JAC-2017-1702 Revised Manuscript 1

Molecular Characterization of Predominant *Streptococcus pneumoniae* Serotypes Causing Invasive Infections in Canada: The SAVE Study, 2011-2015

Alyssa R. Golden[1]*, Heather J. Adam[1,2], James A. Karlowsky[1,2], Melanie Baxter[1], Kimberly A. Nichol[2], Irene Martin[3], Walter Demczuk[3], Paul Van Caeseele[1,4], Jonathan B. Gubbay[5], Brigitte Lefebvre[6], Greg Horsman[7], George Zahariadis[8], David Haldane[9], Rita Gad[10], Gregory German[11], Matthew W. Gilmour[1,3], Michael R. Mulvey[1,3], Daryl J. Hoban[1,2], George G. Zhanel[1], on behalf of the Canadian Antimicrobial Resistance Alliance (CARA)[†]

[1] Department of Medical Microbiology, Max Rady College of Medicine, Rady Faculty of Health Sciences, University of Manitoba. 727 McDermot Avenue, Winnipeg, Manitoba, R3E 3P5, Canada

[2] Clinical Microbiology – Health Sciences Centre, Diagnostic Services Manitoba. MS673-820 Sherbrook Street, Winnipeg, Manitoba, R3A 1R9, Canada

[3] National Microbiology Laboratory – Public Health Agency of Canada. 1015 Arlington Street, Winnipeg, Manitoba, R3E 3R2, Canada

[4] Cadham Provincial Laboratory. 750 William Avenue, Winnipeg, Manitoba, R3E 3J7, Canada

[5] Public Health Ontario. 661 University Avenue, Toronto, Ontario, M5G 1M1, Canada

[6] Laboratoire de santé publique du Québec. 20045 Ch Ste-Marie, Ste-Anne-de-Bellevue, Québec, H9X 3R5, Canada

[7] Saskatchewan Disease Control Laboratory. 5 Research Drive, Regina, Saskatchewan, S4S 0A4, Canada

[8] Newfoundland and Labrador Public Health Laboratory. Dr. Leonard A. Miller Centre – Suite 1, 100 Forest Road, St. John's, Newfoundland and Labrador, A1A 1E3, Canada

[9] Queen Elizabeth II Health Science Centre. 5805 South Street, Halifax, Nova Scotia, B3H 1V8, Canada

[10] New Brunswick Department of Health. 520 King Street, Fredericton, New Brunswick, E3B 5G8, Canada

[11] Health PEI. 16 Garfield Street, Charlottetown, Prince Edward Island, C1A 7N8, Canada

*Address for Correspondence: Alyssa Golden, Department of Medical Microbiology, University of Manitoba; MS673-820 Sherbrook Street, Winnipeg, Manitoba, R3A 1R9, Canada.
Phone: 1 (204) 787-4684. Fax: 1 (204) 787-4699. E-mail: umgoldea@myumanitoba.ca

Intended Category: Original article

Running Title: Molecular characterization of *S. pneumoniae* in Canada

[†] Member laboratories are listed in the Acknowledgements section.

Synopsis

Objectives:

This study characterized the eleven most predominant serotypes of invasive *S. pneumoniae* infections collected by the annual SAVE study in Canada, between 2011 and 2015.

Methods:

A subset of the eleven most predominant serotypes (7F, 19A, 22F, 3, 12F, 11A, 9N, 8, 33F, 15A and 6C) collected by the SAVE study were analyzed using PFGE and MLST, as well as PCR to identify pilus-encoding genes. WGS analyses were performed on a subset of the above isolates plus a random selection of background strains.

Results:

Of the predominant serotypes analyzed, 7F, 33F and 19A were obtained more commonly from children less than 6 years of age, while 15A, 6C, 22F and 11A were more common in adults over 65 years. Pneumococcal pilus PI-1 was identified in antimicrobial susceptible serotype 15A (61/212) and <10% of 6C isolates (16/188). PI-2 was found in serotype 7F (683/701) and two-thirds of 11A isolates (162/241). Only serotype 19A-ST320 possessed both pili. Molecular and phylogenetic analyses identified serotypes 19A, 15A, 6C, 9N and 33F as highly diverse, while 7F, 22F and 11A demonstrated clonality. Antimicrobial resistance determinants were common within diverse serotypes, and usually similar within a clonal complex.

Conclusions:

Despite successful use of conjugate vaccines, *S. pneumoniae* remains a highly diverse organism in Canada. Several predominant serotypes, both antimicrobial susceptible and MDR, have demonstrated rapid clonal expansion or an increase in diversity. As *S. pneumoniae* continues to evolve in Canada, WGS will be a necessary component in the ongoing surveillance of antimicrobial-resistant and expanding clones.

Introduction

Streptococcus pneumoniae is a highly diverse organism capable of causing invasive disease in children, older adults and immunocompromised individuals.[1] The primary virulence factor of this pathogen is the polysaccharide capsule, which is crucial for immune evasion; extensive study of the capsule has led to identification of 97 capsular serotypes.[2] Historically, only a small number of these serotypes have accounted for the majority of invasive pneumococcal disease (IPD).[1] Those serotypes most commonly causing invasive disease were targeted by pneumococcal conjugate vaccines (PCV), which were utilized in Canada beginning in 2002 and have led to widespread success, including significant overall decreases in IPD, particularly those due to vaccine serotypes.[3] However, after using PCVs (7-valent, 13-valent) for over a decade in Canada, serotype prevalence has shifted due to both replacement of vaccine types and vaccine escape through capsular switching events.[3–6]

During the period of PCV development and use, antimicrobial and MDR in *S. pneumoniae* has remained a constant concern, escalated by the worldwide dissemination of resistant and MDR international clones. The Pneumococcal Molecular Epidemiology Network (PMEN) was established in 1997 to both standardize the classification and create a global collection of resistant clones. Prior to PCV-7 use, most vaccine serotypes (4, 6B, 9V, 14, 18C, 19F, 23F) had at least one widely disseminated PMEN clone of concern demonstrating antimicrobial resistance and extensive clonal expansion. Of note were penicillin-resistant Spain^{9V}-3, macrolide-resistant England^{14}-9, and MDR Spain^{23F}-1, Spain^{6B}-2 and Taiwan^{19F}-14.[7] While use of PCV-7 reduced the impact of these particular clones, others became increasingly prevalent in the post-PCV-7 era, including Netherlands^{3}-31, Netherlands^{7F}-39 and MDR serotype 19A-ST320 isolates related to Taiwan^{19F}-14.[5,8] More recently in Canada, PCV-13 use has been associated with a greater prevalence of the highly clonal serotype 22F (ST433), as well as MDR isolates related to Sweden^{15A}-25.[6,9,10]

As noted by Klugman in 2002, genetic analysis of these successful international clones has been crucial in understanding the spread of antimicrobial resistance in *S. pneumoniae*.[11] Previously, subtyping methods were used to identify genetic relatedness and genes or mutations associated with resistance. Many laboratory "gold standards" currently rely on subtyping methods such as PFGE or MLST, despite the limited amount of information they provide.[12] In recent years, WGS has become the method of choice to characterize isolates due to the rapidly decreasing cost, short time to completion and unambiguous examination of the total genetic content of a strain at the single nucleotide level.[12]

The *S. pneumoniae* Serotyping and Antimicrobial Susceptibility: Assessment for Vaccine Efficacy in Canada (SAVE) study is an annual study which began in 2011, after PCV-13 introduction in Canada. The purpose of the current study was to characterize the eleven most predominant serotypes (7F, 19A, 22F, 3, 12F, 11A, 9N, 8, 33F, 15A, 6C) collected by the SAVE study using PFGE, MLST and pilus PCR, as well as WGS analyses to determine population structure, phylogenomic relationships and antimicrobial resistance determinants for a subset of these isolates.

Materials and Methods

Bacterial isolates

From January 2011 to December 2015, *S. pneumoniae* isolated from sterile body sites by participating Canadian provincial public health and hospital laboratories were forwarded to the Public Health Agency of Canada-National Microbiology Laboratory (PHAC-NML) in Winnipeg, Canada. As part of an ongoing collaboration between the Canadian Antimicrobial Resistance Alliance (CARA) and PHAC-NML, PHAC-NML forwarded their collection of invasive isolates of *S. pneumoniae* isolates from eight Canadian provincial laboratories (Saskatchewan, Manitoba, Ontario, Quebec, Nova Scotia, Prince Edward Island, Newfoundland and Labrador, and a portion of isolates collected from New Brunswick) to CARA for antimicrobial susceptibility testing. For the SAVE study, regional analysis were conducted as

Western (Saskatchewan and Manitoba), Central (Ontario and Quebec) and Eastern (New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland and Labrador).

In total, 6207 invasive isolates of *S. pneumoniae* collected as part of SAVE study between 2011 and 2015 were forwarded to the CARA for antimicrobial susceptibility testing. Patient gender and age information was available for 5980 (96.3%) and 6072 (97.8%) of the isolates, respectively. The annual numbers of isolates were: 1379 isolates from 2011, 1285 from 2012, 1138 from 2013, 1210 from 2014, and 1195 from 2015.

Antimicrobial susceptibility testing

Antimicrobial susceptibility testing was performed at the Winnipeg Health Sciences Centre (Winnipeg, Manitoba, Canada) using the CLSI standard broth microdilution method[13,14] with custom-designed, in-house prepared, 96-well microtitre panels containing doubling-dilutions of antimicrobial agents in cation-adjusted Mueller-Hinton broth supplemented with 4% lysed horse blood. All isolates were tested against penicillin, ceftriaxone, cefuroxime, clarithromycin, clindamycin, telithromycin, levofloxacin, moxifloxacin, linezolid, trimethoprim/sulfamethoxazole, doxycycline, tigecycline, chloramphenicol and vancomycin. MICs were interpreted as susceptible, intermediate or resistant using CLSI MIC breakpoints.[13] MDR was defined as resistance to three or more classes of antimicrobial agents (penicillin resistance was defined using the CLSI breakpoint for oral penicillin V, MIC ≥2 mg/L). Isolates resistant to five or more classes of antimicrobials were considered extensively drug resistant (XDR).

