Figure 6.4: Expression of *ssa* following co-culture with human tonsillar epithelium..... 147
Figure 6.5.: Expression of *ssa* by *emm*4 isolates following co-culture with A549s..... 148
Figure 6.6: Copies of *ssa* gene following co-culture with human tonsillar epithelium..... 150
Figure 6.7: Human tonsillar epithelium enhances *ssa* expression in other genotypes..... 152
Figure 7.1: Homology between toxigenic prophage of *emm*3, *emm*4 and *emm*12 161
Figure 7.2: Graphical hypotheses: M4$_{complete}$ vs. M4$_{degraded}$ 165

List of tables

Table 1.1: *S. pyogenes* isolates used in this study 66

Table 1.3: Primers used in this study.....72

Table 3.1: Non-synonymous $M4_{degraded}$ lineage-specific SNPs.....93

Table 3.2: Synonymous $M4_{degraded}$ lineage-specific SNPs 95

Table 3.3: Intergenic $M4_{degraded}$ lineage-specific SNPs.....96

Abbreviations

Φ	prophage
µg	microgram
µL	microliter
µM	micromolar
aa	amino acid
attB	bacterial chromosome attachment site
attP	prophage attachment site
bp	base pair
BrdU	bromodeoxyuridine
BSAC	British Society for Antimicrobial Chemotherapy
CBA	Columbia blood agar
cDNA	complimentary DNA
CDS	coding sequence
CFU	colony forming units
CovRS	Control of virulence response regulator/sensor kinase
CRISPR	clustered regularly interspaced palindromic repeats
dsDNA	double-stranded DNA
DMEM	Dulbecco's modified Eagles' medium
DNase	deoxyribonuclease
ELISA	enzyme linked immunosorbent assay
ECM	extracellular matrix
FCS	foetal calf serum
HA	hyaluronic acid
HBSS	Hanks' balanced salt solution
HTE	Human tonsillar epithelium
HylP	prophage-encoded hyaluronidase
ICE	integrative and conjugative element
IFN	interferon
IgA	immunoglobulin A
IgG	immunoglobulin G
IgM	immunoglobulin M
IL	interleukin
kDa	kilodalton
LB	Luria Bertani/lysogeny broth

mg	milligram
MGE	mobile genetic element
MGW	molecular grade water
mL	millilitre
mM	millimolar
MMR	mismatch-repair
MOI	multiplicity of infection
NET	neutrophil extracellular trap
ng	nanogram
OD_{600}	optical density/absorbance at 600nm
ORF	open reading frame
PAMP	pathogen-associated molecular pattern
PBS	phosphate buffered saline
PRR	pattern recognition receptor
PSGN	post-streptococcal glomerulonephritis
qPCR	quantitative PCR
RD	region of difference
RNase	ribonuclease
RPMI	Roswell Park Memorial Institute media
RT	reverse transcription
RT-qPCR	reverse transcription real-time quantitative PCR
SAPI	staphylococcal pathogenicity island
sbsp.	subspecies
ssDNA	single-stranded DNA
SNP	single nucleotide polymorphism
SPIF	soluble phage inducing factor
SpyCI	*Streptococcus pyogenes* chromosomal island
STSS	streptococcal toxic shock
TCR	T-cell receptor
TCS	two-component regulatory system
THA	Todd-Hewitt agar
THB	Todd-Hewitt broth
TLR	toll-like receptor
U/mL	units per millilitre
w/v	weight per volume
w/w	weight per weight
WGS	whole genome sequencing

Chapter 1

1. Introduction

1.1 *Streptococcus pyogenes*

S. pyogenes is a major Gram-positive pathogen of humans (Carapetis *et al.*, 2005). The bacterium, also known as the Lancefield Group A *Streptococcus,* carries the Lancefield Group A carbohydrate, but *S. pyogenes* is closely related to the Group C and G streptococci, particularly the equine pathogen *S. equi* sbsp. *equi*, with which it shares ~80% DNA sequence similarity (Holden *et al.*, 2009). The zoonotic pathogen *S. equi* sbsp. *zooepidemicus* is thought to be ancestral to *S. equi* sbsp. *equi* and also shares ~80% DNA sequence similarity with *S. pyogenes*. *S. pyogenes* also shares significant DNA sequence similarity with *S. dysgalactiae* sbsp. *equisimilis*, which carries the Group C/G carbohydrates, and occasionally the Group A carbohydrate (Brandt and Spellerberg, 2009; Chochua *et al.*, 2019). Indeed, clinical manifestations of *S. dysgalactiae* sbsp. *equisimilis* infection are not dissimilar to those observed in *S. pyogenes* (Brandt and Spellerberg, 2009). Other clinically significant streptococcal species appear to have contributed to the evolution of *S. pyogenes*. Most notably the canine pathogen *S. canis,* and the human pathogen, the Group B *Streptococcus, S. agalactiae* (Lefébure *et al.*, 2012). There is evidence of a shared pool of mobile genetic elements between *S. pyogenes* and a number of other pyogenic streptococcal species, and indeed the wider *Streptococcaceae* (Canchaya *et al.*, 2003; Desiere *et al.*, 2001; Holden *et al.*, 2009; Lefébure *et al.*, 2012).

1.2 Clinical manifestations of *S. pyogenes*

1.2.1 Tonsillopharyngitis

The primary anatomical niche of *S. pyogenes* is the tonsil, and tonsillopharyngitis is one of the most common manifestations of infection, with approximately 616 million incident cases per annum (Walker *et al.*, 2014). Symptoms include fever, cervical lymphadenopathy and sore throat with patchy purulent exudate (Carapetis *et al.*, 2005) (Figure 1.1). The role of specific virulence factors in the pathogenesis of streptococcal pharyngitis has not been fully elucidated, though a number of studies have indicated a role for prophage-encoded toxins, notably DNases and streptococcal superantigens, in colonisation, propagation and transmission of infection (Al-Shahib *et al.*, 2016; Davies *et al.*, 2019; Kasper *et al.*, 2014; Zeppa *et al.*, 2017). Pharyngitis is typically self-limiting, though if left untreated, may lead to potentially life-threatening complications, notably acute rheumatic fever and rheumatic heart disease (Cunningham, 2000).

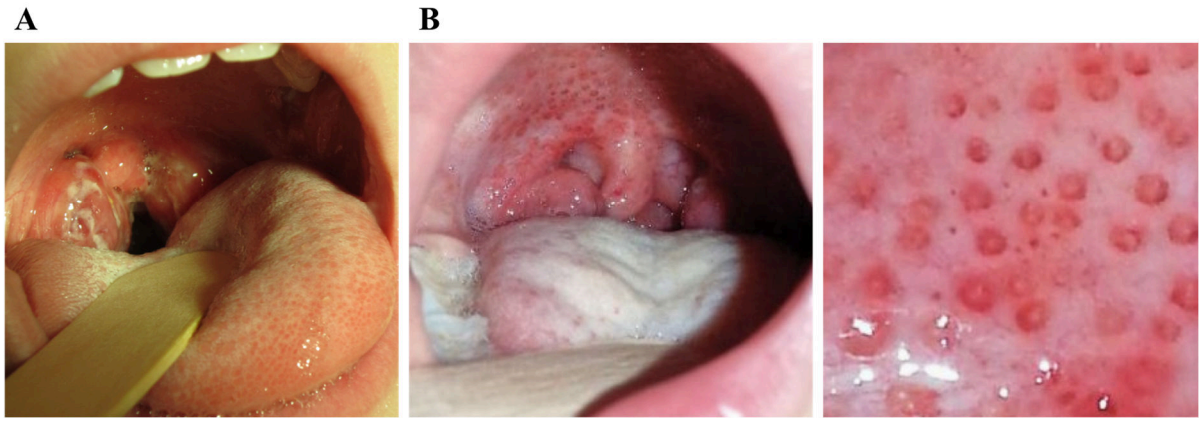

Figure 1.1: Oropharyngeal features of streptococcal tonsillopharyngitis

A) *S. pyogenes* infection of the tonsil is associated with marked inflammation and pain, often with the presence of white patchy, or streaked purulent exudates. B) Erythematous papular palatal petechiae with pale foci (so called 'doughnut lesions'), are occasionally observed, and though not always present, can be pathognomonic (Kobayashi and Iwasaki, 2019). "Strep throat2010.jpg" by James Heilman, MD is licensed under CC BY-SA 3.0.

1.2.2 Scarlet fever

Scarlet fever is one of the classical childhood exanthems, principally transmitted by infectious droplets, and highly contagious. Symptoms of scarlet fever include malaise, fever, cervical lymphadenopathy, accompanying tonsillitis and a pathognomonic finely papular red rash on the face and body, which may desquamate (Basetti *et al.*, 2017). Inflammation and swelling of the tongue (so called 'strawberry tongue') is also typical (Basetti *et al.*, 2017). Scarlet fever is believed to be a toxin-mediated disease, notably by streptococcal superantigens associated with prophage, particularly *ssa*, *speC* and *speA*, indeed, a number of epidemiological and molecular investigations have revealed an association between the genes encoding these toxins and scarlet fever (Chalker *et al.*, 2017; Lynskey *et al.*, 2019a; Silva-Costa *et al.*, 2014; Turner *et al.*, 2016; Walker, 2015). It is thought that these potent factors are secreted at a focus of infection, typically (though not exclusively) the tonsil and spread throughout the body triggering massive expansion of specific T-cell subsets (Proft and Fraser, 2016). The streptococcal superantigens have been shown to play a role in colonisation and infection at these sites (Brouwer *et al.*, 2020; Kasper *et al.*, 2014; Zeppa *et al.*, 2017). With the discovery of antibiotics, and possibly other unknown factors, the incidence of scarlet fever gradually dwindled, as did the severity of disease (Carapetis *et al.*, 2005; Chalker *et al.*, 2017). However, in recent years there has been a dramatic resurgence of this disease in some countries and there have been a number of well-studied and ongoing scarlet fever outbreaks, particularly in the United Kingdom, Hong Kong and mainland China (Brockmann *et al.*, 2018; Chalker *et al.*, 2017; Lynskey *et al.*, 2019a; Walker, 2015). Genotype *emm*4 *S. pyogenes* are often associated with scarlet fever and paediatric infection in many parts of the world and are routinely recovered from throat swabs (Kim *et al.*, 2019; Silva-Costa *et al.*, 2014; Turner *et al.*, 2016; Whitehead *et al.*, 2011).

1.2.3 **Impetigo**

Non-bullous impetigo is another common manifestation of *S. pyogenes* infection, affecting the outermost layer of the integumentary system, the stratum corneum (Stevens and Bryant, 2016). Most often, infection begins with a well-defined papule which progresses to a vesicular rash, surrounded by erythema. Gradually, these lesions form pustules filled with clear liquid which rupture and form a characteristic crust. These lesions may be observed anywhere on the body, but are common on the face and lower extremities (Pereira, 2014). In some parts of the world, particularly parts of Africa and the tropics, impetigo is a seasonal disease (Armitage *et al.*, 2019).

1.2.4 **Erysipelas**

Erysipelas is the radiating infection of the dermis and often efferent lymphatics, characterised by a red, raised rash with clear demarcation between involved and uninvolved tissues, which are most often lower extremities, or more rarely the skin on the face (Stevens and Bryant, 2016). When the lower extremities are affected, infection commonly follows minor damage to the integrity of the skin, such as burns or mild trauma. In facial erysipelas, there is often history of streptococcal pharyngitis (Cunningham, 2000). Erysipelas is perhaps the most superficial manifestation of *S. pyogenes* infection, and prognosis following antibiotic therapy is excellent, with complications, such as the involvement of underlying tissues, being rare (Stevens and Bryant, 2016).

1.2.5 **Cellulitis**

Streptococcal cellulitis refers to the rapidly spreading inflammation of the skin and subcutaneous tissues, which may be accompanied by systemic symptoms such as fever and malaise, in addition to bacteraemia and lymphangitis (Stevens and Bryant, 2016). Although common, if untreated, cellulitis may progress to invasive disease in the surrounding and underlying tissues, and in at least one study has actually been associated with the highest risk of death versus other manifestations (Lamagni *et al.*, 2009). Cellulitis, or indeed other superficial or relatively minor infections of the skin by *S. pyogenes*, may be accompanied by lymphangitis, which presents as red, tender, linear channels following the efferent lymphatics from the focus of infection and leading to regional lymph nodes (Vindenes and McQuillen, 2015) (Figure 1.2).

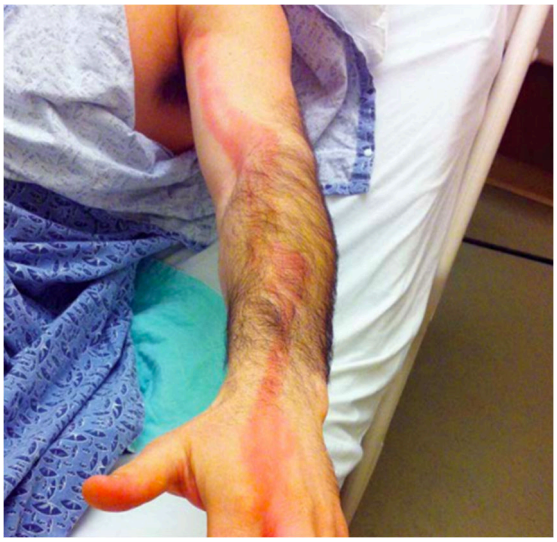

Figure 1.2: Cellulitis and lymphangitis

Skin and soft tissue infections of caused by *S. pyogenes* may be associated with lymphangitis. In this case, cellulitis of the third finger, was accompanied by lymphangitis, seen as a well demarcated linear streak from the focus of infection, toward regional lymph nodes. Reproduced with permission from Vindenes and McQuillen, 2015, Copyright Massachusetts Medical Society.

1.2.6 Necrotising fasciitis

Necrotising fasciitis is a severe, fulminant and highly invasive manifestation of *S. pyogenes* infection involving the subcutaneous tissues, notably the fascia and fat. Extreme pain, profuse swelling, ecchymoses, and fluid-filled, violaceous bullae are typical, with a grey 'dishwater' exudate, which may be accompanied by systemic symptoms (Bisno and Stevens, 1996; Stevens and Bryant, 2017). While necrotising fasciitis may result from direct introduction of *S. pyogenes* by penetrating injury (Hagberg *et al.*, 1997; Sablier *et al.*, 2010), or as a complication of an existing skin and soft tissue infection, often, infection is cryptogenic with no obvious cause (Stevens and Bryant, 2017). There is some evidence that blunt force trauma, bruising or muscle strain may precipitate infection (Lamb *et al.*, 2018a; Nuwayhid *et al.*, 2007), possibly due to an antecedent episode of superficial infection, such as tonsillitis, scarlet fever, or asymptomatic carriage, either in the patient or a close contact (Watts *et al.*, 2019) (Figure 1.3). The link between superficial manifestations and necrotising fasciitis are not completely clear, but recent evidence suggests that these may be linked by metastases of *S. pyogenes* from infectious foci at sites such as the tonsil via. the lymphatics, whereby the bacterium gains access to the circulation, and beyond (Siggins *et al.*, 2020).

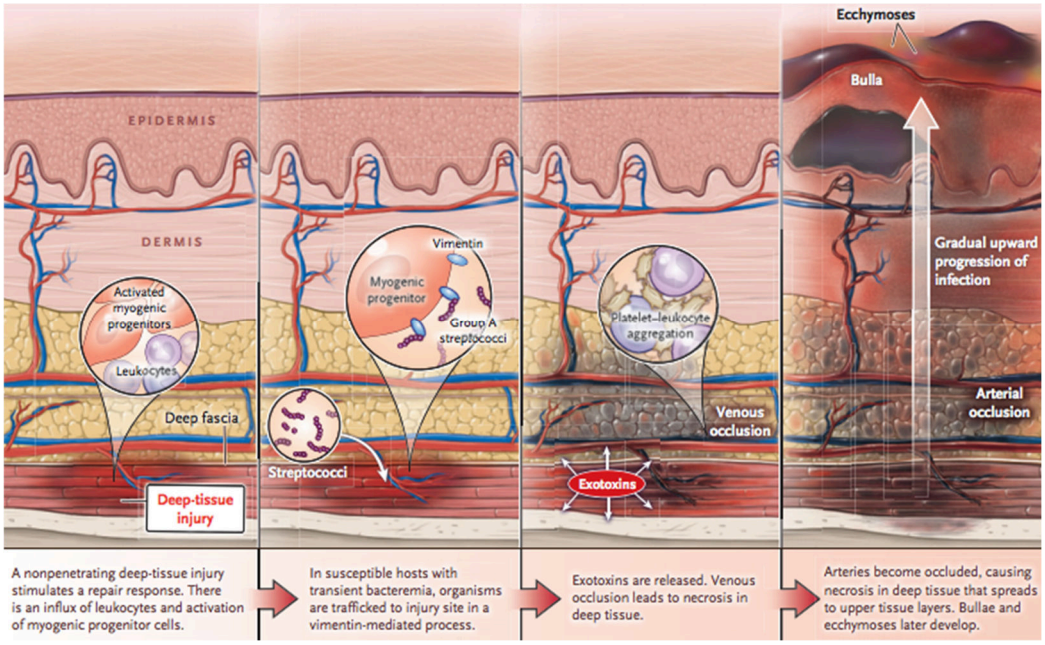

Figure 1.3: *S. pyogenes* and cryptogenic necrotising fasciitis: pathogenesis

Necrotising fasciitis caused by *S. pyogenes* may have no obvious cause, but can be associated with blunt force trauma and bruising. These (sometimes intangibly minor) traumas set in motion molecular cascades that appear to attract streptococci present in the circulation. The subsequent, progressive and devastating destruction of the deep tissues is wrought not by the propensity for *S. pyogenes* (the so-called 'flesh-eating') to consume its host, but rather the production of myriad, highly potent exotoxins by *S. pyogenes* colonising the subcutaneous tissues in necrotising fasciitis. Reproduced with permission from (Stevens and Bryant, 2017), Copyright Massachusetts Medical Society.

1.2.7 Puerperal sepsis

Puerperal sepsis is defined by the World Health Organisation as an infection of the genitourinary tract at any stage between the rupture of membranes or the induction of labour and the 42nd day after delivery in which two or more of the following symptoms are present: fever, vaginal discharge, pelvic pain and/or delayed reduction of uterine size post-partum. Additionally, this may include infections elsewhere in the body during this time, such as haemorrhagic pneumonia (Turner *et al.*, 2013). It remains the leading cause of direct maternal death, and *S. pyogenes* is the most common aetiologic agent (Turner *et al.*, 2013). Genotype *emm*28 *S. pyogenes* are frequently associated with puerperal sepsis, possibly due to the acquisition of the RD2 ICE from the Group B *Streptococcus, S. agalactiae,* which encodes the putative adhesin AspA (Luca-Harari *et al.*, 2009; Sitkiewicz *et al.*, 2011; Zhang *et al.*, 2006). Despite this, mortality attributed to this disease is disproportionately attributable to genotypes *emm*1 and *emm*3 (Luca-Harari *et al.*, 2009). Further, the prophage-encoded streptococcal superantigen *speC* has been linked to this devastating manifestation (Vlaminckx *et al.*, 2003). Risk factors for postpartum sepsis include previous pharyngeal infection or asymptomatic carriage in the patient or a contact, such as family or a healthcare worker.

1.2.8 Streptococcal toxic shock syndrome

Streptococcal toxic shock syndrome (STSS) is a systemic and imminently lethal complication of *S. pyogenes* infection and has a strong association with high mortality and morbidity, with 81% of patients succumbing to illness within 7 days (Lamagni *et al.*, 2009; Walker *et al.*, 2014). STSS is ostensibly toxin-mediated, however, there is some debate over which toxins these might be, but a prevailing theory is that streptococcal superantigens are instrumental in the pathogenesis of this disease. This hypothesis holds that the production of streptococcal superantigens at a focus or foci of infection are released into the circulation and activate large numbers of antigen presenting cells and T-cells, which subsequently release huge quantities of cytokines, leading to systemic shock and multiorgan failure. Studies have implicated the potent chromosomally-encoded superantigen SMEZ and the prophage-encoded SpeA. Superantigenic activity, however, does not necessarily correlate with the genetic predisposition of a specific *emm*-type to cause STSS. Genotype *emm*3 *S. pyogenes,* for example, lack a functional *smeZ* allele and are poorly mitogenic despite their overrepresentation in cases of invasive disease and STSS (Lamagni *et al.*, 2009, 2008; Turner *et al.*, 2012).

1.2.9 Post-infectious immunological sequelae

1.2.9.1 Acute rheumatic fever and rheumatic heart disease

By far the most burdensome manifestation of *S. pyogenes* infection is rheumatic heart disease (Efstratiou and Lamagni, 2016). This syndrome typically follows streptococcal pharyngitis, and is characterised by arthritis and carditis, with or without the presence of neurological sequelae such as Sydenham's chorea (Lee *et al.* 2009). Although a fairly rare manifestation in most affluent countries, rheumatic heart disease is devastating in regions with limited access to healthcare infrastructure and antibiotics. It is estimated that between 15.6-19.6 million people globally live with rheumatic heart disease, and almost 80% of these are in regions of the world with low economic status (Carapetis *et al.*, 2005; Efstratiou and Lamagni, 2016). Acute rheumatic fever typically occurs between 4-8 weeks antecedent to an episode of streptococcal pharyngitis, and approximately 30% of cases of acute rheumatic fever will progress to rheumatic heart disease. The pathogenesis of acute rheumatic fever and rheumatic heart disease are unclear, but a number of mechanisms have been proposed. These include antibody cross-reactivity between the *S. pyogenes* M protein and proteins abundant in cardiac and skeletal muscles, notably myosin and tropomyosin (Cunningham, 2000). Another enduring hypothesis is that hyperencapsulation in specific *emm*-types may contribute to pathology (Smoot *et al.*, 2002b; Stollerman and Dale, 2008).

1.2.9.2 Acute post-streptococcal glomerulonephritis

Acute post-streptococcal glomerulonephritis (APSGN) typically follows an episode of skin or soft tissue infection, such as impetigo or pyoderma by approximately 6 weeks, or 1-2 weeks following pharyngeal infection. There may be a genetic predisposition to PSGN, as ~40% of patients with PSGN have a positive family history

(Rodriguez-Iturbe and Haas, 2016). However, this could conceivably be confounded by intra-familial transmission of the same isolate or clone, or indeed shared socioeconomic factors that might also play a role. Some 470,000 cases of PSGN occur per annum (Carapetis *et al.*, 2005), and is disproportionately burdensome in regions of low economic status, constituting the most common cause of paediatric renal injury in the Middle East, Africa and parts of Australasia (Rodriguez-Iturbe and Musser, 2008). Indeed, 97% of all reported PSGN cases are in regions of low economic growth. The precise pathologic mechanisms of PSGN are not presently known, though the deposition of antigen-antibody complexes in renal glomeruli and is a prevailing theory and similar propositions have been made by multiple lines of scientific inquiry (Cunningham, 2000).

1.2.10 **Asymptomatic carriage**

Asymptomatic carriage of *S. pyogenes* can occur within the oropharynx, the female genital tract, the rectum, and on the skin (Efstratiou and Lamagni, 2016; Walker *et al.*, 2014). Generally, a carrier is defined as a patient who is culture or rapid antigen detection test positive for *S. pyogenes* in the absence of symptoms, and often, where feasible, without evidence of anti-*S. pyogenes* antibodies on serial serological testing (Martin, 2016). Pharyngeal and rectal or vaginal asymptomatic carriage is rare in healthy adults, usually <5% and <1%, respectively (Efstratiou and Lamagni, 2016; Steer *et al.*, 2012). Carriage in the absence of clinical symptoms is more common in children, with a meta-analysis based on the systematic review of 29 studies indicating a pooled prevalence of 12% (95% confidence interval 9-14%) (Shaikh *et al.*, 2010). Other studies corroborate these data, placing asymptomatic carriage at between 15-20% (Martin, 2016). One of the few longitudinal studies on carriage in children deduced that over a four year period, 50% of children were carriers (Shet and Kaplan, 2004). Some studies place a high value on the burden of these patients in driving outbreaks of disease (Cockerill *et al.*, 1997; Participants, 2002; Weiss *et al.*, 1999), whereas others do not (Martin, 2016) (Figure 1.4).

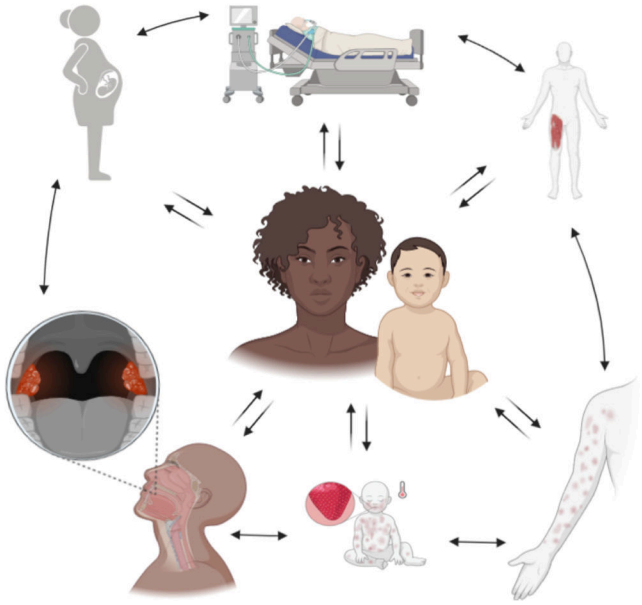

Figure 1.4: Selected transmission and clinical manifestations of *S. pyogenes*

Transmission of *S. pyogenes* can occur via. a number of routes, and can manifest a diversity of clinical manifestations, though not all exposures lead to clinical disease or asymptomatic carriage. (Carapetis *et al.*, 2005; Cunningham, 2000). The role of the asymptomatic carrier or subclinical patient is not well understood, and may be more important in children, but physical association or shared vicinity, either with asymptomatic carriers or patients experiencing a range of sequelae, even the most debilitating, remain potential contributories to the transmission of this highly adaptable and protean pathogen. Post-infectious immunological sequalae are not depicted, but are highly burdensome consequences of transmission and infection. Figure created by the author with BioRender.com.

1.3 Control and prevention of S. *pyogenes* infection

1.3.1 Treatment

Generally, uncomplicated and superficial manifestations of *S. pyogenes* infection, such as tonsillitis, are treated with phenoxymethyl penicillin, though another treatment option, particularly where penicillin may not be tolerated, are macrolide antibiotics such as erythromycin (Pelucchi *et al.*, 2012). To remediate more severe and invasive disease, such as necrotising fasciitis, antibiotics alone are insufficient and early surgical debridement of fascial planes is necessary (Machado, 2011; Wood *et al.*, 1993). In these cases, antibiotics such as clindamycin are administered intravenously, often in conjunction with passive immunotherapy in the form of intravenous immunoglobulin (IVIG) (Carapetis *et al.*, 2014; Raithatha and Bryden, 2012), which may reduce mortality and sequalae associated with toxic shock syndrome (Sriskandan *et al.*, 2006). Due to the increased risk of invasive disease among household contacts and neonates of invasive disease patients, swift antibiotic prophylaxis may be indicated to prevent secondary infection (Carapetis *et al.*, 2014; Mearkle *et al.*, 2017).

1.3.2 **Antimicrobial resistance**

Until very recently, it was thought that *S. pyogenes* resistance toward beta lactam antibiotics was vanishingly rare. However, reduced *in-vitro* susceptibility to these antimicrobials has been observed in numerous diverse *emm*-types owing to mutations within the *pbp2x* penicillin binding protein gene (Musser *et al.*, 2020; Vannice *et al.*, 2019). Although these mutations do not confer full resistance in *S. pyogenes*, similar mutations observed in *S. pneumoniae* forebode that attaining reduced susceptibility in this manner may represent a stepwise strategy in this direction, which is cause for concern (Grebe and Hakenbeck, 1996; Vannice *et al.*, 2019). Inasmuch as this transition toward manifest beta lactam resistance in *S. pyogenes* is, as yet, incomplete, treatment failure may still occur in some 35% of tonsillopharyngitis cases (Sela and Barzilai, 1999). It has been suggested that β-lactamase producing bacteria also present at non-sterile sites might protect *S. pyogenes* from elimination with these antibiotics (Brook and Gober, 2006; Pichichero and Casey, 2007). The ability of *S. pyogenes* to invade host cells, becoming internalised and shielded from antibiotics, particularly penicillin, is now well described (Kaplan *et al.*, 2006; Kreikemeyer *et al.*, 2004; Marouni *et al.*, 2004), and might contribute to treatment failure (Sela and Barzilai, 1999). Resistance to macrolide and tetracycline antibiotics among *S. pyogenes* is comparatively more common, though more so in some regions than others. In Hong Kong and mainland China in particular, the recent resurgence of scarlet fever has been driven, in part, by the acquisition of resistance determinates reducing the efficacy of macrolides (Tse *et al.*, 2012; Walker, 2015). In the United Kingdom, non-susceptibility to macrolides appears to be increasing, and as of 2020, has risen from ~4-6% in the previous five seasons, to ~10%. Reduced susceptibility to tetracycline and clindamycin also appears to be increasing (HPA, 2020). Reduced susceptibility to macrolide and tetracycline antibiotics is typically mediated by mobile genetic elements, and the contribution of these to resistance in *S. pyogenes* is discussed in greater detail in section 1.5.5.1.

1.3.3 **Vaccines**

No commercial vaccine is presently licensed for the prevention of *S. pygoenes* infection. Efforts to develop a safe and efficacious vaccine have been hampered, in large part, due to the extensive global genetic diversity and considerable dispensable gene pool of *S. pyogenes* (Davies *et al.*, 2019; Steer *et al.*, 2009). Concerns regarding crossover between *S. pyogenes* and human antigens theoretically triggering autoimmunity have also mired vaccine efforts (Massell *et al.*, 1969; Steer *et al.*, 2009; Steer *et al.*, 2016) While a number of candidate vaccines have been explored in animal models, arguably the most robustly studied is the 26-valent vaccine based upon the N-terminal sequence of 26 different genotypes (Hu *et al*., 2002; McNeil *et al.*, 2005). The 26-valent vaccine evoked a >4-fold increase in geometric mean antibody titres above baseline in healthy adults, and a significant reduction of bacterial counts in patient sera obtained post-immunisation. Reassuringly, no evidence of autoimmune sequalae was detected, nor were antibodies towards human tissues detected (McNeil *et al.*, 2005). A 30-valent M-protein based vaccine has also shown promise in randomised, controlled phase 1 study (Pastural *et al*., 2020).

1.4 The *S. pyogenes* genome and genotypes

Historically, *S. pyogenes* isolates were subclassified serologically according to reactivity to M protein, also a major virulence factor, expressed as fibrils at the cell surface. The N-terminal region of M is hypervariable and type-specific and as such, can be classified with specific antisera to discern a serotype (Lancefield, 1962). More recently, this system was ratified and expanded by sequencing, wherein the 5' hypervariable sequence of *emm* is sequenced, with an *emm* genotype assigned according to >95% sequence homology with an existing *emm* specific sequence (CDC, 2008; (Efstratiou, 2000). Within *emm* types, diversity between isolates is relatively low, whereas inter-*emm* type diversity is comparatively high (Figure 1.5) and often attributable predominantly to prophage and mobile genetic elements (Beres *et al.*, 2002; Ferretti *et al.*, 2001; Smoot *et al.*, 2002a; Turner *et al.*, 2019b).

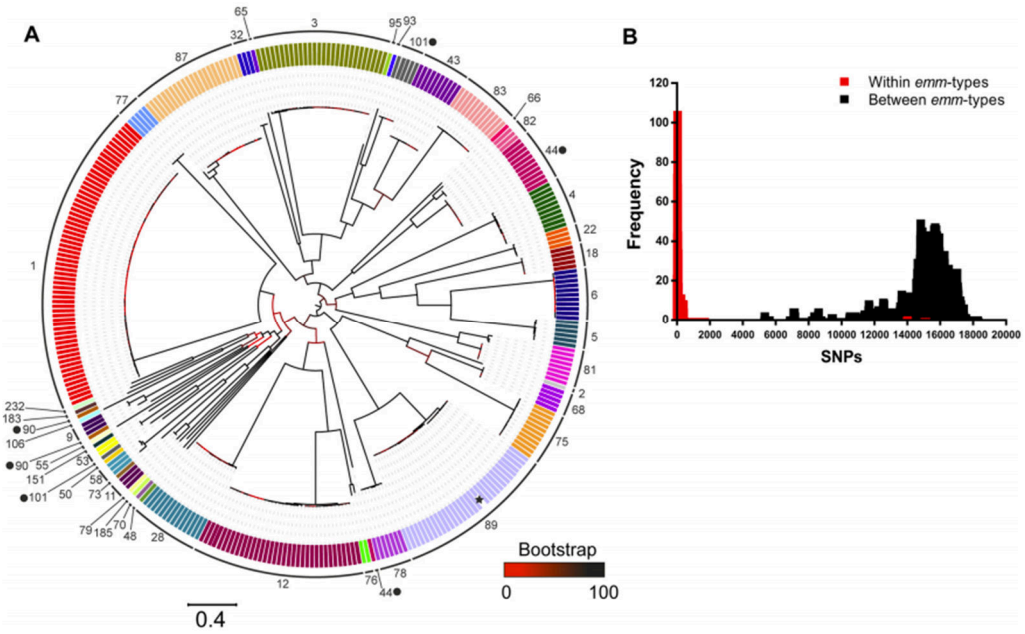

Figure 1.5: Diversity within and between *emm* genotypes

A) *S. pyogenes* isolates predominantly cluster by *emm* genotype, with little diversity observed within genotypes. Conversely, diversity is observed when isolates from different *emm* genotypes are compared, (including the majority of isolates in the present thesis). Maximum likelihood phylogenetic tree constructed from 113,805 core SNPs extracted after genome sequence data from all isolates to *emm*89 reference strain H293 (star). Notably, *emm*44, *emm*90 and *emm*101 did not, and are denoted by two separate lineages (black dots). Branches are shaded according to bootstrapping support scale. *S. pyogenes* genotypes are denoted by coloured boxes corresponding genotype number. B) The number of SNPs both within and between *emm* genotypes reflects similarity inferred by the phylogenetic tree, with a high number of SNPs between *emm* types (black bars), but a much lower number of SNPs when genomes of isolates belonging to the same *emm* genotype were compared (red bars). (Image taken, with permission, from Turner *et al.*, 2019).

Genotyping is an indispensable tool for epidemiological surveillance and has bolstered current understandings of the population biology of *S. pyogenes*. Indeed, non-random associations between *emm*-types and specific disease

manifestations are observed in many parts of the world, as are fairly predictable patterns of *emm*-type dominance. Though this is not always observed, and a number of regions (particularly those in whom the burden of *S. pyogenes* disease is especially high) comprise large and diverse populations of *emm*-types that are often uncommon elsewhere, with little of the genotype competition and dominance observed in much of the rest of the world. This poses obvious problems for efforts to develop an *S. pyogenes* vaccine which hitherto has focused principally on conferring immunity to the *emm* protein of globally prevalent *emm*-types.

1.5 Mobile genetic elements and *S. pyogenes*

The *S. pyogenes* metagenome comprises multiple functional, satellite and cryptic prophage, in addition to integrative and conjugative elements (ICE) and chromosomal islands (SpyCI), or evidence thereof (Beres and Musser, 2007; Panchaud *et al.*, 2009; Rezaei Javan *et al.*, 2019). Although it appears that *S. pyogenes* may be able to undergo transformation in biofilm, and possibly when adherent to nasopharyngeal-associated lymphoid tissue, there is a remarkable paucity of mechanistic evidence for transformation in *S. pyogenes* (Marks *et al.*, 2014a). Thus, a tempting and prevailing hypothesis is that mobile genetic elements (MGE) such as prophage, ICE and SpyCI are the principal vehicles of genetic variation in this organism, and indeed its close relatives (Holden *et al.*, 2009). The majority of diversity between *emm*-types is attributable to prophage and other mobile genetic elements, with most (but not all) isolates lysogenised with at least one prophage, and some containing as many as eight elements integrated into the bacterial chromosome (McShan and Nguyen, 2019).

1.5.1 Lysogenic bacteriophage

S. pyogenes has an intimate and longstanding evolutionary relationship with lysogenic prophage. The majority of these prophage, like many other streptococcal phages, are viruses assigned to the large *Caudovirales* order, with the vast majority belonging to the family *Siphoviridae*. These are tailed, double-stranded DNA viruses with circular genomes, and a similar genomic structure to coliphage Lambda, with contiguous genes comprising discernible modules dedicated to specific functions, in a roughly predictable arrangement (Canchaya *et al.*, 2003; McShan and Nguyen, 2019). Comparative genomics has revealed considerable homology between prophage of *S. pyogenes* and those of other pathogenic streptococci of clinical and veterinary importance, and interestingly, non-pathogenic streptococci (Desiere *et al.*, 2001; Holden *et al.*, 2009). It also seems that the prophage of streptococci conform to the modular evolution model, proposed in the early 1980's as a mechanism by which genetic diversity is achieved in bacterial viruses, with selection apparently acting on functional modules, as opposed to single genes (Botstein, 1980; Desiere *et al.*, 2001; McShan and Nguyen, 2019).

1.5.2 General genomic features of streptococcal phage

The lysogenic phage of *S. pyogenes* conform to a typically lambdoid genomic arrangement, with largely distinct modules dedicated to performing specific functions. Typically, this includes modules dedicated to replication, regulation, lysogeny and lysogenic conversion, capsid and tail fibre assembly, and host cell lysis (Figure 1.6) (Botstein, 1980; Canchaya *et al.*, 2003; Desiere *et al.*, 2001; McShan and Nguyen, 2019).

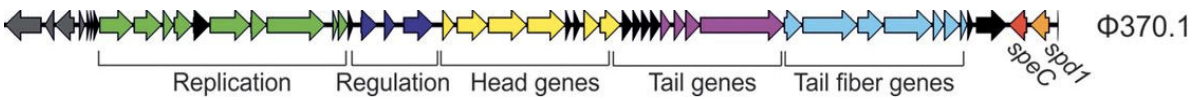

Figure 1.6: Genomic architecture of *S. pyogenes* lysogenic prophage

Lysogenic prophage of *S. pyogenes* genomes exhibit similarities in genetic organisation to prototype lysogenic coliphage lambda (Canchaya *et al.*, 2003). Discrete modules comprising genes dedicated to specific functions, such as replication (green), regulation (dark purple), the formation of progeny virions (yellow; lilac; pale blue), and lysogenic conversion are frequently observed, but are not always obvious. Toxin genes, such as superantigens (*speC*; red) and DNases (*spd1*; orange) are typically situated at distal ends of integrated viruses, though are not always present. Unclassifiable genes (black) are common. This genomic arrangement is exemplified in SF370 prophage Φ370.1 (Canchaya *et al.*, 2002). Figure taken and adapted, with permission, from Dr. Claire. E. Turner Turner *et al.*, 2015).

1.5.2.1 Prophage integration and excision

Like other lysogenic prophage, in *S. pyogenes,* these elements integrate into the bacterial chromosome by site-specific recombination between homologous sequences in the bacterial (*attB*) and phage (*attP*) genomes, which vary in length from as little as 5 bp, to in excess of 100 bp. This process is mediated by prophage-encoded integrase and excisionase proteins, usually, though not exclusively, in conjunction with repressor and antirepressor genes, akin to those found in coliphage Lambda (Figure 1.7). Attachment sites are well conserved between *emm*-types, and there are approximately 20 such sites in the core genome (McShan and Nguyen, 2019). Most often, *attB* sites are found on the lagging strand (3'), relative to the origin of replication, and within open reading frames (ORF) and might not be expected to modulate expression of these genes. To a lesser extent, *attB* sequences also include intergenic regions, and the 5' strand, in which case integration of prophage and other MGE may impact the expression of the target gene by physical separation from native promoters (Scott *et al.*, 2012, 2008).

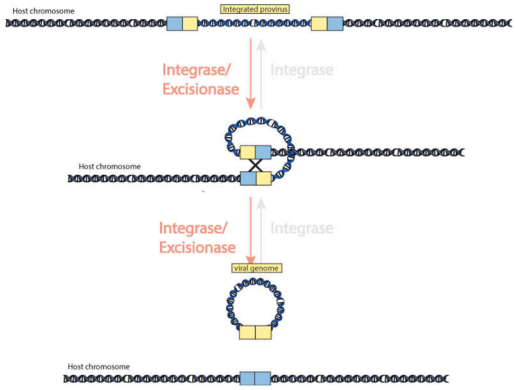

Figure 1.7: Prophage integration and excision mediated by integrase and excisionase

Prophage retain their ability by using site-specific integrase and excisionase genes, which permit the prophage to both exit and enter the bacterial genome at specific attachment site(s). "Viral genome excision" by ViralZone is licensed under CC BY 4.0.

1.5.2.2 Prophage induction

Although prophage induction may occur spontaneously at relatively low levels (Nanda *et al.*, 2015), in a number of bacterial pathogens, prophage excision is controlled, in part, by the SOS response (Oppenheim *et al.*, 2005) (Figure 1.8). Ordinarily, the SOS-response is triggered following DNA damage by oxidative stress and exposure to other agents such as ultraviolet light and antibiotics (Fernández De Henestrosa *et al.*, 2000; Maiques *et al.*, 2006; Michel, 2005; Oppenheim *et al.*, 2005; Zhang *et al.*, 2000). The conserved inducer RecA is subsequently activated, which binds single-stranded DNA, thus forming a nucleoprotein filament, promoting proteolysis of the repressor, LexA and the expression of SOS-response genes. Cleavage of LexA results in the de-repression of the SOS regulon, triggering the response. The lambdoid repressor cI is subject to autocatalytic cleavage following SOS-response initiation, similarly to LexA. Once cI is cleaved, repression by cI is relieved, and genes that mediate excision from the bacterial chromosome, particle assembly and egress are expressed, inducing the lytic cycle. Thus, Lambdoid phage are sensitive to stimuli in the extracellular milieu and inextricably linked to the condition of the bacterial host (Oppenheim *et al.*, 2005). Although this full complement of genes required to initiate the SOS-response are thought to be absent from *S. pyogenes* and other streptococcal species, analogous mechanisms have been described (Boutry *et al.*, 2013; Varhimo *et al.*, 2007; Žgur-Bertok, 2013).

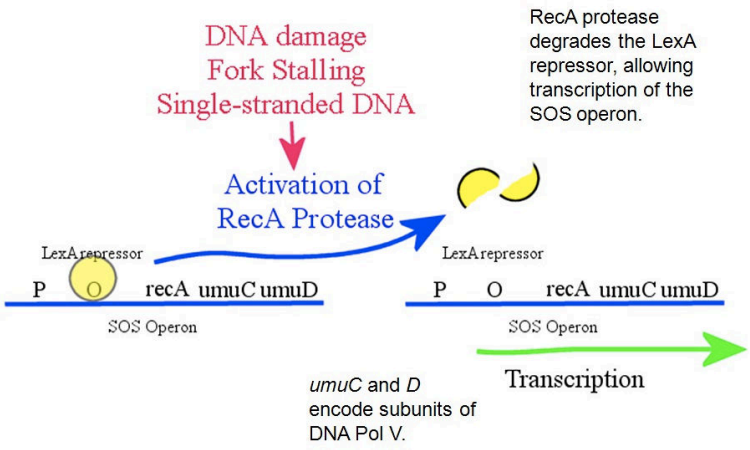

Figure 1.8: The SOS-response

The bacterial SOS-response is triggered by DNA damage, and is intimately linked to the maintenance of lysogeny and induction of lysogenic prophage by the repressor LexA. "SOS response" by Allen Gathman is licensed under CC BY-NC-SA 2.0

1.5.3 Satellite and cryptic prophage

As has been observed in other clinically relevant bacteria, occasionally, prophage elements are found to have decayed, becoming immobile, stably integrated and more permanent features of the bacterial chromosome (Desiere *et al.*, 2001; Rezaei Javan *et al.*, 2019). Satellite prophage are prophage which have lost the necessary components for independent replication, and are reliant on other prophage, and indeed their bacterial hosts for survival (Rezaei Javan *et al.*, 2019). By contrast, cryptic prophages are those which have degraded or decayed, existing as stably integrated elements. Although in some cases they may excise, cryptic prophage do not affiliate with helper phage, do not form infectious particles and do not lyse the host cell (Canchaya *et al.*, 2003). It is thought that these elements are of fundamental importance to the evolution of many bacterial pathogens, and certainly, there is evidence for this in *S. pyogenes* (Banks *et al.*, 2002; Stephen B. Beres *et al.*, 2004; Kaneko *et al.*, 1998; Panchaud *et al.*, 2009).

Defective and decayed prophage in *S. pyogenes* have been explicitly described previously in *emm*1 strain SF370 (Canchaya *et al.*, 2002; Scott *et al.*, 2012), some *emm*4 isolates (Henningham *et al.*, 2014), *emm*5 strain Manfredo (Canchaya *et al.*, 2003), *emm*28 strain MGAS6180 (Green *et al.*, 2005) and some *emm*89 strains (Turner *et al.*, 2015). It seems likely that this is may not be representative of the true frequency of cryptic and satellite prophage in *S. pyogenes* genomes, the detection of which until relatively recently has been constrained by the inability of short-read sequencing to resolve repetitive and paralogous sequences that are typical of prophage encoding regions. In the same vein, many phage genes and proteins remain uncharacterised and undescribed, however, as these are scrutinised more closely, additional layers of both complexity and understanding prophage and *S. pyogenes* are expected.

A recent genomic survey of 50 *S. pyogenes* genomes from a variety of *emm*-types identified 44 unique putative satellite phage, with an average size of ~12.8 kb and an average of 23 genes (Rezaei Javan *et al.*, 2019). Through acting as molecular 'parasites of parasites', interfering with the replication of functional phage elements, satellite prophage may actually benefit the host bacterium and promote survival (Novick, 2003; Novick *et al.*, 2010; Penadés and Christie, 2015). Additional roles for defective elements in bolstering the fitness of the host may include reduced susceptibility to antimicrobials, osmotic, oxidative and acid stressors, enhanced biofilm production, and reduced metabolic burden versus lysogeny with a fully functional element (Wang *et al.*, 2010). Such elements can also provide alternative promoters that maintain the expression of genes and regulons within which they are integrated (Scott *et al.*, 2012). While satellite prophages have been implicated in increased virulence and other phenotypic modulations in a number of other bacterial pathogens (Rezaei Javan *et al.*, 2019), at the time of writing, there does not appear to be substantial mechanistic evidence for the role of prophage satellitism in the biology of *S. pyogenes.*

Despite often being referred to as defective or decayed prophage which have undergone or are undergoing mutational decay, evidence suggests that streptococcal satellite prophage may be conserved for decades, maintaining almost complete nucleotide similarity throughout the genome, even against the highly recombinant genetic scaffold of the pneumococcus (Rezaei Javan *et al.*, 2019). The finding that full-length prophage and satellite prophage among the streptococci are in fact phylogenetically distinct groups, with 93% of all satellite prophage-encoded genes (>70% amino acid sequence similarity) not detected in any full-length elements corroborates this hypothesis (Rezaei Javan *et al.*, 2019) .

1.5.3.1 Lysogeny module

The ability for lysogenic bacteriophages to establish a stable infection within the host cell without immediately inducing lysis is a defining feature, and is regulated by the expression of genes within the lysogeny module. The lysogeny module encodes a site-specific integrase and an excisionase (Figure 1.7), in addition to repressor and antirepressor proteins, such as *cI* and *cro*, respectively (McCullor *et al.*, 2018; McShan and Nguyen, 2019) (Figure 1.6). Among the *Siphovoridae*, integrase enzymes typically utilise a catalytic tyrosine to initiate strand cleavage and site-specific recombination, and may require other additional prophage-, bacterium-encoded cofactors (Groth and Calos, 2004). Although the integrase mediates integration, and the excisionase mediates excision, the loss of function of the integrase, but retention of the excisionase would presumably lead to rapid loss of the vulnerable extrachromosomal element from the population, whereas loss of the excisionase might be expected to lead to a persistent integrative state wherein the excisionase were functionally redundant. There is however some evidence that depending on the degree of homology between the integrase gene of one phage, and the attachment sequences of another, such an element might be able to utilise the enzyme of a prophage that has retained a functional enzyme for this purpose (Groth and Calos, 2004).

The loss of the ability to integrate by deletion events affecting this module can lead to the conversion of a lysogenic element to either a lytic prophage (McCullor *et al.*, 2018) or a permanently integrated prophage, which lacks the ability to exist as an independently replicative entity ((da Silva Duarte *et al.*, 2019; Frígols *et al.*, 2015). Lysogeny is maintained by the expression of *cI*, which suppresses the lytic activator *cro* and this is maintained by autoregulation of *cI* (Bednarz *et al.*, 2014). When transcription of *cI* is not maintained at a sufficient level, transcription of *cro* may occur, which prevents the expression of *cI* and thus induces the lytic cycle (Bednarz *et al.*, 2014). Lysogeny is a highly stable state that may persist for millions of bacterial cell generations (Bednarz *et al.*, 2014; Zong *et al.*, 2010). However, this condition may be interrupted upon induction of the prophage by a number of signals, such as the initiation of the SOS-response (Figure 1.8) (Nanda *et al.*, 2015). During lysogeny, prophage may be largely transcriptionally silent, occasionally with the notable exception of genes believed to confer an advantage to the host bacterium, which promote the retention of such elements in the population (Ventura *et al.*, 2002a). While this phenomenon is mutually beneficial for both the virus and the bacterium, there is some evidence that genes that are subject to silencing may be lost, leading to the domestication of prophage as stably integrated elements or 'remnants' (Bobay *et al.*, 2014; Ventura *et al.*, 2002a).

1.5.3.2 Head and tail fibre genes and capsid assembly

Capsid and tail fibre genes are generally required for the assembly of nascent virus particles and subsequent transmission, and particle formation and encapsidation is a multistep process. Empty viral procapsids are icosahedral shells comprising multiple copies of the major capsid protein arranged as an icosahedral lattice, the assembly of which requires the major capsid protein itself, and often a scaffold and portal protein (Dokland, 1999; Murialdo and Becker, 1978). The scaffold protein is generally required for geometric fidelity, and in its absence, procapsid formation may become dysregulated and aberrant (Dokland, 1999). The mechanism by which the scaffold protein facilitates correct geometric architecture of the procapsid is not known (Aksyuk and Rossmann, 2011). Most dsDNA tailed viruses assemble the procapsid from multiple copies of one major capsid gene and one copy of gene encoding the scaffold protein though occasionally more are observed (Rao and Feiss, 2008). One of the capsids 12 fivefold vertices comprises a dodecameric cone-shaped portal protein, subunits of which are organised radially around a central channel which facilitates the translocation of the viral chromosome during packaging. After packaging, the icosahedral lattice of the procapsid is subject to a conformational change, expanding to become a mature shell. Though this process is not always triggered by packaging, this often occurs when packaging has begun, as the shell is being packaged (Aksyuk and Rossmann, 2011; Fuller *et al.*, 2007). Stringent evolutionary linkage dictates the association between capsid size and genome length (Rao and Feiss, 2008). The formation of infectious particles may not be required for lethal disease mediated by toxins encoded by prophage genes, such as Shiga toxin 2 by enterohaemorrhagic *E. coli* (Balasubramanian *et al.*, 2019). Prophage which lack their own structural genes have been described in *S. pyogenes* (Remmington *et al.* 2020; Canchaya *et al.*, 2002; Turner *et al.*, 2015) and are believed to be widespread among the streptococci (Rezaei Javan *et al.*, 2019).

1.5.3.3 Portal proteins, terminases and DNA packaging

DNA packaging is initiated by the small terminase, and occasionally the portal protein (Rao and Feiss, 2008). For the majority of dsDNA viruses, the viral genome is packaged into capsids following the formation of a packaging motor complex which comprises the phage terminase and portal proteins, and is driven by ATP hydrolysis (Rao and Feiss, 2008). A concatemeric or multimeric template of multiple copies of the viral genome is processed and translocated through the portal protein and into the empty capsid. When one or more copies of the viral genome have been packaged into the procapsid, the DNA is cleaved by the phage terminase (Aksyuk and Rossmann, 2011).

1.5.3.4 Neck and tail genes and completion of capsid assembly

After packaging, the capsid must be closed, and for many bacteriophages, this is followed by tail fibre attachment. In tailed phage, capsid closure is achieved by two general mechanisms: a conformational change which closes the portal channel or by the assembly of connector proteins. In the former, the portal protein serves as the binding site for the join between the tail protein and the capsid (Aksyuk and Rossmann, 2011; Rao and Feiss, 2008). When connector and neck proteins are implemented, these are the site for tail fibre binding (Orlova *et al.*, 2003). In addition to preventing premature egress of DNA due to the extreme pressures within the capsid, these proteins may also mediate the injection of the viral genome on attachment with a new host (Aksyuk and Rossmann, 2011; Orlova *et al.*, 2003; Rao and Feiss, 2008).

1.5.3.5 Tail and tail-fibre genes, assembly and capsid attachment

Most bacteriophage particles feature a proteinaceous tail (Ackermann and Prangishvili, 2012). The *Siphoviridae*, the family to which the majority of the phage of *S. pyogenes* belong, this tail is long, flexible and noncontractile and lacks a sheath. By contrast, the tails of the *Podoviridae* are short and noncontractile, whereas the *Myoviridae* have long and contractile tails (Nobrega *et al.*, 2018). Siphoviral and myoviral tail assembly occurs separately to capsid assembly, and thus, preformed tails are bound to preformed capsids, as opposed to being sequentially attached, as is the case for the *Podoviridae* (Aksyuk and Rossmann, 2011). The tail is essential for onward transmission and infection of new bacterial hosts by progeny virions, and is critical for the recognition of bacterial receptors at the host cell surface (Nobrega *et al.*, 2018). Among the *Siphoviridae* this cylindrical structure comprises multiple copies of the same tail tube protein, at the terminal end of which are the tail spike, tail fibres and a baseplate or initiator complex, (hereafter referred to as the baseplate) (Ackermann and Prangishvili, 2012; Aksyuk and Rossmann, 2011; Nobrega *et al.*, 2018). The baseplate is adorned with proteins that recognise specific receptors at the bacterial cell surface (Farenc *et al.*, 2014; Nobrega *et al.*, 2018). Among the lactococcal siphoviruses, the baseplate contains multiple receptor binding proteins, and this may be the same for those of the streptococci (Farenc *et al.*, 2014). When the receptor binding proteins bind with their targets at the bacterial cell

surface, this induces a conformational change, whereupon the tail polymerises, and the virus binds irreversibly to the host, subsequently injecting its nucleic acid into the bacterium (Nobrega *et al.*, 2018). Often, but not always, the phage of *S. pyogenes* (Canchaya *et al.*, 2003; McShan and Nguyen, 2019) and closely related species (Desiere *et al.*, 2001; Holden *et al.*, 2009; Singh *et al.*, 2014) features a hyaluronidase as part of the tail, which may aid in penetrating the HA capsule expressed by some isolates.

1.5.3.6 Phage-associated hyaluronidases

Though they are not homologous at the DNA sequence level and are immunologically distinct from chromosomally encoded hyaluronidases (Halperin *et al.*, 1987), many of the phage of *S. pyogenes* carry genes encoding hyaluronidases, denoted HylP's (Hynes *et al.*, 1995), including many encapsulated *emm*-types (Beres *et al.*, 2002; Stephen B. Beres *et al.*, 2004; Canchaya *et al.*, 2003; McShan and Nguyen, 2019; Nasser *et al.*, 2014; Smoot *et al.*, 2002a). The HylP enzymes of *S. pyogenes* are capable of degrading both HA and chondroitin sulphate (Mishra *et al.*, 2006; Mishra and Bhakuni, 2009), but lack an N-terminal signal peptide, and do not appear to be actively secreted from the bacterial cell, instead forming part of the phage tail fibre (Hynes *et al.*, 1995; Mishra *et al.*, 2006). It has been proposed that the function of these enzymes is to degrade the HA capsule (Hynes *et al.*, 1995). Serological studies have demonstrated that humans are exposed to HylPs during infection (Halperin *et al.*, 1987), and if one rejects the hypothesis that these play a supplemental role in virulence for the bacterium, perhaps this is due to prophage induction and transmission among streptococci *in-vivo* (Banks *et al.*, 2016; Broudy and Fischetti, 2003; Fischetti, 2007), though what utility activity towards other glycosaminoglycans might confer is unclear, adaptive immunity to phage components is well described (Krut and Bekeredjian-Ding, 2018; Van Belleghem *et al.*, 2019). In the equine pathogen *S. equi* sbsp. *equi*, which shares ~80% genome sequence identity with *S. pyogenes* (Holden *et al.*, 2009), a HylP has been shown to interact with a major human cell surface receptor for HA, CD44, and similar results were achieved with HylP derived from *S. pyogenes* (Singh *et al.*, 2014).

1.5.3.7 Holins and lysins

Virtually all dsDNA bacteriophage encode both a holin and a lysin (Young, 2014). Lysins are peptidoglycan hydrolases that disrupt the cell wall from within, and thus compromise the structural integrity of the bacterium which undergoes lysis, subsequently releasing progeny virions into the extracellular milieu (Pohane and Jain, 2015). Holins are small transmembrane proteins that facilitate the passive diffusion of lysins from cytoplasm to the periplasm, by forming a pore in the cytoplasmic membrane (Savva *et al.*, 2014). Although lysins are only required at the end of the infectious cycle, surprisingly, their transcription commences in the early stages in both lytic and lysogenic phage (Luke *et al.*, 2002; Pohane and Jain, 2015; Ventura *et al.*, 2002b) though often mature proteins are not detected until later stages (Luke *et al.*, 2002; Pohane and Jain, 2015). This is both inference of the ability of endolysin genes to be transcribed by host RNA polymerases and of post-transcriptional regulation of these genes. Suppression of endolysin mRNA translation is, in some cases, achieved by stable hairpin loops

(Pohane and Jain, 2015). Use of host bacterium-derived transcription factors to regulate endolysin expression has also been reported in some phage of both streptococci and other bacterial genera (Berdygulova *et al.*, 2011; Garcia *et al.*, 2002). Endolysin regulation can also be achieved by regulation of holin genes, although holin-independent translocation of endolysin have been described. In the pneumococcus, trafficking of endolysins that lack a signal sequence may be mediated by LytA (Frias *et al.*, 2013). Similarly, holins may act independently of endolysin-mediated lysis. Indeed, in some cases, holins form smaller pores in the cytoplasmic membrane, insufficient in size to permit endolysin translocation, but sufficient to lead to the compromised integrity of the host bacterial cell (Pang *et al.*, 2013). Having induced lysis of their host, and imbued with the capacity to degrade the cell wall (and if necessary, capsule) of prospective bacterial hosts, the lifecycle of the prophage recommences.

1.5.3.8 Lysogenic conversion module

From an evolutionary perspective, prophage must confer some advantage to the host bacterium that would neutralise otherwise deleterious effects of lysogeny on the lysogen (Abedon and LeJeune, 2005). In *S. pyogenes*, prophage often carry genes encoding bacterial virulence factors. Specifically, streptococcal superantigens, DNases and a secreted phospholipase (Proft *et al.*, 2003; Remmington and Turner, 2018; Sitkiewicz *et al.*, 2006). Such genes are located in the lysogenic conversion module. Prophage and similar elements also exist which do not appear to carry any genes with an immediately obvious benefit to the bacterium or potential detriment to human health (McShan and Nguyen, 2019). These integrate into the bacterial chromosome at sites that are fairly conserved, and have occasionally been purported to influence the regulation and expression of genes at their respective attachment sites, notably CRISPR loci (Beres *et al.*, 2002; Nozawa *et al.*, 2011). Presumably, retaining non-toxigenic prophage confers some advantage, perhaps genomic stabilisation (Nakagawa *et al.*, 2003) or immunity to bacteriophage superinfection (Labrie *et al.*, 2010), resistance to environmental stressors (Wang *et al.*, 2010), and enhanced biofilm formation (Carrolo *et al.*, 2010), though perhaps toxigenic phage immediate similar phenotypic modulations also.

1.5.3.9 Paratox

Paratox is a small gene, often contiguous with prophage-encoded toxin genes, that has recently been characterised in *S. pyogenes* (Mashburn-Warren *et al.*, 2012). It appears that the role of paratox is to reduce yet further the frequency at which *S. pyogenes* acquires exogenous DNA by transformation. Experimental evidence suggests that this is achieved through interfering with the binding of XIP with ComR, which ordinarily form a complex involved in the activation of genes that mediate transformation and recombination (Mashburn-Warren *et al.*, 2018). Although it has previously been reported that all sequenced *S. pyogenes* strains contain at least one copy of paratox (Mashburn-Warren *et al.*, 2018, 2012), this seems to have been an oversight (Turner *et al.*, 2015; Zhu *et al.*, 2015b). Intriguingly, deletion of paratox in the *emm*1 background increased the *in-vitro* transformation efficiency of the isolate by >100,000 fold, versus wildtype (Mashburn-Warren *et al.*, 2018).

1.5.4 Integrative and conjugative elements of *S. pyogenes*

Integrative and conjugative elements (ICE) of *S. pyogenes* are increasingly recognised as major virulence determinates and, like bacteriophage, constitute an expedient artifice for the exchange of genetic material with other bacterial species. Of these, one ICE has been studied in some detail and linked to pathogenic processes in *S. pyogenes,* denoted region of difference 2 (RD2). RD2 is an inducible element, comprising approximately 30 genes, is not found in all isolates, but has been identified in *emm*2, *emm*4 and *emm*28, and is best characterised in the latter genotype. This ICE encodes 7 secreted proteins, and 4 hypothetical cell surface-anchored proteins, two of these have been identified and characterised as virulence factors, *Spy1325*/*aspA* and *Spy1336*/*R28* (Green *et al.*, 2005; Kachroo *et al.*, 2019; Zhang *et al.*, 2006). AspA has been identified as an important adhesin required for upper respiratory tract colonisation (Franklin *et al.*, 2013). A homologue of this gene in the Group B *Streptococcus*, *S. agalactiae* was similarly shown to mediate adherence to vaginal epithelial cells (Rego *et al.*, 2016). A second RD2-associated virulence gene, *Spy1336*/*R28*, encodes an additional surface protein which has recently been implicated in increased virulence, also in genotype *emm*28 (Kachroo *et al.*, 2019).

1.5.5 *S. pyogenes* chromosomal islands

Formerly believed to be degraded prophage, but subsequently recognised as distinct biological entities and denoted *S. pyogenes* chromosomal islands (SpyCI), these elements are similar to the staphylococcal pathogenicity islands (SaPI), and like SaPI, may retain mobility in the absence of structural genes encoding capsid proteins by affiliation with auxiliary 'helper' phage (Frígols *et al.*, 2015). A number of *emm*-types variably host these elements integrated into the bacterial chromosome at various sites, most notably *emm*1, *emm*4 and *emm*28, but also *emm*2, *emm*5, *emm*6, *emm*18, *emm*31, *emm*37, *emm*49, *emm*53, *emm*59 and *emm*123 (Nguyen and McShan, 2014; Scott *et al.*, 2008). Homologous elements are also found in other streptococcal species, such as *S. dysgalactiae* sbsp. *equisimilis*, *S. canis* and *S. anginosus* (Nguyen and McShan, 2014). SpyCI are not typically toxigenic, but may regulate gene expression at by growth-phase integration and excision at attachment sites, often the DNA MMR operon (Scott *et al.*, 2008).

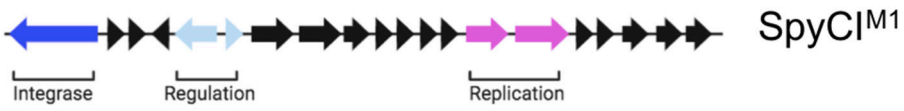

Figure 1.9: Genomic architecture of SpyCI

Similar to prophage, though not phylogenetically so, SpyCI such as SpyCI^{M1} of *emm*1 SF370 do appear to bear some semblance of modular arrangement, though SpyCI are substantially smaller than prophage, as are putative modules. In general, SpyCI genome sequences are dominated by genes of unknown function (black arrows) with few classifiable genes. These elements do not encode toxins, and their true role in the molecular biology of *S. pyogenes* is not well understood, but their prevalence among the streptococci suggests they may fulfil a hitherto undetermined purpose Figure created using EasyFig (Sullivan *et al.*, 2011) and sequence derived from completed reference genome SF370 (Ferretti *et al.*, 2001).

1.5.5.1 Mobile genetic elements, antibiotic resistance and *S. pyogenes*

Although antibiotic resistance in *S. pyogenes* is not exclusively mediated by mobile genetic elements (Cattoir, 2015; Musser *et al.*, 2020; Vannice *et al.*, 2019), in addition to encoding bacterial virulence factors, MGE of *S. pyogenes* may also mediate reduced susceptibility to antimicrobials by carrying resistance genes. Specifically, these elements confer variable resistance to macrolides and/or tetracyclines, by carrying *mef* and *tet* genes, respectively (Giovanetti *et al.*, 2005, 2003). Of the macrolide efflux pump encoding (*mef*) genes, *mefA* is the most common variant and interestingly, is carried on a prophage which in turn harbours a transposon, itself carrying *mefA* (Brenciani *et al.*, 2010; Clancy *et al.*, 1996). Prophage-associated tetracycline resistance in *S. pyogenes* is often mediated by *tetO*, which encodes a ribosomal protection protein (Giovanetti *et al.*, 2003). Although *tetO* is typically found in association with *mefA* on the same transposon, for instance on the archetypal *S. pyogenes* resistance-mediating prophage, ΦM46.1 (Brenciani *et al.*, 2010), *tetO* may be found without *mefA* on other mobile elements (Giovanetti *et al.*, 2012). Similarly, *mefA* may also be found alone, for instance on transposon inserted into other prophages, such as Φ1207.3 or Φ10394.4 (Brenciani *et al.*, 2004). In other streptococcal species, homologous transposons carrying resistance determinants may be found alone, integrated at other sites within the bacterial chromosome and not associated with a prophage (Santagati *et al.*, 2000). Worryingly, *in-vitro* experiments have shown that ΦM46.1 is also able to transduce *S. agalactiae*, *S. gordonii* and *S. suis,* though for now, these remain absent from surveillance data (Giovanetti *et al.*, 2014). ICE are also capable of carrying genes that confer resistance, notably ICE*Sp1116*, which encodes the ribosomal methylase *ermB*, ICE*Sp2905* (actually one ICE integrated into another), which encodes *ermA* and *tetO* and ICE-*emm*12, encoding *tetO* and *tetM* (Giovanetti *et al.*, 2012; Walker, 2015). Unlike carriage of prophage-encoded exotoxins, the incidence of isolates in which MGE-encoded resistance determinates appears to vary considerably geographically, and may reflect prescribing practices. For example, such isolates are dominant in outbreaks of scarlet fever in Hong Kong and mainland China, but are comparatively scarce in the United Kingdom (Lynskey *et al.*, 2019a; Tse *et al.*, 2012; Turner *et al.*, 2016; Walker, 2015).

1.6 *S. pyogenes* and host immunity

1.6.1 Innate immunity

The innate immune system is sufficient for clearance of *S. pyogenes*, and T-cell deficient mice are still able to clear infection (Medina *et al.*, 2001). The potency of innate immunity against *S. pyogenes* is perhaps reflected in the myriad diverse strategies that they have evolved to thwart detection and elimination by this system (Walker *et al.*, 2014). The interface between host and pathogen represents a highly dynamic convergence of tightly coordinated molecular events from both sides. The physiological and anatomical barriers of the host represent the primary interface in the interaction between host and pathogen. Before a response is triggered, the human host is passively protected by the physical barriers of the dry, desquamating, nutrient poor, low pH and high salinity environment of the integumentary system, and also the mucosal membranes, notably within the human

oropharynx. Like the integumentary system, mucosal membranes are prone to sloughing of superficial layers in addition to the secretion of antimicrobial enzymes such as lysozyme and antimicrobial peptides. These barriers are conferred additional protection by native microorganisms, which are highly efficacious in precluding invading pathogens by secreting antimicrobial compounds and other metabolites, sequestering nutrients and competing for space.

1.6.1.1 Immunological recognition of *S. pyogenes*

Invading pathogens are detected by the recognition of specific epitopes known as pathogen-associated molecular patterns (PAMPs), triggering an immune response. These sequences are recognised by receptors known as pathogen recognition receptors (PRRs), which are expressed on a multitude of cell types, including those lining the primary infection sites of *S. pyogenes*, the skin and oropharynx, and effector cells of the immune system. Data is available suggests a multitude of molecular intermediates and considerable functional redundancy in the detection of *S. pyogenes*-specific PAMPs by PRRs.

1.6.1.2 Toll-like receptors and *S. pyogenes*

Toll-like receptors (TLRs) are membrane anchored PRRs that are expressed on most innate immune effector cells and some adaptive immune cells, the stimulation of which activates the transcription factor NFkB and the interferon regulator family (IRF) transcription factors (Kawai and Akira, 2010). The expression of TLRs can also be upregulated on epithelial and endothelial cells (Iwasaki and Medzhitov, 2004). Macrophages and dendritic cells are critical in the host immune response to, and the control of infection and express TLRs both at the cell surface and intracellularly (Goldmann *et al.*, 2004; Loof *et al.*, 2007; Takeda *et al.*, 2003). In *S. pyogenes* infection, as is the case for many other Gram-positive bacteria, TLRs represent an important mechanism for recognition and signalling, and the induction of immunological and inflammatory responses. Recent evidence suggests that the adaptor molecule myeloid differentiation primary response protein (MyD88) is critical in the instigation of a functional immunological response to *S. pyogenes* and indeed a number of TLR are known to signal via. this protein (Loof *et al.*, 2010, 2008). Classically, TLR-2 was thought to recognise a number of Gram-positive bacterial epitopes including lipoteichoic acids and peptidoglycan. However, recent evidence suggests that peptidoglycan is in fact recognised by the intracellular receptors Nod1 and Nod2, and that lipoteichoic acids are synergistically recognised by TLR-1 and TLR-2, but peptidoglycan is not (Travassos *et al.*, 2004). Interestingly, it appears that despite being a major component of the immune response to a number of Gram-positive bacteria, TLR-2 may not be a key PRR in immunity to *S. pyogenes*. Indeed, TLR-2 deficient mice are only marginally more sensitive to *S. pyogenes* infection (Mishalian *et al.*, 2011). Conversely, mice that are deficient of MyD88 exhibit markedly hampered recruitment of neutrophils and dramatically enhanced susceptibility to infection with *S. pyogenes* (Loof *et al.*, 2010, 2008).

1.6.1.3 Other pattern recognition receptors and *S. pyogenes*

S. pyogenes is also recognised by non-TLR PRR during infection. Recently, *S. pyogenes* has been shown to trigger the activation of the NOD-like receptor (NLRP) inflammasome, NLRP3, which in turn leads to the release of IL-1β through the activity of macrophage caspase-1 (Harder *et al.*, 2009). Experimentally, this response has been shown to be induced by the pore-forming cytolysin streptolysin O, which is secreted by *S. pyogenes*. Interestingly, this pathway is independent of TLR- and MyD88-mediated signalling, and unlike the activation of these receptors, proceeds without the requirement for ATP (Harder *et al.*, 2009). PRR stimulation may also occur via. TREM-1, a PRR expressed on myeloid cells, notably neutrophils, monocytes and macrophages, potentiating the pro-inflammatory response to *S. pyogenes* (Colonna and Facchetti, 2003; Horst *et al.*, 2013). Indeed, inhibition of this receptor ameliorated the inflammatory response of macrophages *in-vitro*, and mitigated many of the clinical features of STSS in a murine model of infection, significantly reducing systemic levels of IL-6 and TNF-α (Horst *et al.*, 2013). Most macrophage subsets express scavenger receptors at the surface, and of these, scavenger receptor A (SRA) recognises Gram-positive organisms, specifically by binding LTA, mediating direct, non-opsonic uptake of bacteria to confer resistance to infection (Peiser *et al.*, 2002; Platt and Gordon, 2001; Thomas *et al.*, 2000). SRA readily recognises bacterial ligands at the cell surface of *S. pyogenes*, however expression of M-protein, specifically *emm*5 and *emm*6, has been shown to dramatically dampen this recognition (Areschoug *et al.*, 2008). One report describes a mechanism by which scavenger receptor B (SRB), also known as CD36, recognises some *S. pyogenes* isolates in whole blood, in a fashion that may be mediated by the binding of LDL to the streptococcal collagen-like protein Scl1, acting as an opsonin (Zhou *et al.*, 2016).

1.6.1.4 The inflammatory response

Recognition of PAMPs by PRR induces the inflammatory response; a molecular cascade that if not carefully coordinated, can prove highly problematic for both human hosts and invading bacteria (Mogensen, 2009). Typically, this involves the secretion of inflammatory cytokines such as such as IL-1β, IL-6, IL-8, IL-12 and TNF-α, largely by immune cells, but also by epithelial and endothelial cells. Once initiated, the inflammatory response modulates changes to the microvasculature within and around infectious foci, promoting vascular dilation which in turn engenders increased perfusion, coupled with a concomitant decrease in the rate of flow. The endothelial surface upregulates adhesion molecules in response to these signals, expressing more E- and P-selectin and ICAM-1 which promotes increased binding and diapedesis of neutrophils which follow inflammatory signals to the site of infection. The corollary of these inflammatory events is the increased exposure of effector cells with inflamed or damaged cells and tissues, as well as the offending agent or agents. The induction of a functional inflammatory response by dendritic cells in response to *S. pyogenes* is subject to an absolute requirement of MyD88 (Loof *et al.*, 2008). In mice that are deficient for MyD88, the inflammatory response is not properly initiated and subcutaneous infection with *S. pyogenes* rapidly results in lethal systemic infection and multiorgan failure (Loof *et al.*, 2010). On the other hand, an overzealous inflammatory response, particularly in

response to exotoxins released by *S. pyogenes*, can be equally damaging, particularly when these are released into the circulation causing sepsis and STSS, which is associated with high mortality rates (Lamagni *et al.*, 2008).

1.6.1.5 Phagocytes

The primary phagocytic cells involved in the prevention and control of *S. pyogenes* infection are neutrophils, mast cells, macrophages and dendritic cells (Di Nardo *et al.*, 2008; Loof *et al.*, 2007; Matsui *et al.*, 2011; Valderrama and Nizet, 2018; Voyich *et al.*, 2004). These cells, in addition to producing inflammatory cytokines themselves, follow these signals to infectious or inflammatory foci wherein they begin to trap and phagocytose offending bacteria. Phagocytosed bacteria are subsequently destroyed in intracellular vesicles by the activity of superoxide, hydrogen peroxide and hypochlorous acid following the formation of the phagolysosome by the fusion of the phagosome and lysosome (Uribe-Quero and Rosales, 2017). Both mast cells and neutrophils have been shown to extrude extracellular traps that ensnare and neutralise invading microorganisms (Brinkmann *et al.*, 2004; Goldmann *et al.*, 2008). These traps comprise chromatin, histones and other proteolytic and antimicrobial peptides (Brinkmann *et al.*, 2004). In addition to reducing the mobility of infectious agents and facilitating phagocytosis, they also exert antimicrobial activity independent of uptake-mediated elimination. Until relatively recently, there was some debate about whether NETs were produced *in-vivo*, however this phenomenon has since been demonstrated *in-vivo* using intravital microscopy and other techniques (Alasmari, 2020; Kolaczkowska *et al.*, 2015). Perhaps the most effective illustration of the importance of neutrophils in defending against *S. pyogenes* infection are the myriad extraordinary ways in which the bacterium strives to avoid them (Goldblatt *et al.*, 2019; Lin *et al.*, 2009; Lynskey *et al.*, 2017; Remmington and Turner, 2018; Turner *et al.*, 2009a; Uchiyama *et al.*, 2012; Von Köckritz-Blickwede and Nizet, 2009; Von Pawel-Rammingen, 2012; Voyich *et al.*, 2004). Mast cells employ similar means of pathogen elimination, and have been implicated in host resistance to cutaneous infection with *S. pyogenes* (Di Nardo *et al.*, 2008; Matsui *et al.*, 2011). Paradoxically, there is also evidence that *S. pyogenes* will readily survive and replicate within neutrophils (Medina et al., 2003).

