

Two multi-fragment recombination events resulted in the β -lactam-resistant serotype 11A-ST6521 related to Spain9V-ST156 pneumococcal clone spreading in south-western Europe, 2008 to 2016

Aida González-Díaz^{1,2}, Miguel P Machado³, Jordi Càmara^{1,2}, José Yuste^{2,4}, Emmanuelle Varon⁵, Miriam Domenech⁴, María Del Grosso⁶, José María Marimón^{2,7}, Emilia Cercenado^{2,8}, Nieves Larrosa⁹, María Dolores Quesada¹⁰, Dionisia Fontanals¹¹, Assiya El-Mniai⁵, Meritxell Cubero^{1,2}, João A Carriço³, Sara Martí^{1,2}, Mario Ramirez³, Carmen Ardanuy^{1,2,12}

1. Microbiology Department, Hospital Universitari Bellvitge, IDIBELL-UB, L'Hospitalet de Llobregat, Spain
2. Research Network for Respiratory Diseases (CIBERES), ISCIII, Madrid, Spain
3. Institute of Microbiology, Institute of Molecular Medicine, Faculty of Medicine, University of Lisbon, Lisbon, Portugal
4. Pneumococcal Reference Laboratory, Centro Nacional de Referencia, ISCIII, Madrid, Spain
5. National Reference Centre for Pneumococci, Centre Hospitalier Intercommunal de Créteil, Créteil, France
6. Infection Diseases Department, Istituto Superiore di Sanità, Rome, Italy
7. Biodonostia, Infectious Diseases Area, Respiratory Infection and Antimicrobial Resistance Group, Osakidetza Basque Health Service, Donostialdea Integrated Health Organisation, Microbiology Department, San Sebastian, Spain
8. Clinical Microbiology and Infectious Disease Department, Hospital General Universitario Gregorio Marañón, Madrid, Spain
9. Microbiology Department, Hospital Universitari Vall d'Hebron, UAB, Barcelona, Spain
10. Microbiology Department, Clinical Laboratory North Metropolitan Area, Hospital Universitari Germans Trias i Pujol, UAB, Badalona, Spain
11. Microbiology Department, Hospital Universitari Parc Taulí, Sabadell, Spain
12. Department of Pathology and Experimental Therapeutics, School of Medicine, University of Barcelona, Barcelona, Spain

Correspondence: Carmen Ardanuy (c.ardanuy@bellvitgehospital.cat)

Citation style for this article:

González-Díaz Aida, Machado Miguel P, Càmara Jordi, Yuste José, Varon Emmanuelle, Domenech Miriam, Del Grosso María, Marimón José María, Cercenado Emilia, Larrosa Nieves, Quesada María Dolores, Fontanals Dionisia, El-Mniai Assiya, Cubero Meritxell, Carriço João A, Martí Sara, Ramirez Mario, Ardanuy Carmen. Two multi-fragment recombination events resulted in the β -lactam-resistant serotype 11A-ST6521 related to Spain9V-ST156 pneumococcal clone spreading in south-western Europe, 2008 to 2016. Euro Surveill. 2020;25(16):pii=1900457. <https://doi.org/10.2807/1560-7917.ES.2020.25.16.1900457>

Article submitted on 11 Jul 2019 / accepted on 01 Oct 2019 / published on 23 Apr 2020

Background: The successful pneumococcal clone Spain^{9V}-ST156 (PMEN₃) is usually associated with vaccine serotypes 9V and 14. **Aim:** Our objective was to analyse the increase of a serotype 11A variant of PMEN₃ as cause of invasive pneumococcal disease (IPD) in Spain and its spread in south-western Europe. **Methods:** We conducted a prospective multicentre study of adult IPD in Spain (2008–16). Furthermore, a subset of 61 penicillin-resistant serotype 11A isolates from France, Italy, Portugal and Spain were subjected to whole genome sequencing (WGS) and compared with 238 genomes from the European Nucleotide Archive (ENA). **Results:** Although the incidence of serotype 11A in IPD was stable, a clonal shift was detected from CC62 (penicillin-susceptible) to CC156 (penicillin-resistant). By WGS, three major 11A-CC156 lineages were identified, linked to ST156 (n=5 isolates; France, Italy and Portugal), ST166 (n=4 isolates; France and Portugal) and ST838/6521 (n=52 isolates; France, Portugal and Spain). Acquisition of the 11A capsule allowed to escape vaccine effect. AP200 (11A-ST62) was the donor for ST156 and ST838/6521 but not for ST166. In-depth analysis of ST838/6521 lineage showed two multi-fragment recombination events including four and seven fragments from an 11A-ST62 and an NT-ST344 representative, respectively.

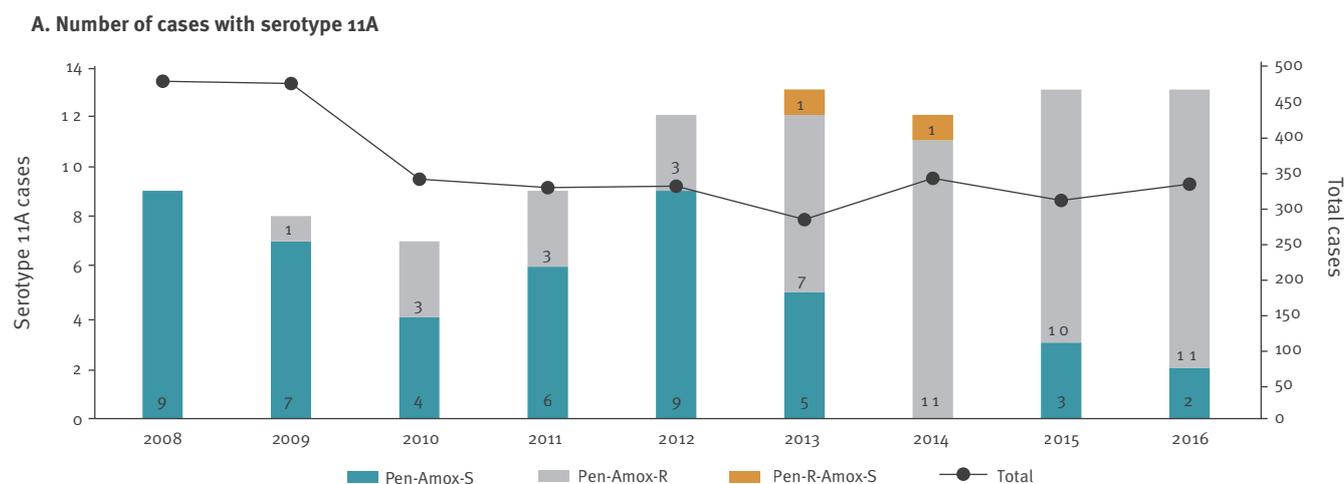
Conclusion: The increase in penicillin-resistant serotype 11A IPD in Spain was linked to the spread of a vaccine escape PMEN₃ recombinant clone. Several recombination events were observed in PMEN₃ acquiring an 11A capsule. The most successful 11A-PMEN₃ lineage spreading in south-western Europe appeared after two multi-fragment recombination events with representatives of two major pneumococcal clones (11A-ST62 and NT-ST344).

