Universidade de Lisboa

Faculdade de Farmácia

EVOLUTION AND ADAPTATION OF *STREPTOCOCCUS PNEUMONIAE* POPULATION IN THE ERA OF EXPANDED CONJUGATE VACCINES

Catarina Isabel Brandão Candeias

Thesis supervised by Professor Raquel Sá-Leão,

and co-supervised by Professor Aida Duarte

Master in Biopharmaceutical Sciences

2018

Universidade de Lisboa

Faculdade de Farmácia

EVOLUTION AND ADAPTATION OF *STREPTOCOCCUS PNEUMONIAE* POPULATION IN THE ERA OF EXPANDED CONJUGATE VACCINES

Catarina Isabel Brandão Candeias

Thesis supervised by Professor Raquel Sá-Leão,

and co-supervised by Professor Aida Duarte

Master in Biopharmaceutical Sciences

2018

The studies presented in this thesis were performed within the Laboratory of Molecular Microbiology of Human Pathogens, at the Instituto de Tecnologia Química e Biológica António Xavier, Universidade Nova de Lisboa, under the supervision of Professor Raquel Sá-Leão, and the internal co-supervision of Professor Aida Duarte.

The studies included in this thesis were selected as poster communications:

Candeias, C., S. Félix, S. Handem, H. de Lencastre and R. Sá-Leão. 2017. **Evolution and adaptation of *Streptococcus pneumoniae* population in the era of conjugate vaccines.** Poster 301 in Molecular Microbiology and Microbial Physiology. MICROBIOTEC17 – Congress of Microbiology and Biotechnology 2017. Porto, Portugal. December 7-9, 2017.

Candeias, C., S. Félix, S. Handem, H. de Lencastre and R. Sá-Leão. 2017. **Clonal changes in pneumococcal population following private use of PCV13 in Portugal.** Abstr: ISPPD-0441. ISPPD-11 – 11th International Symposium on Pneumococci and Pneumococcal Diseases. Melbourne, Australia. April 15-19, 2018.

Abstract

Streptococcus pneumoniae (or pneumococcus) is one of the most important causes of bacterial infections mainly in young children, immunocompromised individuals of all ages and the elderly. It is responsible for meningitis, bacteremia and pneumonia, as well as sinusitis and otitis media. Besides this, the natural way of living of this bacterium is through asymptomatic nasopharyngeal colonization. Colonization rates are generally high among young children, the major reservoir and source of transmission of pneumococci. Once colonization is the first step for pneumococcal disease and is essential for bacteria spreading between individuals, this has been an important focus of research. Most pneumococcus have a polysaccharide capsule that is responsible for the existence of close to 100 variants (or serotypes) and that is considered its main virulence factor. For this reason, the capsule has been the main target for pneumococcal vaccine design. In Portugal, pneumococcal conjugate vaccines (PCVs) were only introduced in the Nacional Immunization Program in 2015.

In the era of multivalent pneumococcal conjugated vaccines, epidemiological and surveillance studies are of major importance to obtain information that will help our understanding of serotype and clonal replacement changes occurring in carriage and disease as a consequence of vaccination. To obtain this information, microbial typing tools are fundamental to provide data about the evolution of bacterial population. Serotyping tools are widely used to address the effectiveness of pneumococcal vaccination. However, it is known that capsular type alone may not give the full picture regarding vaccine impact.

The aim of this study was to address the clonal evolution and adaptation of the pneumococcus after the introduction of the 13-valent pneumococcal conjugate vaccine (PCV13) in Portugal. For that, we used multilocus sequence typing (MLST), the 'gold standard' of molecular typing for the pneumococcus. We analyzed close to 700 nasopharyngeal isolates from children attending day-care centers, collected before introduction of PCV13 (2009 – 2010) and in the years afterwards (2011 – 2016) in two regions of Portugal.

Our focus turned into answering main questions regarding the clonal evolution leading to persistence of some vaccine-types and emergence of non-vaccine types, overtime. The results showed that, regarding the few vaccine-types that are still in circulation, for most cases it is one major clone, that was already in circulation before vaccine introduction, that is responsible for the persistence of that serotype, as it is the case of the clonal complex (CC) 179 associated with serotype 19F. Besides that, whenever there is high clonal diversity and the presence of antimicrobial resistant clones before vaccine introduction, in most cases it is the resistant clone

that is maintained overtime. Regarding non-vaccine types we detected clonal expansion associated with clones that were already in circulation and the emergence of novel, previously undetected clones. In this group of serotypes, antimicrobial resistance did not have such an important role as the one detected for the vaccine-type clones. We found few capsular switch events, but we could not correlate this phenomenon with vaccine impact.

With this study we could conclude that PCV13 impacts on the pneumococcal population beyond capsular type and that this population is able to genetically adapt and evolve to this pressure.

Keywords: *Streptococcus pneumoniae*, colonization, PCV13, MLST, clones.

