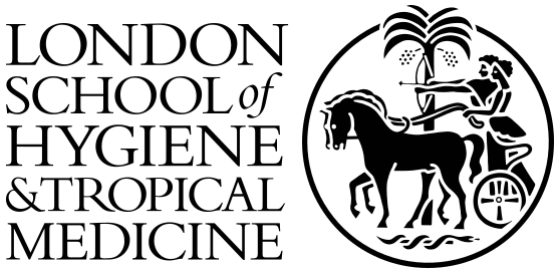

LSHTM Research Online

Ebruke, CN; (2018) Molecular epidemiological and pathogenesis studies of Streptococcus pneumoniae serotype 1 strains from West Africa. MPhil thesis, London School of Hygiene & Tropical Medicine. DOI: https://doi.org/10.17037/PUBS.04647232

Downloaded from: https://researchonline.lshtm.ac.uk/id/eprint/4647232/

DOI: https://doi.org/10.17037/PUBS.04647232

Usage Guidelines:

Please refer to usage guidelines at https://researchonline.lshtm.ac.uk/policies.html or alternatively contact researchonline@lshtm.ac.uk.

Available under license. To note, 3rd party material is not necessarily covered under this license: http://creativecommons.org/licenses/by-nc-nd/3.0/

https://researchonline.lshtm.ac.uk

Molecular epidemiological and pathogenesis studies of *Streptococcus pneumoniae* serotype 1 strains from West Africa

CHINELO NWABUISI EBRUKE MSc (Ibadan), BSc Hons (Benin)

Thesis submitted in accordance with the requirements for the degree of

Master of Philosophy
University of London

JANUARY 2018

Department of Pathogen Molecular Biology

Faculty of Infectious and Tropical Diseases

LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE

Funded by Medical Research Council Unit, The Gambia

Research group affiliation(s): Vaccines and Immunity Theme, Medical Research Council (MRC) Unit, The Gambia, Atlantic Road, Fajara, P.O.Box 273 Banjul, The Gambia, West Africa

Declaration

I, Chinelo Nwabuisi Ebruke, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

For the population structure and antimicrobial resistance patterns of *S. pneumoniae* isolates from The Gambia, pneumococcal studies conducted in The Gambia from 1995 – 2016 were sourced through a literature search as well as on-going studies. The field work clinical work and certain aspects of the laboratory work, specifically pneumococcal serotyping and antimicrobial susceptibility testing, were conducted as part of those studies by members of the Medical Research Council, The Gambia. I received statistical support in analysing the data in this chapter.

Whole genome sequencing of *S. pneumoniae* serotype 1 was performed by staff of the Wellcome Trust Sanger Institute (WTSI) as part of the Pneumococcal African Genome (PAGe) project and Global Pneumococcal Sequencing (GPS) project of which The MRC Gambia Unit is a collaborating institution. I received bioinformatics support to analyse the genomes included in this thesis.

Signed: 

Date: 15th September, 2017

Abstract

In West Africa, *Streptococcus pneumoniae* remains a leading cause of deaths in young children and serotype 1 strains are particularly important in causing invasive pneumococcal disease (IPD) despite being rare in nasopharyngeal carriage. The *S. pneumoniae* ST217 clonal complex, consisting of ST217 and the various locus variants has been shown to be the predominant pneumococcal serotype 1 clone in the sub-region. It is unclear how the recent introduction of pneumococcal conjugate vaccine (PCV) in countries in the sub-region could affect patterns of pneumococcal disease. Improving our understanding of the unique nature of pneumococcal serotype 1 strains within the context of other pneumococcal serotypes circulating in West Africa prior to introduction of PCV would be critical in interpreting any subsequent changes in patterns of disease in this region. This forms the basis for my research studies for this PhD thesis. Epidemiological studies on *S. pneumoniae* strains in The Gambia revealed that a new dominant clone of serotype 1, ST3081, a single locus variant of ST217, emerged and appeared to have spread across the whole country. ST3081 appeared to have replaced ST618, a triple locus variant of ST217 and the previously dominant pneumococcal serotype 1 lineage circulating in The Gambia for over a decade earlier. This thesis, which also evaluated antimicrobial resistance patterns, showed that ST3081 isolates were more resistant to co-trimoxazole than ST618 isolates. In addition, comparative genomic analysis highlighted the role of recombination in driving the evolution of serotype 1 STs in The Gambia. It also revealed important genetic differences between these two predominant STs in The Gambia, ST3081 and ST618, in antimicrobial genes such as *tetM* and virulence genes such as exist within the fucose metabolism operon. These findings would be

useful in informing strategies to improve the monitoring and control of pneumococcal serotype 1 disease.

Table of contents

List of tables

List of figures

List of abbreviations

ATCC	American Type Culture Collection
BA	Blood Agar
bp	Base Pair
Cbp	Choline Binding Protein
CC	Clonal Complex
CFU	Colony Forming Unit
CI	Confidence Interval
CLSI	Clinical and Laboratory Standards Institute
CSF	Cerebrospinal fluid
DNA	Deoxyribonucleic Acid
dNTP	Deoxyribonucleotide Triphosphate
EDTA	Ethylenediaminetetraacetic Acid
EPI	Expanded Programme on Immunization
GBA	Gentamycin Blood Agar
Hib	*Haemophilus influenzae* Type B
HIV	Human Immunodeficiency Virus
IPD	Invasive Pneumococcal Disease
LSHTM	London School of Hygiene and Tropical Medicine
mL	Millilitres
MLST	Multilocus Sequence Typing
mM	Millimolar
mPCR	Multiplex Polymerase Chain Reaction
MRC	Medical Research Council

ng	Nanogram
NP	Nasopharyngeal
NPS	Nasopharyngeal Swab
NT	Non Typeable Pneumococci
NVT	Non Vaccine Serotypes
OM	Otitis Media
OR	Odds Ratio
PBS	Phosphate Buffered Saline
PcpA	Pneumococcal Choline Binding Protein A
PCR	Polymerase Chain Reaction
PCV	Pneumococcal Conjugate Vaccine
PhT	Pneumococcal Histidine Triad Protein
Ply	Pneumolysin
PMEN	Pneumococcal Epidemiology Network
PPSV	Pneumococcal Polysaccharide Vaccine
PsaA	Pneumococcal Surface Adhesin A
PspA	Pneumococcal Surface Protein A
RRL	Regional Reference Laboratory
RT-PCR	Real Time PCR
ß	Beta
SNP	Single Nucleotide Polymorphism
ST	Sequence Type
STGG	Skim Milk-Tryptone-Glucose-Glycerol
TAE	Tris-Acetate-EDTA

UK	United Kingdom
USA	United States of America
UV	Ultra-Violet light
VT	Vaccine serotypes
WGS	Whole Genome Sequencing
WHO	World Health Organisation
WTSI	Wellcome Trust Sanger Institute
α	Alpha
µg	Microgram
µL	Microlitre

Acknowledgements

I would like to express my sincere thanks to my supervisors, Dr Martin Antonio and Professor Brendan Wren. Your painstaking oversight and guidance have been extremely helpful as I made my way through this work. I appreciate working with you.

I acknowledge the invaluable inputs from two post-docs, Drs Brenda Kwambana and Vanessa Terra, who have worked closely with me at different stages of this project. I am grateful for your encouraging words and invaluable input, reviewing drafts of this thesis.

I am thankful to have received the MRC studentship, which provided funding for this work.

My sincere thanks also to Dr Anna Roca, Dr Grant Mackenzie and Dr Effua Usuf, (Epidemiologists);Archibald Worwui and Dr Madikay Senghore (Bioinformaticians); Catherine Okoi (Laboratory Manager) and Ebenezer Foster-Nyarko (Microbiologist) at The MRC Gambia Unit, for their helpful inputs into this work. I also thank Dr Sheikh Jarju, Dr Abdul Ceesay and all other members of the Molecular Microbiology group and the WHO Regional Reference Laboratory at The MRC Gambia Unit. You all have been wonderful in providing support to me throughout my study and for the careful collection of samples I have used for my work.

I acknowledge helpful support from the Pneumococcal African Genome (PAGe) consortium Global Pneumococcal Sequencing (GPS) project in providing the pneumococcal serotype 1 genomes I have used for my work. I also acknowledge the

CLIMB project (Cloud Infrastructure for Microbial Bioinformatics) for providing cloud-based computing, storage and analysis tools for my analysis.

I also thank my wonderful parents and parents-in-law, siblings and friends for your prayers, love, and support.

Special thanks to my dearest family, my husband and children, and to Faye Mendy. Thank you for urging me on through it all. You are the best.

Finally, I am grateful to God Almighty, for His unconditional love and ever present help.

Publications

Publications from this thesis

1. Laura Bricio-Moreno[1*], Chinelo Ebruke[2,3*], Chrispin Chaguza[1,4], Jennifer Cornick[1,4], Brenda Kwambana-Adams[2,3], Marie Yang[1], Grant Mackenzie[2,3], Brendan W. Wren[3], Dean Everett[1,4] Martin Antonio[2,3,5,¶], Aras Kadioglu[1,¶] *In vivo* modelling and comparative genomic analysis reveals the phenotypic and genotypic factors driving serotype 1 replacement in The Gambia. JID (2017) *in press.* *These authors contributed equally.

2. Cornick JE, Tastan Bishop Ö, Yalcin F, Kiran AM, Kumwenda B, Chaguza C, Govindpershad S, Ousmane S, Senghore M, du Plessis M, Pluschke G, Ebruke C, McGee L, Sigaùque B, Collard JM, Bentley SD, Kadioglu A, Antonio M, von Gottberg A, French N, Klugman KP, Heyderman RS, Alderson M, Everett DB; PAGe consortium. The global distribution and diversity of protein vaccine candidate antigens in the highly virulent *Streptococcus pnuemoniae* serotype 1. Vaccine. 2017 Feb 7;35(6):972-980. doi: 10.1016/j.vaccine.2016.12.037. Epub 2017 Jan 9.

3. Chaguza C, Cornick JE, Harris SR, Andam CP, Bricio-Moreno L, Yang M, Yalcin F, Ousmane S, Govindpersad S, Senghore M, Ebruke C, Du Plessis M, Kiran AM, Pluschke G, Sigauque B, McGee L, Klugman KP, Turner P, Corander J, Parkhill J, Collard JM, Antonio M, von Gottberg A, Heyderman RS, French N, Kadioglu A, Hanage WP, Everett DB, Bentley SD; PAGe Consortium. Understanding pneumococcal serotype 1 biology through population genomic analysis. BMC Infect Dis. 2016 Nov 8;16(1):649.

4. Kwambana-Adams B, Hanson B, Worwui A, Agbla S, Foster-Nyarko E, Ceesay F, Ebruke C, Egere U, Zhou Y, Ndukum M, Sodergren E, Barer M, Adegbola R, Weinstock G, Antonio M Rapid replacement by non-vaccine pneumococcal serotypes may mitigate the impact of the pneumococcal conjugate vaccine on nasopharyngeal bacterial ecology. Sci Rep. 2017 Aug 15;7(1):8127. doi: 10.1038/s41598-017-08717-0.

5. Kwambana-Adams BA, Asiedu-Bekoe F, Sarkodie B, Afreh OK, Kuma GK, Owusu-Okyere G, Foster-Nyarko E, Ohene SA, Okot C, Worwui AK, Okoi C, Senghore M, Otu JK, Ebruke C, Bannerman R, Amponsa-Achiano K, Opare D, Kay G, Letsa T, Kaluwa O, Appiah-Denkyira E, Bampoe V, Zaman SM, Pallen MJ, D'Alessandro U, Mwenda JM, Antonio M. An outbreak of pneumococcal meningitis among older children (≥5 years) and adults after the implementation of an infant vaccination programme with the 13-valent pneumococcal conjugate vaccine in Ghana. BMC Infect Dis. 2016 Oct 18;16(1):575.

6. Ebruke C, Roca A, Egere U, Darboe O, Hill PC, Greenwood B, Wren BW, Adegbola RA, Antonio M. 2015. Temporal changes in nasopharyngeal carriage of *Streptococcus pneumoniae* serotype 1 genotypes in healthy Gambians before and after the 7-valent pneumococcal conjugate vaccine. PeerJ 3:e903.

7. Cornick JE, Chaguza C, Harris SR, Yalcin F, Senghore M, Kiran AM, Govindpershad S, Ousmane S, Plessis MD, Pluschke G, Ebruke C, McGee L, Sigaùque B, Collard J-M, Antonio M, von Gottberg A, French N, Klugman KP, Heyderman RS, Bentley SD, Everett DB, Consortium ftP. 2015. Region-specific diversification of the highly virulent serotype 1 *Streptococcus pneumoniae*. Microbial Genomics 1.

Other relevant publications

1. Senghore M, Bayliss SC, Kwambana-Adams BA, Foster-Nyarko E, Manneh J, Dione M, Badji H, Ebruke C, Doughty EL, Thorpe HA, Jasinska AJ, Schmitt CA, Cramer JD, Turner TR, Weinstock G, Freimer NB, Pallen MJ, Feil EJ, Antonio M. Transmission of *Staphylococcus aureus* from Humans to Green Monkeys in The Gambia as Revealed by Whole-Genome Sequencing. Appl Environ Microbiol. 2016 Sep 16;82(19):5910-7.

2. Ebruke C, Dione MM, Walter B, Worwui A, Adegbola RA, Roca A, Antonio M High genetic diversity of *Staphylococcus aureus* strains colonising the nasopharynx of Gambian villagers before widespread use of pneumococcal conjugate vaccines. BMC Microbiol. 2016 Mar 12;16:38. doi: 10.1186/s12866-016-0661-3.

3. Williams TM, Loman NJ, Ebruke C, Musher DM, Adegbola RA, Pallen MJ, Weinstock GM, Antonio M. 2012. Genome analysis of a highly virulent serotype 1 strain of *Streptococcus pneumoniae* from West Africa. PLoS One 7:e26742.

4. Adetifa IM, Antonio M, Okoromah CA, Ebruke C, Inem V, Nsekpong D, Bojang A, Adegbola RA. 2012. Pre-vaccination nasopharyngeal pneumococcal carriage in a Nigerian population: epidemiology and population biology. PLoS One 7:e30548.

5. Hill PC, Townend J, Antonio M, Akisanya B, Ebruke C, Lahai G, Greenwood BM, Adegbola RA. 2010. Transmission of *Streptococcus pneumoniae* in rural Gambian villages: a longitudinal study. Clin Infect Dis 50:1468-1476

International conference posters from this thesis

1. C. Ebruke, A. Worwui, M. Senghore, B. Kwambana-Adams, B. Wren, M. Antonio: "Identifying potential pneumococcal protein vaccine candidates from *Streptococcus pneumoniae* serotype 1 isolated from West Africa", 10th International Symposium on Pneumococci and Pneumococcal Diseases ISPPD-10, Glasgow, Scotland.

2. C. Ebruke, BA. Kwambana, J. Hinds, K.A. Gould, E. Foster-Nyarko, B. Wren and M. Antonio: "Microarray serotyping in the detection of low abundance pneumococcal serotypes within the context of a PCV-7 trial'', 9th International Symposium on Pneumococci and Pneumococcal Diseases ISPPD-9, Hyderabad, India (Travel grant).

3. C. Ebruke**,** A. Roca, U. Egere, P. Hill, B. Greenwood, B. Wren, R. Adegbola and M. Antonio: "Impact of PCV-7 on nasopharyngeal carriage of *Streptococcus pneumoniae* serotype 1 genotypes in healthy Gambians'', 8th International Symposium on Pneumococci and Pneumococcal Diseases ISPPD-8, Iguaçu falls, Brazil (Travel grant).

4. C. Ebruke, O. Secka, BW. Wren, T. Corrah, R. Adegbola and M. Antonio: "Population structure of *Streptococcus pneumoniae* serotype 1 in The Gambia from 1995-2010: a dramatic shift from ST618 to ST3081 lineage'', 8th International Symposium on Pneumococci and Pneumococcal Diseases ISPPD-8, Iguaçu falls, Brazil (Travel grant).

1. Introduction

1.1. Historical perspectives

Streptococcus pneumoniae was first isolated in 1881, independently by George Sternberg, an American scientist and Louis Pasteur, a French scientist from experiments with human saliva injected subcutaneously into rabbits (Austrian, 1981). In 1884, the pneumococcus was found to be the cause of pneumonia by Albert Fraenkel (Austrian, 1960). During this period, the Gram stain was described and used to differentiate pneumococcal pneumonia from other microbes causing pneumonia (Austrian, 1960). These discoveries led to more research aimed at improving our understanding of host defence mechanisms against pneumococcal infections. In 1902, Fred Neufeld discovered capsular swelling with specific antibodies which led to the pneumococcal Quellung reaction, a method that remains the gold standard for differentiating pneumococcus into various serotypes (Austrian, 1981). He had also previously discovered the bile solubility of *S. pneumoniae* which is currently used in the rapid identification of the organism (Austrian, 1981). Building on these discoveries, Frederick Griffith in 1928 demonstrated bacterial transformation when he showed that a non-virulent strain of pneumococcus injected into mice became virulent by exposing it to heat killed virulent strains (Griffith, 1928). This subsequently led to the discovery of deoxyribonucleic acid (DNA) in 1944 as the transforming factor, when Oswald Avery, Colin MacLeod and Maclyn McCarty resolved Griffith's observation (Avery et al., 1944). This discovery has since formed the basis of the field of molecular genetics.

1.2. *Streptococcus pneumoniae* – the organism

S. pneumoniae is a Gram positive, alpha-haemolytic coccal-shaped bacterium that is usually found in pairs, described as lancet-shaped diplococci, but may also occur singly or in short chains (Figure 1.1). A non-motile, non-spore forming, and facultative anaerobic organism, it often appears on blood agar plates as round colonies, sometimes mucoid, ranging from 0.5 to 1.25 µm in diameter with smooth edges. Older cultures, 24-48hours, may appear as "draughtsman" colonies having depressed centres and raised edges. *S. pneumoniae* can be differentiated from other alpha haemolytic *Streptococci* by the optochin susceptibility and bile solubility tests, both of which are positive with *S. pneumoniae*.

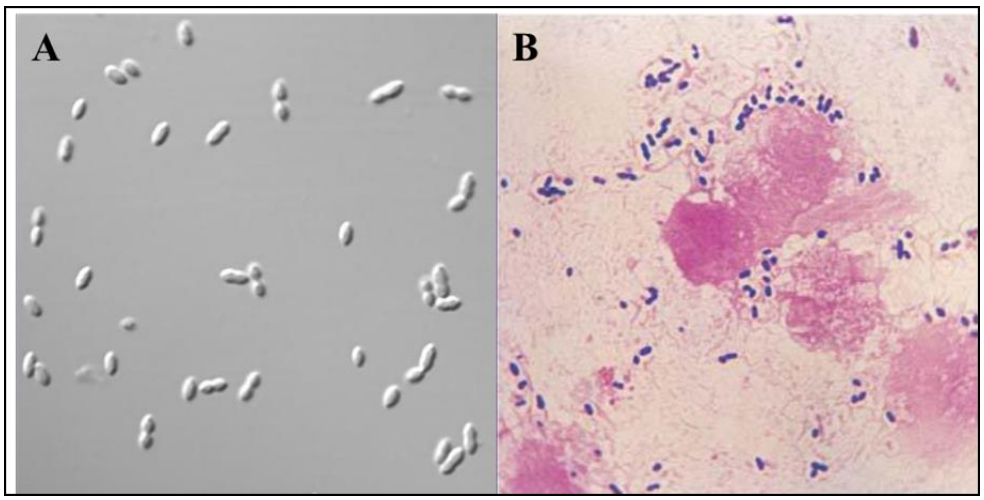

Figure 1.1 Morphological features of *Streptococcus pneumoniae*.

A) An electron micrograph of *S. pneumoniae* cells appearing singly or in pairs. B) A Gram stain showing Gram positive *S. pneumoniae* cells

The pneumococcal cell has a thick cell wall comprised of teichoic acid and peptidoglycan. A prominent feature of the cell surface of most pneumococcal strains is the capsule (White et al., 1938). Encapsulated pneumococci are known to be pathogenic and largely responsible for causing invasive pneumococcal disease to

varying extents based on their capsular type (Brueggemann et al., 2003). Several studies have shown the important role of the capsule in causing disease. The pneumococcal capsule is known to inhibit complement activity and phagocytosis as well as reducing mucus-mediated clearance from the human host (Hyams et al., 2010, Nelson et al., 2007). In addition, molecular evidence for the role of the pneumococcal capsule in virulence was demonstrated by Watson and Musher in their experimental study using mutant strains of pneumococcal serotype 3 that lacked only a polysaccharide capsule (Watson and Musher, 1990, Briles et al., 1992). They showed that the 50% lethal dose for mutant strains was greater than 5×10^7 colony forming units (CFU), in comparison to a 50% lethal dose of 1 CFU for the wild-type strains. Evidence from these studies as well as the observed decline in vaccine type pneumococcal disease following introduction of conjugate vaccines, the formulation of which have been based on capsular polysaccharides of the leading invasive pneumococcal serotypes, highlight the importance of the pneumococcal capsule in virulence of the organism. However, some other studies have shown that factors other than the capsule, such as the genetic background of pneumococcal strains, could also be important for full virulence of pneumococci (Kelly et al., 1994). Even though encapsulated pneumococci are typically known to cause invasive disease, non-encapsulated pneumococci have also been shown to cause non-invasive diseases like conjunctivitis and otitis media and in rare instances, have been responsible for a few cases of invasive disease (Valentino et al., 2014, Hotomi et al., 2016, Park et al., 2014). Based on the capsule, *S. pneumoniae* strains are divided into different serotypes that are serologically and biochemically different. Currently, there are over 95 known serotypes of pneumococci (Calix and Nahm, 2010, Oliver et al., 2013, Park et al., 2015, van Tonder et al., 2015). The pneumococcal cell surface has

several proteins and other factors that aid in evading the host immune system. Earlier reports have demonstrated that the pneumococcus exhibits phase variation in colony morphology with differential expression of the polysaccharide capsule, varying between a transparent and an opaque form. Evidence from these studies suggests that these forms of the pneumococcus play an important role in the ability of the pathogen to adapt to different environments within the human host at different times. The transparent phenotype is seen predominantly in the nasopharynx and has been shown to be the non-virulent form in established mouse models. During this phase, there is a down regulation of the cps genes which results in low levels of capsular polysaccharide being produced and more teichoic acid. This reduced expression of the capsule and higher levels of teichoic acid seen in the transparent phase is thought to facilitate nasopharyngeal carriage as teichoic acid has been shown to be important for adherence of pneumococci to the nasopharynx. In contrast to the transparent phenotype, the opaque phenotype is associated with more expression of the capsule and is known to predominate in blood (Weiser et al., 1994, Weiser et al., 1996, Manso et al., 2014). Given that the serotype specific PCVs, rolled out for routine immunization over the last two decades, have been developed based on the pneumococcal capsular types, the phenomenon of phase shift in pneumococci could have implications for overall effectiveness of these vaccines. The possibility of emergence in nasopharyngeal colonization of non-encapsulated strains of the pneumococcus, demonstrated by some recent studies (Roca et al., 2015), highlights a need for the development of pneumococcal vaccines that are able to provide wider coverage than the currently deployed PCVs. Some on-going efforts in vaccine development that employ use of non-encapsulated whole cell vaccines which express immunogenic proteins or attenuated virulence factors common to encapsulated and

non-encapsulated strains of the pneumococcus appear to be addressing these concerns (Chen et al., 2015, Goncalves et al., 2014).

The first complete pneumococcal genome was published in 2001, a serotype 4 strain comprising of 2.16 Mbps encoding 2236 predicted coding sequences (Tettelin et al., 2001). Subsequently, the genome sequence of several other strains across different serotypes have been published (Lanie et al., 2007, Hahn et al., 2015, Ip et al., 2015, Cornick et al., 2015, Williams et al., 2012). Being a naturally transforming organism, high rates of recombination occur within the genome of pneumococcal strains causing variability even amongst pneumococcal strains belonging to the same serotype (Croucher et al., 2011).

1.3. Epidemiology of *S. pneumoniae* diseases

The epidemiology of *S. pneumoniae* relates in most part to its ability to cause invasive pneumococcal disease (IPD) and these diseases occur in all age groups although the risks are highest in young children and the elderly. IPD is the isolation of *S. pneumoniae* from a normally sterile body site such as blood, cerebrospinal fluid, pleural or pericardial fluid and has been estimated to be responsible for 700,000 to 1 million deaths in young children each year (O'Brien et al., 2009). Global estimates for child deaths indicate that pneumococcal disease case fatality rates are highest in Africa and South Asia (Figure 1.2). The average year to year pneumococcal disease incidence varied between countries from as low as 188 (131-284) to as high as 6387 (4937-7909) per 100,000 in children less than 5 years with the highest incidence rates in Africa (O'Brien et al., 2009).

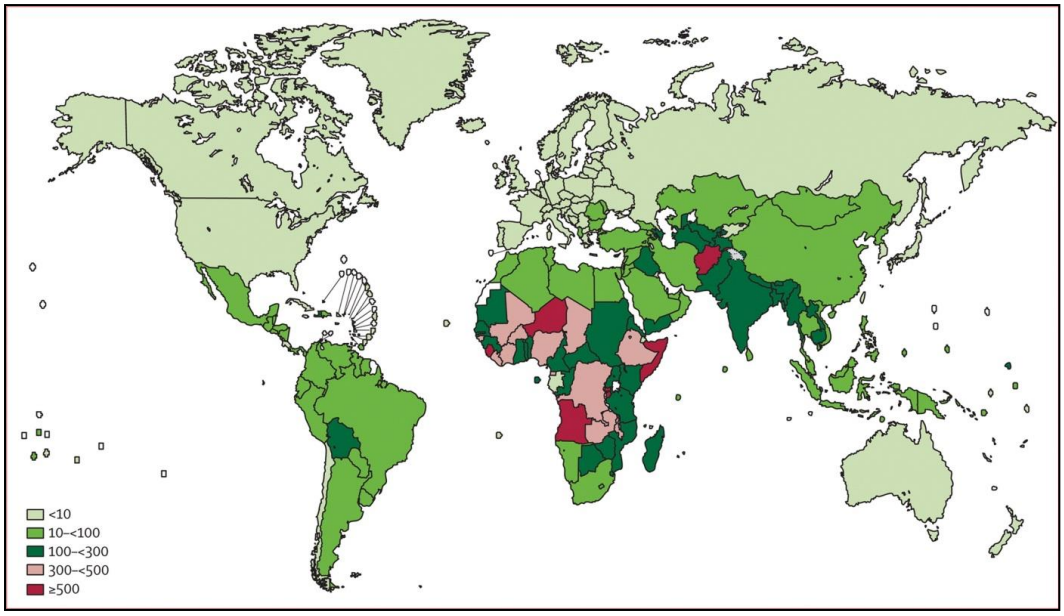

Figure 1.2 Pneumococcal deaths in HIV-negative children aged 1-59 months per 100,000 children. Adapted from (O'Brien et al., 2009).

Studies conducted previously in The Gambia prior to the introduction of pneumococcal conjugate vaccines showed incidence rates of IPD to be 224, 139 and 82 per 100,000 child years for children aged 2-11, 12- 23 and 24-35 months respectively in a peri-urban setting (Usen et al., 1998). Another study done in a rural setting of The Gambia reported even higher IPD incidence rates of 554 and 458/100,000 in children less than 1 and 2 years old respectively (O'Dempsey et al., 1996).

More recently, a population-based surveillance for IPD conducted in rural Gambia over several years during and after the implementation of PCVs showed a reduction in the incidence of IPD from the baseline values. Specifically, reductions in IPD incidence was observed across all age groups with rates dropping from 253 cases per 100,000 to 113 cases per 100,000; 113 to 49 cases per 100,000; 12 to 10 cases per

100000 and 9 to 4 cases per 100,000 in children age 2-23 months, 2-4 years, 5-14 years and adults ≥ 15 years respectively (Mackenzie et al., 2016).

In addition to regional and age-related differences in the epidemiology of IPD, the different clinical syndromes caused by invasive pneumococcal disease also have an influence on its epidemiology. Pneumonia, septicaemia and meningitis have been reported more frequently as clinical syndromes of IPD and are important causes of deaths globally.

1.3.1. Pneumonia

Pneumonia, a lower respiratory tract infection that specifically affects the lungs, is the leading cause of child deaths globally. Pneumonia illness is classified clinically either as a non-severe or a severe pneumonia illness (WHO, 2013). Childhood pneumonia deaths are often due to a severe pneumonia illness with approximately 1 in 5 of these childhood pneumonia deaths caused by *S. pneumoniae* (Black et al., 2010). Global disease burden estimates indicate there were as many as 120 million episodes of pneumonia in 2011, including 14 million episodes of severe pneumonia in children less than 5 years (Walker et al., 2013). These estimates represent widely varying regional incidence rates with reported rates lowest in the European region and highest in the African and Southeast Asian regions. As many as 30% and 39% of all severe pneumonia episodes occurred during this period in the African and southeast Asian regions respectively. The distribution pattern of childhood pneumonia deaths across world regions mirrors rates observed for pneumonia incidence. The highest number of deaths was reported in sub-Saharan Africa, accounting for as much as 43% of all childhood pneumonia deaths in 2011. Estimates for the year 2013, showed that pneumonia was responsible for

approximately 935,000 deaths of children under the age of five years and that this accounted for 15% of all deaths of children under five years old globally (Figure 1.3) (Liu et al., 2015).

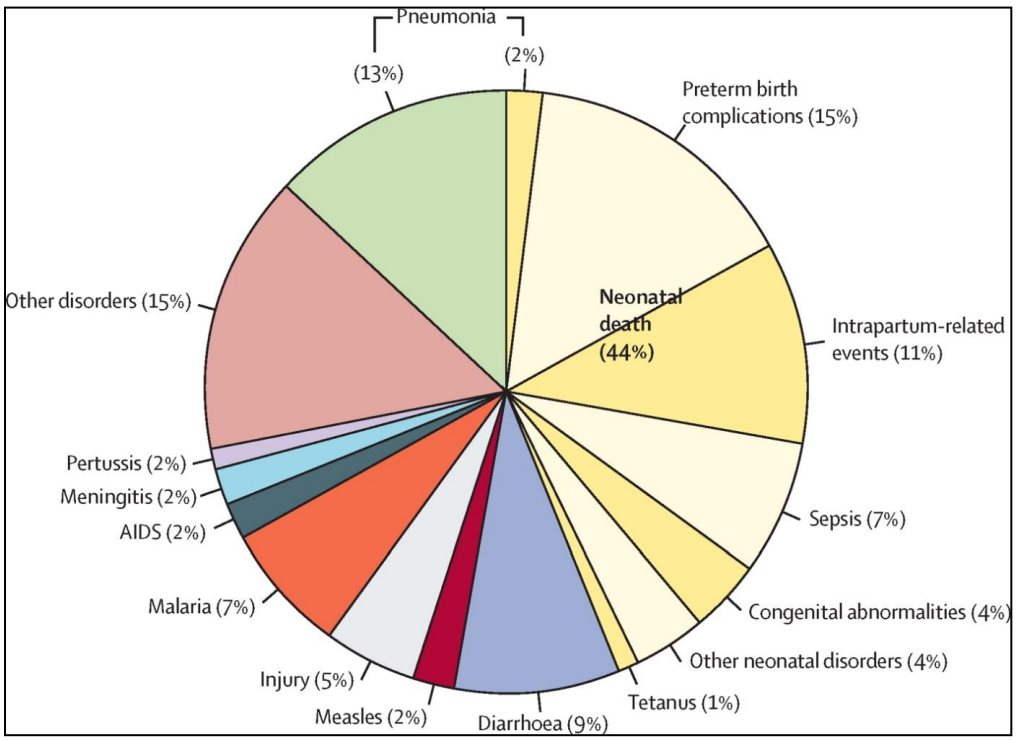

Figure 1.3 Global causes of death in children less than 5 years in 2013. Adapted from (Liu et al., 2015).

Bacterial pathogens have been identified more frequently than other pathogens in severe pneumonia cases with *S. pneumoniae*, *H. influenzae* and *Staphylococcus aureus* the leading bacteria (Adegbola et al., 1994) prior to the introduction of PCVs. Of these bacterial pathogens, *S. pneumoniae* has remained the predominant bacterial cause of pneumonia in aetiology studies conducted both before and after the introduction of routine vaccination with pneumococcal conjugate vaccines with non-vaccine type pneumococci predominating in the post-PCV era (Rudan et al., 2008, Rudan et al., 2013, Carrol et al., 2011).

1.3.2. Meningitis

Bacterial meningitis has been known to be associated with very high fatality, with reported annual mortality incidence in developing countries of 98 per 100,000 in children less than 1 year (Gessner *et al* 2010). Outbreaks of epidemic meningitis are particularly frequent in the African meningitis belt, a region that extends from The Gambia, in the west, to Ethiopia and Eritrea in the east (Leimkugel et al., 2005, Yaro et al., 2006). Predominant bacterial causes of meningitis include *Neisseria meningitidis, S. pneumoniae* and *Haemophilus influenzae* type b (Hib) in the pre-Hib vaccine era. Traditionally, the meningococcus has been known to be a leading cause of epidemic meningitis, but reports from the West African sub-region have indicated that the pneumococcus plays a more significant role in causing meningitis than thought previously (Gessner et al., 2010, O'Brien et al., 2009). The public health significance of bacterial meningitis extends beyond the deaths caused by this disease. Survivors are frequently left with significant sequelae. Some reports indicate that approximately half of all surviving children have life-long neurologic impairment including deafness, learning disabilities and seizures (Edmond et al., 2010, Goetghebuer et al., 2000). A systematic review also showed that the risk of at least one major complication was highest following *S. pneumoniae* meningitis in comparison to outcomes in meningitis caused by *N. meningitidis* or *Hi*b (median (IQR) risks of 24.7% (16.2-35.3%), 7.2% (4.3-11.2%) and 9.5% (7.1-15.3%) respectively) (Edmond et al., 2010).

1.3.3. Sepsis

Sepsis refers to the presence of rapidly dividing bacteria in the blood stream with manifest signs of illness in the individual. Global estimates for incidence of

pneumococcal sepsis is reported be as frequent as 87 per 100,000 persons with highest rates seen in Africa (O'Brien et al., 2009). The case fatality rates for sepsis across all age groups ranged from 22% in the Western Pacific region to 58% in Africa region with higher rates seen in very young children and the elderly (O'Brien et al., 2009).

1.3.4. Non-invasive pneumococcal disease

Pneumococci are also known to cause less severe and non-invasive forms of disease such as otitis media (middle ear infection), sinusitis and bronchitis. Pneumococci have been identified in as many as 57% of middle ear aspirates taken from cases with acute otitis media (Tamir et al., 2015). Other bacteria known to cause otitis media include *H. influenzae* and *Moraxella catarrhalis* (Tamir et al., 2015).

1.4. Risk Factors for invasive *S. pneumoniae* disease

The risk for IPD has been shown to be significantly higher in certain groups or associated with specific factors.

1.4.1. Age

The risk for IPD has been shown to be significantly higher in children younger than two years of age and elderly persons older than 65 years (Robinson et al., 2001, Trotter et al., 2010). This may be related to relative immaturity or reduced competence of the immune system in these individuals to combat infection *S. pneumoniae*. In addition, data from different parts of the world indicate that the influence of age on IPD risk varies across regions. Children in developing countries were reported to have a lower median age for IPD incidence than was seen in children in developed countries (Greenwood, 1999).

1.4.2. Sex

Males have been shown to have a higher risk for IPD (Nuorti et al., 2000). Reports from IPD surveillance across different geographical regions indicate that higher rates of IPD in males have persisted through the pre-PCV era into the post PCV era (de St Maurice et al., 2016, Wagenvoort et al., 2017). A recent study that utilized population and laboratory-based IPD surveillance data from 1998-2013 showed that IPD incidence rates in male subjects were as much as 1.5- 2 times higher than rates seen in females, regardless of race or existing co-morbidities (de St Maurice et al., 2016). Earlier experimental studies using mice models also demonstrated higher susceptibility among males to respiratory and systemic pneumococcal disease (Kadioglu et al., 2011).This relationship remains to be clearly understood. A previous report suggested that this observation might be confounded by gender differences in dietary and alcohol intake in adults (Cortese et al., 1992).

1.4.3. Seasonality

Reports from different geographical regions of the world have provided evidence of the seasonal patterns in IPD incidence. Incidence of IPD peaks in the winter periods in Europe and the United States of America (Ampofo et al., 2008) (Trotter et al., 2010) whilst in The Gambia, the peak occurs in the rainy season (O'Dempsey et al., 1996) and in the cold dry months in Malawi (Gordon et al., 2001).

1.4.4. Socioeconomic factors

Socioeconomic status, as assessed using a variety of factors in different studies, has been shown to be consistently associated with the risk of IPD. Persons in the lower

socioeconomic class have higher risk of IPD. Indicators of low social class such as overcrowding, illiteracy, malnutrition and lower income have all been shown to increase pneumococcal disease incidence (Chen et al., 1998, Nuorti et al., 2000, Savitha et al., 2007, Ciruela et al., 2013).

1.4.5. Chronic medical conditions and immunosuppression

Chronic illnesses or underlying medical conditions such as diabetes mellitus, chronic lung disease including asthma or chronic heart disease, have all been shown to significantly increase the risk of IPD (Ortqvist et al., 2005, Frankel et al., 1996). Additionally, human immunodeficiency virus (HIV) infected persons have been reported to have significantly higher risks for IPD in comparison to HIV negative persons (Nunes et al., 2011, Yin et al., 2012).

1.4.6. Genetic predisposition

Evidence for the influence of genes on IPD risk have come from various research studies including research evaluating the role of the sickle cell gene. The sickle haemoglobin (HbS) is a variant of the normal adult haemoglobin (HbA) and the sickle cell gene is inherited as a recessive trait. Individuals who are homozygous for the gene (HbSS) are regarded as having sickle cell disease (SCD). Sickle cell disease patients are reported to have significantly higher risks, as much as a 600-fold increased risk for IPD and higher IPD incidence than the general population (Greenwood, 1999, Williams et al., 2009). This increased risk in SCD patients is thought to be related to defects in splenic function and complement deficiency. The increased risk for IPD in sickle cell disease patients is important given that the burden of sickle cell disease is known to be highest in Africa. Global estimates of

sickle cell disease burden indicate that in 2010, of 305,800 new-borns delivered annually with the sickle cell disease, 79% occurred in sub-Saharan Africa (Piel et al., 2013). It may not be surprising, therefore, that high rates of pneumococcal nasopharyngeal carriage and invasive disease have been reported in many African countries, in comparison to rates observed in other parts of the world (O'Brien et al., 2009).

The roles of some other genetic factors, such as single nucleotide polymorphisms in selected genes involved in the innate and adaptive immune response, in individual differences in susceptibility to pneumococcal disease have been evaluated by different studies but the results have been largely inconclusive (Brouwer et al., 2009).

1.5. Epidemiology of *S. pneumoniae* carriage

S. pneumoniae is asymptomatically carried in the upper respiratory tract of human beings, known to be its major natural reservoir, hence its dependence on humans for survival. The nasopharynx is the preferred sampling site for pneumococcal carriage studies although some studies have also considered sampling from the oropharynx (Odutola et al., 2013). *S. pneumoniae* is carried simultaneously with a wide range of other organisms in the nasopharynx, constituting a complex microbial community. Some of these organisms include *Haemophilus influenzae, Moraxella catarrhalis, Staphylococcus aureus* and other *Streptococcus* species.

Nasopharyngeal carriage of pneumococci is believed to be an important step that precedes invasive disease and has been used as a measure for efficacy of vaccines (Bogaert et al., 2004a). Carriage rates of *S. pneumoniae* have varied across different regions of the world with higher rates recorded in developing countries compared to

developed countries. For example, relatively low carriage rates of 8.6% in children less than 5 years of age have been reported in Italy (Marchisio et al., 2002). By contrast, higher carriage rates between 85 and 97% have been found in children less than 5 years of age in The Gambia (Hill et al., 2006, Cheung et al., 2009, Obaro et al., 1996). Also similar rates between 50 and 90% have been found across other countries in sub-Saharan Africa (Vallès et al., 2006, Adetifa et al., 2012, Woolfson et al., 1997, Joloba et al., 2001, Feikin et al., 2003).

Pneumococcal serotypes found most frequently in nasopharyngeal carriage are serotypes in groups 19, 6, and 23 while those rarely found include serotypes 1, 5 and 7F (Roca et al., 2011, Feikin et al., 2003, Ba et al., 2014).

1.6. Risk factors for *S. pneumoniae* carriage

Low socio-economic status has been shown to be a risk for pneumococcal carriage in some studies. Overcrowding and poor ventilation often fosters transmission of pneumococci. Also, lack of education has been identified as a risk factor for pneumococcal carriage in some studies. Carriage rates also vary across age groups with a decreasing carriage rate observed with increasing age. For example, in The Gambia and Kenya respectively, carriage rates of 93.4% and 57.7 % was reported in children less than 5 years; 86.3% and 41.0 % in children aged 5 – 15 years and 60.6% and 6.4 % in individuals above 15 years (Abdullahi et al., 2008, Hill et al., 2006). Seasonality is another factor affecting pneumococcal carriage. A recent study conducted in three different geographical regions; South East Asia, Africa and North America showed that *S. pneumoniae* are optimally transmitted during the cooler and driest months (Numminen et al., 2015). Also attendance of day care centres; smoking history of parents, use of wood for cooking fuel , antibiotic therapy have

also been identified as risk factors in different settings (Regev-Yochay et al., 2012, Cardozo et al., 2008).

1.7. Pneumococcal population biology

1.7.1. Definition and importance of pneumococcal population biology

The concept of bacterial population biology deals with study of the differences and similarities of genetic expression of a bacterial species that exists within a local community. It also aims to provide insights into the evolutionary mechanisms behind the emergence of the existing variants of the bacteria over a period of time, including the effects of any interventions in the community on that bacterial population. Some bacterial strains are known to exist as commensals in the human body, whilst other strains have been associated with causing invasive disease. Bacterial population biology also explores the relationships between invasive bacterial strains and the rest of the bacterial population, including the specific factors driving the invasive potential of these bacterial strains. For the purpose of this thesis, this section focuses on the population biology of *Streptococcus pneumoniae*.

1.7.2. Sampling approach and the relationship between pneumococcal carriage and disease

Pneumococcal strains are known to reside in the human nasopharynx as part of the normal bacterial flora, but are also able to cause invasive disease. Nasopharyngeal carriage is known to precede onset of invasive pneumococcal disease (Bogaert et al., 2004a, Gray et al., 1980). Some studies have shown that an understanding of the carriage state is vital for a better understanding of pneumococcal population biology (Brueggemann et al., 2003). This is because the population of pneumococci in the

nasopharynx more accurately represents the natural population of pneumococci circulating in that community. As such, it is thought that pneumococcal disease isolates are best considered within the context of the carriage population circulating in that community over the same time period. In addition, some authors have suggested that this approach helps avoid the risk of over- or under- estimating the invasive potential of pneumococcal isolates seen in invasive disease because it takes into account the level of exposure of the population to that serotype or strain. Also, it would provide better clarity on how pneumococcal disease relates to carriage (Brueggemann et al., 2003). Therefore, an effective approach to the study of pneumococcal population biology would include an understanding of the pneumococcal carriage population, the mechanisms of transmission and how this could potentially inform on the likely patterns of pneumococcal invasive disease within that local community.

1.7.3. Tools for understanding pneumococcal population biology

Streptococcus pneumoniae strains have been characterized by serotyping, a method that distinguishes specific pneumococcal types based on their capsular polysaccharides, which also serve as an important factor in determining the virulence potential of the pneumococcal strain (Briles et al., 1992). About 100 different pneumococcal serotypes have been identified, even though only a handful of these are known to cause disease. However, the advent of molecular typing methods have shown that within each pneumococcal serotype exists clearly distinct genotypes and also allowed for identification of pneumococcal strains with closely related genetic profiles as clones or clonal complexes (Enright and Spratt, 1998). Of these molecular typing methods, the use of MLST and affordability of WGS that gives access to the entire genome of *S. pneumoniae*, have allowed for a more informative description

and a clearer understanding of the evolving population biology of *Streptococcus pneumoniae* and mechanisms for expansion. These and other pneumococcal typing methods are described later in this chapter

1.7.4. Evolution of pneumococci: Molecular mechanisms of change in the pneumococcal population biology

Since its' discovery in the 1900, *Streptococcus pneumoniae*, a highly transforming pathogen, has evolved over time, often times as a response to the selective pressure of interventions such as antibiotic usage and the introduction of conjugate vaccines (Albrich et al., 2004, Chiba et al., 2014, Gladstone et al., 2015b, Croucher et al., 2011). Additionally, it is thought that evolution of the pneumococcus could also be due to secular changes, changes that are not attributable to any specific interventions in the community (Jefferies et al., 2010). A good understanding of the population biology of *Streptococcus pneumoniae* in different geographical regions and across different age groups will provide clearer insights into the mechanisms of evolution and could inform on more effective strategies for control of pneumococcal carriage and disease. Studies into the mechanisms by which the pneumococcus has evolved have shown that this has included recombination of genomic material and point mutation, with recombination events occurring far more commonly (Croucher et al., 2011).

High rates of recombination occurring within the genomes of *S. pneumonaie* have been shown to be one of the main drivers of the changing population structure of *S. pneumoniae*. In addition to exchange of genetic material between pneumococcal strains, recombination exchanges are also known to occur between pneumococci and other streptococci which colonize the human upper respiratory tract, such as

Streptococcus mitis, *Streptococcus infantis* and *Streptococcus oralis* (Donati et al., 2010, Sanguinetti et al., 2012). Some of these exchanges across species have involved virulence genes such as those encoding for pneumolysin, mitilysin and neuraminidase A (Johnston et al., 2010). It would seem obvious therefore, recombination events within pneumococcal strains and between the pneumococcus and other species have contributed to the evolution and structural diversity of the pneumococcus (Kilian et al., 2014).

Studies using whole genome sequencing have also shown that rates of recombination are higher in non-typeable pneumococci which do not express a capsule than in encapsulated (typeable) pneumococci (Chewapreecha et al., 2014b, Croucher et al., 2011). Non encapsulated pneumococci have been described frequently in carriage population and only rarely seen in invasive disease (Chewapreecha et al., 2014a, Marsh et al., 2010). Nevertheless, these non-encapsulated pneumococci are believed to play an important part in the pneumococcal population dynamics through switching between encapsulated and non-encapsulated states and thereby facilitate increases in antibiotic resistance. The importance of genetic recombination in promoting development of antibiotic resistant pneumococcal strains and facilitating pneumococcal virulence has been well-documented (Slager et al., 2014, Griffith, 1928, Dowson et al., 1993). In addition, studies have demonstrated that pneumococcal serotype changes can be explained by recombination events occurring at the capsular locus (Salter et al., 2012, Mostowy et al., 2017). Serotype changes have been reported commonly in the context of concurrent nasopharyngeal carriage of multiple pneumococcal serotypes and frequently occur between vaccine serotypes (Kamng'ona et al., 2015).

1.7.5. Factors driving pneumococcal evolutionary changes

1.7.5.1. Pneumococcal conjugate vaccination

The use of PCVs, from as early as the year 2000 in the United States, has been effective in reducing the incidence of invasive disease and carriage of vaccine serotypes. Prevalence of PCV7 serotypes was seen to decrease from 64% of invasive and 50% of non-invasive isolates in 1999–2000 to 3.8% and 4.2%, respectively, across all age groups, in 2010–2011 (Richter et al., 2013). This decrease in prevalence of vaccine type serotypes following introduction of PCV-7, was also observed in other parts of the world (Gladstone et al., 2015b, Roca et al., 2011, Feikin et al., 2013, Nzenze et al., 2013) However use of PCVs have also resulted in an increase in non-vaccine serotypes, known as serotype replacement (Gladstone et al., 2015b). Another is the phenomenon referred to as capsular switching which occurs when a vaccine serotype is observed to acquire the capsule of a non-vaccine serotype (Wyres et al., 2013). Following introduction of PCV-7 in the USA, a shift in the pneumococcal population structure was observed as a result of the emergence of serotype 19A, a non PVC-7 vaccine serotype, due to capsular switching events with the PCV-7 vaccine serotype 19F, occurring in the PMEN1 lineage of pneumococcus (Croucher et al., 2011). Similarly, capsular switching events have been demonstrated in the switch from serotype 4, a PCV7 vaccine serotype, to serotype 19A (Golubchik et al., 2012). More recently, expansion of the serotype 35B, ST156 lineage was reported in The USA, following the introduction of PCV-13, which excludes serotype 35B. This expansion was due to a capsule switching event occurring between the serotype 35B, ST558 lineage and serotype 9V, ST156 lineage giving rise to the serotype 35B, ST156 lineage (Chochua et al., 2017). The prevalence of serotype 35B appears to have increased in both invasive and non-

invasive disease (Richter et al., 2014, Richter et al., 2013). Furthermore, through capsular switching, vaccine type pneumococcal strains have been able to evade the effects of conjugate vaccines, and sometimes become more virulent in the process. ST 320, originally seen as serotype 19F, a PCV-7 serotype, has since been seen in the post PCV-7 era to exhibit the capsule of serotype 19A, a non PCV-7 serotype. This serotype 19A variant was found to have spread across many parts of the United States and to other parts of the world (Ansaldi et al., 2011, Golubchik et al., 2012). The ST 320 has been shown to be highly antibiotic resistant, possessing the *mefE* and *ermB* genes which confer resistance to macrolide antibiotics (Bowers et al., 2012, Croucher et al., 2013).

Another way in which vaccination has influenced evolution of the pneumococcus is how selective vaccine-induced pressure has contributed towards increases in levels of antibiotic resistance amongst circulating pneumococci. With introduction of conjugate vaccines, the prevalence of penicillin non-susceptible pneumococci among non-vaccine type pneumococci has been noted to increase and become even more diverse (Gertz et al., 2010, Gherardi et al., 2012).

Therefore, introduction of conjugate vaccines has triggered adaptive responses by the pneumococcus that have resulted in increasing levels of evolution of this highly transforming pathogen. Further effects on the pneumococcal population structure of long term use of conjugate vaccines remains unclear. The absence of more effective preventative strategies, such as use of protein antigen based vaccines may lead to further rounds of serotype replacement (Devine et al., 2017, Gladstone et al., 2015b).

1.7.5.2. Antibiotic Usage

The use of antibiotics has been known to induce changes that have contributed to evolution of the pneumococcus, with emergence and spread of multi-drug resistant clones (McGee et al., 2001). The Pneumococcal Molecular Epidemiology Network clone 1 (PMEN1), an *S. pneumoniae* lineage, one of the first described pneumococcal antibiotic resistant clones, has been shown to be resistant to penicillin, chloramphenicol and tetracycline (Croucher et al., 2009). This clone, first identified in Barcelona in 1984, has subsequently been found in parts of Africa, Asia and America (Munoz et al., 1991, Parry et al., 2002, McGee et al., 2001). These evolutionary changes that have occurred in response to antibiotic-induced stress in the pneumococcus, have been mediated through different mechanisms including mutation, horizontal gene transfer and genetic recombination with the pneumococcus acquiring genes that confer resistance to known antibiotics (Slager et al., 2014, Hakenbeck et al., 1999, Dowson et al., 1989, Hanage et al., 2009). Previous studies have described mutations in penicillin binding protein (PBP) genes that conferred resistance to beta lactam antibiotics such as penicillin, amoxicillin and cefotaxime. Additionally, recombination events have been shown to occur between pneumococcal serotypes, such as between serotypes 9V and 23F, 9N and 14, 35C and 17F and 12F and 7F, and have included PBP genes. These recombination events have been associated with the emergence of pneumococcal strains with resistance to other classes of antibiotics (Hanage et al., 2009, Wyres et al., 2013). Recombination events have facilitated the transfer of transposon and integrative conjugative elements (ICEs) that have conferred antibiotic resistant properties across pneumococcal populations. The transposon Tn2010 is known to carry the *mefE* and *ermB* genes within the macrolide-resistant pneumococcal strain, ST 320 (Bowers et

al., 2012, Croucher et al., 2013). Other transposons such as Tn916, Tn5251, Tn5252, Tn5253 and Tn1545 are known to carry genes which confer resistance to chloramphenicol and tetracycline (Rice, 1998, Ayoubi et al., 1991).

The pattern of antibiotic resistance amongst pneumococcal isolates is also known to be influenced by changes in prevalence of specific serotypes following the introduction of pneumococcal conjugate vaccination. In the USA, introduction of PCV7 was seen to precede changes in the sequences of resistance genes against macrolides and β-lactams (Bowers et al., 2012, Pelton et al., 2007).

1.7.5.3. Secular changes

Regardless of the influence of vaccination or other major interventions in a community, temporal fluctuations in the prevalence of circulating serotypes or genotypes of pneumococci have been reported in different geographical settings (Finland and Barnes, 1977, Jefferies et al., 2010). In a study that employed MLST and eBURST analysis of invasive pneumococcal isolates collected over a period of 5 years preceding the introduction of PCV-7 in Scotland, considerable changes in serogroup and clonal distributions were observed throughout the study period. Only 9.8% of the 338 different STs identified were seen to be stable and associated with invasive disease every year of the study period. Many new STs were observed each year and persisted only transiently. The exception to this was the significant increase in serotype 1 ST 306 clone which was noted to persist throughout the 5 year period. Similar to the findings in Scotland, another study that performed large scale MLST characterization of both invasive and carriage pneumococcal isolates collected over a 15 year period prior to introduction of PCV in an African country, fluctuations in the distributions of circulating STs were observed (Brueggemann et al., 2013). The authors noted that there were no major changes in the study area, such as changes in

antimicrobial use, laboratory practice or vaccine uptake that would have accounted for the fluctuations observed during the study period.

It seems plausible therefore that, regardless of the response of the pneumococcus to selective pressure from interventions such as vaccination and antibiotic usage, secular changes in the distribution of circulating genotypes and serotypes would occur as part of the natural evolution of this pathogen.

1.7.6. Regional Differences in Population Biology of *S. pneumoniae*

Studies carried out in different regions of the world have demonstrated some similarities in the population biology of *S. pneumoniae* such as the global spread of some antibiotic resistant strains and the effects of vaccination on prevalence of vaccine type pneumococci. However, these studies have also highlighted important differences between regions in the population of circulating pneumococci, such as differences in the predominant serotypes and sequence types seen in invasive disease and in nasopharyngeal carriage.

Following the introduction of PCV in many European countries, there were significant declines in IPD caused by vaccine serotypes (Tin Tin Htar et al., 2015). Some of the more prevalent serotypes seen in the post-PCV13 era included 24 F, 22 F, 8 and 15A, with serotypes 19A and 3 the more prevalent serotypes seen in PCV10-using countries. Significant reductions in prevalence of serotypes 19A, 7 F, 1 and 6A were observed in countries using PCV13. Similar reductions in prevalence of vaccine serotypes following PCV introduction has been observed in the USA (Richter et al., 2013). Non-PCV13 vaccine types were reported to emerge in the post-PCV13 period, even though most were seen to be less invasive in comparison to the pre-existing vaccine serotypes. Some of the serotypes with higher invasive

potential reported from the UK include serotypes 7F, 19A, 3, 8, and 33F (van Hoek et al., 2014).

In contrast to observations in Western countries, the serotypes with highest prevalence in invasive disease in West Africa were predominantly serotypes 1 and 5 (Adegbola et al., 2006). Serotype 3 was rarely seen in invasive disease but more common in carriage among West African isolates, contrasting with the pattern of serotype 3 epidemiology in some Western countries (Donkor et al., 2013, Martin and Brett, 1996). In addition, a few studies that evaluated the genetic population structure of African pneumococcal isolates revealed high levels of serotype and clonal diversity (Donkor et al., 2013, Chaguza et al., 2017). Serotypes 1 and 5, the leading causes of IPD in sub-Saharan Africa, were seen to be more clonal than other serotypes that were seen more commonly in carriage. Also, the PMEN clones, commonly seen in European and other Western countries, were rarely seen among African pneumococcal isolates (Donkor et al., 2013, Chaguza et al., 2017).