Serotyping

Serotyping was performed using the Quellung reaction[15] using pool, group, type and factor commercial antisera (Statens Serum Institute, Copenhagen, Denmark). Isolates for which a serotype was not determined by a Quellung reaction were confirmed as *S. pneumoniae* by *rpoB* gene sequencing.[16]

PFGE, MLST and Pilus PCR

Ten randomly selected isolates of each of the eleven most common serotypes per year (7F, 19A, 22F, 3, 12F, 11A, 9N, 8, 33F, 15A, 6C; 50 of each serotype, 550 total isolates) were characterized for genetic relatedness by PFGE and MLST. PFGE was performed as previously described.[17,18] Gels were analysed using BioNumerics® Software (Applied Maths Inc, Austin, TX). Isolates with ≥ 80% relatedness were considered a cluster.

MLST was performed on the same 550 isolates using methods and primers previously described at http://pubmlst.org/spneumoniae. Sequences were analysed using Lasergene® SeqMan Pro (DNAStar, Madison, WI). MLST sequence types (STs) were compared to the Pneumococcal Molecular Epidemiology Network (PMEN) database (http://www.sph.emory.edu/PMEN) to identify commonly circulating clones. Minimum spanning trees were generated using PHYLOViZ 2.0 open-source software.[19]

To assess putative virulence, PCR to determine the presence of pneumococcal pili was performed using previously described primers.[20]

Isolate Selection for WGS

A total of 192 isolates were selected for WGS by the Illumina MiSeq platform. An initial 84 isolates from the SAVE study were specifically selected from the above serotypes due to preliminary characterization that indicated MDR, novel MLST STs and/or the potential to be a capsular switch variant. To achieve broader coverage of the diverse pneumococcal population, 78 additional isolates from SAVE were selected as "background". These isolates were selected using a random number generator and included three of each serotype included in the 23-valent pneumococcal polysaccharide vaccine and up to three of other non-vaccine serotypes to total 78 isolates. To include isolates collected prior to PCV-13 introduction, 30 invasive *S. pneumoniae* isolates randomly selected from 2007-2009 were included. These isolates were collected as part of the BESST study;[21] only isolates collected from the same provinces and source as the other 162 isolates were included. To control one of the many variables, all isolates were selected from the ≥ 65-year age category, as this age group had the largest

and most diverse collection of isolates from which to sample. Overall, 44 different serotypes were represented in this analysis.

WGS Data Analysis

Phylogenomic analysis was performed using SNVPhyl, a PHAC-NML custom-built bioinformatics pipeline[22] and reference genome *S. pneumoniae* R6 (NCBI: NC_003098). Briefly, repeat regions of the reference genome were identified using MUMMer v.3.23[23] and collected into a file of locations to be excluded from further analysis. MUMMer was run using a minimum length of 150 and a minimum percent identity of 90. Reads were then mapped to the reference genome using SMALT v.0.7.4 (http://www.sanger.ac.uk/science/tools/smalt), with a k-mer size of 13 and a step size of 6. Variant calling was performed using both FreeBayes v.0.9.20[24] and SAMtools.[25,26] FreeBayes detected variants using a minimum coverage of 10, a minimum mean mapping quality of 30 and an alternate allele proportion of 0.75. SAMtools was used to confirm the variant calls made by FreeBayes. These SNVs were filtered and merged, as previously described, to construct a multiple sequence alignment.[22] Four isolates mapped poorly to the reference strain, and were removed from further analysis. PhyML v.3.0[27] was used to generate a maximum likelihood phylogenetic tree from this alignment, which was then visualized using FigTree software (v.1.4.3, http://tree.bio.ed.ac.uk/software/figtree). Phylogenetic clades were determined using ClusterPicker v.1.2.3 software using the default parameters.[28]

The presence of acquired resistance genes was determined for SAVE isolates using the ResFinder 2.1 program.[29] This open-source software is freely available from the Center for Genomic Epidemiology (https://cge.cbs.dtu.dk/services/ResFinder/), and identified resistance genes for macrolides (*mefA*, *ermB*, *msrD*), tetracyclines (*tetM*) and chloramphenicol (*cat*). Genes with chromosomal mutations conferring resistance were extracted using NCBI BLAST tools[30,31] and compared to those of *S. pneumoniae* R6. Extracted genes were aligned to the reference sequence using the ClustalW2 multiple sequence alignment program [32]. For the penicillin-binding proteins, DNA sequences

were translated into amino acid sequences and examined for mutations in the active site motifs of *pbp1A* (STMK, SRNVP, KTG), *pbp2B* (SVVK, SSNT, KTGTA) and *pbp2X* (STMK, AHSSNV, LKSGT) as previously described [33,34]. Nucleotide sequences of *parC*, *gyrA*, and *folA/P* were examined for previously described mutations that convey fluoroquinolone and trimethoprim/sulfamethoxazole resistance, respectively [35,36].

Statistical Analysis

Differences in serotype distribution between the various demographic parameters were assessed for statistical significance (P <0.05) using a two-tailed Fisher's exact test (α=0.05).

Results

Isolate Demographics

Among the *S. pneumoniae* isolates collected for the SAVE 2011-15 study, the eleven most predominant serotypes accounted for 65.9% (4092/6207) of isolates (Table 1). The serotype distribution of these eleven serotypes was evaluated by region and age group. When compared to the 868 isolates (21.2%) collected from Western Canada, the proportion of serotypes 12F (n=219, 75.3%; P<0.0001) and 8 (n=140, 28.2%; P=0.0124) collected were significantly higher in the West than the other predominant serotypes. Similarly, when compared to the 2,701 isolates (66%) collected from Central Canada, the proportion of serotypes 15A (172, 78.2%; P<0.0001), 11A (205, 76.2%; P=0.0005), 7F (n=514, 73.0%; P=0.0002) and 22F (n=425, 71.1%; P=0.01) were significantly higher than the other predominant serotypes. Only serotype 19A demonstrated a significantly higher proportion overall in Eastern Canada (17.6% versus 12.8%, P=0.0017). Of note in the age group distribution are those ages where pneumococcal disease is most common, particularly young children and older adults. Significant serotypes in children were 7F, which was more frequently isolated from infants less than one year of age (n=30, 4.3%; P=0.03), 33F, which was more commonly isolated from those one year to less than two years of age (n=20, 8.9%; P<0.0001), and 19A, which was more likely isolated from children aged two

years to less than six years (*n*=53, 8.8%; *P*<0.0001). In adults over the age of 65 (*n*=1,499, 36.6%), there were four predominant serotypes that were significantly more prevalent than the other serotypes of interest. These were serotypes 15A (*n*=126, 57.3%; *P*<0.0001), 6C (*n*=116, 56.6%; *P*<0.0001), 22F (*n*=270, 45.2%; *P*<0.0001) and 11A (*n*=117, 43.5%; *P*=0.0266). In the adult population (18-<50), serotype 12F (*n*=129, 44.3%; *P*<0.0001) and 8 (*n*=77, 32.0%; *P*=0.0002) were most frequently obtained. Except for serotypes 15A and 33F, isolates were more commonly obtained from males.

Genetic Characterization

Pneumococcal Pili

Overall, 3878 isolates with a predominant serotype had full PCR results for both PI-1 and PI-2 (Table 2). Several serotypes had little to no association with either pilus, including 22F, 3, 12F, 9N, 8 and 33F. Isolates with PI-1 genes (serotypes 6C, 15A and 19A) were not commonly associated with a MDR phenotype. Of the 16 serotype 6C isolates (8.5%) that harboured PI-1, only one was MDR; however, two-thirds of the same 16 isolates were resistant to trimethoprim/sulfamethoxazole. Despite being a commonly MDR serotype, 15A isolates containing PI-1 (*n*=61, 28.8%) expressed little resistance and were not associated with a MDR phenotype. Serotype 19A demonstrated just over 5% MDR in isolates containing PI-1 (*n*=257, 44.5%). PI-2 was the more common pilus type overall, likely due to 97.4% (683/701) of serotype 7F isolates (the most common serotype over the study period) possessing PI-2. This was the only instance where almost the entire serotype cohort demonstrated one specific genotype (apart from having no pilus genes). Over half of serotype 11A isolates (*n*=162, 67.2%) possessed PI-2. Neither serotype 7F or 11A demonstrated an appreciable amount of MDR when PI-2 genes were present. The clearest association of pneumococcal pili with MDR was found with the dual PI-1/PI-2 genotype demonstrated by serotype 19A. Of the serotype 19A isolates tested, 21.1% demonstrated the dual genotype and 95.9% of these isolates were MDR or XDR.

PFGE, MLST and WGS

Isolates clustered similarly using both PFGE (data not shown) and MLST (Figure 1). Although the number of isolates typed by WGS was much lower, phylogenomic analysis provided a more in-depth perspective on many of the predominant serotypes and their relatedness to other types (Figure 2). Illumina MiSeq sequencing resulted in an average of 539 336 reads/genome with an average genome coverage of 77x. *De novo* assembly yielded an average contig and N50 length of 59 983 bp and 121 439 bp, respectively.