Introduction

Streptococcus pneumoniae is an important human pathogen that usually colonises the upper respiratory tract. Pneumococci can invade sterile sites causing different invasive pneumococcal diseases (IPD) such as bacteraemic pneumonia, meningitis or primary bacteraemia. Furthermore, pneumococci cause other non-invasive diseases such as acute otitis media in children or acute exacerbations of COPD in adults [1,2]. The polysaccharide capsule, which presents one of more than 97 serotypes, is the main virulence factor and its diversity has been associated with differences in invasiveness and mortality. Pneumococcal conjugated vaccines (PCV) include the serotypes more commonly found among IPD. In Spain, three conjugate vaccines had been licensed: the first one in 2001,

FIGURE 1

Total and pneumococcal serotype 11A cases, associated clonal complex and resistance, south-western Europe, 2008-2016 (n = 96)



B. Resistance characteristics

	Penicillin MIC (mg/L)	Amoxicillin MIC (mg/L)	CC	Related STs	Other resistance	Period
Pen-Amox-S	≤ 0.06	≤ 0.5	62	53, 62, 408, 4305, 8904	Erythromycin, Cotrimoxazole	2008-2013, 2015-2016
Pen-R-Amox-S	0.25	0.25-0.5	1010	1010	Cotrimoxazole	2013-2014
Pen-Amox-R	1-4	2-8	156	838, 6521	Cotrimoxazole	2009-2016

Amox: amoxicillin; CC: clonal complex; IPD: invasive pneumococcal disease; MIC: minimum inhibitory concentration; Pen: penicillin; R: resistance; S: susceptible; ST: sequence type.

The recombinant clone ST6521-serotype 11A is spreading in Europe.

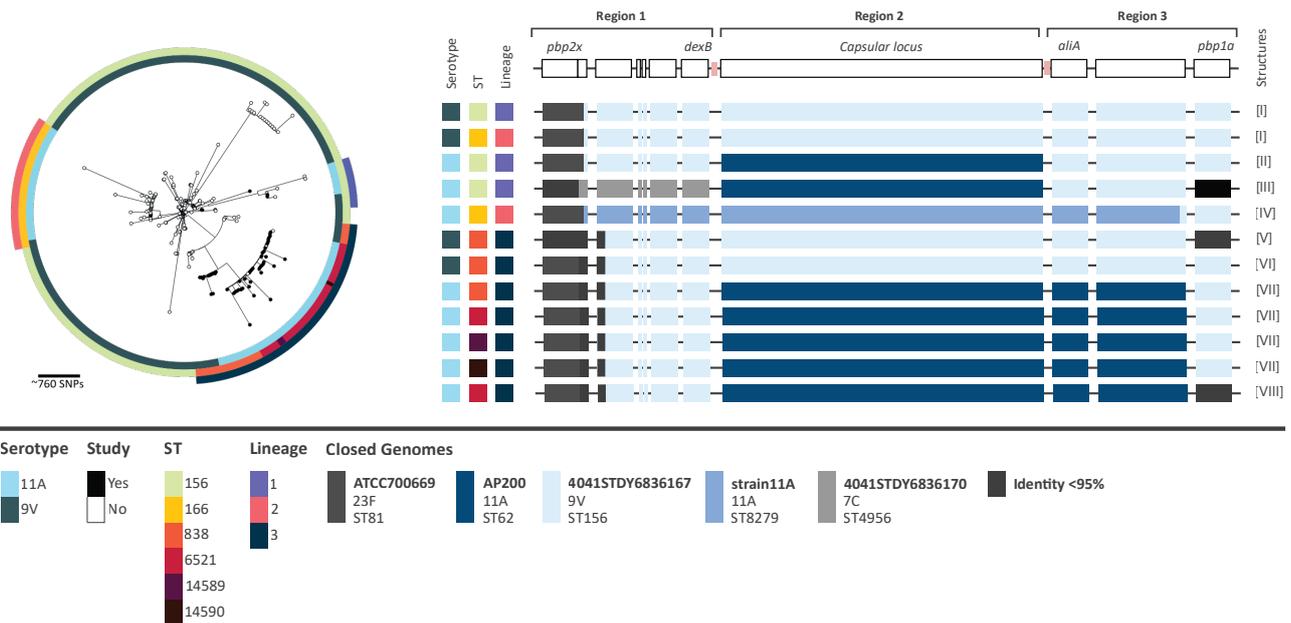
PCV7, targeted serotypes 4, 6B, 9V, 14, 18C, 19F and 23F; the second in 2009, PCV10, added serotypes 1, 5 and 7F and PCV13 in 2010 added serotypes 3, 6A and 19A [3]. The natural ability of *S. pneumoniae* to undergo genetic transformation allows pneumococci to change the capsule (capsular switching) which allows them to escape vaccine pressure. The emergence of new recombinant clones along with serotype replacement has changed the serotype distribution after the introduction of paediatric vaccination [4,5]. Among them, an increase in serotype 11A, not included in PCV13, has been reported worldwide [1,4,6,7]. Many studies suggest that serotype 11A has a low invasive potential [5] being present mainly in child carriers [8] and patients with chronic obstructive pulmonary disease (COPD) [9]. Serotype 11A has classically been linked to the antibiotic-susceptible lineage CC62 [10], with the exception of some macrolide-resistant isolates with the M phenotype (resistance to 14- and 15-membered ring macrolides) [11]. In 2005, the emergence of a penicillin- and amoxicillin-resistant serotype 11A lineage related to CC156 was identified [10]. In our setting,

serotype 11A isolates related to the CC156 clone were the leading cause of β-lactam resistance in the years 2015 and 2016 [12]. Nowadays, in Spain, there are two main serotype 11A sequence types (ST) derived from CC156: ST838, first detected in 2005, which is a single locus variant (SLV) of ST156; and ST6521 which is a double locus variant (DLV) of ST156 emerged in 2009 [10]. These STs are related to clonal complex CC156 also known as the Spain^{9V}-ST156 clone (PMEN3). This clone, originally associated with serotype 9V, has been related to capsular switching since the 1990s when the acquisition of the capsule 14 led to a worldwide spread of this serotype. In fact, most of serotype 14 isolates causing IPD in the pre-PCV era were related to the PMEN3 recombinant lineage [13,14]. Besides serotype 14, PMEN3 recombined with other serotypes showing a high capacity to exchange the capsular locus, resulting in the evasion of the vaccine effects [15].

In this study, we firstly studied the increase of β-lactam-resistant serotype 11A pneumococci as the cause of adult IPD in the framework of a multicentre study in

FIGURE 2

Capsular switching and recombination breakpoints, pneumococcal isolates, south-western Europe



BLAST: basic local alignment search tool; ENA: European Nucleotide Archive; MLST: multilocus sequence typing; NCBI: National Center for Biotechnology Information; SNP: single nucleotide polymorphism; ST: sequence type.

Panel A: The maximum likelihood phylogenetic tree includes the 61 sequenced strains and 238 strains of serotypes 9V and 11A for which genomes were available in ENA and which belonged to ST related to CC156 (ST156, ST166, ST838 or ST6521). The colour of each circle represents different characteristics of the strains. Outer circle: lineages; middle circle: ST; inner circle: serotypes.

Panel B: The analysis of the capsular operon was done with 61 genomes from this study and a selection of 59 from the 238 ENA genomes. The different structures are numbered I to VIII. Colours are related to the closest fully closed genome on NCBI (indicated in the legend with serotype and MLST). Black regions: areas with identity <95% with NCBI fully closed genomes after BLAST search; pink squares: non-analysed flanking transposases.