Resumo

Streptococcus pneumoniae (ou pneumococos) é uma das principais causas de infeções bacterianas, principalmente em crianças, indivíduos imunocomprometidos de todas as idades e idosos. É um agente causativo de meningite, bacteriemia e pneumonia, assim como sinusite e otite. Apesar disto, o estilo de vida natural desta bactéria é através da colonização assintomática da nasofaringe. As taxas de colonização são geralmente elevadas em crianças, os principais reservatórios e fontes de transmissão de pneumococos. Uma vez que a colonização precede sempre a infeção e é essencial para a transmissão entre indivíduos, esta tem sido um foco importante de investigação. A maior parte das estirpes de pneumococos contêm uma cápsula polissacarídea que é responsável pela existência de cerca de 100 variantes (ou serotipos) e que é considerada o seu principal fator de virulência. Por esta razão, a cápsula tem sido utilizada no desenvolvimento de vacinas pneumocócicas. Em Portugal, a vacina pneumocócica polissacarídea conjugada apenas foi introduzida no Plano Nacional de Vacinação em 2015.

Na presente época, em que estão disponíveis vacinas multivalentes, os estudos epidemiológicos e de vigilância são de enorme importância de forma a alcançar toda a informação que ajudará a perceber particularidades como a substituição de serotipos e clones que ocorre em colonização e em infeção como consequência da vacinação. De forma a obter esta informação, ferramentas de tipagem microbiana são fundamentais para fornecer dados acerca da evolução populacional bacteriana. Métodos de tipagem capsular são frequentemente utilizados para testar a eficácia da vacina pneumocócica. No entanto, a cápsula pode não fornecer toda a informação relativamente ao impacto da vacina.

O principal objetivo deste estudo foi verificar a evolução e adaptação clonal do pneumococo após a introdução da vacina pneumocócica conjugada 13-valente (PCV13) em Portugal. Para isso, utilizámos a técnica ‘multilocus sequence typing’ (MLST), que é atualmente o padrão para tipagem molecular de pneumococos. Analisámos cerca de 700 isolados da nasofaringe recolhidas de crianças que frequentam infantários antes da introdução da PCV13 (2009 – 2010) e nos anos seguintes (2011 – 2016) em duas regiões de Portugal.

Após análise dos dados, o nosso objetivo foi responder a algumas questões relacionadas com a evolução clonal que permitiu a persistência de alguns serotipos vacinais e a emergência dos serotipos não vacinais, ao longo do tempo. Os nossos resultados mostram que, relativamente aos poucos serotipos vacinais que ainda estão em circulação, na maioria dos casos é um clone principal, que já estava em circulação antes da introdução da vacina, que é responsável pela persistência dos serotipos, como é o caso do complexo clonal (CC) 179 associado ao serotipo

19F. Para além disto, quando há grande diversidade clonal e a presença de clones resistentes a agentes antimicrobianos antes da introdução da vacina, na maioria dos casos é o clone resistente que é mantido ao longo do tempo. Relativamente aos serotipos não vacinais detetámos expansão clonal associada a clones que já estavam em circulação e emergência de novos clones que não tinham ainda sido detetados. Neste grupo de serotipos, a resistências a agentes antimicrobianos não mostrou ser tão importante como nos clones associados a serotipos vacinais. Também detetámos alguns eventos de alteração da cápsula, no entanto, não conseguimos correlacioná-los com o impacto da vacina.

Com este estudo podemos concluir que a vacina tem um impacto na população de pneumococos que vai para além do serotipo, e que esta população é capaz de se adaptar e evoluir geneticamente em resposta a esta pressão.

Palavras-chave: *Streptococcus pneumoniae*, colonização, PCV13, MLST, clones.

Acknowledgements

The studies present in this master thesis would not have been possible without the contribution of some people who I would like to acknowledge here.

My supervisor, Professor Raquel Sá-Leão, for your time, your guidance and your demand for perfection. Thank you for all the scientific and non-scientific advices and for always turning a negative into a positive, a mistake into a learning opportunity and a problem into a solution. I am very proud and honored to work for and with you, a role model of a scientist and a person.

Professor Aida Duarte, for believing in this project and accepting this co-supervision. Thank you for all the time, all the availability and the suggestions you gave me.

All my colleagues at the Microbial of Human Pathogens Unit, thank you for making me feel part of the team since day one, and for all the help along the way. A special thank you to my little family from the Laboratory of Molecular Microbiology of Human Pathogens. Thank you for being with me through the good, the bad and the ugly. Every single one of you has a special place in my heart. It is a pleasure to work with such wise and kind people, that have been key factors for my scientific and personal growth.

My friends, thank you for all the support and all the patience. CBF colleagues, thank you for all the good moments that we spent together through this journey. Catarina Barros, for you there are no words. Thank you for all the advices and for being always there no matter what. With you all became easier and I can proudly say that this master gave me a friendship for life. Sara Rodrigues, more than a friend, you are family. Besides the distance, you manage to always light my day up and put a smile on my face, even in the cloudiest days. Lucky me for having such an amazing person in my life.

Carlos, thank you for your patience, for being the best at dealing with my bad temper, for being my best friend, my confident, and for loving me unconditionally.

