

Genomic investigation of the most common *Streptococcus pneumoniae* serotypes causing invasive infections in Canada: the SAVE study, 2011–2020

Alyssa R. Golden ^{1*}, Heather J. Adam^{2,3}, James A. Karlowsky^{2,3}, Melanie Baxter³, John Schellenberg³, Irene Martin¹, Walter Demczuk¹, Jessica Minion⁴, Paul Van Caesele⁵, Julianne V. Kus^{6,7}, Allison McGeer⁸, Brigitte Lefebvre⁹, Hanan Smadi¹⁰, David Haldane¹¹, Yang Yu¹², Kristen Mead¹³, Michael R. Mulvey^{1,3}, and George G. Zhanel³, on behalf of the Canadian Antimicrobial Resistance Alliance (CARA)†

¹National Microbiology Laboratory, Public Health Agency of Canada, 1015 Arlington Street, Winnipeg, Manitoba, R3E 3R2, Canada; ²Clinical Microbiology, Shared Health, MS673–820 Sherbrook Street, Winnipeg, Manitoba, R3A 1R9, Canada; ³Department of Medical Microbiology and Infectious Diseases, Max Rady College of Medicine, University of Manitoba, Room 543–745 Bannatyne Avenue, Winnipeg, Manitoba, R3E 0J9, Canada; ⁴Roy Romanow Provincial Laboratory, Saskatchewan Health Authority, 5 Research Drive, Regina, Saskatchewan, S4S 0A4, Canada; ⁵Cadham Provincial Laboratory, Shared Health, 750 William Avenue, Winnipeg, Manitoba, R3E 3J7, Canada; ⁶Public Health Ontario Laboratory, 661 University Avenue, Toronto, Ontario, M5G 1M1, Canada; ⁷Department of Laboratory Medicine and Pathobiology, University of Toronto, 1 King's College Circle—6th Floor, Toronto, Ontario, M5S 1A8, Canada; ⁸Toronto Invasive Bacterial Diseases Network (TIBDN), Department of Microbiology, Mount Sinai Hospital, 600 University Avenue—Suite 171, Toronto, Ontario, M5G 1X5, Canada; ⁹Laboratoire de santé publique du Québec, Institut national de santé publique du Québec, 20045 Ch Ste-Marie, Ste-Anne-de-Bellevue, Québec, H9X 3R5, Canada; ¹⁰Epidemiology and Surveillance Branch, New Brunswick Department of Health, 520 King Street, Fredericton, New Brunswick, E3B 5G8, Canada; ¹¹Department of Pathology and Laboratory Medicine, Queen Elizabeth II Health Science Centre, 1276 South Park Street, Halifax, Nova Scotia, B3H 2Y9, Canada; ¹²Newfoundland and Labrador Public Health Laboratory, Dr. Leonard A. Miller Centre—Suite 1, 100 Forest Road, St. John's, Newfoundland and Labrador, A1A 1E3, Canada; ¹³Provincial Laboratory Services, Queen Elizabeth Hospital, 60 Riverside Drive, Charlottetown, Prince Edward Island, C1A 8T5, Canada

*Corresponding author. E-mail: alyssa.golden@phac-aspc.gc.ca

†Members are listed in the Acknowledgements section.

Objectives: To investigate the lineages and genomic antimicrobial resistance (AMR) determinants of the 10 most common pneumococcal serotypes identified in Canada during the five most recent years of the SAVE study, in the context of the 10-year post-PCV13 period in Canada.

Methods: The 10 most common invasive *Streptococcus pneumoniae* serotypes collected by the SAVE study from 2016 to 2020 were 3, 22F, 9N, 8, 4, 12F, 19A, 33F, 23A and 15A. A random sample comprising ~5% of each of these serotypes collected during each year of the full SAVE study (2011–2020) were selected for whole-genome sequencing (WGS) using the Illumina NextSeq platform. Phylogenomic analysis was performed using the SNVPhyl pipeline. WGS data were used to identify virulence genes of interest, sequence types, global pneumococcal sequence clusters (GPSC) and AMR determinants.

Results: Of the 10 serotypes analysed in this study, six increased significantly in prevalence from 2011 to 2020: 3, 4, 8, 9N, 23A and 33F ($P \leq 0.0201$). Serotypes 12F and 15A remained stable in prevalence over time, while serotype 19A decreased in prevalence ($P < 0.0001$). The investigated serotypes represented four of the most prevalent international lineages causing non-vaccine serotype pneumococcal disease in the PCV13 era: GPSC3 (serotypes 8/33F), GPSC19 (22F), GPSC5 (23A) and GPSC26 (12F). Of these lineages, GPSC5 isolates were found to consistently possess the most AMR determinants. Commonly collected vaccine serotypes 3 and 4 were associated with GPSC12 and GPSC27, respectively. However, a more recently collected lineage of serotype 4 (GPSC192) was highly clonal and possessed AMR determinants.

Conclusions: Continued genomic surveillance of *S. pneumoniae* in Canada is essential to monitor for the appearance of new and evolving lineages, including antimicrobial-resistant GPSC5 and GPSC162.

Introduction

Pneumococcal conjugate vaccines (PCVs) have been effective at reducing the incidence of invasive pneumococcal disease (IPD) in Canada.¹ Beginning in mid-2010, PCV13 (serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9 V, 14, 18C, 19A, 19F and 23F) became the recommended vaccine in Canada for routine childhood immunization.² Since the introduction of this vaccine, the distribution of serotypes across Canada has shifted in favour of non-vaccine serotypes such as 8, 22F and 33F.^{3,4} The annual *Streptococcus pneumoniae* Serotyping and Antimicrobial Susceptibility: Assessment for Vaccine Effectiveness in Canada (SAVE) study began in 2011, with the goal of assessing *S. pneumoniae* serotypes and their antimicrobial susceptibility patterns in Canada after the introduction of the PCV13 vaccine. In 2018, we published data on the first 5 years of the SAVE study, encompassing 2011 to 2015.^{4–6} Over this time period, the 10 most commonly collected serotypes were, in order, 7F, 19A, 22F, 3, 12F, 11A, 9N, 8, 33F and 15A. Molecular and phylogenetic analyses identified many of these serotypes to have high within-serotype diversity, including 19A, 15A, 9N and 33F, of which many clones were antimicrobial-resistant; other serotypes of interest were clonal (7F, 22F and 11A).⁵

Since the publication of this previous genomic study, a new typing scheme has been created for *S. pneumoniae*; the Global Pneumococcal Sequencing Cluster (GPSC) framework is the first international scheme for clustering pneumococcal isolates on the basis of whole-genome sequences (including both core and accessory genome data).⁷ This tool is freely and easily accessible on Pathogenwatch (<https://pathogen.watch>), an online genomic surveillance platform. As the distribution of pneumococcal serotypes in Canada shifts due to antimicrobial and vaccine use, this tool might prove to be valuable in quickly determining changes in the circulating lineages within serotypes.

The purpose of this study was to use whole-genome sequencing (WGS) to investigate the lineages and genomic antimicrobial resistance (AMR) determinants of the 10 pneumococcal serotypes most commonly collected in Canada during the five most recent years of the SAVE study (2016–2020), in the context of the 10-year post-PCV13 period in Canada.

Materials and methods

Bacterial isolates

From January 2011 to December 2020, participating Canadian provincial public health and hospital laboratories forwarded isolates of *S. pneumoniae* from sterile body sites 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 analyses were conducted as Western Canada (Saskatchewan and Manitoba), Central Canada (Ontario and Quebec) and Eastern Canada (New Brunswick, Nova Scotia, Prince Edward Island, and Newfoundland and Labrador).