Serotype 19A, a commonly MDR serotype, was found to be highly variable in this study. Thirteen STs were identified by MLST, indicating high diversity in this group of isolates. The most common types were ST320, a frequently MDR type related to international clone PMEN14, and ST695, associated with susceptibility to all antimicrobials except for clarithromycin. Additionally, a smaller cluster of serotype 19A isolates typed as identical or related to PMEN37, an international clone originally identified in serotype 15B. Of these many STs, the only types to demonstrate the dual PI-1/PI-2 genotype mentioned above were those related to PMEN14. Aside from one large cluster of isolates related to PMEN14, serotype 19A isolates were difficult to pinpoint in the phylogenetic analysis because of their relatedness to several different serotypes, resulting in their distribution throughout many smaller clusters of background isolates. Fourteen of 32 serotype 19A isolates that were sequenced were not ultimately related to PMEN14. However, these isolates were instead related to PMEN clones 1, 3, 9, 21, 25, 30, 32 and 37, and thus related in various degrees to isolates of serotypes 3, 4, 9V, 14, 15A, 15B, 17F, 19F, 21, 23F and 24F. This indicates that serotype 19A isolates likely frequently participate in recombination.

Like serotype 19A, serotype 15A (also frequently MDR) demonstrated numerous STs in this study. Fourteen STs identified by MLST were found to be associated with serotype 15A; however, half of these were identical or related to ST63, a frequently MDR type designated PMEN25. Isolates related to this clone had a common resistance pattern of clarithromycin, clindamycin and doxycycline, with occasional resistance to penicillin or trimethoprim/sulfamethoxazole. The other seven STs were

predominantly susceptible to all antimicrobials and over half of these isolates possessed PI-1 genes. Most serotype 15A isolates clustered together in the phylogenetic analysis, however a few were more highly related to PMEN3 and PMEN30 (serotypes 9V and 21, respectively).

Two other predominant serotypes that demonstrated high diversity were 6C and 9N. Serotype 6C demonstrated 20 different STs by MLST, the most of any serotype in this study. As opposed to a large cluster of closely related types, serotype 6C had a few smaller clusters containing two or three STs, with many others differing by three or more alleles (Figure 1). Despite the large number of STs, only two were related to an international clone: ST338, an international clone (PMEN26) originally associated with serotype 23F, and ST5241, a double-locus variant (DLV) of PMEN26. Much of the MDR in the typed serotype 6C isolates was attributed to this cluster, specifically ST5241, which demonstrated resistance to clarithromycin, clindamycin and doxycycline. Despite being relatively dispersed in the MLST minimum spanning tree, the whole genomes of serotype 6C isolates clustered together, except for the PMEN26-related strain that was more closely related to serogroup 23 background strains. Only one isolate with typing data was positive for PI-1 genes, making it difficult to determine if a pattern exists between PI-1 presence and ST for serotype 6C. Serotype 9N demonstrated 12 STs by MLST, most of which were contained in a large cluster centred on ST66, a single-locus variant (SLV) of PMEN18; phylogeny presented similar results (Figure 1 and 2). Despite this clone being known for its resistance to antimicrobials, most isolates with this type were fully susceptible to all agents. A small number of isolates were variants of PMEN34, commonly associated with serotype 12F. Only one isolate was MDR, with resistance to clarithromycin, clindamycin, doxycycline and penicillin, but it was not related to either international clone.

Serotype 7F, the most commonly isolated serotype in the SAVE study, was also the most clonal of the serotypes studied. All but one isolate typed as ST191 by MLST, identical to the PMEN39 international clone originally isolated in the Netherlands. Interestingly, the ST191 isolates typed by WGS

were the most distantly related to the rest of the population, clustering distantly from all other isolates (Figure 2). One serotype 7F isolate demonstrating resistance to clarithromycin, clindamycin and doxycycline typed as ST63 and clustered accordingly with the MDR 15A isolates in the maximum likelihood tree. Serotype 22F demonstrated similar results to serotype 7F. Over 95% of serotype 22F isolates typed by MLST fell into a cluster founded by ST433. Additionally, in both MLST and WGS there were serotype 22F isolates that clustered with ST63; these isolates typed as ST9352, a SLV of ST63 that demonstrated the common ST63 resistance pattern, with the addition of levofloxacin resistance.

Two serotypes that clustered closely together by both MLST and WGS were 11A and 33F. Serotype 11A was relatively clonal, with most isolates typing as ST62 and just over half exhibiting PI-2 genes. One isolate typed as ST156, identical to commonly PI-1 piliated and antimicrobial resistant international clone PMEN3; the serotype 11A isolate in question also possessed PI-1 genes and demonstrated resistance to penicillin and trimethoprim/sulfamethoxazole. Unlike serotype 11A, serotype 33F demonstrated more diversity in MLST types. Eight STs were identified for serotype 33F, four of which were newly assigned during the study period. Visually, these isolates appeared separated in the minimum spanning tree (Figure 1); however, despite the variability in ST, all serotype 33F isolates clustered closely in the phylogenetic analysis. Interestingly, despite appearing highly related to serotype 11A in the minimum spanning tree, phylogenetic analysis including background serotypes revealed that serotype 33F demonstrated greater relatedness to serotype 18C, a type not studied in detail. The cluster highlighted as serotype 33F in Figure 2 was determined by the ClusterPicker program and is comprised of not only serotype 33F isolates, but also several 18C isolates.

Serotype 3 was an interesting case; by MLST, serotype 3 appeared very clonal. Almost 90% of the isolates tested were ST180 (PMEN31) or a SLV. However, when examining the phylogeny of these isolates, ST180 was comprised of two distinct clades (Figure 2). Clade I was completely susceptible to all antimicrobials, while Clade II demonstrated resistance to chloramphenicol, clarithromycin, doxycycline

297 and often clindamycin. Additionally, one isolated typed as ST177 by both MLST and WGS clustered with
298 other isolates related to PMEN21 (originally represented by serotype 19F).
299 Serotypes 8 and 12F were similar in the fact that they demonstrated two distinct clusters of
300 isolates by MLST – one related to an international clone and the other not. The largest serotype 12F
301 cluster was founded on ST218, an international clone from Denmark (PMEN34), and isolates were often
302 resistant to clarithromycin. In sharp contrast, the other cluster shared no MLST alleles in common with
303 other isolates of the same serotype and was MDR, with resistance to clarithromycin, clindamycin and
304 doxycycline. Phylogenetic analysis determined that ST218 isolates were more related to other
305 international clones, particularly PMEN4, while the MDR isolates clustered near several background
306 serotypes, including 31, 10A and 35F. Unlike serotype 12F, the largest cluster of serotype 8 isolates was
307 not related to an international clone (PMEN33, ST53). Less than 20% of isolates typed by MLST were
308 part of this cluster and no isolates with a sequenced whole genome demonstrated a type related to this
309 clone. Most serotype 8 isolates clustered around founder ST1480, a type that is six alleles different from
310 ST53. Isolates clustered similarly in the maximum likelihood tree, with 38 and 22F representing the most
311 closely related serotypes. Only one serotype 8 isolate tested had an MDR phenotype; this ST63 isolated
312 clustered accordingly in the phylogenetic analysis and had the characteristic ST63 resistance pattern as
313 described above.
314 ***Antimicrobial Resistance Determinants***
315 Of the 162 isolates from SAVE with whole genomes available, 27 (16.7%) demonstrated
316 discrepancies between genotype and phenotype for one or more antimicrobials. After repeating the
317 susceptibility testing in triplicate, two isolates would not grow for repeat testing and were therefore
318 removed from the analysis of resistance genes, leaving 160 SAVE genomes available for this analysis.
319 After retesting, four discrepancies remained. The genes and mutations identified are outlined in Table 3.
320 *Acquired Resistance Genes*

Both *mefA* and *ermB* macrolide resistance genes were common amongst tested isolates. Overall, *ermB* (*n*=48, 30.0%) was more commonly identified than *mefA* (*n*=17, 10.6%) and was found in a larger number of serotypes. Both genes conferred high rates of nonsusceptibility to clarithromycin (94.1 and 97.9% for *mefA* and *ermB*, respectively); *ermB* also provided 81.2% of isolates with nonsusceptibility to clindamycin. Of the two determinants on their own, *ermB* was more commonly associated with MDR isolates. The dual *mefA*+*ermB* genotype was only present in serotype 19A isolates, specifically ST320 isolates demonstrating an XDR phenotype. However, this dual genotype did not necessarily confer full resistance to macrolides and lincosamides, with isolates demonstrating 94.4% and 61.1% nonsusceptibility, respectively. Upon further inspection, six of 18 isolates that contained the dual *mefA*+*ermB* genotype but were clarithromycin susceptible, intermediate, or demonstrated only low-level resistance (1 mg/L) contained a truncated version of the *ermB* gene caused by a premature stop codon at base 642.

The *tetM* gene associated with tetracycline resistance was present in 40.0% (*n*=64) of isolates and conferred 93.7% nonsusceptibility to doxycycline. Four isolates carried an intact *tetM* gene but were not resistant to doxycycline, perhaps indicating a nonfunctional gene. Presence of *tetM* was associated with 19 different serotypes and a large proportion of these isolates were MDR (89.1%). Only a small number of isolates, predominantly serotype 3, Clade II, carried the *cat* gene (*n*=8, 5.0%). However, possession of this gene was consistently associated with resistance to chloramphenicol (100% resistance) and all isolates carrying *cat* were also MDR.