Macrophages, unlike neutrophils and mast cells, are more heterogenous and constitute a self-renewing population of professional phagocytes supplemented by contribution from differentiating monocytes. Macrophages are thus found resident within tissues in the absence of disease, as well as in the circulation (Gordon and Plüddemann, 2017). These phagocytes recognise bacteria by interaction between surface and endosome expressed PRR and SR and their cognate PAMP and DAMP ligands. Compared to other phagocytes, macrophages acidify more strongly and indeed more rapidly (Yates *et al.*, 2005). A number of lines of investigation have demonstrated the fundamental contribution of macrophages to the innate immune response to *S. pyogenes*. Experimental knockdown of these cells has highlighted their role in neutralising *S. pyogenes* in the bloodstream (Goldmann *et al.*, 2004). Further, in mice depleted of macrophage, cutaneous infection with *S. pyogenes* was significantly more likely to disseminate systemically and progress to invasive disease involving multiple internal organs (Mishalian *et al.*, 2011). Despite this, there is evidence that *S. pyogenes* is able to resist phagolysosome acidification and

intracellular processing and readily replicates within these cells (Bastiat-Sempe *et al.*, 2014; Hertzén *et al.*, 2010; Thulin *et al.*, 2006). Dendritic cells are able to directly recognise *S. pyogenes* by TLRs (Loof *et al.*, 2008) and this recognition has been shown to drive dendritic cell maturation *in-vitro* (Veckman *et al.*, 2004) and *in-vivo,* whereupon they release pro-inflammatory cytokines (Loof *et al.*, 2007). Following dendritic cell ablation in mice, cutaneous infection with *S. pyogenes* was more likely to spread to efferent lymph nodes and internal organs (Loof *et al.*, 2007). There is evidence to suggest that the bacterium is also able to inhibit dendritic cell maturation, resist destruction within the phagolysosome and survive within (Cortés and Wessels, 2009).

1.6.2 **Complement**

Complement comprises a highly abundant, hierarchical system of over 30 plasma and cell surface-associated proteins activated by interlinked, cascading pathways (Figure 1.10). Complement is triggered following the recognition of pathogenic microorganisms, in addition to other insults, and culminates in the production of pro-inflammatory effectors, opsonisation of the invading organism or biological entity, and the formation of the membrane attack complex (MAC) and subsequent lysis of the pathogen (Dunkelberger and Song, 2010). Complement is activated by three distinct pathways denoted the classical pathway, the lectin pathway, and the alternative pathway. The classical pathway is activated by the formation of the immune complex by the binding of C1q with the serine proteases C1r and C1s, which subsequently bind to the Fc-region of immunoglobulins, often IgG and IgM, which are sequestered at the surface of invading microorganisms. Further, the activation of the two serine proteases promotes the cleavage of C4 and C2 and the formation of C3 convertase. When C3 convertase is formed, C3 is cleaved to yield C3a and C3b, and this inflection point of complement activation is the bottleneck at which all three complement activation pathways converge. The production of C3b enables this molecule to bind to and identify non-self-entities by decorating their surface, and in turn, this promotes further complement activation, eventually culminating in the production of anaphylatoxins and assembly. In contrast, the lectin pathway is initiated by a similar but distinctly immunoglobulin-independent mechanism wherein PAMPs are recognised via. carbohydrate ligands by mannose-binding lectins. Analogous to the aforementioned C1 complex of the classical pathway, mannose-binding-lectin forms an immunological complex with affiliated serine proteases which cleave C2 and C4, subsequently generating C3 convertase. The final major mechanism by which complement can be activated is the alternative pathway, triggered by spontaneous hydrolysis of C3, which is highly abundant in plasma, and yields a hydrated molecule that is analogous to C3b. The hydrated analogue, denoted C3(H_2O), binds to and facilitates the cleavage of Factor B to yield Bb and Ba by Factor D. By this mechanism, the C3 convertase of the alternative pathway, C3(H_2O)Bb, is formed, promoting the cleavage of C3 and the production of C3a and C3b. The activation of complement again is convergent in the activation of three major effector pathways. Broadly, these are the MAC-mediated lysis of the pathogen, the production of pro-inflammatory mediators, and finally, opsonin mediated clearance of the organism or organisms and priming of phagocytic cells (Dunkelberger and Song, 2010). As complement straddles both the innate and adaptive immunity and involves multiple arms of the human immune system, the mechanisms by which it curtails *S. pyogenes*

colonisation and infection are numerous. Factor B appears to be important to complement mediated immunity towards *S. pyogenes*, and facilitates C3b deposition at the surface of the bacterium via. the alternative pathway (Yuste *et al.*, 2006). Indeed, it seems that the alternative pathway might be the more instrumental activation pathway in the clearance of *S. pyogenes* infection by innate immunity, specifically mediated by C1q, which interestingly, may also exert activity against *S. pyogenes* in a fashion that is independent of complement in both humans and mice. In the same study, both Factor B and C1q-deficient mice exhibited enhanced susceptibility to invasive disease and C1q deficient serum of both human and mice exhibited impaired capacity to induce opsonophagocytic killing of *S. pyogenes* (Yuste *et al.*, 2006). In addition to expressing a multitude of factors which interfere with the potent antibacterial activity of complement at multiple levels (Laabei and Ermert, 2018), *S. pyogenes* appears to take advantage of the complement interference strategies of other pathogens, notably the Sarcoptiforme ectoparasitic mite *Sarcoptes scabei* which secretes a potent serine protease inhibitor, and indeed the bacterium exhibits a particular proclivity for sarcoptic skin lesions (Armitage *et al.*, 2019; Swe *et al.*, 2017).

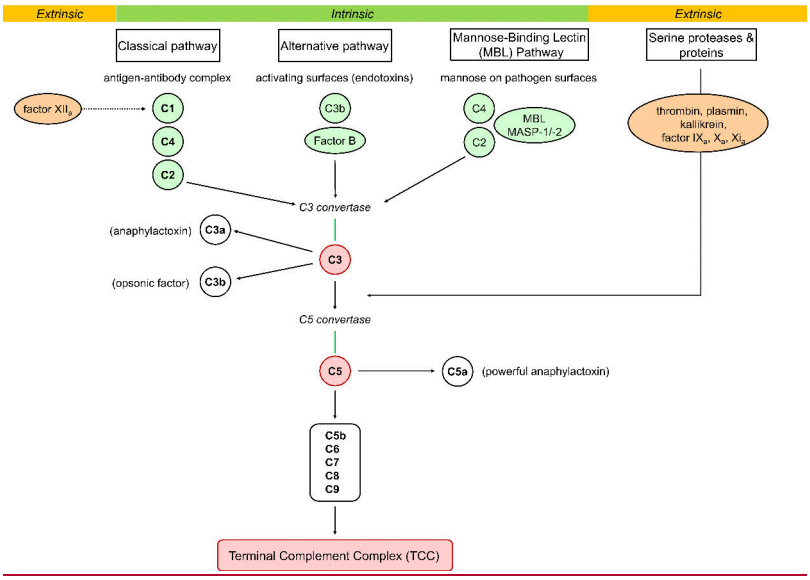

Figure 1.10: Complement pathway

The three iterations of the complement pathway are the alternative, classical and lectin pathways. Though these pathways follow distinct molecular mechanisms, they each converge on the production of convertase C3. The C3 convertase promotes the synthesis and release of chemoattractant C3a and C5a. C5b, C6-9 coalesce to form the membrane attack complex (also known as the terminal complement complex) (Dunkelberger and Song, 2010). "Complement activation.jpg" by Hofman.martijn is licensed under CC BY-SA 4.0.

1.6.3 Antibodies and adaptive immunity

There is substantive evidence that anti-*S. pyogenes* antibodies are produced late in infection or in convalescence following infection with this bacterium (Beachey *et al.*, 1988; Mortensen *et al.*, 2015; Shet *et al.*, 2003), and a lasting adaptive response appears to be mediated by both T cells, notably Th1 and Th17, and immunoglobulins,

predominantly IgG1 and IgG3, and IgA (Bessen and Fischetti, 1988; Fluckiger *et al.*, 1998; Mortensen *et al.*, 2015; Soderholm *et al.*, 2017). The humoral response to *S. pyogenes* infection can be broadly classified as being M protein specific and non-M protein specific. Lasting immunity to specific isoforms of M protein is well-described, and forms the basis for efforts to develop a vaccine (Davies *et al.*, 2019), and abundant anti-M protein directed CD4+ T cells which secrete pro-inflammatory cytokines can be found in tonsillar tissue derived from patients with recurrent streptococcal tonsillitis, versus patients with tonsillar hypertrophy, which does not have an infectious aetiology (Gogos and Federle, 2019). Other proteins, including C5a peptidase, immunogenic secreted protein and SpyCEP similarly induce protective antibodies and are also appealing targets (Turner *et al.*, 2009b; Davies *et al.*, 2019). Antibodies to the streptococcal superantigens SpeA and SpeC are capable of inhibiting their mitogenic potential (Norrby-Teglund *et al.*, 1996), and the peptidase activity of C5a peptidase is abrogated by antibody neutralisation (Cleary *et al.*, 2004). Anti-*S. pyogenes* adaptive immunity is thus multifactorial, and typically begins at the mucosal surfaces of the pharynx. Here, IgA serves to protect against adherence and colonisation of the bacterium in the early stages, binding to the C-repeat region of M-protein (Bessen and Fischetti, 1988; Cunningham, 2000; Fluckiger *et al.*, 1998). Intranasal immunisation of mice with C5a peptidase derivatives was observed to confer resistance of the host to colonisation by *S. pyogenes* mediated by stimulating salivary secretory IgA (Ji *et al.*, 1997). The second arm of protective immunity to *S. pyogenes* is mediated by IgG, which prevents internalisation and invasion, inhibiting the progression to established, deep-set infection and entry to the bloodstream by promoting phagocytosis and clearance in conjunction with complement (Fluckiger *et al.*, 1998). The production of IFN- γ also appears to play a role in the bodies defence against *S. pyogenes*, although children appear to produce inferior quantities of IgG3 and IFN- γ compared to adults (Mortensen *et al.*, 2015), and this may go some way to explaining why children are particularly susceptible to infection, though perhaps recurrent exposures over time contribute to the more robust protection observed in adults, mediated by a mechanism that is independent of opsonisation (Cunningham, 2000).

1.7 *S. pyogenes* virulence factors

1.7.1 M protein and M-like proteins

In addition to forming the basis of the widely used basis for genotyping *S. pyogenes* isolates, the *emm* or M-protein is an important virulence factor (Figure 1.11). M-protein is an α-helical coiled coil that exists as a fibrillar adhesin projecting from the cell surface. Anchored within the peptidoglycan of *S. pyogenes* by a C-terminal LPXTG motif, this reasonably conserved region is followed by 4 heterogenous repeat domains, denoted A-D. These domains are variable in their sequence and size according to the number of heptad repeats. As the protein extends towards the N-terminus, sequence variability increases, ultimately culminating in the terminal ~50 residues of the A repeat domain. This hypervariable region confers serospecificity, and the DNA sequence confers the genotype. Though the M protein has been subject to intense experimental scrutiny over the years, due to the highly variable nature of M, it is often helpful, and indeed important, to interpret such studies in the context of

genetic background. Arguably the best described function of M-protein as a virulence factor is its capacity as an anti-phagocytic molecule, which is achieved by a number of proposed mechanisms. In some cases, this is thought to be mediated by the mature protein forming complexes with fibrinogen and other host-derived proteins, or alternatively, possibly due to the M-protein interfering with the activity of convertases at the cell surface (Carlsson *et al.*, 2005), both mechanisms reportedly prevent or reduce complement deposition. While this has been observed in the M4 and M5 isoforms (Carlsson *et al.*, 2005; Courtney *et al.*, 2006a), this was not observed in M6 (Horstmann *et al.*, 1992). The M1 protein, when in complex with fibrinogen, may be released from the cell surface (possibly by the activity of SpeB), subsequently activating neutrophils, promoting inflammation and vascular leakage (Påhlman *et al.*, 2006). There is also some evidence to suggest that some M-protein isoforms (notably M5) may have superantigenic activity (Bisno, 1991; Robinson and Kehoe, 1992). Another well substantiated mechanism of M-proteins implication in pathogenesis is as an adhesin (Smeesters *et al.*, 2010). Isogenic *S. pyogenes* strains differing only in their *emm* allele exhibit markedly different adhesive properties to various cell types, and indeed bind to these in different ways to elicit this effect (Berkower *et al.*, 1999). The M-protein may also promote invasion of host epithelial cells, and there is some evidence to suggest that this is mediated by binding of fibronectin by M-protein, specifically in this case the M1 isoform (Cue *et al.*, 2001). Despite the putative adhesin and invasin properties of M-protein, it is surprising that it does not appear to play a major role in biofilm formation, as the capacity to propagate these structures can vary considerably within *emm*-types (Afshar *et al.*, 2017; Lembke *et al.*, 2006; Turner *et al.*, 2015). Though the M-protein may not be the critical determinate of biofilm formation, in *emm*5 isolates, the M5 protein does appear to play some role in the initial stages of adherence (Cho and Caparon, 2005). The amino acid sequence homology of certain M-isoforms to human fibrillar proteins is curious if not completely surprising in an organism as adept at interfering and inhibiting host immunity as is *S. pyogenes.* Certainly, a number of distinct M-proteins share significant homology and indeed induce antibodies that are cross-reactive with host proteins such as keratin, myosin, tropomyosin, and laminin (Fischetti, 1989; Robinson and Kehoe, 1992). It stands to reason that this immunological phenomenon may go some way to explaining the association of some *emm*-types with the oft neglected post-infectious sequelae (Bessen *et al.*, 1989; Bisno, 1991; Fischetti, 1989; Robinson and Kehoe, 1992). Chimeric *emm* alleles, the products of splicing events between *emm* and downstream *enn* have been described in a number of *emm* types (DebRoy *et al.*, 2018; Frost *et al.*, 2020).The phenotypic impact of this event on these isolates, and indeed the effect this might have on coverage by the multi-valent M-protein vaccine remains unknown.

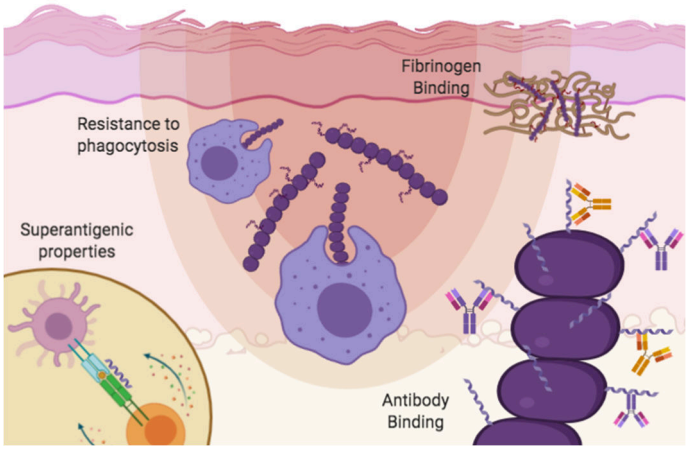

Figure 1.11: M Protein: selected mechanisms of action

The M protein of *S. pyogenes* is capable of interfering with and interacting with the human host through myriad mechanisms. These appear to vary between isoforms, but notable functions include binding immunoglobulins, resisting phagocytosis, binding fibrinogen, and acting as a potent mitogen. Figure created by the author with BioRender.com

1.7.2 Hyaluronic acid capsule

The capsule of *S. pyogenes* is comprised of hyaluronic acid (HA), a high molecular weight glycosaminoglycan, also found in the tissues of many eukaryotic organisms, including in the connective tissues of humans (Crater *et al.*, 1995). Capsule biosynthesis is mediated by the three HA synthetases which together comprise the capsule operon, *hasA, hasB* and *hasC* (Wessels, 2019). Many mechanisms of putative capsule functions have been proposed, with arguably the most prevailing being immune evasion and resistance to opsonophagocytic killing (Dale *et al.*, 1996; Wessels, 2016). Until relatively recently, it was believed that the HA capsule of *S. pyogenes* was an essential facet of *S. pyogenes*, necessary for immune-evasion and virulence (Ashbaugh *et al.*, 1998). However, numerous lines of enquiry have shown that capsule is inessential for *S. pyogenes* to attain full virulence and cause severe disease (Flores *et al.*, 2019, 2012; Galloway-Peña *et al.*, 2016; Henningham *et al.*, 2014; Turner *et al.*, 2019b, 2015). Indeed, loss of capsule may actually be advantageous in certain contexts, enhancing adhesion to epithelial cells and promoting persistence and longer-term shedding (Flores *et al.*, 2014; Kawabata *et al.*, 1999; Turner *et al.*, 2019b, 2015; Vega *et al.*, 2020).

Though there is substantive evidence that capsule is inessential, encapsulation in *S. pyogenes* might be regarded as very much variations on a theme, with modulations to capsule production resulting from mutations in transcriptional regulators (Lynskey *et al.*, 2013; Sumby *et al.*, 2006), polymorphisms within the *hasABC* operon (Ashbaugh *et al.*, 1998; Turner *et al.*, 2019b), loss of the operon (Flores *et al.*, 2019; Turner *et al.*, 2015) or possibly simultaneous degradation of capsule by hyaluronidases secreted by the bacterium (Henningham *et al.*, 2014; Starr and Engleberg, 2006) (Figure 1.12). In this way, expression of capsule varies temporally, and in response to environmental and immunological stimuli during growth and during infection, including strong

upregulation on growth in human blood (Berkower *et al.*, 1999; Biswas and Scott, 2003; Gryllos *et al.*, 2001; Li *et al.*, 2014; Lynskey *et al.*, 2013; Shea *et al.*, 2011; Unnikrishnan *et al.*, 1999; Virtaneva *et al.*, 2005). Experimental restoration of capsule expression in acapsular *emm*-types may enhance whole blood survival, and elimination of capsule expression in encapsulated *emm*-types, perhaps unsurprisingly, appears to reduce whole blood survival in some *emm*-types (Vega *et al.*, 2020), though not others (Henningham *et al.*, 2014). It is interesting however that in genotype *emm*89 isolates, those that have naturally undergone capsule loss survive just as well in whole blood as encapsulated isolates (Turner *et al.*, 2015), though this could conceivably be due to a compensatory of other noteworthy mutations characteristic of contemporaneous *emm*89 isolates. Inasmuch as a growing body of evidence appears to support the dispensability of the *S. pyogenes* capsule, many studies have conversely indicated that hyper-encapsulated strains may exhibit enhanced virulence or propensity to cause severe disease (Schrager *et al.*, 1996; Tamayo *et al.*, 2010). Like many aspects of *S. pyogenes* biology, it seems likely that genetic background may play an appreciable role in defining the consequences of capsule loss, and indeed there is some experimental evidence for this using a small number of *emm*-types (Vega *et al.*, 2020).

Notably acapsular *emm*-types include *emm*4, 22, 28, 87, although increasingly, *emm*-types are being found to comprise both encapsulated and non-encapsulated strains (Flores *et al.*, 2019, 2012; Turner *et al.*, 2019b, 2015). Increasingly, it seems that a number of *emm*-types may compensate for a lack of capsule with increased expression of SLO and NADase (Turner *et al.*, 2019b). Recent adaptation of the infant mouse pneumococcal transmission model to *S. pyogenes* suggests that encapsulated *emm*-types might be better suited to rapid transmission, whereas acapsular *emm*-types may persist and transmit at lower levels, albeit for longer periods of time (Vega *et al.*, 2020). The same study found that the prototype encapsulated *emm*3 strain required a lower index:contact ratio, whereas the prototype acapsular *emm*87 strain required more infected infant mice per uninfected littermates to transmit efficiently.

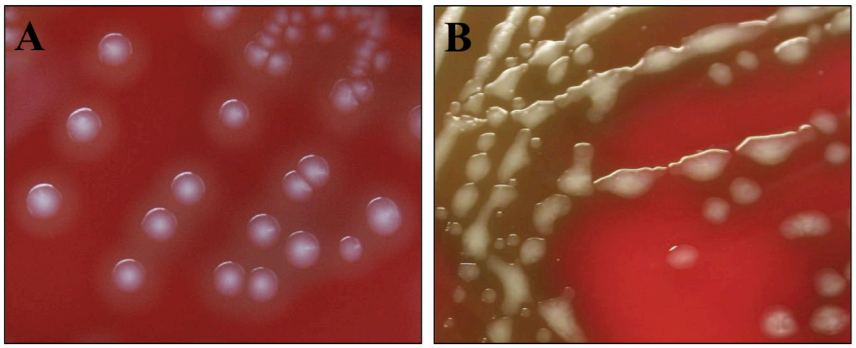

Figure 1.12: Variable expression of hyaluronic acid capsule by *S. pyogenes*

Overexpression of capsule can have marked impact on phenotype. A) An isolate of *S. pyogenes* on Columbia horse blood agar, which is not overexpressing capsule. B) An isolate of *S. pyogenes* over expressing capsule. Colonies are characteristic in their 'snotty' or 'oil-drop' appearance, are glossy, and lack well circumscribed colony margins, surrounded by irregular β-haemolytic halos. Some genotypes, or lineages thereof, are genetically acapsular (Flores *et al.*, 2012; Turner *et al.*, 2019a, 2015), while others are consistently highly encapsulated (Lynskey *et al.*, 2013). Image adapted from Bacteria In Photos & Tamayo *et al.*, 2010.

1.7.3 Streptolysin S

Streptolysin S (SLS) is a potent cytotoxin with high activity against erythrocytes, and a number of other human cells, including platelets, neutrophils, lymphocytes and epithelial and myocardial cells (Dale *et al.*, 2002; Datta *et al.*, 2005). SLS is produced by the *sag* operon that comprises 9 genes, all of which are necessary and sufficient for the production of the active protein (Nizet *et al.*, 2000). The first gene in this operon, *sagA*, produces a precursor molecule which is subsequently modified and processed by the remaining genes immediately downstream (Datta *et al.*, 2005). Though the most obvious role of SLS is as one of two haemolysins, as can be seen when *S. pyogenes* is propagated on blood agar, (Figure 1.12), the toxin also may play a role in the disruption of mucosal barriers and enhancing cellular invasion by driving calpain mediated proteolysis of E-cadherin expressed by epithelial cells (Rios-Doria *et al.*, 2003; Sumitomo *et al.*, 2011). Indeed, isogenic *sagA* mutants were significantly less able to form necrotic lesions and spread systemically in a murine model of disease (Datta *et al.*, 2005).

1.7.4 Streptolysin O and NADase

The secreted cytolysin streptolysin O (SLO) and the co-toxin NAD+ glycohydrolase (NADase) are encoded by the co-transcribed genes *slo* and *nga*, along with the gene encoding the NADase inhibitor, *ifs*. SLO is cytotoxic, and is one of the major haemolysins of *S. pyogenes*, though unlike SLS, SLO is cholesterol dependent, and triggers osmolysis of human cells by the formation of transmembrane pores. SLO also mediates the activity of its co-toxin by facilitating the translocation of NADase across the cell membrane, and by this mechanism, the two toxins act synergistically to induce lysis of human keratinocytes (Bricker *et al.*, 2005; Madden *et al.*, 2001; Michos *et al.*,

2006; Velarde *et al.*, 2017), though NADase is cytotoxic in its own right (Chandrasekaran and Caparon, 2015; Sharma *et al.*, 2016). These two virulence factors may also promote invasion and intracellular survival, possibly by interfering with autophagocytic killing (Hsieh *et al.*, 2018). Interestingly, enhanced expression of the *nga-ifs-slo* locus by the acquisition of a high activity promoter variant has been associated with a number of successful lineages, and thus increased production of these potent cytotoxins appears to confer an appreciable advantage (Nasser *et al.*, 2014; Turner *et al.*, 2019b, 2015; Zhu *et al.*, 2015a).

1.7.5 Chromosomally-encoded hyaluronidases

Hyaluronidases encoded on the core chromosome of *S. pyogenes*, denoted HylA, are generally considered to be spreading factors, serving to expedite streptococci or their toxins and other large molecules through tissue planes in the human host (Hynes and Walton, 2000; Starr and Engleberg, 2006). It has been reported that encapsulated *S. pyogenes* isolates lack a functional *hylA* gene owing to a SNP, resulting in the substitution of a D for a V at position 199 of the protein (Hynes *et al.*, 2009). Subsequent work has suggested that the expression of capsule and the retention of a functional *hylA* genes are mutually exclusive, possible evidence of competitive evolution, though they may be co-expressed in Group C streptococci (Henningham *et al.*, 2014; Maclennan, 1956). The *S. pyogenes* hyaluronidase might serve a possible nutritive function during pathogenesis, wherein *S. pyogenes* may use HA as a carbon source (Starr and Engleberg, 2006). Studies investigating the role of HylA in virulence are scarce, but along with the previous report that a significant number of major encapsulated *emm*-types lack a functional allele, suggest that either HylA are not major virulence factors or that there is some functional redundancy to these enzymes (Henningham *et al.*, 2014; Hynes *et al.*, 2009; Starr and Engleberg, 2006).

1.7.6 Serum opacity factor

Serum opacity factor (SOF) is a pleiotropic virulence factor produced by 45-50% of *emm*-types, and approximately 50% of invasive isolates (Courtney and Pownall, 2010). SOF is so named due to the opacifying effect it exerts on mammalian serum by binding to apolipoprotein moieties of high density lipoprotein (HDL) in human blood, triggering the release of their lipid cargo, which coalesces as droplets (Courtney and Pownall, 2010). Different *emm*-types secrete immunologically distinct isoforms of the mature SOF protein, and hence, the ability of anti-SOF antisera to inhibit the opacification of serum by *S. pyogenes in-vitro* can be used in a fashion similar to the erstwhile M serotyping scheme to subclassify *S. pyogenes* isolates, and indeed SOF-types correlate well with established *emm*-types (Johnson *et al.*, 2006). Some *emm*-types do however carry multiple SOF allelic variants (Beall *et al.*, 2000). There is evidence to suggest that SOF acts as an adhesin, possibly by interacting with integrins and fibulin-1 receptors in epithelial tissues (Courtney *et al.*, 2009; Courtney and Pownall, 2010; Gillen *et al.*, 2008; Schwarz-Linek *et al.*, 2006). There is also evidence for a role in promoting survival in whole blood, conferring resistance to phagocytosis (Courtney and Pownall, 2010). Additionally, SOF may promote bacterial invasion and internalisation by host cells through interacting with a fibronectin intermediate (Gillen *et al.*, 2008).

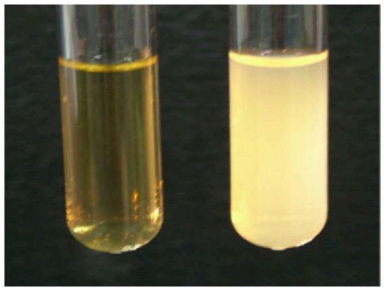

Figure 1.13: Serum opacification by SOF

Overnight incubation of human serum with recombinantly expressed SOF (right-hand vial) produces marked opacification relative to the untreated human serum control (left-hand vial). Image: "Opacification of Human serum by SOF" reproduced with permission from Courtney and Pownall, 2010 as licensed under CC BY-3.0.

1.7.7 Streptococcal pyrogenic exotoxin B

The cysteine protease, streptococcal pyrogenic exotoxin B (SpeB), despite its name, is neither a pyrogen nor an exotoxin, but is nonetheless a highly promiscuous and broadly acting virulence factor of *S. pyogenes* (Gerlach *et al.*, 1994). The myriad contributions of SpeB in the infection biology of *S. pyogenes* are at times paradoxical, and despite extensive study, direct assessments of its precise function remain elusive. This is perhaps due to notable variations in expression between anatomical sites and different environments (Lukomski *et al.*, 1999; Olsen *et al.*, 2012), genetic background (Friães *et al.*, 2015), mutations in transcriptional regulators (Sumby *et al.*, 2006), and indeed the propensity of mature SpeB to degrade not only itself, but also many other virulence factors (Nelson, *et al.*, 2011). The production of SpeB is tightly controlled, and begins with the synthesis of the zymogen, SpeBz, which is cleaved to yield an active enzyme, denoted SpeBm, by mature SpeB (Carroll and Musser, 2011). Ordinarily, this might constitute somewhat of a causality dilemma, though fortunately for the bacterium, the serine protease HtrA (DegP) has also been shown to activate the zymogen in this way, and perhaps unsurprisingly has been shown to be essential for the production of SpeBm (Lyon and Caparon, 2004). There appears to be a clear role for SpeB in the proteolysis of bacterial factors which are often susceptible to degradation or modification. Bacterial substrates include M-protein (Raeder *et al.*, 1998), SmeZ (Nooh *et al.*, 2006), EndoS (Allhorn *et al.*, 2008) and Sda1 (Walker *et al.*, 2007), all of which are potent mediators of virulence. This represents an interesting paradigm for the regulation of virulence by *S. pyogenes* wherein redundant factors are degraded, or desirable factors are cleaved and activated by SpeB. Mature SpeB also degrades a litany of eukaryotic factors, degrading immunoglobulins (Collin *et al.*, 2002; Collin and Olsén, 2001; Eriksson and Norgren, 2003), and chemokines (Egesten *et al.*, 2009). SpeB also appears to promote tissue damage and epithelial cell apoptosis (Tamura *et al.*, 2004; Tsai *et al.*, 1999), and may trigger inflammation (Kapur *et al.*, 1993), promote vasodilation, and also induce pain (Herwald *et al.*, 1996). While some studies have drawn an association between increased expression of SpeB and enhanced virulence (Lukomski *et al.*, 1999, 1998), others report an inverse relationship or indeed report

having found no relationship at all (C. D. Ashbaugh *et al.*, 1998; Ashbaugh and Wessels, 2001; Kansal *et al.*, 2000; Talkington *et al.*, 1993). It is important to note however, that early studies, particularly those conducted prior to the widespread application of WGS, may be confounded by the association of other mutations and variants with invasive (or non-invasive) disease, and therefore might benefit from reappraisal with hindsight (Horstmann *et al.*, 2011; Ikebe *et al.*, 2010; Lynskey *et al.*, 2019b; Shea *et al.*, 2011; Sumby *et al.*, 2006; Turner *et al.*, 2009).

1.7.8 **Immunoglobulin degrading enzymes**

As *S. pyogenes* major anatomical niches are mucosal surfaces, and occasionally sterile sites and the bloodstream, in addition to evading innate immunity, the bacterium must also avoid elimination by adaptive immune effectors. While the cysteine protease SpeB is capable of degrading immunoglobulins *in-vitro*, this does not appear to extrapolate to more physiologically relevant conditions (Persson *et al.*, 2013). In actuality, it seems unlikely that SpeB is able to degrade immunoglobulins with any appreciable efficiency, owing to the promiscuity and non-specificity of this enzyme (Persson *et al.*, 2013). Surprisingly given its typical habitat, *S. pyogenes* seems to lack a specific IgA-degrading enzyme (Von Pawel-Rammingen, 2012), but has demonstrable capacity to degrade IgG and prevent effector functions by activity of the cysteine protease IdeS (IgG degrading enzyme of *S. pyogenes*) (Agniswamy *et al.*, 2004; Lei *et al.*, 2002; Vincents *et al.*, 2008; Von Pawel-Rammingen, 2012). There are two allelic variants of *ideS* (also known as *mac*), encoding two functionally divergent complexes with overlapping substrate specificities, but importantly, quite distinct roles in IgG-mediated immune interference (Lei *et al.*, 2002). The substrate for complex I is a region within the IgG molecule, the cleavage of which inactivates the immunoglobulin, though this protease has some limited capacity for IgM cleavage also (Agniswamy *et al.*, 2004). Complex II on the other hand acts by blocking the phagocyte-bound Fc receptor CD16 (Agniswamy *et al.*, 2004). Curiously, IdeS also appears to interfere with the antibacterial action of neutrophils by arresting ROS production (Von Pawel-Rammingen, 2012). Antibodies to IdeS are found in low concentrations in acute-phase sera, but are abundant in convalescent sera of patients with both superficial and invasive *S. pyogenes* infection, indicating the production of both complexes *in-vivo* (Lei *et al.*, 2002, 2001). Further, a cofactor of IdeS, cystatin C, has been identified, which greatly increases its efficiency, and may sustain its immune-evading properties even at distal sites where the enzymatic concentration of the endopeptidase might be otherwise dampened by human protease inhibitors (Vincents *et al.*, 2008). IgG effector functions are further stymied by the activity of the glycan hydrolase, EndoS. EndoS has specificity for the N-glycan moiety of IgG, enabling *S. pyogenes* to evade opsonophagocytic killing through the dual action of impeding Fc-receptor binding and induction of complement activation pathways (Allhorn *et al.*, 2008; Collin *et al.*, 2002).

1.7.9 **Immunoglobulin binding proteins**

Although *S. pyogenes* appears to lack an IgA degrading protease, the capacity for some genotypes to sequester IgA at the cell surface has been described, occurring via. a hitherto uncharacterised N-terminal binding domain

of M-protein (Johnsson *et al.*, 1994). In pattern C, D and E isolates, the M-like protein *enn* can facilitate the sequestration of IgA at the cell surface *in-vitro*. The M22 protein encoded by *emm*22 also reportedly binds IgA (Carlsson *et al.*, 2003). The highly conserved bacterial cell surface-associated protein SibA has been shown to bind both the Fc- and Fab-regions of IgA, as well as IgM and IgG. There is also some evidence that SibA can also be secreted, possibly acting as a decoy to bind immunoglobulins (Fagan *et al.*, 2001). Interaction with IgG has been reported in some genotypes, mediated again by M protein. Although the M protein of some genotypes may bind all IgG subclasses, others can only bind specific isoforms such as IgG3 (Reglinski and Sriskandan, 2014; Retnoningrum and Cleary, 1994). Remarkably, the M1 protein appears to differentially bind IgG in response to concentrations of the immunoglobulin in the extracellular milieu (Nordenfelt *et al.*, 2012). The authors of this elegant study demonstrated that in the throat, where IgG concentrations in saliva are low, IgG was preferentially bound to M1 by the Fc-region. However, in the blood, where there are high concentrations of IgG, binding preferentially occurred via. Fab, and were less protective (Figure 1.14). Genotypes belonging to patterns D and E produce an M-like protein, denoted Mrp, the C-terminal region of which features two leucine zipper domains, thought to mediate binding to immunoglobulins of this class (Hamada *et al.*, 2015; Krebs *et al.*, 1996).

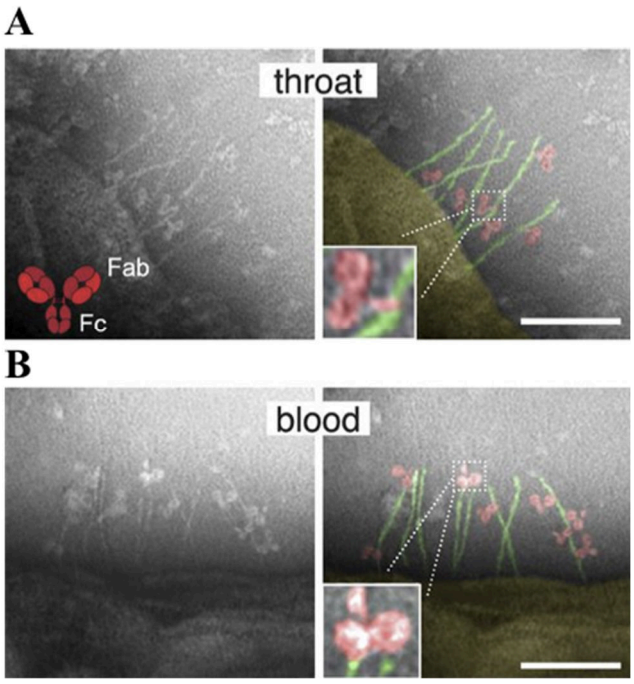

Figure 1.14: Differential, concentration-dependent binding of IgG by M1

Negatively stained electron micrographs showing differential IgG binding by M1 A). IgG (false-coloured red) binding by M1(false-coloured green) in the low-IgG conditions of the throat occurs preferentially by the Fab-domain, and are protective. B) In the blood, where IgG (false-coloured red) concentrations are high, M1 (false-coloured green) preferentially sequesters IgG by the Fc-domain, in the reverse orientation, reducing protection. Scale bars: 100nm. Image reproduced and adapted with permissions from 'Antibodies at the surface of *S. pyogenes* isolated from a patient with invasive disease' in Nordenfelt *et al.*, 2012, under CC BY-NC-SA 3.0.

1.7.10 **SpyCEP**

The *S. pyogenes* cell-envelope protease (SpyCEP) has broad truncating activity against a number of ELR+ chemokines, that is, chemokines such as IL-8 which feature a Glu-Leu-Arg motif at their N-terminus (Goldblatt *et al.*, 2019). These chemokines are fundamental is promoting neutrophil recruitment, diapedesis and margination to the site of infection in tissues (Bachelerie *et al.*, 2014; Zlotnik and Yoshie, 2012). In order for chemokines such as IL-8 to elicit the activation of this response, they are sequestered by micromolar affinity by glycosaminoglycans at the cell surface of leukocytes, as part of immune surveillance (Bachelerie *et al.*, 2014; Goldblatt *et al.*, 2019; Zlotnik and Yoshie, 2012). In *S. pyogenes* infection, in spite of often overwhelming bacterial burden, there can be a remarkable dearth of neutrophils at infectious foci (Edwards *et al.*, 2005; Zinkernagel *et al.*, 2008). This has been shown to be mediated by the cleavage of IL-8 and other chemokines by SpyCEP, resulting in a dimerised, less potent, and less efficacious molecule that is significantly less able to induce leukocyte recruitment and chemotaxis (Edwards *et al.*, 2005; Zinkernagel *et al.*, 2008; Turner *et al.*, 2009; Goldblatt *et al.*, 2019) (Figure 1.15). Indeed, cleaved IL-8 molecules are thus vastly impaired in their capacity to bind with cognate receptors (CXCR1 and CXCR2) on the leukocyte surface, markedly dampening the induction and recruitment of neutrophils (Goldblatt *et al.*, 2019). In addition, experimental evidence suggests that truncated IL-8 may be less able to be sequestered by the neutrophil glycocalyx, which maintains local concentrations of chemokines in close proximity to their cognate receptors (Goldblatt *et al.*, 2019). The corollary of these events *in-vivo* is a dramatic reduction in the capacity for IL-8, and other ELR+ chemokines to signal to neutrophils and promote their migration to the site of infection, in many cases, leaving *S. pyogenes* to proliferate, destroying host tissues and migrating to distal anatomical sites, unhindered by these crucial effectors of innate immunity (Edwards *et al.*, 2005; Goldblatt *et al.*, 2019; Turner *et al.*, 2009; Zinkernagel *et al.*, 2008).

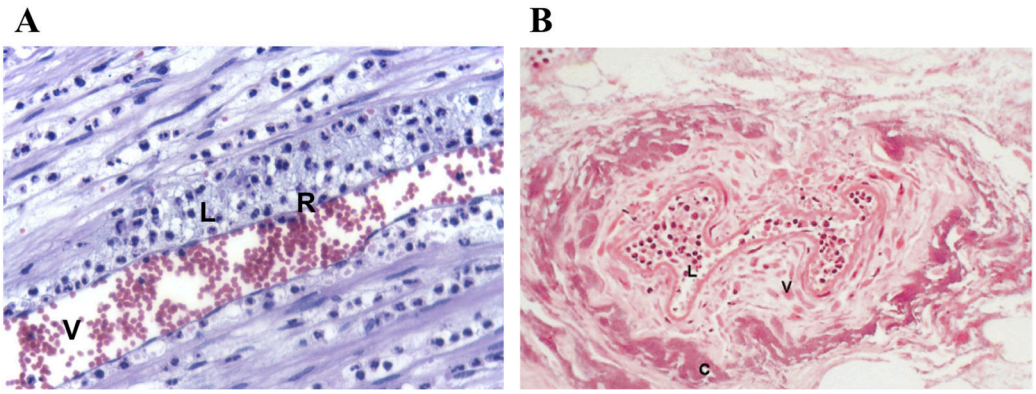

Figure 1.15: Neutrophil diapedesis and margination is abrogated by SpyCEP

A) Neutrophils follow inflammatory signals released by phagocytes and affected tissues in acute inflammation, 'rolling' 'R' along the luminal surface of the blood vessel 'V', before traversing the endothelial barrier, entering the tissues 'L' and travelling to the site of inflammation (not visible). B) *S. pyogenes* cell envelope protease (SpyCEP) cleaves ELR+ chemokines, notably IL-8, markedly reducing the capacity efficiency at which leukocytes 'L' to travel to infectious foci on receipt of these signals, to tackle *S. pyogenes* at infectious foci 'C' and control infection (Edwards *et al.*, 2005; Turner *et al.*, 2009). Image A: "Margination of neutrophils in acute inflammation.jpg" by Department of Pathology, Calicut Medical college is licensed under CC BY-SA 4.0. Image B taken from Edwards *et al.* 2005. N.b. Images are intended to illustrate a comparison, but are unrelated, differentially stained and magnified, and represent different tissues, and conditions.

1.7.11 **C5a peptidase**

Similarly to SpyCEP, the C5a peptidase of *S. pyogenes* (ScpA) is a highly conserved cell envelope protease that reduces the recruitment of neutrophils and other phagocytes to infectious foci (Yinduo *et al.*, 1996; Ji *et al.*, 1997; Lynskey *et al.*, 2017). This is achieved by degrading the chemotaxin C5a, and despite its name, C3 and C3a also (Cleary *et al.*, 1992; Lynskey *et al.*, 2017). By degrading C3, the downstream antibacterial effectors of this cascade, including C5a, are inhibited (Lynskey *et al.*, 2017). In crippling the complement cascade, *S. pyogenes* is able to reduce opsonisation, phagocytosis and clearance and is imbued with an enhanced ability to survive in whole human blood, disseminate to distal anatomical sites and persist (Ji *et al.*, 1998, 1997; Lynskey *et al.*, 2017; Yinduo *et al.*, 1996). Auxiliary functions of ScpA include promoting bacterial adhesion to epithelial and endothelial cells, and surprisingly, there is evidence of a further role in pathogenesis that is independent of either C3 or C5, which may or may not be related to the role of this virulence factor as an adhesin (Lynskey *et al.*, 2017).

1.7.12 **Streptococcal inhibitor of complement**

The streptococcal inhibitor of complement (SIC) is a secreted protein that interacts and interferes with the complement cascade and formation of the membrane attack complex (Åkessont *et al.*, 1996; Fernie-King *et al.*, 2001). Subsequent studies have shown that SIC also inhibits the activity of secreted eukaryotic antimicrobial molecules, including lysozyme and LL-37 (Fernie-King *et al.*, 2002; Pence *et al.*, 2010). SIC is also rapidly

internalised by human cells, wherein it disrupts the cytoskeleton (Hoe *et al.*, 2002). There are in excess of 300 alleles encoding SIC variants, and these variants have been shown to occur at high frequency during outbreaks, possibly due to selective pressures exerted by secreted immunoglobulins at the mucosal surface of the oropharynx (Hoe *et al.*, 1999). Indeed, it seems that pharyngeal carriage and passage of *emm*1 isolates in mice can yield a large number of allelic *sic* variants, though conversely, this level of variation was not observed by passaging the same isolate *in-vitro* (Hoe *et al.*, 1999). SIC variants were similarly not identified in a separate study, passaging isolates in a murine model of invasive infection (Stockbauer *et al.*, 1998). Although SIC is best studied in genotype *emm*1 *S. pyogenes*, it has subsequently been occasionally detected in a number of other *emm*-types (Ma *et al.*, 2002).

1.7.13 **Secreted phospholipase**

S. pyogenes isolates carrying a prophage encoding the streptococcal superantigen *speK* often also carry the contiguous gene *sla*, which encodes a secreted phospholipase A_2. This virulence factor is arguably most thoroughly studied in *emm*3, wherein it has been found to strongly promote colonisation and infection the pharynges in cynomolgus macaques and mediate adhesion to human epithelial cells *in-vitro* (Sitkiewicz *et al.*, 2006). SLA may play a crucial role in the notably invasive phenotype of *emm*3 infection, contributing to the clinical features of STSS, notably leukostasis, endothelial injury and vascular leakage, via. the arachidonic acid pathway (Oda *et al.*, 2017). Isogenic *sla* mutants also exhibit an attenuation of virulence in murine infection models, producing smaller infectious foci (Sitkiewicz *et al.*, 2006).

1.7.14 **DNases of *S. pyogenes***

Another major mechanism of virulence in *S. pyogenes* is the production of nucleic acid degrading enzymes, DNases. There are presently 8 described DNases, 6 encoded by genes carried on prophage; *sda1, sda2*, *spd1*, *spd3*, *spd4, sdn* and 2 which are encoded by genes within the bacterial chromosome; *spdB* and *spnA* (Remmington and Turner, 2018). The majority of the DNases in *S. pyogenes* are secreted enzymes, however *spnA* is cell wall-anchored, featuring an LPXTG motif (Hasegawa *et al.*, 2010). At least two DNase genes are produced by all *S. pyogenes* isolates (Remmington and Turner, 2018), inference that these enzymes confer an evolutionary advantage. Perhaps the best supported role for DNases in infectious processes is the degradation of NETs (Buchanan *et al.*, 2006; Sumby *et al.*, 2005c; Walker *et al.*, 2007), and it is in this capacity that these virulence factors are believed to enhance virulence in both vertebrate and invertebrate models (Buchanan *et al.*, 2006; Sumby *et al.*, 2005; Hasegawa *et al.*, 2010; Chalmers *et al.*, 2017) (Figure 1.16). DNases also appear to have a role in the propagation of infection in the nasopharynx (Al-Shahib *et al.*, 2016; Sumby *et al.*, 2005c). Similarly, there is evidence to suggest that these nucleases may serve to dampen host immunity by preventing TLR-9 signalling and the recognition of unmethylated CpG-rich bacterial DNA by macrophages, degrading it before it can be detected (Uchiyama *et al.*, 2012). When one considers the nature of pyogenic streptococcal infections, it

is reasonable to surmise that in the midst of an infectious focus, wherein *S. pyogenes* is likely to be deprived of nutrients and possibly facing imminent threat of elimination by host immunity, DNases might present a useful escape mechanism – liquefying the extracellular milieu and facilitating escape and spread. Indeed early *ex-vivo* experiments would appear to support this, though recent proteomic analyses on aspirates derived from invasive infections caused by multiple *emm*-types, did not detect any DNases (Sherry, Johnson and Tillett, 1949; Edwards *et al.*, 2018).

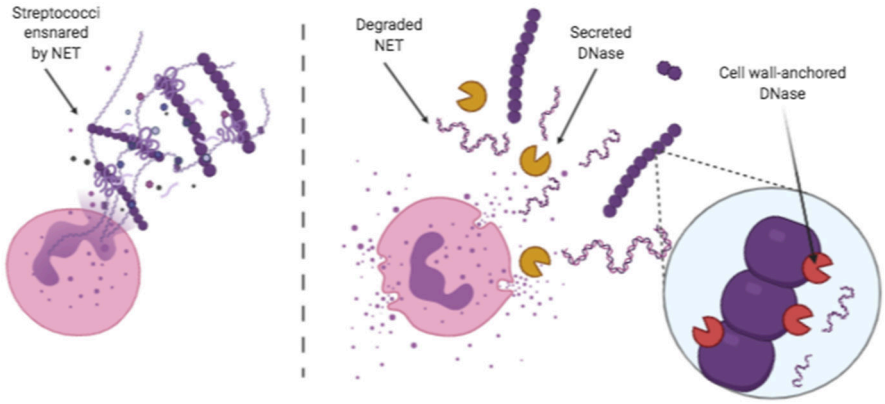

Figure 1.16: Streptococcal DNases promote expedition of bacteria ensnared in NETs

By producing both cell wall-anchored and secreted DNases, *S. pyogenes*, and indeed other streptococcal species, are imbued with a potent mediator of neutrophil interference, by degrading neutrophil extracellular traps both those deposited in the immediate extracellular environment around the bacterium, but also beyond infectious foci by diffusing through the extracellular milieu (Remmington and Turner, 2018). Figure created by the author with BioRender.com.

1.7.15 **Streptococcal superantigens**

There are presently 13 described streptococcal superantigens in *S. pyogenes*, 5 of these are encoded on the bacterial chromosome: *smeZ, speG*, *speJ*, *speQ* and *speR.* The remaining 8 are encoded by genes carried on prophage: *ssa, speA, speC, speH, speK, speI, speL* and *speM* (Commons *et al.*, 2014; Reglinski *et al.*, 2019). These exist as multiple alleles within and between *emm*-types and other streptococcal species (Commons *et al.*, 2014)

At the cellular level, superantigens bypass ordinary epitope processing and presentation, instead binding irreversibly to the variable region of the β-chain on specific T-cell subsets. This binding triggers massive expansion of these subsets, and precipitates the release of large quantities of proinflammatory cytokines. The precise role of streptococcal superantigens in infection is debated, but their carriage is presumably advantageous to the bacterium in some way given that every isolate tested to date has encodes at least one allele. In experimental models of superficial disease, streptococcal superantigens in *S. pyogenes* have been demonstrated to promote nasopharyngeal colonisation (Kasper *et al.*, 2014; Zeppa *et al.*, 2017), reduce immunoglobulin secretion, impair

chemotactic responses, and induce B cell lysis (Davies *et al.*, 2019) (Figure 1.17). In systems designed to emulate invasive manifestations, it would appear that these toxins mediate many of the clinical features of STSS and systemic inflammation (Sriskandan *et al.*, 1996). On a larger scale, streptococcal superantigens have been recurrently epidemiologically associated with scarlet fever and streptococcal pharyngitis (Chalker *et al.*, 2017; Silva-Costa *et al.*, 2014) and occasionally rheumatic fever (Commons *et al.*, 2014).

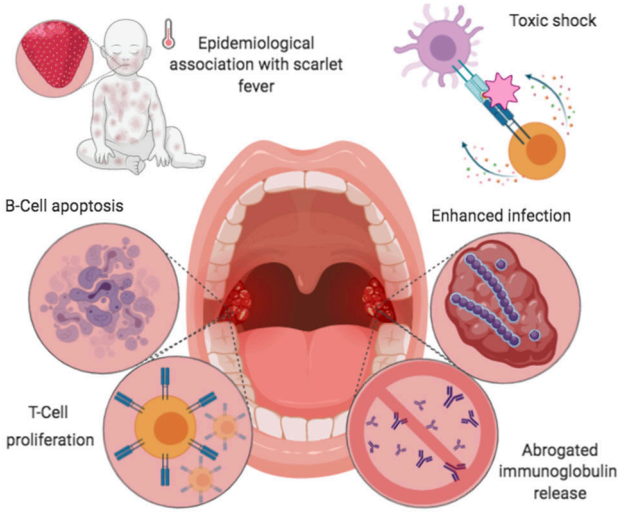

Figure 1.17: Streptococcal superantigens: selected mechanisms of action

The precise role of the streptococcal superantigens in the pathogenesis of disease in *S. pyogenes* is not completely clear. As potent stimulators of T-cell subsets, thus promoting copious release of pro-inflammatory cytokines, it is thought that these molecules are major mediators of STSS (Commons *et al.*, 2014). Streptococcal superantigens also appear to be able to modulate host immunity to their benefit, inducing B-Cell death, T-Cell proliferation, and markedly reducing the release of antibodies in the human tonsil (Davies *et al.*, 2019). They may also play a role in colonisation and infection of these tissues (Afshar *et al.*, 2017; Brouwer *et al.*, 2020). Figure created by the author with BioRender.com.

1.8 Transcriptional regulation in *S. pyogenes*

1.8.1 Two-component regulatory systems

To date, 13 two-component regulatory systems (TCS) have been described in *S. pyogenes* (Kreikemeyer *et al.*, 2003). As in other bacteria, these are typically a transmembrane sensor histidine kinase, and a cognate cytoplasmic DNA-binding response element, which is regulated by phosphorylation. The sensor kinase possesses auto-kinase activity, which is enhanced or suppressed in response to environmental stimuli (Jung *et al.*, 2012).

1.8.1.1 CovRS

The TCS CovRS (control of virulence regulator/sensor kinase) is a cell membrane-associated histidine kinase comprised of two proteins; the membrane bound sensor kinase CovS and the cytoplasmic response regulator

CovR. To exert control over its regulatory targets, CovS undergoes autophosphorylation at its cytoplasmic domain, promoting the subsequent phosphorylation of CovR, thereby increasing its affinity for specific binding sites within the promoters of target gene (Horstmann *et al.*, 2018). These targets are many, and it is thought that CovRS regulates approximately 15% of the *S. pyogenes* genome, mostly acting to repress target genes (Graham *et al.*, 2002). A large number of these regulatory targets play pivotal roles in virulence and infection, among these are the cysteine protease SpeB, the HA capsule biosynthesis operon, SIC, NADase and co-transcribed streptolysin O, SpeA, the DNases *spd3* and *sda*, SpyCEP and C5a peptidase (Federle, McIver and Scott, 1999; Graham *et al.*, 2002; Sumby *et al.*, 2006; Turner *et al.*, 2009). Indeed, there is evidence for the longitudinal coordination of the CovRS regulome throughout pharyngeal infection (Virtaneva *et al.*, 2005).

A number of factors have been shown to signal through CovRS, notably extracellular magnesium, and the human antimicrobial peptide LL-37 (Figure 1.18), although at concentrations below those required to exert antibacterial activity, and apparently independent of this (Gryllos *et al.*, 2008, 2007). Virulence modulating homologues of CovRS are found in many other pathogenic streptococci of clinical and veterinary significance, including *S. agalactiae*, *S. dysgalactiae* sbsp. *equisimilis* and *S. suis,* although in the latter*, covR* is not accompanied by a cognate sensor kinase, and is an orphan regulator (Jiang *et al.*, 2008; Pan *et al.*, 2009). A homologous system is present in *S. equi* sbsp. *zooepidemicus,* though its phenotypic influence remains as yet unexplored (Velineni and Timoney, 2015). In *S. pyogenes,* CovRS itself is negatively regulated by the orphan regulator RocA (Biswas and Scott, 2003; Lynskey *et al.*, 2019b).

Mutations in this regulator are thought to arise during invasive infection, possibly due to pressure from host immunity, and are detected in as many as 40% of invasive *S. pyogenes* isolates (Ikebe *et al.*, 2010; Li *et al.*, 2014; Shea *et al.*, 2011; Walker *et al.*, 2007). Because of the sheer breadth of genes regulated by CovRS, and variations thereof between *emm*-types, it is challenging to make general assessments on the overall impact of loss of function mutations in *covR* and/or *covS.* Most sources agree, however, that CovRS mutants exhibit increased virulence, and are associated with increased mortality and morbidity in humans (Cole *et al.*, 2006; Sumby *et al.*, 2006; Walker *et al.*, 2007; Turner *et al.*, 2009).

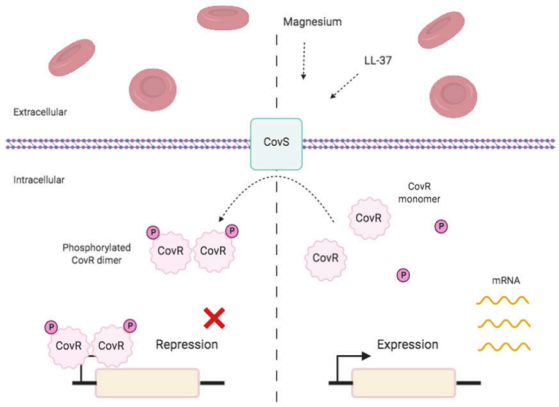

Figure 1.18: Transcriptional regulation by CovRS

CovRS comprises two components, CovS, the sensor kinase, and CovR, the response regulator. CovRS regulates ~15% of the *S. pyogenes* genome, and is largely repressive in its influence. This is achieved by phosphorylation of the cytoplasmic effacing domain of CovS, which in turn, promotes the phosphorylation of CovR. CovS is receptive to signalling by host-derived factors, such as LL-37 and extracellular magnesium (Gryllos *et al.*, 2008). The phosphorylation of CovR enhances the affinity for CovR and binding sites within the streptococcal genome, which include a large number of virulence factors, such as *speB*, *cepA* and *speA* (Sumby *et al.*, 2006; Turner *et al.*, 2009), and in encapsulated isolates, the HA capsule, giving encapsulated isolates a glossy, mucoid appearance (Gryllos *et al.*, 2008). Figure created by the author with BioRender.com

1.8.1.2 Ihk/Irr

In the context of virulence and pathogenesis, evidence suggests that Ihk/Irr (*isp* adjacent histidine kinase/*isp* adjacent response regulator) is involved in the transcriptional response by *S. pyogenes* following phagocytosis by macrophages and neutrophils, and resistance to killing therein (Hertzén *et al.*, 2012; Voyich *et al.*, 2004, 2003). In support of these observations, this TCS has been found to regulate a number of genes pertaining to resistance to ROS and neutrophil-associated antimicrobial peptides in addition to genes involved in cell wall biogenesis (Voyich *et al.*, 2004). Well characterised, specific virulence factors in the Ihk/Irr regulome are surprisingly few, but include the DNases *spd3* and *spdB,* the haemolysin SLS and *fbp* (fibronectin binding protein), as revealed by global transcription studies (Voyich *et al.*, 2004). Orthologues of *Ihk/Irr* are found in *S. suis* and have similarly been implicated in resistance to killing by macrophages and enhanced survival in environments with low pH and abundant oxidative stressors *in-vitro* (Han *et al.*, 2012).

1.8.1.3 FasBCAX

Evidence suggests that the fibronectin/fibrinogen-binding/haemolytic activity/streptokinase (*fas*) operon mediates in *S. pyogenes* the transition from colonisation and adhesion stages of infection, to those of dissemination and spread. The system comprises four genes, encoding a putative response regulator, two predicted sensor kinases, and a small regulatory RNA (sRNA) effector molecule, denoted *fasA*, *fasB*, *fasC*, and *fasX*, respectively (Kreikemeyer *et al.*, 2001; Miller *et al.*, 2014). The hypothesis that the *fasBCAX* operon promotes the two opposing phenotypes of colonisation and dissemination is largely born of the relationship of the regulator with

adhesins such as Prtf1 and Prtf2 (Danger *et al.*, 2015b), and minor collagen-binding minor pilus protein, Cpa (Liu *et al.*, 2012), while at the same time regulating the expression of the well-studied spreading factor, streptokinase (Ramirez-Peña *et al.*, 2010). Indeed, fasX is able to bind to mRNAs and prevent translation and protein expression, as is the case for Cpa, and this is achieved by obstructing access for ribosomes at ribosomal binding sites on the transcript (Liu *et al.*, 2012). FasX may also bind to streptokinase mRNAs to promote stability, thereby increasing transcript abundance and promoting a ~10-fold increase in the production of this virulence factor (Ramirez-Peña *et al.*, 2010). The regulatory influence of this operon extends to a number of virulence genes, may be growth-phase dependent (Kreikemeyer *et al.*, 2001), and as is the case with other regulatory networks in *S. pyogenes*, appears to exhibit variability between isolates and genotypes (Danger *et al.*, 2015a; Perez *et al.*, 2009).

1.8.2 Standalone Transcriptional Regulators

1.8.2.1 Rgg/RopB

Rgg, also known as RopB (regulator of protease B) has been shown to modulate various processes in *S. pyogenes*. The regulation of *speB* by Rgg is well demonstrated, and it is required for expression of this toxin (Olsen *et al.*, 2012). As such, a number of lines of investigation have demonstrated a central role in virulence and disease association, notably pharyngitis (Carroll *et al.*, 2011; Olsen *et al.*, 2012). Other virulence genes that are directly or indirectly regulated by Rgg include the secreted DNases *spdB* and *spd3* and the gene encoding NADase, *nga* (Anbalagan and Chaussee, 2013; Dmitriev *et al.*, 2006). Additional work is indicative of a role in metabolism also, specifically the metabolism of carbon sources in some *emm*-types (Dmitriev *et al.*, 2006).

1.8.2.2 Mga

The Mga (multi-gene regulator of GAS) regulon includes approximately 10% of the *S. pyogenes* genome, and includes the regulators RofA and Rgg/RopB. In sensing carbohydrates and other sugars in the extracellular milieu, Mga plays a critical role in regulating the expression of a number of virulence factors, modulating phenotypes that promote survival in soft tissues and human blood, in addition to the formation of biofilms (Cho and Caparon, 2005; Graham *et al.*, 2005). Other signals reportedly include temperature, environmental gases, salinity and iron levels (Hondorp and McIver, 2007; McIver *et al.*, 1995; Podbielski *et al.*, 1992). Because of the sensitivity of Mga to a multitude of signals, it has been proposed that Mga allows *S. pyogenes* to respond to its extracellular environment and new niches, adjusting to a transcriptional profile that is optimal for survival (Cho and Caparon, 2005; Graham *et al.*, 2005; Musser and DeLeo, 2005). With such a strong association with such a large portion of the genome, including other regulators, it is perhaps unsurprising that multiple lines of investigation have associated *Mga* mutants with attenuated virulence (Kihlberg *et al.*, 1995; Terao *et al.*, 2001). Longitudinal transcriptome studies of pharyngeal infection in cynomolgus macaques found Mga activity to be at its greatest when symptoms were most severe (Virtaneva *et al.*, 2005). Mga may exert its influence through responding to the

abundance of sugars by phosphorylating conserved histidine residues in target sequences, similarly to other carbohydrate activated virulence regulators in other bacterial pathogens, notably the AtxA virulence regulator, an Mga homologue in *Bacillus anthracis* (Hondorp and McIver, 2007).

1.8.2.3 RofA and RofA-like protein type regulators

RofA and RofA-like protein type regulators (RALPs) constitute a family of four stand-alone transcriptional regulators, RofA, Nra, RALP-3 and RALP-4 (Granok *et al.*, 2000). RofA has multiple functions, including sensing oxidative stressors and environmental gases, mostly acting as a positive regulator of target genes (Fogg and Caparon, 1997). Mutations in *rofA* have been linked to increased toxin production, including SLS, M-protein and SpeA in some *emm*-types (Beckert *et al.*, 2001; Lynskey *et al.*, 2019a). RofA mutants have also been shown to exhibit reduced adherence to- and invasion of human epithelial cells, possibly due to modulations in the production of toxins and adhesins, including *emm* and *hasB* (Beckert *et al.*, 2001). A homologue of *rofA,* denoted *nra*, exhibits ~60% DNA sequence similarity to *rofA* and also has regulatory function, though conversely, largely as a negative influencer (Podbielski *et al.*, 1999). A survey of over 100 isolates detected either *rofA* or *nra* in all strains, and similarly to RofA, Nra similarly regulates a number of genes involved in adhesion and virulence, including those that are also regulated by RofA, including *speA*, *speB* and the *mga* regulon (Luo *et al.*, 2008). Both regulators are thought to regulate the biosynthesis of pilus, and the inactivation of *nra* is associated with attenuated virulence (Lizano *et al.*, 2008; Luo *et al.*, 2008).

1.9 Mobile genetic elements & *S. pyogenes*: impact on host phenotype

As the prophage complement within a given *emm*-type is fairly conserved, it is perhaps not surprising that their loss and/or gain can lead to striking phenotypic changes as well as changes to the circulating bacterial population and sometimes devastating epidemiological shifts. Particularly with the application of WGS, the population biology of *S. pyogenes* is being scrutinised with increasing resolution, and a number of studies have successfully identified the emergence of nascent successful lineages and have demonstrated relationships between these events and epidemiological surveillance data. Often in *S. pyogenes*, such events are catalysed by changes in prophage complement and therefore the constellation of phage-encoded virulence factors in a given *emm*-type or successful clone (Beres *et al.*, 2006; Davies *et al.*, 2015; Al-Shahib *et al.*, 2016; Afshar *et al.*, 2017; Brouwer *et al.*, 2020).

1.9.1 Enhanced toxin expression at the host-pathogen interface

Multiple lines of scientific enquiry have demonstrated that prophage-encoded DNases and superantigens are expressed *in-vivo* (Kazmi *et al.*, 2001; Sriskandan *et al.*, 1999; Virtaneva *et al.*, 2005) and on co-culture with human pharyngeal cells (Broudy, Pancholi and Fischetti, 2001, 2002; Banks, Lei and Musser, 2003) models of infection, and an appealing hypothesis is that these virulence factors contribute to increased fitness of the host bacterium. The streptococcal superantigen SpeA has been shown to dysregulate the function of tonsillar immunity,

specifically driving a proliferative T-cell phenotype, triggering B-cell apoptosis, and abrogating the release of IgA, IgM and IgG. Taken together, the data suggest that expression of SpeA may confer a survival advantage to superantigen expressing bacteria at the host-pathogen interface (Davies *et al.*, 2019).

1.9.2 Enhanced nasopharyngeal shedding

A nationwide increase in invasive disease in the United Kingdom was attributed to the emergence of a novel lineage within the *emm*3 *S. pyogenes* population (Al-Shahib *et al.*, 2016). Following genomic analyses, it was established that the acquisition of an additional *speC-spd1* encoding prophage, ΦUK-M3.1, and the simultaneous loss of two canonical *emm*3 prophage, the non-toxigenic Φ315.1, and the *ssa*-encoding prophage, Φ315.2, had occurred among isolates belonging to the emergent lineage. Isolates belonging to the emergent lineage were characterised by significantly greater DNase activity, in addition to prolonged infection and enhanced nasopharyngeal shedding in a murine model of disease (Al-Shahib *et al.*, 2016) (Figure 1.19). Interestingly, a similar *emm*3 variant had emerged twice previously in North America, and was also associated with transient increases in invasive disease (Beres *et al.*, 2004). There also appears to be a role for synergy between prophage-associated toxins, specifically Spd1 and SpeC, in promoting nasopharyngeal infection by genotype *emm*12 isolates, and a reliance on host-derived factors for their activation and potentiation (Brouwer *et al.*, 2020).

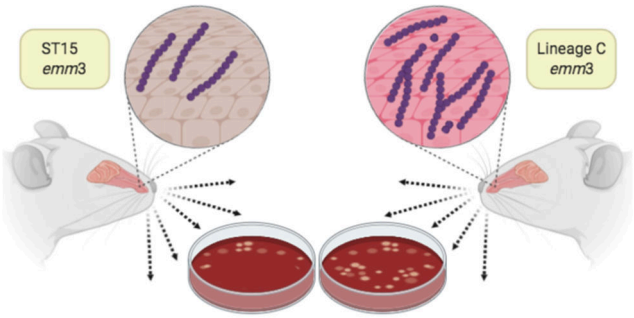

Figure 1.19: Enhanced nasopharyngeal infection associated with epidemic *emm*3 lineage

A sudden upsurge in invasive *S. pyogenes* infections in the United Kingdom in 2008/2009 was associated with an emergent lineage of *emm*3, denoted Lineage C. Lineage C *emm*3 had lost two of the canonical *emm*3 prophage, including one carrying *ssa*, but had gained a new prophage carrying *speC*/*spd1* (Al-Shahib *et al.*, 2016). The lineage, though no more virulent than ancestral strains, Lineage C was found to be associated with enhanced nasopharyngeal infection and airborne shedding in murine models. This was attributed to the acquisition of the novel prophage, specifically, the DNase, Spd1 (Afshar *et al.*, 2017). Figure created by the author with BioRender.com.

1.9.3 Clonal expansion and extinction

Modulations to the complement of mobile genetic elements present in the genome of the *S. pyogenes* can lead to striking changes to bacterial population dynamics. Acquisition of the *speA*-encoding prophage Φ315.5 in *emm*3 *S. pyogenes* was shown to be associated with clonal expansion, whereas isolates lacking this prophage were more

likely to undergo clonal extinction (Beres *et al.*, 2004). A possible explanation for the increased fitness of these isolates is that genotype *emm*3 isolates that are positive for *speA* have been associated with recurrent episodes of pharyngitis, and might therefore be more likely to persist in a population (Musser *et al.*, 1992). In Hong Kong and mainland China, an ongoing nationwide upsurge in scarlet fever has been attributed to multiple distinct lineages of genotype *emm*12 *S. pyogenes* expressing high levels of the streptococcal superantigen SSA (Walker, 2015). A significant association between scarlet fever cases and isolates carrying two copies of *ssa* on two separate prophages, ΦHKU.vir (*speC, spd1* and *ssa*) and ΦHKU.ssa (*ssa* alone) was found, in addition to macrolide resistance genes, also affiliated with mobile genetic elements. In the United Kingdom, the dramatic resurgence of scarlet fever in recent years was attributed to a novel M1T1 variant expressing high levels of the prophage-encoded superantigen SpeA (Lynskey *et al.*, 2019a). Genome sequencing has also demonstrated that acquiring novel allelic variants of specific prophage-encoded virulence factors can also favour the emergence of successful lineages with a significantly different phenotype to parental isolates (Banks *et al.*, 2002; Barnett *et al.*, 2019; Wilkening and Federle, 2017).

1.9.4 **Remodelling of the bacterial chromosome**

Not all prophage and prophage-like elements encode genes with an immediately obvious role in virulence. In fact, a number of these elements either integrate at sites which result in pronounced changes to the genome and phenotype of the host bacterium. In genotype *emm*3, large-scale genomic inversions have been reported, mediated in part by homologous recombination between prophage genes, yielding not only novel variants of prophage by virulence factor translocation, but large scale rearrangements of the bacterial chromosome (Nakagawa *et al.*, 2003). Although this report is not complimented by epidemiological analyses, it was hypothesised that the resultant *emm*3 variant was associated with a coincident resurgence in rheumatic fever and invasive infections in Japan occurring from the 1990's onwards. Later work in North America found no association between this variant and any infection type, but demonstrated that the generation of either genomic conformation occurs frequently throughout growth *in-vitro* (Beres *et al.*, 2004).

1.9.5 **Regulation of gene expression at attachment sites**

The SpyCI of *S. pyogenes* are reported to regulate the expression of the DNA MMR operon, through a dynamic process of integration and excision. During exponential growth, SpyCI is in its excised state, permitting the transcription of the MMR operon by a common promoter upstream of *mutS*. During stationary phase, the element is more stably integrated, blocking transcription of genes downstream of *mutS*, and promoting a more mutable phenotype (Nguyen and McShan, 2014; Scott *et al.*, 2012) (Figure 1.20). In subsequent work, isolates cured of this element not only were unable to attain the mutator phenotype, but were found to have significant transcriptional modulations to a large number of genes, including increased expression of many pertaining to virulence, notably *emm, nga*, *slo* and the capsule biosynthesis operon (Hendrickson *et al.*, 2015). Whether these

changes would be observed following natural loss of the element *in-vivo* is unclear. Further investigation also illuminated a mechanism by which these chromosomal islands can rescue aberrant and irreversible induction of a mutator phenotype in the event of their decay. In an *emm*5 isolate in which the integrase of the SpyCIM5 element is a pseudogene preventing excision and integration from the bacterial chromosome, it was hypothesised that the isolate would exhibit a persistent mutator phenotype. However, this was found not to be the case and closer scrutiny revealed that in fact, the integrase ORF serves as an alternative cryptic promoter for this operon, enabling the transcription of *mutL* and other MMR genes, despite their permanent separation from the native promoter (Scott *et al.*, 2012). Prophage that do not appear to carry any virulence genes can also be found integrated into the CRISPR loci of *emm*3 isolates, which may impact the expression of these genes (Beres *et al.*, 2002). Presumably, this confers some advantage to the bacterium, as although this has at times been lost by epidemic *emm*3 lineages, these have not persisted and this appears to be a transient, less stable phenomenon (Al-Shahib *et al.*, 2016; Stephen B. Beres *et al.*, 2004).

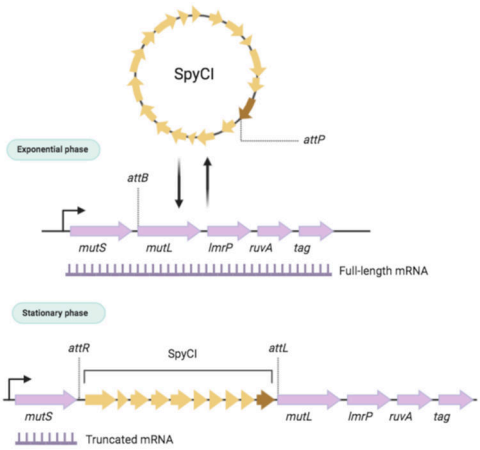

Figure 1.20: SpyCI mediated regulation of gene expression in *S. pyogenes*

The growth-phase dependent regulation of the DNA MMR operon in an SF370 *emm*1 isolate was attributed to the mobility of the *S. pyogenes* chromosomal island SpyCI^{M1}. SpyCI^{M4}, ordinarily integrated within the operon between *mutS* and *mutL*, was excised during exponential growth, and the operon transcribed. During stationary phase, the element was stably integrated and did not excise, separating genes downstream of *mutL* from the native promoter, and blocking transcription. This was shown to promote a more mutable phenotype, ostensibly to overcome challenges, such as overcrowding, or nutrient deprivation, by increasing the frequency at which successful escape mutants might emerge (Nguyen and McShan, 2014; Scott *et al.*, 2008). Figure created by the author with BioRender.com.

1.9.6 Enhanced adhesion and colonisation

In a humanised murine model of nasopharyngeal infection, it was demonstrated that the prophage-encoded streptococcal superantigen *speA* was essential for colonisation with a genotype *emm*18 isolate (Kasper *et al.*, 2014), providing additional insight into the benefit of carrying these elements. Additional insight into the role of

superantigens in the propagation of infection is provided by a separate study by the same authors, which demonstrates that superantigens are instrumental in the manipulation of specific T cells to promote colonisation in *S. pyogenes,* using SpeA and heterologous expression of staphylococcal enterotoxin B as a paradigm (Zeppa *et al.*, 2017). Very recent work suggests that among *emm*12 isolates, SpeC and Spd1, which are invariably contiguous on the same prophage, may act in synergy to bolster nasopharyngeal infection in mice (Brouwer *et al.*, 2020). These findings echo earlier, similar reports on the possible importance of *speC*/*spd1*-associated prophage and nasopharyngeal infection that was observed among an epidemic *emm*3 lineage, as aforementioned (Afshar *et al.*, 2017). The streptococcal adhesin AspA, encoded by a gene present on the ICE RD2 appears to plays a major role in adhesion and the propagation of upper respiratory tract infection (Franklin *et al.*, 2013). Isogenic *aspA* mutants are significantly less able to adhere to epithelial cells *in-vitro*, and more prone to clearance from the upper respiratory tract of mice. It was also shown that AspA potentiates resistance to phagocytosis by macrophages.

1.9.7 **Enhanced virulence**

It has been suggested that the resurgence of *S. pyogenes* infections globally, coupled with the observation that the bacterium appeared to be gaining virulent potential, might be explicable by the acquisition of *speA* by *emm*1 and *emm*3 genotypes, among others (Marshall and Kelch, 1986). The RD2 encoded Spy1336/R28 has been implicated in increased virulence. Remarkably, a single nucleotide insertion in the intergenic region between *R28* and an upstream transcriptional regulator *Spy1337* was found to significantly enhance the binding of the regulator to the intergenic homopolymeric nucleotide tract and subsequently increase the expression of not only *R28*, but a number of other chromosomally-encoded virulence factors (Kachroo *et al.*, 2019). Isolates with this single nucleotide insertion were significantly more virulent in a murine model of necrotising myositis, and were more resistant to killing by human PMN leukocytes (Kachroo *et al.*, 2019) (Figure 1.21). The insertion and enhanced virulence were not associated with any specific population or location, however, suggesting that possibly, this phenotype was not overtly associated with overall enhanced fitness. Given the association between *emm*28, the female genitourinary tract and puerperal sepsis, perhaps a less virulent, more transmission and colonisation oriented phenotype is favourable (Luca-Harari *et al.*, 2009). Yet more recently, and indeed conversely, a transposon mutant library of *emm*28 *S. pyogenes* screening genes that contributed to fitness during the first 24-hours of colonisation in a cynomolgus macaque model, did not identify any RD2 associated genes (Zhu *et al.*, 2020). This might be due to the timeframe of sampling (only genes contributing at 24 hours), perhaps issues with recapitulation between models (notably mice, cell culture and macaques), or due to heterogeneity within the inoculum wherein wild type strains within the inocula compensating for abolished expression of secreted proteins in mutant strains harbouring transposons.