1.8. *S. pneumoniae* serotype 1 disease

S. pneumoniae serotype 1 is one the leading causes of IPD in different parts of the world (Adegbola et al., 2006, Holliman et al., 2007, Hausdorff et al., 2000, Chiou et al., 2008). It has been associated with pneumonia, meningitis, sepsis, empyema and rarer diseases such as peritonitis and salpingitis (Sirotnak et al., 1996, Westh et al., 1990, Byington et al., 2002, Eltringham et al., 2003). In The Gambia, prior to the introduction of pneumococcal vaccines, it was the leading cause of IPD responsible for 20% of all cases (Adegbola et al., 2006). Other African countries have reported similar trends including Kenya (Brueggemann et al., 2013), Mozambique (Vallès et al., 2006), Togo, Burkina Faso (Traore et al., 2009) and Niger (Collard et al., 2013).

In developed countries including Spain, Denmark, Portugal, France, Switzerland and the UK, serotype 1 also plays a significant role in IPD (McChlery et al., 2005, Kronenberg et al., 2006, Konradsen and Kaltoft, 2002, Henriques Normark et al., 2001, Hanquet et al., 2010, Aguiar et al., 2010). Some studies have attributed the increase in serotype 1 IPD to use of PCV-7 and tagged it a replacement non-vaccine serotype whilst other studies have attributed this to secular trends as many countries monitoring serotype 1 IPD over several years have seen annual fluctuations without the influence of vaccines (Hanquet et al., 2010). Interestingly, serotype 1 is one of few pneumococcal serotypes associated with outbreaks and has been responsible for several epidemics in different regions of the world (Leimkugel et al., 2005, Dagan et al., 2000, Staples et al., 2015). Across the meningitis belt, it has been responsible for highly lethal meningitis outbreaks in as seen recently in Ghana (Kwambana-Adams et al., 2016) and previously in Ghana and Togo with features similar to outbreaks caused by *Neisseria meningitidis* (Yaro et al., 2006, Leimkugel et al., 2005). A review of cases in the African meningitis belt revealed that serotype 1 was responsible for 59-79% of cases (Gessner et al., 2010). Outbreaks of serotype 1 have also been reported in men's shelters in Boston (DeMaria et al., 1980) and Paris (Mercat et al., 1991); in closed communities in Israel (Dagan et al., 2000) and in Australia (Lai et al., 2013, Gratten et al., 1993, Staples et al., 2015). In contrast to the pattern of a high potential for causing severe disease with fatal outcomes in developing country settings, pneumococcal serotype 1 has been reported in some developed countries to cause non-severe disease with non-fatal outcomes (Sjostrom et al., 2006).

Other interesting characteristics of pneumococcal serotype 1 include a high susceptibility to many antibiotics (Henriques Normark et al., 2001, Antonio et al.,

2008) and in some studies affecting mostly older children and young adults (Collard et al., 2013, Traore et al., 2009).

It has also been shown to have distinct geographical differences which may be due to its rarity in nasopharyngeal carriage (Brueggemann and Spratt, 2003). In a phylogenetic analysis conducted on 166 invasive serotype 1 isolates obtained from 14 countries, three lineages were observed. Lineage A isolates were exclusively from Europe and North America with ST306 mainly from Europe and ST227 from North America, England and Canada. Lineage B was predominantly from Africa, majorly ST217 and lineage C from Chile, mostly ST615 (Brueggemann and Spratt, 2003). This geographic distinction has been supported by various studies, for example, the ST217 clonal complex is the predominant clone across many countries in the African meningitis belt such as The Gambia, Ghana, Burkina Faso, Togo and Niger (Antonio et al., 2008, Leimkugel et al., 2005, Yaro et al., 2006). Three of the licensed pneumococcal vaccines include protection against serotype 1 and are discussed in detail in section 1.9.

1.9. *S. pneumoniae* serotype 1 carriage

Although pneumococcal serotype 1 is one of the leading serotypes causing invasive pneumococcal disease (IPD) in many regions of the world, particularly in sub-Saharan Africa, it is rarely found in carriage studies. In The Gambia, serotype 1 was found in 0.5% of 2478 isolates in a pneumococcal carriage study (Hill et al., 2006). This finding was similar to a report from Mozambique, where serotype 1 was the leading serotypes among invasive isolates (40%), with only one (0.5%) serotype 1 isolate found among 192 carriage isolates (Vallès et al., 2006). In Kenya, no serotype

1 isolate was recovered from carriage despite being the leading cause of invasive disease (Brueggemann et al., 2013). Pneumococcal serotype 1 carriage rates reported in developed countries are largely similar to rates reported in developing country settings. In Switzerland, serotype 1 was found in 26 (1.7%) out of 1540 carriage isolates (Kronenberg et al., 2006) while in a study conducted in the UK comparing invasive and carriage isolates collected in the same region over the same time period, no serotype 1 isolate was found in carriage despite been found in invasive disease (Sleeman et al., 2006). It is unclear why the carriage rates of serotype 1 are so rare. One possible reason could be that commonly used culture methods are not sensitive enough to detect it in carriage. A study showed improved detection of *S. pneumoniae* in saliva samples compared to the nasopharynx using molecular methods (Krone et al., 2015). Another explanation could be due to a presumably short duration of carriage. In a longitudinal study carried out in Australia with 2-4 weeks sampling, only one child carried serotype 1 at 2 consecutive examinations. Serotype 1 was also found in low and high densities and co-carried with other serotypes in 34% of samples tested (Smith-Vaughan et al., 2009). The authors suggested that serotype 1 carriage may be underestimated due to masking by other serotypes in co-colonization (Smith-Vaughan et al., 2009). Prevalence of serotype 1 carriage has been reported to be high during outbreaks caused by this serotype. In an outbreak of serotype 1 pneumococcal pneumonia in a men's shelter in Boston in 1978, 10 of 104 (10%) residents of the shelter were carriers of pneumococcal serotype 1 (DeMaria et al., 1980) compared to 1 – 2% or less in the absence of an outbreak. Similarly, in 1938 in Worcester Mass, the pneumonia outbreak that occurred was largely due to serotype 1 and the carriage rate was as high as 10% compared to being found only once or twice in 500 cultures in a normal population (Smillie et al., 1938). However,

this was not the case in another pneumococcal pneumonia outbreak in two men's shelters in Paris from 1988 - 1989 with only one of the 57 serotypes isolated from carriers being serotype 1 (1.8%) (Mercat et al., 1991). The authors suggested that the low carriage rate could have been due to the timing of sampling as the nasopharyngeal carriage study was conducted at the end of the outbreak period.

1.10. Pathogenesis and virulence factors

Pathogenicity is defined as the ability of an organism or microbe to cause harm or disease in a host (Pirofski and Casadevall, 2012). Pathogenicity is thought to depend on possession by the microbe of certain virulence factors that mediate the disease outcome in the host. Virulence, on the other hand, is defined as the relative capacity of the microbe to cause disease or harm in the host. Whilst pathogenicity is often considered in terms of the presence or absence of this ability in the microbe, i.e. whether or not the microbe is pathogenic or not, virulence often refers to the extent or degree of damage or pathology caused by the microbe to the host. It has been suggested that both pathogenicity and virulence are inherent microbial properties that are usually expressed in the context of a susceptible host, thus highlighting the importance of the host-pathogen interaction in the expression of these properties by the microbe (Casadevall and Pirofski, 2001). This section discusses some important aspects of the pathogenicity and virulence factors of *S. pneumoniae.*

S. pneumoniae is spread from person to person by air-borne droplets and often resides asymptomatically in the upper respiratory tract of humans. From the nasopharynx, the pneumococcus may spread to contiguous organs such as the ears, sinuses or it may get carried down the trachea and bronchi to the lungs (Bogaert et al., 2004a). It is also capable of penetrating the mucosal barrier of the respiratory

tract to gain access to the bloodstream and thereby spreading to other more distant parts of the body such as the lining of the abdominal cavity, bones and their joint spaces or the cerebrospinal fluid by penetrating the blood brain barrier (Figure 1.4). When it gains access to these parts of the body that are normally sterile, it can result in invasive disease such as pneumonia, septicaemia, arthritis, osteomyelitis and meningitis.

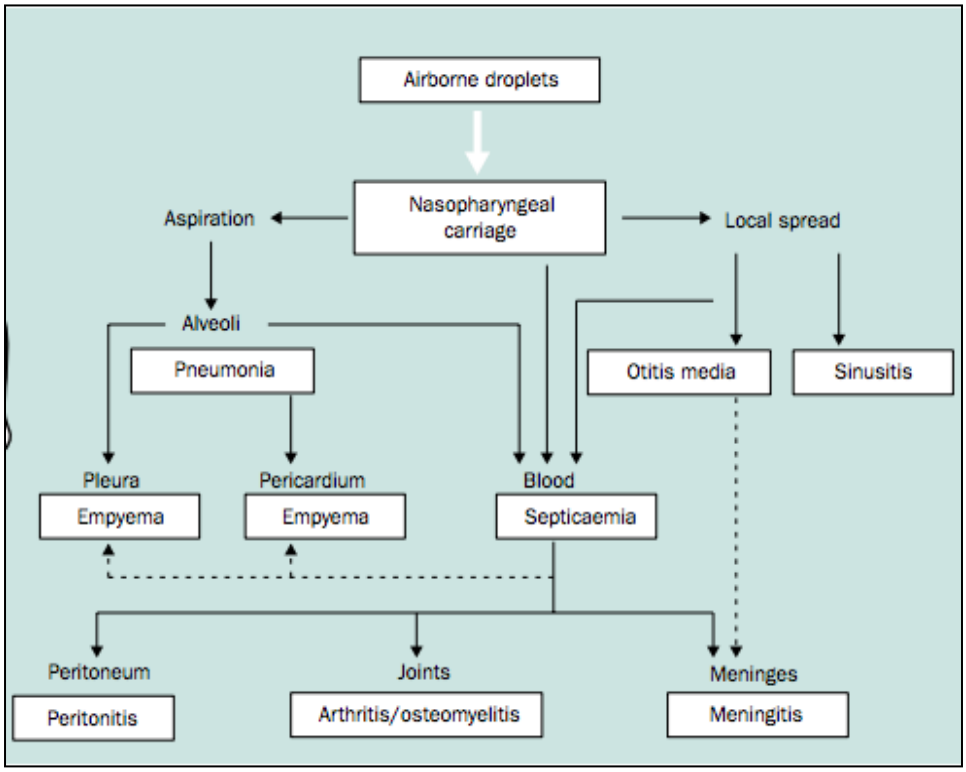

Figure 1.4 Pathogenic route for *S. pneumoniae* infections. Adapted from (Bogaert et al., 2004a).

In order to achieve this access, it possesses various attributes encoded by virulence factors that aid its establishment in host cells and facilitate its ability to cause disease. *S. pneumoniae* has several virulence factors described, the numbers of which appear to have increased in recent years with the aid of new technologies such as signature tagged mutagenesis (STM), transposon sequencing (Tn-seq), gene

expression assays such as microarray and deep RNA sequencing (Hava and Camilli, 2002, van Opijnen et al., 2009, Orihuela et al., 2004, van Opijnen and Camilli, 2012, Mann et al., 2012). The advent of whole genome sequencing and functional characterisation of these genes in animal models have also facilitated the discovery of further virulence genes. Some virulence genes are present in almost all pneumococcal isolates while others vary from one pneumococcus to the other suggesting that different strains of *S. pneumoniae* may vary in their virulence properties. This variation across pneumococcal strains is consistent with the fact that *S. pneumoniae* is highly transformable and can readily uptake exogenous DNA from closely related species. Several virulence genes are found on the pneumococcal cell surface and others in the cytoplasm, (Fig 1.5).

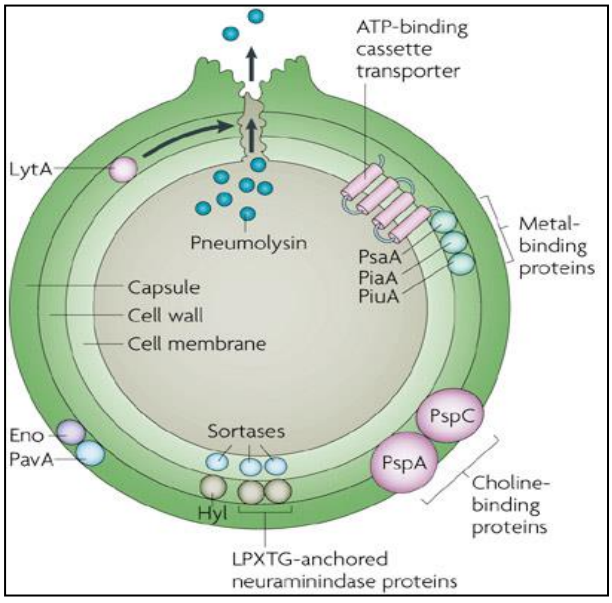

Figure 1.5 *S. pneumoniae* cell showing important virulence factors.

LytA, autolysin; Eno, enolase; PavA, pneumococcal adhesion and virulence A; Hyl, hyaluronate lyase; PspA and PspC, pneumococcal surface protein A and C; PiuA, pneumococcal iron uptake A; PiaA, pneumococcal iron acquisition A; PsaA, pneumococcal surface antigen. Adapted from (Kadioglu et al., 2008).

1.10.1. Capsule

Demonstration of the virulence properties of the pneumococcal polysaccharide capsule, a major virulence factor of *S. pneumoniae*, dates back to early experiments conducted by Fred Griffith, who observed that a non-encapsulated avirulent strain became virulent following co infection with a heat killed virulent encapsulated strain in mice (Griffith, 1928). The capsule contributes towards pathogenicity by disrupting several aspects of complement system and interferes with neutrophil mediated immunity leading to reduced opsonophagocytic capacity. Additionally, it is able to gain access to epithelial surfaces by avoiding entrapment in the nasal mucus (Hyams et al., 2010, Giebink et al., 1977). It has been reported that the capsule of different pneumococcal serotypes differs substantially in their ability to cause disease (Sandgren et al., 2004, Kelly et al., 1994) due to differences in their ability to evade host defence mechanisms. Melin *et al.* observed that serotypes 1 and 5, serotypes commonly associated with invasive disease, were particularly resistant to complement deposition and opsonophagocytic killing, whilst other serotypes such as 6B and 23F, associated with carriage, were more sensitive to deposition of C3 and opsonophagocytosis (Melin et al., 2010). Also, animal experiments have provided evidence of a strong association between capsular type and ability to kill mice (Briles et al., 1992).

S. pneumoniae has over 95 different capsular polysaccharides, all of which demonstrate a capsule locus showing similar organization with genes encoding specific capsular types flanked by genes common to all types and this locus has been published for 90 pneumococcal serotypes (Bentley et al., 2006). Amongst closely related serotypes based on *cps* genes, pneumococcal serotype 1 appears unique in its' lack of the rhamnose biosynthesis gene, the result of frame shift mutations within the

gene (Mavroidi et al., 2007). The peculiarity of the *cps* locus of pneumococcal serotype 1 is further demonstrated by the presence of an unusual sugar AAT-Galp, a component of its' pneumococcal teichoic and lipoteichoic acids. The only other organism reported to have this sugar within its capsule is *Bacteroides fragilis* (Aanensen et al., 2007). Most pneumococcal serotypes possess negatively charged polysaccharides but serotype 1 has both positive and negatively charged polysaccharides in a repeating unit structure referred to as a zwitterionic motif, which allows serotype 1 to function as a T-cell dependent antigen. In mice, this has been implicated in the formation of abscesses (Tzianabos et al., 1993, Velez et al., 2009).

1.10.2. Cell wall

The cell wall is known for its ability to induce an inflammatory response and the presence of teichoic acid facilitates this process (Tomasz and Saukkonen, 1989). It activates the alternative complement pathway (Winkelstein and Tomasz, 1978) and stimulates the production of platelet activating factor (Cabellos et al., 1992). It also facilitates attachment to the host endothelial cells resulting in loss of barrier integrity, effects that are mediated through the production of cytokines (Geelen et al., 1993). Experimental animals injected with purified cell wall or its products have been shown to exhibit signs that are similar to several pneumococcal diseases (Carlsen et al., 1992, Tuomanen et al., 1985).

1.10.3. Pneumolysin

Pneumolysin, one of the most studied pneumococcal proteins, is a highly conserved pneumococcal protein on account of which it has been proposed for use in the

formulation of protein vaccines (Briles et al., 2003, Kuo et al., 1995). It is a 53 kDa cytoplasmic protein and forms part of a larger group of proteins of pathogenic Gram positive bacteria known as cholesterol-dependent cytolysins. First cloned in 1983, the importance of this gene in virulence has been demonstrated in several animal infection models (Ogunniyi et al., 2007, Berry et al., 1989b). It has been shown to be cytotoxic and inhibits ciliary movements on respiratory epithelium (Feldman et al., 1990). It also activates the classical complement pathway (Mitchell et al., 1990). Different alleles have been observed amongst different pneumococcal serotypes due to mutations occurring within the gene. Clonal expansion of specific clones carrying *ply* alleles expressing reduced or no haemolytic activity has been observed particularly with serotype 1 ST306 (Kirkham et al., 2006). In addition to serotype 1, serotypes 7F and 8 both predominantly invasive disease-causing serotypes have been found to have these alleles and have been associated with disease outbreaks (Jefferies et al., 2007, Staples et al., 2015). Interestingly, it has been observed that the allelic variation of *ply* responsible for reduced haemolytic activity has no effect in reducing the susceptibility of *S. pneumoniae* to pneumolysin based vaccines (Harvey et al., 2011).

1.10.4. Pili

Pili, filamentous structures found on the surface of many Gram-positive bacteria, are encoded within pathogenicity islands which are a group of mobile genetic elements acquired by bacteria through horizontal gene transfer. Pili were first observed on the surface of pneumococci in 2006 (Barocchi et al., 2006). PI-1 is encoded by the RlrA accessory region and has been shown to influence colonization, virulence and the inflammatory response in mouse models. Strains containing PI-1 were more virulent,

exhibited enhanced adherence to lung epithelial cells and had a competitive advantage over strains mutant strains lacking this factor (Barocchi et al., 2006). A second type identified, PI-2, was also involved in adherence (Bagnoli et al., 2008). It was observed that PI-2 was associated with pneumococcal multilocus sequence type (MLST) belonging serotypes 1, 2, 7F, 19A and 19F. It was noted that strains belonging to clonal complex (CC) 271 were found to contain both PI-1 and P1-2 (Bagnoli et al., 2008).

1.10.5. *S. pneumoniae* surface proteins

The pneumococcal cell surface is made up of several proteins which contribute significantly to the virulence of the organism. They are marked by one of three motifs; a choline binding domain, a lipoprotein domain or the LPXTG cell wall anchor, (Bergmann and Hammerschmidt, 2006) (Figure 1.6). Some of the well-studied pneumococcal surface proteins thought to be potential pneumococcal vaccine candidates are discussed briefly below.

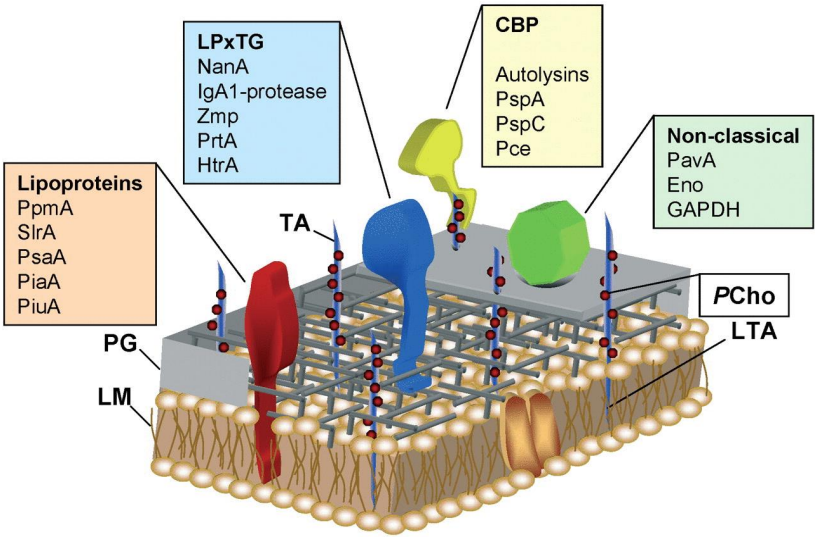

Figure 1.6 Different classes of *S. pneumoniae* surface-exposed proteins.

The cell wall is made up of lipoteichoic acid (LTA), a phospholipid membrane (LM), peptidoglycan (PG), teichoic acid (TA) and phosphoryl choline (PCho) which anchors choline binding proteins to the cell wall (CBP). Adapted from (Bergmann and Hammerschmidt, 2006).

1.10.5.1. Choline binding proteins (Cbps)

Choline binding proteins are anchored by phophorylcholine on teichoic and lipoteichoic acids on the cell wall (Hakenbeck et al., 2009). *S. pneumoniae* produces different Cbps including CbpA, PspA (pneumococcal surface protein A), PcpA (pneumococcal choline-binding protein A), cell wall lytic enzyme; N-acetylmuramoyl-L-alanine amidase (LytA; autolysin) and these are described briefly below. Other Cbps including CbpD, CbpE, CbpG, a b-N-acetylglucosamidase (LytB) and a b-N-acetylmuramidase (LytC; lysozyme) have been shown to play a role in nasopharyngeal colonization. CbpG has also been shown to play a role in sepsis (Gosink et al., 2000).

Choline binding protein A **(**CpbA), also known as PspC (pneumococcal surface protein C) plays a role in adherence by binding to cytokine activated pneumocytes and endothelial cells. This is thought to be an important process for pneumococci transiting from colonization to invasive disease (Rosenow et al., 1997). It has also been shown to contribute to sepsis. Mice infected with PspC mutants showed a significant increase in survival than the wild type (Iannelli et al., 2004). Its locus is highly polymorphic due to the high variability observed in the proline rich region. However, the DNA encoding for signal peptides is conserved across strains (Brooks-Walter et al., 1999).

Pneumococcal surface protein (PspA) is an important pneumococcal surface protein shown to be required for full virulence of pneumococci (Ren et al., 2003). It is a lactoferrin binding protein that varies in molecular size ranging from 67 kDa to 99

kDa. It inhibits complement deposition and activation and interferes with uptake into phagocytes. Variations within the gene sequence have given rise to different PspA types (Crain et al., 1990); even though conserved epitopes remain that allow vaccination within a single type to confer cross protection to other types (Ren et al., 2003).

Pneumococcal choline-binding protein A (PcpA) was identified as a virulence factor in a large scale signature tagged mutagenesis experiment (Hava and Camilli, 2002). The gene *pcp*A encodes a 79 kDa protein containing a C terminal choline binding domain (Sanchez-Beato et al., 1998). It is involved in adherence of *S. pneumoniae* to human nasopharyngeal and lung epithelial cells (Khan et al., 2012).

Autolysin (LytA) is a cell wall degrading enzyme that causes autolysis of the bacterial cell. Though normally inactive, it can be activated under conditions such as nutrient starvation or penicillin starvation. It releases degradation products such as peptidoglycan and teichoic acid, which can cause inflammation and release of bacterial proteins such as pneumolysin (Berry and Paton, 2000, Berry et al., 1989a).

1.10.5.2. Liporoteins

PsaA (pneumococcal surface adhesion A) is part of an ABC transporter operon in which PsaA is a substrate binding lipoprotein, PsaB, the ATP-binding protein and PsaC the permease likely involved in transporting manganese and zinc into the cytoplasm of pneumococcus. It is thought to function as an adhesin. Strains of the pneumococcus lacking this gene have been found to be avirulent in animal model tests (Sampson et al., 1994, Tseng et al., 2002).

PiuA (pneumococcal iron uptake A) and PiaA (pneumococcal iron acquisition A) are lipoprotein components of two iron ABC transporters (Brown et al., 2001). They have been found in all pneumococcal isolates tested and are required for virulence as demonstrated in murine infection models (Brown et al., 2001).

Putative proteinase maturation protein A (PpmA) and streptococcal lipoprotein rotamase A (SlrA) are members of the family of peptidyl-propyl isomerases and are thought to be involved in secretion and activation of cell surface molecules. PpmA has been shown to be important in pneumococcal virulence (Overweg et al., 2000) and SlrA in nasopharyngeal colonization (Hermans et al., 2006).

1.10.5.3. LPXTG-anchored surface proteins

These are proteins covalently anchored to the cell wall through a carboxyl-terminal motif, LPXTG and are recognized by a sortase enzyme. They include neuraminidases (NanA and NanB)**,** hyaluronate lyase and serine protease.

NanA and NanB, two enzymes implicated in neuraminidase activity have been shown to be essential for colonization and infection of the upper and lower respiratory tract. Neuraminidase facilitates the survival of *S. pneumoniae* in the respiratory tract and blood by enabling attachment to epithelial cells. (Manco et al., 2006).

Hyaluronate lyase (Hyl), a 107 kDa protein attached to the cell wall breaks down hyaluronic acid, an important component of the mammalian connective tissue (Berry et al., 1994). It has been thought to facilitate bacterial penetration through the host tissue (Berry et al., 1994) Mutations in *hyl* did not have an impact on virulence in a mouse intraperitioneal challenge model using a serotype 2 strain (D39). However a

double mutation of h*yl* and *ply* showed significant reduced virulence (Berry and Paton, 2000).

Serine protease (PrtA) was discovered during an immunological screening of convalescent patient sera (Zysk et al., 2000). The role of PrtA in pneumococcal virulence has been demonstrated in animal models, where survival in mice was significantly longer when infected with a PrtA negative mutant strain compared to the wild type (Bethe et al., 2001). Although the exact function of this protein is yet to be fully determined, it is known to be co-regulated with a number of virulence genes including *ply* (Bethe et al., 2001).

1.10.5.4. Other *S. pneumoniae* surface proteins

Surface proteins not having any of the three anchor motifs described above include pneumococcal histidine triad (Pht) proteins, pneumococcal adherence and virulence factor A, enolase and pneumococcal adhesion and virulence A (PavA).

The Pht family proteins, characteristic of a conserved motif HxxHHxH, were discovered while utilizing whole genome sequences of pneumococci to identify proteins which protect against pneumococcal infections (Wizemann et al., 2001). This family consists of four proteins (PhtA, PhtB, PhtD, PhtE) and function by inhibiting complement deposition via the complement regulator, factor H (Ogunniyi et al., 2009). PhtD is one of the pneumococcal proteins currently included in protein vaccines undergoing clinical trials (Leroux-Roels et al., 2014).

Enolase is a plasminogen receptor which mediates pneumococcal attachment to human epithelial and endothelial cell (Bergmann et al., 2013). PavA is an important pneumococcal surface protein involved in virulence and adhesion (Holmes et al.,

2001, Pracht et al., 2005). Knock out mutants of this gene were less attenuated in mouse sepsis and meningitis models (Holmes et al., 2001, Pracht et al., 2005).

1.11. Models utilized in studying the pathogenesis of *S. pneumoniae*

1.11.1. Experimental human pneumococcal carriage

Humans are the natural host of *S. pneumoniae* and thus the development of an experimental human pneumococcal carriage model offers the potential for more accurate descriptions of mechanisms of protection against pneumococcal carriage. This human model experiment involves intranasal inoculation of a known quantity of *S. pneumoniae* to healthy adult volunteers (Gritzfeld et al., 2013). This model has been used in evaluating the suitability of candidate proteins for pneumococcal protein vaccines (McCool et al., 2002). It has also been used in determining appropriate methods for pneumococcal carriage detection, which will be useful in epidemiological and vaccine efficacy studies (Gritzfeld et al., 2013, Gladstone et al., 2015a). Limitations of this model include ethical issues related to safety and logistical challenges with selection of volunteers, inoculums preparation and completeness of follow up.

1.11.2. Animal models

Due to the limitations associated with experimental human models, the use of animal models in studying the role of *S. pneumoniae* in nasopharyngeal carriage and invasive disease. These models are discussed briefly below.

1.11.2.1. Mammalian models

Some mammalian models that have been used in pneumococcal research include rats, mice, rabbits, chinchilla, gerbil and guinea pigs with the latter three animal models used mostly in the study of middle ear infections (Chiavolini et al., 2008). However, for most other studies of bacterial pathogens, mice and rat models tend to be used more frequently, possibly due to their advantage of having tightly controlled immune system. Compared to rat models, mice are cheaper, easier to handle and better suited for screening of drugs and vaccines, and are therefore used more frequently than rat models. Additionally, the availability of the mouse whole genome sequence and technologies to induce mutations in the mouse genome makes it a more attractive model to study host pathogen interactions. Outbred mice are used to mimic the natural variation in response to infection seen in humans. Routes of administration may vary based on the intended experiments and may include, intranasal, intravenous, intraperitoneal and intra cerebral routes. Also, animal models have been developed for different invasive disease conditions including pneumonia, meningitis, sepsis, otitis media as well as carriage models (Chiavolini et al., 2008). Variations between mice in their susceptibility or resistance to pneumococcal infection have been shown highlighting the role of the genetic background of hosts in pneumococcal infections (Preston et al., 2004, Gingles et al., 2001).

Animal models have been widely applied in pneumococcal research. Some of these applications include virulence testing of phenotypic pneumococcal strains (Williams et al., 2012), testing novel drugs and efficacy of antimicrobials (Sullivan et al., 1993), pneumococcal vaccine assessments (Khan and Pichichero, 2013) and evaluating mechanisms of invasive pneumococcal disease pathogenesis (Ibrahim et al., 2004). Despite the usefulness of mammalian models, their use is associated with

a number of limitations including ethical challenges associated with use of mammals in research, the huge costs involved and the specialized training required to conduct animal experiments.

1.11.2.2. Galleria mellonella infection model

An alternative to mammalian models of infection is the use of invertebrate hosts, such as nematodes or insects. Nematodes or insects that could be used include *Caenorhabditis elegans*, *Drosophila melanogaster* and the larvae of the Greater wax moth, *Galleria mellonella* (Tan et al., 1999, Jander et al., 2000). Insects, like mammals, possess a complex innate immune system. Cells within the hemolymph are able to phagocytose or encapsulate invading microbes. They are also able to produce lysozyme and small antibacterial peptides.

G. mellonella has been used extensively as an infection model in the study of bacterial infections including *S. pneumoniae Proteus*, *Pseudomonas aeruginosa*, *Burkholderia cepacia*, *Acinetobacter*, Group A *Streptococcus*, *Staphylococcus aureus*, *Campylobacter*, as well as fungal infections such as *Cryptococcus*, *Aspergillus* and *Candida* (Evans and Rozen, 2012, Jander et al., 2000, Gao et al., 2010, Seed and Dennis, 2008). Use of a *G. mellonella* model offers some advantages over other insect models. The *G. mellonella* is relatively large, about 2 cm long weighing approximately 250 mg and thus they are easier to handle and for injecting clearly defined doses of bacteria. The *G. mellonella* model also allows for infection studies to be conducted at temperatures between 15°C and 37°C, hence thereby mimicking a mammalian environment unlike other insects such as the drosophila whose maximum temperature only allows for experimentation at 25°C (Rejasse et al., 2012). In comparison to mammalian models, *G. mellonella* models are cheaper,

do not require specialized facilities for maintenance, are easier to use, do not require licences and are more ethically acceptable. They can be used as an initial tool in the screening of large numbers of strains thus reducing the number of strains needed for subsequent use in mouse models. Additionally, studies have shown good correlation between mice and *G. mellonella* virulence studies (Jander et al., 2000). Cells are injected via the first right pro-leg into the haemocoel and recent studies are exploring the mouth route. Evans *et al* showed for the first time that the *G. mellonella* model is a good tool for *S. pneumoniae* research. They demonstrated, for example, that a pneumolysin mutant was less virulent in the *G. mellonella* model than the wild type as demonstrated in several mice experiments (Evans and Rozen, 2012).

1.12. Laboratory tools for pneumococcal characterization.

1.12.1. Culture based identification of *S. pneumoniae*

A range of specimen types is collected for identification of *S. pneumoniae*. These include nasopharyngeal swabs (NPS), sputum and specimens collected from normally sterile sites of the body such as blood, cerebrospinal fluid, joint fluid, pleural and lung aspirate.

Laboratory processing of these specimens includes microscopy, standard culture, and testing for antibiotic sensitivity. Microscopic examination of specimens requires carrying out Gram stain. A Gram-positive result reveals lancet-shaped diplococcic and is indicative of *S. pneumoniae*. Specimens are cultured on selective defibrinated blood agar, mostly sheep or horse blood, supplemented with 5 µg/ml of gentamycin to prevent the growth of non-pneumococcal isolates. Alpha haemolytic colonies are confirmed as *S. pneumoniae* using optochin and bile solubility tests.

The optochin test is used in the presumptive identification of alpha-haemolytic streptococci. *S. pneumoniae* cells are sensitive to optochin (ethylhydrocupreine hydrochloride) though some strains are optochin resistant. Optochin sensitive *S. pneumoniae* surrounding a disk impregnated with optochin are lysed due to changes in surface tension, thus creating a clear zone of inhibition.

The bile solubility test is used to differentiating *S. pneumoniae* from other alpha haemolytic Streptococci. Bile salts lower the tension between the bacterial cell and the medium thus enhancing the pneumococcus' autolytic process resulting in the lysis of the cell wall. This effect is observed by clearing of the initially turbid bile solution.

Pneumococcal antimicrobial susceptibility testing is often interpreted using standard interpretation guidelines such as the Clinical and Laboratory Institute (CLSI) guidelines.

Limitations of culture based laboratory diagnosis of pneumococcus, particularly in resource- constrained settings such as in sub Saharan Africa, include non-availability of the test at the point of care, lack of appropriate specimen collection equipment, attendant challenges with specimen contamination and storage, lengthy waiting time for culture results and prior antibiotics use which reduces pneumococcal detection rates in samples collected. This has led to the development of molecular based methods to complement standard culture techniques.

1.12.2. *S. pneumoniae* serotyping

S. pneumoniae has a polysaccharide capsule, the basis for which it has been classified into more than 95 different types, also known as serotypes (Henrichsen,

1995, Oliver et al., 2013). These serotypes all play different roles with some serotypes more frequently associated with causing invasive disease and others more often seen in carriage (Sandgren et al., 2004). Given the significant contribution of this pathogen to serious diseases worldwide and with the advent of serotype-specific pneumococcal vaccines, pneumococcal serotyping is therefore required for effective monitoring of vaccine impact in communities and thus informs policy makers on priorities for the development of appropriate new vaccine formulations. It is also helpful for determining current and emerging pneumococcal serotypes thus facilitating our understanding of any regional or global changes in pneumococcal epidemiology. Additionally, pneumococcal serotyping is useful in the detection of multiple serotypes (Turner et al., 2011), shown to be common in nasopharyngeal carriage, even though the ability and extent to which multiple serotypes are detected will depend on the methods employed. Pneumococcal serotyping is performed either by phenotypic or genotypic methods. Phenotypic methods are based on biochemical assays and a few are described briefly below.

1.12.2.1. Quellung Reaction

This is the gold standard technique for pneumococcal serotyping and is widely used by different reference laboratories. New methods of pneumococcal serotyping that are developed need to be standardized against the Quellung reaction, which has both high sensitivity and specificity. The Quellung reaction was first described in 1902 by German bacteriologist, Ferdinand Neufeld (Austrian, 1981). The test is a biochemical reaction that involves binding of polysaccharide antigens of the pneumococcal capsule to antibodies in the serotype specific antisera. This results in the formation of a complex molecule with the

pneumococcal capsule appearing swollen when viewed under a microscope as a result of increased surface tension (Habib et al., 2014). Limitations of this method include the high cost of antiserum and the need for trained personnel for accurate interpretation of results, in comparison to the latex agglutination method. The test is also time consuming.

1.12.2.2. Latex agglutination

Agglutination tests were first discovered by Herbert Edward Durham and Max Von Gruber in 1896, whilst working with the pneumococcus, with the reaction called the Gruber- Durham reaction at the time. This principle has subsequently been applied to the detection of numerous other organisms.

In this method, specific antibodies are attached to latex particles, which when in contact with a specific pneumococcal capsular antigen results in a visible agglutination reaction and a clearing of the background suspension. Latex agglutination serotyping has been shown to have good agreement with Quellung reaction (Lalitha et al., 1996). Also, it is fast and relatively easy to perform, does not require specialized equipment and is an inexpensive test. It is therefore a suitable method of choice for resource poor countries, and high throughput laboratories. Commercially prepared latex agglutination kits are available but often expensive. On the other hand, in-house preparations of latex reagents have also been used but often require extensive laboratory quality control procedures to prevent misinterpretation of results. In The Gambia, a locally prepared antibody coated latex agglutination test is used on colonies from primary culture plates suspended in saline (Adegbola et al., 2006). The WHO recommended method for latex agglutination involves selecting unique or morphologically different strains of *S. pneumoniae* and serotyping these

strains (Satzke et al., 2013). However, to increase detection of multiple serotypes, the latex sweep agglutination method is used. This method involves serotyping a plate sweep of pneumococcal colonies. Various studies have demonstrated the suitability of this method particularly in large studies for improved detection of multiple serotypes and it was recently supported by the PneuCarriage study which is the largest study evaluating different pneumococcal serotyping methods (Turner et al., 2011, Satzke et al., 2015).

1.12.2.3. Other phenotypic assays

Other phenotypic assays for pneumococcal serotyping that have been developed include enzyme-linked immunosorbent assays (ELISA) and an enzyme linked immunoassay (EIA) (Lankinen et al., 2004); immune blot assays including dot blot and colony blot (Bogaert et al., 2004b) and bead based assays on flow cytometry or luminex platforms (Sheppard et al., 2011). However, these methods are generally expensive, time-consuming, limited in the number of serotypes detected or lacking specificity through cross reactions or have a combination of these factors.

1.12.3. Molecular typing of *S. pneumoniae*

The use of molecular tools in the study of *S. pneumoniae* has significantly improved our understanding of the behaviour of this pathogen. Numerous molecular methods have been employed in typing *S. pneumoniae*; a few are briefly described below.

1.12.3.1. PCR based serotyping

Production of the pneumococcal capsule is controlled by capsular polysaccharide genes located at the *cps* locus flanked by conserved *dex*B and *ali*A genes and

sequence data for the region is available for over 90 serotypes of *S. pneumoniae* (Bentley et al., 2006). The differences within the *cps* locus form the basis for differentiation of pneumococcal serotypes by most of the genotypic methods.

The first multiplex PCR (mPCR) assay used in molecular serotyping was a seven step sequential assay capable of detecting 33 serotypes/ serogroups was developed in 2006 from invasive pneumococcal disease (Pai et al., 2006). This assay was tested in The Gambia and its utility amongst carriage isolates and in the detection of multiple serotypes was well demonstrated (Antonio et al., 2009). This method utilizes PCR primer pools in a sequential manner and visualized by agarose gel electrophoresis based on amplicon sizes. It has since been modified to suit different geographical areas including Latin America (Dias et al., 2007), Europe (Iraurgui et al., 2010) and Africa (Morais et al., 2007). It has also been modified to use on a range of clinical specimens (Njanpop Lafourcade et al., 2010). One of the limitations of this method is cross reactivity within some serotypes as documented by Pai et al. For example, between serotype 6A and 6B which is because they share identical sequences within the cpsA loci differing in only one single nucleotide polymorphism in the *wciP* gene hence making it difficult to develop type specific primers. Other examples include serotypes 7F and 7C cross reacting with serotypes 7A and 7B as well as 9V, 38, 11A, 12F, 22F, 33A, 33F, 35F and 18C (Pai et al., 2006). Identifying these individual serotypes would require further testing by serological methods of serotyping.

Real time PCR (RT-PCR) was developed following the advent of mPCR. RT-PCR is relatively faster and more sensitive than mPCR and offers more specificity as it requires hybridization to a probe in addition to amplification primers (Azzari et al.,

2010, Pimenta et al., 2013). Its limitation is that it is more expensive compared to mPCR and also has limited scope in multiplexing.

Other methods include a PCR-Restriction Fragment Length Polymorphism (PCR-RFLP) assay in which digested PCR amplicon fragments are analysed using gel electrophoresis and serotype is assigned a unique RFLP pattern (Batt et al., 2005),combining PCR and electrospray ionization mass spectrometry (PCR/ESI-MS) (Massire et al., 2012), PCR fragment analysis and automated fluorescent capillary electrophoresis (FAF-mPCR) (Selva et al., 2012), PCR and reverse line blot hybridization (mPCR/RLB) (O'Sullivan et al., 2011), a high throughput method using nanofluidic real time PCR (Dhoubhadel et al., 2014) and sequence-typing which utilizes a single primer pair to serotype pneumococcal isolates (Leung et al., 2012).

1.12.3.2. Molecular serotyping by microarray

Microarrays are ordered sets of DNA molecules of known sequence fixed on a physical support. This technology allows for the simultaneous detection of thousands of genes in a small sample through specific binding to an array of high density probes. A microarray for molecular serotyping of *S. pneumoniae* with the ability to detect over 90 serotypes has been developed by the Bacterial Microarray Group (BμGS) at St George's University of London. The BμGS SP-CPSv1.1.0 microarray is a custom designed microarray on the Agilent SurePrint platform and comprised primarily of 60mer oligonucleotides probes (Newton et al., 2011). It was designed with multiple reporters representing all 432 known capsular polysaccharide synthesis genes (Bentley et al., 2006). It also has additional probes incorporated to distinguish between serotypes with similar sets of *cps* genes for example 6A vs 6B;

7A vs 7F; 9A vs 9V; 11A vs 11D; 12F vs 44 etc. In addition, it has incorporated probes for the entire genome of two fully sequenced pneumococcal strains, TiGr4 and R6. It has probes for antibiotic resistant genes alpA3, cat, ermB, ermC, mefA, sat4, tetK, tetL, tetM and tetO. It also has reporter elements for detecting other pathogens often found co-colonizing the nasopharynx namely *Streptococcus agalactiae*, *Streptococcus equi*, *Streptococcus pyogenes*, *Staphylococcus aureus*, *Haemophilus influenza*, *Neisseria meningitides*, *Moraxella catarrhalis*, *Bordetella pertussis* and *Corynebacterium diphtheriae*. This technique involves several steps of sample processing and labelling, hybridization, washing, image acquisition and data analysis (Newton et al., 2011). Data analysis is automated using empirical Bayesian statistical models (Newton et al., 2011).

In addition to comprehensive molecular serotyping for 93 known pneumococcal serotypes, this method is able to detect multiple serotype carriage and the ability to determine the relative abundance of each serotype to as little as 1% (Turner et al., 2011). This makes it a very attractive method when compared to all other methods of serotyping. It also has the potential to detect novel serotypes, co-carriage of *S. pneumoniae* with other pathogens and antibiotic-resistance genes. The microarray method has been utilized in several studies (Turner et al., 2011, Kandasamy et al., 2015). Limitations of the microarray method include the need for technical expertise particularly for interpretation of any unusual findings, the relatively high cost of equipment and reagents thus limiting its availability to a small number of research laboratories.

1.12.3.3. Pulsed-field gel electrophoresis (PFGE)

PFGE has been used extensively for investigating outbreaks and on-going disease surveillance due to its high epidemiological concordance and high discriminatory power. It is also relatively inexpensive and readily reproducible. This process involves cleaving genomic DNA with restriction endonucleases resulting in restriction fragments, which are separated on an agarose gel by 'pulsed-field' electrophoresis. This is visualised on the gel as bands to give a PFGE pattern (Lefevre et al., 1993). However, this method is not without limitations as it is time consuming, labour intensive and technically demanding. It could also be prone to significant subjectivity in interpretation of findings thus requiring continuous quality control to minimize this risk.

1.12.3.4. Multiple-locus variable-number tandem repeat analysis (MLVA)

MLVA is a method which targets short tandem repeats across the bacterial genome, has a scheme developed for *S. pneumoniae* targeting 17 distinct loci with a dedicated web-based database (Koeck et al., 2005). It has been applied in several epidemiological studies (Koeck et al., 2005, Yaro et al., 2006) but is not widely used as it is expensive and time consuming.

1.12.3.5. Multilocus enzyme electrophoresis (MLEE)

MLEE, another method used in epidemiological typing of different organisms, is based on the relative mobility under electrophoresis of large numbers of intracellular enzymes with unique mobility patterns associated with each strain (Selander et al., 1986). It has a lower resolution than multilocus sequence typing (MLST), described below, and is therefore not as widely used. Despite the usefulness of these tools,

significant gaps remain in their ability to provide global epidemiological understanding of bacteria due in part to non-standardized approaches as results obtained were frequently not comparable between laboratories (Maiden, 2006, Maiden et al., 1998).

1.12.3.6. Multilocus sequence typing (MLST)

Multilocus sequence typing, proposed in 1998 (Maiden et al., 1998) as a genetic-based method that could improve typing, was also applied in the study of population biology and in evolution studies. It involves sequencing of approximately 450-500 bp internal fragments of several housekeeping genes typically using capillary Sanger sequencing. Each of the sequences is trimmed, assembled and submitted to the MLST database and assigned an allele number. A combination of seven allele numbers gives rise to an allelic profile called a sequence type (ST). The MLST database contains reference allele sequences, sequence types as well as epidemiological data for all strains submitted. The MLST scheme for *S. pneumoniae* uses internal fragments of seven housekeeping genes namely, *aroE* (shikimate dehydrogenase), *gdh* (glucose-6-phosphate dehydrogenase), *gki* (glucose kinase), *recP* (transketolase), *spi* (signal peptidase I), *xpt* (xanthine phosphoribosyltransferase), *ddl* (D-alanine-D-alanine ligase).

The advantage of MLST over other genotyping methods lies in its reproducibility and portability. Materials can be exchanged between laboratories and it is accessible electronically, making it appropriate for use in evolutionary and epidemiological studies (Enright and Spratt, 1998, Maiden, 2006). The application of MLST has significantly improved our understanding of *S. pneumoniae*, as it has been used to characterize invasive and carriage strains (Antonio et al., 2008, Enright and Spratt,

1998, Brueggemann et al., 2003), track epidemic outbreaks (Antonio et al., 2008, Leimkugel et al., 2005), antimicrobial resistant strains (Pletz et al., 2004, Siira et al., 2009) and to understand transmission dynamics (Hill et al., 2010).

Despite its usefulness, MLST lacks the ability to adequately differentiate between bacterial strains since it is limited to variations in housekeeping genes. A more complete understanding of the role of *S. pneumoniae* in causing diseases would require an evaluation of other components of the bacterium such as virulence genes. With the advent of second-generation sequencing technologies, MLST can now be assigned from whole-genome sequence information, rather than sequencing each locus separately. This provides a more cost effective approach and MLST data can be used to correlate with other genomic features.

1.12.3.7. Whole genome sequencing

Whole genome sequencing (WGS) makes available the complete DNA sequence of bacterial cells.

Sanger sequencing, also known as the chain termination or dideoxy method developed by Fredrick Sanger in the 1970s (Sanger et al., 1977) was one of the earlier methods of WGS. Improvements on the Sanger sequencing method led to the automated Sanger sequencing technique on the ABI platform by applied Biosystems. Sanger sequencing is performed with components required for DNA replication, a single stranded DNA template, a primer, polymerase deoxynucleoside triphosphates (dNTPs) and in addition, special dideoxynucleoside triphosphates (ddNTPs). DNA elongation which occurs by the incorporation of regular dNTPs by the DNA polymerase is terminated once ddNTPs is incorporated. This is due to the lack of a 3′-OH group needed for the formation of phosphodiester bonds between two

nucleotides (figure 1.7). The ddNTPs are labelled with dyes that allow for their detection following capillary gel electrophoresis.

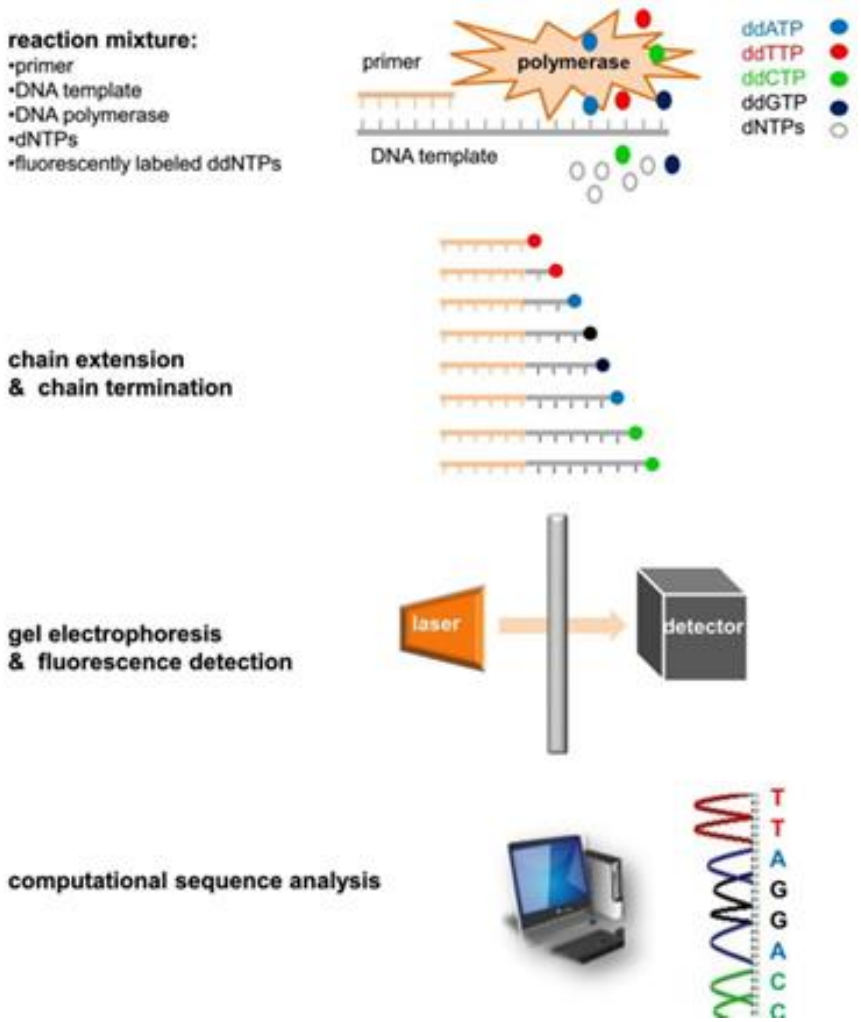

Figure 1.7 Automated Sanger sequencing workflow

The advantages of Sanger sequencing techniques over the new methods of sequencing are the production of high quality sequence data and long reads of about 900bp.

Since the inception of Sanger sequencing, several other sequencing technologies have been developed. These newer methods have a higher throughput and eliminate

electrophoresis as well as bacterial cloning. The disadvantage for these newer methods is that relatively short reads are produced. Some of these newer technologies include pyrosequencing (454) by Roche Inc., Branford, CT, USA; Sequencing by Oligo Ligation Detection (SOLiD) by Life technologies Corporation, Grand Island, NY, USA); MySeq and HiSeq by Illumina Inc., San Diego, CA, USA, The Ion Torrent Personal Genome Machine (PGM^{TM}) by Life Technologies Corporation, Grand Island, NY, USA; the PacBio RS II Single Molecule Real-time Sequencing (SMRT) by Pacific Biosciences Inc., Menlo Park, CA 94025 (Metzker, 2010) and MinION by Oxford Nanopore Technologies Ltd, Oxford, UK. Characteristics of these methods are summarized in table 1.1.

Table 1.1 Summary of sequencing technologies

	Automated Sanger Sequencing	454	Illumina	SoLiD	Ion Torrent	PacBio	Nanopore
Company	Applied Biosystems Inc.	Roche diagnostics	Illumina Inc	Life technologies	Life technologies	Pacific biosciences	Oxford nanopore technologies
Year		2005	2006	2007	2010	2011	
Library preparation	Yes	Yes	Yes	Yes	No	No	Minimal
Amplification method		Emulsion PCR	Bridge amplification	Emulsion PCR	Emulsion PCR	No amplification	No amplification
Sequencing technology	Chain termination	Sequencing by synthesis - pyrosequencing	Sequencing by synthesis - reversible dye termination	Sequencing by ligation – oligonucleotide probe ligation	Sequencing by synthesis – ion semi-conductor sequencing	Sequencing by synthesis - single molecule real time sequencing	Sequencing by synthesis - single molecule sequencing
Detection method	Fluorescent emission from dye terminator nucleotides	Light emission from secondary reactions initiated by pyrophosphate release	Fluorescent emission from incorporated dye-labelled nucleotides	Fluorescent emission from ligated dye-labelled oligonucleotides	Proton detection		Nanopore-based detection of single molecules
Read lengths/run	~ 900 bases	~700 bases	~250 bases on Miseq and ~125 bases on Hiseq	50 – 75 bases	~200 bases		
Maximum output/run		0.7 gb	15gb on Miseq and 1500 gb on Hiseq	120gb	100gb	1gb	1gb
Number of reads/run	1	1×10^6	3×10^8 on Miseq and 5×10^9 on Hiseq	1×10^9	6×10^7	1×10^6	6×10^4

Time/run	2h	24-48h	27h on Miseq and 240h on Hiseq	14 days	2 – 5h	1-2h	48-72h
Raw error rate	0.3	1	0.8	0.01	1	12.9	34
Advantages	Long reads and high quality sequencing	Longest reads, greater accuracy and precision compared to other NGS methods	Widest range of platforms ranging from Miseq to Hiseq X ten useful for population based studies. Very high throughput	High throughput and accuracy	Reduced turnaround time, flexible workflow and cheaper price compared to other NGS methods	Longer reads and faster run times compared to most NGS methods.	Portable devices. Easy sample preparation, no amplification step required prior to sequencing.
Limitations	Expensive, time consuming for large project and lower throughput compared to NGS methods	High cost of reagents, complex sample preparation and high error rates in homopolymers repeats. Discontinuation of support from Roche in 2016.		Complex sample preparation, very long run times and short read lengths. Requires huge computational infrastructure and expert support to analyse raw data.	Difficulty in reading homopolymer repeats	High error rates, lower throughput compared to NGS methods and high cost per base	High error rate. Requirement to optimize the speed of DNA translocation through the nanopore to ensure measurement of the ionic current change.