My family, my heroes. I cannot thank you enough for everything you do, for always turning the impossible possible. You helped me get up when I was down, and you shared my happiness when I was up. Thank you for believing in me. Besides the distance my heart is always where

you are. I live and I work to make you proud. This thesis is dedicated to you. *A minha família, os meus heróis. Não consigo agradecer-vos o suficiente por tudo o que fazem, por sempre tornarem possível o impossível. Vocês levantam-me quando estou em baixo, e partilham a minha felicidade quando estou em cima. Obrigada por acreditarem em mim. Apesar da distância o meu coração está sempre onde vocês estão. Eu vivo e trabalho para que se orgulhem de mim. Esta tese é dedicada a vocês.*

Abbreviations

CC – clonal complex

ChoP – phosphorylcholine

DCC – day-care center

DLV – double locus variant

IPD – invasive pneumococcal disease

MLST – multilocus sequence typing

NIP – national immunization program

NVT – non-vaccine type

PCV – pneumococcal conjugate vaccine

PCV7 – 7-valent pneumococcal conjugate vaccine

PCV10 – 10-valent pneumococcal conjugate vaccine

PCV13 – 13-valent pneumococcal conjugate vaccine

PFGE – pulse-field gel electrophoresis

PMEN – pneumococcal molecular epidemiology network

PPV – pneumococcal polysaccharide vaccine

PPV23 – 23-valent pneumococcal polysaccharide vaccine

PsaA – pneumococcal surface adhesion A

rPAF – receptor for platelet-activating factor

SLV – single locus variant

ST – sequence type

STGG – skim milk tryptone-glucose-glycerol

TAE – tris-acetate EDTA buffer

TE – tris-EDTA buffer

TLV – triple locus variant

TSA – tryptic soy agar

VT – vaccine-type

WHO – world health organization

Table of contents

1 Introduction 1

1.1 *Streptococcus pneumoniae*: overview and characteristics 1

1.2 Pneumococcal disease 1

1.3 The polysaccharide capsule 2

1.4 Pneumococcal colonization 2

1.5 Antimicrobial resistance 3

1.6 Pneumococcal vaccination 4

1.6.1 23-valent pneumococcal polysaccharide vaccine (PPV23) 4

1.6.2 7-valent pneumococcal conjugate vaccine (PCV7) 5

1.6.3 10-valent pneumococcal conjugate vaccine (PCV10) 6

1.6.4 13-valent pneumococcal conjugate vaccine (PCV13) 6

1.7 Genetic diversity 6

1.8 The Portuguese scenario 7

1.9 Typing methodologies 8

1.9.1 Capsular typing 9

1.9.2 Molecular typing 10

2 Objective 11

3 Materials and Methods 13

3.1 Study design 13

3.2 Multilocus Sequence Typing analysis 13

3.2.1 DNA extraction 14

3.2.2 MLST.....14

3.2.3 Sequence analysis.....14

3.3 Data analysis.....15

4 Results and Discussion.....17

4.1 MLST of pneumococcal isolates.....17

4.2 Evolution of vaccine-types.....17

4.3 Non-vaccine types analysis.....27

4.4 Capsular switch events.....37

5 Conclusions.....41

6 References.....43

7 Anexes.....49

Index of Tables and Figures

Table 1: Data from epidemiological studies performed in Portugal from 1996 to 2016 (10, 28). 8

Table 2: Summary table for the tested MLST isolates. 13

Table 3: Allelic profiles obtained in this study by serotype and geographic area. 19

Table 4: Clonal evolution of vaccine-types overtime in the urban and rural regions. 23

Table 5: Resistant samples and correspondent antimicrobial resistance profiles of vaccine-types, overtime. 24

Table 6: Clonal evolution of non-vaccine types overtime in the urban and rural regions. 28

Table 7: Resistant samples and correspondent antimicrobial resistance profile of non-vaccine types, overtime. 31

Figure 1: MLST scheme for allele and ST assignment. 15

Figure 2: Clonal lineages of carriage isolates in Portugal (2009-2016). 18

1 Introduction

1.1 *Streptococcus pneumoniae*: overview and characteristics

Two independent microbiologists, George M. Sternberg and Louis Pasteur, firstly described, in 1881, in human saliva, the presence of lancet-shaped pairs of coccoid bacteria, and detected its virulence by inoculating human saliva into rabbits (1). After being identified with several names, the name *Streptococcus pneumoniae* was established in 1974 due to its propensity to cause pulmonary disease and grow as chains of cocci in liquid media (1). *Streptococcus pneumoniae* (or pneumococcus) is a lancet-shaped, diplococci, facultative anaerobic, Gram-positive bacterium of major importance due to its role on morbidity, mortality and health care associated costs worldwide (2). This bacterium has been the subject of intense research and it was of great importance to several biological discoveries like the genetic transformation and the discovery of antibody-mediated opsonization (3).

1.2 Pneumococcal disease

We can differentiate pneumococcal disease in invasive pneumococcal disease (IPD) when the pneumococcus is isolated from naturally sterile body fluids; and non-IPD (4). Bacteremia (blood stream infection) and meningitis (infection of the meninges that cover the brain and the spinal cord) are part of the invasive group and sinusitis and otitis media (infection of the middle-ear space) are characteristic of the non-invasive group (5). In terms of pneumococcal pneumonia (known as lung infection) we can have both scenarios (bacteremic or non-bacteremic pneumonia) being the non-bacteremic pneumonia the most prevalent form of infection in adults (2, 6).