Serotyping

Serotyping was performed using the Quellung reaction⁸ using pool, group, type and factor commercial antisera (SSI Diagnostica; Statens Serum Institut, Copenhagen, Denmark). Isolates for which a serotype was not determined by a Quellung reaction were confirmed as *S. pneumoniae* by *rpoB* gene sequencing.⁹

WGS isolate selection and analysis

The 10 most common *S. pneumoniae* serotypes collected by the SAVE study in the most recent 5-year period (2016–2020) were 3, 22F, 9N, 8, 4, 12F, 19A, 33F, 23A and 15A. A random sample of ~5% of the total of each of these serotypes collected each study year (2011–2020) were selected for WGS. Overall, 345 isolates were sequenced by the PHAC–NML DNA Core Facility using the Illumina NextSeq platform. Quality control was assessed using the FastQC tool (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and contigs were assembled using Shovill (<https://github.com/tseemann/shovill#readme>; Galaxy v.1.0.4+galaxy).¹⁰ Sequencing generated an average of 993 715 reads per genome and average genome coverage of 92 times. *De novo* assembly yielded average contig and N50 lengths of 51 730 and 139 418 bp, respectively.

Phylogenomic analysis was performed using the PHAC–NML built SNVPhyl pipeline to identify core single nucleotide variants (SNVs).¹¹ For each serotype cohort, the isolate with the oldest collection date was selected as the internal reference for the pipeline. Trees generated through the SNVPhyl pipeline were visualized using FigTree v.1.4.4 (<http://tree.bio.ed.ac.uk/software/figtree/>) and in-house developed R Shiny scripts. Trees were rooted on the oldest common outlier. If SNV differences for major clusters are not noted in the text, they can be found in the provided figures.

WGS data were used to identify the presence of pneumococcal pilus genes (PI-1, PI-2), as well as genes encoding resistance to: macrolides (*ermB*, *mefA/E*), tetracyclines (*tetM*), trimethoprim/sulfamethoxazole (*folA*, *folP*), β -lactams (*pbp1a*, *pbp2b*, *pbp2x*) and fluoroquinolones (*gyrA*, *parC*).^{12–14} Predicted resistance phenotypes for these genes and alterations can be found in Table S1 (available as Supplementary data at JAC Online). Pneumococcal surface protein A (*pspA*) families for serotype 3 were identified using the EMBOSS ‘primersearch’ function (up to 10% mismatch) and previously described primers.^{15,16}

MLST alleles and sequence types (ST) were identified *in silico* and queried against the PubMLST *S. pneumoniae* MLST database (<http://pubmlst.org/spneumoniae/>). Isolates were considered part of a clonal cluster (CC) if they shared ≥ 5 of 7 MLST alleles in common. Additionally, the genomic surveillance platform Pathogenwatch (<https://pathogen.watch>) was used to assign isolates to a GPSC.⁷ GPSCs that were commonly identified in this study, as well as associated serotypes, MLST STs/CCs and collection year, are listed in Table S2.

Statistical analysis

Trends in serotype distribution over the 2011 to 2020 time period were assessed for statistical significance using the Cochran–Armitage test for trend, with a *P* value of <0.05 considered significant.

Nucleotide sequence accession numbers

The whole-genome data reported in this study have been deposited in the NCBI Sequence Read Archive as part of BioProject PRJNA510562.

Results

In total, 14,138 invasive isolates of *S. pneumoniae* collected as part of SAVE study between 2011 and 2020 were forwarded to

the CARA for antimicrobial susceptibility testing. Biological sex and age information was available for 13 475 (95.3%) and 13 893 (98.2%) of the isolates, respectively. The annual numbers of isolates were: 1379 isolates from 2011, 1285 from 2012, 1160 from 2013, 1274 from 2014, 1262 from 2015, 1284 from 2016, 1598 from 2017, 1930 from 2018, 1893 from 2019 and 1073 from 2020.

The annual prevalence from 2011 to 2020 of the 10 serotypes analysed by WGS in this study is presented in Table 1.

PCV13 serotypes—3, 4, 19A

Of the 62 serotype 3 isolates analysed in this study, 57 belonged to GPSC12/CC180. This lineage was identified from across Canada in all study years (Figure 1, Table S1). Despite being the same GPSC and MLST clone, these samples fell into three distinct phylogenetic clades (Figure 1). The first and largest cluster represents the previously described ‘clade I’ lineage,¹⁷ which was temporally and geographically distributed, possessed only sporadic AMR determinants and was associated with *pspA* family 2 variants. Isolates within this clade were highly related, demonstrating an average of only 26.2 SNVs difference. The second cluster represented the internationally recognized ‘clade II’ lineage,¹⁷ which varied from clade I in that isolates had increased presence of AMR determinants and were associated with *pspA* family 1 variants. Almost half (45%) of clade II isolates possessed *ermB*, *tetM*, *cat* and alterations in *pbp2x* (Table 2). Isolates were slightly more variable within this clade, averaging 45.0 SNVs difference. While the third cluster demonstrated similar characteristics to clade I, it was significantly different enough (>500 SNVs) to be considered unrelated.

PCV13 vaccine serotype 4 increased significantly over the study period, representing $\leq 2.2\%$ of collected isolates from 2011 to 2016 before steadily increasing up to 12.6% in 2020 (Table 1). Serotype 4 isolates were most commonly part of GPSC27 (18/27), representing two distinct MLST clusters: CC205 and CC244 (nine each). GPSC27/CC205 isolates, labelled cluster 1 in Figure 2, were geographically diverse and differed by an average of ~ 30 SNVs. GPSC27/CC244 isolates were predominantly collected from western Canada and separated into two subclusters. Cluster 2A was collected from 2019 and 2020, while cluster 2B was collected from 2013–2017; these clusters differed by an average of only 3.4 and 7.5 SNVs, respectively. GPSC27 isolates did not possess AMR determinants (Table 2), however, they did possess PI-1. In contrast, three isolates of serotype 4/GPSC162/ST2759 possessed *ermB*, *tetM* and *cat* (cluster 3). These isolates were all collected from central Canada in 2020 and were highly clonal (3.1 SNVs difference).

In contrast, PCV13 serotype 19A decreased substantially in prevalence from 13.0% of isolates collected in 2011 to 3.4% in 2020 (Table 1). Serotype 19A isolates in this study belonged to three main lineages: GPSC27/CC695 (13/43), GPSC1/CC320 (12/43) and GPSC4/CC199 (10/43), each with very different AMR characteristics (Table 2, Figure S1). GPSC4/CC199 isolates were geographically spread, predominantly isolated from 2015 or prior and generally possessed a combination of alterations in *pbp2b* and *pbp2x* that were predicted to confer a penicillin-intermediate phenotype. GPSC27/CC695 were collected from central and eastern Canada and were temporally distributed, although many of

the more recently collected serotype 19A isolates fell into this lineage. This cluster was highly clonal, demonstrating an average of only 7.1 SNVs difference between isolates. Most possessed *mefA/E* (8/13) and all had PI-1. Comparatively, GPSC1/CC320 isolates possessed almost all AMR determinants: *mefA/E*, *ermB*, *tetM*, *folA* I100L, *folP* insertions and alterations in key motifs of all three PBPs (Table 2); this was also the only lineage in the current study that possessed both PI-1 and PI-2. GPSC1/CC320 isolates were geographically spread, generally from the earlier years of the SAVE study (9/12 were from 2015 and prior) and were also highly related, with an average of 10.4 SNVs difference between isolates (Figure S1).

Other vaccine serotypes—8, 9N, 12F, 22F, 33F

Serotypes 8, 9N and 33F all significantly increased in prevalence from 2011 to 2020 (Table 1). Serotype 8 isolates predominantly belonged to GPSC98/CC1268 in this study (22/30). This lineage was geographically diverse across Canada, predominantly collected from 2016 and later. A smaller number of serotype 8 isolates were part of GPSC3/CC53 (7/30); most of these isolates were collected prior to 2016. Serotype 8 isolates were only sporadically associated with AMR determinants (Table 2, Figure S2). Serotype 33F was also part of GPSC3, with isolates commonly falling into one of two subclusters: CC100 or CC717 (Figure S3). CC100 isolates commonly possessed *mefA/E* (11/15) and *folP* alterations (15/15), while CC717 isolates all possessed *ermB* and *tetM* (Table 2). GPSC3/CC100 were temporally and geographically diverse, while GPSC3/CC717 were generally collected from central Canada in 2011–2015. Serotype 9N isolates were either part of GPSC16/CC66 (25/32) or GPSC32/CC405 (7/32). GPSC16/CC66 was geographically and temporally diverse but demonstrated significant variability within the cluster (Figure S4). GPSC32/CC405, was also geographically diverse but was collected only from 2015 onwards. Both lineages only demonstrated sporadic presence of AMR determinants (Table 2).