Of these methods, Illumina sequencing is currently the most widely used sequencing platform with a wide variety of sequencing instruments ranging from the lower throughput equipment such as the Miseq to ultra-high throughput equipment such as the Hiseq X10 suitable for population wide studies. The different Illumina sequencing equipment provide more options for cost, runtime and read length. Illumina sequencing workflow involves library preparation in which DNA is fragmented and ligated to adaptors at the 5' and 3' end of the DNA fragment. The ligated DNA fragments are loaded into a flow cell surface which has complementary oligonucleotides to the adaptors ligated to the DNA fragments. Each fragment undergoes bridge amplification to form distinct clonal clusters. This process is known as cluster generation. Sequencing by synthesis technology is utilized making use of a reversible dye termination method. In this method, four reversible terminator bound and fluorescently labelled dNTPs are present during each cycle of sequencing. Once incorporated, the DNA elongation is terminated and fluorescence captured (Figure 1.8)

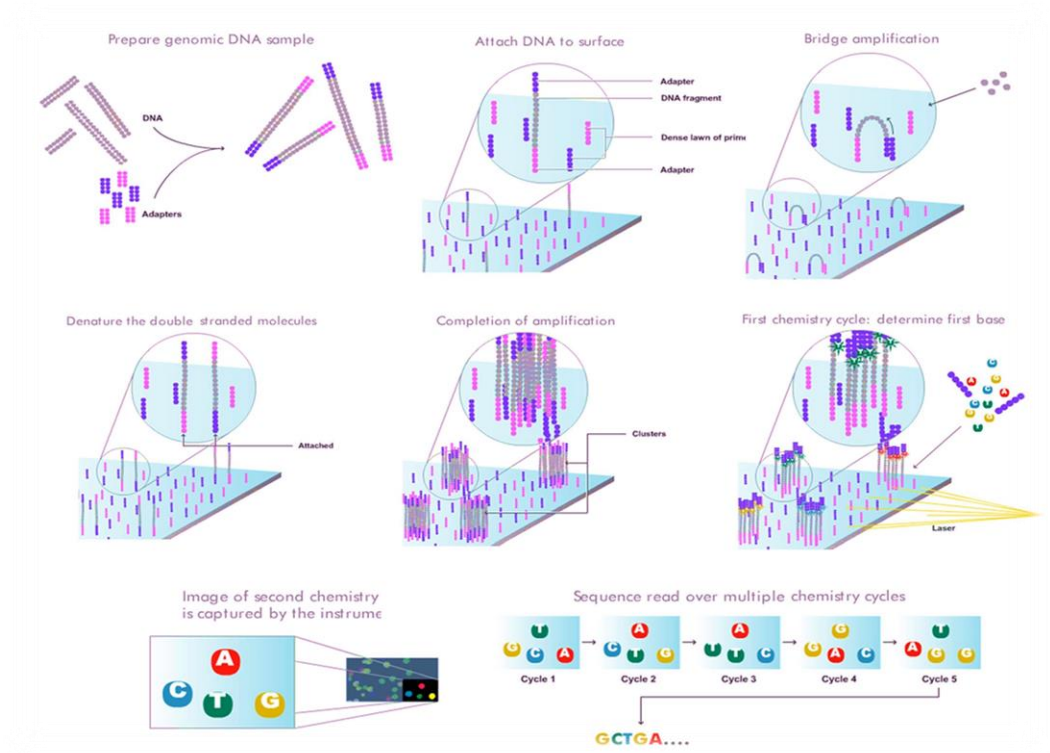

Figure 1.8 Illumina sequencing work flow

These methods produce large data sets, which need to be assembled into continuous DNA sequences (contigs), with several genome assemblers being put use. In addition, several computer programs are now readily available for processing sequence data.

This availability of whole genome sequence of different strains of *S. pneumoniae* in addition to mutagenesis and functional characterization of knock out mutants using various animal models and assays is fast revolutionizing the understanding of pathogens and their interactions with host cells.

For pneumococcal serotyping, it is hoped that next generation sequencing will fill the gaps experienced in the application of existing serotyping methods such as challenges with long turnaround time and limitations with multiplexing for multiple serotype detection. Whole genome sequencing offers an advantage over existing

serotyping methods as serotypes of pneumococcal isolates have been deduced by mapping raw sequence reads against reference capsular locus sequencing (Everett et al., 2012) and *in-silico* methods have been used to differentiate non-typeable *S. pneumoniae* (Chewapreecha et al., 2014a). Another advantage of this method is the potential it holds for discovering novel pneumococcal serotypes and detecting capsular switching. However, a limitation of this method is its' requirement for computational capacities, complex bioinformatics pipelines and expertise and longer preparation times. A modification of this method, based on target enrichment coupled with tagged sample pooling for the identification and serotyping of *S. pneumoniae* by using next generation sequencing (Liyanapathirana et al., 2014) offers some improvement in terms of cost. Nevertheless, this modification still requires significant bioinformatics expertise and has reduced specificity of primers in identifying closely related serotypes.

Other applications of WGS include characterizing multiple antibiotic resistant pandemic strains (Fani et al., 2011, Croucher et al., 2009), understanding the role of virulence factors (Paton and Giammarinaro, 2001) and in reverse vaccinology to identify vaccine candidates (Wizemann et al., 2001, Paton and Giammarinaro, 2001, Di Guilmi and Dessen, 2002). However, up to 30% of predicted coding sequences encode proteins of unknown function and are often referred to as hypothetical proteins. Also, little is known regarding the genetic attributes of invasive strains such as pneumococcal serotype 1 in causing disease. Continued whole genome sequencing and functional characterization of genes across strains from different global locations will be required to fully understand the genetic basis of *S. pneumoniae* pathogenesis, which will inform development of appropriate and improved preventative and therapeutic strategies.

1.13. Treatment and Control of *S. pneumoniae* infection

1.13.1. Antibiotics

Antibiotic therapy forms the mainstay of treatment of pneumococcal disease. Even though variations in antibiotic treatment offered across different clinical settings exist, the WHO recommends the use of penicillin as first line antibiotic treatment (WHO, 2013). Recommendations for antibiotic treatment options in areas with high levels of proven penicillin resistance include the use of 3rd generation cephalosporins such as ceftriaxone and cefotaxime (Bradley et al., 2011).

A major concern with the use of antibiotics is the development of antimicrobial resistance. This is particularly worrisome with the pneumococcus, a naturally transforming organism, which has the ability to transfer and spread resistant genes between pneumococcal strains (Croucher et al., 2011). The pneumococcus becomes resistant to antibiotics by altering itself to block the action of the antibiotic. Antibiotics work against bacteria by either interfering with cell wall synthesis, nucleic acid synthesis or with protein synthesis. Beta lactam antibiotics, such as penicillins known for their characteristics beta lactam ring in their chemical structure interferes with cross linking of peptidoglycan of the pneumococcal cell wall by binding to transpeptidase enzyme thus blocking its action and inhibiting cell wall synthesis. The pneumococcus develops resistance to penicillin by alterations to penicillin binding proteins through transformation thereby inhibiting the action of penicillin. Macrolides, tetracyclines and chloramphenicol interfere with microbial protein synthesis by binding to either the 50s or 30s ribosomal subunits thereby inhibiting the synthesis of the polypeptide chain. Pneumococcus develops resistance to these drugs by methylating the ribosomal subunits via genes usually found on

plasmids such as the *erm* genes. By doing this, the pneumococcus is able to prevent the antibiotic from binding to the ribosomal subunit. In addition, pneumococci develop resistance to antibiotics via efflux pumps encoded for by transposable elements. This way, the pneumococcus is able to pump out antibiotics once it has gained entry into the cell.

Factors contributing to increases in antibiotic resistance include widespread abuse of antibiotics and clonal spread of antimicrobial resistant strains. In Iceland for example, increase in antibiotics resistance was observed over several years. Initially, expansion of an international multidrug resistant clone Spain^{6B}-2 was observed and this declined after a public health intervention. However, in subsequent years, an expansion of a different international multidrug resistant clone Taiwan^{19F}-14 was observed to cause higher antibiotic resistance rates than had been described previously (Hjalmarsdottir and Kristinsson, 2014). A study conducted in four African countries, namely Senegal, Ivory Coast, Tunisia and Morocco reported high resistance to penicillin (mostly intermediate), erythromycin, tetracycline and cotrimoxazole (Benbachir et al., 2001). In Nigeria, Iroezindu *et al* showed that there was high prevalence of resistance to commonly prescribed drugs including penicillin, cotrimoxazole and ceftriaxone (Iroezindu et al., 2014) while in Ghana, intermediate resistance to penicillin was reported (Dayie et al., 2013). In The Gambia, resistance to tetracycline, cotrimoxazole chloramphenicol and penicillin have been documented (Adegbola et al., 2006).

Implications of antimicrobial resistance among pneumococci, particularly for developing countries where the burden of disease is highest, include higher costs of treatment, as the alternative drugs to which the pneumococcus is sensitive are frequently more expensive, and non–availability of such drugs, both factors of which

contribute to increase poor treatment outcomes and deaths from IPD in these settings. In addition, with very high pneumococcal carriage rates in developing countries, the likelihood of spread of these antimicrobial resistant genes is high and would further worsen the existing challenges with effective control of pneumococcal disease.

Interestingly, pneumococcal serotype 1 has been shown to be susceptible to most of the antibiotics tested with the exception of cotrimoxazole (Konradsen and Kaltoft, 2002, Porat et al., 2001). The Pneumococcal Molecular Epidemiology Network (PMEN), established in 1997 to enhance global surveillance of antibiotics resistant strains (McGee et al., 2001) currently has 43 clones in the network, four of which are serotype 1 strains. All four serotype1 strains have been reported to be penicillin susceptible (Porat et al., 2001, Brandileone et al., 1998, Henriques Normark et al., 2001). This high susceptibility to antibiotics of serotype 1 relative to lower antibiotic susceptibility patterns for other pneumococcal serotypes may be attributed to its rarity in nasopharyngeal carriage, which reduces the probability for recombination events to occur. In a study comparing antibiotic susceptibility patterns between serotypes mostly found in invasive disease (serotype 1 and 5) and those found mostly in nasopharyngeal carriage (serotype 6B and 23F), it was observed that serotypes 1 and 5 were susceptible to all antibiotics tested including penicillin, erythromycin and tetracycline but not to trimethoprim-sulfamethoxazole (SXT). In contrast, a significant proportion of the strains belonging to serotype 6B (56%) and 23F (21%) were resistant to penicillin and were associated with multidrug resistance (Porat et al., 2001).

1.13.2. Vaccines

1.13.2.1. History of *S. pneumoniae* vaccines

The first pneumococcal vaccine trial was conducted by Sir Almroth E. Wright, a British physician in 1911. He tested a killed whole cell pneumococcal vaccine on South African miners who had a high incidence of pneumococcal infections at the time. His intervention resulted in a significant reduction in pneumonia incidence amongst the miners (Wright et al., 1914). That trial paved the way for many other trials that led to the licensure of two hexavalent polysaccharide vaccines in the United States which were later withdrawn in 1954. The advent of antibiotics proved to be very effective against pneumococcal infections during this time. The interest in pneumococcal vaccines was rekindled in the 1970s because of the emergence of pneumococcal strains with reduced susceptibility to available antibiotics at the time (Kazanjian, 2004). This led to the development of the pneumococcal polysaccharide vaccine and its use were championed by Robert Austrian and others. Firstly, a 14-valent vaccine was licensed in The United States in 1977 and later replaced by a 23-valent vaccine in 1983 (Grabenstein and Klugman, 2012). Conjugate vaccines, with increased immunogenicity, were later developed and licensed in the United States in 2000 and have since been introduced to national childhood vaccination programs in many other countries around the world.

1.13.2.2. Current licensed *S. pneumoniae* vaccines

Two different types of pneumococcal vaccine are currently licensed, namely, the polysaccharide vaccines (PPSV) and the conjugate vaccines (PCV) (Table 1.1).

These vaccines are licensed for use in different parts of the world today with different dosing schedules and across different age groups.

Table 1.2 Current licensed pneumococcal vaccines and serotypes included

Vaccines	Pneumococcal serotypes included											
PCV-7	4	6B	9V	14	18C	19F	23F					
PCV-10	4	6B	9V	14	18C	19F	23F	1	5	7F		
PCV-13	4	6B	9V	14	18C	19F	23F	1	5	7F	3	6A	19A
PPSV-23	4	6B	9V	14	18C	19F	23F	1	5	7F	3	33F	19A
	2	8	9N	10A	11A	12F	15B	17F	20	22F			

PCV, pneumococcal conjugate vaccine; PPSV, pneumococcal polysaccharide vaccine

1.13.2.3. Pneumococcal polysaccharide vaccines

Pneumovax23, produced by Merck, Whitehouse Station, NJ, USA is a pneumococcal conjugate vaccine containing purified capsular polysaccharides from 23 different serotypes of *S. pneumoniae*, including serotype 1. It is recommended for use in adults over 65 years of age and for younger adults with increased risk of an IPD. This vaccine has been shown to be very useful in immune competent individuals (Butler et al., 1993). The vaccine is currently used in many developed countries but is yet to be widely used in many developing countries, including any African country, probably due to the relatively high costs.

Pneumococcal polysaccharide vaccines induce a B-cell response with the production of anticapsular opsonising immunoglobulin but no T-cell response and hence no

memory, limiting its usefulness in children less than 2 years of old (Douglas et al., 1983). Due to the fact that this vaccine is not usually immunogenic in children less than 2 years, it is not recommended for use in this age group. Given the significantly high risk of pneumococcal invasive disease in this age group, the need for an effective vaccine led to the development of pneumococcal conjugate vaccines.

1.13.2.4. Pneumococcal conjugate vaccines.

Several pneumococcal conjugate vaccines have undergone clinical trials around the world but the first licensed PCV vaccine, available in 2000, was a 7-valent pneumococcal conjugate vaccine, Prevenar 7 developed Wyeth, New York, NY, USA. It contains capsular polysaccharides from seven serotypes (Table 1.1) conjugated with an alum-bound adjuvant to a non-toxic mutant of diphtheria toxin (CRM197) acting as a protein carrier. Selection of the seven serotypes included in PCV-7 was based on the serotypes with the highest prevalence in invasive disease in developed countries particularly in the United States of America. PCV-7 did not include serotype 1.as it was not one of the leading serotypes in invasive disease in The USA, even though it was one of the two predominant serotypes in IPD in several African countries. Improving on the performance of the PCV-7, a 13-valent conjugate vaccine, PCV-13, containing all seven serotypes included in PCV-7 and six additional serotypes including serotype 1, was licensed and has been in use in many countries since 2010 (Centers for Disease and Prevention, 2010). Also available is a 10-valent conjugate vaccine, PCV- 10 Synflorix, developed by GlaxoSmithKline, London, UK containing all the serotypes contained in PCV7 and serotypes 1, 5 and 7F. The conjugate vaccines are able to induce both B-cell and T-cell responses, which make them suitable for children less than 2 years old.

Currently, several African countries have introduced PCV-10 or PCV-13 into their national immunization programs.

The introduction of pneumococcal conjugate vaccines has led to a significant decrease in rates of pneumococcal disease caused by vaccine serotypes (Kaplan et al., 2004, Roca et al., 2013, Roca et al., 2011). The administration of these vaccines also had indirect effects or "herd effects", with reductions in invasive pneumococcal disease incidence amongst the non-vaccinated population, mostly attributed to the reduction of transmission of vaccine serotypes in communities as a result of decreased carriage (Roca et al., 2011). In addition to reducing rates of invasive disease and carriage, use of the conjugate vaccines have led to reductions in antimicrobial resistance is some settings as many drug resistant clones identified prior to the advent of conjugate vaccines are serotypes covered by these vaccines (Dagan, 2009).

The effectiveness of PCVs amongst Africans has been demonstrated by various studies. In The Gambia, a trial of the PCV-9 vaccine involving 17437 children age 6-51 weeks old showed a 15% reduction in hospitalisation and a 16% reduction in all-cause mortality as well as a 7% and 37% vaccine efficacy against both clinical and radiological pneumonia (Cutts et al., 2005). In another trial in The Gambia, use of the PCV-7 showed a significant reduction in the carriage prevalence of vaccine type pneumococci amongst healthy participants (Roca et al., 2011, Roca et al., 2013). The PCV-7 was introduced in the Gambian Expanded Program on Immunization (EPI) in 2009, however this vaccine lacked pneumococcal serotype 1 and 5, which were the leading cause of IPD in the Gambia (Adegbola et al., 2006). The serotype 1 and 5-containing PCV13 was subsequently introduced to the national immunization programme in 2011.

The effect of PCVs on serotype 1 is still being monitored. Some countries have reported an increase in pneumococcal serotype 1 following the use of PCV-7 (Nunes et al., 2008). Trials in The Gambia and in South Africa, of the 9-valent PCV-9, which contains serotype 1, showed limited vaccine efficacy against disease caused by serotype 1, even though the size of these studies were small (Cutts et al., 2005).

In spite of the proven effectiveness of currently available pneumococcal conjugate vaccines, there are a few limitations associated with their use. One of the more serious concerns is due to the fact that protection is limited to vaccine serotypes with the risk of replacement in the population with non-vaccine serotypes in carriage and disease (Byington et al., 2005, Park et al., 2008). In England and Wales, an increase in non-PCV13 invasive pneumococcal disease was observed in children younger than 5 and adults over 45 years old (Waight et al., 2015). In the same region, an increase in non-PCV 13 serotypes was also observed in nasopharyngeal carriage with significant differences observed in serotypes 21, 23B, 33F and 35F (Gladstone et al., 2015b). In the USA, replacement in non-PCV13 vaccine types has been documented in invasive disease and carriage mostly due to serotypes 21 and 15B/C (Lee et al., 2014). In Africa, serotype replacement in pneumococcal nasopharyngeal carriage has been documented in Kenya after introduction of PCV-10 (Hammitt et al., 2014) but not in Gambia in a village randomized trial of PCV-7 (Roca et al., 2011, Roca et al., 2013). Other limitations include the high cost of production which translates to high vaccine costs and the complexities of conjugation with over 90 pneumococcal serotypes involved. There are also concerns with possible increases in incidence of disease caused by other respiratory pathogens such as *Staphylococcus*

aureus. These limitations have indicated the need for further research into alternative pneumococcal vaccines.

1.13.2.5. Future pneumococcal vaccines

One of the options for alternative pneumococcal vaccines that have been widely explored is the use of proteins that are highly conserved in the pneumococcus in producing pneumococcal protein vaccines. This is expected to address concerns related to limited serotype coverage seen with the use of PCVs. Five pneumococcal proteins leading in the development of pneumococcal protein vaccines either individually or in combination are pneumolysin (Ply), pneumococcal histidine triad protein D (PhtD), pneumococcal surface protein (PspA), pneumococcal, pneumococcal surface adhesion A (PsaA) and PcpA (Leroux-Roels et al., 2014, Prymula et al., 2014, Bologa et al., 2012, Berglund et al., 2014, Chen et al., 2015). The GlaxoSmithKline (GSK) Biologicals protein vaccine containing Ply and PhtD is currently undergoing a phase II clinical trial in The Gambia (Odutola et al., 2015). Other types of vaccines under investigation include the use of inactivated unencapsulated whole cell pneumococci. The whole cell pneumococcal vaccines would offer some cost advantage over conjugate vaccines due to their lower cost of production (Goncalves et al., 2014). Other considerations alternative pneumococcal vaccines include use of DNA vaccines, which are simpler to manufacture, offer lower production costs and are better stability during transportation, conditions which are favourable to developing countries (Miyaji et al., 2002, Vadesilho et al., 2012).

1.13.3. Other control measures

Other control measures against pneumococcal infections include reductions in air pollution from cigarette smoking and use of biomass fuels such as wood burning stoves, and overcrowding all of which have been shown to be risk factors for pneumococcal infections (Reisman et al., 2014, Nuorti et al., 2000). Also, provision of adequate nutrition and improving public health awareness on pneumococcal infections would facilitate more effective control of pneumococcal infections.

Active surveillance is also an important tool in the control of pneumococcal infections. It provides information on the burden of disease, prevailing serotypes, antimicrobial use and resistance patterns and information on vaccine usage and effectiveness. This improves local awareness and informs national and international policy makers on effective treatment and prevention strategies.

1.14. Project aims and objectives

The significant burden of *S. pneumoniae*, a leading cause of severe bacterial illness and deaths globally, on health systems at regional and country levels, underscores the need for continued efforts at effective control and prevention of this pathogen. Regional differences in the population structure and epidemiology of *S. pneumoniae*, as well as recent changes such as the introduction of pneumococcal conjugate vaccines and changing trends in antibiotic use, highlight the importance of on-going surveillance at regional and country levels. This thesis aimed to describe the population structure of *S. pneumoniae* in The Gambia in the context of PCV vaccination and the patterns of antibiotics resistance among pneumococcal isolates. Given the importance of *S. pneumoniae* serotype 1 as a leading cause of IPD in the

West Africa sub-region, this thesis aimed to understand in particular, the characteristics of invasive and carriage serotype 1 strains in West Africa, particularly the clonal complex consisting of sequence type 217 (ST217) and the locus variants

The specific objectives were to:

- Determine the population structure of pneumococci in The Gambia in both invasive pneumococcal disease and nasopharyngeal carriage over the last two decades, including the pre- and post-PCV era. This aim would be used to test the hypothesis that sustained vaccination is associated with changes in the predominant serotypes and sequence types circulating in The Gambia.

- Determine the antibiotic resistant patterns and assess the impact of vaccination on the prevalence of antibiotic-resistant *S. pneumoniae* in nasopharyngeal carriage and invasive disease isolates in rural and urban Gambia. This aim would be used to test the hypothesis that the introduction of PCVs would result in reductions in levels of antibiotic resistance among pneumococcal isolates circulating in The Gambia as the known antibiotic resistant strains are included in the vaccine.

- To describe the evolution of pneumococcal serotype 1 strains in West Africa.

2. Materials and methods

2.1. Bacterial strains

S. pneumoniae isolates obtained from patients and healthy volunteers in the West African sub-region were included in this study. A detailed description of the countries from which these pneumococcal isolates were obtained, the rationale for selection, age groups and associated metadata has been provided in sections 3.2, 4.2 and 5.2.

2.2. Primers used in this study

A list of the primers used for molecular serotyping and multilocus sequence typing (MLST) is provided below (Table 2.1). All primers were obtained from Metabion International, Germany.

Table 2.1 List of primers used in this study

*Primer	Sequence (5′ to 3′)	Size (bp)	Purpose	Reference
cpsA F	GCA GTA CAG CAG TTT GTT GGA CTG ACC	190	*S. pneumoniae* detection	(Pai et al., 2006)
cpsA R	GAA TAT TTT CAT TAT CAG TCC CAG TC	190	*S. pneumoniae* detection	
wzy F	CTC TAT AGA ATG GAG TAT ATA AAC TAT GGT TA	280	*S. pneumoniae* serotype 1detection	
wzy R	CCA AAG AAA ATA CTA ACA TTA TCA CAA TAT TGG C	280	*S. pneumoniae* serotype 1detection	
aroE F	GCC TTT GAG GCG ACA GC	405	MLST	(Enright and Spratt, 1998)
aroE R	TGC AGT TCA (G/A)AA ACA T(A/T)T TCT AA	405	MLST	
gdh F	ATG GAC AAA CCA GC(G/A/T/C) AG(C/T) TT	460	MLST	
gdh R	GCT TGA GGT CCC AT(G/A) CT(G/A/T/C) CC	460	MLST	

gki F	GGC ATT GGA ATG GGA TCA CC	483	MLST
gki R	TCT CCC GCA GCT GAC AC	483	MLST
recP F	GCC AAC TCA GGT CAT CCA GG	450	MLST
recP R	TGC AAC CGT AGC ATT GTA AC	450	MLST
spi F	TTA TTC CTC CTG ATT CTG TC	474	MLST
spi R	GTG ATT GGC CAG AAG CGG AA	474	MLST
xpt F	TTA TTA GAA GAG CGC ATC CT	486	MLST
xpt R	AGA TCT GCC TCC TTA AAT AC		MLST
ddl F	TGC (C/T)CA AGT TCC TTA TGT GG	441	MLST
ddl R	CAC TGG GT(G/A) AAA CC(A/T) GGC AT	441	MLST

*The primers are named after the gene; *cpsA*, capsular biosynthesis gene A; *wzy*, O-antigen polysaccharide polymerase; *aroE*, shikimate dehydrogenase; *gdh*, glucose-6-phosphate dehydrogenase; *gki*, glucose kinase; *recP*, transketolase; *spi*, signal peptidase I; *xpt*, xanthine phosphoribosyltransferase); *ddl*, D-alanine-D-alanine ligase.

2.3. Culture and storage conditions for *S. pneumoniae* isolates

All carriage isolates of *S. pneumoniae* were obtained by culturing nasopharyngeal swabs (NPS) stored in Skim Milk Tryptone Glucose Glycerol (STGG) directly, or from frozen glycerol stocks of pneumococci. Invasive isolates were obtained from frozen glycerol stocks of pneumococci. Gentamicin sheep blood agar (GBA) plates consisting of 39 g/L Columbia base agar (Oxoid, Basingstoke, UK), 5% (v/v) defribinated sheep blood (TCS, Biosciences Ltd, Botolph Claydon, UK) and 5 µg/ml gentamicin (Oxoid, Basingstoke, UK) were used for both direct culture and reviving of stored isolates. All cultured plates were incubated for 18-24 hours at 37ºC in 5% CO_2.

2.3.1. Optochin sensitivity

Pneumococci were confirmed by morphological characteristics and optochin sensitivity. Presumptive pneumococcal colonies were picked from the primary plate and streaked onto 5% sheep blood Columbia agar (BA) plates. A 5 µg optochin disc (Oxoid, Basingstoke, UK) was placed in the centre of each streak and incubated overnight at 37ºC in 5% CO_2. Zones of inhibition greater than 14 mm indicated susceptibility. *S. pneumoniae* ATCC 49619 and *Streptococcus pyogenes* ATCC 49399 strains were used as positive and negative controls respectively.

2.3.2. Storage of *S. pneumoniae* cells

All confirmed *S. pneumoniae* cells were stored in 15% glycerol broth at -70ºC in line with standard protocols (Satzke et al., 2013), until retrieved for further processing.

2.3.3. Antimicrobial susceptibility testing

2.3.3.1. Disk diffusion method

Antimicrobial susceptibility testing was performed using the Kirby-Bauer disc diffusion method for pneumococcal isolates following the Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2011). *S. pneumoniae* ATCC 49619 was included as a control strain in each batch of susceptibility testing. Inocula were prepared from fresh cultures (18-24 hour) of confirmed pneumococcal isolates from BA plates. Pneumococcal colonies were emulsified in 2 ml normal saline of turbidity equivalent to 0.5 McFarland standard, and spread evenly onto Mueller Hinton agar plates supplemented with 5% sheep blood using sterile cotton tipped swabs (Fisher Scientific UK Ltd, Loughborough, UK). Using sterile forceps, antibiotic discs of

prescribed concentrations were placed onto the inoculated plate. Antibiotics tested included oxacillin (1 µg), chloramphenicol (30 µg), erythromycin (5 µg), cotrimoxazole (1.25 µg/23.75 µg), tetracycline (30 µg) and cefotaxime (30 µg) purchased from Oxoid, Basingstoke, UK. Inoculated plates were incubated for 16-18 hours at 37°C in 5% $C0_2$. Zone of inhibition sizes were measured and interpreted using the CLSI interpretative chart (CLSI, 2017) (Table 2.2).

2.3.3.2. Epsilometer (E-test) agar diffusion method

To determine minimum inhibitory concentration (MIC), E-test strips were used. The strips were applied using sterile forceps onto inoculated agar plates (described section 2.3.4.1) facing upwards towards the opening of the plate. The MIC value was read following overnight incubation where the edge of the inhibition ellipse intersected the strip. MICs were determined for penicillin and cefotaxime and results interpreted using the CLSI interpretative chart (CLSI, 2017) (Table 2.2).

Table 2.2 Clinical and Laboratory Standards Institute interpretative charts of pneumococcal disk diffusion zone size

	Sensitive	**Intermediate**	**Resistant**
Disk diffusion	**Mm**	**mm**	**mm**
Chloramphenicol	≥ 21	-	≤ 20
Erythromycin	≥ 21	16-20	≤ 15
Co-trimoxazole	≥ 19	16-18	≤ 15
Tetracycline	≥ 28	25-27	≤ 24
MIC	**µg/ml**	**µg/ml**	**µg/ml**
Penicillin	≤0.06	0.12-1	≥2
Cefotaxime	≤ 1	2	≥4

2.4. DNA extraction and analysis

2.4.1. Extraction of genomic DNA

One millilitre of sterile phosphate buffered saline (PBS) was added to an 18-24 hour pneumococcal cultured blood agar plate and the cells scrapped with a sterile L-shape plastic spreader to make a suspension of all the pneumococcal colonies. The suspension was collected and centrifuged at 7500 rpm for 10 min. The pellet was resuspended in 20 mg/ml lyzozyme in lysis buffer (20 mM Tris-Cl, pH8.0, 2 mM sodium EDTA, 1.2% Triton X-100). DNA was extracted using the DNeasy Blood & Tissue Kit, (Qiagen, Crawley, UK), following manufacturer's Gram positive protocol. DNA was eluted in 200 µl of elution buffer and stored at -20°C. DNA was quantified using the nanodrop for microarray analysis. In preparation for whole genome sequencing, DNA was quantified using picogreen and DNA integrity assessed by agarose gel electrophoresis.

2.4.2. Agarose gel electrophoresis

Agarose gels were prepared by dissolving agarose (Sigma-Aldrich, Poole, UK) in 1 × TAE buffer (40 mM Tris Acetate, 1 mM EDTA, pH 8) to a final concentration of 1% (w/v). Ethidium bromide (Sigma-Aldrich, Poole, UK) was added to a concentration of 0.1 µg/ml. DNA samples mixed with DNA loading buffer at a 5:1 ratio were loaded onto the gel in an electrophoresis tank containing 1 × TAE buffer. DNA fragments were separated at 120 V for 1 hour and visualised by UV illumination using a gel documentation system (Gel Doc 2000; Bio-Rad, UK). Band sizes were estimated by comparing with quantitative DNA markers, Quick-Load®

100 bp DNA Ladder and 50 ng NEB Lambda DNA-Hind III Digest (New England Biolabs (UK) Ltd).

2.5. Pneumococcal serotyping

2.5.1. Latex agglutination

Latex agglutination was performed as described previously (Adegbola et al., 2006, Roca et al., 2011). Briefly, pneumococcal cell suspensions were made in 2 ml normal saline (1.0 McFarland) from an overnight blood agar plate. 20 µl of the suspension was dispensed into 10 separate wells of a serotyping tray and an equal amount of each of the main group (A, B, C, D, E, F, G, H, I and Omni) latex antisera (Statens Serum Institute, Copenhagen, Denmark) added to the wells. The mixture was rocked for a maximum of two minutes and agglutination observed in the wells. This procedure was then repeated with the subgroups under each main group and further typed with corresponding serum factor types. The pneumococcal serotype was determined with the aid of a chart provided by Staten Serum Institute, Denmark.

2.5.2. Molecular serotyping by block PCR

All serotype 1 isolates were confirmed by molecular serotyping as described previously (Pai et al., 2006). All reagents were purchased from Qiagen Ltd, UK except primers from Metabion International AG, Germany. Glycerol stocks of cells were used as template in a multiplex PCR reaction that comprised of:

PCR cycling conditions were as follows:

95°C hold for 15 min, followed by 35 cycles of

94°C for 30 sec

54°C for 90 sec

72°C for 60 sec

72°C final extension for 10 min.

Products were visualized by agarose gel electrophoresis.

2.6. Multilocus sequence typing

MLST was performed on all *S. pneumoniae* serotype 1 carriage and invasive (2004 - 2014) isolates in this study. Isolates were recovered from glycerol stocks by plating out on 5% BA plates and incubated at 37°C in 5% CO_2 for 18 hours. A single colony from each isolate was picked, streaked and incubated at 37°C for 18 hours. Glycerol stocks were prepared and these were used as templates for amplification of the seven housekeeping genes (*aroE, gdh, gki, recP, spi, xpt,* and *ddl*) targeted for MLST. The reaction mixture contained:

10X Buffer with 1.5 mm $MgCl_2$ 2.5 µl	
10 mM dNTPs	0.5 µl
12.5 mM primers	1 µl each
5 U/µl *Taq* Polymerase	0.25 µl
Template (glycerol stocks)	1 µl

Sterile DNA free water	To 25 µl

PCR cycling conditions were as follows:

5 min hold at 95°C, followed by 30 cycles of

95°C for 30 sec,

55°C for 30 sec

72°C for 1 minute and a final extension of

72°C for 10 min.

Products were visualized by agarose gel electrophoresis. PCR products were purified using Exonuclease I and Shrimp Alkaline Phosphatase (USB, Staufen, Germany). The reaction mixture contained:

Exonuclease I	0.5 µl
Shrimp Alkaline Phosphatase	0.5 µl
PCR product	8 µl
Sterile DNA free water	To 10 µl

Cycling conditions were as follows:

37°C for 45 min

80°C for 20 min

Sequencing was done on both strands with BigDye Terminator Cycle Sequencing kit (Applied Biosystems, UK). The reaction mixture contained:

5X Buffer	2.075 µl
BigDye	0.5 µl
5 µM primer	0.32 µl
Cleaned PCR products	2 µl
Sterile DNA free water	To 10 µl

PCR cycling conditions were as follows:

1 min hold at 96°C, followed by 30 cycles of

96°C for 10 sec

50°C for 5 sec

60°C for 4 min

The labelled fragments were separated by size using 3130xl Genetic Analyser (Applied Biosystems, UK). Sequences were edited, complementary sense and antisense fragments were aligned using the Lasergene software (version 7.1; DNASTRAR, US) and submitted to the MLST database website (www.mlst.net). Existing or novel allele type numbers were assigned and sequence type numbers defined by the MLST database.

2.7. Whole genome sequencing, assembly and annotation of *S. pneumoniae* serotype 1 genomes.

Whole genome sequencing (WGS) of pneumococcal serotype 1 isolates used in this thesis was performed at the Wellcome Trust Sanger Institute (WTSI) Cambridge, UK. Sequencing was performed by staff of the WTSI as part of the Pneumococcal African Genome (PAGe) consortium and the Global Pneumococcal Sequencing (GPS) Project of which The MRC Gambia Unit is a collaborating institution.

Multiple paired-end DNA sequencing using index tagged libraries was performed using the Illumina Genome Analyser GAII (Illumina, USA) as described previously (Cornick et al., 2015). Published sequence reads generated have been deposited in the European Nucleotide Archive (ENA) (http://www.ebi.ac.uk/ena/) under study number ERP000156 (Cornick et al., 2015).

All sequence analysis was conducted on the Genomics Virtual Lab on the Cloud Infrastructure for Microbial Genomics (CLIMB) running on Ubuntu (Connor et al., 2016). Paired end sequencing reads in fastq format were downloaded unto the server and subjected to the nullabor pipeline (Seemann T) which is one of the bioinformatics tools preloaded on the CLIMB server.

With the nullabor pipeline, the following analysis were performed:

- Sequencing reads were trimmed and low quality bases and reads removed using Trimmomatic (Bolger et al., 2014).
- Species identification was performed per isolate by a k-mer analysis against a database of known genomes using Kraken version 0.10.5-beta (Wood and Salzberg, 2014).

- De novo assemblies were generated using MEGA-HIT version 1.0.3 (Li et al., 2015) and annotation performed using Prokka version 1.12-beta (Seemann, 2014).
- MLST profiles were determined for each isolate from generated assemblies using the Seeman's MLST program (Seemann T).
- Antibiotic resistance genes were detected from generated assemblies using abricate version 0.3 (Seemann T) against a database of known antimicrobial resistance genes.
- Single nucleotide variants (SNVs) were determined from the sequencing reads by mapping them to the reference genome- spn1041 (Genbank: CACE00000000) using Snippy version 3 (Seemann T) and multiple sequence alignment of SNVs in the core genome of all isolates generated using snippy-core. A maximum likelihood and phylogenetic tree was generated using FastTree version 2.1.8 Double precision (No SSE3)~ OpenMP (8 threads) and the distance matrix based on pairwise SNV difference between the isolates was calculated.

Using the multiple sequence alignment generated from the nullabor pipeline, recombination events was determined using Genealogies Unbiased By recomBinations In Nucleotide Sequences (Gubbins) algorithm (Croucher et al., 2015). With Gubbins, a phylogenetic tree was constructed based on putative point mutations occurring outside regions having high densities of base substitutions. Recombination events were displayed in a heatmap.

A phylogenetic tree was constructed using RAxML v7.0.4 (Stamatakis, 2006) with the generalized time reversible (GTR) model with Gamma heterogeneity among nucleotide sites and 100 bootstrap replicates. Visualization and annotation with associated metadata of study isolates was performed using FigTree version 1.4.2 and Interactive Tree of Life (iTOL) software (Letunic and Bork, 2016).

Visualization of the evolutionary signals from the core and accessory regions of the genomes was performed using PANINI (Pangenome Neighbor Identification for Bacterial Populations), a web based tool that integrates with Microreact (Abudahab et al., 2017)

2.8. Ethical approval

Approval for the studies conducted in this thesis was obtained from the Joint Medical Research Council (MRC)/Gambia Government Ethics Committee and the Ethics Committee of the London School of Hygiene & Tropical Medicine, UK. Community and individual consent was obtained from study participants and the conduct of the trial was guided by a Data Safety and Monitoring Board.

2.9. Data analysis

All statistical analyses were carried out in STATA (version 11, Stata Corporation, College Station TX) and significance testing performed Chi-square test. A *p*-value of less than 0.05 was taken to indicate statistical significance. To determine the geographical relationship of serotype 1 and 5 clones to those found in other regions of the world, cluster analysis of allelic profiles of serotype 1 and 5 isolates from this

study and those from the MLST database was performed using geoBURST (http://goeBURST.phyloviz.net).

3. Population structure of *Streptococcus pneumoniae* in The Gambia from pre-and post-introduction of pneumococcal conjugate vaccines

3.1. Introduction

Streptococcus pneumoniae remains a leading cause of serious infections and deaths in young children and the elderly globally (O'Brien et al., 2009). As global efforts to prevent and control pneumococcal disease gain momentum, a good understanding of the population structure of *Streptococcus pneumoniae*, including the circulating pneumococcal serotypes and genotypes would be critical in guiding the development of appropriate interventions. Some studies have reported changes in circulating serotypes causing invasive disease and nasopharyngeal carriage in different parts of the world (Ubukata et al., 2015, Devine et al., 2017). However, most of these studies have been conducted in developed country settings with few in developing countries and there remain large gaps in our understanding of the pneumococcal population structure in developing countries where the burden of disease is disproportionately highest. In addition, changes in recent times, such as the increasingly widespread introduction of routine pneumococcal vaccination and changing patterns of HIV prevalence, could have an impact on the circulating pneumococcal serotypes and genotypes.

The PCV- 7, the first licensed pneumococcal conjugate vaccine rolled out across many countries including countries in Sub-Saharan Africa, recorded huge successes in the reduction of IPD and nasopharyngeal carriage caused by vaccine serotypes (Waight et al., 2015, Mackenzie et al., 2016, Roca et al., 2013). Despite these successes, the introduction of PCV-7 has been associated with an increase in IPD caused by non-vaccine serotypes (Olarte et al., 2017, Del Amo et al., 2016).

Additionally, serotypes 1 and 5 were not included in PCV-7; these two serotypes have been known to be leading serotypes causing IPD in sub-Saharan Africa (Adegbola et al., 2006, Everett et al., 2012, Brueggemann et al., 2013). Reports from The Gambia indicated that approximately one in every five cases of IPD were attributable to serotype 1 in the pre-PCV era (Adegbola et al., 2006). There have also been reports of deadly epidemic outbreaks in the African meningitis belt caused by pneumococcal serotype 1 (Kwambana-Adams et al., 2016, Leimkugel et al., 2005, Yaro et al., 2006). Consequently, the use of increased valence vaccines that contain serotypes 1 and 5, such as PCV 10 (PCV7 serotypes plus serotypes 1, 5 and 7F) and PCV13 (PCV-7 plus serotypes 1, 5, 3, 6A, 19A, 7F) would seem appropriate for sub-Saharan Africa. In The Gambia, PCV-7 was introduced into its national immunization programme in 2009, with a switch to PCV-13 in 2011.

Given the importance of pneumococcal serotype 1 as a leading cause of IPD in The Gambia, monitoring the population structure of invasive disease-causing serotypes before and after the introduction of PCV-13 into routine immunization in The Gambia would be vital in any efforts to assess vaccine impact. Similarly, it would be both appropriate and critical to monitor any changes in the pneumococcal population in nasopharyngeal carriage, as nasopharyngeal carriage is known to precede invasive disease and pneumococcal vaccination could have an effect on serotypes seen in carriage.

The aims of this chapter are to undertake a longitudinal study to determine the population structure of *S. pneumoniae* in The Gambia from January 1995 - December 2016, providing a description of the circulating serotypes and associated genotypes and in particular pneumococcal serotype 1. It also aims to assess any changes in the circulating pneumococcal serotypes following the introduction of the

pneumococcal conjugate vaccine in The Gambia. These aims will be used to explore the hypothesis that sustained vaccination can cause serotype replacement and that inclusion of further serotypes in formulation may not be as effective as expected.

3.2. Methods

Pneumococcal studies conducted in The Gambia from 1995- 2016 were sourced through a literature search as well as on-going unpublished studies (personal communication). Eight main datasets contributing to various studies conducted during the pre-defined period were identified. Of these, four studies were selected for the current analysis as they were conducted across different regions in The Gambia, covering a larger geographical spread (Figure 3.1), and also included study population of all age groups (Table 3.1).

Figure 3.1 Map of The Gambia showing the administrative regions. Nasopharyngeal pneumococcal carriage isolates were obtained from Sibanor in the Western Region and Basse in the Upper River Region of The Gambia. Invasive isolates were obtained from the MRC clinic in the Greater Banjul area embedded in the Western Region and the health centres in the Central River Region and the Upper River Region. Map taken from Ezilon (www.exilon.com)

Table 3.1 Selected published pneumococcal studies conducted in The Gambia from 1996 – 2016

	Study name			
Study[$]	**PNI**	**SVT**	**PSP**	**PCSU**
Characteristics				
Location	MRC Ward, Fajara	Sibanor, Foni	Basse & Bansang	Basse
Region*	WR	WR	URR	URR
Setting	Urban & rural	Rural	Rural	Rural
Invasive/carriage	Invasive	Carriage	Invasive	Carriage
Context	Hospital surveillance	PCV-7 vaccine trial - CSS	Population based surveillance	Population based CSS
Period of collection	Jan 1995-Dec 2014	Dec 2003 - Feb 2009	May 2008 - Dec 2016	May - Aug 2009
Age group	All age groups	All age groups	≥ 2 months	All age groups, 6-10 years in some
Specimen[#]	Blood, CSF, LA, PF, JFA	NPS	Blood, CSF, LA, PF, JFA	NPS

*WR; Western Region, CRR; Central River Region, URR; Upper River Region.

[#]CSF, cerebrospinal fluid; LA, lung aspirate; PF, pleural fluid; JFA, joint fluid aspirates.

$ PNI study; pneumococcal invasive disease study, SVT; pneumococcal carriage study conducted in the Western Region, PSP; pneumococcal surveillance project, PCSU study; pneumococcal carriage study conducted in The Upper River Region of The Gambia.

The IPD studies selected included the pneumococcal invasive disease (PNI) study, a hospital-based pneumococcal surveillance at the Medical Research Council (MRC) hospital in the Western region of The Gambia (Figure 3.1) and the pneumococcal surveillance project (PSP), a population-based surveillance of IPD investigating

effectiveness of the introduction of PCV (Mackenzie et al., 2016). This population-based surveillance was conducted in the Upper River Region (URR) and the Central River Region (CRR), both rural parts of The Gambia (Figure 3.1). For both studies, clinical specimens were collected from patients presenting at peripheral and referral health facilities or at the MRC ward with suspected pneumonia, septicaemia or meningitis. Specimens included blood, cerebrospinal fluid (CSF), lung aspirates, pleural aspirates or a combination of these. When more than one isolate was obtained during a single episode of illness, for example from blood and CSF, only one isolate was included in the analysis. Pneumococcal isolates obtained from these studies were serotyped as part of previous studies (Mackenzie et al., 2016, Adegbola et al., 2006) or sent to the World Health Organization (WHO) Regional Reference Lab (RRL) for invasive bacterial disease for the West Africa region, located at the MRC Gambia Unit, for pneumococcal serotyping by latex agglutination and antimicrobial susceptibility testing. All isolates were genotyped by multi-locus sequence typing (MLST).

Two nasopharyngeal carriage studies were included in this analysis. Firstly, large pneumococcal carriage studies (SVT studies) conducted in 21 selected rural villages in the Western region of The Gambia (Figure 3.1), which have been described previously (Hill et al., 2006, Roca et al., 2011). Briefly, a pre-vaccination cross sectional survey was conducted between December 2003 and May 2004 in which nasopharyngeal swab (NPS) samples were collected from subjects of all age groups (Hill et al., 2006). Following this, a single-blinded, cluster-randomised (by village) trial to evaluate the impact of PCV-7 on pneumococcal carriage was conducted in the study villages (Roca et al., 2011). In one group of 11 villages, all individuals over the age of 30 months received one dose of PCV-7 whilst subjects in this age

group resident in 10 control villages received one dose of Serogroup C meningococcal conjugate vaccine. All children less than 30 months of age in both study groups and infants born during the course of the trial received PCV-7. NPS were collected in three different cross-sectional surveys (from 4-6 months, 12-14 months and up to 30 months) (Roca et al., 2011).

Secondly, a cross-sectional population-based nasopharyngeal carriage study conducted in the Upper River Region of The Gambia between May and August 2009 (Usuf et al., 2015). All ages were sampled but in some households, only those ages 6-10 were sampled as this was the age group that participated in an earlier PCV- 9 trial (Cutts et al., 2005) conducted in The Gambia from 2000 – 2004.

3.3. Results

3.3.1. Epidemiology background of isolates

3.3.1.1. Western region

A total of 6371 NPS samples were collected from the Western Region, of which 52.19% (3325/6371) were positive for *S. pneumoniae*. Of the 3325 NPS positive for *S. pneumoniae*, a total of 3807 *S. pneumoniae* carriage isolates obtained and were analysed in this study (Roca et al., 2011). The percentage of males and females carrying *S. pneumoniae* were 50.37% (1675/3325) and 49.53% (1648/3325) respectively. The age distribution of individuals carrying *S. pneumoniae* was 26.80% (891/3325), 35.58% (1183/3325) and 37.56% (1249/3325) for ages < 5 years, 5 <15years and ≥ 15years respectively.

A total of 890 invasive pneumococcal isolates collected from 1995 – 2015 in the Western Region. However, 11 isolates were non-viable after culture and were

dropped from further analysis. Hence, a total of 879 pneumococcal isolates were analysed in this study. A summary of the specimen types and brief description of patient characteristics is provided in Table 3.2.

Table 3.2 Summary of patient characteristics of all invasive isolates analysed from the Western Region, The Gambia (N=879)

Characteristic	Category	Isolates N (%)
Age group (years)	<5 years	392 (44.60)
	5 - <15 years	107 (12.17)
	≥15 years	248 (28.21)
	Unknown	132 (15.02)
Gender	Female	325 (36.97)
	Male	463 (52.67)
	Unknown	91 (10.35)
Specimen	Blood	647 (73.61)
	CSF	121 (13.77)
	LA	56 (6.37)
	Others	55 (6.25)
Total		879

CSF, cerebrospinal fluid; LA, lung aspirate. Other specimen include pleural fluid and joint fluid aspirates.

3.3.1.2. Upper River region

A total of 1228 *S. pneumoniae* carriage isolates obtained from 1117/2933 NPS samples collected from the Upper River Region were analysed in this study (Usuf et al., 2015). Of the 1117 individuals carrying *S. pneumoniae*, 50.58% were males and

49.42% females. The age distribution of individuals carrying *S. pneumoniae* was 33.21% (371/1117), 53.80% (601/1117) and 12.98% (145/1117) for ages < 5 years, 5 <15years and ≥ 15years respectively.

A total of 400 pneumococcal isolates were collected from 2008 – 2016 in the Upper River Region but only 394 viable isolates were analysed in this study. Patient characteristics are shown in Table 3.3.

Table 3.3 Summary of patient characteristics of all invasive isolates analysed from the Upper River Region, The Gambia

Characteristics	**Category**	**Isolates N (%)**
Age group (years)	<5 years	326 (82.74)
	5 - <15 years	33 (8.38)
	≥15 years	39 (7.36)
	Unknown	6 (1.52)
Gender	Female	161 (40.86)
	Male	230 (58.38)
	Unknown	3 (0.76)
Specimen	Blood	286 (72.59)
	CSF	27 (6.85)
	LA	69 (17.51)
	Others	12 (3.05)
Total		394

CSF, cerebrospinal fluid; LA, lung aspirate. Other specimen include pleural fluid and joint fluid aspirates

3.3.2. Overall serotype distribution of pneumococcal study isolates

3.3.2.1. Western Region

In the Western region of The Gambia, 53 pneumococcal serotypes/groups were found causing IPD. Of these 53 serotypes, the leading ten serotypes were 1, 5, 14, 6A, 23F, 19A, 12F, 6B, 19F and 7F accounting for 73.17% (643/879) of all pneumococcal isolates (Table 5.4). PCV-7, PCV-10 and PCV-13 serotypes accounted 25.6% (225/879), 61.56% (541/879) and 76.12% (669/879) of all pneumococcal isolates causing IPD.

Among children <5 years old, the leading serotypes causing IPD in the Western region were serotypes 14 (15.56%, n=61/392), 5 (14.03%, n=55/392) and 1 (13.52%, n=53/392). Among children 5 - <15 years, serotypes 1 (36.45%, 39/107), 5 (14.95%, n=16/107) and 14 (6.54%, n=7/107) were the leading causes of IPD and among patients ≥15 years, serotypes 1 (23.39%, n=58/248), 5 (12.10%, n=30/248) and 7F (5.65%, 14/248) were the leading causes of IPD.

Of 3807 pneumococcal carriage isolates obtained, the leading serotypes obtained were 3, 15B/C, 6A, 11, 35B/C, 23F, 19F, 6B, 16A/F and 34, all accounting for 49.76% (1895/3807). PCV7, PCV-10 and PCV-13 serotypes accounted for 22.25% (847/3807), 24.4% (929/3807) and 41.81% (1592/3807) of all pneumococcal carriage isolates obtained.

Serotypes 1 and 5 were rarely found among nasopharyngeal carriage pneumococcal isolates (0.89% and 0.45% respectively; Figure 3.2) even though they were the leading serotypes amongst invasive isolates (19.23% and 14.11% respectively). In contrast, serotypes 11, 35B/C and 34 were predominant in carriage but rarely found

among invasive isolates (Figure 3.2), whilst serotypes 6B and 23F were commonly found among both invasive and carriage isolates.

Table 3.4 Age distribution of pneumococcal serotypes among invasive (1995-2015) and carriage (2003-2009) isolates collected from the Western Region of The Gambia

Serotype	Invasive isolates						Carriage isolates					
	<5 years	5-<15 years	≥15 years	Unknown	All ages (n)	All ages (%)	< 5 years	5-<15 years	≥ 15 years	Unknown	All ages (n)	All ages (%)
1	53	39	58	19	169	19.23	12	11	11	0	34	0.89
5	55	16	30	23	124	14.11	4	10	3	0	17	0.45
14	61	7	12	4	84	9.56	40	22	18	0	80	2.1
6A	51	5	4	12	72	8.19	109	47	49	0	205	5.38
23F	24	6	8	7	45	5.12	87	42	37	0	166	4.36
19A	21	2	10	6	39	4.44	52	37	30	0	119	3.13
12F	10	1	12	8	31	3.53	5	8	11	0	24	0.63
6B	16	3	5	7	31	3.53	74	47	40	0	161	4.23
19F	12	5	6	2	25	2.84	83	33	47	0	163	4.28
7F	7	1	14	1	23	2.62	4	18	9	0	31	0.81
3	2	1	8	6	17	1.93	60	152	126	1	339	8.9
9V	7	3	4	3	17	1.93	41	38	29	0	108	2.84
25	1	0	13	1	15	1.71	0	3	0	0	3	0.08
46	7	1	5	1	14	1.59	0	4	8	0	12	0.32
4	3	4	6	0	13	1.48	18	37	38	0	93	2.44
9L	6	0	4	2	12	1.37	12	6	14	0	32	0.84
10A/C	4	1	3	2	10	1.14	29	40	56	0	125	3.28
18C	7	1	0	2	10	1.14	9	35	32	0	76	2
9A	8	0	2	0	10	1.14	17	13	8	0	38	1
2	6	0	1	2	9	1.02	0	6	4	0	10	0.26
15B/C	2	0	5	2	9	1.02	78	69	61	0	208	5.46
16A/F	1	1	7	0	9	1.02	30	61	54	0	145	3.81
11	1	1	7	0	9	1.02	45	78	67	0	190	4.99
35B/C	0	1	1	2	4	0.46	40	67	75	0	182	4.78
34	1	0	2	1	4	0.45	39	59	38	0	136	3.57
NT	0	0	0	0	0	0	19	46	67	0	132	3.47
Others	26	8	21	19	74	8.41	167	391	422	0	980	25.77
Total	**392**	**107**	**248**	**132**	**879**	**100**	**1,074**	**1,379**	**1,353**	**1**	**3,807**	**100**

3.3.2.2. Upper River Region

Of 45 different serotypes found causing IPD in the Upper River Region (URR), the leading serotypes were 1, 5, 12F, 23F, 14, 35B, 9V, 6A, 6B and 24A/B, accounting for 72.85% (287/394) of all pneumococcal isolates (Table 3.5). PCV-7, PCV-10 and PCV-13 serotypes accounted for 15.48% (61/394), 52.79% (208/394) and 58.63% (231/394) respectively of all invasive pneumococcal isolates.

Among children > 5 years old, the leading serotypes causing invasive disease in the URR were serotypes 1 (19.63%, n=64/326); 12F (15.03%, n=49/326) and 5 (12.88%, n=42/326). Among children 5 - <15 years, serotypes 1 (42.42%, n=14/33), 5 (30.30%, n=10/33) and 35B (9.09%, n=3/33) were predominant. Whilst in the age group ≥15 years, serotypes 1 (31.03%, n=9/29), 12F (17.24%, n=5/29) and 5 (17.24%, n=5/29) were the leading causes of IPD.

Seventy-five different pneumococcal serotypes were found among the 1228 pneumococcal carriage isolates collected. Serotypes 6A, 23F, 3, 19F, 34, 15B, 14, 11A, 19A and NT were the leading serotypes found among all pneumococcal nasopharyngeal carriage isolates in the URR (Table 3.5). PCV-7, PCV-10 and PCV-13 serotypes accounted for 23.04% (283/1228), 25.33% (311/1228) and 48.7% (598/1228) respectively of all pneumococcal carriage isolates.

Similar to the Western region, serotypes 1 and 5 were rarely found in nasopharyngeal carriage (0.98% and 0.90% respectively- Figure 3.3) despite being the leading causes of IPD (21.7% and 14.46% respectively).

Table 3.5 Age distribution of pneumococcal serotypes among invasive (2008-2016) and carriage (2009) isolates collected from the Upper River Region of The Gambia.