Spain. Secondly, using whole genome sequencing (WGS), we analysed the spread of the recombinant clone S11A-ST6521 including neighbouring countries in south-western Europe which also detected penicillin-resistant serotype 11A isolates.

Methods

Study design and bacterial characterisation

This study was initially conducted in the framework of a laboratory-based multicentre study of adult (≥ 18 years-old) IPD patients, involving six Spanish hospitals. An IPD episode was defined as the isolation of *S. pneumoniae* from a normally sterile body fluid and only one isolate per episode was included [16]. We included IPD episodes caused by serotype 11A from 2008 to 2016.

Serotyping by dot blot assay or Quellung reaction was done at the Spanish Reference Laboratory for Pneumococci (SRLP). The antibiotic susceptibility to seven antimicrobials (amoxicillin, cefotaxime, cotrimoxazole, erythromycin, levofloxacin, penicillin and

tetracycline) was tested by microdilution following the European Committee on Antimicrobial Susceptibility Testing (EUCAST) guidelines [17]. Genotyping was performed by pulse field gel electrophoresis (PFGE) and/or multilocus sequence typing (MLST) as described previously [18,19].

Whole-genome sequencing and assembly

To evaluate the spread of the serotype 11A multidrug-resistant clone, a total of 61 penicillin-resistant (≥ 0.12 mg/L) serotype 11A isolates were selected for WGS. Of these, 21 were collected from the multicentre study and the remaining 40 were obtained from the Reference Laboratory for Pneumococci or Pneumococcal Reference Groups from France (n=17), Italy (n=4), Portugal (n=4) and Spain (n=16). Metadata including demographic characteristics, WGS-derived data and antibiotic minimum inhibitory concentrations (MIC) are summarised in Supplementary Table S1. A diagram of the study design and a flowchart of the bioinformatic analysis are also summarised in Supplementary Figure S1.

Bacteria were grown overnight in 5% sheep blood agar at 37 °C +5% CO₂. DNA was extracted using the QIAamp DNA Blood Mini Kit (Qiagen, Hilden, Germany) and quantified using the QuantiFluor dsDNA System (Promega, Wisconsin, United States (US)). Illumina paired-end libraries (2 × 150 bp) were prepared with the Nextera XT kit and sequenced on Illumina MiSeq Platform (Illumina Inc., San Diego, US). Read quality assessment and genome assembly was done using the INNUca v3.2 pipeline (<https://github.com/B-UMMI/INNUca>) through `ummidock/innuca:3.2-01`. Firstly, a quality control of the reads was performed using `FastQC vo.11.5` (<http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) and reads were cleaned and trimmed with `Trimmomatic vo.36` [20]. The genome was assembled using `SPAdes v3.11.0` [21] and subsequently polished using `Pilon v1.18` [22]. The in silico MLST was determined using `Seemann's MLST v2.11` (<https://github.com/tseemann/mlst>). The final assemblies were annotated by `Prokka v1.12` [23] through `ummidock/prokka:1.12` Docker image. Reads were deposited at the European Nucleotide Archive (ENA) with accession numbers summarised in Supplementary Table S1.

Comparative analysis by whole genome sequencing

Molecular antibiotic resistance mechanisms

The in silico analysis of gene mutations involved in antibiotic resistance (*pbp1a*, *pbp2x*, *pbp2b*, *parC*, *parE*, *gyrA*, *folA* and *folP*) was done manually using Geneious R9 (Biomatters, Auckland, New Zealand) and the R6 genome (NC_003098) as reference. The acquired resistance mechanisms were screened using `Abricate vo.8.0` (<https://github.com/tseemann/abricate>) through `flowcraft/abricate:0.8.0-3` Docker image for the Comprehensive Antibiotic Resistance Database (CARD) [24] and `ResFinder` [25] databases.

Phylogenetic analysis

To better explore the diversity and relationship with other pneumococci with related serotype and ST to the penicillin-resistant 11A isolates included in this study, 47,529 available *S. pneumoniae* genomes deposited in ENA were downloaded on 29 October 2018 using `getSeqENA v1.3` (<https://github.com/B-UMMI/getSeqENA>) with `Aspera Connect v3.7.2.141527` (<https://aspera-software.com/software/clients/connect/>). In case of large fastq file sizes, they were first downsampled for an estimated depth of coverage of 100 × with the `sample_fastq` script commit `01acodb` available at https://github.com/jacarrico/sample_fastq and making use of `Seqtk v1.2-r94`, (<https://github.com/lh3/seqtk>). The serotype was in silico deduced using `Seroba v1.0.1` [26]. Among the 47,529 available genomes, we selected 238 with serotypes 9V or 11A and with one of the following MLST: ST156, ST166, ST838 or ST6521 (Supplementary Table S2). Phylogenetic analysis was performed by constructing an assembly-based core genome-single nucleotide polymorphism (SNP) phylogenetic tree with

the default parameters of *Parsnp* from the Harvest suite [27] with the exception of parameter 'x' which identifies and removes recent recombination using `PhilPack` [28]. Isolate ERR2681167 was used as reference. Phylogenetic tree visualisation was done using `Microreact` [29].

Capsular operon analysis

Three contiguous regions were analysed: region 1 (*pbp2x* to *dexB*), region 2 (capsular locus) and region 3 (*aliA* to *pbp1a*). Because the capsular operon is flanked by transposases, which are difficult to assemble using the small reads generated by Illumina methodology, these regions were located in different contigs, so it was not possible to analyse the whole region from *pbp2x* to *pbp1a*. A BLASTn search on the National Center for Biotechnology (NCBI) website (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was performed to determine which deposited fully closed genome presented the highest identity with the isolates included in the study. In addition, region 2 was compared with reference capsular operons (1813/39-11A (CR931653) and 980/60-9V (CR931648)) previously described [30].

Identification of recombination regions

To determine the genomic differences between ST838 and ST6521 isolates, a maximum-likelihood phylogenetic tree was constructed with the 68 isolates belonging to these two ST, 52 from this study and 16 from ENA, using `RaxML` [31]. The 9V-ST838 (ERR2303060) draft assembled genome was concatenated by mapping the contigs against 4041STDY6836167 (NZ_LS483448), which is a closed serotype 9V, ST156 genome using `Mauve` [32]. The 9V-ST838 concatenated genome was used as a reference and the 68 reads belonging to lineage 3 isolates were mapped using `Snippy 3.1` (<https://github.com/tseemann/snippy>). An assembly-free core-SNP alignment was done with `Snippy's` core module (`snippy-core`). The alignment was used to identify the recombinant regions through the evaluation of the density of base substitution using `Gubbins v2.3.4` [33] and the results were visualised in `Phandango` [34]. The allelic variation of fragment exchange in recombination events were manually examined with `Geneious R9` and the highest identity with *S. pneumoniae* genomes available on NCBI was explored using BLASTn search (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>).

Core and accessory genome

The core genes, which are shared by all isolates, and accessory genes of the 68 isolates belonging to ST838 and ST6521 were analysed using `Roary v3.12.0` [35], with a minimum percentage identity of 80% for BLASTp. Identification of accessory genes segregating each ST was done by `Scoary v1.16.16` [36].