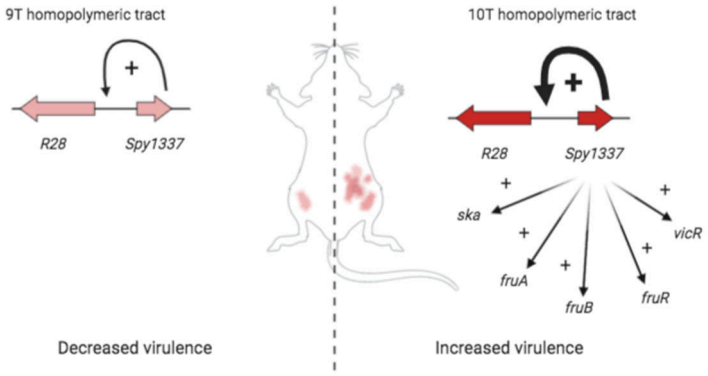

Figure 1.21: Enhanced virulence associated with the insertion of a single nucleotide

The insertion of a single nucleotide within intergenic sequence between R28 and an upstream transcriptional regulator was associated with a dramatic enhancement in the virulent potential of *emm*28 isolates. Transcriptional changes extended far beyond the mobile genetic element, RD2, to a number of chromosomal genes, including well-studied virulence factors such as streptokinase. This was attributed to enhanced binding of the transcriptional regulator Spy1337 to the intergenic sequence (Kachroo *et al.*, 2019). Figure created by the author with BioRender.com.

1.10 Conclusions and research justifications

S. pyogenes is a protean, and highly dynamic human pathogen, with a seemingly boundless capacity for complex interplay with host immunity through the expression of a vast armamentarium of virulence factors, evoking a similarly broad spectrum of clinical manifestations (Carapetis *et al.*, 2005). Unusually, *S. pyogenes* is not prominent in concerning trends of antimicrobial resistance like a number of other clinically significant Gram-positive bacterial agents such as *Staphylococcus aureus, Enterococcus* spp., and *S. pneumoniae* (Cattoir, 2015). Unlike these microorganisms however, the epidemiology of *S. pyogenes* in recent years has been tempestuous, particularly in the United Kingdom and the far east, which have seen the dramatic resurgence of scarlet fever and antecedent increases in invasive disease (Afshar *et al.*, 2017; Al-Shahib *et al.*, 2016; Banks *et al.*, 2002; Ben Zakour *et al.*, 2015; Chalker *et al.*, 2017; Efstratiou and Lamagni, 2016; Lynskey *et al.*, 2019a; Tse *et al.*, 2012; Walker, 2015). These events are being retrospectively delineated with increasing resolution by the application of WGS, providing a fascinating but worrisome insight into this highly adaptable and often unpredictable pathogen, the behaviour of which appears to be inextricably linked to the acquisition, loss and regulation of mobile genetic elements, which certainly deserve further study.

1.10.1 **Hypotheses**

Genome analyses of *S. pyogenes* isolates submitted to the British Society for Antimicrobial Chemotherapy from multiple sites across England and Wales was recently undertaken by this laboratory (Turner *et al.*, 2019b). The genotype *emm*4 isolates in this collection were found to cluster with either of the two completed reference genomes, MGAS10750 (Beres and Musser, 2007) and MEW427 (Jacoba *et al.*, 2016), forming two small lineages

(Remmington *et al.* 2020) in corroboration with a recent description (DebRoy *et al.*, 2018). Gubbins SNP analysis revealed four areas of putative recombination within regions encoding four mobile genetic elements in those isolates clustering with MEW427. These included the prophage Φ427.1/10750.1, Φ427.2/10750.2 and Φ427.3/10750.3, encoding *speC*/*spd1*, *spd3* and *ssa*, respectively, and the integrative and conjugative element SpyCI, also known as Φ10750.4. The ability of prophage to excise and thus elicit a gene dosing effect may play an important role in the expression of cognate virulence factors in *S. pyogenes* (Banks *et al.*, 2016; Broudy *et al.*, 2002, 2001) and other bacteria (Balasubramanian *et al.*, 2019). Φ10750.4 has been shown to regulate the expression of the DNA MMR operon in a number of *emm*-types (Scott *et al.*, 2012). This is achieved a manner that is growth-phase dependent, regulated by the dynamic integration and excision of the element from the bacterial chromosome (Hendrickson *et al.*, 2015; Nguyen and McShan, 2014; Scott *et al.*, 2012). It was therefore hypothesised that these elements may be unable to excise from the chromosome in some genotype *emm*4 isolates, but not others, and thus investigating the molecular biology of these isolates might serve as a valuable paradigm for understanding the regulation and expression of prophage-encoded virulence factors in *S. pyogenes* and the impact on bacterial phenotype.

1.10.2 **Project aims**

- To characterise the specific prophage genes and modules that have been subject to deletion or loss in MEW427-like isolates relative to MGAS10750-like isolates
- To establish whether the excision potential of prophage elements in MEW427-like isolates was different to those in MGAS10750-like isolates
- To establish when the prophage-encoded virulence factors are expressed
- To determine whether or not the expression of cognate bacterial factors encoded by the prophage of genotype *emm*4 is subject to a gene dosing effect
- To characterise the lineages in our *emm*4 collection and determine any potential phenotypic impact of prophage gene loss on these isolates
- To investigate whether environmental stimuli are able to induce the differential expression of prophage-encoded virulence factors in these isolates

Chapter 2

2. Materials and Methods

2.1 Materials

2.1.1 *S. pyogenes* isolates

British Society for Antimicrobial Chemotherapy (BSAC) isolates were collected by the society's Bacteraemia Resistance Surveillance Program, from multiple hospitals across England, between 2001 and 2011 (Turner *et al.*, 2019b). The programme collected 344 isolates in total, of which *emm*4 (n=10) comprised ~2.9% and all were included in the present work. A smaller sample (n=5) of *emm*3 and *emm*12 and a single *emm*6 were also selected from the collection. Genotype *emm*89 H293 was isolated in 1995 by the Central Public Health Laboratory (Colindale, UK) (Sriskandan *et al.*, 2000), is well-characterised (Lamb *et al.*, 2018a; Reglinski *et al.*, 2015; Sriskandan *et al.*, 2000; Turner *et al.*, 2009; Turner *et al.*, 2012; Unnikrishnan *et al.*, 2002), and has been WGS to completion, revealing no integrated prophage (Turner *et al.*, 2015). Genotype *emm*1 H305 is an M1T1 scarlet fever isolate collected in the 1880s, and is also robustly characterised, including for superantigenicity (Lamb *et al.*, 2018; Sriskandan *et al.*, 2001, 1999; Turner *et al.*, 2009; Unnikrishnan *et al.*, 1999).

Table 2.1: *S. pyogenes* isolates used in this study

Isolate	*emm*-type*	Subtype†	Region	Date isolated
BSAC_bs192[a]	4	4	Sheffield	2001
BSAC_bs400[a]	4	4	Manchester	2002
BSAC_bs468[a]	4	4	Southampton	2002
BSAC_bs472[a]	4	4	London	2002
BSAC_bs696[a]	4	4	London	2003
BSAC_bs916[a]	4	4	Manchester	2004
BSAC_bs105[a]	4	4	Birmingham	2005
BSAC_bs134[a]	4	4	Chester	2006

Isolate	*emm*-type*	Subtype†	Region	Date isolated
BSAC_bs138[a]	4	4	Newcastle	2006
BSAC_bs180[a]	4	4	Birmingham	2008
BSAC_bs25[b]	3	3.2	Kent	2001
BSAC_bs515[b]	3	3.2	Kent	2003
BSAC_bs821[b]	3	3.1	Cambridge	2004
BSAC_bs975[b]	3	3.17	Truro	2004
BSAC_bs181[b]	3	3.1	Bristol	2008
BSAC_bs567[b]	12	12	Bristol	2003
BSAC_bs155[b]	12	12	Birmingham	2007
BSAC_bs184[b]	12	12	Chester	2008
BSAC_bs230[b]	12	12	Bristol	2010
BSAC_bs2438[d]	6	6.5	Southampton	2010
H293[c]	89	89	London	1995
H305[d]	1	1	London	1880s

***** Isolates were genotyped using *emm* gene sequences derived from *de novo* assembled sequence data, per CDC guidelines (Turner et al., 2019b). Genotypes were designated based on the DNA sequence encoding the hypervariable N-terminal region of M protein, which in turn determines serospecificity, with a submission query match of >90 bp to determine *emm*-genotype. † *emm*-subtypes were assigned according to 100% sequence similarity of the 180 bp subtype-encoding sequence of *emm* for each isolate, which includes the signal sequence and a portion of the mature protein, per CDC guidelines.

[a] Used in Chapters 1, 2, 3 and 4. Used in Chapter 6[b]. [c] Used in chapters 4 and 6,[d] used in Chapter 5.

2.1.2 **Bacteriological growth media**

- Columbia Horse Blood agar (Oxoid, Basingstoke, UK or VWR)
- Todd-Hewitt Broth (THB) (Oxoid, Basingstoke, UK)
- Todd-Hewitt Agar (THA)
  - THB 2% (w/v) bacteriological agar (Oxoid)
- Luria Bertani Broth (Oxoid)

- Luria Bertani Agar (LBA)
  - LB 2% (w/v) bacteriological agar

2.1.3 Antibiotics

- Ampicillin (Sigma)
- Penicillin/streptomycin (Sigma)

2.1.4 Human cell lines

Immortalised human tonsillar keratinocytes (E6/E7-HTE) (Ximbio, London), were isolated from routine bilateral tonsillectomy specimens, had been immortalised by transfection with retroviral vectors encoding Human papillomavirus type 16 oncogenes E6 and E7, previously (Grayson *et al.*, 2018; Spanos *et al.*, 2008) and kindly provided by Professor Craig Murdoch, University of Sheffield. A459s are an immortalised human pneumocyte cell line, that have been widely applied to infection work with *S. pyogenes* (LaPenta *et al.*, 1994; Lieber *et al.*, 1976; Lynskey *et al.*, 2017).

2.1.5 Human cell culture media and supplements

- Cell culture PBS (Gibco, Fisher Scientific)
- Ham's F12 Nutrient Mixture (Gibco, Fisher Scientific)
- Dulbecco's Modified Eagle's Medium (DMEM) (Gibco, Fisher Scientific)
- RPMI (Gibco, Fisher Scientific)
- Foetal calf serum (FCS) (Gibco, Fisher Scientific)
- Adenine (Sigma)
- Human recombinant insulin (Sigma)
- 3, 3, 5- Tri-iodothyronine (Sigma)
- Apo-transferrin (Sigma)
- Epidermal growth factor (Sigma)
- Modified Green's (adenine- and flavin-enriched) Media
  - DMEM 66% (v/v)
  - Ham's F12 21.6% (v/v)
  - FCS 10% (v/v)
  - Adenine 0.025µg/mL
  - Human recombinant insulin (Sigma) 5µg/mL
  - 3, 3, 5- Tri-iodothyronine (Sigma)
  - Apo-transferrin (Sigma) 1.36ng/mL/5µg/mL
  - Epidermal growth factor (Sigma) 1.36ng/mL/5µg/mL

2.1.6 **Enzymes**

- Turbo DNA-free® DNase (Ambion, Cambridgeshire, UK)
- Transcriptor reverse transcriptase (Roche Applied Science, Burgess Hill, UK)
- SYBR® Green Jumpstart™ Taq Readymix™ (Sigma)
- OneTaq Quick-Load 2X Master mix (New England Biolabs)
- Phusion® High-Fidelity DNA Polymerase (New England Biolabs)
- Mutanolysin (Sigma)
- Lysozyme (Sigma)
- Proteinase K (Sigma)
- T4 DNA Ligase (New England Biolabs)
- Restriction enzymes (New England Biolabs): BamHI
- 1X Trypsin-EDTA solution (Insight Biotechnology)

2.1.7 **Standard solutions and buffers**

- Phosphate buffered saline (PBS)
  - 137 mM NaCl
  - 2.7 mM KCl
  - 10 mM Na_2HPO_4
  - 2 mM KH_2PO_4

- Tris-acetate buffer (TAE) 50X
  - 2 M Tris (w/v)
  - 57.1 mL/L Glacial acetic acid (w/v)
  - 50 mM EDTA (w/v) dissolved at pH 8

- SET Buffer
  - 75 mM NaCl
  - 25 mM EDTA
  - 20 mM Tris pH 7.5

- Sodium acetate pH 5.5 buffer
  - 0.1% (v/v) Sodium dodecyl sulphate (SDS)
  - 30 mM Na-Acetate pH 5.5

- DNA loading buffer
  - 0.25% (w/v) bromophenol blue (Sigma)

- 0.25% (w/v) xylene cyanole (Sigma)
- 30% (v/v) glycerol (AnalR)

2.1.8 Kits

- Monarch® DNA Gel Extraction Kit (New England Biolabs)
  - Gel dissolving buffer
  - DNA wash buffer
  - DNA elution buffer
  - DNA clean-up columns

- Monarch® PCR DNA Clean Up Kit (New England Biolabs)
  - DNA clean up binding buffer
  - DNA wash buffer
  - Elution Buffer

- Monarch® DNA Plasmid Miniprep Kit (New England Biolabs)
  - Plasmid resuspension buffer
  - Plasmid lysis buffer
  - Plasmid neutralisation buffer
  - Plasmid wash buffer 1
  - Plasmid wash buffer 2
  - DNA elution buffer

- pCR2.1 TA cloning kit (Invitrogen)
  - Linearised pCR®2.1 (25 ng/µL in 10mM Tris-HCl, 1mM EDTA, pH 8)
  - ExpressLink™ T4 DNA Ligase (5.0 Weiss units/µL)
  - 5X ExpressLink™ T4 DNA Ligase Buffer (50 mM Tris-HCl, pH 7.6 , 50 mM $MgCl_2$, 5 mM ATP, 5 mM dithiothreitol, 25 % (w/v) polyethylene glycol-8000)
  - 10X PCR Buffer (100 mM Tris-HCl, pH 8.3 (at 42°C), 500 mM KCl, 25 mM $MgCl_2$, 0.01% (w/v) gelatine)
  - 50 mM dNTP's (12.5 mM, dATP, dCTP, dGTP and dTTP, pH 8).
  - DNA template control 0.1µg/µL (10 mM Tris-HCl, 1 mM EDTA, pH 8
  - Control PCR Primers each at 0.1 µg/µL (10 mM Tris-HCl, 1 mM EDTA, pH 8)
  - SOC medium (2% tryptone (w/v), 0.5% (w/v) yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM)
  - TOP10 *E. coli* cells
  -

- 5-bromo-2'-deoxyuridine (BrdU) ELISA (Colourimetric) (Roche)
  - BrdU Labelling Reagent (1000X)
  - FixDenat (Ready-to-Use)
  - Anti-BrdU-peroxidase antibody
  - Antibody Dilution Solution (Ready-to-Use)
  - Washing Buffer (10X PBS)
  - Substrate Solution TMB (Ready-to-Use)

2.1.9 **Miscellaneous reagents**

- Aqueous phenol (Sigma)
- Phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma)
- X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) (Bioline)
- RNaseIn RNase Inhibitor (Sigma)
- All primers were supplied by Sigma (Dorset, UK)
- 100 mM dNTPs (ATP, TTP, CTP, GTP: 25 mM each (Promega, Southampton, UK)

2.1.10 **Primers**

Table 2.2: Primers used in this study

Primer	**Sequence (5'-3')**	**Function**
10750.1 A	GCATCCAGACTATTCCATTC	Φ10750.1/Φ427.1 induction
10750.1 B	CGTATGATGTTCAATCTAGGATAG	
10750.1 C	TGCGTCAACAGTTATTGTCG	
10750.1 D	ACATTAGCCTCGTTCACGC	
10750.2 A	ATCAACTAAGGCAGCTTCTG	Φ10750.2/Φ427.2 induction
10750.2 B	CGGAACTCTTGACTACACCTC	
10750.2 C	AACAAACCTTGCCAAGTACG	
10750.2 D	CCATCTCTGTAACAGTCAAATG	
10750.3 A	CCAATCAAGAAGGCTGTAATG	Φ10750.3/Φ427.3 induction
10750.3 B	GCACCTGGAGCAATATTTG	
10750.3 C	TACAGAAGGATATCGTAACGGG	
10750.3 D	TTGCAAGTCGTCTCATTCAAG	
SpyCIM4 A	CGAGAACTTCCGGTAATTC	$SpyCI^{M4}$ induction
SpyCIM4 B	CGAATATCAGCATGACTTTG	
SpyCIM4 C	AGCATCCAAGACCAATGG	
SpyCIM4 D	CTTCAAGCAATGACAACCC	
MutL_F	GTCTCAATTTCCCCACCAGTAG	*mutL* transcript detection
MutL_R	CAAATTGCAGCTGGTGAAG	
MutS_F	CTTGAAGCGGGGTCATATTC	*mutS* transcript detection
MutS_R	GTTGGTGCTAAGACCATATTTGC	
GYRAF	AGCGAGACAGATGTCATTGCTCAG	Construction of qPCR.1 (Salim *et al.*, 2005)
GYRAR	CCAGTCAAACGACGCAAACG	Construction of qPCR.1 (Salim *et al.*, 2005)

Primer	Sequence (5'-3')	Function
SSAF[a]	GCACCATAGAAATCAAATTGAAG	Construction of qPCR.1
SSAR	ATATAGCCTGTCTCGTACGGAG	Construction of qPCR.1
SPD3F[a]	GGCGATCTTGACAATCTGC	Construction of qPCR.1
SPD3R	AATTATGCCAGCCAGGAGG	Construction of qPCR.1
SPECF	CATAGATGTACTCACCGGTGTG	Construction of qPCR.2
SPD1R	TCATGGGTAATGAATCGTGG	Construction of qPCR.2
gyrA-RT_F	AGCGAGACAGATGTCATTGCTCAG	qPCR (Salim *et al.*, 2005)
gyrA-RT_R	CCAGTCAAACGACGCAAACG	qPCR (Salim *et al.*, 2005)
speC-RT_F	CATAGATGTACTCACCGGTGTG	qPCR
speC-RT_R	GTTCCGAAATGTCTTATGAGGC	qPCR
spd3-RT_F[a]	CGACTGAGACACCAGGAATC	qPCR
spd3-RT_R	AATTATGCCAGCCAGGAGG	qPCR
ssa-RT_F[a]	AATATAGCCTGTCTCGTACGGAG	qPCR
ssa-RT_R	GTATGTACGGAGGTGTTACTGAGC	qPCR
M13_R	CAGGAAACAGCTATGACC	pCR2.1 sequencing

[a] modified from original sequence(s) following Sanger sequencing of completed plasmid

2.2 Methods

2.2.1 Bacterial strains and media

All *S. pyogenes* isolates were stored in suspension with THB + 20-30% (v/v) glycerol at -80°C. Frozen stocks were recovered by inoculation of CBA plates (Oxoid, Basingstoke, UK) to achieve individual colonies, and incubated overnight at 37°C in an atmosphere supplemented with 5% (v/v) CO_2 throughout the project. For standard liquid culture, sweep inocula (2-3 colonies) were propagated statically in THB (Oxoid) overnight, unless otherwise stated.

E. coli stocks were maintained in suspension in LB +20-30% (v/v) glycerol and 50μg/mL ampicillin, frozen at -80°C. Stocks were recovered by inoculation of LB agar plates supplemented with 50 μg/mL ampicillin and

incubation at 37°C, under aerobic conditions. For liquid culture, LB broth (50 µg/mL ampicillin) was inoculated and incubated at 37°C under aerobic conditions, with shaking at 200-220 rpm.

2.2.2 *S. pyogenes* DNA extraction

S. pyogenes DNA extraction was performed using the method of Pospiech and Neumann (Pospiech and Neumann, 1995). Bacterial suspensions were pelleted by centrifugation at 2,000 x *g* for 10 minutes and the supernatant discarded. The bacterial pellet was resuspended in 1 mL sterile PBS with 20 U/mL mutanolysin (Sigma) and 200 µg/mL lysozyme (Sigma) and incubated at 37°C for 10 minutes. The suspension was then centrifuged at 200 x *g* for 10 minutes, the supernatant discarded and the pellet resuspended in 500 µL SET buffer (5 mM NaCl, 25 mM EDTA, 20 mM Tris pH 7.5) with 1% (v/v) SDS and 0.5 mg/mL proteinase K (Sigma). The suspension was then incubated with occasional mixing at 55ºC for two hours. After incubation, 1/3 volume of 5 M NaCl and 1 volume of chloroform was then added and the suspension incubated at room temperature for 30 minutes. The suspension was then centrifuged for 16,000 x *g* for 15 minutes, and the aqueous phase was transferred to a fresh tube, and the DNA precipitated by the addition of 1 volume of isopropanol. The DNA pellet was recovered by centrifugation at 16,000 x *g* for 5 minutes, and washed with 70% (v/v) ethanol before being allowed to dry and subsequent resuspension in 20-200 µL molecular grade water (Honeywell). DNA was quantified and quality measured by nano drop.

2.2.3 *S. pyogenes* RNA Extraction

S. pyogenes were collected by centrifugation at 2000 *x* g for 10 minutes, and lysed in 500 µL sodium acetate buffer (0.1% (v/v) SDS, 30 mM Na-Acetate pH 5.5) and 500 µL hot acidic aqueous phenol (Sigma) at 65°C for 10 minutes with occasional mixing. The suspension was centrifuged again at maximum speed for 10 minutes. The aqueous phase was subsequently aspirated and transferred to a fresh tube containing 500 µL phenol:chloroform:isoamyl alcohol (25:24:1) (Sigma). After centrifugation for a further 5 minutes at maximum speed, the aqueous phase was aspirated and transferred to a fresh tube containing 500 µL isopropanol and 500 µL 3M sodium acetate buffer pH 7 and stored at -20°C overnight to precipitate RNA. The following morning, to complete RNA precipitation, tubes were centrifuged at 16,000 *x* g for 30 minutes at 4°C, before being washed in freshly prepared, pre-chilled (-20°C) 80% (v/v) ethanol. Finally, once dry, RNA pellets were resuspended in 50 µl of MGW and stored at -80°C. The concentration and quality of the RNA was measured by nanodrop prior to use.

2.2.4 PCR and agarose gel electrophoresis

Phusion® High-Fidelity DNA Polymerase (New England Biolabs) was used for all standard (non-real time) PCR. Primers were reconstituted with molecular grade water to 100 µM and stored at - 20ºC, and used at a final of 10

μM, unless otherwise stated. Reaction mixtures were gently vortexed to mix constituents, and briefly pulse-centrifuged to concentrate reaction mixtures. For PCR reactions without incorporated loading dye, DNA loading buffer (0.25% (w/v) bromophenol blue (Sigma), 0.25% (w/v) xylene cyanole (Sigma), 30% (v/v) glycerol (AnalR) was added before loading onto 0.5-2% (w/v) agarose gels stained with SYBR® Safe (Sigma) and underwent electrophoresis in agarose gels in TAE at 120V for ~30 minutes. Gels were visualised by UV transillumination. DNA ladder 1kb plus (Invitrogen) was used to infer product size. Extension and annealing times and temperatures were modified as appropriate for each reaction.

2.2.5 **Detection of prophage excision by PCR**

To detect prophage and SpyCI excision, four primers were designed using a method adapted from Aziz *et al.* (Aziz *et al.*, 2005). These were designed to detect each *emm*4 prophage or SpyCI^{M4} in its integrated and extra-chromosomal state. Primer combination A and B and combination C and D flanked the attachment sites at each end of the prophage-encoding region, where the viral genome and the bacterial chromosome meet, such that amplification by these primer combinations could only occur if the element was integrated. Combining primers B and C, the prophage was detectable in its excised state, amplifying across a junction of viral DNA unified by re-circularisation. Excision was confirmed by amplification with primer pair A and D, which amplified across the attachment site in the bacterial chromosome, but was only able to do so if the attachment site was unoccupied by prophage (Figure 2.1).

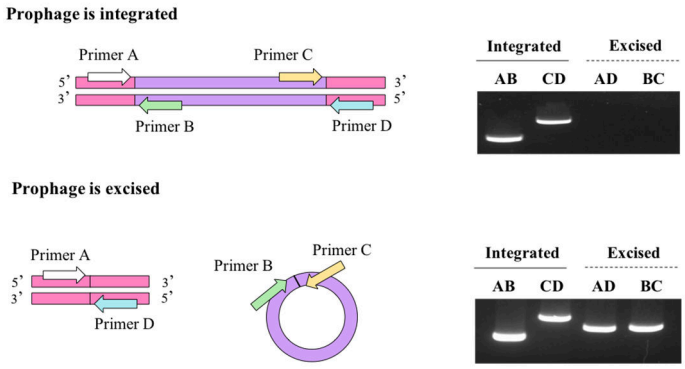

Figure 2.1: Schematic of prophage conformation PCR

Primer combination A (white arrow) and B (green arrow), and combination C (yellow arrow) and D (blue arrow) amplify across the junction at which the bacterial chromosome (pink) and the prophage genome (purple) meet when the prophage is integrated, *attR* and *attL,* respectively. Prophage absence, or excision of the prophage from the bacterial chromosome, and thus the vacation of the attachment site, permits amplification with primers A and D. The circularised, extra-chromosomal conformation of the excised prophage (purple circle) is detectable by amplification with primers B and C across *attP,* the junction at which the 3' and 5' ends of the linearised prophage meet, following prophage induction.

2.2.6 **Purification of PCR products**

The Monarch® PCR and DNA Clean-up Kit (New England Biolabs) was used to purify and concentrate all PCR and ligation products, per manufacturers guidelines. All centrifugation steps were carried out at 16,000 x *g*. PCR amplicon was mixed with DNA Binding Buffer at a ratio of 2:1 or 5:1, dependent on product size. Binding Buffer/PCR reaction mixtures were transferred to spin columns and centrifuged for 1 minute, before discarding flow-through, and transferring the column to a collection tube. Columns were washed twice with 200µL DNA Wash Buffer, and flow-through discarded. DNA was eluted in 20-200µL molecular grade water into a clean tube by centrifugation. Eluents were quantitated by nano drop and stored at -20°C, unless otherwise stated.

2.2.7 **Purification of DNA following electrophoresis**

PCR products were excised from the agarose gel under UV transillumination from 0.9% (w/v) agarose gels, using a clean razor blade, and purified using the Monarch® DNA Gel Extraction Kit (New England Biolabs). Excised slices were dissolved in 4 volumes Gel Dissolving Buffer per 1 volume agarose, and incubating at 50°C for 5-10 minutes, with occasional vortex. The mixture was transferred to a binding column, which had been inserted into a collection tube, and bound by centrifugation for 1 minute at 16,000 x *g*, with flow-through discarded. The column was washed twice with 200 µL DNA Wash Buffer, spinning for 1 minute at 16,000 x *g* following each wash, and discarding subsequent flow-through. Washed columns were transferred to clean 1.5 mL tubes and DNA eluted in 10-20 µL molecular grade water after incubating at room temperature for 1 minute. Eluents were quantitated by nano drop and stored at -20°C, unless elsewise described.

2.2.8 **Plasmid purification**

Overnight 5 mL cultures of TOP10 *E. coli* carrying the desired plasmid were harvested by centrifugation, supernatants discarded and pellets resuspended in 200 µL Plasmid Resuspension Buffer, and thoroughly emulsified by pipetting. 200 µL Plasmid Lysis Buffer was added to the suspension, which was mixed by inversion at room temperature for ~1 minute. Lysis was stopped by addition of 400 µL Plasmid Neutralisation Buffer, and for 2 minutes, again gently mixed by inversion until colour change and precipitation were observed. Lysate was centrifuged for 5 minutes at 16,000 x *g*, with this centrifugal force applied to all subsequent spins. Supernatant was aspirated, transferred to a binding column inserted into a collection tube, and spun for 1 minute, flow-through was discarded. The column was then subject to two washing steps, once with 200 µL Wash Buffer 1 followed by a 1-minute spin, and then a second, using Wash Buffer 2 at a volume of 400 µL, and centrifugation for 1 minute. Columns were transferred to clean 1.5 mL tubes and DNA eluted in 30-50 µL molecular grade water. Eluents were quantitated by nano drop and stored at -20°C prior to use.

2.2.9 **Construction of plasmids for quantitative real-time PCR**

A qPCR method to accurately detect and measure the copy number of RNA transcripts or target genes was developed previously and has since been adapted and applied to both *S. pyogenes* and *Staphylococcus aureus* (Subramenium *et al.*, 2015; Lynskey *et al.*, 2019a, 2013; Sharma *et al.*, 2019; Turner *et al.*, 2009). These method uses a plasmid constructed to contain concatenated regions of each target gene. The copy number of this plasmid can be determined and serial dilutions provide a standard curve against which copies of corresponding target genes in samples can be measured.

To measure copies of the housekeeping gene *gyrA*, and target prophage genes *ssa* and *spd3,* primers were designed to amplify 100-220bp regions of each Beginning with *ssa* and *gyrA*, followed by subsequent genes, fragments were amplified with Phusion® High-Fidelity DNA Polymerase (New England Biolabs) to produce blunt-ended templates, purified with Monarch® PCR purification Kit (New England Biolabs). These were subject to overnight blunt-end ligation with T4 ligase (New England Biolabs) at 14°C. Primers gyrA-F and ssa-R primers were then used to amplify from the ligation reaction a fragment comprising one copy of *gyrA* and one copy of *ssa* ligated together, and purified from the agarose electrophoresis gel using the Monarch® DNA Gel Extraction Kit (New England Biolabs). To the *gyrA-ssa* fragment, *spd3* was ligated, and the process repeated until all three genes were successfully ligated. An additional three DNase genes were also ligated to this fragment (*sdn*, *spd4* and *spdB*) but were ultimately not used due to a change in direction of the project (Figure 2.2). The completed fragment was subjected to PCR with OneTaq Quick-Load 2X Master mix (New England Biolabs) to achieve fresh TA-overhangs,purified using the Monarch® PCR purification Kit (New England Biolabsand ligated into pCR2.1 using the TA Cloning® Kit (Invitrogen), to yield a plasmid hereafter referred to as qPCR.1. This plasmid was then used to transform TOP10 *E. coli* (Invitrogen)by heat-shock, following manufacturers guidelines.

LB agar plates supplemented with 50 µg/mL ampicillin and spread with X-Gal (Sigma) were used to select for transformants, identifiable as white colonies. 5 mL aliquots of LB broth supplemented with 50 µg/mL ampicillin were inoculated with individual putative transformants using a sterile pipette tip, and incubated overnight. Plasmid DNA was extracted from pelleted cultures the following morning using the Monarch® DNA Plasmid Miniprep Kit (New England Biolabs). External Sanger sequencing (Eurofins Scientific) confirmed successful transformation of TOP10 *E. coli* with qPCR.1. Two of the original primers were subsequently modified to accommodate unintentional sequence variations in the insert, introduced during the cloning process; SPD3, and SSAF, to yield spd3-RT_F, and ssa-RT_F, respectively. A plasmid carrying *speC* and *spd1* was already available, qPCR.2, and this was used to detect transcription of these genes. The copy number of both plasmids was calculated using the size of the plasmid and the concentration of purified DNA. Finally, qPCR.1 and qPCR.2 were linearised using BamHI (New England Biolabs), and diluted to 2 x 10^8, respectively, in coliphage lambda DNA (Sigma), before storing at -20°C, until required.

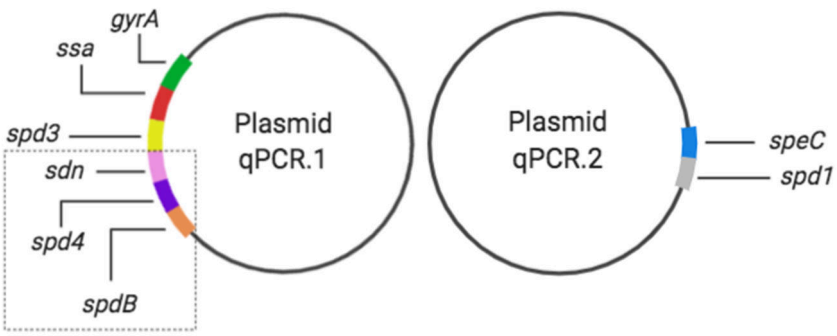

Figure 2.2: Schematic of the plasmids constructed for qPCR quantification

Simplified schematic representation of plasmids constructed to quantify qPCR data. Each contains a single copy of a fragment corresponding to an *emm*4 gene of interest; *gyrA*, *ssa* and *spd3* cloned into qPCR.1; and *speC* and *spd1* cloned into qPCR.2. Additional genes were cloned in to qPCR.1 prior to a change in project focus from genotype *emm*3 *S. pyogenes* (dotted square).

2.2.10 qPCR primer matrices

Primer matrices for qPCR primers designed to amplify each target gene were conducted to determine optimal primer concentrations for qPCR, to achieve a standard with an R^2=1, a low negative control, and PCR efficiencies for each run that were within ≤5% of one another, to permit precise comparisons of each gene across separate experiments and replicates. This was accomplished using serial dilutions of the qPCR.1 and qPCR.2, and subjecting these to qPCR with reactions containing alternating combinations of forward qPCR primer, diluted 1:5, 1:10 and 1:20, paired with a corresponding dilution of a reverse qPCR primer. Consequently, amplification of the plasmid standards with reactions containing each possible combination of forward and reverse primers for each target gene, to identify the most suitable combinations for all qPCR primers.

2.2.11 Reverse transcription and quantitative real-time PCR

RNA samples were DNase treated with Turbo DNA-*free* (Ambion) for 30 minutes at 37°C to remove genomic DNA. Concentration and quality of RNA was determined by nanodrop, and between 1-5 µg was converted to cDNA by reverse transcription with Transcriptor reverse transcriptase (Roche Applied Science). The reaction was then heated to 65°C with 4 µL of 10 µM random hexaoligos (Sigma), before transferring immediately to ice. Following the addition of 8µL RT Buffer, 0.5µL reverse transcriptase, 4µL 10mM dNTPs and 0.5µL RNAseIn (RNAse inhibitor, Sigma), reverse transcription proceeded at 25°C for 10 minutes, 55°C for 30 minutes and 85°C for 5 minutes. After reverse transcription, molecular grade H_2O was added to achieve a final concentration of 40 ng/ µL of cDNA. Corresponding reverse transcription negative reactions, without reverse transcriptase and RNaseIn (RNase Inhibitor, Sigma Aldrich), were prepared in parallel to each sample in order to confirm the

absence of genomic DNA contamination following DNase treatment. DNA was also standardised to 40 ng/µL for qPCR to detect gene copies, where indicated.

SYBR® Green Jumpstart Taq Readymix (Sigma) was used for all qPCR, with 200ng cDNA template per reaction, measured in triplicate. Forward and reverse real-time qPCR primers were diluted appropriately (either 1:10 or 1:20) from 100 µM stocks and added to a master mix of 12.5 µL SYBR® Green Jumpstart Taq Readymix (Sigma), 7 µL $MgCl_2$, 10 µL molecular grade water. For each run, serial dilutions of the appropriate plasmid (2×10^8 to 2×10^1 copies per reaction) were included to generate a standard, to which values were normalised and subsequently expressed as either target gene copies, or target transcript copies, per 10,000 copies of *gyrA,* with cycling parameters specific to each target gene(s) (Figure 2.3). Each reaction was measured in triplicate.

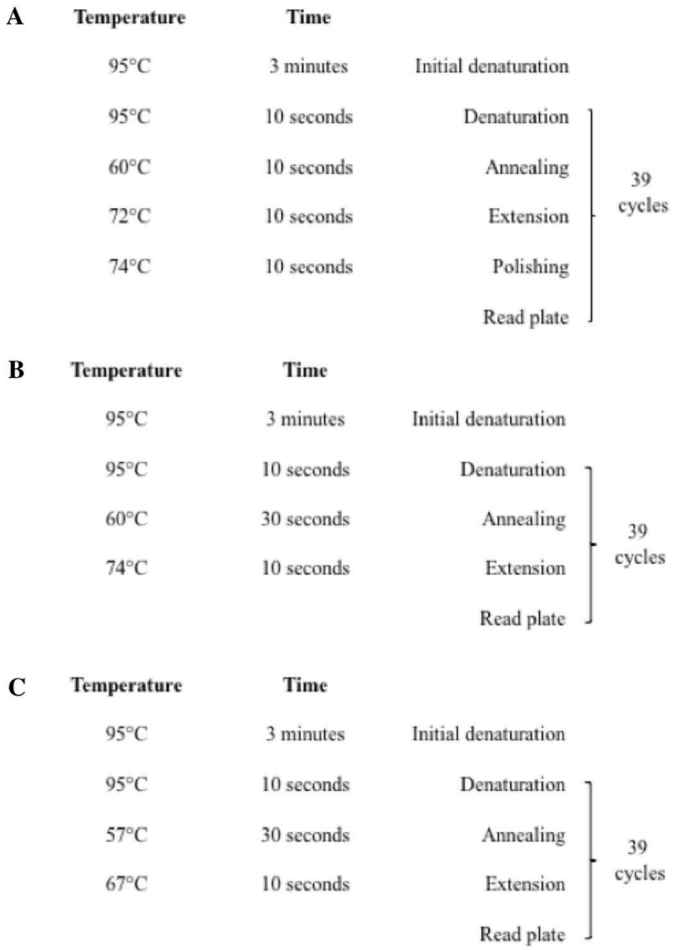

Figure 2.3: Target gene qPCR cycling parameters

qPCR cycling parameters for quantitation of transcript and gene copies A) Cycling parameters to quantitate *speC* and *gyrA* transcript/gene copies, which have identical parameters B) Cycling parameters for the quantitation of *spd3* transcript/gene copies. C) Cycling parameters for the quantitation of *ssa* transcript/gene copies.

2.2.12 **Maintenance of human cell lines**

All cell culture media was prepared in batches of 500 mL and stored at 4°C for up to 1 month to maintain cells prior to infection experiments. Human tonsil epithelial cells were sustained using Green's flavin and adenine rich media (Allen-Hoffmann and Rheinwald, 1984; Grayson *et al.*, 2018) (recipe provided in materials section 2.1.5). The standard preparation of Green's Media was modified by the exclusion of penicillin/streptomycin, amphotericin B, hydrocortisone, and cholera toxin, as prior work with Green's Media in this laboratory, and a collaborating laboratory previously found these constituents to be either unnecessary or deleterious to the applications described here. With these modifications, HTE cells grew well, without visible cytopathic effect across multiple experiments. A549 cells were sustained using RPMI +10% (v/v) FCS without antibiotics, as has been previously described (LaPenta *et al.*, 1994; Lieber *et al.*, 1976; Lynskey *et al.*, 2017).

Chapter 3

3. Bioinformatic characterisation of *emm*4 isolates

3.1 Introduction

Although *emm*4 isolates (n=10) constitute just 2.9% of the BSAC bacteraemia collection (n=344), genotype *emm*4 *S. pyogenes* are a major disease causing *emm*-type in high income countries globally (Efstratiou and Lamagni, 2016), occasionally even superseding the consistently dominant *emm*1 genotype (Kim *et al.*, 2019; Whitehead *et al.*, 2011). This *emm*-type is genetically acapsular (Flores *et al.*, 2012), and a frequent cause of both invasive and non-invasive clinical manifestations, though associations between *emm*4, scarlet fever and paediatric infection have been reported (Flores *et al.*, 2019; Galloway-Peña *et al.*, 2016; Jaggi *et al.*, 2005; Kim *et al.*, 2019; Silva-Costa *et al.*, 2014; Turner *et al.*, 2016; Whitehead *et al.*, 2011). Based on two completed *emm*4 reference genome sequences, MGAS10750 and MEW427, *emm*4 isolates are lysogenised by three prophages; a prophage encoding the superantigen *speC* and the DNase *spd1*, a second prophage encoding an additional DNase, *spd3,* and a third prophage encoding a further streptococcal superantigen, *ssa*, referred to therein and hereafter as Φ10750.1/Φ427.1, Φ10750.2/Φ427.2 and Φ10750.3/Φ427.3, respectively (Beres and Musser, 2007; Jacoba *et al.*, 2016). These prophage are also found in clinical *emm*4 isolates globally (Athey *et al.*, 2016, 2014; Beres and Musser, 2007; Chalker *et al.*, 2017; Chochua *et al.*, 2017; Jacoba *et al.*, 2016; Kapatai *et al.*, 2017; Turner *et al.*, 2019b, 2017, 2016). An *S. pyogenes* chromosomal island (SpyCI) is also common in *emm*4, typically integrated into the DNA MMR operon (Nguyen and McShan, 2014), and appears to regulate the expression of adjacent genes through a dynamic growth-phase dependent pattern of integration and excision in some genotypes (Hendrickson *et al.*, 2015; Nguyen and McShan, 2014; Scott *et al.*, 2012).

Recent WGS of a longitudinal sample of *S. pyogenes* bloodstream isolates in the United Kingdom was undertaken in collaboration with the BSAC and included ten *emm*4 isolates (Turner *et al.*, 2019b). Gubbins analysis indicated multiple SNP clusters in those isolates clustering with reference genome MEW427, indicative of possible regions of recombination. The majority of the remaining five isolates in this small collection clustered with reference genome MGAS10750, and did not exhibit SNP clustering. Four of the six SNP clusters were located within each of the four mobile genetic element-encoding regions in these isolates (Figure 3.1).

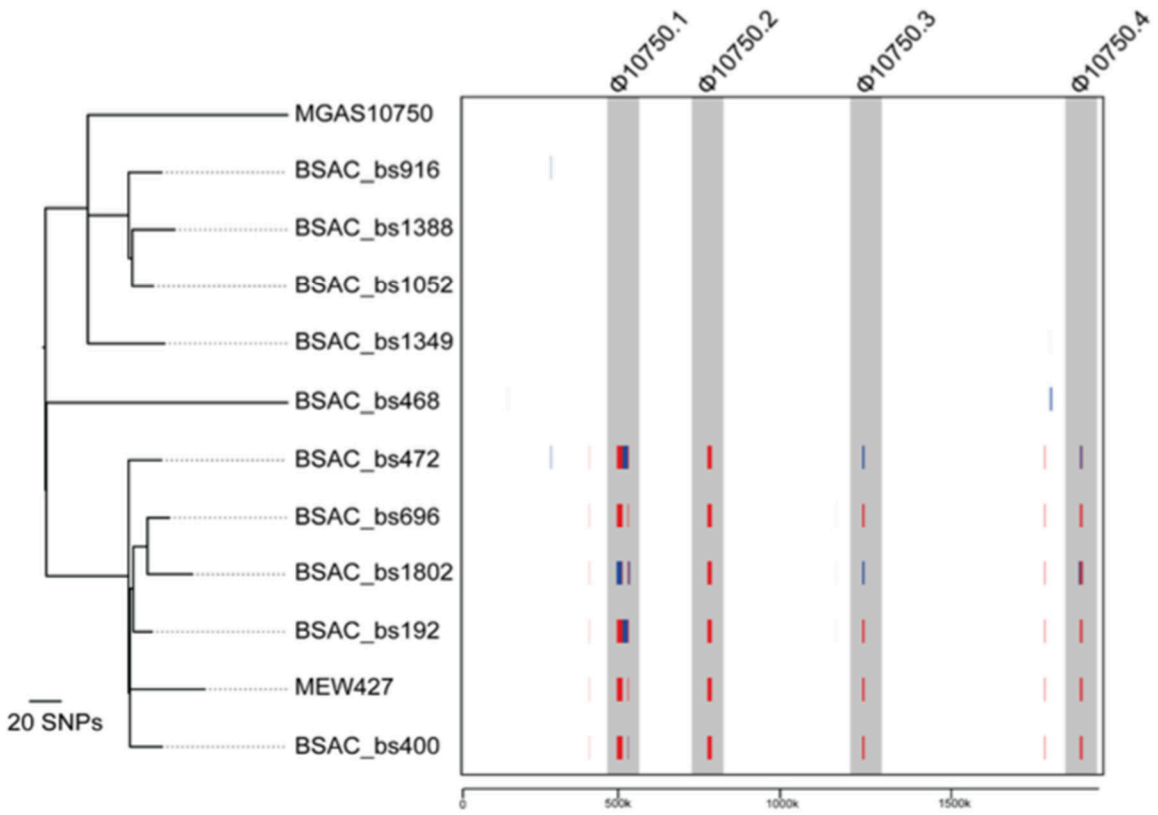

Figure 3.1. Putative regions of recombination within MEW427-like BSAC isolates

Whole genome sequence data for BSAC *emm*4 isolates (n=10) were mapped to completed reference genome strains MGAS10750 and MEW427. Gubbins analysis indicated putative regions of recombination in half (n=5) of these isolates, which clustered with MEW427. These regions were not identified in isolates clustering with MGAS10750 (n=4) and BSAC_bs468. The regions of predicted recombination were within the three prophage and a SpyCI element typical of the genotype, denoted here as Φ10750.1, Φ10750.2, Φ10750.3, and SpyCI^{M4} (Φ10750.4), respectively. Blue lines indicate SNP clustering specific to a single isolate, red lines indicate SNP clustering present in multiple isolates. Two additional clusters of SNPs are visible in isolates associated with MEW427 and correspond to *emm* and *spy0485*, a putative microcin immunity factor. The phylogenetic tree (left) was generated from 556 polymorphic sites following the exclusion of putative recombinant regions. The scale bar adjacent to the phylogenetic tree represents the number of SNPs, whereas the scalebar below the boxed region (right) indicates genomic position, relative to completed reference genome MGAS10750. Figure taken from Remmington *et al.* 2020.

3.2 Chapter aims and hypotheses

The aims of this Chapter were to investigate the bioinformatic features of the genotype *emm*4 isolates within the BSAC collection, in particular, with a view to establish characteristics that set those *emm*4 isolates clustering with reference genome MEW427 from those clustering with MGAS10750. The BSAC *emm*4 were also characterised using *de novo* assembled WGS data to inform and aid in the interpretation of phenotypic work using these isolates in subsequent chapters. It was hypothesised that putative SNP clustering within prophage-encoding regions might

indicate regions of recombination within isolates clustering with MEW427, though additional lineage-specific features were expected, as the phylogenetic tree (Figure 3.1) is drawn based on core SNPs, excluding putative regions of recombination. Isolate specific genetic features were also examined.

3.3 Methods

3.3.1 Genomic comparison of reference genomes

Publicly available completed *emm*4 whole genome reference sequences for MGAS10750 (Beres and Musser, 2007) and MEW427 (Jacoba *et al.*, 2016) were compared using the Artemis Comparison Tool (ACT) release 18.0.3 (Rutherford *et al.*, 2000).

3.3.2 Bioinformatic comparison of prophage-encoding regions

De novo assemblies for BSAC *emm*4 isolates was performed previously using Velvet (Turner *et al.*, 2019b) and was repeated with SPAdes (Bankevich *et al.*, 2012) to clarify prophage regions. The DNA sequences from all annotated coding regions for all three prophage and SpyCI^{M4} were extracted from MGAS10750 and were used to BLAST *de novo* assemblies to determine gene presence or absence. Genes with >98% identity over 100% length were classified as present. Amino acid sequences for each gene were subject to BLAST to determine putative function, and these were categorised as either integrase, replication, regulation, structural (encompassing capsid, neck, and tail fibre genes), hyaluronidase, infection (holin/lysin), and superantigen/DNase, and uncharacterised (hypothetical proteins).

3.3.3 Variable factor typing of BSAC *emm*4 isolates

Sequences were obtained for the 13 superantigen (*ssa*, *speA, speC, speG*, *speH, speI, speJ, speK, speL, speM, speQ, speR* and *smeZ*) and 8 DNase genes (*spnA, spdB, sda1, sda2, spd1 spd3, spd4* and *sdn*) found in *S. pyogenes*, from publicly available WGS data. The presence or absence of these genes in the BSAC *emm*4 isolates was determined by BLAST analysis against the *de novo* assemblies.

3.3.4 Lineage specific SNPs

Reads from *de novo* assembled WGS data constructed using Velvet and SPAdes obtained for each BSAC *emm*4 isolate were mapped to completed *emm*4 reference genome MGAS10750 using SMALT, and SNPs were identified. These data were generated by the project supervisor, Dr. Claire E. Turner, and analysed by the author.

3.3.5 **Pangenome analysis**

De novo assembled WGS data were annotated using Prokka (Seemann, 2014), and the pangenome was estimated using Roary (Sitto and Battistuzzi, 2020). This data was generated by the project supervisor, Dr. Claire E. Turner, but analysed by the author to identify gene presence and absence, and also to provide possible evidence of phage other than the canonical *emm*4 prophage.

3.4 Results

3.4.1 **Gene loss within prophage-encoding regions of reference genome MEW427**

Gubbins analysis had indicated putative areas of recombination within prophage-encoding regions of *emm*4 isolates clustering with the reference genome MEW427 (Jacoba *et al.*, 2016), but not those clustering with reference genome MGAS10750 (Beres *et al.*, 2006). The completed reference genomes were therefore compared, revealing that putative recombinant sequences were in fact regions of gene loss within prophage- and SpyCI-encoding regions (Figure 3.2). The large (~49kb) erythromycin resistance-associated ICE 10750-RD.2 present in MGAS10750 (Beres and Musser, 2007), was absent from MEW427, as expected (Galloway-Peña *et al.*, 2018).

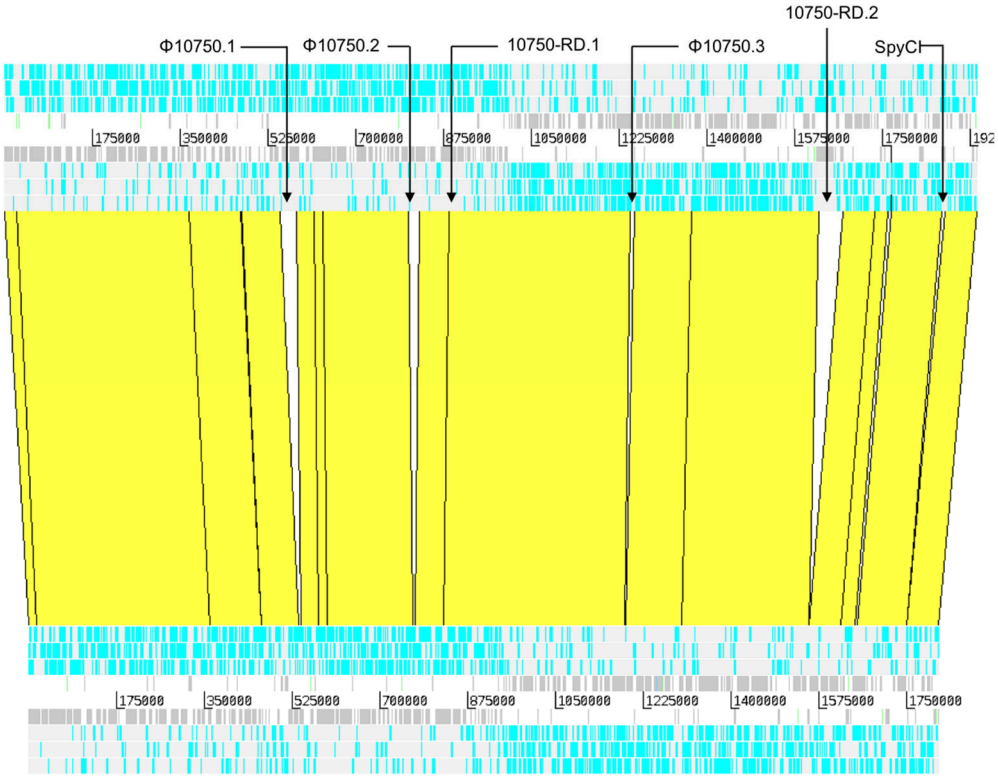

Figure 3.2: Comparison of *emm*4 reference genomes MGAS10750 and MEW427

Completed *emm*4 reference genomes MGAS10750 (top sequence) and MEW427 (bottom sequence) were compared in ACT. This comparison revealed regions within MGAS10750 that were absent in MEW427. On closer examination, these regions were determined to be mobile genetic element-encoding sequences. Specifically, regions of the *speC*/*spd1*-, *spd3*- and *ssa*-encoding prophage of MGAS10750, denoted Φ10750.1, Φ10750.2 and Φ10750.3, respectively. Portions of these were absent in the corresponding elements of MEW427, Φ427.1, Φ427.2, and Φ427.3, respectively. Further, the *emm*4 SpyCI of MGAS10750 was also found to contain a region comprising multiple genes that was not present in the corresponding SpyCI of MEW427. The approximate location of each mobile genetic element within *emm*4 reference genome MGAS10750 are indicated with black arrows. 10750-RD.2 is a large (~49kb) integrative and conjugative element associated with the erythromycin resistance gene *ermA* present in MGAS10750 but absent from MEW427 (Beres and Musser, 2007).

3.4.2 Bioinformatic features of MGAS10750 and MEW427 prophage

BLAST analysis of the three prophage-encoding regions of MGAS10750 and MEW427 revealed a typically lambdoid genomic arrangement, with discrete modules dedicated to specific functions. Relative to the prophage of MGAS10750, those derived from MEW427 had smaller genomes, with variable loss of structural genes and site-specific integrase genes across the three phage, but consistent absence of replicative and regulatory modules. Intriguingly, prophage-associated virulence genes, *speC* and the contiguous *spd1*, *spd3*, and *ssa*, were not lost, and were present at full-length associated with Φ10750.1/Φ427.1, Φ10750.2/Φ427.2 and Φ10750.3/Φ427.3, respectively (Figure 3.3). The *speC*/*spd1*-associated Φ10750.1 was observed to have undergone the most dramatic reduction in size, reducing in size from 63 genes/43,457bp at full length, to just 16 genes/13,487bp in Φ427.1, a

loss of ~78% of its original gene content (Figure 3.3A). The *spd3*-associated prophage Φ10750.2 comprised 59 genes/37,606bp at full-length, but having lost ~50% of its genes, in Φ427.2, consisted of 29 genes or 20,018bp (Figure 3.3B). The *ssa* prophage, Φ427.3, exhibited the least gene loss, with a ~26% reduction in genome size, reducing from 53 genes (or 35,302bp) at full-length, to 35 genes (or 26,301bp) relative to Φ10750.3 (Figure 3.3C).

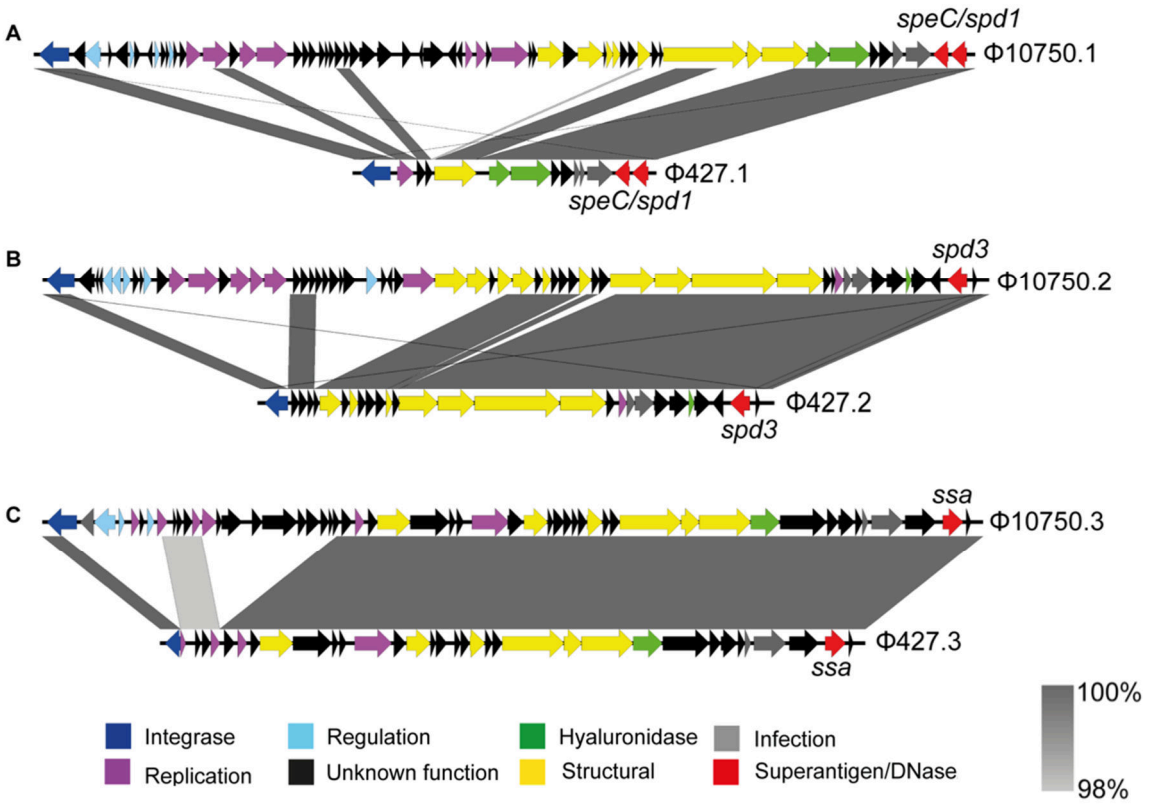

Figure 3.3: Comparison of gene content between MGAS10750 and MEW427 prophage

The three prophage encoding regions of reference genomes MGAS10750 and were compared, revealing gene loss in the prophage of MEW427, relative to those of MGAS10750, which were full-length. Recurrently, genetic modules pertaining to regulatory and replicative functions were lost. A) The canonical *speC/spd1 emm*4 prophage Φ10750.1 (top) comprises 63 genes, while the corresponding element in MEW427 (bottom) carries just 16, having lost ~78% of its gene content. B) The *spd3*-associated *emm*4 prophage of Φ10750.2 (top) contained 59 genes, whereas the corresponding MEW427, Φ427.2, contained 29, having lost ~50% gene content. C) The *ssa*-associated prophage Φ10750.3 of MGAS10750 (top) was observed to have lost ~26% of its genetic content, decreasing in size from 53 genes to 35 in the corresponding region derived from completed reference genome MEW427, Φ427.3. Scale bars indicate DNA sequence similarity between coding sequences. Gene function was determined by BLAST analysis of amino acid sequence and placed into eight categories based on predicted function: integrase (navy), regulatory (pale blue), replication (purple), structural (yellow), infection (grey), hyaluronidase (green) and superantigen/DNase (red). Image taken from Remmington *et al.*, 2020.

3.4.3 Superantigen and DNase typing of BSAC *emm*4 isolates

To determine how similar the BSAC *emm*4 isolates might be to the reference genomes, first BLAST analysis was used against the *de novo* assemblies to confirm the presence of the variable superantigen and DNase genes (Figure

3.4). All ten BSAC *emm*4 carried prophage-associated superantigen genes *speC* and *ssa*, and the chromosomally-encoded *smeZ*. Although the *speC* allele was identical among all ten BSAC *emm*4, these isolates carried a slightly different allele to MGAS10750, owing to a missense mutation of C:A at 13bp, translating to the replacement of histidine to aspartic acid, at 5 aa. The *ssa* and *smeZ* alleles of the BSAC *emm*4 were 100% identical to those of MGAS10750. Isolate BSAC_bs192 carried an additional prophage-associated superantigen *speK*. All BSAC *emm*4 were positive for the prophage-associated DNases *spd1* and *spd3*. BSAC_bs1802 also carried an additional prophage-associated DNase, *sdn*. Chromosomally-encoded *spdB* and *spnA* were detected in all isolates, and exhibited 100% DNA sequence similarity between the BSAC *emm*4 and MGAS10750.

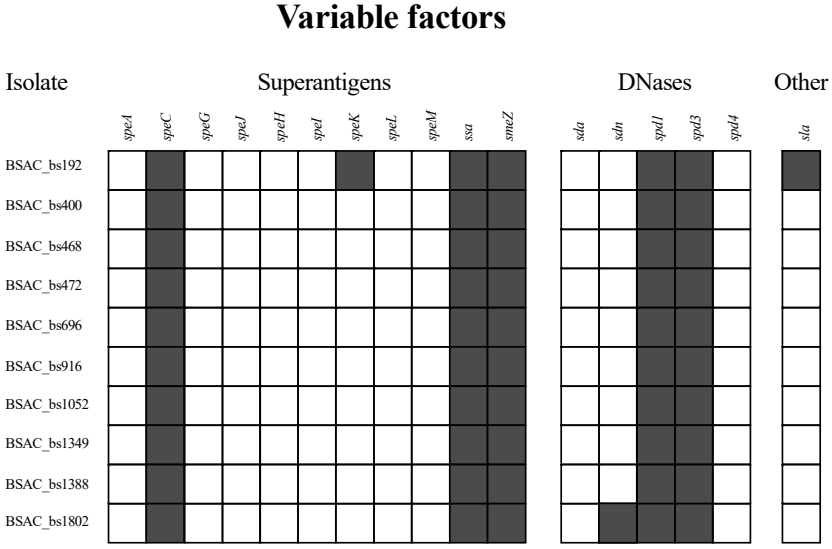

Figure 3.4 Variable factors present in BSAC *emm*4 isolates

Superantigen and DNase typing of BSAC *emm*4 isolates revealed a variable factor complement consistent with that of completed *emm*4 reference genomes MGAS10750 and MEW427. All isolates were positive for *speC*, *spd3*, *ssa* and *smeZ*. BSAC_bs192 uniquely carried *speK*/*sla*-. and BSAC_bs1802, carried the DNase *sdn*.

3.4.4 Bioinformatic features of BSAC *emm*4 prophage

To establish the prophage content associated with the variable factors of the BSAC *emm*4 isolates, DNA sequences were extracted from annotated coding regions of the three canonical *emm*4 prophage present in MGAS10750, Φ10750.1, Φ10750.2, Φ10750.3. SpyCI^{M4} was extracted along with the flanking *mutL* and *mutS,* and gene content determined identically. These were used to BLAST *de novo* assemblies for the BSAC *emm*4 isolates, using a cut off of ≥99% sequence similarity over 100% length.

Five BSAC *emm*4 isolates, BSAC_bs468, BSAC_bs916, BSAC_bs1052, BSAC_bs1349, BSAC_bs1388, were lysogenised by three prophages that appeared to be as complete or near-complete (92-100% gene presence) as were Φ10750.1, Φ10750.2 and Φ10750.3 (Figure 3.5). Four of the five isolates that are more closely related to

MEW427 and carried the same predicted recombination (BSAC_bs192, BSAC_bs400, BSAC_bs472 and BSAC_bs696), were found to carry prophage that exhibited marked gene loss across all three elements (21-66% gene presence across Φ427.1, Φ427.2, and Φ427.3). Patterns of gene loss within these isolates was highly similar to that observed in the prophage of MEW427, and predominantly related to loss of functional modules dedicated to prophage replication, genetic regulation, lysogeny and to a lesser and more variable extent, structural and tail fibre genes. Isolates clustering with MGAS10750, without prophage degradation, were thus denoted $M4_{complete}$, and those clustering with MEW427 and exhibiting gene loss were denoted $M4_{degraded.}$

3.4.4.1 Loss of functional prophage-encoded integrases in $M4_{degraded}$ isolates

Integrase genes of both Φ10750.1 and Φ427.1, denoted therein *spy0560* and *02750*, respectively, were identical, and were identified in the corresponding prophage of all BSAC *emm*4 isolates. However, relative to the 1,140bp integrase gene of Φ10750.2, *spy0831*, the corresponding gene in Φ427.2, *03820*, was lacking the first 264bp of the gene, including the start codon. Identical deletion events were observed in the Φ427.2 elements of all $M4_{degraded}$, relative to $M4_{complete}$ BSAC *emm*4, in which this gene was full-length, as it is in MGAS10750. While Φ10750.3 carried a full-length 1,143bp integrase gene, *spy1328*, the Φ427.3 integrase, *05920*, comprised just 547bp, lacking the initial 546bp, including the start codon. The same deletion was found in all $M4_{degraded}$.

3.4.4.2 Loss of prophage regulatory modules in $M4_{degraded}$ isolates

Regulatory genes were identified as those predicted to influence gene expression, including, where present, genes that maintain lysogeny or promote induction, but not those that are associated with the replication process. Φ10750.1 carries four identifiable putative regulatory genes, *spy0562, spy0565, spy0569* and *spy0571*, however, these are entirely absent in Φ427.1 of MEW427, and all $M4_{degraded}$ BSAC *emm*4. Φ10750.2 carries four genes with putative regulatory function, which were absent in Φ427.2 of MEW427 and MEW427-like isolates. Only two genes could be assigned putative regulatory function in Φ10750.3; *spy1326* and *spy1327*. These were not present in Φ427.3 of MEW427 or BSAC *emm*4. The Φ10750.3/Φ427.3 element of BSAC_bs1802, which is only partially degraded, carries both *spy1326* and *spy1327*.

3.4.4.3 Loss of prophage replicative modules in $M4_{degraded}$ isolates

Genes with putative replicative function, was defined as those predicted to be involved in the replication of the prophage genome, and those encoding proteins that facilitate packaging of DNA into nascent capsids. Φ427.1 did not carry any full-length genes to which a putative replicative function could be ascribed. One gene from this module, which included 3 classifiable genes, *spy0575, spy0577* and *spy0578*, was partially present in Φ427.1. This gene, *0275,* appeared to be the product of the truncation of *spy0575.*

Φ10750.2 carried a five-gene module with putative replicative function, comprising *spy0843-spy0847*. All were absent in Φ427.2, and MEW427-like BSAC *emm*4 including BSAC_bs1802. Φ10750.3 encoded five genes with

putative replicative function, *spy1327, spy1323, spy1321, spy1320, spy1317.* Of these, only *spy1321* is present in Φ427.3*,* wherein it shares 100% DNA sequence similarity with *spy1321* over the full length. This was not detected in the majority of $M4_{degraded}$ BSAC *emm*4, however, one $M4_{degraded}$ isolate, BSAC_bs696, retained *spy1317*. Unusual MEW427-like isolate BSAC_bs1802, which carries a near full-length Φ10750.3/Φ427.3, lacked only one gene from the putative replicative module, *spy1321*.

3.4.4.4 Variable structural gene loss across all three *emm*4 prophage

Thirteen genes carried by Φ10750.1 were classified as having a possible structural function, encoding proteins dedicated to the assembly of phage particles, including major and minor capsid proteins, tail fibre, head, neck, head and neck joining, and portal proteins. All of these were detected in Φ10750.1 of the $M4_{complete}$ BSAC *emm*4 but were either absent or had been subject to deletion or possible splicing in MEW427 and $M4_{degraded}$ BSAC *emm*4. Φ427.1 of MEW427 carried an annotated 1,968bp gene denoted *02270* which had arisen from the fusion of regions corresponding to *spy0609*, *spy0612* and *spy0614* of Φ10750.1*,* wherein they are predicted to encode tail fibre genes. Φ10750.2 carried eleven genes predicted to encode structural proteins, *spy0860, spy0861, spy0864, spy0865, spy0869-spy0871* and *spy0874-spy0877*, all found in $M4_{complete}$. Only two were lost in Φ427.2 of MEW427 and $M4_{degraded}$, *spy0860* and *spy0861*, encoding a phage terminase and portal protein, respectively. Putative tail protein encoding gene *spy0874* in Φ10750.2 was also partially deleted in Φ427.2 with the loss of the first 229 aa/1800bp of the gene. Seven genes encoding proteins with predicted structural function were identified in Φ10750.3 and these were all present in Φ427.3 prophage of MEW427 and $M4_{degraded}$.

3.4.4.5 Unusual prophage of BSAC_bs1802

BLAST and pangenome analyses of BSAC_bs1802 indicated that *speC/spd1* may be carried on a different phage than Φ10750.1/Φ427.1. Although the exact composition of this prophage could not be determined from the *de novo* assembly as it was divided over multiple contigs, subsequent long-read sequencing confirmed that the prophage was most closely related to *emm*87 prophage Φ743.4 (Athey *et al.*, 2016), hereafter Φ1802.1 (Remmington *et al.*, 2020). The 3' regions of Φ1802.1 and Φ743.4 integrase genes are highly similar to that of Φ10750.1/ Φ427.1. The *spd3*- and *ssa*-associated prophage of BSAC_bs1802 were similar to corresponding reference genome prophage, and exhibited more variable gene loss of ~49% and ~94% within Φ427.2, and Φ427.3, respectively.

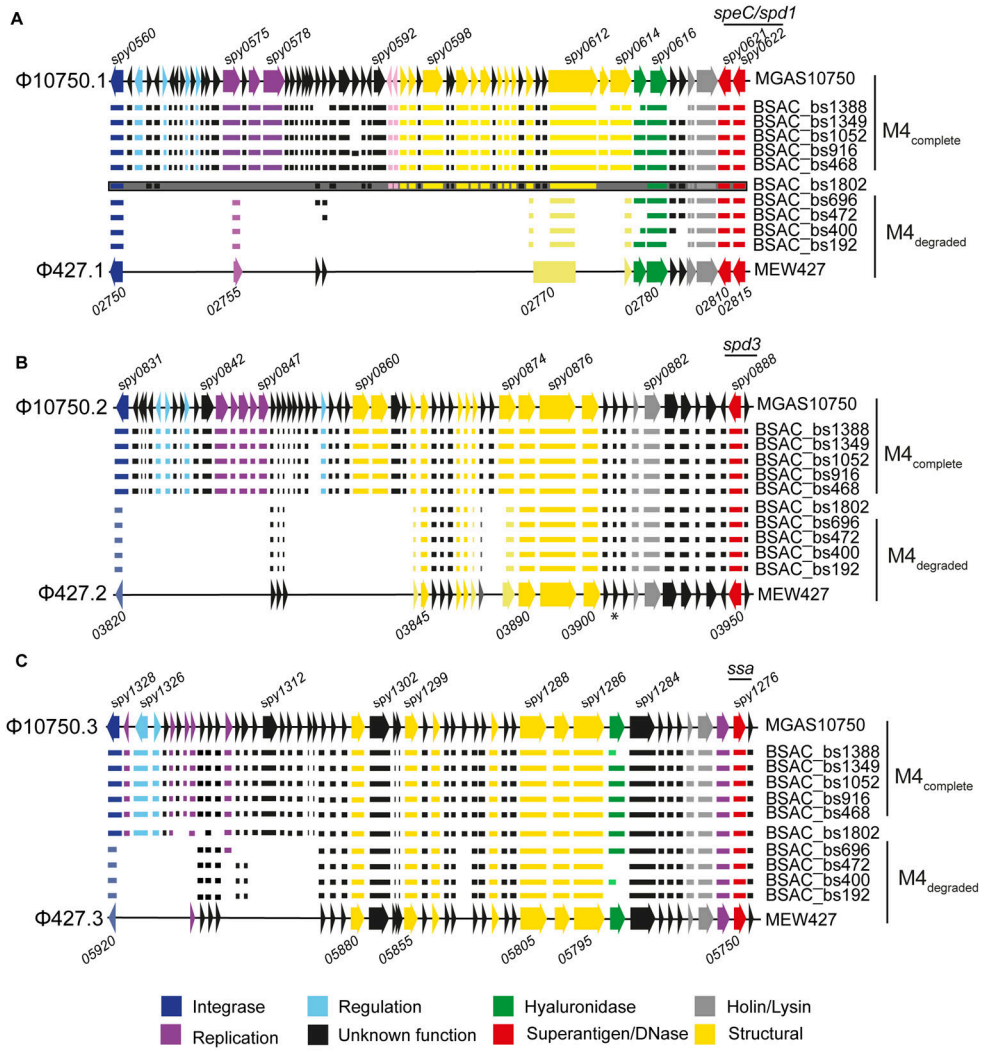

Figure 3.5: Bioinformatic comparison of prophage-encoding regions for *emm*4 isolates

Gene presence and absence across the three canonical prophage of *emm*4 *S. pyogenes*. Genes are represented by coloured arrows in reference genomes MGAS10750 and MEW427, and coloured lines in BSAC isolates. A) Φ10750.1 B) Φ10750.2 C) Φ10750.3 were present at full-length in MGAS10750 and $M4_{complete}$ isolates BSAC_bs468, BSAC_bs916, BSAC_bs1052, BSAC_bs1349 and BSAC_bs1388. Similar corresponding prophage were detected in MEW427; A) Φ427.1 B) Φ427.2 C) Φ427.3, however these were found to have undergone marked gene loss across all three elements. Near identical degraded prophage were identified in the majority of $M4_{degraded}$ isolates; BSAC_bs192, BSAC_bs400, BSAC_bs472 and BSAC_bs696. A greater number of prophage-associated genes were detected in BSAC_bs1802 compared to the other $M4_{degraded}$, and the *speC*/*spd1*-encoding prophage of this $M4_{degraded}$ isolate (Φ1802.1) (indicated with grey box) was dissimilar to Φ10750.1/Φ427.1, with some DNA sequence homology, but assembled with genes that were not detected in MGAS10750 or MEW427. Φ10750.1 gene *spy0619* was split into two separate genes in MEW427 and the BSAC isolates, but not MGAS10750. Locus numbers are indicated above and below reference genome sequences in 5 gene increments for MEW427. Annotated genes present in reference sequences but not BSAC *emm*4 sequences are indicated with an asterisk. Amino acid sequences for CDS were subject to BLAST and assigned COG classifications and colour codes per figure key. BSAC isolate numbers flank the sequences on the right-hand side.

3.4.5 Bioinformatic features of MGAS10750 and MEW427 SpyCI

In addition to typically harbouring three prophage, *emm*4 *S. pyogenes* are also host to a SpyCI, usually integrated within the DNA MMR operon (Scott *et al.*, 2012), hereafter SpyCIM4. SpyCIM4 was compared for MGAS10750 and MEW427, revealing gene loss within the MEW427 variant (
Figure 3.6). A number of SpyCIM4 genes could not be classified by BLAST, and this element does not encode any known bacterial virulence factors. There was, however, a notable absence of genes with putative replicative function in MEW427 SpyCIM4 compared to MGAS10750.

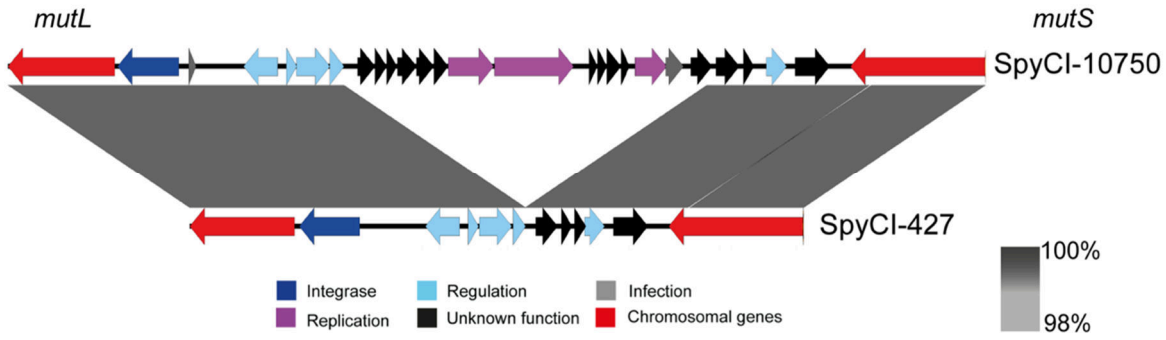

Figure 3.6.: Comparison of gene content between SpyCIM4 of MGAS10750 and MEW427

SpyCI-encoding regions derived from completed *emm*4 reference genomes MGAS10750 (top) and MEW427 (bottom) were compared, revealing gene loss, largely restricted to genes with replicative function (purple). Integrase genes and genes dedicated to regulation (pale blue) were unaffected. The holin/lysin of the MGAS10750 variant was not present in the MEW427 variant. A number of unclassifiable genes were also absent in the MEW427 variant, when compared to MGAS10750. Shaded grey bar indicates sequence similarity.

3.4.5.1 Gene loss within $M4_{degraded}$ $SpyCI^{M4}$

As was observed in MEW427 a large region in the centre of SpyCIM4 of the $M4_{complete}$ isolates had been lost in the corresponding element of $M4_{degraded}$ (Figure 3.7). This included predominantly genes of unknown function, but also a holin/lysin, *spy1916,* and four genes to which putative replicative function were assigned; two putative DNA primases, *spy1909*, and *spy1910*, and an ATPase-encoding gene, *spy1915*, were uniformly absent in MEW427 and $M4_{degraded}$ BSAC *emm*4 SpyCIM4. The SpyCI of BSAC_bs1802 was unlike SpyCIM4, but was ~99% similar to the SpyCI of *emm*28 and *emm*77 (Remmington *et al.*, 2020).