The prevalence of serotype 12F remained stable over the study period (Table 1). Isolates fell predominantly into either GPSC26/CC989 (6/30) or GPSC32/CC218 (22/30). GPSC26/CC989 isolates were collected from central Canada from 2015 onwards. These samples all possessed *tetM* and *folP* alterations, with some also possessing *cat* (3/6) or *folA* I100L (5/6) (Table 2). GPSC32/CC218 divided further into two subclusters: the first predominantly contained older isolates collected from western Canada that possessed no resistance genes. The second cluster contained temporally diverse isolates collected from western Canada that possessed *mefA/E*. These two clusters demonstrated an average SNV difference of 29 and 6.7 SNVs, respectively (Figure S5).

Serotype 22F isolates appeared quite clonal, with 56/60 isolates falling into GPSC19/CC433, a lineage that is broadly spread across Canada. Despite this, isolates in this cluster varied from each other by an average of 178 SNVs (Figure S6). A small subcluster of these isolates possessed *mefA/E*, predominantly collected from central Canada in 2015 or later. Notably, serotype 22F isolates that were unrelated to GPSC19/CC433 possessed alterations in *pbp2b* and *pbp2x*; two of these were part of GPSC9/CC63 and contained alterations in the QRDR regions of *gyrA* and *parC* (Table 2).

Table 1. Trends in the prevalence of the 10 serotypes analysed in this study

Serotype	Prevalence, % (n)										Trend
	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	
3	7.6 (105)	8.2 (106)	7.8 (90)	7.8 (99)	8.0 (101)	10.4 (133)	9.6 (154)	12.3 (237)	11.3 (214)	10.7 (115)	Increasing
4	2.2 (30)	1.8 (23)	1.1 (13)	0.9 (12)	1.3 (17)	1.2 (16)	3.8 (60)	3.4 (66)	6.2 (118)	12.6 (135)	Increasing
8	2.8 (39)	3.3 (42)	4.5 (52)	3.9 (50)	4.9 (62)	6.1 (78)	4.8 (76)	5.8 (111)	5.9 (111)	8.1 (87)	Increasing
9N	2.6 (36)	3.3 (43)	4.0 (46)	4.2 (53)	5.5 (69)	5.5 (70)	6.6 (106)	5.6 (109)	7.1 (135)	7.0 (75)	Increasing
12F	4.5 (62)	4.0 (52)	5.0 (58)	4.1 (52)	5.6 (71)	5.1 (66)	4.6 (73)	4.6 (89)	4.4 (83)	5.1 (55)	NS
15A	4.2 (58)	3.7 (48)	2.8 (32)	3.4 (43)	3.2 (40)	3.3 (42)	4.2 (67)	3.8 (73)	4.1 (77)	3.7 (40)	NS
19A	13.0 (179)	9.6 (123)	10.3 (120)	7.1 (90)	7.5 (95)	5.6 (72)	5.0 (80)	5.1 (98)	3.5 (67)	3.4 (37)	Decreasing
22F	7.8 (107)	11.2 (144)	10.9 (126)	9.5 (121)	8.6 (108)	8.9 (114)	8.7 (139)	9.2 (177)	9.5 (180)	6.2 (66)	NS ^c
23A	2.2 (31)	3.5 (45)	2.8 (32)	3.2 (41)	3.1 (39)	3.3 (43)	5.2 (83)	3.4 (65)	4.2 (79)	2.7 (29)	Increasing
33F	1.8 (25)	3.1 (40)	4.4 (51)	3.6 (46)	5.2 (66)	4.0 (51)	4.6 (74)	3.5 (67)	4.5 (86)	3.1 (33)	Increasing
Total	1379	1285	1160	1274	1262	1284	1598	1930	1893	1073	

NS, not significant.
^cP=0.0519, borderline decreasing.

Non-vaccine serotypes—15A, 23A

The prevalence of serotype 15A remained stable over the study period (Table 1). The largest proportion of isolates belonged to GPSC9/CC63 (16/22); these isolates were geographically and temporally diverse and also diverse within their phylogenetic cluster (~124 SNVs difference) (Figure S7). All GPSC9/CC63 isolates possessed *ermB* and *tetM*, as well as numerous alterations in PBPs (Table 2). A small cluster demonstrated alterations in key motifs of all three PBPs and were predicted to possess full resistance to penicillin (5/16). The remainder possessed alterations in only *pbp2b* and *pbp2x* and were predicted to demonstrate a penicillin-intermediate phenotype. Four serotype 15A isolates collected more recently (2018–2020) were GPSC6/ST3811. These isolates were very different from GPSC9/CC63 in that they possessed no AMR determinants.

Overall, serotype 23A increased significantly over the study period, from 2.2% in 2011 to 4.2% in 2019 (although prevalence did drop back to 2.7% in 2020) (Table 1). This serotype was split between GPSC7/CC42 (11/20) and GPSC5/CC338 (9/20). Isolates within GPSC7/CC42 were geographically and temporally spread, highly variable within their cluster (203 SNVs difference) and only possessed sporadic AMR determinants (Figure 3). GPSC5/CC338 isolates were collected from 2015–2020 in central Canada. Six of these isolates possessed *ermB*, *tetM*, alterations in all PBPs, and were closely related (average of 26.6 SNVs). The remainder possessed the I100L alteration in *folA* and alterations in *pbp2x* only.

Discussion

This study investigated the genomic lineages and AMR determinants of the 10 most commonly identified pneumococcal serotypes in Canada during the five most recent years of the SAVE study (2016–2020). Despite only five additional years elapsing since our previous characterization, the 10 most common serotypes collected have changed.⁵ In 2011–2015, the 10 most common serotypes were, in order, 7F, 19A, 22F, 3, 12F, 11A, 9N, 8, 33F and 15A.⁵ Comparing this to the 10 most common types from the five most recent years (3, 22F, 9N, 8, 4, 12F, 19A, 33F, 23A and 15A), 7F is no longer a common type and 19A has dropped in ranking, probably due to PCV13 use. Non-PCV13 serotypes 8 and 9N have also increased in rank order. PCV13 serotype 3 became the most common serotype over the 5-year timespan while serotype 4 has suddenly become common. Even these small changes underscore how variable the distribution of pneumococcal serotypes can be under the pressure of antimicrobial and conjugate vaccine use.

In a recent study of international pneumococcal genomics, Lo *et al.* noted that the 10 most prevalent international lineages causing non-vaccine serotype IPD in the PCV13 era were GPSC3, 19, 11, 5, 38, 26, 36, 48, 9 and 7.¹⁸ The serotypes analysed in our current study represent four of these GPSCs: GPSC3 (serotypes 8 and 33F), GPSC19 (serotype 22F), GPSC5 (serotype 23A) and GPSC26 (serotype 12F). GPSC3/serotype 8 is associated with CC53, a lineage that was uncommonly identified in this study (eight of 30 sequenced isolates, predominantly from earlier years). In contrast, GPSC98 represented two-thirds of serotype 8 isolates in this study. Although defined as a GPSC of