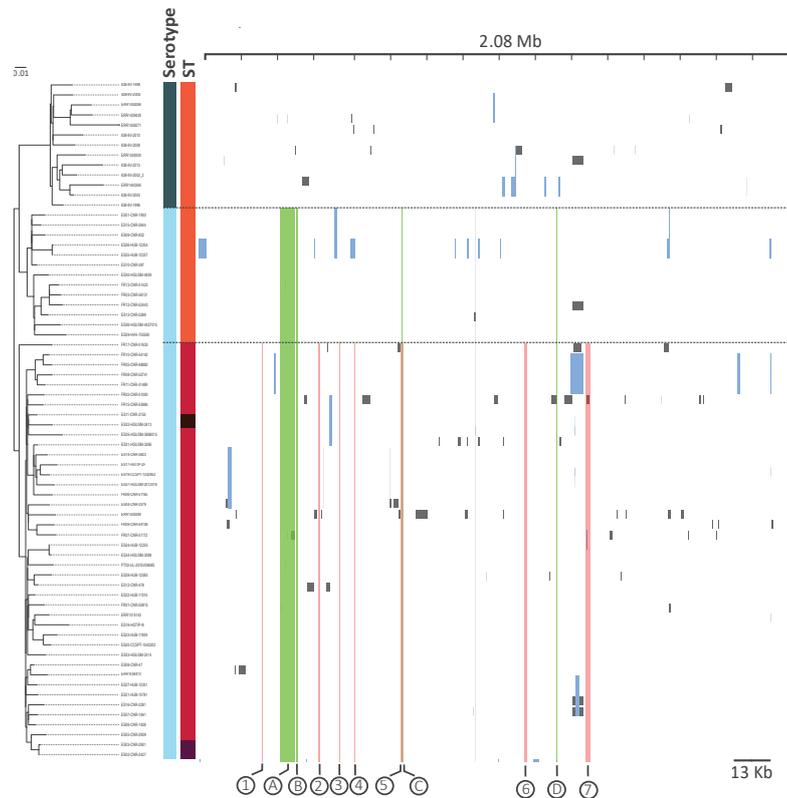
Ethical statement

This project has been approved by the Clinical Research Ethics Committee of the Hospital de Bellvitge (PR153/18). Written informed consent was not required as this was an observational study with isolates

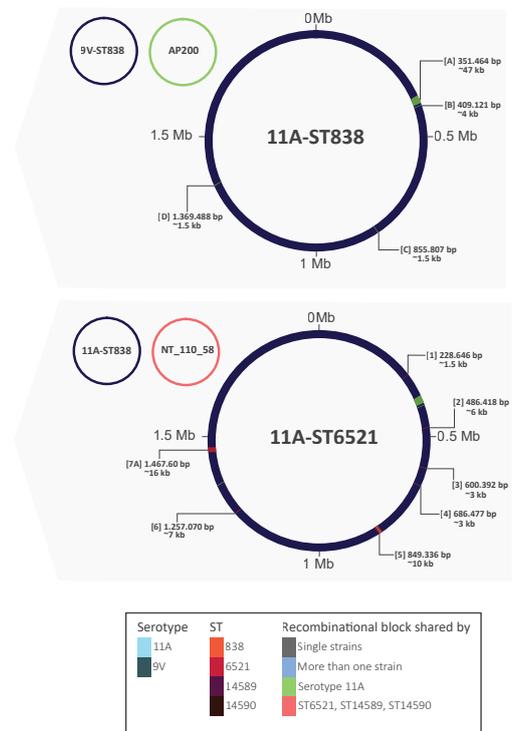
FIGURE 3

Analysis of recombination in lineage 3 (serotypes 9V and 11A), pneumococcal isolates, south-western Europe (n = 68)

A. Phylogenetic incorporating serotype, ST and recombination events detected by Gubbins



B. Multi-fragment recombination (1 to 7 and A to D) between receptor (9V-ST838) and two putative donors (11A-ST62 and NT-ST344)



ST: sequence type.

Panel A. Green blocks: recombination segments for which putative donor is 11A-ST62 (AP200) and which are shared by strains of ST838, ST6521, ST14589 and ST14590; these blocks are labelled with capital letters (A to D). Red blocks: recombination segments which putative donor is NT-ST344 which are shared by strains ST6521, ST14589 and ST14590; these blocks are labelled with numbers (1 to 7). Grey blocks: recombination segments present in a single strain. Light blue blocks: recombination segments shared by small groups of strains.

Panel B. Dark blue: receptor 9V-ST838; green: putative receptor 11A-ST62; red: putative receptor NT-ST344. Coordinates refer to the ones of the genome of strain 4041STDY6836167. The approximate recombining fragment size is indicated.

obtained as part of the normal microbiological routine. Patient confidentiality was always protected; all data were anonymised and protected according to national normative.

Results

A clonal shift has occurred among serotype 11A isolates causing adult invasive pneumococcal disease in Spain

Among 3,200 adult IPD episodes detected during the study period, 96 were caused by serotype 11A isolates. We observed a non-significant increase in the incidence of serotype 11A IPD from 0.24 episodes per 100,000 persons per year in 2008 to 0.36 episodes per 100,000 persons per year in 2016 ($p=0.22$; incidence risk ratio = 0.66; 95% confidence interval (CI): 0.37–1.20). However, a clonal replacement linked to an

increase in penicillin-resistant isolates was observed. While in 2008, no penicillin-resistant serotype 11A isolates were detected, 11 of 13 IPD episodes due to serotype 11A were caused by penicillin-resistant isolates in 2016 (Figure 1). The PFGE/MLST analysis revealed two major clonal complexes: the penicillin-susceptible CC62 and the penicillin-resistant CC156. Isolates of CC62 were mostly antibiotic-susceptible (including to penicillin) except for several cotrimoxazole- (26 of 44; 59.1%) or erythromycin-resistant (13 of 44; 29.5%) isolates. Macrolide resistance was due to the M phenotype, as previously described among CC62 isolates in Spain [11]. On the other hand, isolates of CC156 showed a common resistance pattern, being penicillin-resistant (MIC range: 1–4 mg/L) and cotrimoxazole-resistant (MIC range: 1/19 –>2/38 mg/L). Moreover, CC156 isolates were mostly amoxicillin-resistant with a MIC range from 2 to 8 mg/L (Figure 1).

Whole genome sequencing of penicillin-resistant serotype 11A from south-western Europe

For an in-depth genomic analysis of penicillin-resistant serotype 11A isolates related to CC156, a collection of 61 penicillin-resistant serotype 11A isolates isolated from patients with IPD in France, Italy, Portugal and Spain were subjected to WGS (Supplementary Figure S1). These isolates were recovered from 25 regions of the four European countries: France (n=9), Italy (n=2), Portugal (n=3), and Spain (n=11) (Supplementary Figure S2).