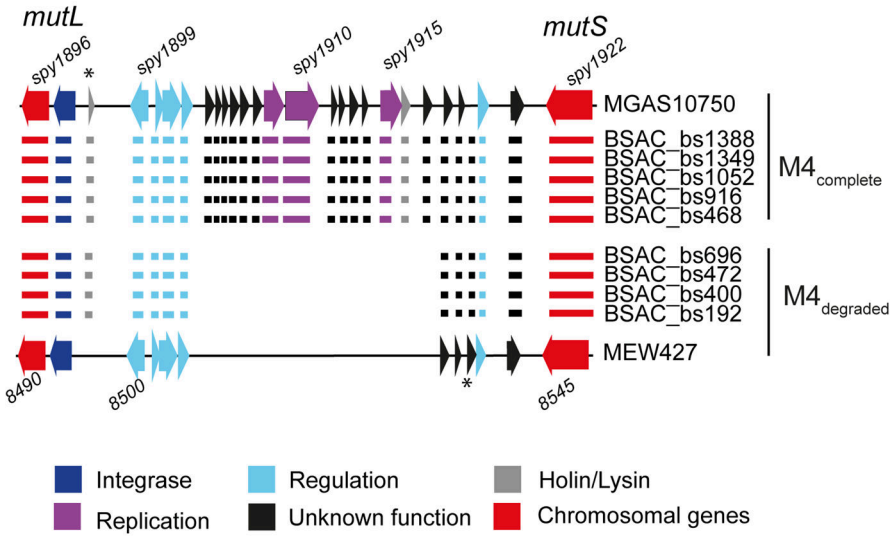

Figure 3.7: Bioinformatic comparison of BSAC SpyCIM4

The SpyCIM4 element is integrated between *mutL* and *mutS*, in the DNA MMR operon (red arrows). At full-length, in reference genome MGAS10750, SpyCIM4 comprises 25 genes, all of which were detected in $M4_{complete}$ isolates. The SpyCIM4 of $M4_{degraded}$ isolates comprised just 11 genes, having lost ~56% of its gene content. The pattern of gene loss was common to all $M4_{degraded}$ isolates. BSAC_bs1802 did carry a SpyCI element, however this was highly similar (~99%) to a different element, and was not compared. Locus numbers are provided above and below the reference genome SpyCIM4 sequences, and MEW427 is labelled at 5-gene increments. A gene that was annotated in MGAS10750 and all BSAC *emm*4 isolates but was not annotated in MEW427, and a gene that was annotated in MEW427 and all *emm*4, but not annotated in MGAS10750 are indicated with asterisks.

3.4.6 $M4_{degraded}$ lineage-specific SNPs

As well as the regions of gene loss within prophages and SpyCIM4, there were also a number of lineage specific core genome SNPs detected in all $M4_{degraded}$ BSAC *emm*4 isolates (including BSAC_bs1802), relative to completed reference genome MGAS10750 and all $M4_{complete}$ isolates. Of these, 19 were nonsynonymous, (Table 3.1), 10 were synonymous (Table 3.2), and 5 were situated within intergenic sequences (Table 3.3). Analysis of lineage-specific SNPs raised some interesting questions on possible phenotypes that might have been attributable to $M4_{degraded}$, particularly, though not exclusively, non-synonymous SNPs. Two fibronectin-binding proteins, *spy0114* and *spy0115,* carried a non-synonymous SNP each, Glu319Gly and Met11191Ile, respectively. A premature stop codon after 12aa in *Spy0196*, which encodes a carbonic anhydrase, was noted in $M4_{degraded}$ isolates. A gene encoding a putative helix-turn-helix, an apparently uncharacterised transcriptional regulator, *spy0642*, also featured a premature stop codon after 176aa in $M4_{degraded}$. All $M4_{degraded}$ isolates in the BSAC collection carried a non-synonymous SNP in *rgg/ropB* (Thre104Ile). A non-synonymous SNP was also observed in a Pbp1a family penicillin binding protein *spy1846* (Gly741Ser).

Table 3.1: Non-synonymous $M4_{degraded}$ lineage-specific SNPs

CDS[a]	Putative function	Size (bp)	Size (aa)	Nucleotide change[b]	Amino acid change[c]
Spy0114	Fibronectin-binding protein	1,620	539	G:C 955bp	Glutamine to Glycine 319 aa
Spy0115	Fibronectin-binding protein	4,104	1,367	G:A 3,573bp	Methionine to Isoleucine 1,191 aa
Spy0023	Phospho-ribosylformylglycinamidine synthase	3,774	1,257	C:A 1,645bp	Leucine to Isoleucine 549 aa
Spy0196	Carbonic anhydrase	501	166	C:T 37bp	Stop after 12aa
Spy0435	NADPH-dependent quinone reductase	666	221	A:G 301bp	Serine to Glycine 101 aa
Spy0464	Putative dehydrogenase	987	328	C:T 221bp	Threonine to Methionine 74 aa
Spy0642	Helix-turn-helix transcriptional regulator	1,485	494	G:T 529bp	Stop after 176 aa
Spy0996	Ribose-phosphate diphosphokinase	987	326	C:T 809bp	Proline to Leucine 270 aa
Spy1077	Lipoate protein ligase	1,020	339	T:C 956bp	Isoleucine to Threonine 319 aa
Spy1080	Glycine cleavage system protein H	333	110	G:A 215bp	Arginine to Histidine 72 aa

Table 3.1 continued

CDS[a]	Putative function	Size (bp)	Size (aa)	Nucleotide change[b]	Amino acid change[c]
Spy1085	Phosphopantothenoyl cysteine decarboxylase	546	181	G:A 517bp	Alanine to Threonine 173 aa
Spy1130	Cation diffusion facilitator family transporter	1,236	411	C:T 668bp	Alanine to Valine 223 aa
Spy1181	GTPase	1,314	437	G:A 752bp	Serine to Asparagine 251 aa
Spy1419	Sugar transport system permease	999	332	G:A 41bp	Glycine to Aspartic acid 14 aa
Spy1507	Tagatose-bisphosphate aldolase	978	325	C:T 572bp	Serine to Leucine 191 aa
Spy1573	Aspartate aminotransferase	1,251	416	T:A 19bp	Tyrosine to Asparagine 5aa
Spy1807	C5a peptidase	3,498	1,165	C:A 646bp	Leucine to Isoleucine 216aa
Spy1831	Rgg1/RopB	843	280	C:T 311bp	Threonine to Isoleucine 104 aa
Spy1846	Pbp1a family penicillin binding protein	2,337	778	G:A 2,221	Glycine to Serine 741 aa

[a] CDS as annotated in completed *emm*4 reference genome MGAS10750, [b] nucleotide change and position relative to the beginning of the CDS as annotated in MGAS10750, [c] amino acid change relative to the beginning of the putative protein as annotated in MGAS10750

Table 3.2: Synonymous $M4_{degraded}$ lineage-specific SNPs

CDS[a]	Putative function	Size (bp)	Size (aa)	Nucleotide Change[b]
Spy0065	30S Ribosomal protein S5	495	164	C:T 177 bp
Spy0192	ABC transporter ATP-binding protein	1,788	595	T:C 1,287 bp
Spy0467	Glycosyl-transferase	990	329	C:T 603 bp
Spy0731	Glutamine-hydrolyzing carbamoyl-phosphate synthase small subunit	1,083	360	T:C 246 bp
Spy0821	Single-stranded-DNA-specific exonuclease	2,211	736	C:T 1,221 bp
Spy1009	FAD:Protein FMN transferase	939	312	C:T 681 bp
Spy1082	NADH-dependent flavin oxidoreductase	1,200	399	G:A 738 bp
Spy1222	Phosphoenolpyruvate-protein phosphotransferase	1,734	577	C:T 1,581 bp
Spy1810	Mrp	1,176	388	T:C 75 bp
Spy1863	Urocanate hydratase	2,031	676	C:T 1,035 bp

[a] CDS as annotated in completed *emm*4 reference genome MGAS10750, [b] nucleotide change and position relative to the beginning of the CDS as annotated in MGAS10750

Table 3.3: Intergenic $M4_{degraded}$ lineage-specific SNPs

Flanking CDS[a]		Length	Nucleotide change[b]
Spy0216	*Spy0217*	89bp	C:T 23bp
Ribonuclease M5	16S dimethyl-adenosine transferase		
Spy0354	*Spy0355*	237bp	A:G 226bp
Hypothetical	Hypothetical		
Spy0397	*Spy0398*	1,212bp	G:A 1,164bp
Hypothetical	Blp-family class III bacteriocin		
Spy1636	*Spy1637*	169bp	A:G 75bp
Helix-turn-helix transcriptional regulator	Helix-turn-helix transcriptional regulator		
Spy1766	*Spy1767*	253bp	G:A 145bp
PolC-type DNA polymerase III	Proline-tRNA ligase (ProS)		

[a] CDS as annotated in completed *emm*4 reference genome MGAS10750, [b] nucleotide change and position relative to the beginning of the CDS as annotated in MGAS10750

3.4.6.1 Regulator sequence analysis for BSAC *emm*4

Sequences for transcriptional regulators derived from WGS data for BSAC *emm*4 were analysed and compared to both *emm*89 strain H293, and both completed *emm*4 reference genomes, MGAS10750 and MEW427. These included genes encoding the transcriptional regulators RofA, CovRS, Rgg1/RopB, Rgg2/MutR, Rgg3, Rgg4/ComR, FasBCAX, RivR and RocA. None of the BSAC *emm*4 were found to carry nonsynonymous SNPs in *rgg2*/*mutR, rgg3, rgg4*/*comR, rivR* or *rocA* relative to MGAS10750, MEW427 and H293. However, some variation was observed in a small number of the isolates in the regulators *rofA*, *covRS*, *rgg1*/*ropB*, and *fasB* of the *fasBCAX* operon.

Specifically, two $M4_{degraded}$ isolates carried nonsynonymous SNPs in *rofA*. Relative to the *rofA* allele of MGAS10750, BSAC_bs192 carried a G:A substitution at 38bp, resulting in the change from Gly to Glu at 13aa.

The *rofA* allele of BSAC_bs400 featured a premature stop codon after 177 aa, due to a G:T substitution at 532bp. Three isolates carried nonsynonymous sequence variations in the CovRS locus, BSAC_bs468, BSAC_bs1388 and BSAC_bs1802. Isolate BSAC_bs1802 was the only isolate carrying a nonsynonymous SNP in the response *covR*, with this allele featuring a missense substitution of G:A at 272bp, translating to a Gly to Asp at 91aa in the mature protein. Two isolates carried SNPs in the gene encoding the sensor kinase *covS*, BSAC_bs468, and BSAC_bs1388, though these SNPs were not identical. BSAC_bs468 carried two missense mutations in *covS*, one at A:C 640bp, translating to Thr to Pro 214aa within the histidine kinase/adenylyl cyclase/methyl-binding proteins/phosphatases (HAMP) domain, and a further missense mutation at G:A 1,135bp, resulting in Ala to Thr at 379aa. The *covS* allele of BSAC_bs1388 carries a deletion at 79bp which had led to a frameshift and the introduction of a premature stop codon after 36aa.

$M4_{degraded}$ BSAC *emm*4 carried a SNP in *ropB*, with a C:T missense mutation at 311bp, translating to Threonine to Isoleucine at 104 aa. This SNP was not found in MGAS10750, $M4_{complete}$ isolates, MEW427, or H293. A singular MEW427-like isolate, BSAC_bs192, carried an additional SNP in this gene, in which a substitution of T:A at 459bp resulted in the introduction of a premature stop codon after 153aa, with RopB/Rgg expected to be 281aa at full-length in the other BSAC *emm*4, H293, MGAS10750 and MEW427. One isolate, BSAC_bs1349, carried a non-synonymous SNP in *fasB*, a gene encoding a histidine protein kinase and forming part of the *fasBCAX* operon. In BSAC_b1349, *fasB* carries a missense substitution T:C at 226bp, resulting in the substitution of Serine for Proline at 76 aa.

3.4.7 Additional genetic characteristics of $M4_{complete}$ and $M4_{degraded}$

Analysis of the bacterial pangenome revealed a number of lineage-specific SNPs. A novel *emm*4 allele was detected in MEW427 and all $M4_{degraded}$ BSAC *emm*4, but not MGAS10750 nor $M4_{complete}$ BSAC *emm*4. This allele was the product of a fusion event between *emm*4, which encodes the *emm*4 M-protein, and the downstream *emm*-like *enn*, specifically, the 5' region of *emm*4 and the downstream 3' region of *enn*. This fusion was predicted to yield a novel M-protein variant. MGAS10750 carries the canonical *emm*4 gene, encoding M4, whereas MEW427 chimeric *emm*/*enn* gene encodes the 350aa isoform. The five $M4_{complete}$ BSAC *emm*4 isolates also encode the canonical M4, whereas the five $M4_{degraded}$ BSAC *emm*4 carry a chimeric allele, which encodes a distinct isoform of 392aa in length. The chimeric gene of MEW427 differs from that of the BSAC *emm*4 MEW427-like isolates by a 126bp deletion, which is predicted to yield a protein that is thus 42aa shorter. Both the MEW427 variant ($M4^{C350}$) and the MEW427-like BSAC *emm*4 variant ($M4^{C392}$) were recently identified in a localised population of *emm*4 isolates at a single hospital in Houston, Texas, USA (DebRoy *et al.*, 2018), which also detected four other variants of the chimeric protein. While *emm*4 with the canonical *emm* and *enn* genes are classically considered pattern E *S. pyogenes* ('generalists') according to the arrangement of the FCT region and predicted tissue tropism (Bessen *et al.*, 1996), genotype *emm*4 isolates carrying the novel allele may constitute a novel *emm*-pattern type (DebRoy *et al.*, 2018).

A 1,020bp gene encoding a putative microcin C7 self-immunity protein of 339 aa, denoted *spy0485*, was detected in MGAS10750 and $M4_{complete}$ isolates. This gene is predicted to encode an LD-carboxypeptidase, and may thus be responsible for catalysing the hydrolysis of tetrapeptide moieties in peptidoglycan (Barendt *et al.*, 2011). This gene shared 100% DNA sequence similarity across MGAS10750 and $M4_{complete}$ BSAC *emm*4, but was not found in MEW427, nor $M4_{degraded}$ BSAC *emm*4. When present, the *spy0485* gene appeared to form a four-gene cluster, along with three other genes located immediately downstream, *spy0486, spy0487* and *spy0488*, all of which are predicted to encode three histidyl-tRNA synthetases of 201bp/66aa, 486bp/161aa and 417bp/138aa in size, respectively. These genes were each detected in MGAS10750 and the $M4_{complete}$ isolates, and shared 100% DNA sequence similarity in each case. However, these genes were absent in MEW427 and $M4_{degraded}$ isolates. A BLAST database search revealed that both the presence and absence of this cluster between the flanking genes *spy0484* and *spy0489* are common in *S. pyogenes* genomes of multiple major *emm*-types, and neither variation is unique to *emm*4.

3.4.8 **Non-canonical prophage among *emm*4 isolates resolved by long-read sequencing**

For some *emm*4 isolates, there was evidence of prophage genes not associated with the canonical *emm*4 prophages, Φ10750.1/ Φ427.1, Φ10750.2/ Φ427.2 or Φ10750.3/Φ427.3, albeit these had not been resolved using short-read sequencing. Virulence factor typing had indicated that BSAC_bs192 and BSAC_bs1802, carried *speK/sla*, and *sdn*, respectively. Pangenome analyses suggested the presence of non-canonical *emm*4 prophage genes in these two isolates, as well as in BSAC_bs472 and BSAC_bs1388, and prophage had also not assembled within a single contig for some isolates due to the introduction of contig breaks. These isolates were subject to long-read Nanopore sequencing, revealing additional prophage assembled within single contigs, that are not typically part of the *emm*4 prophage complement. Interestingly, these appeared not to have undergone degradation (Remmington *et al.*, 2020) (Figure 3.8). BSAC_bs1802 carried an additional prophage (hereafter Φ1802.2), which was associated with *sdn*, and most closely related to *emm*87 NGAS743 prophage Φ743.3 (Athey *et al.*, 2016) integrated 12 genes downstream of Φ1802.1. BSAC_bs192 carried a *speK/sla*-associated prophage (hereafter Φ192.1), integrated between spy1388 (*yesN*) and spy1389 (*msrA*). Both BSAC_bs472 (integrated at the same site as Φ192.1) and BSAC_bs1388 were each lysogenised by a non-toxigenic prophage; hereafter Φ472.1 and Φ1388.1, respectively, integrated between *spy0673* (*pheT*) and *spy0674*. Φ192.1, Φ472.1 and Φ1388.1, all were most closely related to an undesignated *emm*28 prophage of MGAS292984 (Kachroo *et al.*, 2019; Remmington *et al.*, 2020), although the lysogeny module of Φ192.1 was near-identical to an undesignated *ssa*-associated prophage of *emm*75 isolate NCTC13751. Comparison of BSAC_bs400, which was not subject to long-read sequencing, with the completed sequence of BSAC_bs1388 revealed an additional prophage in this isolate, Φ400.1. Φ400.1 did not carry known toxin genes, and shared ~68% sequence similarity with Φ1388.1 (Figure 3.8).

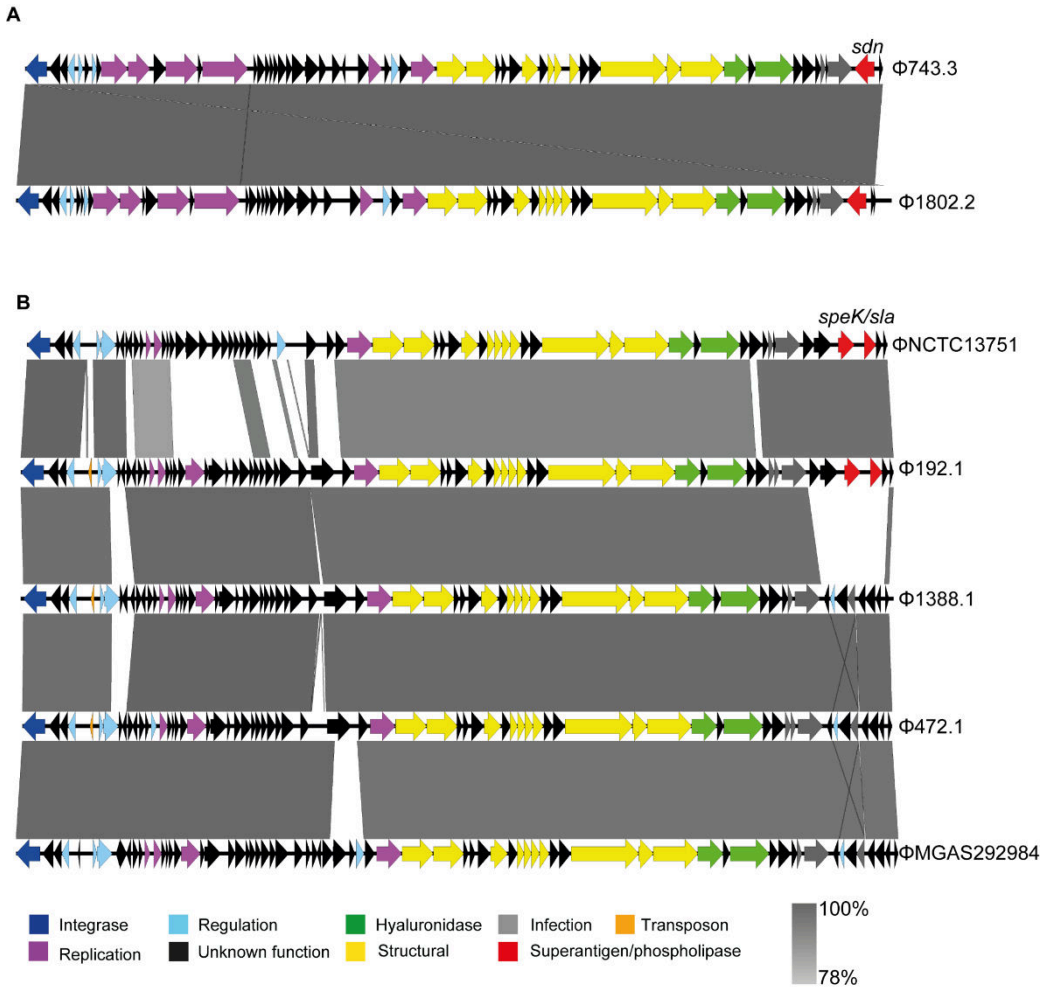

Figure 3.8: Additional non-canonical prophage identified by long-read sequencing

A) Long-read sequencing confirmed the presence of an *sdn*-associated prophage within BSAC_bs1802. (Φ1802.2), which was near-identical to *emm*87 NGAS743 Φ743.3 (Athey *et al.*, 2016). B) BSAC_bs192 was confirmed to harbour a *speK*/*sla* prophage (Φ192.1), which was closely related to an unnamed *speK*/*sla*-associated prophage of *emm*75 NCTC13751. Φ192.1 was highly similar non-toxigenic prophage resolved by long-read sequencing in BSAC_bs1388, (Φ1388.1), and BSAC_bs472 (Φ472.1). Although the lysogeny module of Φ192.1 was almost identical to that of the NCTC13751 prophage, Φ192.1, Φ427.1 and Φ1388.1 were most closely related to a non-toxin-encoding *emm*28 prophage in isolate MGAS29284 (Kachroo *et al.*, 2019). Predicted gene function was determined by BLAST, and assigned COG classification; integrase; navy, regulation; pale blue, hyaluronidase; green, infection (holin/lysin); grey, replication; purple, structural; yellow, superantigen/phospholipase; red, transposon; orange. Shaded grey scale bar indicates DNA sequence similarity (%). Figure taken from Remmington *et al.* 2020.

3.5 Discussion

A small population of *emm*4 isolates was collected by the BSAC from multiple locations in the United Kingdom between 2001 and 2011 (Turner *et al.*, 2019b). The BSAC *emm*4 were subject to typing for variable factors, specifically superantigens and DNases, revealing a generally consistent complement of these virulence factors

that was common to all ten. All ten isolates, like both MGAS10750 and MEW427, carried a *speC*/*spd1* prophage integrated at the 3' *attB* site of the putative dipeptidase gene*, pepD*/*Spy_0623*; an *spd3*-encoding prophage at the 3' *attB* site of the dTDP-glucose-4,6-dehydratase, *Spy_0890*; an *ssa*-associated prophage at the 5' *attB* site of the DNA-binding protein encoding gene, *hup/Spy_1329,* and the *emm*4 SpyCI element at the 5' *attB* of the DNA MMR protein *mutL*/*Spy_1896*. All ten isolates also carried the chromosomal DNases *spdB* and *spnA*, and the chromosomal superantigen *smeZ*.

Intriguingly, these isolates were largely divided into two clades, associated with either of two completed *emm*4 reference genomes, MGAS10750 and MEW427. Gubbins analyses identified clusters of SNPs, putative regions of recombination, within prophage-encoding regions of completed *emm*4 reference genome MEW427 and BSAC *emm*4 isolates that were more closely related to MEW427, BSAC_bs192, BSAC_bs400, BSAC_bs472, BSAC_bs696, and BSAC_bs1802. However, these regions were not observed in those isolates that were more closely related to MGAS10750, BSAC_bs468, BSAC_bs916, BSAC_bs1052, BSAC_bs1349 and BSAC_bs1388, nor MGAS10750 itself. Subsequent analysis revealed that these regions were not recombinant sequence, but rather regions of gene loss within MEW427 and the BSAC *emm*4 isolates clustering with this reference. The gene loss was found to have occurred within prophage-encoding regions, specifically, the three canonical genotype *emm*4 prophage Φ10750.1/Φ427.1, Φ10750.2/Φ427.2, Φ10750.3/Φ427.3, and SpyCI^{M4}. Φ427.1 was found to have lost ~70% of its gene content relative to the corresponding element, Φ10750.1, in MGAS10750. Φ427.2 was found to have lost ~50% of its gene content relative to the full-length Φ10750.2, and Φ427.3, was found to have lost ~26% of its genome when compared to the MGAS10750 prophage, Φ10750.3. The MEW427 SpyCIM4, was observed to have only ~44% of the genes that were found in the full-length element in MGAS10750. Similar patterns of loss were identified in four of the five BSAC *emm*4 clustering with MEW427. However, one isolate, BSAC_bs1802, despite clustering with MEW427, carried a different *speC*/*spd1* prophage to the other BSAC isolates (Φ1802.1) and a near-complete, only partially degraded Φ10750.3. Φ427.2 of BSAC_bs1802 was identical to the that of other BSAC *emm*4 clustering with MEW427. Gene loss, when observed, was noted to principally affect prophage genetic modules involved in genetic regulation, prophage excision/integration and replication, and to a less and more variable extent, structural genes required for the assembly of infectious particles. The loss of functional integrase genes in particular would be expected to preclude excision and particle formation.

Other M4$_{degraded}$ lineage-specific SNPs were present, that might impact bacterial phenotype, particularly those which were non-synonymous Two fibronectin-binding proteins, *spy0114* and *spy0115,* carried a non-synonymous SNP each, (Glu319Gly and Met11191Ile, respectively) that were considered in their capacity to modulate adhesion among the *emm*4 isolates to fibronectin *in-vitro* or *-in-vivo* (Brouwer *et al.*, 2016). A non-synonymous SNP was detected in *Spy0196,* a putative carbonic anhydrase encoding gene, leading to the introduction of a premature stop codon after 12 aa. These enzymes are perhaps best studied in *S. pneumoniae* and *S. mutans*, wherein they appear to mediate survival in ambient air, and fluctuations in other atmospheric and environmental

conditions (Burghout *et al.*, 2010), and thus might thus alter bacterial growth. All $M4_{degraded}$ and MEW427 carried a naturally occurring Rgg/RopB variant (Thr104Ile), but this was not predicted to affect activity based on comparison to published sequences (Ikebe *et al.*, 2010). Gene *Spy1846,* encoding a putative penicillin binding protein, also carried a non-synonymous SNP in $M4_{degraded}$ (Gly741Ser). However, none of the BSAC *emm*4, which had previously been subject to rigorous and well-standardised antimicrobial sensitivity testing by the BSAC, demonstrated a phenotype of reduced sensitivity or resistance to β-lactams (Turner *et al.*, 2019b). This variant was unrelated to penicillin binding protein SNPs to which reduced sensitivity to β-lactams in *S. pyogenes* was recently described (Musser *et al.*, 2020; Vannice *et al.*, 2019). A naturally occurring C5a peptidase variant was also observed (Leu216Ile), but this appeared to fall within a repeat region. Three isolates carried variant CovRS alleles; BSAC_bs468 (CovS; Thr241Pro and Ala379Thr), BSAC_1388 (CovS; 36aa-STOP) and BSAC_bs1802 (CovR; Gly91Asp), which may have a number of far-reaching, isolate-specific consequences depending on the impact of the SNP, including enhanced virulence and changes in toxin expression (Federle *et al.*, 1999; Galloway-Peña *et al.*, 2018; Graham *et al.*, 2002; Sumby *et al.*, 2006; Turner *et al.*, 2009).

The presence of an equal number of $M4_{complete}$ and $M4_{degraded}$ *emm*4 isolates in the BSAC collection, coupled with the data in this Chapter on prophage content, prompted the examination of the larger *emm*4 population. This proved highly illuminating, and the initial observation that the isolates in this small population formed two lineages clustering with either reference genome, was recapitulated by further work carried out by project supervisor Dr. Claire E. Turner (Remmington *et al.*, 2020). This confirmed that the international population of *emm*4 *S. pyogenes*, ratified by the inclusion of all available *emm*4 whole genome sequence data, comprised two broad lineages, one of which was characterised by cryptic prophage (Figure 3.9).

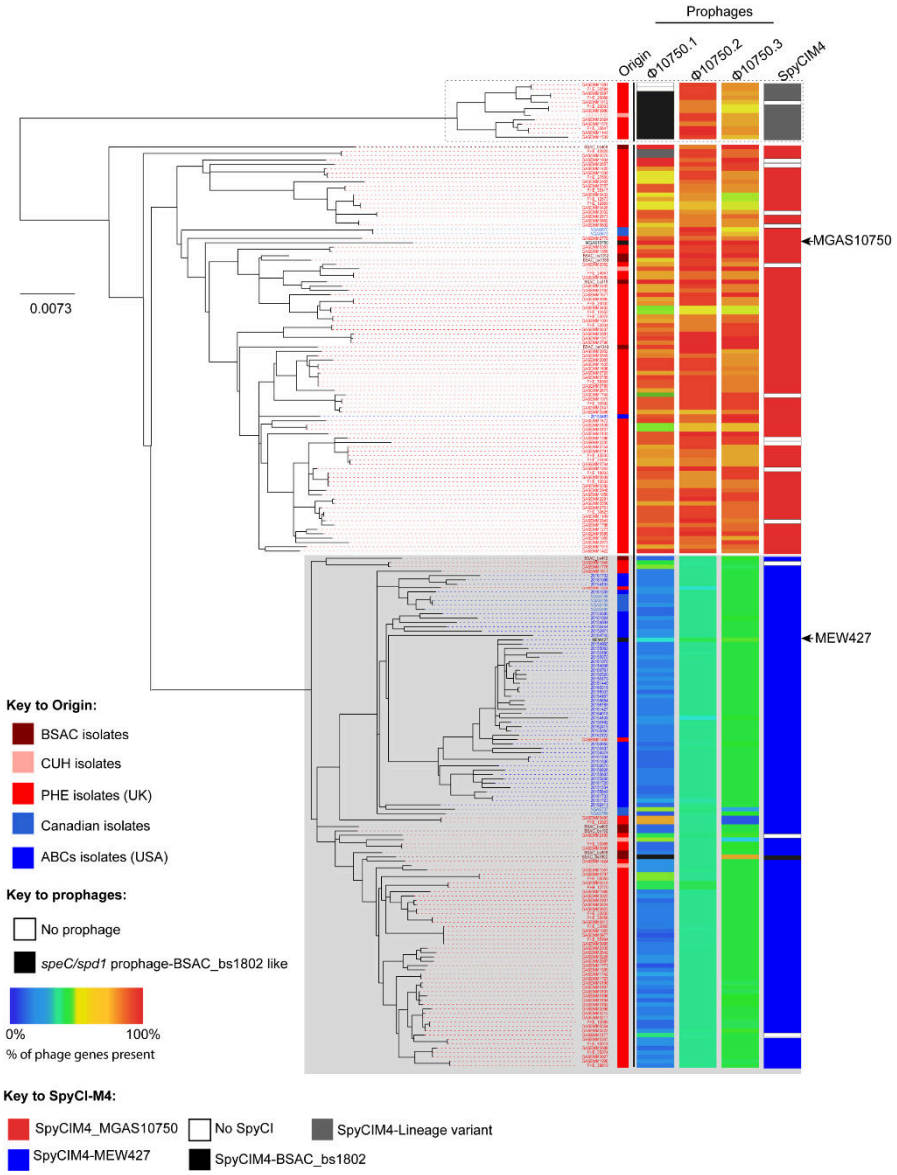

Figure 3.9: International *emm*4 lineage associated with cryptic prophage

A mid-rooted phylogenetic tree was constructed using core SNPs following mapping of publicly available short-read WGS data from 223 international *emm*4 isolates, and completed reference genomes MGAS10750 and MEW427, indicated with black arrows. Lineages were defined as isolates clustering with MGAS10750 (unshaded), those clustering with MEW427 (shaded), and a small lineage comprising UK 13 isolates with partially degraded, or variant MGE relative to the rest of the population (dotted box). Prophage gene content was determined for each isolate by BLAST for each canonical *emm*4 prophage Φ10750.1 (column 1), Φ10750.2 (column 2), and Φ10750.3 (column 3), and used to calculate a percentage, expressed as colour scale per figure key. SpyCIM4 variants corresponded with these respective lineages (column 4). Those associated with the $M4_{degraded}$ lineage carrying the degraded variant (blue bar), and those associated with $M4_{complete}$ carrying the full-length $SpyCI^{M4}$ (red bar), as determined by sequence alignment. BSAC_bs1802 SpyCI, unique among the *emm*4 population, is shown as a black bar. Where elements were absent, this is indicated with a white bar. Isolate names are linked to corresponding branches by dotted lines, and colour coded by geographical origin, per figure key (UK; red, and North America; blue). Scale bar represents substitutions per site. Figure taken from Remmington *et al.* 2020.

Intriguingly, the North American population of *emm*4 *S. pyogenes* comprised predominantly $M4_{degraded}$ isolates, whereas both $M4_{complete}$ and $M4_{degraded}$ isolates were found in the United Kingdom. Although it is clear from the nature of the BSAC collection as bloodstream isolates that full-length prophage are inessential for *emm*4 to attain full virulence, the geographic distribution of these two lineages raised interesting questions on the difference between $M4_{complete}$ and $M4_{degraded}$ isolates. There may be other factors responsible, in addition to, or independent of prophage degradation, notably a chimeric *emm* gene, formed by the splicing of *emm* and the *emm*-like *enn* (DebRoy *et al.*, 2018). This chimeric gene was indeed detected in the MEW427-like BSAC *emm*4, and as previously reported, MEW427, and chimeric *emm* genes have subsequently been identified in multiple *emm*-types (Frost *et al.*, 2020).

Resolving prophage regions in bacterial genomes with short-read sequencing methods is difficult due to the homologous and mosaic nature of prophage. Unless an element is located within a single contig, discerning gene presence and absence is challenging. It is reassuring, however, that for all isolates, bar BSAC_bs1802, patterns of gene loss were identical, or near identical to the corresponding completed reference genome sequences. Detecting additional phage that are 'atypical' of the genotype is also a noted constraint, which is especially concerning given the potentially calamitous epidemiological consequences that loss and gain of prophage in *S. pyogenes* can evoke (Al-Shahib *et al.*, 2016; Stephen B. Beres *et al.*, 2004; Davies *et al.*, 2015; Tse *et al.*, 2012).

Long-read sequencing of four BSAC *emm*4 post hoc confirmed the gene content of Φ1802.1, as had been achieved by repeating assembly using alternative software (Remmington *et al.*, 2020). Long-read sequencing also demonstrated the presence of four non-canonical and hitherto unidentified *emm*4 prophage in these isolates, which was not achieved with short-read sequencing; Φ192.1 and Φ1802.2, carrying *speK/sla* and *sdn,* respectively, and Φ472.1, Φ1388.1 and, Φ400.1, which carried no toxin genes. Interestingly, these prophage did not appear to have undergone substantive gene loss. The similarity of these prophage to existing elements of other genotypes infers the exchange of prophage between and within genotypes, in this case *emm*28 and *emm*75, and this was supported in the literature (Brüssow and Desiere, 2001; Desiere *et al.*, 2001; Canchaya *et al.*, 2003). Given the human-restricted nature of *S. pyogenes*, such exchanges likely occurs in humans, and can be recapitulated in murine models of infection (Broudy and Fischetti, 2003). However, it seems likely that streptococci with less restricted host-ranges, such as the emerging pathogen *S. dysgalactiae* sbsp. *equisimilis* (Chochua *et al.*, 2019), may augment accessibility of this arsenal to substantially more discerning species, such as *S. pyogenes* or the equine pathogen *S. equi* sbsp. *equi* (Holden *et al.*, 2009; Turner *et al.*, 2019a). As long-read sequencing becomes more affordable and widely available, it seems likely this technology will further illuminate the role prophage play in the biology of *S. pyogenes* and related species.

The data in this chapter led to the hypothesis that prophage of $M4_{degraded}$ *emm*4, having undergone marked gene loss in prophage- and $SpyCI^{M4}$-encoding regions would exhibit differential prophage excision, with $M4_{degraded}$

prophage unable to exit the bacterial chromosome. If proven, this could conceivably impact the growth, toxin expression and overall virulence of these isolates. These hypotheses were explored in subsequent chapters.

3.5.1 Chapter 3: Key findings

- Substantial gene loss was observed within prophage- and SpyCI-encoding regions in completed *emm*4 reference genome MEW427 but not MGAS10750
- Gene loss primarily affected prophage modules concerned with genetic regulation, replication and to a lesser extent, structural genes and thus may impact prophage excision and integration
- Cognate virulence genes were not lost
- The patterns of gene loss observed in MEW427 were also observed in half of the BSAC *emm*4 isolates, specifically those clustering phylogenetically with MEW427
- BSAC *emm*4 that clustered phylogenetically with MGAS10750 did not exhibit gene loss in prophage-encoding regions
- The global *emm*4 population includes an international lineage broadly associated with gene loss within prophage- and SpyCI-encoding regions

Chapter 4

4. Prophage induction and gene expression

4.1 Introduction

Lysogenic toxin-encoding bacteriophage are important contributories to the molecular biology and pathogenesis of number of Gram-positive and Gram-negative bacteria, and frequently encode bacterial toxins (Canchaya *et al.*, 2003; Holmes, 2000; Krüger and Lucchesi, 2015; McShan and Nguyen, 2019; Sakaguchi *et al.*, 2005; Xia and Wolz, 2014). Much of their lifecycle is spent as a stably integrated element, a state which is under tight regulatory control (Oppenheim *et al.*, 2005). Although integrated prophages are transmitted to daughter cells by vertical gene transfer, and perhaps by recombination (Turner *et al.*, 2015), to achieve high-frequency transmission to new hosts, prophage must excise from the bacterial chromosome and enter the lytic cycle, eventually forming prophage particles.

Prophage induction may occur in response to external stimuli and the initiation of the SOS-response by the host bacterium. In many bacterial species, this response is triggered by DNA damage or interference with DNA synthesis. Antibiotics such as ciprofloxacin and other fluoroquinolones, chemotherapeutic agents such as mitomycin C, in addition to other chemical and physical agents including hydrogen peroxide and ultraviolet light, have well-described potential to elicit this effect (Banks *et al.*, 2003; Fernández De Henestrosa *et al.*, 2000; Maiques *et al.*, 2006; Michel, 2005; Oppenheim *et al.*, 2005; Zhang *et al.*, 2000). These insults can damage DNA in a variety of ways, such as creating double-strand breaks, and inducing crosslinks. In a number of bacterial species, the highly conserved DNA repair and maintenance protein and inducer RecA is activated by DNA damage, binding ssDNA to form a nucleoprotein filament and promoting self-cleavage of the downstream repressor, LexA. This cleavage subsequently de-represses SOS genes and triggers the SOS-response. The phage-encoded repressor cI is also cleaved on the initiation of the SOS-response, resulting in the relief of its repressive activity over genes which are involved in prophage induction, and interlacing DNA damage with prophage induction, and inextricably linking the fate of the lysogenic virus with the lysogen (Oppenheim *et al.*, 2005). Prophage induction may also occur at comparatively lower frequencies throughout bacterial growth (Alexeeva *et al.*, 2018; Cortes *et al.*, 2019; Nanda *et al.*, 2015), possibly due to low-level initiation of the SOS-response as the corollary of accumulated DNA damage sustained during binary fission (Nanda *et al.*, 2015).

The classical genetic apparatus associated with the SOS-response have hitherto not been identified in a number of the streptococci, including *S. pyogenes* (Gasc *et al.*, 1980; Varhimo *et al.*, 2007). However, analogous compensatory systems are reported (Boutry *et al.*, 2013; Varhimo *et al.*, 2007; Žgur-Bertok, 2013), and the prophage of *S. pyogenes* and other streptococci are inducible with DNA-damaging agents (Banks *et al.*, 2003;

Euler *et al.*, 2016; Hyder and Streitfeld, 1978; Ingrey *et al.*, 2003; López *et al.*, 2014). There is also evidence for spontaneous prophage induction in *S. pyogenes* (Banks *et al.*, 2003), and other streptococci of clinical interest (Alexeeva *et al.*, 2018; Brassil *et al.*, 2020; Carrolo *et al.*, 2010). It seems likely that the two occur simultaneously, particularly *in-vivo*, wherein one might expect myriad immunological, anatomical and possibly pharmacological inducers will coincide with the expansion of bacterial populations, itself leading to the accumulation of genetic and biochemical spontaneous inducers.

Prophage excision, both induced and spontaneous, appears to play a role in the transcriptional control of cognate virulence factors in *S. pyogenes* and other bacteria. Occasionally, but not exclusively, this may be due to a gene dosing effect, or evidence thereof (Aertsen *et al.*, 2005; Balasubramanian *et al.*, 2019; Banks *et al.*, 2003; Broudy *et al.*, 2001; Broudy and Fischetti, 2003; Cortes *et al.*, 2019; Ingrey *et al.*, 2003). There is also evidence that other, less well-studied mobile genetic elements of *S. pyogenes* can impact the expression of genes beyond the elements themselves (Eraso *et al.*, 2020; Kachroo *et al.*, 2019), and that this influence can be mediated by their capacity to excise and integrate (Nguyen and McShan, 2014; Scott *et al.*, 2012), as is observed among the SpyCI. The regulation, control and utilisation of prophage-borne toxins is thus far from clear, particularly in the context of infection. Even less well described are the roles of degraded, decayed or 'defective' prophage.

4.2 Chapter hypotheses and aims

The present chapter sought to characterise the inducibility of the prophage and the SpyCI of *emm*4 *S. pyogenes*. Specifically, to determine if, between the $M4_{complete}$ and $M4_{degraded}$ isolates, differential induction could be demonstrated. It was further hypothesised that $M4_{complete}$ isolates might express higher levels of prophage-encoded virulence factors, relative to $M4_{degraded}$ isolates, possibly, in the case of the former, owing to an inability to leave the chromosome, replicate, and elicit a gene dosing effect. On the part of SpyCIM4, it was hypothesised that permanent integration, if observed, would abolish expression of *mutS*.

4.3 Methods

4.3.1 Prophage and SpyCI induction using mitomycin C

Overnight liquid cultures of all 10 BSAC *emm*4 were diluted 1:10 and grown to OD_{600} 0.3, whereupon these were split into two tubes; one aliquot was treated with mitomycin C (Sigma) (0.2 μg/mL) and the second served as an untreated control, as previously described (Banks, Lei and Musser, 2003). Cultures were incubated for an additional 3 hours at 37°C in an atmosphere enriched with 5% (v/v) CO_2, before the bacterial pellet was harvested and frozen at -20°C for DNA extraction the following day (for full protocol, see main methods section 2.2.2), before application of the prophage induction PCR(s) (Figure 2.1). Cultures were initially also treated with

hydrogen peroxide (Sigma) (0.5mM), however, this was found to induce prophage less strongly and more variably than mitomycin C, which consistently and strongly induced phage.

4.3.2 Spontaneous prophage and SpyCI induction experiments

For spontaneous prophage induction experiments, overnight cultures of all 10 BSAC *emm*4 *S.* were washed once in fresh THB (Oxoid), diluted 1:10 and OD_{600} measured hourly. Bacteria were harvested by centrifugation at 3, 5 and 8 hours, based on the correspondence of these timepoints with early-, mid- and late-log stages of bacterial growth, as determined by growth curve (these data are shown later in Chapter 5, Figure 5.1). To detect SpyCI excision, cultures of BSAC_bs1349 and BSAC_bs696 were harvested by centrifugation at 30 minutes intervals for the first 3 hours, and hourly thereafter, with the final timepoint being 8 hours. All pellets were immediately frozen at -20°C and DNA extracted and standardised the next day before being subject to PCR with prophage/SpyCI induction primers.

4.3.3 Detection of gene expression and copy number by RT-qPCR and qPCR

Synthesis of cDNA and qPCR was carried out also as described in the main methods (section 2.2.11), without modification. Transcript copies of *speC*, *spd3* and *ssa* were measured by qPCR using a method to include a standard curve, generated using a plasmid (per section 2.2.9 of main methods) and transcript copies of the target gene were thus normalised using housekeeping gene *gyrA* transcripts. DNA extraction was performed on thawed samples, (as described in methods section 2.2.2) and was standardised to 40 ng/µL (200ng per reaction) for qPCR to detect gene copies of *speC*, using the aforementioned qPCR method, and *speC* gene copies were normalised to copies of housekeeping gene *gyrA* (for cycling parameters, see main methods Figure 2.3).

4.3.4 Semi-quantitative detection of *mutS* and *mutL* transcripts

RNA was extracted and converted to cDNA for all 10 BSAC *emm*4 after 3 hours growth in THB. cDNA from this timepoint was selected as this corresponds with the stage of growth at which, according to published literature (Nguyen and McShan, 2014; Scott *et al.*, 2012), excision of SpyCIM4 was purportedly lowest, and expression of *mutS* thus abolished. For each isolate, 50 ng of cDNA per reaction, and the corresponding RT-negative were probed by standard PCR to measure *mutL* and *mutS* transcripts, semi-quantitatively, by agarose gel electrophoresis and UV transillumination.

4.4 Results

4.4.1 Prophage of M4$_{complete}$ are inducible but prophage of M4$_{degraded}$ are not

To test experimental induction of *emm*4 prophage, liquid cultures were treated with mitomycin C and DNA was extracted. Primers were used to detect integrated prophage, and where present, excised phage. As hypothesised, Φ1075.1 of M4$_{complete}$ isolates were induced spontaneously in untreated samples (*c.* 6 hours growth), and were more strongly induced following treatment with mitomycin C (Figure 4.1A). Contrarily, Φ427.1 of M4$_{degraded}$ were not induced spontaneously in the untreated cultures, nor cultures treated with mitomycin C (Figure 4.4B). The exception to this was unusual M4$_{degraded}$ BSAC_bs1802. For this isolate, the unique *speC*/*spd1*-encoding Φ1802.1 and was strongly induced following treatment with mitomycin C (Figure 4.1C).

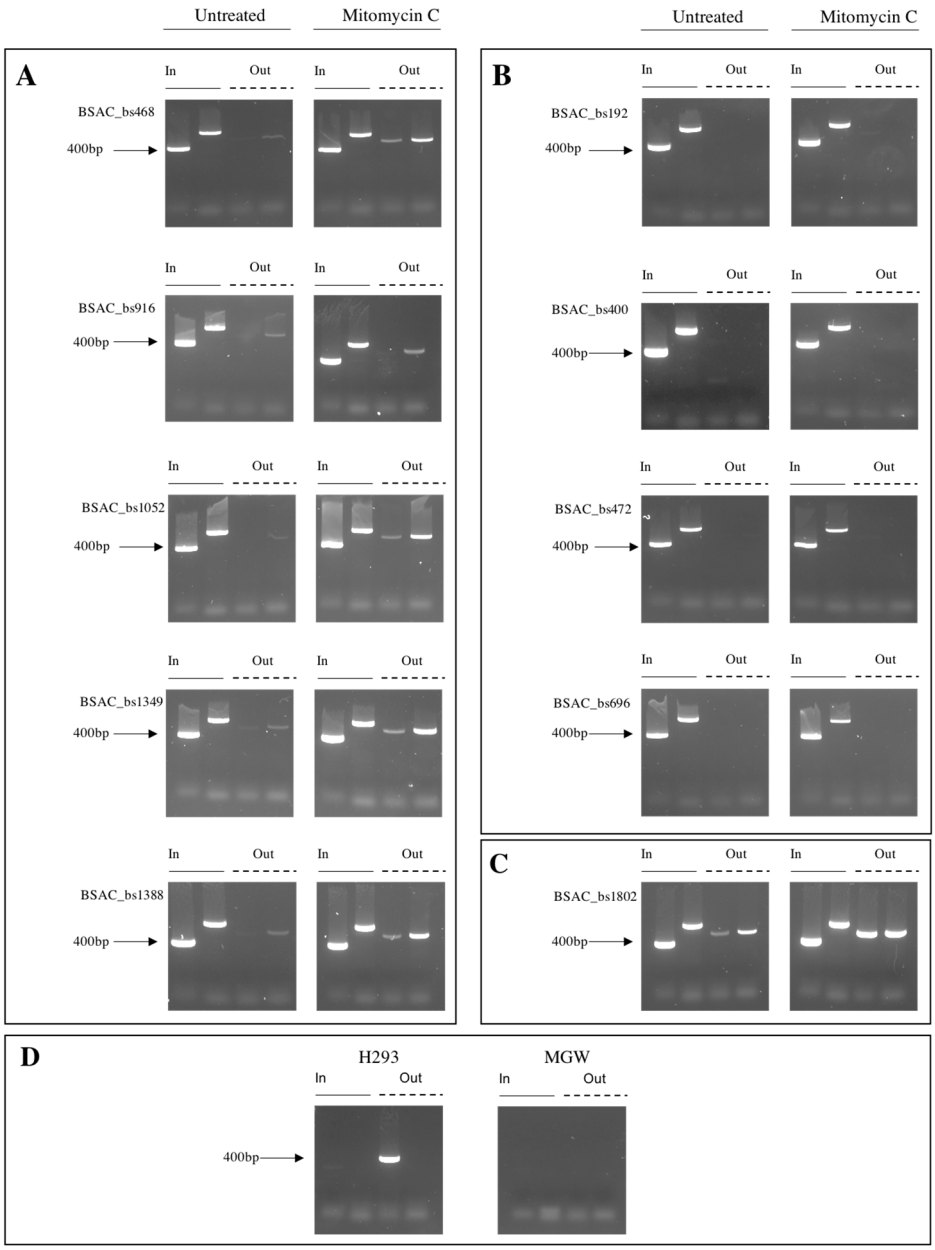

Figure 4.1. Φ10750.1/Φ427.1 induction in liquid culture using mitomycin C

Amplification with the two primer combinations denoted 'In' indicate an integrated prophage, whereas amplification with the two primer combinations denoted 'out' indicate excision of this prophage. A) Φ10750.1 in all five $M4_{complete}$ was detected, and found to be induced spontaneously during growth in liquid culture, and was more strongly induced with mitomycin C. B) The corresponding Φ427.1 of $M4_{degraded}$ was detected in its integrated conformation, however, this prophage was not observed in its extra-chromosomal state in untreated cultures, nor could the excised prophage be detected in DNA samples extracted from bacterial cultures exposed to mitomycin C. C) Φ1802.1 of BSAC_bs1802 was detected as an integrated and excised element in both culture conditions. D) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No bands were observed in MGW control. DNA was standardised to 40 ng/µL prior to PCR.

The same DNA samples were probed for integration and excision of Φ10750.2/Φ427.2. It was only in DNA extracted from $M4_{complete}$ that the excised prophage could be detected (Figure 4.2A). In these cases, the full-length, excised prophage was detected weakly in DNA extracted from untreated cultures, and was more strongly induced in those treated with mitomycin C. Conversely, Φ427.2 of $M4_{degraded}$ was not detected as having excised under either culture conditions (Figure 4.2B).

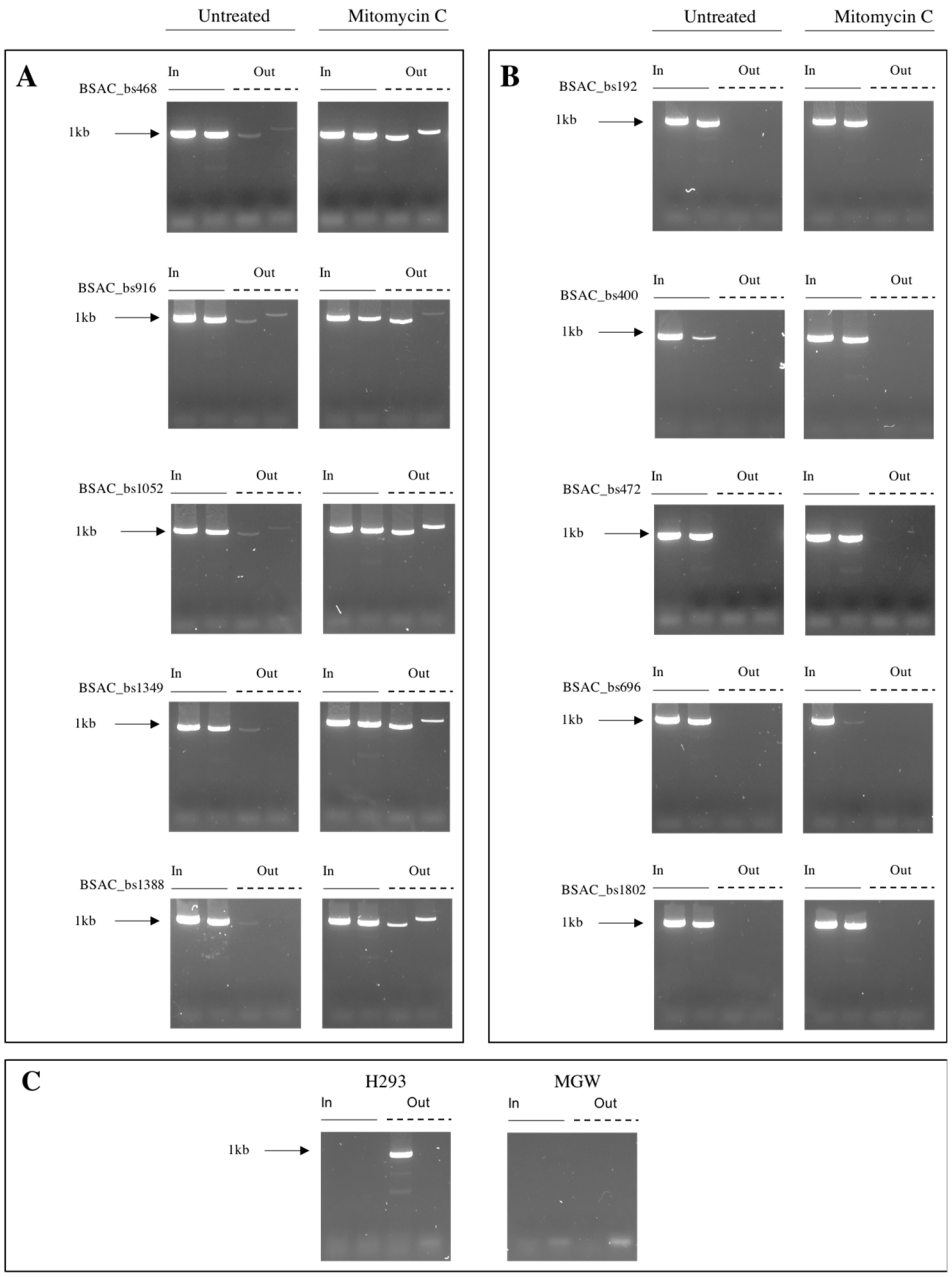

Figure 4.2.: Φ10750.2/Φ427.2 induction in liquid culture using mitomycin C

Two primer combinations were used to detect integrated prophage 'in', and a further two were used to detect prophage in their extra-chromosomal conformation following excision from the bacterial chromosome ('out'). A) Excision of Φ10750.2 was detected in all five $M4_{complete}$ isolates and could be induced to excise more strongly with mitomycin C. B) Φ427.2 in all four $M4_{degraded}$ and BSAC_bs1802 was not induced under either culture condition. C) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No bands were observed in the MGW negative. DNA samples were standardised to 40 ng/μL prior to PCR.

Primers were also designed to detect the *emm*4 prophage encoding the streptococcal superantigen *ssa*, Φ10750.3/Φ427.3. These were applied to the same standardised DNA samples to determine the excision potential of these prophage. Full-length $M4_{complete}$ Φ10750.3 could be induced by mitomycin C, and were found to be induced to a lesser extent in untreated cultures (Figure 4.3A). Conversely, Φ427.3 of $M4_{degraded}$ isolates appeared unable to excise from the chromosome under either culture conditions (Figure 4.3B). Unusual isolate BSAC_bs1802, carrying a partially degraded Φ10750.3, was inducible under both culture conditions (Figure 4.3C).

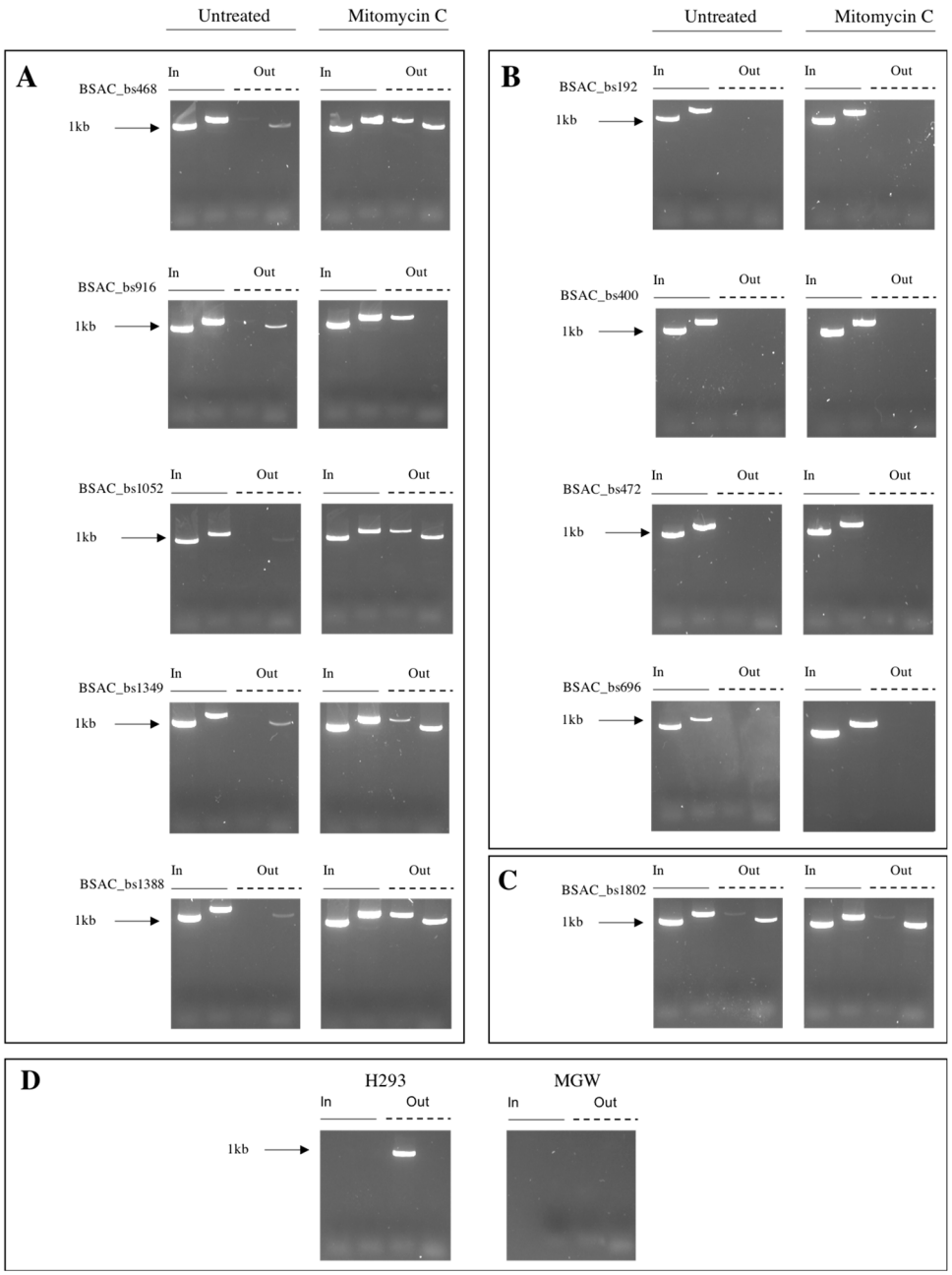

Figure 4.3. Φ10750.3/Φ427.3 induction in liquid culture using mitomycin C

Prophage excision ('in') was detected with two pairs of primers, and by alternating the combination of these primers, the excised prophage ('out'), where present, could be detected. A) PCR on DNA extracted from cultures of all five $M4_{complete}$ isolates demonstrated that Φ10750.3 was weakly induced in untreated cultures, and to a comparatively stronger degree following treatment with mitomycin C. B) Φ427.3 in $M4_{degraded}$ isolates was not detected in its induced, excised conformation under either culture conditions. C) Φ1802.3 of BSAC_bs1802 exhibits only very minor gene loss, and was strongly induced under both conditions. D) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No bands were observed in MGW negative. DNA was standardised to 40 ng/μL prior to testing.

4.4.2 **Prophage induction during *S. pyogenes* growth in THB**

Prophage may be induced spontaneously during bacterial growth (Alexeeva *et al.*, 2018; Cortes *et al.*, 2019; Nanda *et al.*, 2015). It was unclear whether this may occur in $M4_{complete}$ or $M4_{degraded}$, or if these events might be associated with a particular timepoint among isolates from either lineage. To establish this, DNA was extracted from bacteria harvested at early-, mid-, and late-log and tested for prophage induction by PCR. As expected, Φ10750.1 of $M4_{complete}$ were weakly spontaneously induced during bacterial growth (Figure 4.4A), however, Φ427.1 of $M4_{degraded}$ were not induced at any sampling time (Figure 4.4B). Φ1802.1 of BSAC_bs1802 was found to be strongly induced throughout growth (Figure 4.4C). It did not appear that induction of $M4_{complete}$ was occurring more so at one timepoint than others, though possibly amplification with primer combination AD was more readily detected at 3 hours.

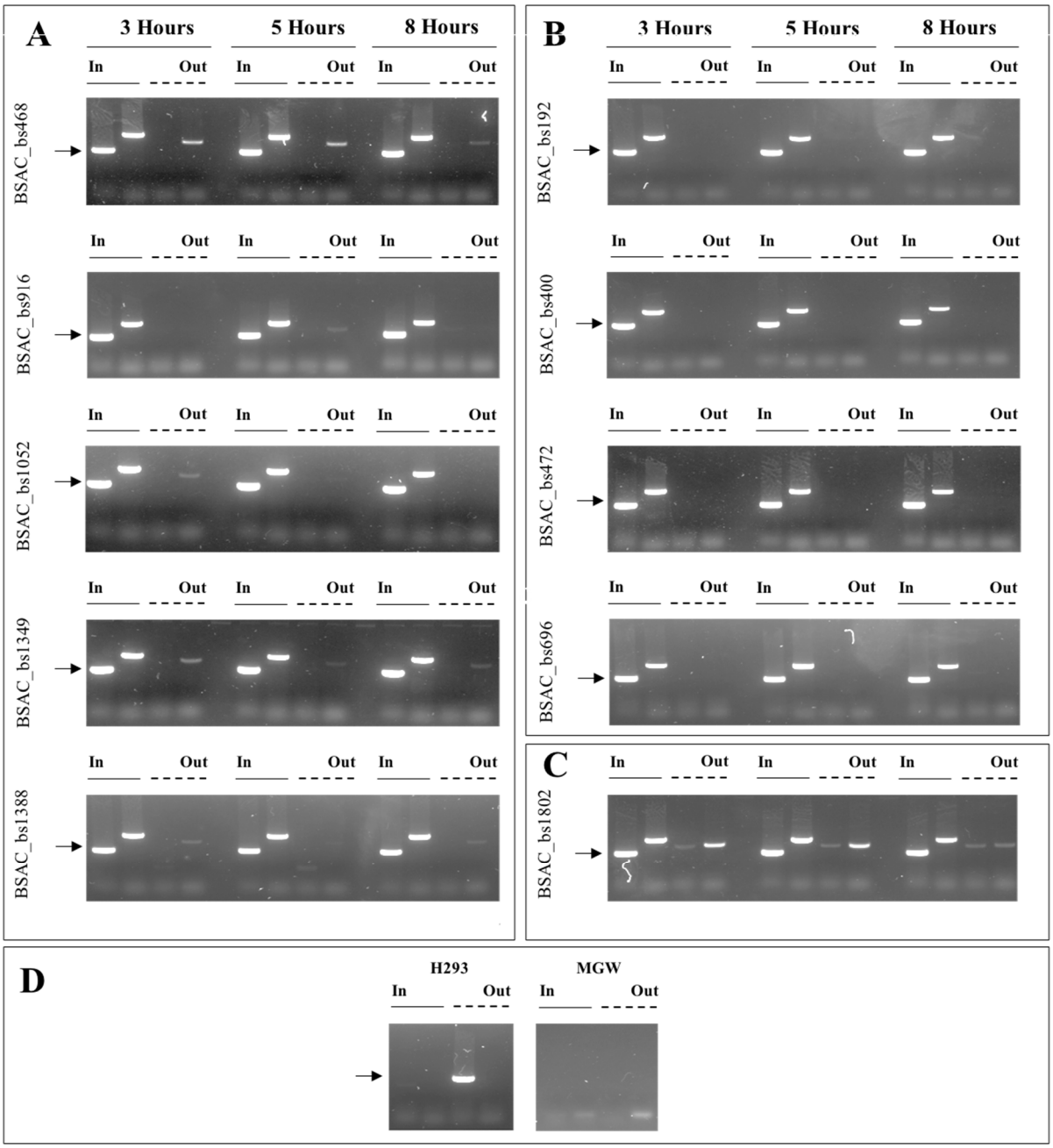

Figure 4.4.: Spontaneous Φ10750.1/Φ427.1 induction during bacterial growth

Amplification with two primer combinations was used to detect integrated prophage ('In'), and different combinations of these same primers were used to produce a further two combinations was used to detect excised prophage ('out'). A) Φ10750.1 was weakly induced at least one timepoint per isolate. B) Φ427.1 of $M4_{degraded}$ did not exhibit the capacity spontaneous induction. C) $M4_{degraded}$ BSAC_bs1802 carries a different *speC*/*spd1*-encoding prophage to the other BSAC *emm*4 and this was detectable as being strongly induced throughout growth. DNA samples were standardised to 40 ng/µL before PCR. D) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No amplification occurred in the MGW negative.

Primers for Φ10750.2/Φ427.2 were applied to the same standardised DNA samples. As was expected, Φ10750.2 was weakly spontaneously induced during growth, though it did not appear that induction occurred at any specific time, and did not increase with later timepoints (Figure 4.5A). Φ427.2 was not induced at any stage (Figure 4.5B).

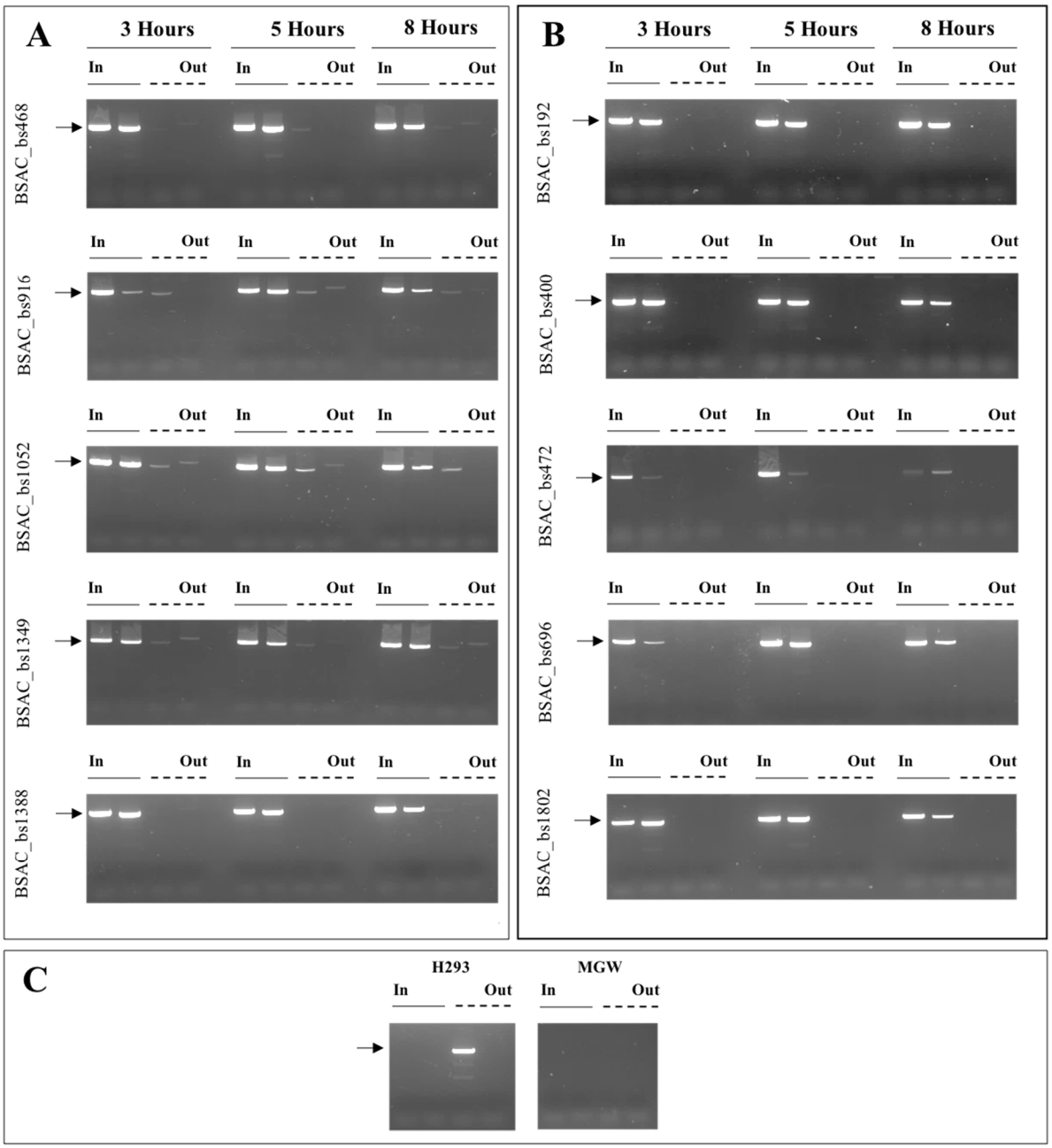

Figure 4.5.: Spontaneous Φ10750.2/Φ427.2 induction during bacterial growth

Two primer pairs were used to detect the prophage in its integrated state within the bacterial chromosome ('In'), and different combinations of these same primers were used to detect its extra-chromosomal conformation ('out'). A) Φ10750.2 induction was observed throughout growth in all bar one $M4_{complete}$, BSAC_bs1388. B) Φ427.3 of $M4_{degraded}$ could not be detected as induced at any of the three time points. C) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No amplification occurred in MGW negative. DNA was standardised to 40 ng/µL prior to PCR.

By applying the Φ10750.3/Φ427.3 phage conformation PCR to the same DNA samples, it was found, that prophage of $M4_{complete}$ isolates were induced throughout growth, and were similar across all timepoints (Figure

4.6A). Conversely, Φ427.3 was not induced during bacterial growth at any of the three timepoints (Figure 4.6B).but Φ10750.3 of BSAC_bs1802 was spontaneously induced (Figure 4.6C).

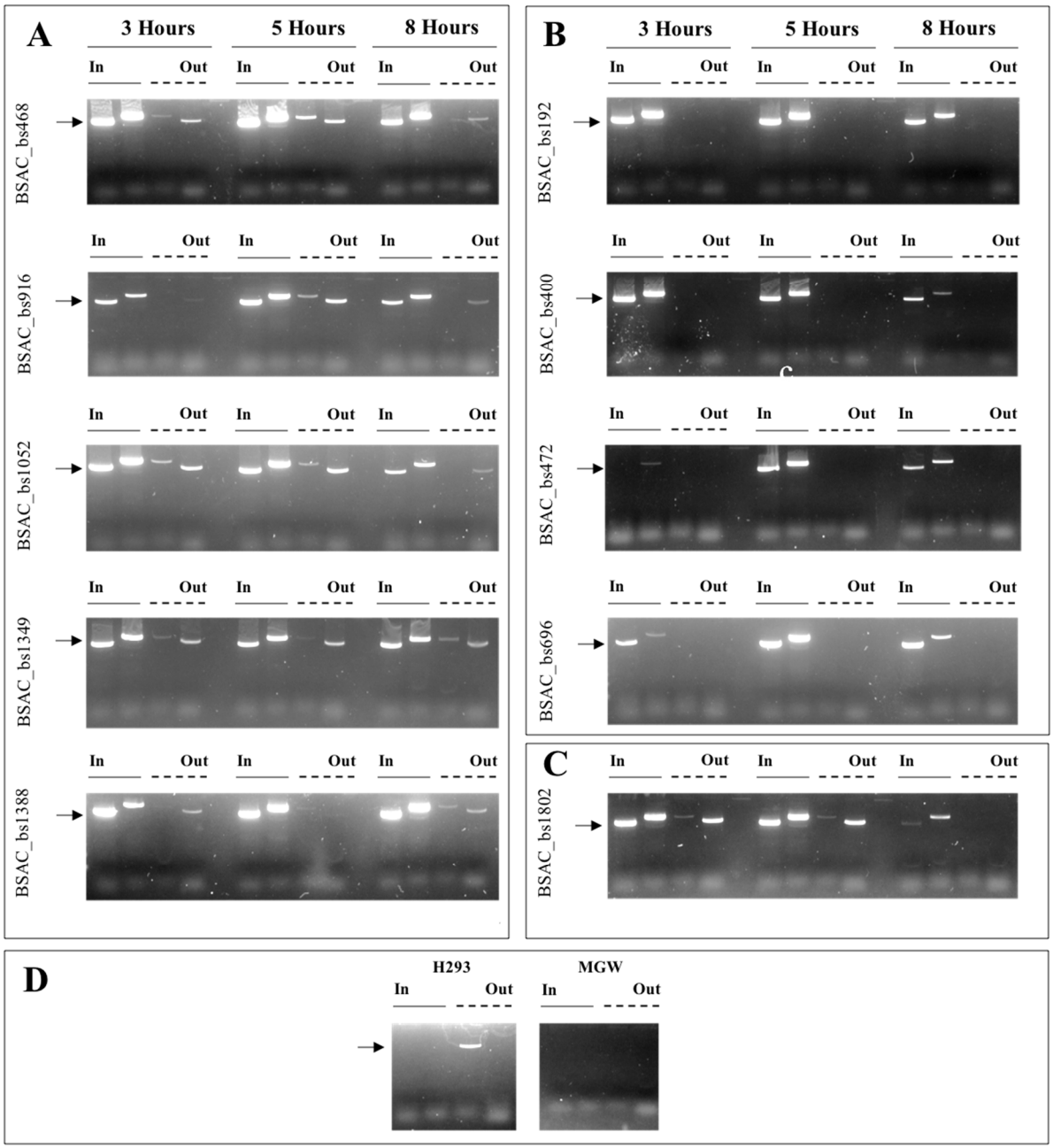

Figure 4.6.: Spontaneous Φ10750.3/Φ427.3 induction during bacterial growth

Two primer pairs were used to identify the integrated prophage element ('in'), and two pairs were used to detect the excised prophage ('out'). A) $M4_{complete}$ Φ10750.3 was strongly spontaneously induced throughout growth. B) Φ427.3 of $M4_{degraded}$ was not induced at any time point. C) Φ10750.3/Φ427.3 of BSAC_bs1802 was detectable as a spontaneously induced element. D) H293 DNA was used as an internal control, as it has no prophage, and this was confirmed by amplification with AD. No amplification was observed in MGW negative. DNA was standardised to 40 ng/μL prior to PCR.

4.4.3 **Prophage toxin gene expression during bacterial growth**

Prophage induction in *S. pyogenes* may increase the expression of cognate virulence factors in *S. pyogenes* (Banks *et al.*, 2016; Broudy *et al.*, 2002, 2001). Having demonstrated that the of $M4_{complete}$ isolates exhibit full capacity for excision across all three canonical genotype *emm*4 prophage, and confirmed that the corresponding prophage of $M4_{degraded}$ isolates could not be induced, it was apposite to examine what impact this might have on toxin gene transcription, if any. RNA samples collected from the same experiments as the timepoint DNA samples were thus probed for expression of *speC*, *spd3* and *ssa*, encoded by Φ10750.1/Φ427.1, Φ10750.2/Φ427.2, Φ10750.3/Φ427.3, respectively.

At three and five hours, *speC* expression was similar between $M4_{complete}$ and $M4_{degraded}$ isolates (p=0.395, Mann Whitney, and p=>0.9999, and Mann Whitney, respectively). However, at eight hours, the expression of *speC* by all *emm*4 isolates increased, and began to diverge, with $M4_{complete}$ isolates expressing *speC* more highly than $M4_{degraded}$ isolates though this did not reach significance (p=0.0952, Mann-Whitney) (Figure 4.7A). Intriguingly, the increased expression did not appear to be associated with a precipitant increase in *speC* gene copies (Figure 4.7B). Two isolates did exhibit an increase in gene copies at eight hours and were found to carry naturally occurring variants of CovS; BSAC_bs1388 (36aa-STOP) and CovR; BSAC_bs1802 (Gly91Asp), respectively. BSAC_bs468 (Thr241Pro and Ala379Thr), did not exhibit the same increase.