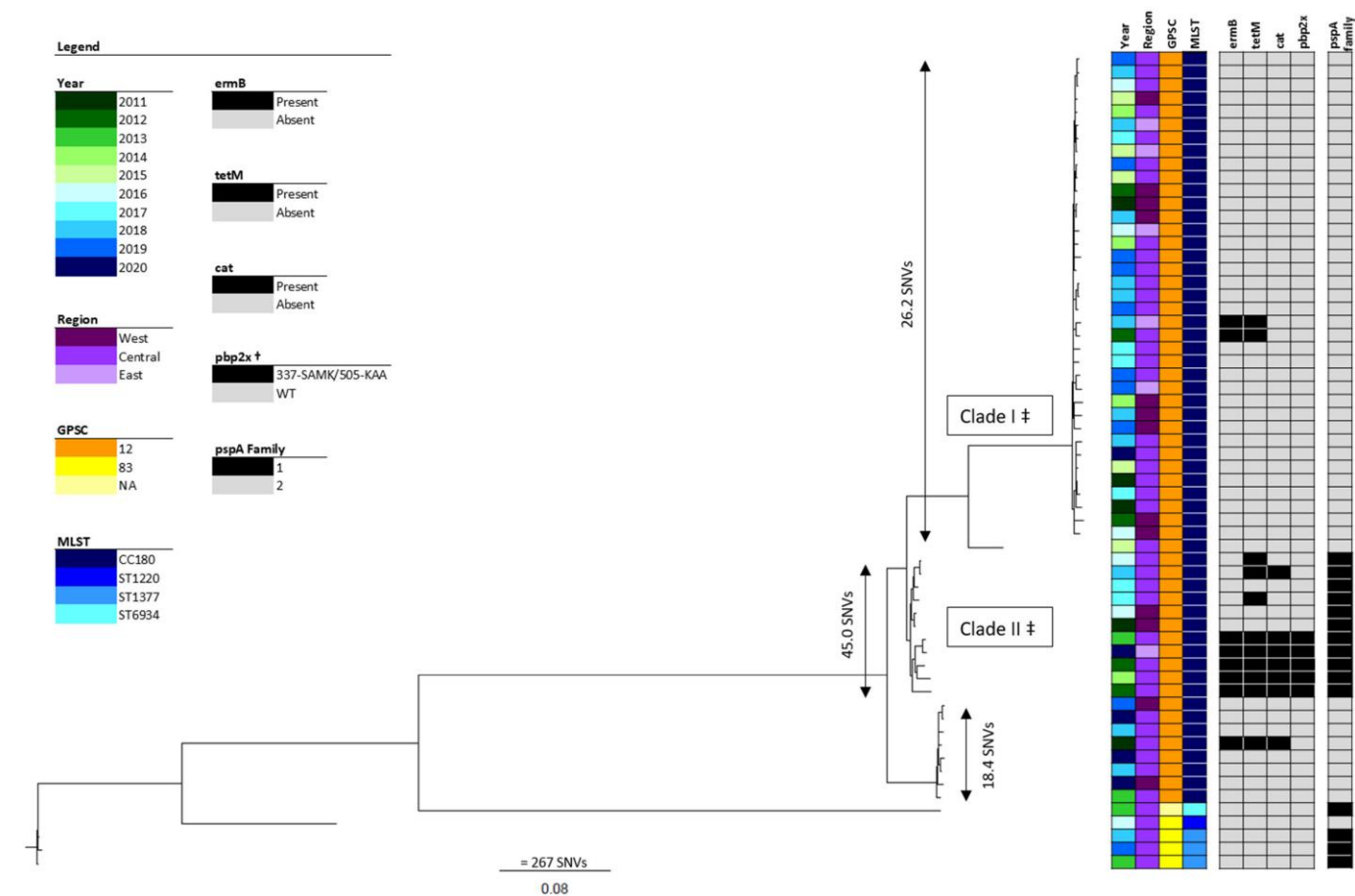


Figure 1. Maximum likelihood core SNV phylogenetic tree of *S. pneumoniae* serotype 3 isolates collected from across Canada, 2011–2020. * A total of 3,335 sites were used in the phylogeny, including 86.2% of the core genome. SC11–1155-P was used as a mapping reference. SNV density filtering was set to remove regions with >5 SNVs per 500 base pairs. † *pbp2x* alterations 337-SAMK/505-KAA are not predicted to confer penicillin resistance. ‡ Clades previously described by Azarian *et al.*¹⁷ This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

‘intermediate’ prevalence by Gladstone *et al.*,⁷ GPSC98 isolates have been deposited to Pathogenwatch by nine other countries on five continents. Based on the isolates analysed in this study, it can be postulated that the increased prevalence of serotype 8 over time in Canada is due to GPSC98 and not the globally prevalent lineage GPSC3, although further study is required. Our study is in contrast to European studies of serotype 8, that have demonstrated GPSC3/CC53 to be the most prevalent and increasing lineage.^{19,20} However, Lo *et al.* noted that major lineages can be associated with different serotypes in different countries; the apt example given was that GPSC3 was associated with serotype 8 in South Africa but serotype 33F in the USA.¹⁸ As Canada shares a border with the USA and, therefore, often demonstrates similar pneumococcal serotype trends, it is logical that GPSC3 in Canada is also more commonly associated with serotype 33F.

In Canada, extensive characterization of serotype 22F has been performed in the post-PCV13 era.^{21–25} The serotype has remained highly clonal over several years, with a consistent >90% of isolates belonging to GPSC19/CC433. The current study was no different, but did note a subcluster of GPSC19/CC433 possessing

mefA/E, predominantly collected from 2015 onwards; the expansion of this specific clone in Canada has previously been identified by Demczuk *et al.*²³ and Duvvuri *et al.*²² As the SAVE study has recently noted a significant increase in macrolide-resistant serotype 22F,²⁵ it is likely that this clone has continued to expand since these initial studies.

The fourth most common GPSC associated with non-vaccine type IPD in the Lo *et al.* paper was GPSC5.¹⁸ This study noted that GPSC5 was associated with significant resistance; 93% of isolates were penicillin-resistant and 86% were non-susceptible to two or more classes of antimicrobials. The study also noted that GPSC5 was associated with two common penicillin-resistant serotypes, 35B/D in Malawi and South Africa and serotype 19A in Israel.¹⁸ Pathogenwatch has GPSC5 records for 16 different serotypes; recent genomic studies in Brazil and Argentina also noted a plethora of serotypes associated with GPSC5.^{26,27} In our current study, GPSC5 was associated with serotype 23A. Like GPSC5 isolates in the literature, GPSC5/serotype 23A collected from the SAVE study were antimicrobial-resistant, possessing *ermB*, *tetM* and alterations in multiple PBPs (although not predicted to demonstrate full penicillin resistance by CLSI oral penicillin