Chromosomal Mutations

Alterations in key motifs of PBPs were discovered in 75/160 (46.9%) isolates with whole genomes available (Table 3), particularly in serotypes 19A and 15A. The most common single alteration was a lone Thr451Ala mutation in the SSNT motif of *pbp2B*; this mutation was found in a variety of serotypes and was associated with increased penicillin-nonsusceptibility. Conversely, isolates with solely

a *pbp2X* alteration were fully susceptible to penicillin. No isolate was found to contain mutations in *pbp1A* alone. The highest penicillin MICs were most commonly associated with mutations in all three PBPs (100% nonsusceptibility overall, Table 3). The most common set of alterations was Thr371Ser in STMK and Pro432Thr in SRNVP of *pbp1A*, Thr451Ala in SSNT and Ala624Gly in KTGTA of *pbp2B* and Thr338Ala in STMK and Leu546Val in LKSGT of *pbp2X*. This pattern of alterations was exclusively associated with ST320-19A isolates with MDR/XDR phenotypes.

Mutations in the quinolone-resistance determining regions (QRDR) of *parC* and *gyrA* were uncommon, with only 14/160 isolates (8.8%) demonstrating alterations (Table 3). Three isolates exhibited alterations in the QRDR of *parC*, two with Ser79Phe and one with Ser79Tyr. These mutations conferred 66.7% nonsusceptibility to levofloxacin, the fluoroquinolone that preferentially targets *parC*. Similarly, three isolates were identified with Ser81Phe mutations in *gyrA*. These isolates demonstrated 33.3% susceptibility to moxifloxacin, which preferentially targets *gyrA*. Eight isolates were determined to have mutations in the QRDR regions of both *parC* and *gyrA*. Four isolates contained Ser79Phe and Ser81Phe mutations in *parC* and *gyrA*, respectively, two contained Ser79Tyr and Ser81Phe and two contained Ser79Phe and Ser81Leu. This last set of isolates were the ST9352-22Fs previously discussed as being related to ST63. Overall, half of the isolates demonstrating a QRDR mutation were also MDR or XDR (Table 3). Other than serotype 22F as mentioned above, mutation patterns in *parC* and *gyrA* were not specific to serotype. However, mutations in one or both genes were essential for fluoroquinolone resistance as isolates with neither mutation were fully susceptible to both levofloxacin and moxifloxacin.

Mutations in *folA* and localized insertions in *folP* were identified in 47/160 isolates (29.4%) (Table 3). Although none of these isolates exhibited the Ile100Leu mutation in *folA* alone, ten isolates contained solely a *folP* insertion and provided either intermediate resistance or susceptibility to trimethoprim/sulfamethoxazole. Dual alterations of both *folA* and *folP* were more commonly identified than a single mutation (37/47). The combination of both mutations conferred 100% nonsusceptibility to

trimethoprim/sulfamethoxazole and were more commonly associated with MDR (83.8%). Interestingly, there were seven different *folP* insertions of one or two amino acids between codons 59 and 69 that were associated with varying levels of resistance, MDR and serotype specificity. The most common insertion consisted of an extra Glu-Ile (EI) after codon 66 of *folP*. This alteration was present in combination with *folA*-Ile100Leu in 17 ST320-19A isolates, was associated with the highest trimethoprim/sulfamethoxazole MICs, and isolates were either MDR or XDR. These serotype 19A isolates were the same as those mentioned above which demonstrated the most common set of alterations in all three PBPs.

Discussion

The use of conjugate vaccines to combat IPD has resulted in dramatic shifts in serotype distribution throughout Canada.[4] This study characterized the most predominant serotypes circulating in Canada in the years following PCV-13 introduction. The findings of this study indicate that *S. pneumoniae* remains a highly diverse organism, with several serotypes, both susceptible and MDR, demonstrating either clonal expansion or an increase in diversity.

This study identified associations between several predominant serotypes and vaccine-eligible age groups: serotypes 7F, 19A and 33F in children less than six, and 15A, 6C, 22F and 11A in adults over the age of 65. A recent review of worldwide serotype distribution data corroborated many of these findings, with slight variations based on geographic location and specimen source.[37] In particular, serotypes 6C and 33F were more commonly noted here than in the systematic review. While serotype 33F has been noted for its high invasive capacity in children, 6C has been more commonly isolated as a nasopharyngeal carrier with a lower capacity for invasion.[38] The high prevalence of serotype 6C IPD isolated from adults over 65, particularly in Central Canada, may indicate the occurrence of an outbreak, and that older adults are more vulnerable to IPD caused by this serotype than children.

Interestingly, serotypes 12F and 8 were common in adults aged 18-<50. Serotype 12F was most common in Western Canada, where 75.3% of 12F isolates were collected. A study performed in Alaska noted that this serotype is not normally associated with nasopharyngeal carriage in healthy people, but is instead a common cause of IPD outbreaks;[39] reports have described outbreaks of serotype 12F IPD in the United States and most recently, Winnipeg, Manitoba, Canada in 2008-11.[39–41] As serotype 12F collection has remained high in Western Canada since this time, it is possible that the outbreak has continued. Interestingly, two serotype 12F isolates analysed in the current study were MDR, as opposed to solely possessing macrolide resistance like many ST218 isolates.[41] As serotype 12F is prone to causing outbreaks, these new MDR clones are of particular interest for future study. Serotype 8 was also common in adults aged 18-<50, and has been similarly noted in Spain to be generally susceptible to antimicrobials and largely isolated from adult patients. Interestingly, Spain has experienced clonal expansion of MDR serotype 8 strains related to PMEN25 (ST63).[42] These isolates were resistant to macrolides, lincosamides, tetracyclines, and fluoroquinolones; initially restricted to HIV-positive patients in Madrid, this clone spread through adults in nine other regions.[42] One such isolate was collected and analyzed in the current study; however, this isolate did not demonstrate fluoroquinolone resistance. MDR serotype 8-ST63 should be monitored closely in Canada as it could become a strain of concern in normally healthy adult patients.

In general, serotypes that demonstrated high levels of diversity in this study were also those that had the highest rates of MDR and demonstrated strong associations with a specific resistance pattern.[43] Serotypes 6C, 15A, 19A and 33F were commonly resistant to antimicrobials, with resistance mediated by the acquisition of foreign resistance determinants, specifically *ermB* and *tetM.* A recent study by Croucher *et al*. identified a correlation between serotype diversity and the total number of recombination events experienced;[44] the diversity of the above serotypes would indicate high recombination frequencies and thus increased chances of obtaining acquired resistance determinants.

Of the diverse serotypes noted above, 19A appeared to participate most frequently in recombination involving the capsule. In this study alone, serotype 19A was the donor strain for multiple putative recombinations involving serotypes 14, 15A, 15B, 19F, 21 and 23F. Several whole genome studies of *S. pneumoniae* in other countries have noted similar recombinations, including Bulgaria, Germany, Russia and the United States.[45–48] Importantly, penicillin resistance, a particularly common trait of serotype 19A, is mediated by point mutations in PBP genes located on either side of the capsular polysaccharide operon of *S. pneumoniae*. A study of isolates from an East Asian population noted that recombination has facilitated the consistent spread of β-lactam resistance amongst the pneumococcal population. Similarly, *folA* genes demonstrated manifestations of recombination; interestingly, there was no association between *folP* insertions and recombination.[49]

Despite the variability within serotype 19A strains, most studies have focussed on the ST320 clone. A Canadian study followed the development of serotype 19A from 1993-2008 and concluded that the emergence of ST320 was the combinatory result of vaccine selection pressure, antimicrobial pressure and the propensity of *S. pneumoniae* to undergo recombination.[5] This strain was originally identified as a vaccine escape recombinant in the post-PCV-7 era and has continued to be a successful clone well into the use of PCV-13, despite serotype 19A being included in the formulation.[45] Approximately 25% of serotype 19A isolates typed by MLST in the current study were ST320. Despite possessing the dual *mefA/ermB* genotype, *tetM,* alterations in all three PBPs, *folA* mutations, *folP* insertions and both pneumococcal pili, ST320 isolates collected in this study were rarely fluoroquinolone resistant. The high fitness cost associated with mutations in DNA replication enzymes ensures that, at least for the time being, ST320 isolates are susceptible to at least one antimicrobial class.

One of the most clonal serotypes described in this study was 7F; almost all isolates tested were ST191 or a variant related to PMEN39 and few were resistant to antimicrobials. Once the most common serotype isolated by the SAVE study, use of PCV-13 resulted in a dramatic decrease in prevalence of

serotype 7F.[43,50] Studies have estimated the specific PCV-13 vaccine effectiveness for serotype 7F to be over 90%;[51] the clonal nature of this serotype, and thus the lack of serotype variability may have contributed to the success of opsonophagocytic killing of serotype 7F in PCV-13. A similar serotype of interest in this study was 22F, ranking as the most commonly collected serotype in SAVE 2015.[43] The predominant serotype 22F clone was ST433, a finding that has been noted in many other countries, including Japan, Sweden and the United States.[52–54] Many serotype 22F isolates were resistant to clarithromycin;[43] in this study, resistance was found to be mediated by either *mefA* or *ermB*, though a similar Canadian study noted *mefA* to be the most common macrolide resistance determinant in serotype 22F.[10] As serotype 22F shares many properties with vaccine-success serotype 7F, it is possible that 22F will react similarly to vaccine use when PCV-15 (PCV-13 plus 22F and 33F) becomes available.

