
[All ETDs from UAB](#)

[UAB Theses & Dissertations](#)

2010

Evolution of Pneumococcal Serogroup 6

Preston Eugene Bratcher
University of Alabama at Birmingham

Follow this and additional works at: <https://digitalcommons.library.uab.edu/etd-collection>

Recommended Citation

Bratcher, Preston Eugene, "Evolution of Pneumococcal Serogroup 6" (2010). *All ETDs from UAB*. 1244.
<https://digitalcommons.library.uab.edu/etd-collection/1244>

This content has been accepted for inclusion by an authorized administrator of the UAB Digital Commons, and is provided as a free open access item. All inquiries regarding this item or the UAB Digital Commons should be directed to the [UAB Libraries Office of Scholarly Communication](#).

EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6

by

PRESTON EUGENE BRATCHER

MOON H. NAHM, COMMITTEE CHAIR
WILLIAM H. BENJAMIN, JR.
DAVID E. BRILES
SUSAN K. HOLLINGSHEAD
DAVID G. PRITCHARD
HUI WU

A DISSERTATION

Submitted to the graduate faculty of The University of Alabama at Birmingham,
in partial fulfillment of the requirements for the degree of
Doctor of Philosophy

BIRMINGHAM, ALABAMA

2010

Copyright by
Preston Eugene Bratcher
2010

EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6

PRESTON EUGENE BRATCHER

MICROBIOLOGY

ABSTRACT

Serogroup 6 of *Streptococcus pneumoniae* has been known to contain three serotypes, named 6A, 6B and 6C, with highly homologous capsule gene loci. The 6A and 6B capsule gene loci consistently differ from each other by only one nucleotide in the *wciP* gene. The 6A capsule gene locus has a galactosyltransferase (*wciN_α*), which has been replaced with a glucosyltransferase (*wciN_β*) in the 6C capsule gene locus. We considered that a new serotype, 6D, would be possible if the glucosyltransferase found in 6C strains was combined with the *wciP* gene from a 6B strain. We demonstrate that this gene combination yields a viable pneumococcal strain and that the capsular polysaccharide (PS) from this strain has the predicted chemical structure and serological similarity to the capsular PS of the 6B serotype. The new 6D serotype can be distinguished from serotype 6B strains with a set of mAbs, and by using these mAbs, we were able to discover two 6D isolates obtained from Korea. Therefore, the novel 6D serotype is structurally unique and clinically relevant. After the discovery of 6C and the novel 6D, it became necessary to reexamine the evolution relationship of serogroup 6. In addition to supporting the conclusions from previous studies, the results presented in this study have allowed us to build a hypothetical model for serogroup 6 evolution. By analyzing sequences of a diverse set of 6C strains, we hypothesize that the *cps* loci from all these strains have a single origin, due to their low level of diversity. While the source of the foreign gene incorporated into a serogroup 6 strain to create 6C is still unknown,

our results provide evidence that the source of this gene in 6D was from a 6C strain. Furthermore, through MLST analysis we have examined the mechanism of the well-documented expansion of serotype 6C post common use of the PCV-7 vaccine.

Keywords: *Streptococcus pneumoniae*, serogroup, 6, evolution, capsule

TABLE OF CONTENTS

	<i>Page</i>
ABSTRACT	iii
LIST OF TABLES	vi
LIST OF FIGURES	vii
LIST OF ABBREVIATIONS	ix
INTRODUCTION	1
PRODUCTION OF A UNIQUE PNEUMOCOCCAL CAPSULE SEROTYPE BELONGING TO SEROGROUP 6	8
IDENTIFICATION OF NATURAL PNEUMOCOCCAL ISOLATES EXPRESSING SEROTYPE 6D BY GENETIC, BIOCHEMICAL, AND SEROLOGICAL CHARACTERIZATION	37
EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6 POST-DISCOVERY OF THE NOVEL 6C AND 6D SEROTYPES	56
SUMMARIZING CONCLUSION	83
LIST OF GENERAL REFERENCES	87

LIST OF TABLES

<i>Table</i>	<i>Page</i>
PRODUCTION OF A UNIQUE PNEUMOCOCCAL CAPSULE SEROTYPE BELONGING TO SEROGROUP 6	
1 List of PCR Primers	13
EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6 POST-DISCOVERY OF THE NOVEL 6C AND 6D SEROTYPES	
1 List of PCR Primers	61
2 Strain data	71

LIST OF FIGURES

Table Page

INTRODUCTION

1 Structure of Serogroup 6 Capsular Polysaccharide Repeating Units.....	4
2 Capsule Polysaccharide Synthesis Loci of Serogroup 6 Members.....	5

PRODUCTION OF A UNIQUE PNEUMOCOCCAL CAPSULE SEROTYPE BELONGING TO SEROGROUP 6

1 Diagram of the <i>wciN</i> region exchange experiment.....	15
2 GLC-MS chromatograms showing carbohydrate composition of capsular PS from pneumococcal strains expressing 6B and 6X1 serotypes and before and after periodate treatment	21
3a Proposed structure of the hydrated form of the repeating unit of 6X1 capsular PS.....	23
3b Mass spectrum of the repeating units.....	23
3c Daughter ions of the ion of 683.3 amu shown in 3B	23
4 Ability of various capsular PSs (2 mg ml ⁻¹) to inhibit binding of mAb to ELISA plates (y axis) after the PSs were hydrolysed for various time periods (x axis).....	26
5a/b Ability of lysates of various pneumococcal strains to inhibit the binding of mAbs Hyp6BM7 (a) and Hyp6BM8 (b) to 6B PS-coated ELISA plates.....	26

IDENTIFICATION OF NATURAL PNEUMOCOCCAL ISOLATES EXPRESSING SEROTYPE 6D BY GENETIC, BIOCHEMICAL, AND SEROLOGICAL CHARACTERIZATION

1 Flow cytometry profiles of various pneumococcal isolates stained with mAbs.....	44
--	----

LIST OF FIGURES (CONTINUED)

2 Binding of mAbs (y -axis) to 6B PS-coated ELISA plates in the presence of varying amounts of bacterial lysates (x -axis)46

3 Amount of antigenicity (y -axis) remaining after alkali hydrolysis for different time periods (x -axis)49

4a/b GLC profile of serotype 6B PS (a) and the PS purified from a pneumococcal isolate MNZ21 (b)51

EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6 POST-DISCOVERY OF THE NOVEL 6C AND 6D SEROTYPES

1 Capsule gene loci of a 6A strain, 6B strain, 6C strain, and 6D strain.....66

2 Sequence diversity of the $wciN_{\beta}$ and its flanking regions for serotype 6C and 6D strains63

3 Neighbor joining trees of unique cps profiles.....67

4 Two models for serogroup 6 evolution.....76

LIST OF ABBREVIATIONS

Amu – atomic mass units

CDC – Center for Disease Control and Prevention

c.f.u.- colony forming units

COPD – Chronic Obstructive Pulmonary Disease

cps – capsular polysaccharide synthesis

DNA – Deoxyribonucleic Acid

dNTP - deoxynucleotide triphosphate

ELISA – enzyme-linked immunosorbant assay

FACS – fluorescence activated cell sorting

FBS – fetal bovine serum

Fig – figure

GLC-MS – gas/liquid chromatography – mass spectrometry

mAb – monoclonal antibody

Min – minute

MLST – multilocus sequence typing

MS-MS – tandem mass spectrometry

PBS – phosphate buffered saline

PCR – polymerase chain reaction

PCV-7 – Heptavalent Pneumococcal Conjugate Vaccine

PS – polysaccharide

SDS - sodium dodecyl sulfate

LIST OF ABBREVIATIONS (CONTINUED)

SNP – single nucleotide polymorphism

ST – sequence type

THY – Todd Hewitt Broth with yeast extract (0.5%)

UAB – University of Alabama at Birmingham

$wciN_{\alpha}$ – $wciN_{6A}$

$wciN_{\beta}$ – $wciN_{6C}$

INTRODUCTION

Streptococcus pneumoniae is a gram positive bacterium commonly found asymptotically colonizing the human nasopharynx, which is its natural niche as an obligate human bacterium. *S. pneumoniae* can, however, invade further into the body, causing diseases such as otitis media, pneumonia, sepsis, and meningitis (Kaijser, 1979). Pneumococcal disease is most common in populations with weak immune systems, mainly children, the elderly, and the immunocompromised (Lynch & Zhanel, 2009). In order to persist and disseminate from its natural reservoir, the pneumococcus has evolved several colonization/virulence factors which facilitate the persistence in and dissemination from its natural reservoir. Amongst these factors are antibiotic resistance, the ability to form biofilms, an IgA protease, three surface-bound exoglycosidases, a pneumolysin, natural competence, and the polysaccharide capsule (Kadioglu *et al.*, 2008). The last of these is thought to be one of the most important of these factors, if not the most important, as unencapsulated variants are avirulent in animal models of disease and are unable to colonize the upper respiratory tract in animal models (Magee & Yother, 2001; Morona *et al.*, 2004). Additionally, the majority of disease isolates possess a capsule.

The pneumococcus coats itself with a polysaccharide capsule, which is thought to function by shielding the bacterium from recognition by the host. This includes reducing inflammation in the host and a reduction in opsonophagocytosis (Avery & Dubos, 1931; Tuomanen *et al.*, 1985). Each bacterium produces a capsule of a single structure (referred to as a serotype), and at least 93 unique structures have been discovered so far (Bratcher *et al.*, 2009; Calix & Nahm, 2010; Henrichsen, 1995; Park *et al.*, 2007b). Some

of these structures are very closely related, both serologically and structurally, and these closely related types are said to belong to the same serogroup.

For 91 of the 93 known serotypes, all genes responsible for the synthesis of the polysaccharide capsule are found at a single locus in the genome between the *dexB* and *aliA* genes. This locus is referred to as the capsular polysaccharide synthesis (*cps*) locus. Downstream of the *dexB* gene are four capsule regulatory genes, *wzg*, *wzh*, *wzd*, and *wze* (or *cpsABCD*), which are common to the 91 serotypes. Downstream of the regulatory genes are the genes responsible for building the specific capsule repeating unit, a flippase (*wzx*), a polymerase (*wzy*), as well as genes necessary for synthesizing necessary precursors for building the capsule repeating unit. Interestingly, these central region, capsule-specific genes have a lower G+C content than the average for the pneumococcal genome and the capsule regulatory genes, suggesting they may have originated from a foreign source. Downstream of the central region and upstream of *aliA*, 40 serotypes have four genes responsible for rhamnose synthesis, and these genes all have the average pneumococcal G+C content (Bentley *et al.*, 2006).

Because the pneumococcal capsule is shielding the bacteria's surface, it is the main antigen presented to the host, and it has been shown that natural colonization results in antibodies specific to the capsule type present (Musher *et al.*, 1997). It may be because of this capsule-specific protective response that pneumococci have diversified the structures of their capsular polysaccharides to such an extent. This same strategy of raising antibodies against the capsule is used today for immunizations to protect against pneumococcal infections. Currently used vaccines contain purified capsule to create serotype-specific protection in immunized people. PPV-23 (Pneumovax) contains

capsular polysaccharide from 23 different serotypes, but is only effective in adults, as children do not respond well to T cell-Independent antigens (Barrett, 1985). For children, polysaccharide conjugated to a diphtheria toxin is used to induce protective immunity, and the conjugate vaccine, PCV-7 (Prevnar), contains capsular polysaccharide from seven serotypes (Rennels *et al.*, 1998). PCV-10 (Synflorix), which is licensed in Europe, contains ten polysaccharides, eight of which are conjugated to non-typeable *Haemophilus influenzae* protein D (Prymula & Schuerman, 2009). Recently, a 13-valent conjugate vaccine was licensed in the United States (2010).

The 93 known serotypes are not equally prevalent in carriage and infections. Therefore, vaccines only incorporate polysaccharides from the most prevalent serotypes in order to increase effectiveness in reducing disease. For example, polysaccharide from serotype 6B is included in PPV23, PCV-7, PCV-10, and PCV-13. Both established members of serogroup 6, 6A and 6B, are prevalent in both carriage and disease, and immunization with 6B polysaccharide is thought to produce a response that is cross-protective against 6A infection (Gould *et al.*, 1987; Inostroza *et al.*, 1998; Vakevainen *et al.*, 2001). Because of the prevalence of serogroup 6 isolates worldwide, this serogroup has been well studied epidemiologically, biochemically, and genetically. Biochemically, the repeating units of the two polysaccharides have closely related structures; the 6A PS repeating unit has the structure of [\rightarrow 2) galactose (1 \rightarrow 3) glucose (1 \rightarrow 3) rhamnose (1 \rightarrow 3) ribitol (5 \rightarrow phosphate)] (Figure 1). The 6B repeating unit differs only in the rhamnose-ribitol linkage, which is (1 \rightarrow 4) in 6B (Kenne *et al.*, 1979; Rebers & Heidelberger, 1961). Genetically, they are also closely related, with the only conserved difference between the

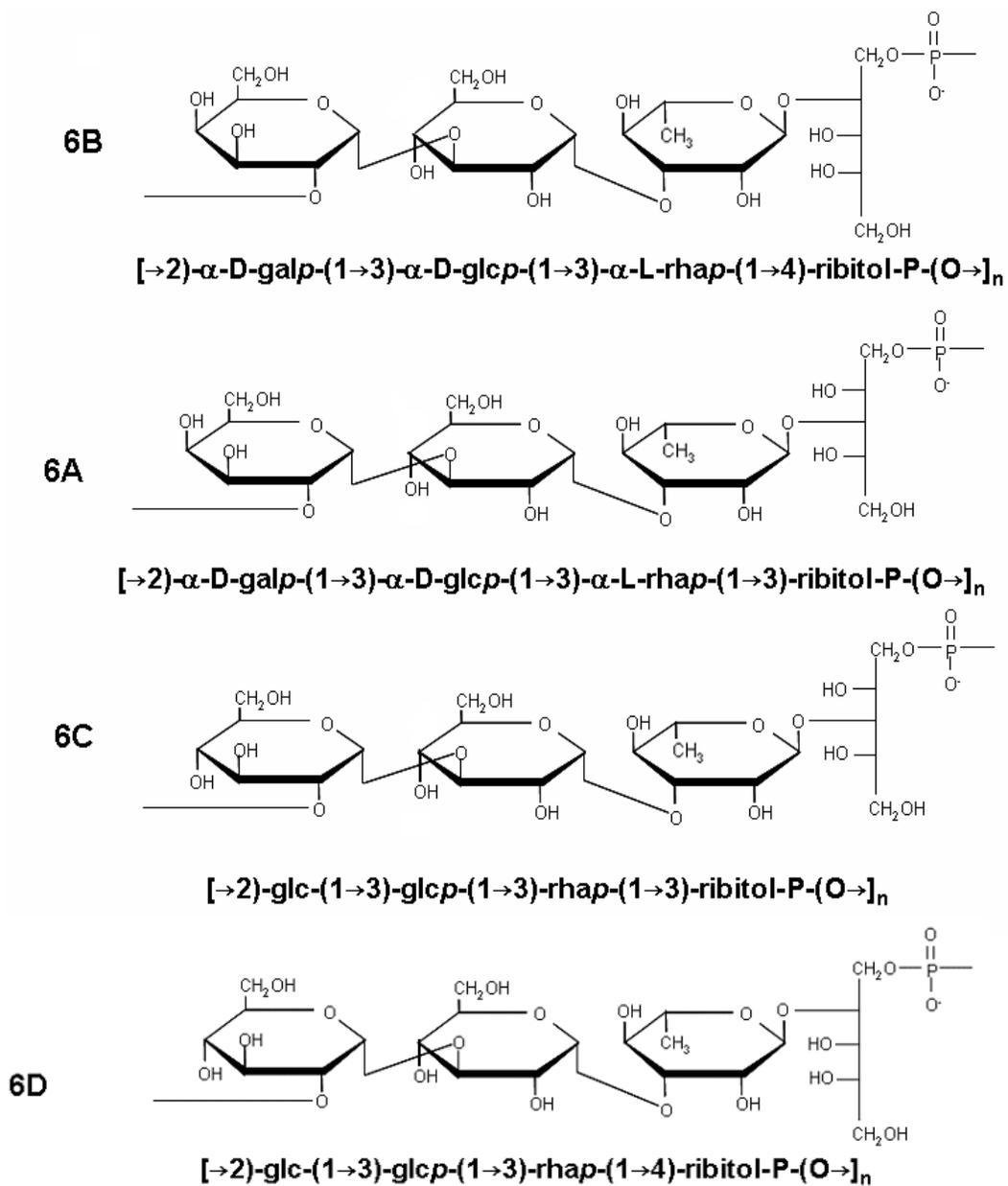


Figure 1. Structure of Serogroup 6 Capsular Polysaccharide Repeating Units. Sugars are abbreviated as follows: glucose (glc), galactose (gal) and rhamnose (rha).

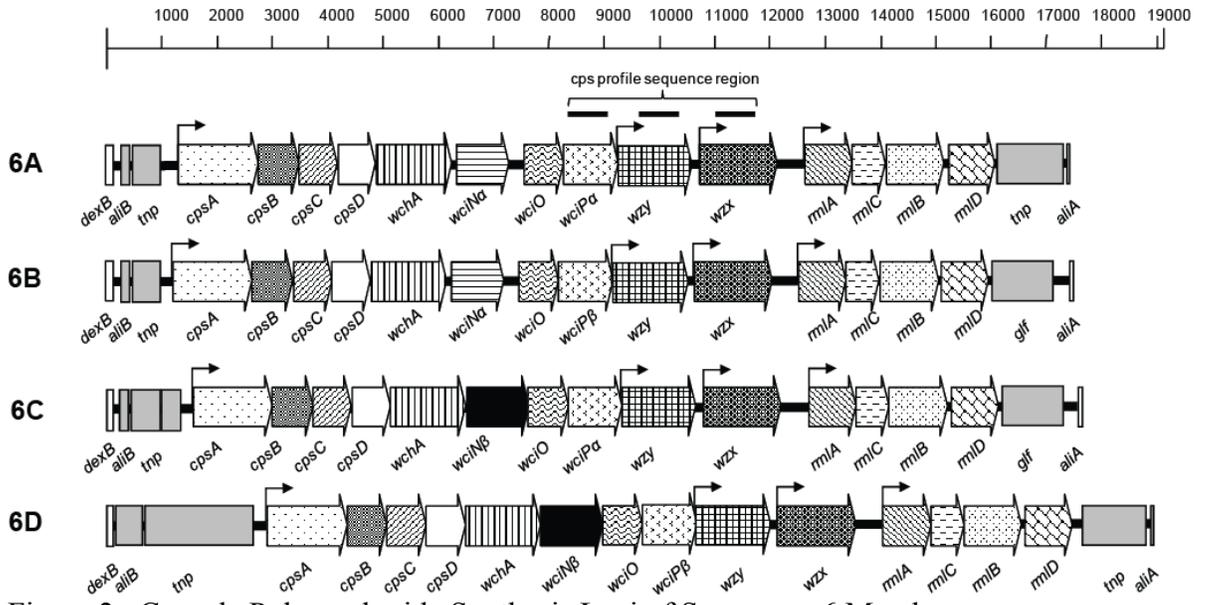


Figure 2. Capsule Polysaccharide Synthesis Loci of Serogroup 6 Members

two serotypes being a SNP in the *wciP* gene, which codes for the rhamnosyltransferase (Figure 2). All 6A strains contain a G at nt 584 of *wciP* [resulting in a serine at residue 195 (referred to as *wciP_α*)], while all 6B strains have an A at this location [resulting in an asparagine at residue 195 (referred to as *wciP_β*)]. It is interesting to note that there seems to be two divergent classes of serogroup 6 sequences, which have around 5.4% divergence between them, with both 6A and 6B equally containing what will be referred to as class 1 sequences and 6B alone containing class 2 sequences. In addition to the sequence divergence between the two classes, the majority of class 2 sequences contain a putatively nonfunctional region of ~300 bps referred to as the INDEL sequence (Mavroidi *et al.*, 2004).

In 2006, a study to validate a multiplex serotyping assay led to the discovery of a subtype within the collection of 6A strains which had a different monoclonal antibody binding profile than 6A, while genetically containing *wciP_α* (Lin *et al.*, 2006). Biochemical analysis revealed that this subset has a unique repeating unit structure, with a glucose replacing the galactose residue in the 6A repeating unit (Figure 1), and therefore this subtype was declared a new serotype, 6C (Park *et al.*, 2007b). Further study showed that, genetically, these 6C strains are different from 6A in that they have replaced the *wciN* gene found in 6A and 6B [referred to herein as *wciN_α* (previously called *wciN_{6A}*)], with a gene from a foreign source that has only 50% homology to *wciN_α* [referred to as *wciN_β* (previously called *wciN_{6C}*)] (Park *et al.*, 2007a) (Figure 2). Since the discovery and characterization of this new serotype, several studies examining the epidemiology of 6C have been published, which collectively have led to the conclusion that serotype 6C is increasing in prevalence in populations vaccinated with 6B (Carvalho

Mda *et al.*, 2009; Leach *et al.*, 2009; Nahm *et al.*, 2009; Park *et al.*, 2008; Tocheva *et al.*, 2010). The discovery of this new serotype has raised several questions which we have sought to answer in the research presented within this dissertation. By examining the sequence variation in a locationally diverse set of 6C clinical isolates, we are able to gain insights into relationships amongst the 6C strains, the number of 6C creation events, and the mechanism of 6C expansion.