	Invasive isolates						Carriage isolates				
Serotype	**<5 years**	**5-<15 years**	**≥15 years**	**Unknown**	**All ages (n)**	**All ages (%)**	**< 5 years**	**5-<15 years**	**≥ 15 years**	**All ages (n)**	**All ages (%)**
1	64	14	9	0	87	22.08	7	1	4	12	0.98
5	42	10	5	0	57	14.47	2	4	5	11	0.9
12F	49	0	5	1	55	13.96	5	0	4	9	0.73
23F	16	0	0	0	16	4.06	45	11	37	93	7.57
14	13	1	0	1	15	3.81	21	4	15	40	3.26
35B	12	3	0	0	15	3.81	13	6	9	28	2.28
9V	9	2	2	0	13	3.3	5	1	5	11	0.9
6A	8	2	0	0	10	2.54	87	8	66	161	13.11
6B	9	1	0	0	10	2.54	6	2	5	13	1.06
24B/F	8	0	1	0	9	2.28	1	1	9	11	0.89
46	7	0	0	1	8	2.03	1	1	1	3	0.24
3	7	0	0	0	7	1.78	17	14	59	90	7.33
15B/C	6	0	0	0	6	1.52	25	4	23	52	4.23
19A	3	0	3	0	6	1.52	13	3	20	36	2.93
2	4	0	0	2	6	1.52	2	0	4	6	0.49
38	5	0	1	0	6	1.52	1	1	7	9	0.73
40	5	0	0	0	5	1.27	0	2	8	10	0.81
7B	5	0	0	0	5	1.27	0	1	1	2	0.16
11B	4	0	0	0	4	1.01	1	1	3	5	0.41
13	4	0	0	0	4	1.01	7	8	17	32	2.61
25F	4	0	0	0	4	1.01	0	1	1	2	0.16
9L	3	0	1	0	4	1.01					
19F	3	0	0	0	3	0.76	43	8	36	87	7.08
34	1	0	0	0	1	0.25	17	6	33	56	4.56
11A	1	0	0	0	1	0.25	7	3	27	37	3.01
NT	0	0	0	0	0	0	8	5	22	35	2.85
Others	34	0	2	1	37	9.42	73	56	248	377	30.71
Total	326	33	29	6	394	100	407	152	669	1228	100

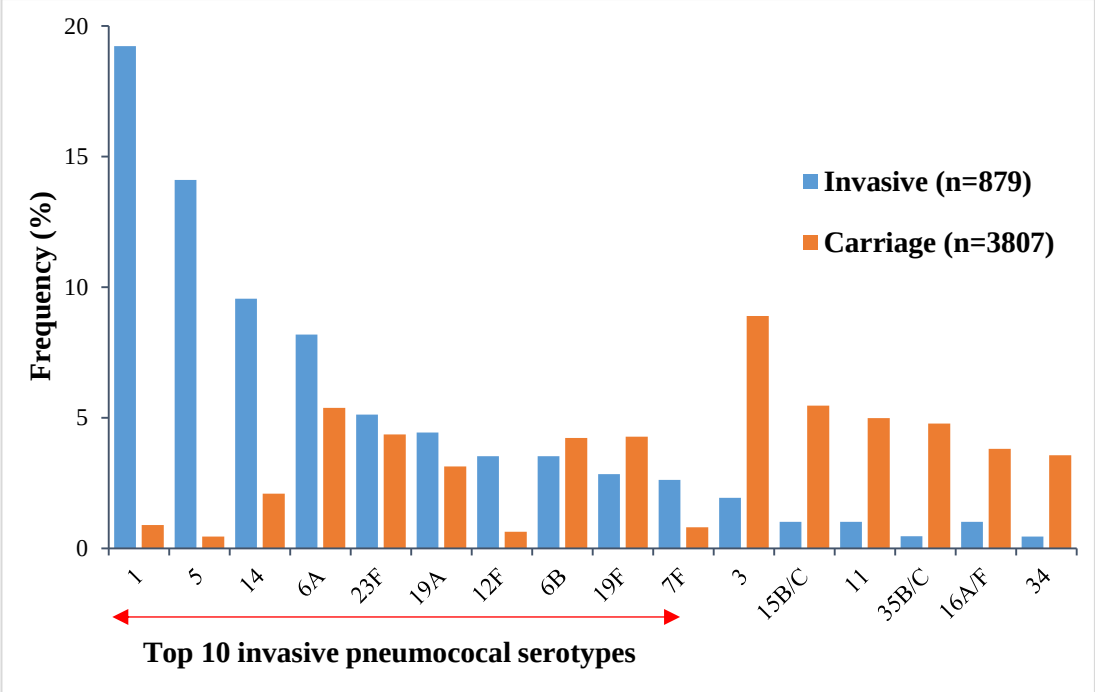

Figure 3.2 Pneumococcal serotype distribution of invasive and carriage isolates from the Western Region of The Gambia from (1995-2015).
The red arrow shows the leading invasive serotypes while the other serotypes are the leading serotypes in nasopharyngeal carriage.

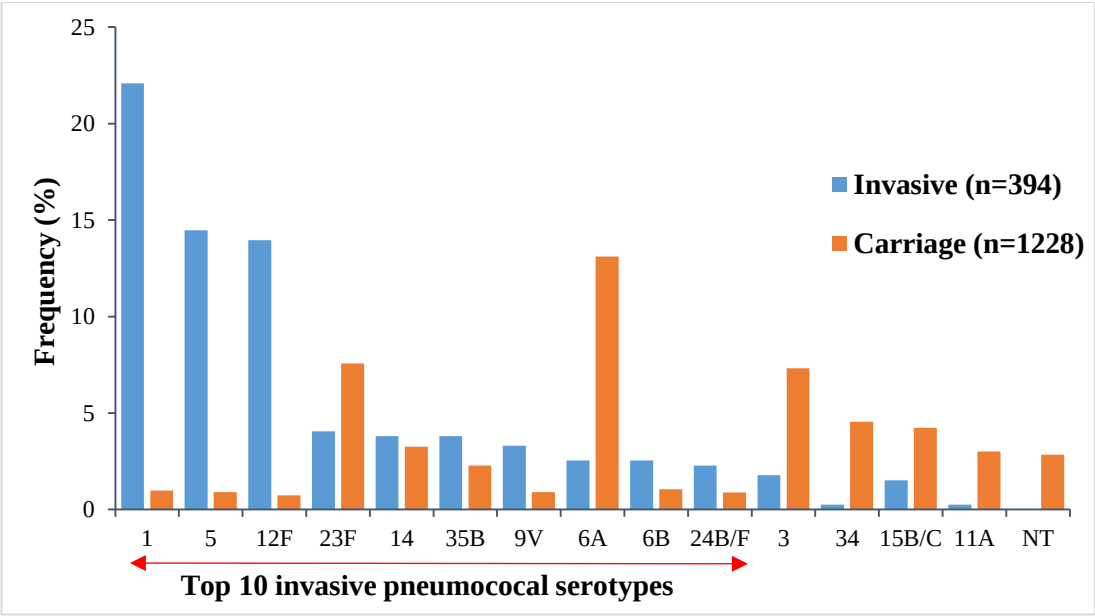

Figure 3.3 Pneumococcal serotype distribution of invasive and carriage isolates from the Upper River Region of The Gambia from (1995-2016).
The red arrow shows the leading invasive serotypes while the other serotypes are the leading serotypes in nasopharyngeal carriage.

3.3.3. Yearly and monthly distribution of invasive pneumococcal serotypes

Distribution of pneumococcal serotypes varied from year to year, most notable were the yearly fluctuations of serotypes 1 and 5. Prior to the introduction of PCV-7 in The Gambia, pneumococcal serotypes 1 and 5 were the leading causes of IPD and remained consistently the leading serotypes post introduction of PCV-7 and PCV-13. The slight exception to this pattern was in the post PCV-13 era when serotype 12F was the second leading cause of IPD after serotype 1, with serotype 5 was the third leading serotype.

The prevalence of serotype 1 varied from as high as 40% and 32% in 2007 and 1997 to as low as 3% to 8% in 2006 and 2015 respectively. The prevalence of serotype 12F increased from zero in 2006 and 2007 pre PCV-7 to 26% and 21% in 2011 and 2015 post PCV-7. The reverse of this distribution pattern was observed for serotypes 4, 6B, 18C and 19F which decreased from the pre-PCV introduction to the post PCV-7 introduction. The prevalence of serotype 14 decrease post PCV-7 vaccination but has re-emerged post PCV-13 introduction. Other non-vaccine types causing IPD in 2015 and 2016 were serotypes 9V and 23F (Figure 3.4).

Pneumococcal non-vaccine serotypes 12F and 35B were rarely detected in the PCV-7 era, but were subsequently observed in the post PCV-13 era, possibly as replacement serotypes (Figure 3.4).

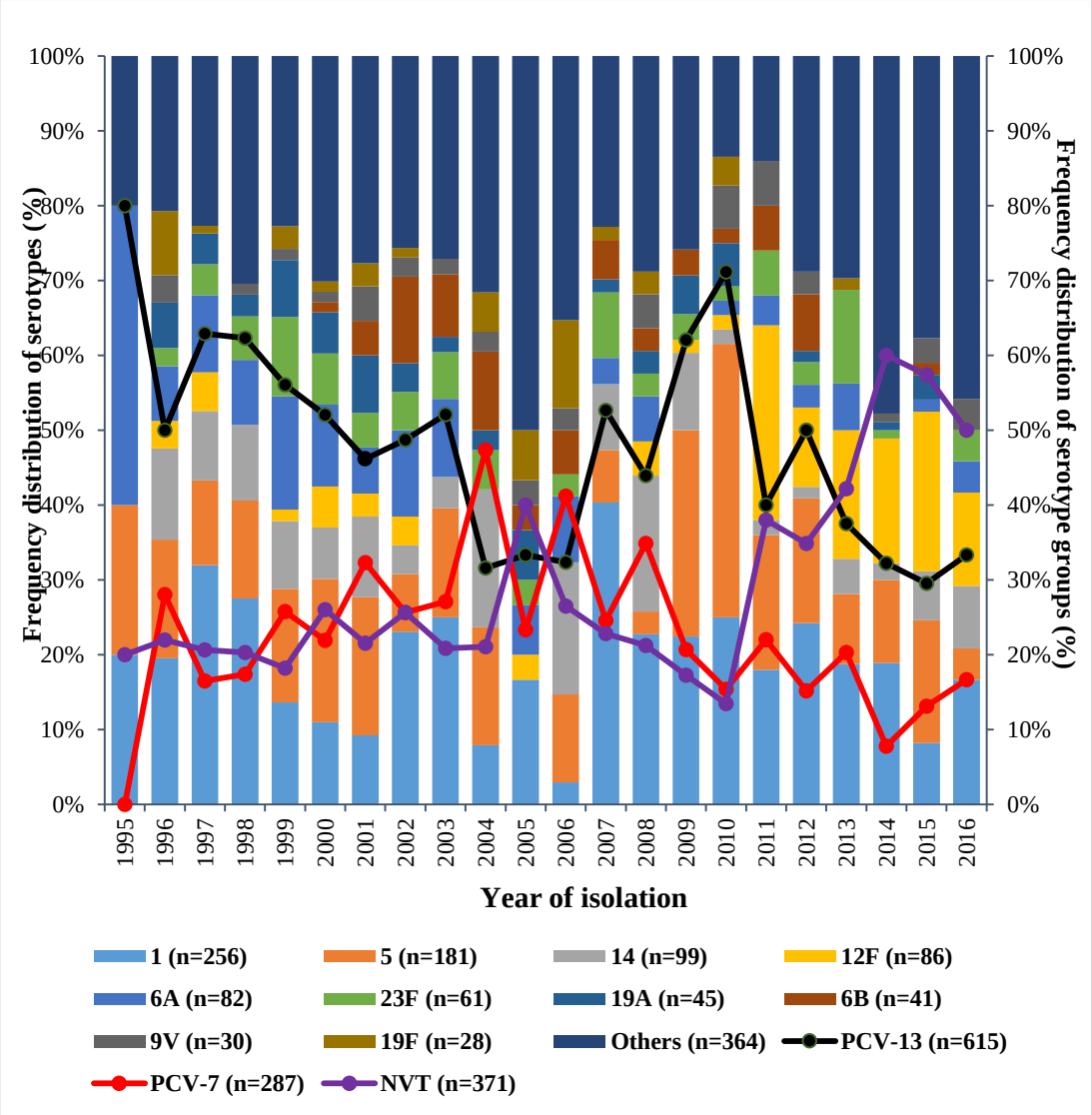

Figure 3.4 Yearly distribution of invasive isolates from The Gambia.
Each serotype is represented as a percentage of all invasive pneumococcal isolates obtained in each year as indicated on the primary Y-axis. PCV-7 vaccine serotypes, the additional six serotypes making up PCV-13 serotypes and the non-vaccine serotypes (NVT) are presented on the secondary Y-axis as a percentage of all invasive isolates yearly. PCV-7 was introduced in The Gambia national immunisation scheme in 2009 and PCV-13 in 2011.

The majority of pneumococcal isolates were obtained between December and May each year. Peaks in the numbers of isolates occurred from March to May (hot dry season) mostly as serotypes 1 and 5 (Figure 3.5).

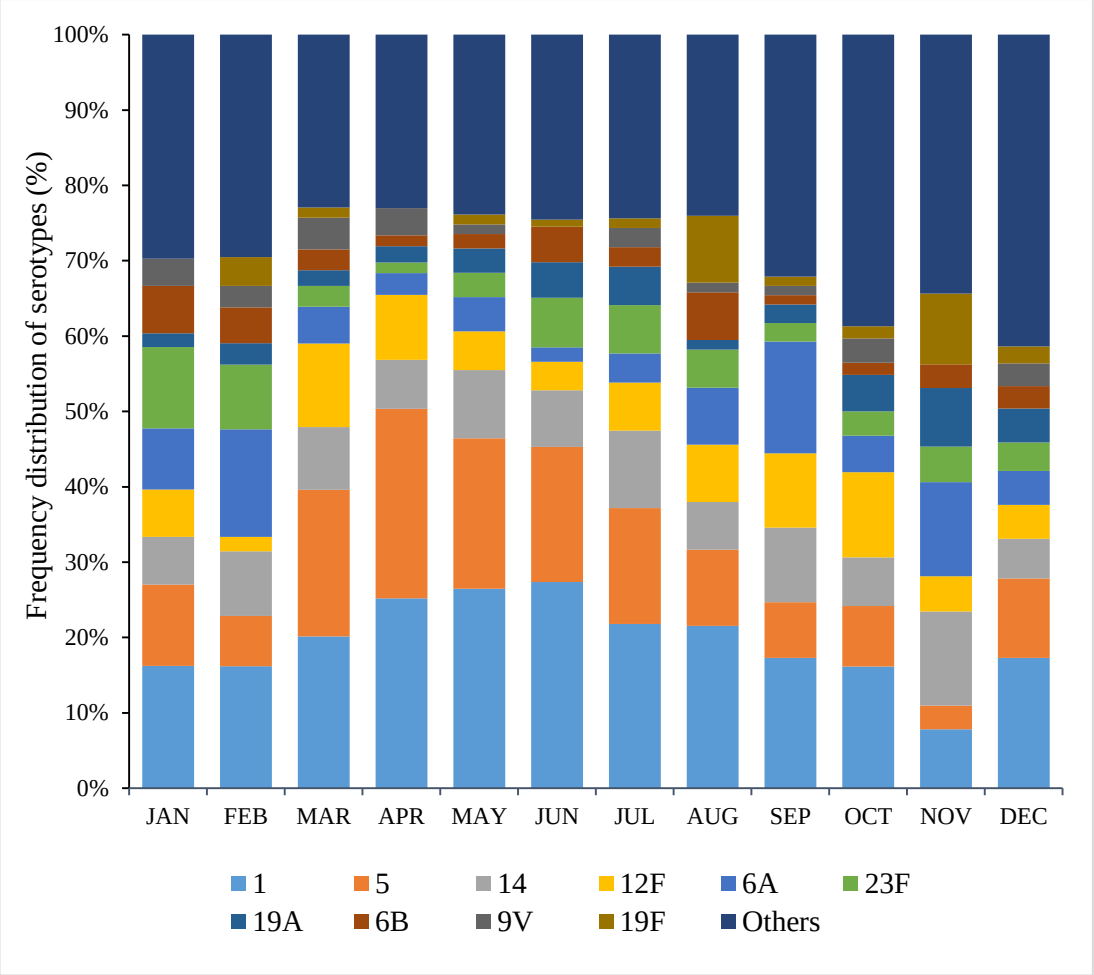

Figure 3.5 Monthly distribution of top 10 invasive isolates from The Gambia.
Each serotype is represented as a percentage of all invasive pneumococcal isolates obtained in each year as indicated on the X-axis.

3.3.4. Multi Locus Sequence Type analysis

3.3.4.1. Nasopharyngeal carriage

MLST analysis was conducted on 113 of the 1228 randomly selected pneumococcal carriage isolates collected from the URR. Sixty-six (66) different STs were obtained with ST925 (all serotype 19A) and ST913 (all serotype 6A) the leading STs observed (5.35% and 3.5% respectively). There were only two serotype 1 isolates found, both of which belonged to ST3081.

In the WR, MLST was conducted on 32 of 34 serotype 1 carriage isolates obtained. ST3081 (17/32) was the leading ST followed by ST618 (12/32). Only one ST217 isolate was found. All ST618 isolates were found in 2004 and 2007 while all ST3081 isolates were found in 2007 and 2008.

3.3.4.2. Invasive pneumococcal disease

MLST analysis was performed on 877 of 1273 randomly selected pneumococcal invasive isolates collected. Overall, 201 STs were found causing IPD with the leading STs ST3081_serotype 1 (11.40%, n=99/877), ST618_serotype 1 (10.72%, n=94/877), ST3404_serotype 5 (7.75%, n=68/877), ST989_serotype 12F (7.30%, n=64/877), ST289_serotype 5 (3.88%, n=34/877), ST802_serotype 23F (2.05%, n=18/877), ST847_serotype 19A (2.05%, n=18/877), ST63_serotype 14 (1.60%, n=14/877), ST217_serotype 1 (1.48%, n=13/877), ST3339_serotype 5 (1.37%, n=12/877) all accounting for 51% of all cases of IPD over the period 1995 to 2016 (Figure 3.6).

The distribution of ST3081, ST618 and ST217; the three predominant serotype 1 STs varied over the study period. ST618 was the predominant and stable ST associated with serotype 1 IPD from 1995- 2006 and was last observed in The Gambia in 2007. In contrast, ST3081 was first observed in The Gambia in 2007 and thereafter became the predominant ST of serotype 1. ST217, although observed in lower numbers was seen along with ST618 in the earlier years and with ST3018 in the more recent years (Figure 3.6). Distribution of the three leading pneumococcal serotype 5 STs, ST3404, ST289 and ST3339 varied from year to year. However, unlike serotype 1, all three STs occurred simultaneously for most of the time. Also, the prevalence of ST289 increased in the post PCV-13 era and ST3404 started to decline post PCV-7 (Table 3.6).

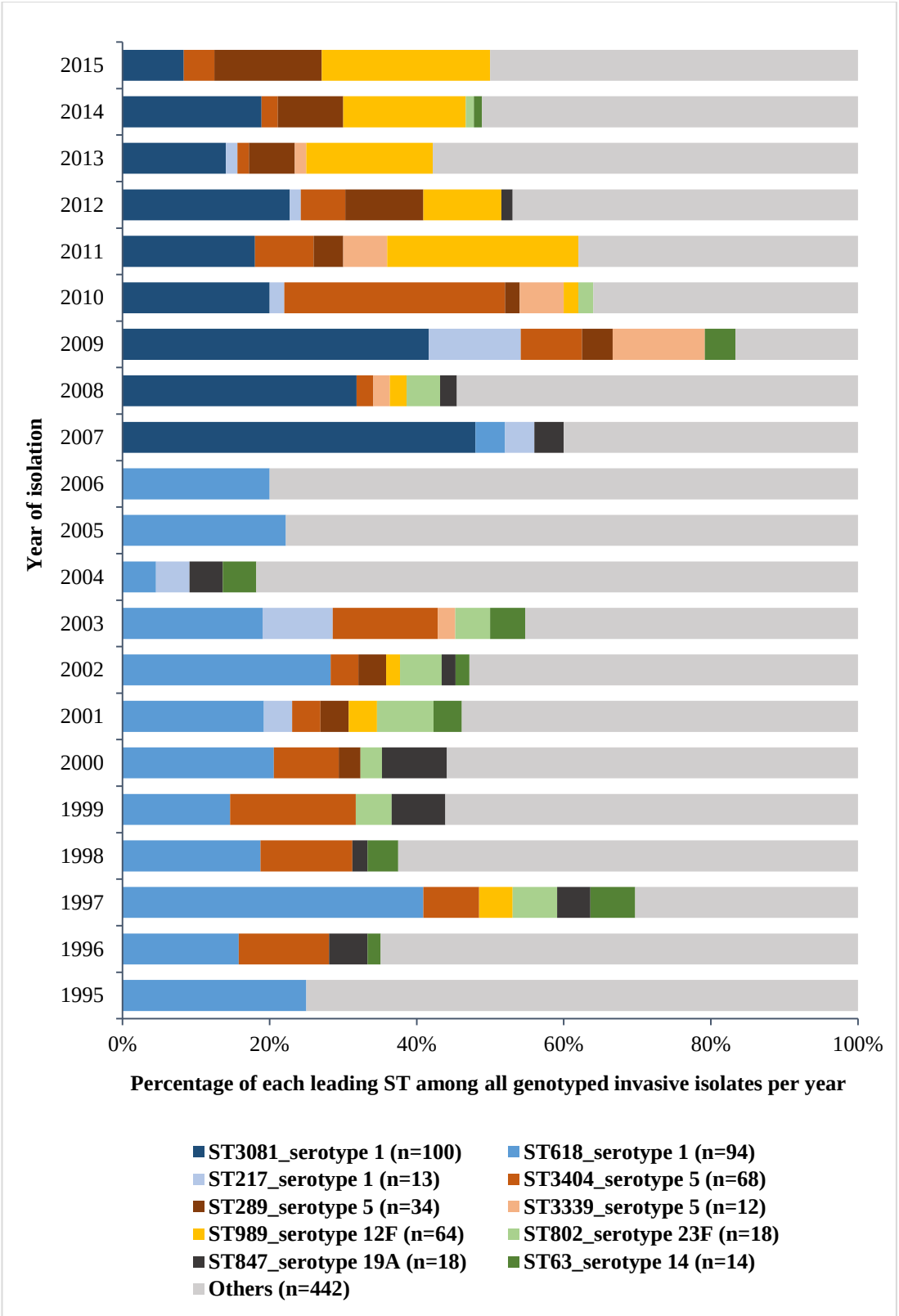

Figure 3.6 Yearly distribution of sequence types from The Gambia over two decades.

Each ST is represented as a percentage of all genotyped invasive pneumococcal isolates obtained in each year as indicated on the Y-axis. PCV-7 was introduced in The Gambia national immunization scheme in 2009 and PCV-13 in 2011.

The ST989, belonging to serotype 12F, was observed to increase in the post PCV-13 era, whilst ST847, ST63 and ST802, belonging to serotypes 19A, 14 and 23F respectively, were not observed post PCV-13 introduction. However, serotype 14 and 23F isolates observed in the post PCV-13 period were mostly due to ST2447 and ST1526 respectively (Table 3.6).

Table 3.6 Sequence types of pneumococcal serotypes in the pre and post introduction of PCV-7 and PCV-13 in The Gambia

Serotype	ST	Pre-PCV7	Post PCV7	Post-PCV13
1	217	10	1	2
	618	94	0	0
	3081	29	17	54
5	289	5	1	28
	3404	37	16	13
14	63	12	0	1
	3321	11	0	0
	2447	2	0	8
19A	847	17	0	1
	9769	0	0	1
23F	802	16	1	1
	1526	1	0	8
9V	280	5	2	3
	5719	0	0	3

Pre-PCV7, 1995 – August 2009; Post PCV-7, September 2009 – May, 2011; Post PCV-13, June 2011 – 2016.

Use of the minimum spanning tree showed a clear geographical clustering of the Gambian serotype 1 STs compared with serotype 1 STs from other parts of Africa and Asia obtained from the MLST database (http://pubmlst.org/spneumoniae/). In addition, these STs were different from STs obtained from the other parts of the world where the clonal complex 306 predominates (Figure 3.4). In contrast to the

pattern seen for serotype 1 STs, Gambian serotype 5 STs were seen to cluster together with STs from Africa and other parts of the world.

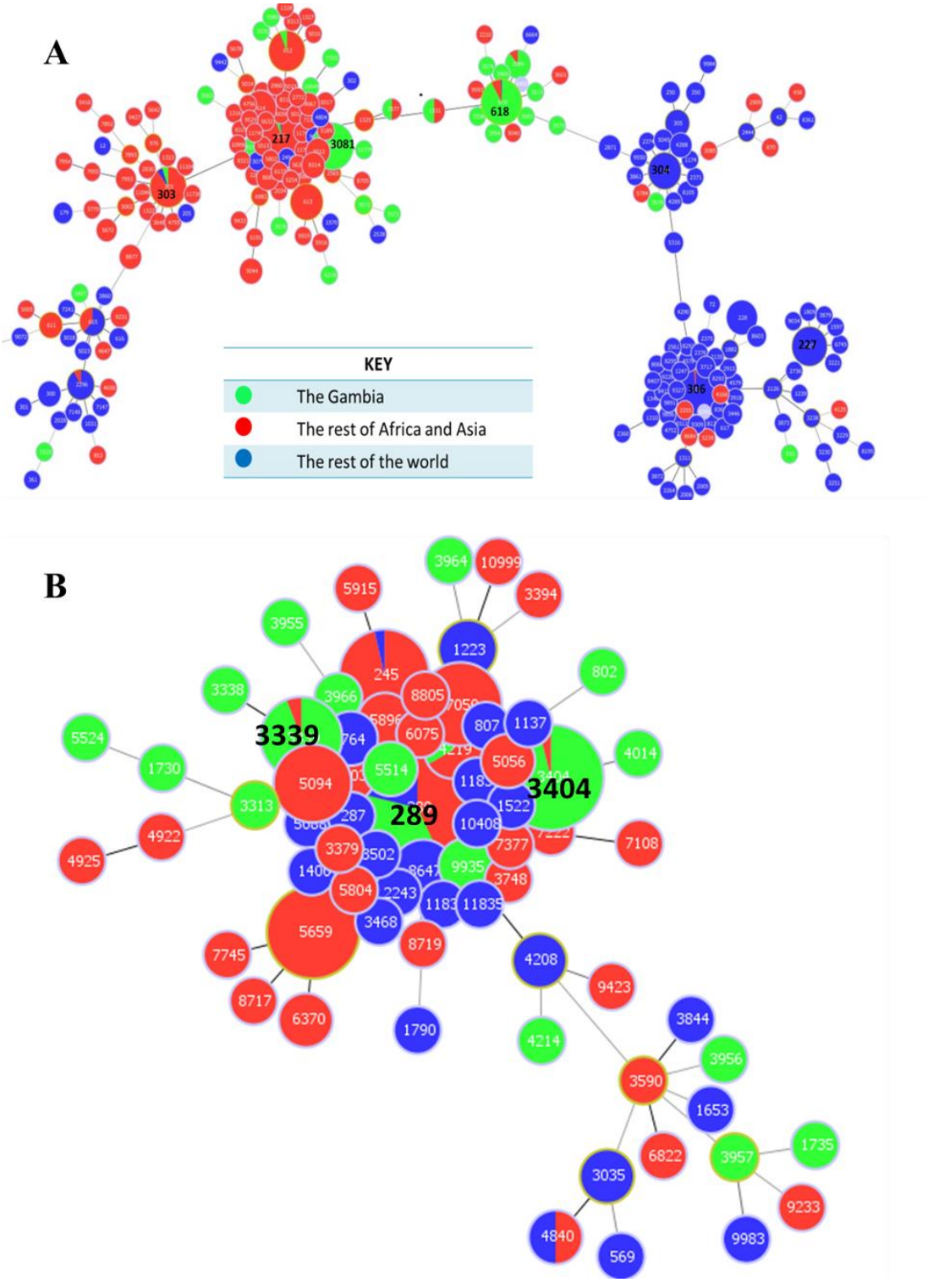

Figure 3.7 geoBURST analysis comparing Gambian pneumococcal serotype 1 (A) and serotype 5 (B) isolates to the rest of Africa, Asia and globally.
Each circle represents a single sequence type (ST). The area of each circle is proportional to the number of isolates included in the analysis.

3.4. Discussion

This study provides data on the prevailing serotypes and sequence types of *S. pneumoniae* in The Gambia over a 20 year period. Importantly, these isolates were collected during the pre, and for a few years, in the post-PCV-13 vaccination era, and as such provides valuable background data that will be relevant for a robust interpretation of any changes in the circulating pneumococcal serotypes observed following the introduction of routine vaccination with the pneumococcal conjugate vaccines in The Gambia.

Results from this study indicate that there was a reduction in IPD caused by vaccine type serotypes following the introduction of PCV, with the exception of serotypes 1 and 5. The reduction of vaccine serotypes after the introduction of PCVs has been reported elsewhere, including in other parts of Africa (Mackenzie et al., 2016, Cohen et al., 2016). Pneumococcal serotype 1 has been known to be one of the leading causes of IPD in The Gambia and other parts of West Africa. Given the introduction of PCV-13; a serotype 1 containing conjugate vaccine, in The Gambia it was unexpected that this strain persisted after vaccination. This might be related to the fact that serotype 1 has been reported to cause IPD mostly amongst older children and young adults than it does in children less than 5 years old (Kwambana-Adams et al., 2016, Brueggemann et al., 2013). Since PCV is given to infants at 2, 3 and 4 months of age, it is plausible that a longer post-PCV period of observation than has been covered by this study might be required to observe any appreciable effects of the vaccine in the older population who do not receive this vaccine but remain the age group that are most at risk for pneumococcal serotype 1 IPD. This finding may also point to the need for a review of current target age groups for pneumococcal vaccination, as has been reported in a few studies previously (Klugman et al., 2011).

In Ghana, despite the introduction of PCV-13 in the national immunization program in 2012, a deadly outbreak caused by serotype 1 was reported in 2016 with most of the cases seen among older children and young adults (Kwambana-Adams et al., 2016). Earlier trials of the 9-valent conjugate vaccine (PCV-9), a serotype 1-containing pneumococcal vaccine, in South Africa and in The Gambia both reported findings that were inconclusive on the vaccine's protection against serotype 1 disease (Klugman et al., 2011, Cutts et al., 2005).

All serotype 1 isolates in this study belonged to the ST217 clonal complex, the same clonal complex to which serotype 1 strains were isolated in other parts of Africa (Brueggemann and Spratt, 2003). Reports have shown that most serotype 1 isolates obtained from South and East Africa mostly comprise of ST217 (Brueggemann et al., 2013, du Plessis et al., 2016), with West African isolates mostly comprising of ST303 isolates (Leimkugel et al., 2005, Kwambana-Adams et al., 2016). In The Gambia, however, two major STs were observed to be circulating, ST618 which has been reported in other parts of West Africa and ST3081 reported only in neighbouring Senegal within Africa and in Oman in the MLST database. Our findings indicate that ST3081 emerged as a new dominant clone in 2007 prior to the introduction of PCVs in The Gambia and subsequently spread across the entire country, causing IPD. In addition, ST3081 was seen to have replaced ST618, which had been the dominant lineage among serotype 1 isolates circulating in The Gambia for over a decade prior. This change in dominant genotypes was also observed among serotype 1 isolates obtained from the nasopharynx of healthy individuals in The Gambia with the time period of the switch also seen to be similar (Ebruke et al., 2015), an indication that this change had occurred both in invasive disease and carriage states as well as across both rural and urban settings in The Gambia. It may

be that the observed change in dominant circulating genotypes is reflecting adaptive changes by the pneumococcus in response to changes in the human population and human activity, which yet remain unclear.

The clonal replacement observed in The Gambia has also been reported in Brazil where temporal switching between two lineages of serotype 1 and wide dispersion of a Swedish clone was observed (Chiou et al., 2008). A possible explanation for this replacement in dominant clones could be genetic recombination or point mutations as occurs in the evolution of certain microbes, with recombination the more likely scenario in the case of *S. pneumoniae* (Henriques-Normark et al., 2008, Feil et al., 2000, Spratt et al., 2001). This could be attributed to particular characteristics of the host community, such as antibiotics usage, vaccine selective pressure or a virulence advantage of the pneumococcal lineage (Spratt and Greenwood, 2000). A study conducted in Portugal among healthy carriers revealed an emergence and expansion of a serotype 1 lineage soon after the introduction of PCV-7 (Nunes et al., 2008). However, a carriage study of serotype 1 in the context of a PCV-7 vaccine trial in The Gambia did not show substantial evidence that the observed changes in genotype were as a result of selective pressure through vaccination (Ebruke et al., 2015). The effect of vaccination as a contributing factor to the observed changes in genotypes of serotype 1 remains to be proven. It becomes critically important, therefore to exploit the use of more informative techniques such as whole genome sequencing to determine the absolute genetic differences between these two STs and thus improve our understanding of the key mechanisms involved in clonal replacement.

Results from this study show that pneumococcal serotype 1 has a geographical structure with strains from Africa and Asia clustering together and that these strains

cluster differently from pneumococcal serotype 1 strains from other parts of the world, as has been reported in some other studies (Brueggemann and Spratt, 2003). A short duration of nasopharyngeal carriage of pneumococcal serotype 1 strains, resulting in reduced spread of the organism during travel across geographical regions, may be a possible explanation for the observed geographic clustering. It would appear that among *S. pneumoniae* serotype 1 strains in West Africa, there is an expansion of the ST217 clonal complex, as has been reported not only in The Gambia, but also in Ghana, Niger and Burkina Faso. This spread seems to be responsible for the deadly epidemics observed in these geographical areas (Antonio et al., 2008, Leimkugel et al., 2005, Yaro et al., 2006), including a recent outbreak of pneumococcal serotype 1 meningitis in 2016 in Ghana (Kwambana-Adams et al., 2016). This study provides critical background data that will be useful in interpreting any subsequent observed changes in the circulating pneumococcal serotype 1 strains and thus help improve our knowledge of the patterns of spread and the phylogenetic structure of this serotype.

Findings from this study show that pneumococcal serotype 5, like serotype 1, is a major cause of IPD in The Gambia and rarely found in nasopharyngeal carriage is consistent with reports from other studies (Brueggemann et al., 2013). All serotype 5 isolates included in this study belonged to the ST289 clonal complex dominated by ST289 and ST3404. In contrast to serotype 1, serotype 5 STs fluctuated throughout the study period and at 2016 both STs were found co-occurring. Additionally, unlike the predominant serotype 1 STs that showed a clear geographical clustering, the same dominant serotype 5 ST289 has been reported in various regions of the world and across different continents. Further studies of serotype 5 will be useful in

understanding the reasons behind these differences in distribution between these two important pneumococcal serotypes and the public health relevance.

Another noteworthy finding from this study is that different genotypes of vaccine serotypes emerged post vaccination. This may suggest that within a given serotype there is significant genetic and immunogenic variation. Among serotype 14 isolates, ST63 dominated pre-vaccination but ST2447, a SLV of ST63, now dominates post vaccination. Similarly, among serotype 23F, ST802 was occurring mostly pre PCV-7 vaccination, however post-vaccination, ST1526 was the predominant ST. Given the fact that current PCVs in use has been shown to be effective against serotypes included in the vaccine, it is not known to what extent the changes in the circulating STs might impact on the effectiveness of currently used pneumococcal conjugate vaccines.

Replacement of vaccine serotypes with non-vaccine serotypes following introduction of pneumococcal conjugate vaccines has been noted as one of the outcomes of PCV use and this effect has been reported in several countries (Janoir et al., 2016, Del Amo et al., 2016). In the USA, serotype 19A became a major replacement serotype in the Post PCV-7 era (Pilishvili et al., 2010). In this study, we report that pneumococcal serotype 12F emerged as an important replacement serotype. In particular, this serotype was hardly seen in the pre-vaccination period but subsequently became one of the leading causes of IPD amongst study isolates in the post-PCV period. Similar to pneumococcal serotype 1, serotype 12F was rarely detected in nasopharyngeal carriage and all isolates belonged to a single dominant clonal complex, ST989. Reports from other parts of the world have shown that serotype 12F is able to cause epidemic outbreaks similar to serotype 1 (Deng et al., 2016). This is of particular importance given the fact that serotype 12F is not

included in PCV13 currently in routine use in The Gambia. In addition, this study shows the emergence of non-vaccine type serotype 35B which has also been reported in other parts of the world (Olarte et al., 2017). These findings on emergence of non-vaccine pneumococcal serotypes that have epidemic potential may suggest that the impact of currently used PCV may be far from optimal.

Findings from this study highlight the need for the continuous population-based post-vaccine surveillance of the circulating pneumococcal serotypes to provide a robust evidence base for the timely review of existing policies on the prevention and control of pneumococcal disease. New vaccines that contain coverage against the predominant circulating serotypes are needed. In addition, given the importance of serotypes 1, 5 and 12F, emphasis on the closer monitoring of any changes in the molecular epidemiology of these strains would be helpful for refinements to be made in appropriate intervention guidelines that will aid better preparedness for epidemic outbreaks.

4. Antibiotic susceptibility patterns of *Streptococcus pneumoniae* in The Gambia from pre-and post-introduction of pneumococcal conjugate vaccines

4.1. Introduction

Antibiotic resistance is a global problem and is now gaining more attention worldwide (Laxminarayan et al., 2013, Laxminarayan et al., 2016). The increasing global attention this problem has attracted is due in part to a growing recognition of its potential for severe adverse effects on health systems including reductions in the effectiveness of existing treatments for severe illnesses with associated increasing deaths. In resource poor countries, the challenge with antibiotic resistance is particularly worrying as unavailability of alternative drugs are often unaffordable by majority of those affected. This frequently leads to delays in treatment that further worsen treatment outcomes. For *S. pneumoniae*, the problem of antibiotic resistance is of particular concern due to the disproportionately heavy burden of pneumococcal disease and associated higher mortality rates in developing countries in comparison to developed countries (O'Brien et al., 2009).

Antibiotic resistance in *S. pneumoniae* isolates is mediated through mechanisms employed by the bacteria to evade the effects of antibiotic treatment; some of these mechanisms have been discussed briefly in section 1.13.1. Antibiotic resistance in *S. pneumoniae* was first described in the 1960s and has since evolved over time (Hansman, 1967). The PMEN network was set up to monitor global patterns of antibiotic resistance among *S. pneumoniae* isolates (McGee et al., 2001).

The introduction of pneumococcal vaccines has been shown to reduce resistance among pneumococcal isolates (Kyaw et al., 2006). Serotypes included in the 1st

licenced conjugate vaccine PCV-7 were largely responsible for multidrug resistance (Whitney et al., 2000). The reduction in carriage and transmission of these vaccine serotypes that followed introduction of PCV-7, along with reductions in antibiotic usage, led to a corresponding decrease in the prevalence of antibiotic resistance among circulating pneumococcal isolates(Kyaw et al., 2006). However, with serotype replacement, non-vaccine serotypes subsequently emerged, some of which were antibiotic resistant, such as the post PCV-7 emergence of serotype 19A in the USA (Pelton et al., 2007). The introduction of PCV-13, with its wider valency including additional six serotypes, would appear to have stemmed this problem (Tomczyk et al., 2016, Dagan et al., 2015). However, as the possibility that emerging non-vaccine serotypes could develop resistance to antibiotics remains, the problem of antibiotic resistance continues to pose a challenge for healthcare systems.

It is therefore important to continue to monitor the effect of vaccination and antimicrobial resistance globally and at country levels to help inform effective strategies for prevention and control of antibiotic resistance. PCV-7 was introduced in The Gambia in 2009, with a switch to PCV-13 in 2011. In the previous chapter, described the serotypes and genotypes among invasive and carriage isolates in The Gambia collected over two decades over the pre- and post- PCV era.

The aim of this chapter is to describe the antibiotic resistant patterns and assess the impact of vaccination on the prevalence of antibiotic-resistant *S. pneumoniae* in nasopharyngeal carriage and invasive disease isolates in rural and urban Gambia. In addition, to determine if the emerging non-vaccine serotypes are susceptible or resistant to antibiotics used commonly in The Gambia. This aim would be used to test the hypothesis that the introduction of PCVs would result in reductions in levels

of antibiotic resistance among pneumococcal isolates circulating in The Gambia as the known antibiotic resistant strains are included in the vaccine.

4.2. Methods

Parent studies and sources of pneumococcal isolates used and analysed in this chapter have been described in section 3.2 of this thesis. These studies were conducted in The Gambia over the period 1995- 2015.

Briefly, invasive pneumococcal isolates were obtained from a hospital-based pneumococcal surveillance at the MRC hospital in the Western region of The Gambia and from a population-based surveillance of IPD in the Upper and Central River Regions investigating effectiveness of the introduction of PCV in The Gambia. Nasopharyngeal carriage isolates included in this chapter were obtained from large pneumococcal carriage studies conducted on healthy participants residing in 21 villages in the Western region of The Gambia.

As part of these studies, antibiotic susceptibility testing was conducted on these isolates by the MRC Unit, The Gambia, of which I contributed to. Antibiotic susceptibility testing was not performed on carriage isolates collected from the Upper River Region and so these isolates were excluded from analyses done in this chapter.

Study isolates were tested by the Kirby-Bauer disk diffusion and Epsilometer (E-test) agar diffusion methods for pneumococcal isolates following the Clinical and Laboratory Standards Institute (CLSI) guidelines (CLSI, 2017) as described in section 2.3.3. Antibiotics tested by disk diffusion included, chloramphenicol (30 µg), erythromycin (5 µg), cotrimoxazole (1.25 µg/23.75 µg) and tetracycline (30 µg). E-

test to determine MICs was performed for penicillin and cefotaxime. All results were interpreted using the CLSI guidelines as summarized in the table 4.1 (CLSI, 2017).

Table 4.1 Clinical and Laboratory Standards Institute interpretative charts of pneumococcal disk diffusion zone size and minimum inhibitory concentration

	Sensitive	**Intermediate**	**Resistant**
Disk diffusion	**mm**	**Mm**	**mm**
Chloramphenicol	≥ 21	-	≤ 20
Erythromycin	≥ 21	16-20	≤ 15
Co-trimoxazole	≥ 19	16-18	≤ 15
Tetracycline	≥ 28	25-27	≤ 24
MIC	**µg/ml**	**µg/ml**	**µg/ml**
Penicillin	≤0.06	0.12-1	≥2
Cefotaxime	≤ 1	2	≥4

4.3. Results

4.3.1. Epidemiological background of study isolates

Of 1,290 invasive pneumococcal isolates collected from the Western and Upper River Regions of The Gambia over the study period, antibiotic susceptibility testing was performed on 1055 (82%) isolates and analysed in this study. Characteristics of the patients from whom these isolates were collected are shown in Table 4.2.

Table 4.2 Summary of patient characteristics from whom invasive isolates were collected from 1995 – 2015 and included in this study

Characteristics	Category	Isolates N (%)
Age group (years)	<5 years	540 (51.2)
	5 - <15 years	128 (12.1)
	≥15 years	257 (24.4)
	Unknown	130 (12.3)
Gender	Female	393 (37.3)
	Male	572 (54.2)
	Unknown	90 (8.5)
Specimen	Blood	769 (72.9)
	CSF	132 (12.5)
	LA	94 (8.9)
	Others	60 (5.7)
Total		1055

CSF, cerebrospinal fluid; LA, Lung aspirate

Of 3,807 pneumococcal carriage isolates obtained from studies conducted in the Western Region of The Gambia, antibiotic susceptibility testing was available for 2,884 isolates (76%) and analysed in this study. The percentage of males and females carrying *S. pneumoniae* were 49.7% (1434/2884) and 50.2% (1449/2884) respectively.

4.3.2. Antibiotic susceptibility patterns among invasive pneumococcal isolates

Pneumococcal invasive isolates collected from patients with IPD in The Gambia were highly susceptible to four of the six antibiotics tested; cefotaxime (100%),

penicillin (99.8%), erythromycin (96.4%) and chloramphenicol (81.9%). Reduced susceptibility was observed for tetracycline (40.8%) and co-trimoxazole (26.2%) (Table 4.3).

Table 4.3 Overall distribution of antimicrobial susceptibility patterns of invasive pneumococcal isolates in The Gambia from 1995 – 2015 (n =1055 [for erythromycin, n=473])

Antibiotics	Resistant n (%)	Intermediate n (%)	Susceptible n (%)
Co-trimoxazole	715 (67.77)	64 (6.07)	276 (26.16)
Chloramphenicol	191 (18.1)	-	864 (81.9)
Tetracycline	493 (46.73)	132 (12.51)	430 (40.76)
Erythromycin	7 (1.48)	10 (2.11)	456 (96.41)
Penicillin	0	2 (0.19)	1053 (99.81)
Cefotaxime	0	0	1055 (100)

Disk diffusion test was performed for co-trimoxazole, chloramphenicol, tetracycline and erythromycin. E-test was done for Penicillin and cefotaxime

Monitoring the resistance patterns of invasive isolates from The Gambia over the pre and post PCV era, resistance to co-trimoxazole appeared to increase over time from a prevalence of 60.77% in the pre-PCV era to 95.15% in the post-PCV13 era (Table 4.4). Whereas for tetracycline, chloramphenicol and erythromycin, a fluctuating pattern with a slight decrease from the pre-PCV 7 era to post introduction of PCV-7 and subsequently increasing in the post PCV-13 era (Table 4.4) was observed. The only two isolates with intermediate resistance to penicillin were found in the Pre-PCV7 and Post-PCV 7 era.

Table 4.4 Antimicrobial resistance of invasive pneumococcal isolates distributed by study periods (pre-PCV-7, post-PCV-7 and post PCV-13) in The Gambia

	Number (%) of resistant invasive isolates

Antibiotics	Pre-PCV7 (n=831)	Post-PCV7 (n=59)	Post-PCV13 (n=165)
Co-trimoxazole			
Intermediate	62 (7.46)	2 (3.39)	0
Resistant	505 (60.77)	53 (89.83)	157 (95.15)
Chloramphenicol			
Resistant	145 (17.45)	9 (15.25)	37 (22.42)
Tetracycline			
Intermediate	99 (11.91)	11 (18.64)	22 (13.33)
Resistant	396 (47.65)	22 (37.29)	75 (45.45)
Erythromycin	Pre-PCV7 (n=249)	Post-PCV7 (n=59)	Post-PCV13 (n=165)
Resistant	2 (0.80)	1 (1.69)	4 (2.42)

Pre-PCV7, 1995 – August 2009; Post PCV-7, September 2009 – May, 2011; Post PCV-13, June 2011 – 2015.

Stratifying by vaccine types, no significant differences were found between PCV-7, PCV-13 and non-vaccine serotypes for resistance to co-trimoxazole and erythromycin (p=0.086 and 0.606 respectively). However, prevalence of resistance to chloramphenicol was significantly higher among NVT, and resistance to tetracycline significantly higher among PCV-7 serotypes, in comparison to other vaccine types (p = 0.002 and p<0.001, respectively) (Table 4.5).

Table 4.5 Antimicrobial resistance of invasive isolates (n =1055 [for erythromycin, n=473]) stratified by PCV vaccine and non-vaccine types

	Number (%) of resistant invasive isolates		
Antibiotics	PCV-7 (n=255)	PCV-13 (n=544)	NVT (n=256)
Co-trimoxazole			
Intermediate	20 (7.84)	24 (4.41)	20 (7.81)
Resistant	170 (66.67)	384 (70.59)	161 (62.89)

Chloramphenicol			
Resistant	33 (12.94)	95 (17.46)	63 (24.61)
Tetracycline			
Intermediate	19 (7.45)	88 (16.18)	25 (9.77)
Resistant	156 (61.18)	214 (39.34)	123 (48.05)
Erythromycin	PCV-7 (n=227)	PCV-13 (n=118)	NVT (n=128)
Resistant	2 (1.69)	2 (0.88)	3 (2.34)

PCV-7 vaccine serotypes (PCV-7), the additional six serotypes making up PCV-13 serotypes (PCV-13) and the non-vaccine serotypes (NVT) are all other serotypes not contained in PCV-13.

4.3.3. Yearly antibiotics susceptibility patterns

The year by year data for co-trimoxazole resistance showed a noticeable increase in 2010 post introduction of PCV-7 and subsequently remained high over the rest of the period under review. A substantial increase in chloramphenicol resistance was observed in 2002 and 2003, prior to the introduction of PCVs in The Gambia which was followed by lower resistance levels over the subsequent years. Resistance to tetracycline appeared to fluctuate from year to year throughout the study period, mostly between 20% - 60% prevalence rates whilst resistance to erythromycin was generally very low (below 10%) all through the review period with no erythromycin-resistant isolates seen in 2014 and 2015 (Figure 4.1).

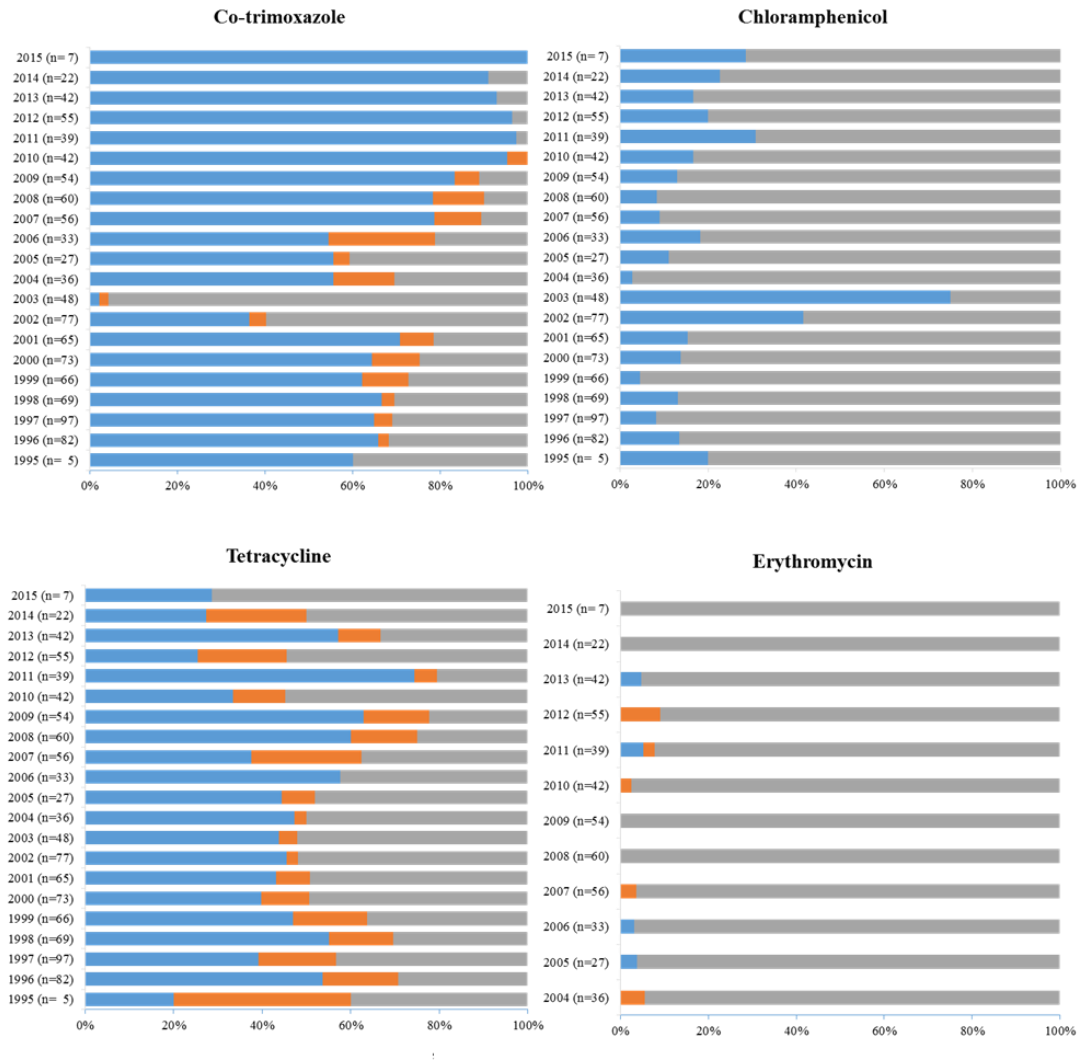

Figure 4.1 Yearly antimicrobial resistance patterns among 1055 invasive study isolates collected from 1995- 2015.
For erythromycin, antibiotics data included was from 2004 -2015, n=473.
PCV-7 was introduced in The Gambia national immunization scheme in 2009 and PCV-13 in 2011.
The bar colours blue, orange and grey represent percentage of resistant, intermediate and resistant isolates respectively per year.

4.3.4. Antibiotic resistance patterns by pneumococcal serotypes

Antibiotic resistance patterns for the leading pneumococcal serotypes seen among invasive isolates in The Gambia as described in section 3.3.3, all of which are included in PCV-13 except serotype 12F, are shown in Figure 4.2. Pneumococcal serotype 7F showed high levels of resistance to all but one of the antibiotics tested. Of 25 serotype 7F isolates analysed, 96% were resistant to tetracycline and co-

trimoxazole, while 72% were resistant to chloramphenicol. Of 147 serotype 5 isolates tested, 90%, 26.4% and 12.2% were resistant to co-trimoxazole, tetracycline and chloramphenicol respectively. In addition, the two isolates that showed intermediate resistance to penicillin were both serotype 5 isolates (not shown in Fig 4.2). Of 234 pneumococcal serotype 1 isolates tested, 61.1%, 34.6% and 12% were resistant to co-trimoxazole, tetracycline and chloramphenicol respectively. Resistance to erythromycin was low (<10%) among all the leading pneumococcal serotypes (Fig. 4.2). Of 110 study isolates tested for erythromycin, two serotype 1 isolates were resistant.

Among NVTs, resistance among serotype 12F isolates (n=50) were 92%, 88% and 66% to co-trimoxazole, tetracycline and chloramphenicol respectively. Other non-vaccine serotypes not shown in Figure 4.2 include serotype 46, with 70.6%, 70.6% and 41.8% of 17 isolates tested found to be resistant to co-trimoxazole, tetracycline and chloramphenicol respectively.

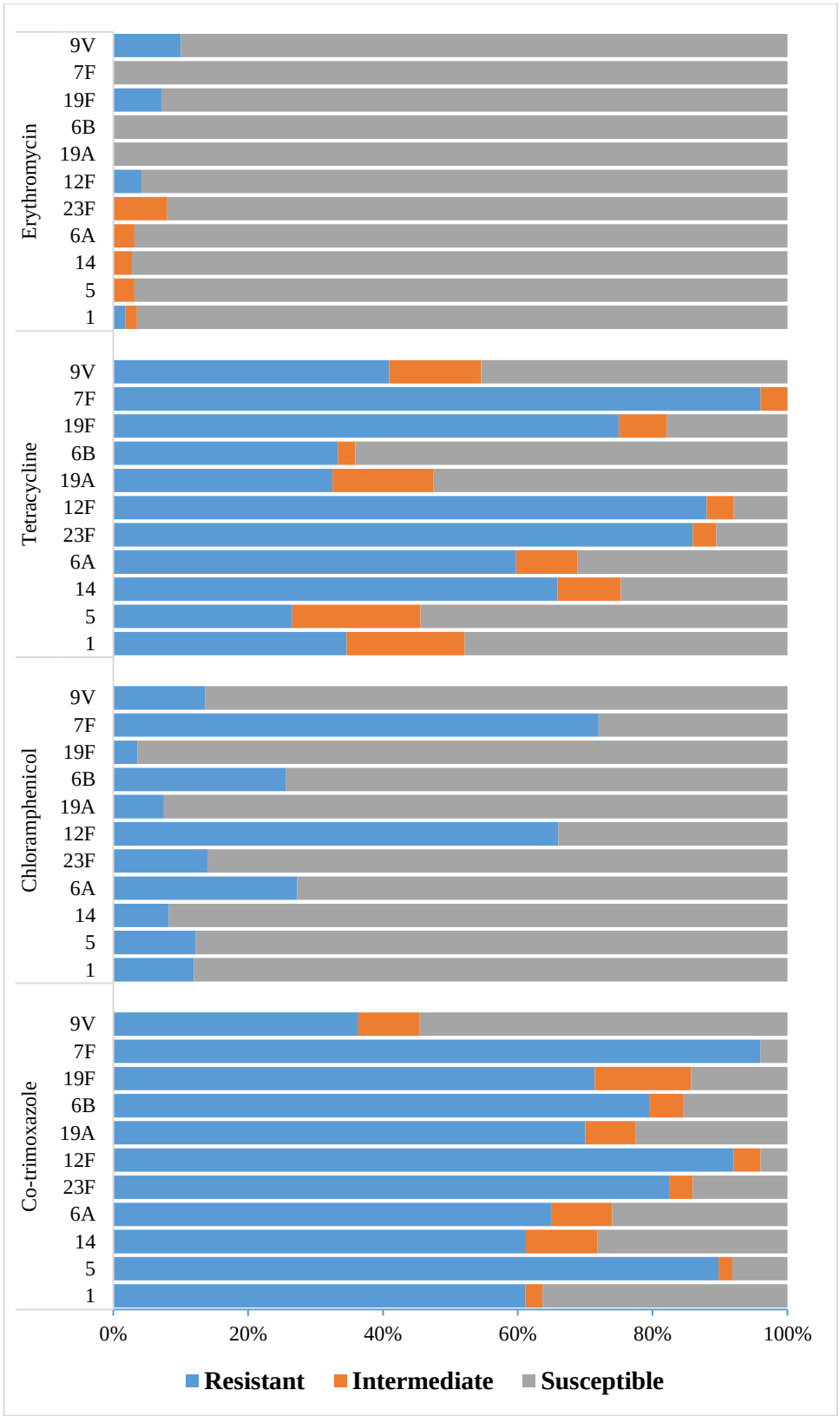

Figure 4.2 Antibiotics resistance patterns of the leading pneumococcal serotypes causing IPD in The Gambia.