The main risk groups for pneumococcal disease are young children, due to poor natural humoral immunity and, consequently, the predisposition to viral respiratory infections, the elderly, due to weakened immune system, and immunocompromised patients of all ages (4, 7). In developing countries, pneumococcal disease is especially important due to the millions of deaths it is responsible for (5).

For invasive disease to occur, the local presence of inflammatory factors (for example, interleukin 1 and tumor necrosis factor) is required in order to produce an inflammatory cascade (7). This, in turn, will result in different type and number of epithelial and endothelial target cells receptors, like the platelet-activating-factor receptor that is upregulated allowing the binding of the pneumococcus, promoting its own internalization and transcellular migration through respiratory epithelium and vascular endothelium, leading to invasion (7).

1.3 The polysaccharide capsule

Most pneumococcus are encapsulated, meaning that they are surrounded by a complex and highly immunogenic polysaccharide capsule (2). This polysaccharide capsule is recognized as the main virulence factor of this bacteria mainly due to its anti-phagocytic activity, but also because it protects the cell wall, the main cause of inflammatory response in pneumococcal infection (7). The capsule will protect the pneumococcus from opsonophagocytosis by the host's neutrophils and macrophages, and other immune factors like neutrophil-extracellular traps (3). Non-encapsulated strains are less virulent due to increased access of antibodies and complement to the pneumococcal surface and, consequently, an increased clearance by the host's immune system. However, these are frequently associated with superficial infections (6). The phenomenon of nontypeability has been described associated with diverse mechanisms besides the lack of capsule, like mutation in the locus responsible for capsule production (*cps*) that inhibit capsule production or novel capsules (8).

The pneumococcal capsule is responsible for the existence of close to 100 capsule variants, the serotypes, that are distinguished through the expression of polysaccharides with different structure and antigenicity (5, 9, 10). When serotypes show immunological cross-reactivity in animal models, they can be defined as belonging to the same serogroup (11). However, the existence of such diversity does not correlate directly with invasive disease, once only a relatively small portion of serotypes is mainly found in most pneumococcal infections (2). Before vaccine introduction, only 6 to 11 serotypes among the almost 100 variants, were described as responsible for more than 70% of all IPD occurring in children worldwide (9).

1.4 Pneumococcal colonization

Besides the health risk associated with this pathogen and the high burden of invasive pneumococcal disease, its natural lifestyle is through asymptomatic nasopharyngeal colonization of healthy people (12). Pneumococci usually share the respiratory tract with other bacteria like *Haemophilus influenzae, Moraxella catarrhalis* and *Staphylococcus aureus*, and can be isolated from the nasopharynx of 5 to 90% of healthy people (2, 4, 7). The importance of colonization relies on the fact that if there is no colonization, pneumococcal disease with the homologous strain will not occur, so preventing nasopharyngeal colonization is a valid strategy to prevent pneumococcal disease (7). It is also important to refer that colonization allows for evolution of the species, once horizontal gene transfer with other pneumococci and closely related species that are coexisting (co-colonization) can occur (12, 13).

As for disease, the pneumococcal capsule plays an important role on colonization allowing evasion attachment to airway mucus which is known as an expulsion mechanisms of the host (6). The pneumococcus binds to cell-surface carbohydrates due to the presence of cell-wall-associated surface proteins (like the pneumococcal surface adhesion A, PsaA) that also have a role on non-specific, physicochemical interactions with the resting epithelium (7). The pneumococcal enzyme neuraminidase is also important for colonization because it cleaves terminal sugars from human glycoconjugates from mucin present in the host's mucus, what can consequently uncover receptors important for adherence (7). Several receptor-ligand interactions have been proposed as essential for the adherence process to the host epithelial surface of the nasopharynx. For example, the phosphorylcholine (ChoP) that is a bacterial adhesin used to mimic the natural ligand of the receptor for platelet-activating factor (rPAF), allowing the bacteria to use this receptor that is present in the epithelial surface of the human nasopharynx (5).

Asymptomatic carriage rates and duration differ with factors like age, geographical area, socioeconomic factors, overcrowding environments and when upper respiratory infections are present (2, 4). Infants and young children are the main reservoir and source of transmission by respiratory droplets, especially when talking about children attending day-care centers (DCCs), as it is known that their colonization rates can surpass the 60%, they can be colonized several times in their life and normally with high pneumococcal density (10, 13). Attending DCCs is known to be a risk factor for pneumococcal colonization and transmission, as it is a crowded environment where immature immune systems and poor hygienic behaviors are very present (14).

1.5 Antimicrobial resistance

Before the introduction of pneumococcal vaccines, the mostly used treatment to fight pneumococcal disease was penicillin, however, this rapidly led to the appearance of resistance in 1965 and multidrug resistant strains in 1977 and, consequently, treatment failure (15). Development of pneumococcal resistance to the most common used antibiotics like penicillins, cephalosporins, co-trimoxazole and macrolides is now a very important and serious problem all over the world (9). It has been showed that, by 2015, 30% of the cases of invasive pneumococcal disease were due to pneumococcus resistant to at least one antimicrobial agent (16).