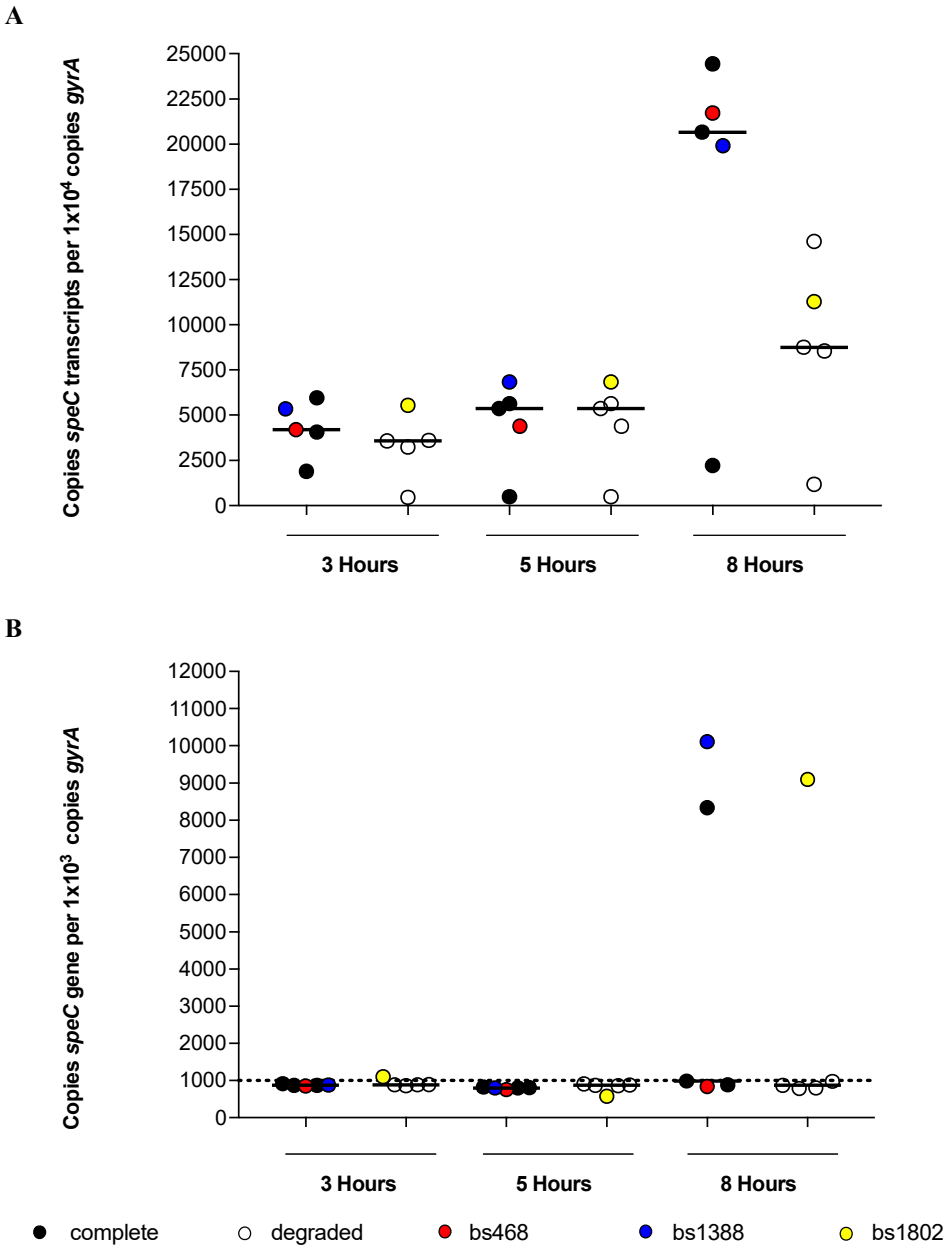

Figure 4.7.: Expression of *speC* and gene copies during growth in THB

A) Expression of *speC* in THB was similar for all *emm*4 isolates at three (p=0.3095, Mann-Whitney) and five hours (p=>0.9999, Mann-Whitney), but higher at eight hours, though there was a difference in expression between complete (black circles) and degraded isolates (white circles), with $M4_{complete}$ expressing *speC* more highly (p=0.0952, Mann-Whitney). B) DNA extracted from the same samples was probed for copies of the *speC* gene by qPCR. An increase in *speC* gene copy numbers was not observed in either $M4_{complete}$ (black circles) or $M4_{degraded}$ (white circles) at eight hours (p=0.3095, Mann-Whitney). Variant *covS* isolates BSAC_bs468 and BSAC_bs1388 are highlighted as a red circles and blue circles, respectively. *covR* variant BSAC_bs1802 is indicated as yellow circles. Data are the mean of three independent RNA/DNA samples, measured in triplicate and normalised and expressed as copies of *speC* transcript or gene copy per 10,000 copies of housekeeping gene *gyrA* transcripts or gene copy. Horizontal bars indicate the median. Dotted line represents the point at which copies of *speC* and *gyrA* are approximately equivalent.

Expression of *spd3*, carried by Φ10750.2/Φ427.2, was similarly expressed at comparable levels between $M4_{complete}$ and $M4_{complete}$ isolates at three (p=>0.9999, Mann Whitney), five (p=>0.9999, Mann Whitney), and eight hours (p=>0.9999). A single $M4_{complete}$ isolate expressed *spd3* more highly than the other isolates at 8 hours, BSAC_bs1349 (Figure 4.8).

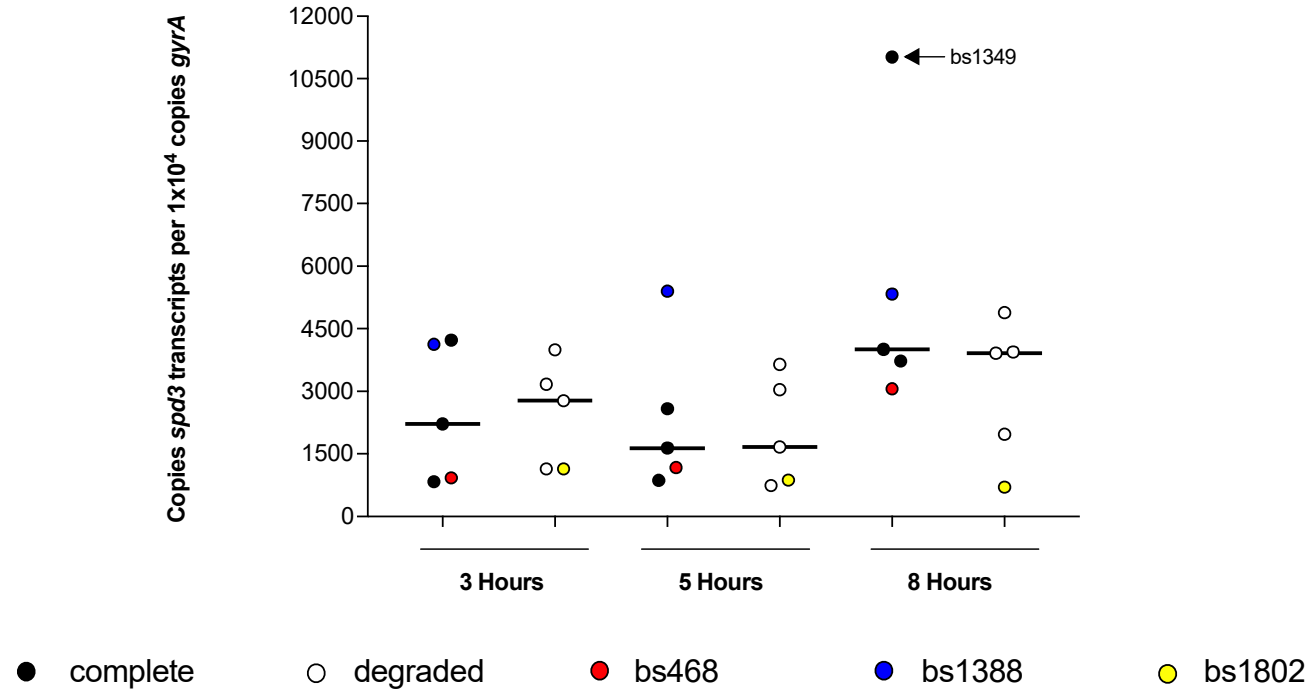

Figure 4.8: Expression of *spd3* during growth in THB

Expression of the DNase *spd3* was comparable between $M4_{complete}$ (black circles) and $M4_{degraded}$ isolates (white circles). There was no major difference between expression of *spd3* in these two groups at three (p=>0.9999, Mann-Whitney), five (p=>0.9999, Mann-Whitney), or eight hours (p=>0.9999, Mann-Whitney). Variant *covS* isolates BSAC_bs468 and BSAC_bs1388 are indicated as a red and a blue circle, respectively. Variant *covR* isolate BSAC_bs1802 is represented by a yellow circle. Isolate BSAC_bs139 is highlighted with an arrow. Data are the mean of three independent RNA samples measured in triplicate, and normalised to 10,000 copies of *gyrA* transcripts. Horizontal bars indicate the median.

Expression of *ssa* also showed similar patterns of transcriptions throughout growth with a subtle elevation observed for all isolates at five hours (

Figure 4.9), although expression of *ssa* between both $M4_{complete}$ and $M4_{degraded}$ were comparable at three (p=0.8413, Mann Whitney), five (p=>0.9999, Mann Whitney) and eight hours (p=0.0952, Mann Whitney). At eight hours, a single $M4_{complete}$ isolate and a single $M4_{degraded}$ isolate expressed *ssa* more highly than the other BSAC *emm*4, BSAC_bs1349 and BSAC_bs192, respectively. Isolates with variant CovRS alleles did not appear to express *ssa* more highly than other *emm*4 isolates.

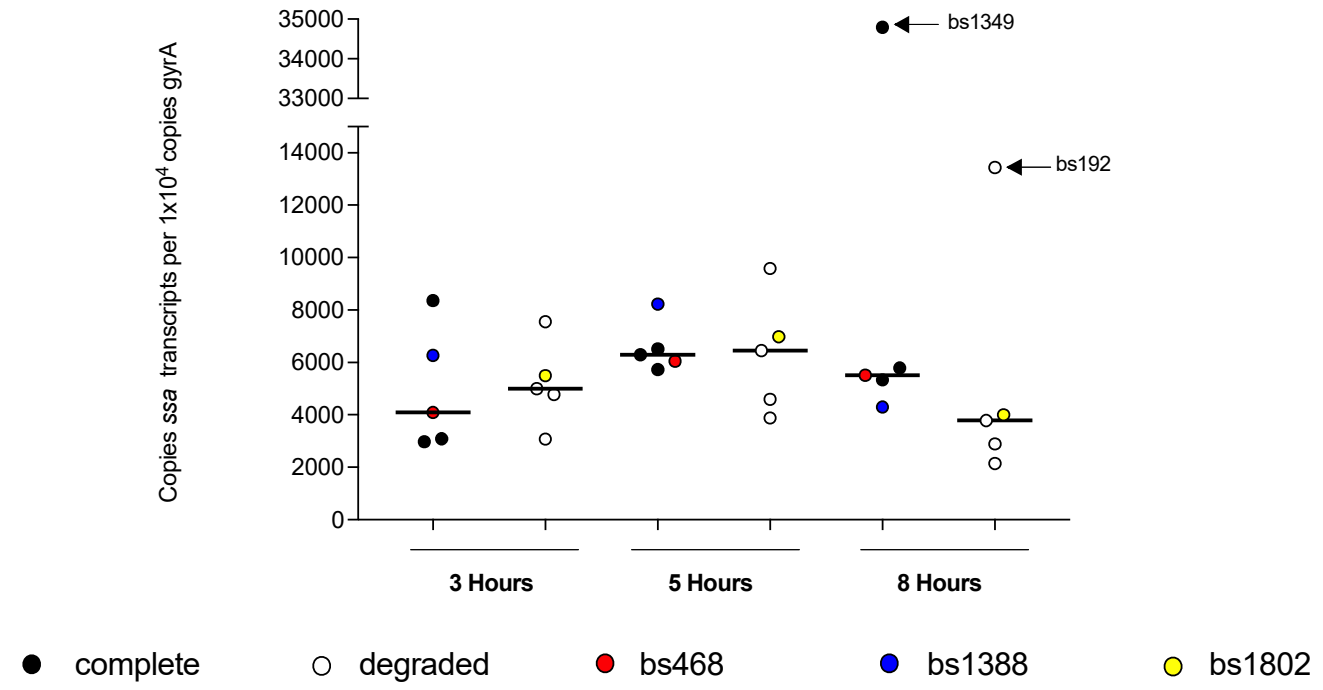

Figure 4.9.: Expression of *ssa* during growth in THB

The expression of *ssa* in THB was similar for $M4_{complete}$ (black circles) and $M4_{degraded}$ (white circles) at three (p=0.8413, Mann-Whitney), five p=>0.9999, Mann-Whitney) and eight hours (p=0.0952, Mann-Whitney). Variant *covS* isolates BSAC_bs468 and BSAC_bs1388 are indicated with a red and a blue circle, respectively. Variant *covR* isolate BSAC_bs1802 represented with a yellow circle. BSAC_bs1394 and BSAC_bs192 expressed *ssa* more highly than did the other isolates at eight hours, and are highlighted with arrows. Data are the mean of three independent RNA samples measured in triplicate and normalised per 10,000 copies of *gyrA* transcripts. Horizontal bars indicate the median.

4.4.4 SpyCIM4 of $M4_{complete}$ and $M4_{degraded}$ are not inducible

SpyCI of $M4_{complete}$ are full-length, whereas $M4_{degraded}$ were observed to have undergone gene loss, in a fashion similar to that described here between lysogenic prophage. An analogous integration and excision PCR was designed and applied to SpyCIM4, the mobility of which was believed to regulate the expression of the DNA MMR operon in *S. pyogenes* (Nguyen and McShan, 2014; Scott *et al.*, 2012). The applicability of this regulatory mechanism between SpyCI and expression of the DNA MMR operon was investigated by MSc student Samuel Haywood, co-supervised by the author and Dr. Claire E. Turner. Unexpectedly, neither the SpyCIM4 of $M4_{complete}$ or $M4_{degraded}$ were inducible with mitomycin C (Figure 4.10A), nor were they found to be mobile during exponential growth (Figure 4.10B), despite the $M4_{complete}$ SpyCIM4 not having undergone degradation.

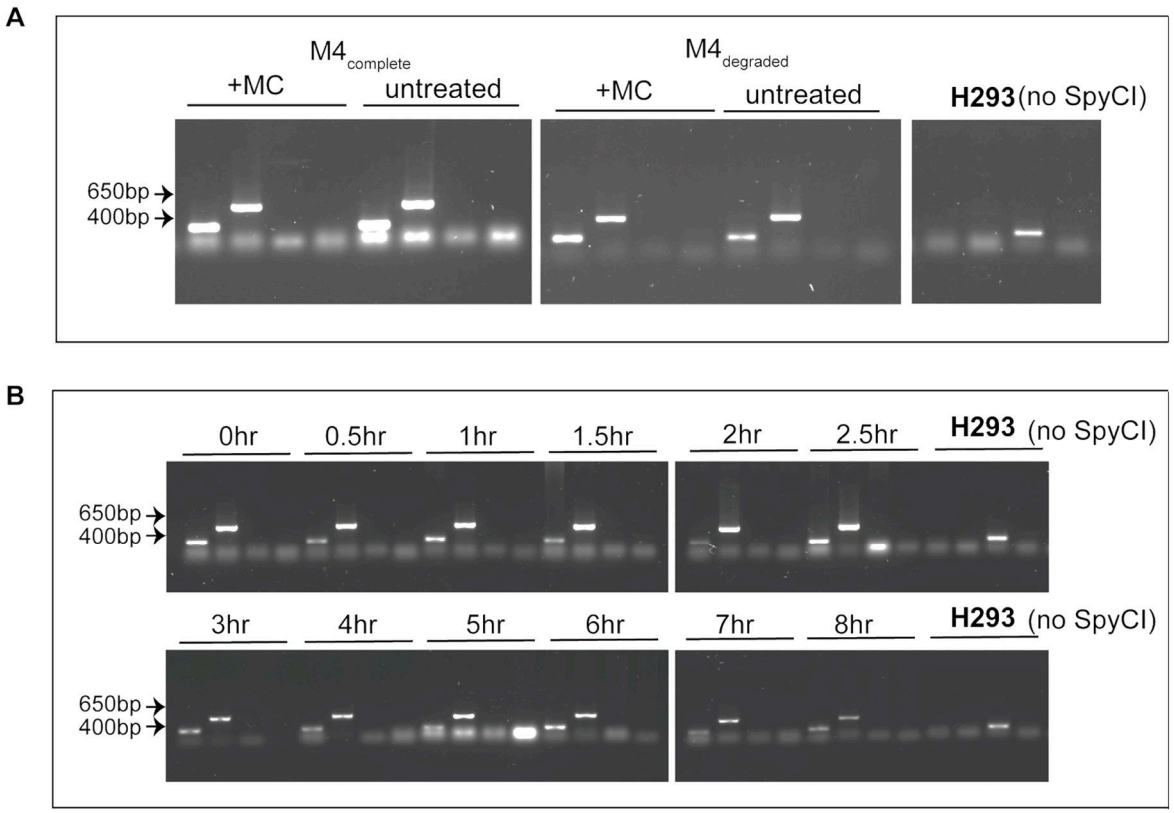

Figure 4.10: SpyCI^{M4} induction with mitomycin C and during growth in THB

A) M4complete isolate BSAC_bs1349 and M4degraded isolate BSAC_bs696 were treated with mitomycin C to induce SpyCI. SpyCI could not be detected in its excised conformation for $M4_{complete}$, nor $M4_{degraded}$. B) Both isolates were grown for 8 hours, and at no timepoint was SpyCI detected as an excised element in either $M4_{complete}$ or $M4_{degraded}$. H293 has no SpyCI was used as an internal control, with amplification only occurring to confirm the absence of SpyCI at the attachment site. DNA was standardised to 40 ng/μL prior to PCR. 650bp and 400bp arrows indicate position relative to 1kb plus DNA ladder (Invitrogen).

SpyCIM4 of both $M4_{complete}$ and $M4_{degraded}$ were unable to excise, which was not expected. It was thus hypothesised that the permanent integration of SpyCIM4, would, as described in published reports, block transcription of *mutL*, and the rest of the DNA mismatch repair (MMR) operon (Nguyen and McShan, 2014; Scott *et al.*, 2012). Transcripts for both *mutL* and *mutS* at 3-hours growth were therefore measured semi-quantitively for all BSAC *emm*4. Transcripts for both genes were detected in all isolates (Figure 4.11), suggesting that in *emm*4, SpyCIM4 may not regulate this operon by a dynamic process of integration and excision.

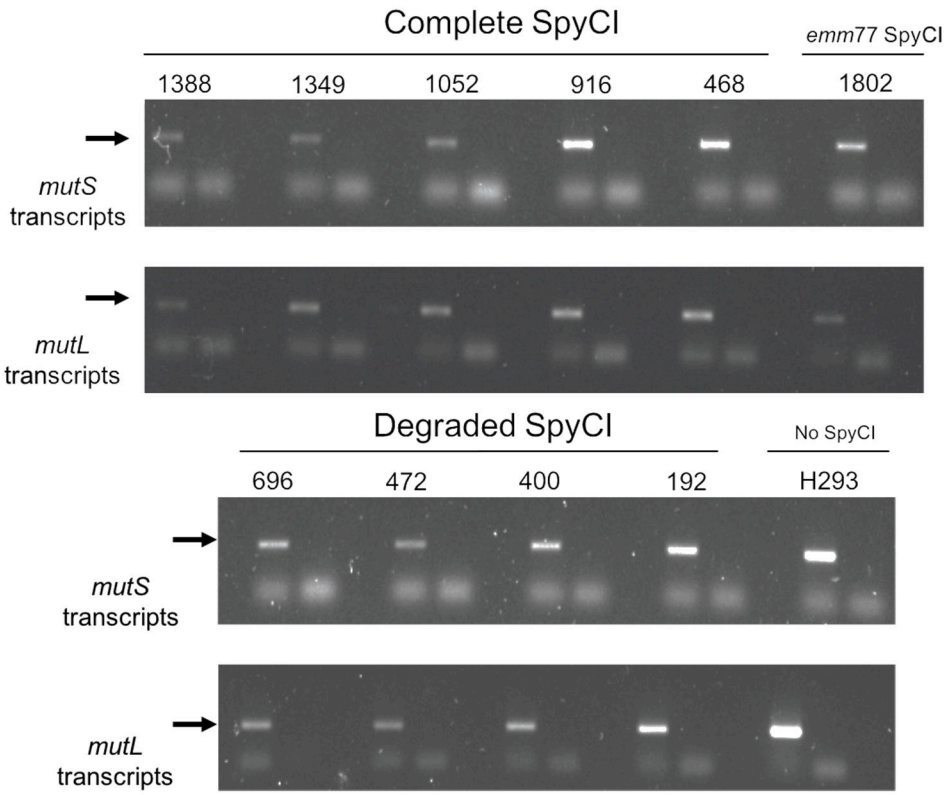

Figure 4.11: *mutS* and *mutL* expression during bacterial growth

RNA was extracted from all BSAC *emm*4 after 3 hours growth, DNase treated, and converted to cDNA by reverse transcription. cDNA was probed semi-quantitatively for *mutS* and *mutL* transcripts by PCR. For all isolates, *mutS* and *mutL* transcripts were detected. For each isolate, an RT-negative was prepared in parallel (lacking reverse transcriptase and RNaseIn) to serve as a control for contaminating DNA, and loaded after each corresponding sample and neither cDNA nor genomic DNA were detected. H293, which has no SpyCI, was included as an internal control. No transcripts were detected in the MGW negative control. Arrows indicate ~400bp on 1kb plus DNA ladder (Invitrogen).

4.5 Discussion

Although the constellation of canonical lysogenic prophage and SpyCIM4 superficially appear to be the same between *emm*4 isolates, the gene content of these prophage is variable between $M4_{complete}$ and $M4_{degraded}$ *emm*4 isolates. Prophages could only be induced to excise from the bacterial chromosome among $M4_{complete}$ isolates. Φ427.1 was found to be stably integrated under the same conditions in which Φ10750.1 was inducible, in spite of the former retaining a functional integrase. This might suggest that integrase retention alone is not sufficient to abolish excision potential. It is possible that other gene losses in this prophage are responsible for this observation, particularly regulatory and replicative prophage modules. Φ427.2 could not be induced to excise experimentally, nor was it found to be induced during growth in THB at any stage. Conversely, Φ10750.2, was induced under identical conditions. In addition to having lost regulatory and replicative modules, Φ427.2 was predicted to lack a functional integrase, any and all of which might be expected to impact excision potential. Φ427.3 was similarly

decayed, having lost identifiable replicative and regulatory modules, in addition to harbouring an integrase that was predicted to be a defective. This element could not be induced in $M4_{degraded}$ isolates, but the corresponding Φ10750.3, was inducible with mitomycin C and during exponential growth. Spontaneous induction of prophage did not seem to correlate convincingly with specific timepoint as might be expected if spontaneous prophage induction is due to the accumulation of DNA damage over successive generations (Cortes *et al.*, 2019), though this may occur at later stages of growth not sampled or perhaps dynamically within a comparatively smaller subset of the larger bacterial population.

It was interesting to observe that levels of transcription of *spd3* and *ssa* attained by *emm*4 during growth in THB were generally comparable between both lineages. CovRS represses *spd3* in *emm*1 (Sumby *et al.*, 2006) and repression by CovRS is known to be maximal at early- and mid-log, but relieved at later stages of growth (Graham *et al.*, 2005). Despite this, CovRS variant *emm*4 isolates were not made obvious by their expression of *spd3.* There is also, however, evidence that *spd3* may be regulated by Rgg1/RopB (Anbalagan and Chaussee, 2013). Although the veracity of the *emm*4 CovRS variants was not confirmed, regulatory variation between genotypes may also have a part to play in these findings. Transcription of *speC* was found to be greater in $M4_{complete}$ than $M4_{degraded}$ isolates at eight hours, an increase of ~4-fold observed in the former, versus ~2-fold in the latter, from the previous timepoint of five hours. The differential increase in the expression of *speC* between the two types of *emm*4 isolates might translate to differential SpeC expression.

A thought-provoking observation was that although in $M4_{complete}$ isolates, Φ10750.1 was detected by PCR as having excised spontaneously during growth, an increase in *speC* gene copies was not generally observed. This suggests that while induction may occur, it is not necessarily followed by extra-chromosomal replication of the element, demonstrating that prophage induction, but not replication and induction of the phage lytic cycle may be required for the enhancement of *speC* expression. A similar phenomenon has been observed in enterohaemorrhagic *E. coli* 0157, wherein expression of shiga toxin (Stx1 and/or Stx2) and the elicitation of the lethal effect in a murine model of shigellosis, requires only prophage induction, not necessarily replication and the assembly of phage particles (Balasubramanian *et al.*, 2019). CovS variant BSAC_bs1388 (36aa-STOP), and CovR variant BSAC_bs1802 (Gly91Asp) exhibited a marked increase in *speC* gene copies, suggesting the *speC*/*spd1* prophage of these isolates may have been undergoing replication, though the association with CovRS may be merely coincidental, and it must be noted that in BSAC_bs1802, this prophage is unlike any other in the *emm*4 population both in our collection, nor indeed the international population (Remmington *et al.*, 2020). The prophage conformation PCR detected excision in these two isolates at eight hours, but not as strongly as might be expected.

The expression of *spd3* was similar between $M4_{complete}$ and $M4_{degraded}$ isolates, and did not appear to be growth-phase dependent, despite the capacity for Φ10750.2 to induce, whilst Φ427.2 could not. Expression of *ssa* was

also very similar between the *emm*4 isolates, regardless of whether or not the prophage was degraded, and despite Φ10750.3 being inducible, while Φ427.3 was not. It was curious that neither *spd3* nor *ssa* appeared to be growth-phase dependent, as are many other virulence genes in *S. pyogenes*, including those which are encoded by prophage (Chaussee *et al.*, 2008; Turner *et al.*, 2009; Unnikrishnan *et al.*, 1999). This suggests that perhaps, to attain enhanced expression of these genes, in *emm*4, excision of one or more of these prophage may be unnecessary, as appears to be the case for *speA* in *emm*3, albeit under different conditions and against a very different genetic background (Banks *et al.*, 2003).

Inasmuch as these observations held true for the majority of the *emm*4 prophage, one isolate was discordant; BSAC_bs1802. Despite clustering with MEW427 by core SNPs, this isolate was dissimilar to the other BSAC isolates belonging to the $M4_{degraded}$ lineage, and indeed, the international *emm*4 population at large, in a number of ways. Φ1802.1, (the *speC*/*spd1*-associated prophage of this isolate) is most similar (~99%) to prophage present in *emm*28 MGAS29284 (Kachroo *et al.*, 2019), and could be induced and notably, a subtle increase in intensity in induction of Φ1802.1 was observed at 8-hours. Gene copies of *speC* also were markedly higher at this timepoint, which was not observed for the majority of the other *emm*4 isolates. Intriguingly, this did not appear to evoke an appreciable enhancement in *speC* transcription for this isolate. Φ427.2 of BSAC_bs1802 is identical to that of the other $M4_{degraded}$ BSAC isolates, and similarly, could not be induced to excise, and expression of *spd3* was unremarkable. Φ10750.3 of BSAC_bs1802 was readily inducible, though again, expression of *ssa* by this isolate was unremarkable.

Published reports that describe the growth-phase dependent control of DNA MMR operon gene expression by SpyCI integration and excision (Nguyen and McShan, 2014; Scott *et al.*, 2012) were, unexpectedly, not supported. It was interesting to note that in both $M4_{complete}$ and $M4_{degraded}$ with complete and degraded SpyCI, respectively, the element did not was not induced by mitomycin C, nor was SpyCI induced during bacterial growth at any timepoint tested. Further, transcripts for both *mutS* and *mutL* were readily detected in all BSAC *emm*4 at 3 hours growth, when excision of the element and expression of the operon were purportedly minimal according to existing reports (Nguyen and McShan, 2014; Scott *et al.*, 2012). These data thus suggest that expression of the DNA MMR operon may not be regulated in this way in genotype *emm*4.

Taken together, the data in this chapter clarify that for the large part, $M4_{degraded}$ isolates, which exhibit marked gene loss in prophage-encoding regions, have been rendered immobile; are not induced during bacterial growth in THB; nor can they be induced experimentally with mitomycin C. It was hypothesised that this may be due to the loss of functional prophage modules dedicated to regulatory and replicative modules, and owing in part to loss of function mutations to prophage-encoded integrase genes in Φ427.2 and Φ427.3. Yet, with the possible exception of *speC*, this did not appear to evoke differential expression of *spd3* and *ssa* between isolates of the two lineages. These data might not necessarily reflect expression in more realistic and infection-relevant systems, for

example, expression and induction might be appreciably different on co-culture with human cells, or using *ex-vivo* or *in-vivo* models (Banks *et al.*, 2003; Broudy *et al.*, 2002, 2001; Sitkiewicz *et al.*, 2010; Virtaneva *et al.*, 2005; Zhu *et al.*, 2020, 2017). The possible phenotypic consequences of these observations in more infection-relevant conditions were therefore explored in subsequent chapters.

4.5.1 Chapter 4: Key findings

- Degraded prophage present in the genomes of $M4_{degraded}$ isolates could not be induced experimentally, nor were they induced throughout bacterial growth
- Conversely, complete or full-length prophage of $M4_{complete}$ isolates were inducible experimentally and were weakly induced during growth
- $M4_{complete}$ isolates expressed more *speC* in THB at eight hours than did $M4_{degraded}$ isolates
- Increased expression of *speC* at eight hours by $M4_{complete}$ was not associated with an increase in *speC* copy number
- Expression of *spd3* and *ssa* was comparable between $M4_{complete}$ and $M4_{degraded}$ throughout bacterial growth and did not appear to be dependent on growth-phase
- $SpyCI^{M4}$ could not be induced in $M4_{complete}$ or $M4_{degraded}$ and did not appear to regulate the expression of DNA-mismatch repair genes, as previously reported

Chapter 5

5. Phenotypic characterisation of *emm*4 isolates

5.1 Introduction

Emergent *S. pyogenes* lineages are often characterised by marked phenotypic modulations relative to their ancestral clones. Frequently, such lineages are associated with the loss and/or acquisition, or differential regulation of bacterial virulence factors encoded by mobile genetic elements (Afshar *et al.*, 2017; Al-Shahib *et al.*, 2016; Kachroo *et al.*, 2019; Lynskey *et al.*, 2019a; Nasser *et al.*, 2014; Turner *et al.*, 2019b, 2015; Zhu *et al.*, 2015b, 2015a). There have been a number of broad phenotypic *S. pyogenes* studies incorporating singular or small numbers of *emm*4 isolates (Flores *et al.*, 2012; Galloway-Peña *et al.*, 2018; Henningham *et al.*, 2014; Loh *et al.*, 2013; Menschner *et al.*, 2020), and studies to characterise specific molecular mechanisms within the genotype (Galloway-Peña *et al.*, 2018; Vega *et al.*, 2020). However, genotype *emm*4 *S. pyogenes*, in spite of their disease burden and global prevalence (Efstratiou and Lamagni, 2016), have not been subject to robust phenotypic characterisation hitherto. This is certainly the case when one reflects on the many largescale studies that have greatly enhanced our understanding of other major *emm*-types, and associated lineages thereof (Walker *et al.*, 2007; Tse *et al.*, 2012; Venturini *et al.*, 2013; Turner *et al.*, 2015; Walker, 2015; Afshar *et al.*, 2017; Kachroo *et al.*, 2019).

Work in Chapter 3 determined that genotype *emm*4 isolates in the BSAC collection were associated either with the $M4_{complete}$ lineage, or the $M4_{degraded}$ lineage. Further, it was demonstrated that those isolates associated with the $M4_{degraded}$ lineage, were incapable of prophage induction, while the prophage of $M4_{complete}$ could be induced to excise. With the exception of *speC* transcription, which was elevated at 8-hours growth, and higher in $M4_{complete}$, than $M4_{degraded}$, growth-phase dependent expression of prophage toxin genes *spd3* and *ssa* was not observed, and transcription of *spd3* and *ssa* were similar between isolates of both lineages. Analysis of pangenome data and regulator sequence analyses also revealed a number of additional lineage- and isolate-specific SNPs that could conceivably contribute to a differential phenotype that may be associated with either groups of isolates, such as the fusion of *emm* and *enn* among $M4_{degraded}$ isolates, or naturally occurring CovRS variants, as described in Chapter 3.

5.2 Chapter hypotheses and aims

With a clear phylogenetic distinction between $M4_{complete}$ and $M4_{degraded}$ isolates, and the demonstrable for $M4_{complete}$ prophage to induce, while $M4_{degraded}$ could not, it was hypothesised that these differences might evoke a polarising phenotype between isolates belonging to either lineage. The phenotypic consequences of additional lineage- and

isolate-specific SNPs, beyond those pertaining to prophage gene content, were also considered. As such, the work described in this chapter sought to characterise these isolates using a number of well-described models and systems that have been applied to characterise *S. pyogenes*, and other pathogenic streptococci.

5.3 Methods

5.3.1 Lancefield assay

Overnight *S. pyogenes* cultures were diluted 1:5 in fresh THB (Oxoid) and grown to an OD_{600} of 0.3. Cultures were briefly vortexed, and diluted 1:10,000 in sterile PBS (Gibco) to achieve inocula of approximately 50-100 CFU/mL and 30 µL aliquots were plated on THA to determine CFU/mL for confirmation. Blood was freshly drawn by venepuncture from two healthy human donors, and 170µL of heparinised blood was transferred to 1.5mL tubes and inoculated with 30 µL *S. pyogenes*/PBS suspension, in triplicate, per isolate. Samples were placed in a rotary for gentle mixing, and incubated at 37°C for 3 hours. At the end of the incubation, 10µL was plated in triplicate from each of the three tubes, for all *emm*4 isolates and used to inoculate THA to determine CFU/mL. A multiplication factor was calculated by dividing the mean final CFU/mL per tube, by the CFU/mL of the corresponding inocula (Lancefield, 1959).

5.3.2 Superantigenicity assay

The superantigenic activity of $M4_{complete}$ and $M4_{degraded}$ *emm*4 isolates was measured by BrdU incorporation ELISA kit (Roche) per manufacturers guidelines (methods section 2.1.8). Overnight cultures for all BSAC *emm*4 isolates were grown in RPMI +10% (v/v) FCS (Gibco), incubated overnight at 37°C in an atmosphere enriched with 5% (v/v) CO_2. Genotype *emm*1 strain H305 was included as a well characterised, highly superantigenic internal control (Sriskandan *et al.*, 1999; Turner *et al.*, 2012; Unnikrishnan *et al.*, 2002). Overnight cultures were centrifuged at 10,000 x *g* for 15 minutes and supernatants filter sterilised through 0.2 µ filters (Sartorius, AG, VWR).

Whole blood was drawn from healthy donors and diluted 1:1 with pre-warmed 37°C PBS (Gibco). Ficoll-Paque (Amersham Pharmacia Biotech) was overlaid with an equal volume of blood:PBS and centrifuged at 800 x *g* for 35 minutes at room temperature. The buffy coat was carefully aspirated with a sterile Pasteur pipette and washed once with pre-warmed 37°C Hanks' Balanced Salt Solution (Gibco) at 800 x *g* for 10 minutes at room temperature. The resultant human PBMC pellet was resuspended in pre-warmed RPMI (+10% (v/v) FCS) with 50 U/mL penicillin and 50µg/mL of streptomycin (Sigma) and seeded in 96-well plates (Corning) at $2x10^5$ cells per well, in a volume of 200 µL. 20 µL of sterilised culture supernatant was added to triplicate wells and incubated for 120 hours incubation in a humidified environment supplemented with 5% (v/v) CO_2 at 37°C. A negative control of

fresh RPMI +10% (v/v) FCS was included as a negative control; absorbance values of which were deducted from sample absorbances, which were then expressed as the mean of technical triplicates.

5.3.3 *Galleria mellonella* virulence assay

Overnight cultures of *emm*4 isolates were washed twice in pre-warmed PBS to remove secreted toxins and waste products. Groups of 10 wax moth larvae (*Galleria mellonella*) (LiveFoods Direct Ltd, Sheffield, UK) were infected with 10^5-10^6 CFU per larvae, with one of each of the ten BSAC *emm*4 *S. pyogenes* by injection with an ultra-fine insulin needle (BD) at the base of the left hind proleg, as standardised previously (Afshar *et al.*, 2017; Loh *et al.*, 2013; Olsen *et al.*, 2011). Larvae were then incubated in petri dishes lined with filter paper at 37°C, in an atmosphere supplemented with 5% (v/v) CO_2. Larvae were scored for survival every 24 hours for 7 days. An uninfected PBS control group was included per experiment.

5.3.4 Gram's crystal violet adhesion assay

Bacterial adhesion was assayed using an established method, with modification (O'Toole, 2011), which has been applied to *S. pyogenes* (Turner *et al.*, 2015). Overnight cultures of BSAC *emm*4 isolates and *emm*6 control strain BSAC_bs2438 were diluted 1:10 in fresh THB, and 150 µL was pipetted into triplicate wells, per isolate. Sterile THB was used as a negative control, triplicate averages of which was deducted from triplicate averages of test wells and the internal control. Plates were incubated for 24 or 48 hours, as indicated, at 37°C in an atmosphere supplemented with 5% (v/v) CO_2. Isolates were assayed for adhesion to ECM components by pre-treating 96-well polystyrene cell culture plates (Corning) with 50µg/mL of either fibronectin (Sigma), fibrinogen (Sigma), and type IV collagen (Sigma), overnight at 4°C. Treated plates were washed once with sterile PBS prior to inoculation and incubation, as above. Following incubation, supernatants and non-adherent bacteria were discarded, and wells washed twice with water, before staining with 120 µL 0.1% (w/v) Gram's Crystal Violet (Sigma) at room temperature for 10 minutes. Excess stain was discarded, and plates washed twice in water before being blotted onto paper towels. Gram's stain was solubilised by the addition of 120 µL 100% (v/v) ethanol (Sigma), and incubated at room temperature for 10 minutes. Absorbance was measured at OD_{600} by spectrophotometry, samples were measured in triplicate. Where values are expressed as % difference of the internal control, this was achieved by dividing the sample mean OD_{600} by the positive control mean OD_{600}, and multiplying by 100 to achieve a percentage.

5.4 Results

5.4.1 Growth in THB

To determine whether the two *emm*4 lineages exhibited a difference in growth characteristics, BSAC *emm*4 isolates were grown in THB (Figure 5.1). Isolates were cultured exponentially and OD_{600} determined by

spectrophotometry each hour for eight hours. Aliquots were taken at 3, 5 and 8 hours (determined to be early-, mid- and late-log cultures, respectively) to determine CFU/mL. No significant difference was observed between $M4_{complete}$ and $M4_{degraded}$ isolates in liquid culture (p=0.6905, Mann-Whitney) or by plating for viable counts (p=>0.9999, Kruskal-Wallis) suggesting similar growth patterns across all isolates (Figure 5.1).

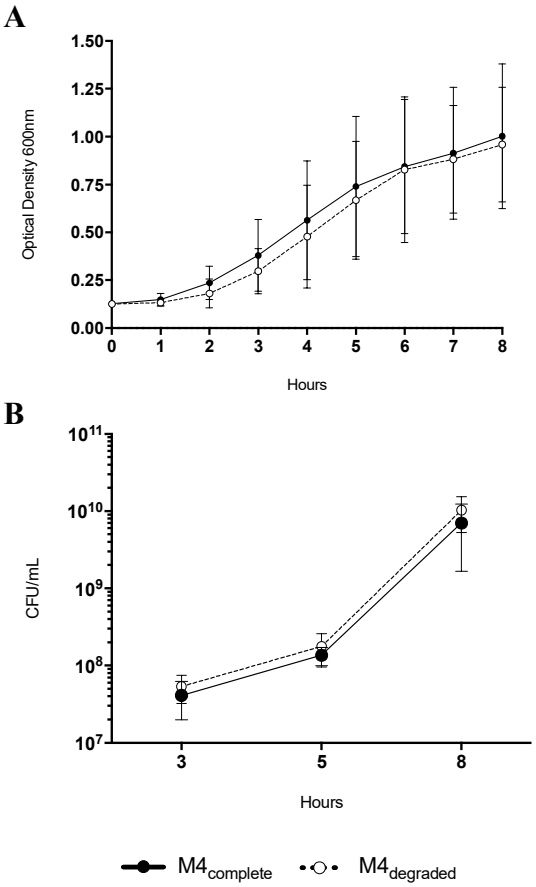

Figure 5.1. Growth characteristics of *emm*4 isolates

A) Exponential growth in THB was not significantly different between $M4_{complete}$ (solid black line, black circles) and $M4_{degraded}$ isolates (dotted black line, white circles) (p=0.6905, Mann-Whitney on Area under curve values for complete and degraded isolates). Data represent the mean (±standard deviation) from three biological repeats of five $M4_{complete}$ isolates, and five $M4_{degraded}$ isolates. B) CFU/mL similarly determined that there was no significant difference in CFU/mL in liquid culture between $M4_{complete}$ and $M4_{degraded}$ (p=>0.9999, Kruskal-Wallis). Data represent the mean of five $M4_{complete}$ isolates (solid black line, black circles), and five $M4_{degraded}$ isolates (dotted black line, white circles).

5.4.2 **Lancefield assay: survival of *emm*4 in whole human blood**

Although growth in bacteriological media was similar between isolates associated with either lineage, it was important to ascertain whether these observations might be expanded to more physiologically relevant conditions.

With this in mind, isolates were grown in whole human blood, a well described phenotypic assay to measure *S. pyogenes* replication *ex-vivo* (Lancefield, 1959; Reglinski *et al.*, 2016; Todd, 1927). The isolates from either *emm*4 lineage survived comparably well in whole human blood on the first (Figure 5.2A) and second occasions (Figure 5.2B). In each case one isolate grew slightly better than the other *emm*4, on both occasions this was an $M4_{degraded}$ isolate, though not the same isolate; BSCA_bs696 in the first experiment, and BSAC_bs400 in the second experiment.

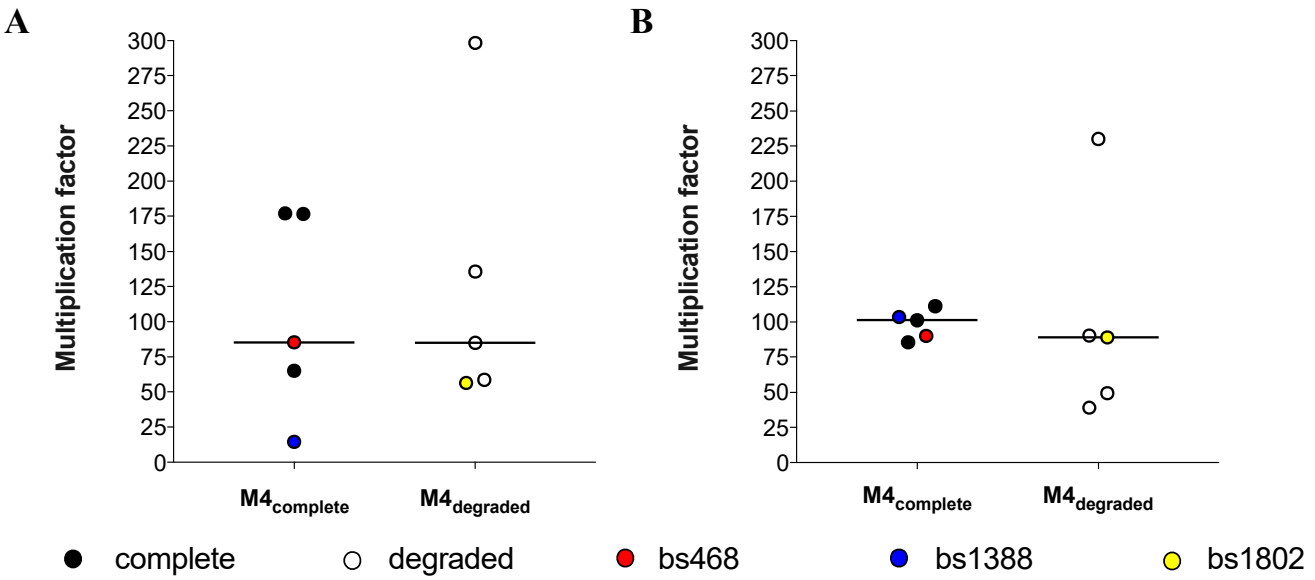

Figure 5.2. Survival of *emm*4 isolates in whole human blood.

A) There was no significant difference in the survival $M4_{complete}$ (black circles) and $M4_{degraded}$ isolates (white circles) (p=>0.9999, Mann-Whitney). Strains were grown in triplicate using a single healthy donor and multiplication factor was calculated by dividing the mean final CFU/mL per isolate, by the CFU/mL of the corresponding inoculum after three hours incubation. B) The experiment was repeated with a second healthy donor, and yielded similar results; again, no significant difference was observed between $M4_{complete}$ (black circles) and $M4_{degraded}$ isolates (white circles) (p=>0.4206, Mann-Whitney). Variant *covS* isolates BSAC_bs468 and BSAC_bs1388 are indicated with a red and a blue circle, respectively. Variant *covR* isolate BSAC_bs1802 is indicated with a yellow circle. Horizontal bars indicate the median.

5.4.3 Superantigenicity of *emm*4 culture supernatants

Culture supernatant filtrates harvested from fresh overnight cultures of the BSAC *emm*4 isolates grown in RPMI +10% (v/v) FCS were assayed for their mitogenicity towards human lymphocytes. The supernatants collected from $M4_{complete}$ isolates were comparable in superantigenic activity to those collected from $M4_{degraded}$ isolates, and were not significantly different (p=0.4206, Mann-Whitney) (Figure 5.3).

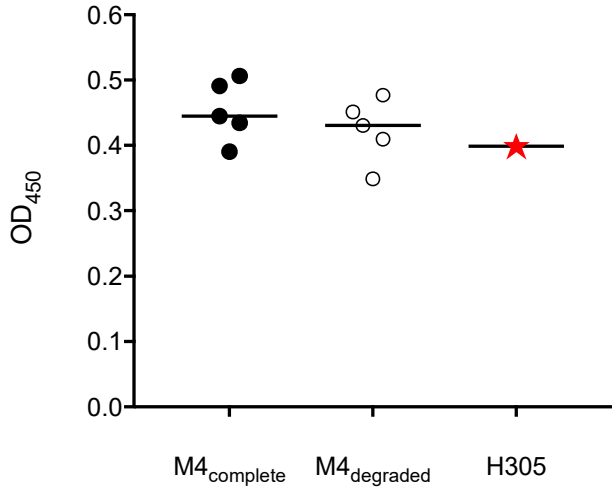

Figure 5.3. Superantigenic activity of *emm*4 isolates

Superantigenic activity of overnight culture supernatants from *emm*4 isolates towards human peripheral blood mononuclear cells. There was no significant difference in superantigenic activity in overnight culture supernatant filtrates from $M4_{complete}$ (black circles) and $M4_{degraded}$ (white circles) isolates (p=0.4206, Mann-Whitney), *emm*1 strain H305 (red star) was included as a well-characterised positive control. Isolates were grown overnight at 37°C with 5% (v/v) CO_2 in RPMI media, supplemented with 10% (v/v) FCS. Human lymphocytes were isolated by Ficoll-Paque gradient, and stimulated with fresh filtered culture supernatants. Mitogenicity of human T-lymphocytes was measured by ELISA to detect BrdU incorporation, and expressed as OD_{450}. Data represent the mean OD_{450}, minus the mean OD_{450} of the sterile media negative control. Lymphocytes were harvested from one donor and samples were measured in triplicate. Horizontal bars indicate the median.

5.4.4 *Galleria mellonella* virulence assay with *emm*4 isolates

To assay for general differences in virulence between $M4_{complete}$ and $M4_{degraded}$, the *Galleria mellonella* infection model was used. This model is well described for assaying general virulence in *S. pyogenes* (Afshar *et al.*, 2017; Chalmers *et al.*, 2017; Loh *et al.*, 2013; Olsen *et al.*, 2011). Larvae were infected at 10^5-10^6 CFU per larvae, and unexpectedly, $M4_{complete}$ isolates were found to be significantly more virulent than $M4_{degraded}$ isolates (p=0.0016, Mantel-Cox log rank test) (Figure 5.4A). The experiment was repeated in its entirety with a second batch of larvae, using a similar dose, and again, significantly enhanced virulence was observed in $M4_{complete}$ isolates relative to $M4_{degraded}$ isolates (p=0.0014, Mantel-Cox log rank test; Figure 5.4B). When the experiment was repeated a third time using a lower dose (10^4-10^5 CFU per larvae), the findings were recapitulated, again demonstrating the enhanced virulence of $M4_{complete}$ isolates (p=0.0012, Mantel-Cox log rank test; Figure 5.4C).

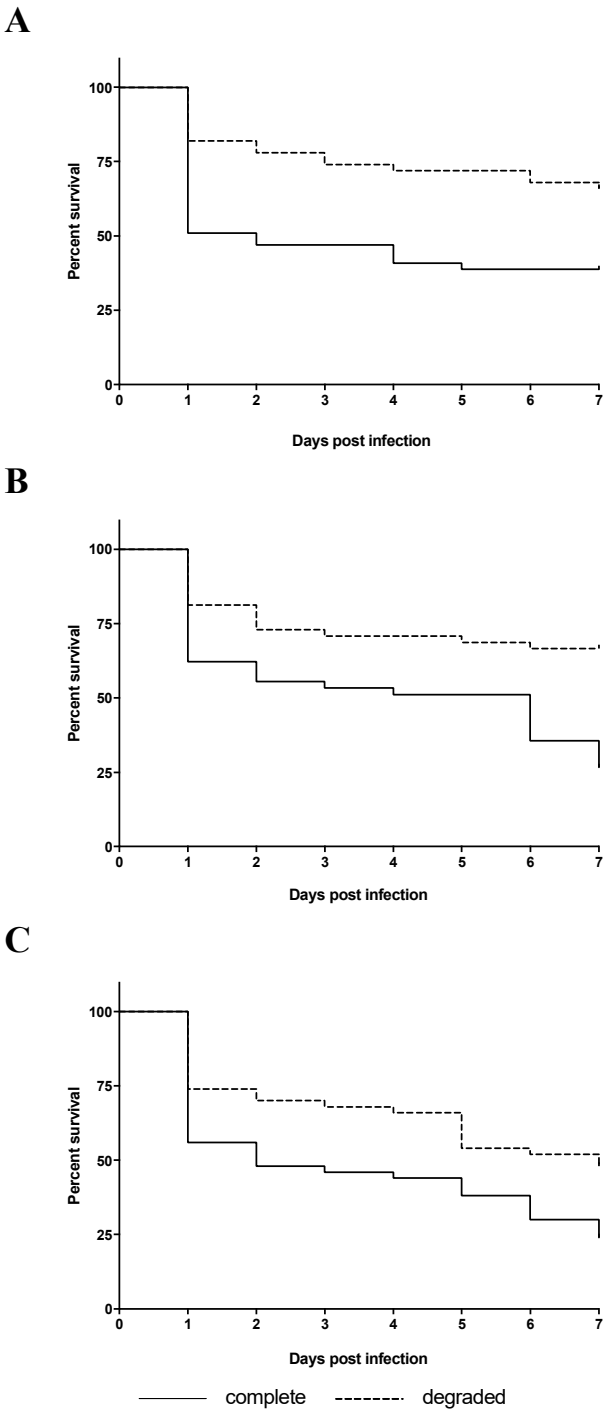

Figure 5.4: *Galleria mellonella* virulence assay

A) $M4_{complete}$ isolates (solid black line) are significantly more virulent than $M4_{degraded}$ (dotted black line) in a wax moth larvae model of invasive disease (p=0.0016, Mantel-Cox Log-rank test). 50 larvae per group (ten larvae per isolate) were infected with 10^5-10^6 CFU per larvae, incubated at 37°C, and scored for survival every 24 hours for seven days post infection. B) Repeating the experiment, a second time on a second batch of larvae, yielded similar results (p=0.0014, Mantel-Cox Log-rank test). C) The entire experiment was repeated with a third batch of *G. mellonella* larvae but a lower dose of 10^4-10^5 CFU per larvae and yielded similar results, again demonstrating increased virulence associated with $M4_{complete}$ *emm*4 isolates, compared to $M4_{degraded}$ isolates (p=0.0012, Mantel-Cox Log-rank test).

5.4.5 **Biofilm formation and adhesion of *emm*4 isolates**

The capacity for biofilm formation and adhesion to a number of substrata were explored in a project co-supervised by the author, in collaboration with MSc student Jordan Finlay, using standard crystal violet adhesion assays in 96-well plates (Corning). Genotype *emm*4 *S. pyogenes* were found to be capable of forming robust biofilms after 24-hours. There were no major differences observed in the adhesion to polystyrene by isolates of either lineage, though $M4_{complete}$ isolates did slightly exceed the adhesion of $M4_{degraded}$ at both 24- and 48-hours. At both timepoints, a single $M4_{complete}$ isolate, BSAC_bs1349, demonstrated higher adhesion than did the other isolates, this was most obvious after 24-hours, though this difference narrowed after 48-hours, as did the difference between $M4_{complete}$ and $M4_{degraded}$ (Figure 5.5).

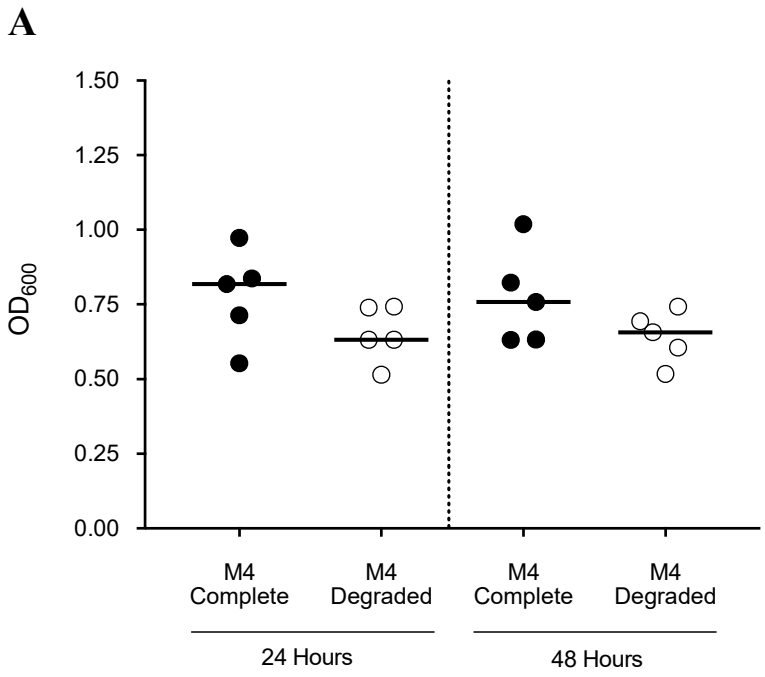

Figure 5.5: Adhesion of *emm*4 isolates to polystyrene

Adhesion of *emm*4 isolates to polystyrene 96-well plates when cultured in THB, as determined by Gram's crystal violet adhesion assay. At both 24- and 48-hours, $M4_{complete}$ isolates (black circles) adhered marginally better to polystyrene than did $M4_{degraded}$ isolates (white circles). This was not statistically significant at 24- ($p=0.2222$, Mann-Whitney) or 48-hours ($p=0.2222$, Mann-Whitney). The mean of triplicate negative control wells (sterile THB) was deducted from OD_{600} of test wells. Data represent the mean of four experimental repeats, measured in triplicate on each occasion. Horizontal bars indicate the data median.

S. pyogenes is exquisitely attuned to the human host, and expresses a litany of virulence factors and adhesins to attain colonisation, carriage and infection (Brouwer *et al.*, 2016). Thus, the adhesion of genotype *emm*4 isolates to physiologically important substrates, fibronectin, type IV collagen and fibrinogen, was tested. The crystal violet assay was applied to 96-well plates pre-coated with each substrate, and optical densities were expressed as

percentage change relative to internal positive control *emm*6 BSAC_bs2438, which was included on each 96-well plate. This isolate was shown in prior work in this laboratory to adhere robustly, consistently and reproducibly under a variety of conditions. After 24-hours incubation, *emm*4 isolates of both lineages adhered comparably to wells coated with fibronectin and type IV collagen, albeit being less adherent than the *emm*6. However, enhanced adhesion, (comparable to, or in excess of the control *emm*6 isolate) was observed when *emm*4 were cultured in plates pre-treated with fibrinogen. A single $M4_{degraded}$ isolate, BSAC_bs1802, adhered better than did the other *emm*4 isolates of both lineages in wells treated with fibrinogen (Figure 5.6).

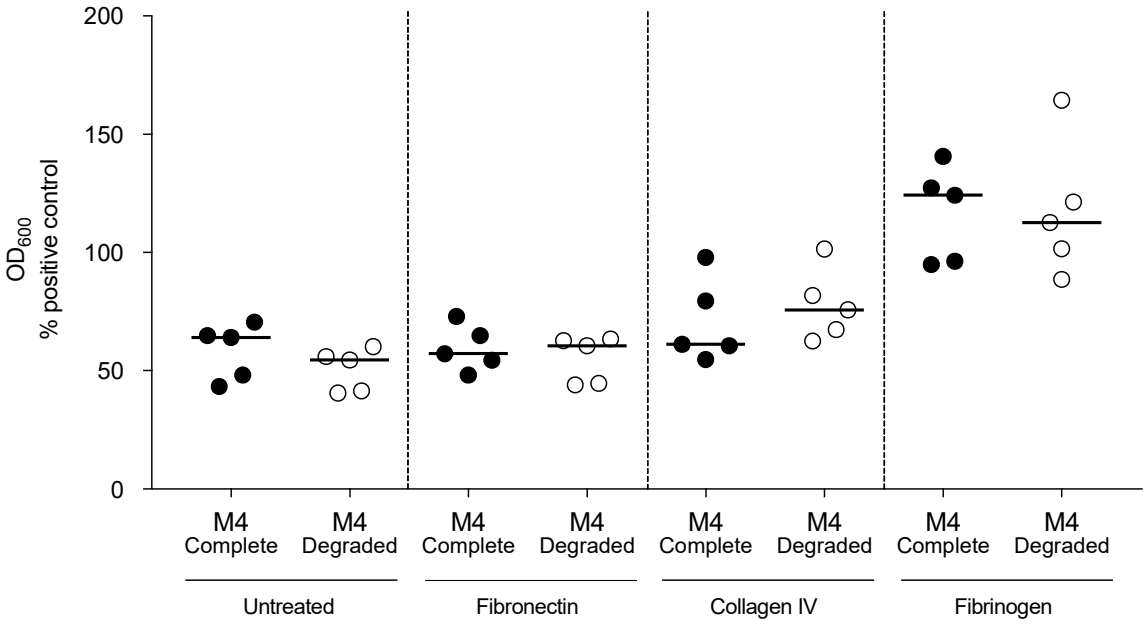

Figure 5.6: Adhesion of *emm*4 isolates to eukaryotic substrata

Polystyrene plates were pre-coated with fibronectin, collagen IV and fibrinogen and *emm*4 isolates assayed for adhesion, relative to OD_{600} internal control *emm*6 isolate BSAC_bs2438 after 24-hours incubation in THB. Samples were measured in triplicate, and OD_{600} of negative control wells (sterile THB) was deducted from all sample and positive control wells. The percentage OD_{600} of positive control was calculated by dividing triplicate means of sample wells by triplicate mean of the positive the control, multiplied by 100 to yield a percentage. Data indicate a single experimental repeat. Horizontal bars indicate data median.

5.5 Discussion

The data in this Chapter successfully characterised *emm*4 isolates by growth in THB and whole human blood, and in terms of virulence in the *G. mellonella* model. Growth of $M4_{degraded}$ and $M4_{complete}$ isolates was very similar both in THB, and in whole human blood. Superantigenic activity, although challenging, irreproducible, and thus open to interpretation, appeared to demonstrate comparable superantigenicity of overnight RPMI +10% (v/v) FCS supernatant filtrates derived from isolates belonging to either *emm*4 lineage. Perhaps most intriguingly, $M4_{complete}$ were significantly more virulent than $M4_{degraded}$ in a *G. mellonella* model of infection.

There was no demonstrable growth phenotype associated with isolates from either lineage when measured by optical density or CFU/mL at 3-, 5-, and 8-hours of growth in THB (Figure 5.1), indicating that none of the lineage-specific (or isolate-specific) SNPs contribute significantly growth under these conditions. These data would also appear to preclude the possibility for spontaneous prophage induction and bacterial lysis, if correlated with a specific timepoint, to be sufficient to evoke a drop in viable bacterial counts. This was expected, as data in Chapter 4 suggested this would not be the case.

Despite being genetically acapsular, *emm*4 isolates survived and grew well in whole human blood, in corroboration with the burgeoning body of published literature that the HA capsule is inessential for whole blood survival, certainly in specific genotypes (Flores *et al.*, 2012; Henningham *et al.*, 2014; Turner *et al.*, 2015), and both $M4_{complete}$ and $M4_{degraded}$ isolates survived comparably in both donors. There was some variation between the two donors used, but this did not appear to be due to specific immunity to M4.The two isolates carrying variant CovS isoforms, BSAC_bs468 (Thr241Pro and Ala379Thr), BSAC_bs1388 (36aa-STOP), and BSAC_bs1802, and the isolate carrying a variant *covR* allele (Gly91Asp), did not exhibit enhanced growth relative to the other *emm*4 isolates. The *emm*/*enn* splicing event identified in $M4_{degraded}$ isolates in Chapter 3 does not appear to enhance or diminish survival in whole human blood, at least in this population of *emm*4 isolates, and using these donors.

The superantigenicity of overnight culture supernatants collected from *emm*4 isolates was not significantly different between $M4_{complete}$ and $M4_{degraded}$ isolates (Figure 5.3). This may be accurate, however, the BrdU ELISA implemented here inherently lacks the power to discriminate between the activity of each superantigen present in the respective isolates and therefore can only detect overall activity. Regardless of this fact, it was unexpectedly challenging to generate reproducible data using the Roche BrdU ELISA kit on multiple attempts, and it did not seem to be amenable to the intended application. Certainly, the observation that internal control H305 appeared less mitogenic than *emm*4 isolates, was noted with suspicion.

Previous studies suggest that recombinant SMEZ appears to be more potent than SpeC, while both appear to be more potent still than SSA (Afshar *et al.*, 2017; Nooh *et al.*, 2006; Proft and Fraser, 2016; Reglinski *et al.*, 2019). Very recent work describes a mechanism by which the release and activity of SSA is dependent on SLO mediated glutathione efflux (Brouwer *et al.*, 2020). As such, the activity of SSA might be further diminished under the conditions described here, due to a possible paucity of glutathione, posing further potential technical difficulties in extracting meaningful data from the assay, particularly in the absence of isogenic mutants. Future work may thus benefit from the supplementation of culture supernatants with this antioxidant. CovS variant isolates BSAC_bs468 (Thr241Pro and Ala379Thr) and BSAC_bs1388 (36aa-STOP), and CovR variant BSAC_bs1802 (Gly91Asp) were not significantly more superantigenic than other *emm*4 isolates. The major regulator CovRS has been reported to repress the expression of prophage-encoded virulence factors, including the streptococcal superantigen *speA* (Sumby *et al.*, 2006). Although genotype *emm*4 isolates do not typically carry *speA*, nor was it

detected in any of the BSAC *emm*4 isolates, it was possible that against the *emm*4 background, CovRS regulated the expression of *ssa*, *smeZ* and/or *speC* in RPMI +10% (v/v) FCS, in which overnight cultures for this specific assay were grown, might vary.

The observation that $M4_{complete}$ isolates were more virulent than $M4_{degraded}$ isolates was unexpected, although the majority of larval death occurred within 24-hours of challenge in all three experiments. Although *G. mellonella* larvae are insensitive to streptococcal superantigens, lacking an adaptive immune system, this model is sensitive to secreted nucleases and a number of other streptococcal virulence factors (Huang *et al.*, 2015; Olsen *et al.*, 2011). Data in Chapter 4 demonstrated a difference in the expression of *speC* at 8-hours growth in THB, which is co-transcribed with the DNase *spd1* (Broudy *et al.*, 2002). It was therefore conceivable that a similar phenomenon was occurring during *G. mellonella*, which were assayed for 7 days. A number of subsets of haemocytes in *G. mellonella* also have phagocytic capacity, and extrude structures similar to NETs (Altincicek *et al.*, 2008; Olsen *et al.*, 2011), so it is possible that enhanced DNase activity, inferred by polarised expression of *speC* among $M4_{complete}$ isolates relative to $M4_{degraded}$ isolates, could explain these findings. Though equally, this could also be due to chromosomally-encoded, secreted DNases, which were not measured.
Empirical evidence of *S. pyogenes* as being heavily reliant on fomite persistence and environmental reservoirs such as children's toys and other inert surfaces, are both scant and conflicting (Efstratiou and Lamagni, 2016; Ingham *et al.*, 2006; Marks *et al.*, 2014b). Typically, transmission requires contact, or shared vicinity with members of the community, such as family members or healthcare workers, who may be colonised and carrying *S. pyogenes* or are experiencing clinical disease, whereby exposure to infectious respiratory droplets or colonised tissue can be achieved (Afshar *et al.*, 2017; Flores *et al.*, 2017; Galloway-Peña *et al.*, 2016; Turner *et al.*, 2013). As a human-restricted pathogen, however, the capacity for *S. pyogenes* adhere both to and persist within the host is undeniable (Ryan and Juncosa, 2016), so it was apposite to attempt to probe these associations. In this case with *emm*4 isolates appeared to exhibit a predilection for fibrinogen, relative to other tested substrata, and certainly in preference to polystyrene, relative to other substrates. The reason for this is unclear, however, it has been suggested that a number of M4- and M-like proteins of *emm*4 may mediate fibrinogen binding, with variable success (Courtney *et al.*, 2006b; Podbielski *et al.*, 1996). The acapsular nature of *emm*4 may also play a role (Flores *et al.*, 2012). In *S. pneumoniae*, prophage induction has been associated with enhanced biofilm formation (Carrolo *et al.*, 2010). Under the conditions explored herein, if this does contribute to the adhesion of *emm*4 with inducible prophage, it does not appear to be a major mechanism.

Taken together, the data in this Chapter, and those before it, raise a number of questions regarding how *emm*4 isolates, both $M4_{complete}$ and $M4_{degraded}$, might behave in more faithful reproductions of the host-pathogen interface. These were explored in the next Chapter, using a human tonsil co-culture model as a paradigm.

5.5.1 **Chapter 5: Key findings**

- Both $M4_{complete}$ and $M4_{degraded}$ isolates grew comparably in THB and survived and multiplied comparably in whole human blood
- Adhesion to polystyrene and eukaryotic substrata was also comparable between both $M4_{complete}$ and $M4_{degraded}$
- Using a crude model to assay overall virulence, *G. mellonella* infection experiments suggested that $M4_{complete}$ isolates may be more virulent than $M4_{degraded}$
- Otherwise, $M4_{complete}$ and $M4_{degraded}$ appeared phenotypically very similar

Chapter 6

6. Modulation of superantigen expression by HTE co-culture

6.1 Introduction

S. pyogenes is a dynamic pathogen capable of causing a variety of human diseases both at distinct anatomical sites such as the skin or tonsil, and systemically, surviving in blood and lymph and even within immune cells such as neutrophils and macrophages (Armitage *et al.*, 2019; Carapetis *et al.*, 2005; Cunningham, 2000; Lynskey *et al.*, 2015; O'Neill *et al.*, 2016; Siggins *et al.*, 2020; Valderrama and Nizet, 2018). In order to economise, the bacterium must adapt, regulate and coordinate the expression of specific necessary genes at each stage of infection and within each niche, and minimise the expression of genes that are not useful (Banks *et al.*, 2016; Broudy *et al.*, 2002, 2001; Kachroo *et al.*, 2019; Sitkiewicz *et al.*, 2010; Virtaneva *et al.*, 2005; Zhu *et al.*, 2017). This Chapter will focus on the primary infection site of *S. pyogenes,* the human tonsil, and the role of prophage-encoded virulence factors at this critical interface between host and pathogen. Streptococcal superantigens appear to play a role in host colonisation (Afshar *et al.*, 2017; Zeppa *et al.*, 2017) and modulating immunity in the human tonsil to benefit the pathogen (Davies *et al.*, 2019). The role of the streptococcal DNases, the putative pathogenic influence of which was, until relatively recently, largely thought to be constrained to promoting the expedition of ensnared streptococci from NETs (Buchanan *et al.*, 2006; Derré-Bobillot *et al.*, 2013; Ma *et al.*, 2017; Sumby *et al.*, 2005a), may also promote infection and colonisation in the nasopharynx (Afshar *et al.*, 2017), and may even act in synergy with the streptococcal superantigens to this end (Brouwer *et al.*, 2020).

It was previously reported that co-culture of *S. pyogenes* with human pharyngeal cells induced toxigenic prophage and increased expression of SpeC and Spd1, as inferred by plaque enumeration, Western blot, and visualisation of phage particles (Broudy *et al.*, 2002, 2001). The former study described (though did not demonstrate) that a similar inducing effect had been observed following co-culture of *S. pyogenes* and a breast cancer (mcf-7) cell line. A study published the following year, using *emm*3 strain MGAS315, infers prophage induction by PCR to detect toxin genes in purified phage DNA extracts, derived from supernatant filtrates recovered following co-culture, using *speB* as a control for contaminating genomic DNA (Banks *et al.*, 2003). Of the five canonical *emm*3 prophage, co-culture appears to increase the induction of Φ315.4 and Φ315.6, carrying *speK*/*sla* and *sdn*, respectively. Evidence of the remaining prophage in filtrates were undetectable. Subsequent RT-qPCR analyses demonstrated that transcription of *sdn* was enhanced, whereas *speK* transcripts were found to decrease, while transcripts of *sla* increased. This was mirrored in quantities of these toxins detected in filtrates, as determined by Western blot. Notably caveats to these experiments are that MGAS315 does not carry an *speC*/*spd1* prophage (Al-Shahib *et al.*, 2016; Beres *et al.*, 2002), the authors do not measure *ssa* transcripts, nor SSA toxin expression, and were unable to source an antibody for Sdn or Spd4.

As the published literature stands, the authors suggest that a specific molecule secreted by human cells is responsible. This is referred to specifically in the earlier two of the three studies, and hereafter, as SPIF (soluble phage inducing factor) (Broudy *et al.*, 2002, 2001). The identity of SPIF is not known, and at the time of writing, identification and characterisation of the molecule does not appear to have continued. Earlier work suggests that SPIF may be a soluble molecule of <10kDa in size that is resistant to protease digestion, heating to 100°C, and storage at -20°C over a number of months (Broudy *et al.*, 2001).

In Chapter 4, it was shown that in THB, spontaneous prophage induction did not convincingly correlate with a specific bacterial growth timepoint. Additionally, with the exception of *speC* transcription, which was raised at late-log, transcription of *spd3* and *ssa* were similar between lineages, and across timepoints in THB. The elevated transcription at 8-hours of *speC*, as observed in Chapter 4, did not appear to be associated with an increase in *speC* gene copies. Reflecting on these published data, and the data described in the present work, it was apposite to determine if tonsil cells were able to induce Φ10750.1/Φ427.1, Φ10750.2/Φ427.2, and Φ10750.3/Φ427.3, and further, to determine if these cells could enhance the expression of cognate toxins.

6.2 Chapter hypotheses and aims

Work described in this Chapter continued prior efforts to demonstrate a polarising phenotype between $M4_{complete}$ like and $M4_{degraded}$ isolates. These isolates were studied in a more physiologically- and infection-relevant system, one designed to emulate the conditions of the host-pathogen interface between *S. pyogenes* and its primary anatomical niche; the human tonsil. Under these conditions, the hypothesis that differential prophage induction between $M4_{complete}$ and $M4_{degraded}$ described in previous chapters could elicit a polarising phenotype of prophage-associated toxin expression was tested, to unexpected and insightful effect. The co-culture model was also applied to other major *emm*-types.

6.3 Methods

6.3.1 Storage and recovery of HTE cells and A549s

All human cells were maintained in liquid nitrogen at -80°C until required. For recovery, cells were thawed swiftly by gentle swirling in a water bath at 37°C. Thawed 1 mL suspensions were then transferred to sterile 15 mL tubes and washed in 9 mL pre-warmed media, and supernatant discarded. The pellet was resuspended in either 5 mL or 10 mL of pre-warmed media and transferred to T25 or T75 flasks (Corning), respectively. Flasks were incubated in a humidified environment at 37 °C, supplemented with 5% (v/v) CO_2 for 24 hours, before the media was changed to remove any residual DMSO. For splitting and seeding, cells were washed twice with pre-warmed PBS

and trypsinised using 1X Trypsin-EDTA solution (Insight Biotechnology) as and when appropriate, with digestion blocked using 10% (v/v) FCS.

6.3.2 *S. pyogenes* co-culture with HTE cells and A549s

Lifted and resuspended HTE cells were live-dead stained using Trypan Blue (Sigma) enumerated using an Improved Neubauer counting chamber (Sigma) and seeded into 6-well plates (Corning) at 1.25 x 10^5 cells/mL (2.5 x 10^5 cells/well) in Green's (flavin and adenine rich) Media (prepared per methods section 2.1.5), without antibiotics and incubated overnight to confluence. Overnight cultures of *S. pyogenes* were washed twice with pre-warmed RPMI +10% (v/v) FCS, and adjusted to OD_{600}1.0 in pre-warmed RPMI +10% (v/v) FCS, and diluted by a further 1:5 to achieve an approximate MOI of 100 (1 x 10^7 CFU). An aliquot of each was retained for plating for CFU/mL enumeration. Prior to use, HTE monolayers were washed twice in pre-warmed RPMI +10% (v/v) FCS, which was then discarded to remove residual Green's Media and cellular debris. 2 mL of the adjusted *S. pyogenes* suspension was added to each well containing either HTE monolayers, or empty wells intended to incubate 2 mL of the adjusted *S. pyogenes* RPMI +10% (v/v) FCS suspension alone. For RNA extraction, two HTE-seeded wells and two empty wells were used per isolate to ensure extraction of sufficient quantities of RNA. Single wells proved sufficient for DNA extraction, for both HTE-seeded wells, as well as those containing the adjusted bacterial/ RPMI +10% (v/v) FCS suspension alone. Once inoculated, plates were incubated for 6-hours in a humidified atmosphere at 37°C, supplemented with 5% (v/v) CO_2. At the end of the experiment, 100 µL aliquots were removed for plating and CFU/mL enumeration from each well. For RNA extraction, isolate- and culture condition-matched wells were homogenised by pipetting, pooled and subsequently processed as a single sample, immediately, as previously described in the main methods (section 2.2.3). Wells from which DNA was to be extracted were homogenised, pelleted by centrifugation, and frozen at -20°C until extraction proceeded as described in the main methods (section 2.2.2).

6.4 Results

6.4.1 Growth of *S. pyogenes* in RPMI +10% (v/v) FCS and following HTE co-culture

HTE wells seeded at 1.25 x 10^5 cells per mL (2.5 x 10^5 /well) were inoculated at an intended bacterial multiplicity of infection of 100, and a range of between 1 x 10^6 and 1.16 x 10^8 CFU/mL was achieved across experiments. All co-culture experiments were carried out using RPMI +10% (v/v) FCS, following noted and undesirable differences in fold-change between CFU/mL of bacterial inocula and CFU/mL post-incubation, between *S. pyogenes* cultured either Green's or cultured in RPMI +10% (v/v) FCS (Figure 6.1). Both lineages multiplied comparably in RPMI +10% (v/v) FCS or when subject to co-culture with HTE cells.