After in silico MLST analysis, six different ST were found: ST156, ST166 (SLV of 156), ST838 (SLV of ST156), ST6521 (SLV of ST838 and DLV of ST156), a new SLV of ST6521 (ST14589) and a new DLV of ST6521 (ST14590). Supplementary Figure S3 shows the phenotypic (MIC) and genotypic (resistome) profiles of the 11A-CC156 isolates and their association with the ST. The allelic profiles of their penicillin-binding proteins (PBP) are described in Supplementary Table S3, a name was assigned to each allele based on the previous definition by the US Centers for Disease Control and Prevention (CDC) [37,38]. The differences between ST in resistance to β -lactams were linked to changes in PBP. A common PBP1A allele (allele 15) identical to that of PMEN1 (American Type Culture Collection 700669) was shared by all but three isolates (the latter presented allele 7 and a new allele (NEW1) close to allele 96 (96% identity)). We found the changes at the STMK373 motif (T371A allele 15 and T371S in allele NEW1) that had been previously described [15]. All but two isolates presenting allele 15, presented PBP2X allele 18 and a new allele (NEW1) related to allele 36 (98.9% identity). Both alleles differed in only three amino acids and have the 337S A MK substitution. Two distinct PBP2B alleles were found segregating lineages ST156 and ST166 from ST838 and ST6521. All isolates presented the 443SSN A change in PBP2B. Moreover, isolates with allele 76 had an additional 10 changes between residues 590 and 641 related to increased amoxicillin MIC [15].

Macrolide-resistance was associated with the presence of integrative conjugative elements. All four ST166 isolates carried the Tn6002 transposon (AY898750) described previously in *S. pneumoniae*. This is a Tn916-like structure containing the *tet(M)* gene and the macrolide-aminoglycoside-streptothricin (MAS) element *erm(B)*, conferring resistance to tetracycline and to erythromycin and clindamycin, macrolide-lincosamide-streptogramin b (MLS_B) phenotype, respectively. One 11A-ST6521 isolate carried a mobile element, the *mef(E)* gene in the macrolide efflux genetic assembly (MEGA) element. All isolates were resistant to cotrimoxazole, a characteristic of the Spain^{9V}-ST156 (PMEN3) clone. All but one had the I100L amino acid substitution in the dihydrofolate reductase known to confer resistance to trimethoprim. In addition, all isolates had a single insertion (S₆₂SSY) or double insertion

(S₆₂SSYS; in ST166 isolates) in the dihydropteroate synthase, known to confer resistance to sulfamethoxazole.

Different capsular switching events identified among the major serotype 11A lineages

Of the 238 selected genomes from ENA, 198 were serotype 9V (n=183 ST156, n=2 ST166 and n=13 ST838) and the remaining 40 were serotype 11A (n=5 ST156, n=32 ST166 and n=3 ST6521). A phylogenetic tree was constructed using these 238 genomes and the 61 genomes of the present study (Figure 2A, Supplementary Table S2). Three major lineages, including both 9V and 11A isolates, were defined including one or two closely related ST: lineage 1 (ST156), lineage 2 (ST166) and lineage 3 (ST838 and ST6521). To identify putative recombination events in the capsular locus we used all isolates, except in lineage 1, which accounted for the highest number of genomes (n=188). In order to simplify the analysis of the capsular switching event in this lineage we selected all five 11A-ST156 isolates from this study and the four phylogenetically closest 9V-ST156 isolates (Supplementary Table S2).

The analysis of the capsular locus and the flanking regions (region 1: *pbp2x-dexB*; region 2: capsular locus; region 3: *aliA-pbp1a*) in all lineages revealed eight different structures (numbered I to VIII; Figure 2B). Structure definition was based on per cent identity with complete genomes available in NCBI, where sequence stretches with identity lower than 95% were coloured in black. Among serotype 9V, three different structures were found (I, V and VI), the main differences among them were the region downstream of *pbp2x* in structures V and VI, and *pbp1a* in structure V. On the other hand, five different structures (II, III, IV, VII and VIII) were found among serotype 11A isolates. Two of them (II and III) were detected in lineage 1. Structure III presented different *pbp2x* and *pbp1a* genes and a different sequence in region 1, with high identity with the 4041STDY6836167 genome (NZ_LS483448), a serotype 7C-ST4956 isolate. All lineage 2 isolates had structure IV, which was closely related to that of an 11A isolate named strain11A (NZ_CP018838). However, the capsular locus (region 2) of structure IV had only 92% identity with the 11A capsule locus of the reference isolate 1813/39-11A [30]. Structure IV showed high identity (>99%) with isolate PMP1342 (MF140334.1), a genetic variant of the 11A locus, with the exception of the *wzg* gene (identical to that of the 980/60-9V reference isolate). Finally, all but one of the lineage 3 isolates shared structure VII. In this structure the DNA fragment integrated from AP200 (11A-ST62) included the capsular operon (region 2) and a portion of the region 3 but not the *pbp1a*, which was similar to that of 4041STDY6836167, the 9V-ST156 isolate (Figure 2B). Structure VIII was found in a single isolate and differed from structure VII in the *pbp1a* gene. There were amino acid changes in some genes of the capsular locus in all structures which are summarised in Supplementary Table S4.

Several results of recombination events are present in the successful 11A-ST6521 lineage

A scan for possible recombination events across the entire genome and identification of the possible donors was done with the 68 isolates of lineage 3, corresponding to the most successful emerging penicillin-resistant serotype 11A isolates. Besides several possible recombinations identified among single isolates or minor clusters, we found two major recombination events (Figure 3). We considered as the donor the isolate which provided the capsule to the new recombinant isolate and as the recipient the isolate which incorporated the capsule sequence in its genetic background. The first recombination event involved a 9V-ST838 isolate as the recipient and a serotype 11A-ST62 isolate related to AP200 (NC_014494) as the donor. In this event, the capsule which allow to escape the vaccine effect (ca 16 kb, included in region A) and three additional regions (B–D) of the AP200-like isolate were incorporated into the 9V-ST838 genome (represented in green in Figure 3). Region A was a ca 50 kb recombination region which on closer analysis appeared to be discontinuous. This region included four segments of the AP200-like isolate (26 kb, 12 kb, 4.5 kb and 0.5 kb) interspersed with two small segments of the 9V-ST838 isolate (2.5 kb and 4 kb), including the *pbp1a* gene. There was an additional 2 kb segment (between the 12 kb and 4.5 kb segments of the AP200-like isolate) different from both isolates which could be characteristic of the actual donor. Region A could therefore have resulted from a multi-fragment sequence integration. On the other hand, the recombination in the B, C and D regions incorporated fragments of ca 4 kb, 1.5 kb and 1.5 kb, respectively. A second multi-fragment recombination event possibly involved the recombinant 11A-ST838 isolate as recipient and a representative of the worldwide disseminated non-typable ST344 lineage (Figure 3). We identified seven fragments with high similarity to the NT-ST344 isolate NT_110_58 (NZ_CP007593) in the recipient genome (11A-ST838). One of these fragments included the *aroE* gene, explaining the variation in the MLST profile.

Allelic variation between ST838 and ST6521 in accessory genome and core genes

In a previous study, we showed different behaviours in evasion of host defenses and in biofilm formation between the two major ST of serotype 11A lineage 3 (11A-ST838 and 11A-ST6521) [10]. In lineage 3, core genes represented around 70% of the total genes detected. Among the 1,707 core genes, the region A multi-fragment recombination introduced 25 allelic variants into the recipient genome, in addition to the serotype 11A specific *cps* locus genes which are not part of the core genome. Furthermore, the second recombination event between 11A-ST838 and NT_110_58 incorporated 37 allelic variants into the seven exchanged fragments differentiating 11A-ST6521 from 11A-ST838. Most of the core genes presented just one allele, while others presented two or more alleles. Most of the allelic differences among core genes did not segregate

serotypes, ST or sub-lineages, with the exception of the allelic differences incorporated in the recombination events. The 62 allelic variations potentially introduced by recombination are summarised in Supplementary Table S5. The analysis of the accessory genome performed with Scoary revealed one gene that segregated ST6521 from ST838 ($p < 0.01$). This gene was a Gcn5-related N-acetyltransferase (spnnt_RS02275) which was truncated in the ST838 isolates following to a single nucleotide deletion.