Attending day-care centers has been linked to carriage of antibiotic resistant strains. This is mainly due to frequent use of antibiotics to treat upper respiratory infections in children, to the

immaturity of the immune system and crowding effect (16, 17). High rates of resistant disease-causing bacteria are correlated with high rates, and inappropriate use, of antibiotics (17). Besides that, pneumococci seem to be highly adaptable, evolving into new drug resistant clones, even after vaccination, and this may be due to a high frequency of transformation and events like capsular switching, "de novo" acquisition of resistance, introduction of new clones or expansion of specific resistant clones (15, 18). The latter case was already described for Portugal as the major factor for antimicrobial resistance maintenance among children attending day-care centers (19).

With the main goal of standardize the nomenclature of the major clones responsible for antimicrobial resistance distribution worldwide, the Pneumococcal Molecular Epidemiology Network (PMEN) was established (20). PMEN clones are clones that have been identified as having a large geographic distribution, are well established through time in a specific country, and are epidemiologically relevant (20). The nomenclature of each clone identifies the country in which the clone was first isolated, the serotype first associated to the clone and the sequence type (21).

1.6 Pneumococcal vaccination

To overcome the high pneumococcal disease burden, vaccination strategies have been implemented overtime using the polysaccharide capsule as the primary target. Due to the large number of serotypes that exist these vaccines, which target only a subset of capsules, are of limited intrinsic coverage (6). Pneumococcal vaccines started to gain interest in 1911, however, only in the 1960s, after ineffectiveness of penicillin treatment, vaccine development boosted (2). Nowadays there are two types of vaccines licensed for use in humans: pneumococcal polysaccharide vaccines (PPVs) that contain purified polysaccharide antigens, producing antibodies without immune memory; and pneumococcal conjugate vaccines (PCVs) that also contain purified polysaccharide antigens but conjugated with an immunogenic protein carrier. The latter is more effective in young children due to the production of immunity B-cells memory (2, 22).

1.6.1 23-valent pneumococcal polysaccharide vaccine (PPV23)

The first vaccine, licensed in the US in 1983, was the 23-valent pneumococcal polysaccharide vaccine (PPV23, PNEUMOVAX23, Merck) that contains capsular polysaccharides of 23 serotypes (1, 2, 3, 4, 5, 6B, 7F, 8, 9N, 9V, 10A, 11A, 12F, 14, 15B, 17F,

18C, 19F, 19A, 20, 22F, 23F, and 33F) (6). This vaccine is poorly immunogenic due to the fact that capsule proteins are T-cell independent antigens that only stimulate mature B-lymphocytes and not T-lymphocytes, therefore is only recommended for children older than 2 years and adults over 65 years who are at risk (4).

1.6.2 7-valent pneumococcal conjugate vaccine (PCV7)

The 7-valent pneumococcal conjugate vaccine (PCV7; Prevenar7, Pfizer; serotypes 4, 6B, 9V, 14, 18C, 19F and 23F) was developed and licensed in the United States in 2000 and in June 2001 in Portugal. It targeted the serotypes that were responsible for 70-90% of IPD in the USA. Its use led to a significant decrease in incidence and mortality due to IPD caused by vaccine types (VTs) (6). This effect was observed not only in vaccinated children but also in non-vaccinated children and adults, a consequence of herd immunity (4, 6, 22, 23). Ideally, eliminating carriage of PCV7 serotypes in vaccinated children would potentially abolish the source of spread in the community, ultimately leading to a decrease in the transmission rate to unvaccinated children and other age groups (23). However, serotype replacement appeared, with the emergence of serotypes that were not covered by the vaccine (non-vaccine types – NVTs) together with the rise of non-PCV7 IPD, especially concerning serotype 19A that also accounted for major antibiotic resistance (4, 24, 25). These NVTs took advantage of the niche that used to be occupied by VTs and that, after vaccine implementation, was left empty (15).

An epidemiological carriage study was conducted in the US, at three time points: 2001, 2004 and 2007 (26). Vaccine coverage, pneumococcal carriage and, specifically, carriage of NVTs increased throughout the years in almost all ages (26). By 2007, the most commonly carried serotypes were 19A, 6A, 15B/C, 35B and 11A (26). Regarding carriage of VTs the authors observed a significant decrease from 36% in 2001 to 15% in 2004 and 3% in 2007 (26). With this data they could demonstrate serotype replacement within 7 years of PCV7 use (26). A review regarding pneumococcal surveillance data from Canada, South Africa, Australia, USA, England and Wales from 1998 to 2009 showed a decrease of IPD caused by VTs, not only in children under 5 years of age but also in individuals older than the target group for vaccination, demonstrating also the herd effect (27). On the other hand, IPD cases due to NVTs increased also in the two distinct groups (27).

1.6.3 10-valent pneumococcal conjugate vaccine (PCV10)

The 10-valent pneumococcal conjugate vaccine (PCV10; Synflorix, GlaxoSmithKline; serotypes 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F and 23F) was licensed in 2008 and became available in Portugal in April 2009, however vaccine usage was very low, with an estimated use of 13% in the target group in 2009 (data from INE/National Statistics Institute) (28).