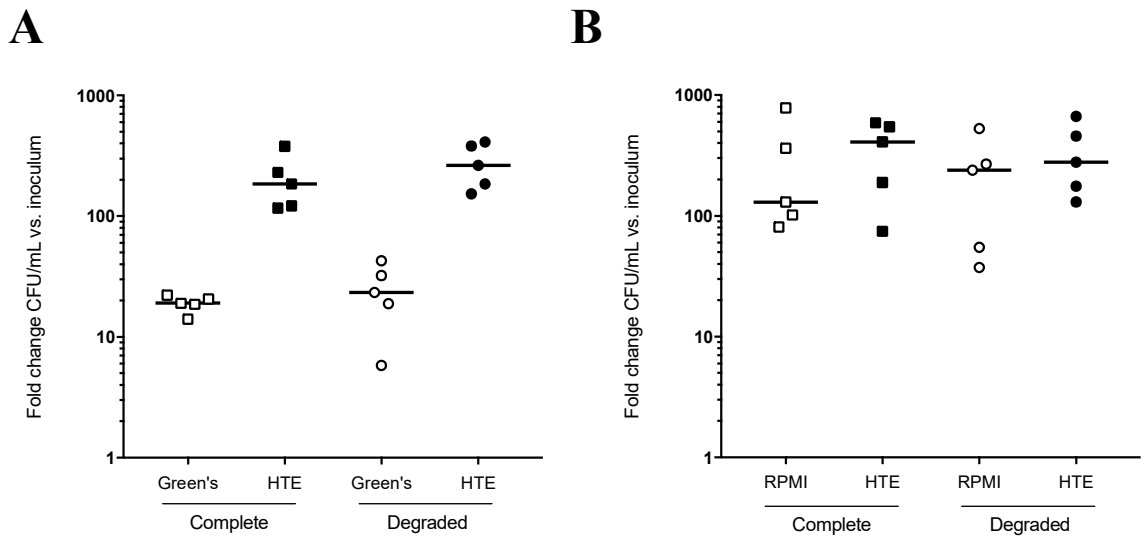

Figure 6.1: Differential growth of *S. pyogenes* in Green's versus RPMI +10% (v/v) FCS

Fold-change in CFU/mL from the bacterial inoculum to CFU/mL recovered post-incubation of *S. pyogenes* with either HTE cells and cell culture media, or *S. pyogenes* and cell culture media (without HTE cells) using two types of media: Green's Media or RPMI +10% (v/v) FCS. A) A marked fold-increase in CFU/mL from the inoculum was observed between wells that contained Green's Media and *S. pyogenes* alone, compared to wells containing *S. pyogenes* co-cultured with HTE monolayers in Green's Media. B) The fold-change from bacterial inoculum CFU/mL to post-incubation CFU/mL was considerably less between *S. pyogenes* cultured in well containing RPMI +10% (v/v) FCS alone, and wells containing *S. pyogenes* and HTE monolayers in RPMI +10% (v/v) FCS. Data represent the mean of duplicate averages. Horizontal bars indicate the data median.

6.4.2 **Human tonsillar epithelium does not enhance prophage induction**

To emulate more infection-relevant conditions, and to determine if human cells could induce the *emm*4 prophage, prophage induction was probed by PCR using standardised DNA extracted from the BSAC isolates following incubation in RPMI +10% (v/v) FCS, with or without the presence of a confluent monolayer of HTE cells. Unfortunately, PCR was unsuccessful for the majority of the samples, which may have been due to a high level of eukaryotic DNA extracted from HTE cells. The potential for HTE cells to internalise and sequestrate phage was also considered. Though efforts were made to homogenise co-culture wells, eukaryotic cells in these wells were not subject to a lysis step. Thus, representative isolates were selected from $M4_{complete}$ and $M4_{degraded}$ isolates for which successful PCR was achieved.

Prophage induction was not detected among the prophage of the $M4_{degraded}$ isolate, as expected, but unusual $M4_{degraded}$ isolate BSAC_bs1802, for which Φ1802.1 and the near-complete Φ10750.3 of this isolate were found to be induced under both conditions. Intriguingly, for those isolates with complete prophage, although induction was detected, this did not appear to be consistently enhanced by the presence of HTE cells (Figure 6.2).

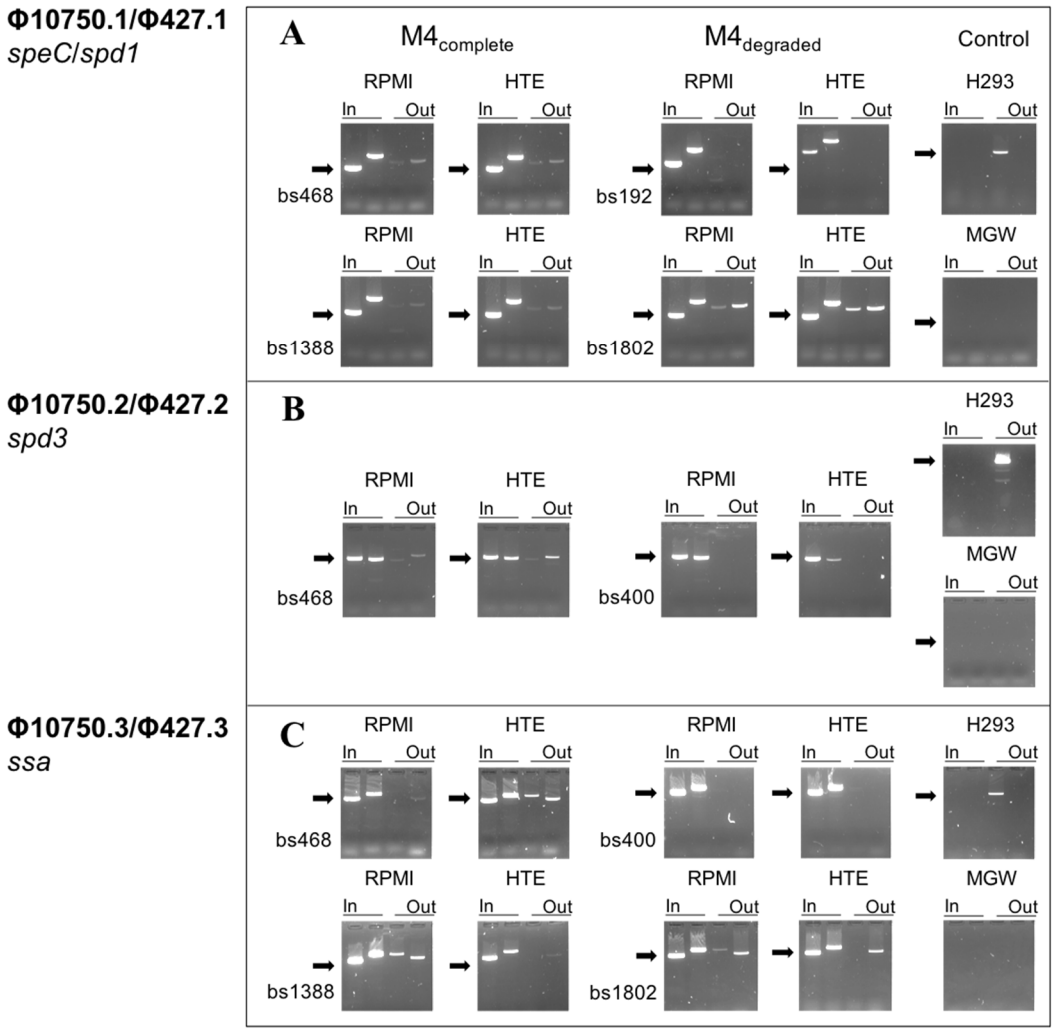

Figure 6.2: Prophage induction following co-culture with human tonsillar epithelium

DNA was extracted from BSAC *emm*4 isolates grown in RPMI +10% (v/v) FCS alone, or in RPMI +10% (v/v) FCS in the presence of a confluent monolayer of human tonsillar epithelium (HTE), and probed by PCR for integration (In) and excision (Out) of Φ10750.1/Φ427.1, Φ10750.2/Φ427.2, Φ10750.3/Φ427.3. A) Φ10750.1 of $M4_{complete}$ BSAC_bs468 and BSAC_bs1388 were induced comparatively under both conditions. Φ427.1 of $M4_{degraded}$ BSAC_bs192 was not induced, but Φ1802.1 was induced strongly under both conditions. Arrow indicates ~400bp. B) Φ10750.2 of $M4_{complete}$ BSAC_bs468 was induced under both conditions tested, whereas Φ427.2 of $M4_{degraded}$ BSAC_bs400 was not. Arrow indicates ~1kb C) Φ10750.3 of $M4_{complete}$ isolates BSAC_bs468 and BSAC_bs1388 were induced under both conditions tested, as was unusual $M4_{degraded}$ Φ1802.3. $M4_{degraded}$ BSAC_bs400, was not induced. Arrow indicates ~1kb. DNA was standardised to 40 ng/μL prior to PCR. Water served as a negative control, and genotype *emm*89 isolate H293 which lacks prophage was used as an internal control.

6.4.3 **Modulation of superantigen expression following HTE co-culture**

Prophage induction was not generally enhanced when $M4_{complete}$ and $M_{4degraded}$ isolates were co-cultured with HTE cells. Although an increase in prophage induction and replication may elicit an enhancement in gene expression through the provision of a greater quantity of templates for transcription, this may not be the sole mechanism by which prophage-associated toxins are regulated (Banks *et al.*, 2003). Thus, transcripts were measured to further explore the relationship between prophage induction and gene expression in genotype *emm*4 *S. pyogenes* by converting extracted RNA samples to cDNA (per section 2.2.11 of the main methods) and applying the previously described qPCR method, normalised to a standard curve (per 2.2.9 section of the main methods). Surprisingly, the expression of *speC* was significantly reduced following HTE co-culture ($p=0.0200$, Wilcoxon signed rank test) (Figure 6.3A), although with no major difference between $M4_{complete}$ and $M4_{degraded}$ isolates (Figure 6.3B). Expression of *speC* by CovRS variant *emm*4 isolates, BSAC_bs468, BSAC_bs1388 and BSAC_bs1802, were the lowest of all isolates, both in RPMI +10% (v/v) FCS and following co-culture with HTE cells.

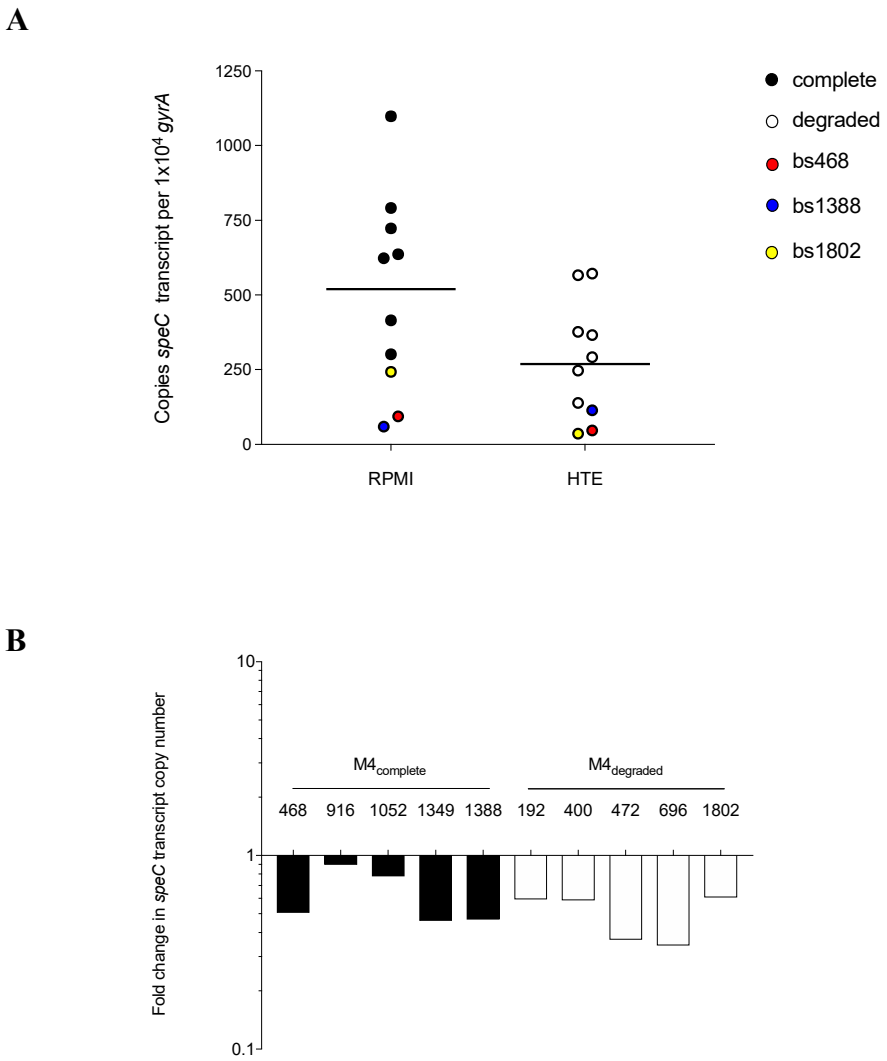

Figure 6.3. Expression of *speC* following co-culture with human tonsillar epithelium

A) Copies of *speC* transcripts were lower in RPMI +10% (v/v) FCS alone (black circles) than following co-culture with HTE cells (white circles) in all *emm*4 isolates (p=0.002, Wilcoxon signed rank test). Variant *covS* isolates BSAC_bs468 (red circle) and BSAC_bs1388 (blue circle), and variant *covR* isolate BSAC_bs1802 (yellow circle) expressed little *speC* under either culture conditions. Data represent RNA samples measured in triplicate and normalised to copies of *speC* transcript per 10,000 transcripts of housekeeping gene *gyrA*. Horizontal bars indicate the median. B) Data represented as fold-change in *speC* per 10,000 *gyrA* following co-culture with HTE cells. Similar fold change was observed between $M4_{complete}$ isolates (black bars) and $M4_{degraded}$ isolates (white bars).

Expression of *spd3* was also measured for all 10 BSAC *emm*4 isolates by RT-qPCR, however, transcripts were below the limit of detection for all isolates both in RPMI +10% (v/v) FCS alone, and following co-culture with HTE cells (data not shown). When expression of *ssa* was measured, however, this was found to be significantly and uniformly enhanced among *emm*4 isolates following co-culture, relative to culture in RPMI +10% (v/v) FCS alone (p=0.0020, Wilcoxon signed rank test) (Figure 6.4A). Degraded isolate BSAC_bs192 expressed more *ssa* than all the other isolates following HTE co-culture. CovRS variant isolates BSAC_bs468 (CovS; Thr241Pro and

Ala379Thr), BSAC_1388 and BSAC_bs1802 (CovR; Gly91Asp) expressed *ssa* transcript copies than other *emm*4 in RPMI +10% (v/v) FCS alone. However, following co-culture, BSAC_bs1388 and BSAC_bs1802 did not respond as did the other *emm*4 by enhancing *ssa* transcription. Transcript copies of *ssa* was similar between $M4_{complete}$ and $M4_{degraded}$ isolates. The fold-change in *ssa* transcript copies among all isolates ranged between 1.2- and 9-fold (Figure 6.4B)

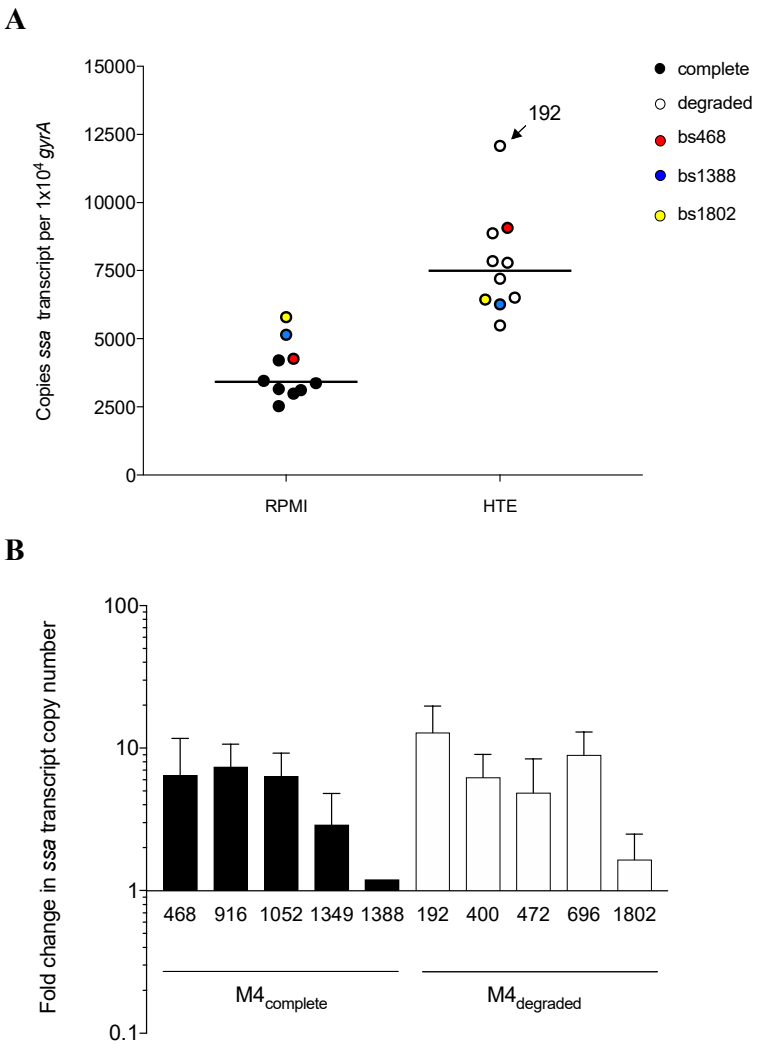

Figure 6.4: Expression of *ssa* following co-culture with human tonsillar epithelium A)

Transcription of *ssa* was found to be enhanced among all *emm*4 isolates of both lineages by HTE cells (p=0.0020, Wilcoxon signed rank test). The single *covR* variant isolate BSAC_bs1802 is denoted with a yellow circle. Isolates with variant *covS* alleles, BSAC_bs468 and BSAC_bs1388, are indicated as red and blue circles, respectively. Data represent the mean of three individual experiments and horizontal bars indicate data median. B) Data presented as average fold-change in *ssa* transcript copies between *emm*4 isolates cultured in RPMI +10% (v/v) FCS alone, and following co-culture with HTE cells. No major difference was observed between $M4_{complete}$ isolates (black bars) and $M4_{degraded}$ isolates (white bars). Data represent the mean (+ standard error) of three independent RNA samples measured in triplicate and normalised to copies of *ssa* transcript per 10,000 copies of housekeeping gene *gyrA* transcripts.

6.4.4 **Expression of *ssa* was also enhanced by A549 cells**

To determine if the phenomenon reported in the present work was tonsil-specific, the co-culture experiment was repeated with A549 cells, a pneumocyte cell line, and a smaller sample from the same 10 *emm*4 isolates: BSAC_bs192, BSAC_bs696, BSAC_bs916, BSAC_bs1052. A459 cells did elicit an increase in *ssa* transcript copy number among the isolates tested (Figure 6.5A). Neither $M4_{complete}$ nor $M4_{degraded}$ isolates expressed *ssa* more highly than the other in RPMI +10% (v/v) FCS alone, or following A549 co-culture, and fold-change ranged from 2- to 7-fold (Figure 6.5B).

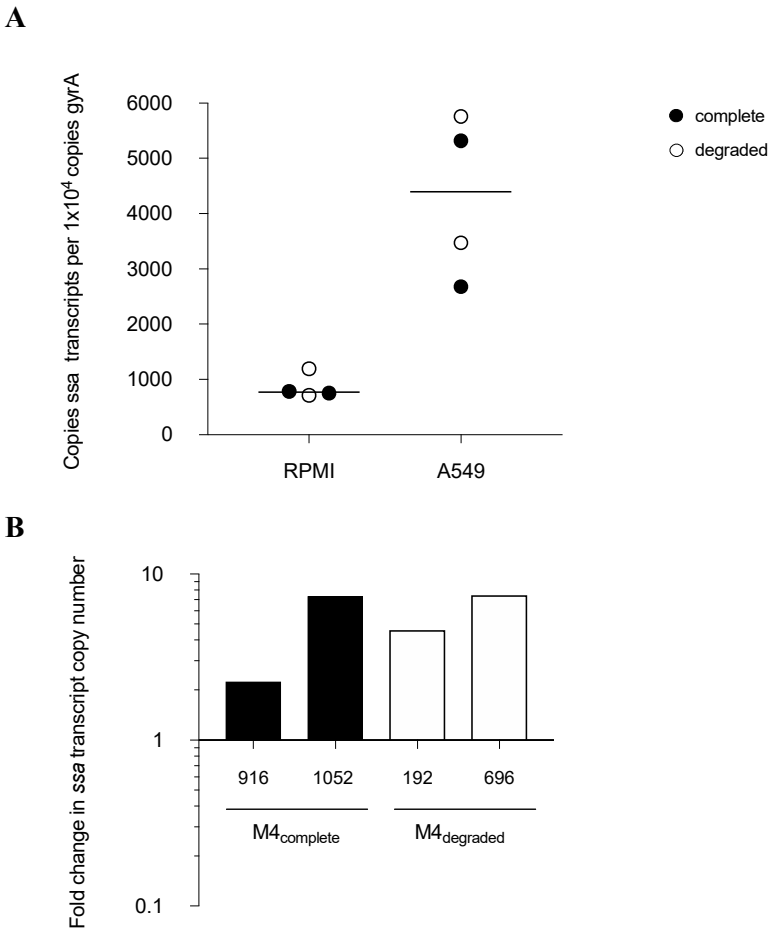

Figure 6.5.: Expression of *ssa* by *emm*4 isolates following co-culture with A549s

A) Co-culture of *emm*4 isolates with A549 cells (white circles) also enhanced the expression of *ssa,* relative to culture in RPMI +10% (v/v) FCS alone (black circles). Data represent singular RNA samples extracted from two $M4_{complete}$ (BSAC_bs192 and BSAC_bs696) isolates and two $M4_{degraded}$ (BSAC_bs916 and BSAC_bs1052) isolates, measured in triplicate and normalised to copies of *ssa* transcript per 10,000 copies of housekeeping gene *gyrA* transcripts. Horizontal bars indicate the data median. B) Data expressed as mean fold change in *ssa* transcript copies from culture in RPMI +10% (v/v) FCS alone in $M4_{complete}$ (black bars) and $M4_{degraded}$ (white bars), to *ssa* transcript copies detected following co-culture with HTE cells.

6.4.5 **Enhanced *ssa* expression was not associated with an increase in gene copies**

Enhanced expression of prophage-encoded virulence factors may be achieved following prophage induction, as a consequence of prophage induction (Banks *et al.*, 2003; Broudy *et al.*, 2002, 2001), though hitherto has not been supported by the present work. Following the increase in *ssa* transcription that was observed among both $M4_{complete}$ and $M4_{degraded}$ isolates with HTE-co-culture, gene copies of *ssa* were measured to determine if this was associated with a concomitant increase in prophage replication. Φ10750.3 can be induced to excise, however, as yet, it does not appear that $M4_{degraded}$ are capable of excision, including on co-culture with HTE cells. Gene copies were similar under both conditions for all *emm*4 isolates ($p=0.1602$, Wilcoxon signed rank test), and surprisingly, an increase in gene copies was not observed among $M4_{complete}$ isolates relative to $M4_{degraded}$. One CovS variant $M4_{complete}$ isolate, BSAC_bs1388 (36aa-STOP), did exhibit a marked increase in *ssa* gene copies upon exposure to HTE, yet surprisingly, had exhibited the lowest fold-change in *ssa* transcripts of all *emm*4, ~1.2-fold. The unusual $M4_{degraded}$ isolate BSAC_bs1802, which carried a naturally occurring CovR variant (Gly91Asp) also appeared to demonstrate an increase in *ssa* gene copies, despite not having responded to HTE co-culture by appreciably enhancing expression of *ssa* (~1.5 fold-change).

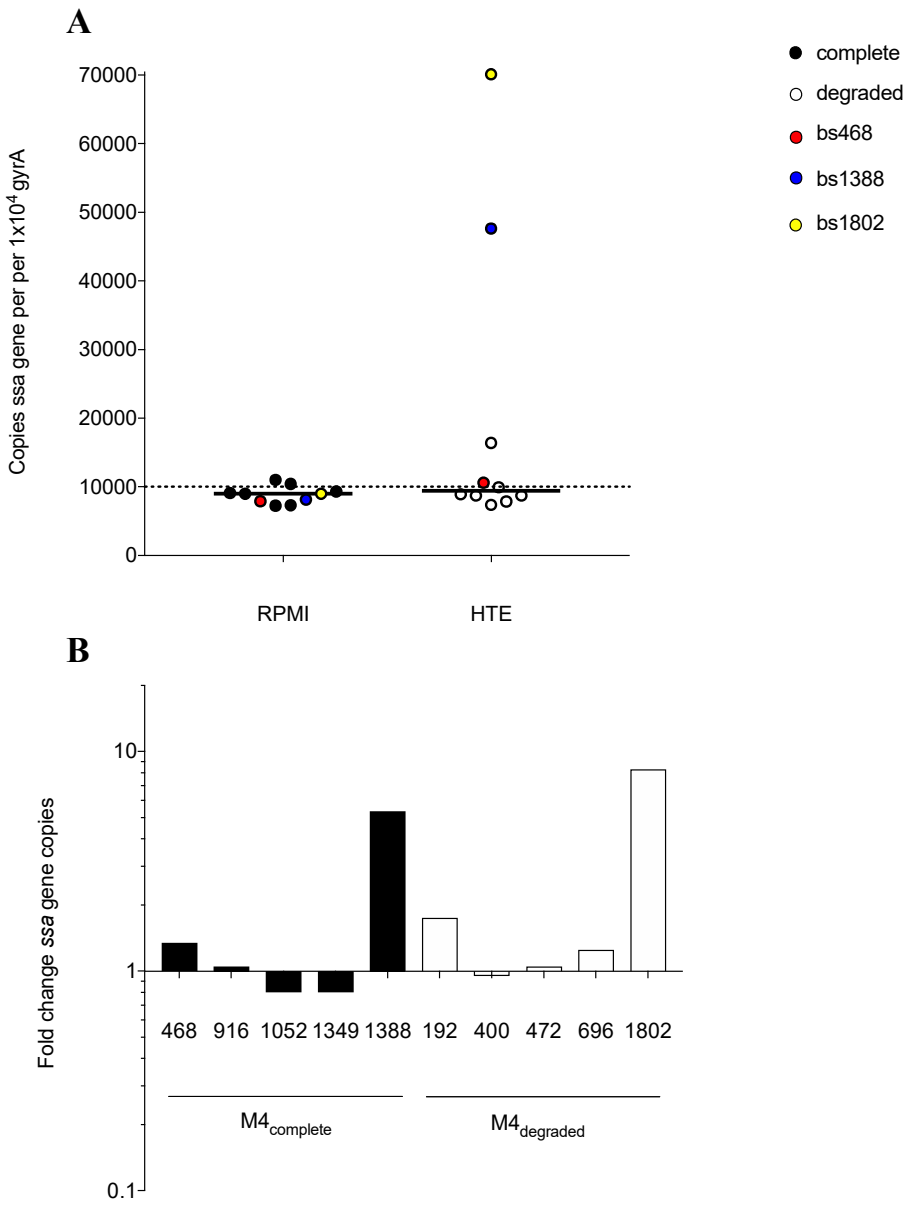

Figure 6.6: Copies of *ssa* gene following co-culture with human tonsillar epithelium

A) The enhanced expression of *ssa* was not associated with an increase in gene copy number among *emm*4 isolates (p=0.1602, Wilcoxon signed rank test). BSAC_bs1802 (yellow circle) which carries a naturally occurring *covR* allele, and an isolate carrying a naturally occurring *covS* allele, BSAC_bs1388, (blue circle) demonstrated an increase in *ssa* gene copies. An increase in *ssa* gene copy number was not observed of *covS* variant isolate BSAC_bs468 (red circle). Dotted line indicates the point at which copies of *gyrA* and *ssa* are approximately equivalent, 10,000 copies. Data represent the mean of calculated averages for all ten *emm*4 isolates measured in triplicate, across two experiments, and normalised to 10,000 copies *gyrA*. B) Data presented as mean fold-change in *ssa* gene copies between culture in RPMI +10% (v/v) FCS alone, and following co-culture with human tonsillar epithelial cells. There was no difference in gene copy number between $M4_{complete}$ (black bars) and $M4_{degraded}$ (white bars). Data represent the mean of two independent DNA samples measured in triplicate and normalised to copies of *ssa* per 10,000 copies of housekeeping gene *gyrA*.

6.4.6 **Expression of *ssa* is enhanced by human tonsillar epithelium among other genotypes**

Having found that *ssa* expression was enhanced among genotype *emm*4 isolates following co-culture with HTE cells, the co-culture model was applied to other genotypes which are frequently lysogenised by prophage carrying *ssa.* Genotypes *emm*3 and *emm*12 are two such *emm*-types (Beres *et al.*, 2004; Chalker *et al.*, 2017; Turner *et al.*, 2017, 2019). Isolates belonging to these genotypes in the BSAC collection were thus screened for *ssa* by BLAST to detect the gene among *emm*3 and *emm*12 isolates. Enhanced expression of *ssa* was observed among *emm*3 isolates following co-culture (p=0.0312, Wilcoxon signed rank test). Similarly, *emm*12 isolates also exhibited enhanced expression of *ssa* following co-culture (p=0.0312, Wilcoxon signed rank test). Expression of *ssa* in RPMI +10% (v/v) FCS alone was highest among *emm*4 isolates, and *emm*12 isolates exhibited a higher baseline expression than did *emm*3 (Figure 6.7A). Expression of *ssa* by genotype *emm*4 isolates following HTE co-culture was higher than both *emm*3 and *emm*12 following co-culture. The expression of *ssa* by genotype *emm*3 isolates increased following co-culture with HTE cells, between 1- to 5-fold. A similar response was observed among *emm*12 isolates, increasing between 1- to 6-fold. A fold-change of between 1- and 13-fold was observed among *emm*4 isolates (Figure 6.4B).

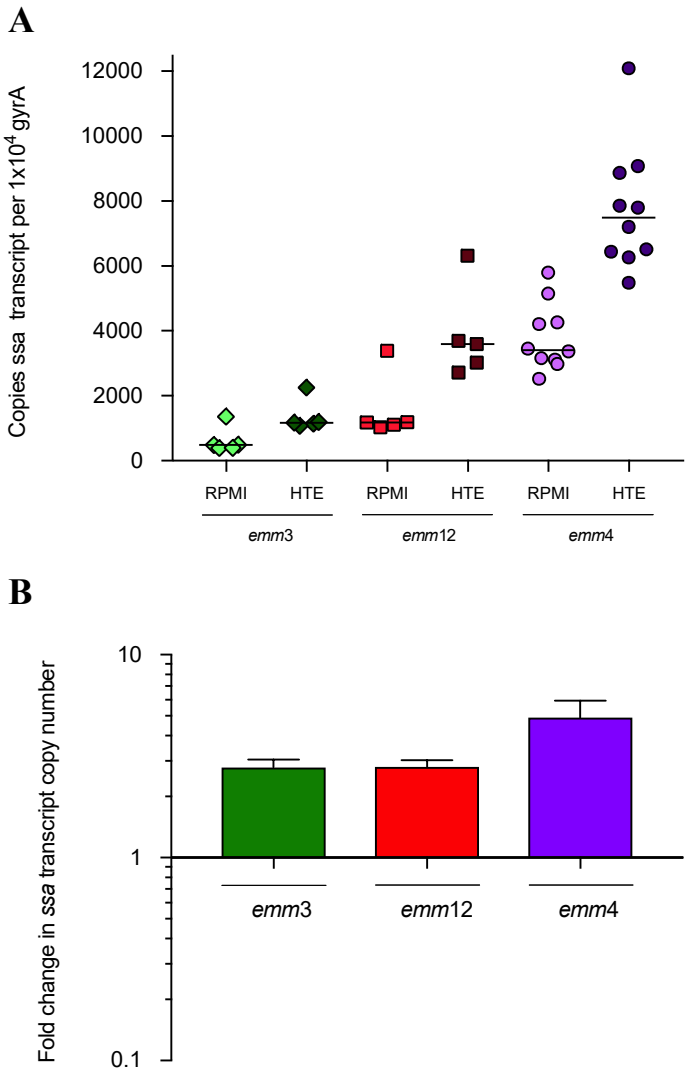

Figure 6.7: Human tonsillar epithelium enhances *ssa* expression in other genotypes

A) Genotype *emm*3 (green diamonds) exhibited a borderline significant enhancement of *ssa* expression following co-culture (p=0.0312, Wilcoxon signed rank test), as did *emm*12 isolates (red squares) (p=0.0312, Wilcoxon signed rank test), as was previously observed in *emm*4, (reproduced from fig. 6.4 for reference; purple circles). Data represent the mean of three independent RNA samples measured in triplicate and normalised to transcript copies of *ssa* per 10,000 transcript copies of *gyrA.* Horizontal bars indicate the data median. B) Data represented as fold-change in *ssa* transcript copy number, per 10,000 transcript copies of *gyrA*. Genotype *emm*4 isolates (reproduced from fig. 6.4; purple bar) increased *ssa* expression by 1-13-fold following co-culture. Genotype *emm*3 (red bar) and *emm*12 isolates (green bar) increased *ssa* expression by 1-5-fold, and 1-6-fold, respectively. Error bars indicate mean (+standard error of the mean).

6.5 Discussion

The data in this Chapter demonstrated that co-culture of *emm*4 isolates with human tonsillar epithelial cells did not seem to trigger the induction of prophage. The finding that both HTE cells and A549s were able to enhance the expression of the streptococcal superantigen *ssa* is novel, and thought provoking. Particularly when contrasted

with the reduction of *speC* expression. There was a noted dearth of evidence to support the hypotheses of Chapter 4 that the enhanced expression of prophage-associated toxins may be associated with a concomitant increase in prophage induction, and an increase in gene copies following co-culture. In fact, in the case of Φ10750.3/Φ427.3, it appeared that neither prophage induction nor replication were necessary to achieve this effect. This was exemplified by the comparable levels of transcription attainable by $M4_{complete}$ and $M4_{degraded}$, the latter harbouring cryptic elements. The finding that HTE cells have the capacity to enhance the expression of *ssa* in both genotype *emm*3 and *emm*12, as well as in *emm*4, is reassuring and novel.

PCR on co-culture DNA extracts demonstrated that while $M4_{complete}$ prophage Φ10750.1, Φ10750.2, and Φ10750.3 could be induced in RPMI +10% (v/v) FCS alone, or on co-culture with HTE cells, the veracity of $M4_{degraded}$ Φ427.1, Φ427.2, and Φ427.3 as non-inducible, cryptic prophage was strengthened. Unexpected difficulties were experienced when applying the prophage conformation PCR to DNA samples extracted from co-culture experiments, possibly due to the failure to preclude eukaryotic DNA on extraction. However, PCR was successful for a number of $M4_{complete}$ and $M4_{degraded}$ isolates. Φ10750.1 induction under both conditions was comparable, as was induction of Φ10750.2. One $M4_{complete}$ isolate for which the Φ10750.3 phage induction PCR was successful, BSAC_bs468, suggested that induction was in fact stronger in RPMI +10% (v/v) FCS compared to co-culture conditions, whereas BSAC_bs1388 appeared to show the opposite, and demonstrated the greatest increase in *ssa* gene copies.

It had also been reported that human pharyngeal cells can enhance the expression of prophage-associated toxins, specifically *speC* and *spd1* (Broudy *et al.*, 2002, 2001), which as aforementioned, in the *emm*4 population, are encoded by Φ10750.1/ Φ427.1, and are among the canonical constellation of prophage in this genotype (Stephen B Beres *et al.*, 2006; Chalker *et al.*, 2017; Jacoba *et al.*, 2016; Turner *et al.*, 2019b, 2017). Expression of *speC* was in fact found to decrease following co-culture with HTE cells, and this was observed equally in both $M4_{complete}$ and $M4_{degraded}$, decreasing on average ~0.5 fold. The DNase *spd3* appeared to be neither expressed in RPMI +10% (v/v) FCS alone or following HTE co-culture, which was unexpected, as transcription at 3, 5 and 8 hours in THB was clearly detectable. This raises the possibility that *spd3* expression may be repressed by a component of RPMI +10% (v/v) FCS, perhaps FCS in particular. Indeed, all published *S. pyogenes* co-culture studies, of which, at the time of writing, there have been three (Broudy, Pancholi and Fischetti, 2001, 2002; Banks, Beres and Musser, 2003), have co-cultured in cell culture media without serum. . This may go some way to explain discrepancies in these data. However, it is also possible that *spd3* is not expressed early in colonisation and infection in response to exposure to the tonsillar epithelium or in the absence of other immunological or anatomical stimuli which are lacking in this model. It would be very interesting to explore this phenomenon further using whole tonsil explants, as has been achieved previously (Davies *et al.*, 2019).

In further support of an alternative hypothesis of a gene dosing effect in the elicitation of the enhancement of prophage-associated toxins, measurements of *ssa* gene copies confirmed that not only was induction of prophage

unnecessary to achieve an enhancement in transcription of the gene, but that replication of the element was neither necessary nor sufficient. This is a novel finding in the case of *ssa*-associated prophage and *emm*4, but corroborates existing literature that has demonstrated that a robust increase in toxin gene expression may be achieved, irrespective of, or indeed, in the absence of prophage induction, and may be associated with increased toxin production. Though presently, the literature is far from replete with studies investigating the role of prophage in toxin expression by *S. pyogenes*, the present work marries well with observations made of the *speA*-encoding *emm*3 prophage Φ315.5 following co-culture with human cells (Banks *et al.*, 2003). Unlike *speA*, however, *ssa* transcription does not appear to be under the constitutive influence of CovRS (Sumby *et al.*, 2006). However, there may be scope for hypotheses on the role of CovRS as a possible regulator of *ssa* expression and induction of the associated prophage, however. Indeed, two CovRS variant *emm*4 isolates exhibited a marked increase in *ssa* gene copies following co-culture, $M4_{complete}$ BSAC_bs1388 (CovS; 36aa-STOP), and $M4_{degraded}$ BSAC_1802 (CovR; Gly91Asp), despite not having responded appreciably to co-culture by enhancing *ssa* expression. While these data may simply be coincidental, there is some evidence that the transcriptional regulator Rgg/RopB may influence prophage excision, as demonstrated using an *emm*5 (Dmitriev *et al.*, 2006), but this has not been reported hitherto for CovRS, or indeed any other major regulators of *S. pyogenes*. The Rgg/RopB allele for BSAC_bs1388 and BSAC_bs1802 are identical to the other MGAS10750- and $M4_{degraded}$ isolates, respectively, and is not predicted to affect activity (Ikebe *et al.*, 2010).

It appears that at the transcriptional level, both $M4_{complete}$ and $M4_{degraded}$ *emm*4 are appreciably similar with regards to the expression of prophage-associated toxins under conditions designed to emulate the host-pathogen interface. The findings that HTE modulates the expression of the streptococcal superantigen *ssa* in clinical isolates belonging to at least three major *emm*-types is novel and thought provoking, and certainly warrants further study. Preliminary, inferential data which appear to indicate the influence of CovRS over prophage induction and replication would also constitute an exciting and hitherto unappreciated role for this major regulator in the infection biology of *S. pyogenes*, as would the involvement of CovRS in the expression of *ssa* at the host-pathogen interface.

6.5.1 **Chapter 6: Key findings**

- *S. pyogenes* isolates were found to modulate expression of superantigens following co-culture with HTE cells:
    - Expression of *speC* decreased following co-culture of *emm*4 isolates with HTE cells
    - Expression of *ssa* was enhanced following co-culture of *emm*4 isolates with HTE cells but this was not due to an increase in gene copies and appeared to be independent of prophage induction
- The enhancing effect on *ssa* expression was not specific to HTE cells and was also observed following co-culture of *emm*4 isolates with A549 cells
- Expression of *ssa* is not controlled by prophage induction and replication in *emm*4 isolates

- Genotype *emm*4 isolates with naturally occurring CovRS variant alleles exhibited substantially increased *ssa* gene copies following co-culture versus 'wildtype' *emm*4 isolates
- Genotype *emm*3 and *emm*12 isolates also exhibited an enhancement in *ssa* expression following co-culture with HTE cells
- Genotype *emm*4 isolates expressed more *ssa* than did *emm*3 and *emm*12 isolates following HTE co-culture

Chapter 7

7. Discussion

7.1 Genotype *emm*4 prophage

7.1.1 Two clades within clinical *emm*4 isolates

A small collection of clinical *emm*4 isolates collected by BSAC (Turner *et al.*, 2019b) were subject to WGS as part of prior work carried out by this laboratory. Preliminary analyses carried out by project supervisor Dr. Claire E. Turner revealed that these ten isolates broadly clustered evenly with completed reference genome MGAS10750 (Beres and Musser, 2007), or MEW427 (Jacoba *et al.*, 2016),. These data also suggested that those clustering with MEW427 had undergone multiple regions of recombination, which has in recent years been shown, not only to be a source of diversity in *S. pyogenes*, but also a mechanism by which successful new and unusual variants may emerge (Nasser *et al.*, 2014; Sumby *et al.*, 2005b; Turner *et al.*, 2019b, 2015), though not, at the time of writing, in *emm*4.

7.1.2 Prophage degradation within MEW427 clade

Comparison of the two completed reference genomes showed that these putative recombinant regions in fact represented variably sized regions of sequence loss within the prophages and SpyCI of MEW427, relative to the corresponding elements of MGAS10750. Prophage genes involved in prophage excision, regulation and replication, and to a lesser extent, tail fibre and capsid protein encoding genes, had been deleted or subject to marked sequence loss, in a similar pattern across the three prophage, albeit to varying extents. Φ427.1, Φ427.2, Φ427.3 having undergone ~78%, ~50%, and ~26% gene loss, respectively, relative to Φ10750.1, Φ10750.2, and Φ10750.3. Interestingly, the prophage-associated superantigen and DNase genes were retained in each case. Similar gene loss as observed in the ten *emm*4 BSAC isolates. Those clustering with MEW427 exhibited gene loss in Φ427.1, Φ427.2, Φ427.3, and SpyCIM4. BSAC isolates associated with the MGAS10750 clade carried near-identical, full-length copies of these elements. For Φ427.2 and Φ427.3, the integrase genes, required for prophage excision/integration, were truncated and predicted to be non-functional. Φ427.1 carried a full-length and presumably functional integrase, but was subject to such pronounced gene loss elsewhere on the element, that were it able to excise, it would be unable to form progeny virions. Degraded, elements are described in the literature, including occasionally in *S. pyogenes* (Canchaya *et al.*, 2003, 2002; Green *et al.*, 2005; Turner *et al.*, 2015), and recent a recent genomic survey suggests that such elements are common components of the streptococcal genome (Rezaei Javan *et al.*, 2019).

7.1.3 $M4_{degraded}$ prophage are cryptic

Prophage induction was probed among the $M4_{complete}$ and $M4_{degraded}$ BSAC isolates, cultured under a variety of conditions over the course of the project, including exponential growth, following treatment with mitomycin C, in RPMI +10% (v/v) FCS, and following co-culture with HTE cells. Full-length prophage of $M4_{complete}$ isolates were induced under all conditions tested, whereas those of $M4_{degraded}$ isolates, (with the exception of unusual $M4_{degraded}$ isolate BSAC_bs1802), were not. These immobile prophage of $M4_{degraded}$ isolates were thus designated cryptic prophage.

7.1.4 A novel international *emm*4 lineage characterised by cryptic prophage

Additional work to place the BSAC *emm*4 within an international context by project supervisor Dr. Claire E. Turner ratified initial hypotheses: *emm*4 isolates generally clustered with either MGAS10750 or MEW427 (Remmington *et al.* 2020). Crucially, this demonstrated that within the wider *emm*4 population, represented by publicly available WGS data for 233 international *emm*4 isolates, was a lineage broadly associated with pronounced gene loss within prophage- and chromosomal island-encoding regions (Remmington *et al.* 2020). Curiously, while the population of North America almost exclusively comprised $M4_{degraded}$ isolates, the UK population was more diverse, with both lineages apparently co-existing (Remmington *et al.* 2020). More powerful and largescale epidemiological and genomic studies would be warranted to investigate this inquisitive finding, which could not be explained by consultation of clinical data, where available.

There are a number of possibilities as to why prophage crypticism had occurred within $M4_{degraded}$ isolates. It was not possible to determine the temporal nature of the lineages emergence, but perhaps *emm*4 without inducible prophage are more resistant to environmental stressors, and prophage inducers, such as antibiotics (Allen *et al.*, 2011; Maiques *et al.*, 2006; Zhang *et al.*, 2000), extremes of pH (Miller-Ensminger *et al.*, 2020), host-derived factors such as reactive oxygen species (Henningham *et al.*, 2015), or even exposure to UV light and fluctuations in temperature (Yue *et al.*, 2012), as has been demonstrated in a litany of bacterial species harbouring cryptic prophage. One interesting study demonstrated that *S. pneumoniae*, through the production of hydrogen peroxide, was able to exploit the lysogenic nature of *Staphylococcus aureus* competitors by triggering induction of prophage and lysis of lysogens (Selva *et al.*, 2009). The primary anatomical niche of *S. pyogenes*, the human tonsil, is certainly shared with a diverse population of niche competitors, including *S. pneumoniae* (Adegbola *et al.*, 2014). Neutralising the imminent threat of lysis by inactivating or degraded prophage represents a logical hypothesis for the pervasive prophage decay observed in $M4_{degraded}$ isolates. Induction of *S. pyogenes* prophage has been demonstrated *in-vivo* using murine models of nasopharyngeal infection (Broudy and Fischetti, 2003), and as aforementioned, in response to a factor secreted by human cells (Banks *et al.*, 2003; Broudy *et al.*, 2002, 2001). The capacity for rescue of these cryptic elements by 'helper prophage', as is observed among satellite prophage (Rezaei Javan *et al.*, 2019), phage-inducible chromosomal islands (Penadés and Christie, 2015), are certainly precluded by the incapacity of these prophage to exit the genome. However, vertical transmission is virtually

assured, and perhaps cryptic prophage are still transmitted by recombination. A hypothesis for which there may be some support among contemporary *emm*89 isolates, which frequently (though not characteristically) harbour a degraded *speC*/*spd1* prophage that may have been acquired by recombination; ΦM89.1 (Turner *et al.*, 2015).

7.2 Redundancy of prophage mobility regulation among *emm*4 *S. pyogenes*

7.2.1 Prophage induction is not associated with enhanced toxin transcription

It had been hypothesised that the induction of prophage, and subsequent replication may elicit a gene dosing effect, providing a greater number of toxin gene templates for transcription, thus enhancing toxin production (gene duplication and amplification) as in other bacterial species (Andersson and Hughes, 2009; Sandegren and Andersson, 2009). Growth-phase experiments to detect spontaneous prophage induction and toxin transcription among $M4_{complete}$ isolates did not correlate with specific timepoints. Indeed, the expression of *ssa* and *spd3* was very similar between $M4_{complete}$ and $M4_{degraded}$ isolates, and did not vary by growth phase, regardless of the fact that prophage clearly inducible in the former, and demonstrably cryptic in the latter. Interestingly, *speC* expression was elevated at 8-hours growth in THB among both $M4_{complete}$ and $M4_{degraded}$. Despite this, when prophage induction and *speC* gene copies were measured, prior reports that induction of an undesignated M6 prophage enhanced SpeC and Spd1 expression (Broudy *et al.*, 2002, 2001) were not supported, and surprisingly, an increase in gene copies did not appear to correlate with prophage induction. Certainly, when cultured in THB, it seemed that transcription of prophage-associated toxin genes was irrespective of induction of the associated prophage.

7.2.2 Genotype *emm*4 DNA MMR operon is not regulated by $SpyCI^{M4}$

The SpyCI are, compared to prophage, vastly understudied atoxigenic viral parasites. At the time of writing, the only perceptible role of SpyCI are as growth-phase-dependent regulators of the DNA MMR operon (Nguyen and McShan, 2014; Scott *et al.*, 2012). The finding that SpyCIM4 of either lineage could not be induced during growth, or when cultured with mitomycin C, was therefore unanticipated, given that in $M4_{complete}$ isolates, SpyCIM4 carry the necessary genetic apparatus to permit mobility, while those of $M4_{degraded}$ isolates do not. Surprisingly, transcripts for both *mutL* and *mutS* were nevertheless abundantly detected for all isolates in cDNA samples synthesised from RNA extracted from *emm*4 isolates grown to the timepoint at which SpyCI was purportedly at its most stably integrated, and expression of *mutS* at its lowest (Scott *et al.*, 2012). It appears that in *emm*4, this operon is not regulated by the mobility of SpyCIM4. One study reports that experimental curing of SpyCI from *emm*1 isolates had dramatic consequences for global transcription, including a number of well-studied virulence factors; perhaps indicating that elements fulfil a hitherto unelucidated role in streptococcal biology (Hendrickson *et al.*, 2015).

7.3 Genotype *emm*4 at the host-pathogen interface

7.3.1 Phenotypic similarities between $M4_{complete}$ and $M4_{degraded}$

The phylogenetic distinction between $M4_{complete}$ and $M4_{degraded}$ isolates and the body of evidence attesting to the importance of prophage induction in toxin expression notwithstanding (Balasubramanian *et al.*, 2019; Broudy *et al.*, 2002, 2001; Ingrey *et al.*, 2003), a polarising phenotype could not be demonstrated in most cases between the two lineages. Indeed, it appeared that these isolates were rather similar in their capacity to survive in whole human blood, may be similarly superantigenic, and exhibit comparable capacity for adhesion.

One distinguishing characteristic of $M4_{degraded}$ was a significant attenuation in virulence of $M4_{degraded}$ isolates in the *G. mellonella* infection model. *G. mellonella* lack adaptive immunity, and are insensitive to streptococcal superantigens (Tsai *et al.*, 2016). However, the immune system of *G. mellonella* does include phagocytic haemocytes, which extrude chromatin-rich matrices akin to NETs (Altincicek *et al.*, 2008), which are sensitive to streptococcal DNases (Chalmers *et al.*, 2017). As such, this finding might be attributable to differences in DNases activity. Higher levels of *speC* expression, and thus, inferentially, the co-transcribed DNase *spd1*, were attainable by $M4_{complete}$, relative to $M4_{degraded}$ isolates after culture for 8-hours in THB, it was conceivable that DNases activity might explain this attenuation. However, chromosomally-encoded DNases might equally play a role. Intriguingly, the acquisition of an *speC*/*spd1*-encoding prophage, ΦUK-M3.1, by an epidemic genotype *emm*3 ST15 lineage in the United Kingdom (Al-Shahib *et al.*, 2016), and associated with enhanced DNase activity, was insufficient to confer enhanced virulence in *G. mellonella* (Afshar *et al.*, 2017). The chimeric *emm*4 allele that was identified in the present work, and in subsequent works (DebRoy *et al.*, 2018; Frost *et al.*, 2020), could also have some influence, by interfering with phagocytosis or perhaps other pathogenic mechanisms attributed to M and M-like proteins (Frost *et al.*, 2020; Smeesters *et al.*, 2010). M1 deletion mutants have previously exhibited attenuated virulence in the *G. mellonella* model, a phenotype that was reversed on complementation (Loh *et al.*, 2013). Other phenotypes that have been ascribed to emergent lineages, such as enhanced nasopharyngeal shedding (Afshar *et al.*, 2017), or enhanced colonisation and infection (Brouwer *et al.*, 2020; Kasper *et al.*, 2014). These cannot be observed in this crude model of systemic infection, which is intended to assay overall virulence, and is useful in that regard. However, it was clear from the nature of the BSAC *emm*4, which were all bloodstream isolates, that both lineages are able to cause severe disease and attain full virulence.

Non-synonymous SNPs identified in genes encoding fibronectin-binding proteins, and other lineage-specific SNPs did not affect adhesion of either $M4_{complete}$ or $M4_{degraded}$ isolates. Indeed, neither lineage appeared to have been imbued with enhanced adhesion and biofilm formation on polystyrene, fibronectin, type IV collagen, or fibrinogen, though interesting predilection for *emm*4 adhesion to the latter was observed, perhaps due to the acapsular nature of the genotype (Flores *et al.*, 2012). In *S. pneumoniae*, prophage induction may bolster biofilm

propagation (Carrolo *et al.*, 2010), but the capacity for $M4_{complete}$ lysogens to form biofilms through a similar mechanism was not observed.

7.3.2 **Human tonsillar epithelium modulates superantigen expression**

Previous studies have shown that a soluble factor secreted by human cells induces prophage and/or enhances expression of cognate virulence factors (Banks *et al.*, 2003; Broudy *et al.*, 2002, 2001). When *emm*4 were co-cultured with HTE cells, induction of Φ10750.1 among $M4_{complete}$ isolates was, though achievable, similar to induction in RPMI +10% (v/v) FCS alone. Φ427.1 of $M4_{degraded}$ were not induced under either culture condition. Despite the differential induction of these prophage, the impact of HTE co-culture was the same: *speC* transcript copies were significantly lower on co-culture than in RPMI +10% (v/v) FCS alone. This contrasts with seminal literature on the topic which suggest that co-culture with pharyngeal cells induces prophage and enhances toxin expression (Broudy *et al.*, 2002, 2001). Substantial differences in the methodologies implemented therein and herein notwithstanding, it did not appear that HTE cells increased expression of *speC*, nor did they appear to appreciably induce $M4_{complete}$ prophage Φ10750.1; indeed, *speC* expression decreased in spite of prophage induction. Induction of Φ10750.2 did not increase following HTE co-culture either, and again, Φ427.2 was shown to be cryptic. Transcripts of *spd3* could not be detected in isolates of either lineage in RPMI +10% (v/v) FCS or on exposure to HTE cells, which lent support to the revised hypothesis that induction may be inconsequential to the transcription of cognate toxin genes.

Induction of Φ10750.3 in co-culture experiments were conflicting. One $M4_{complete}$ isolate exhibited strong induction in response to co-culture, while another the opposite. Φ427.3 was not induced under either co-culture conditions, as expected. However, co-culture with HTE cells elicited a significant enhancement in transcript copies of *ssa* in both $M4_{complete}$ and $M4_{degraded}$ isolates alike, and again, this response appeared irrespective of the capacity for Φ10750.3/Φ427.3 to induce. Measurements of *ssa* gene copies demonstrated that the enhancement in *ssa* transcription was both independent of induction of the associated element, and subsequent prophage replication; gene dosage. Experimentation with *emm*3 and *emm*12 isolates showed that HTE-mediated enhancement of *ssa* was not genotype specific, which was reassuring. While fold-change in *ssa* transcript copies was similar between *emm*3 and *emm*12, baseline expression and expression following co-culture were higher in both *emm*4 and *emm*12 relative to *emm*3, perhaps due to homology between Φ10750.3 and ΦHKU.ssa. Indeed, ΦHKU.ssa is a chimera of Φ10750.3 and the *emm*12 Φ9429.1, which carries *speC*/*spd1*, and although in *emm*4, toxin expression appears irrespective of prophage induction, these phage may carry other genes that might influence expression. ΦHKU.ssa does not carry *speC*/*spd1*, instead the lysogenic conversion module is near-identical to that of Φ10750.3, as is the distal lysogeny module; ΦHKU.ssa in turn exhibits appreciable homology with Φ315.2 (Figure 7.1). Unlike both ΦHKU.ssa and Φ10750.3/Φ427.3, which integrate at the same 5' *attB*, the DNA-binding protein HU (Stephen B. Beres *et al.*, 2006; Davies *et al.*, 2015), the Φ315.2 *attB* is a 3' tmRNA

(Beres *et al.*, 2002), which could perhaps impact how the expression the toxin gene carried by the this phage is controlled.

There is of course scope for the regulation of prophage toxins by bacterial regulatory networks, which may also contribute to variation between genotypes (Anbalagan and Chaussee, 2013; Sumby *et al.*, 2006). Although the present work raises questions, it marries well with descriptions of co-culture experiments using an MGAS315 *emm*3 strain with pharyngeal cells (Banks *et al.*, 2003). This study reported that co-culture neither triggered the induction of the *speA*-associated Φ315.5, nor was this necessary to enhance the transcription of the superantigen. The authors went on to demonstrate greater quantities of SpeA in supernatant filtrates from co-culture conditions, relative to those from *S. pyogenes* grown in DMEM alone. It seems plausible that the expression of *ssa* may follow a similar theme, though among genotype *emm*1 isolates, *speA* expression is regulated, at least in part, by bacterial regulatory networks, notably CovRS and RofA (Lynskey *et al.*, 2019a; Sumby *et al.*, 2006). Indeed, there was evidence in the present work that CovRS may play a role in the regulatory networks which mediates the transcriptional response of *ssa* following expose to HTE cells, and induction of the associated prophage. Additional prophage-associated regulatory mechanisms may also yet be discovered.

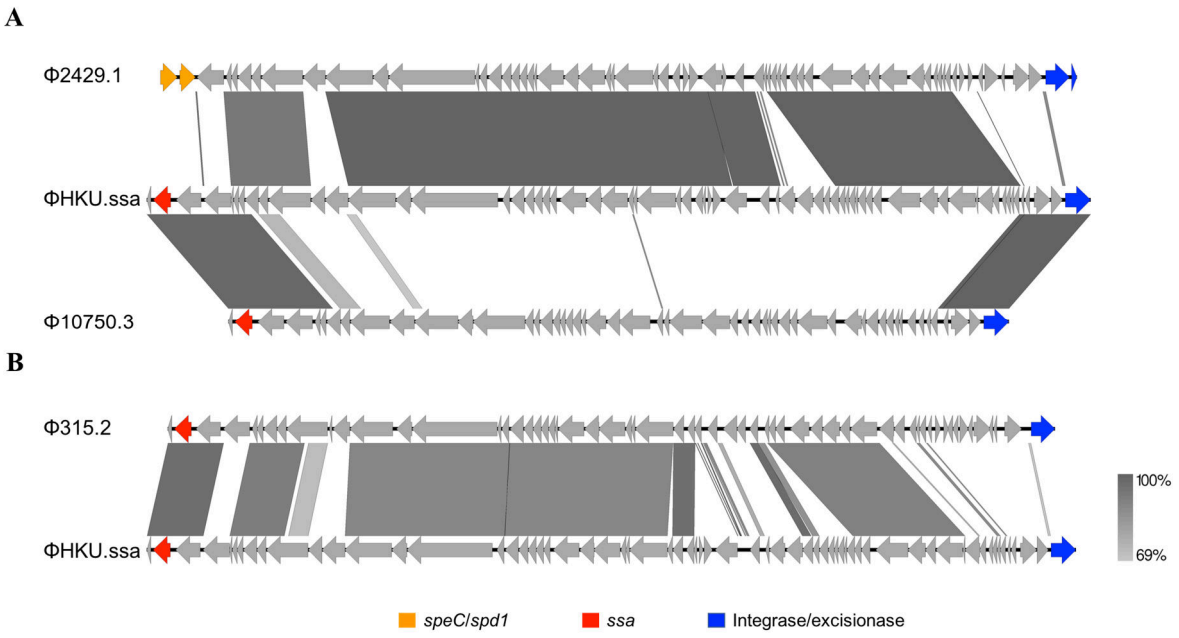

Figure 7.1: Homology between toxigenic prophage of *emm*3, *emm*4 and *emm*12

A) Genotype *emm*12 prophage ΦHKU.ssa is a chimera of *speC/spd1*-associated *emm*12 prophage Φ2429.1 and *ssa*-associated *emm*4 prophage Φ10750.3. The majority of the sequence is Φ2429.1-derived. However, flanking lysogenic conversion and lysogeny modules of ΦHKU.ssa are near-identical to those of Φ10750.3. B) Genotype *emm*3 prophage Φ315.2 shares considerable homology with ΦHKU.ssa, notably the lysogenic conversion module, including *ssa*. Genes are colour coded per figure key: orange; *speC/spd1*, red; *ssa*, navy; integrase and excisionase genes. Remaining genes are shown grey. Shaded scale bar indicates % DNA sequence similarity. Prophage sequences were extracted from publicly available completed reference genomes (Stephen B. Beres *et al.*, 2006; Davies *et al.*, 2015) and compared using EasyFig (Sullivan *et al.*, 2011).

7.4 Bacterial regulation of prophage induction

7.4.1 Possible role for CovRS in prophage replication and toxin expression

Expected variation between genotypes in the control of gene expression by major regulators of *S. pyogenes* notwithstanding, isolates carrying naturally occurring CovS variants BSAC_bs468 (Thr241Pro and Ala379Thr) and BSAC_bs1388 (36aa-STOP), and CovR variant isolate BSAC_bs1802 (Gly91Asp), did not exhibit major differential transcription of *speC* or *spd3* under any conditions tested, nor did these variants appear phenotypically different to other *emm*4 over the course of the project. While *speC* and *ssa* are, at the time of writing, not known to be under the influence of this regulator, a genotype *emm*1 isolate was previously shown to highly express Spd3 following CovRS mutation (Sumby *et al.*, 2006). This was not observed in these isolates. Gene copies of *speC* and *ssa* were measured as part of two experiments; growth in THB and following HTE co-culture, respectively. BSAC_1388 and BSAC_bs1802 exhibited dramatic increases in gene copies in both experiments, compared to other *emm*4. This was not observed for the vast majority of the other *emm*4, apart from BSAC_bs1349, which did exhibit an increase in *speC* gene copies after 8 hours growth in THB, but not in co-culture experiments. These dramatic increases in gene copy number did not appear, however, to correlate with prophage induction at the corresponding timepoint of 8-hours growth in THB, nor following co-culture for either BSAC_bs1388 and BSAC_bs1802. It is possible that this is due to the induction of these prophage and the formation of prophage particles, which were subsequently extracted with the rest of the bacterial culture in both experiments. It was particularly interesting to note that despite having undergone marked replication, transcription of *ssa* following HTE co-culture for these two isolates did not change appreciably, relative to RPMI +10% (v/v) FCS alone, and certainly did not exhibit as robust of an enhancement in *ssa* expression as did the other *emm*4. This could certainly explain why the gene copies of *speC* and *ssa* in either case were so disproportionately in excess of *gyrA* gene copies, to which they were normalised. The creation of isogenic mutants would have verified these observations, but were unfortunately not successful, but efforts to investigate these observations are ongoing.

7.5 Conclusions

The relationship between prophage mobility and toxin expression seems substantially more complex than was anticipated, and at the host-pathogen interface, the mediators and regulators of these events are largely unknown. Perhaps unsurprisingly, considering the human-restricted nature of *S. pyogenes*, there appears an important role for host-derived factors, which have the potential to impact virulence factor expression both at transcriptional and post-translational levels (Banks *et al.*, 2003; Broudy *et al.*, 2002, 2001; Brouwer *et al.*, 2020). Given their prevalence and apparent epidemiological importance (Al-Shahib *et al.*, 2016; Stephen B. Beres *et al.*, 2004; Davies *et al.*, 2015; Tse *et al.*, 2012), a better understanding of toxigenic prophage and their contribution to the biology of *S. pyogenes* might greatly improve our efforts to understand, monitor, prevent and control this highly lysogenised, highly adaptable and notoriously protean pathogen.

7.1 Future work

The enhancement of *ssa* transcription in three major *emm* types by HTE cells is novel and exciting. This work requires validation to determine the impact these transcriptional events have on protein expression. Principally, this would involve the recombinant expression and purification of SSA, and the generation of an antibody to quantitate SSA by Western blot. Protein activity assays such as the highly sensitive tritiated thymidine incorporation assay may be necessary, as SSA is poorly mitogenic compared to other streptococcal superantigens such as SpeC and SMEZ (Afshar *et al.*, 2017). The superantigenicity of this superantigen has very recently been measured using IL-2 as a marker of lymphocyte proliferation, and may prove valuable (Brouwer *et al.*, 2020). The apparent thiol-mediated regulation of SSA activity described very recently should be considered for future work with this superantigen, particularly *in-vitro* experiments (Brouwer *et al.*, 2020).

In the context of host-pathogen interactions, the factor secreted by human cells that elicited the effects on toxin expression described herein and elsewhere (Banks *et al.*, 2003; Broudy *et al.*, 2002) remains unknown, and the identification of this factor and its mechanism of action would be a tantalising and impactful discovery. The present work determined that the *ssa* inducing factor was not specific to HTE cells, and this was demonstrated similarly in preliminary experiments with A549 cells. Prior publications have used pharyngeal cells, and a breast cancer cell line, thus, taken together, it seems clear that the inducer is not tonsil specific (Broudy *et al.*, 2001). However, it would be interesting to explore with immortalised or primary cell lines the host restriction of this effect using zoonotic streptococci such as *S. dysgalactiae* sbsp. *equisimilis*, which occasionally also appear to carry *ssa* (Igwe *et al.*, 2003), and myriad other superantigens also present in *S. pyogenes* and the equine restricted *S. equi* sbsp. *equi* (Holden *et al.*, 2009; Turner *et al.*, 2019b). It would be also valuable and insightful to explore whether these observations are recapitulated using whole human tonsil explants, as has been achieved previously (Davies *et al.*, 2019), and may offer a more faithful reproduction of the anatomy of the tonsil and its associated immunological components.

Animal models may also prove insightful in investigating the attenuated virulence associated with $M4_{degraded}$ *emm*4, though could be, in equal measure, both avoided or complimented by epidemiological analyses and NHS Demographic Batch Tracing, where possible, of patients from which *emm*4 are isolated, as has been achieved previously (Afshar *et al.*, 2017). More powerful genomic analyses, such as BEAST analyses, may prove insightful in identifying the temporal nature of $M4_{degraded}$ emergence. The latter, however, may be due to complex sociological factors such a prescribing practice, the delineation of which may not be feasible or conclusive.

If ratified, hypotheses on CovRS and potential influence on prophage induction and replication would be impactful, and constitute an exciting and hitherto unappreciated role for this well-studied major regulator. There was also evidence in the present work that CovRS may play a role in the regulation of *ssa* expression at the host-pathogen interface. It would be particularly interesting to investigate whether this mechanism, if proven, extends

to prophage that encode toxins which are known to be under its regulatory control in other genotypes, such as *speA* (Sumby *et al.*, 2006), and could extend to other bacterial regulators of prophage-encoded toxins, such as RofA and Rgg1/RopB (Dmitriev *et al.*, 2006; Lynskey *et al.*, 2019a). Efforts to investigate the relationship between CovRS and the transcriptional response of *ssa* among *emm*4 isolate following exposure to HTE cells are ongoing. This response may also involve other regulatory networks and other genes, however, and the creation of a transposon mutant library and repeating co-culture experiments using a more high-throughput format may prove illuminating.

The benefits, if indeed there are any, of harbouring degraded prophage in *emm*4 *S pyogenes* are unclear, and the emergence of $M4_{degraded}$ may be due to a number of anatomical, immunological, pharmacological, epidemiological and sociological factors, or indeed the combination of multiple variables. The clear capacity for both *emm*4 lineages to co-exist in the UK suggested that $M4_{degraded}$ were at least as fit as $M4_{complete}$, though the relative paucity of the former in North America is intriguing, and may offer clues on the emergence of this lineage. Data in the present work suggested that toxin expression was very similar between both lineages, and may be irrespective of prophage. $M4_{degraded}$ were also found to be less virulent than $M4_{complete}$. Perhaps $M4_{degraded}$ isolates are therefore better suited to colonisation or longer-term carriage, resilience to exogenous antibacterial factors *in-vivo*, or tolerance to geographically variable prescribing practices, as opposed to achieving success by more obvious and reductive notions of infection and pathogenesis (
Figure 7.2).

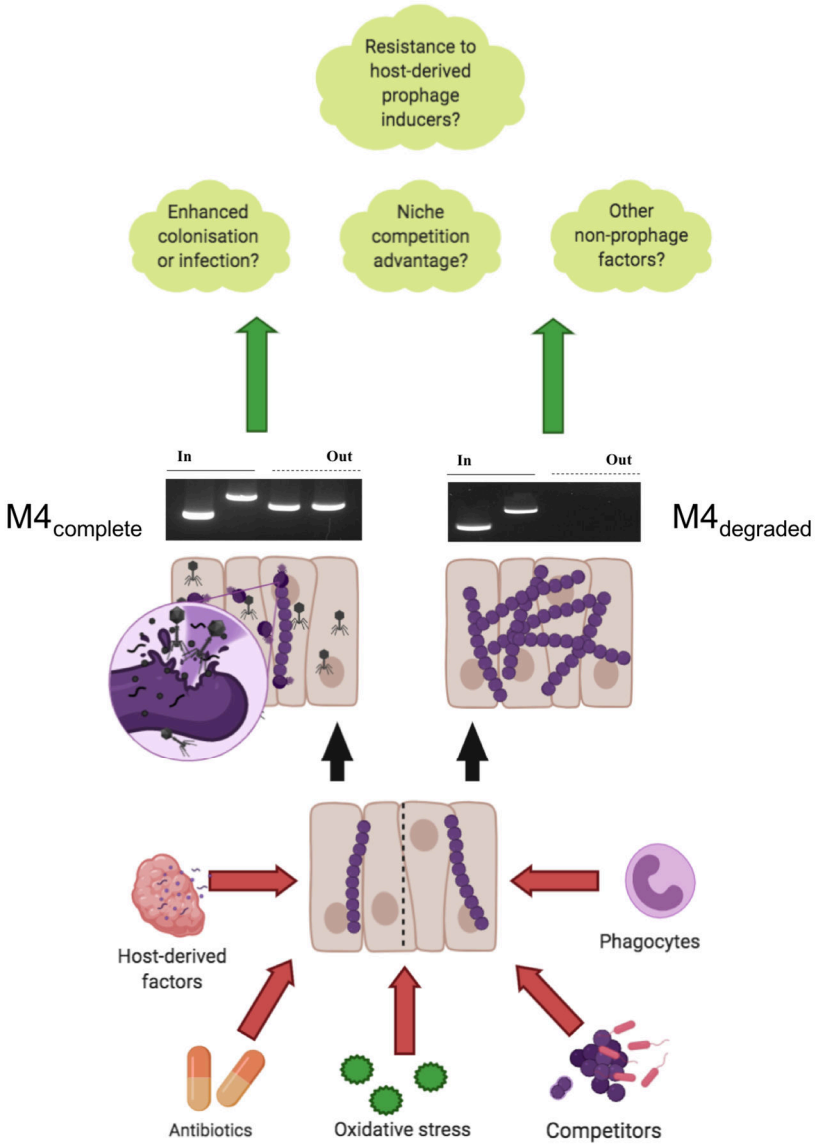

Figure 7.2: Graphical hypotheses: $M4_{complete}$ vs. $M4_{degraded}$

Aside from the enhanced virulence of $M4_{complete}$ isolates relative to $M4_{degraded}$ isolates, the two lineages appeared to be very similar. Factors contributing to the emergence of $M4_{degraded}$ are as yet unknown, but may reflect a number of different exogenous factors such as prescribing practices, approaches to infection, and niche competition, and may be mediated by prophage domestication. Further phenotypes may become apparent in more complex models of infection, and following more powerful epidemiological and genomic analyses than were available in the present work. Figure created by the author with BioRender.com.

8. Appendix

Publications

Remmington, A., & Turner, C. E. (2018). The DNases of pathogenic Lancefield streptococci. *Microbiology. 164*(3), 242–250. https://doi.org/10.1099/mic.0.000612

Remmington, A., Haywood, S., Edgar, J., Green, L. R., de Silva, T., & Turner, C. E. (2020). Cryptic prophages within a *Streptococcus pyogenes* genotype *emm*4 lineage. *Microbial genomics*, 10.1099/mgen.0.000482. Advance online publication. https://doi.org/10.1099/mgen.0.000482

Posters and presentations

Remmington, A., Turner, C. E. "Identification of a lineage within the global population of *Streptococcus pyogenes* genotype *emm*4 characterized by prophage degradation", in *Streptococcal Biology: Molecular Mechanisms at the Streptococcal-Host Interface, Gordon Research Conference* (2018) Newry, ME, USA. (Poster)

Remmington, A., Turner, C. E. "Identification of a lineage within the global population of *Streptococcus pyogenes* genotype *emm*4 characterized by prophage degradation", in *Microbiology Society Annual Conference* (2018) Birmingham, UK. (Poster).