4.3.5. Antibiotic resistance patterns by pneumococcal sequence types

Antibiotic resistant patterns by sequence type is shown in figure 4.3 for the leading STs among invasive isolates in The Gambia.

The two leading STs observed in this study were ST618 and ST3081, both of which are serotype 1 STs, with our report in the previous chapter indicating that ST3081 had replaced ST618 as the leading serotype 1 ST circulating in The Gambia. In this study, resistance patterns to chloramphenicol and tetracycline were similar for both STs. However, 93.9% (77/82) of ST3081 isolates were resistant to co-trimoxazole in comparison to only 26.6% (25/94) of ST618 isolates (Figure 4.3). Yearly distribution of serotype 1 STs in The Gambia showed that from 1995 – 2005, most serotype 1 isolates had high susceptibility to co-trimoxazole, mostly due to ST618. However, from 2006 co-trimoxazole resistance among serotype 1 isolates were mostly due to ST3081 (Figure 4.4). In addition, both serotype 1 isolates found to be resistant to erythromycin were ST3081 isolates (Figure 3). Among ST217 isolates, a serotype 1 genotype observed to be occurring during periods of occurrence of both ST618 and ST3081, all isolates were resistant to tetracycline (100%). In addition, resistance to chloramphenicol among ST217 isolates was much higher than was observed among ST618 and ST3081 isolates (Figure 4.3).

Among the serotype 5 predominant STs, ST289 and ST3404, resistance to co-trimoxazole was high, but was low for chloramphenicol and tetracycline. In addition, one isolate each of these two STs showed intermediate resistance to erythromycin. Among ST847 isolates, the predominant serotype 19A ST in The Gambia, all isolates were resistant or showed intermediate resistance to co-trimoxazole, but all isolates belonging to this ST were sensitive to chloramphenicol. Among serotype 14 STs, all ST63 were resistant to tetracycline while ST3321 isolates showed only

intermediate resistance. Both ST802 of serotype 23F and ST3324 of serotype 6A, showed high resistance to co-trimoxazole and tetracycline but lower for chloramphenicol. However, only one 23F ST1526 isolate showed intermediate resistance to erythromycin.

All serotype 12F isolates included in this study belong to ST989 in The Gambia and so results by ST is similar to what was reported in section 4.3.4.

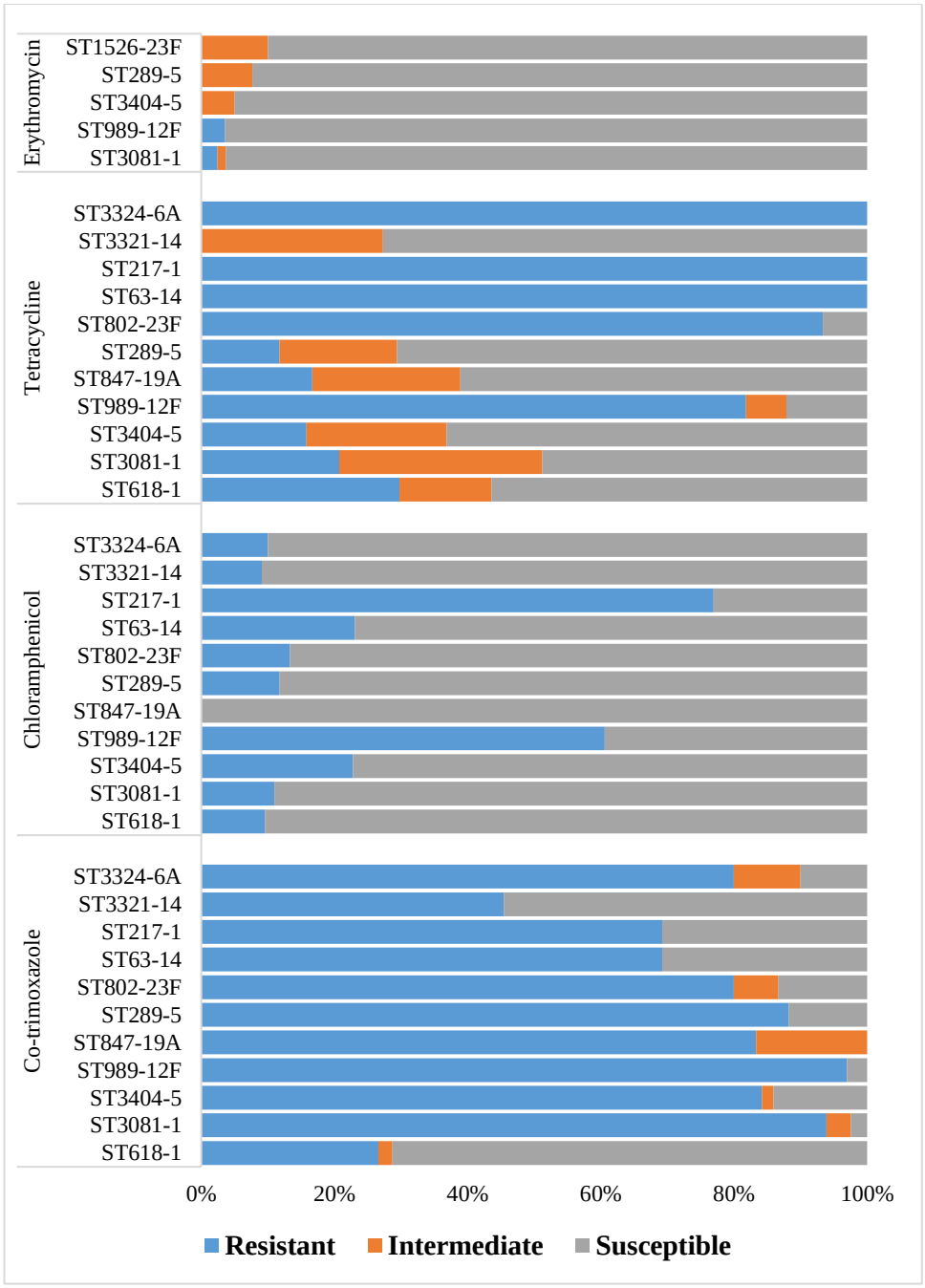

Figure 4.3 Antibiotics resistance patterns of the leading pneumococcal sequence types causing IPD in The Gambia.

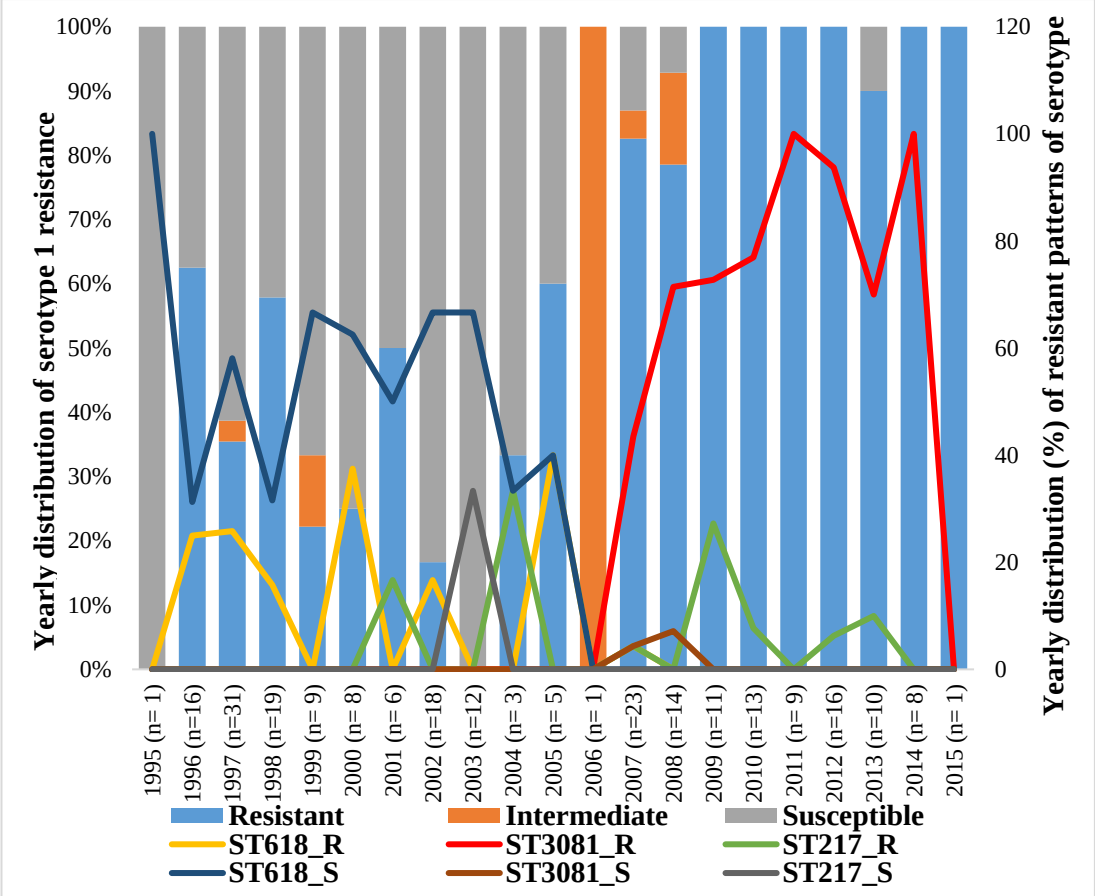

Figure 4.4 Yearly antimicrobial resistance patterns among *S. pneumoniae* serotype 1 study isolates collected from 1995- 2015
PCV-7 was introduced in The Gambia national immunization scheme in 2009 and PCV-13 in 2011.

4.3.6. Antibiotic susceptibility patterns among pneumococcal carriage isolates

As was observed among pneumococcal invasive isolates, nasopharyngeal carriage isolates were highly susceptible to erythromycin (91.2%) and chloramphenicol (88.3%). Reduced susceptibility was observed for tetracycline (41.5%) and co-trimoxazole (33.4%) (Table 4.6).

Table 4.6 Antibiotic susceptibility patterns among pneumococcal carriage isolates (n=2,884)

Antibiotics	Resistant N (%)	Intermediate N (%)	Susceptible N (%)
Co-trimoxazole	1,551 (53.78)	369 (12.79)	964 (33.43)
Chloramphenicol	338 (11.72)	-	2,546 (88.28)
Tetracycline	1,282 (44.45)	404 (14.01)	1,198 (41.54)
Erythromycin	40 (1.39)	215 (7.45)	2,629 (91.16)

4.3.7. Antibiotic resistance patterns among carriage isolates stratified by pneumococcal serotypes

Stratifying by pneumococcal serotypes (Table 4.7), resistance patterns among carriage isolates included in this study were similar to those for pneumococcal invasive isolates reported above. Across all 26 carriage serotypes represented, resistance to co-trimoxazole or tetracycline was consistently high or moderate. Among serotypes 1 and 5 carriage isolates included in this study, 79% and 64%, respectively, were resistant to co-trimoxazole, with little (4%) or no (0%) resistance to chloramphenicol and erythromycin respectively. In contrast, serotype 12F isolates demonstrated high levels of resistance to three of the four antibiotics tested (Tetracycline 100%, co-trimoxazole 83% and chloramphenicol 83%). Similarly, among serotype 7F isolates, high to moderate levels of resistance to these three antibiotics was observed, with 76%, 56% and 32% resistance prevalence to Tetracycline, co-trimoxazole and chloramphenicol respectively. (Table 4.7).

Table 4.7 Antibiotic Susceptibility of pneumococcal carriage isolates from The Gambia stratified by serotypes

		Chloramphenicol		Erythromycin			Co-trimoxazole			Tetracycline		
Serotype	Total (n)	R (%)	S (%)	R (%)	I (%)	S (%)	R (%)	I (%)	S (%)	R (%)	I (%)	S (%)
1	28	4	96	0	4	96	79	7	14	32	43	25
2	7	29	71	0	14	86	29	14	57	57	0	43
3	213	11	89	1	5	93	31	5	64	39	14	47
4	75	15	85	1	13	85	35	17	48	31	29	40
5	14	0	100	0	7	93	64	21	14	43	0	57
11	134	11	89	3	9	88	60	12	28	31	22	46
12	18	83	17	0	22	78	83	11	6	100	0	0
14	52	8	92	0	12	88	69	17	13	73	8	19
25	3	0	100	0	0	100	33	0	67	0	100	0
34	113	6	94	1	6	93	48	29	23	17	23	60
46	6	17	83	0	0	100	100	0	0	100	0	0
6A	166	17	83	1	13	87	67	22	11	70	4	26
6B	126	13	87	0	4	96	79	15	6	34	21	45
7F	25	32	68	0	0	100	56	28	16	76	16	8
9A	21	5	95	0	0	100	95	0	5	48	19	33
9L	26	8	92	0	0	100	54	19	27	15	31	54
9V	76	20	80	1	16	83	93	3	4	37	17	46
NT	128	10	90	5	4	91	48	8	44	48	7	45
10A/C	62	0	100	0	2	98	27	13	60	15	15	71
15B/C	105	8	92	1	5	94	89	4	8	34	14	51
16A/F	66	0	100	0	3	97	73	6	21	77	3	20
18C	56	13	88	2	9	89	36	21	43	55	9	36
19A	87	7	93	1	7	92	68	18	14	47	11	41
19F	119	10	90	0	7	93	83	9	8	71	8	20
23F	117	14	86	3	11	85	78	11	11	85	5	10
35B/C	95	9	91	2	15	83	29	14	57	31	18	52

Table 4.8 Antimicrobial resistance of nasopharyngeal carriage isolates (n =2884) stratified by PCV vaccine and non-vaccine types

	Number (%) of resistant carriage isolates		
Antibiotics	**PCV-7 (n=621)**	**PCV-13 (n= 533)**	**NVT (n= 1730)**
Co-trimoxazole			
Intermediate	79 (12.72)	76 (14.26)	214 (12.37)
Resistant	442 (71.18)	281 (52.72)	828 (47.86)
Chloramphenicol			
Resistant	81 (13.04)	66 (12.38)	191 (11.04)
Tetracycline			
Intermediate	86 (13.85)	62 (11.63)	256 (14.80)
Resistant	347 (55.88)	275 (51.59)	660 (38.15)
Erythromycin			
Intermediate	59 (9.50)	40 (7.50)	116 (6.71)
Resistant	7 (1.13)	5 (0.94)	28 1.62)

PCV-7 vaccine serotypes (PCV-7), the additional six serotypes making up PCV-13 serotypes (PCV-13) and the non-vaccine serotypes (NVT) are all other serotypes not contained in PCV-13.

4.4. Discussion

This study provides the patterns of antibiotic susceptibility among invasive and carriage pneumococcal isolates in The Gambia over a 20 year period including the pre-PCV and PCV era. Results from this study indicate that susceptibility of both invasive and nasopharyngeal carriage pneumococcal isolates in The Gambia to commonly used antibiotics was generally high, with 99.8% of invasive isolates susceptible to penicillin and over 80% of invasive and carriage isolates susceptible to chloramphenicol. The exceptions to this pattern of high susceptibility of study isolates to the commonly used antibiotics were the high levels of non-susceptibility

to co-trimoxazole and tetracycline with as much as 68% of invasive isolates and 54% of carriage isolates resistant to co-trimoxazole. This study also showed that levels of co-trimoxazole resistance increased over time from the pre-PCV era into the PCV era, and the levels of antibiotic resistance differed by pneumococcal serotypes and sequence types. These results provide important background data for continued monitoring of antibiotic resistance patterns following the introduction of pneumococcal vaccines in The Gambia.

Findings of high levels of susceptibility of pneumococcal isolates in The Gambia to commonly used antibiotics, and in particular a very low prevalence of intermediate resistance (0.2%) of invasive pneumococci to penicillin, is at variance with reports of high rates of penicillin resistance in other parts of Africa. In contrast to our findings, much higher rates of up to 28% and 55.3% respectively of full and intermediate resistance to penicillin have been reported from invasive isolates in North West Nigeria (Lliyasu et al., 2015) and similarly high rates (>20%) of penicillin non-susceptible pneumococci have been reported previously in many other African countries (Benbachir et al., 2001, El Mdaghri et al., 2012, Wasfy et al., 2005, Vallès et al., 2006). Findings from this study, which included pneumococcal isolates collected from the Western, Upper and Central regions of The Gambia over time periods spanning the pre- and post- PCV era, indicate that the prevalence of penicillin resistant invasive pneumococci has remained very low in The Gambia post PCV introduction. In the USA, decreases in the prevalence of antibiotic resistant pneumococci was observed following introduction of PCV (Richter et al., 2014). The decrease in prevalence of penicillin-resistant pneumococci was thought to be due to decreases in antibiotic use in the USA that followed introduction of PCV. Reasons why the rates of penicillin-resistant pneumococci have remained low in The

Gambia are still unclear but could be due to differences in use of antibiotics between countries, as access to antibiotics may be generally poorer in The Gambia than it is in other countries. Difficulties in access to antibiotics in The Gambia might be related to a combination of factors such as higher drug costs, less availability and poorer disposable incomes in many families, as has been suggested previously (Adegbola et al., 2006, Adetifa et al., 2012). Some studies have shown that use of antibiotics could drive increases in rates of antibiotic resistance among pneumococci. Low rates of 1.4% for penicillin resistant and 10.9% for reduced sensitivity pneumococci among carriage isolates obtained from young children have also been reported from Niger during the pre-PCV era although little is known of the patterns of antibiotic use in that setting (Ousmane et al., 2017).

In this study, prevalence of non-susceptibility to co-trimoxazole throughout the study period was 73.8% and 66.6%, among invasive and carriage pneumococci respectively. These high rates of non-susceptibility to co-trimoxazole are consistent with reports from other parts of Africa and elsewhere including from Kenya 98.6% (Kobayashi et al., 2017), Malawi 96% (Cornick et al., 2014) and India, 81.8% (Jain et al., 2005). Importantly, results from this study indicate that the high levels of resistance to co-trimoxazole cut across all the predominant serotypes and sequence types seen in The Gambia. Evidence of non-susceptibility to co-trimoxazole in many settings, along with reports of poor clinical response to co-trimoxazole treatment, formed the basis for current treatment guidelines for pneumonia, in particular that amoxicillin be used in place of co-trimoxazole as first line treatment (WHO, 2012). Our results show that the pattern of non-susceptibility to co-trimoxazole has persisted and increased in the post PCV era in The Gambia and provide evidence that

the drug may not be appropriate for treatment of pneumococcal diseases in The Gambia.

Reports from other parts of the world have shown a decrease in resistant pneumococci following introduction of PCV, driven largely by decreases in antibiotic-resistant vaccine serotypes (Kyaw et al., 2006, Tomczyk et al., 2016). In the current study, other than a significant increase in resistance to co-trimoxazole from a prevalence of 60.8% in the pre-PCV-7 era to 95.2% in the post PCV-13 era, our results did not show any difference in prevalence of resistance to the commonly used antibiotics between the periods before and after introduction of PCV in The Gambia. These results contrast with reports from The USA and France which showed that overall rates of antibiotic resistance among pneumococci decreased following introduction of the PCV-13 (Richter et al., 2014, Janoir et al., 2016). Reductions in the proportions of pneumococcal isolates resistant to penicillin, erythromycin, cotrimoxazole and tetracycline were reported in France (Janoir et al., 2016). Observed decreases in prevalence of antibiotic resistant pneumococci in these two countries were noted to be due, in large part, to declines in the PCV-13 vaccine serotype 19A, which had been shown previously to be multi-drug resistant. The absence of an appreciable reduction in levels of resistance to most of the antibiotics included in the current study, might be due to the relatively short period of observation (2011-2015) following introduction of PCV-13 in the current study. Continued surveillance for a longer period of time following PCV-13 introduction in The Gambia would allow for better appreciation of the true effect of PCV-13 on antibiotic resistance patterns.

In addition, this study showed that among invasive pneumococci, the predominant non vaccine serotype in The Gambia, serotype 12F, demonstrated high levels of

resistance to three of the antibiotics tested, including co-trimoxazole and chloramphenicol. In contrast, isolates belonging to serotypes 1 and 5, two of the leading PCV-13 vaccine types in The Gambia, demonstrated high levels of resistance to only 2 of the antibiotics tested. Serotype 12F has been shown to be an important emerging non-vaccine serotype that could play a role in driving antibiotic resistance patterns. Reports from France showed that non-PCV-13 serotypes increased significantly following introduction of PCV-13, accounting for more than half of all IPD in the country and that serotype 12F was the leading non-PCV serotype consistently associated with resistance to multiple antibiotics, including co-trimoxazole, tetracycline, erythromycin and chloramphenicol (Janoir et al., 2016). Results from the current study highlight the need for continued monitoring of emerging pneumococcal serotypes circulating in The Gambia and their antibiotic resistance patterns, to inform on more effective control measures against IPD in the post PCV-13 era.

Most of the resistance to co-trimoxazole seen in serotype 1 was due to ST3081, whereas susceptibility of ST618 serotype 1 isolates was high. In the previous chapter, it was shown that ST3081 had replaced ST618 as the predominant serotype 1 ST circulating in The Gambia. The finding in this chapter that the high levels of ST3081 resistant strains appeared to have emerged over a period of time would suggest that antibiotic pressure might have played a role in the replacement of ST618 by ST3081 as the dominant serotype 1 clone.

In conclusion, this study showed that levels of antibiotic non-susceptibility among invasive pneumococci in The Gambia have fluctuated over the years following nationwide introduction of PCV. Importantly, non-susceptibility to most of the commonly used antibiotics and in particular resistance to penicillin, has remained

low in comparison to rates observed in other parts of Africa and some other developing countries. This finding could have implications for continued use of penicillin in this region. Also, this study has highlighted the need for continued monitoring of emerging non-vaccine type pneumococci in the post PCV era, and that these NVT could play an important role in patterns of antibiotic resistance among pneumococci in The Gambia.

5. Comparative genomic analysis of *S. pneumoniae* serotype 1 strains from West Africa

5.1. Introduction

Studies conducted and described in previous chapters of this thesis on the population structure of *S. pneumoniae* in The Gambia among invasive and carriage isolates highlight the importance of pneumococcal serotype 1 in The Gambia and provide a description of some important characteristics of *S. pneumoniae* serotype 1. These characteristics include a clonal replacement of dominant clones, i.e. ST3018 replacing ST618 and specific differences in antimicrobial resistance patterns among circulating serotype 1 STs. A better understanding of the genomics characteristics of these STs will be useful in broadening our understanding on in the evolution of *S. pneumoniae* in The Gambia and West Africa.

Evidence for geographical clustering of pneumococcal serotype 1 has been shown in studies performed on isolates obtained across different regions of the world (Brueggemann and Spratt, 2003, Cornick et al., 2015). A phylogenetic study utilizing MLST reported that of 166 pneumococcal serotype 1 isolates from 16 locations around the world, three distinct lineages were observed, with lineage A strains from Europe and North America, lineage B mostly from Africa and Israel and lineage C predominantly from Chile (Brueggemann and Spratt, 2003). More recently, the Pneumococcal African Genomics (PAGe) Consortium reported similar findings of geographical clustering of pneumococcal serotype 1 from WGS of 448 isolates from 27 countries across different parts of the world, including 11 in Africa (Cornick et al., 2015). The phylogenetic studies grouped these isolates into four lineages, three of which (A-C) had been previously described (Brueggemann and

Spratt, 2003) of which African serotype 1 isolates were found in lineage B. The newly described lineage D comprised of Asian isolates. In a further classification, the African lineage B phylogeny was grouped into six distinct clades (labelled i-vi) which clustered by geography. Pneumococcal serotype 1 strains from The Gambia collected over a five year period (2003- 2008) were grouped into 3 distinct clades (i, iii and v) and this was dependent on ST. ST3081 isolates were grouped in clade v along with STs from South Africa and Mozambique while ST618 isolates grouped into clade i.

As demonstrated in the studies described above, whole genome sequencing (WGS) which is now more readily available and affordable than previously, offers the potential to evaluate bacterial genomes on a small scale as well as in large epidemiological studies (Chewapreecha et al., 2014c, Laabei et al., 2014, Chewapreecha et al., 2014a). *S. pneumoniae*, a naturally transforming organism, is known to have evolved over time, its' evolution mediated by different genetic mechanisms including point mutation, recombination and horizontal gene transfer. These evolutionary changes could be due to selective pressures from various factors including antibiotics usage and vaccination with PCV (Croucher et al., 2011). WGS offers the opportunity to understand more clearly, the relative importance of these mechanisms in the evolution of *S. pneumoniae* in different geographical settings. PCV 7 was introduced in The Gambia in 2009, and this was later replaced by PCV 13 in 2011. It remains unclear how this intervention may have affected evolution of *S. pneumoniae*, in particular the predominant pneumococcal serotypes and sequence types circulating in The Gambia.

Given the importance of pneumococcal serotype 1 as a leading cause of IPD in The Gambia, this chapter aims to describe the phylogenetic structure of pneumococcal

serotype 1 strains collected over a 20 year-period, 1995 – 2015, including historical pneumococcal serotype 1 isolates and serotype 1 isolates collected in the post PCV era in The Gambia. Employing recombination and pan genome analysis, it also aims to understand the evolution of STs in West Africa. It is hoped that findings from this chapter will provide clearer insights into the pattern of replacement observed previously in The Gambia among the predominant pneumococcal serotype 1 strains.

5.2. Methods

Two hundred and fifty one *S. pneumoniae* serotype 1 isolates were randomly selected from pneumococcal serotype 1 isolates obtained from invasive disease and nasopharyngeal carriage studies. These isolates were predominantly recovered from The Gambia but other context isolates recovered from West African countries including Togo, Senegal, Niger and Benin were also included.

All invasive isolates (188 isolates) were obtained from the World Health Organization Regional Reference Laboratory (WHO RRL) hosted at The MRC Unit The Gambia. Gambian isolates were sent in to the WHO RRL from a hospital-based pneumococcal surveillance at the MRC hospital in the Greater Banjul Area and from a population-based surveillance of PCV effectiveness in the Upper River Region of The Gambia (as described in section 3.2). Other invasive disease isolates were sent to the WHO RRL from major referral hospitals in Togo from Centre Hospitalier Universitaire - Tokoin of Lome, Senegal from Hopital d'Enfant Albert Royal, Dakar, Ghana from Kumasi Teaching Hospital, Kumasi, Niger from Centre Hospitalier Universitaire - Laboratoire Hopital National, NE Niamey and Benin from Service National de Laboratoire Sante Publique. All nasopharyngeal carriage isolates (63 isolates) were obtained from carriage studies on healthy individuals from the

Western Region and Upper River Region of The Gambia (study description in section 3.2). A summary of the characteristics of patients from whom 251 *S. pneumoniae* serotype 1 isolates were collected and analysed in this chapter are described in Table 5. 1.

Table 5.1 Summary of patient characteristics from whom isolates analysed in this study were collected (N=251)

Characteristic	Category	Isolates N (%)
Age group (years)	<5 years	150 (59.76)
	5 - <15 years	56 (22.31)
	≥15 years	27 (10.76)
	Unknown	18 (7.17)
*Context	Hospital	57 (22.71)
	Surveillance	131 (52.19)
	Carriage	63 (25.10)
Gender	Female	83 (33.07)
	Male	104 (41.43)
	Unknown	64 (25.50)
Specimen	Blood	110 (43.82)
	CSF	43 (17.13)
	LA	26 (10.36)
	NPS	63 (25.10)
	Others	9 (3.39)
Country	The Gambia	219 (87.25)
	Senegal	13 (5.18)
	Togo	9 (3.59)
	Niger	7 (2.79)
	Ghana	2 (0.80)
	Benin	1 (0.40)

N, number of isolates; CSF, Cerebrospinal fluid; LA, Lung aspirate; NPS, nasopharyngeal swabs. Other specimen include pleural fluid and joint fluid aspirates.*Context: Hospital surveillance conducted from 1996-2015 (MRC, Clinic, Western Region, The Gambia); 2007-2010 (Tokoin Teaching Hospital, Togo); 2010 (Hopital d'Enfant Albert Royal, Dakar, Senegal); 2009-2010 (Centre Hospitalier Universitaire - Laboratoire Hopital National, NE Niamey, Niger); population based surveillance into PCV effectiveness conducted in the Upper River region of The Gambia, 2008-2015, pneumococcal carriage studies conducted in the Western region of The Gambia from 2003-2009 and South West Nigeria in 2009.

Study isolates comprised of 13 sequence types all within the ST217 clonal complex, namely, ST3081, ST618, ST303, ST217, ST2084, ST612, ST3575, ST3581, ST10649, ST11779, ST12197, ST12310 and ST3579 (Table 5.2).

Table 5.2 Multilocus sequence type profiles of *S. pneumoniae* serotype 1 isolates from West Africa

		Sequence types							Specimen	
ST	N (%)	*aroE*	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*	Invasive	Carriage
3081	142 (56.57)	10	18	4	1	7	232	9	98	44
618	64 (25.50)	13	8	4	1	7	19	14	50	14
303	10 (3.98)	10	5	4	1	7	19	9	8	2
217	16 (6.37)	10	18	4	1	7	19	9	14	2
2084	3 (1.20)	13	8	4	2	7	19	14	3	0
612	3 (1.20)	10	18	4	1	7	19	31	3	0
3575	3 (1.20)	10	191	4	1	7	19	9	3	0
3581	2 (0.80)	13	8	4	115	7	19	14	2	0
10649	2 (0.80)	10	338	4	1	7	19	9	2	0
11779	2 (0.80)	338	18	4	1	7	232	9	2	0
12197	2 (0.80)	338	18	4	1	7	232	9	2	0
12310	2 (0.40)	10	18	4	1	2	232	9	0	1
3579	1 (0.40)	13	8	4	5	7	250	14	1	0
Total	**251**								**188**	**63**

aroE, shikimate dehydrogenase; *gdh*, glucose-6-phosphate dehydrogenase; *gki*, glucose kinase; *recP*, transketolase; *spi*, signal peptidase I; *xpt*, xanthine phosphoribosyltransferase); *ddl*, D-alanine-D-alanine ligase.

DNA was extracted and quantified as described in section 2.4 and sent to WTSI for whole genome sequencing as described in section 2.7. Paired end sequencing reads were analysed using the nullabor pipeline (section 2.7).

5.3. Results

5.3.1. Genomic characterisation

Of the 251 serotype 1 genomes analysed using the nullabor pipeline, the median depth of coverage was 220, ranging from 63 – 460. The average GC content was 40% ranging from 37.8 to 42.7. Sequencing reads were mapped to reference genome, SPN1041 [Genbank: CACE00000000]. The size of draft assemblies ranged from 2022554 – 2375186 bp and the number of contigs from 128 – 539.

5.3.2. Phylogenetic analysis

A maximum likelihood phylogenetic tree using RAxML was constructed from core genome alignments of 251 serotype 1 isolates utilizing the Gubbins algorithm. The tree was constructed based on putative point mutations occurring outside recombination regions for better evolutionary inference. The tree was rooted at midpoint and visualized using ITOL. *S. pneumoniae* serotype 1 isolates included in this study belonging to the same ST clustered on monophyletic branches on the phylogenetic tree. The serotype 1 strains grouped into four clades (figure 5.1).

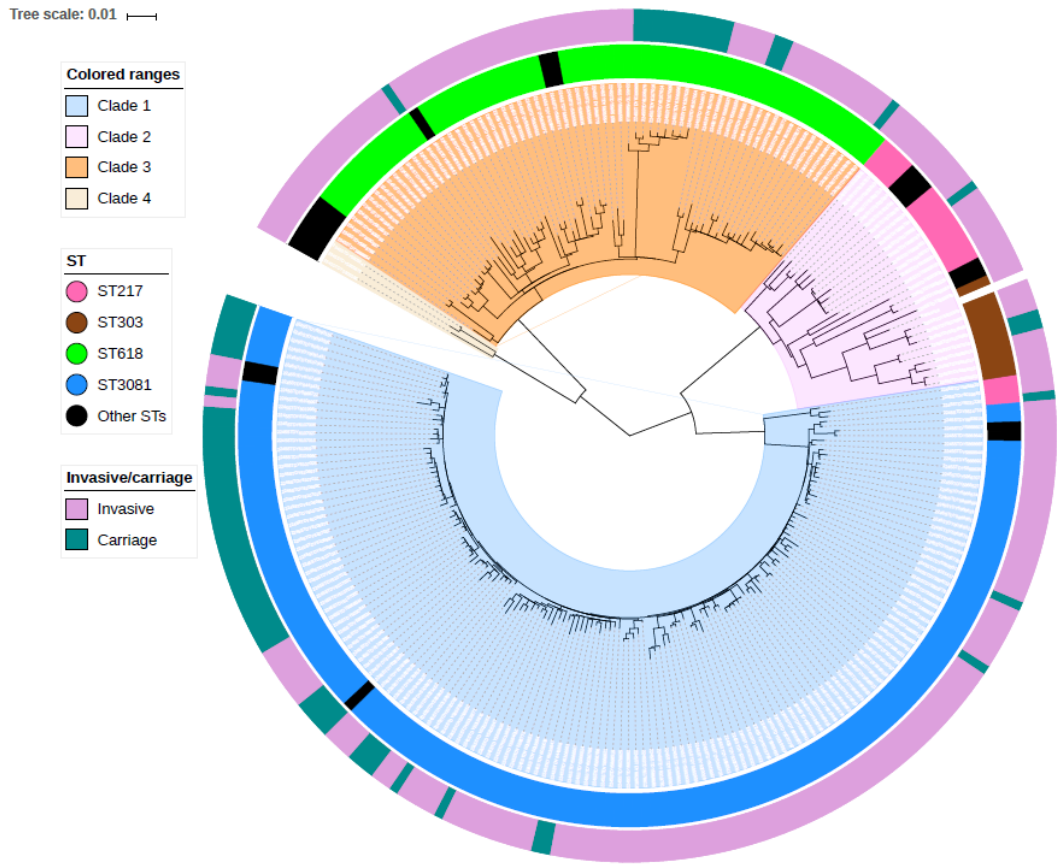

Figure 5.1 Maximum-likelihood phylogenetic tree of CC 217 *S. pneumoniae* serotype 1 isolates.
Branch lengths and labels are coloured by clades. The internal ring after the labels represent the STs and the outer ring represents indicate if invasive or carriage isolates as represented by the key on the left.

The shortest branch lengths on the tree were observed in clade 1 indicative of a newly emerging clone. This clade comprised of 2 sub-clades, the first sub-clade consisting of all ST3081 and its' single locus variants ST12917, ST12310 and ST11779 from both invasive disease and nasopharyngeal carriage. All isolates in this clade came from The Gambia and neighbouring country Senegal. Three ST217 isolates, two from invasive disease and one from carriage, recovered from The Gambia formed a subclade within the ST3081 dominated clade 1. Within clade 1 nasopharyngeal carriage isolates clustered more closely with each other compared to invasive disease isolates.

Clade 2 comprised of the remaining ST217 isolates from both invasive disease and nasopharyngeal carriage and its single locus variants, ST303, ST10649, and ST3575. Isolates from other West African countries were mostly found in clade 2. Within clade 2, all ST303 clustered together and as observed in clade 1 the carriage isolates clustered together on the phylogenetic tree. Clusters of ST217 isolates were separated by ST3575 isolates.

Clade 3 comprised of all ST618 isolates and its' single locus variants ST2084, ST3581 and double locus variant ST 3579 isolated from both invasive disease and nasopharyngeal carriage. West African isolates included in this study were also found in this clade. The branch lengths in clade 3 were longer and varied substantially than those from clade 1. This is a reflection of the epidemiological data showing that ST618 (Clade 3) were in existence several years before the emergence of ST3081 (Clade 1) in The Gambia. Within clade 3, all the different STs clustered together as well as nasopharyngeal carriage isolates.

Clade 4 comprised of all three ST612 isolates from The Gambia and Senegal. ST612 is a single locus variant of ST217.

No clear clustering of isolates was observed with other associated patient characteristics such as age group, gender and specimen and hence these are not shown in figure 5.1.

5.3.3. Recombination analysis

To understand the evolutionary characteristics occurring within pneumococcal serotype 1 isolates from West Africa included in this study, regions of recombination were predicted from the core genome alignments of all study serotype 1 isolates

using Gubbins. Sequence reads were mapped against a serotype 1 reference sequence, Sp1041 with Genbank accession number CACE00000000.

Unique patterns of recombination blocks were observed in each of the four clades within this study (Figure 5.2). Clade 1 consisting mostly of ST3081 isolates had 12 recombination blocks and was placed on a branch that had accumulated 1142 SNPs with 1065 occurring inside recombination regions. Clade 2 consisting of ST217 and ST303 isolates had 26 combination blocks occurring and placed on a branch that had accumulated 1932 SNPs occurring of which 1819 occurred inside recombination regions. Clade 3 isolates consisting of mostly ST618 isolates had the highest number of recombination blocks (33) occurring suggesting that it has been the oldest ST and had more time for recombination to occur. Isolates within clade 3 were placed on a branch that had accumulated a total of 3464 SNPs of which 97% were found inside recombination regions (3370/3464). Clade 4 isolates consisting of ST612 isolates had 14 recombination blocks occurring and was placed on a branch that had accumulates a total of 760 SNPs with 645 occurring within recombination regions.

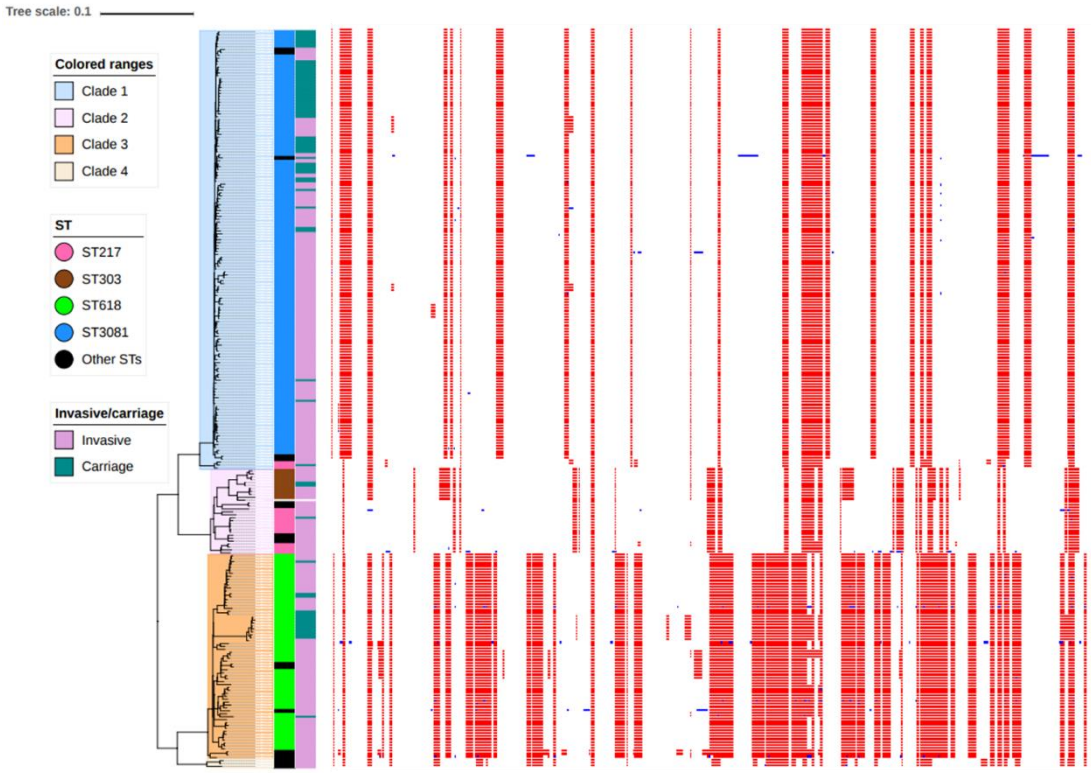

Figure 5.2 Presence and absence of recombination blocks within CC217 *S. pneumoniae* serotype 1.
The heat map on displays the presence and absence recombination blocks next to the phylogenetic tree. Red segments represent present while white segments represent absent. The blue blocks represent recombination events occurring in a single isolate.

Regions within the genomes where recombination events were occurring include genes within the capsular loci, genes coding for antibiotics resistance, phage, restriction modification systems, transport proteins, hypothetical proteins and well known pneumococcal virulence genes. Specific examples include the pneumococcal capsular region in which recombination events were occurring in all four clades. The recombination blocks overlapped with different genes within this region that were clade specific. For example, the recombination block overlapping *wzx*, *gla* and *udg* was specific to clade 1 isolates only, while recombination blocks overlapping *wzy* to *wzg* were unique to clade 2, 3 and 4 (Figure 5.3).

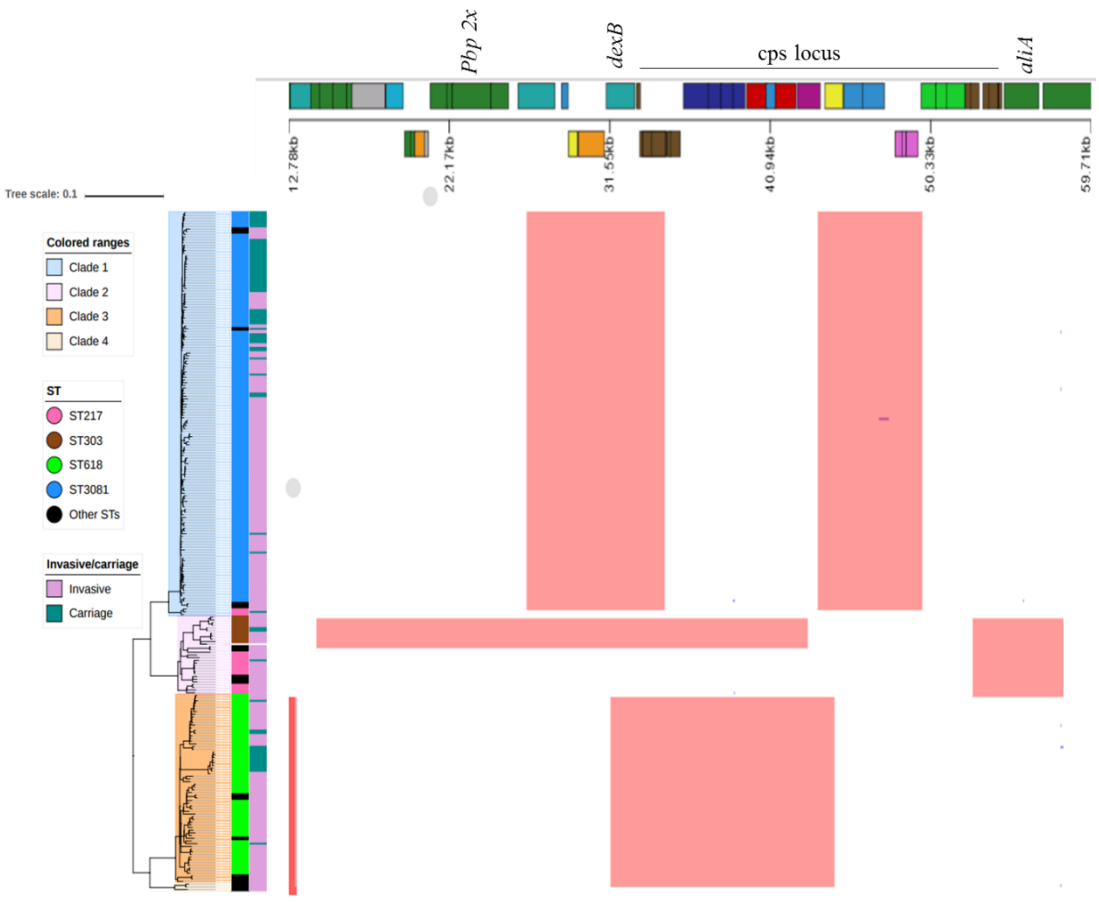

Figure 5.3 Recombination blocks occurring within the pneumococcal capsular region.

The heat map on displays the presence and absence recombination blocks next to the phylogenetic tree. Red segments represent present while white segments represent absent. The blue blocks represent recombination events occurring in a single isolate.

Recombination blocks overlapping with antibiotics resistance genes including *pbp1b* and *gyrA* found in a clade 1 ST12310 isolate *pbp2A* in all four clades in clade specific recombination blocks, *pbp2x* in clade 2 ST303 isolates only, as well as *ciaR*, *ciaH*, *pbX*, *mraW* and *mraY*. The gene *folP* in clade 1 among 5 ST3081 isolates. Known virulence genes that play a role in nasopharyngeal carriage *zmpB*, *nanA* and *nanB* were seen in clade 2 and 3, but not in clade 1. The gene *cbpA* in all 4 clades, *pepA* and *pepO* in clade 3 only. Previous studies showed that an important difference between ST3081 and ST618 isolates was the presence of two different types of

fucose metabolism operons which is consistent with what has also been observed among the current study isolates that includes more serotype 1 isolates collected over a longer period than in previous reports. Recombination was observed occurring within this region among only clade 2 and 3 isolates, but not clade 1 isolates. In addition, the carriage cluster of ST618 isolates in clade 2 had a unique recombination block occurring within this region.

5.3.4. Pan genome analysis

Pangenome analysis was conducted using Roary software. Gene presence or absence table was inputted into Panini to explore the accessory genome of serotype 1 isolates included in this study. Output from Panini was sent to the webtool Microreact and visualised in the context of the phylogeny and associated metadata. The accessory genome content of isolates clustered isolates from the three main clades (clade 1, 2 and 3) into three distinct clusters Clade 3 isolates made up of mostly ST618 isolates clustered farther away from clade 1 and 2 clusters, a reflection of MLST types and core genome content (figure 5.4).Genes contributing to the clustering of the accessory genomes are listed in Appendix 1.2

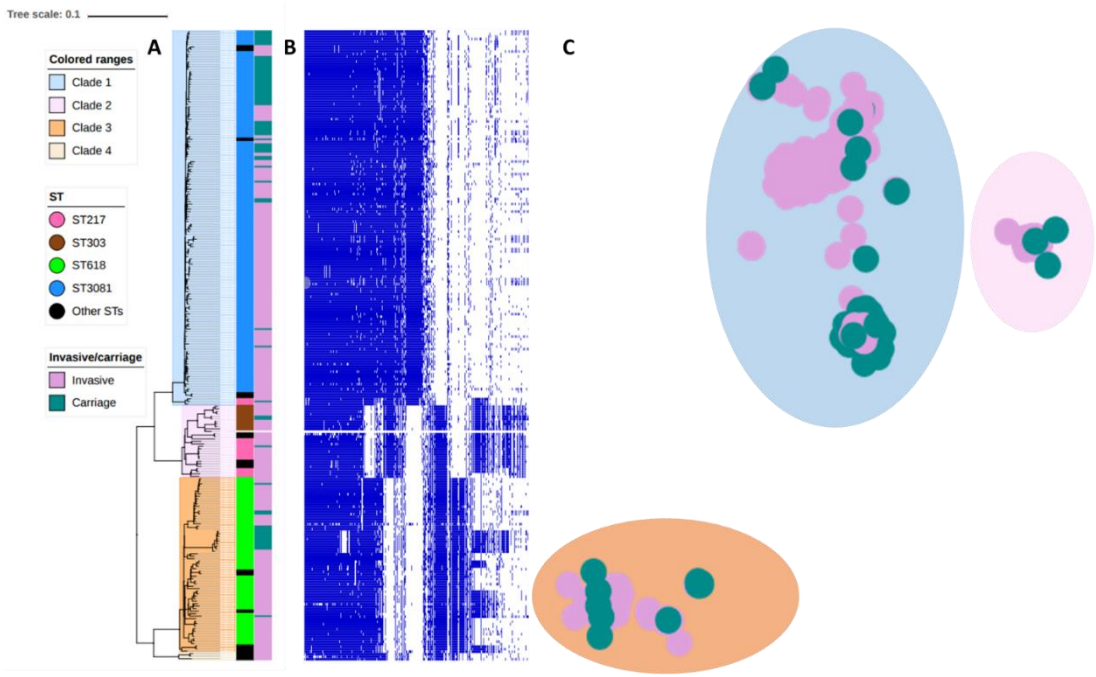

Figure 5.4 Pneumococcal accessory region of pneumococcal serotype 1 genomes
Panel A: Pneumococcal serotype 1 phylogeny based on the core region. Panel B: Heat map showing the presence and absence of accessory genes within the dataset. Blue represents genes present and white represents genes absent. C. Output from PANINI on the pneumococcal serotype 1 accessory region. The outer circles are coloured by clade and the inner circles by invasive disease or carriage.

Among clade 1 isolates two main clusters were observed, one made up of mostly invasive isolates and the other made up of mostly nasopharyngeal carriage isolates (figure 5.4). Differences observed among carriage and invasive isolates were due to large numbers of uncharacterized proteins which were unique to carriage isolates, but not found among invasive isolates within the clades. Characterized proteins found included some conserved proteins among the ST618 carriage isolates such as chloramphenicol acetyl transferase. Others were genes that are part of toxin-antitoxin systems including *pezT-2* (toxin PezT) and *xre* (HTH-type transcriptional regulator Xre).

Other accessory genes contributing to differences observed among serotype 1 isolates in this study determined by ABRICATE software in the nullabor pipeline

include genes coding for antimicrobial resistance (Figure 5.5). Of interest, the *tetM* responsible for resistance to tetracycline was found among clades 2 and 3 isolates but not in clade 1. Two alleles of *tetM* were found; all clade 2 isolates harboured the *tetM2* (Genbank accession number X90939) and clade 3 isolates harboured the *tetM12* allele (Genbank accession number FR671418). Among clade 3 isolates, the *tetM12* allele was found mostly among carriage isolates.

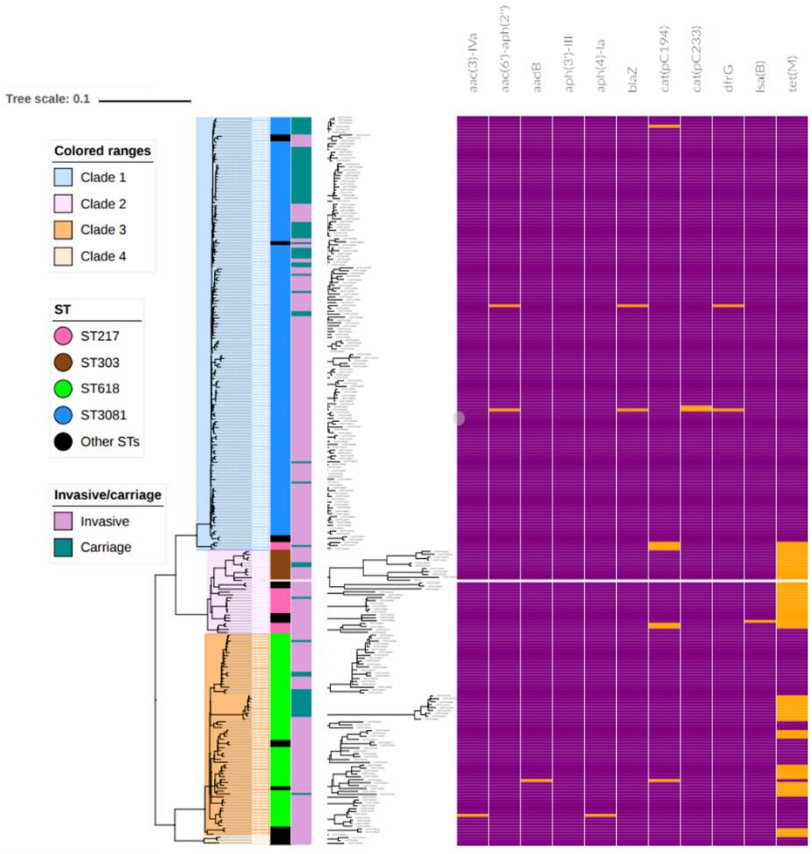

Figure 5.5 antibiotics resistance genes
The left panel is the pneumococcal serotype 1 phylogeny base on the core region. The heat map shows the presence and absence of resistance genes. Orange blocks represent genes present, while purple blocks represent genes absent.

5.4. Discussion

This study highlights some important characteristics that aid our understanding of the evolution of pneumococcal serotype 1 among pneumococcal isolates from across

the West Africa sub-region. It highlights the important role of recombination in evolution of serotype 1 STs in the sub-region as well as associated genes such as antimicrobial resistance genes, virulence genes and colonization factors associated with these processes. It also highlights important genomic differences between the major circulating serotype 1 STs (ST618 and ST3081) that could explain the clonal replacements observed in The Gambia.

Phylogenetic analysis grouped study isolates into four clades with ST3081 belonging to Clade 1 and ST618 belonging to Clade 3, and separated on long branches. The ST3081 Clade had shorter branch lengths than the ST618 Clade, indicative of a more recently evolved clade and is consistent with the epidemiological data discussed in chapter 3. This finding is also consistent with findings from earlier studies conducted on African serotype 1 isolates which included a subset of isolates included in this dataset. In those earlier reports, ST3081 isolates grouped with Clade v and ST618 isolates grouped with other Asian clones in Clade i (Cornick et al., 2015). With the inclusion of more historical serotype 1 isolates and serotype 1 isolates in the post-2011 period sampled during a global serotype 1 study, a greater degree of diversification in terms of branch lengths occurring among Clade 1 isolates are seen in comparison to what had been reported earlier. In addition, this study revealed clusters of carriage isolates on the phylogenetic tree in different clades, providing an opportunity to explore differences between invasive and carriage strains, which had not been possible in the previous study due to its smaller sample size and shorter period of coverage. Other important findings from the phylogenetic analysis included the separation of ST217 isolates into 2 different clades in this study, Clades 1 and 2. Three isolates found in Clade 1, were observed to share the most common recent ancestor with ST3081 isolates and in Clade 2, shared the most common recent

ancestor with ST303 isolates. Studies conducted on a large collection of 226 ST217 isolates from nine African and four Asian countries grouped these STs into 5 clusters (Chaguza et al., 2016). In that study, Gambian ST217 isolates were found in two clusters, SC3-SEA along with samples from Mozambique and Malawi, where Clade 1 isolates in the current study belong and the SC2-WA with isolates from West Africa and some parts of Asia where Clade 2 isolates in the current study belonged (Chaguza et al., 2016). This may indicate some inter-country transmission of clones. This possibility should be considered given the increase in international travel across countries with increasing human activity in recent years and also since The Gambia is a known tourist destination, thereby attracting more frequent international travels. These factors may have contributed to the observed diversity of pneumococcal serotype 1 strains in the West Africa sub-region.