1.6.4 13-valent pneumococcal conjugate vaccine (PCV13)

The 13-valent pneumococcal conjugate vaccine (PCV13; Prevenar 13, Pfizer; serotypes 1, 3, 4, 5, 6A, 6B, 7F, 9V, 14, 18C, 19A, 19F and 23F) was licensed in 2009 in the US, and became available in Portugal in January 2010, replacing PCV7 (4, 10). It contains the polysaccharide antigens individually conjugated to a nontoxic diphtheria carrier protein (cross-reactive material 197) (9, 22). PCV13 had no reimbursement by the state until its entrance in the National Immunization Program (NIP) in August 2015. Since then it is given to all children in the target group free of charge.

It is known that pneumococcal conjugate vaccines impact on infection, colonization and, as they target vaccine types that represent the majority of antibiotic-resistant strains, they also have the potential to decrease the prevalence of antibiotic resistance among the pneumococcal population (18). In the US, between 1998 and 2015, use of PCVs led to a decline of 90% and 50% in invasive pneumococcal disease in children younger than 5 years of age and in adults, respectively (15). At the end of 2015, the World Health Organization (WHO) estimated a 37% coverage for pneumococcal vaccines implemented in 129 countries. This is still insufficient, particularly in what regards low and middle income countries (29). In June 2017, 141 countries had already implemented PCVs in their National Immunization Program, and 16 countries had plans to do so (30).

1.7 Genetic diversity

As described above, the pneumococcal polysaccharide capsule has been the main focus of study, however, it is important to highlight that a specific serotype can include several genetically divergent clones (25). Advances in genetic typing methodologies allowed to uncover that 20 to 30% of the pneumococcal genome is accessory genome (31). This variability leads to protein content differences, which may ultimately impact on colonization and virulence abilities between different isolates, even within the same serotype or genotype (3).

Consequently, efforts are being made to understand if indeed isolates from the most invasive serotypes have increased ability to cause disease or if instead, this is related only with specific virulent clones belonging to each serotype (32). *S. pneumoniae* is naturally transformable, horizontal exchange of DNA occurs frequently and any locus can suffer recombination (3). Consequently, it also has the ability to switch the polysaccharide capsule locus with other pneumococci (capsular switching), what, for the serotypes covered by the vaccine, can lead to carriage of a serotype not targeted by the vaccine as an escape mechanism (15).

1.8 The Portuguese scenario

In Portugal, different studies on colonization dynamics using nasopharyngeal samples of children attending day-care centers, have provided data from almost 20 years, covering different periods that enable to detect vaccine usage and corresponding impact on serotype and clonal distribution, as well as antimicrobial use and resistance overtime (10).

Cross-sectional colonization studies with children attending DCCs in Lisbon began early in the 1990s (17, 33). PCV7 impact on colonization was evaluated with cross-sectional studies performed in seven time-points. The results showed that PCV7 use increased significantly overtime, besides never being introduced in the NIP; the percentage of pneumococcal carriage remained high and stable (between 60 and 70%) and the same was true for antibiotic consumption; and serotype replacement occurred overtime (10, 34, 35). Furthermore, between 2009 and 2016, six cross-sectional sampling periods were conducted in two regions of Portugal (Oeiras, an urban region, and Montemor-o-Novo, a rural region) to evaluate the impact of the private use of PCV13 (28). Table 1 summarizes all the years in which epidemiological studies took place and the respective number of participating DCCs, the mean age of children that were sampled, the percentages of carriage, antibiotic use and PCV7 and PCV13 use among the participants.

Using pulse field gel electrophoresis (PFGE) and multilocus sequence typing (MLST), clonal evolution and spread of internationally disseminated clones have been studied. However, the majority of previous studies addressed only antimicrobial resistant isolates (17, 19, 35-37).

Table 1: Data from epidemiological studies performed in Portugal from 1996 to 2016 (10, 28).

Year		No. of DCCs	Mean age (years)	Pneumococcal carriage (%)	Antibiotic use (%)		PCV7 use (%)			PCV13 use (%)	
					Sampling	Previous month	0-6 years	0-2 years	3-6 years	0-2 years	3-6 years
1996		7	3.6	46.9	8.3	26.6	-	-	-	-	-
1997		12	3.5	46.8	7.0	29.0	-	-	-	-	-
1998		12	3.5	58.1	8.2	22.5	-	-	-	-	-
1999		15	3.3	62.9	7.0	23.8	-	-	-	-	-
2001		10	3.5	65.6	5.8	23.5	-	-	-	-	-
2002		13	3.5	66.9	7.0	23.0	10.7	27.6	-	-	-
2003		13	3.5	72.4	4.8	16.4	21.6	37.6	-	-	-
2006		11	3.5	67.9	4.1	19.3	57.2	78.8	-	-	-
2007		11	3.5	61.7	7.0	19.8	71.5	84.7	-	-	-
2009/10	Urban	9	3.3	62.1	3.4	15.0	-	78.8	75.9	-	-
	Rural	16	3.6	59.8	2.4	11.2	-	91.1	81.1	-	-
2011/12	Urban	6	3.2	62.4	6.8	21.1	-	15.2	63.9	60.0	12.0
	Rural	16	3.5	62.8	2.9	7.1	-	10.9	76.5	75.4	9.2
2015/16	Urban	7	2.8	61.6	4.1	16.9	-	0.0	0.0	85.8	79.8
	Rural	15	3.5	59.5	3.4	13.1	-	0.0	0.0	96.3	86.7