A type that demonstrated little diversity by MLST was serotype 3, with most isolates belonging to the predominant ST180 clonal complex (PMEN31). However, phylogenetic analysis of serotype 3 revealed two different clades of isolates with specific antimicrobial resistance patterns. A recent phylogenetic analysis of a small international collection of CC180 isolates indicated that most were unaffected by recombination, having little diversity and appearing “frozen” from an evolutionary standpoint (clade I). However, other CC180 isolates in this collection exhibited significant accumulation of genetic variation, although little antimicrobial resistance was seen (clade II).[55] In the current study, ST180 isolates belonging to clade II often possessed three acquired resistance determinants conferring resistance to four different antimicrobials; notably, the clade II group was the only cluster of isolates in this study to consistently possess *cat*. As the previously discussed study collected isolates from 1993-2007, it is possible that the later collection date of the clade II isolates in this study (2011-14) allowed increased time to acquire resistance genes through recombination events. A more recent study by Azarian *et al*. included isolates collected from 24 different countries from 1993-2014. It was determined that 19% of CC180 isolates belonged to clade II and that approximately 26% of clade II isolates

possessed *ermB* and *tetM*, in comparison to one-half of ST180 isolates belonging to clade II and 100% possessing *ermB*/*tetM* in the current study.[56] Interestingly, this study also determined that clade I and clade II ST180 isolates differed in their surface protein antigens, most notably pspA. Clade I isolates possessed family 2 pspA variants, while clade II isolates possessed family 1 variants.[56] The prevalence of clade II isolates, frequent antimicrobial resistance and different antigen profiles indicates the need for additional screening of serotype 3 isolates in Canada. However, as MLST does not discriminate between isolates of the same ST, WGS will be crucial in separating these very different clades of ST180.

The presence of pneumococcal pili has been previously described as a clonal property.[20,57] In this study, there was a clear correlation between pilus presence and several predominant clones. This included serotype 7F-ST191 (PI-2), 19A-ST416 and ST695 (PI-1) and 19A-ST320 (PI-1 and PI-2), all of which have been previously noted in studies performed in Italy, Portugal and the United States.[20,35,57,58] Though other studies have also observed the lack of pili in MDR serotype 15A-ST63 isolates,[35] this is one of few studies to note that susceptible 15A-ST58 isolates often possess PI-1. Interestingly, this study had comparatively more PI-2 piliated serotype 11A isolates than a recent Active Bacterial Core surveillance study in the United States.[35] The American study found that only 38% of serotype 11A-ST62 isolates contained PI-2 in comparison to over 60% in the current study. In general, piliation in ST62 isolates has been variable depending on the study; Zahner *et al*. noted that this variability in PI-2 presence indicates that piliation is not essential for serotype 11A to cause invasive disease.[20] A recent study illustrated an overall decline in pneumococcal pilus frequency, as many piliated types were contained in PCV-7.[59] It is reasonable to assume that the frequency of pili will decrease even more with PCV-13 use, as the prevalence of disease caused by commonly isolated and piliated serotypes 7F and 19A should decrease after a number of years.

This study is limited by the lack of participation of all Canadian provinces. As Alberta and British Columbia do not submit isolates, the regional analyses may be skewed due to underrepresentation of

the Western region. Additionally, submission of IPD isolates to the PHAC-NML is voluntary and passive, which restricts the reporting of incidence data. Lastly, the small sample size included in the WGS analysis is only a very small portion of *S. pneumoniae* isolates collected by the SAVE study. Inclusion of more isolates of interest and more background strains would allow for better representation of the breadth of genetic diversity in the Canadian pneumococcal population.

The observations made in this study indicate that *S. pneumoniae* is a pathogen of high genetic variability, and therefore worthy of further genetic surveillance. *S. pneumoniae* has demonstrated the capacity to propagate highly successful clones, such as ST320, ST433 and ST191, while also participating in frequent recombination to increase genetic diversity and spread antimicrobial resistance genes. Importantly, this study illustrated the increased ability of WGS to discriminate between closely related isolates, in comparison to PFGE and MLST. As *S. pneumoniae* continues to evolve in Canada, WGS will be crucial to differentiate virulent clones and outbreak strains and in the ongoing surveillance of antimicrobial resistance.

Acknowledgements

We would like to thank the following Canadian Public Health Laboratory Network (CPHLN) laboratories for their participation in this study: Saskatchewan Disease Control Laboratory; Regina, Saskatchewan; Cadham Provincial Laboratory, Winnipeg, Manitoba; Public Health Ontario Laboratories, Toronto, Ontario; Laboratoire de santé publique du Quebec, Ste-Anne-de-Bellevue; Queen Elizabeth II Health Science Centre, Halifax, Nova Scotia; New Brunswick Regional Hospitals; Queen Elizabeth Hospital, Charlottetown, Prince Edward Island; and Newfoundland Public Health Laboratory, St. John's, Newfoundland. We would also like to thank PHAC-NML and the members of CARA for their efforts and support of this project.

The data in this paper were previously presented in part at the ASM Microbe meeting, June 1-5, 2017 in New Orleans, LA.

Funding

This work was supported, in part, by the University of Manitoba; Diagnostic Services Manitoba; the National Microbiology Laboratory; Pfizer Canada; and Merck Canada Inc.

Transparency Declaration

G. G. Z. has received research grants from Pfizer Canada and Merck Canada. All other authors have no declarations.

Disclaimer

The opinions expressed in this paper are those of the authors, and do not necessarily represent those of Pfizer Canada or Merck Canada Inc.

References

1. Lynch J, Zhanel G. *Streptococcus pneumoniae*: epidemiology, risk factors, and strategies for prevention. *Semin Respir Crit Care Med* 2009; **30**: 189–209.

2. Geno KA, Gilbert GL, Song JY, *et al.* Pneumococcal capsules and their types: past, present, and future. *Clin Microbiol Rev* 2015; **28**: 871–99.

3. Bettinger JA, Scheifele DW, Kellner JD, *et al.* The effect of routine vaccination on invasive pneumococcal infections in Canadian children, Immunization Monitoring Program, Active 2000-2007. *Vaccine* 2010; **28**: 2130–6.

4. Demczuk WHB, Martin I, Griffith A, *et al.* Serotype distribution of invasive *Streptococcus pneumoniae* in Canada after the introduction of the 13-valent pneumococcal conjugate vaccine, 2010–2012. *Can J Microbiol* 2013; **59**: 778–88.

5. Pillai DR, Shahinas D, Buzina A, *et al.* Genome-wide dissection of globally emergent multi-drug resistant serotype 19A *Streptococcus pneumoniae*. *BMC Genomics* 2009; **10**: 642.

6. Golden AR, Adam HJ, Gilmour MW, *et al.* Assessment of multidrug resistance, clonality and virulence in non-PCV-13 *Streptococcus pneumoniae* serotypes in Canada, 2011-13. *J Antimicrob Chemother* 2015; **70**: 1960–4.

7. McGee L, McDougal L, Zhou J, *et al.* Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae* defined by the pneumococcal molecular epidemiology network. *J Clin Microbiol* 2001; **39**: 2565–71.

8. Beall B, McEllistrem MC, Gertz RE, *et al.* Pre- and postvaccination clonal compositions of invasive pneumococcal serotypes for isolates collected in the United States in 1999, 2001, and 2002. *J Clin Microbiol* 2006; **44**: 999–1017.

9. Golden AR, Adam HJ, Zhanel GG. Invasive *Streptococcus pneumoniae* in Canada, 2011-2014: Characterization of new candidate 15-valent pneumococcal conjugate vaccine serotypes 22F and 33F.

Vaccine 2016; **34**: 2527–30.

10. Demczuk WHB, Martin I, Hoang L, *et al.* Phylogenetic analysis of emergent *Streptococcus pneumoniae* serotype 22F causing invasive pneumococcal disease using whole genome sequencing. *PLoS One* 2017; **12**: e0178040.

11. Klugman KP. The successful clone: the vector of dissemination of resistance in *Streptococcus pneumoniae*. *J Antimicrob Chemother* 2002; **50**: 1–6.

12. Gilmour MW, Graham M, Reimer A, *et al*. Public health genomics and the new molecular epidemiology of bacterial pathogens. *Public Health Genomics* 2013; **16**: 25–30.

13. Clinical and Laboratory Standards Institute. *Performance Standards for Antimicrobial Susceptibility Testing. Twenty-seventh Informational Supplement M100-S27*. CLSI, Wayne, PA, USA; 2017.

14. Clinical and Laboratory Standards Institute. *Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically: Tenth Edition M07-A10*. CLSI, Wayne, PA, USA; 2015.

15. Austrian R. The Quellung reaction, a neglected microbiologic technique. *Mt Sinai J Med* **43**: 699–709.

16. Drancourt M, Roux V, Fournier P-E, *et al*. *rpoB* gene sequence-based identification of aerobic Gram-positive cocci of the genera Streptococcus, Enterococcus, Gemella, Abiotrophia, and Granulicatella. *J Clin Microbiol* 2004; **42**: 497–504.

17. Louie M, Louie L, Papia G, *et al*. Molecular analysis of the genetic variation among penicillin-susceptible and penicillin-resistant *Streptococcus pneumoniae* serotypes in Canada. *J Infect Dis* 1999; **179**: 892–900.

18. McEllistrem MC, Stout JE, Harrison LH. Simplified protocol for pulsed-field gel electrophoresis analysis of *Streptococcus pneumoniae*. *J Clin Microbiol* 2000; **38**: 351–3.

19. Nascimento M, Sousa A, Ramirez M, *et al*. PHYLOViZ 2.0: providing scalable data integration and visualization for multiple phylogenetic inference methods. *Bioinformatics* 2017; **33**: 128–9.