Due to the fact that, genetically, there are two genes responsible for the structural differences amongst the members of serogroup 6 (*wciN* and *wciP*), and that each of these genes has two “versions”(α and β), we hypothesized that a fourth member could exist which contained *wciN β* and *wciP β* (Figure 2). The resulting polysaccharide produced from this strain would have a 1 \rightarrow 4 rhamnose-ribitol linkage while lacking galactose and having two glucoses in the repeating unit (Figure 1). The work contained in this dissertation sought to answer the questions of whether this genetic combination produces a capsular polysaccharide, what would be the structure of the capsule if one were produced, and does this serotype exist in nature.

Although the evolution of serogroup 6 has been examined previously, the discovery of a new member of serogroup 6 has prompted us to take another look at the genetic evolution of these serotypes. By combining our data with that of the previous study, we are able to compose a model of the evolution of the serogroup 6 cps loci as well as form hypotheses based on the data as to the mechanisms controlling this evolution.

PRODUCTION OF A UNIQUE PNEUMOCOCCAL CAPSULE SEROTYPE
BELONGING TO SEROGROUP 6

by

PRESTON E. BRATCHER, IN HO PARK, SUSAN K. HOLLINGSHEAD, MOON H.
NAHM

Microbiology. 2009 Feb; 155(Pt 2): 576-83.

Copyright

2009

by

Society for General Microbiology

Used by permission

Format adapted and errata corrected for dissertation

Abstract

Serogroup 6 of *Streptococcus pneumoniae* contains three serotypes, named 6A, 6B and 6C, with highly homologous capsule gene loci. The 6A and 6B capsule gene loci consistently differ from each other by only one nucleotide in the *wciP* gene. The 6A capsule gene locus has a galactosyltransferase, which has been replaced with a glucosyltransferase in the 6C capsule gene locus. We considered that a new serotype named '6X1' would be possible if the galactosyltransferase of the 6B capsule gene locus is replaced with the glucosyltransferase of 6C. We demonstrate that this gene transfer yields a viable pneumococcal strain and that the capsular polysaccharide (PS) from this strain has the predicted chemical structure and serological similarity to the capsular PS of the 6B serotype. The new strain (i.e. serotype 6X1) is typed as 6B by the quellung reaction, but it can be distinguished from 6B strains with mAbs to 6B PS. Reexamination of 264 pneumococcal isolates that had been previously typed as 6B with classical typing methods revealed no isolates expressing serotype 6X1. Nevertheless, this study shows that this capsular PS is biochemically possible and could exist/emerge in nature.

Introduction

Streptococcus pneumoniae is a major human pathogen commonly responsible for pneumonia, bacteremia, meningitis and otitis media, especially among young children and older adults (Fedson, 1988). The most prominent virulence factor of the pneumococcus is the capsular polysaccharide (PS), which coats the surface of the bacterium to block antibodies and complement from binding to surface moieties and being recognized by phagocytic cells (Avery & Dubos, 1931). To avoid host immunity, *S.*

pneumoniae, as a species, can express at least 91 distinct capsules which are chemically and serologically distinct (Henrichsen, 1995; Park *et al.*, 2007b). Antibodies against this PS have been shown to provide serotype-specific protection from infection, and vaccines against pneumococcus incorporate capsular PSs of the most prevalent strains (Cole, 1913). For example, serogroup 6 strains are very common in invasive pneumococcal disease and the current vaccines are formulated to protect against serogroup 6 infections (Hausdorff *et al.*, 2000).

Serogroup 6 has classically contained two serotypes, 6A and 6B (Kamerling, 2000). These two serotypes produce capsules with very similar structures. Both of these PSs have repeating units composed of galactose–glucose–rhamnose–ribitol–phosphate, but the linkage between rhamnose and ribitol is 1→3 for 6A and 1→4 for 6B (Kamerling, 2000). The two serotypes also have very similar capsule gene loci, which are about 17 kb in size and contain all the genes for capsule biosynthesis (Jiang *et al.*, 2001; Park *et al.*, 2007a). The only genetic difference between 6A and 6B serotypes has been attributed to one nucleotide of *wciP*, which encodes a rhamnosyltransferase responsible for the rhamnose–ribitol linkage (Aanensen *et al.*, 2007; Mavroidi *et al.*, 2004). WciP with a codon for serine at residue 195 is associated with serotype 6A, but a codon for asparagine is associated with 6B (Mavroidi *et al.*, 2004).

Recently, a new serotype, 6C, was discovered using two mAbs (Lin *et al.*, 2006; Park *et al.*, 2007b). Although 6C was previously typed as 6A with classical serotyping tools, chemical analysis revealed that the galactose of the 6A PS is replaced by a glucose residue in 6C. Genetic studies of the 6A and 6C capsule gene loci have shown that *wciN* is responsible for the difference, as substitution of *wciN*_{6C} (*wciN* of 6C) for *wciN*_{6A} (*wciN*

of 6A) through homologous recombination results in a serotype switch from 6A to 6C (Park *et al.*, 2007a).

The discovery of 6C led to the logical suggestion that recombination might also produce a new member of the serogroup 6 family, herein labeled '6X1', which would have *wciP* of 6B and *wciN*_{6C}. The designation 6X1 is used to distinguish this artificial serotype from the next naturally found serogroup 6 member, which, if discovered, might be designated 6D. Chemically, 6X1 PS may have glucose instead of galactose and have a 1→4 rhamnose–ribitol linkage. However, it was not clear whether the theoretical 6X1 serotype was indeed biologically feasible, nor was it clear whether it might already exist in nature. Therefore, we have produced a serotype 6X1 strain and examined our laboratory collection of pneumococcal isolates for a possible 6X1 strain.

Methods

Bacterial strains and culture

To determine whether 6X1 exists in nature, 264 pneumococcal isolates that were previously serotyped as '6B' by classical means were reserotyped for serotypes 6B or 6X1 using mAbs. The isolates were a part of our laboratory collection of 6B isolates, which have originated from Africa, Asia, Australia, South America, North America and Europe. In addition to these, TIGR6A, TIGR6AX and TIGR6C, which are isogenic strains of TIGR4 expressing the 6A-type capsule, no capsule, and 6C-type capsule (Park *et al.*, 2007a), were used as assay controls or as a source of DNA. The source of the TIGR4 strain was the authors' laboratory (Tettelin *et al.*, 2001). Additional TIGR4 variants, TIGR6B, TIGR6BX and TIGR6X1, were prepared as described below. All

bacteria were grown in Todd–Hewitt broth (BD Biosciences) supplemented with 0.5% yeast extract (THY) and kept frozen at –80°C until used. The TIGR6X1 strain will be available upon request from the University of Alabama at Birmingham (UAB) to all qualified investigators for research purposes.

PCR and DNA sequencing

PCR mixtures contained 38.8 microliters sterile water, 2 microliters of each primer (5 pmol/microliter), 2 microliters 10 mM dNTP, 5 microliters 10x LA *Taq* buffer solution (Takara Biochemical), and 0.2 microliters LA *Taq* polymerase (2.5 U/microliter; Takara Biomedical). As template, either chromosomal DNA isolated with a Wizard genomic DNA purification kit (Promega) or colonies grown on blood agar plates were used. Thermal cycling conditions varied depending on the primer set used. PCR products were analyzed by electrophoresis in 1% agarose gels. The primers used are listed in Table 1. PCR products were purified using the Wizard PCR Clean-up System (Promega), and the DNA sequencing was performed by the Genomics Core Facility at UAB. DNA sequences were analyzed with Lasergene v5.1 software (DNASTAR) and were compared with the previously reported sequences of the 6B and 6C *cps* loci in GenBank (accession nos CR931639 and EF538714, respectively) (Mavroidi *et al.*, 2004).

Construction of TIGR6X1 by replacing *wciN*_{6A} of TIGR6B with *wciN*_{6C}

Our strategy for creating TIGR6X1 is described in Fig. 1. First, TIGR6B expressing serotype 6B was prepared by inserting the 6B capsule gene locus region into the TIGR4 genetic background using the Janus-cassette system, as previously described in detail (Park *et al.*, 2007a). Second, the *wciN* gene was removed from TIGR6B (Park *et*

Table 1. List of PCR primers

Primer	Sequence (5'-3')	Source or reference
Forward primers		
5103	AAACATGACATCAATTACA	Park <i>et al.</i> (2007a)
5106	TACCATGCAGGGTGAATGT	Park <i>et al.</i> (2007a)
5108	ATGGTGAGAGATATTTGTCAC	Park <i>et al.</i> (2007a)
5113	GGGAAAATAAAAAATAGGTCGGG	Park <i>et al.</i> (2007a)
5114	TTAGTGACGGAGGCAGGTGAA	This study
5129	TCCTACTTACAGCAACTTCTCGTG	This study
5138	AAAGCTATGTCGCCTGCTAAAAAAGCAATGGCTA	This study
Reverse primers		
3102	CTGGCATGTCATCTTTAGAAAA	This study
3104	CCTGAAAACAATACTACTT	This study
3105	TTCCCATCTCTAAACATTCTCCT	This study
3141	GGCGACATAGCTTTTCTTTCAATATCTT	This study

al., 2007a) by transforming it with cassette 1 and selecting for kanamycin-resistant isolates. Cassette 1 has the Janus cassette, which contains a kanamycin-resistance gene (*kanA^R*) and a streptomycin-sensitivity gene (*rpsL⁺*), and two flanking regions designed for homologous recombination to the 6B capsule gene locus (Park *et al.*, 2007a). A kanamycin-resistant strain was obtained and back-crossed into TIGR6B three times using genomic DNA from this resistant strain. Back-crossing was performed in order to minimize the possibility of unwanted mutations in the TIGR6B background. The resulting strain, which was labeled TIGR6BX, lost *wciN* and did not produce capsular PS. To insert *wciN_{6C}*, TIGR6BX was transformed with Cassette 2. Cassette 1 and Cassette 2 were prepared from the genomic DNAs of TIGR6AX and CHPA388, respectively, using primer set 5113 and 3102 (Park *et al.*, 2007a). While Cassette 2 contained a part of *wciP* in addition to *wciN_{6C}*, it did not contain the *wciP* codon responsible for distinguishing 6A and 6B serotypes (Fig. 1). After selection for streptomycin resistance and back-crossing against TIGR6BX three times, a streptomycin-resistant strain was produced, designated TIGR6X1. When the capsule gene locus of TIGR6X1 was sequenced from *wchA* to *wciP*, the sequence showed that *wciN_{6B}* had been replaced with the *wciN_{6C}* gene, as intended. For this sequencing, primer sets 5114–3141, 5138–3104 and 5106–3105 were used to produce amplicons, and primers 5103, 5108 and 5129 were used in sequencing (see Table 1 for primer sequences). The TIGR6X1 sequence has been deposited in GenBank (accession no. EU714777). TIGR6X1 was morphologically indistinguishable from TIGR6B when grown on blood agar plates. Also, TIGR6X1 grew as well as other pneumococcal strains in THY broth (data not shown).

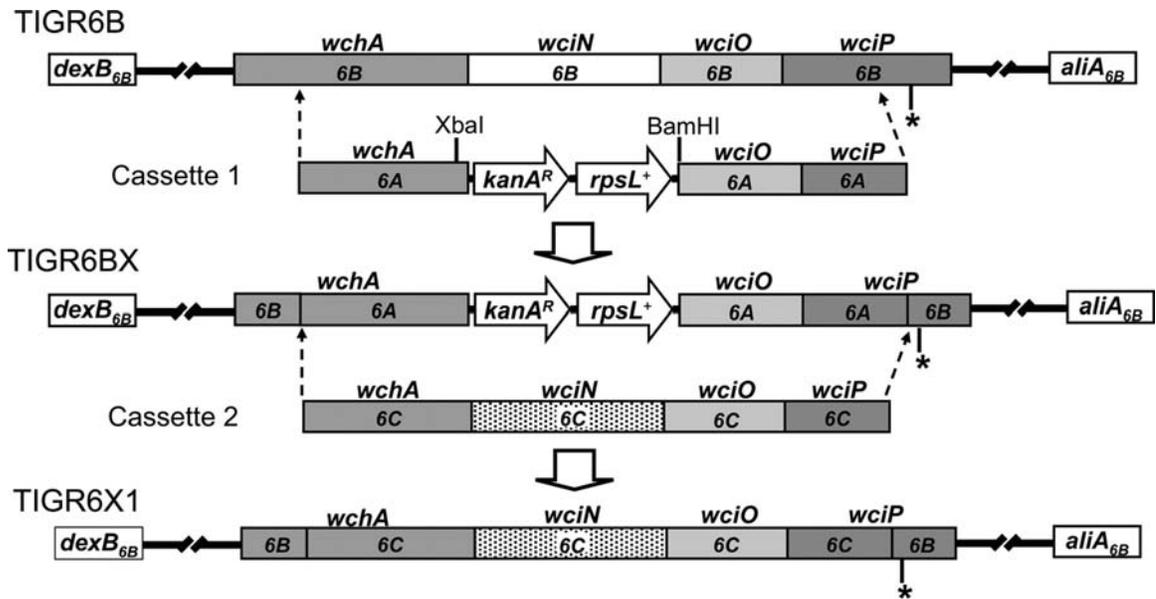


Fig. 1. Diagram of the *wciN* region exchange experiment; an asterisk at the *wciP* gene indicates the critical codon (serine versus asparagine) associated with the 6A/6B serotype difference (Mavroidi *et al.*, 2004). *kanA*^R, kanamycin-resistance gene; *rpsL*⁺, streptomycin-sensitivity gene.

Quellung reaction

Bacterial colonies from blood agar plates were suspended in a small volume of PBS (0.137 M NaCl, 2.31 mM KH₂PO₄, 7.69 mM Na₂HPO₄ and 2.15 mM KCl), and 2 microliters of this broth was combined with 2 microliters of serum and 2 microliters of methylene blue dye solution (3 mg/ml methylene blue, 1.5 mg/ml NaCl in sterile water) on a glass microscope slide. After adding a coverslip, mixtures were examined under bright-field microscopy using a 100X oil-immersion lens. The rabbit antisera specific for serotypes 6A and 6B were prepared by the Centers for Disease Control and Prevention (CDC). Factor 6a serum made by Staten Seruminstitut was purchased from Mira Vista Diagnostics.

Inhibition ELISA to distinguish serotypes 6B and 6X1

The two serotypes were distinguished using an inhibition-type ELISA. Briefly, the wells of ELISA plates (Corning Costar) were coated at 37°C with 5 microgram/ml 6B capsular PS (ATCC) overnight in PBS. After washing the plates three times with PBS containing 0.05% Tween 20, 50 microliters of a previously diluted bacterial culture supernatant (or lysate) was added to the wells along with 50 microliters of anti-6B mAb. Pneumococcal lysates were prepared by growing pneumococci overnight in 1 ml THY broth without shaking and then incubating the tubes for 15 min at 37°C with a lysis buffer (0.1% sodium deoxycholate, 0.01% SDS, 0.15 M sodium citrate in deionized water). Culture supernatants of 6B-specific hybridomas Hyp6BM7 and Hyp6BM8 were used at dilutions of 1:50 and 1:100, respectively. These hybridomas were produced from a fusion of myeloma cells with spleen cells isolated from mice immunized with 6B PS (Sun *et al.*, 2001). After 30 min incubation in a humid incubator at 37°C, the plates were washed

three times and incubated for 30 min with alkaline phosphatase-conjugated goat anti-mouse immunoglobulin (Sigma). The plates were washed three times, and then 100 microliters of paranitrophenyl phosphate substrate (Sigma) in diethanolamine buffer at a concentration of 1 mg/ml was added, and allowed to incubate at room temperature for 1–2 h. The A_{405} was read with a microplate reader (BioTek Instruments).

Purification of capsular PS

Capsular PS expressing serotype 6X1 was purified in two different ways. One method was to purify the PS by ethanol precipitation, ion-exchange chromatography and molecular mass-sizing chromatography, as described previously (Park *et al.*, 2007b). The other method, which is faster than the first method, was to purify capsular PS after removing protoplasts, as described below. TIGR6X1 was grown in 1 L THY broth without shaking until the culture reached $OD_{600} \sim 0.4$. The culture was then centrifuged at 15,000 *g* for 10 min. The cell pellet was washed twice with 1 L PBS, resuspended in 30 ml protoplast buffer [20 % sucrose, 5 mM Tris/HCl (pH 7.4), 2.5 mM Mg_2SO_4 in deionized water] with mutanolysin (Sigma) at a concentration of 20 U/ml, and allowed to incubate overnight at room temperature. The next day, the bacterial cells were examined under a phase-contrast microscope to ensure that ‘protoplasting’ had occurred, and then protoplasts were removed by centrifugation at 27,000 *g* for 15 min. The supernatant was sterilized through a 0.22 micron pore-size filter, diluted 1:1 in deionized water, and loaded onto a DEAE-Sepharose column (Amersham Biosciences) with a 2 ml bed volume. The column was washed with 4 ml 50 mM ammonium acetate, and the PS was eluted from the column with 4 ml 500 mM ammonium acetate. After lyophilization, the eluted PS was loaded on a Sephacryl S-300 HR column (Amersham Biosciences) with a

bed volume of 130 ml and the PS was eluted with 10 mM Tris/HCl (pH 7.4). The fractions were tested for the presence of PS by an inhibition assay using Hyp6BM8. The first 10 ml of fractions, which contains most of the PS, were pooled and lyophilized.

Monosaccharide composition analysis of PS

A 1 mg sample of lyophilized capsular PS prepared by the protoplast method was dissolved in 500 milliliters of 1 M methanolic HCl and incubated at 80°C for 16 h. After evaporating the methanolic HCl using a nitrogen stream, the sample was washed twice by dissolving in 250 microliters of methanol and drying under a nitrogen stream. The sample was then incubated with 200 microliters of Tri-Sil Reagent (Pierce Biotech) to trimethylsilylate all the residues. The reaction products were analyzed on a gas-liquid chromatograph/mass spectrometer (Varian 4000) fitted with a 15 m (0.25 mm diameter) VF-5 capillary column. Column temperature was maintained at 100°C for 5 min and then increased to 275°C at 20°C/min, and finally held at 275°C for 5 min. The effluent was analyzed by MS using the electron impact ionization mode. The area of each monosaccharide peak in GLC-MS was determined using Varian MS Workstation v6.5 software.

Analysis of PS by MS-MS

Intact capsular PSs prepared by the ethanol-precipitation method were hydrolyzed to their repeating units before analysis by MS. A 2 mg sample of PS was hydrolyzed in 1 ml 10 mM NaOH at 85°C for 120 hours, followed by another hydrolysis with 50 mM NaOH at 85°C for 120 hours. At the end of hydrolysis, all samples were neutralized with 0.1 M HCl.

MS-MS was performed at the Mass Spectrometry Shared Facility at the UAB with a Micromass Q-TOF2 mass spectrometer equipped with an electrospray ion source. The samples, dissolved in distilled water, were injected into the mass spectrometer with running buffer (50:50 acetonitrile: water containing 0.1% formic acid) at a rate of 1 microliter/min using a Harvard syringe pump. The injected sample was negatively ionized with electrospray and detected with a time-of-flight mass spectrometer. For MS-MS, the parent ion was fragmented into daughter ions by energizing it to either 35 or 40 eV before collision with argon gas. The daughter ions were analyzed with a time-of-flight mass spectrometer. The MS-MS spectra were processed using the Max-Ent3 module of MassLynx 2.5.

Oxidation and reduction of PS

Capsular PSs were dissolved in 80 mM sodium acetate buffer (pH 4) at a concentration of 1 mg/ml. Sodium periodate was added to the PS solution to a final concentration of 40 mM and the reaction mixture was incubated in the dark at 4°C for 72 hours. Excess periodate was destroyed by adding ethylene glycol. To determine the intact monosaccharides of the oxidized capsular PS, PS was then lyophilized and analyzed using GLC-MS as described above. To investigate the glycosidic bonds, the sample was reduced with sodium borohydride or sodium tetradeuteroborate as previously described (Park *et al.*, 2007b), before being subjected to MS-MS as described above. 6X1 PS was prepared by the protoplast method, and 6B PS was obtained from the ATCC.

Hydrolytic stability assay

A 0.9 ml volume of PS (2 mg/ml) in water was mixed with 0.1 ml 0.1 M NaOH, and this solution was split into two Eppendorf tubes and incubated at 85°C. At the

indicated times, 0.1 ml was removed from these samples, neutralized with 0.1 M HCl, and then stored at 4°C until used in the inhibition ELISA. Using the same buffers and incubation conditions described for the inhibition ELISA above, plates were coated with 100 microliters of 6A, 6B, 6C or 6X1 PS (5 microgram/ml). The ELISA was performed with the hydrolyzed samples on plates coated with their respective PSs. For 6A and 6C PSs, Hyp6AG1 was used as the primary antibody (as performed in Park *et al.*, 2007b), and for 6B and 6X1, Hyp6BM8 was used (as described above). Data shown are the average of samples run in duplicate.