A longitudinal study inside one specific DCC was performed between 1998 and 1999, to understand the dynamics of colonization, analyzing specific persistence and transmission events, and gain some insights regarding fitness abilities of individual clones of *S. pneumoniae* and *H. influenza* (14). Besides the large number of different genotypes identified and the constant introduction of novel clones in the population, the authors detected a high level of cross-transmission, mainly due to the spread of pre-existing clones (14).

Co-colonization was also addressed using a molecular serotyping microarray (12). The authors observed that vaccinated children not only have lower co-colonization rates and, consequently, decreased opportunities for horizontal gene transfer between strains, but when present, VTs often act as minor serotype when in co-colonization with NVTs (12).

1.9 Typing methodologies

With the introduction of expanded conjugate vaccines, *S. pneumoniae* has changed its epidemiology and, for that, new serotyping methods and the application of genotyping tools had to be developed (15). Epidemiological studies based on nasopharyngeal colonization have major importance to track vaccine impact in different settings (18).

For that reason, the WHO has recommended the application of specific methods for different aims (38). To understand colonization in children, swabbing the nasopharynx with swabs made from calcium alginate, rayon, Dacron or nylon materials is the preferred method of sampling (38). Regarding specimen transport and storage, skim milk tryptone-glucose-glycerol (STGG) medium (39) is recommended, together with transport on wet ice or cold conditions and storage at -20°C. Selective pneumococci isolation should be made using a selective medium of trypticase soy agar (TSA) base with sheep, horse or goat blood supplemented with gentamicin and with an overnight incubation at 35-37°C in a CO_2 enriched atmosphere (38). Pneumococci should be identified by optochin (ethylhydrocupreine) susceptibility (disc diffusion method – when the diameter of the inhibition zone is 14 mm or superior we have a positive identification due to optochin's inhibition of the pneumococcal ATPase), and bile solubility (colonies of *S. pneumoniae* lyse or dissolve in the presence of 10% sodium deoxycholate and turbid solutions of the organism turn clear due to the activation of the autolytic enzyme N-acetylmuramyl-L-alanine amidase encoded by the *lytA* gene, that will split the muramic acid and alanine bound in the peptidoglycan) (3, 40). The Quellung reaction remains the recommended method for pneumococcal serotyping as it is highly sensitive and specific (38, 41).

1.9.1 Capsular typing

Nowadays, pneumococcal capsular typing can be performed using antisera to detect specific capsular epitopes in an antibody-antigen manner, or using molecular typing techniques that analyze the sequence of the capsular gene (41). The "gold standard" method for serotyping is the aforementioned Quellung reaction, where antibodies from rabbit antiserum bind and react to the pneumococcal polysaccharide capsule and, under a microscope, the bacteria show up as opaque and enlarged due to the change in the refractile index of light passing through the capsule (41, 42). Besides that, PCR-based serotyping strategies have been designed to overcome the disadvantages associated with serological testing, like the high cost of antisera, subjectivity in interpretation and technical expertise requirements (41, 43). The production of capsule is controlled by capsular polysaccharide synthesis genes located at the *cps* locus, which is situated between the conserved *dexB* and *aliA* genes (44). The differentiation of serotypes using PCR-based methods is only possible due to the presence of serotype-specific genes present in the central part of the locus, once the first four genes are conserved in most serotypes (43). This serotyping strategy also allows for simultaneous detection of different capsule types (multiplex PCR) decreasing immensely the time and reagent used (41). Non-encapsulated

strains will produce a null result when typed with Quellung reaction, and that is why they are termed non-typeable, so, specific molecular approaches should be used to confirm that it is pneumococcus and not any closely related streptococcal species (8).

1.9.2 Molecular typing

As it was already described, there is high diversity between different serotypes and even within the same serotype, so adopting genotyping methodologies is important so we can understand the clonal diversity of the pneumococcal population (32). Molecular methods have major importance, particularly, to assess the impact of vaccines and antibiotics in the population, the relation between different isolates, as well as sources and routes of bacterial transmission (42). Usage of molecular typing methods can also clarify about local and global epidemiology, that is if the isolates of a localized disease outbreak belong or not to different strains, and the relationship between strains causing disease in one geographic area with those isolated worldwide (45). PFGE used to be the molecular method of excellence to type *S. pneumoniae*. However comparison of results between different studies was difficult and the interpretation of results could be ambiguous (37). The development of a portable high-resolution molecular typing procedure with high discriminatory power, that requires only some expertise and with good reproducibility, like MLST, was a step forward towards simpler genotyping analysis (32, 42). MLST procedure for *S. pneumoniae* is based on the sequence of internal fragments of seven housekeeping genes that are considered stable and less prone to recombination than other regions of the genome: *ddl* (D-alanine-D-alanine ligase), *spi* (signal peptidase I), *aroE* (shikimate dehydrogenase), *recP* (transkelotase), *gdh* (glucose-6-phosphate dehydrogenase), *gki* (glucose kinase) and *xpt* (xanthine phosphoribosyltransferase) (32, 42).