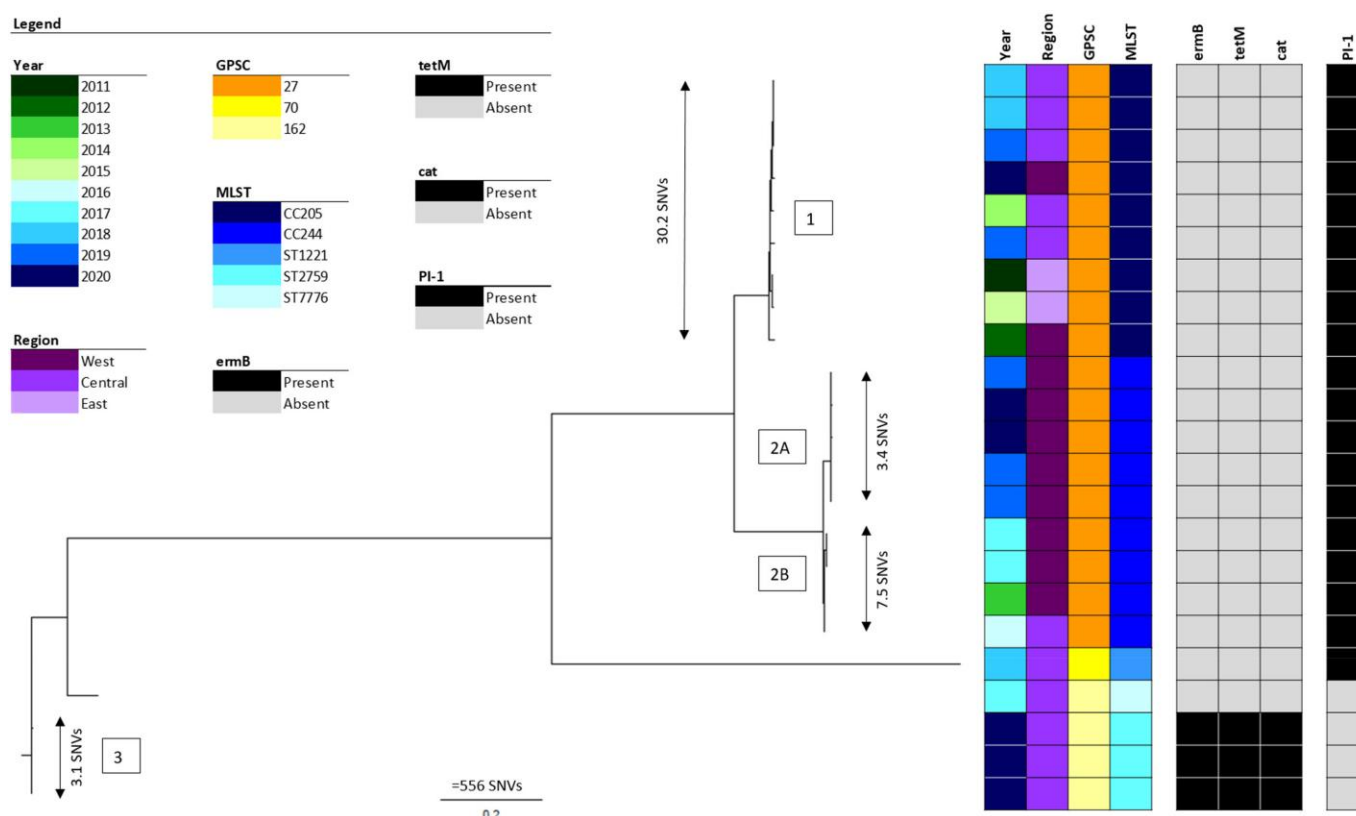


Figure 2. Maximum likelihood core SNV phylogenetic tree of *S. pneumoniae* serotype 4 isolates collected from across Canada, 2011–2020. *A total of 2779 sites were used in the phylogeny, including 80.4% of the core genome. SC11-4519-P was used as a mapping reference. SNV density filtering was set to remove regions with >5 SNVs per 500 base pairs. This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

breakpoints). These isolates were collected beginning in 2015, suggesting that this lineage has contributed to the increasing prevalence of serotype 23A in Canada over time. Numerous studies from other countries such as South Korea, Greece, Japan and the USA have also noted increases in serotype 23A over time,^{28–32} some describing penicillin resistance and MDR in their cohorts.^{29,32} Comprehensive WGS studies of serotype 23A have not been performed in Canada. It will be important to investigate serotype 23A in the future to monitor the expansion of this commonly penicillin-resistant lineage.

A final non-vaccine type lineage noted in our study was GPSC26, which has been described as the majority global lineage of serotype 12F.¹⁸ GPSC26/serotype 12F isolates have been noted for their non-susceptibility to chloramphenicol, tetracycline and trimethoprim/sulfamethoxazole.^{18,26} Although some serotype 12F isolates in Canada belonged to this lineage and demonstrated the characteristic resistance pattern, most serotype 12F isolates were GPSC32/CC218. In North America, serotype 12F has been associated with outbreaks and disease in disadvantaged populations. For example, in a recent study of disease clusters disproportionately affecting persons experiencing homelessness (PEH), Beall *et al.* identified serotype 12F as the most common serotype causing IPD in PEH in western USA.³³ In the current study, two separate clades of GPSC32/CC218 were identified, one with and one without the characteristic

presence of *mefA/E*. This resistant clade was originally identified in a 2010 outbreak of IPD that affected PEH and people with substance use disorders in Manitoba, Canada.^{34,35} The lineage has successfully expanded in western and central Canada as it has continued to be collected throughout the SAVE study. After investigating the Pathogenwatch public database, it has become more evident that this lineage may be restricted to Canada and the USA, as the only publicly available GPSC32/CC218 isolates demonstrating macrolide resistance were from the USA.

Although non-vaccine type lineages are important to monitor to determine the future landscape of pneumococcal disease, it is also important to discuss serotypes 3 and 4. These PCV13 serotypes have increased in prevalence in Canada since 2011. It is known that the PCV13-generated immune response to serotype 3 is reduced, probably due to abundant capsular production overwhelming the protective capacity of vaccine-generated antibodies.³⁶ This has contributed to the steady presence of serotype 3 in Canada since PCV13 introduction, in comparison to the sharp decreases in prevalence seen with some other PCV13 serotypes (e.g. 6A, 7F, 19A).³⁷ Serotype 3 isolates in Canada have consistently been part of the international global MLST clone CC180 (GPSC12); however, MLST alone cannot differentiate between the recognized clades of GPSC12/CC180 circulating worldwide. Initially, CC180 isolates were shown to belong to a single, rarely recombining clade.³⁸ More recently, an

Table 2. Antimicrobial resistance gene profiles of common pneumococcal serotypes, as well as their various global pneumococcal sequencing clusters and STs

Serotype (n)	GPSC (n)	MLST CC/ST (n)	Antimicrobial resistance gene profile(s) (n) ^{a,b}
3 (62)	12 (57)	CC180 (57)	<i>ermB</i> — <i>tetM</i> (4); <i>ermB</i> — <i>tetM</i> — <i>cat</i> (2); <i>ermB</i> — <i>tetM</i> — <i>cat</i> — <i>pbp2x</i> SAMK/WT/KAA/WT (4); <i>ermB</i> — <i>tetM</i> — <i>cat</i> — <i>pbp2x</i> SAMK/WT/KAA/WT— <i>parC</i> D83N (1); WT (46)
			WT (1)
			WT (3)
			WT (1)
4 (23)	27 (18)	ST205 (9)	WT (9)
			WT (9)
			WT (1)
			<i>ermB</i> — <i>tetM</i> — <i>cat</i> (3)
8 (30)	3 (6)	CC53 (6)	WT (1)
			<i>tetM</i> (1); WT (5)
			WT (22)
			<i>pbp1a</i> WT/WT/WT/TSQY— <i>pbp2b</i> WT/YSSN/QLQP/GKT (1)
9N (32)	16 (25)	CC66 (25)	WT (1)
			<i>gyrA</i> S81F (1); <i>ermB</i> — <i>tetM</i> (2); <i>folP</i> INS_RP[60] (1); <i>mefAE</i> (2); <i>parC</i> S79F (2); WT (17)
			WT (7)
			<i>mefAE</i> — <i>tetM</i> — <i>cat</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64] (1); <i>tetM</i> — <i>cat</i> — <i>folA</i> I100L— <i>folP</i> INS_R[61] (2); <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_R[61] (2); <i>tetM</i> — <i>folP</i> INS_R[61] (1); WT (1)
12F (30)	32 (22)	CC218 (22)	<i>mefAE</i> (12); WT (10)
			<i>ermB</i> — <i>tetM</i> (1); <i>ermB</i> — <i>tetM</i> — <i>pbp2x</i> WT/WT/WT/IKSG (1)
			WT (4)
			<i>ermB</i> — <i>tetM</i> — <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (5); <i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> WT/WT/KEA/WT (10)
15A (22)	9 (16)	CC63 (16)	WT (1)
			WT (1)
			<i>ermB</i> — <i>mefAE</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_EI[67]— <i>pbp1a</i> SSMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/AIDTK/WT— <i>pbp2x</i> SAMK/WT/EDT/VKSG (12)
			<i>folP</i> INS_EI[67]— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/WT/WT (2); <i>mefAE</i> — <i>folA</i> I100L— <i>folP</i> INS_SS[63]— <i>pbp1a</i> SSMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1); <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/WT/WT (5); WT (2)
19A (43)	4 (10)	CC199 (10)	<i>ermB</i> — <i>tetM</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT (1); <i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> WT/WT/KEA/WT (1)
			<i>ermB</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64]— <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/YSSN/QLQP/GKT— <i>pbp2x</i> WT/WT/WT/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_SY[64]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1)
			WT (2)
			<i>ermB</i> — <i>tetM</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT (1); <i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> WT/WT/KEA/WT (1)
19A (43)	10 (2)	CC276 (2)	<i>ermB</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64]— <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/YSSN/QLQP/GKT— <i>pbp2x</i> WT/WT/WT/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_SY[64]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1)
			<i>ermB</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64]— <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/YSSN/QLQP/GKT— <i>pbp2x</i> WT/WT/WT/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_SY[64]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1)
			<i>ermB</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64]— <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/YSSN/QLQP/GKT— <i>pbp2x</i> WT/WT/WT/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_SY[64]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1)
			<i>ermB</i> — <i>tetM</i> — <i>folA</i> I100L— <i>folP</i> INS_SY[64]— <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/YSSN/QLQP/GKT— <i>pbp2x</i> WT/WT/WT/VKSG (1); <i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_SY[64]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> SAMK/WT/KEA/VKSG (1)