The pneumococcus is a highly transforming pathogen and it has been shown that recombination events play a major role in its evolution (Croucher et al., 2011, Chewapreecha et al., 2014b). This study showed that recombination plays an important role in the evolution of serotype 1 STs with over 95% of SNPs found occurring inside recombination regions. Studies have shown that different pneumococcal lineages differ in their rates of recombination (Chewapreecha et al., 2014a, Croucher et al., 2013) and the current study has shown more recombination events occurring within Clade 3 isolates than seen in Clade 1 isolates. This may suggest that these events occurring within Clade 3, which is mostly made up of ST618 isolates, have led to distortion or loss of genes that compromised the fitness of ST618 isolates, thereby giving ST3081 a fitness advantage over ST618.

Important regions within the genome where recombination events which were clade specific were found occurring in all clades in this study include the capsule locus. It

has been shown that recombination events occurring in the pneumococcal capsule locus and flanking regions including the penicillin binding proteins (PBPs) have led to capsule switching and development of penicillin non-susceptibility (Brueggemann et al., 2007, Golubchik et al., 2012). With the recent introduction in several African countries of the serotype 1-containing PCV13, findings from this study further emphasize the need for continuous post vaccination surveillance, in particular, for capsule switching events. In this study recombination events overlapping pbp2x which is found upstream of *dex*B flanking the capsule region was unique to Clade 2 isolates, specifically ST303 isolates. Other genes undergoing recombination among study isolates in which changes occurring within genes could affect antibiotic susceptibility included *cia*R, *cia*H, *mra*Y and *mra*W, all of which are involved in cell wall synthesis (Chewapreecha et al., 2014c). In chapter 4, we reported only 2 of 1055 (0.19%) pneumococcal isolates in The Gambia demonstrating intermediate resistance to penicillin with neither of the two intermediate resistant isolates being serotype 1. However, this low level of penicillin non-susceptibility calls for on-going surveillance as changes over time in the patterns of antibiotic usage or changes in the affordability and availability of antibiotics in The Gambia may drive changes in the levels of resistance among the predominant pneumococci in circulation.

In addition, differences in recombination blocks at the capsule locus between ST3081 and ST618 may offer some explanation on the observed clonal replacement seen in The Gambia in the recent past. It has been shown that capsule thickness provides some virulence advantage to pneumococci in invasive disease. It remains a possibility, therefore, that these observed differences between ST3081 and ST618 in recombination events at the capsule locus, may have an influence on the integrity of the pneumococcal capsule and virulence of these STs in invasive disease. Further

research along these lines will provide better clarity on the roles of these findings in the observed replacement between these serotype 1 STs.

Additionally, regions with well characterized pneumococcal colonization factors such as ZmpB, NanA and NanB were seen in this study to have recombination occurring in isolates in Clade 3, but not in Clade 1 isolates. Similarly, amino acid sequence variability with ST clustering, i.e. different alleles were found between ST3081 and ST618 isolates in this study. In established models of infection, ST3081 isolates were seen to be better adapted for nasopharyngeal carriage than ST618 isolates (Brico-Monero, Ebruke et al., submitted manuscript). It is unclear if these differences in allele types have implications for a fitness advantage to ST3081. In a different protein, pneumolysin, differences in an amino acid allele between ST618 and ST3081 isolates was associated with increased haemolytic activity which has been shown to an important mechanism during epidemic outbreaks (Brico-Monero, Ebruke et al., submitted manuscript).

Within the accessory region, an important hub for exploring evolutionary mechanisms, some important observations were made from this study. Distinct clade-specific clusters were observed within the accessory region of isolates. Previous studies conducted on Gambian serotype 1 isolates, which represent a subset of isolates included in the current study, have highlighted significant differences between the predominant Gambian serotype 1 clones, ST618 and ST3081. These differences include the fucose metabolism operon, an important virulence factor of *S. pneumoniae*, and other important virulence factors such as XerD and FtsK (Cornick et al., 2015). In the current study, which has included a larger sample size and a longer period of sampling, the findings have remained consistent with these earlier reports. This study showed that recombination events were found occurring within

these operons in Clades 2 and 3 isolates but not in Clade 1 isolates. It is plausible that differences in operon types and patterns of recombination among pneumococcal serotype 1 strains may have contributed to clonal replacements observed in The Gambia. However, the precise role in fitness of these operons and their relative importance in the ecological fitness of the pneumococcus are yet to be fully understood.

Also noted in the accessory region of isolates in this study, were clade specific differences in antibiotic resistance genes. The *tet*M genes, carried on Tn916 transposon and responsible for resistance to tetracycline, were found among Clades 2 and 3 isolates with differences in alleles but was not seen at all in Clade 1, ST3081 isolates. In chapter 4, we report tetracycline resistance in both ST618 and ST3081 isolates even though the level of resistance was higher in ST618 isolates. It is therefore not clear why the *tet*M gene was completely absent among ST3081 isolates. However, the Int-Tn_1, a transposase from transposon Tn916 which confirms tetracycline resistance was conserved among all study isolates. In chapter 4, phenotypic differences in co-trimoxazole was noted between the predominant STs. In the current study however, the genes associated with resistance to co-trimoxazole including *folP* and f*olA* were not present. In addition, recombination events were found occurring in only five ST3081 isolates within the location of *folP* gene. Studies have shown a poor correlation between phenotypic resistance to cotrimoxazole and stains undergoing recombination in that region (Chewapreecha et al., 2014a).

In phylogenetic analysis, we mentioned clusters of carriage isolates observed among the clades in this study, differences in accessory region are largely due to uncharacterized portions which will need further studies to enable a clear understand

their roles. Of the notable known genes are pneumococcal *zet*a toxin, *pez*T and *xre* which is a part of the toxin-antitoxin systems. These systems help maintain the integrity of the cell and the antitoxin helps to supress the action of the toxin. However, under circumstances of stress, the antitoxin can be degraded and allows the toxin to function. The *pez*T toxin present in ST618 carriage isolates but absent from invasive ST618 isolates and ST3081 invasive and carriage isolates has been shown to inhibit peptidoglycan wall synthesis (Mutschler et al., 2011).

A few limitations to this study that should be considered in the interpretation of the findings include the following, other pneumococcal serotypes or closely related streptococci have not been included in this analyses. Some of these serotypes or bacterial species may have played a role in the observed clonal replacement and evolution of pneumococcal serotype 1 isolates cells since any observed differences in the predominant STs may depend on competition with STs of other serotypes or bacterial species that are known to co-exist with *S. pneumoniae* in the human nasopharynx. Additionally, mutagenesis experiments followed by animal infection experiments to ascertain if the selected determinants have a role in fitness of the STs, were not carried out. These experiments would have further strengthened the evidence for findings from the current study.

In conclusion, this study has highlighted the possible roles of recombination, antibiotic resistance, virulence genes and colonization factors as possible key factors involved in the evolution of pneumococcal serotype 1 in the West Africa sub-region. These findings highlight important differences in circulating clones in West Africa of the predominant serotype 1 strains, ST3081 and ST618. Continuous surveillance would be helpful in providing additional evidence that will aid improvements in our

understanding of the evolutionary mechanisms behind observed changes among the predominant pneumococcal sequence types in circulation.

6. Discussion and future work

6.1. Background

This thesis describes the population structure and antibiotic resistance patterns of *S. pneumoniae* strains from the West African sub-region over two decades. It provides a description of the evolutionary biology of pneumococcal serotype 1, a leading cause of IPD in the sub-region, within the context of other pneumococcal serotypes over a period prior to and including the introduction of PCV in the West Africa sub-region. This approach, to describe the molecular epidemiology of pneumococcal serotype 1 within the context of other serotypes, has provided a baseline that allows fora clearer understanding of the characteristics of pneumococcal serotype 1. Earlier studies have shown that evaluating the relationships between the predominant pneumococcal serotypes and genotypes circulating within a geographical setting in both invasive disease and carriage over the same time period provides a better understanding of the evolution of *S. pneumoniae* in that community (Brueggemann et al., 2003, Varon et al., 2015, Hanage et al., 2005).The importance of *S. pneumoniae* serotype 1 relates to its attributes. Various epidemiological studies carried out in different geographical locations on carriage and invasive strains at different times have provided evidence that pneumococcal serotype 1 is one of the leading causes of IPD globally and in the West African sub-region in particular, but paradoxically is rarely found in nasopharyngeal carriage (Adegbola et al., 2006, Brueggemann et al., 2013, Collard et al., 2013, Hill et al., 2006, Kronenberg et al., 2006). In addition, *S. pneumoniae* serotype 1 has higher susceptibility to antimicrobial agents compared to other non-pneumococcal serotypes 1 isolates (Antonio et al., 2008). Furthermore, higher valency pneumococcal conjugate vaccines that have included this important serotype (PCV 10 and PCV 13) were only

recently introduced over the last few years and are now increasingly being used in routine immunization schedules across different parts of the world including in the West African sub-region. These reasons provide the background for this PhD. This study therefore characterized invasive and nasopharyngeal carriage pneumococci, and in particular pneumococcal serotype 1 isolates, providing relevant description in epidemiological and genotypic contexts, including patterns of antimicrobial resistance over a 20 year period pre- and post-PCV era.

6.1. Research findings and implication for future studies

In this thesis, while studying the population structure of pneumococcal serotypes, I have shown that use of PCVs in the Gambia reduced vaccine type serotypes with the exception of serotype 1 and 5. I have also shown clonal replacements occurring within serotype 1 STs prior to the introduction of PCV-7 and demonstrated serotype replacement by serotypes 12F and 35B following PCV-13 introduction. This study also exploited use of whole genome sequencing to reveal important differences between the predominant serotype 1 STs which provide insights into possible roles of recombination in the evolution of pneumococcal serotype 1 in the West Africa sub-region.

6.1.1. *S. pneumoniae* population structure and vaccination

In this thesis, a reduction in prevalence of vaccine type pneumococcal serotypes following the introduction of PCV was observed, the exception being serotypes 1 and 5 both of which were rare in nasopharyngeal carriage but persisted into the post-PCV era as predominant pneumococcal serotypes in invasive disease. This finding is consistent with the report by Mackenzie et al (Mackenzie et al, 2016) and other

reports (Cohen et al., 2016, Tin Tin Htar et al., 2015, von Mollendorf et al., 2017). In the population-based surveillance report by Mackenzie et al, which was carried out in the rural part of The Gambia which makes up a subset of isolates included in this study, there was no evidence of a significant reduction on serotype 1 disease by following PCV13 introduction, with the episodes of serotype 1 IPD remaining the same throughout the duration of the surveillance, whilst temporal fluctuations were seen for serotype 5 disease. The study reported that serotype 1 IPD was more prevalent in children aged 2-4years than it was in those age less than 2years. However, the rather short baseline period (16months) prior to PCV-7 introduction and only a 2-year post PCV-13 observation period might have limited the ability of the study to detect any appreciable effects on serotype 1 disease. Findings from the current study, which included pneumococcal isolates from both urban and rural parts of The Gambia and covered a slightly longer period of post PCV-13 observation than reported by Mackenzie et al, provide further evidence for the need for cautious optimism regarding expected benefits from PCV-13 on serotype 1 disease. These findings may also re-echo concerns that have been expressed previously regarding the need for a booster dose or consideration for alternative immunization schedules that would include children in the older age groups who are more at risk for serotype 1 disease (Kwambana-Adams et al., 2016; Brueggemann et al., 2013, Ebruke et al., 2015, Klugman et al, 2011).

The emergence of non-vaccine type pneumococci, following introduction of PCV has been reported from other settings (Janoir et al., 2016, Del Amo et al., 2016). In particular, serotype 12F which was noted in this study to be an important cause of IPD in the post-PCV period, has been reported to cause epidemic outbreaks similar

to pneumococcal serotype 1in other parts of the world (Deng et al., 2016). Earlier studies on serotype 12F suggests this serotype has certain characteristics that are similar to pneumococcal serotype 1 including a high potential for causing epidemics and rarity in nasopharyngeal carriage (Zulz et al., 2013). The emergence of non-vaccine pneumococcal serotypes that have epidemic potential suggest the possibility that expected gains in the ability of PCV-13 to cause reductions in future epidemic outbreaks due to pneumococcal serotype 1, may be diluted. Continuous population based post-vaccine surveillance of the circulating pneumococcal serotypes, and in particular monitoring of serotypes 1, 5 and 12F, would be needed to accurately assess the impact of PCV-13 and guide further development of effective interventions.

Implications of the emergence of different genotypes of vaccine serotypes in the post PCV era seen in this study remain to be seen. The predominant STs for isolates belonging to serotypes 14 and 23F in the post-PCV era differed from the predominant STs of these serotypes that circulated in the pre-PCV period. It is yet unclear if these changes are related to the introduction of PCV or if they represent secular changes unrelated to the vaccine. This finding might be indicative of the potential for significant genetic and immunogenic variation that exists within a pneumococcal serotype, of which the impact on effectiveness of currently used pneumococcal conjugate vaccines remains to be seen.

6.1.2. Antibiotic resistance patterns

With over 99% of invasive isolates susceptible to penicillin, findings from this study contrasts with much higher rates of penicillin non-susceptibility (from above 20% to as much as 55%) that have been reported from other parts of Africa (Benbachir et al.,

2001, El Mdaghri et al., 2012, Wasfy et al., 2005, Vallès et al., 2006). However, the low rates of penicillin non-susceptibility reported in the current study are consistent with findings from an earlier study in The Gambia (Adegbola et al., 2006), carried out in the pre-PCV era. Findings from this study, which included pneumococcal isolates collected through the pre- and post- PCV era, would indicate that the prevalence of penicillin non-susceptible pneumococci has remained low in The Gambia and may have decreased further over the recent few years. Access to antibiotics in The Gambia is generally poor in comparison to other countries and stems from a combination of many factors including higher drug costs and poorer disposable incomes in many families, and could explain the low rates of antibiotic resistance in The Gambia (Adegbola et al., 2006, Adetifa et al., 2012).

Previous reports have provided evidence of decreases in the prevalence of resistant pneumococci following introduction of PCV in different parts of the world, driven largely by decreases in antibiotic-resistant vaccine serotypes (Richter et al., 2014, Janoir et al., 2016). In this study, the prevalence of resistant pneumococci to commonly used antibiotics in The Gambia remained largely unchanged from the pre- to the post-PCV era, with the exception of co-trimoxazole. However, the post PCV-13 period of observation included in this study (2011-2015) might have been inadequate to allow for a full assessment of the vaccine impact on prevalence of antibiotic resistant strains and on-going surveillance would allow the appreciation of the true effect of PCV-13 on antibiotic resistance patterns. The high rates of non-susceptibility to co-trimoxazole reported in this study align with studies in other parts of Africa and other developing country settings, where similarly high prevalence of co-trimoxazole resistant pneumococci, as high as 98.6%, have been reported (Kobayashi et al., 2017, Cornick et al., 2014, Jain et al., 2005). However,

with adoption in The Gambia of revised treatment guidelines, that recommend use of amoxicillin in place of co-trimoxazole as first line treatment for pneumonia (WHO, 2012), it is not unlikely that changes in the patterns of antibiotic use may follow over the years and with it, changes in the patterns of selective antibiotic pressure on circulating pneumococci in The Gambia. It remains to be seen if these changes will happen and if they may affect the prevalence of co-trimoxazole resistant pneumococci, as has been reported elsewhere (Richter et al., 2014). Furthermore, WGS findings from this study revealed evidence of recombination events at locations within the capsule locus that could potentially lead to capsule switching and development of penicillin non-susceptibility. With the recent introduction of PCV in The Gambia, and reports from elsewhere that capsule switching have followed introduction of pneumococcal vaccines, cautious monitoring of the prevailing serotypes and any evidence of capsule switching and any changes in antibiotic resistance patterns would be appropriate to help clarify the roles of these findings in The Gambia and similar settings.

Importantly, this study showed that the predominant emerging non-vaccine serotype in The Gambia, serotype 12F, demonstrated high levels of resistance to three of the four commonly used antibiotics included in this study. In contrast, isolates belonging to serotypes 1 and 5, the two leading vaccine serotypes seen in IPD in The Gambia, were seen to demonstrate high levels of resistance to fewer antibiotics in comparison to the NVT serotype 12F. Previous reports indicate that serotype 12F is an important emerging non-vaccine serotype that could play an important role in driving antibiotic resistance patterns in the post-PCV13 era (Janoir et al., 2016). This further buttresses the need for ongoing monitoring of the patterns of emerging pneumococcal serotypes and their antibiotic resistance profiles in the post PCV-13 era. in the sub-region.

However, the absence of any of the internationally disseminated antibiotic resistant clones, such as the PMEN1 lineage, ST81(Croucher et al., 2011), would be helpful for effective antibiotic treatment in this setting but would nevertheless require on-going surveillance. It remains unclear why these internationally disseminated antibiotic resistant strains were absent in The Gambia but their absence may have contributed to the observed low levels of antibiotic resistance among pneumococci in The Gambia.

6.1.3. *S. pneumoniae* serotype 1 sequence types and clonal replacement

In The Gambia, the two major serotype 1 STs observed to be circulating, ST618 and ST3081, had been reported previously in other parts of West Africa and in neighbouring Senegal within Africa and in Oman in the MLST database. This study highlights the change in predominant clones of invasive and carriage serotype 1 strains circulating in The Gambia, with ST3081 emerging as a new dominant clone in 2007, replacing ST618 which had predominated in The Gambia in the decade prior. These changes in dominant circulating genotypes might reflect adaptive changes by the pneumococcus in response to changes in the human population and human activity, such as antibiotic usage or vaccination. A carriage study of serotype 1 in the context of a PCV-7 vaccine trial in The Gambia did not show any clear evidence that the observed changes in genotype were the results of vaccine selective pressure (Ebruke et al., 2015). Findings from this study indicate that the two predominant serotype 1 STs differed in their antibiotic resistance patterns with higher levels of antibiotic resistance to co-trimoxazole seen amongst ST3081 isolates.

The availability of informative techniques such as whole genome sequencing has given us the opportunity to understanding of the key mechanisms involved in these clonal replacements observed in pneumococcal serotype 1. Below are important findings from the current study that may have possible roles in the evolution of pneumococcal serotype 1 among pneumococcal isolates from across the West Africa sub-region. We report the role of recombination in the evolution of *S. pneumoniae* in the sub-region. Clade and ST specific recombination events were occurring in the capsular locus and this suggests that this may have led to distortion of genes which may have resulted in a fitness advantage for ST3081 over ST618. Another important implication of the recombination events noted to have occurred in the pneumococcal capsule locus, is how these events may impact on capsule switching and the development of penicillin resistance, as has been suggested in other studies (Brueggemann et al., 2007, Golubchik et al., 2012).

In addition, observations in this study that the tetM genes, carried on Tn916 transposon and responsible for resistance to tetracycline, were found only among ST618 isolates but not in ST3081 isolates as well as other resistance genes found, may be indicative of differences in the selective pressure of antibiotics on evolution of these STs in The Gambia. These findings further emphasize the need for continuous post vaccination surveillance of antibiotic resistance.

Other findings in the current study include differences between the predominant serotype 1 STs in recombination events that involved well known colonization factors, such as ZmpB, NanA, NanB and in addition to important virulence factors such as the fucose metabolism operon (Cornick et al., 2015) reported by previous studies. Animal studies conducted on ST618 and ST3081 isolates showed that ST3081 was more virulent in well-established mouse models of infection and better

adapted for nasopharyngeal carriage. These findings provide additional evidence of possible mechanisms that could explain the observed clonal replacements seen in serotype 1 STs in The Gambia.

6.2. Future work

6.2.1. Continued epidemiological studies of *S. pneumoniae* population structure in West Africa

In the light of recent introduction of pneumococcal conjugate vaccines in The Gambia and in other African countries, continued surveillance of the serotypes and sequence types in circulation would be an important follow-up to this work. This will help to confirm the findings described in the current study and identify any further changes that may occur over time. Surveillance over a longer period of time would allow for a more robust assessment of the impact of PCV-13 on the predominant serotypes and sequence types in circulation, such as the emergence of non-vaccine serotypes, and would also provide better clarity on how the patterns of antibiotic resistance may change in the long term, within countries in the region.

6.2.2. Genomic characterization of *S. pneumoniae* serotypes

Genomic studies on the circulating strains of *S. pneumoniae* in the broader context of other pneumococcal serotypes and sequence types, as well as other bacterial species that co-exist with pneumococci in nasopharyngeal carriage, would be appropriate as a follow-up to this research study. This would enable a more robust assessment of the evolutionary mechanisms of the predominant pneumococcal strains in circulation that takes into account relationships with other pneumococcal serotypes and other relevant bacterial species. Results from these studies could provide greater insights

into the key evolutionary mechanisms and potentially inform on development of newer strategies that will be relevant for effective control of pneumococcal disease in the West Africa region.

6.2.3. Optimizing currently available interventions for control of pneumococcal disease in West Africa

Efforts that aimed at a review of the timing of currently used PCVs, particularly in African countries where the burden of pneumococcal serotype 1 disease is in older children and adults, could result in greater levels of reduction in deaths from this serotype. Research to assess optimal timing for additional booster doses of currently used PCVs would be appropriate.

In addition to on-going efforts to develop more effective pneumococcal vaccines and efforts at encouraging appropriate antibiotic use in treatment of pneumococcal disease, efforts to engage local and regional health authorities in intensifying public health campaigns on risk reduction strategies for control of pneumococcal disease in countries within the sub-region could have immediate impact on the prevalence of pneumococcal disease. Enlightenment campaigns that address known risk factors such as overcrowding and malnutrition, as well as the provision of improved medical care for persons with increased risk for pneumococcal disease including sickle cell anaemia and HIV positive persons, could bring result in immediate outcomes in reducing illness and deaths from pneumococcal disease.

6.3. Concluding remarks

This study adds to our understanding of the population structure of invasive and carriage pneumococcal serotypes from the West African region, and the patterns of antibiotic resistance over the pre and post-PCV era. It also provides some clarity into some possible mechanisms behind evolution of the predominant circulating pneumococcal sequence types in the region. Knowledge gained from this study will be useful in informing appropriate preventive public health decisions in the region

References

AANENSEN, D. M., MAVROIDI, A., BENTLEY, S. D., REEVES, P. R. & SPRATT, B. G. 2007. Predicted functions and linkage specificities of the products of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J Bacteriol,* 189**,** 7856-76.

ABDULLAHI, O., NYIRO, J., LEWA, P., SLACK, M. & SCOTT, J. A. 2008. The descriptive epidemiology of *Streptococcus pneumoniae* and *Haemophilus influenzae* nasopharyngeal carriage in children and adults in Kilifi district, Kenya. *Pediatr Infect Dis J,* 27**,** 59-64.

ABUDAHAB, K., PRADA, J. M., YANG, Z., BENTLEY, S. D., CROUCHER, N. J., CORANDER, J. & AANENSEN, D. M. 2017. PANINI: Pangenome Neighbor Identification for Bacterial Populations. *bioRxiv*.

ADEGBOLA, R. A., FALADE, A. G., SAM, B. E., AIDOO, M., BALDEH, I., HAZLETT, D., WHITTLE, H., GREENWOOD, B. M. & MULHOLLAND, E. K. 1994. The etiology of pneumonia in malnourished and well-nourished Gambian children. *Pediatr Infect Dis J,* 13**,** 975-82.

ADEGBOLA, R. A., HILL, P. C., SECKA, O., IKUMAPAYI, U. N., LAHAI, G., GREENWOOD, B. M. & CORRAH, T. 2006. Serotype and antimicrobial susceptibility patterns of isolates of *Streptococcus pneumoniae* causing invasive disease in The Gambia 1996-2003. *Trop Med Int Health,* 11**,** 1128-35.

ADETIFA, I. M., ANTONIO, M., OKOROMAH, C. A., EBRUKE, C., INEM, V., NSEKPONG, D., BOJANG, A. & ADEGBOLA, R. A. 2012. Pre-vaccination nasopharyngeal pneumococcal carriage in a Nigerian population: epidemiology and population biology. *PLoS One,* 7**,** e30548.

AGUIAR, S. I., BRITO, M. J., GONCALO-MARQUES, J., MELO-CRISTINO, J. & RAMIREZ, M. 2010. Serotypes 1, 7F and 19A became the leading causes of pediatric invasive pneumococcal infections in Portugal after 7 years of heptavalent conjugate vaccine use. *Vaccine,* 28**,** 5167-73.

ALBRICH, W. C., MONNET, D. L. & HARBARTH, S. 2004. Antibiotic selection pressure and resistance in *Streptococcus pneumoniae* and Streptococcus pyogenes. *Emerg Infect Dis,* 10**,** 514-7.

AMPOFO, K., BENDER, J., SHENG, X., KORGENSKI, K., DALY, J., PAVIA, A. T. & BYINGTON, C. L. 2008. Seasonal invasive pneumococcal disease in children: role of preceding respiratory viral infection. *Pediatrics,* 122**,** 229-37.

ANSALDI, F., CANEPA, P., DE FLORENTIIS, D., BANDETTINI, R., DURANDO, P. & ICARDI, G. 2011. Increasing incidence of *Streptococcus pneumoniae* serotype 19A and emergence of two vaccine escape recombinant ST695 strains in Liguria, Italy, 7 years after implementation of the 7-valent conjugated vaccine. *Clin Vaccine Immunol,* 18**,** 343-5.

ANTONIO, M., HAKEEM, I., AWINE, T., SECKA, O., SANKAREH, K., NSEKPONG, D., LAHAI, G., AKISANYA, A., EGERE, U., ENWERE, G., ZAMAN, S. M., HILL, P. C., CORRAH, T., CUTTS, F., GREENWOOD, B. M. & ADEGBOLA, R. A. 2008. Seasonality and outbreak of a predominant *Streptococcus pneumoniae* serotype 1 clone from The Gambia: expansion of ST217 hypervirulent clonal complex in West Africa. *BMC Microbiol,* 8**,** 198.

ANTONIO, M., HAKEEM, I., SANKAREH, K., CHEUNG, Y. B. & ADEGBOLA, R. A. 2009. Evaluation of sequential multiplex PCR for direct detection of

multiple serotypes of *Streptococcus pneumoniae* from nasopharyngeal secretions. *J Med Microbiol,* 58**,** 296-302.

AUSTRIAN, R. 1960. The Gram stain and the etiology of lobar pneumonia, an historical note. *Bacteriol Rev,* 24**,** 261-5.

AUSTRIAN, R. 1981. Pneumococcus: the first one hundred years. *Rev Infect Dis,* 3**,** 183-9.

AVERY, O. T., MACLEOD, C. M. & MCCARTY, M. 1944. Studies on the Chemical Nature of the Substance Inducing Transformation of Pneumococcal Types : Induction of Transformation by a Desoxyribonucleic Acid Fraction Isolated from Pneumococcus Type Iii. *J Exp Med,* 79**,** 137-58.

AYOUBI, P., KILIC, A. O. & VIJAYAKUMAR, M. N. 1991. Tn5253, the pneumococcal omega (cat tet) BM6001 element, is a composite structure of two conjugative transposons, Tn5251 and Tn5252. *J Bacteriol,* 173**,** 1617-22.

AZZARI, C., MORIONDO, M., INDOLFI, G., CORTIMIGLIA, M., CANESSA, C., BECCIOLINI, L., LIPPI, F., DE MARTINO, M. & RESTI, M. 2010. Realtime PCR is more sensitive than multiplex PCR for diagnosis and serotyping in children with culture negative pneumococcal invasive disease. *PLoS One,* 5**,** e9282.

BA, F., SECK, A., BA, M., THIONGANE, A., CISSE, M. F., SECK, K., NDOUR, M., BOISIER, P. & GARIN, B. 2014. Identifying an appropriate PCV for use in Senegal, recent insights concerning *Streptococcus pneumoniae* NP carriage and IPD in Dakar. *BMC Infect Dis,* 14**,** 627.

BAGNOLI, F., MOSCHIONI, M., DONATI, C., DIMITROVSKA, V., FERLENGHI, I., FACCIOTTI, C., MUZZI, A., GIUSTI, F., EMOLO, C., SINISI, A., HILLERINGMANN, M., PANSEGRAU, W., CENSINI, S., RAPPUOLI, R., COVACCI, A., MASIGNANI, V. & BAROCCHI, M. A. 2008. A Second Pilus Type in *Streptococcus pneumoniae* Is Prevalent in Emerging Serotypes and Mediates Adhesion to Host Cells. *Journal of Bacteriology,* 190**,** 5480-5492.

BAROCCHI, M. A., RIES, J., ZOGAJ, X., HEMSLEY, C., ALBIGER, B., KANTH, A., DAHLBERG, S., FERNEBRO, J., MOSCHIONI, M., MASIGNANI, V., HULTENBY, K., TADDEI, A. R., BEITER, K., WARTHA, F., VON EULER, A., COVACCI, A., HOLDEN, D. W., NORMARK, S., RAPPUOLI, R. & HENRIQUES-NORMARK, B. 2006. A pneumococcal pilus influences virulence and host inflammatory responses. *Proceedings of the National Academy of Sciences of the United States of America,* 103**,** 2857-2862.

BATT, S. L., CHARALAMBOUS, B. M., MCHUGH, T. D., MARTIN, S. & GILLESPIE, S. H. 2005. Novel PCR-Restriction Fragment Length Polymorphism Method for Determining Serotypes or Serogroups of *Streptococcus pneumoniae* Isolates. *Journal of Clinical Microbiology,* 43**,** 2656-2661.

BENBACHIR, M., BENREDJEB, S., BOYE, C. S., DOSSO, M., BELABBES, H., KAMOUN, A., KAIRE, O. & ELMDAGHRI, N. 2001. Two-year surveillance of antibiotic resistance in *Streptococcus pneumoniae* in four African cities. *Antimicrob Agents Chemother,* 45**,** 627-9.

BENTLEY, S. D., AANENSEN, D. M., MAVROIDI, A., SAUNDERS, D., RABBINOWITSCH, E., COLLINS, M., DONOHOE, K., HARRIS, D., MURPHY, L., QUAIL, M. A., SAMUEL, G., SKOVSTED, I. C., KALTOFT, M. S., BARRELL, B., REEVES, P. R., PARKHILL, J. &

SPRATT, B. G. 2006. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet,* 2**,** e31.

BERGLUND, J., VINK, P., TAVARES DA SILVA, F., LESTRATE, P. & BOUTRIAU, D. 2014. Safety, immunogenicity, and antibody persistence following an investigational *Streptococcus pneumoniae* and *Haemophilus influenzae* triple-protein vaccine in a phase 1 randomized controlled study in healthy adults. *Clin Vaccine Immunol,* 21**,** 56-65.

BERGMANN, S. & HAMMERSCHMIDT, S. 2006. Versatility of pneumococcal surface proteins. *Microbiology,* 152**,** 295-303.

BERGMANN, S., SCHOENEN, H. & HAMMERSCHMIDT, S. 2013. The interaction between bacterial enolase and plasminogen promotes adherence of *Streptococcus pneumoniae* to epithelial and endothelial cells. *Int J Med Microbiol,* 303**,** 452-62.

BERRY, A. M., LOCK, R. A., HANSMAN, D. & PATON, J. C. 1989a. Contribution of autolysin to virulence of *Streptococcus pneumoniae*. *Infect Immun,* 57**,** 2324-30.

BERRY, A. M., LOCK, R. A., THOMAS, S. M., RAJAN, D. P., HANSMAN, D. & PATON, J. C. 1994. Cloning and nucleotide sequence of the *Streptococcus pneumoniae* hyaluronidase gene and purification of the enzyme from recombinant Escherichia coli. *Infect Immun,* 62**,** 1101-8.

BERRY, A. M. & PATON, J. C. 2000. Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. *Infect Immun,* 68**,** 133-40.

BERRY, A. M., YOTHER, J., BRILES, D. E., HANSMAN, D. & PATON, J. C. 1989b. Reduced virulence of a defined pneumolysin-negative mutant of *Streptococcus pneumoniae*. *Infect Immun,* 57**,** 2037-42.

BETHE, G., NAU, R., WELLMER, A., HAKENBECK, R., REINERT, R. R., HEINZ, H. P. & ZYSK, G. 2001. The cell wall-associated serine protease PrtA: a highly conserved virulence factor of *Streptococcus pneumoniae*. *FEMS Microbiol Lett,* 205**,** 99-104.

BLACK, R. E., COUSENS, S., JOHNSON, H. L., LAWN, J. E., RUDAN, I., BASSANI, D. G., JHA, P., CAMPBELL, H., WALKER, C. F., CIBULSKIS, R., EISELE, T., LIU, L., MATHERS, C., CHILD HEALTH EPIDEMIOLOGY REFERENCE GROUP OF, W. H. O. & UNICEF 2010. Global, regional, and national causes of child mortality in 2008: a systematic analysis. *Lancet,* 375**,** 1969-87.

BOGAERT, D., DE GROOT, R. & HERMANS, P. W. 2004a. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis,* 4**,** 144-54.

BOGAERT, D., VEENHOVEN, R. H., SLUIJTER, M., SANDERS, E. A., DE GROOT, R. & HERMANS, P. W. 2004b. Colony blot assay: a useful method to detect multiple pneumococcal serotypes within clinical specimens. *FEMS Immunol Med Microbiol,* 41**,** 259-64.

BOLGER, A. M., LOHSE, M. & USADEL, B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics,* 30**,** 2114-20.

BOLOGA, M., KAMTCHOUA, T., HOPFER, R., SHENG, X., HICKS, B., BIXLER, G., HOU, V., PEHLIC, V., YUAN, T. & GURUNATHAN, S. 2012. Safety and immunogenicity of pneumococcal protein vaccine candidates: monovalent choline-binding protein A (PcpA) vaccine and

bivalent PcpA-pneumococcal histidine triad protein D vaccine. *Vaccine,* 30**,** 7461-8.

BOWERS, J. R., DRIEBE, E. M., NIBECKER, J. L., WOJACK, B. R., SAROVICH, D. S., WONG, A. H., BRZOSKA, P. M., HUBERT, N., KNADLER, A., WATSON, L. M., WAGNER, D. M., FURTADO, M. R., SAUBOLLE, M., ENGELTHALER, D. M. & KEIM, P. S. 2012. Dominance of multidrug resistant CC271 clones in macrolide-resistant *Streptococcus pneumoniae* in Arizona. *BMC Microbiol,* 12**,** 12.

BRADLEY, J. S., BYINGTON, C. L., SHAH, S. S., ALVERSON, B., CARTER, E. R., HARRISON, C., KAPLAN, S. L., MACE, S. E., MCCRACKEN, G. H., JR., MOORE, M. R., ST PETER, S. D., STOCKWELL, J. A., SWANSON, J. T., PEDIATRIC INFECTIOUS DISEASES, S. & THE INFECTIOUS DISEASES SOCIETY OF, A. 2011. Executive summary: the management of community-acquired pneumonia in infants and children older than 3 months of age: clinical practice guidelines by the Pediatric Infectious Diseases Society and the Infectious Diseases Society of America. *Clin Infect Dis,* 53**,** 617-30.

BRANDILEONE, M. C., DI FABIO, J. L., VIEIRA, V. S., ZANELLA, R. C., CASAGRANDE, S. T., PIGNATARI, A. C. & TOMASZ, A. 1998. Geographic distribution of penicillin resistance of *Streptococcus pneumoniae* in Brazil: genetic relatedness. *Microb Drug Resist,* 4**,** 209-17.

BRILES, D. E., CRAIN, M. J., GRAY, B. M., FORMAN, C. & YOTHER, J. 1992. Strong association between capsular type and virulence for mice among human isolates of *Streptococcus pneumoniae*. *Infect Immun,* 60**,** 111-6.

BRILES, D. E., HOLLINGSHEAD, S. K., PATON, J. C., ADES, E. W., NOVAK, L., VAN GINKEL, F. W. & BENJAMIN, W. H., JR. 2003. Immunizations with pneumococcal surface protein A and pneumolysin are protective against pneumonia in a murine model of pulmonary infection with *Streptococcus pneumoniae*. *J Infect Dis,* 188**,** 339-48.

BROOKS-WALTER, A., BRILES, D. E. & HOLLINGSHEAD, S. K. 1999. The pspC gene of *Streptococcus pneumoniae* encodes a polymorphic protein, PspC, which elicits cross-reactive antibodies to PspA and provides immunity to pneumococcal bacteremia. *Infect Immun,* 67**,** 6533-42.

BROUWER, M. C., DE GANS, J., HECKENBERG, S. G., ZWINDERMAN, A. H., VAN DER POLL, T. & VAN DE BEEK, D. 2009. Host genetic susceptibility to pneumococcal and meningococcal disease: a systematic review and meta-analysis. *Lancet Infect Dis,* 9**,** 31-44.

BROWN, J. S., GILLILAND, S. M. & HOLDEN, D. W. 2001. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. *Mol Microbiol,* 40**,** 572-85.

BRUEGGEMANN, A. B., GRIFFITHS, D. T., MEATS, E., PETO, T., CROOK, D. W. & SPRATT, B. G. 2003. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone-specific differences in invasive disease potential. *J Infect Dis,* 187**,** 1424-32.

BRUEGGEMANN, A. B., MUROKI, B. M., KULOHOMA, B. W., KARANI, A., WANJIRU, E., MORPETH, S., KAMAU, T., SHARIF, S. & SCOTT, J. A. 2013. Population genetic structure of *Streptococcus pneumoniae* in Kilifi, Kenya, prior to the introduction of pneumococcal conjugate vaccine. *PLoS One,* 8**,** e81539.

BRUEGGEMANN, A. B., PAI, R., CROOK, D. W. & BEALL, B. 2007. Vaccine escape recombinants emerge after pneumococcal vaccination in the United States. *PLoS Pathog,* 3**,** e168.

BRUEGGEMANN, A. B. & SPRATT, B. G. 2003. Geographic distribution and clonal diversity of *Streptococcus pneumoniae* serotype 1 isolates. *J Clin Microbiol,* 41**,** 4966-70.

BUTLER, J. C., BREIMAN, R. F., CAMPBELL, J. F., LIPMAN, H. B., BROOME, C. V. & FACKLAM, R. R. 1993. Pneumococcal polysaccharide vaccine efficacy. An evaluation of current recommendations. *JAMA,* 270**,** 1826-31.

BYINGTON, C. L., SAMORE, M. H., STODDARD, G. J., BARLOW, S., DALY, J., KORGENSKI, K., FIRTH, S., GLOVER, D., JENSEN, J., MASON, E. O., SHUTT, C. K. & PAVIA, A. T. 2005. Temporal trends of invasive disease due to *Streptococcus pneumoniae* among children in the intermountain west: emergence of nonvaccine serogroups. *Clin Infect Dis,* 41**,** 21-9.

BYINGTON, C. L., SPENCER, L. Y., JOHNSON, T. A., PAVIA, A. T., ALLEN, D., MASON, E. O., KAPLAN, S., CARROLL, K. C., DALY, J. A., CHRISTENSON, J. C. & SAMORE, M. H. 2002. An epidemiological investigation of a sustained high rate of pediatric parapneumonic empyema: risk factors and microbiological associations. *Clin Infect Dis,* 34**,** 434-40.

CABELLOS, C., MACINTYRE, D. E., FORREST, M., BURROUGHS, M., PRASAD, S. & TUOMANEN, E. 1992. Differing roles for platelet-activating factor during inflammation of the lung and subarachnoid space. The special case of *Streptococcus pneumoniae*. *J Clin Invest,* 90**,** 612-8.

CALIX, J. J. & NAHM, M. H. 2010. A new pneumococcal serotype, 11E, has a variably inactivated wcjE gene. *J Infect Dis,* 202**,** 29-38.

CARDOZO, D. M., NASCIMENTO-CARVALHO, C. M., ANDRADE, A.-L. S. S., SILVANY-NETO, A. M., DALTRO, C. H. C., BRANDÃO, M.-A. S., BRANDÃO, A. P. & BRANDILEONE, M.-C. C. 2008. Prevalence and risk factors for nasopharyngeal carriage of *Streptococcus pneumoniae* among adolescents. *Journal of Medical Microbiology,* 57**,** 185-189.

CARLSEN, B. D., KAWANA, M., KAWANA, C., TOMASZ, A. & GIEBINK, G. S. 1992. Role of the bacterial cell wall in middle ear inflammation caused by *Streptococcus pneumoniae*. *Infect Immun,* 60**,** 2850-4.

CARROL, E. D., MANKHAMBO, L. A., GUIVER, M., BANDA, D. L., GROUP, I. P. D. S., DENIS, B., DOVE, W., JEFFERS, G., MOLYNEUX, E. M., MOLYNEUX, M. E., HART, C. A. & GRAHAM, S. M. 2011. PCR improves diagnostic yield from lung aspiration in Malawian children with radiologically confirmed pneumonia. *PLoS One,* 6**,** e21042.

CASADEVALL, A. & PIROFSKI, L. 2001. Host-pathogen interactions: the attributes of virulence. *J Infect Dis,* 184**,** 337-44.

CENTERS FOR DISEASE, C. & PREVENTION 2010. Licensure of a 13-valent pneumococcal conjugate vaccine (PCV13) and recommendations for use among children - Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR Morb Mortal Wkly Rep,* 59**,** 258-61.

CHAGUZA, C., CORNICK, J. E., ANDAM, C. P., GLADSTONE, R. A., ALAERTS, M., MUSICHA, P., PENO, C., BAR-ZEEV, N., KAMNG'ONA, A. W., KIRAN, A. M., MSEFULA, C. L., MCGEE, L., BREIMAN, R. F., KADIOGLU, A., FRENCH, N., HEYDERMAN, R. S., HANAGE, W. P., BENTLEY, S. D. & EVERETT, D. B. 2017. Population genetic structure,

antibiotic resistance, capsule switching and evolution of invasive pneumococci before conjugate vaccination in Malawi. *Vaccine,* 35**,** 4594-4602.

CHAGUZA, C., CORNICK, J. E., HARRIS, S. R., ANDAM, C. P., BRICIO-MORENO, L., YANG, M., YALCIN, F., OUSMANE, S., GOVINDPERSAD, S., SENGHORE, M., EBRUKE, C., DU PLESSIS, M., KIRAN, A. M., PLUSCHKE, G., SIGAUQUE, B., MCGEE, L., KLUGMAN, K. P., TURNER, P., CORANDER, J., PARKHILL, J., COLLARD, J. M., ANTONIO, M., VON GOTTBERG, A., HEYDERMAN, R. S., FRENCH, N., KADIOGLU, A., HANAGE, W. P., EVERETT, D. B., BENTLEY, S. D. & CONSORTIUM, P. A. 2016. Understanding pneumococcal serotype 1 biology through population genomic analysis. *BMC Infect Dis,* 16**,** 649.

CHEN, A., MANN, B., GAO, G., HEATH, R., KING, J., MAISSONEUVE, J., ALDERSON, M., TATE, A., HOLLINGSHEAD, S. K., TWETEN, R. K., BRILES, D. E., TUOMANEN, E. I. & PATON, J. C. 2015. Multivalent Pneumococcal Protein Vaccines Comprising Pneumolysoid with Epitopes/Fragments of CbpA and/or PspA Elicit Strong and Broad Protection. *Clin Vaccine Immunol,* 22**,** 1079-89.

CHEN, F. M., BREIMAN, R. F., FARLEY, M., PLIKAYTIS, B., DEAVER, K. & CETRON, M. S. 1998. Geocoding and linking data from population-based surveillance and the US Census to evaluate the impact of median household income on the epidemiology of invasive *Streptococcus pneumoniae* infections. *Am J Epidemiol,* 148**,** 1212-8.

CHEUNG, Y. B., ZAMAN, S. M., NSEKPONG, E. D., VAN BENEDEN, C. A., ADEGBOLA, R. A., GREENWOOD, B. & CUTTS, F. T. 2009. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian children who participated in a 9-valent pneumococcal conjugate vaccine trial and in their younger siblings. *Pediatr Infect Dis J,* 28**,** 990-5.

CHEWAPREECHA, C., HARRIS, S. R., CROUCHER, N. J., TURNER, C., MARTTINEN, P., CHENG, L., PESSIA, A., AANENSEN, D. M., MATHER, A. E., PAGE, A. J., SALTER, S. J., HARRIS, D., NOSTEN, F., GOLDBLATT, D., CORANDER, J., PARKHILL, J., TURNER, P. & BENTLEY, S. D. 2014a. Dense genomic sampling identifies highways of pneumococcal recombination. *Nat Genet,* 46**,** 305-9.

CHEWAPREECHA, C., HARRIS, S. R., CROUCHER, N. J., TURNER, C., MARTTINEN, P., CHENG, L., PESSIA, A., AANENSEN, D. M., MATHER, A. E., PAGE, A. J., SALTER, S. J., HARRIS, D., NOSTEN, F., GOLDBLATT, D., CORANDER, J., PARKHILL, J., TURNER, P. & BENTLEY, S. D. 2014b. Dense genomic sampling identifies highways of pneumococcal recombination. *Nat Genet,* 46**,** 305-309.

CHEWAPREECHA, C., MARTTINEN, P., CROUCHER, N. J., SALTER, S. J., HARRIS, S. R., MATHER, A. E., HANAGE, W. P., GOLDBLATT, D., NOSTEN, F. H., TURNER, C., TURNER, P., BENTLEY, S. D. & PARKHILL, J. 2014c. Comprehensive identification of single nucleotide polymorphisms associated with beta-lactam resistance within pneumococcal mosaic genes. *PLoS Genet,* 10**,** e1004547.

CHIAVOLINI, D., POZZI, G. & RICCI, S. 2008. Animal models of *Streptococcus pneumoniae* disease. *Clin Microbiol Rev,* 21**,** 666-85.

CHIBA, N., MOROZUMI, M., SHOUJI, M., WAJIMA, T., IWATA, S., UBUKATA, K. & INVASIVE PNEUMOCOCCAL DISEASES SURVEILLANCE STUDY, G. 2014. Changes in capsule and drug resistance of Pneumococci after introduction of PCV7, Japan, 2010-2013. *Emerg Infect Dis,* 20**,** 1132-9.

CHIOU, A. C., ANDRADE, S. S., ALMEIDA, S. C., ZANELLA, R. C., ANDRADE, A. L. & BRANDILEONE, M. C. 2008. Molecular assessment of invasive *Streptococcus pneumoniae* serotype 1 in Brazil: evidence of clonal replacement. *J Med Microbiol,* 57**,** 839-44.

CHOCHUA, S., METCALF, B. J., LI, Z., WALKER, H., TRAN, T., MCGEE, L. & BEALL, B. 2017. Invasive Serotype 35B Pneumococci Including an Expanding Serotype Switch Lineage, United States, 2015-2016. *Emerg Infect Dis,* 23**,** 922-930.

CIRUELA, P., SOLDEVILA, N., HERNANDEZ, S., SELVA, L., DE SEVILLA, M. F., GARCIA-GARCIA, J. J., MORAGA, F., PLANES, A. M., MUNOZ-ALMAGRO, C. & DOMINGUEZ, A. 2013. Risk factors for invasive pneumococcal disease in a community with a high proportion of non vaccine serotypes. *Vaccine,* 31**,** 960-6.

CLSI 2011. Performance Standards for Antimicrobial Susceptibility Testing; Twenty-First Informational Supplement. *CLSI document M100-S21.* Wayne, PA: Clinical and Laboratory Standards Institute.

CLSI 2017. Performance Standards for Antimicrobial Susceptibility Testing. 27th ed. *CLSI supplement M100.* Wayne, PA: Clinical and Laboratory Standards Institute; 2017.

COHEN, R., BISCARDI, S. & LEVY, C. 2016. The multifaceted impact of pneumococcal conjugate vaccine implementation in children in France between 2001 to 2014. *Hum Vaccin Immunother,* 12**,** 277-84.

COLLARD, J. M., ALIO SANDA, A. K. & JUSOT, J. F. 2013. Determination of pneumococcal serotypes in meningitis cases in Niger, 2003-2011. *PLoS One,* 8**,** e60432.

CONNOR, T. R., LOMAN, N. J., THOMPSON, S., SMITH, A., SOUTHGATE, J., POPLAWSKI, R., BULL, M. J., RICHARDSON, E., ISMAIL, M., THOMPSON, S. E., KITCHEN, C., GUEST, M., BAKKE, M., SHEPPARD, S. K. & PALLEN, M. J. 2016. CLIMB (the Cloud Infrastructure for Microbial Bioinformatics): an online resource for the medical microbiology community. *Microb Genom,* 2**,** e000086.

CORNICK, J. E., CHAGUZA, C., HARRIS, S. R., YALCIN, F., SENGHORE, M., KIRAN, A. M., GOVINDPERSHAD, S., OUSMANE, S., PLESSIS, M. D., PLUSCHKE, G., EBRUKE, C., MCGEE, L., SIGAÙQUE, B., COLLARD, J.-M., ANTONIO, M., VON GOTTBERG, A., FRENCH, N., KLUGMAN, K. P., HEYDERMAN, R. S., BENTLEY, S. D., EVERETT, D. B. & CONSORTIUM, F. T. P. 2015. Region-specific diversification of the highly virulent serotype 1 *Streptococcus pneumoniae*. *Microbial Genomics,* 1.

CORNICK, J. E., HARRIS, S. R., PARRY, C. M., MOORE, M. J., JASSI, C., KAMNG'ONA, A., KULOHOMA, B., HEYDERMAN, R. S., BENTLEY, S. D. & EVERETT, D. B. 2014. Genomic identification of a novel co-trimoxazole resistance genotype and its prevalence amongst *Streptococcus pneumoniae* in Malawi. *J Antimicrob Chemother,* 69**,** 368-74.

CORTESE, M. M., WOLFF, M., ALMEIDO-HILL, J., REID, R., KETCHAM, J. & SANTOSHAM, M. 1992. High incidence rates of invasive pneumococcal

disease in the White Mountain Apache population. *Arch Intern Med,* 152**,** 2277-82.

CRAIN, M. J., WALTMAN, W. D., 2ND, TURNER, J. S., YOTHER, J., TALKINGTON, D. F., MCDANIEL, L. S., GRAY, B. M. & BRILES, D. E. 1990. Pneumococcal surface protein A (PspA) is serologically highly variable and is expressed by all clinically important capsular serotypes of *Streptococcus pneumoniae*. *Infect Immun,* 58**,** 3293-9.

CROUCHER, N. J., FINKELSTEIN, J. A., PELTON, S. I., MITCHELL, P. K., LEE, G. M., PARKHILL, J., BENTLEY, S. D., HANAGE, W. P. & LIPSITCH, M. 2013. Population genomics of post-vaccine changes in pneumococcal epidemiology. *Nat Genet,* 45**,** 656-63.

CROUCHER, N. J., HARRIS, S. R., FRASER, C., QUAIL, M. A., BURTON, J., VAN DER LINDEN, M., MCGEE, L., VON GOTTBERG, A., SONG, J. H., KO, K. S., PICHON, B., BAKER, S., PARRY, C. M., LAMBERTSEN, L. M., SHAHINAS, D., PILLAI, D. R., MITCHELL, T. J., DOUGAN, G., TOMASZ, A., KLUGMAN, K. P., PARKHILL, J., HANAGE, W. P. & BENTLEY, S. D. 2011. Rapid pneumococcal evolution in response to clinical interventions. *Science,* 331**,** 430-4.

CROUCHER, N. J., PAGE, A. J., CONNOR, T. R., DELANEY, A. J., KEANE, J. A., BENTLEY, S. D., PARKHILL, J. & HARRIS, S. R. 2015. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res,* 43**,** e15.

CROUCHER, N. J., WALKER, D., ROMERO, P., LENNARD, N., PATERSON, G. K., BASON, N. C., MITCHELL, A. M., QUAIL, M. A., ANDREW, P. W., PARKHILL, J., BENTLEY, S. D. & MITCHELL, T. J. 2009. Role of conjugative elements in the evolution of the multidrug-resistant pandemic clone *Streptococcus pneumoniae*Spain23F ST81. *J Bacteriol,* 191**,** 1480-9.

CUTTS, F. T., ZAMAN, S. M., ENWERE, G., JAFFAR, S., LEVINE, O. S., OKOKO, J. B., OLUWALANA, C., VAUGHAN, A., OBARO, S. K., LEACH, A., MCADAM, K. P., BINEY, E., SAAKA, M., ONWUCHEKWA, U., YALLOP, F., PIERCE, N. F., GREENWOOD, B. M. & ADEGBOLA, R. A. 2005. Efficacy of nine-valent pneumococcal conjugate vaccine against pneumonia and invasive pneumococcal disease in The Gambia: randomised, double-blind, placebo-controlled trial. *Lancet,* 365**,** 1139-46.

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

DAGAN, R., GRADSTEIN, S., BELMAKER, I., PORAT, N., SITON, Y., WEBER, G., JANCO, J. & YAGUPSKY, P. 2000. An outbreak of *Streptococcus pneumoniae* serotype 1 in a closed community in southern Israel. *Clin Infect Dis,* 30**,** 319-21.

DAGAN, R., JUERGENS, C., TRAMMEL, J., PATTERSON, S., GREENBERG, D., GIVON-LAVI, N., PORAT, N., GURTMAN, A., GRUBER, W. C. & SCOTT, D. A. 2015. Efficacy of 13-valent pneumococcal conjugate vaccine (PCV13) versus that of 7-valent PCV (PCV7) against nasopharyngeal colonization of antibiotic-nonsusceptible *Streptococcus pneumoniae*. *J Infect Dis,* 211**,** 1144-53.

DAYIE, N. T., ARHIN, R. E., NEWMAN, M. J., DALSGAARD, A., BISGAARD, M., FRIMODT-MOLLER, N. & SLOTVED, H. C. 2013. Penicillin

resistance and serotype distribution of *Streptococcus pneumoniae* in Ghanaian children less than six years of age. *BMC Infect Dis,* 13**,** 490.

DE ST MAURICE, A., SCHAFFNER, W., GRIFFIN, M. R., HALASA, N. & GRIJALVA, C. G. 2016. Persistent Sex Disparities in Invasive Pneumococcal Diseases in the Conjugate Vaccine Era. *J Infect Dis,* 214**,** 792-7.

DEL AMO, E., ESTEVA, C., HERNANDEZ-BOU, S., GALLES, C., NAVARRO, M., SAUCA, G., DIAZ, A., GASSIOT, P., MARTI, C., LARROSA, N., CIRUELA, P., JANE, M., SA-LEAO, R., MUNOZ-ALMAGRO, C. & CATALAN STUDY GROUP OF INVASIVE PNEUMOCOCCAL, D. 2016. Serotypes and Clonal Diversity of *Streptococcus pneumoniae* Causing Invasive Disease in the Era of PCV13 in Catalonia, Spain. *PLoS One,* 11**,** e0151125.

DEMARIA, A., JR., BROWNE, K., BERK, S. L., SHERWOOD, E. J. & MCCABE, W. R. 1980. An outbreak of type 1 pneumococcal pneumonia in a men's shelter. *JAMA,* 244**,** 1446-9.

DENG, X., PEIRANO, G., SCHILLBERG, E., MAZZULLI, T., GRAY-OWEN, S. D., WYLIE, J. L., ROBINSON, D. A., MAHMUD, S. M. & PILLAI, D. R. 2016. Whole-Genome Sequencing Reveals the Origin and Rapid Evolution of an Emerging Outbreak Strain of *Streptococcus pneumoniae* 12F. *Clin Infect Dis,* 62**,** 1126-32.

DEVINE, V. T., CLEARY, D. W., JEFFERIES, J. M., ANDERSON, R., MORRIS, D. E., TUCK, A. C., GLADSTONE, R. A., O'DOHERTY, G., KURUPARAN, P., BENTLEY, S. D., FAUST, S. N. & CLARKE, S. C. 2017. The rise and fall of pneumococcal serotypes carried in the PCV era. *Vaccine,* 35**,** 1293-1298.