Results

6X1 capsular PS is structurally different from 6B capsular PS

Recently, we have demonstrated the structure of 6C capsular PS by identifying its monosaccharide composition by GLC-MS, and identifying the sequence of monosaccharides and their glycosidic linkages by MS-MS (Park *et al.*, 2007b). We used similar approaches to show that TIGR6X1 produces a capsular PS that is chemically different from that of TIGR6B. We first determined the monosaccharide/ribitol composition of TIGR6B and TIGR6X1 capsular PS by GLC-MS. The chromatogram for 6B showed peaks for ribitol, rhamnose, galactose and glucose, as expected for 6B capsular PS (Fig. 2). However, the 6X1 chromatogram did not show galactose peaks, although it showed peaks for ribitol, rhamnose and glucose (Fig. 2). Thus, TIGR6X1 produces a capsular PS different from that of TIGR6B. We then analyzed the monosaccharide composition after treating TIGR6B and TIGR6X1 PSs with sodium periodate, which selectively destroys ribitol and monosaccharides with vicinal glycols. As

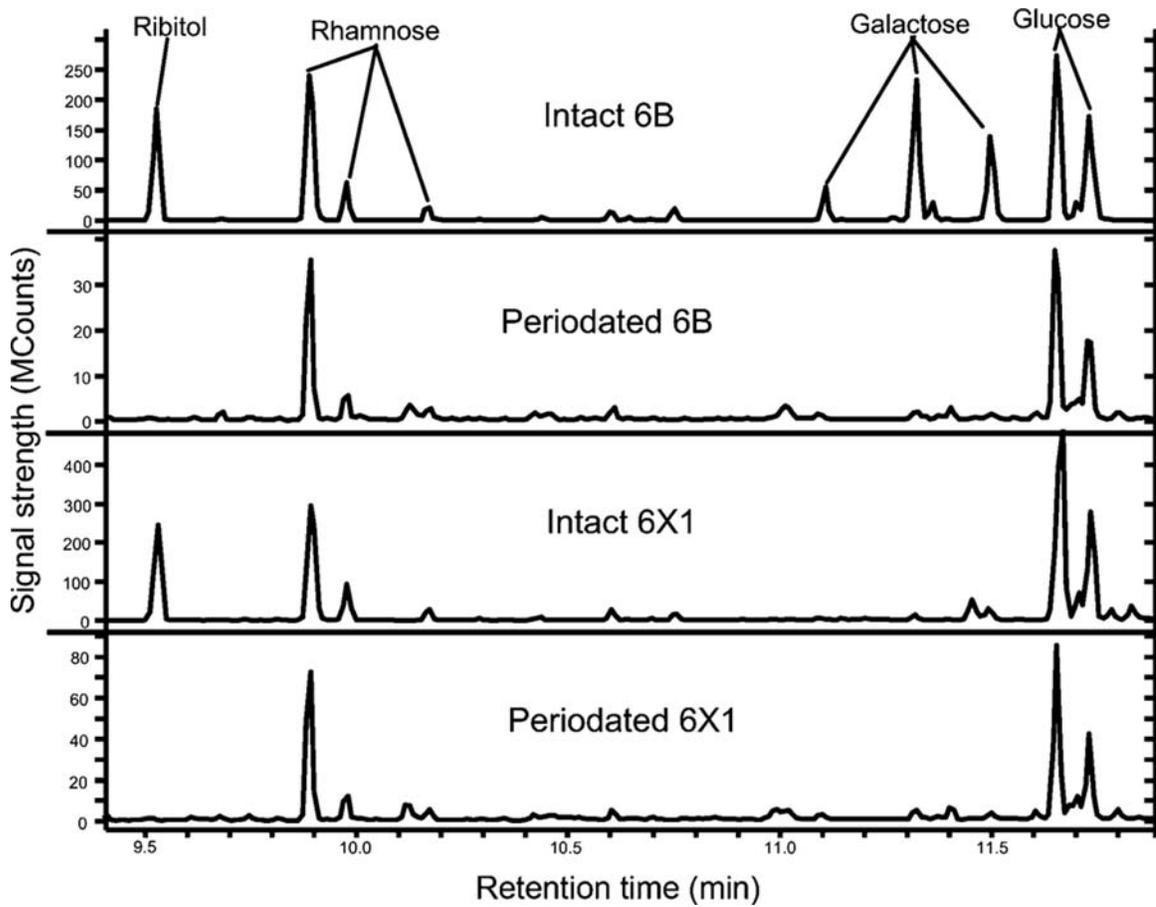


Fig. 2. GLC-MS chromatograms showing carbohydrate composition of capsular PS from pneumococcal strains expressing 6B and 6X1 serotypes and before and after periodate treatment. The monosaccharides are identified in the top chromatogram. Each panel is labeled with a description of the sample.

expected for 6B PS, periodate treatment of TIGR6B PS eliminated the peaks for ribitol and galactose (Fig. 2), but the treatment did not alter the rhamnose and glucose peaks. In contrast, periodate treatment of TIGR6X1 PS extinguished the ribitol peak and reduced the glucose peaks, such that the glucose peak area became similar to the rhamnose peak area. This suggested that the repeating unit of 6X1 PS has two glucose residues, whereas the repeating unit of 6B PS has one glucose residue and one galactose residue.

To determine whether the monosaccharide sequence of the 6X1 PS is as proposed in Fig. 3(a), we performed an alkali hydrolysis which breaks the phosphodiester bonds and produces repeating units. As previously observed for 6C PS (Park *et al.*, 2007b), the hydrolysis yields two types of repeating units of identical mass, one with the phosphate ion linked to ribitol (labeled forward fragmentation) and another linked to glucose (labeled reverse fragmentation). The phosphate ion endows the repeating unit with a negative charge. When the alkali hydrolysis product was analyzed for negative ions by MS-MS, the results showed two prominent peaks of 683 and 701 amu (Fig. 3b), which were identical to the anhydrous and hydrated masses, respectively, of the predicted repeating unit of 6X1 PS (Fig. 3a). The peak at 260.902 amu was absent in other MS-MS attempts and may represent a contaminant.

We then subjected the ion of 683 amu (i.e. the intact repeating unit) to argon collision and identified its daughter ions by MS-MS analysis. We found daughter ions at 521, 359 and 213 amu, which respectively represent daughter ions that have lost the first glucose, the second glucose and the rhamnose (Fig. 3c). Also we observed peaks at 549, 403 and 241 amu, which also correspond to the daughter ions formed after reverse fragmentation by losing ribitol, ribitol–rhamnose and ribitol–rhamnose–glucose 2,

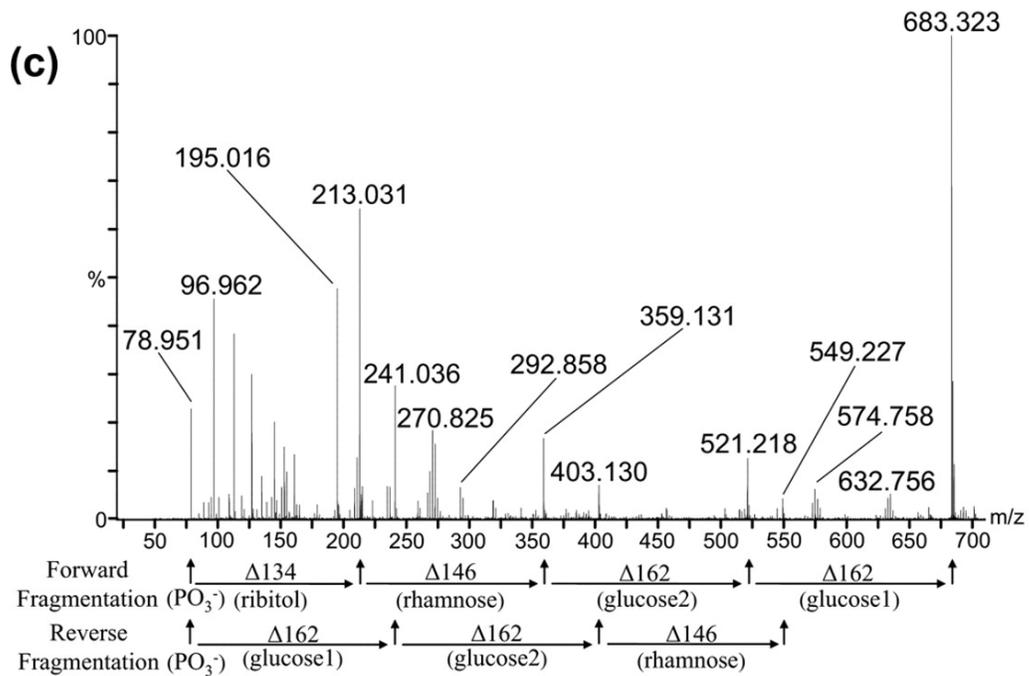
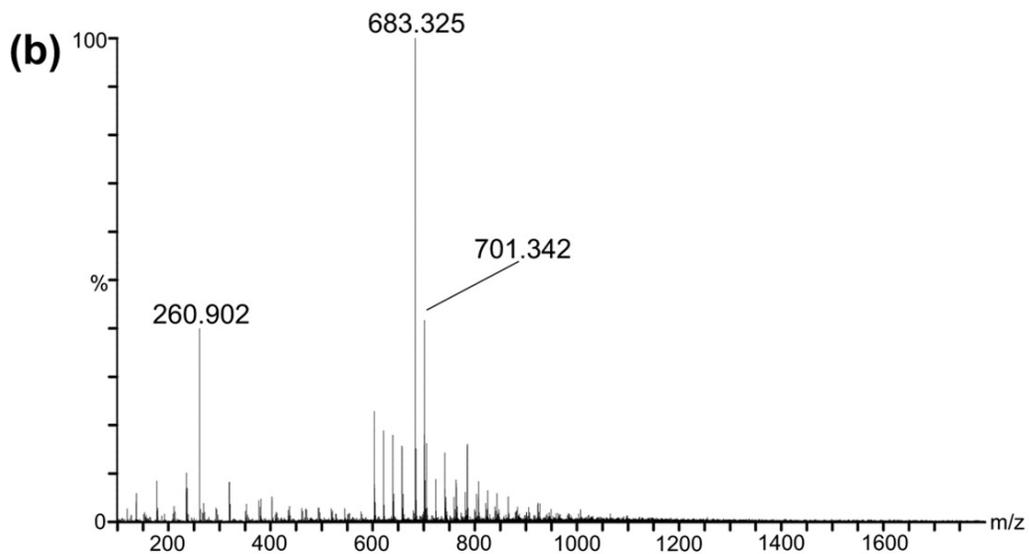
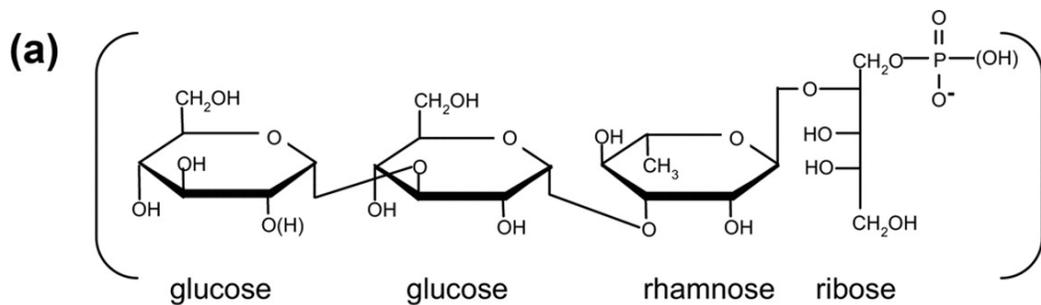


Fig. 3. (a) Proposed structure of the hydrated form of the repeating unit of 6X1 capsular PS. The calculated molecular mass is 701 amu. (b) Mass spectrum of the repeating units. The peaks at 683.3 m/z and 701.3 m/z correspond, respectively, to the anhydrous and hydrated forms of the repeating units. (c) Daughter ions of the ion of 683.3 amu shown in (b). Daughter ions are identified at the bottom. The peaks at 270.825, 574.758 and 632.756 amu and their satellite peaks (separated by 2 amu due to chloride isotopes) represent sodium chloride salt clusters (Hao *et al.*, 2001). The peaks at 270.825 represent $(\text{NaCl})_4\text{Cl}^-$. Peaks at 574.758 amu probably represent another salt cluster, $(\text{NaCl})_9\text{Cl}^-$, with a water molecule, like salt clusters with organic solvent molecules (Zhou & Hamburger, 1996). The peaks at 632.7 amu have one more NaCl (i.e. 58 amu) than the peaks at 574.758 amu.

respectively (Fig. 3c). Three peaks with 113, 127 and 145 amu were absent in other MS-MS analyses and may represent contaminants. Thus, the monosaccharide sequence of the 6X1 PS repeating unit is glucose 1–glucose 2–rhamnose–ribitol, as proposed in Fig. 3(a). The two glucose residues were labeled 1 and 2 for clarity.

To determine the linkages between the residues of 6X1 PS, we examined the periodate-treated 6X1 repeating units with MS-MS as we have done for 6C PS (Park *et al.*, 2007b). This showed that ribitol and glucose 1 are cleaved by periodate, while glucose 2 and rhamnose are not. The masses of daughter ions showed that the phosphodiester bond is made to the second position of glucose 1 and that all other glycosidic bonds are the same as in 6B PS (data not shown). The MS-MS studies supported the proposed structure shown in Fig. 3(a).

During the alkali hydrolysis experiments for MS, we observed that the 6X1 PS was very resistant to alkali hydrolysis. To measure resistance to hydrolysis, we examined the ability of 6A, 6B, 6C and 6X1 PSs to inhibit binding of Hyp6BM8 (for 6B and 6X1 PSs) or Hyp6AG1 (for 6A and 6C PSs) to target PS after alkali hydrolysis for various time periods. 6A and 6C PSs completely lost their ability to inhibit after only 1 hour of hydrolysis. However, 6B PS lost 90% of its inhibitory ability in 8 hours and more than 100 hours of hydrolysis was needed for 6X1 PS to lose 90% of its inhibitory ability (Fig. 4). Thus, 6X1 PS is much more resistant than 6A and 6C PS to alkali hydrolysis, and may be more resistant than 6B PS.

6X1 is serologically similar to, but distinct from, 6B

When the serological properties of TIGR6X1 were examined by the quellung reaction using polyclonal rabbit antisera, it was found to react with factor serum 6a and

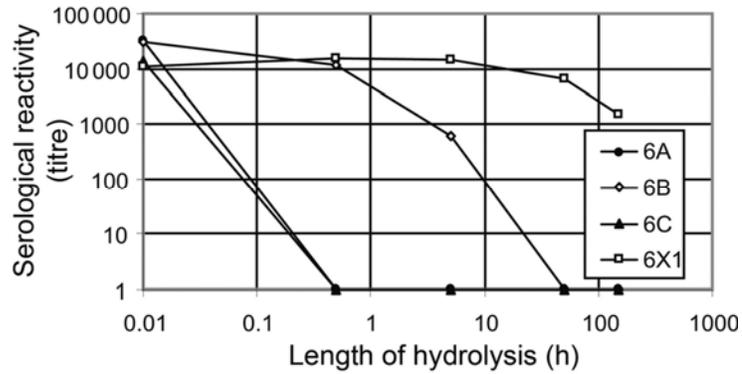


Fig. 4. Ability of various capsular PSs (2 mg/ml) to inhibit binding of mAb to ELISA plates (y axis) after the PSs were hydrolyzed for various time periods (x axis); ‘titer’ indicates the dilution of a sample necessary to inhibit binding by 50%. For 6A and 6C PSs, ELISA plates were coated with 6A PS and mAb Hyp6AG1 was used. For 6B and 6X1 PSs, ELISA plates were coated with 6B PS and mAb Hyp6BM8 was used.

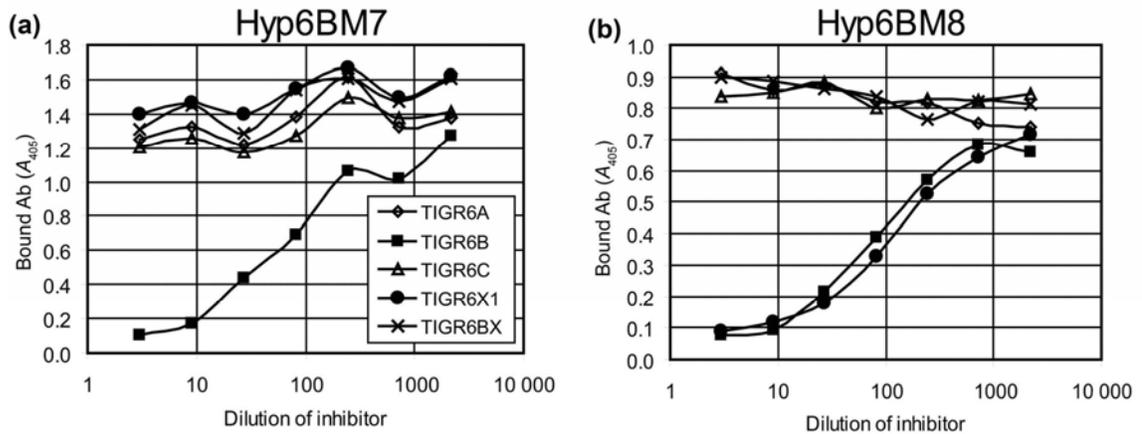


Fig. 5. Ability of lysates of various pneumococcal strains to inhibit the binding of mAbs Hyp6BM7 (a) and Hyp6BM8 (b) to 6B PS-coated ELISA plates. The amount of mAb bound (y axis) is shown in the presence of varying amounts of inhibitors (x axis). Lysates of pneumococcal strains were used as inhibitors. The pneumococcal strains are TIGR6A (◊), TIGR6B (■), TIGR6C (△), TIGR6X1 (●) and TIGR6BX (x). TIGR6BX does not produce capsular PS.

was typed as 6B. When we examined TIGR6X1 PS for binding various mAbs to 6A and 6B PS using an inhibition ELISA, we found it to be reactive with many mAbs to 6B PS. For instance, TIGR6X1 inhibited Hyp6BM8 binding to 6B PS. These observations clearly demonstrated that 6X1 PS is serologically very close to 6B PS. However, we also found a mAb specific to 6B PS (Hyp6BM7) but not reactive with 6X1 PS (Fig. 5). Thus, 6X1 PS is serologically distinct from 6B PS.

Pneumococcal isolates expressing serotype 6X1 PS were not found in nature

The serological studies described above showed that if pneumococcal isolates expressing 6X1 PS are present in nature, they would have been typed as serotype 6B. To look for the presence of serotype 6X1 isolates in nature, we examined 264 pneumococcal isolates that had been previously typed as serotype 6B using an inhibition ELISA capable of distinguishing between the 6B and 6X1 serotypes (Fig. 5). These 6B isolates came from six continents [North America (109), South America (94), Europe (24), Asia (17), Africa (12) and Australia (8)], and were isolated from patients with bacteremia (38), meningitis (19), pneumonia (40) and otitis media (17), as well as from healthy carriers (44) (106 isolates were from patients for whom the diagnosis is unknown). None of the 264 6B strains exhibited the antibody-binding profile of 6X1. Thus, the prevalence of serotype 6X1, if it exists, is much less than that of 6B.

Discussion

The actual synthesis of the capsular PS requires cooperation among many different gene products. For instance, a new repeating unit made by a new glycosyltransferase must be compatible with the existing flippase as well as the

polymerase before it can be expressed as a new capsule. Thus, to show that serotype 6X1 is possible, we produced a ‘TIGR6X1’ strain by inserting *wciN*_{6C} into a 6B capsule gene locus, and demonstrated that the new strain produces a capsular PS with the predicted structure, displays serological similarity to 6B, and can grow as well as other members of serogroup 6 in various growth conditions. Thus, serotype 6X1 could exist in nature.

Just as 6C was previously typed as ‘6A’ by the classical typing method (Lin *et al.*, 2006; Park *et al.*, 2007b), the quellung reaction method typed the new 6X1 strain as serotype 6B. Thus, to identify natural isolates expressing serotype 6X1, we re-examined the isolates that had been classically defined as 6B using mAbs. Despite our testing more than 250 such isolates, we did not find 6X1 isolates in nature. Additionally, *wciN*_{6C} was not detected among the CDC isolates that were serotyped as ‘6B’ with the classical typing method (Dr B. Beall, CDC, personal communication). Thus, pneumococcal isolates expressing serotype 6X1 may not exist in nature. If serotype 6X1 exists in nature, its prevalence is extremely low.