20. Zähner D, Gudlavalleti A, Stephens DS. Increase in pilus islet 2-encoded pili among *Streptococcus*

574 *pneumoniae* isolates, Atlanta, Georgia, USA. *Emerg Infect Dis* 2010; **16**: 955–62.
575 21. Adam HJ, Karlowsky JA, Nichol KA, *et al.* Baseline epidemiology of *Streptococcus pneumoniae*
576 serotypes in Canada prior to the introduction of the 13-valent pneumococcal vaccine. *Microb Drug*
577 *Resist* 2012; **18**: 176–82.
578 22. Petkau A, Mabon P, Sieffert C, *et al.* SNVPhyl: a single nucleotide variant phylogenomics pipeline for
579 microbial genomic epidemiology. *Microb Genomics* 2017; **3**, doi: 10.1099/mgen.0.000116.
580 23. Kurtz S, Phillippy A, Delcher AL, *et al.* Versatile and open software for comparing large genomes.
581 *Genome Biol* 2004; **5**: R12.
582 24. Garrison E, Marth G. Haplotype-based variant detection from short-read sequencing. 2012:
583 arXiv:1207.3907.
584 25. Li H. A statistical framework for SNP calling, mutation discovery, association mapping and population
585 genetical parameter estimation from sequencing data. *Bioinformatics* 2011; **27**: 2987–93.
586 26. Li H, Handsaker B, Wysoker A, *et al.* The Sequence Alignment/Map format and SAMtools.
587 *Bioinformatics* 2009; **25**: 2078–9.
588 27. Guindon S, Dufayard JF, Lefort V, *et al*. New algorithms and methods to estimate maximum-
589 likelihood phylogenies: Assessing the performance of PhyML 3.0. *Syst Biol* 2010; **59**: 307–21.
590 28. Ragonnet-Cronin M, Hodcroft E, Hué S, *et al.* Automated analysis of phylogenetic clusters. *BMC*
591 *Bioinformatics* 2013; **14**: 317.
592 29. Zankari E, Hasman H, Cosentino S, *et al.* Identification of acquired antimicrobial resistance genes. *J*
593 *Antimicrob Chemother* 2012; **67**: 2640–4.
594 30. Camacho C, Coulouris G, Avagyan V, *et al.* BLAST+: architecture and applications. *BMC Bioinformatics*
595 2009; **10**: 421.
596 31. Cock PJA, Grüning BA, Paszkiewicz K, *et al*. Galaxy tools and workflows for sequence analysis with
597 applications in molecular plant pathology. *PeerJ* 2013; **1**: e167.

598 32. Larkin MA, Blackshields G, Brown NP, *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* 2007;
599 **23**: 2947–8.
600 33. Nichol KA, Zhanel GG, Hoban DJ. Penicillin-binding protein 1A, 2B, and 2X alterations in Canadian
601 isolates of penicillin-resistant *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 2002; **46**: 3261–
602 4.
603 34. Hakenbeck R, Brückner R, Denapaite D, *et al*. Molecular mechanisms of β-lactam resistance in
604 *Streptococcus pneumoniae*. *Future Microbiol* 2012; **7**: 395–410.
605 35. Metcalf BJ, Gertz RE, Gladstone RA, *et al.* Strain features and distributions in pneumococci from
606 children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA.
607 *Clin Microbiol Infect* 2016; **22**: 60.e9-60.e29.
608 36. Brueggemann AB, Coffman SL, Rhomberg P, *et al.* Fluoroquinolone resistance in *Streptococcus*
609 *pneumoniae* in United States since 1994-1995. *Antimicrob Agents Chemother* 2002; **46**: 680–8.
610 37. Cui YA, Patel H, O'Neil WM, *et al*. Pneumococcal serotype distribution: A snapshot of recent data in
611 pediatric and adult populations around the world. *Hum Vaccines Immunother* 2017; **13**: 1229–41.
612 38. Yildirim I, Hanage WP, Lipsitch M, *et al.* Serotype specific invasive capacity and persistent reduction
613 in invasive pneumococcal disease. *Vaccine* 2010; **29**: 283–8.
614 39. Zulz T, Wenger JD, Rudolph K, *et al.* Molecular characterization of *Streptococcus pneumoniae*
615 serotype 12F isolates associated with rural community outbreaks in Alaska. *J Clin Microbiol* 2013; **51**:
616 1402–7.
617 40. Schillberg E, Isaac M, Deng X, *et al.* Outbreak of invasive *Streptococcus pneumoniae* serotype 12F
618 among a marginalized inner-city population in Winnipeg, Canada, 2009-2011. *Clin Infect Dis* 2014; **59**:
619 651–7.
620 41. Deng X, Peirano G, Schillberg E, *et al.* Whole-genome sequencing reveals the origin and rapid
621 evolution of an emerging outbreak strain of *Streptococcus pneumoniae* 12F. *Clin Infect Dis* 2016; **62**:

1126–32.

42. Ardanuy C, De La Campa AG, García E, *et al.* Spread of *Streptococcus pneumoniae* serotype 8-ST63 multidrug-resistant recombinant clone, Spain. *Emerg Infect Dis* 2014; **20**: 1848–56.

43. Adam HJ, Golden AR, Karlowsky JA, *et al.* Analysis of multi-drug resistance in the predominant *Streptococcus pneumoniae* serotypes in Canada: the SAVE study, 2011-2015. *J Antimicrob Chemother* 2017: In Press.

44. Croucher NJ, Kagedan L, Thompson CM, *et al.* Selective and genetic constraints on pneumococcal serotype switching. *PLoS Genet* 2015; **11**: e1005095.

45. Croucher NJ, Chewapreecha C, Hanage WP, *et al.* Evidence for soft selective sweeps in the evolution of pneumococcal multidrug resistance and vaccine escape. *Genome Biol Evol* 2014; **6**: 1589–602.

46. Mayanskiy N, Savinova T, Alyabieva N, *et al.* Antimicrobial resistance, penicillin-binding protein sequences, and pilus islet carriage in relation to clonal evolution of *Streptococcus pneumoniae* serotype 19A in Russia, 2002–2013. *Epidemiol Infect* 2017; **145**: 1–12.

47. Setchanova LP, Alexandrova A, Dacheva D, *et al*. Dominance of multidrug-resistant Denmark(14)-32 (ST230) clone among *Streptococcus pneumoniae* serotype 19A isolates causing pneumococcal disease in Bulgaria from 1992 to 2013. *Microb Drug Resist* 2015; **21**: 35–42.

48. Makarewicz O, Lucas M, Brandt C, *et al.* Whole genome sequencing of 39 invasive *Streptococcus pneumoniae* sequence type 199 isolates revealed switches from serotype 19A to 15B. *PLoS One* 2017; **12**: e0169370.

49. Chewapreecha C, Harris SR, Croucher NJ, *et al.* Dense genomic sampling identifies highways of pneumococcal recombination. *Nat Genet* 2014; **46**: 305–9.

50. Streptococcus and STI Unit, Bacteriology and Enteric Diseases Program NML. National Laboratory Surveillance of Invasive Streptococcal Disease in Canada - Annual Summary 2014. Available at: https://www.canada.ca/en/public-health/services/publications/drugs-health-products/national-

646 laboratory-surveillance-invasive-streptococcal-disease-canada-annual-summary-2014.html.
647 51. Andrews NJ, Waight PA, Burbidge P, *et al.* Serotype-specific effectiveness and correlates of
648 protection for the 13-valent pneumococcal conjugate vaccine: A postlicensure indirect cohort study.
649 *Lancet Infect Dis* 2014; **14**: 839–46.
650 52. Nakano S, Fujisawa T, Ito Y, *et al.* Serotypes, antimicrobial susceptibility, and molecular epidemiology
651 of invasive and non-invasive *Streptococcus pneumoniae* isolates in paediatric patients after the
652 introduction of 13-valent conjugate vaccine in a nationwide surveillance study condu. *Vaccine* 2016; **34**:
653 67–76.
654 53. Chang Q, Stevenson AE, Croucher NJ, *et al.* Stability of the pneumococcal population structure in
655 Massachusetts as PCV13 was introduced. *BMC Infect Dis* 2015; **15**: 68.
656 54. Galanis I, Lindstrand A, Darenberg J, *et al.* Effects of PCV7 and PCV13 on invasive pneumococcal
657 disease and carriage in Stockholm, Sweden. *Eur Respir J* 2016; **47**: 1208–18.
658 55. Croucher NJ, Mitchell AM, Gould KA, *et al.* Dominant role of nucleotide substitution in the
659 diversification of serotype 3 pneumococci over decades and during a single infection. *PLoS Genet* 2013;
660 **9**: e1003868.
661 56. Azarian T, Mitchell P, Pollard A, *et al.* Global emergence and population dynamics of divergent
662 serotype 3 ST180 pneumococci post-pneumococcal conjugate vaccine (PCV). In: *ASM Microbe 2017*.
663 New Orleans, Louisiana, 2017. Poster ID: Saturday-EES LB5.
664 57. Aguiar SI, Serrano I, Pinto FR, *et al*. The presence of the pilus locus is a clonal property among
665 pneumococcal invasive isolates. *BMC Microbiol* 2008; **8**: 41.
666 58. Del Grosso M, Camilli R, D'Ambrosio F, *et al.* Increase of pneumococcal serotype 19A in Italy is due to
667 expansion of the piliated clone ST416/CC199. *J Med Microbiol* 2013; **62**: 1220–5.
668 59. Aguiar SI, Melo-Cristino J, Ramirez M. Use of the 13-valent conjugate vaccine has the potential to
669 eliminate pilus carrying isolates as causes of invasive pneumococcal disease. *Vaccine* 2012; **30**: 5487–90.

670

671

672 **Table 1:** Demographic information for the eleven most predominant serotypes collected by the SAVE 2011-2015 study.