DHOUBHADEL, B. G., YASUNAMI, M., NGUYEN, H. A., SUZUKI, M., VU, T. H., THI THUY NGUYEN, A., DANG, D. A., YOSHIDA, L. M. & ARIYOSHI, K. 2014. Bacterial load of pneumococcal serotypes correlates with their prevalence and multiple serotypes is associated with acute respiratory infections among children less than 5 years of age. *PLoS One,* 9**,** e110777.

DI GUILMI, A. M. & DESSEN, A. 2002. New approaches towards the identification of antibiotic and vaccine targets in *Streptococcus pneumoniae*. *EMBO Rep,* 3**,** 728-34.

DIAS, C. A., TEIXEIRA, L. M., CARVALHO MDA, G. & BEALL, B. 2007. Sequential multiplex PCR for determining capsular serotypes of pneumococci recovered from Brazilian children. *J Med Microbiol,* 56**,** 1185-8.

DONATI, C., HILLER, N. L., TETTELIN, H., MUZZI, A., CROUCHER, N. J., ANGIUOLI, S. V., OGGIONI, M., DUNNING HOTOPP, J. C., HU, F. Z., RILEY, D. R., COVACCI, A., MITCHELL, T. J., BENTLEY, S. D., KILIAN, M., EHRLICH, G. D., RAPPUOLI, R., MOXON, E. R. & MASIGNANI, V. 2010. Structure and dynamics of the pan-genome of *Streptococcus pneumoniae* and closely related species. *Genome Biol,* 11**,** R107.

DONKOR, E. S., ADEGBOLA, R. A., WREN, B. W. & ANTONIO, M. 2013. Population biology of *Streptococcus pneumoniae* in West Africa: multilocus sequence typing of serotypes that exhibit different predisposition to invasive disease and carriage. *PLoS One,* 8**,** e53925.

DOUGLAS, R. M., PATON, J. C., DUNCAN, S. J. & HANSMAN, D. J. 1983. Antibody response to pneumococcal vaccination in children younger than five years of age. *J Infect Dis,* 148**,** 131-7.

DOWSON, C. G., COFFEY, T. J., KELL, C. & WHILEY, R. A. 1993. Evolution of penicillin resistance in *Streptococcus pneumoniae*; the role of Streptococcus mitis in the formation of a low affinity PBP2B in S. pneumoniae. *Mol Microbiol,* 9**,** 635-43.

DOWSON, C. G., HUTCHISON, A., BRANNIGAN, J. A., GEORGE, R. C., HANSMAN, D., LINARES, J., TOMASZ, A., SMITH, J. M. & SPRATT, B. G. 1989. Horizontal transfer of penicillin-binding protein genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. *Proc Natl Acad Sci U S A,* 86**,** 8842-6.

DU PLESSIS, M., ALLAM, M., TEMPIA, S., WOLTER, N., DE GOUVEIA, L., VON MOLLENDORF, C., JOLLEY, K. A., MBELLE, N., WADULA, J., CORNICK, J. E., EVERETT, D. B., MCGEE, L., BREIMAN, R. F., GLADSTONE, R. A., BENTLEY, S. D., KLUGMAN, K. P. & VON GOTTBERG, A. 2016. Phylogenetic Analysis of Invasive Serotype 1 Pneumococcus in South Africa, 1989 to 2013. *J Clin Microbiol,* 54**,** 1326-34.

EBRUKE, C., ROCA, A., EGERE, U., DARBOE, O., HILL, P. C., GREENWOOD, B., WREN, B. W., ADEGBOLA, R. A. & ANTONIO, M. 2015. Temporal changes in nasopharyngeal carriage of *Streptococcus pneumoniae* serotype 1 genotypes in healthy Gambians before and after the 7-valent pneumococcal conjugate vaccine. *PeerJ,* 3**,** e903.

EDMOND, K., CLARK, A., KORCZAK, V. S., SANDERSON, C., GRIFFITHS, U. K. & RUDAN, I. 2010. Global and regional risk of disabling sequelae from bacterial meningitis: a systematic review and meta-analysis. *Lancet Infect Dis,* 10**,** 317-28.

EL MDAGHRI, N., JILALI, N., BELABBES, H., JOUHADI, Z., LAHSSOUNE, M. & ZAID, S. 2012. Epidemiological profile of invasive bacterial diseases in children in Casablanca, Morocco: antimicrobial susceptibilities and serotype distribution. *East Mediterr Health J,* 18**,** 1097-101.

ELTRINGHAM, G., KEARNS, A., FREEMAN, R., CLARK, J., SPENCER, D., EASTHAM, K., HARWOOD, J. & LEEMING, J. 2003. Culture-Negative Childhood Empyema Is Usually Due to Penicillin-Sensitive *Streptococcus pneumoniae* Capsular Serotype 1. *Journal of Clinical Microbiology,* 41**,** 521-522.

ENRIGHT, M. C. & SPRATT, B. G. 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology,* 144 (Pt 11)**,** 3049-60.

EVANS, B. A. & ROZEN, D. E. 2012. A *Streptococcus pneumoniae* infection model in larvae of the wax moth Galleria mellonella. *Eur J Clin Microbiol Infect Dis*.

EVERETT, D. B., CORNICK, J., DENIS, B., CHEWAPREECHA, C., CROUCHER, N., HARRIS, S., PARKHILL, J., GORDON, S., CARROL, E. D., FRENCH, N., HEYDERMAN, R. S. & BENTLEY, S. D. 2012. Genetic characterisation of Malawian pneumococci prior to the roll-out of the PCV13 vaccine using a high-throughput whole genome sequencing approach. *PLoS One,* 7**,** e44250.

FANI, F., LEPROHON, P., LEGARE, D. & OUELLETTE, M. 2011. Whole genome sequencing of penicillin-resistant *Streptococcus pneumoniae* reveals

mutations in penicillin-binding proteins and in a putative iron permease. *Genome Biol,* 12**,** R115.

FEIKIN, D. R., DAVIS, M., NWANYANWU, O. C., KAZEMBE, P. N., BARAT, L. M., WASAS, A., BLOLAND, P. B., ZIBA, C., CAPPER, T., HUEBNER, R. E., SCHWARTZ, B., KLUGMAN, K. P. & DOWELL, S. F. 2003. Antibiotic resistance and serotype distribution of *Streptococcus pneumoniae* colonizing rural Malawian children. *Pediatr Infect Dis J,* 22**,** 564-7.

FEIKIN, D. R., KAGUCIA, E. W., LOO, J. D., LINK-GELLES, R., PUHAN, M. A., CHERIAN, T., LEVINE, O. S., WHITNEY, C. G., O'BRIEN, K. L., MOORE, M. R. & SEROTYPE REPLACEMENT STUDY, G. 2013. Serotype-specific changes in invasive pneumococcal disease after pneumococcal conjugate vaccine introduction: a pooled analysis of multiple surveillance sites. *PLoS Med,* 10**,** e1001517.

FEIL, E. J., ENRIGHT, M. C. & SPRATT, B. G. 2000. Estimating the relative contributions of mutation and recombination to clonal diversification: a comparison between Neisseria meningitidis and *Streptococcus pneumoniae*. *Res Microbiol,* 151**,** 465-9.

FELDMAN, C., MITCHELL, T. J., ANDREW, P. W., BOULNOIS, G. J., READ, R. C., TODD, H. C., COLE, P. J. & WILSON, R. 1990. The effect of *Streptococcus pneumoniae* pneumolysin on human respiratory epithelium in vitro. *Microb Pathog,* 9**,** 275-84.

FINLAND, M. & BARNES, M. W. 1977. Changes in occurrence of capsular serotypes of *Streptococcus pneumoniae* at Boston City Hospital during selected years between 1935 and 1974. *J Clin Microbiol,* 5**,** 154-66.

FRANKEL, R. E., VIRATA, M., HARDALO, C., ALTICE, F. L. & FRIEDLAND, G. 1996. Invasive pneumococcal disease: clinical features, serotypes, and antimicrobial resistance patterns in cases involving patients with and without human immunodeficiency virus infection. *Clin Infect Dis,* 23**,** 577-84.

GAO, W., CHUA, K., DAVIES, J. K., NEWTON, H. J., SEEMANN, T., HARRISON, P. F., HOLMES, N. E., RHEE, H. W., HONG, J. I., HARTLAND, E. L., STINEAR, T. P. & HOWDEN, B. P. 2010. Two novel point mutations in clinical Staphylococcus aureus reduce linezolid susceptibility and switch on the stringent response to promote persistent infection. *PLoS Pathog,* 6**,** e1000944.

GEELEN, S., BHATTACHARYYA, C. & TUOMANEN, E. 1993. The cell wall mediates pneumococcal attachment to and cytopathology in human endothelial cells. *Infect Immun,* 61**,** 1538-43.

GERTZ, R. E., JR., LI, Z., PIMENTA, F. C., JACKSON, D., JUNI, B. A., LYNFIELD, R., JORGENSEN, J. H., CARVALHO MDA, G., BEALL, B. W. & ACTIVE BACTERIAL CORE SURVEILLANCE, T. 2010. Increased penicillin nonsusceptibility of nonvaccine-serotype invasive pneumococci other than serotypes 19A and 6A in post-7-valent conjugate vaccine era. *J Infect Dis,* 201**,** 770-5.

GESSNER, B. D., MUELLER, J. E. & YARO, S. 2010. African meningitis belt pneumococcal disease epidemiology indicates a need for an effective serotype 1 containing vaccine, including for older children and adults. *BMC Infect Dis,* 10**,** 22.

GHERARDI, G., D'AMBROSIO, F., VISAGGIO, D., DICUONZO, G., DEL GROSSO, M. & PANTOSTI, A. 2012. Serotype and clonal evolution of penicillin-nonsusceptible invasive *Streptococcus pneumoniae* in the 7-valent

pneumococcal conjugate vaccine era in Italy. *Antimicrob Agents Chemother,* 56**,** 4965-8.

GIEBINK, G. S., VERHOEF, J., PETERSON, P. K. & QUIE, P. G. 1977. Opsonic requirements for phagocytosis of *Streptococcus pneumoniae* types VI, XVIII, XXIII, and XXV. *Infect Immun,* 18**,** 291-7.

GINGLES, N. A., ALEXANDER, J. E., KADIOGLU, A., ANDREW, P. W., KERR, A., MITCHELL, T. J., HOPES, E., DENNY, P., BROWN, S., JONES, H. B., LITTLE, S., BOOTH, G. C. & MCPHEAT, W. L. 2001. Role of genetic resistance in invasive pneumococcal infection: identification and study of susceptibility and resistance in inbred mouse strains. *Infect Immun,* 69**,** 426-34.

GLADSTONE, R. A., GRITZFELD, J. F., COUPLAND, P., GORDON, S. B. & BENTLEY, S. D. 2015a. Genetic stability of pneumococcal isolates during 35 days of human experimental carriage. *Vaccine,* 33**,** 3342-5.

GLADSTONE, R. A., JEFFERIES, J. M., TOCHEVA, A. S., BEARD, K. R., GARLEY, D., CHONG, W. W., BENTLEY, S. D., FAUST, S. N. & CLARKE, S. C. 2015b. Five winters of pneumococcal serotype replacement in UK carriage following PCV introduction. *Vaccine,* 33**,** 2015-21.

GOETGHEBUER, T., WEST, T. E., WERMENBOL, V., CADBURY, A. L., MILLIGAN, P., LLOYD-EVANS, N., ADEGBOLA, R. A., MULHOLLAND, E. K., GREENWOOD, B. M. & WEBER, M. W. 2000. Outcome of meningitis caused by *Streptococcus pneumoniae* and *Haemophilus influenzae* type b in children in The Gambia. *Trop Med Int Health,* 5**,** 207-13.

GOLUBCHIK, T., BRUEGGEMANN, A. B., STREET, T., GERTZ, R. E., JR., SPENCER, C. C., HO, T., GIANNOULATOU, E., LINK-GELLES, R., HARDING, R. M., BEALL, B., PETO, T. E., MOORE, M. R., DONNELLY, P., CROOK, D. W. & BOWDEN, R. 2012. Pneumococcal genome sequencing tracks a vaccine escape variant formed through a multi-fragment recombination event. *Nat Genet,* 44**,** 352-5.

GONCALVES, V. M., DIAS, W. O., CAMPOS, I. B., LIBERMAN, C., SBROGIO-ALMEIDA, M. E., SILVA, E. P., CARDOSO, C. P., JR., ALDERSON, M., ROBERTSON, G., MAISONNEUVE, J. F., TATE, A., ANDERSON, P., MALLEY, R., FRATELLI, F. & LEITE, L. C. 2014. Development of a whole cell pneumococcal vaccine: BPL inactivation, cGMP production, and stability. *Vaccine,* 32**,** 1113-20.

GORDON, M. A., WALSH, A. L., CHAPONDA, M., SOKO, D., MBVWINJI, M., MOLYNEUX, M. E. & GORDON, S. B. 2001. Bacteraemia and mortality among adult medical admissions in Malawi--predominance of non-typhi salmonellae and *Streptococcus pneumoniae*. *J Infect,* 42**,** 44-9.

GOSINK, K. K., MANN, E. R., GUGLIELMO, C., TUOMANEN, E. I. & MASURE, H. R. 2000. Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. *Infect Immun,* 68**,** 5690-5.

GRABENSTEIN, J. D. & KLUGMAN, K. P. 2012. A century of pneumococcal vaccination research in humans. *Clin Microbiol Infect,* 18 Suppl 5**,** 15-24.

GRATTEN, M., MOREY, F., DIXON, J., MANNING, K., TORZILLO, P., MATTERS, R., ERLICH, J., HANNA, J., ASCHE, V. & RILEY, I. 1993. An outbreak of serotype 1 *Streptococcus pneumoniae* infection in central Australia. *Med J Aust,* 158**,** 340-2.

GRAY, B. M., CONVERSE, G. M., 3RD & DILLON, H. C., JR. 1980. Epidemiologic studies of *Streptococcus pneumoniae* in infants: acquisition, carriage, and infection during the first 24 months of life. *J Infect Dis,* 142**,** 923-33.

GREENWOOD, B. 1999. The epidemiology of pneumococcal infection in children in the developing world. *Philos Trans R Soc Lond B Biol Sci,* 354**,** 777-85.

GRIFFITH, F. 1928. The Significance of Pneumococcal Types. *J Hyg (Lond),* 27**,** 113-59.

GRITZFELD, J. F., WRIGHT, A. D., COLLINS, A. M., PENNINGTON, S. H., WRIGHT, A. K., KADIOGLU, A., FERREIRA, D. M. & GORDON, S. B. 2013. Experimental human pneumococcal carriage. *J Vis Exp*.

HABIB, M., PORTER, B. D. & SATZKE, C. 2014. Capsular Serotyping of *Streptococcus pneumoniae* Using the Quellung Reaction. *Journal of Visualized Experiments : JoVE***,** 51208.

HAHN, C., HARRISON, E. M., PARKHILL, J., HOLMES, M. A. & PATERSON, G. K. 2015. Draft Genome Sequence of the *Streptococcus pneumoniae* Avery Strain A66. *Genome Announc,* 3.

HAKENBECK, R., GREBE, T., ZAHNER, D. & STOCK, J. B. 1999. beta-lactam resistance in *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-binding proteins. *Mol Microbiol,* 33**,** 673-8.

HAKENBECK, R., MADHOUR, A., DENAPAITE, D. & BRUCKNER, R. 2009. Versatility of choline metabolism and choline-binding proteins in *Streptococcus pneumoniae* and commensal streptococci. *FEMS Microbiol Rev,* 33**,** 572-86.

HAMMITT, L. L., AKECH, D. O., MORPETH, S. C., KARANI, A., KIHUHA, N., NYONGESA, S., BWANAALI, T., MUMBO, E., KAMAU, T., SHARIF, S. K. & SCOTT, J. A. 2014. Population effect of 10-valent pneumococcal conjugate vaccine on nasopharyngeal carriage of *Streptococcus pneumoniae* and non-typeable *Haemophilus influenzae* in Kilifi, Kenya: findings from cross-sectional carriage studies. *Lancet Glob Health,* 2**,** e397-405.

HANAGE, W. P., FRASER, C., TANG, J., CONNOR, T. R. & CORANDER, J. 2009. Hyper-recombination, diversity, and antibiotic resistance in pneumococcus. *Science,* 324**,** 1454-7.

HANAGE, W. P., KAIJALAINEN, T. H., SYRJANEN, R. K., AURANEN, K., LEINONEN, M., MAKELA, P. H. & SPRATT, B. G. 2005. Invasiveness of serotypes and clones of *Streptococcus pneumoniae* among children in Finland. *Infect Immun,* 73**,** 431-5.

HANQUET, G., KISSLING, E., FENOLL, A., GEORGE, R., LEPOUTRE, A., LERNOUT, T., TARRAGÓ, D., VARON, E. & VERHAEGEN, J. 2010. Pediatric Pneumococcal Serotypes in 4 European Countries. *Emerging Infectious Diseases,* 16**,** 1428-1439.

HARVEY, R. M., OGUNNIYI, A. D., CHEN, A. Y. & PATON, J. C. 2011. Pneumolysin with low hemolytic activity confers an early growth advantage to *Streptococcus pneumoniae* in the blood. *Infect Immun,* 79**,** 4122-30.

HAUSDORFF, W. P., BRYANT, J., PARADISO, P. R. & SIBER, G. R. 2000. Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. *Clin Infect Dis,* 30**,** 100-21.

HAVA, D. L. & CAMILLI, A. 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. *Mol Microbiol,* 45**,** 1389-406.

HENRICHSEN, J. 1995. Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol,* 33**,** 2759-62.

HENRIQUES-NORMARK, B., BLOMBERG, C., DAGERHAMN, J., BATTIG, P. & NORMARK, S. 2008. The rise and fall of bacterial clones: *Streptococcus pneumoniae*. *Nat Rev Microbiol,* 6**,** 827-37.

HENRIQUES NORMARK, B., KALIN, M., ORTQVIST, A., AKERLUND, T., LILJEQUIST, B. O., HEDLUND, J., SVENSON, S. B., ZHOU, J., SPRATT, B. G., NORMARK, S. & KALLENIUS, G. 2001. Dynamics of penicillin-susceptible clones in invasive pneumococcal disease. *J Infect Dis,* 184**,** 861-9.

HERMANS, P. W., ADRIAN, P. V., ALBERT, C., ESTEVAO, S., HOOGENBOEZEM, T., LUIJENDIJK, I. H., KAMPHAUSEN, T. & HAMMERSCHMIDT, S. 2006. The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. *J Biol Chem,* 281**,** 968-76.

HILL, P. C., AKISANYA, A., SANKAREH, K., CHEUNG, Y. B., SAAKA, M., LAHAI, G., GREENWOOD, B. M. & ADEGBOLA, R. A. 2006. Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian villagers. *Clin Infect Dis,* 43**,** 673-9.

HILL, P. C., TOWNEND, J., ANTONIO, M., AKISANYA, B., EBRUKE, C., LAHAI, G., GREENWOOD, B. M. & ADEGBOLA, R. A. 2010. Transmission of *Streptococcus pneumoniae* in rural Gambian villages: a longitudinal study. *Clin Infect Dis,* 50**,** 1468-76.

HJALMARSDOTTIR, M. A. & KRISTINSSON, K. G. 2014. Epidemiology of penicillin-non-susceptible pneumococci in Iceland, 1995-2010. *J Antimicrob Chemother,* 69**,** 940-6.

HOLLIMAN, R. E., LIDDY, H., JOHNSON, J. D. & ADJEI, O. 2007. Epidemiology of invasive pneumococcal disease in Kumasi, Ghana. *Trans R Soc Trop Med Hyg,* 101**,** 405-13.

HOLMES, A. R., MCNAB, R., MILLSAP, K. W., ROHDE, M., HAMMERSCHMIDT, S., MAWDSLEY, J. L. & JENKINSON, H. F. 2001. The pavA gene of *Streptococcus pneumoniae* encodes a fibronectin-binding protein that is essential for virulence. *Mol Microbiol,* 41**,** 1395-408.

HOTOMI, M., NAKAJIMA, K., HIRAOKA, M., NAHM, M. H. & YAMANAKA, N. 2016. Molecular epidemiology of nonencapsulated *Streptococcus pneumoniae* among Japanese children with acute otitis media. *J Infect Chemother,* 22**,** 72-7.

HYAMS, C., CAMBERLEIN, E., COHEN, J. M., BAX, K. & BROWN, J. S. 2010. The *Streptococcus pneumoniae* capsule inhibits complement activity and neutrophil phagocytosis by multiple mechanisms. *Infect Immun,* 78**,** 704-15.

IANNELLI, F., CHIAVOLINI, D., RICCI, S., OGGIONI, M. R. & POZZI, G. 2004. Pneumococcal surface protein C contributes to sepsis caused by *Streptococcus pneumoniae* in mice. *Infect Immun,* 72**,** 3077-80.

IBRAHIM, Y. M., KERR, A. R., MCCLUSKEY, J. & MITCHELL, T. J. 2004. Role of HtrA in the Virulence and Competence of *Streptococcus pneumoniae*. *Infection and Immunity,* 72**,** 3584-3591.

IP, M., MA, H., LI, C., TSUI, S. & ZHOU, H. 2015. Draft Genome Sequences of Two *Streptococcus pneumoniae* Serotype 19F Sequence Type 271 Clinical Isolates with Low- and High-Level Cefotaxime Resistance. *Genome Announc,* 3.

IRAURGUI, P., TORRES, M. J., GANDIA, A., VAZQUEZ, I., CABRERA, E. G., OBANDO, I., GARNACHO, J. & AZNAR, J. 2010. Modified sequential multiplex PCR for determining capsular serotypes of invasive pneumococci recovered from Seville. *Clin Microbiol Infect,* 16**,** 1504-7.

IROEZINDU, M. O., CHIMA, E. I., ISIGUZO, G. C., MBATA, G. C., ONYEDUM, C. C., ONYEDIBE, K. I. & OKOLI, L. E. 2014. Sputum bacteriology and antibiotic sensitivity patterns of community-acquired pneumonia in hospitalized adult patients in Nigeria: a 5-year multicentre retrospective study. *Scand J Infect Dis,* 46**,** 875-87.

JAIN, A., KUMAR, P. & AWASTHI, S. 2005. High nasopharyngeal carriage of drug resistant *Streptococcus pneumoniae* and *Haemophilus influenzae* in North Indian schoolchildren. *Trop Med Int Health,* 10**,** 234-9.

JANDER, G., RAHME, L. G. & AUSUBEL, F. M. 2000. Positive correlation between virulence of Pseudomonas aeruginosa mutants in mice and insects. *J Bacteriol,* 182**,** 3843-5.

JANOIR, C., LEPOUTRE, A., GUTMANN, L. & VARON, E. 2016. Insight Into Resistance Phenotypes of Emergent Non 13-valent Pneumococcal Conjugate Vaccine Type Pneumococci Isolated From Invasive Disease After 13-valent Pneumococcal Conjugate Vaccine Implementation in France. *Open Forum Infect Dis,* 3**,** ofw020.

JEFFERIES, J. M., JOHNSTON, C. H., KIRKHAM, L. A., COWAN, G. J., ROSS, K. S., SMITH, A., CLARKE, S. C., BRUEGGEMANN, A. B., GEORGE, R. C., PICHON, B., PLUSCHKE, G., PFLUGER, V. & MITCHELL, T. J. 2007. Presence of nonhemolytic pneumolysin in serotypes of *Streptococcus pneumoniae* associated with disease outbreaks. *J Infect Dis,* 196**,** 936-44.

JEFFERIES, J. M., SMITH, A. J., EDWARDS, G. F., MCMENAMIN, J., MITCHELL, T. J. & CLARKE, S. C. 2010. Temporal analysis of invasive pneumococcal clones from Scotland illustrates fluctuations in diversity of serotype and genotype in the absence of pneumococcal conjugate vaccine. *J Clin Microbiol,* 48**,** 87-96.

JOHNSTON, C., HINDS, J., SMITH, A., VAN DER LINDEN, M., VAN ELDERE, J. & MITCHELL, T. J. 2010. Detection of large numbers of pneumococcal virulence genes in streptococci of the mitis group. *J Clin Microbiol,* 48**,** 2762-9.

JOLOBA, M. L., BAJAKSOUZIAN, S., PALAVECINO, E., WHALEN, C. & JACOBS, M. R. 2001. High prevalence of carriage of antibiotic-resistant *Streptococcus pneumoniae* in children in Kampala Uganda. *Int J Antimicrob Agents,* 17**,** 395-400.

KADIOGLU, A., CUPPONE, A. M., TRAPPETTI, C., LIST, T., SPREAFICO, A., POZZI, G., ANDREW, P. W. & OGGIONI, M. R. 2011. Sex-based differences in susceptibility to respiratory and systemic pneumococcal disease in mice. *J Infect Dis,* 204**,** 1971-9.

KADIOGLU, A., WEISER, J. N., PATON, J. C. & ANDREW, P. W. 2008. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol,* 6**,** 288-301.

KAMNG'ONA, A. W., HINDS, J., BAR-ZEEV, N., GOULD, K. A., CHAGUZA, C., MSEFULA, C., CORNICK, J. E., KULOHOMA, B. W., GRAY, K., BENTLEY, S. D., FRENCH, N., HEYDERMAN, R. S. & EVERETT, D. B. 2015. High multiple carriage and emergence of *Streptococcus pneumoniae* vaccine serotype variants in Malawian children. *BMC Infect Dis,* 15**,** 234.

KANDASAMY, R., GURUNG, M., THAPA, A., NDIMAH, S., ADHIKARI, N., MURDOCH, D. R., KELLY, D. F., WALDRON, D. E., GOULD, K. A., THORSON, S., SHRESTHA, S., HINDS, J. & POLLARD, A. J. 2015. Multi-serotype pneumococcal nasopharyngeal carriage prevalence in vaccine naive Nepalese children, assessed using molecular serotyping. *PLoS One,* 10**,** e0114286.

KAPLAN, S. L., MASON, E. O., JR., WALD, E. R., SCHUTZE, G. E., BRADLEY, J. S., TAN, T. Q., HOFFMAN, J. A., GIVNER, L. B., YOGEV, R. & BARSON, W. J. 2004. Decrease of invasive pneumococcal infections in children among 8 children's hospitals in the United States after the introduction of the 7-valent pneumococcal conjugate vaccine. *Pediatrics,* 113**,** 443-9.

KAZANJIAN, P. 2004. Changing interest among physicians toward pneumococcal vaccination throughout the twentieth century. *J Hist Med Allied Sci,* 59**,** 555-87.

KELLY, T., DILLARD, J. P. & YOTHER, J. 1994. Effect of genetic switching of capsular type on virulence of *Streptococcus pneumoniae*. *Infect Immun,* 62**,** 1813-9.

KHAN, M. N. & PICHICHERO, M. E. 2013. CD4 T cell memory and antibody responses directed against the pneumococcal histidine triad proteins PhtD and PhtE following nasopharyngeal colonization and immunization and their role in protection against pneumococcal colonization in mice. *Infect Immun,* 81**,** 3781-92.

KHAN, M. N., SHARMA, S. K., FILKINS, L. M. & PICHICHERO, M. E. 2012. PcpA of *Streptococcus pneumoniae* mediates adherence to nasopharyngeal and lung epithelial cells and elicits functional antibodies in humans. *Microbes Infect,* 14**,** 1102-10.

KILIAN, M., RILEY, D. R., JENSEN, A., BRUGGEMANN, H. & TETTELIN, H. 2014. Parallel evolution of *Streptococcus pneumoniae* and Streptococcus mitis to pathogenic and mutualistic lifestyles. *MBio,* 5**,** e01490-14.

KIRKHAM, L. A., JEFFERIES, J. M., KERR, A. R., JING, Y., CLARKE, S. C., SMITH, A. & MITCHELL, T. J. 2006. Identification of invasive serotype 1 pneumococcal isolates that express nonhemolytic pneumolysin. *J Clin Microbiol,* 44**,** 151-9.

KLUGMAN, K. P., MADHI, S. A., ADEGBOLA, R. A., CUTTS, F., GREENWOOD, B. & HAUSDORFF, W. P. 2011. Timing of serotype 1 pneumococcal disease suggests the need for evaluation of a booster dose. *Vaccine,* 29**,** 3372-3.

KOBAYASHI, M., CONKLIN, L. M., BIGOGO, G., JAGERO, G., HAMPTON, L., FLEMING-DUTRA, K. E., JUNGHAE, M., CARVALHO, M. D., PIMENTA, F., BEALL, B., TAYLOR, T., LASERSON, K. F., VULULE, J., VAN BENEDEN, C., KIM, L., FEIKIN, D. R., WHITNEY, C. G. & BREIMAN, R. F. 2017. Pneumococcal carriage and antibiotic susceptibility patterns from two cross-sectional colonization surveys among children aged <5 years prior to the introduction of 10-valent pneumococcal conjugate vaccine - Kenya, 2009-2010. *BMC Infect Dis,* 17**,** 25.

KOECK, J. L., NJANPOP-LAFOURCADE, B. M., CADE, S., VARON, E., SANGARE, L., VALJEVAC, S., VERGNAUD, G. & POURCEL, C. 2005. Evaluation and selection of tandem repeat loci for *Streptococcus pneumoniae* MLVA strain typing. *BMC Microbiol,* 5**,** 66.

KONRADSEN, H. B. & KALTOFT, M. S. 2002. Invasive pneumococcal infections in Denmark from 1995 to 1999: epidemiology, serotypes, and resistance. *Clin Diagn Lab Immunol,* 9**,** 358-65.

KRONE, C. L., WYLLIE, A. L., VAN BEEK, J., ROTS, N. Y., OJA, A. E., CHU, M. L., BRUIN, J. P., BOGAERT, D., SANDERS, E. A. & TRZCINSKI, K. 2015. Carriage of *Streptococcus pneumoniae* in aged adults with influenza-like-illness. *PLoS One,* 10**,** e0119875.

KRONENBERG, A., ZUCS, P., DROZ, S. & MUHLEMANN, K. 2006. Distribution and invasiveness of *Streptococcus pneumoniae* serotypes in Switzerland, a country with low antibiotic selection pressure, from 2001 to 2004. *J Clin Microbiol,* 44**,** 2032-8.

KUO, J., DOUGLAS, M., REE, H. K. & LINDBERG, A. A. 1995. Characterization of a recombinant pneumolysin and its use as a protein carrier for pneumococcal type 18C conjugate vaccines. *Infect Immun,* 63**,** 2706-13.

KWAMBANA-ADAMS, B. A., ASIEDU-BEKOE, F., SARKODIE, B., AFREH, O. K., KUMA, G. K., OWUSU-OKYERE, G., FOSTER-NYARKO, E., OHENE, S. A., OKOT, C., WORWUI, A. K., OKOI, C., SENGHORE, M., OTU, J. K., EBRUKE, C., BANNERMAN, R., AMPONSA-ACHIANO, K., OPARE, D., KAY, G., LETSA, T., KALUWA, O., APPIAH-DENKYIRA, E., BAMPOE, V., ZAMAN, S. M., PALLEN, M. J., D'ALESSANDRO, U., MWENDA, J. M. & ANTONIO, M. 2016. An outbreak of pneumococcal meningitis among older children (>/=5 years) and adults after the implementation of an infant vaccination programme with the 13-valent pneumococcal conjugate vaccine in Ghana. *BMC Infect Dis,* 16**,** 575.

KYAW, M. H., LYNFIELD, R., SCHAFFNER, W., CRAIG, A. S., HADLER, J., REINGOLD, A., THOMAS, A. R., HARRISON, L. H., BENNETT, N. M., FARLEY, M. M., FACKLAM, R. R., JORGENSEN, J. H., BESSER, J., ZELL, E. R., SCHUCHAT, A., WHITNEY, C. G. & ACTIVE BACTERIAL CORE SURVEILLANCE OF THE EMERGING INFECTIONS PROGRAM, N. 2006. Effect of introduction of the pneumococcal conjugate vaccine on drug-resistant *Streptococcus pneumoniae*. *N Engl J Med,* 354**,** 1455-63.

LAABEI, M., RECKER, M., RUDKIN, J. K., ALDELJAWI, M., GULAY, Z., SLOAN, T. J., WILLIAMS, P., ENDRES, J. L., BAYLES, K. W., FEY, P. D., YAJJALA, V. K., WIDHELM, T., HAWKINS, E., LEWIS, K., PARFETT, S., SCOWEN, L., PEACOCK, S. J., HOLDEN, M., WILSON, D., READ, T. D., VAN DEN ELSEN, J., PRIEST, N. K., FEIL, E. J., HURST, L. D., JOSEFSSON, E. & MASSEY, R. C. 2014. Predicting the virulence of MRSA from its genome sequence. *Genome Res,* 24**,** 839-49.

LAI, J. Y., COOK, H., YIP, T. W., BERTHELSEN, J., GOURLEY, S., KRAUSE, V., SMITH, H., LEACH, A. J. & SMITH-VAUGHAN, H. 2013. Surveillance of pneumococcal serotype 1 carriage during an outbreak of serotype 1 invasive pneumococcal disease in central Australia 2010-2012. *BMC Infect Dis,* 13**,** 409.

LALITHA, M. K., PAI, R., JOHN, T. J., THOMAS, K., JESUDASON, M. V., BRAHMADATHAN, K. N., SRIDHARAN, G. & STEINHOFF, M. C. 1996. Serotyping of *Streptococcus pneumoniae* by agglutination assays: a cost-effective technique for developing countries. *Bull World Health Organ,* 74**,** 387-90.

LANIE, J. A., NG, W. L., KAZMIERCZAK, K. M., ANDRZEJEWSKI, T. M., DAVIDSEN, T. M., WAYNE, K. J., TETTELIN, H., GLASS, J. I. & WINKLER, M. E. 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. *J Bacteriol,* 189**,** 38-51.

LANKINEN, K. S., RINTAMAKI, S., SYRJANEN, R., KILPI, T., RUUTU, P. & LEINONEN, M. 2004. Type-specific enzyme immunoassay for detection of pneumococcal capsular polysaccharide antigens in nasopharyngeal specimens. *J Microbiol Methods,* 56**,** 193-9.

LAXMINARAYAN, R., DUSE, A., WATTAL, C., ZAIDI, A. K., WERTHEIM, H. F., SUMPRADIT, N., VLIEGHE, E., HARA, G. L., GOULD, I. M., GOOSSENS, H., GREKO, C., SO, A. D., BIGDELI, M., TOMSON, G., WOODHOUSE, W., OMBAKA, E., PERALTA, A. Q., QAMAR, F. N., MIR, F., KARIUKI, S., BHUTTA, Z. A., COATES, A., BERGSTROM, R., WRIGHT, G. D., BROWN, E. D. & CARS, O. 2013. Antibiotic resistance-the need for global solutions. *Lancet Infect Dis,* 13**,** 1057-98.

LAXMINARAYAN, R., MATSOSO, P., PANT, S., BROWER, C., ROTTINGEN, J. A., KLUGMAN, K. & DAVIES, S. 2016. Access to effective antimicrobials: a worldwide challenge. *Lancet,* 387**,** 168-75.

LEE, G. M., KLEINMAN, K., PELTON, S. I., HANAGE, W., HUANG, S. S., LAKOMA, M., DUTTA-LINN, M., CROUCHER, N. J., STEVENSON, A. & FINKELSTEIN, J. A. 2014. Impact of 13-Valent Pneumococcal Conjugate Vaccination on Carriage in Young Children in Massachusetts. *J Pediatric Infect Dis Soc,* 3**,** 23-32.

LEFEVRE, J. C., FAUCON, G., SICARD, A. M. & GASC, A. M. 1993. DNA fingerprinting of *Streptococcus pneumoniae* strains by pulsed-field gel electrophoresis. *J Clin Microbiol,* 31**,** 2724-8.

LEIMKUGEL, J., ADAMS FORGOR, A., GAGNEUX, S., PFLUGER, V., FLIERL, C., AWINE, E., NAEGELI, M., DANGY, J. P., SMITH, T., HODGSON, A. & PLUSCHKE, G. 2005. An outbreak of serotype 1 *Streptococcus pneumoniae* meningitis in northern Ghana with features that are characteristic of Neisseria meningitidis meningitis epidemics. *J Infect Dis,* 192**,** 192-9.

LEROUX-ROELS, G., MAES, C., DE BOEVER, F., TRASKINE, M., RUGGEBERG, J. U. & BORYS, D. 2014. Safety, reactogenicity and immunogenicity of a novel pneumococcal protein-based vaccine in adults: a phase I/II randomized clinical study. *Vaccine,* 32**,** 6838-46.

LETUNIC, I. & BORK, P. 2016. Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res,* 44**,** W242-5.

LEUNG, M. H., BRYSON, K., FREYSTATTER, K., PICHON, B., EDWARDS, G., CHARALAMBOUS, B. M. & GILLESPIE, S. H. 2012. Sequetyping: serotyping *Streptococcus pneumoniae* by a single PCR sequencing strategy. *J Clin Microbiol,* 50**,** 2419-27.

LI, D., LIU, C. M., LUO, R., SADAKANE, K. & LAM, T. W. 2015. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics,* 31**,** 1674-6.

LIU, L., OZA, S., HOGAN, D., PERIN, J., RUDAN, I., LAWN, J. E., COUSENS, S., MATHERS, C. & BLACK, R. E. 2015. Global, regional, and national

causes of child mortality in 2000-13, with projections to inform post-2015 priorities: an updated systematic analysis. *Lancet,* 385**,** 430-40.

LIYANAPATHIRANA, V., ANG, I., TSANG, D., FUNG, K., NG, T. K., ZHOU, H. & IP, M. 2014. Application of a target enrichment-based next-generation sequencing protocol for identification and sequence-based prediction of pneumococcal serotypes. *BMC Microbiol,* 14**,** 60.

LLIYASU, G., HABIB, A. G. & MOHAMMAD, A. B. 2015. Antimicrobial Susceptibility Pattern of Invasive Pneumococcal Isolates in North West Nigeria. *J Glob Infect Dis,* 7**,** 70-4.

MACKENZIE, G. A., HILL, P. C., JEFFRIES, D. J., HOSSAIN, I., UCHENDU, U., AMEH, D., NDIAYE, M., ADEYEMI, O., PATHIRANA, J., OLATUNJI, Y., ABATAN, B., MUHAMMAD, B. S., FOMBAH, A. E., SAHA, D., PLUMB, I., AKANO, A., EBRUKE, B., IDEH, R. C., KUTI, B., GITHUA, P., OLUTUNDE, E., OFORDILE, O., GREEN, E., USUF, E., BADJI, H., IKUMAPAYI, U. N., MANJANG, A., SALAUDEEN, R., NSEKPONG, E. D., JARJU, S., ANTONIO, M., SAMBOU, S., CEESAY, L., LOWE-JALLOW, Y., JASSEH, M., MULHOLLAND, K., KNOLL, M., LEVINE, O. S., HOWIE, S. R., ADEGBOLA, R. A., GREENWOOD, B. M. & CORRAH, T. 2016. Effect of the introduction of pneumococcal conjugate vaccination on invasive pneumococcal disease in The Gambia: a population-based surveillance study. *Lancet Infect Dis,* 16**,** 703-11.

MAIDEN, M. C. 2006. Multilocus sequence typing of bacteria. *Annu Rev Microbiol,* 60**,** 561-88.

MAIDEN, M. C., BYGRAVES, J. A., FEIL, E., MORELLI, G., RUSSELL, J. E., URWIN, R., ZHANG, Q., ZHOU, J., ZURTH, K., CAUGANT, D. A., FEAVERS, I. M., ACHTMAN, M. & SPRATT, B. G. 1998. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci U S A,* 95**,** 3140-5.

MANCO, S., HERNON, F., YESILKAYA, H., PATON, J. C., ANDREW, P. W. & KADIOGLU, A. 2006. Pneumococcal neuraminidases A and B both have essential roles during infection of the respiratory tract and sepsis. *Infect Immun,* 74**,** 4014-20.

MANN, B., VAN OPIJNEN, T., WANG, J., OBERT, C., WANG, Y.-D., CARTER, R., MCGOLDRICK, D. J., RIDOUT, G., CAMILLI, A., TUOMANEN, E. I. & ROSCH, J. W. 2012. Control of Virulence by Small RNAs in *Streptococcus pneumoniae*. *PLoS Pathogens,* 8**,** e1002788.

MANSO, A. S., CHAI, M. H., ATACK, J. M., FURI, L., DE STE CROIX, M., HAIGH, R., TRAPPETTI, C., OGUNNIYI, A. D., SHEWELL, L. K., BOITANO, M., CLARK, T. A., KORLACH, J., BLADES, M., MIRKES, E., GORBAN, A. N., PATON, J. C., JENNINGS, M. P. & OGGIONI, M. R. 2014. A random six-phase switch regulates pneumococcal virulence via global epigenetic changes. *Nat Commun,* 5.

MARCHISIO, P., ESPOSITO, S., SCHITO, G. C., MARCHESE, A., CAVAGNA, R., PRINCIPI, N. & HERCULES PROJECT COLLABORATIVE, G. 2002. Nasopharyngeal carriage of *Streptococcus pneumoniae* in healthy children: implications for the use of heptavalent pneumococcal conjugate vaccine. *Emerg Infect Dis,* 8**,** 479-84.

MARSH, R., SMITH-VAUGHAN, H., HARE, K. M., BINKS, M., KONG, F., WARNING, J., GILBERT, G. L., MORRIS, P. & LEACH, A. J. 2010. The

nonserotypeable pneumococcus: phenotypic dynamics in the era of anticapsular vaccines. *J Clin Microbiol,* 48**,** 831-5.

MARTIN, D. R. & BRETT, M. S. 1996. Pneumococci causing invasive disease in New Zealand, 1987-94: serogroup and serotype coverage and antibiotic resistances. *N Z Med J,* 109**,** 288-90.

MASSIRE, C., GERTZ, R. E., SVOBODA, P., LEVERT, K., REED, M. S., POHL, J., KREFT, R., LI, F., WHITE, N., RANKEN, R., BLYN, L. B., ECKER, D. J., SAMPATH, R. & BEALL, B. 2012. Concurrent Serotyping and Genotyping of Pneumococci by Use of PCR and Electrospray Ionization Mass Spectrometry. *Journal of Clinical Microbiology,* 50**,** 2018-2025.

MAVROIDI, A., AANENSEN, D. M., GODOY, D., SKOVSTED, I. C., KALTOFT, M. S., REEVES, P. R., BENTLEY, S. D. & SPRATT, B. G. 2007. Genetic relatedness of the *Streptococcus pneumoniae* capsular biosynthetic loci. *J Bacteriol,* 189**,** 7841-55.

MCCHLERY, S. M., SCOTT, K. J. & CLARKE, S. C. 2005. Clonal analysis of invasive pneumococcal isolates in Scotland and coverage of serotypes by the licensed conjugate polysaccharide pneumococcal vaccine: possible implications for UK vaccine policy. *Eur J Clin Microbiol Infect Dis,* 24**,** 262-7.

MCCOOL, T. L., CATE, T. R., MOY, G. & WEISER, J. N. 2002. The immune response to pneumococcal proteins during experimental human carriage. *J Exp Med,* 195**,** 359-65.

MCGEE, L., MCDOUGAL, L., ZHOU, J., SPRATT, B. G., TENOVER, F. C., GEORGE, R., HAKENBECK, R., HRYNIEWICZ, W., LEFEVRE, J. C., TOMASZ, A. & KLUGMAN, K. P. 2001. Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae* defined by the pneumococcal molecular epidemiology network. *J Clin Microbiol,* 39**,** 2565-71.

MELIN, M., TRZCINSKI, K., ANTONIO, M., MERI, S., ADEGBOLA, R., KAIJALAINEN, T., KAYHTY, H. & VAKEVAINEN, M. 2010. Serotype-related variation in susceptibility to complement deposition and opsonophagocytosis among clinical isolates of *Streptococcus pneumoniae*. *Infect Immun,* 78**,** 5252-61.

MERCAT, A., NGUYEN, J. & DAUTZENBERG, B. 1991. An outbreak of pneumococcal pneumonia in two men's shelters. *Chest,* 99**,** 147-51.

METZKER, M. L. 2010. Sequencing technologies - the next generation. *Nat Rev Genet,* 11**,** 31-46.

MITCHELL, T. J., MENDEZ, F., PATON, J. C., ANDREW, P. W. & BOULNOIS, G. J. 1990. Comparison of pneumolysin genes and proteins from *Streptococcus pneumoniae* types 1 and 2. *Nucleic Acids Res,* 18**,** 4010.

MIYAJI, E. N., FERREIRA, D. M., LOPES, A. P., BRANDILEONE, M. C., DIAS, W. O. & LEITE, L. C. 2002. Analysis of serum cross-reactivity and cross-protection elicited by immunization with DNA vaccines against *Streptococcus pneumoniae* expressing PspA fragments from different clades. *Infect Immun,* 70**,** 5086-90.

MORAIS, L., CARVALHO MDA, G., ROCA, A., FLANNERY, B., MANDOMANDO, I., SORIANO-GABARRO, M., SIGAUQUE, B., ALONSO, P. & BEALL, B. 2007. Sequential multiplex PCR for identifying pneumococcal capsular serotypes from South-Saharan African clinical isolates. *J Med Microbiol,* 56**,** 1181-4.

MOSTOWY, R. J., CROUCHER, N. J., DE MAIO, N., CHEWAPREECHA, C., SALTER, S. J., TURNER, P., AANENSEN, D. M., BENTLEY, S. D., DIDELOT, X. & FRASER, C. 2017. Pneumococcal capsule synthesis locus cps as evolutionary hotspot with potential to generate novel serotypes by recombination. *Mol Biol Evol*.

MUNOZ, R., COFFEY, T. J., DANIELS, M., DOWSON, C. G., LAIBLE, G., CASAL, J., HAKENBECK, R., JACOBS, M., MUSSER, J. M., SPRATT, B. G. & ET AL. 1991. Intercontinental spread of a multiresistant clone of serotype 23F *Streptococcus pneumoniae*. *J Infect Dis,* 164**,** 302-6.

MUTSCHLER, H., GEBHARDT, M., SHOEMAN, R. L. & MEINHART, A. 2011. A novel mechanism of programmed cell death in bacteria by toxin-antitoxin systems corrupts peptidoglycan synthesis. *PLoS Biol,* 9**,** e1001033.

NELSON, A. L., ROCHE, A. M., GOULD, J. M., CHIM, K., RATNER, A. J. & WEISER, J. N. 2007. Capsule enhances pneumococcal colonization by limiting mucus-mediated clearance. *Infect Immun,* 75**,** 83-90.

NEWTON, R., HINDS, J. & WERNISCH, L. 2011. Empirical Bayesian models for analysing molecular serotyping microarrays. *BMC Bioinformatics,* 12**,** 88.

NJANPOP LAFOURCADE, B. M., SANOU, O., VAN DER LINDEN, M., LEVINA, N., KARANFIL, M., YARO, S., TAMEKLOE, T. A. & MUELLER, J. E. 2010. Serotyping pneumococcal meningitis cases in the African meningitis belt by use of multiplex PCR with cerebrospinal fluid. *J Clin Microbiol,* 48**,** 612-4.

NUMMINEN, E., CHEWAPREECHA, C., TURNER, C., GOLDBLATT, D., NOSTEN, F., BENTLEY, S. D., TURNER, P. & CORANDER, J. 2015. Climate induces seasonality in pneumococcal transmission. *Sci Rep,* 5**,** 11344.

NUNES, M. C., VON GOTTBERG, A., DE GOUVEIA, L., COHEN, C., MOORE, D. P., KLUGMAN, K. P. & MADHI, S. A. 2011. The impact of antiretroviral treatment on the burden of invasive pneumococcal disease in South African children: a time series analysis. *AIDS,* 25**,** 453-62.

NUNES, S., SA-LEAO, R., PEREIRA, L. C. & LENCASTRE, H. 2008. Emergence of a serotype 1 *Streptococcus pneumoniae* lineage colonising healthy children in Portugal in the seven-valent conjugate vaccination era. *Clin Microbiol Infect,* 14**,** 82-4.

NUORTI, J. P., BUTLER, J. C., FARLEY, M. M., HARRISON, L. H., MCGEER, A., KOLCZAK, M. S. & BREIMAN, R. F. 2000. Cigarette smoking and invasive pneumococcal disease. Active Bacterial Core Surveillance Team. *N Engl J Med,* 342**,** 681-9.

NZENZE, S. A., SHIRI, T., NUNES, M. C., KLUGMAN, K. P., KAHN, K., TWINE, R., DE GOUVEIA, L., VON GOTTBERG, A. & MADHI, S. A. 2013. Temporal changes in pneumococcal colonization in a rural African community with high HIV prevalence following routine infant pneumococcal immunization. *Pediatr Infect Dis J,* 32**,** 1270-8.

O'BRIEN, K. L., WOLFSON, L. J., WATT, J. P., HENKLE, E., DELORIA-KNOLL, M., MCCALL, N., LEE, E., MULHOLLAND, K., LEVINE, O. S., CHERIAN, T., HIB & PNEUMOCOCCAL GLOBAL BURDEN OF DISEASE STUDY, T. 2009. Burden of disease caused by *Streptococcus pneumoniae* in children younger than 5 years: global estimates. *Lancet,* 374**,** 893-902.

O'DEMPSEY, T. J., MCARDLE, T. F., LLOYD-EVANS, N., BALDEH, I., LAWRENCE, B. E., SECKA, O. & GREENWOOD, B. 1996. Pneumococcal disease among children in a rural area of west Africa. *Pediatr Infect Dis J,* 15**,** 431-7.

O'SULLIVAN, M. V. N., ZHOU, F., SINTCHENKO, V., KONG, F. & GILBERT, G. L. 2011. Multiplex PCR and Reverse Line Blot Hybridization Assay (mPCR/RLB). *Journal of Visualized Experiments : JoVE***,** 2781.

OBARO, S. K., ADEGBOLA, R. A., BANYA, W. A. & GREENWOOD, B. M. 1996. Carriage of pneumococci after pneumococcal vaccination. *Lancet,* 348**,** 271-2.

ODUTOLA, A., ANTONIO, M., OWOLABI, O., BOJANG, A., FOSTER-NYARKO, E., DONKOR, S., ADETIFA, I., TAYLOR, S., BOTTOMLEY, C., GREENWOOD, B. & OTA, M. 2013. Comparison of the prevalence of common bacterial pathogens in the oropharynx and nasopharynx of gambian infants. *PLoS One,* 8**,** e75558.

ODUTOLA, A., OTA, M. O., OGUNDARE, E. O., ANTONIO, M., OWIAFE, P., WORWUI, A., GREENWOOD, B., ALDERSON, M., TRASKINE, M., VERLANT, V., DOBBELAERE, K. & BORYS, D. 2015. Reactogenicity, safety and immunogenicity of a protein-based pneumococcal vaccine in Gambian children aged 2-4 years: a phase II randomized study. *Hum Vaccin Immunother***,** 0.

OGUNNIYI, A. D., GRABOWICZ, M., MAHDI, L. K., COOK, J., GORDON, D. L., SADLON, T. A. & PATON, J. C. 2009. Pneumococcal histidine triad proteins are regulated by the Zn2+-dependent repressor AdcR and inhibit complement deposition through the recruitment of complement factor H. *FASEB J,* 23**,** 731-8.

OGUNNIYI, A. D., LEMESSURIER, K. S., GRAHAM, R. M., WATT, J. M., BRILES, D. E., STROEHER, U. H. & PATON, J. C. 2007. Contributions of pneumolysin, pneumococcal surface protein A (PspA), and PspC to pathogenicity of *Streptococcus pneumoniae* D39 in a mouse model. *Infect Immun,* 75**,** 1843-51.

OLARTE, L., KAPLAN, S. L., BARSON, W. J., ROMERO, J. R., LIN, P. L., TAN, T. Q., HOFFMAN, J. A., BRADLEY, J. S., GIVNER, L. B., MASON, E. O. & HULTEN, K. G. 2017. Emergence of Multidrug-Resistant Pneumococcal Serotype 35B among Children in the United States. *J Clin Microbiol,* 55**,** 724-734.

OLIVER, M. B., VAN DER LINDEN, M. P., KUNTZEL, S. A., SAAD, J. S. & NAHM, M. H. 2013. Discovery of *Streptococcus pneumoniae* serotype 6 variants with glycosyltransferases synthesizing two differing repeating units. *J Biol Chem,* 288**,** 25976-85.

ORIHUELA, C. J., RADIN, J. N., SUBLETT, J. E., GAO, G., KAUSHAL, D. & TUOMANEN, E. I. 2004. Microarray Analysis of Pneumococcal Gene Expression during Invasive Disease. *Infection and Immunity,* 72**,** 5582-5596.

ORTQVIST, A., HEDLUND, J. & KALIN, M. 2005. *Streptococcus pneumoniae*: epidemiology, risk factors, and clinical features. *Semin Respir Crit Care Med,* 26**,** 563-74.

OUSMANE, S., DIALLO, B. A., OUEDRAOGO, R., SANDA, A. A., SOUSSOU, A. M. & COLLARD, J. M. 2017. Serotype Distribution and Antimicrobial Sensitivity Profile of *Streptococcus pneumoniae* Carried in Healthy Toddlers before PCV13 Introduction in Niamey, Niger. *PLoS One,* 12**,** e0169547.

OVERWEG, K., KERR, A., SLUIJTER, M., JACKSON, M. H., MITCHELL, T. J., DE JONG, A. P., DE GROOT, R. & HERMANS, P. W. 2000. The putative proteinase maturation protein A of *Streptococcus pneumoniae* is a conserved surface protein with potential to elicit protective immune responses. *Infect Immun,* 68**,** 4180-8.

PAI, R., GERTZ, R. E. & BEALL, B. 2006. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. *J Clin Microbiol,* 44**,** 124-31.

PARK, I. H., GENO, K. A., SHERWOOD, L. K., NAHM, M. H. & BEALL, B. 2014. Population-based analysis of invasive nontypeable pneumococci reveals that most have defective capsule synthesis genes. *PLoS One,* 9**,** e97825.

PARK, I. H., GENO, K. A., YU, J., OLIVER, M. B., KIM, K. H. & NAHM, M. H. 2015. Genetic, biochemical, and serological characterization of a new pneumococcal serotype, 6H, and generation of a pneumococcal strain producing three different capsular repeat units. *Clin Vaccine Immunol,* 22**,** 313-8.

PARK, S. Y., MOORE, M. R., BRUDEN, D. L., HYDE, T. B., REASONOVER, A. L., HARKER-JONES, M., RUDOLPH, K. M., HURLBURT, D. A., PARKS, D. J., PARKINSON, A. J., SCHUCHAT, A. & HENNESSY, T. W. 2008. Impact of conjugate vaccine on transmission of antimicrobial-resistant *Streptococcus pneumoniae* among Alaskan children. *Pediatr Infect Dis J,* 27**,** 335-40.

PARRY, C. M., DUONG, N. M., ZHOU, J., MAI, N. T., DIEP, T. S., THINH LE, Q., WAIN, J., VAN VINH CHAU, N., GRIFFITHS, D., DAY, N. P., WHITE, N. J., HIEN, T. T., SPRATT, B. G. & FARRAR, J. J. 2002. Emergence in Vietnam of *Streptococcus pneumoniae* resistant to multiple antimicrobial agents as a result of dissemination of the multiresistant Spain(23F)-1 clone. *Antimicrob Agents Chemother,* 46**,** 3512-7.

PATON, J. C. & GIAMMARINARO, P. 2001. Genome-based analysis of pneumococcal virulence factors: the quest for novel vaccine antigens and drug targets. *Trends Microbiol,* 9**,** 515-8.

PELTON, S. I., HUOT, H., FINKELSTEIN, J. A., BISHOP, C. J., HSU, K. K., KELLENBERG, J., HUANG, S. S., GOLDSTEIN, R. & HANAGE, W. P. 2007. Emergence of 19A as virulent and multidrug resistant Pneumococcus in Massachusetts following universal immunization of infants with pneumococcal conjugate vaccine. *Pediatr Infect Dis J,* 26**,** 468-72.