Despite the fact that it is not detected in nature, the 6X1 serotype could emerge in nature by one of two mechanisms. One mechanism involves a mutation of the *wciP* gene of 6C, since the only difference between 6A and 6B serotypes appears to be one nucleotide in the *wciP* gene. The mutation rate for pneumococci is $\sim 1 \times 10^{-8}$ (del Campo *et al.*, 2005; Gould *et al.*, 2007; Morosini *et al.*, 2003) and a chronic obstructive pulmonary disease (COPD) patient with stable pneumonia may have 2.6×10^8 c.f.u. of pneumococci per milliliter of sputum (Sethi *et al.*, 2007). Thus, the correct mutation should arise in almost all cases of 6C pneumonia and often in other 6C infections with lower bacterial load. The alternative mechanism involves the lateral gene transfer of *wciN* from a 6C

strain into a 6B strain as we have done here. This situation could actually occur in nature, since carriage of multiple pneumococcal serotypes can be relatively high among children (Gratten *et al.*, 1986; Hill *et al.*, 2008) and serotypes 6B and 6C are fairly common in some parts of the world (e.g. Brazil) (Lin *et al.*, 2006; Park *et al.*, 2007a). Furthermore, homologous recombination would occur easily, since the 17 kb capsule gene loci of 6B and 6C are almost identical except for the *wciN* gene. These considerations strongly suggest that the circumstances for creating serotype 6X1 exist in nature.

Given that the circumstances for creating 6X1 do exist in nature, it is interesting to consider reasons for its absence. It is possible that 6C could have appeared so recently that there may not have been enough time for 6X1 to appear. Alternatively, there could be natural immune barriers against 6X1, but we found that pre-immune human sera do not kill or opsonize TIGR6X1 (data not shown). The most likely explanation is that there has not been enough biological pressure to select 6X1 over 6A, 6B or 6C. The need for selection pressure was recently demonstrated by an increase in the prevalence of serotypes 19A and 6C following the use of the conjugate vaccine (Nahm *et al.*, 2009; Park *et al.*, 2008). In the absence of a survival advantage, the 6X1 serotype may have appeared in nature (as it should in almost every case of a 6C infection, as mentioned above) but has not been propagated as a consequence of competition with more abundant 6A, 6B or 6C serotypes. In an analogous manner, antibiotic-resistant strains survive and propagate when antibiotics are used clinically, and disappear when antibiotics are discontinued (Katsunuma *et al.*, 2007).

6A PS was included in the original 14-valent PS vaccine, but it was replaced with 6B PS when the 23-valent vaccine became available in 1983, because 6B PS can elicit

antibodies that cross-react with 6A PS and is much more resistant to hydrolytic breakdown than 6A PS (Zon *et al.*, 1982). In this study, we also discovered that 6X1 PS is as chemically stable as 6B PS and much more resistant to hydrolysis than 6C PS. Since 6X1 PS would likely elicit antibodies cross-reactive with 6C PS, 6X1 may be more useful as a vaccine component than 6C PS.

As our knowledge of pneumococcal capsule genetics, biochemistry and serology advances, we may wish to produce pneumococcal strains that express artificially created capsular PS for various reasons. We propose that these artificially created strains should be named in a systematic manner to avoid their names being confused with those of natural strains. We chose to use 'X' to denote such experimental strains. We chose '6X1' as the name for our strain because this experimental strain belongs to serogroup 6 and this is the first experimental strain within serogroup 6. This systematic approach should be applicable to any new strains created experimentally.

It is important to understand the evolution of the capsule gene locus, which encodes the most important virulence factor of pneumococci. Even when only two serotypes were known, the evolution of serogroup 6 was extensively studied (Mavroidi *et al.*, 2004; Robinson *et al.*, 2002). Serogroup 6 has become more interesting following the discovery of serotype 6C. Now, the serogroup would be even more interesting for evolution studies with a potential serotype, serotype 6X1.

Acknowledgments

The work was funded by a grant AI-031473 from the NIH and P. E. B. is supported by a training grant (T32 GM008111) from the NIH. UAB has applied for a

patent covering the creation of 6X1. We thank Dr B. Beall and Mr R. Gertz at the CDC for assistance with the quellung reaction tests. We also thank Rob Cartee for performing GLC-MS and Marion Kirk for performing MS-MS analysis.

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**, 7856–7876.

Avery, O. T. & Dubos, R. (1931). The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med* **54**, 73–89.

Cole, R. (1913). Treatment of pneumonia by means of specific serums. *JAMA* **61**, 663–666.

del Campo, R., Morosini, M. I., de la Pedrosa, E. G., Fenoll, A., Munoz-Almagro, C.,

Maiz, L., Baquero, F. & Canton, R. (2005). Population structure, antimicrobial resistance, and mutation frequencies of *Streptococcus pneumoniae* isolates from cystic fibrosis patients. *J Clin Microbiol* **43**, 2207–2214.

Fedson, D. S. (1988). Pneumococcal vaccine. In *Vaccines*, pp. 271–299. Edited by S. A. Plotkin & E. A. Mortimer, Jr. Philadelphia, PA: W. B. Saunders.

Gould, C. V., Sniegowski, P. D., Shchepetov, M., Metlay, J. P. & Weiser, J. N.

(2007). Identifying mutator phenotypes among fluoroquinolone-resistant strains of *Streptococcus pneumoniae* using fluctuation analysis. *Antimicrob Agents Chemother* **51**, 3225–3229.

Gratten, M., Gratten, H., Poli, A., Carrad, E., Raymer, M. & Koki, G. (1986). Colonisation of *Haemophilus influenzae* and *Streptococcus pneumoniae* in the upper respiratory tract of neonates in Papua New Guinea: primary acquisition, duration of carriage, and relationship to carriage in mothers. *Biol Neonate* **50**, 114–120.

Hao, C., March, R. E., Croley, T. R., Smith, J. C. & Rafferty, S. P. (2001). Electrospray ionization tandem mass spectrometric study of salt cluster ions. Part 1 – investigations of alkali metal chloride and sodium salt cluster ions. *J Mass Spectrom* **36**, 79–96.

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**, 100–121.

Henrichsen, J. (1995). Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* **33**, 2759–2762.

Hill, P. C., Cheung, Y. B., Akisanya, A., Sankareh, K., Lahai, G., Greenwood, B. M. & Adegbola, R. A. (2008). Nasopharyngeal carriage of *Streptococcus pneumoniae* in Gambian infants: a longitudinal study. *Clin Infect Dis* **46**, 807–814.

Jiang, S. M., Wang, L. & Reeves, P. R. (2001). Molecular characterization of *Streptococcus pneumoniae* type 4, 6B, 8, and 18C capsular polysaccharide gene clusters. *Infect Immun* **69**, 1244–1255.

Kamerling, J. P. (2000). Pneumococcal polysaccharides: a chemical view. In *Streptococcus Pneumoniae Molecular Biology and Mechanisms of Disease*, pp. 81–114. Edited by A. Tomasz. Larchmont, NY: Mary Ann Liebert.

Katsunuma, Y., Hanazumi, M., Fujisaki, H., Minato, H., Hashimoto, Y. & Yonemochi, C. (2007). Influence of avilamycin administration and its subsequent withdrawal on emergence and disappearance of antimicrobial resistance in enterococci in the intestine of broiler chickens. *J Appl Microbiol* **102**, 1159–1166.

Lin, J., Kalsoft, M. S., Brandao, A. P., Echaniz-Aviles, G., Brandileone, M. C., Hollingshead, S. K., Benjamin, W. H. & Nahm, M. H. (2006). Validation of a multiplex pneumococcal serotyping assay with clinical samples. *J Clin Microbiol* **44**, 383–388.

Mavroidi, A., Godoy, D., Aanensen, D. M., Robinson, D. A., Hollingshead, S. K. & Spratt, B. G. (2004). Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J Bacteriol* **186**, 8181–8192.

Morosini, M. I., Baquero, M. R., Sanchez-Romero, J. M., Negri, M. C., Galan, J. C., del Campo, R., Perez-Diaz, J. C. & Baquero, F. (2003). Frequency of mutation to rifampin resistance in *Streptococcus pneumoniae* clinical strains: *hexA* and *hexB* polymorphisms do not account for hypermutation. *Antimicrob Agents Chemother* **47**, 1464–1467.

Nahm, M. H., Lin, J., Finkelstein, J. A. & Pelton, S. I. (2009). Expansion of the newly discovered pneumococcal serotype 6C in the nasopharynx following introduction of pneumococcal conjugate vaccine. *J Infect Dis* in press

Park, I. H., Park, S., Hollingshead, S. K. & Nahm, M. H. (2007a). Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* **75**, 4482–4489.

Park, I. H., Pritchard, D. G., Cartee, R., Brandao, A., Brandileone, M. C. & Nahm, M. H. (2007b). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**, 1225–1233.

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. & other authors (2008). Impact of conjugate vaccine on transmission of antimicrobial-resistant *Streptococcus pneumoniae* among Alaskan children. *Pediatr Infect Dis J* **27**, 335–340.

Robinson, D. A., Briles, D. E., Crain, M. J. & Hollingshead, S. K. (2002). Evolution and virulence of serogroup 6 pneumococci on a global scale. *J Bacteriol* **184**, 6367–6375.

Sethi, S., Sethi, R., Eschberger, K., Lobbins, P., Cai, X., Grant, B. J. & Murphy, T. F. (2007). Airway bacterial concentrations and exacerbations of chronic obstructive pulmonary disease. *Am J Respir Crit Care Med* **176**, 356–361.

Sun, Y., Hwang, Y. & Nahm, M. H. (2001). Avidity, potency, and cross-reactivity of monoclonal antibodies to pneumococcal capsular polysaccharide serotype 6B. *Infect Immun* **69**, 336–344.

Tettelin, H., Nelson, K. E., Paulsen, I. T., Eisen, J. A., Read, T. D., Peterson, S., Heidelberg, J., DeBoy, R. T., Haft, D. H. & other authors (2001). Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. *Science* **293**, 498–506.

Zhou, S. & Hamburger, M. (1996). Formation of sodium cluster ions in electrospray mass spectrometry. *Rapid Commun Mass Spectrom* **10**, 797–800.

Zon, G., Szu, S. C., Egan, W., Robbins, J. D. & Robbins, J. B. (1982). Hydrolytic stability of pneumococcal group 6 (type 6A and 6B) capsular polysaccharides. *Infect Immun* **37**, 89–103.

IDENTIFICATION OF NATURAL PNEUMOCOCCAL ISOLATES EXPRESSING
SEROTYPE 6D BY GENETIC, BIOCHEMICAL, AND SEROLOGICAL
CHARACTERIZATION

by

PRESTON E. BRATCHER, KYUNG-HO KIM, JIN H. KANG, JUNG Y. HONG,
MOON H. NAHM

Microbiology. 2010 Feb; 156(Pt 2): 555-60.

Copyright

2010

by

Society for General Microbiology

Used by permission

Format adapted for dissertation

Abstract

The recently discovered pneumococcal serotype 6C was created when the original *wciN* gene in the 6A capsule gene locus was naturally replaced with a new gene. Since the capsule gene loci of 6A and 6B serotypes may differ by only one base pair in the *wciP* gene, it was speculated that a new serotype ‘6D’ would be possible if the new *wciN* gene were inserted into the 6B capsule gene locus. Although pneumococci expressing serotype 6D could be produced in the laboratory, initial searches for natural pneumococcal isolates expressing serotype 6D were unsuccessful. However, we now report the discovery of two naturally occurring pneumococcal isolates from Korea that have the serological, genetic and biochemical features predicted for serotype 6D.

Introduction

Streptococcus pneumoniae (the pneumococcus) is a significant human pathogen (Fedson, 1988), and expresses on its surface serotype-specific capsular polysaccharides (PSs), which greatly increase its virulence (Avery & Dubos, 1931). However, antibodies to the capsule can abrogate the virulence and provide serotype-specific protection (Cole, 1913). *S. pneumoniae* as a species can produce more than 90 different capsule types (Park *et al.*, 2007b), but not all capsule types are equally prevalent in human diseases. For instance, prior to the use of conjugate vaccines, several serotypes such as 6A, 6B, 14, 19A, 19F and 23F were responsible for a large number of invasive pneumococcal diseases (Robinson *et al.*, 2001) and consequently have been extensively studied.

Serotypes 6A and 6B have been grouped into serogroup 6 because of their serological similarities. Also, serotypes 6A and 6B consistently differ in their 17 kb

capsule gene loci by only one nucleotide in the *wciP* gene (Mavroidi *et al.*, 2004). This difference produces 6A WciP with a serine at residue 195 but 6B WciP with an asparagine at the same residue (Mavroidi *et al.*, 2004). Serotypes 6A and 6B produce capsular PSs with almost identical structures except for a different rhamnose–ribitol linkage (Kamerling, 2000). Recently, we discovered a new member of serogroup 6 which we named 6C (Park *et al.*, 2007b). Serotype 6C is distinguished serologically from serotypes 6A and 6B by its unique monoclonal antibody (mAb) binding profile (Lin *et al.*, 2006). Subsequent genetic studies have shown that the capsule gene loci of serotypes 6A and 6B have *wciN_{6A}* (named alternatively as *wciN_α* here), but that the 6C capsule gene locus has *wciN_{6C}* (or *wciN_β*) (Park *et al.*, 2007a). Finally, biochemical studies revealed that 6C PS has a glucose residue in place of the galactose residue present in the 6A or 6B PSs (Park *et al.*, 2007b).

Following the discovery of serotype 6C, we postulated the presence of a new serotype, which could be created by mutating the critical nucleotide in the *wciP* gene of the 6C capsule gene locus or by inserting the *wciN_β* gene into the 6B capsule gene locus (Bratcher *et al.*, 2009). This new serotype, which logically could have been named 6D, was provisionally named 6X1 because the new serotype had not been found in nature (Bratcher *et al.*, 2009). Nevertheless, our group continued the search and discovered natural isolates expressing serotype 6D in Korea; we describe our findings below.

Methods

Bacterial strains and culture

TIGR6A, TIGR6B, TIGR6C and TIGR6X1, which are isogenic strains of TIGR4 expressing the four different capsule serotypes of serogroup 6, were used as controls for the assays (Bratcher *et al.*, 2009). Fourteen serogroup 6 isolates, including MNZ21 and MNZ22, were obtained from the nasopharyngeal cultures of 14 healthy children (less than 5 years old) attending a day-care centre in Jeju Island in Korea in 2008 and were subsequently transferred to the Nahm laboratory for further analysis. These isolates showed the typical α -hemolytic colony morphology of *S. pneumoniae* on blood agar plates, were susceptible to optochin, and were bile soluble. All bacteria were grown in Todd–Hewitt broth (BD Biosciences) supplemented with 0.5% yeast extract (THY) and kept frozen at -80°C until used.

Flow cytometry

Aliquots of frozen bacteria were thawed, washed, resuspended in FACS buffer (PBS containing 3% FBS and 0.1% sodium azide) and incubated with culture supernatants of hybridomas (diluted 1:10 in FACS buffer) for 20 min at room temperature with shaking. After washing, the bacteria were incubated with fluorescein-conjugated goat antibody to mouse immunoglobulin for 20 min at room temperature with shaking. After washing away unbound goat antibody, the bacteria were resuspended in FACS buffer containing Syto9 (160 nM) and examined with a flow cytometer (FACSCalibur, Becton Dickinson). The data were then analyzed with the Cell Quest program. Isotype-matched negative controls were used to identify negative staining and their fluorescence signals were less than 20 units (data not shown).

PCR and DNA sequencing

PCR mixtures contained 38.8 microliters of sterile water, 2 microliters of each 5 pmol/microliter primer, 2 microliters of 10 mM dNTPs, 5 microliters of 10X buffer solution and 0.2 microliters of LA Taq polymerase (2.5 U/microliter, Takara Biomedical). For the template, either chromosomal DNA isolated with the Wizard genomic DNA purification kit (Promega) or colonies grown on blood agar plates were used. Thermal cycling conditions were previously described (Park *et al.*, 2007a). PCR products were analyzed by electrophoresis in 1% agarose gel. The primers used here were the forward primer for *wciN* (5106), TACCATGCAGGGTGGGAATGT; the reverse primer for *wciN* (3101), CCATCCTTCGAGTATTGC; the forward primer for *wciP* (5108), ATGGTGAGAGATATTTGTAC; and the reverse primer for *wciP* (3107), AGCATGATGGTATATAAGCC. PCR products were purified using the Wizard PCR Clean-up System (Promega), and the DNA sequencing was performed by the genomics core facility at the University of Alabama at Birmingham. DNA sequences were analyzed with Lasergene v. 5.1 software (DNASTAR).

Inhibition ELISA

Capsular PSs were distinguished using an inhibition-type ELISA. Briefly, the wells of ELISA plates (Corning Costar Corp.) were coated at 37°C with 5 microgram/ml of 6B capsular PS (ATCC, Manassas, VA, USA) overnight in PBS. After washing the plates three times with PBS containing 0.05% Tween 20, 50 microliters of a previously diluted bacterial culture supernatant (or lysate) was added to the wells along with 50 microliters of an anti-6B mAb. Pneumococcal lysates were prepared by growing pneumococci overnight in 1 ml THY broth without shaking and then incubating the tubes for 15 min at 37°C with 110 microliters of a lysis buffer (0.1% sodium deoxycholate,

0.01% SDS and 0.15 M sodium citrate in deionized water). Culture supernatants of 6B-specific hybridomas Hyp6BM7 and Hyp6BM8 were used at dilutions of 1:50 and 1:100, respectively. These hybridomas were produced from the fusion of myeloma cells with spleen cells isolated from mice immunized with 6B PS (Sun *et al.*, 2001). After 30 min incubation in a humid incubator at 37°C, the plates were washed three times and incubated for 30 min with alkaline phosphatase-conjugated goat anti-mouse immunoglobulin (Sigma). The plates were washed three times. Then 100 microliters of *p*-nitrophenyl phosphate substrate (Sigma) in diethanolamine buffer at a concentration of 1 mg/ml was added, and the plates were allowed to incubate at room temperature for 1–2 hours so that an optimal absorbance (maximal A_{405} of 1.5) was reached. The A_{405} was read with a microplate reader (BioTek Instruments).

Hydrolytic stability assay

A 0.9 ml sample of 2 mg/ml PS in water was mixed with 0.1 ml 0.1 M NaOH. This solution was then divided between two Eppendorf tubes and incubated at 85°C. Capsular PSs from serotypes 6A, 6B, 6C and 6X1 were described before (Bratcher *et al.*, 2009). Capsular PS from a 2 l culture of MNZ21 was purified using ethanol precipitation after lysis and removal of cell debris as previously described (Park *et al.*, 2007b). The precipitate was then subjected to molecular mass sizing chromatography (Sephacryl S-500HR, 500 ml column volume) and fractions positive for PS were pooled and lyophilized. At the indicated times, 0.1 ml was removed from these samples, neutralized with 0.1 M HCl and then stored at 4°C until use in the inhibition ELISA. Using the same buffers and incubation conditions as described for the inhibition ELISA above, plates were coated with 100 microliters of 5 microgram/ml of 6A, 6B, 6C, 6X1 or MNZ21 PS

(Bratcher *et al.*, 2009). The ELISA was performed with the hydrolyzed samples on plates coated with their respective PSs. For 6A and 6C PSs, Hyp6AG1 was used as the primary antibody (as performed by Park *et al.*, 2007b); for 6B and MNZ21 PSs, Hyp6BM8 was used (as described above). Data shown are the mean of samples run in duplicate.

Monosaccharide composition analysis of PS

Capsular PS from MNZ21 was purified as described above. 6B PS from ATCC was used as a control. A 15 microgram sample of lyophilized capsular PS was dissolved in 400 microliters of 1.5 M methanolic HCl and incubated at 80°C for 16 hours. After evaporating the methanolic HCl in a vacuum centrifuge, the sample was redissolved in 250 microliters of methanol, transferred to a glass GC vial insert and again dried in a vacuum centrifuge. The sample was then trimethylsilylated with 50 microliters of Tri-Sil Reagent (Pierce Biotech). The reaction products were analyzed as described by Baker *et al.* (2006) using a gas-liquid chromatograph (HP5890, Hewlett Packard) fitted with a 30 m HP-1 wide-bore fused-silica column coated with a 0.88 micron layer of cross-linked methylsilicone gum. The column temperature was maintained at 100°C for 5 min and then increased to 275°C at a rate of 20°C/min. Finally, it was held at 275°C for 5 min.