Discussion

Current prevention of pneumococcal infections is based on vaccination and vaccines are serotype-dependent, offering protection against a subset of known serotypes. Only 13 serotypes are included in the conjugate vaccine with the highest available valency, PCV13; its widespread use changed the worldwide serotype distribution [1]. During a multicentre study, a clonal replacement was observed among invasive pneumococci of serotype 11A (not included in PCV13), associated with the emergence of β -lactam resistance. This new lineage, related to Spain^{9V}-ST156, was penicillin- and amoxicillin-resistant hampering the treatment of severe infections. Besides IPD, this amoxicillin-resistant variant of serotype 11A has in Spain caused acute exacerbations of COPD patients [9] and otitis media in children [39], infections for which amoxicillin usually is the first-choice antibiotic therapy.

We analysed by WGS 61 isolates, from Spain and other south-western European countries where penicillin-resistant 11A isolates have also been detected. Through this bioinformatic analysis, we tried to reconstruct the recent dynamics of the Spain^{9V}-ST156 clone in south-western Europe. We showed that the 11A-ST6521 lineage was present in France, Portugal and Spain. Two additional 11A lineages (ST156 and ST166) were identified among French, Italian and Portuguese isolates. However, the β -lactam MIC were lower than those of ST838 and ST6521, in which high amoxicillin MIC are related to amino acid changes in the transpeptidase domain of PBP2B (residues 590–641), as described [15].

Two different recombinant structures were detected in lineage 1 (11A-ST156) for which we could not discard a common origin. Probably, there was a first recombination with an AP200-like isolate, followed by a second recombination that incorporated a fragment between *pbp2x* and *dexB* which presents high identity with a 7C-ST1623 isolate. But this was not unequivocally shown by our analysis. Moreover, a second lineage (11A-ST166) was detected in France and Portugal, as well as in several isolates deposited in the ENA, showing yet different recombination events. This lineage presented a variant of the serotype 11A capsular operon previously identified in Fiji [40]. The presence of a double insertion in the dihydropteroate synthase and the Tn6002 transposon carrying *tet(M)* and *erm(B)* was a hallmark of all four ST166 isolates suggesting

that another horizontal DNA transfer event was pivotal in conferring antimicrobial resistance to lineage 2 (11A-ST166). Although the number of isolates included in the present work was low, the detection of this lineage in France and Portugal and its multidrug resistance deserve further surveillance.

Another recombination event involved lineage 3 (11A-ST838 giving rise to 11A-ST6521). Besides capsule and PBPs, other regions involving single isolates, minor or major clusters, accounted for the genetic diversity of this lineage. The presence of changes in a single isolate could suggest a patho-adaptative process in the course of invasive disease that was not enough to confer the ability to interpatient spread [41]. On the other hand, the presence of the same change in most isolates from the same ST suggests that it may have a beneficial fitness effect on their spread.

The careful analysis of ST838 and ST6521 serotype 11A isolates identified two subsequent multi-fragment recombination events. The first one occurred between an AP200-like isolate [42] and a 9V-ST838. The major consequence of this recombination was the capsular switch that allowed this lineage to escape vaccine-induced immunity. However, in a scenario of only vaccine pressure, a spread of 11A-ST62 could have been hypothesised. Therefore, it seems that the genetic characteristics of the Spain^{9V}-ST156 clone such as the β -lactam resistance in an area with high-level antibiotic consumption were determinants for its success. Probably both vaccine and antibiotic pressure could favour this drift. Besides the capsular operon, three additional fragments from the putative AP200-like donor were integrated which may have resulted in still unidentified but important phenotypic changes. The second event occurred with an ST344 isolate, a non-encapsulated worldwide disseminated clone identified in nasopharyngeal colonisation and non-invasive infections, particularly conjunctivitis [43,44]. In this recombination, seven fragments of the ST344 representative were acquired by 11A-ST838. Although we could not exclude the possibility that these multi-fragment acquisitions occurred in different events from other pneumococci, our results strongly suggest that the two multi-fragment recombinations occurred with two isolates related to AP200 and NT_110_58. The ability of PMEN3 to acquire multiple large genomic regions has been described previously, with a recombinant having acquired 5.3% of its genome from a PMEN1 representative [45]. Recently, it has been suggested that large recombination events are favoured in environments allowing stable cell-to-cell contact, such as colonisation or biofilm-associated infections [46]. Serotype 11A is currently a major serotype colonising children [7], and the serotype replacement in colonisation which occurred after PCV introduction could have favoured the widespread emergence of these 11A-PMEN3 variants. The incorporation of further genetic material from an ST344 representative could offer clues for the spread of 11A-ST6521. Among this acquired DNA,

there are genes involved in biofilm formation, such as *lytB* [47]. Possibly, this second recombination could be behind the higher capacity to form biofilms and to evade the host immune system of the recombinant 11A-ST6521 lineage which was previously described [10]. These traits can represent a patho-adaptive advantage of colonisation and also of causing biofilm-related diseases such as otitis media or acute exacerbations of COPD [9,10], possibly contributing to the resilience of PMEN3 lineages as a cause of pneumococcal disease in the PCV era.

Conclusion

We describe a clonal shift among serotype 11A isolates causing IPD in Spain and the spread of a recombinant clone through different European countries. The first event allowed the β -lactam resistant clone to remain as cause of pneumococcal diseases and to escape the current vaccine. The second event gave adaptative advantages to cause disease. Probably both vaccine and antibiotic pressure could favour this drift. Further studies are needed to determine how much this lineage has increased and what impact the multidrug resistance of this vaccine escape recombinant clone will have on the therapy of IPD and other pneumococcal diseases in south-western Europe. Moreover, the ability of the Spain^{9V}-ST156 clone to evolve, resulting in new recombinant lineages with new serotypes, implies a need to monitor this invasive clone and its possible drifts in the future.

Acknowledgements

This study was supported by grants from Fondo de Investigaciones Sanitarias de la Seguridad Social (PI14/00627; PI18/00339, INT 15/0186; INT16/0117) and from Centro de Investigación Biomédica en Red (CIBER) de Enfermedades Respiratorias (CIBERES) CBo6/06/0037), an initiative of the Instituto de Salud Carlos III, Madrid, Spain and ONEIDA project (LISBOA-01-0145-FEDER-016417) co-funded by FEEL – ‘Fundos Europeus Estruturais e de Investimento’ from ‘Programa Operacional Regional Lisboa 2020’ and by national funds from FCT – ‘Fundação para a Ciência e a Tecnologia’. Financial support was also provided by the European Regional Development Fund (ERDF). We thank CERCA Programme /Generalitat de Catalunya for institutional support.

Conflict of interest

CA received research funding from Pfizer, unrelated to the present study. MR received honoraria for serving on the speaker's bureau of Pfizer and for consulting for GlaxoSmithKline and Merck Sharp and Dohme. All other authors declare no conflicts of interest.