Serotype (n)	n (%) of Isolates Associated with Demographic													
	Region			Age Group (years)								Gender		
	West	Central	East	<1	1 - <2	2 - <6	6 - <18	18 - <50	50 - <65	≥65	NG[a]	Male	Female	NG[a]
7F (704)	89 (12.6)	514 (73.0)	101 (14.3)	30 (4.3)	4 (0.6)	22 (3.1)	44 (6.3)	216 (30.7)	185 (26.3)	184 (26.1)	19 (2.7)	368 (52.3)	314 (44.6)	22 (3.1)
19A (603)	100 (16.6)	397 (65.8)	106 (17.6)	11 (1.8)	27 (4.5)	53 (8.8)	21 (3.5)	107 (17.7)	164 (27.2)	198 (32.8)	22 (3.6)	306 (50.7)	276 (45.8)	21 (3.5)
22F (598)	88 (14.7)	425 (71.1)	85 (14.2)	17 (2.8)	25 (4.2)	26 (4.3)	7 (1.2)	92 (15.4)	144 (24.1)	270 (45.2)	17 (2.8)	305 (51.0)	268 (44.8)	25 (4.2)
3 (494)	94 (19.0)	337 (68.2)	63 (12.8)	13 (2.6)	6 (1.2)	23 (4.7)	9 (1.8)	75 (15.2)	143 (28.9)	202 (40.9)	23 (4.7)	253 (51.2)	221 (44.7)	20 (4.0)
12F (291)	219 (75.3)	63 (21.6)	9 (3.1)	6 (2.1)	7 (2.4)	4 (1.4)	9 (3.1)	129 (44.3)	74 (25.4)	57 (19.6)	5 (1.7)	154 (52.9)	128 (44.0)	9 (3.1)
11A (269)	46 (17.1)	205 (76.2)	18 (6.7)	5 (1.9)	14 (5.2)	4 (1.5)	7 (2.6)	43 (16.0)	77 (28.6)	117 (43.5)	2 (0.7)	136 (50.6)	127 (47.2)	6 (2.2)
9N (242)	43 (17.8)	160 (66.1)	39 (16.1)	5 (2.1)	3 (1.2)	1 (0.4)	4 (1.7)	41 (16.9)	94 (38.8)	84 (34.7)	10 (4.1)	128 (52.9)	108 (44.6)	6 (2.5)
8 (241)	68 (28.2)	140 (58.1)	33 (13.7)	8 (3.3)	2 (0.8)	2 (0.8)	10 (4.1)	77 (32.0)	72 (29.9)	64 (26.6)	6 (2.5)	140 (58.1)	91 (37.8)	10 (4.1)
33F (225)	47 (20.9)	157 (69.8)	21 (9.3)	7 (3.1)	20 (8.9)	8 (3.6)	8 (3.6)	39 (17.3)	55 (24.4)	81 (36.0)	7 (3.1)	98 (43.6)	119 (52.9)	8 (3.6)
15A (220)	31 (14.1)	172 (78.2)	17 (7.7)	4 (1.8)	12 (5.5)	8 (3.6)	0	28 (12.7)	38 (17.3)	126 (57.3)	4 (1.8)	96 (43.6)	120 (54.5)	4 (1.8)
6C (205)	43 (21.0)	131 (63.9)	31 (15.1)	3 (1.5)	5 (2.4)	6 (2.9)	3 (1.5)	26 (12.7)	43 (21.0)	116 (56.6)	3 (1.5)	119 (58.0)	81 (39.5)	5 (2.4)
Total (4092)	868 (21.2)	2701 (66.0)	523 (12.8)	109 (2.7)	125 (3.1)	157 (3.8)	122 (3.0)	873 (21.3)	1089 (26.6)	1499 (36.6)	118 (2.9)	2103 (51.4)	1853 (45.3)	136 (3.3)

673 [a] NG, information not given.

674

675

676 **Table 2:** Pneumococcal pilus presence demonstrated by the eleven most predominant *S. pneumoniae*

677 serotypes collected by the SAVE 2011-2015 study.

Serotype (n*)	Genotype	% with Genotype (n)	% with Genotype that are MDR (n)
7F (701)	PI-1	0	0
	PI-2	97.4 (683)	0.3 (2)
	Dual	0	0
	None	2.6 (18)	5.6 (1)
19A (578)	PI-1	44.5 (257)	5.1 (13)
	PI-2	0.9 (5)	40.0 (2)
	Dual	21.1 (122)	95.9 (117)
	None	33.6 (194)	9.3 (18)
22F (584)	PI-1	0	0
	PI-2	0	0
	Dual	0	0
	None	100 (584)	1.0 (6)
3 (480)	PI-1	0.4 (2)	0
	PI-2	0	0
	Dual	0	0
	None	99.6 (478)	2.5 (12)
12F (276)	PI-1	0	0
	PI-2	0	0
	Dual	0	0
	None	100 (276)	1.4 (4)
11A (241)	PI-1	1.2 (3)	0
	PI-2	67.2 (162)	0
	Dual	0	0
	None	31.5 (76)	2.6 (2)
9N (198)	PI-1	0.5 (1)	0
	PI-2	0	0
	Dual	0	0
	None	99.5 (197)	0.5 (1)

8 (217)	PI-1	0.5 (1)	0
	PI-2	0	0
	Dual	0	0
	None	99.5 (216)	0.5 (1)
33F (203)	PI-1	0	0
	PI-2	0	0
	Dual	0.5 (1)	0
	None	99.5 (202)	6.9 (14)
15A (212)	PI-1	28.8 (61)	0
	PI-2	0	0
	Dual	0	0
	None	71.2 (151)	63.6 (96)
6C (188)	PI-1	8.5 (16)	6.3 (1)
	PI-2	0	0
	Dual	0	0
	None	91.5 (172)	3.5 (6)
All (3878)	PI-1	8.8 (341)	4.1 (14)
	PI-2	21.9 (850)	0.5 (4)
	Dual	3.2 (123)	95.1 (117)
	None	66.1 (2564)	6.3 (161)

678 *, n with complete results for both PCR reactions. Isolates that maintained double positive or double
679 negative results after repeating were excluded.

680 **Table 3:** Resistance genes identified in 160[a] *S. pneumoniae* isolates from the SAVE study sequenced using whole genome sequencing.

681

Antimicrobial Class	Resistance Gene	Count (%)	S/I/R (n)	%S	%NS	Serotypes	%MDR
β-Lactam	*pbp2B* only	17 (10.6)	3/14/0	17.6	82.4	6ABC(4), 7F(1), 8(1), 10A (1), 15A(5), 19A(1), 22F(2), 23B(2)	58.8 (10)
	pbp2X only	10 (6.3)	10/0/0	100	0	3(4), 5(1), 11A(1), 12F(1), 15B(1), 16F(1), 19A(1)	50.0 (5)
	1A+2B	1 (0.6)	0/1/0	0	100	24F(1)	0
	1A+2X	1 (0.6)	0/1/0	0	100	35B(1)	0
	2B+2X	5 (3.1)	1/3/1	20.0	80.0	6C(2), 15A(2), 19A(1)	80.0 (4)
	1A+2B+2X	41 (25.6)	0/9/32	0	100	6B(1), 9V(3), 15AB(7), 19AF(23), 23F(1), 29(1), 35B(5)	85.4 (35)
	None	85 (53.1)	82/3/0	96.5	3.5	-	11.8 (10)
Macrolide/	*mefA* only	17 (10.6)	1/2/14[b]	5.9	94.1	6ABC(5), 9V(2), 14(2), 12F(1), 15B(1), 19A(1), 29(1), 35B(4)	29.4 (5)
Lincosamide/	*ermB* only	48 (30.0)	1/1/46[b]	2.1	97.9	3(6), 6BC(3), 7F (1), 8(1), 9N(1), 11A(2), 12F(2), 15AB(14), 17F(2), 19A(7), 22F(3), 23AF(2), 24F(1), 33F(3)	81.3 (39)
Streptogramin			9/0/39[c]	18.8	81.2		
	Dual	18 (11.2)	1/2/15[b]	5.6	94.4	19A(17), 19F(1)	100 (18)
			7/0/11[c]	38.9	61.1		
	None	77 (48.1)	76/0/1	98.7	1.3	-	2.6 (2)
			77/0/0	100	0		

682

683

Tetracycline	*tetM*	64 (40.0)	4*/0/60	6.3	93.7	3(6), 6BC(3), 7F(1), 8(1), 9N(1), 10A(1), 11A(1), 12F(2), 15AB (14), 17F(2), 19AF(25), 22F(2), 23F(1), 24F(1), 33F(3)	89.1 (57)
	None	96 (60.0)	93/2/1	96.9	3.1	-	7.3 (7)
Fluoroquinolone	*parC* S79 only	3 (1.9)	1/0/2[d]	33.3	66.7	19A(2), 22F(1)	66.7 (2)
	gyrA S81 only	3 (1.9)	2/1/0[e]	66.7	33.3	9N(1), 19A(1), 35B(1)	33.3 (1)
	both	8 (5.0)	0/0/8[d]	0	100	6A(2), 11A(1), 19A(1), 22F(2), 23F(2)	50.0 (4)
			0/4/4[e]	0	100		
	None	146 (91.3)	146/0/0[d]	100	0	-	38.4 (56)
			146/0/0[e]	100	0		
Trimethoprim/	*folA* I100L only	0	-	-	-	-	-
Sulfamethoxazole	*folP* mutation only	10 (6.3)	2/7/1	20.0	80.0	5(1), 15BC(4), 18C(1), 19A(1), 23B(1), 24F(1), 33F(1)	10.0 (1)
	both	37 (23.1)	0/1/36	0	100	5(1), 6ABC(4), 9V(3), 10A(1), 11A(2), 15AB(2), 19AF(21), 23F(1), 35B(2)	83.8 (31)
	None	113 (70.6)	112/0/1	99.1	0.9	-	28.3 (32)
Chloramphenicol	*cat*	8 (5.0)	0/0/8	0	100	3(5), 15B(1), 19A(1), 23F(1)	100 (8)
	None	152 (95.0)	152/0/0	100	0	-	36.8 (56)

S, susceptible; I, intermediate; R, resistant; NS, non-susceptible. [a], isolates without full susceptibility results were excluded from the analysis. [b], susceptibility to clarithromycin. [c], susceptibility to clindamycin. [d], susceptibility to levofloxacin. [e], susceptibility to moxifloxacin. *Discrepant isolates that possessed *tetM* but were susceptible to doxycycline.