Results and Discussion

Serological findings

Studies with TIGR6X1 showed that it binds to a mAb (Hyp6BM8), but that it does not react with other mAbs reacting with serotypes 6A, 6B and 6C (Bratcher *et al.*, 2009). We used flow cytometry to test the two isolates (MNZ21 and MNZ22) and four laboratory strains expressing serotypes 6A, 6B, 6C and 6X1 for their ability to bind to this

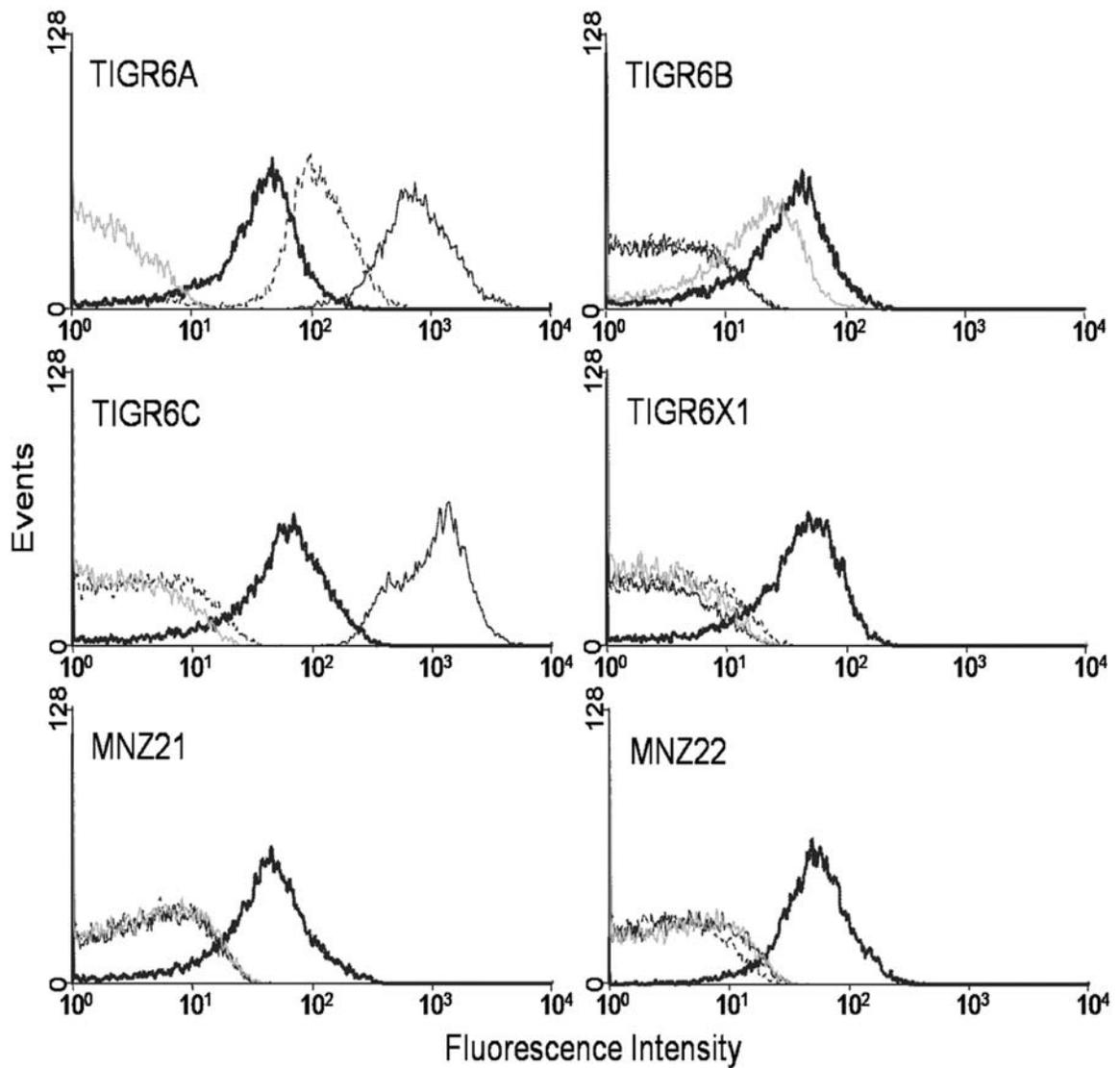


Fig. 1. Flow cytometry profiles of various pneumococcal isolates stained with mAbs. The mAbs used were Hyp6AG1 (thin black line), Hyp6AM3 (dashed line), Hyp6BM8 (thick black line), and Hyp6BM7 (thin grey line). Studies with controls showed the staining threshold to be about 20 fluorescence units (data not shown).

panel of mAbs. As shown in Fig. 1, we were able to directly demonstrate that MNZ21, MNZ22 and TIGR6X1 bound to Hyp6BM8, but did not bind to the three other mAbs. In contrast, the three control strains bound Hyp6BM8 (thick black line) as well as at least one of the three other mAbs [Hyp6AM3 (dashed line) stained TIGR6A, Hyp6AG1 (thin black line) stained TIGR6A and TIGR6C, and Hyp6BM7 (thin grey line) stained TIGR6B] in the expected manner (Fig. 1). Although Hyp6BM8 bound to all the strains shown in Fig. 1, its binding is selective to serogroup 6 as it did not stain a 9V strain (data not shown). These mAb binding patterns clearly show that the two isolates were serologically comparable to TIGR6X1 but distinct from serotypes 6A, 6B and 6C.

To further confirm that the two clinical isolates behave serologically as one would expect TIGR6X1 to behave, we examined these strains using an inhibition-type immunoassay which tests the ability of the PS to inhibit the binding of a mAb to 6B PS. Although Hyp6BM8 can bind to all members of serogroup 6 PS (Fig. 1), this mAb preferentially binds to 6B PS quite strongly and it was previously shown that TIGR6A and TIGR6C lysates did not inhibit its binding to 6B PS (Bratcher *et al.*, 2009), although they are able to inhibit the binding of Hyp6BM8 to 6A PS (data not shown). We confirmed here again that TIGR6A and TIGR6C lysates did not inhibit the binding of Hyp6BM8 to 6B PS (Fig. 2) even though the lysates of TIGR6A and TIGR6C contained 6A and 6C capsular PS respectively (demonstrated by 6A and 6C PS assays; data not shown). In contrast, binding of Hyp6BM8 to 6B PS could be inhibited by TIGR6B lysate as well as by the lysates of the three strains TIGR6X1, MNZ21 and MNZ22 (Fig. 2a). This indicates that MNZ21 and MNZ22 do not belong to serotypes 6A and 6C. The two clinical isolates could be distinguished from serotype 6B with Hyp6BM7, which

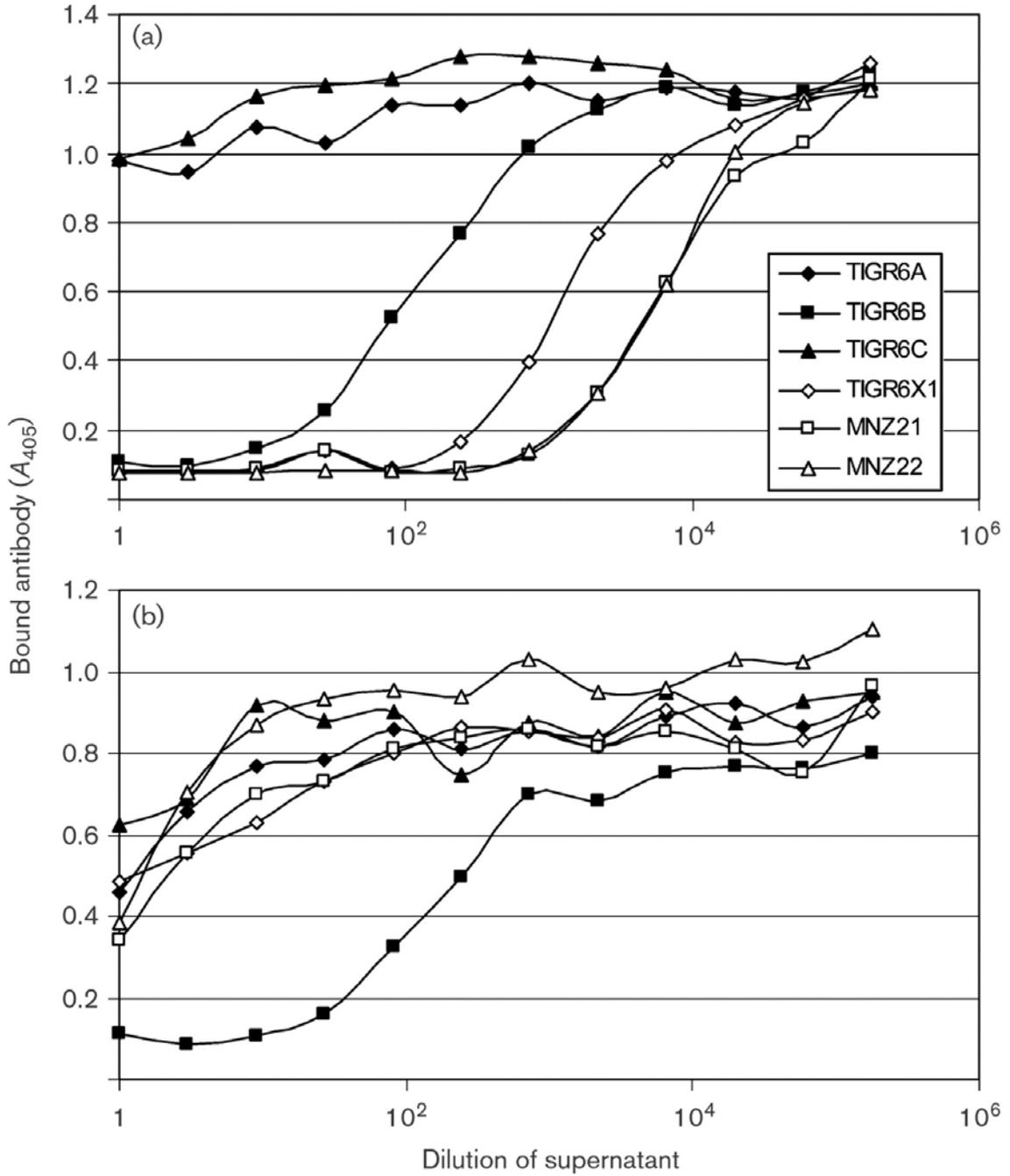


Fig. 2. Binding of mAbs (y -axis) to 6B PS-coated ELISA plates in the presence of varying amounts of bacterial lysates (x -axis). The mAbs used were Hyp6BM8 (a) and Hyp6BM7 (b). The bacterial isolates studied are listed in the figure key.

specifically reacts with 6B PS only. Its ability to bind to 6B PS could be inhibited with TIGR6B lysate, but not with the lysates of the three strains TIGR6X1, MNZ21 and MNZ22 (Fig. 2b). In addition, TIGR6X1, MNZ21 and MNZ22 PSs failed to inhibit the binding of Hyp6AG1 and Hyp6AM3 to 6A PS-coated plates (data not shown). These serological studies clearly show that MNZ21 and MNZ22 behave like TIGR6X1.

Genetic findings

TIGR6X1 was created genetically by inserting the *wciN_β* gene into a 6B capsule gene locus (Bratcher *et al.*, 2009). Thus, *wciN_β* and *wciP_{6B}* are the two genetic features of 6X1. To investigate the presence of these two genetic features in isolates MNZ21 and MNZ22, we PCR amplified the *wciN* and *wciP* regions as we described before (Bratcher *et al.*, 2009) and determined the nucleotide sequences of the amplicons. The sequences have been deposited in GenBank (accession no. GQ848645 for MNZ21 and GQ848646 for MNZ22). The sequences from the two isolates were identical, suggesting that they share a single origin. Their *wciN* sequences are strikingly different from the *wciN_α* sequence (GenBank accession no. CR931638), but are highly homologous [99.6 % identity (1650/1656 bp) in the region sequenced using the PCR product from primers 5106 and 3101] to the *wciN_β* sequence of strain CHPA388 (Park *et al.*, 2007a) in GenBank (accession no. EF538714). A characteristic of 6B *wciP* is the presence of an A at position 584 (based on the sequence of *wciP* in a previous publication: Mavroidi *et al.*, 2004), which creates a codon for asparagine at residue 195 of the 6B WciP protein. In contrast, *wciP* from 6A or 6C has G at the corresponding position, creating a codon encoding serine (Mavroidi *et al.*, 2004). The two new isolates had A at the corresponding

position (GenBank accession no. GQ848645 and GQ848646). Thus, isolates MNZ21 and MNZ22 have the two genetic features of the 6X1 serotype.

In addition, we performed multi-locus sequence typing (MLST) analysis with MNZ21 and MNZ22 using the established method for *S. pneumoniae* (Enright & Spratt, 1998). The experiment showed that the two strains have identical sequences for the 3199 bp involved in MLST and both belong to ST282. As these isolates were collected from the same day-care centre in Jeju island in Korea, we suspect that these isolates are two members of the same clone.

Biochemical findings

Biochemical features distinguishing the PSs of the four members in the serogroup 6 are (a) rhamnose–ribitol linkage and (b) the presence of glucose instead of galactose. The 6A and 6B PSs contain galactose while 6X1 and 6C PSs do not. Capsular PSs from 6B and 6X1 have a 1→4 rhamnose–ribitol linkage, whereas the PSs from 6A and 6C have a 1→3 rhamnose–ribitol linkage. The two rhamnose–ribitol linkages can be readily distinguished because the 1→3 linkage is very sensitive to alkali hydrolysis, but the 1→4 linkage is not (Zon *et al.*, 1982).

To distinguish between the two rhamnose–ribitol linkage types, we investigated capsular PS from MNZ21 to determine its susceptibility to alkali hydrolysis. As shown in Fig. 3, serotype 6A and 6C PSs retained less than 1% of their original antigenicity after 1.5 hours of alkali hydrolysis. In contrast, 6B, 6X1, and MNZ21 PSs retained significantly more of their antigenicity after 0.5 and 1.5 hours. Thus, MNZ21 PS is as resistant to alkali hydrolysis as are the 6B and 6X1 PSs and therefore must not have the hydrolytically unstable 1→3 linkage.

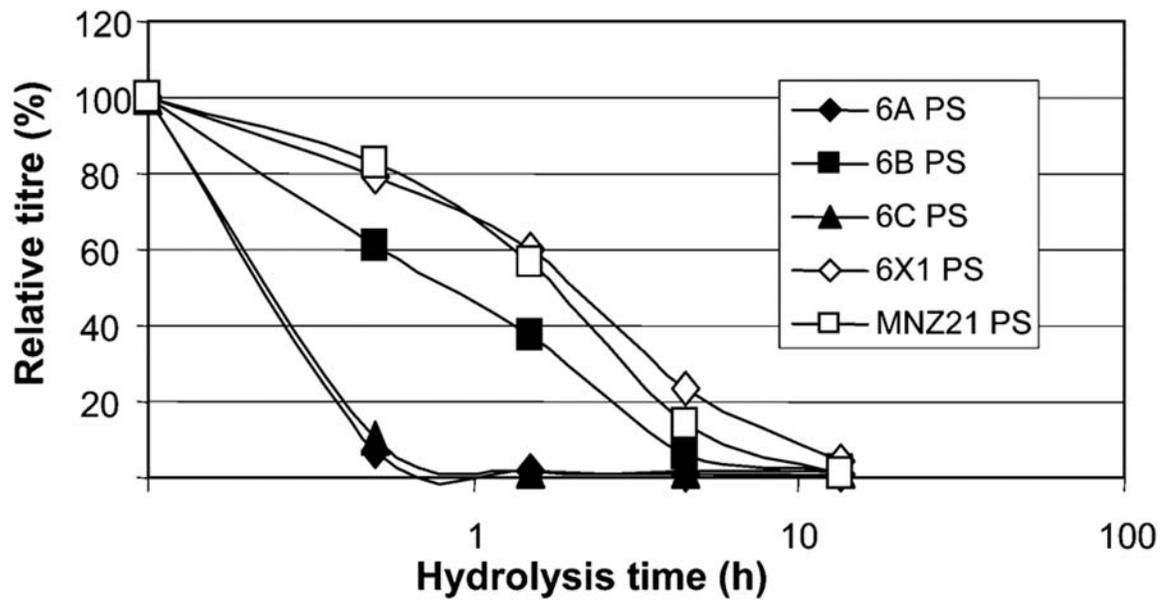


Fig. 3. Amount of antigenicity (y -axis) remaining after alkali hydrolysis for different time periods (x -axis). The PSs studied are listed in the figure key.

Studies of 6X1 PS showed that the product of *wciN_β* inserts glucose instead of galactose (Bratcher *et al.*, 2009). To determine that the *wciN_β* gene in the natural isolates is functional, we purified capsular PS from MNZ21 and determined its component carbohydrates (Fig. 4b). When we analyzed 6B PS for component carbohydrates, we found peaks for ribitol, rhamnose, galactose and glucose (marked in Fig. 4a). Also, the relative peak heights for the four carbohydrates were comparable. For instance, the major rhamnose peak was as tall as the dominant glucose peak. When we examined the component carbohydrates of MNZ21 PS (Fig. 4b), we found peaks for ribitol, rhamnose and glucose, but we did not find the galactose peaks which precede the glucose peaks in 6B PS. Also, in contrast to 6B PS (Fig. 4a), the major glucose peak of MNZ21 PS was much taller than the major rhamnose peak (Fig. 4b). This is also consistent with the conclusion that MNZ21 PS has one rhamnose but two glucose residues per repeating unit. Thus, the *wciN_β* gene in MNZ21 PS is functional, and MNZ21 capsular PS has glucose instead of galactose.

The above serological, genetic and biochemical studies clearly show that capsule type 6X1 can be expressed among natural pneumococcal isolates. Also, since the submission of our manuscript, pneumococcal isolates possessing the genetic characteristics of serotype 6X1 were discovered in Fijian islands (Jin *et al.*, 2009). In view of these results, we suggest that the serotype designation 6D should be used in place of 6X1. Furthermore, finding natural isolates expressing the 6D capsule type in two very distant parts of the world suggests that this serotype is widespread throughout the world and that typing for 6D should be included in the serotyping scheme for pneumococci.

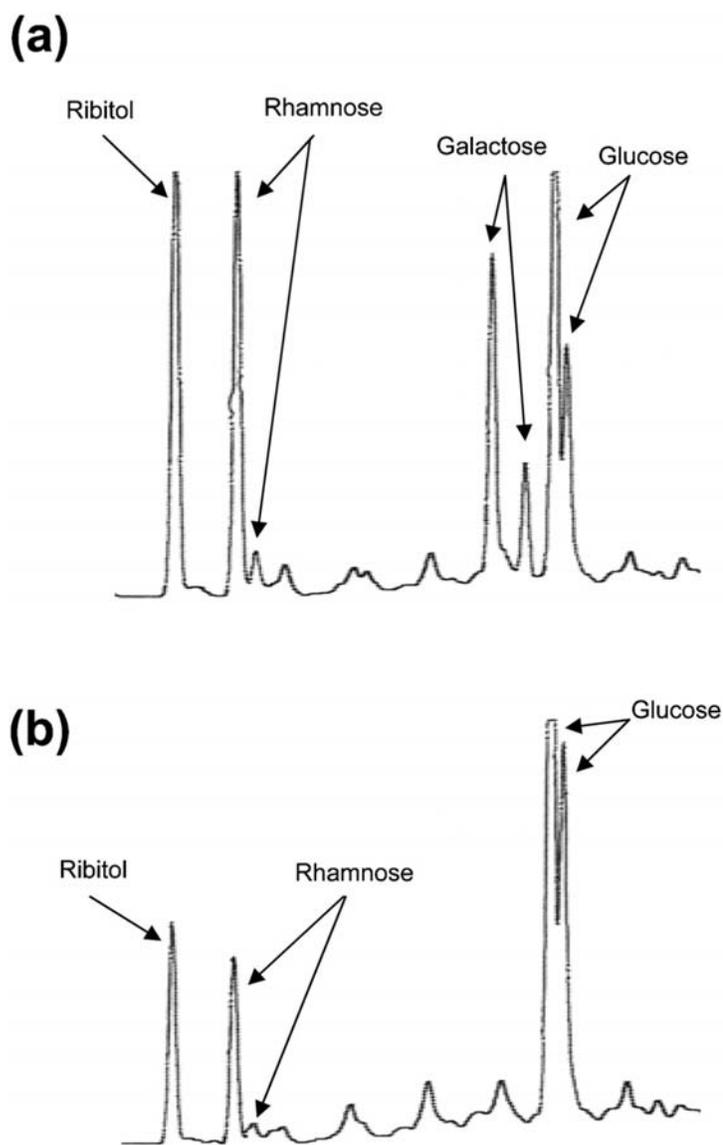


Fig. 4. GLC profile of serotype 6B PS (a) and the PS purified from a pneumococcal isolate MNZ21 (b). The GLC instrument used here identified two peaks for rhamnose, two peaks for galactose and two peaks for glucose (Dr D. Pritchard, personal communication); these peaks are labeled in (a) along with the ribitol peak. The elution times (in minutes) for the labeled peaks (from left to right) are: (a) 10.325 (for ribitol), 10.692 and 10.805 (for rhamnose), 12.219 and 12.414 (for galactose), and 12.585 and 12.585 (for glucose); (b) 10.323, 10.690, 10.803, 12.582 and 12.655.

The discovery of serotype 6D completes the search for the potential members of serogroup 6, but it now raises new questions to be studied. For instance, the prevalence of the 6D serotype should now be investigated in different parts of the world, among different disease types, and among different age groups. As the use of pneumococcal vaccines may influence the prevalence of 6D as it did the prevalence of 6C (Nahm *et al.*, 2009; Park *et al.*, 2008), discovery of the 6D serotype will also have an important impact on future studies of pneumococcal conjugate vaccines.

Acknowledgments

We thank Dr D. Pritchard, S. Dong, J. Yu and I. Park for discussions and assistance. The work was supported with NIH funding (AI-31473) to M. H. N.

References

Avery, O. T. & Dubos, R. (1931). The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med* **54**, 73–89.

Baker, J. R., Liu, C., Dong, S. & Pritchard, D. G. (2006). Endopeptidase and glycosidase activities of the bacteriophage B30 lysin. *Appl Environ Microbiol* **72**, 6825–6828.