Authors' contributions

Study design: AGD, JC, MR, CA. Strain collection: AGD, JY, EV, MDG, JMM, EC, NL, MDQ, DF, MR, CA. Experimental work: AGD, JY, EV, MDG, EC, NL, JMM, MDQ, DF, MD, AEM, MC. Bioinformatic analysis: AGD, MPM, JAC, MR, CA. Data analysis: AGD, MPM, SM, JC, MR, CA. Manuscript drafting:

References

1. Càmara J, Marimón JM, Cercenado E, Larrosa N, Quesada MD, Fontanals D, et al. Decrease of invasive pneumococcal disease (IPD) in adults after introduction of pneumococcal 13-valent conjugate vaccine in Spain. *PLoS One*. 2017;12(4):e0175224. <https://doi.org/10.1371/journal.pone.0175224> PMID: 28384325
2. Grau I, Ardanuy C, Cubero M, Benitez MA, Liñares J, Pallares R. Declining mortality from adult pneumococcal infections linked to children's vaccination. *J Infect*. 2016;72(4):439-49. <https://doi.org/10.1016/j.jinf.2016.01.011> PMID: 26868606
3. Càmara J, Ardanuy C. Pneumococcal disease and conjugate vaccines. *Enferm Infecc Microbiol Clin*. 2018;36(10):605-6. <https://doi.org/10.1016/j.eimc.2018.07.012> PMID: 30220517
4. Golden AR, Adam HJ, Karlowsky JA, Baxter M, Nichol KA, Martin I, et al. Molecular characterization of predominant *Streptococcus pneumoniae* serotypes causing invasive infections in Canada: the SAVE study, 2011-15. *J Antimicrob Chemother*. 2018;73(7) suppl. 7:i120-31. <https://doi.org/10.1093/jac/dky157> PMID: 29982573
5. Sá-Leão R, Pinto F, Aguiar S, Nunes S, Carriço JA, Frazão N, et al. Analysis of invasiveness of pneumococcal serotypes and clones circulating in Portugal before widespread use of conjugate vaccines reveals heterogeneous behavior of clones expressing the same serotype. *J Clin Microbiol*. 2011;49(4):1369-75. <https://doi.org/10.1128/JCM.01763-10> PMID: 21270219
6. Richter SS, Diekema DJ, Heilmann KP, Dohrn CL, Riahi F, Doern GV. Changes in pneumococcal serotypes and antimicrobial resistance after introduction of the 13-valent conjugate vaccine in the United States. *Antimicrob Agents Chemother*. 2014;58(11):6484-9. <https://doi.org/10.1128/AAC.03344-14> PMID: 25136018
7. Lindstrand A, Galanis I, Darenberg J, Morfeldt E, Naucler P, Blennow M, et al. Unaltered pneumococcal carriage prevalence due to expansion of non-vaccine types of low invasive potential 8 years after vaccine introduction in Stockholm, Sweden. *Vaccine*. 2016;34(38):4565-71. <https://doi.org/10.1016/j.vaccine.2016.07.031> PMID: 27473304
8. Hanage WP, Kajjalainen TH, Syrjänen RK, Auranen K, Leinonen M, Mäkelä PH, et al. Invasiveness of serotypes and clones of *Streptococcus pneumoniae* among children in Finland. *Infect Immun*. 2005;73(1):431-5. <https://doi.org/10.1128/IAI.73.1.431-435.2005> PMID: 15618181
9. Shoji H, Vázquez-Sánchez DA, Gonzalez-Díaz A, Cubero M, Tubau F, Santos S, et al. Overview of pneumococcal serotypes and genotypes causing diseases in patients with chronic obstructive pulmonary disease in a Spanish hospital between 2013 and 2016. *Infect Drug Resist*. 2018;11:1387-400. <https://doi.org/10.2147/IDR.S165093> PMID: 30214260
10. Aguinagalde L, Corsini B, Domenech A, Domenech M, Càmara J, Ardanuy C, et al. Emergence of amoxicillin-resistant variants of SpainV-ST156 pneumococci expressing serotype 11A correlates with their ability to evade the host immune response. *PLoS One*. 2015;10(9):e0137565. <https://doi.org/10.1371/journal.pone.0137565> PMID: 26368279
11. Ardanuy C, Fenoll A, Berón S, Calatayud L, Liñares J. Increase of the M phenotype among erythromycin-resistant *Streptococcus pneumoniae* isolates from Spain related to the serotype 14 variant of the SpainV-3 clone. *Antimicrob Agents Chemother*. 2006;50(9):3162-5. <https://doi.org/10.1128/AAC.00269-06> PMID: 16940119
12. González-Díaz A, Càmara J, Ercibengoa M, Cercenado E, Larrosa N, Quesada MD, et al. Emerging non-13-valent pneumococcal conjugate vaccine (PCV13) serotypes causing adult invasive pneumococcal disease in the late-PCV13 period in Spain. *Clin Microbiol Infect*. 2019;S1198-743X(19)30589-0. <https://doi.org/10.1016/j.cmi.2019.10.034> PMID: 31756452
13. Coffey TJ, Daniels M, Enright MC, Spratt BG. Serotype 14 variants of the Spanish penicillin-resistant serotype 9V clone of *Streptococcus pneumoniae* arose by large recombinational replacements of the cpsA-ppb1a region. *Microbiology*. 1999;145(Pt 8):2023-31. <https://doi.org/10.1099/13500872-145-8-2023> PMID: 10463168
14. Enright MC, Fenoll A, Griffiths D, Spratt BG. The three major Spanish clones of penicillin-resistant *Streptococcus pneumoniae* are the most common clones recovered in recent cases of meningitis in Spain. *J Clin Microbiol*. 1999;37(10):3210-6. <https://doi.org/10.1128/JCM.37.10.3210-3216.1999> PMID: 10488179
15. Càmara J, Cubero M, Martín-Galiano AJ, García E, Grau I, Nielsen JB, et al. Evolution of the β -lactam-resistant *Streptococcus pneumoniae* PMEN3 clone over a 30 year period in Barcelona, Spain. *J Antimicrob Chemother*. 2018;73(11):2941-51. <https://doi.org/10.1093/jac/dky305> PMID: 30165641
16. Centers for Disease Control and Prevention (CDC). National Notifiable Diseases Surveillance System (NNDSS). Invasive Pneumococcal Disease (IPD) (*Streptococcus pneumoniae*). Atlanta: CDC. [Accessed: 2 Apr 2020]. Available from: <http://www.cdc.gov/nndss/conditions/invasive-pneumococcal-disease/case-definition/2010/>.
17. The European Committee on Antimicrobial Susceptibility Testing (EUCAST). Breakpoint tables for interpretation of MICs and zone diameters. Version 9.0. 2019. [Accessed: 2 Apr 2020]. Available from: <http://www.eucast.org>.
18. Tenover FC, Arbeit RD, Goering RV, Mickelsen PA, Murray BE, Persing DH, et al. Interpreting chromosomal DNA restriction patterns produced by pulsed-field gel electrophoresis: criteria for bacterial strain typing. *J Clin Microbiol*. 1995;33(9):2233-9. <https://doi.org/10.1128/JCM.33.9.2233-2239.1995> PMID: 7494007
19. Enright MC, Spratt BG. A multilocus sequence typing scheme for *Streptococcus pneumoniae*: identification of clones associated with serious invasive disease. *Microbiology*. 1998;144(Pt 11):3049-60. <https://doi.org/10.1099/00221287-144-11-3049> PMID: 9846740
20. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 2014;30(15):2114-20. <https://doi.org/10.1093/bioinformatics/btu170> PMID: 24695404
21. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol*. 2012;19(5):455-77. <https://doi.org/10.1089/cmb.2012.0021> PMID: 22506599
22. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One*. 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963> PMID: 25409509
23. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 2014;30(14):2068-9. <https://doi.org/10.1093/bioinformatics/btu153> PMID: 24642063
24. Jia B, Raphenya AR, Alcock B, Waglechner N, Guo P, Tsang KK, et al. CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res*. 2017;45(D1):D566-73. <https://doi.org/10.1093/nar/gkw1004> PMID: 27789705
25. Zankari E, Hasman H, Cosentino S, Vestergaard M, Rasmussen S, Lund O, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012;67(11):2640-4. <https://doi.org/10.1093/jac/dks261> PMID: 22782487
26. Epping L, van Tonder AJ, Gladstone RA, Bentley SD, Page AJ, Keane JA, et al. SeroBA: rapid high-throughput serotyping of *Streptococcus pneumoniae* from whole genome sequence data. *Microb Genom*. 2018;4(7). <https://doi.org/10.1099/mgen.0.000186> PMID: 29870330
27. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. *Genome Biol*. 2014;15(11):524. <https://doi.org/10.1186/s13059-014-0524-x> PMID: 25410596
28. Bruen TC, Philippe H, Bryant D. A simple and robust statistical test for detecting the presence of recombination. *Genetics*. 2006;172(4):2665-81. <https://doi.org/10.1534/genetics.105.048975> PMID: 16489234
29. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, Glasner C, et al. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom*. 2016;2(11):e000093. <https://doi.org/10.1099/mgen.0.000093> PMID: 28348833
30. Bentley SD, Aanensen DM, Mavroidi A, Saunders D, Rabinowitz E, Collins M, et al. Genetic analysis of the capsular biosynthetic locus from all 90 pneumococcal serotypes. *PLoS Genet*. 2006;2(3):e31. <https://doi.org/10.1371/journal.pgen.0020031> PMID: 16532061
31. Kozlov AM, Darriba D, Flouri T, Morel B, Stamatakis A. RAxML-NG: a fast, scalable and user-friendly tool for maximum likelihood phylogenetic inference. *Bioinformatics*. 2019;35(21):4453-5. <https://doi.org/10.1093/bioinformatics/btz305> PMID: 31070718
32. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res*. 2004;14(7):1394-403. <https://doi.org/10.1101/gr.2289704> PMID: 15231754
33. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, Bentley SD, et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using