Remmington, A., Turner, C. E. "*Streptococcus pyogenes*: Investigating Unusual Phenotypes", in *The Drugs Don't Work Symposium* (2017) Sheffield, UK. (Poster)

9. References

Abedon, S.T., LeJeune, J.T., 2005. Why Bacteriophage Encode Exotoxins and other Virulence Factors. Evolutionary Bioinformatics 1, 117693430500100. https://doi.org/10.1177/117693430500100001

Ackermann, H.W., Prangishvili, D., 2012. Prokaryote viruses studied by electron microscopy. Archives of Virology 157, 1843–1849. https://doi.org/10.1007/s00705-012-1383-y

Adegbola, R.A., DeAntonio, R., Hill, P.C., Roca, A., Usuf, E., Hoet, B., Greenwood, B.M., 2014. Carriage of *Streptococcus pneumoniae* and other respiratory bacterial pathogens in low and lower-middle income countries: a systematic review and meta-analysis. PloS one 9, e103293. https://doi.org/10.1371/journal.pone.0103293

Aertsen, A., Faster, D., Michiels, C.W., 2005. Induction of Shiga toxin-converting prophage in *Escherichia coli* by high hydrostatic pressure. Applied and Environmental Microbiology 71, 1155–1162. https://doi.org/10.1128/AEM.71.3.1155-1162.2005

Afshar, B., Turner, C.E., Lamagni, T.L., Smith, K.C., Al-Shahib, A., Underwood, A., Holden, M.T.G., Efstratiou, A., Sriskandan, S., 2017. Enhanced nasopharyngeal infection and shedding associated with an epidemic lineage of emm3 Group A *Streptococcus*. Virulence 8, 1390–1400. https://doi.org/10.1080/21505594.2017.1325070

Agniswamy, J., Lei, B., Musser, J.M., Sun, P.D., 2004. Insight of host immune evasion mediated by two variants of Group A *Streptococcus* Mac protein. Journal of Biological Chemistry 279, 52789–52796. https://doi.org/10.1074/jbc.M410698200

Åkessont, P., Sjöholm, A.G., Björck, L., 1996. Protein SIC, a novel extracellular protein of *Streptococcus pyogenes* interfering with complement function. Journal of Biological Chemistry 271, 1081–1088. https://doi.org/10.1074/jbc.271.2.1081

Aksyuk, A.A., Rossmann, M.G., 2011. Bacteriophage assembly. Viruses 3, 172–203. https://doi.org/10.3390/v3030172

Al-Shahib, A., Underwood, A., Afshar, B., Turner, C.E., Lamagni, T., Sriskandan, S., Efstratiou, A., 2016. Emergence of a novel lineage containing a prophage in emm/M3 Group A *Streptococcus* associated with upsurge in invasive disease in the UK. Microbial Genomics 2, 1–9. https://doi.org/10.1099/mgen.0.000059

Alasmari, S.Z., 2020. In Vivo Imaging of Neutrophil Extracellular Traps (NETs): Visualization Methods and Outcomes. BioMed Research International, 1–6. https://doi.org/10.1155/2020/4192745

Alexeeva, S., Guerra Martínez, J.A., Spus, M., Smid, E.J., 2018. Spontaneously induced prophages are abundant in a naturally evolved bacterial starter culture and deliver competitive advantage to the host. BMC Microbiology 18, 1–16. https://doi.org/10.1186/s12866-018-1229-1

Allen-Hoffmann, B.L., Rheinwald, J.G., 1984. Polycyclic aromatic hydrocarbon mutagenesis of human epidermal keratinocytes in culture. Proceedings of the National Academy of Sciences of the United States of America 81, 7802–7806. https://doi.org/10.1073/pnas.81.24.7802

Allen, H.K., Looft, T., Bayles, D.O., Humphrey, S., Levine, U.Y., Alt, D., Stanton, T.B., 2011. Antibiotics in

feed induce prophages in swine fecal microbiomes. mBio 2, 1–9. https://doi.org/10.1128/mBio.00260-11

Allhorn, M., Olsén, A., Collin, M., 2008. EndoS from *Streptococcus pyogenes* is hydrolyzed by the cysteine proteinase SpeB and requires glutamic acid 235 and tryptophans for IgG glycan-hydrolyzing activity. BMC Microbiology 8, 1–10. https://doi.org/10.1186/1471-2180-8-3

Altincicek, B., Stötzel, S., Wygrecka, M., Preissner, K.T., Vilcinskas, A., 2008. Host-Derived Extracellular Nucleic Acids Enhance Innate Immune Responses, Induce Coagulation, and Prolong Survival upon Infection in Insects. The Journal of Immunology 181, 2705–2712. https://doi.org/10.4049/jimmunol.181.4.2705

Anbalagan, S., Chaussee, M.S., 2013. Transcriptional Regulation of a Bacteriophage Encoded Extracellular DNase (Spd-3) by Rgg in *Streptococcus pyogenes*. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0061312

Andersson, D.I., Hughes, D., 2009. Gene amplification and adaptive evolution in bacteria. Annual Review of Genetics 43, 167–195. https://doi.org/10.1146/annurev-genet-102108-134805

Areschoug, T., Waldemarsson, J., Gordon, S., 2008. Evasion of macrophage scavenger receptor A-mediated recognition by pathogenic streptococci. European Journal of Immunology 38, 3068–3079. https://doi.org/10.1002/eji.200838457

Armitage, E.P., Senghore, E., Darboe, S., Barry, M., Camara, J., Bah, S., Marks, M., Cerami, C., Roca, A., Antonio, M., Turner, C.E., de Silva, T.I., 2019. High burden and seasonal variation of paediatric scabies and pyoderma prevalence in the Gambia: A cross-sectional study. PLoS Neglected Tropical Diseases 13, 1–16. https://doi.org/10.1371/journal.pntd.0007801

Ashbaugh, C. D., Alberti, S., Wessels, M.R., 1998. Molecular analysis of the capsule gene region of Group A *Streptococcus*: The hasAB genes are sufficient for capsule expression. Journal of Bacteriology 180, 4955–4959. https://doi.org/10.1128/jb.180.18.4955-4959.1998

Ashbaugh, C D., Warren, H.B., Carey, V.J., Wessels, M.R., 1998. Molecular analysis of the role of the Group A Streptococcal cysteine protease, hyaluronic acid capsule, and M protein in a murine model of human invasive soft-tissue infection. Journal of Clinical Investigation 102, 550–560. https://doi.org/10.1172/JCI3065

Ashbaugh, C.D., Wessels, M.R., 2001. Absence of a cysteine protease effect on bacterial virulence in two murine models of human invasive Group A streptococcal infection. Infection and Immunity 69, 6683–6688. https://doi.org/10.1128/IAI.69.11.6683-6686.2001

Ashwinkumar Subramenium, G., Viszwapriya, D., Iyer, P.M., Balamurugan, K., Karutha Pandian, S., 2015. covR Mediated Antibiofilm Activity of 3-Furancarboxaldehyde Increases the Virulence of Group A *Streptococcus*. PLOS ONE 10, 1–19. https://doi.org/10.1371/journal.pone.0127210

Athey, T.B.T., Teatero, S., Li, A., Marchand-Austin, A., Beall, B.W., Fittipaldia, N., 2014. Deriving Group A *Streptococcus* typing information from short-read whole-genome sequencing data. Journal of Clinical Microbiology 52, 1871–1876. https://doi.org/10.1128/JCM.00029-14

Athey, T.B.T., Teatero, S., Sieswerda, L.E., Gubbay, J.B., Marchand-Austin, A., Li, A., Wasserscheid, J.,

Dewar, K., McGeer, A., Williams, D., Fittipaldi, N., 2016. High incidence of invasive Group A *Streptococcus* disease caused by strains of uncommon emm types in Thunder Bay, Ontario, Canada. Journal of Clinical Microbiology 54, 83–92. https://doi.org/10.1128/JCM.02201-15

Aziz, R.K., Edwards, R.A., Taylor, W.W., Low, D.E., McGeer, A., Kotb, M., 2005. Mosaic Prophages with Horizontally Acquired Genes Account for the Emergence and Diversification of the Globally Disseminated M1T1 Clone of *Streptococcus pyogenes*. Journal of Bacteriology 187, 3311 LP – 3318. https://doi.org/10.1128/JB.187.10.3311-3318.2005

Bachelerie, F., Ben-Baruch, A., Burkhardt, A.M., Combadiere, C., Farber, J.M., Graham, G.J., Horuk, R., Sparre-Ulrich, A.H., Locati, M., Luster, A.D., Mantovani, A., Matsushima, K., Murphy, P.M., Nibbs, R., Nomiyama, H., Power, C.A., Proudfoot, A.E.I., Rosenkilde, M.M., Rot, A., Sozzani, S., Thelen, M., Yoshie, O., Zlotnik, A., 2014. International union of pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. Pharmacological Reviews 66, 1–79. https://doi.org/10.1124/pr.113.007724

Balasubramanian, S., Osburne, M.S., BrinJones, H., Tai, A.K., Leong, J.M., 2019. Prophage induction, but not production of phage particles, is required for lethal disease in a microbiome-replete murine model of enterohemorrhagic E. *coli* infection. PLoS Pathogens 15, 1–27. https://doi.org/10.1371/journal.ppat.1007494

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A.A., Dvorkin, M., Kulikov, A.S., Lesin, V.M., Nikolenko, S.I., Pham, S., Prjibelski, A.D., Pyshkin, A. V., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M.A., Pevzner, P.A., 2012. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. Journal of Computational Biology 19, 455–477. https://doi.org/10.1089/cmb.2012.0021

Banks, D.J., Beres, S.B., Musser, J.M., 2002. The fundamental contribution of phages to GAS evolution, genome diversification and strain emergence. Trends in Microbiology 10, 515–521. https://doi.org/10.1016/S0966-842X(02)02461-7

Banks, D.J., Lei, B., Musser, J.M., 2016. Prophage Induction and Expression of Prophage-Encoded Virulence Factors in Group A *Streptococcus* Serotype M3 Strain MGAS315 71, 7079–7086. https://doi.org/10.1128/IAI.71.12.7079

Barendt, S.M., Sham, L.-T., Winkler, M.E., 2011. Characterization of mutants deficient in the L,D-carboxypeptidase (DacB) and WalRK (VicRK) regulon, involved in peptidoglycan maturation of *Streptococcus pneumoniae* serotype 2 strain D39. Journal of bacteriology 193, 2290–2300. https://doi.org/10.1128/JB.01555-10

Barnett, T.C., Bowen, A.C., Carapetis, J.R., 2019. The fall and rise of Group A *Streptococcus* diseases. Epidemiology and Infection 147, 1–6. https://doi.org/10.1017/S0950268818002285

Basetti, S., Hodgson, J., Rawson, T.M., Majeed, A., 2017. Scarlet fever: A guide for general practitioners. London Journal of Primary Care 9, 77–79. https://doi.org/10.1080/17571472.2017.1365677

Bastiat-Sempe, B., Love, J.F., Lomayesva, N., Wessels, M.R., 2014. Streptolysin O and NAD-glycohydrolase prevent phagolysosome acidification and promote Group A *Streptococcus* survival in macrophages. mBio

5, 1–11. https://doi.org/10.1128/mBio.01690-14

Beachey, E.H., Bronze, M., Dale, J.B., Kraus, W., Poirier, T., Sargent, S., 1988. Protective and autoimmune epitopes of streptococcal M proteins. Vaccine 6, 192–196. https://doi.org/10.1016/S0264-410X(88)80027-6

Beall, B., Gherardi, G., Lovgren, M., Facklam, R.R., Forwick, B.A., Tyrrell, G.J., 2000. emm and sof gene sequence variation in relation to serological typing of opacity-factor-positive Group A streptococci. Microbiology 146, 1195–1209. https://doi.org/10.1099/00221287-146-5-1195

Beckert, S., Kreikemeyer, B., Podbielski, A., 2001. Group A streptococcal rofA gene is involved in the control of several virulence genes and eukaryotic cell attachment and internalization. Infection and Immunity 69, 534–537. https://doi.org/10.1128/IAI.69.1.534-537.2001

Bednarz, M., Halliday, J.A., Herman, C., Golding, I., 2014. Revisiting bistability in the lysis/lysogeny circuit of bacteriophage lambda. PLoS ONE 9. https://doi.org/10.1371/journal.pone.0100876

Ben Zakour, N.L., Davies, M.R., You, Y., Chen, J.H.K., Forde, B.M., Stanton-Cook, M., Yang, R., Cui, Y., Barnett, T.C., Venturini, C., Ong, C.L.Y., Tse, H., Dougan, G., Zhang, J., Yuen, K.Y., Beatson, S.A., Walker, M.J., 2015. Transfer of scarlet fever-associated elements into the Group A *Streptococcus* M1T1 clone. Scientific Reports 5, 1–7. https://doi.org/10.1038/srep15877

Berdygulova, Z., Westblade, L.F., Florens, L., Koonin, E. V., Chait, B.T., Ramanculov, E., Washburn, M.P., Darst, S.A., Severinov, K., Minakhin, L., 2011. Temporal regulation of gene expression of the thermus *thermophilus* bacteriophage P23-45. Journal of Molecular Biology 405, 125–142. https://doi.org/10.1016/j.jmb.2010.10.049

Beres, S.B., Musser, J.M., 2007. Contribution of exogenous genetic elements to the Group A *Streptococcus* metagenome. PLoS ONE 2. https://doi.org/10.1371/journal.pone.0000800

Beres, S.B., Richter, E.W., Nagiec, M.J., Sumby, P., Porcella, S.F., DeLeo, F.R., Musser, J.M., 2006. Molecular genetic anatomy of inter- and intraserotype variation in the human bacterial pathogen Group A *Streptococcus*. Proceedings of the National Academy of Sciences of the United States of America 103, 7059–7064. https://doi.org/10.1073/pnas.0510279103

Beres, S.B., Sylva, G.L., Barbian, K.D., Lei, B., Hoff, J.S., Mammarella, N.D., Liu, M.Y., Smoot, J.C., Porcella, S.F., Parkins, L.D., Campbell, D.S., Smith, T.M., McCormick, J.K., Leung, D.Y.M., Schlievert, P.M., Musser, J.M., 2002. Genome sequence of a serotype M3 strain of Group A *Streptococcus*: Phage-encoded toxins, the high-virulence phenotype, and clone emergence. Proceedings of the National Academy of Sciences of the United States of America 99, 10078–10083. https://doi.org/10.1073/pnas.152298499

Beres, S.B., Sylva, G.L., Sturdevant, D.E., Granville, C.N., Liu, M., Ricklefs, S.M., Whitney, A.R., Parkins, L.D., Hoe, N.P., Adams, G.J., Low, D.E., DeLeo, F.R., McGeer, A., Musser, J.M., 2004. Genome-wide molecular dissection of serotype M3 Group A *Streptococcus* strains causing two epidemics of invasive infections. Proceedings of the National Academy of Sciences of the United States of America 101, 11833–11838. https://doi.org/10.1073/pnas.0404163101

Berkower, C., Ravins, M., Moses, A.E., Hanski, E., 1999. Expression of different Group A streptococcal M

proteins in an isogenic background demonstrates diversity in adherence to and invasion of eukaryotic cells. Molecular Microbiology 31, 1463–1475. https://doi.org/10.1046/j.1365-2958.1999.01289.x

Bessen, D.E., Jones, K.F., Fischetti, V.A., 1989. Evidence for two distinct classes of streptococcal M protein and their relationship to rheumatic fever. Journal of Experimental Medicine 169, 269–283. https://doi.org/10.1084/jem.169.1.269

Bessen, D.E., Sotir, C.M., Readdy, T.L., Hollingshead, S.K., 1996. Genetic Correlates of Throat and Skin Isolates of Group A Streptococci. The Journal of Infectious Diseases 173, 896–900. https://doi.org/10.1093/infdis/173.4.896

Bisno, A.L., Stevens, D.L., 1996. Streptococcal infections of skin and soft tissues. The New England journal of medicine 334, 240–245. https://doi.org/10.1056/NEJM199601253340407

Bisno, A.L., 1991. Group A streptococcal infections and acute rheumatic fever. N Engl J Med 325, 293–302. https://doi.org/10.1186/1476-069X-12-14

Biswas, I., Scott, J.R., 2003. Identification of rocA, a positive regulator of covR expression in the Group A *Streptococcus*. Journal of Bacteriology 185, 3081–3090. https://doi.org/10.1128/JB.185.10.3081-3090.2003

Bobay, L.M., Touchon, M., Rocha, E.P.C., 2014. Pervasive domestication of defective prophages by bacteria. Proceedings of the National Academy of Sciences of the United States of America 111, 12127–12132. https://doi.org/10.1073/pnas.1405336111

Botstein, D., 1980. a Theory of Modular Evolution for Bacteriophages. Annals of the New York Academy of Sciences 354, 484–491. https://doi.org/10.1111/j.1749-6632.1980.tb27987.x

Boutry, C., Delplace, B., Clippe, A., Fontaine, L., Hols, P., 2013. SOS response activation and competence development are antagonistic mechanisms in *Streptococcus thermophilus*. Journal of Bacteriology 195, 696–707. https://doi.org/10.1128/JB.01605-12

Brandt, C.M., Spellerberg, B., 2009. Human Infections Due to *Streptococcus dysgalactiae* Subspecies *equisimilis* . Clinical Infectious Diseases 49, 766–772. https://doi.org/10.1086/605085

Brassil, B., Mores, C.R., Wolfe, A.J., Putonti, C., 2020. Characterization and spontaneous induction of urinary tract *Streptococcus anginosus* prophages. Journal of General Virology 101, 685–691. https://doi.org/10.1099/jgv.0.001407

Brenciani, A., Bacciaglia, A., Vignaroli, C., Pugnaloni, A., Varaldo, P.E., Giovanetti, E., 2010. Φm46.1, the main *Streptococcus pyogenes* element carrying *mef(A)* and *tet(O)* genes. Antimicrobial Agents and Chemotherapy 54, 221–229. https://doi.org/10.1128/AAC.00499-09

Brenciani, A., Ojo, K.K., Monachetti, A., Menzo, S., Roberts, M.C., Varaldo, P.E., Giovanetti, E., 2004. Distribution and molecular analysis of *mef(A)*-containing elements in tetracycline-susceptible and -resistant *Streptococcus pyogenes* clinical isolates with efflux-mediated erythromycin resistance. Journal of Antimicrobial Chemotherapy 54, 991–998. https://doi.org/10.1093/jac/dkh481

Bricker, A.L., Carey, V.J., Wessels, M.R., 2005. Role of NADase in virulence in experimental invasive Group A streptococcal infection. Infection and Immunity 73, 6562–6566. https://doi.org/10.1128/IAI.73.10.6562-

6566.2005

Brinkmann, V., Reichard, U., Goosmann, C., Fauler, B., Uhlemann, Y., Weiss, D.S., Weinrauch, Y., Zychlinsky, A., 2004. Neutrophil Extracellular Traps Kill Bacteria. Science 303, 1532–1535. https://doi.org/10.1126/science.1092385

Brockmann, S.O., Eichner, L., Eichner, M., 2018. Constantly high incidence of scarlet fever in Germany. The Lancet Infectious Diseases 18, 499–500. https://doi.org/10.1016/S1473-3099(18)30210-X

Brook, I., Gober, A.E., 2006. Recovery of interfering and β-lactamase-producing bacteria from Group A β-haemolytic streptococci carriers and non-carriers. Journal of Medical Microbiology 55, 1741–1744. https://doi.org/10.1099/jmm.0.46796-0

Broudy, T.B., Fischetti, V.A., 2003. In Vivo Lysogenic Conversion of Tox− *Streptococcus pyogenes* to Tox+ with Lysogenic Streptococci or Free Phage. Society 71, 3782–3786. https://doi.org/10.1128/IAI.71.7.3782

Broudy, T.B., Pancholi, V., Fischetti, V.A., 2002. The In Vitro Interaction of Streptococcus pyogenes with Human Pharyngeal Cells Induces a Phage-Encoded Extracellular DNase. Society 70, 2805–2811. https://doi.org/10.1128/IAI.70.6.2805

Broudy, T.B., Pancholi, V., Fischetti, V.A., 2001. Induction of lysogenic bacteriophage and phage-associated toxin from Group A streptococci during coculture with human pharyngeal cells. Infection and Immunity 69, 1440–1443. https://doi.org/10.1128/IAI.69.3.1440-1443.2001

Brouwer, S., Barnett, T.C., Ly, D., Kasper, K.J., De Oliveira, D.M.P., Rivera-Hernandez, T., Cork, A.J., McIntyre, L., Jespersen, M.G., Richter, J., Schulz, B.L., Dougan, G., Nizet, V., Yuen, K.-Y., You, Y., McCormick, J.K., Sanderson-Smith, M.L., Davies, M.R., Walker, M.J., 2020. Prophage exotoxins enhance colonization fitness in epidemic scarlet fever-causing *Streptococcus pyogenes*. Nature Communications 11, 5018. https://doi.org/10.1038/s41467-020-18700-5

Brouwer, S., Barnett, T.C., Rivera-Hernandez, T., Rohde, M., Walker, M.J., 2016. *Streptococcus pyogenes* adhesion and colonization. FEBS Letters 590, 3739–3757. https://doi.org/10.1002/1873-3468.12254

Brüssow, H., Desiere, F., 2001. Comparative phage genomics and the evolution of siphoviridae: Insights from dairy phages. Molecular Microbiology 39, 213–223. https://doi.org/10.1046/j.1365-2958.2001.02228.x

Buchanan, J.T., Simpson, A.J., Aziz, R.K., Liu, G.Y., Kristian, S.A., Kotb, M., Feramisco, J., Nizet, V., 2006. DNase expression allows the pathogen Group A *Streptococcus* to escape killing in neutrophil extracellular traps. Current Biology 16, 396–400. https://doi.org/10.1016/j.cub.2005.12.039

Burghout, P., Cron, L.E., Gradstedt, H., Quintero, B., Simonetti, E., Bijlsma, J.J.E., Bootsma, H.J., Hermans, P.W.M., 2010. Carbonic Anhydrase Is Essential for *Streptococcus pneumoniae.* Growth in Environmental Ambient Air. Journal of Bacteriology 192, 4054 LP – 4062. https://doi.org/10.1128/JB.00151-10

Canchaya, C., Desiere, F., McShan, W.M., Ferretti, J.J., Parkhill, J., Brüssow, H., 2002. Genome analysis of an inducible prophage and prophage remnants integrated in the *Streptococcus pyogenes* strain SF370. Virology 302, 245–258. https://doi.org/10.1006/viro.2002.1570

Canchaya, C., Proux, C., Fournous, G., Bruttin, A., Brussow, H., 2003. Prophage Genomics. Microbiology and Molecular Biology Reviews 67, 473–473. https://doi.org/10.1128/mmbr.67.3.473.2003

Carapetis, J.R., Jacoby, P., Carville, K., Ang, S.-J.J., Curtis, N., Andrews, R., 2014. Effectiveness of clindamycin and intravenous immunoglobulin, and risk of disease in contacts, in invasive Group A streptococcal infections. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 59, 358–365. https://doi.org/10.1093/cid/ciu304

Carapetis, J.R., Steer, A.C., Mulholland, E.K., Weber, M., 2005. The global burden of Group A streptococcal diseases. Lancet Infectious Diseases 5, 685–694. https://doi.org/10.1016/S1473-3099(05)70267-X

Carlsson, F., Berggård, K., Stålhammar-Carlemalm, M., Lindahl, G., 2003. Evasion of phagocytosis through cooperation between two ligand-binding regions in *Streptococcus pyogenes* M protein. Journal of Experimental Medicine 198, 1057–1068. https://doi.org/10.1084/jem.20030543

Carlsson, F., Sandin, C., Lindahl, G., 2005. Human fibrinogen bound to *Streptococcus pyogenes* M protein inhibits complement deposition via the classical pathway. Molecular Microbiology 56, 28–39. https://doi.org/10.1111/j.1365-2958.2005.04527.x

Carroll, R.K., Musser, J.M., 2011. From transcription to activation: How Group A *Streptococcus*, the flesh-eating pathogen, regulates SpeB cysteine protease production. Molecular Microbiology 81, 588–601. https://doi.org/10.1111/j.1365-2958.2011.07709.x

Carroll, R.K., Shelburne, S.A., Olsen, R.J., Suber, B., Sahasrabhojane, P., Kumaraswami, M., Beres, S.B., Shea, P.R., Flores, A.R., Musser, J.M., 2011. Naturally occurring single amino acid replacements in a regulatory protein alter streptococcal gene expression and virulence in mice. Journal of Clinical Investigation 121, 1956–1968. https://doi.org/10.1172/JCI45169

Carrolo, M., Frias, M.J., Pinto, F.R., Melo-Cristino, J., Ramirez, M., 2010. Prophage spontaneous activation promotes DNA release enhancing biofilm formation in *Streptococcus pneumoniae*. PLoS ONE 5. https://doi.org/10.1371/journal.pone.0015678

Cattoir, V., 2015. Mechanisms of antibiotic resistance - *Streptococcus pyogenes*. Frontiers in Microbiology 6, 1–45. https://doi.org/10.3389/fmicb.2015.00034

Chalker, V., Jironkin, A., Coelho, J., Al-Shahib, A., Platt, S., Kapatai, G., Daniel, R., Dhami, C., Laranjeira, M., Chambers, T., Guy, R., Lamagni, T., Harrison, T., Chand, M., Johnson, A.P., Underwood, A., Ramsay, M., Fry, N., Purohit, A., Brown, R., 2017. Genome analysis following a national increase in Scarlet Fever in England 2014. BMC Genomics 18, 4–13. https://doi.org/10.1186/s12864-017-3603-z

Chalmers, C., Khemlani, A.H.J., Sohn, C.R., Loh, J.M.S., Tsai, C., Proft, T., 2017. *Streptococcus pyogenes* nuclease A (SpnA) mediated virulence does not exclusively depend on nuclease activity. 1, 42-48 Journal of Microbiology, Immunology and Infection. https://doi.org/10.1016/j.jmii.2017.09.006

Chandrasekaran, S., Caparon, M.G., 2015. The *Streptococcus pyogenes*NAD+ glycohydrolase modulates epithelial cell PARylation and HMGB1 release. Cellular Microbiology 17, 1376–1390. https://doi.org/10.1111/cmi.12442

Chaussee, M.A., Dmitriev, A. V., Callegari, E.A., Chaussee, M.S., 2008. Growth phase-associated changes in the transcriptome and proteome of *Streptococcus pyogenes*. Archives of Microbiology 189, 27–41. https://doi.org/10.1007/s00203-007-0290-1

Cho, K.H., Caparon, M.G., 2005. Patterns of virulence gene expression differ between biofilm and tissue communities of *Streptococcus pyogenes*. Molecular Microbiology 57, 1545–1556. https://doi.org/10.1111/j.1365-2958.2005.04786.x

Chochua, S., Metcalf, B.J., Li, Z., Rivers, J., Mathis, S., 2017. Population and Whole Genome Invasive Group A Streptococci Recovered in the United States in 2015. American Society for Microbiology 8, 1422–17.

Chochua, S., Rivers, J., Mathis, S., Li, Z., Velusamy, S., McGee, L., Beneden, C. Van, Li, Y., Metcalf, B.J., Beall, B., 2019. Emergent invasive Group A *Streptococcus dysgalactiae* subsp. Equisimilis, United States, 2015–2018. Emerging Infectious Diseases 25, 1543–1547. https://doi.org/10.3201/eid2508.181758

Clancy, J., Petitpas, J., Dib-Hajj, F., Yuan, W., Cronan, M., Kamath, A. V., Bergeron, J., Retsema, J.A., 1996. Molecular cloning and functional analysis of a novel macrolide-resistance determinant, mefA, from *Streptococcus pyogenes*. Molecular Microbiology 22, 867–879. https://doi.org/10.1046/j.1365-2958.1996.01521.x

Cleary, P.P., Matsuka, Y. V., Huynh, T., Lam, H., Olmsted, S.B., 2004. Immunization with C5a peptidase from either Group A or B streptococci enhances clearance of Group A streptococci from intranasally infected mice. Vaccine 22, 4332–4341. https://doi.org/10.1016/j.vaccine.2004.04.030

Cleary, P.P., Prahbu, U., Dale, J.B., Wexler, D.E., Handley, J., 1992. Streptococcal C5a peptidase is a highly specific endopeptidase. Infection and Immunity 60, 5219–5223. https://doi.org/10.1128/iai.60.12.5219-5223.1992

Cockerill, F.R., MacDonald, K.L., Thompson, R.L., Roberson, F., Kohner, P.C., Besser-Wiek, J., Manahan, J.M., Musser, J.M., Schlievert, P.M., Talbot, J., Frankfort, B., Steckelberg, J.M., Wilson, W.R., Osterholm, M.T., 1997. An outbreak of invasive Group A streptococcal disease associated with high carriage rates of the invasive clone among school-aged children. Journal of the American Medical Association 277, 38–43. https://doi.org/10.1001/jama.277.1.38

Cole, J.N., McArthur, J.D., McKay, F.C., Sanderson-Smith, M.L., Cork, A.J., Ranson, M., Rohde, M., Itzek, A., Sun, H., Ginsburg, D., Kotb, M., Nizet, V., Chhatwal, G.S., Walker, M.J., 2006. Trigger for Group A streptococcal M1T1 invasive disease. The FASEB Journal 20, 1745–1747. https://doi.org/10.1096/fj.06-5804fje

Collin, M., Olsén, A., 2001. Effect of SpeB and EndoS from *Streptococcus pyogenes* on human immunoglobulins. Infection and Immunity 69, 7187–7189. https://doi.org/10.1128/IAI.69.11.7187-7189.2001

Collin, M., Svensson, M.D., Sjo, A.G., Jensenius, J.C., Sjo, U., Olse, A., 2002. EndoS and SpeB from *Streptococcus pyogenes* Inhibit Immunoglobulin-Mediated Opsonophagocytosis. 70, 6646–6651. https://doi.org/10.1128/IAI.70.12.6646

Colonna, M., Facchetti, F., 2003. TREM-1 (Triggering Receptor Expressed on Myeloid Cells): A New Player in Acute Inflammatory Responses. The Journal of Infectious Diseases 187, S397–S401. https://doi.org/10.1086/374754

Commons, R.J., Smeesters, P.R., Proft, T., Fraser, J.D., Robins-Browne, R., Curtis, N., 2014. Streptococcal

superantigens: Categorization and clinical associations. Trends in Molecular Medicine 20, 48–62. https://doi.org/10.1016/j.molmed.2013.10.004

Cortés, M.G., Wessels, M.R., 2009. Inhibition of Dendritic Cell Maturation by Group A *Streptococcus* . The Journal of Infectious Diseases 200, 1152–1161. https://doi.org/10.1086/605696

Cortés, M.G., Krog, J., Balázsi, G., 2019. Optimality of the spontaneous prophage induction rate. Journal of Theoretical Biology 483. https://doi.org/10.1016/j.jtbi.2019.110005

Courtney, H.S., Hasty, D.L., Dale, J.B., 2006a. Anti-phagocytic mechanisms of *Streptococcus pyogenes*: Binding of fibrinogen to M-related protein. Molecular Microbiology 59, 936–947. https://doi.org/10.1111/j.1365-2958.2005.04977.x

Courtney, H.S., Hasty, D.L., Dale, J.B., 2006b. Anti-phagocytic mechanisms of *Streptococcus pyogenes*: Binding of fibrinogen to M-related protein. Molecular Microbiology 59, 936–947. https://doi.org/10.1111/j.1365-2958.2005.04977.x

Courtney, H.S., Li, Y., Twal, W.O., Argraves, W.S., 2009. Serum opacity factor is a streptococcal receptor for the extracellular matrix protein fibulin-1. Journal of Biological Chemistry 284, 12966–12971. https://doi.org/10.1074/jbc.M901143200

Courtney, H.S., Pownall, H.J., 2010. The structure and function of serum opacity factor: A unique streptococcal virulence determinant that targets high-density lipoproteins. Journal of Biomedicine and Biotechnology 2010. https://doi.org/10.1155/2010/956071

Crater, D.L., Dougherty, B.A., Van de Rijn, I., 1995. Molecular characterization of hasC from an operon required for hyaluronic acid synthesis in Group A streptococci. Demonstration of UDP-glucose pyrophosphorylase activity. Journal of Biological Chemistry 270, 28676–28680. https://doi.org/10.1074/jbc.270.48.28676

Cue, D., Lam, H., Cleary, P.P., 2001. Genetic dissection of the *Streptococcus pyogenes* M1 protein: Regions involved in fibronectin binding and intracellular invasion. Microbial Pathogenesis 31, 231–242. https://doi.org/10.1006/mpat.2001.0467

Cunningham, M.W., 2000. Pathogenesis of Group A Streptococcal Infections. Clinical Microbiology Reviews 13, 470 LP – 511. https://doi.org/10.1128/CMR.13.3.470

da Silva Duarte, V., Giaretta, S., Campanaro, S., Treu, L., Armani, A., Tarrah, A., Oliveira de Paula, S., Giacomini, A., Corich, V., 2019. A Cryptic Non-Inducible Prophage Confers Phage-Immunity on the *Streptococcus thermophilus* M17PTZA496. Viruses 11. https://doi.org/10.3390/v11010007

Dale, J.B., Chiang, E.Y., Hasty, D.L., Courtney, H.S., 2002. Antibodies against a synthetic peptide of SagA neutralize the cytolytic activity of streptolysin S from Group A streptococci. Infection and Immunity 70, 2166–2170. https://doi.org/10.1128/IAI.70.4.2166-2170.2002

Dale, J.B., Washburn, R.G., Marques, M.B., Wessels, M.R., 1996. Hyaluronate capsule and surface M protein in resistance to opsonization of Group A streptococci. Infection and Immunity 64, 1495–1501. https://doi.org/10.1128/iai.64.5.1495-1501.1996

Danger, J.L., Cao, T.N., Cao, T.H., Sarkar, P., Treviño, J., Pflughoeft, K.J., Sumby, P., 2015a. The small

regulatory RNA FasX enhances Group A *Streptococcus* virulence and inhibits pilus expression via serotype-specific targets. Molecular microbiology 96, 249–262. https://doi.org/10.1111/mmi.12935

Danger, J.L., Makthal, N., Kumaraswami, M., Sumby, P., 2015b. The FasX Small Regulatory RNA Negatively Regulates the Expression of Two Fibronectin-Binding Proteins in Group A *Streptococcus*. Journal of Bacteriology 197, 3720 LP – 3730. https://doi.org/10.1128/JB.00530-15

Datta, V., Myskowski, S.M., Kwinn, L.A., Chiem, D.N., Varki, N., Kansal, R.G., Kotb, M., Nizet, V., 2005. Mutational analysis of the Group A streptococcal operon encoding streptolysin S and its virulence role in invasive infection. Molecular Microbiology 56, 681–695. https://doi.org/10.1111/j.1365-2958.2005.04583.x

Davies, F.J., Olme, C., Lynskey, N.N., Turner, C.E., Sriskandan, S., 2019. Streptococcal superantigen-induced expansion of human tonsil T cells leads to altered T follicular helper cell phenotype, B cell death and reduced immunoglobulin release. Clinical and Experimental Immunology 197, 83–94. https://doi.org/10.1111/cei.13282

Davies, M.R., Holden, M.T., Coupland, P., Chen, J.H.K., Venturini, C., Barnett, T.C., Zakour, N.L. Ben, Tse, H., Dougan, G., Yuen, K.-Y., Walker, M.J., 2015. Emergence of scarlet fever *Streptococcus pyogenes emm*12 clones in Hong Kong is associated with toxin acquisition and multidrug resistance. Nature Genetics 47, 84–87. https://doi.org/10.1038/ng.3147

Davies, M.R., McIntyre, L., Mutreja, A., Lacey, J.A., Lees, J.A., Towers, R.J., Duchêne, S., Smeesters, P.R., Frost, H.R., Price, D.J., Holden, M.T.G., David, S., Giffard, P.M., Worthing, K.A., Seale, A.C., Berkley, J.A., Harris, S.R., Rivera-Hernandez, T., Berking, O., Cork, A.J., Torres, R.S.L.A., Lithgow, T., Strugnell, R.A., Bergmann, R., Nitsche-Schmitz, P., Chhatwal, G.S., Bentley, S.D., Fraser, J.D., Moreland, N.J., Carapetis, J.R., Steer, A.C., Parkhill, J., Saul, A., Williamson, D.A., Currie, B.J., Tong, S.Y.C., Dougan, G., Walker, M.J., 2019. Atlas of Group A streptococcal vaccine candidates compiled using large-scale comparative genomics. Nature Genetics 51, 1035–1043. https://doi.org/10.1038/s41588-019-0417-8

DebRoy, S., Li, X., Kalia, A., Galloway-Pena, J., Shah, B.J., Fowler, V.G., Flores, A.R., Shelburne, S.A., 2018. Identification of a chimeric *emm* gene and novel *emm* pattern in currently circulating strains of *emm*4 Group A *Streptococcus*. Microbial genomics 4, 1–6. https://doi.org/10.1099/mgen.0.000235

Derré-Bobillot, A., Cortes-Perez, N.G., Yamamoto, Y., Kharrat, P., Couvé, E., Da Cunha, V., Decker, P., Boissier, M.-C., Escartin, F., Cesselin, B., Langella, P., Bermúdez-Humarán, L.G., Gaudu, P., 2013. Nuclease A (Gbs0661), an extracellular nuclease of *Streptococcus* agalactiae, attacks the neutrophil extracellular traps and is needed for full virulence. Molecular microbiology 89, 518–531. https://doi.org/10.1111/mmi.12295

Desiere, F., McShan, W.M., Van Sinderen, D., Ferretti, J.J., Brüssow, H., 2001. Comparative genomics reveals close genetic relationships between phages from dairy bacteria and pathogenic streptococci: Evolutionary implications for prophage-host interactions. Virology 288, 325–341. https://doi.org/10.1006/viro.2001.1085

Di Nardo, A., Yamasaki, K., Dorschner, R.A., Lai, Y., Gallo, R.L., 2008. Mast Cell Cathelicidin Antimicrobial

Peptide Prevents Invasive Group A *Streptococcus* Infection of the Skin . The Journal of Immunology 180, 7565–7573. https://doi.org/10.4049/jimmunol.180.11.7565

Dmitriev, A. V., McDowell, E.J., Kappeler, K. V., Chaussee, M.A., Rieck, L.D., Chaussee, M.S., 2006. The Rgg regulator of *Streptococcus pyogenes* influences utilization of nonglucose carbohydrates, prophage induction, and expression of the NAD-glycohydrolase virulence operont. Journal of Bacteriology 188, 7230–7241. https://doi.org/10.1128/JB.00877-06

Dokland, T., 1999. Scaffolding proteins and their role in viral assembly. Cellular and molecular life sciences : CMLS 56, 580–603. https://doi.org/10.1007/s000180050455

Dunkelberger, J.R., Song, W.C., 2010. Complement and its role in innate and adaptive immune responses. Cell Research 20, 34–50. https://doi.org/10.1038/cr.2009.139

Edwards, R.J., Pyzio, M., Gierula, M., Turner, C.E., Abdul-Salam, V.B., Sriskandan, S., 2018. Proteomic analysis at the sites of clinical infection with invasive *Streptococcus pyogenes*. Scientific Reports 8, 1–9. https://doi.org/10.1038/s41598-018-24216-2

Edwards, R.J., Taylor, G.W., Ferguson, M., Murray, S., Rendell, N., Wrigley, A., Bai, Z., Boyle, J., Finney, S.J., Jones, A., Russell, H.H., Turner, C., Cohen, J., Faulkner, L., Sriskandan, S., 2005. Specific C-Terminal Cleavage and Inactivation of Interleukin-8 by Invasive Disease Isolates of *Streptococcus pyogenes* . The Journal of Infectious Diseases 192, 783–790. https://doi.org/10.1086/432485

Efstratiou, A., 2000. Group A streptococci in the 1990s. Journal of Antimicrobial Chemotherapy 45, 3–12. https://doi.org/10.1093/jac/45.suppl_1.3

Efstratiou, A., Lamagni, T., 2016. Epidemiology of *Streptococcus pyogenes*. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 1–27.

Egesten, A., Olin, A.I., Linge, H.M., Yadav, M., Mörgelin, M., Karlsson, A., Collin, M., 2009. SpeB of *Streptococcus pyogenes* differentially modulates antibacterial and receptor activating properties of human chemokines. PLoS ONE 4, 1–9. https://doi.org/10.1371/journal.pone.0004769

Eraso, J.M., Kachroo, P., Olsen, R.J., Beres, S.B., Zhu, L., Badu, T., Shannon, S., Cantu, C.C., Saavedra, M.O., Kubiak, S.L., Porter, A.R., DeLeo, F.R., Musser, J.M., 2020. Genetic heterogeneity of the Spy1336/R28-Spy1337 virulence axis in *Streptococcus pyogenes* and effect on gene transcript levels and pathogenesis. PloS one 15, e0229064. https://doi.org/10.1371/journal.pone.0229064

Eriksson, A., Norgren, M., 2003. Cleavage of antigen-bound immunoglobulin G by SpeB contributes to streptococcal persistence in opsonizing blood. Infection and Immunity 71, 211–217. https://doi.org/10.1128/IAI.71.1.211-217.2003

Euler, C.W., Juncosa, B., Ryan, P.A., Deutsch, D.R., McShan, W.M., Fischetti, V.A., 2016. Targeted curing of all lysogenic bacteriophage from *Streptococcus pyogenes* using a novel counter-selection technique. PLoS ONE 11, 1–22. https://doi.org/10.1371/journal.pone.0146408

Fagan, P.K., Reinschied, D., Gottschalk, B., Chhatwal, G.S., 2001. Identification and characterization of a novel secreted immunoglobulin binding protein from Group A *Streptococcus*. Infection and Immunity 69, 4851–4857. https://doi.org/10.1128/IAI.69.8.4851-4857.2001

Farenc, C., Spinelli, S., Vinogradov, E., Tremblay, D., Blangy, S., Sadovskaya, I., Moineau, S., Cambillau, C., 2014. Molecular Insights on the Recognition of a *Lactococcus lactis* Cell Wall Pellicle by the Phage 1358 Receptor Binding Protein. Journal of Virology 88, 7005–7015. https://doi.org/10.1128/jvi.00739-14

Federle, M.J., McIver, K.S., Scott, J.R., 1999. A response regulator that represses transcription of several virulence operons in the Group A *Streptococcus*. Journal of Bacteriology 181, 3649–3657. https://doi.org/10.1128/jb.181.12.3649-3657.1999

Fernández De Henestrosa, A.R., Ogi, T., Aoyagi, S., Chafin, D., Hayes, J.J., Ohmori, H., Woodgate, R., 2000. Identification of additional genes belonging to the LexA regulon in *Escherichia* coli. Molecular Microbiology 35, 1560–1572. https://doi.org/10.1046/j.1365-2958.2000.01826.x

Fernie-King, B.A., Seilly, D.J., Davies, A., Lachmann, P.J., 2002. Streptococcal inhibitor of complement inhibits two additional components of the mucosal innate immune system: Secretory leukocyte proteinase inhibitor and lysozyme. Infection and Immunity 70, 4908–4916. https://doi.org/10.1128/IAI.70.9.4908-4916.2002

Fernie-King, B.A., Seilly, D.J., Willers, C., Würzner, R., Davies, A., Lachmann, P.J., 2001. Streptococcal inhibitor of complement (SIC) inhibits the membrane attack complex by preventing uptake of C567 onto cell membranes. Immunology 103, 390–398. https://doi.org/10.1046/j.1365-2567.2001.01249.x

Ferretti, J.J., McShan, W.M., Ajdic, D., Savic, D.J., Savic, G., Lyon, K., Primeaux, C., Sezate, S., Suvorov, A.N., Kenton, S., Lai, H.S., Lin, S.P., Qian, Y., Jia, H.G., Najar, F.Z., Ren, Q., Zhu, H., Song, L., White, J., Yuan, X., Clifton, S.W., Roe, B.A., McLaughlin, R., 2001. Complete genome sequence of an M1 strain of *Streptococcus pyogenes*. Proceedings of the National Academy of Sciences of the United States of America 98, 4658–4663. https://doi.org/10.1073/pnas.071559398

Fischetti, V.A., 2007. In vivo acquisition of prophage in *Streptococcus pyogenes*. Trends in Microbiology 15, 297–300. https://doi.org/10.1016/j.tim.2007.05.003

Fischetti, V.A., 1989. Streptococcal M protein: Molecular design and biological behavior. Clinical Microbiology Reviews 2, 285–314. https://doi.org/10.1128/CMR.2.3.285

Flores, A.R., Chase McNeil, J., Shah, B., Van Beneden, C., Shelburne, S.A., 2019. Capsule-negative *emm* types are an increasing cause of pediatric Group A streptococcal infections at a large pediatric hospital in Texas. Journal of the Pediatric Infectious Diseases Society 8, 244–250. https://doi.org/10.1093/jpids/piy053

Flores, A.R., Jewell, B.E., Fittipaldi, N., Beres, S.B., Musser, J.M., 2012. Human disease isolates of serotype M4 and M22 Group A *Streptococcus* lack genes required for hyaluronic acid capsule biosynthesis. mBio 3, 4–8. https://doi.org/10.1128/mBio.00413-12

Flores, A.R., Jewell, B.E., Olsen, R.J., Shelburne, S.A., Fittipaldi, N., Beres, S.B., Musser, J.M., 2014. Asymptomatic carriage of Group A *Streptococcus* is associated with elimination of capsule production. Infection and Immunity 82, 3958–3967. https://doi.org/10.1128/IAI.01788-14

Flores, A.R., Luna, R.A., Runge, J.K., Shelburne, S.A., Baker, C.J., 2017. Cluster of fatal Group A streptococcal *emm*87 infections in a single family: Molecular basis for invasion and transmission. Journal of Infectious Diseases 215, 1648–1652. https://doi.org/10.1093/infdis/jix177

Fluckiger, U., Jones, K.F., Fischetti, V.A., 1998. Immunoglobulins to Group A streptococcal surface molecules decrease adherence to and invasion of human pharyngeal cells. Infection and Immunity 66, 974–979. https://doi.org/10.1128/iai.66.3.974-979.1998

Fogg, G.C., Caparon, M.G., 1997. Constitutive expression of fibronectin binding in *Streptococcus pyogenes* as a result of anaerobic activation of rofA. Journal of Bacteriology 179, 6172–6180. https://doi.org/10.1128/jb.179.19.6172-6180.1997

Franklin, L., Nobbs, A.H., Bricio-Moreno, L., Wright, C.J., Maddocks, S.E., Sahota, J.S., Ralph, J., O'Connor, M., Jenkinson, H.F., Kadioglu, A., 2013. The AgI/II Family Adhesin AspA Is Required for Respiratory Infection by *Streptococcus pyogenes*. PLoS ONE 8. https://doi.org/10.1371/journal.pone.0062433

Friães, A., Pato, C., Melo-Cristino, J., Ramirez, M., 2015. Consequences of the variability of the CovRS and RopB regulators among *Streptococcus pyogenes* causing human infections. Scientific Reports 5, 1–12. https://doi.org/10.1038/srep12057

Frias, M.J., Melo-Cristino, J., Ramirez, M., 2013. Export of the pneumococcal phage SV1 lysin requires choline-containing teichoic acids and is holin-independent. Molecular Microbiology 87, 430–445. https://doi.org/10.1111/mmi.12108

Frígols, B., Quiles-Puchalt, N., Mir-Sanchis, I., Donderis, J., Elena, S.F., Buckling, A., Novick, R.P., Marina, A., Penadés, J.R., 2015. Virus Satellites Drive Viral Evolution and Ecology. PLoS Genetics 11, 1–19. https://doi.org/10.1371/journal.pgen.1005609

Frost, H.R., Davies, M.R., Delforge, V., Lakhloufi, D., Sanderson-Smith, M., Srinivasan, V., Steer, A.C., Walker, M.J., Beall, B., Botteaux, A., Smeesters, P.R., 2020. Analysis of Global Collection of Group A *Streptococcus* Genomes Reveals that the Majority Encode a Trio of M and M-Like Proteins. mSphere 5, e00806-19. https://doi.org/10.1128/mSphere.00806-19

Fuller, D.N., Raymer, D.M., Rickgauer, J.P., Robertson, R.M., Catalano, C.E., Anderson, D.L., Grimes, S., Smith, D.E., 2007. Measurements of single DNA molecule packaging dynamics in bacteriophage lambda reveal high forces, high motor processivity, and capsid transformations. Journal of molecular biology 373, 1113–1122. https://doi.org/10.1016/j.jmb.2007.09.011

Galloway-Peña, J., Clement, M.E., Kuinkel, B.K.S., Ruffin, F., Flores, A.R., Levinson, H., Shelburne, S.A., Moore, Z., Fowler, V.G., 2016. Application of whole-genome sequencing to an unusual outbreak of invasive Group A streptococcal disease. Open Forum Infectious Diseases 3, 1–4. https://doi.org/10.1093/ofid/ofw042

Galloway-Peña, J., DebRoy, S., Brumlow, C., Li, X., Tran, T.T., Horstmann, N., Yao, H., Chen, K., Wang, F., Pan, B.F., Hawke, D.H., Thompson, E.J., Arias, C.A., Fowler, V.G., Bhatti, M.M., Kalia, A., Flores, A.R., Shelburne, S.A., 2018. Hypervirulent Group A *Streptococcus* emergence in an acaspular background is associated with marked remodeling of the bacterial cell surface. PLoS ONE 13, 1–23. https://doi.org/10.1371/journal.pone.0207897

Garcia, M., Pimentel, M., Moniz-Pereira, J., 2002. Expression of mycobacteriophage Ms6 lysis genes is driven by two σ70-like promoters and is dependent on a transcription termination signal present in the leader

RNA. Journal of Bacteriology 184, 3034–3043. https://doi.org/10.1128/JB.184.11.3034-3043.2002

Gasc, A.M., Sicard, N., Claverys, J.P., Sicard, A.M., 1980. Lack of SOS repair in *Streptococcus pneumoniae*. Mutation research 70, 157–165. https://doi.org/10.1016/0027-5107(80)90155-4

Gerlach, D., Reichardt, W., Fleischer, B., Schmidt, K.H., 1994. Separation of mitogenic and pyrogenic activities from so-called erythrogenic toxin type B (Streptococcal Proteinase). Zentralblatt fur Bakteriologie 280, 507–514. https://doi.org/10.1016/S0934-8840(11)80510-4

Gillen, C.M., Courtney, H.S., Schulze, K., Rohde, M., Wilson, M.R., Timmer, A.M., Guzman, C.A., Nizet, V., Chhatwal, G.S., Walker, M.J., 2008. Opacity factor activity and epithelial cell binding by the serum opacity factor protein of *Streptococcus pyogenes* are functionally discrete. Journal of Biological Chemistry 283, 6359–6366. https://doi.org/10.1074/jbc.M706739200

Giovanetti, E., Brenciani, A., Lupidi, R., Roberts, M.C., Varaldo, P.E., 2003. Presence of the *tet*(*O*) gene in erythromycin- and tetracycline-resistant strains of *Streptococcus pyogenes* and linkage with either the *mef(A)* or the *erm(A)* gene. Antimicrobial Agents and Chemotherapy 47, 2844–2849. https://doi.org/10.1128/AAC.47.9.2844-2849.2003

Giovanetti, E., Brenciani, A., Morroni, G., Tiberi, E., Pasquaroli, S., Mingoia, M., Varaldo, P.E., 2014. Transduction of the *Streptococcus pyogenes* bacteriophage Φm46.1, carrying resistance genes *mef(A)* and *tet(O)*, to other *Streptococcus* species. Frontiers in Microbiology 5, 1–9. https://doi.org/10.3389/fmicb.2014.00746

Giovanetti, E., Brenciani, A., Tiberi, E., Bacciaglia, A., Varaldo, P.E., 2012. ICESp2905, the *erm(TR)-tet(O)* element of *Streptococcus pyogenes*, is formed by two independent integrative and conjugative elements. Antimicrobial Agents and Chemotherapy 56, 591–594. https://doi.org/10.1128/AAC.05352-11

Giovanetti, E., Brenciani, A., Vecchi, M., Manzin, A., Varaldo, P.E., 2005. Prophage association of *mef(A)* elements encoding efflux-mediated erythromycin resistance in *Streptococcus pyogenes*. Journal of Antimicrobial Chemotherapy 55, 445–451. https://doi.org/10.1093/jac/dki049

Gogos, A., Federle, M.J., 2019. Modeling *Streptococcus pyogenes* pharyngeal colonization in the mouse. Frontiers in Cellular and Infection Microbiology 9, 1–12. https://doi.org/10.3389/fcimb.2019.00137

Goldblatt, J., Lawrenson, R.A., Muir, L., Dattani, S., Hoffland, A., Tsuchiya, T., Kanegasaki, S., Sriskandan, S., Pease, J.E., 2019. A Requirement for Neutrophil Glycosaminoglycans in Chemokine:Receptor Interactions Is Revealed by the Streptococcal Protease SpyCEP. The Journal of Immunology 202, 3246–3255. https://doi.org/10.4049/jimmunol.1801688

Goldmann, O., Rohde, M., Chhatwal, G.S., Medina, E., 2004. Role of Macrophages in Host Resistance to Group A Streptococci. Infection and Immunity 72, 2956–2963. https://doi.org/10.1128/IAI.72.5.2956-2963.2004

Goldmann, O., Thulin, P., Heinemann, K., Norrby-teglund, A., Rohde, M., Medina, E., 2008. Phagocytosis-independent antimicrobial activity of mast cells by means of extracellular trap formation. Blood 111, 3070–3080. https://doi.org/10.1182/blood-2007-07-104018.The

Gordon, S., Plüddemann, A., 2017. Tissue macrophages: Heterogeneity and functions. BMC Biology 15, 1–18.

https://doi.org/10.1186/s12915-017-0392-4

Graham, M.R., Smoot, L.M., Migliaccio, C.A.L., Virtaneva, K., Sturdevant, D.E., Porcella, S.F., Federle, M.J., Adams, G.J., Scott, J.R., Musser, J.M., 2002. Virulence control in Group A *Streptococcus* by a two-component gene regulatory system: Global expression profiling and in vivo infection modeling. Proceedings of the National Academy of Sciences of the United States of America 99, 13855–13860. https://doi.org/10.1073/pnas.202353699

Graham, M.R., Virtaneva, K., Porcella, S.F., Barry, W.T., Gowen, B.B., Johnson, C.R., Wright, F.A., Musser, J.M., 2005. Group A *Streptococcus* transcriptome dynamics during growth in human blood reveals bacterial adaptive and survival strategies. American Journal of Pathology 166, 455–465. https://doi.org/10.1016/S0002-9440(10)62268-7

Granok, A.B., Parsonage, D., Ross, R.P., Caparon, M.G., 2000. The RofA binding site in *Streptococcus pyogenes* is utilized in multiple transcriptional pathways. Journal of Bacteriology 182, 1529–1540. https://doi.org/10.1128/JB.182.6.1529-1540.2000

Grayson, A.K., Hearnden, V., Bolt, R., Jebreel, A., Colley, H.E., Murdoch, C., 2018. Use of a Rho kinase inhibitor to increase human tonsil keratinocyte longevity for three-dimensional, tissue engineered tonsil epithelium equivalents. Journal of tissue engineering and regenerative medicine 12, e1636–e1646. https://doi.org/10.1002/term.2590

Grebe, T., Hakenbeck, R., 1996. Penicillin-binding proteins 2b and 2x of *Streptococcus pneumoniae* are primary resistance determinants for different classes of beta-lactam antibiotics. Antimicrobial agents and chemotherapy 40, 829–834. https://doi.org/10.1128/AAC.40.4.829

Green, N.M., Zhang, S., Porcella, S.F., Nagiec, M.J., Barbian, K.D., Beres, S.B., LeFebvre, R.B., Musser, J.M., 2005. Genome Sequence of a Serotype M28 Strain of Group A *Streptococcus*: Potential New Insights into Puerperal Sepsis and Bacterial Disease Specificity . The Journal of Infectious Diseases 192, 760–770. https://doi.org/10.1086/430618

Groth, A.C., Calos, M.P., 2004. Phage integrases: Biology and applications. Journal of Molecular Biology 335, 667–678. https://doi.org/10.1016/j.jmb.2003.09.082

Gryllos, I., Cywes, C., Shearer, M.H., Cary, M., Kennedy, R.C., Wessels, M.R., 2001. Regulation of capsule gene expression by Group A *Streptococcus* during pharyngeal colonization and invasive infection. Molecular Microbiology 42, 61–74. https://doi.org/10.1046/j.1365-2958.2001.02635.x

Gryllos, I., Grifantini, R., Colaprico, A., Jiang, S., DeForce, E., Hakansson, A., Telford, J.L., Grandi, G., Wessels, M.R., 2007. Mg2+ signalling defines the Group A streptococcal CsrRS (CovRS) regulon. Molecular Microbiology 65, 671–683. https://doi.org/10.1111/j.1365-2958.2007.05818.x

Gryllos, I., Tran-Winkler, H.J., Cheng, M.F., Chung, H., Bolcome, R., Lu, W., Lehrer, R.I., Wessels, M.R., 2008. Induction of Group A *Streptococcus* virulence by a human antimicrobial peptide. Proceedings of the National Academy of Sciences of the United States of America 105, 16755–16760. https://doi.org/10.1073/pnas.0803815105

Hagberg, C., Radulescu, A., Rex, J.H., 1997. Necrotizing Fasciitis Due to Group A *Streptococcus* after an

Accidental Needle-Stick Injury. New England Journal of Medicine 337, 1699. https://doi.org/10.1056/NEJM199712043372318

Halperin, S.A., Ferrieri, P., Gray, E.D., Kaplan, E.L., Wannamaker, L.W., 1987. Antibody Response to Bacteriophage Hyaluronidase in Acute Glomerulonephritis After Group A Streptococcal Infection. Journal of Infectious Diseases 155, 253–261. https://doi.org/10.1093/infdis/155.2.253

Hamada, S., Kawabata, S., Nakagawa, I., 2015. Molecular and genomic characterization of pathogenic traits of Group A *Streptococcus pyogenes*. Proceedings of the Japan Academy Series B: Physical and Biological Sciences 91, 539–559. https://doi.org/10.2183/pjab.91.539

Han, H., Liu, C., Wang, Q., Xuan, C., Zheng, B., Tang, J., Yan, J., Zhang, J., Li, M., Cheng, H., Lu, G., Gao, G.F., 2012. The two-component system Ihk/Irr contributes to the virulence of *Streptococcus suis* serotype 2 strain 05ZYH33 through alteration of the bacterial cell metabolism. Microbiology (United Kingdom) 158, 1852–1866. https://doi.org/10.1099/mic.0.057448-0

Harder, J., Franchi, L., Muñoz-Planillo, R., Park, J.-H., Reimer, T., Núñez, G., 2009. Activation of the Nlrp3 Inflammasome by *Streptococcus pyogenes* Requires Streptolysin O and NF-κB Activation but Proceeds Independently of TLR Signaling and P2X7 Receptor . The Journal of Immunology 183, 5823–5829. https://doi.org/10.4049/jimmunol.0900444

Hasegawa, T., Minami, M., Okamoto, A., Tatsuno, I., Isaka, M., Ohta, M., 2010. Characterization of a virulence-associated and cell-wall-located DNase of *Streptococcus pyogenes*. Microbiology 156, 184–190. https://doi.org/10.1099/mic.0.031955-0

Hendrickson, C., Euler, C.W., Nguyen, S. V., Rahman, M., McCullor, K.A., King, C.J., Fischetti, V.A., McShan, W.M., 2015. Elimination of chromosomal Island SpyCIM1 from *Streptococcus pyogenes* strain SF370 reverses the mutator phenotype and alters global transcription. PLoS ONE 10, 1–23. https://doi.org/10.1371/journal.pone.0145884

Henningham, A., Döhrmann, S., Nizet, V., Cole, J.N., 2015. Mechanisms of Group A *Streptococcus* resistance to reactive oxygen species. FEMS microbiology reviews 39, 488–508. https://doi.org/10.1093/femsre/fuu009

Henningham, A., Yamaguchi, M., Aziz, R.K., Kuipers, K., Buffalo, C.Z., Dahesh, S., Choudhury, B., Van Vleet, J., Yamaguchi, Y., Seymour, L.M., Ben Zakour, N.L., He, L., Smith, H. V., Grimwood, K., Beatson, S.A., Ghosh, P., Walker, M.J., Nizet, V., Colea, J.N., 2014. Mutual exclusivity of hyaluronan and hyaluronidase in invasive Group A *Streptococcus*. Journal of Biological Chemistry 289, 32303–32315. https://doi.org/10.1074/jbc.M114.602847

Hertzén, E., Johansson, L., Kansal, R., Hecht, A., Dahesh, S., Janos, M., Nizet, V., Kotb, M., Norrby-Teglund, A., 2012. Intracellular *Streptococcus pyogenes* in human macrophages display an altered gene expression profile. PloS one 7, 1–10. https://doi.org/10.1371/journal.pone.0035218

Hertzén, E., Johansson, L., Wallin, R., Schmidt, H., Kroll, M., Rehn, A.P., Kotb, M., Mörgelin, M., Norrby-Teglund, A., 2010. M1 protein-dependent intracellular trafficking promotes persistence and replication of *Streptococcus pyogenes* in macrophages. Journal of Innate Immunity 2, 534–545.

https://doi.org/10.1159/000317635

Herwald, H., Collin, M., Müller-Esterl, W., Björck, L., 1996. Streptococcal cysteine proteinase releases kinins: A novel virulence mechanism. Journal of Experimental Medicine 184, 665–673. https://doi.org/10.1084/jem.184.2.665

Hoe, N.P., Ireland, R.M., DeLeo, F.R., Gowen, B.B., Dorward, D.W., Voyich, J.M., Liu, M., Burns, E.H., Culnan, D.M., Bretscher, A., Musser, J.M., 2002. Insight into the molecular basis of pathogen abundance: Group A *Streptococcus* inhibitor of complement inhibits bacterial adherence and internalization into human cells. Proceedings of the National Academy of Sciences of the United States of America 99, 7646–7651. https://doi.org/10.1073/pnas.112039899

Hoe, N.P., Nakashima, K., Lukomski, S., Grigsby, D., Liu, M., Kordari, P., Dou, S.J., Pan, X., Vuopio-Varkila, J., Salmelinna, S., McGeer, A., Low, D.E., Schwartz, B., Schuchat, A., Naidich, S., De Lorenzo, D., Fu, Y.X., Musser, J.M., 1999. Rapid selection of complement-inhibiting protein variants in Group A *Streptococcus* epidemic waves. Nature Medicine 5, 924–929. https://doi.org/10.1038/11369

Holden, M.T.G., Heather, Z., Paillot, R., Steward, K.F., Webb, K., Ainslie, F., Jourdan, T., Bason, N.C., Holroyd, N.E., Mungall, K., Quail, M.A., Sanders, M., Simmonds, M., Willey, D., Brooks, K., Aanensen, D.M., Spratt, B.G., Jolley, K.A., Maiden, M.C.J., Kehoe, M., Chanter, N., Bentley, S.D., Robinson, C., Maskell, D.J., Parkhill, J., Waller, A.S., 2009. Genomic evidence for the evolution of *Streptococcus equi* host restriction, increased virulence, and genetic exchange with human pathogens. PLoS Pathogens 5. https://doi.org/10.1371/journal.ppat.1000346

Holmes, R.K., 2000. Biology and molecular epidemiology of diphtheria toxin and the *tox* gene. The Journal of infectious diseases 181 Suppl, S156-67. https://doi.org/10.1086/315554

Hondorp, E.R., McIver, K.S., 2007. The Mga virulence regulon: Infection where the grass is greener. Molecular Microbiology 66, 1056–1065. https://doi.org/10.1111/j.1365-2958.2007.06006.x

Horst, S.A., Linnér, A., Beineke, A., Lehne, S., Höltje, C., Hecht, A., Norrby-Teglund, A., Medina, E., Goldmann, O., 2013. Prognostic value and therapeutic potential of TREM-1 in *Streptococcus pyogenes*-induced sepsis. Journal of Innate Immunity 5, 581–590. https://doi.org/10.1159/000348283

Horstmann, N., Sahasrabhojane, P., Suber, B., Kumaraswami, M., Olsen, R.J., Flores, A., Musser, J.M., Brennan, R.G., Shelburne, S.A., 2011. Distinct single amino acid replacements in the control of virulence regulator protein differentially impact streptococcal pathogenesis. PLoS Pathogens 7(10). https://doi.org/10.1371/journal.ppat.1002311

Horstmann, N., Tran, C.N., Brumlow, C., DebRoy, S., Yao, H., Nogueras Gonzalez, G., Makthal, N., Kumaraswami, M., Shelburne, S.A., 2018. Phosphatase activity of the control of virulence sensor kinase CovS is critical for the pathogenesis of Group A *Streptococcus*. PLoS Pathogens 14, 1–26. https://doi.org/10.1371/journal.ppat.1007354

Horstmann, R.D., Sievertsen, H.J., Leippe, M., Fischetti, V.A., 1992. Role of fibrinogen in complement inhibition by streptococcal M protein. Infection and Immunity 60, 5036–5041. https://doi.org/10.1128/iai.60.12.5036-5041.1992

Hsieh, C.L., Huang, H.M., Hsieh, S.Y., Zheng, P.X., Lin, Y.S., Chiang-Ni, C., Tsai, P.J., Wang, S.Y., Liu, C.C., Wu, J.J., 2018. NAD-glycohydrolase depletes intracellular NAD+ and inhibits acidification of autophagosomes to enhance multiplication of Group A *Streptococcus* in endothelial cells. Frontiers in Microbiology 9, 1–11. https://doi.org/10.3389/fmicb.2018.01733

Hsueh, S.-C., Huang, Y.-T., Hsueh, P.-R., 2019. *Streptococcus pyogenes* cutaneous infection following sandfly bites. Journal of microbiology, immunology, and infection = Wei mian yu gan ran za zhi. https://doi.org/10.1016/j.jmii.2019.09.001

Hu, M.C., Walls, M.A., Stroop, S.D., Reddish, M.A., Beall, B., Dale, J.B., 2002. Immunogenicity of a 26-valent Group A streptococcal vaccine. Infection and immunity 70, 2171–2177. https://doi.org/10.1128/iai.70.4.2171-2177.2002

Huang, X., Li, D., Xi, L., Mylonakis, E., 2015. *Galleria mellonella* Larvae as an Infection Model for *Penicillium marneffei*. Mycopathologia 180, 159–164. https://doi.org/10.1007/s11046-015-9897-y

Hyder, S.L., Streitfeld, M.M., 1978. Transfer of erythromycin resistance from clinically isolated lysogenic strains of *Streptococcus pyogenes* via their endogenous phage. The Journal of infectious diseases 138, 281–286. https://doi.org/10.1093/infdis/138.3.281

Hynes, W., Johnson, C., Stokes, M., 2009. A single nucleotide mutation results in loss of enzymatic activity in the hyaluronate lyase gene of *Streptococcus pyogenes*. Microbial Pathogenesis 47, 308–313. https://doi.org/10.1016/j.micpath.2009.09.008

Hynes, W.L., Hancock, L., Ferretti, J.J., 1995. Analysis of a second bacteriophage hyaluronidase gene from Strptococcus *pyogenes*: Evidence for a third hyaluronidase involved in extracellular enzymatic activity. Infection and Immunity 63, 3015–3020. https://doi.org/10.1128/iai.63.8.3015-3020.1995

Hynes, W.L., Walton, S.L., 2000. Hyaluronidases of Gram-positive bacteria. FEMS Microbiology Letters 183, 201–207. https://doi.org/10.1016/S0378-1097(99)00669-2

Igwe, E.I., Shewmaker, P.L., Facklam, R.R., Farley, M.M., Van Beneden, C., Beall, B., 2003. Identification of superantigen genes *speM, ssa,* and *smeZ* in invasive strains of beta-hemolytic Group C and G streptococci recovered from humans. FEMS Microbiology Letters 229, 259–264. https://doi.org/10.1016/S0378-1097(03)00842-5

Ikebe, T., Ato, M., Matsumura, T., Hasegawa, H., Sata, T., Kobayashi, K., Watanabe, H., 2010. Highly frequent mutations in negative regulators of multiple virulence genes in Group A streptococcal toxic shock syndrome isolates. PLoS Pathogens 6, 1–17. https://doi.org/10.1371/journal.ppat.1000832

Ingham, S.C., Wadhera, R.K., Chu, C.H., DeVita, M.D., 2006. Survival of *Streptococcus pyogenes* on foods and food contact surfaces. Journal of Food Protection 69, 1159–1163. https://doi.org/10.4315/0362-028X-69.5.1159

Ingrey, K.T., Ren, J., Prescott, J.F., 2003. A fluoroquinolone induces a novel mitogen-encoding bacteriophage in *Streptococcus canis*. Infection and Immunity 71, 3028–3033. https://doi.org/10.1128/IAI.71.6.3028-3033.2003

Iwasaki, A., Medzhitov, R., 2004. Toll-like receptor control of the adaptive immune responses. Nature Immunology 5, 987–995. https://doi.org/10.1038/ni1112

Jacoba, K.M., Spilkera, T., LiPuma, J.J., Dawida, S.R., Watsona, M.E., 2016. Complete genome sequence of *emm*4 *Streptococcus pyogenes* MEW427, a throat isolate from a child meeting clinical criteria for pediatric autoimmune neuropsychiatric disorders associated with *Streptococcus* (PANDAS). Genome Announcements 4, 1–2. https://doi.org/10.1128/genomeA.00127-16

Jaggi, P., Tanz, R.R., Beall, B., Shulman, S.T., 2005. Age influences the *emm* type distribution of pediatric Group A streptococcal pharyngeal isolates. Pediatric Infectious Disease Journal 24, 1089–1092. https://doi.org/10.1097/01.inf.0000190023.89759.96

Ji, Y., Carlson, B., Kondagunta, A., Cleary, P.P., 1997. Intranasal immunization with C5a peptidase prevents nasopharyngeal colonization of mice by the Group A *Streptococcus*. Infection and Immunity 65, 2080–2087. https://doi.org/10.1128/iai.65.6.2080-2087.1997

Ji, Y., Schnitzler, N., Demaster, E., Cleary, P., 1998. Impact of M49, Mrp, Enn, and C5a peptidase proteins on colonization of the mouse oral mucosa by *Streptococcus pyogenes*. Infection and Immunity 66, 5399–5405. https://doi.org/10.1128/iai.66.11.5399-5405.1998

Jiang, S.M., Ishmael, N., Hotopp, J.D., Puliti, M., Tissi, L., Kumar, N., Cieslewicz, M.J., Tettelin, H., Wessels, M.R., 2008. Variation in the group B *Streptococcus* CsrRS regulon and effects on pathogenicity. Journal of Bacteriology 190, 1956–1965. https://doi.org/10.1128/JB.01677-07

Johnson, D.R., Kaplan, E.L., VanGheem, A., Facklam, R.R., Beall, B., 2006. Characterization of Group A streptococci (*Streptococcus pyogenes*): Correlation of M-protein and *emm*-gene type with T-protein agglutination pattern and serum opacity factor. Journal of Medical Microbiology 55, 157–164. https://doi.org/10.1099/jmm.0.46224-0

Johnsson, E., Andersson, G., Lindahl, G., Hedén, L.O., 1994. Identification of the IgA-binding region in streptococcal protein Arp. The Journal of Immunology 153, 3557 LP – 3564.

Jung, K., Fried, L., Behr, S., Heermann, R., 2012. Histidine kinases and response regulators in networks. Current Opinion in Microbiology 15, 118–124. https://doi.org/10.1016/j.mib.2011.11.009

Kachroo, P., Eraso, J.M., Beres, S.B., Olsen, R.J., Zhu, L., Nasser, W., Bernard, P.E., Cantu, C.C., Saavedra, M.O., Arredondo, M.J., Strope, B., Do, H., Kumaraswami, M., Vuopio, J., Gröndahl-Yli-Hannuksela, K., Kristinsson, K.G., Gottfredsson, M., Pesonen, M., Pensar, J., Davenport, E.R., Clark, A.G., Corander, J., Caugant, D.A., Gaini, S., Magnussen, M.D., Kubiak, S.L., Nguyen, H.A.T., Long, S.W., Porter, A.R., DeLeo, F.R., Musser, J.M., 2019. Integrated analysis of population genomics, transcriptomics and virulence provides novel insights into *Streptococcus pyogenes* pathogenesis. Nature Genetics 51, 548–559. https://doi.org/10.1038/s41588-018-0343-1

Kaneko, J., Kimura, T., Narita, S., Tomita, T., Kamio, Y., 1998. Complete nucleotide sequence and molecular characterization of the temperate staphylococcal bacteriophage φPVL carrying Panton-Valentine leukocidin genes. Gene 215, 57–67. https://doi.org/10.1016/S0378-1119(98)00278-9

Kansal, R.G., McGeer, A., Low, D.E., Norrby-Teglund, A., Kotb, M., 2000. Inverse relation between disease severity and expression of the streptococcal cysteine protease, SpeB, among clonal M1T1 isolates recovered from invasive Group A streptococcal infection cases. Infection and Immunity 68, 6362–6369. https://doi.org/10.1128/IAI.68.11.6362-6369.2000

Kapatai, G., Coelho, J., Platt, S., Chalker, V.J., 2017. Whole genome sequencing of Group A *Streptococcus*: Development and evaluation of an automated pipeline for *emm* gene typing. PeerJ 2017, 1–15. https://doi.org/10.7717/peerj.3226

Kaplan, E.L., Chhatwal, G.S., Rohde, M., 2006. Reduced Ability of Penicillin to Eradicate Ingested Group A Streptococci from Epithelial Cells: Clinical and Pathogenetic Implications. Clinical Infectious Diseases 43, 1398–1406. https://doi.org/10.1086/508773

Kapur, V., Majesky, M.W., Li, L.L., Black, R.A., Musser, J.M., 1993. Cleavage of interleukin 1β (IL-1β) precursor to produce active IL-1β by a conserved extracellular cysteine protease from *Streptococcus pyogenes*. Proceedings of the National Academy of Sciences of the United States of America 90, 7676–7680. https://doi.org/10.1073/pnas.90.16.7676

Kasper, K.J., Zeppa, J.J., Wakabayashi, A.T., Xu, S.X., Mazzuca, D.M., Welch, I., Baroja, M.L., Kotb, M., Cairns, E., Cleary, P.P., Haeryfar, S.M.M., McCormick, J.K., 2014. Bacterial Superantigens Promote Acute Nasopharyngeal Infection by *Streptococcus pyogenes* in a Human MHC Class II-Dependent Manner. PLoS Pathogens 10. https://doi.org/10.1371/journal.ppat.1004155

Kawabata, S., Kuwata, H., Nakagawa, I., Morimatsu, S., Sano, K., Hamada, S., 1999. Capsular hyaluronic acid of Group A streptococci hampers their invasion into human pharyngeal epithelial cells. Microbial Pathogenesis 27, 71–80. https://doi.org/10.1006/mpat.1999.0283

Kawai, T., Akira, S., 2010. The role of pattern-recognition receptors in innate immunity: Update on toll-like receptors. Nature Immunology 11, 373–384. https://doi.org/10.1038/ni.1863

Kazmi, S.U., Kansal, R., Aziz, R.K., Hooshdaran, M., Norrby-Teglund, A., Low, D.E., Halim, A.B., Kotb, M., 2001. Reciprocal, temporal expression of speA and speB by invasive M1T1 Group A streptococcal isolates *in vivo*. Infection and Immunity 69, 4988–4995. https://doi.org/10.1128/IAI.69.8.4988-4995.2001

Kihlberg, B.M., Cooney, J., Caparon, M.G., Olsén, A., Björck, L., 1995. Biological properties of a *Streptococcus pyogenes* mutant generated by Tn916 insertion in *mga*. Microbial Pathogenesis 19, 299–315. https://doi.org/10.1016/S0882-4010(96)80003-9

Kim, Seungwook, Lee, S., Park, H., Kim, Sunjoo, 2019. Predominance of *emm*4 and antibiotic resistance of *Streptococcus pyogenes* in acute pharyngitis in a southern region of Korea. Journal of Medical Microbiology 68, 1053–1058. https://doi.org/10.1099/jmm.0.001005

Kobayashi, A., Iwasaki, H., 2019. “Doughnut” Lesions. New England Journal of Medicine 380, e11. https://doi.org/10.1056/nejmicm1810556

Kolaczkowska, E., Jenne, C.N., Surewaard, B.G.J., Thanabalasuriar, A., Lee, W.Y., Sanz, M.J., Mowen, K., Opdenakker, G., Kubes, P., 2015. Molecular mechanisms of NET formation and degradation revealed by intravital imaging in the liver vasculature. Nature Communications 6, 1–13. https://doi.org/10.1038/ncomms7673

Krebs, B., Kaufhold, A., Boyle, M.D.P., Podbielski, A., 1996. Different alleles of the fcrA/mrp gene of *Streptococcus pyogenes* encode M-related proteins exhibiting an identical immunoglobulin-binding pattern. Medical Microbiology and Immunology 185, 39–47. https://doi.org/10.1007/s004300050013

Kreikemeyer, B., Boyle, M.D., Buttaro, B.A., Heinemann, M., Podbielski, A., 2001. Group A streptococcal growth phase-associated virulence factor regulation by a novel operon (*Fas*) with homologies to two-component-type regulators requires a small RNA molecule. Molecular microbiology 39, 392–406. https://doi.org/10.1046/j.1365-2958.2001.02226.x

Kreikemeyer, B., Klenk, M., Podbielski, A., 2004. The intracellular status of *Streptococcus pyogenes*: Role of extracellular matrix-binding proteins and their regulation. International Journal of Medical Microbiology 294, 177–188. https://doi.org/10.1016/j.ijmm.2004.06.017

Kreikemeyer, B., McIver, K.S., Podbielski, A., 2003. Virulence factor regulation and regulatory networks in *Streptococcus pyogenes* and their impact on pathogen-host interactions. Trends in Microbiology 11, 224–232. https://doi.org/10.1016/S0966-842X(03)00098-2

Krüger, A., Lucchesi, P.M.A., 2015. Shiga toxins and stx phages: highly diverse entities. Microbiology. 161, 451–462. https://doi.org/10.1099/mic.0.000003

Krut, O., Bekeredjian-Ding, I., 2018. Contribution of the Immune Response to Phage Therapy. The Journal of Immunology 200, 3037–3044. https://doi.org/10.4049/jimmunol.1701745

Laabei, M., Ermert, D., 2018. Catch Me if you can: *Streptococcus pyogenes* complement evasion strategies. Journal of Innate Immunity 11, 3–12. https://doi.org/10.1159/000492944

Labrie, S.J., Samson, J.E., Moineau, S., 2010. Bacteriophage resistance mechanisms. Nature Reviews Microbiology 8, 317–327. https://doi.org/10.1038/nrmicro2315

Lamagni, T.L., Neal, S., Keshishian, C., Alhaddad, N., George, R., Duckworth, G., Vuopio-Varkila, J., Efstratiou, A., 2008. Severe *Streptococcus pyogenes* infections, United Kingdom, 2003-2004. Emerging Infectious Diseases 14, 201–209. https://doi.org/10.3201/eid1402.070888

Lamagni, T.L., Neal, S., Keshishian, C., Powell, D., Potz, N., Pebody, R., George, R., Duckworth, G., Vuopio-Varkila, J., Efstratiou, A., 2009. Predictors of death after severe *Streptococcus pyogenes* infection. Emerging Infectious Diseases 15, 1304–1307. https://doi.org/10.3201/eid1508.090264

Lamb, L. E., Siggins, M.K., Scudamore, C., Macdonald, W., Turner, C.E., Lynskey, N.N., Tan, L.K.K., Sriskandan, S., 2018. Impact of contusion injury on intramuscular *emm*1 Group A *Streptococcus* infection and lymphatic spread. Virulence 9, 1074–1084. https://doi.org/10.1080/21505594.2018.1482180

Lamb, L E, Zhi, X., Alam, F., Pyzio, M., Scudamore, C.L., Wiles, S., Sriskandan, S., 2018. Modelling invasive Group A streptococcal disease using bioluminescence. BMC Microbiology 18, 60. https://doi.org/10.1186/s12866-018-1200-1

Lancefield, R.C., 1959. Persistence of type-specific antibodies in man following infection with Group A streptococci. The Journal of experimental medicine 110, 271–292. https://doi.org/10.1084/jem.110.2.271

LaPenta, D., Rubens, C., Chi, E., Cleary, P.P., 1994. Group A streptococci efficiently invade human respiratory epithelial cells. Proceedings of the National Academy of Sciences of the United States of America 91, 12115–12119. https://doi.org/10.1073/pnas.91.25.12115

Lefébure, T., Richards, V.P., Lang, P., Pavinski-Bitar, P., Stanhope, M.J., 2012. Gene repertoire evolution of *Streptococcus pyogenes* inferred from phylogenomic analysis with *Streptococcus canis* and *Streptococcus dysgalactiae*. PLoS ONE 7. https://doi.org/10.1371/journal.pone.0037607

Lei, B., DeLeo, F.R., Hoe, N.P., Graham, M.R., Mackie, S.M., Cole, R.L., Liu, M., Hill, H.R., Low, D.E., Federle, M.J., Scott, J.R., Musser, J.M., 2001. Evasion of human innate and acquired immunity by a bacterial homolog of CD11b that inhibits opsonophagocytosis. Nature Medicine 7, 1298–1305. https://doi.org/10.1038/nm1201-1298

Lei, B., DeLeo, F.R., Reid, S.D., Voyich, J.M., Magoun, L., Liu, M., Braughton, K.R., Ricklefs, S., Hoe, N.P., Cole, R.L., Leong, J.M., Musser, J.M., 2002. Opsonophagocytosis-inhibiting Mac protein of Group A *Streptococcus*: Identification and characteristics of two genetic complexes. Infection and Immunity 70, 6880–6890. https://doi.org/10.1128/IAI.70.12.6880-6890.2002

Lembke, C., Podbielski, A., Hidalgo-Grass, C., Jonas, L., Hanski, E., Kreikemeyer, B., 2006. Characterization of biofilm formation by clinically relevant serotypes of Group A streptococci. Applied and Environmental Microbiology 72, 2864–2875. https://doi.org/10.1128/AEM.72.4.2864-2875.2006

Li, J., Liu, G., Feng, W., Zhou, Y., Liu, M., Wiley, J.A., Lei, B., 2014. Neutrophils select hypervirulent CovRS mutants of M1T1 Group A *Streptococcus* during subcutaneous infection of mice. Infection and Immunity 82, 1579–1590. https://doi.org/10.1128/IAI.01458-13

Lieber, M., Smith, B., Szakal, A., Nelson-Rees, W., Todaro, G., 1976. A continuous tumor-cell line from a human lung carcinoma with properties of type II alveolar epithelial cells. International journal of cancer 17, 62–70. https://doi.org/10.1002/ijc.2910170110

Lin, A., Loughman, J.A., Zinselmeyer, B.H., Miller, M.J., Caparon, M.G., 2009. Streptolysin S inhibits neutrophil recruitment during the early stages of *Streptococcus pyogenes* infection. Infection and Immunity 77, 5190–5201. https://doi.org/10.1128/IAI.00420-09

Liu, Z., Treviño, J., Ramirez-Peña, E., Sumby, P., 2012. The small regulatory RNA FasX controls pilus expression and adherence in the human bacterial pathogen Group A *Streptococcus*. Molecular microbiology 86, 140–154. https://doi.org/10.1111/j.1365-2958.2012.08178.x

Lizano, S., Luo, F., Tengra, F.K., Bessen, D.E., 2008. Impact of orthologous gene replacement on the circuitry governing pilus gene transcription in streptococci. PLoS ONE. 3(10): e3450. https://doi.org/10.1371/journal.pone.0003450

Loh, J.M.S., Adenwalla, N., Wiles, S., Proft, T., 2013. *Galleria mellonella* larvae as an infection model for Group A *Streptococcus*. Virulence 4, 419–428. https://doi.org/10.4161/viru.24930

Loof, T.G., Goldmann, O., Gessner, A., Herwald, H., Medina, E., 2010. Aberrant inflammatory response to

Streptococcus pyogenes in mice lacking myeloid differentiation factor 88. American Journal of Pathology 176, 754–763. https://doi.org/10.2353/ajpath.2010.090422

Loof, T.G., Goldmann, O., Medina, E., 2008. Immune recognition of *Streptococcus pyogenes* by dendritic cells. Infection and Immunity 76, 2785–2792. https://doi.org/10.1128/IAI.01680-07

Loof, T.G., Rohde, M., Chhatwal, G.S., Jung, S., Medina, E., 2007. The Contribution of Dendritic Cells to Host Defenses against *Streptococcus pyogenes* . The Journal of Infectious Diseases 196, 1794–1803. https://doi.org/10.1086/523647

López, E., Domenech, A., Ferrándiz, M.J., Frias, M.J., Ardanuy, C., Ramirez, M., García, E., Liñares, J., De La Campa, A.G., 2014. Induction of prophages by fluoroquinolones in *Streptococcus pneumoniae*: Implications for emergence of resistance in genetically-related clones. PLoS ONE 9, 1–8. https://doi.org/10.1371/journal.pone.0094358

Luca-Harari, B., Darenberg, J., Neal, S., Siljander, T., Strakova, L., Tanna, A., Creti, R., Ekelund, K., Koliou, M., Tassios, P.T., Van Der Linden, M., Straut, M., Vuopio-Varkila, J., Bouvet, A., Efstratiou, A., Schalén, C., Henriques-Normark, B., Jasir, A., 2009. Clinical and microbiological characteristics of severe *Streptococcus pyogenes* disease in Europe. Journal of Clinical Microbiology 47, 1155–1165. https://doi.org/10.1128/JCM.02155-08

Luke, K., Radek, A., Liu, X.P., Campbell, J., Uzan, M., Haselkorn, R., Kogan, Y., 2002. Microarray analysis of gene expression during bacteriophage T4 infection. Virology 299, 182–191. https://doi.org/10.1006/viro.2002.1409

Lukomski, S., Burns, E.H., Wyde, P.R., Podbielski, A., Rurangirwa, J., Moore-Poveda, D.K., Musser, J.M., 1998. Genetic inactivation of an extracellular cysteine protease (SpeB) expressed by *Streptococcus pyogenes* decreases resistance to phagocytosis and dissemination to organs. Infection and Immunity 66, 771–776. https://doi.org/10.1128/iai.66.2.771-776.1998

Lukomski, S., Montgomery, C.A., Rurangirwa, J., Geske, R.S., Barrish, J.P., Adams, G.J., Musser, J.M., 1999. Extracellular cysteine protease produced by *Streptococcus pyogenes* participates in the pathogenesis of invasive skin infection and dissemination in mice. Infection and Immunity 67, 1779–1788.