Bratcher, P. E., Park, I. H., Hollingshead, S. K. & Nahm, M. H. (2009). Production of a unique pneumococcal capsule serotype belonging to serogroup 6. *Microbiology* **155**, 576–583.

Cole, R. (1913). Treatment of pneumonia by means of specific serums. *JAMA* **61**, 663–666.

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**, 3049–3060.

Fedson, D. S. (1988). Pneumococcal vaccine. In *Vaccines*, pp. 271–299. Edited by S. A. Plotkin & E. A. Mortimer, Jr. Philadelphia: W.B. Saunders.

Jin, P., Kong, F., Xiao, M., Oftadeh, S., Zhou, F., Liu, C., Russell, F. & Gilbert, G. L. (2009). First report of putative *Streptococcus pneumoniae* serotype 6D among nasopharyngeal isolates from Fijian children. *J Infect Dis* **200**, 1375–1380.

Kamerling, J. P. (2000). Pneumococcal polysaccharides: a chemical view. In *Streptococcus Pneumoniae Molecular Biology and Mechanisms of Disease*, pp. 81–114. Edited by A. Tomasz. Larchmont: Mary Ann Liebert.

Lin, J., Kaltoft, M. S., Brandao, A. P., Echaniz-Aviles, G., Brandileone, M. C., Hollingshead, S. K., Benjamin, W. H. & Nahm, M. H. (2006). Validation of a multiplex pneumococcal serotyping assay with clinical samples. *J Clin Microbiol* **44**, 383–388.

Mavroidi, A., Godoy, D., Aanensen, D. M., Robinson, D. A., Hollingshead, S. K. & Spratt, B. G. (2004). Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J Bacteriol* **186**, 8181–8192.

Nahm, M. H., Lin, J., Finkelstein, J. A. & Pelton, S. I. (2009). Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine. *J Infect Dis* **199**, 320–325.

Park, I. H., Park, S., Hollingshead, S. K. & Nahm, M. H. (2007a). Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* **75**, 4482–4489.

Park, I. H., Pritchard, D. G., Cartee, R., Brandao, A., Brandileone, M. C. & Nahm, M. H. (2007b). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**, 1225–1233.

Park, I. H., Moore, M. R., Treanor, J. J., Pelton, S. I., Pilishvili, T., Beall, B., Shelly, M. A., Mahon, B. E. & Nahm, M. H. (2008). Differential effects of pneumococcal vaccines against serotypes 6A and 6C. *J Infect Dis* **198**, 1818–1822.

Robinson, K. A., Baughman, W., Rothrock, G., Barrett, N. L., Pass, M., Lexau, C., Damaske, B., Stefonek, K., Barnes, B. & other authors (2001). Epidemiology of invasive *Streptococcus pneumoniae* infections in the United States, 1995–1998: opportunities for prevention in the conjugate vaccine era. *JAMA* **285**, 1729–1735.

Sun, Y., Hwang, Y. & Nahm, M. H. (2001). Avidity, potency, and cross-reactivity of monoclonal antibodies to pneumococcal capsular polysaccharide serotype 6B. *Infect Immun* **69**, 336–344.

Zon, G., Szu, S. C., Egan, W., Robbins, J. D. & Robbins, J. B. (1982). Hydrolytic stability of pneumococcal group 6 (type 6A and 6B) capsular polysaccharides. *Infect Immun* **37**, 89–103.

EVOLUTION OF PNEUMOCOCCAL SEROGROUP 6 POST-DISCOVERY OF THE
NOVEL 6C AND 6D SEROTYPES

by

PRESTON E. BRATCHER, IN HO PARK, MELISSA B. OLIVER, MARIA HORTAL,
ANNALISA PANTOSTI, SUSAN K. HOLLINGSHEAD, MOON H. NAHM

In preparation for *Journal of Bacteriology*

Format adapted for dissertation

Abstract

Streptococcus pneumoniae strains expressing serogroup 6 capsules are frequently detected in both asymptomatic colonization of the nasopharynx and in pneumococcal infections. Consequently, the evolutionary origins of serogroup 6 have been extensively studied. For many years, serogroup 6 had been known to contain serotypes 6A and 6B, whose *cps* loci consistently differ by only a single base pair in *wciP*. Recent studies have discovered two new serotypes, named serotypes 6C and 6D, which have *wciN_β* in place of the *wciN_α* present in the 6A and 6B *cps* loci. In view of these developments, we reinvestigated the genetic diversity of serogroup 6 *cps* loci and genomic backgrounds by examining the capsule gene profiles and multilocus sequence types of a diverse collection of serogroup 6 isolates. We confirmed that there are two distinct classes of serogroup 6 *cps* loci: class 1 sequences (possessed by all serogroup 6 members) and class 2 sequences (serotype 6B isolates containing an INDEL sequence). Neighbor joining phylogenetic analysis of class 1 *cps* loci showed that serotypes 6A, 6B, and 6C isolates formed 3 distinct clusters, suggesting that although their *cps* loci may have had similar origins, they have remained genetically independent. The *cps* locus of a serotype 6D strain was associated with the class 1 6B cluster, and seems to have been created by a genetic recombination between 6C and 6B *cps* loci. Few isolates are placed outside the 4 main clusters (class 1 6A, 6B, 6C, and class 2 6B) and even fewer belong to a cluster dominated by a different serotype. These exceptions seem to be the result of recombination or somatic mutation in the *wciP* gene. In order to examine the origins of serotype 6C, we examined the *wciN_β* and flanking regions from our collection, and found that there was very little diversity in this region, suggesting a single origin of the serotype 6C *cps* locus. The multi-locus sequence typing studies of serotype 6C isolates suggest

that, after its creation, the 6C *cps* has been incorporated into many different pneumococcal genomic backgrounds, but may preferentially enter serotype 6A genomic background.

Introduction

Streptococcus pneumoniae (pneumococcus) is a common colonizer of the human nasopharynx, yet it is an important human pathogen responsible for a multitude of diseases, mainly in children, the elderly, and the immune-compromised (Lynch & Zhanel, 2009). Using a polysaccharide capsule, of which there are 93 structurally distinct types (Bratcher *et al.*, 2010; Calix & Nahm, 2010; Henrichsen, 1995; Park *et al.*, 2007b), this bacterium is able to shield its surface from recognition by the host immune system, thereby making the capsule a potent colonization/virulence factor (Avery & Dubos, 1931; Bogaert *et al.*, 2004). Some capsule types (serotypes) are more prevalent in disease than others, and serogroup 6 strains are more commonly isolated in infections than the majority of other serotypes. Vaccination strategies in use today target the pneumococcal capsule from the most prevalent serotypes and most pneumococcal vaccines, including the widely used 7-valent conjugate vaccine (PCV-7), contain serotype 6B polysaccharide.

Because of its clinical importance, serogroup 6 evolution has been previously studied in detail. Serogroup 6 capsular polysaccharide synthesis loci (*cps*) encode 14 ORFs and range in size from ~17 to ~19 kbps due to variations in the non-coding regions found at either end (Figure 1) (Mavroidi *et al.*, 2004; Mavroidi *et al.*, 2007; Park *et al.*, 2007a). Like most serotypes, the central *cps* regions from all serogroup 6 strains have a lower G+C content than the surrounding genes/genomic average, and three of these genes (*wciP*, *wzy*, and *wzx*) are highly specific to serogroup 6 (Mavroidi *et al.*, 2004). Using

DNA sequences of selected parts of these serogroup 6 specific genes to create a “*cps* profile”, Mavroidi et al. showed that the only consistent genetic difference between 6A and 6B was a single nonsynonymous polymorphism in the *wciP* gene (*wciP_α* of 6A has a G at nt 584 while *wciP_β* of 6B has an A) and that serogroup 6 *cps* loci can be divided into two distinct classes (Mavroidi *et al.*, 2004). Class 1 6A and 6B strains do not have an INDEL, while the class 2 strains with divergent *cps* profiles have an INDEL and are all serotype 6B. The *cps* loci from different classes have 5.4% sequence divergence whereas those of class 1 differ by only 1-2% (Mavroidi *et al.*, 2004).

Since these studies have been published, two new serogroup 6 members have been discovered. One new serotype discovered in 2007, 6C, is serologically similar to 6A and has a glucose residue replacing the galactose residue in 6A PS (Park *et al.*, 2007b). The 6C *cps* locus is very homologous (98%) to the 6A *cps* except that it contains a unique *wciN* gene (referred to as *wciN_β*) which has no sequence homology (50%) with other pneumococcal genes, including *wciN_α* of 6A and 6B (Figure 1) (Park *et al.*, 2007a). While PCV-7 has been shown to reduce both the occurrence of carriage and invasive disease of the vaccine related serotype 6A, the serotype 6C seems to be able to evade this cross-reactive protection allotted by the vaccine and is, therefore, increasing in prevalence (Carvalho Mda *et al.*, 2009; Leach *et al.*, 2009; Nahm *et al.*, 2009; Park *et al.*, 2008; Tocheva *et al.*, 2010). Also, discovered recently was serotype 6D (Bratcher *et al.*, 2010; Jin *et al.*, 2009), which can be genetically distinguished from serotype 6C by the *wciP* gene (Bratcher *et al.*, 2009). Serotype 6C contains *wciP_α* whereas serotype 6D has *wciP_β*. In view of these changes in serogroup 6, we have reinvestigated the genetic evolution of serogroup 6 by studying the origins of these new serotypes as well as the mechanism of expansion of serotype 6C.

Methods

PCR, sequencing, and sequence analysis

Genomic DNA was purified from pneumococci using phenol-chloroform extraction and PCR-amplified as described (Mavroidi *et al.*, 2004) using an appropriate pair of primers and a PCR mixture. Primers are shown in Table 1 and the PCR reaction contained 37.75 microliters of sterile water, 1 microliter of genomic DNA, 2 microliters of each primer (5 pmol), 2 microliters of 10mM dNTP, 5 microliters of 10X LA Taq buffer solution, and 0.25 microliters of LA Taq polymerase (2.5 U/microliter; Takara Bio Inc, Shiga, Japan). The DNA sequence of PCR product was determined by the Genomics Core Facility at UAB.

DNA sequences of selected parts of *wciP*, *wzy*, and *wzx* genes were subjected to *cps* profiling studies using approaches described previously (Mavroidi *et al.*, 2004). Alleles were assigned according to the designations previously used, and new alleles were given arbitrarily numbered designations as shown in Figure 3C. The sequences of *wciP*, *wzy*, and *wzx* were concatenated and the concatenated sequences were subjected to neighbor joining analysis (Saitou & Nei, 1987) to investigate evolutionary relationship among *cps* loci of different serogroup 6 isolates. Average pairwise distances were calculated and all trees were drawn using MEGA4 (Tamura *et al.*, 2007).

PCR amplicons were subjected to Multilocus Sequence Typing (MLST) analysis as previously described (Enright & Spratt, 1998). Known alleles were then identified using the pneumococcal MLST website (<http://spneumoniae.mlst.net>) and new allele numbers assigned to new sequences by the database curator. All strains with MLST data

Table 1. List of PCR Primers

Primer names	Location of the primers	Primer sequence	Source or reference
Forward primers			
5117	<i>wchA</i>	5'-ATCAAGTGGTATTGGAAGCGGG	This study
5386	<i>wciN_β</i>	5'-CTGCTTTCCAAAGAGTTCG	This study
5106	<i>wchA</i>	5'-TACCATGCAGGGTGGAATGT	(Park <i>et al.</i> , 2007a)
5108	<i>wciP</i>	5'-ATGGTGAGAGATATTTGTCAC	(Mavroidi <i>et al.</i> , 2004)
5140	<i>wzy</i>	5'-CCTAAAGTGGAGGGAATTTTCG	(Mavroidi <i>et al.</i> , 2004)
5141	<i>wzx</i>	5'-TTCGAATGGGAATTCAATGG	(Mavroidi <i>et al.</i> , 2004)
Reverse Primers			
3386	<i>wciN_β</i>	5'-TAATATACCTATCAACTCCACCGC	This study
3102	<i>wciP</i>	5'-CTGGCATGTCATCTTTAGAAAA	This study
3101	<i>wciO</i>	5'-CCATCCTTCGAGTATTGC	(Park <i>et al.</i> , 2007a)
3107	<i>wciP</i>	5'-AGCATGATGGTATATAAGCC	(Mavroidi <i>et al.</i> , 2004)
3143	<i>wzy</i>	5'-CCTCCCATATAACGAGTGATG	(Mavroidi <i>et al.</i> , 2004)
3144	<i>wzx</i>	5'-GCGAGCCAAATCGGTAAGTA	(Mavroidi <i>et al.</i> , 2004)

listed in Table 2 have been inserted into the online pneumococcal MLST database.

Evolutionary relations among MLST types were performed with eBURST; clonal clusters were defined by strains sharing 6 alleles, and the clusters were assigned a number according to the predicted founder's sequence type.

Pneumococcal isolates

Fifty-seven isolates were collected during the years 1999-2008 in 4 different continents (Table 2). Twenty-four isolates were 6A, 25 were 6C, 6 were 6B, and 2 were 6D. These 57 isolates were subjected to *cps* profiling studies as well as MLST studies. To supplement *cps* profiling studies, we studied an additional 12 6B isolates from the same collection. The resulting panel of 69 isolates included 18 isolates from Asia, 18 from Europe, 19 from S. America, and 14 from N. America. Full information on the additional 12 6B isolates is provided as a supplementary table.

Results

***wciN_β* and flanking region sequence variations amongst 6C and 6D isolates**

When the *wciN_β* gene and flanking region sequences (total 2782 bases) from 26 6C isolates and 2 6D isolates (Figure 2) which were obtained from 4 different continents over a 10 year period were examined, the sequences showed little variation (16 variable sites out of 1125 (1.42 %) with no insertions or deletions. Furthermore, 9 clinical isolates collected from three different continents (Europe, N. America, and S. America) over an 8 year period have the exactly same nucleotide sequences in this region (Figure 2, first 9 sequences). If the 6C *cps* locus was created multiple times, we would expect this diversity to be much greater, especially since we expect that the *wciN_β* gene came from a foreign source.

Our previous study suggested that the insertion of *wciN_β* may have been facilitated by two clearly identifiable flanking regions (identified in gray in Figure 2), which are about 300 bases and 110 bases long in the 5' and 3' regions respectively (Park *et al.*, 2007a). These regions were defined by comparison to a 6A *cps* locus in GenBank (CR931638). Upon reanalysis, the 5' flanking region is very similar to the corresponding region of pneumococcal serotype 33F *cps* and the 3' flanking region is very close to the corresponding region of a class 2 6B strain *cps* as shown at the bottom of Figure 2. Outside these regions, however, the 6C locus is more closely related to the 6A locus used in the previous analysis. Given that the level of diversity in the entire *wciN_β* and flanking region sequenced is very low amongst 6C strains, it may be that multiple recombination events were necessary to create the 6C locus, including incorporation of foreign DNA (containing *wciN_β*) and subsequent recombination events which made the 6C locus functional. Alternatively, the entire *wciN_β* and flanking region could have come from a single foreign source of unknown identity. No matter the source of the *wciN_β* and flanking region of 6C strains, based on the extreme homology between this region in 6C and 6D strains, we suggest that the *wciN_β* and flanking regions in the 6D *cps* locus shares origins with this region in the 6C *cps* locus.

Analysis of the capsule gene loci of 6C and 6D isolates

Recently, isolates expressing serotype 6D were discovered in nature. In order to examine the evolutionary relationship of serotype 6D with the other 3 members of serogroup 6, we determined the sequence of the entire *cps* locus of a serotype 6D isolate (GenBank accession no. HM171374). As illustrated in Figure 1, the 6D *cps* is bound by *dexB* and *aliA* as are the other capsule gene loci (Garcia *et al.*, 2000), and has

transposase-like regions at each end. Between these transposase-like regions, the 6D *cps* has the 14 functional genes found in the other serogroup 6 *cps* loci and the sequence of the 6D *cps* locus is almost identical (98.6% identity in the 14,933 bps that include all the open reading frames) to that of 6C (GenBank accession no. EF538714) except for the known difference in *wciP* and differences in the non-coding regions flanking the 14 genes. We therefore conclude that the sources for the coding regions of the 6D *cps* locus were previously existing pneumococcal strains.

To study evolutionary relationship of serotype 6C with the serotypes 6A and 6B, we determined the *cps* profiles of 57 isolates listed in Table 2 along with 12 additional 6B isolates from these same locations. The *cps* profile was determined exactly as described previously by sequencing a portion of three genes (*wciP*, *wzy*, and *wzx*) that are found in all serogroup 6 capsule gene loci (Mavroidi *et al.*, 2004). Our study identified 11 new alleles (shown in Figure 3): 4 for *wciP*, 5 for *wzy*, and 2 for *wzx*. When we concatenated the three sequences (*wciP*, *wzy*, and *wzx*) of each isolate and determined the evolutionary tree for the 69 isolates using the neighbor joining method (Figure 3A), the class 2 isolates (i.e. isolates with the INDEL) could be clearly separated from the class 1 isolates (isolates without INDEL) as previously described (Mavroidi *et al.*, 2004) with the pairwise genetic distances between them being 5.54% or greater (class 2 vs 6A/6B class 1). When class 1 isolates are examined, they formed 3 distinct clusters as seen in Figure 3A. Although the two 6D isolates are in the cluster for 6B isolates, each cluster is made up of primarily one serotype, and each cluster was herein named for the dominant serotype. The genetic distance between serotype 6A and 6B clusters was 0.259%, but the genetic distance of the 6C cluster from the 6A and 6B clusters was 0.779%.

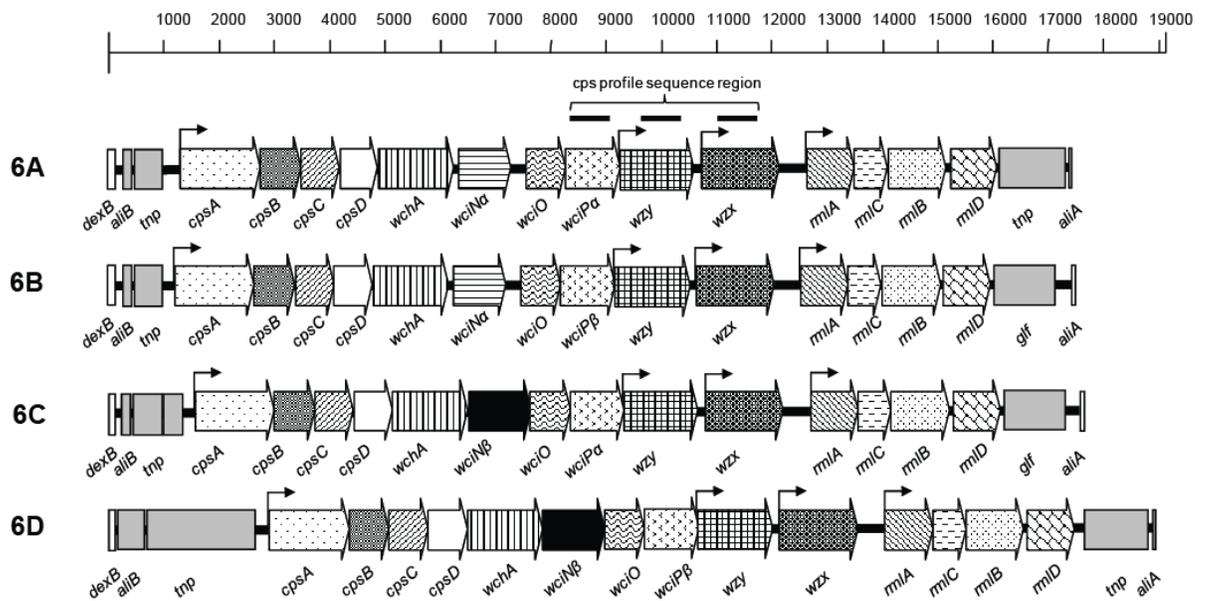


Figure 1. Capsule gene loci of a 6A strain (Genbank accession no. CR931638), 6B strain (Genbank accession no. CR931639), 6C strain (Genbank accession no. EF538714), and 6D strain (Genbank accession no. HM171374). All ORFs involved in capsule synthesis are shown as horizontal arrows, and their direction indicates the transcriptional orientation. The three regions of *wciP*, *wzy*, and *wzx* used for *cps* profiling are shown as three black bars above the 6A *cps* diagram. *wciN* and *wciP* alleles are indicated with α and β . The gene fragment from the beginning of *wciN β* to the end of *wzx* is about 4,130 bases long.