- Gubbins. *Nucleic Acids Res.* 2015;43(3):e15. <https://doi.org/10.1093/nar/gku1196> PMID: 25414349
34. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, Harris SR. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics.* 2018;34(2):292-3. <https://doi.org/10.1093/bioinformatics/btx610> PMID: 29028899
 35. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics.* 2015;31(22):3691-3. <https://doi.org/10.1093/bioinformatics/btv421> PMID: 26198102
 36. Brynildsrud O, Bohlin J, Scheffer L, Eldholm V. Rapid scoring of genes in microbial pan-genome-wide association studies with Scoary. *Genome Biol.* 2016;17(1):1-9. <https://doi.org/10.1186/s13059-016-1108-8>
 37. Centers for Disease Control and Prevention (CDC). Streptococcus Laboratory. minimum inhibitory concentrations (MICs) for β -lactam Antibiotics Predicted by Penicillin Binding Protein Gene Types. Atlanta: CDC. [Accessed: 4 Sep 2019]. Available from: <https://www.cdc.gov/streplab/pneumococcus/mic.html#>
 38. Li Y, Metcalf BJ, Chochua S, Li Z, Gertz RE Jr, Walker H, et al. Validation of β -lactam minimum inhibitory concentration predictions for pneumococcal isolates with newly encountered penicillin binding protein (PBP) sequences. *BMC Genomics.* 2017;18(1):621. <https://doi.org/10.1186/s12864-017-4017-7> PMID: 28810827
 39. Morales M, Ludwig G, Ercibengoa M, Esteva C, Sanchez-Encinales V, Alonso M, et al. Changes in the serotype distribution of *Streptococcus pneumoniae* causing otitis media after PCV13 introduction in Spain. *PLoS One.* 2018;13(12):e0209048. <https://doi.org/10.1371/journal.pone.0209048> PMID: 30562385
 40. Manna S, Ortika BD, Dunne EM, Holt KE, Kama M, Russell FM, et al. A novel genetic variant of *Streptococcus pneumoniae* serotype 11A discovered in Fiji. *Clin Microbiol Infect.* 2018;24(4):428.e1-7. <https://doi.org/10.1016/j.cmi.2017.06.031> PMID: 28736074
 41. Croucher NJ, Mitchell AM, Gould KA, Inverarity D, Barquist L, Feltwell T, 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(10):e1003868. <https://doi.org/10.1371/journal.pgen.1003868> PMID: 24130509
 42. Camilli R, Bonnal RJP, Del Grosso M, Iacono M, Corti G, Rizzi E, et al. Complete genome sequence of a serotype 11A, ST62 *Streptococcus pneumoniae* invasive isolate. *BMC Microbiol.* 2011;11(1):25. <https://doi.org/10.1186/1471-2180-11-25> PMID: 21284853
 43. Keller LE, Robinson DA, McDaniel LS. Nonencapsulated *Streptococcus pneumoniae*: Emergence and pathogenesis. *MBio.* 2016;7(2):e01792. <https://doi.org/10.1128/mBio.01792-15> PMID: 27006456
 44. Hilty M, Wüthrich D, Salter SJ, Engel H, Campbell S, Sá-Leão R, et al. Global phylogenomic analysis of nonencapsulated *Streptococcus pneumoniae* reveals a deep-branching classic lineage that is distinct from multiple sporadic lineages. *Genome Biol Evol.* 2014;6(12):3281-94. <https://doi.org/10.1093/gbe/evu263> PMID: 25480686
 45. Wyres KL, Lambertsen LM, Croucher NJ, McGee L, von Gottberg A, Linares J, et al. The multidrug-resistant PMEN1 pneumococcus is a paradigm for genetic success. *Genome Biol.* 2012;13(11):R103. <https://doi.org/10.1186/gb-2012-13-11-r103> PMID: 23158461
 46. Cowley LA, Petersen FC, Junges R, Jimson D Jimenez M, Morrison DA, Hanage WP. Evolution via recombination: Cell-to-cell contact facilitates larger recombination events in *Streptococcus pneumoniae*. *PLoS Genet.* 2018;14(6):e1007410. <https://doi.org/10.1371/journal.pgen.1007410> PMID: 29897968
 47. Moscoso M, García E, López R. Biofilm formation by *Streptococcus pneumoniae*: role of choline, extracellular DNA, and capsular polysaccharide in microbial accretion. *J Bacteriol.* 2006;188(22):7785-95. <https://doi.org/10.1128/JB.00673-06> PMID: 16936041

Any supplementary material referenced in the article can be found in the online version.

This article is copyright of the authors or their affiliated institutions, 2020.

[License, supplementary material and copyright](#)

This is an open-access article distributed under the terms of the Creative Commons Attribution (CC BY 4.0) Licence. You may share and adapt the material, but must give appropriate credit to the source, provide a link to the licence and indicate if changes were made.