Luo, F., Lizano, S., Bessen, D.E., 2008. Heterogeneity in the polarity of Nra regulatory effects on streptococcal pilus gene transcription and virulence. Infection and Immunity 76, 2490–2497. https://doi.org/10.1128/IAI.01567-07

Lynskey, N.N., Banerji, S., Johnson, L.A., Holder, K.A., Reglinski, M., Wing, P.A.C., Rigby, D., Jackson, D.G., Sriskandan, S., 2015. Rapid Lymphatic Dissemination of Encapsulated Group A Streptococci via Lymphatic Vessel Endothelial Receptor-1 Interaction. PLoS Pathogens 11, 1–19. https://doi.org/10.1371/journal.ppat.1005137

Lynskey, N.N., Goulding, D., Gierula, M., Turner, C.E., Dougan, G., Edwards, R.J., Sriskandan, S., 2013. RocA Truncation Underpins Hyper-Encapsulation, Carriage Longevity and Transmissibility of Serotype M18 Group A Streptococci. PLoS Pathogens 9, 1–15. https://doi.org/10.1371/journal.ppat.1003842

Lynskey, N.N., Jauneikaite, E., Li, H.K., Zhi, X., Turner, C.E., Mosavie, M., Pearson, M., Asai, M., Lobkowicz,

L., Chow, J.Y., Parkhill, J., Lamagni, T., Chalker, V.J., Sriskandan, S., 2019a. Emergence of dominant toxigenic M1T1 *Streptococcus pyogenes* clone during increased scarlet fever activity in England: a population-based molecular epidemiological study. The Lancet Infectious Diseases 19, 1209–1218. https://doi.org/10.1016/S1473-3099(19)30446-3

Lynskey, N.N., Reglinski, M., Calay, D., Siggins, M.K., Mason, J.C., Botto, M., Sriskandan, S., 2017. Multi-functional mechanisms of immune evasion by the streptococcal complement inhibitor C5a peptidase. PLoS Pathogens 13, 1–29. https://doi.org/10.1371/journal.ppat.1006493

Lynskey, N.N., Velarde, J.J., Finn, M.B., Dove, S.L., Wessels, M.R., 2019b. RocA binds CsrRS to modulate CsrRS-mediated gene regulation in Group A *Streptococcus*. mBio 10, 1–15. https://doi.org/10.1128/mBio.01495-19

Lyon, W.R., Caparon, M.G., 2004. Role for Serine Protease HtrA (DegP) of *Streptococcus pyogenes* in the Biogenesis of Virulence Factors SpeB and the Hemolysin Streptolysin S. Infection and Immunity 72, 1618–1625. https://doi.org/10.1128/IAI.72.3.1618-1625.2004

Ma, F., Guo, X., Fan, H., 2017. Extracellular Nucleases of *Streptococcus equi* subsp. *zooepidemicus* Degrade Neutrophil Extracellular Traps and Impair Macrophage Activity of the Host. Applied and environmental microbiology 83. https://doi.org/10.1128/AEM.02468-16

Ma, X., Kikuta, H., Ishiguro, N., Yoshioka, M., Ebihara, T., Murai, T., Kobayashi, I., Kobayashi, K., 2002. Association of the prtF1 gene (encoding fibronectin-binding protein F1) and the sic gene (encoding the streptococcal inhibitor of complement) with *emm* types of Group A streptococci isolated from Japanese children with pharyngitis. Journal of Clinical Microbiology 40, 3835–3837. https://doi.org/10.1128/JCM.40.10.3835-3837.2002

Machado, N.O., 2011. Necrotizing fasciitis: The importance of early diagnosis, prompt surgical debridement and adjuvant therapy. North American Journal of Medical Sciences 3, 107–118. https://doi.org/10.4297/najms.2011.3107

Maclennan, A.P., 1956a. The production of capsules, hyaluronic acid and hyaluronidase by Group A and group C streptococci. Journal of general microbiology 14, 134–142. https://doi.org/10.1099/00221287-14-1-134

Madden, J.C., Ruiz, N., Caparon, M., 2001. Cytolysin-mediated translocation (CMT): A functional equivalent of type III secretion in Gram-positive bacteria. Cell 104, 143–152. https://doi.org/10.1016/S0092-8674(01)00198-2

Maiques, E., Úbeda, C., Campoy, S., Salvador, N., Lasa, Í., Novick, R.P., Barbé, J., Penadés, J.R., 2006. β-lactam antibiotics induce the SOS response and horizontal transfer of virulence factors in *Staphylococcus aureus*. Journal of Bacteriology 188, 2726–2729. https://doi.org/10.1128/JB.188.7.2726-2729.2006

Marks, Laura R., Mashburn-Warren, L., Federle, M.J., Hakansson, A.P., 2014. *Streptococcus pyogenes* biofilm growth *in vitro* and *in vivo* and its role in colonization, virulence, and genetic exchange. Journal of Infectious Diseases 210, 25–34. https://doi.org/10.1093/infdis/jiu058

Marks, Laura R, Reddinger, R.M., Hakansson, A.P., 2014. Biofilm formation enhances fomite survival of *Streptococcus pneumoniae* and *Streptococcus pyogenes*. Infection and immunity 82, 1141–1146. https://doi.org/10.1128/IAI.01310-13

Marouni, M.J., Barzilai, A., Keller, N., Rubinstein, E., Sela, S., 2004. Intracellular survival of persistent Group A streptococci in cultured epithelial cells. International Journal of Medical Microbiology 294, 27–33. https://doi.org/10.1016/j.ijmm.2004.01.001

Martin, J., 2016. The *Streptococcus pyogenes* Carrier State. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 1–13.

Mashburn-Warren, L., Goodman, S.D., Federle, M.J., Prehna, G., 2018. The conserved mosaic prophage protein paratox inhibits the natural competence regulator ComR in *Streptococcus pyogenes*. Scientific Reports 8, 1–15. https://doi.org/10.1038/s41598-018-34816-7

Mashburn-Warren, L., Morrison, D.A., Federle, M.J., 2012. The cryptic competence pathway in *Streptococcus pyogenes* is controlled by a peptide pheromone. Journal of Bacteriology 194, 4589–4600. https://doi.org/10.1128/JB.00830-12

Massell, B.F., Honikman, L.H., Amezcua, J., 1969. Rheumatic Fever Following Streptococcal Vaccination: Report of Three Cases. JAMA 207, 1115–1119. https://doi.org/10.1001/jama.1969.03150190037007

Matsui, H., Sekiya, Y., Takahashi, Tetsufumi, Nakamura, M., Imanishi, K., Yoshida, H., Murayama, S.Y., Takahashi, Takashi, Tsuchimoto, K., Uchiyama, T., Ubukata, K., 2011. Dermal mast cells reduce progressive tissue necrosis caused by subcutaneous infection with *Streptococcus pyogenes* in mice. Journal of Medical Microbiology 60, 128–134. https://doi.org/10.1099/jmm.0.020495-0

McCullor, K., Postoak, B., Rahman, M., King, C., Michael McShana, W., 2018. Genomic sequencing of high-efficiency transducing streptococcal bacteriophage A25: Consequences of escape from lysogeny. Journal of Bacteriology 200. https://doi.org/10.1128/JB.00358-18

McIver, K.S., Heath, A.S., Scott, J.R., 1995. Regulation of virulence by environmental signals in Group A streptococci: Influence of osmolarity, temperature, gas exchange, and iron limitation on *emm* transcription. Infection and Immunity 63, 4540–4542. https://doi.org/10.1128/iai.63.11.4540-4542.1995

McNeil, S.A., Halperin, S.A., Langley, J.M., Smith, B., Warren, A., Sharratt, G.P., Baxendale, D.M., Reddish, M.A., Hu, M.C., Stroop, S.D., Linden, J., Fries, L.F., Vink, P.E., Dale, J.B., 2005. Safety and Immunogenicity of 26-Valent Group A *Streptococcus* Vaccine in Healthy Adult Volunteers. Clinical Infectious Diseases 41, 1114–1122. https://doi.org/10.1086/444458

McShan, W.M., Nguyen, S. V., 2019. The bacteriophages of *Streptococcus pyogenes*. *Streptococcus pyogenes* : Basic Biology to Clinical Manifestations 1–22. https://doi.org/10.1128/microbiolspec.GPP3-0059-2018.Correspondence

Mearkle, R., Saavedra-Campos, M., Lamagni, T., Usdin, M., Coelho, J., Chalker, V., Sriskandan, S., Cordery, R., Rawlings, C., Balasegaram, S., 2017. Household transmission of invasive Group A *Streptococcus* infections in England: a population-based study, 2009, 2011 to 2013. Euro surveillance : bulletin Europeen sur les maladies transmissibles = European communicable disease bulletin 22. https://doi.org/10.2807/1560-7917.ES.2017.22.19.30532

Medina, E., Goldmann, O., Rohde, M., Lengeling, A., Chhatwals, G.S., 2001. Genetic Control of Susceptibility to Group A Streptococcal Infection in Mice. The Journal of Infectious Diseases 184, 846–852. https://doi.org/10.1086/323292

Medina, E., Rohde, M., Chhatwal, G.S., 2003. Intracellular survival of *Streptococcus pyogenes* in polymorphonuclear cells results in increased bacterial virulence. Infection and Immunity 71, 5376–5380. https://doi.org/10.1128/IAI.71.9.5376-5380.2003

Menschner, L., Falke, U., Konrad, P., Toepfner, N., Berner, R., 2020. Survival of Group A *Streptococcus* (GAS) is Enhanced Under Desiccated Culture Conditions. Current Microbiology. https://doi.org/10.1007/s00284-020-01967-8

Michel, B., 2005. After 30 years of study, the bacterial SOS response still surprises us. PLoS Biology 3, 1174–1176. https://doi.org/10.1371/journal.pbio.0030255

Michos, A., Gryllos, I., Håkansson, A., Srivastava, A., Kokkotou, E., Wessels, M.R., 2006. Enhancement of streptolysin O activity and intrinsic cytotoxic effects of the Group A streptococcal toxin, NAD-glycohydrolase. Journal of Biological Chemistry 281, 8216–8223. https://doi.org/10.1074/jbc.M511674200

Miller-Ensminger, T., Garretto, A., Stark, N., Putonti, C., 2020. Mimicking prophage induction in the body: induction in the lab with pH gradients. PeerJ 8, e9718. https://doi.org/10.7717/peerj.9718

Miller, E.W., Cao, T.N., Pflughoeft, K.J., Sumby, P., 2014. RNA-mediated regulation in Gram-positive pathogens: an overview punctuated with examples from the Group A *Streptococcus*. Molecular microbiology 94, 9–20. https://doi.org/10.1111/mmi.12742

Mishalian, I., Ordan, M., Peled, A., Maly, A., Eichenbaum, M.B., Ravins, M., Aychek, T., Jung, S., Hanski, E., 2011. Recruited Macrophages Control Dissemination of Group A *Streptococcus* from Infected Soft Tissues . The Journal of Immunology 187, 6022–6031. https://doi.org/10.4049/jimmunol.1101385

Mishra, P., Akhtar, M.S., Bhakuni, V., 2006. Unusual structural features of the bacteriophage-associated hyaluronate lyase (hylp2). Journal of Biological Chemistry 281, 7143–7150. https://doi.org/10.1074/jbc.M510991200

Mishra, P., Bhakuni, V., 2009. Self-assembly of bacteriophage-associated hyaluronate lyase (HYLP2) into an enzymatically active fibrillar film. Journal of Biological Chemistry 284, 5240–5249. https://doi.org/10.1074/jbc.M806730200

Mogensen, T.H., 2009. Pathogen recognition and inflammatory signaling in innate immune defenses. Clinical Microbiology Reviews 22, 240–273. https://doi.org/10.1128/CMR.00046-08

Mortensen, R., Nissen, T.N., Blauenfeldt, T., Christensen, J.P., Andersen, P., Dietrich, J., 2015. Adaptive Immunity against *Streptococcus pyogenes* in Adults Involves Increased IFN-γ and IgG3 Responses Compared with Children . The Journal of Immunology 195, 1657–1664. https://doi.org/10.4049/jimmunol.1500804

Murialdo, H., Becker, A., 1978. Head morphogenesis of complex double-stranded deoxyribonucleic acid bacteriophages. Microbiological reviews 42, 529–576.

Musser, J.M., Beres, S.B., Zhu, L., Olsen, R.J., Vuopio, J., Hyyryläinen, H.-L., Gröndahl-Yli-Hannuksela, K., Kristinsson, K.G., Darenberg, J., Henriques-Normark, B., Hoffmann, S., Caugant, D.A., Smith, A.J., Lindsay, D.S.J., Boragine, D., Palzkill, T., 2020. Reduced *in vitro* susceptibility of *Streptococcus pyogenes* to beta-lactam antibiotics associated with mutations in the *pbp2x* gene is geographically widespread . Journal of Clinical Microbiology. https://doi.org/10.1128/jcm.01993-19

Musser, J.M., DeLeo, F.R., 2005. Toward a genome-wide systems biology analysis of host-pathogen interactions in Group A *Streptococcus*. American Journal of Pathology 167, 1461–1472. https://doi.org/10.1016/S0002-9440(10)61232-1

Musser, J.M., Gray, B.M., Schlievert, P.M., Pichichero, M.E., 1992. Erratum: *Streptococcus pyogenes* pharyngitis: Characterization of strains by multilocus enzyme genotype, M and T protein serotype, and pyrogenic exotoxin gene probing (Journal of Clinical Microbiology (1992) 30 (602)). Journal of Clinical Microbiology 30, 1905. https://doi.org/10.1128/JCM.30.7.1905-1905.1992

Nakagawa, I., Kurokawa, K., Yamashita, A., Nakata, M., Tomiyasu, Y., Okahashi, N., Kawabata, S., Yamazaki, K., Shiba, T., Yasunaga, T., Hayashi, H., Hattori, M., Hamada, S., 2003. Genome Sequence of an M3 Strain of *Streptococcus pyogenes* reveals a large-scale genomic rearrangement in invasive strains and new insights into phage evolution. Genome 1042–1055. https://doi.org/10.1101/gr.1096703.1

Nanda, A.M., Thormann, K., Frunzke, J., 2015. Impact of spontaneous prophage induction on the fitness of bacterial populations and host-microbe interactions. Journal of Bacteriology 197, 410–419. https://doi.org/10.1128/JB.02230-14

Nasser, W., Beres, S.B., Olsen, R.J., Dean, M.A., Rice, K.A., Long, S.W., Kristinsson, K.G., Gottfredsson, M., Vuopio, J., Raisanen, K., Caugant, D.A., Steinbakk, M., Low, D.E., McGeer, A., Darenberg, J., Henriques-Normark, B., Van Beneden, C.A., Hoffmann, S., Musser, J.M., 2014. Evolutionary pathway to increased virulence and epidemic Group A *Streptococcus* disease derived from 3,615 genome sequences. Proceedings of the National Academy of Sciences of the United States of America 111. https://doi.org/10.1073/pnas.1403138111

Nelson, D.C., Garbe, J., Collin, M., 2011. The cysteine proteinase SpeB from *Streptococcus pyogenes* – a potent modifier of immunologically important host and bacterial proteins . Biological Chemistry 392. https://doi.org/10.1515/bc-2011-208

Nguyen, S. V., McShan, W.M., 2014. Chromosomal islands of *Streptococcus pyogenes* and related streptococci: Molecular switches for survival and virulence. Frontiers in Cellular and Infection Microbiology 4, 1–14. https://doi.org/10.3389/fcimb.2014.00109

Nizet, V., Beall, B., Bast, D.J., Datta, V., Kilburn, L., Low, D.E., De Azavedo, J.C.S., 2000. Genetic locus for streptolysin S production by Group A *Streptococcus*. Infection and Immunity 68, 4245–4254. https://doi.org/10.1128/IAI.68.7.4245-4254.2000

Nobrega, F.L., Vlot, M., de Jonge, P.A., Dreesens, L.L., Beaumont, H.J.E., Lavigne, R., Dutilh, B.E., Brouns, S.J.J., 2018. Targeting mechanisms of tailed bacteriophages. Nature Reviews Microbiology 16, 760–773. https://doi.org/10.1038/s41579-018-0070-8

Nooh, M.M., Aziz, R.K., Kotb, M., Eroshkin, A., Chuang, W.J., Proft, T., Kansal, R., 2006. Streptococcal mitogenic exotoxin, SmeZ, is the most susceptible M1T1 streptococcal superantigen to degradation by the streptococcal cysteine protease, SpeB. Journal of Biological Chemistry 281, 35281–35288. https://doi.org/10.1074/jbc.M605544200

Nordenfelt, P., Waldemarson, S., Linder, A., Mörgelin, M., Karlsson, C., Malmström, J., Björck, L., 2012. Antibody orientation at bacterial surfaces is related to invasive infection. Journal of Experimental Medicine 209, 2367–2381. https://doi.org/10.1084/jem.20120325

Norrby-Teglund, A., Kaul, R., Low, D.E., McGeer, A., Newton, D.W., Andersson, J., Andersson, U., Kotb, M., 1996. Plasma from patients with severe invasive Group A streptococcal infections treated with normal polyspecific IgG inhibits streptococcal superantigen-induced T cell proliferation and cytokine production. The Journal of Immunology 156, 3057 LP – 3064.

Novick, R.P., 2003. Mobile genetic elements and bacterial toxinoses: The superantigen-encoding pathogenicity islands of *Staphylococcus aureus*. Plasmid 49, 93–105. https://doi.org/10.1016/S0147-619X(02)00157-9

Novick, R.P., Christie, G.E., Penadés, J.R., 2010. The phage-related chromosomal islands of Gram-positive bacteria. Nature Reviews Microbiology 8, 541–551. https://doi.org/10.1038/nrmicro2393

Nozawa, T., Furukawa, N., Aikawa, C., Watanabe, T., Haobam, B., Kurokawa, K., Maruyama, F., Nakagawa, I., 2011. CRISPR inhibition of prophage acquisition in *Streptococcus pyogenes*. PLoS ONE 6(5), 19543. https://doi.org/10.1371/journal.pone.0019543

Nuwayhid, Z.B., Aronoff, D.M., Mulla, Z.D., 2007. Blunt Trauma as a Risk Factor for Group A Streptococcal Necrotizing Fasciitis. Annals of Epidemiology 17, 878–881. https://doi.org/10.1016/j.annepidem.2007.05.011

O'Neill, A.M., Thurston, T.L., Holden, D.W., 2016. Cytosolic Replication of Group A *Streptococcus* in Human. mBio 7, 1–12. https://doi.org/10.1128/mBio.00020-16.Invited

O'Toole, G.A., 2011. Microtiter dish biofilm formation assay. Journal of visualized experiments : JoVE. https://doi.org/10.3791/2437

Oda, M., Domon, H., Kurosawa, M., Isono, T., Maekawa, T., Yamaguchi, M., Kawabata, S., Terao, Y., 2017. *Streptococcus pyogenes* phospholipase A2 induces the expression of adhesion molecules on human umbilical vein endothelial cells and aorta of mice. Frontiers in Cellular and Infection Microbiology 7, 1–9. https://doi.org/10.3389/fcimb.2017.00300

Olsen, R.J., Ebru Watkins, M., Cantu, C.C., Beres, S.B., Musser, J.M., 2011. Virulence of serotype M3 Group A *Streptococcus* strains in wax worms (*Galleria mellonella* larvae). Virulence 2, 111–119. https://doi.org/10.4161/viru.2.2.14338

Olsen, R.J., Laucirica, D.R., Watkins, M.E., Feske, M.L., Garcia-Bustillos, J.R., Vu, C., Cantu, C., Shelburne, S.A., Fittipaldi, N., Kumaraswami, M., Shea, P.R., Flores, A.R., Beres, S.B., Lovgren, M., Tyrrell, G.J., Efstratiou, A., Low, D.E., Van Beneden, C.A., Musser, J.M., 2012. Polymorphisms in regulator of protease B (RopB) alter disease phenotype and strain virulence of serotype M3 Group A *Streptococcus*. Journal of Infectious Diseases 205, 1719–1729. https://doi.org/10.1093/infdis/jir825

Oppenheim, A.B., Kobiler, O., Stavans, J., Court, D.L., Adhya, S., 2005. Switches in Bacteriophage Lambda Development. Annual Review of Genetics 39, 409–429. https://doi.org/10.1146/annurev.genet.39.073003.113656

Orlova, E. V., Gowen, B., Dröge, A., Stiege, A., Weise, F., Lurz, R., Van Heel, M., Tavares, P., 2003. Structure of a viral DNA gatekeeper at 10 Å resolution by cryo-electron microscopy. EMBO Journal 22, 1255–1262. https://doi.org/10.1093/emboj/cdg123

Påhlman, L.I., Mörgelin, M., Eckert, J., Johansson, L., Russell, W., Riesbeck, K., Soehnlein, O., Lindbom, L., Norrby-Teglund, A., Schumann, R.R., Björck, L., Herwald, H., 2006. Streptococcal M Protein: A Multipotent and Powerful Inducer of Inflammation. The Journal of Immunology 177, 1221–1228. https://doi.org/10.4049/jimmunol.177.2.1221

Pan, X., Ge, J., Li, M., Wu, B., Wang, C., Wang, J., Feng, Y., Yin, Z., Zheng, F., Cheng, G., Sun, W., Ji, H., Hu, D., Shi, P., Feng, X., Hao, X., Dong, R., Hu, F., Tang, J., 2009. The orphan response regulator CovR: A globally negative modulator of virulence in *Streptococcus suis* serotype 2. Journal of Bacteriology 191, 2601–2612. https://doi.org/10.1128/JB.01309-08

Panchaud, A., Guy, L., Collyn, F., Haenni, M., Nakata, M., Podbielski, A., Moreillon, P., Roten, C.A.H., 2009. M-protein and other intrinsic virulence factors of *Streptococcus pyogenes* are encoded on an ancient pathogenicity island. BMC Genomics 10, 1–11. https://doi.org/10.1186/1471-2164-10-198

Pang, T., Fleming, T.C., Pogliano, K., Young, R., 2013. Visualization of pinholin lesions *in vivo*. Proceedings of the National Academy of Sciences of the United States of America 110. https://doi.org/10.1073/pnas.1222283110

Participants, T.P. of I.G.A.S.I.W., 2002. Prevention of Invasive Group A Streptococcal Disease among Household Contacts of Case Patients and among Postpartum and Postsurgical Patients: Recommendations from the Centers for Disease Control and Prevention. Clinical Infectious Diseases 35, 950–959. https://doi.org/10.1086/342692

Pastural, É., McNeil, S.A., MacKinnon-Cameron, D., Ye, L., Langley, J.M., Stewart, R., Martin, L.H., Hurley, G.J., Salehi, S., Penfound, T.A., Halperin, S., Dale, J.B., 2020. Safety and immunogenicity of a 30-valent M protein-based Group A streptococcal vaccine in healthy adult volunteers: A randomized, controlled phase I study. Vaccine 38, 1384–1392. https://doi.org/10.1016/j.vaccine.2019.12.005

Peiser, L., Mukhopadhyay, S., Gordon, S., 2002. Scavenger receptors in innate immunity. Current Opinion in Immunology 14, 123–128. https://doi.org/10.1016/S0952-7915(01)00307-7

Pelucchi, C., Grigoryan, L., Galeone, C., Esposito, S., Huovinen, P., Little, P., Verheij, T., 2012. Guideline for the management of acute sore throat: ESCMID Sore Throat Guideline Group. Clinical Microbiology and Infection 18, 1–28. https://doi.org/10.1111/j.1469-0691.2012.03766.x

Penadés, J.R., Christie, G.E., 2015. The Phage-Inducible Chromosomal Islands: A Family of Highly Evolved Molecular Parasites. Annual Review of Virology 2, 181–201. https://doi.org/10.1146/annurev-virology-031413-085446

Pence, M.A., Rooijakkers, S.H.M., Cogen, A.L., Cole, J.N., Hollands, A., Gallo, R.L., Nizet, V., 2010. Streptococcal inhibitor of complement promotes innate immune resistance phenotypes of invasive M1T1 Group A *Streptococcus*. Journal of Innate Immunity 2, 587–595. https://doi.org/10.1159/000317672

Pereira, L.B., 2014. Impetigo. Anais Brasileiros de Dermatologia 89, 293–299. https://doi.org/10.1590/abd1806-4841.20142283

Perez, N., Treviño, J., Liu, Z., Ho, S.C.M., Babitzke, P., Sumby, P., 2009. A genome-wide analysis of small regulatory RNAs in the human pathogen Group A *Streptococcus*. PLoS ONE 4. https://doi.org/10.1371/journal.pone.0007668

Persson, H., Vindebro, R., Von Pawel-Rammingen, U., 2013. The streptococcal cysteine protease SpeB is not a natural immunoglobulin-cleaving enzyme. Infection and Immunity 81, 2236–2241. https://doi.org/10.1128/IAI.00168-13

Pichichero, M.E., Casey, J.R., 2007. Systematic review of factors contributing to penicillin treatment failure in *Streptococcus pyogenes* pharyngitis. Otolaryngology - Head and Neck Surgery 137, 851–857. https://doi.org/10.1016/j.otohns.2007.07.033

Platt, N., Gordon, S., 2001. Is the class A macrophage scavenger receptor (SR-A) multifunctional? - The mouse's tale. Journal of Clinical Investigation 108, 649–654. https://doi.org/10.1172/JCI200113903

Podbielski, A., Peterson, J.A., Cleary, P., 1992. Surface protein-CAT reporter fusions demonstrate differential gene expression in the wr regulon of *Streptococcus pyogenes*. Molecular Microbiology 6, 2253–2265. https://doi.org/10.1111/j.1365-2958.1992.tb01401.x

Podbielski, A., Schnitzler, N., Beyhs, P., Boyle, M.D., 1996. M-related protein (Mrp) contributes to Group A streptococcal resistance to phagocytosis by human granulocytes. Molecular microbiology 19, 429–441. https://doi.org/10.1046/j.1365-2958.1996.377910.x

Podbielski, A., Woischnik, M., Leonard, B.A.B., Schmidt, K.H., 1999. Characterization of *nra*, a global negative regulator gene in Group A streptococci. Molecular Microbiology 31, 1051–1064. https://doi.org/10.1046/j.1365-2958.1999.01241.x

Pohane, A.A., Jain, V., 2015. Insights into the regulation of bacteriophage endolysin: Multiple means to the same end. Microbiology. 161, 2269–2276. https://doi.org/10.1099/mic.0.000190

Pospiech, A., Neumann, B., 1995. A versatile quick-prep of genomic DNA from gram-positive bacteria. Trends in Genetics. 11, 217–218. https://doi.org/10.1016/s0168-9525(00)89052-6

Proft, T., Fraser, J.D., 2016. Streptococcal Superantigens: Biological properties and potential role in disease. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 1–41.

Proft, T., Sriskandan, S., Yang, L., Fraser, J.D., 2003. Superantigens and streptococcal toxic shock syndrome. Emerging Infectious Diseases 9, 1211–1218. https://doi.org/10.3201/eid0910.030042

Raeder, R., Woischnik, M., Podbielski, A., Boyle, M.D.P., 1998. A secreted streptococcal cysteine protease can cleave a surface- expressed M1 protein and alter the immunoglobulin binding properties. Research in Microbiology 149, 539–548. https://doi.org/10.1016/S0923-2508(99)80001-1

Raithatha, A.H., Bryden, D.C., 2012. Use of intravenous immunoglobulin therapy in the treatment of septic shock, in particular severe invasive Group A streptococcal disease. Indian journal of critical care medicine : peer-reviewed, official publication of Indian Society of Critical Care Medicine. https://doi.org/10.4103/0972-5229.94433

Ramirez-Peña, E., Treviño, J., Liu, Z., Perez, N., Sumby, P., 2010. The Group A *Streptococcus* small regulatory RNA FasX enhances streptokinase activity by increasing the stability of the *ska* mRNA transcript. Molecular microbiology 78, 1332–1347. https://doi.org/10.1111/j.1365-2958.2010.07427.x

Rao, V.B., Feiss, M., 2008. The Bacteriophage DNA Packaging Motor. Annual Review of Genetics 42, 647–681. https://doi.org/10.1146/annurev.genet.42.110807.091545

Reglinski, M., Gierula, M., Lynskey, N.N., Edwards, R.J., Sriskandan, S., 2015. Identification of the *Streptococcus pyogenes* surface antigens recognised by pooled human immunoglobulin. Scientific reports 5, 15825. https://doi.org/10.1038/srep15825

Reglinski, M., Lynskey, N.N., Sriskandan, S., 2016. Modification of the classical Lancefield assay of Group A streptococcal killing to reduce inter-donor variation. Journal of microbiological methods 124, 69–71. https://doi.org/10.1016/j.mimet.2016.03.015

Reglinski, M., Sriskandan, S., 2014. *Streptococcus pyogenes*, Molecular Medical Microbiology: Second Edition. Elsevier Ltd. https://doi.org/10.1016/B978-0-12-397169-2.00038-X

Reglinski, M., Sriskandan, S., Turner, C.E., 2019. Identification of two new core chromosome-encoded superantigens in *Streptococcus pyogenes*; speQ and speR. Journal of Infection 78, 358–363. https://doi.org/10.1016/j.jinf.2019.02.005

Rego, S., Heal, T.J., Pidwill, G.R., Till, M., Robson, A., Lamont, R.J., Sessions, R.B., Jenkinson, H.F., Race, P.R., Nobbs, A.H., 2016. Structural and functional analysis of cell wall-anchored polypeptide adhesin BspA in *Streptococcus* agalactiae. Journal of Biological Chemistry 291, 15985–16000. https://doi.org/10.1074/jbc.M116.726562

Remmington, A., Haywood, S., Edgar, J., Turner, C., 2020. Cryptic prophages within a *Streptococcus pyogenes* genotype *emm*4 lineage. Microbial Genomics. https://doi.org/10.1101/2020.05.19.103838

Remmington, A., Turner, C.E., 2018. The DNases of pathogenic lancefield streptococci. Microbiology. 164, 242–250. https://doi.org/10.1099/mic.0.000612

Retnoningrum, D.S., Cleary, P.P., 1994. M12 protein from *Streptococcus pyogenes* is a receptor for immunoglobulin G3 and human albumin. Infection and Immunity 62, 2387–2394. https://doi.org/10.1128/iai.62.6.2387-2394.1994

Rezaei Javan, R., Ramos-Sevillano, E., Akter, A., Brown, J., Brueggemann, A.B., 2019. Prophages and satellite prophages are widespread in *Streptococcus* and may play a role in pneumococcal pathogenesis. Nature Communications 10, 1–14. https://doi.org/10.1038/s41467-019-12825-y

Rios-Doria, J., Day, K.C., Kuefer, R., Rashid, M.G., Chinnaiyan, A.M., Rubin, M.A., Day, M.L., 2003. The role of calpain in the proteolytic cleavage of E-cadherin in prostate and mammary epithelial cells. Journal of Biological Chemistry 278, 1372–1379. https://doi.org/10.1074/jbc.M208772200

Robinson, J.H., Kehoe, M.A., 1992. Group A streptococcal M proteins: virulence factors and protective antigens. Immunology Today 13, 362–367. https://doi.org/10.1016/0167-5699(92)90173-5

Rodriguez-Iturbe, B., Musser, J.M., 2008. The current state of poststreptococcal glomerulonephritis. Journal of the American Society of Nephrology 19, 1855–1864. https://doi.org/10.1681/ASN.2008010092

Rutherford, K., Parkhill, J., Crook, J., Horsnell, T., Rice, P., Rajandream, M.-A., Barrell, B., 2000. Artemis: sequence visualization and annotation . Bioinformatics 16, 944–945. https://doi.org/10.1093/bioinformatics/16.10.944

Ryan, P.A., Juncosa, B., 2016. Group A Streptococcal Adherence. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 1–25.

Sablier, F., Slaouti, T., Drèze, P.-A., El Fouly, P.E., Allemeersch, D., Van Melderen, L., Smeesters, P.R., 2010. Nosocomial transmission of necrotising fasciitis. Lancet (London, England) 375, 1052. https://doi.org/10.1016/S0140-6736(10)60052-9

Sakaguchi, Y., Hayashi, T., Kurokawa, K., Nakayama, K., Oshima, K., Fujinaga, Y., Ohnishi, M., Ohtsubo, E., Hattori, M., Oguma, K., 2005. The genome sequence of *Clostridium botulinum* type C neurotoxin-converting phage and the molecular mechanisms of unstable lysogeny. Proceedings of the National Academy of Sciences of the United States of America 102, 17472 LP – 17477. https://doi.org/10.1073/pnas.0505503102

Salim, K.Y., Cvitkovitch, D.G., Chang, P., Bast, D.J., Handfield, M., Hillman, J.D., de Azavedo, J.C.S., 2005. Identification of Group A *Streptococcus* Antigenic Determinants Upregulated *in vivo*. Infection and Immunity 73, 6026 LP – 6038. https://doi.org/10.1128/IAI.73.9.6026-6038.2005

Sandegren, L., Andersson, D.I., 2009. Bacterial gene amplification: Implications for the evolution of antibiotic resistance. Nature Reviews Microbiology 7, 578–588. https://doi.org/10.1038/nrmicro2174

Santagati, M., Iannelli, F., Oggioni, M.R., Stefani, S., Pozzi, G., 2000. Characterization of a genetic element carrying the macrolide efflux gene *mef(A)* in *Streptococcus pneumoniae*. Antimicrobial Agents and Chemotherapy 44, 2585–2587. https://doi.org/10.1128/AAC.44.9.2585-2587.2000

Savva, C.G., Dewey, J.S., Moussa, S.H., To, K.H., Holzenburg, A., Young, R., 2014. Stable micron-scale holes are a general feature of canonical holins. Molecular Microbiology 91, 57–65. https://doi.org/10.1111/mmi.12439

Schrager, H.M., Rheinwald, J.G., Wessels, M.R., 1996. Hyaluronic acid capsule and the role of streptococcal entry into keratinocytes in invasive skin infection. Journal of Clinical Investigation 98, 1954–1958. https://doi.org/10.1172/JCI118998

Schwarz-Linek, U., Höök, M., Potts, J.R., 2006. Fibronectin-binding proteins of Gram-positive cocci. Microbes and Infection 8, 2291–2298. https://doi.org/10.1016/j.micinf.2006.03.011

Scott, J., Nguyen, S. V., King, C.J., Hendrickson, C., McShan, W.M., 2012. Phage-like *Streptococcus pyogenes* chromosomal islands (SpyCI) and mutator phenotypes: Control by growth state and rescue by a SpyCI-encoded promoter. Frontiers in Microbiology 3, 1–15. https://doi.org/10.3389/fmicb.2012.00317

Scott, J., Thompson-Mayberry, P., Lahmamsi, S., King, C.J., McShan, W.M., 2008. Phage-associated mutator phenotype in Group A *Streptococcus*. Journal of Bacteriology 190, 6290–6301. https://doi.org/10.1128/JB.01569-07

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. https://doi.org/10.1093/bioinformatics/btu153

Sela, S., Barzilai, A., 1999. Why do we fail with penicillin in the treatment of Group A *Streptococcus* infections? Annals of medicine. https://doi.org/10.3109/07853899908995895

Selva, L., Viana, D., Regev-Yochay, G., Trzcinski, K., Corpa, J.M., Lasa, Í., Novick, R.P., Penadés, J.R., 2009. Killing niche competitors by remote-control bacteriophage induction. Proceedings of the National Academy of Sciences of the United States of America 106, 1234–1238. https://doi.org/10.1073/pnas.0809600106

Shaikh, N., Leonard, E., Martin, J.M., 2010. Prevalence of streptococcal pharyngitis and streptococcal carriage in children: A meta-analysis. Pediatrics 126. https://doi.org/10.1542/peds.2009-2648

Sharma, H., Turner, C.E., Siggins, M.K., El-Bahrawy, M., Pichon, B., Kearns, A., Sriskandan, S., 2019. Toxic Shock Syndrome Toxin 1 Evaluation and Antibiotic Impact in a Transgenic Model of Staphylococcal Soft Tissue Infection. mSphere 4. https://doi.org/10.1128/mSphere.00665-19

Sharma, O., O'Seaghdha, M., Velarde, J.J., Wessels, M.R., 2016. NAD+-Glycohydrolase Promotes Intracellular Survival of Group A *Streptococcus*. PLoS Pathogens 12, 1–21. https://doi.org/10.1371/journal.ppat.1005468

Shea, P.R., Beres, S.B., Flores, A.R., Ewbank, A.L., Gonzalez-Lugo, J.H., Martagon-Rosado, A.J., Martinez-Gutierrez, J.C., Rehman, H.A., Serrano-Gonzalez, M., Fittipaldi, N., Ayers, S.D., Webb, P., Willey, B.M., Low, D.E., Musser, J.M., 2011. Distinct signatures of diversifying selection revealed by genome analysis of respiratory tract and invasive bacterial populations. Proceedings of the National Academy of Sciences of the United States of America 108, 5039–5044. https://doi.org/10.1073/pnas.1016282108

Sherry, S., Johnson, A., Tillett, W.S., 1949. The action of streptococcal desoxyribose nuclease *in vitro* and on purulent pleural exudations of patients. The Journal of clinical investigation 28, 1094–1104. https://doi.org/10.1172/jci102142

Shet, A., Kaplan, E., 2004. Addressing the Burden of Group A Streptococcal Disease in India. Indian Journal of Pediatrics 71, 41–48. https://doi.org/10.1007/BF02725655

Shet, A., Kaplan, E.L., Johnson, D.R., Cleary, P.P., 2003. Immune Response to Group A Streptococcal C5a Peptidase in Children: Implications for Vaccine Development. The Journal of Infectious Diseases 188, 809–817. https://doi.org/10.1086/377700

Siggins, M.K., Lynskey, N.N., Lamb, L.E., Johnson, L.A., Huse, K.K., Pearson, M., Banerji, S., Turner, C.E., Woollard, K., Jackson, D.G., Sriskandan, S., 2020. Extracellular bacterial lymphatic metastasis drives *Streptococcus pyogenes* systemic infection. Nature Communications 11, 4697. https://doi.org/10.1038/s41467-020-18454-0

Silva-Costa, C., Carriço, J.A., Ramirez, M., Melo-Cristino, J., 2014. Scarlet fever is caused by a limited number of *Streptococcus pyogenes* lineages and is associated with the exotoxin genes *ssa, speA* and *speC*. The Pediatric infectious disease journal 33, 306–310. https://doi.org/10.1097/INF.0000000000000088

Singh, S.K., Bharati, A.P., Singh, N., Pandey, P., Joshi, P., Singh, K., Mitra, K., Gayen, J.R., Sarkar, J., Akhtar, M.S., 2014. The prophage-encoded hyaluronate lyase has broad substrate specificity and is regulated by the N-terminal domain. Journal of Biological Chemistry 289, 35225–35236. https://doi.org/10.1074/jbc.M113.507673

Sitkiewicz, I., Green, N.M., Guo, N., Bongiovanni, A.M., Witkin, S.S., Musser, J.M., 2010. Adaptation of Group A *Streptococcus* to human amniotic fluid. PLoS ONE 5. https://doi.org/10.1371/journal.pone.0009785

Sitkiewicz, I., Green, N.M., Guo, N., Mereghetti, L., Musser, J.M., 2011. Lateral gene transfer of streptococcal ICE element RD2 (region of difference 2) encoding secreted proteins. BMC Microbiology 11. https://doi.org/10.1186/1471-2180-11-65

Sitkiewicz, I., Nagiec, M.J., Sumby, P., Butler, S.D., Cywes-Bentley, C., Musser, J.M., 2006. Emergence of a bacterial clone with enhanced virulence by acquisition of a phage encoding a secreted phospholipase A2. Proceedings of the National Academy of Sciences of the United States of America 103, 16009–16014. https://doi.org/10.1073/pnas.0607669103

Sitto, F., Battistuzzi, F.U., 2020. Estimating Pangenomes with Roary. Molecular Biology and Evolution 37, 933–939. https://doi.org/10.1093/molbev/msz284

Smeesters, P.R., McMillan, D.J., Sriprakash, K.S., 2010. The streptococcal M protein: A highly versatile molecule. Trends in Microbiology 18, 275–282. https://doi.org/10.1016/j.tim.2010.02.007

Smoot, J.C., Barbian, K.D., Van Gompel, J.J., Smoot, L.M., Chaussee, M.S., Sylva, G.L., Sturdevant, D.E., Ricklefs, S.M., Porcella, S.F., Parkins, L.D., Beres, S.B., Campbell, D.S., Smith, T.M., Zhang, Q., Kapur, V., Daly, J.A., Veasy, L.G., Musser, J.M., 2002a. Genome sequence and comparative microarray analysis of serotype M18 Group A *Streptococcus* strains associated with acute rheumatic fever outbreaks. Proceedings of the National Academy of Sciences of the United States of America 99, 4668–4673. https://doi.org/10.1073/pnas.062526099

Smoot, J.C., Korgenski, E.K., Daly, J.A., Veasy, L.G., Musser, J.M., 2002b. Molecular analysis of Group A *Streptococcus* type *emm*18 isolates temporally associated with acute rheumatic fever outbreaks in Salt Lake City, Utah. Journal of Clinical Microbiology 40, 1805–1810. https://doi.org/10.1128/JCM.40.5.1805-1810.2002

Soderholm, A.T., Barnett, T.C., Sweet, M.J., Walker, M.J., 2017. Group A streptococcal pharyngitis: immune responses involved in bacterial clearance and GAS-associated immunopathologies. Journal of Leukocyte Biology jlb.4MR0617-227RR. https://doi.org/10.1189/jlb.4mr0617-227rr

Spanos, W.C., Geiger, J., Anderson, M.E., Harris, G.F., Bossler, A.D., Smith, R.B., Klingelhutz, A.J., Lee, J.H., 2008. Deletion of the PDZ motif of HPV16 E6 preventing immortalization and anchorage-independent growth in human tonsil epithelial cells. Head & Neck 30, 139—147. https://doi.org/10.1002/hed.20673

Sriskandan, S., Ferguson, M., Elliot, V., Faulkner, L., Cohen, J., 2006. Human intravenous immunoglobulin for experimental streptococcal toxic shock: bacterial clearance and modulation of inflammation. The Journal of antimicrobial chemotherapy 58, 117–124. https://doi.org/10.1093/jac/dkl173

Sriskandan, S., Moves, D., Lee Buttery, K., Krausz, T., Evans, T.J., Polak, J., Cohen, J., 1996. Streptococcal pyrogenic exotoxin A release, distribution, and role in a murine model of fasciitis and multiorgan failure due to *Streptococcus pyogenes*. Journal of Infectious Diseases 173, 1399–1407. https://doi.org/10.1093/infdis/173.6.1399

Sriskandan, S., Unnikrishnan, M., Krausz, T., Cohen, J., 2000. Mitogenic factor (MF) is the major DNase of serotype M89 *Streptococcus pyogenes*. Microbiology 146, 2785–2792. https://doi.org/https://doi.org/10.1099/00221287-146-11-2785

Sriskandan, S., Unnikrishnan, M., Krausz, T., Cohen, J., 1999. Molecular analysis of the role of streptococcal pyrogenic exotoxin A (SPEA) in invasive soft-tissue infection resulting from *Streptococcus pyogenes*. Molecular Microbiology 33, 778–790. https://doi.org/10.1046/j.1365-2958.1999.01525.x

Sriskandan, S., Unnikrishnan, M., Krausz, T., Dewchand, H., Van Noorden, S., Cohen, J., Altmann, D.M., 2001. Enhanced Susceptibility to Superantigen-Associated Streptococcal Sepsis in Human Leukocyte Antigen-DQ Transgenic Mice. The Journal of Infectious Diseases 184, 166–173. https://doi.org/10.1086/322018

Starr, C.R., Engleberg, N.C., 2006. Role of hyaluronidase in subcutaneous spread and growth of Group A *Streptococcus*. Infection and Immunity 74, 40–48. https://doi.org/10.1128/IAI.74.1.40-48.2006

Steer, Andrew C., Batzloff, M.R., Mulholland, K., Carapetis, J.R., 2009. Group A streptococcal vaccines: Facts versus fantasy. Current Opinion in Infectious Diseases 22, 544–552. https://doi.org/10.1097/QCO.0b013e328332bbfe

Steer, A.C., Carapetis, J.R., Dale, J.B., Fraser, J.D., Good, M.F., Guilherme, L., Moreland, N.J., Mulholland, E.K., Schodel, F., Smeesters, P.R., 2016. Status of research and development of vaccines for *Streptococcus pyogenes*. Vaccine 34, 2953–2958. https://doi.org/10.1016/j.vaccine.2016.03.073

Steer, A.C., Lamagni, T., Curtis, N., Carapetis, J.R., 2012. Invasive Group A streptococcal disease:

Epidemiology, pathogenesis and management. Drugs 72, 1213–1227. https://doi.org/10.2165/11634180-000000000-00000

Steer, A.C, Law, I., Matatolu, L., Beall, B.W., Carapetis, J.R., 2009. Global *emm* type distribution of Group A streptococci: systematic review and implications for vaccine development. The Lancet. Infectious diseases 9, 611–616. https://doi.org/10.1016/S1473-3099(09)70178-1

Stevens, D.L., Bryant, A.E., 2017. Necrotizing Soft-Tissue Infections. New England Journal of Medicine 377, 2253–2265. https://doi.org/10.1056/NEJMra1600673

Stevens, D.L., Bryant, A.E., 2016. Impetigo, Erysipelas and Cellulitis. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 1–18.

Stockbauer, K.E., Grigsby, D., Pan, X., Fu, Y.X., Perea Mejia, L.M., Cravioto, A., Musser, J.M., 1998. Hypervariability generated by natural selection in an extracellular complement-inhibiting protein of serotype M1 strains of Group A *Streptococcus*. Proceedings of the National Academy of Sciences of the United States of America 95, 3128–3133. https://doi.org/10.1073/pnas.95.6.3128

Stollerman, G.H., Dale, J.B., 2008. The importance of the Group A *Streptococcus* capsule in the pathogenesis of human infections: a historical perspective. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America 46, 1038–1045. https://doi.org/10.1086/529194

Sullivan, M.J., Petty, N.K., Beatson, S.A., 2011. Easyfig: a genome comparison visualizer. Bioinformatics (Oxford, England) 27, 1009–1010. https://doi.org/10.1093/bioinformatics/btr039

Sumby, P., Barbian, K.D., Gardner, D.J., Whitney, A.R., Welty, D.M., Long, R.D., Bailey, J.R., Parnell, M.J., Hoe, N.P., Adams, G.G., Deleo, F.R., Musser, J.M., 2005a. Extracellular deoxyribonuclease made by Group A *Streptococcus* assists pathogenesis by enhancing evasion of the innate immune response. Proceedings of the National Academy of Sciences of the United States of America 102, 1679–1684. https://doi.org/10.1073/pnas.0406641102

Sumby, P., Porcella, S.F., Madrigal, A.G., Barbian, K.D., Virtaneva, K., Ricklefs, S.M., Sturdevant, D.E., Graham, M.R., Vuopio-Varkila, J., Hoe, N.P., Musser, J.M., 2005b. Evolutionary Origin and Emergence of a Highly Successful Clone of Serotype M1 Group A *Streptococcus* Involved Multiple Horizontal Gene Transfer Events . The Journal of Infectious Diseases 192, 771–782. https://doi.org/10.1086/432514

Sumby, P., Whitney, A.R., Graviss, E.A., DeLeo, F.R., Musser, J.M., 2006. Genome-wide analysis of Group A streptococci reveals a mutation that modulates global phenotype and disease specificity. PLoS Pathogens 2, 0041–0049. https://doi.org/10.1371/journal.ppat.0020005

Sumby, P., Zhang, S., Whitney, A.R., Falugi, F., Grandi, G., Graviss, E.A., DeLeo, F.R., Musser, J.M., 2005c. A chemokine-degrading extracellular protease made by Group A *Streptococcus* alters pathogenesis by enhancing evasion of the innate immune response. Infection and Immunity 76, 978–985. https://doi.org/10.1128/IAI.01354-07

Sumitomo, T., Nakata, M., Higashino, M., Jin, Y., Terao, Y., Fujinaga, Y., Kawabata, S., 2011. Streptolysin S

contributes to Group A streptococcal translocation across an epithelial barrier. Journal of Biological Chemistry 286, 2750–2761. https://doi.org/10.1074/jbc.M110.171504

Swe, P.M., Christian, L.D., Lu, H.C., Sriprakash, K.S., Fischer, K., 2017. Complement inhibition by *Sarcoptes scabiei* protects *Streptococcus pyogenes* - An in vitro study to unravel the molecular mechanisms behind the poorly understood predilection of S. *pyogenes* to infect mite-induced skin lesions. PLoS Neglected Tropical Diseases 11, 1–20. https://doi.org/10.1371/journal.pntd.0005437

Takeda, K., Kaisho, T., Akira, S., 2003. Toll -Like Receptors . Annual Review of Immunology 21, 335–376. https://doi.org/10.1146/annurev.immunol.21.120601.141126

Talkington, D.F., Schwartz, B., Black, C.M., Todd, J.K., Elliott, J., Breiman, R.F., Facklam, R.R., 1993. Association of phenotypic and genotypic characteristics of invasive *Streptococcus pyogenes* isolates with clinical components of streptococcal toxic shock syndrome. Infection and Immunity 61, 3369–3374.

Tamayo, E., Montes, M., García-Medina, G., García-Arenzana, J.M., Pérez-Trallero, E., 2010. Spread of a highly mucoid *Streptococcus pyogenes emm*3/ST15 clone. BMC Infectious Diseases 10, 2–6. https://doi.org/10.1186/1471-2334-10-233

Tamura, F., Nakagawa, R., Akuta, T., Okamoto, S., Hamada, S., Maeda, H., Kawabata, S., Akaike, T., 2004. Proapoptotic effect of proteolytic activation of matrix metalloproteinases by *Streptococcus pyogenes* thiol proteinase (*Streptococcus* pyrogenic exotoxin B). Infection and Immunity 72, 4836–4847. https://doi.org/10.1128/IAI.72.8.4836-4847.2004

Terao, Y., Kawabata, S., Kunitomo, E., Murakami, J., Nakagawa, I., Hamada, S., 2001. Fba, a novel fibronectin-binding protein from *Streptococcus pyogenes*, promotes bacterial entry into epithelial cells, and the *fba* gene is positively transcribed under the Mga regulator. Molecular Microbiology 42, 75–86. https://doi.org/10.1046/j.1365-2958.2001.02579.x

Thomas, C.A., Li, Y., Kodama, T., Suzuki, H., Silverstein, S.C., El Khoury, J., 2000. Protection from lethal Gram-positive infection by macrophage scavenger receptor-dependent phagocytosis. Journal of Experimental Medicine 191, 147–155. https://doi.org/10.1084/jem.191.1.147

Thulin, P., Johansson, L., Low, D.E., Gan, B.S., Kotb, M., McGeer, A., Norrby-Teglund, A., 2006. Viable Group A streptococci in macrophages during acute soft tissue infection. PLoS Medicine 3, 371–379. https://doi.org/10.1371/journal.pmed.0030053

Todd, E.W., 1927. A Method of Measuring the Increase or Decrease of the Population of Hæmolytic Streptococci in Blood. British journal of experimental pathology.

Travassos, L.H., Girardin, S.E., Philpott, D.J., Blanot, D., Nahori, M.A., Werts, C., Boneca, I.G., 2004. Toll-like receptor 2-dependent bacterial sensing does not occur via peptidoglycan recognition. EMBO Reports 5, 1000–1006. https://doi.org/10.1038/sj.embor.7400248

Tsai, C.J.Y., Loh, J.M.S., Proft, T., 2016. *Galleria mellonella* infection models for the study of bacterial diseases and for antimicrobial drug testing. Virulence 7, 214–229. https://doi.org/10.1080/21505594.2015.1135289

Tsai, P.J., Lin, Y.S., Kuo, C.F., Lei, H.Y., Wu, J.J., 1999. Group A *Streptococcus* induces apoptosis in human

epithelial cells. Infection and Immunity 67, 4334–4339. https://doi.org/10.1128/iai.67.9.4334-4339.1999

Tse, H., Bao, J.Y.J., Davies, M.R., Maamary, P., Tsoi, H.W., Tong, A.H.Y., Ho, T.C.C., Lin, C.H., Gillen, C.M., Barnett, T.C., Chen, J.H.K., Lee, M., Yam, W.C., Wong, C.K., Ong, C.L.Y., Chan, Y.W., Wu, C.W., Ng, T., Lim, W.W.L., Tsang, T.H.F., Tse, C.W.S., Dougan, G., Walker, M.J., Lok, S., Yuen, K.Y., 2012. Molecular characterization of the 2011 Hong Kong scarlet fever outbreak. Journal of Infectious Diseases 206, 341–351. https://doi.org/10.1093/infdis/jis362

Turner, C.E., Abbott, J., Lamagni, T., Holden, M.T.G., David, S., Jones, M.D., Game, L., Efstratiou, A., Sriskandan, S., 2015. Emergence of a new highly successful acapsular Group A *Streptococcus* clade of genotype *emm*89 in the United Kingdom. mBio 6, 1–11. https://doi.org/10.1128/mBio.00622-15

Turner, C.E., Bedford, L., Brown, N.M., Judge, K., Török, M.E., Parkhill, J., Peacock, S.J., 2017. Community outbreaks of Group A *Streptococcus* revealed by genome sequencing. Scientific Reports 7, 1–9. https://doi.org/10.1038/s41598-017-08914-x

Turner, C.E., Bubba, L., Efstratiou, A., 2019a. Pathogenicity Factors in Group C and G Streptococci. Microbiology Spectrum 7, 264–274. https://doi.org/10.1128/microbiolspec.gpp3-0020-2018

Turner, C.E., Dryden, M., Holden, M.T.G., Davies, F.J., Lawrenson, R.A., Farzaneh, L., Bentley, S.D., Efstratiou, A., Sriskandan, S., 2013. Molecular analysis of an outbreak of lethal postpartum sepsis caused by *Streptococcus pyogenes*. Journal of Clinical Microbiology 51, 2089–2095. https://doi.org/10.1128/JCM.00679-13

Turner, C.E., Holden, M.T.G., Blane, B., Horner, C., Peacock, S.J., Sriskandan, S., 2019b. The emergence of successful *Streptococcus pyogenes* lineages through convergent pathways of capsule loss and recombination directing high toxin expression. mBio 10, 1–20. https://doi.org/10.1128/mBio.02521-19

Turner, C.E., Kurupati, P., Jones, M.D., Edwards, R.J., Sriskandan, S., 2009. Emerging Role of the Interleukin-8 Cleaving Enzyme SpyCEP in Clinical *Streptococcus pyogenes* Infection . The Journal of Infectious Diseases 200, 555–563. https://doi.org/10.1086/603541

Turner, C. E., Kurupati, P., Wiles, S., Edwards, R.J., Sriskandan, S., 2009. Impact of immunization against SpyCEP during invasive disease with two streptococcal species: *Streptococcus pyogenes* and *Streptococcus* equi. Vaccine 27, 4923–4929. https://doi.org/10.1016/j.vaccine.2009.06.042

Turner, C.E., Pyzio, M., Song, B., Lamagni, T., Meltzer, M., Chow, J.Y., Efstratiou, A., Curtis, S., Sriskandan, S., 2016. Scarlet fever upsurge in england and molecular-genetic analysis in North-West London, 2014. Emerging Infectious Diseases 22, 1075–1078. https://doi.org/10.3201/eid2206.151726

Turner, C.E., Sommerlad, M., McGregor, K., Davies, F.J., Pichon, B., Chong, D.L.W., Farzaneh, L., Holden, M.T.G., Spratt, B.G., Efstratiou, A., Sriskandan, S., 2012. Superantigenic Activity of *emm*3 *Streptococcus pyogenes* Is Abrogated by a Conserved, Naturally Occurring smeZ Mutation. PLoS ONE 7, 1–11. https://doi.org/10.1371/journal.pone.0046376

Uchiyama, S., Andreoni, F., Schuepbach, R.A., Nizet, V., Zinkernagel, A.S., 2012. DNase Sda1 allows invasive

M1T1 Group A *Streptococcus* to prevent TLR9-dependent recognition. PLoS Pathogens 8, 1–12. https://doi.org/10.1371/journal.ppat.1002736

Unnikrishnan, M., Altmann, D.M., Proft, T., Wahid, F., Cohen, J., Fraser, J.D., Sriskandan, S., 2002. The Bacterial Superantigen Streptococcal Mitogenic Exotoxin Z Is the Major Immunoactive Agent of *Streptococcus pyogenes* . The Journal of Immunology 169, 2561–2569. https://doi.org/10.4049/jimmunol.169.5.2561

Unnikrishnan, M., Cohen, J., Sriskandan, S., 1999. Growth-phase-dependent expression of virulence factors in an M1T1 clinical isolate of *Streptococcus pyogenes*. Infection and Immunity 67, 5495–5499. https://doi.org/10.1128/iai.67.10.5495-5499.1999

Uribe-Quero, E., Rosales, C., 2017. Control of phagocytosis by microbial pathogens. Frontiers in Immunology 8, 1–23. https://doi.org/10.3389/fimmu.2017.01368

Valderrama, J.A., Nizet, V., 2018. Group A *Streptococcus* encounters with host macrophages. Future Microbiology 13, 119–134. https://doi.org/10.2217/fmb-2017-0142

Van Belleghem, J.D., Dąbrowska, K., Vaneechoutte, M., Barr, J.J., Bollyky, P.L., 2019. Interactions between bacteriophage, bacteria, and the mammalian immune system. Viruses 11. https://doi.org/10.3390/v11010010

Vannice, K.S., Ricaldi, J., Nanduri, S., Fang, F.C., Lynch, J.B., Bryson-Cahn, C., Wright, T., Duchin, J., Kay, M., Chochua, S., Van Beneden, C.A., Beall, B., 2019. *Streptococcus pyogenes* pbp2x Mutation Confers Reduced Susceptibility to β-Lactam Antibiotics. Clinical Infectious Diseases 2017–2020. https://doi.org/10.1093/cid/ciz1000

Varhimo, E., Savijoki, K., Jalava, J., Kuipers, O.P., Varmanen, P., 2007. Identification of a novel streptococcal gene cassette mediating SOS mutagenesis in *Streptococcus* uteris. Journal of Bacteriology 189, 5210–5222. https://doi.org/10.1128/JB.00473-07

Vega, L.A., Sanson, M.A., Shah, B.J., Flores, A.R., 2020. The strain-dependent effect of capsule on transmission and persistence in an infant mouse model of Group A *Streptococcus* infection. Infection and Immunity 1–39. https://doi.org/10.1128/iai.00709-19

Velarde, J.J., O'Seaghdha, M., Baddal, B., Bastiat-Sempe, B., Wessels, M.R., 2017. Binding of NAD Glycohydrolase to Streptolysin O Stabilizes Both Toxins and Promotes Virulence of Group A *Streptococcus* mBio 8, e01382-17. https://doi.org/10.1128/mBio.01382-17

Velineni, S., Timoney, J.F., 2015. Capsular hyaluronic acid of equine isolates of *Streptococcus zooepidemicus* is upregulated at temperatures below 35°C. Equine Veterinary Journal 47, 333–338. https://doi.org/10.1111/evj.12272

Ventura, M., Bruttin, A., Canchaya, C., Brüssow, H., 2002a. Transcription analysis of *Streptococcus thermophilus* phages in the lysogenic state. Virology 302, 21–32. https://doi.org/10.1006/viro.2002.1571

Ventura, M., Foley, S., Bruttin, A., Chennoufi, S.C., Canchaya, C., Brüssow, H., 2002b. Transcription mapping

as a tool in phage genomics: The case of the temperate *Streptococcus thermophilus* phage Sfi21. Virology 296, 62–76. https://doi.org/10.1006/viro.2001.1331

Venturini, C., Ong, C. lynn Y., Gillen, C.M., Ben-Zakour, N.L., Maamary, P.G., Nizet, V., Beatson, S.A., Walker, M.J., 2013. Acquisition of the Sda1-encoding bacteriophage does not enhance virulence of the serotype M1 *Streptococcus pyogenes* strain SF370. Infection and Immunity 81, 2062–2069. https://doi.org/10.1128/IAI.00192-13

Vincents, B., Vindebro, R., Abrahamson, M., von Pawel-Rammingen, U., 2008. The Human Protease Inhibitor Cystatin C Is an Activating Cofactor for the Streptococcal Cysteine Protease IdeS. Chemistry and Biology 15, 960–968. https://doi.org/10.1016/j.chembiol.2008.07.021

Vindenes, T., McQuillen, D., 2015. Acute Lymphangitis. New England Journal of Medicine 372, 649–649. https://doi.org/10.1056/nejmicm1406569

Virtaneva, K., Porcella, S.F., Graham, M.R., Ireland, R.M., Johnson, C.A., Ricklefs, S.M., Babar, I., Parkins, L.D., Romero, R.A., Corn, G.J., Gardner, D.J., Bailey, J.R., Parnell, M.J., Musser, J.M., 2005. Longitudinal analysis of the Group A *Streptococcus* transcriptome in experimental pharyngitis in cynomolgus macaques. Proceedings of the National Academy of Sciences of the United States of America 102, 9014–9019. https://doi.org/10.1073/pnas.0503671102

Vlaminckx, B.J.M., Mascini, E.M., Schellekens, J., Schouls, L.M., Paauw, A., Fluit, A.C., Novak, R., Verhoef, J., Schmitz, F.J., 2003. Site-Specific Manifestations of Invasive Group A Streptococcal Disease: Type Distribution and Corresponding Patterns of Virulence Determinants. Journal of Clinical Microbiology 41, 4941–4949. https://doi.org/10.1128/JCM.41.11.4941-4949.2003

Von Köckritz-Blickwede, M., Nizet, V., 2009. Innate immunity turned inside-out: Antimicrobial defense by phagocyte extracellular traps. Journal of Molecular Medicine 87, 775–783. https://doi.org/10.1007/s00109-009-0481-0

Von Pawel-Rammingen, U., 2012. Streptococcal IdeS and its impact on immune response and inflammation. Journal of Innate Immunity 4, 132–140. https://doi.org/10.1159/000332940

Voyich, J.M., Musser, J.M., DeLeo, F.R., 2004. *Streptococcus pyogenes* and human neutrophils: A paradigm for evasion of innate host defense by bacterial pathogens. Microbes and Infection 6, 1117–1123. https://doi.org/10.1016/j.micinf.2004.05.022

Voyich, J.M., Sturdevant, D.E., Braughton, K.R., Kobayashi, S.D., Lei, B., Virtaneva, K., Dorward, D.W., Musser, J.M., DeLeo, F.R., 2003. Genome-wide protective response used by Group A *Streptococcus* to evade destruction by human polymorphonuclear leukocytes. Proceedings of the National Academy of Sciences of the United States of America 100, 1996–2001. https://doi.org/10.1073/pnas.0337370100

Walker, M.J., 2015. Emergence of scarlet fever *Streptococcus pyogenes emm*12 clones in Hong Kong is associated with toxin acquisition and multidrug resistance. Nature genetics 47, 84–87. https://doi.org/10.1038/ng.3147

Walker, M.J., Barnett, T.C., McArthur, J.D., Cole, J.N., Gillen, C.M., Henningham, A., Sriprakash, K.S.,

Sanderson-Smith, M.L., Nizet, V., 2014. Disease manifestations and pathogenic mechanisms of Group A *Streptococcus*. Clinical Microbiology Reviews 27, 264–301. https://doi.org/10.1128/CMR.00101-13

Walker, M.J., Hollands, A., Sanderson-Smith, M.L., Cole, J.N., Kirk, J.K., Henningham, A., McArthur, J.D., Dinkla, K., Aziz, R.K., Kansal, R.G., Simpson, A.J., Buchanan, J.T., Chhatwal, G.S., Kotb, M., Nizet, V., 2007. DNase Sda1 provides selection pressure for a switch to invasive Group A streptococcal infection. Nature Medicine 13, 981–985. https://doi.org/10.1038/nm1612

Wang, X., Kim, Y., Ma, Q., Hong, S.H., Pokusaeva, K., Sturino, J.M., Wood, T.K., 2010. Cryptic prophages help bacteria cope with adverse environments. Nature Communications 1. https://doi.org/10.1038/ncomms1146

Watts, V., Balasegaram, S., Brown, C.S., Mathew, S., Mearkle, R., Ready, D., Saliba, V., Lamagni, T., 2019. Increased risk for invasive Group A *Streptococcus* disease for household contacts of scarlet fever cases, England, 2011-2016. Emerging Infectious Diseases 25, 529–537. https://doi.org/10.3201/eid2503.181518

Weiss, K., Laverdière, M., Lovgren, M., Delorme, J., Poirier, L., Béliveau, C., 1999. Group A *Streptococcus* carriage among close contacts of patients with invasive infections. American Journal of Epidemiology 149, 863–868. https://doi.org/10.1093/oxfordjournals.aje.a009902

Wessels, M.R., 2019. Capsular Polysaccharide of Group A *Streptococcus*. Microbiology Spectrum 7, 1–21. https://doi.org/10.1128/microbiolspec.gpp3-0050-2018

Wessels, M.R., 2016. Cell Wall and Surface Molecules: Capsule. *Streptococcus pyogenes*: Basic Biology to Clinical Manifestations 53–63.

Whitehead, B.D., Smith, H. V., Nourse, C., 2011. Invasive Group A streptococcal disease in children in Queensland. Epidemiology and Infection 139, 623–628. https://doi.org/10.1017/S0950268810001378

Wilkening, R. V., Federle, M.J., 2017. Evolutionary Constraints Shaping *Streptococcus pyogenes*–Host Interactions. Trends in Microbiology 25, 562–572. https://doi.org/10.1016/j.tim.2017.01.007

Wood, T.F., Potter, M.A., Jonasson, O., 1993. Streptococcal toxic shock-like syndrome. The importance of surgical intervention. Annals of surgery 217, 109–114. https://doi.org/10.1097/00000658-199302000-00003

Xia, G., Wolz, C., 2014. Phages of *Staphylococcus aureus* and their impact on host evolution. Infection, Genetics and Evolution 21, 593–601. https://doi.org/10.1016/j.meegid.2013.04.022

Yates, R.M., Hermetter, A., Russell, D.G., 2005. The kinetics of phagosome maturation as a function of phagosome/lysosome fusion and acquisition of hydrolytic activity. Traffic 6, 413–420. https://doi.org/10.1111/j.1600-0854.2005.00284.x

Yinduo, J.I., McLandsborough, L., Kondagunta, A., Cleary, P.P., 1996. C5a peptidase alters clearance and trafficking of Group A streptococci by infected mice. Infection and Immunity 64, 503–510. https://doi.org/10.1128/iai.64.2.503-510.1996

Young, R., 2014. Phage lysis: Three steps, three choices, one outcome. Journal of Microbiology 52, 243–258. https://doi.org/10.1007/s12275-014-4087-z

Yue, W.-F., Du, M., Zhu, M.-J., 2012. High Temperature in Combination with UV Irradiation Enhances

Horizontal Transfer of stx2 Gene from E. *coli* O157:H7 to Non-Pathogenic E. coli. PLOS ONE 7, 1–7. https://doi.org/10.1371/journal.pone.0031308

Yuste, J., Ali, S., Sriskandan, S., Hyams, C., Botto, M., Brown, J.S., 2006. Roles of the Alternative Complement Pathway and C1q during Innate Immunity to *Streptococcus pyogenes* . The Journal of Immunology 176, 6112–6120. https://doi.org/10.4049/jimmunol.176.10.6112

Zeppa, J.J., Kasper, K.J., Mohorovic, I., Mazzuca, D.M., Haeryfar, S.M.M., McCormick, J.K., 2017. Nasopharyngeal infection by *Streptococcus pyogenes* requires superantigen-responsive Vβ-specific T cells. Proceedings of the National Academy of Sciences of the United States of America 114, 10226–10231. https://doi.org/10.1073/pnas.1700858114

Žgur-Bertok, D., 2013. DNA Damage Repair and Bacterial Pathogens. PLoS Pathogens 9, 9–12. https://doi.org/10.1371/journal.ppat.1003711

Zhang, S., Green, N.M., Sitkiewicz, I., LeFebvre, R.B., Musser, J.M., 2006. Identification and characterization of an antigen I/II family protein produced by Group A *Streptococcus*. Infection and Immunity 74, 4200–4213. https://doi.org/10.1128/IAI.00493-06

Zhang, X., McDaniel, A.D., Wolf, L.E., Keusch, G.T., Waldor, M.K., Acheson, D.W.K., 2000. Quinolone Antibiotics Induce Shiga Toxin–Encoding Bacteriophages, Toxin Production, and Death in Mice. The Journal of Infectious Diseases 181, 664–670. https://doi.org/10.1086/315239

Zhou, L., Liu, L., Yang, J., Li, Y., Bai, W., Liu, N., Li, W., Gao, Y., Xu, L., Liu, Z., Han, R., 2016. LDL acts as an opsonin enhancing the phagocytosis of Group A *Streptococcus* by monocyte and whole human blood. Medical Microbiology and Immunology 205, 155–162. https://doi.org/10.1007/s00430-015-0436-8

Zhu, L., Charbonneau, A.R.L., Waller, A.S., Olsen, R.J., Beres, S.B., Musser, J.M., 2017. Novel Genes Required for the Fitness of *Streptococcus pyogenes* in Human Saliva . mSphere 2, 1–13. https://doi.org/10.1128/mspheredirect.00460-17

Zhu, L., Olsen, R.J., Beres, S.B., Saavedra, O., Kubiak, S.L., 2020. Genome-Wide Screens Identify Group A *Streptococcus* Surface Proteins Promoting Female Genital Tract Colonization and Virulence. The American Journal of Pathology 190. https://doi.org/10.1016/j.ajpath.2019.12.003

Zhu, L., Olsen, R.J., Nasser, W., Beres, S.B., Vuopio, J., Kristinsson, K.G., Gottfredsson, M., Porter, A.R., DeLeo, F.R., Musser, J.M., 2015a. A molecular trigger for intercontinental epidemics of Group A *Streptococcus*. Journal of Clinical Investigation 125, 3545–3559. https://doi.org/10.1172/JCI82478

Zhu, L., Olsen, R.J., Nasser, W., De La Riva Morales, I., Mussera, J.M., 2015b. Trading capsule for increased cytotoxin production: Contribution to virulence of a newly emerged clade of *emm*89 *Streptococcus pyogenes*. mBio 6, 1–10. https://doi.org/10.1128/mbio.01378-15

Zinkernagel, A.S., Timmer, A.M., Pence, M.A., Locke, J.B., Buchanan, J.T., Turner, C.E., Mishalian, I., Sriskandan, S., Hanski, E., Nizet, V., 2008. The IL-8 Protease SpyCEP/ScpC of Group A *Streptococcus* Promotes Resistance to Neutrophil Killing. Cell Host and Microbe 4, 170–178. https://doi.org/10.1016/j.chom.2008.07.002

Zlotnik, A., Yoshie, O., 2012. The Chemokine Superfamily Revisited. Immunity 36, 705–716. https://doi.org/10.1016/j.immuni.2012.05.008