PIEL, F. B., HAY, S. I., GUPTA, S., WEATHERALL, D. J. & WILLIAMS, T. N. 2013. Global burden of sickle cell anaemia in children under five, 2010-2050: modelling based on demographics, excess mortality, and interventions. *PLoS Med,* 10**,** e1001484.

PILISHVILI, T., LEXAU, C., FARLEY, M. M., HADLER, J., HARRISON, L. H., BENNETT, N. M., REINGOLD, A., THOMAS, A., SCHAFFNER, W., CRAIG, A. S., SMITH, P. J., BEALL, B. W., WHITNEY, C. G., MOORE, M. R. & ACTIVE BACTERIAL CORE SURVEILLANCE/EMERGING INFECTIONS PROGRAM, N. 2010. Sustained reductions in invasive pneumococcal disease in the era of conjugate vaccine. *J Infect Dis,* 201**,** 32-41.

PIMENTA, F. C., ROUNDTREE, A., SOYSAL, A., BAKIR, M., DU PLESSIS, M., WOLTER, N., VON GOTTBERG, A., MCGEE, L., CARVALHO MDA, G.

& BEALL, B. 2013. Sequential triplex real-time PCR assay for detecting 21 pneumococcal capsular serotypes that account for a high global disease burden. *J Clin Microbiol,* 51**,** 647-52.

PIROFSKI, L. A. & CASADEVALL, A. 2012. Q and A: What is a pathogen? A question that begs the point. *BMC Biol,* 10**,** 6.

PLETZ, M. W., MCGEE, L., JORGENSEN, J., BEALL, B., FACKLAM, R. R., WHITNEY, C. G. & KLUGMAN, K. P. 2004. Levofloxacin-resistant invasive *Streptococcus pneumoniae* in the United States: evidence for clonal spread and the impact of conjugate pneumococcal vaccine. *Antimicrob Agents Chemother,* 48**,** 3491-7.

PORAT, N., TREFLER, R. & DAGAN, R. 2001. Persistence of two invasive *Streptococcus pneumoniae* clones of serotypes 1 and 5 in comparison to that of multiple clones of serotypes 6B and 23F among children in southern Israel. *J Clin Microbiol,* 39**,** 1827-32.

PRACHT, D., ELM, C., GERBER, J., BERGMANN, S., ROHDE, M., SEILER, M., KIM, K. S., JENKINSON, H. F., NAU, R. & HAMMERSCHMIDT, S. 2005. PavA of *Streptococcus pneumoniae* modulates adherence, invasion, and meningeal inflammation. *Infect Immun,* 73**,** 2680-9.

PRESTON, J. A., BEAGLEY, K. W., GIBSON, P. G. & HANSBRO, P. M. 2004. Genetic background affects susceptibility in nonfatal pneumococcal bronchopneumonia. *Eur Respir J,* 23**,** 224-31.

PRYMULA, R., PAZDIORA, P., TRASKINE, M., RUGGEBERG, J. U. & BORYS, D. 2014. Safety and immunogenicity of an investigational vaccine containing two common pneumococcal proteins in toddlers: a phase II randomized clinical trial. *Vaccine,* 32**,** 3025-34.

REGEV-YOCHAY, G., ABULLAISH, I., MALLEY, R., SHAINBERG, B., VARON, M., ROYTMAN, Y., ZIV, A., GORAL, A., ELHAMDANY, A., RAHAV, G., RAZ, M. & PALESTINIAN-ISRAELI COLLABORATIVE RESEARCH STUDY, G. 2012. *Streptococcus pneumoniae* carriage in the Gaza strip. *PLoS One,* 7**,** e35061.

REISMAN, J., RUDOLPH, K., BRUDEN, D., HURLBURT, D., BRUCE, M. G. & HENNESSY, T. 2014. Risk Factors for Pneumococcal Colonization of the Nasopharynx in Alaska Native Adults and Children. *J Pediatric Infect Dis Soc,* 3**,** 104-11.

REJASSE, A., GILOIS, N., BARBOSA, I., HUILLET, E., BEVILACQUA, C., TRAN, S., RAMARAO, N., STENFORS ARNESEN, L. P. & SANCHIS, V. 2012. Temperature-dependent production of various PlcR-controlled virulence factors in Bacillus weihenstephanensis strain KBAB4. *Appl Environ Microbiol,* 78**,** 2553-61.

REN, B., SZALAI, A. J., THOMAS, O., HOLLINGSHEAD, S. K. & BRILES, D. E. 2003. Both family 1 and family 2 PspA proteins can inhibit complement deposition and confer virulence to a capsular serotype 3 strain of *Streptococcus pneumoniae*. *Infect Immun,* 71**,** 75-85.

RICE, L. B. 1998. Tn916 family conjugative transposons and dissemination of antimicrobial resistance determinants. *Antimicrob Agents Chemother,* 42**,** 1871-7.

RICHTER, S. S., DIEKEMA, D. J., HEILMANN, K. P., DOHRN, C. L., RIAHI, F. & DOERN, G. V. 2014. Changes in pneumococcal serotypes and antimicrobial resistance after introduction of the 13-valent conjugate vaccine in the United States. *Antimicrob Agents Chemother,* 58**,** 6484-9.

RICHTER, S. S., HEILMANN, K. P., DOHRN, C. L., RIAHI, F., DIEKEMA, D. J. & DOERN, G. V. 2013. Pneumococcal serotypes before and after introduction of conjugate vaccines, United States, 1999-2011(1.). *Emerg Infect Dis,* 19**,** 1074-83.

ROBINSON, K. A., BAUGHMAN, W., ROTHROCK, G., BARRETT, N. L., PASS, M., LEXAU, C., DAMASKE, B., STEFONEK, K., BARNES, B., PATTERSON, J., ZELL, E. R., SCHUCHAT, A., WHITNEY, C. G. & ACTIVE BACTERIAL CORE SURVEILLANCE /EMERGING INFECTIONS PROGRAM, N. 2001. Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995-1998: Opportunities for prevention in the conjugate vaccine era. *JAMA,* 285**,** 1729-35.

ROCA, A., BOJANG, A., BOTTOMLEY, C., GLADSTONE, R. A., ADETIFA, J. U., EGERE, U., BURR, S., ANTONIO, M., BENTLEY, S., KAMPMANN, B. & PNEUMO13 STUDY, G. 2015. Effect on nasopharyngeal pneumococcal carriage of replacing PCV7 with PCV13 in the Expanded Programme of Immunization in The Gambia. *Vaccine,* 33**,** 7144-51.

ROCA, A., DIONE, M. M., BOJANG, A., TOWNEND, J., EGERE, U., DARBOE, O., HOWIE, S. R., HILL, P. C., ADEGBOLA, R. A., GREENWOOD, B. M. & ANTONIO, M. 2013. Nasopharyngeal carriage of pneumococci four years after community-wide vaccination with PCV-7 in The Gambia: long-term evaluation of a cluster randomized trial. *PLoS One,* 8**,** e72198.

ROCA, A., HILL, P. C., TOWNEND, J., EGERE, U., ANTONIO, M., BOJANG, A., AKISANYA, A., LITCHFIELD, T., NSEKPONG, D. E., OLUWALANA, C., HOWIE, S. R., GREENWOOD, B. & ADEGBOLA, R. A. 2011. Effects of Community-Wide Vaccination with PCV-7 on Pneumococcal Nasopharyngeal Carriage in The Gambia: A Cluster-Randomized Trial. *PLoS Med,* 8**,** e1001107.

ROSENOW, C., RYAN, P., WEISER, J. N., JOHNSON, S., FONTAN, P., ORTQVIST, A. & MASURE, H. R. 1997. Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of *Streptococcus pneumoniae*. *Molecular Microbiology,* 25**,** 819-829.

RUDAN, I., BOSCHI-PINTO, C., BILOGLAV, Z., MULHOLLAND, K. & CAMPBELL, H. 2008. Epidemiology and etiology of childhood pneumonia. *Bull World Health Organ,* 86**,** 408-16.

RUDAN, I., O'BRIEN, K. L., NAIR, H., LIU, L., THEODORATOU, E., QAZI, S., LUKSIC, I., FISCHER WALKER, C. L., BLACK, R. E., CAMPBELL, H. & CHILD HEALTH EPIDEMIOLOGY REFERENCE, G. 2013. Epidemiology and etiology of childhood pneumonia in 2010: estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. *J Glob Health,* 3**,** 010401.

SALTER, S. J., HINDS, J., GOULD, K. A., LAMBERTSEN, L., HANAGE, W. P., ANTONIO, M., TURNER, P., HERMANS, P. W., BOOTSMA, H. J., O'BRIEN, K. L. & BENTLEY, S. D. 2012. Variation at the capsule locus, cps, of mistyped and non-typable *Streptococcus pneumoniae* isolates. *Microbiology,* 158**,** 1560-9.

SAMPSON, J. S., O'CONNOR, S. P., STINSON, A. R., THARPE, J. A. & RUSSELL, H. 1994. Cloning and nucleotide sequence analysis of psaA, the *Streptococcus pneumoniae* gene encoding a 37-kilodalton protein

homologous to previously reported Streptococcus sp. adhesins. *Infect Immun,* 62**,** 319-24.

SANCHEZ-BEATO, A. R., LOPEZ, R. & GARCIA, J. L. 1998. Molecular characterization of PcpA: a novel choline-binding protein of *Streptococcus pneumoniae*. *FEMS Microbiol Lett,* 164**,** 207-14.

SANDGREN, A., SJOSTROM, K., OLSSON-LILJEQUIST, B., CHRISTENSSON, B., SAMUELSSON, A., KRONVALL, G. & HENRIQUES NORMARK, B. 2004. Effect of clonal and serotype-specific properties on the invasive capacity of *Streptococcus pneumoniae*. *J Infect Dis,* 189**,** 785-96.

SANGER, F., NICKLEN, S. & COULSON, A. R. 1977. DNA sequencing with chain-terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America,* 74**,** 5463-5467.

SANGUINETTI, L., TOTI, S., REGUZZI, V., BAGNOLI, F. & DONATI, C. 2012. A novel computational method identifies intra- and inter-species recombination events in Staphylococcus aureus and *Streptococcus pneumoniae*. *PLoS Comput Biol,* 8**,** e1002668.

SATZKE, C., DUNNE, E. M., PORTER, B. D., KLUGMAN, K. P., MULHOLLAND, E. K. & PNEUCARRIAGE PROJECT, G. 2015. The PneuCarriage Project: A Multi-Centre Comparative Study to Identify the Best Serotyping Methods for Examining Pneumococcal Carriage in Vaccine Evaluation Studies. *PLoS Med,* 12**,** e1001903; discussion e1001903.

SATZKE, C., TURNER, P., VIROLAINEN-JULKUNEN, A., ADRIAN, P. V., ANTONIO, M., HARE, K. M., HENAO-RESTREPO, A. M., LEACH, A. J., KLUGMAN, K. P., PORTER, B. D., SA-LEAO, R., SCOTT, J. A., NOHYNEK, H., O'BRIEN, K. L. & GROUP, W. H. O. P. C. W. 2013. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae*: updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine,* 32**,** 165-79.

SAVITHA, M. R., NANDEESHWARA, S. B., PRADEEP KUMAR, M. J., UL-HAQUE, F. & RAJU, C. K. 2007. Modifiable risk factors for acute lower respiratory tract infections. *Indian J Pediatr,* 74**,** 477-82.

SEED, K. D. & DENNIS, J. J. 2008. Development of Galleria mellonella as an Alternative Infection Model for the Burkholderia cepacia Complex. *Infection and Immunity,* 76**,** 1267-1275.

SEEMANN, T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics,* 30**,** 2068-9.

SEEMANN T, G. D. S. A., BULACH DM, SCHULTZ MB, KWONG JC, HOWDEN BP. *Nullarbor* [Online]. Available: https://github.com/tseemann/nullarbor [Accessed 8 Feb 2017].

SELANDER, R. K., CAUGANT, D. A., OCHMAN, H., MUSSER, J. M., GILMOUR, M. N. & WHITTAM, T. S. 1986. Methods of multilocus enzyme electrophoresis for bacterial population genetics and systematics. *Appl Environ Microbiol,* 51**,** 873-84.

SELVA, L., DEL AMO, E., BROTONS, P. & MUNOZ-ALMAGRO, C. 2012. Rapid and easy identification of capsular serotypes of *Streptococcus pneumoniae* by use of fragment analysis by automated fluorescence-based capillary electrophoresis. *J Clin Microbiol,* 50**,** 3451-7.

SHEPPARD, C. L., HARRISON, T. G., SMITH, M. D. & GEORGE, R. C. 2011. Development of a sensitive, multiplexed immunoassay using xMAP beads

for detection of serotype-specific *Streptococcus pneumoniae* antigen in urine samples. *J Med Microbiol,* 60**,** 49-55.

SIIRA, L., RANTALA, M., JALAVA, J., HAKANEN, A. J., HUOVINEN, P., KAIJALAINEN, T., LYYTIKAINEN, O. & VIROLAINEN, A. 2009. Temporal trends of antimicrobial resistance and clonality of invasive *Streptococcus pneumoniae* isolates in Finland, 2002 to 2006. *Antimicrob Agents Chemother,* 53**,** 2066-73.

SIROTNAK, A. P., EPPES, S. C. & KLEIN, J. D. 1996. Tuboovarian abscess and peritonitis caused by *Streptococcus pneumoniae* serotype 1 in young girls. *Clin Infect Dis,* 22**,** 993-6.

SJOSTROM, K., SPINDLER, C., ORTQVIST, A., KALIN, M., SANDGREN, A., KUHLMANN-BERENZON, S. & HENRIQUES-NORMARK, B. 2006. Clonal and capsular types decide whether pneumococci will act as a primary or opportunistic pathogen. *Clin Infect Dis,* 42**,** 451-9.

SLAGER, J., KJOS, M., ATTAIECH, L. & VEENING, J. W. 2014. Antibiotic-induced replication stress triggers bacterial competence by increasing gene dosage near the origin. *Cell,* 157**,** 395-406.

SLEEMAN, K. L., GRIFFITHS, D., SHACKLEY, F., DIGGLE, L., GUPTA, S., MAIDEN, M. C., MOXON, E. R., CROOK, D. W. & PETO, T. E. 2006. Capsular serotype-specific attack rates and duration of carriage of *Streptococcus pneumoniae* in a population of children. *J Infect Dis,* 194**,** 682-8.

SMILLIE, W. G., WARNOCK, G. H. & WHITE, H. J. 1938. A Study of a Type I Pneumococcus Epidemic at the State Hospital at Worcester, Mass. *Am J Public Health Nations Health,* 28**,** 293-302.

SMITH-VAUGHAN, H., MARSH, R., MACKENZIE, G., FISHER, J., MORRIS, P. S., HARE, K., MCCALLUM, G., BINKS, M., MURPHY, D., LUM, G., COOK, H., KRAUSE, V., JACUPS, S. & LEACH, A. J. 2009. Age-specific cluster of cases of serotype 1 *Streptococcus pneumoniae* carriage in remote indigenous communities in Australia. *Clin Vaccine Immunol,* 16**,** 218-21.

SPRATT, B. G. & GREENWOOD, B. M. 2000. Prevention of pneumococcal disease by vaccination: does serotype replacement matter? *Lancet,* 356**,** 1210-1.

SPRATT, B. G., HANAGE, W. P. & FEIL, E. J. 2001. The relative contributions of recombination and point mutation to the diversification of bacterial clones. *Curr Opin Microbiol,* 4**,** 602-6.

STAMATAKIS, A. 2006. RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics,* 22**,** 2688-90.

STAPLES, M., GRAHAM, R. M., JENNISON, A. V., ARIOTTI, L., HICKS, V., COOK, H., KRAUSE, V., GIELE, C. & SMITH, H. V. 2015. Molecular characterization of an Australian serotype 1 *Streptococcus pneumoniae* outbreak. *Epidemiol Infect,* 143**,** 325-33.

SULLIVAN, M. C., COOPER, B. W., NIGHTINGALE, C. H., QUINTILIANI, R. & LAWLOR, M. T. 1993. Evaluation of the efficacy of ciprofloxacin against *Streptococcus pneumoniae* by using a mouse protection model. *Antimicrobial Agents and Chemotherapy,* 37**,** 234-239.

TAMIR, S. O., ROTH, Y., DALAL, I., GOLDFARB, A., GROTTO, I. & MAROM, T. 2015. Changing trends of acute otitis media bacteriology in central Israel in the pneumococcal conjugate vaccines era. *Pediatr Infect Dis J,* 34**,** 195-9.

TAN, M. W., MAHAJAN-MIKLOS, S. & AUSUBEL, F. M. 1999. Killing of Caenorhabditis elegans by Pseudomonas aeruginosa used to model mammalian bacterial pathogenesis. *Proc Natl Acad Sci U S A,* 96**,** 715-20.

TETTELIN, H., NELSON, K. E., PAULSEN, I. T., EISEN, J. A., READ, T. D., PETERSON, S., HEIDELBERG, J., DEBOY, R. T., HAFT, D. H., DODSON, R. J., DURKIN, A. S., GWINN, M., KOLONAY, J. F., NELSON, W. C., PETERSON, J. D., UMAYAM, L. A., WHITE, O., SALZBERG, S. L., LEWIS, M. R., RADUNE, D., HOLTZAPPLE, E., KHOURI, H., WOLF, A. M., UTTERBACK, T. R., HANSEN, C. L., MCDONALD, L. A., FELDBLYUM, T. V., ANGIUOLI, S., DICKINSON, T., HICKEY, E. K., HOLT, I. E., LOFTUS, B. J., YANG, F., SMITH, H. O., VENTER, J. C., DOUGHERTY, B. A., MORRISON, D. A., HOLLINGSHEAD, S. K. & FRASER, C. M. 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science,* 293**,** 498-506.

TIN TIN HTAR, M., CHRISTOPOULOU, D. & SCHMITT, H. J. 2015. Pneumococcal serotype evolution in Western Europe. *BMC Infect Dis,* 15**,** 419.

TOMASZ, A. & SAUKKONEN, K. 1989. The nature of cell wall-derived inflammatory components of pneumococci. *Pediatr Infect Dis J,* 8**,** 902-3.

TOMCZYK, S., LYNFIELD, R., SCHAFFNER, W., REINGOLD, A., MILLER, L., PETIT, S., HOLTZMAN, C., ZANSKY, S. M., THOMAS, A., BAUMBACH, J., HARRISON, L. H., FARLEY, M. M., BEALL, B., MCGEE, L., GIERKE, R., PONDO, T. & KIM, L. 2016. Prevention of Antibiotic-Nonsusceptible Invasive Pneumococcal Disease With the 13-Valent Pneumococcal Conjugate Vaccine. *Clin Infect Dis,* 62**,** 1119-25.

TRAORE, Y., TAMEKLO, T. A., NJANPOP-LAFOURCADE, B. M., LOURD, M., YARO, S., NIAMBA, D., DRABO, A., MUELLER, J. E., KOECK, J. L. & GESSNER, B. D. 2009. Incidence, seasonality, age distribution, and mortality of pneumococcal meningitis in Burkina Faso and Togo. *Clin Infect Dis,* 48 Suppl 2**,** S181-9.

TROTTER, C. L., WAIGHT, P., ANDREWS, N. J., SLACK, M., EFSTRATIOU, A., GEORGE, R. & MILLER, E. 2010. Epidemiology of invasive pneumococcal disease in the pre-conjugate vaccine era: England and Wales, 1996-2006. *J Infect,* 60**,** 200-8.

TSENG, H. J., MCEWAN, A. G., PATON, J. C. & JENNINGS, M. P. 2002. Virulence of *Streptococcus pneumoniae*: PsaA mutants are hypersensitive to oxidative stress. *Infect Immun,* 70**,** 1635-9.

TUOMANEN, E., LIU, H., HENGSTLER, B., ZAK, O. & TOMASZ, A. 1985. The induction of meningeal inflammation by components of the pneumococcal cell wall. *J Infect Dis,* 151**,** 859-68.

TURNER, P., HINDS, J., TURNER, C., JANKHOT, A., GOULD, K., BENTLEY, S. D., NOSTEN, F. & GOLDBLATT, D. 2011. Improved detection of nasopharyngeal cocolonization by multiple pneumococcal serotypes by use of latex agglutination or molecular serotyping by microarray. *J Clin Microbiol,* 49**,** 1784-9.

TZIANABOS, A. O., ONDERDONK, A. B., ROSNER, B., CISNEROS, R. L. & KASPER, D. L. 1993. Structural features of polysaccharides that induce intra-abdominal abscesses. *Science,* 262**,** 416-9.

UBUKATA, K., CHIBA, N., HANADA, S., MOROZUMI, M., WAJIMA, T., SHOUJI, M., IWATA, S. & INVASIVE PNEUMOCOCCAL DISEASES SURVEILLANCE STUDY, G. 2015. Serotype Changes and Drug Resistance in Invasive Pneumococcal Diseases in Adults after Vaccinations in Children, Japan, 2010-2013. *Emerg Infect Dis,* 21**,** 1956-65.

USEN, S., ADEGBOLA, R., MULHOLLAND, K., JAFFAR, S., HILTON, S., OPARAUGO, A., OMOSIGHO, C., LAHAI, G., CORRAH, T., PALMER, A., SCHNEIDER, G., WEBER, M. & GREENWOOD, B. 1998. Epidemiology of invasive pneumococcal disease in the Western Region, The Gambia. *Pediatr Infect Dis J,* 17**,** 23-8.

USUF, E., BADJI, H., BOJANG, A., JARJU, S., IKUMAPAYI, U. N., ANTONIO, M., MACKENZIE, G. & BOTTOMLEY, C. 2015. Pneumococcal carriage in rural Gambia prior to the introduction of pneumococcal conjugate vaccine: a population-based survey. *Trop Med Int Health,* 20**,** 871-9.

VADESILHO, C. F., FERREIRA, D. M., MORENO, A. T., CHAVEZ-OLORTEGUI, C., MACHADO DE AVILA, R. A., OLIVEIRA, M. L., HO, P. L. & MIYAJI, E. N. 2012. Characterization of the antibody response elicited by immunization with pneumococcal surface protein A (PspA) as recombinant protein or DNA vaccine and analysis of protection against an intranasal lethal challenge with *Streptococcus pneumoniae*. *Microb Pathog,* 53**,** 243-9.

VALENTINO, M. D., MCGUIRE, A. M., ROSCH, J. W., BISPO, P. J., BURNHAM, C., SANFILIPPO, C. M., CARTER, R. A., ZEGANS, M. E., BEALL, B., EARL, A. M., TUOMANEN, E. I., MORRIS, T. W., HAAS, W. & GILMORE, M. S. 2014. Unencapsulated *Streptococcus pneumoniae* from conjunctivitis encode variant traits and belong to a distinct phylogenetic cluster. *Nat Commun,* 5**,** 5411.

VALLÈS, X., FLANNERY, B., ROCA, A., MANDOMANDO, I., SIGAÚQUE, B., SANZ, S., SCHUCHAT, A., LEVINE, M., SORIANO-GABARRÓ, M. & ALONSO, P. 2006. Serotype distribution and antibiotic susceptibility of invasive and nasopharyngeal isolates of *Streptococcus pneumoniae* among children in rural Mozambique. *Tropical Medicine & International Health,* 11**,** 358-366.

VAN HOEK, A. J., SHEPPARD, C. L., ANDREWS, N. J., WAIGHT, P. A., SLACK, M. P., HARRISON, T. G., LADHANI, S. N. & MILLER, E. 2014. Pneumococcal carriage in children and adults two years after introduction of the thirteen valent pneumococcal conjugate vaccine in England. *Vaccine,* 32**,** 4349-55.

VAN OPIJNEN, T., BODI, K. L. & CAMILLI, A. 2009. Tn-seq; high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. *Nature methods,* 6**,** 767-772.

VAN OPIJNEN, T. & CAMILLI, A. 2012. A fine scale phenotype–genotype virulence map of a bacterial pathogen. *Genome Research,* 22**,** 2541-2551.

VAN TONDER, A. J., BRAY, J. E., ROALFE, L., WHITE, R., ZANCOLLI, M., QUIRK, S. J., HARALDSSON, G., JOLLEY, K. A., MAIDEN, M. C., BENTLEY, S. D., HARALDSSON, A., ERLENDSDOTTIR, H., KRISTINSSON, K. G., GOLDBLATT, D. & BRUEGGEMANN, A. B. 2015. Genomics reveals the worldwide distribution of multidrug-resistant serotype 6E pneumococci. *J Clin Microbiol*.

VARON, E., COHEN, R., BECHET, S., DOIT, C. & LEVY, C. 2015. Invasive disease potential of pneumococci before and after the 13-valent pneumococcal conjugate vaccine implementation in children. *Vaccine,* 33**,** 6178-85.

VELEZ, C. D., LEWIS, C. J., KASPER, D. L. & COBB, B. A. 2009. Type I *Streptococcus pneumoniae* carbohydrate utilizes a nitric oxide and MHC II-dependent pathway for antigen presentation. *Immunology,* 127**,** 73-82.

VON MOLLENDORF, C., TEMPIA, S., VON GOTTBERG, A., MEIRING, S., QUAN, V., FELDMAN, C., CLOETE, J., MADHI, S. A., O'BRIEN, K. L., KLUGMAN, K. P., WHITNEY, C. G. & COHEN, C. 2017. Estimated severe pneumococcal disease cases and deaths before and after pneumococcal conjugate vaccine introduction in children younger than 5 years of age in South Africa. *PLoS One,* 12**,** e0179905.

WAGENVOORT, G. H., SANDERS, E. A., VLAMINCKX, B. J., DE MELKER, H. E., VAN DER ENDE, A. & KNOL, M. J. 2017. Sex differences in invasive pneumococcal disease and the impact of pneumococcal conjugate vaccination in the Netherlands, 2004 to 2015. *Euro Surveill,* 22.

WAIGHT, P. A., ANDREWS, N. J., LADHANI, S. N., SHEPPARD, C. L., SLACK, M. P. & MILLER, E. 2015. Effect of the 13-valent pneumococcal conjugate vaccine on invasive pneumococcal disease in England and Wales 4 years after its introduction: an observational cohort study. *Lancet Infect Dis,* 15**,** 535-43.

WALKER, C. L., RUDAN, I., LIU, L., NAIR, H., THEODORATOU, E., BHUTTA, Z. A., O'BRIEN, K. L., CAMPBELL, H. & BLACK, R. E. 2013. Global burden of childhood pneumonia and diarrhoea. *Lancet,* 381**,** 1405-16.

WASFY, M. O., PIMENTEL, G., ABDEL-MAKSOUD, M., RUSSELL, K. L., BARROZO, C. P., KLENA, J. D., EARHART, K. & HAJJEH, R. 2005. Antimicrobial susceptibility and serotype distribution of *Streptococcus pneumoniae* causing meningitis in Egypt, 1998-2003. *J Antimicrob Chemother,* 55**,** 958-64.

WATSON, D. A. & MUSHER, D. M. 1990. Interruption of capsule production in Streptococcus pneumonia serotype 3 by insertion of transposon Tn916. *Infect Immun,* 58**,** 3135-8.

WEISER, J. N., AUSTRIAN, R., SREENIVASAN, P. K. & MASURE, H. R. 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. *Infect Immun,* 62**,** 2582-9.

WEISER, J. N., MARKIEWICZ, Z., TUOMANEN, E. I. & WANI, J. H. 1996. Relationship between phase variation in colony morphology, intrastrain variation in cell wall physiology, and nasopharyngeal colonization by *Streptococcus pneumoniae*. *Infect Immun,* 64**,** 2240-5.

WESTH, H., SKIBSTED, L. & KORNER, B. 1990. *Streptococcus pneumoniae* infections of the female genital tract and in the newborn child. *Rev Infect Dis,* 12**,** 416-22.

WHITE, B., BARNES, L. A. & ROBINSON, E. S. 1938. *The biology of pneumococcus; the bacteriological, biochemical, and immunological characters and activities of Diplococcus pneumoniae, [by] Benjamin White, Ph. D., with the collaboration of Elliott Stirling Robinson and Laverne Almon Barnes,* New York, The Commonwealth fund ;.

WHITNEY, C. G., FARLEY, M. M., HADLER, J., HARRISON, L. H., LEXAU, C., REINGOLD, A., LEFKOWITZ, L., CIESLAK, P. R., CETRON, M.,

ZELL, E. R., JORGENSEN, J. H., SCHUCHAT, A. & ACTIVE BACTERIAL CORE SURVEILLANCE PROGRAM OF THE EMERGING INFECTIONS PROGRAM, N. 2000. Increasing prevalence of multidrug-resistant *Streptococcus pneumoniae* in the United States. *N Engl J Med,* 343**,** 1917-24.

WHO 2013. *Pocket Book of Hospital Care for Children: Guidelines for the Management of Common Childhood Illnesses.* 2nd ed. Geneva.

WILLIAMS, T. M., LOMAN, N. J., EBRUKE, C., MUSHER, D. M., ADEGBOLA, R. A., PALLEN, M. J., WEINSTOCK, G. M. & ANTONIO, M. 2012. Genome analysis of a highly virulent serotype 1 strain of *Streptococcus pneumoniae* from West Africa. *PLoS One,* 7**,** e26742.

WILLIAMS, T. N., UYOGA, S., MACHARIA, A., NDILA, C., MCAULEY, C. F., OPI, D. H., MWARUMBA, S., MAKANI, J., KOMBA, A., NDIRITU, M. N., SHARIF, S. K., MARSH, K., BERKLEY, J. A. & SCOTT, J. A. 2009. Bacteraemia in Kenyan children with sickle-cell anaemia: a retrospective cohort and case-control study. *Lancet,* 374**,** 1364-70.

WINKELSTEIN, J. A. & TOMASZ, A. 1978. Activation of the alternative complement pathway by pneumococcal cell wall teichoic acid. *J Immunol,* 120**,** 174-8.

WIZEMANN, T. M., HEINRICHS, J. H., ADAMOU, J. E., ERWIN, A. L., KUNSCH, C., CHOI, G. H., BARASH, S. C., ROSEN, C. A., MASURE, H. R., TUOMANEN, E., GAYLE, A., BREWAH, Y. A., WALSH, W., BARREN, P., LATHIGRA, R., HANSON, M., LANGERMANN, S., JOHNSON, S. & KOENIG, S. 2001. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. *Infect Immun,* 69**,** 1593-8.

WOOD, D. E. & SALZBERG, S. L. 2014. Kraken: ultrafast metagenomic sequence classification using exact alignments. *Genome Biol,* 15**,** R46.

WOOLFSON, A., HUEBNER, R., WASAS, A., CHOLA, S., GODFREY-FAUSSETT, P. & KLUGMAN, K. 1997. Nasopharyngeal carriage of community-acquired, antibiotic-resistant *Streptococcus pneumoniae* in a Zambian paediatric population. *Bull World Health Organ,* 75**,** 453-62.

WRIGHT, A., PARRY MORGAN, W., COLEBROOK, L. & DODGSON, R. W. 1914. Observations ON PROPHYLACTIC INOCULATION AGAINST PNEUMOCOCCUS INFECTIONS, AND ON THE RESULTS WHICH HAVE BEEN ACHIEVED BY IT. *The Lancet,* 183**,** 1-10.

WYRES, K. L., LAMBERTSEN, L. M., CROUCHER, N. J., MCGEE, L., VON GOTTBERG, A., LINARES, J., JACOBS, M. R., KRISTINSSON, K. G., BEALL, B. W., KLUGMAN, K. P., PARKHILL, J., HAKENBECK, R., BENTLEY, S. D. & BRUEGGEMANN, A. B. 2013. Pneumococcal capsular switching: a historical perspective. *J Infect Dis,* 207**,** 439-49.

YARO, S., LOURD, M., TRAORE, Y., NJANPOP-LAFOURCADE, B. M., SAWADOGO, A., SANGARE, L., HIEN, A., OUEDRAOGO, M. S., SANOU, O., PARENT DU CHATELET, I., KOECK, J. L. & GESSNER, B. D. 2006. Epidemiological and molecular characteristics of a highly lethal pneumococcal meningitis epidemic in Burkina Faso. *Clin Infect Dis,* 43**,** 693-700.

YIN, Z., RICE, B. D., WAIGHT, P., MILLER, E., GEORGE, R., BROWN, A. E., SMITH, R. D., SLACK, M. & DELPECH, V. C. 2012. Invasive

pneumococcal disease among HIV-positive individuals, 2000-2009. *AIDS,* 26**,** 87-94.

ZULZ, T., WENGER, J. D., RUDOLPH, K., ROBINSON, D. A., RAKOV, A. V., BRUDEN, D., SINGLETON, R. J., BRUCE, M. G. & HENNESSY, T. W. 2013. Molecular characterization of *Streptococcus pneumoniae* serotype 12F isolates associated with rural community outbreaks in Alaska. *J Clin Microbiol,* 51**,** 1402-7.

ZYSK, G., BONGAERTS, R. J., TEN THOREN, E., BETHE, G., HAKENBECK, R. & HEINZ, H. P. 2000. Detection of 23 immunogenic pneumococcal proteins using convalescent-phase serum. *Infect Immun,* 68**,** 3740-3.

Appendix

Appendix 1.1 List of study isolates analysed in chapter 5

Isolate ID	ST	Country	Source	Invasivecarried	Isolation year	Gender
10395_5_30	3081	The Gambia	Lung Aspirate	Invasive	2010	M
10395_5_62	217	Niger	CSF	Invasive	2012	M
10395_5_7	303	Togo	CSF	Invasive	2009	M
10395_5_9	303	Togo	CSF	Invasive	2009	F
11658_1_23	618	The Gambia	Blood	Invasive	1997	F
11658_1_25	618	The Gambia	Blood	Invasive	1999	M
11658_1_32	3081	The Gambia	Blood	Invasive	2011	M
11658_1_33	618	Ghana	CSF	Invasive	2010	F
11658_1_34	618	Togo	CSF	Invasive	2007	F
11658_1_35	618	The Gambia	Lung Aspirate	Invasive	2010	F
11658_1_36	618	Senegal	CSF	Invasive	2010	F
11658_1_37	618	Senegal	CSF	Invasive	2010	M
11658_1_38	618	The Gambia	Blood	Invasive	2010	F
11658_1_40	618	The Gambia	CSF	Invasive	2011	F
11658_1_41	3575	Senegal	CSF	Invasive	2010	M
11658_1_42	618	The Gambia	Blood	Invasive	2011	F
11658_1_46	618	The Gambia	Blood	Invasive	2010	F
11658_1_48	618	The Gambia	Pleural Aspirate	Invasive	2010	F
11658_1_50	618	Togo	CSF	Invasive	2008	F
11658_1_53	618	Benin	CSF	Invasive	2010	M
11658_1_54	618	Senegal	CSF	Invasive	2011	M
11658_1_56	618	Senegal	CSF	Invasive	2010	M
11658_1_57	618	The Gambia	Blood	Invasive	2010	M
11658_1_58	612	Senegal	CSF	Invasive	2010	F

11658_1_69	217	Niger	CSF	Invasive	2011	M
11658_1_73	618	The Gambia	CSF	Invasive	2001	M
11658_1_74	618	The Gambia	Blood	Invasive	1996	M
11658_1_75	618	The Gambia	Blood	Invasive	1996	
11658_1_79	618	The Gambia	CSF	Invasive	2007	F
11658_1_81	3579	The Gambia	Blood	Invasive	1998	M
11658_1_85	618	The Gambia	Blood	Invasive	1998	M
11658_1_88	612	The Gambia	Blood	Invasive	1999	M
11658_1_90	618	The Gambia	Blood	Invasive	1996	F
11658_8_20	618	Nigeria	CSF	Invasive	2010	F
11658_8_23	618	The Gambia	Blood	Invasive	1996	M
11791_1_27	303	Togo	CSF	Invasive	2010	M
11791_1_30	618	Senegal	CSF	Invasive	2010	F
11791_1_32	618	Senegal	CSF	Invasive	2010	F
11791_1_33	618	Togo	CSF	Invasive	2009	M
11791_1_34	618	Togo	CSF	Invasive	2007	M
11791_1_36	618	The Gambia	Pleural Aspirate	Invasive	2010	M
11791_1_38	3575	The Gambia	Lung Aspirate	Invasive	1996	F
11791_1_40	303	The Gambia	Blood	Invasive	2010	F
11791_1_42	618	The Gambia	Blood	Invasive	1997	M
11791_1_44	618	The Gambia	Blood	Invasive	2003	M
11791_1_45	618	The Gambia	Blood	Invasive	2002	F
11791_1_51	612	The Gambia	Joint fluid aspirate	Invasive	1998	
11791_1_55	618	The Gambia	Blood	Invasive	1996	F
11791_1_56	618	The Gambia	Blood	Invasive	1996	M
11940_1_10	618	The Gambia	Blood	Invasive	2011	F
11940_1_12	618	The Gambia	Blood	Invasive	2000	F

14410_4_1	3081	The Gambia	Lung Aspirate	Invasive	2010	M
14410_4_10	3081	The Gambia	Blood	Invasive	2013	F
14410_4_11	3081	The Gambia	Blood	Invasive	2011	M
14410_4_20	3081	The Gambia	Lung Aspirate	Invasive	2012	F
14410_4_27	11779	The Gambia	Blood	Invasive	2013	F
14410_4_31	3081	The Gambia	Blood	Invasive	2012	M
14410_4_4	3081	The Gambia	Blood	Invasive	2013	F
14410_4_43	3081	The Gambia	Blood	Invasive	2012	F
14410_4_51	3081	The Gambia	Blood	Invasive	2012	F
14410_4_52	3081	The Gambia	CSF	Invasive	2012	M
14410_4_54	3081	The Gambia	Blood	Invasive	2013	M
14410_4_56	3081	The Gambia	Blood	Invasive	2012	F
14410_4_60	11779	The Gambia	Blood	Invasive	2013	F
14410_4_61	3081	The Gambia	CSF	Invasive	2012	F
14410_4_63	217	The Gambia	Blood	Invasive	2012	F
14410_4_64	3081	The Gambia	CSF	Invasive	2012	F
14410_4_71	3081	The Gambia	Blood	Invasive	2010	F
14410_4_81	3081	The Gambia	Lung Aspirate	Invasive	2011	M
14410_4_82	3081	The Gambia	Lung Aspirate	Invasive	2012	M
14410_4_85	3081	The Gambia	Blood	Invasive	2013	F
14410_5_18	3081	The Gambia	Blood	Invasive	2013	M
14410_5_21	3081	The Gambia	Blood	Invasive	2010	F
14410_5_24	3081	The Gambia	Lung Aspirate	Invasive	2010	M
14410_5_30	3081	The Gambia	Blood	Invasive	2012	F
14410_5_35	3081	The Gambia	Lung Aspirate	Invasive	2009	M
14410_5_39	3081	The Gambia	Lung Aspirate	Invasive	2008	M
14410_5_42	3081	The Gambia	Lung Aspirate	Invasive	2009	M

14410_5_43	3081	The Gambia	Blood	Invasive	2009	F
14410_5_44	3081	The Gambia	Blood	Invasive	2009	M
14410_5_45	3081	The Gambia	Lung Aspirate	Invasive	2009	M
14410_5_47	3081	The Gambia	Blood	Invasive	2010	M
14410_5_48	3081	The Gambia	Blood	Invasive	2009	M
14410_5_49	10649	The Gambia	Blood	Invasive	2008	M
14410_5_50	3081	The Gambia	Lung Aspirate	Invasive	2010	F
14410_5_67	3081	The Gambia	CSF	Invasive	2012	M
14410_5_7	3081	The Gambia	Lung Aspirate	Invasive	2013	F
14410_6_1	10649	The Gambia	Blood	Invasive	2013	M
14410_6_14	3081	The Gambia	Blood	Invasive	2011	M
14410_6_19	3081	The Gambia	Blood	Invasive	2011	M
14410_6_26	3081	The Gambia	Lung Aspirate	Invasive	2010	M
14410_6_30	3081	The Gambia	Blood	Invasive	2013	F
14410_6_33	3081	The Gambia	Blood	Invasive	2011	M
14410_6_38	3081	The Gambia	Blood	Invasive	2009	M
14410_6_39	3081	The Gambia	Blood	Invasive	2012	M
14410_6_4	3081	The Gambia	Blood	Invasive	2013	M
14410_6_41	3081	The Gambia	Blood	Invasive	2011	F
14410_6_48	217	The Gambia	Blood	Invasive	2013	M
14410_6_56	3081	The Gambia	Blood	Invasive	2013	M
14410_6_64	3081	The Gambia	Lung Aspirate	Invasive	2011	F
14410_6_76	618	The Gambia	Blood	Invasive	2003	M
14410_6_8	3081	The Gambia	Blood	Invasive	2012	F
14410_6_80	618	The Gambia	Blood	Invasive	2002	F
14410_6_81	3581	The Gambia	Pleural Aspirate	Invasive	2002	F
14410_6_82	3081	The Gambia	Blood	Invasive	2008	M

14410_6_9	3081	The Gambia	Joint fluid aspirate	Invasive	2009	M
14410_7_1	3081	The Gambia	Blood	Invasive	2007	M
14410_7_11	3081	The Gambia	CSF	Invasive	2008	F
14410_7_12	217	The Gambia	Blood	Invasive	2003	M
14410_7_13	2084	The Gambia	Blood	Invasive	2002	M
14410_7_14	217	The Gambia	Blood	Invasive	2003	M
14410_7_19	618	The Gambia	Blood	Invasive	2003	F
14410_7_25	618	The Gambia	Blood	Invasive	2002	F
14410_7_26	618	The Gambia	Blood	Invasive	2003	
14410_7_33	618	The Gambia	Blood	Invasive	2002	
14410_7_38	618	The Gambia	Blood	Invasive	2002	F
14410_7_39	618	The Gambia	Blood	Invasive	2002	F
14410_7_46	3081	The Gambia	Blood	Invasive	2008	F
14410_7_52	3081	The Gambia	Lung Aspirate	Invasive	2007	M
14410_7_54	217	The Gambia	Blood	Invasive	2007	M
14410_7_6	3581	The Gambia	CSF	Invasive	2002	M
14410_7_60	618	The Gambia	Blood	Invasive	2002	F
14410_7_61	217	The Gambia	Blood	Invasive	2003	M
14410_7_69	3081	The Gambia	Blood	Invasive	2008	F
14410_7_75	618	The Gambia	Blood	Invasive	2003	F
14410_7_76	217	The Gambia	Blood	Invasive	2003	M
14410_7_79	2084	The Gambia	Blood	Invasive	2002	M
14410_7_83	3081	The Gambia	Blood	Invasive	2007	M
14410_7_9	217	The Gambia	CSF	Invasive	2003	F
14520_6_61	303	Ghana	Blood	Invasive	2011	F
14520_6_63	3081	Senegal	Joint fluid aspirate	Invasive	2012	M
14520_6_73	3081	Senegal	CSF	Invasive	2012	F

14520_6_79	217	Togo	CSF	Invasive	2012	M
14520_6_81	303	Niger	CSF	Invasive	2009	
14520_6_82	3081	The Gambia	Blood	Invasive	2010	M
14520_7_5	3081	Senegal	CSF	Invasive	2010	M
14520_7_6	3081	Senegal	CSF	Invasive	2010	M
14520_7_77	3081	The Gambia	Blood	Invasive	2012	M
14520_7_78	3081	The Gambia	Blood	Invasive	2012	M
14520_7_79	3081	Senegal	Blood	Invasive	2012	F
14520_7_81	303	Niger	CSF	Invasive	2012	M
14520_7_89	303	Niger	CSF	Invasive	2010	M
15378_2_30	3081	The Gambia	Lung Aspirate	Invasive	2010	F
15378_2_49	3081	The Gambia	Joint fluid aspirate	Invasive	2008	F
15378_2_50	3081	The Gambia	Blood	Invasive	2010	M
15531_4_50	3081	The Gambia	Lung Aspirate	Invasive	2011	M
15531_4_64	3081	The Gambia	Joint fluid aspirate	Invasive	2008	F
15531_4_65	3081	The Gambia	Joint fluid aspirate	Invasive	2008	F
15531_4_89	3081	The Gambia	Lung Aspirate	Invasive	2011	F
15531_4_91	217	The Gambia	Blood	Invasive	2009	M
15531_5_10	3081	The Gambia	Blood	Invasive	2010	
15531_5_15	3081	The Gambia	Blood	Invasive	2008	
15531_5_31	3081	The Gambia	CSF	Invasive	2004	
15531_5_35	3575	The Gambia	Blood	Invasive	1999	
15531_5_40	3081	The Gambia	Blood	Invasive	2012	
15531_5_53	3081	The Gambia	Lung Aspirate	Invasive	2012	
15531_6_18	3081	The Gambia	Blood	Invasive	2014	M
15531_6_22	3081	The Gambia	Blood	Invasive	2014	F
15531_6_24	3081	The Gambia	Lung Aspirate	Invasive	2014	M

15531_6_29	3081	The Gambia	Blood	Invasive	2014	F
15531_6_32	3081	The Gambia	Blood	Invasive	2014	F
15531_6_34	3081	The Gambia	Lung Aspirate	Invasive	2014	F
15531_6_37	3081	The Gambia	Blood	Invasive	2014	F
15531_6_45	3081	The Gambia	Blood	Invasive	2014	F
15531_6_54	3081	The Gambia	Blood	Invasive	2014	M
15531_6_58	3081	The Gambia	Blood	Invasive	2014	M
15531_6_60	3081	The Gambia	Blood	Invasive	2014	M
15531_6_63	3081	The Gambia	Blood	Invasive	2014	M
15531_6_70	3081	The Gambia	Blood	Invasive	2014	M
15531_6_93	3081	The Gambia	NP swab	Carriage	2009	F
15531_7_22	3081	The Gambia	NP swab	Carriage	2009	M
15531_7_32	618	The Gambia	Blood	Invasive	2005	
15531_7_36	618	The Gambia	Blood	Invasive	2005	
15531_7_40	618	The Gambia	Blood	Invasive	2004	
15873_2_57	3081	The Gambia	Blood	Invasive	2014	M
15873_2_59	3081	The Gambia	Blood	Invasive	2007	F
15873_3_10	3081	The Gambia	Lung Aspirate	Invasive	2012	M
15873_3_12	217	Niger	CSF	Invasive	2010	M
15873_3_15	2084	The Gambia	Blood	Invasive	2003	F
15873_3_9	217	Togo	CSF	Invasive	2011	M
15873_4_58	3081	The Gambia	Blood	Invasive	2014	M
15873_4_59	3081	The Gambia	Lung Aspirate	Invasive	2014	F
15873_5_23	3081	The Gambia	CSF	Invasive	2012	M
15873_5_66	3081	The Gambia	Blood	Invasive	2014	F
16399_5_13	217	The Gambia	NP swab	Carriage	2009	M
16399_7_50	3081	The Gambia	NP swab	Carriage	2009	M

16399_7_63	3081	The Gambia	NP swab	Carriage	2009	M
16399_7_80	3081	The Gambia	NP swab	Carriage	2009	M
16453_2_3	3081	The Gambia	NP swab	Carriage	2009	F
17794_7_153	12310	The Gambia	NP swab	Carriage	2013	M
17794_8_1	3081	The Gambia	NP swab	Carriage	2009	M
17794_8_114	3081	The Gambia	NP swab	Carriage	2009	M
17794_8_93	3081	The Gambia	NP swab	Carriage	2010	F
19183_3_26	618	The Gambia	NP swab	Carriage	2007	
19183_3_27	618	The Gambia	NP swab	Carriage	2007	
19183_3_28	618	The Gambia	NP swab	Carriage	2007	
19183_3_29	618	The Gambia	NP swab	Carriage	2007	
19183_3_30	3081	The Gambia	NP swab	Carriage	2007	
19183_3_31	3081	The Gambia	NP swab	Carriage	2007	
19183_3_32	618	The Gambia	NP swab	Carriage	2007	
19183_3_35	3081	The Gambia	NP swab	Carriage	2008	
19183_3_36	3081	The Gambia	NP swab	Carriage	2007	
19183_3_37	3081	The Gambia	NP swab	Carriage	2008	
19183_3_38	3081	The Gambia	NP swab	Carriage	2008	
19183_3_39	3081	The Gambia	NP swab	Carriage	2007	
19183_3_40	3081	The Gambia	NP swab	Carriage	2007	
19183_3_41	303	The Gambia	NP swab	Carriage	2007	
19183_3_42	3081	The Gambia	NP swab	Carriage	2008	
19183_3_43	3081	The Gambia	NP swab	Carriage	2007	
19183_3_44	3081	The Gambia	NP swab	Carriage	2007	
19183_3_48	3081	The Gambia	NP swab	Carriage	2008	
19183_3_49	3081	The Gambia	NP swab	Carriage	2007	
19183_3_50	618	The Gambia	NP swab	Carriage	2008	

19183_3_53	217	The Gambia	NP swab	Carriage	2007	
19183_3_54	618	The Gambia	NP swab	Carriage	2007	
19183_3_55	3081	The Gambia	NP swab	Carriage	2008	
19183_3_56	3081	The Gambia	NP swab	Carriage	2007	
19183_3_57	3081	The Gambia	NP swab	Carriage	2008	
19183_3_58	3081	The Gambia	NP swab	Carriage	2007	
19183_3_60	3081	The Gambia	NP swab	Carriage	2008	
19183_3_61	3081	The Gambia	NP swab	Carriage	2007	
19183_3_62	618	The Gambia	NP swab	Carriage	2007	
19183_3_63	3081	The Gambia	NP swab	Carriage	2008	
19183_3_64	3081	The Gambia	NP swab	Carriage	2008	
19183_3_65	3081	The Gambia	NP swab	Carriage	2007	
19183_3_66	618	The Gambia	NP swab	Carriage	2007	
19183_3_67	618	The Gambia	NP swab	Carriage	2007	
19183_3_68	3081	The Gambia	NP swab	Carriage	2007	
19183_3_69	3081	The Gambia	NP swab	Carriage	2008	
19183_3_71	3081	The Gambia	NP swab	Carriage	2007	
19183_3_72	618	The Gambia	NP swab	Carriage	2007	
19183_3_73	618	The Gambia	NP swab	Carriage	2007	
19183_3_74	618	The Gambia	NP swab	Carriage	2007	
19183_3_75	3081	The Gambia	NP swab	Carriage	2007	
19183_3_76	303	The Gambia	NP swab	Carriage	2007	
19183_3_77	3081	The Gambia	NP swab	Carriage	2007	
19183_3_78	618	The Gambia	NP swab	Carriage	2007	
19183_3_83	3081	The Gambia	NP swab	Carriage	2007	
19183_3_84	3081	The Gambia	NP swab	Carriage	2008	
19183_3_85	3081	The Gambia	NP swab	Carriage	2008	

19183_3_86	3081	The Gambia	NP swab	Carriage	2007	
19183_3_87	3081	The Gambia	NP swab	Carriage	2007	
19183_3_89	3081	The Gambia	NP swab	Carriage	2008	
19341_1_158	3081	The Gambia	Lung Aspirate	Invasive	2015	M
19341_1_164	3081	The Gambia	Blood	Invasive	2015	M
19341_1_165	3081	The Gambia	Blood	Invasive	2015	M
19341_1_2	12197	The Gambia	CSF	Invasive	2015	F
20402_4_106	3081	The Gambia	Blood	Invasive	2015	M
20402_4_156	3081	The Gambia	NP swab	Carriage	2014	F
20402_4_181	3081	The Gambia	NP swab	Carriage	2011	F
20402_4_2	12197	The Gambia	CSF	Invasive	2015	F

Appendix 1.2 List of genes unique to clades in the accessory region of the genome

Unique to clade 1		
Gene	**Non-unique Gene name**	**Annotation**
prtR		HTH-type transcriptional regulator PrtR
group_3915		hypothetical protein
bspRIM		Modification methylase BspRI
sugB		Trehalose transport system permease protein SugB
group_3901		hypothetical protein
fucA		L-fuculose phosphate aldolase
rnhB		Ribonuclease HII
rafA_1		Alpha-galactosidase
group_223		hypothetical protein
group_3197		hypothetical protein
araP		L-arabinose transport system permease protein AraP
group_136		hypothetical protein
rhaB		L-Rhamnulokinase
group_3905		hypothetical protein
repN		Replication initiation protein
group_3907		hypothetical protein
group_3909		hypothetical protein
group_3910		hypothetical protein
group_3911		hypothetical protein
group_660		hypothetical protein
group_3914		hypothetical protein
hsdR		Type-1 restriction enzyme R protein
Int-Tn_1		Transposase from transposon Tn916
group_3898		hypothetical protein
sftA		DNA translocase SftA
group_3921		hypothetical protein
group_3922		hypothetical protein
hmo		4-hydroxymandelate oxidase
Unique to clade 3		
group_808		putative HTH-type transcriptional regulator
group_1811	*recX*	Regulatory protein RecX
bglA_2		Aryl-phospho-beta-D-glucosidase BglA
group_1789		Arylsulfatase
glcA		Glucan endo-1,3-beta-glucosidase A1
group_514	*glnP_1*	Glutamine transport system permease protein GlnP
group_4999	*mutX_1*	8-oxo-dGTP diphosphatase
xerC_1		Tyrosine recombinase XerC

tcyB		hypothetical protein
Unique to clade 1 ad 2		
glnP_2		putative glutamine ABC transporter permease protein GlnP
mutX_2		8-oxo-dGTP diphosphatase
gutB		Sorbitol dehydrogenase
group_807		putative HTH-type transcriptional regulator
butA		Aminoglycoside 3'-phosphotransferase
leuA_1		2-isopropylmalate synthase
recX		Regulatory protein RecX
bglH_1		Aryl-phospho-beta-D-glucosidase BglH
group_397	*natA_1*	ABC transporter ATP-binding protein NatA
msmE_2		Multiple sugar-binding protein
eabC		Blood-group-substance endo-1,4-beta-galactosidase
Unique to clade 2		
capA		Capsule biosynthesis protein CapA
Unique to clade 2 ad 3		
group_410	*hmo*	4-hydroxymandelate oxidase
group_539	*ecfT_1*	Energy-coupling factor transporter transmembrane protein EcfT
msbA_2		putative ABC transporter ATP-binding protein
doc		Toxin Doc
group_572	*fucA*	L-fuculose phosphate aldolase
group_542	*ykoD_1*	Putative HMP/thiamine import ATP-binding protein YkoD
fucU		L-fucose mutarotase
group_431	*manX_3*	PTS system mannose-specific EIIAB component
group_1497	*eabC*	Blood-group-substance endo-1,4-beta-galactosidase
soxS		Regulatory protein SoxS
group_3293	*manZ_3*	Mannose permease IID component
levE		Fructose-specific phosphotransferase enzyme IIB component
group_1499	*rhaB*	L-Rhamnulokinase
group_3356	*rnhB*	Ribonuclease HII
manY_2		Mannose permease IIC component
group_428	*fucI*	L-fucose isomerase
irtA		Iron import ATP-binding/permease protein IrtA
group_873	*adhB*	Alcohol dehydrogenase 2

Unique to clade 1 ad 3		
bglA		6-phospho-beta-glucosidase BglA
group_3507		hypothetical protein
group_582		hypothetical protein
group_140		hypothetical protein
group_1758		hypothetical protein
licC_2		Lichenan permease IIC component
lacF_1		Lactose-specific phosphotransferase enzyme IIA component
licB_1		Lichenan-specific phosphotransferase enzyme IIB component
group_3918		hypothetical protein
licR_1		putative licABCH operon regulator
group_75		hypothetical protein
group_1021		hypothetical protein
group_139		hypothetical protein
ricR		Copper-sensing transcriptional repressor RicR
iscS_1		Cysteine desulfurase IscS