Continued

Table 2. Continued

Serotype (n)	GPSC (n)	MLST CC/ST (n)	Antimicrobial resistance gene profile(s) (n) ^{a,b}
22F (60)	11 (1)	ST193 (1)	WT (1)
	14 (1)	ST17183 (1)	<i>mefAE</i> — <i>tetM</i> — <i>folP</i> INS_PG[61]— <i>pbp1a</i> SAMK/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/WT/WT (1)
	18 (1)	ST1201 (1)	WT (1)
	27 (13)	CC695 (13)	<i>ermB</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2x</i> SAMK/WT/WT/WT (1); <i>mefAE</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2x</i> SAMK/WT/WT/WT (8); <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2x</i> SAMK/WT/WT/WT (4)
	324 (1)	ST292 (1)	WT (1)
	9 (2)	ST9352 (2)	<i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>gyrA</i> S81L— <i>parC</i> S79F (2)
	19 (56)	CC433 (56)	<i>ermB</i> — <i>mefAE</i> (1); <i>ermB</i> — <i>tetM</i> (2); <i>mefAE</i> (17); <i>parC</i> S79Y (1); WT (35)
	215 (2)	ST6541 (2)	<i>ermB</i> — <i>tetM</i> — <i>pbp2b</i> WT/SSNA/WT/WT— <i>pbp2x</i> WT/WT/KAA/WT (2)
	5 (9)	CC338 (9)	<i>ermB</i> — <i>tetM</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/SVESK/WT (1); <i>ermB</i> — <i>tetM</i> — <i>pbp1a</i> WT/WT/WT/NTGY— <i>pbp2b</i> WT/SSNA/SVESK/WT— <i>pbp2x</i> WT/WT/EDA/VKSG (5); <i>folA</i> I100L— <i>pbp2b</i> WT/SSNA/SVESK/WT (3)
	7 (11)	CC42 (11)	<i>ermB</i> (1); <i>folA</i> I100L— <i>folP</i> INS_CS[63] (1); WT (9)
33F (23)	3 (23)	ST60 (1)	WT (1)
		CC100 (14)	<i>folA</i> I100L— <i>folP</i> INS_VE[66] (1); <i>folP</i> INS_VE[66] (2); <i>mefAE</i> — <i>folA</i> I100L— <i>folP</i> INS_VE[66] (2); <i>mefAE</i> — <i>folP</i> INS_VE[66] (9)
		CC673 (3)	<i>ermB</i> — <i>tetM</i> — <i>folP</i> INS_R[61] (1); WT (2)
		CC717 (4)	<i>ermB</i> — <i>tetM</i> (4)
		ST9348 (1)	<i>ermB</i> — <i>tetM</i> (1)

GPSC, global pneumococcal sequencing cluster; ST, sequence type; CC, clonal cluster; NA, not assigned; WT, wild-type. Predicted resistance phenotypes for these genes and alterations can be found in Table S1.

^aPBPs only listed in the table if there is an alteration of interest in one or more motifs. Wild-type amino acid motifs are as follows: *pbp1a*: 370-STMK/466-SSN/557-KTG/574-TSQF;

pbp2b: 386-SSVK/443-SSNT/565-QLQPT/615-KTG;

pbp2x: 337-STMK/395-SSN/505-KDA/546-LKSG.

^bAmino acid insertions in *folP* noted in the 58-RPGSSVYEI-66 region. The table lists the extra amino acids and [estimated point of insertion].

important secondary clade was identified that diverged from pre-PCV13 isolates. These isolates were found to be more highly recombinant and also commonly antimicrobial-resistant due to the presence of *ermB*, *tetM* and *cat*.¹⁷ Approximately 20% of the GPSC12/CC180 isolates in our current study were determined to be part of clade II and 45% of those possessed AMR determinants. While the resistance profile matched that of isolates in other studies, the prevalence of clade II in Canada is lower than these previous studies describe. A study from England and Wales described clade II as expanding to represent half of the serotype 3 isolates collected since 2014.³⁹ Similarly, Azarian *et al.* observed that clade II made up more than half of the post-PCV13 serotype 3 CC180 population in North America.¹⁷ We did not see this level of clade II in Canada in the present study; however, significantly more sequencing of recent serotype 3 isolates is required to confidently conclude whether CC180

clade II is increasing in Canada over time as it is in other countries.

In contrast to the consistent presence of serotype 3 in the pneumococcal landscape, serotype 4 has just recently started to increase again as a cause of IPD in Canada. The full 2020 surveillance report of pneumococcal disease (all Canadian provinces and territories) noted that serotype 4 only rose past serotype 3 in 2020 to become the most common type associated with IPD.³⁷ This increase may be partially attributed to outbreaks of disease. Beall *et al.* identified serotype 4 as the second most common serotype causing IPD in PEH in western USA.³³ Another recent study identified serotype 4/CC244 (GPSC27) as the cause of significant IPD in PEH and persons with substance use disorders in Alberta, Canada.⁴⁰ This study also noted that CC244 was more common in Alberta, with CC205 common nationally. In our current study, we noted similarly that serotype 4/CC205 was

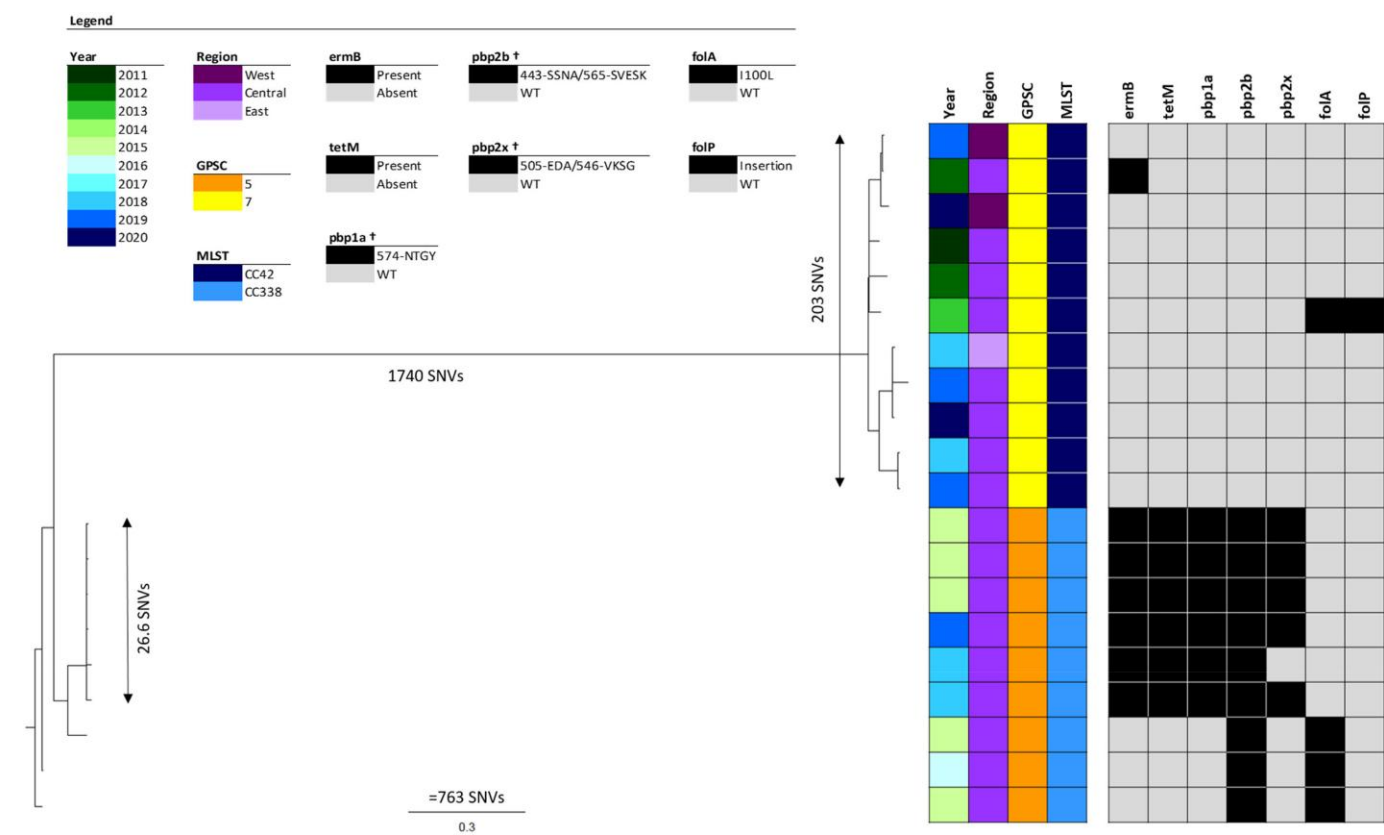


Figure 3. Maximum likelihood core SNV phylogenetic tree of *S. pneumoniae* serotype 23A isolates collected from across Canada, 2011–2020. * A total of 2544 sites were used in the phylogeny, including 84.8% of the core genome. SC11-4141-P was used as a mapping reference. SNV density filtering was set to remove regions with >5 SNVs per 500 base pairs. † The various combinations of pbp alterations present in this serotype were predicted to confer penicillin-intermediate phenotype. This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