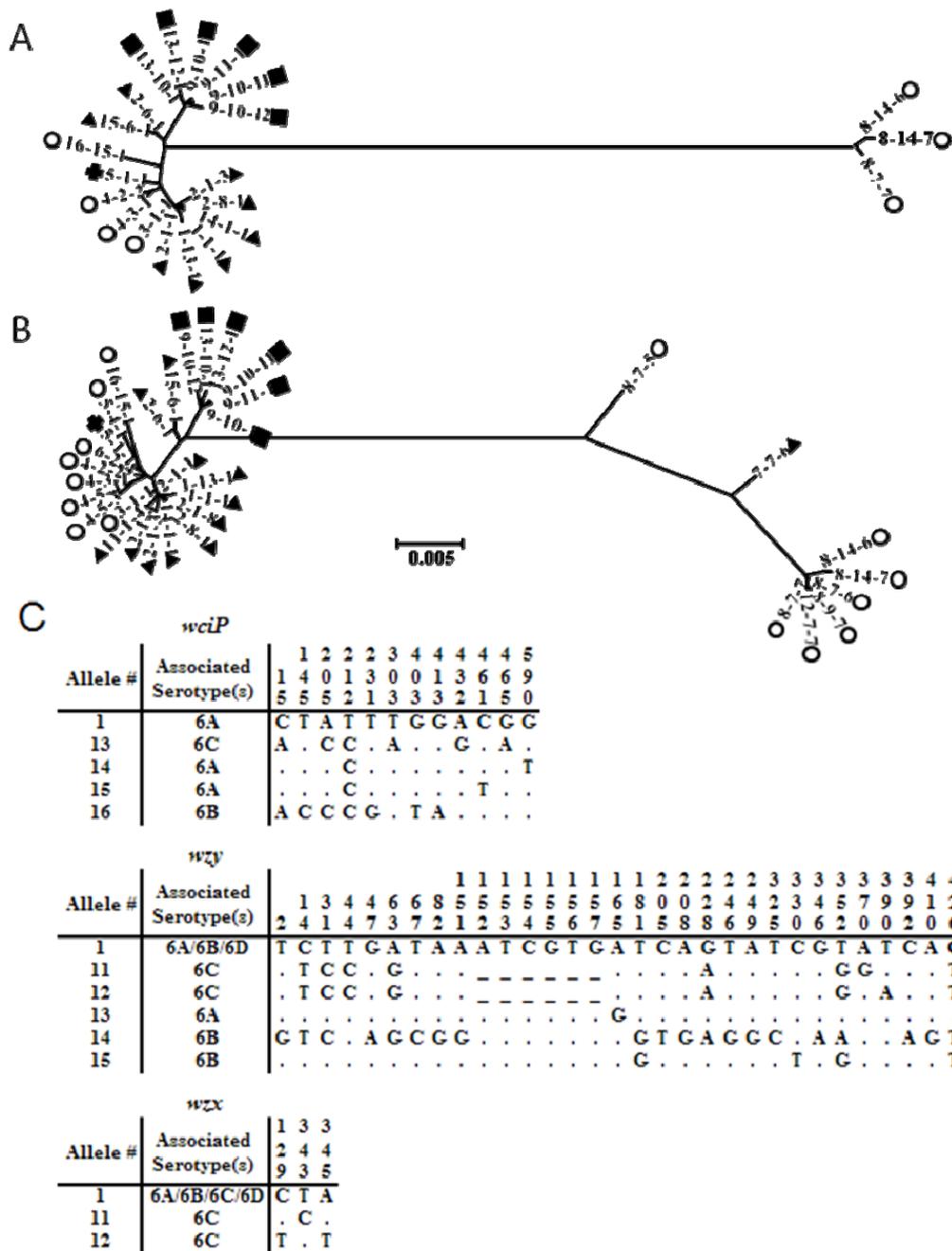


Figure 3. Neighbor joining trees of unique *cps* profiles. (A) Tree constructed with *cps* profile data from our study (see Table 2 for frequency of *cps* profiles and corresponding strain data). 24 6A, 18 6B, 25 6C, and 2 6D isolates were studied and 22 unique *cps* profiles were found. Symbols used to indicate serotypes are: 6A (solid triangle), 6B (open circle), 6C (solid square), and 6D (solid cross). Note that three class 2 6B isolates are shown in the right of the tree and all class 1 strains are shown in the left. (B) Tree

constructed with cps profile data from both our current study and a published study (Mavroidi *et al.*, 2004). Class 2 isolates are shown in the right and two cross-over strains (profiles 8-7-5 and 7-7-6) are shown in the middle. The bar represents genetic distance of 0.5%. (C) The polymorphic nucleotides are shown for all new alleles of the cps genes (*wciP*, *wzy*, and *wzx*). Allele 1 was shown on the top for comparison. Note that *wzy* alleles 11 and 12, which are associated with serotype 6C, have 6 base deletions (positions 152-157).

Interestingly, there are two class 1 profiles (2-6-1/15-6-1) that are more closely related to the class 2 cluster than the class 1 6A and 6B clusters (Figure 3). These strains have four, unique SNPs which separate their *wzy* allele from other class 1 6A and 6B alleles, and the class 2 sequences happen to have the same nucleotides at those locations. The 6C cluster is separated from the class 1 6A and 6B clusters for this same reason (allele 10 shares the four, unique SNPs mentioned above). However, the 6C cluster is separated from even the closely related 2-6-1/15-6-1 profiles because its *wzy* alleles have a 6 base pair deletion (Figure 3C). In contrast, no other *wzy* alleles have the deletion. The serotype specific clustering pattern did not change when the published *cps* profile data (Mavroidi *et al.*, 2004) is merged with our data (Figure 3B).

While most isolates fall into their own serotype specific clusters, there are exceptional isolates. One exception is the two serotype 6D isolates, which expressed profile 5-1-1, which falls into the class 1 6B cluster. Due to this and the fact that there is an exact sequence match of the 6D *wciN_β* and flank region with several 6C strains, we suggest that the *cps* locus of our 6D strains was most likely generated when 6C and 6B *cps* recombined at a location between *wciN* and *wciP*. More 6D isolates should be examined to see if the 6D *cps* locus has been created independently on several occasions.

Analysis of the Genomic Background of 6C Clinical Isolates

Since the introduction of conjugate vaccines, the prevalence of 6C has greatly increased in several parts of the world. To investigate whether all or selected clones of serotype 6C were increasing in prevalence, we performed MLST analysis on a global set of 6C isolates. Table 2 shows their sequence types (ST) along with the clonal complex to which they belong, which were identified by performing eBURST analysis using

resources available at the MLST website (<http://www.mlst.net>). Our results suggest that serotype 6C cps is associated with STs in multiple clonal complexes, as others have found (Carvalho Mda *et al.*, 2009; Jacobs *et al.*, 2009; Nunes *et al.*, 2009). Thus, 6C cps has moved into many different genetic backgrounds of pneumococci.

Inspection of the data in Table 2 suggested that 6A and 6C isolates often have identical STs and this association has been observed by others (Carvalho Mda *et al.*, 2009; Jacobs *et al.*, 2009). To investigate this issue more systematically, the ST associations between serotype 6C and other serotypes were investigated using the online MLST database (<http://spneumoniae.mlst.net>, accessed on 5/6/10). The website contained 433 6A STs, 480 6B STs, 80 6C STs, and 447 19A STs. The STs of 19A serotype was investigated because it is the most commonly found serotype in the nasopharynx (Huang *et al.*, 2009). Eighteen 6C isolates shared ST with 6A (18/80) but only 4 with 6B (4/80) and 1 with 19A (1/80). This indicates that 6C isolates share genetic backgrounds with 6A more than with other serotypes including 6B or 19A ($p < 0.0022$ by Fisher's exact test). In comparison, 6A shares a similar amount of STs with 6B (11/433) and 19A (6/433) [$p < 0.33$ by Fisher's exact test]. Similar results are seen when comparing shared clonal clusters as opposed to shared STs. While there may be confounders in the database beyond our control (e.g., sampling bias), this finding further suggests that the genetic backgrounds of 6C strains are more closely related to those of 6A isolates than those of other serotypes.

Discussion

A previous study of *cps* profiles of serogroup 6 isolates revealed two classes of *cps* profiles (named classes 1 and 2) and two major clusters within class 1, one cluster

Table 2. Strain data

	Strain	Location	Year Isolated	Serotype	Clonal Complex	Sequence Type	cps profile data			Source
							wciP	wzy	wzx	
1	MNZ664	S. America	1999	6A	CC176	4622	2	1	1	This study
2	MNZ680	S. America	2001	6A	CC176	4598	2	8	1	This study
3	MNZ681	S. America	2001	6A	CC176	4600	2	1	1	This study
4	MNZ616	S. America	2004	6A	CC176	4598	2	8	1	This study
5	MNZ631	S. America	2003	6A	CC176	4598	2	1	1	This study
6	MNZ632	S. America	2002	6A	CC176	4598	2	1	1	This study
7	MNZ208	Europe	1999	6A	*NPF1	2611	2	6	1	This study
8	MNZ446	N. America	2007	6A	CC2090	376	2	6	1	This study
9	MNZ428	N. America	2007	6A	CC2090	1538	2	6	1	This study
10	MNZ677	S. America	1999	6A	CC315	1093	2	6	1	This study
11	MNZ218	Europe	2001	6A	CC395	327	15	6	1	This study
12	MNZ459	N. America	2007	6A	CC460	460	2	6	1	This study
13	MNZ497	N. America	2004	6A	CC460	460	1	1	1	This study
14	MNZ479	N. America	2004	6A	CC460	460	1	13	1	This study
15	MNZ226	Europe	2003	6A	CC473	4595	2	1	3	This study
16	MNZ239	Europe	2008	6A	CC473	813	2	1	1	This study
17	MNZ471	N. America	2007	6A	CC473	1876	14	1	1	This study
18	MNZ462	N. America	2007	6A	CC473	1876	14	1	1	This study
19	MNZ212	Europe	1999	6A	CC490	4363	2	1	1	This study
20	MNZ222	Europe	2002	6A	CC690	690	14	1	1	This study
21	MNZ19	Asia	2008	6A	CC81	282	2	1	1	This study
22	MNZ28	Asia	2008	6A	CC81	81	2	1	1	This study
23	MNZ13	Asia	2008	6A	CC81	81	2	1	1	This study
24	MNZ17	Asia	2008	6A	Singleton	4620	2	1	1	This study
25	MNZ31	Asia	2008	6B	CC176	4604	16	15	1	This study
26	MNZ59	Asia	2008	6B	CC176	4605	16	15	1	This study
27	MNZ02	Asia	2008	6B INDEL	CC180	1624	8	14	7	This study
28	MNZ39	Asia	2008	6B INDEL	CC90	90	8	7	7	This study
29	MNZ60	Asia	2008	6B INDEL	CC90	4606	8	14	7	This study
30	MNZ62	Asia	2008	6B INDEL	CC90	90	8	14	7	This study
31	MNZ440	N. America	2007	6C	CC1379	1379	13	10	1	This study
32	MNZ613	S. America	2002	6C	CC1379	1379	9	10	1	This study
33	MNZ472	N. America	2007	6C	CC1379	1379	13	10	1	This study
34	MNZ595	S. America	2001	6C	CC1379	4599	13	10	1	This study
35	MNZ480	N. America	2004	6C	CC1390	1390	13	12	1	This study
36	MNZ488	N. America	2004	6C	CC1390	1390	13	10	1	This study
37	MNZ435	N. America	2007	6C	CC1390	1390	13	10	1	This study
38	MNZ458	N. America	2007	6C	CC1439	4602	13	10	1	This study
39	MNZ219	Europe	2002	6C	CC1715	1715	9	10	1	This study
40	MNZ240	Europe	2008	6C	CC176	2689	9	10	1	This study
41	MNZ23	Asia	2008	6C	CC176	4596	9	10	12	This study
42	MNZ34	Asia	2008	6C	CC176	4596	9	10	12	This study
43	MNZ38	Asia	2008	6C	CC176	4597	9	10	12	This study
44	MNZ604	S. America	1999	6C	CC3854	4601	9	10	1	This study
45	MNZ678	S. America	1999	6C	CC3854	4621	9	10	1	This study
46	MNZ597	S. America	2002	6C	CC3854	4601	9	10	1	This study
47	MNZ683	S. America	2001	6C	CC3854	4601	9	10	1	This study
48	MNZ592	S. America	2003	6C	CC3854	4601	9	10	1	This study
49	MNZ221	Europe	2002	6C	CC395	395	9	10	1	This study
50	MNZ213	Europe	1999	6C	CC395	395	9	10	1	This study
51	MNZ432	N. America	2007	6C	CC473	473	9	10	11	This study
52	MNZ209	Europe	1999	6C	Singleton	4603	9	11	1	This study
53	MNZ227	Europe	2004	6C	Singleton	3531	9	11	1	This study
54	MNZ16	Asia	2008	6C	Singleton	855	9	10	11	This study
55	MNZ18	Asia	2008	6C	Singleton	855	9	10	11	This study
56	MNZ21	Asia	2008	6D	CC81	282	5	1	1	Bratcher et al., 2009
57	MNZ22	Asia	2008	6D	CC81	282	5	1	1	Bratcher et al., 2009

primarily containing serotype 6A and the other 6B. This study was performed before serotypes 6C and 6D were discovered, and therefore provided incomplete evolutionary relationship of serogroup 6 *cps* loci. We have now studied *cps* profiles of more isolates including those of serotypes 6C and 6D. In addition to confirming the two classes, our *cps* profile studies show 3 distinct clusters within class 1 and each cluster can be associated with one of the three serotypes; 6A, 6B, and 6C. Serogroup 6 members are in a unique situation when compared to other serotypes; they have the ability to switch their capsule type by obtaining a single point mutation at a specific location. Given this fact and the observation that their *cps* profiles form serotype-specific clusters, we conclude that serotype switching amongst serogroup 6 members by a point mutation in *wciP* is a very rare event.

Genetic recombination is common among pneumococci and recombination between *cps* loci with similar sequences is thought to be frequent. For instance, serogroup 6 *cps* have a long (17 kb) stretch of DNA that are highly homologous (~95-98%) and recombination between serogroup 6 *cps* should be very common. Indeed, there are examples of recombination between isolates in classes 1 and 2 (Figure 3, (Mavroidi *et al.*, 2004)). In the past, several *cps* profiles were found that were clear mosaics, showing evidence of recombination between class 1 and class 2 sequences. As the class 1 isolates have three genetic clusters that are associated with a single serotype, we suggest that recombination with crossovers occurring within the *cps* profile region is very rare. Further evidence to support this comes from the fact that the 6C cluster maintained two genetic markers (*wciN_β* and *wzy*) that are separated by about 4 kb. Again, this suggests

then during serotype-switching events by recombination, a large portion, containing the serogroup 6 specific genes, is transferred.

Our studies shed new insights into the origins of serotype 6C. Previously we proposed that *wciN_β* gene of an unknown origin was inserted into the 6A *cps* with the help of two flanking regions with limited homology to a serotype 6A strain. As an alternative hypothesis we also considered a recombination event involving *wchA* gene of serotype 33F *cps* and class 2 serotype 6B *cps*, because about 500 bases in the 5' end of *wciN_β* are closely (98.9% identity) related to *wchA* gene of 33F (based on Genbank CR931702) and about 230 bases in the 3' end of *wciN_β* are very homologous (96%) to *cps6bQ* ORF present in class 2 (but not class 1) serotype 6A *cps* (as described in figure 2). Given this evidence and the unexpected presence of “short” *wzy* allele that is specifically associated with serotype 6C, we suggest it is possible that the creation of a serotype 6C strain involved many recombination events in order to form a functional 6C *cps* locus. However, until we know the source of the *wciN_β* gene we can only speculate, and it is also possible that this entire region, including the “short” *wzy* allele, was acquired from this foreign source.

The source of the foreign gene segment is unclear at the moment but would likely originate from bacteria in the nasopharyngeal tract where pneumococci normally reside. As we found limited heterogeneity in the sequences of serotype 6C *cps* loci, we believe that the 6C *cps* locus was created once. Based on the frequency of each profile and minimizing diversity required for the cluster, we believe the founder may have had *cps* profile of 9-10-1 and the consensus *wciN* and flank region (Figure 2); 5 out of the 9 strains that possess this sequence have the 9-10-1 profile. While the knowledge of the 6C

founder sequence may help in its identification, the bacterial gene pool in the nasopharynx is very large and identification of the exact source(s) of the 6C (or any serogroup 6) *cps* would not be simple.

While the origin of serotype 6C *cps* requires additional studies, the *cps* of the two Korean 6D isolates appears to have been created by a genetic recombination between serotypes 6B and 6C. Their *cps* may have recombined between *wciN* and *wciP* but more 6D isolates should be studied before we can conclude all 6D isolates arose in this manner. Interestingly, the Fijian 6D isolates have the same *wciP* allele 5 (Jin *et al.*, 2009) as the Korean 6D isolates. This suggests that 6D serotype may have been created once and spread throughout the world including to two widely separated areas: Korea and Fiji. Given its potentially wide distribution, it is interesting to note that serotype 6D is very rare and that it was not found in several large screens of clinical isolates (Bratcher *et al.*, 2009; Carvalho Mda *et al.*, 2009; Hermans *et al.*, 2008; Jacobs *et al.*, 2010). One of the potential reasons is that vaccination with the 6B PS has provided cross-protection against 6D and has prevented the expansion of the 6D serotype.

Although serogroup 6 *cps* are highly homologous, serotype 6C is associated with *cps* profiles that are 6C specific. However, we and others (Carvalho Mda *et al.*, 2009; Jacobs *et al.*, 2009) have shown that serotype 6C cannot be associated with a specific sequence type. While there are multiple potential explanations, we favor that entire *cps* can move among different pneumococcal backgrounds. This may happen because transposase-like regions are often found at either end of *cps*. The exact function of these regions is still not well understood. However, an interesting observation is that serotype 6C may preferentially share its STs with those of serotype 6A but not 6B or other

serotypes (e.g., serotype 19A). More studies are needed to explain how pneumococcal *cps* can move among different genetic backgrounds. Regardless of the mechanism, the observation that the 6C *cps* locus is expanding in diverse genetic backgrounds supports the hypothesis that the expansion of serotype 6C is due to selective pressure based on serotype (most likely from vaccination), not due to a 6C strain possessing a more advantageous background (e.g. one with multidrug resistance properties).

Our studies significantly clarify the evolutionary relationships of serogroup 6 (Figure 4). After examination of the *cps* profiles of a diverse set of serogroup 6 strains along with the complete sequences of several serogroup 6 *cps* loci, we confirm the hypothesis that the serogroup 6 *cps* loci were created twice (suggested by class 1 and class 2 (INDEL-containing) sequences). The class 1 sequences later diverged from each other, and the resulting sequences formed two distinct, serotype-specific clusters. We suggest that these came about from 6A and 6B remaining genetically separated, accumulating spontaneous mutations which became serotype specific. While this model works fine for remaining in a single genetic background or transferring into another serogroup background, it requires that when 6A and 6B serotype switch to each other, this must occur by transformation of the majority of the serogroup 6 specific genes (not a portion, and not by point mutation). Later, the 6C *cps* locus was generated by combining a foreign gene, *wciN_β*, with portions of the pneumococcal *cps* loci of several different strains (possibly including a class 1 6A strain, class 2 6B strain, and a 33F strain). Given the sequence similarity between all 6C strains and the two 6D strains we have examined, the source of *wciN_β* in 6D was from a 6C strain, while the rest of the 6D *cps* locus seems to have come from various serogroup 6 strains. The above hypothesis is outlined in

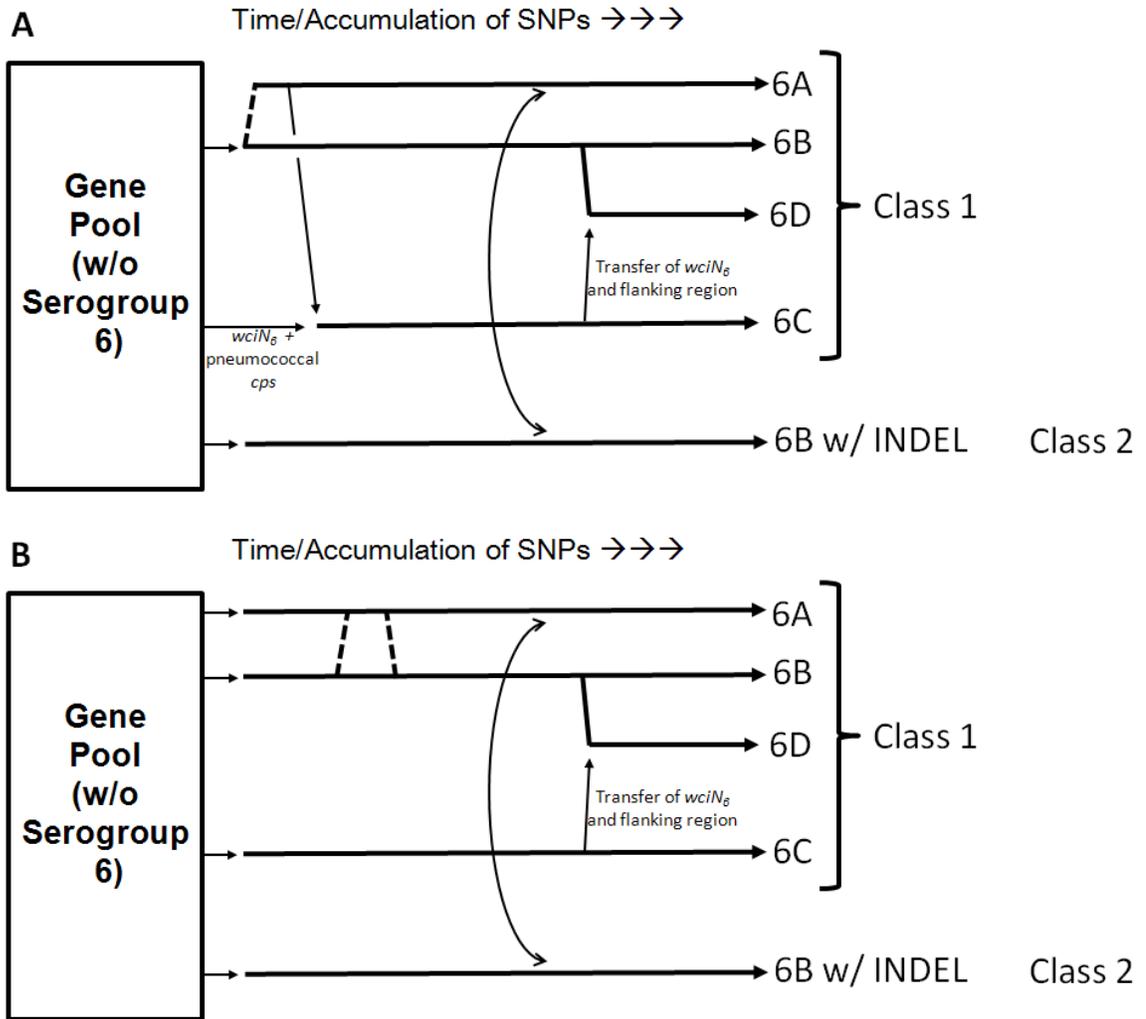


Figure 4. Two models for serogroup 6 evolution. The nature of “gene pool” from where the casule genes have originated is undefined at the moment. Genetic recombination events are shown with thin arrows. Mutations are shown with dotted lines.