689 **Figure 1:** Minimum spanning tree (generated by PHYLOViZ 2.0) of MLST sequence types demonstrated by the eleven most predominant *S.*
690 *pneumoniae* serotypes collected by the SAVE 2011-2015 study. Green outlines indicate a group founder; light blue outlines indicate relatedness
691 to founder; numbers indicate the number of differences between the MLST profiles of the two connected nodes (≤ 2 indicates the two nodes are
692 part of a cluster). Clusters with relation to PMEN international clones are listed along with the representative serotype for that clone.
693
694
695
696

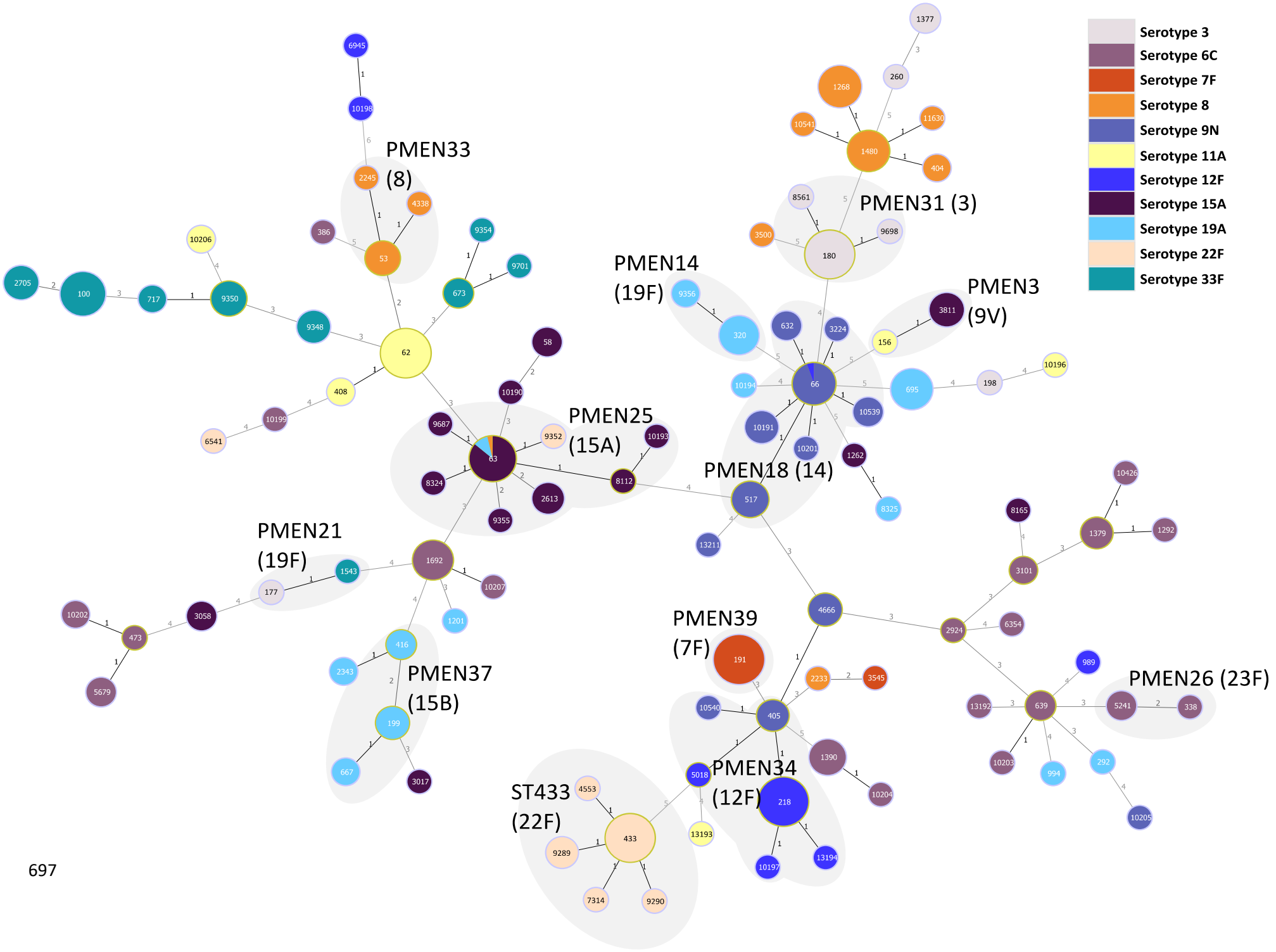
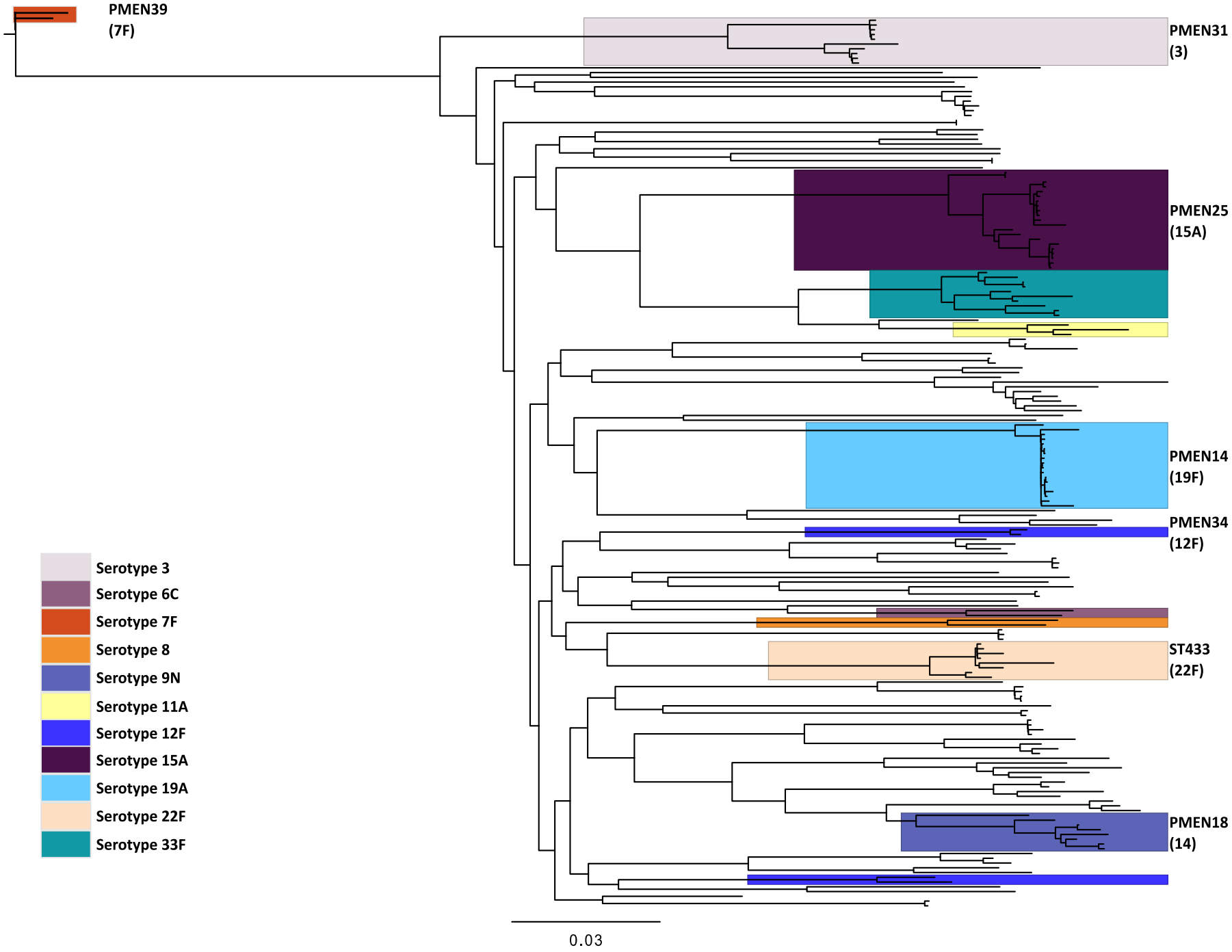

Serotype 3
Serotype 6C
Serotype 7F
Serotype 8
Serotype 9N
Serotype 11A
Serotype 12F
Serotype 15A
Serotype 19A
Serotype 22F
Serotype 33F
PMEN33 (8)
PMEN31 (3)
PMEN14 (19F)
PMEN3 (9V)
PMEN25 (15A)
PMEN18 (14)
PMEN21 (19F)
PMEN39 (7F)
PMEN26 (23F)
PMEN37 (15B)
PMEN34 (12F)
ST433 (22F)

Figure 2: Maximum likelihood tree (generated using PhyML and visualized with FigTree) of 162 isolates from the SAVE 2011-2014 study and 30 background isolates from the BESST 2007-2009 study. Clusters (as delineated by ClusterPicker) containing predominant serotypes collected by the SAVE study are coloured, and relation to PMEN international clones is listed along with the representative serotype for that clone.


PMEN39
(7F)
PMEN31
(3)
PMEN25
(15A)
PMEN14
(19F)
PMEN34
(12F)
ST433
(22F)
PMEN18
(14)
Serotype 3
Serotype 6C
Serotype 7F
Serotype 8
Serotype 9N
Serotype 11A
Serotype 12F
Serotype 15A
Serotype 19A
Serotype 22F
Serotype 33F
0.03