geographically diverse. Although Alberta is not part of the SAVE study, the geographic specificity of CC244 to Alberta confirms our finding that CC244 is a western Canadian clone. Of particular interest in our current study is the presence of GPSC162/ST2759, isolates that possessed *ermB*, *tetM* and *cat* and were highly clonal (3.1 SNVs). GPSC162 was defined as an ‘intermediate’ GPSC by Gladstone *et al*⁷ and Pathogenwatch records for this lineage were generally associated with Asia. A recent genomic investigation of pneumococcal disease in Russia listed all collected serotype 4 isolates as belonging to GPSC162, however, no resistance information was provided.⁴¹ As our study only sequenced a small percentage of each serotype collected each study year and does not receive any epidemiological information, it is possible these isolates were part of an outbreak related to travel. It will be important to monitor this lineage closely, as an outbreak-associated serotype with known drug resistance could pose significant problems.

This study is limited by the participation of only eight out of 10 Canadian provinces. Alberta and British Columbia, Canada’s two most westerly provinces, do not submit isolates to the SAVE study, therefore, regional analyses may underrepresent western Canada. Only a select number of isolates were included for WGS analysis in this study, which is a small proportion of the isolates collected by the SAVE study in any given year. By only including

a small sampling, it is possible that outbreaks or regional serotype variations may have been missed in our analysis. Last, the total number of isolates received as part of SAVE 2020 was much lower than usual, probably due to the intervention strategies put in place to cope with the COVID-19 pandemic. The preliminary incidence rate of IPD in Canada for 2020 was 5.9 cases per 100 000 population,³⁷ representing a sharp decrease from the 10.1 cases per 100 000 population reported in 2019. Only time will tell whether this decrease had a marked effect on the pneumococcal serotype distribution in Canada.

In conclusion, the landscape of pneumococcal serotypes in Canada has shifted since the last genomic investigation of the SAVE study. Commonly collected pneumococcal serotypes from across Canada in 2016 to 2020 (8, 12F, 22F, 23A, 33F) represent some of the most prevalent international lineages associated with non-vaccine serotype IPD. PCV13 serotype 3/GPSC12 continues to be responsible for a significant amount of pneumococcal disease in Canada, separating into two distinct clades on the basis of rate of recombination. Serotype 4 also increased in prevalence, and was predominantly associated with GPSC27. Two specific lineages of note were serotype 4/GPSC162/ST2759 and serotype 23A/GPSC5/CC338. Both lineages have been more recently collected in Canada and were associated with increased AMR determinants in comparison to other clones of the same

serotype. Continued genomic surveillance is required to monitor new and evolving lineages of *S. pneumoniae* in Canada.

Acknowledgements

CARA principal members include: Dr Heather Adam, Melanie Baxter, Dr Alyssa Golden, Dr James Karlowsky, Dr Philippe Lagacé-Wiens, Dr Andrew Walkty and Dr George Zhanel.

We would like to thank the following Canadian Public Health Laboratory Network (CPHLN) laboratories for their participation in this study: Roy Romanow Provincial Laboratory, Regina, Saskatchewan; Cadham Provincial Laboratory, Winnipeg, Manitoba; Public Health Ontario Laboratories, Toronto, Ontario; Laboratoire de santé publique du Québec, 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 the Newfoundland Public Health Laboratory, St. John's, Newfoundland and Labrador.

Funding

The SAVE study was supported in part by the University of Manitoba; Shared Health Manitoba, the National Microbiology Laboratory-Public Health Agency of Canada (PHAC-NML); Merck Canada Inc. and Pfizer Canada Inc.

Transparency declaration

G.G.Z. has received research grants from Merck Canada Inc. and Pfizer Canada Inc. All other authors do not have conflicts to declare.

Disclaimer

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

Supplementary data

Figures S1 to S7 and Tables S1 and S2 are available as [Supplementary data](#) at JAC Online.

References

- 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. <https://doi.org/10.1016/j.vaccine.2009.12.026>
- Desai S, McGeer A, Quach-Thanh C *et al.* Update on the use of conjugate pneumococcal vaccines in childhood: an Advisory Committee Statement (ACS) National Advisory Committee on Immunization (NACI). *Can Commun Dis Rep* 2010; **36**: 1–21. <https://doi.org/10.14745/ccdr.v36i00a12>
- Demczuk WHB, Martin I, Desai S *et al.* Serotype distribution of invasive *Streptococcus pneumoniae* in adults 65 years of age and over after the introduction of childhood 13-valent pneumococcal conjugate vaccination programs in Canada, 2010–2016. *Vaccine* 2018; **36**: 4701–7. <https://doi.org/10.1016/j.vaccine.2018.06.018>
- Adam HJ, Golden AR, Karlowsky JA *et al.* Analysis of multidrug resistance in the predominant *Streptococcus pneumoniae* serotypes in Canada: the SAVE study, 2011–15. *J Antimicrob Chemother* 2018; **73**(Suppl 7): vii12–9. <https://doi.org/10.1093/jac/dky158>
- Golden AR, Adam HJ, Karlowsky JA *et al.* Molecular characterization of predominant *Streptococcus pneumoniae* serotypes causing invasive infections in Canada: the SAVE study, 2011–15. *J Antimicrob Chemother* 2018; **73**(Suppl 7): vii20–31. <https://doi.org/10.1093/jac/dky157>
- Karlowsky JA, Adam HJ, Golden AR *et al.* Antimicrobial susceptibility testing of invasive isolates of *Streptococcus pneumoniae* from Canadian patients: the SAVE study, 2011–15. *J Antimicrob Chemother* 2018; **73**(Suppl 7): vii5–11. <https://doi.org/10.1093/jac/dky156>
- Gladstone RA, Lo SW, Lees JA *et al.* International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. *EBioMedicine* 2019; **43**: 338–46. <https://doi.org/10.1016/j.ebiom.2019.04.021>
- Austrian R. The Quellung reaction, a neglected microbiologic technique. *Mt Sinai J Med* 1976; **43**: 699–709.
- 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. <https://doi.org/10.1128/JCM.42.2.497-504.2004>
- Bankevich A, Nurk S, Antipov D *et al.* SPADeS: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–77. <https://doi.org/10.1089/cmb.2012.0021>
- Petkau A, Mabon P, Sieffert C *et al.* SNVPhyl: a single nucleotide variant phylogenomics pipeline for microbial genomic epidemiology. *Microb Genom* 2017; **3**: e000116.
- Metcalfe BJ, Gertz RE, Gladstone RA *et al.* Strain features and distributions in pneumococci from children with invasive disease before and after 13-valent conjugate vaccine implementation in the USA. *Clin Microbiol Infect* 2016; **22**: 60.e9–60.e29. <https://doi.org/10.1016/j.cmi.2015.08.027>
- Kim L, McGee L, Tomczyk S *et al.* Biological and epidemiological features of antibiotic-resistant *Streptococcus pneumoniae* in pre- and post-conjugate vaccine eras: a United States perspective. *Clin Microbiol Rev* 2016; **29**: 525–52. <https://doi.org/10.1128/CMR.00058-15>
- Demczuk W, Martin I, Griffith A *et al.* Linear regression equations to predict β -lactam, macrolide, lincosamide, and fluoroquinolone MICs from molecular antimicrobial resistance determinants in *Streptococcus pneumoniae*. *Antimicrob Agents Chemother* 2022; **66**: e0137021. <https://doi.org/10.1128/AAC.01370-21>
- Rice P, Longden I, Bleasby A. EMBOSS: the European molecular biology open software suite. *Trends Genet* 2000; **16**: 276–7. [https://doi.org/10.1016/S0168-9525\(00\)00204-2](https://doi.org/10.1016/S0168-9525(00)00204-2)
- Blumental S, Granger-Farbois A, Moisi JC *et al.* Virulence factors of *Streptococcus pneumoniae*. Comparison between African and French invasive isolates and implication for future vaccines. *PLoS ONE* 2015; **10**: e0133885. <https://doi.org/10.1371/journal.pone.0133885>
- Azarian T, Mitchell PK, Georgieva M *et al.* Global emergence and population dynamics of divergent serotype 3 CC180 pneumococci. *PLoS Pathog* 2018; **14**: e1007438. <https://doi.org/10.1371/journal.ppat.1007438>
- Lo SW, Gladstone RA, van Tonder AJ *et al.* Pneumococcal lineages associated with serotype replacement and antibiotic resistance in childhood invasive pneumococcal disease in the post-PCV13 era: an international whole-genome sequencing study. *Lancet Infect Dis* 2019; **19**: 759–69. [https://doi.org/10.1016/S1473-3099\(19\)30297-X](https://doi.org/10.1016/S1473-3099(19)30297-X)
- Sanz JC, Rodríguez-Avil I, Ríos E *et al.* Increase of serotype 8, ST53 clone, as the prevalent strain of *Streptococcus pneumoniae* causing invasive disease in Madrid, Spain (2012–2015). *Enferm Infecc Microbiol Clin (Engl Ed)* 2020; **38**: 105–10. <https://doi.org/10.1016/j.eimc.2019.05.006>
- Hansen CB, Fuursted K, Valentiner-Branth P *et al.* Molecular characterization and epidemiology of *Streptococcus pneumoniae* serotype 8 in Denmark. *BMC Infect Dis* 2021; **21**: 421. <https://doi.org/10.1186/s12879-021-06103-w>