Figure 4A. Alternatively, it is possible that each cluster in the *cps* profile tree represents a unique capture of the serogroup 6 genes from the gene pool (Figure 4B). Neither model can be proven, and we will not know which model is more accurate until the foreign sources have been defined. These findings show that the evolution of capsule gene loci is very complex and can occur in various ways. Perhaps it is because pneumococci is an expert at recombination and can create new capsule types so easily that *S. pneumoniae* can express so many different serotypes.

Acknowledgments

This work was supported with NIH funding (AI-31473) to M.H.N. and P.E.B. was partially supported by a training grant (T32 GM008111) from the NIH.

References

Avery, O. T. & Dubos, R. (1931). The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med* **54**, 73-89.

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

Bratcher, P. E., Park, I. H., Hollingshead, S. K. & Nahm, M. H. (2009). Production of a unique pneumococcal capsule serotype belonging to serogroup 6. *Microbiology* **155**, 576-583.

Bratcher, P. E., Kim, K. H., Kang, J. H., Hong, J. Y. & Nahm, M. H. (2010). Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical, and serological characterization. *Microbiology* **156**, 555-560.

Calix, J. J. & Nahm, M. H. (2010). A new pneumococcal serotype, 11E, has variably inactivated wcjE gene. *J Infect Dis* **In Print**.

Carvalho Mda, G., Pimenta, F. C., Gertz, R. E., Jr. & other authors (2009). PCR-based quantitation and clonal diversity of the current prevalent invasive serogroup 6 pneumococcal serotype, 6C, in the United States in 1999 and 2006 to 2007. *J Clin Microbiol* **47**, 554-559.

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), 3049-3060.

Garcia, E., Llull, D., Munoz, R., Mollerach, M. & Lopez, R. (2000). Current trends in capsular polysaccharide biosynthesis of *Streptococcus pneumoniae*. *Res Microbiol* **151**, 429-435.

Henrichsen, J. (1995). Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* **33**, 2759-2762.

Hermans, P. W., Blommaart, M., Park, I. H., Nahm, M. H. & Bogaert, D. (2008). Low prevalence of recently discovered pneumococcal serotype 6C isolates among healthy Dutch children in the pre-vaccination era. *Vaccine* **26**, 449-450.

Huang, S. S., Hinrichsen, V. L., Stevenson, A. E. & other authors (2009). Continued impact of pneumococcal conjugate vaccine on carriage in young children. *Pediatrics* **124**, e1-11.

Jacobs, M. R., Bajaksouzian, S., Bonomo, R. A. & other authors (2009). Occurrence, distribution, and origins of *Streptococcus pneumoniae* Serotype 6C, a recently recognized serotype. *J Clin Microbiol* **47**, 64-72.

Jacobs, M. R., Dagan, R., Bajaksouzian, S., Windau, A. R. & Porat, N. (2010).

Validation of factor antiserum 6d for serotyping *Streptococcus pneumoniae* serotype 6C.

J Clin Microbiol.

Jin, P., Kong, F., Xiao, M., Oftadeh, S., Zhou, F., Liu, C., Russell, F. & Gilbert, G. L.

(2009). First Report of Putative *Streptococcus pneumoniae* Serotype 6D among

Nasopharyngeal Isolates from Fijian Children. *J Infect Dis* **200**, 1375-1380.

Leach, A. J., Morris, P. S., McCallum, G. B., Wilson, C. A., Stubbs, L., Beissbarth,

J., Jacups, S., Hare, K. & Smith-Vaughan, H. C. (2009). Emerging pneumococcal

carriage serotypes in a high-risk population receiving universal 7-valent pneumococcal

conjugate vaccine and 23-valent polysaccharide vaccine since 2001. *BMC Infect Dis* **9**,

121.

Lynch, J. P., 3rd & Zhanel, G. G. (2009). *Streptococcus pneumoniae*: epidemiology,

risk factors, and strategies for prevention. *Semin Respir Crit Care Med* **30**, 189-209.

Mavroidi, A., Godoy, D., Aanensen, D. M., Robinson, D. A., Hollingshead, S. K. &

Spratt, B. G. (2004). Evolutionary genetics of the capsular locus of serogroup 6

pneumococci. *J Bacteriol* **186**, 8181-8192.

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**, 7841-7855.

Nahm, M. H., Lin, J., Finkelstein, J. A. & Pelton, S. I. (2009). Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine. *J Infect Dis* **199**, 320-325.

Nunes, S., Valente, C., Sa-Leao, R. & de Lencastre, H. (2009). Temporal trends and molecular epidemiology of recently described serotype 6C of *Streptococcus pneumoniae*. *J Clin Microbiol* **47**, 472-474.

Park, I. H., Park, S., Hollingshead, S. K. & Nahm, M. H. (2007a). Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* **75**, 4482-4489.

Park, I. H., Pritchard, D. G., Cartee, R., Brandao, A., Brandileone, M. C. & Nahm, M. H. (2007b). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**, 1225-1233.

Park, I. H., Moore, M. R., Treanor, J. J., Pelton, S. I., Pilishvili, T., Beall, B., Shelly, M. A., Mahon, B. E. & Nahm, M. H. (2008). Differential effects of pneumococcal vaccines against serotypes 6A and 6C. *J Infect Dis* **198**, 1818-1822.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406-425.

Tamura, K., Dudley, J., Nei, M. & Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. *Mol Biol Evol* **24**, 1596-1599.

Tocheva, A. S., Jefferies, J. M., Christodoulides, M., Faust, S. N. & Clarke, S. C. (2010). Increase in serotype 6C pneumococcal carriage, United Kingdom. *Emerg Infect Dis* **16**, 154-155.

SUMMARIZING CONCLUSION

The discovery and characterization of 6D isolates ends the search for a potential serogroup 6 member. Its capsular polysaccharide contains the predicted repeating unit made up of [\rightarrow 2) glucose (1 \rightarrow 3) glucose (1 \rightarrow 3) rhamnose (1 \rightarrow 4) ribitol (5 \rightarrow phosphate)]. As we are able to differentiate the four serogroup 6 members by monoclonal antibody binding profiles, we were able to screen our library for potential 6D isolates. After screening a large number of globally diverse strains, we were only able to find two 6D isolates that were obtained from the nasopharynx of children attending a daycare in Korea. Genetically, the two Korean isolates were created by recombination between a 6B and 6C strain, as shown by containing *wciN β* with a *wciP β* allele belonging to a 6B strain, not a *wciP α* in a 6C strain that had obtained the necessary mutation to switch to *wciP β* . As our results along with the results of others who have screened for 6D in clinical isolates show, the prevalence of 6D seems to be very low, and, so far, undetectable in highly vaccinated populations (Bratcher *et al.*, 2009; Carvalho Mda *et al.*, 2009; Hermans *et al.*, 2008; Jacobs *et al.*, 2010). Regardless, it may be important to monitor 6D prevalence in the future to see if it is selected for by vaccination, like 6C.

From closely examining several genetic properties of a diverse set of serotype 6C strains, we were able to come to a number of conclusions. Firstly, by comparing the sequences of the *wciN β* and flanking regions from these strains, we noticed that there was a small amount of diversity (~1% of sites were variable) with no insertions, deletions, or regions of increased amounts of variability. This suggested that there was a single origin of the *cps* locus of all serotype 6C strains. This suggestion was further supported by looking at the diversity in the *cps* profiles of these strains, for which there were even less

variable sites (~0.4%). As we studied a global set of strains, we have concluded that there was a single functional incorporation of *wciN_β* from an unknown foreign source which generated the original 6C *cps* locus. Since this original 6C strain was created, 6C has expanded into different genetic backgrounds, as opposed to one strain clonally expanding due to a selective advantage in the background, such as antibiotic resistance. This leads us to conclude that the positive selection for serotype 6C strains is mainly influenced by vaccination.

After analysis of the *cps* profiles of a set of our serogroup 6 strains along with those strains analyzed in a previous study, we came to the same conclusion that there may have been two independent serogroup 6 *cps* locus creation events, leading to the two diverse types of sequences (class 1 and class 2). In addition, amongst the class 1 sequences, phylogenetic trees show that three clusters form which are serotype specific (except for one 6B strain in the 6A cluster which formed by point mutation and the 6D strain). Interestingly, although we know serotype switching from one serogroup 6 member to another occurs in nature, as evidenced by isolation of multiple serogroup 6 serotypes in the same genetic background, the previous study only showed two sequences that were recombination events within the *cps* profiling region (our study found none). Therefore, we conclude that serotype switching by point mutation or recombination within *cps* profiling area (*wciP* to *wzx*) is a very rare event, and when serotype switching does occur, a large portion of the *cps* locus is transferred. We can hypothesize that serotype switching by point mutation should be a relatively common occurrence in any serogroup 6 infection, but as this event is rarely detectable, we believe that there may be a

lack of selection for 6A over 6B after exposure to 6B (or vice versa); if this selection did commonly occur, then we would not see *cps* profile clusters that were serotype specific.

As the vaccines only allot protection against a portion of the serotypes found in nature, it is important to continue monitoring serotype prevalence in the future. This is especially the case in the present time, since we are starting to witness serotype replacement caused by the use of the conjugate vaccine (Huang *et al.*, 2005; Mera *et al.*, 2008; Singleton *et al.*, 2007). As far as serogroup 6 is concerned, immunization with 6B is resulting in a decrease in both 6B and 6A prevalence in infections (Carvalho Mda *et al.*, 2009) and now, after conjugate vaccine use, in carriage as well (Leach *et al.*, 2009). The reverse effect is being seen for 6C, as both in vitro experiments (Park *et al.*, 2008) and epidemiological data suggest that the vaccines do not protect against 6C. Since only a small handful of 6D strains have been isolated, it is too early to tell whether vaccination will have an effect on 6D prevalence, but it is interesting to note that detection of 6D strains has been limited to China, Fiji, Australia, Finland, and Korea, and even at these locations it is very rare (<http://spneumoniae.mlst.net/>)(Jin *et al.*, 2009)(unpublished data).

The studies of serogroup 6 isolates have given us a rare insight into the evolution of the pneumococcal capsule. For one, 6A and 6B serotypes are so closely related that they can switch by changing a single nucleotide in the *wciP* gene. This observation has allowed us to gain insight into the influence of selection on serotype switching amongst serogroup 6 strains in their natural environment in the human nasopharynx. In addition, the discovery of the new serogroup 6 members 6C and 6D have provided an example of the formation of new serotypes by incorporation of a foreign gene into the *cps* locus of an existing serotype. Overall, this study suggests that the *cps* locus of *S. pneumoniae* is very

flexible, as evidenced by the multiple creations of serogroup 6 *cps* loci and the replacement of a single gene to form a new serotype. These observations highlight the need to monitor serotype prevalence as well as improve our methods for detection of new serotypes, especially in the age of vaccines.

GENERAL REFERENCES

(2010). Licensure of a 13-valent pneumococcal conjugate vaccine (PCV13) and recommendations for use among children - Advisory Committee on Immunization Practices (ACIP), 2010. *MMWR* **59**, 258-261.

Avery, O. T. & Dubos, R. (1931). The protective action of a specific enzyme against type III pneumococcus infection in mice. *J Exp Med* **54**, 73-89.

Barrett, D. J. (1985). Human immune responses to polysaccharide antigens: An analysis of bacterial polysaccharide vaccines in infants. *AdvPediatr* **32**, 139.

Bentley, S. D., Aanensen, D. M., Mavroidi, A. & other authors (2006). Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *Pub Lib of Sci Genet* **2**, e31.

Bratcher, P. E., Kim, K. H., Kang, J. H., Hong, J. Y. & Nahm, M. H. (2009). Identification of natural pneumococcal isolates expressing serotype 6D by genetic, biochemical, and serological characterization. *Microbiology*.

Bratcher, P. E., Park, I. H., Hollingshead, S. K. & Nahm, M. H. (2009). Production of a unique pneumococcal capsule serotype belonging to serogroup 6. *Microbiology* **155**, 576-583.

Calix, J. J. & Nahm, M. H. (2010). A new pneumococcal serotype, 11E, has variably inactivated wcjE gene. *J Infect Dis* **In Print**.

Carvalho Mda, G., Pimenta, F. C., Gertz, R. E., Jr. & other authors (2009). PCR-based quantitation and clonal diversity of the current prevalent invasive serogroup 6 pneumococcal serotype, 6C, in the United States in 1999 and 2006 to 2007. *J Clin Microbiol* **47**, 554-559.

Gould, G. A., Rhind, G. B., Morgan, A. D., Williamson, G. & Calder, M. A. (1987). Pneumococcal serotypes in sputum isolates during acute respiratory illness in Edinburgh. *Thorax* **42**, 589-592.

Henrichsen, J. (1995). Six newly recognized types of *Streptococcus pneumoniae*. *J Clin Microbiol* **33**, 2759-2762.

Hermans, P. W., Blommaart, M., Park, I. H., Nahm, M. H. & Bogaert, D. (2008). Low prevalence of recently discovered pneumococcal serotype 6C isolates among healthy Dutch children in the pre-vaccination era. *Vaccine* **26**, 449-450.

Huang, S. S., Platt, R., Rifas-Shiman, S. L., Pelton, S. I., Goldmann, D. & Finkelstein, J. A. (2005). Post-PCV7 changes in colonizing pneumococcal serotypes in 16 Massachusetts communities, 2001 and 2004. *Pediatrics* **116**, e408-413.

Inostroza, J., Trucco, O., Prado, V., Vinet, A. M., Retamal, G., Ossa, G., Facklam, R. R. & Sorensen, R. U. (1998). Capsular serotype and antibiotic resistance of *Streptococcus pneumoniae* isolates in two Chilean cities. *Clin Diagn Lab Immunol* **5**, 176-180.

Jacobs, M. R., Dagan, R., Bajaksouzian, S., Windau, A. R. & Porat, N. (2010). Validation of factor antiserum 6d for serotyping *Streptococcus pneumoniae* serotype 6C. *J Clin Microbiol*.

Jin, P., Kong, F., Xiao, M., Oftadeh, S., Zhou, F., Liu, C., Russell, F. & Gilbert, G. L. (2009). First Report of Putative *Streptococcus pneumoniae* Serotype 6D among Nasopharyngeal Isolates from Fijian Children. *J Infect Dis* **200**, 1375-1380.

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.

Kaijser, B. (1979). Pneumococcal infections and the possible need for a vaccine. *Scand J Infect Dis* **11**, 179-184.

Kenne, L., Lindberg, B. & Madden, J. K. (1979). Structural studies of the capsular antigen from *Streptococcus pneumoniae* type 26. *Carbohydr Res* **73**, 175-182.

Leach, A. J., Morris, P. S., McCallum, G. B., Wilson, C. A., Stubbs, L., Beissbarth, J., Jacups, S., Hare, K. & Smith-Vaughan, H. C. (2009). Emerging pneumococcal carriage serotypes in a high-risk population receiving universal 7-valent pneumococcal conjugate vaccine and 23-valent polysaccharide vaccine since 2001. *BMC Infect Dis* **9**, 121.

Lin, J., Kaltoft, M. S., Brandao, A. P., Echaniz-Aviles, G., Brandileone, M. C., Hollingshead, S. K., Benjamin, W. H. & Nahm, M. H. (2006). Validation of a multiplex pneumococcal serotyping assay with clinical samples. *J Clin Microbiol* **44**, 383-388.

Lynch, J. P., 3rd & Zhanel, G. G. (2009). Streptococcus pneumoniae: epidemiology, risk factors, and strategies for prevention. *Semin Respir Crit Care Med* **30**, 189-209.

Magee, A. D. & Yother, J. (2001). Requirement for capsule in colonization by Streptococcus pneumoniae. *Infect Immun* **69**, 3755-3761.

Mavroidi, A., Godoy, D., Aanensen, D. M., Robinson, D. A., Hollingshead, S. K. & Spratt, B. G. (2004). Evolutionary genetics of the capsular locus of serogroup 6 pneumococci. *J Bacteriol* **186**, 8181-8192.

Mera, R., Miller, L. A., Fritsche, T. R. & Jones, R. N. (2008). Serotype replacement and multiple resistance in *Streptococcus pneumoniae* after the introduction of the conjugate pneumococcal vaccine. *Microb Drug Resist* **14**, 101-107.

Morona, J. K., Miller, D. C., Morona, R. & Paton, J. C. (2004). The effect that mutations in the conserved capsular polysaccharide biosynthesis genes *cpsA*, *cpsB*, and *cpsD* have on virulence of *Streptococcus pneumoniae*. *J Infect Dis* **189**, 1905-1913.

Musher, D. M., Groover, J. E., Reichler, M. R., Riedo, F. X., Schwartz, B., Watson, D. A., Baughn, R. E. & Breiman, R. F. (1997). Emergence of antibody to capsular polysaccharides of *Streptococcus pneumoniae* during outbreaks of pneumonia: association with nasopharyngeal colonization. *Clin Infect Dis* **24**, 441-446.

Nahm, M. H., Lin, J., Finkelstein, J. A. & Pelton, S. I. (2009). Increase in the prevalence of the newly discovered pneumococcal serotype 6C in the nasopharynx after introduction of pneumococcal conjugate vaccine. *J Infect Dis* **199**, 320-325.

Park, I. H., Park, S., Hollingshead, S. K. & Nahm, M. H. (2007a). Genetic basis for the new pneumococcal serotype, 6C. *Infect Immun* **75**, 4482-4489.

Park, I. H., Pritchard, D. G., Cartee, R., Brandao, A., Brandileone, M. C. & Nahm, M. H. (2007b). Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. *J Clin Microbiol* **45**, 1225-1233.

Park, I. H., Moore, M. R., Treanor, J. J., Pelton, S. I., Pilishvili, T., Beall, B., Shelly, M. A., Mahon, B. E. & Nahm, M. H. (2008). Differential effects of pneumococcal vaccines against serotypes 6A and 6C. *J Infect Dis* **198**, 1818-1822.

Prymula, R. & Schuerman, L. (2009). 10-valent pneumococcal nontypeable Haemophilus influenzae PD conjugate vaccine: Synflorix. *Expert Rev Vaccines* **8**, 1479-1500.

Rebers, P. A. & Heidelberger, M. (1961). The specific polysaccharide of type VI pneumococcus. II. The repeating unit. *JAmChemSoc* **83**, 3056-3059.

Rennels, M. B., Edwards, K. M., Keyserling, H. L. & other authors (1998). Safety and immunogenicity of heptavalent pneumococcal vaccine conjugated to CRM197 in United States infants. *Pediatrics* **101**, 604-611.

Singleton, R. J., Hennessy, T. W., Bulkow, L. R., Hammitt, L. L., Zulz, T., Hurlburt, D. A., Butler, J. C., Rudolph, K. & Parkinson, A. (2007). Invasive pneumococcal disease caused by nonvaccine serotypes among alaska native children with high levels of 7-valent pneumococcal conjugate vaccine coverage. *Jama* **297**, 1784-1792.

Tocheva, A. S., Jefferies, J. M., Christodoulides, M., Faust, S. N. & Clarke, S. C. (2010). Increase in serotype 6C pneumococcal carriage, United Kingdom. *Emerg Infect Dis* **16**, 154-155.

Tuomanen, E., Tomasz, A., Hengstler, B. & Zak, O. (1985). The relative role of bacterial cell wall and capsule in the induction of inflammation in pneumococcal meningitis. *J Infect Dis* **151**, 535-540.

Vakevainen, M., Eklund, C., Eskola, J. & Kayhty, H. (2001). Cross-reactivity of antibodies to type 6B and 6A polysaccharides of *Streptococcus pneumoniae*, evoked by pneumococcal conjugate vaccines, in infants. *J Infect Dis* **184**, 789-793.