- 21 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. <https://doi.org/10.1016/j.vaccine.2016.03.058>
- 22 Duvvuri VR, Deng X, Teatero S et al. Population structure and drug resistance patterns of emerging non-PCV-13 *Streptococcus pneumoniae* serotypes 22F, 15A, and 8 isolated from adults in Ontario, Canada. *Infect Genet Evol* 2016; **42**: 1–8. <https://doi.org/10.1016/j.meegid.2016.04.007>
- 23 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. <https://doi.org/10.1371/journal.pone.0178040>
- 24 Golden AR, Adam HJ, Baxter M et al. Whole genome characterization of *Streptococcus pneumoniae* from respiratory and blood cultures collected from Canadian hospitals before and after PCV-13 implementation in Canada: focus on serotypes 22F and 33F from CANWARD 2007–2018. *Vaccine* 2021; **39**: 5474–83. <https://doi.org/10.1016/j.vaccine.2021.08.061>
- 25 Golden AR, Fear T, Baxter M et al. Invasive pneumococcal disease caused by serotypes 22F and 33F in Canada: the SAVE study 2011–2018. *Diagn Microbiol Infect Dis* 2021; **101**: 115447. <https://doi.org/10.1016/j.diagmicrobio.2021.115447>
- 26 Galetti P, Lo SW, Hawkins PA et al. Population genetic structure, serotype distribution and antibiotic resistance of *Streptococcus pneumoniae* causing invasive disease in children in Argentina. *Microb Genom* 2021; **7**: 000636.
- 27 Almeida SCG, Lo SW, Hawkins PA et al. Genomic surveillance of invasive *Streptococcus pneumoniae* isolates in the period pre-PCV10 and post-PCV10 introduction in Brazil. *Microb Genom* 2021; **7**: 000635.
- 28 Jung YH, Choe YJ, Lee CY et al. Impact of national pneumococcal vaccination program on invasive pneumococcal diseases in South Korea. *Sci Rep* 2022; **12**: 15833. <https://doi.org/10.1038/s41598-022-20363-9>
- 29 Kim GR, Kim E-Y, Kim SH et al. Serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* causing invasive pneumococcal disease in Korea between 2017 and 2019 after introduction of the 13-valent pneumococcal conjugate vaccine. *Ann Lab Med* 2023; **43**: 45–54. <https://doi.org/10.3343/alm.2023.43.1.45>
- 30 Xirogianni A, Marmaras N, Georgakopoulou T et al. Pneumococcal meningitis in Greece: a retrospective serotype surveillance study in the post-PCV13 era (2010–2020). *Vaccine* 2022; **40**: 5079–87. <https://doi.org/10.1016/j.vaccine.2022.07.015>
- 31 Yanagihara K, Kosai K, Mikamo H et al. Serotype distribution and antimicrobial susceptibility of *Streptococcus pneumoniae* associated with invasive pneumococcal disease among adults in Japan. *Int J Infect Dis* 2021; **102**: 260–8. <https://doi.org/10.1016/j.ijid.2020.10.017>
- 32 Suaya JA, Mendes RE, Singa HL et al. *Streptococcus pneumoniae* serotype distribution and antimicrobial nonsusceptibility trends among adults with pneumonia in the United States, 2009–2017. *J Infect* 2020; **81**: 557–66. <https://doi.org/10.1016/j.jinf.2020.07.035>
- 33 Beall B, Chochua S, Li Z et al. Invasive pneumococcal disease clusters disproportionately impact persons experiencing homelessness, injecting drug users, and the western United States. *J Infect Dis* 2022; **226**: 332–41. <https://doi.org/10.1093/infdis/jiac058>
- 34 Schillberg E, Isaac M, Deng X et al. Outbreak of invasive *Streptococcus pneumoniae* serotype 12F among a marginalized inner-city population in Winnipeg, Canada, 2009–2011. *Clin Infect Dis* 2014; **59**: 651–7. <https://doi.org/10.1093/cid/ciu366>
- 35 Deng X, Peirano G, Schillberg E et al. Whole-genome sequencing reveals the origin and rapid evolution of an emerging outbreak strain of *Streptococcus pneumoniae* 12F. *Clin Infect Dis* 2016; **62**: 1126–32. <https://doi.org/10.1093/cid/ciw050>
- 36 Choi EH, Zhang F, Lu YJ et al. Capsular polysaccharide (CPS) release by serotype 3 pneumococcal strains reduces the protective effect of anti-type 3 CPS antibodies. *Clin Vaccine Immunol* 2016; **23**: 162. <https://doi.org/10.1128/CI.00591-15>
- 37 Golden AR, Griffith A, Demczuk WHB et al. Invasive pneumococcal disease surveillance in Canada, 2020. *Can Commun Dis Rep* 2022; **48**: 396–406. <https://doi.org/10.14745/ccdr.v48i09a04>
- 38 Croucher NJ, Mitchell AM, Gould KA et al. Dominant role of nucleotide substitution in the diversification of serotype 3 pneumococci over decades and during a single infection. *PLoS Genet* 2013; **9**: e1003868. <https://doi.org/10.1371/journal.pgen.1003868>
- 39 Groves N, Sheppard CL, Litt D et al. Evolution of *Streptococcus pneumoniae* serotype 3 in England and Wales: a major vaccine evader. *Genes (Basel)* 2019; **10**: E845. <https://doi.org/10.3390/genes10110845>
- 40 Kellner JD, Ricketson LJ, Demczuk WHB et al. Whole-genome analysis of *Streptococcus pneumoniae* serotype 4 causing outbreak of invasive pneumococcal disease, Alberta, Canada. *Emerg Infect Dis* 2021; **27**: 1867. <https://doi.org/10.3201/eid2707.204403>
- 41 Egorova E, Kumar N, Gladstone RA et al. Key features of pneumococcal isolates recovered in central and northwestern Russia in 2011–2018 determined through whole-genome sequencing. *Microb Genom* 2022; **8**: mgen000851.