To understand MLST results in a specific epidemiological context, software tools directed to evolutionary analysis should be used (46). PHILOViZ is capable of handling large and multiple data sets like the allelic data from MLST and the corresponding isolate information, and is an efficient, user-friendly and freely available software (46). The goeBURST algorithm present in PHILOViZ represents the relationship between the isolates analyzed, using the number of differences in each allelic profile, and consequently, joining similar isolates in clonal complexes (CCs) (47). A clonal complex definition is based on the idea that a more fitted genotype (founder) will increase its frequency in the population, and will gradually diversify, by point mutation or recombinational replacement, creating a cluster of phylogenetically closely related strains (48).

2 Objective

In Portugal, pneumococcal colonization patterns in young children have been studied for over 15 years (19, 28, 34, 35, 49, 50). However, no studies were performed on how PCV13 impacts on pneumococci asymptomatically carried, in terms of clonal diversity, evolution, capsular switch, emergence, persistence or disappearance of clones. Previous studies using MLST, described variations that occurred following PCV7 introduction, but their focus was on the drug-resistant pneumococci (19, 36, 49, 51). On the other hand, for pneumococcal disease, several studies regarding major clones in circulation before and after the widespread use of conjugate vaccines are published. In Portugal, between 2001 and 2003, the contributions of genotype versus serotype to invasive disease potential were addressed. The authors described that some serotypes that were not included in PCV13 (8, 9N, 9L, 12B and 20) presented enhanced invasive disease potential (31). This is of importance because of the already described serotype replacement phenomenon, once non-PCV13 serotypes are occupying the niches left by the decreasing vaccine-types, and emerging in frequency (28). Besides that, the authors identified, between clones within the same serotype, heterogeneity in invasive disease potential (31).

In this way, the aim of this project was to determine the impact of PCV13 on pneumococcal evolution in carriage over time. For this end we used a collection of nasopharyngeal samples obtained from children attending day-care centers (DCCs), in an urban and a rural region of Portugal, over three periods of time: 2009 and 2010 (pre-PCV13 period), 2011 and 2012 (early-PCV13 period), and 2015 and 2016 (late-PCV13 period). These samples were tested by multilocus sequence typing (MLST), the state of the art technique for molecular typing of *S. pneumoniae*, and analyzed together with the corresponding serotyping and antimicrobial resistance data.

3 Materials and Methods

3.1 Study design

The first steps for any study regarding pneumococcal colonization are sampling and characterization of positive samples. This often includes assignment of capsular type and antimicrobial resistance patterns. This project is focused on genotyping of pneumococci using MLST. However, it is important to understand the process of collection and characterization of the samples. In preparation for this study, I was involved in the characterization of all nasopharyngeal samples collected in 2016 from children attending day-care centers in Oeiras and Montemor-o-Novo. This collection was obtained under the framework of an ongoing surveillance initiative and the results have been recently presented (28).

To evaluate the impact of PCV13 on pneumococcal population structure in carriage, a total of 657 nasopharyngeal samples obtained from children attending day-care centers of Oeiras and Montemor-o-Novo, from three periods of time, were selected for MLST analysis. The criteria for sample selection was: at least 20% of the strains for each serotype, year and location, chosen randomly. A summary highlighting the strains studied by MLST in the different PCV13 coverage periods and the different regions is presented in Table 2.

Table 2: Summary table for the tested MLST isolates.

Isolates	Pre-PCV13		Early-PCV13		Late-PCV13		Total
	Urban	Rural	Urban	Rural	Urban	Rural	
Total*	677	453	457	390	391	254	2622
Tested for MLST**	158	113	114	99	101	72	657

* Total number of samples from the collection;

** Number of samples chosen for MLST, after applying our criteria of at least 20% of the strains for each serotype, year and location, chosen randomly.

3.2 Multilocus Sequence Typing analysis

To evaluate clonal changes among pneumococci colonizing healthy children following the introduction of PCV13, three-time periods of sampling were considered: 2009 and 2010, the pre-PCV13 era; 2011 and 2012, the early-PCV13 era; and 2015 and 2016, the late-PCV13 era.

3.2.1 DNA extraction

Total DNA extraction was done using the boiled cells method (52). Briefly, samples were plated onto TSA plates supplemented with 5% sheep blood and incubated at 37°C in a 5% CO_2 atmosphere overnight. On the next day, a loop full of pneumococci was suspended in 500 µL of TE (10 mM Tris.HCl and 1 mM EDTA, pH 8) and heated for 10 min at 95°C. A 1:10 dilution with sterilized water was performed and maintained at 4°C. These samples were used as template DNA in PCR amplifications.

3.2.2 MLST

MLST was performed as previously described (32). A PCR amplification from the extracted DNA of internal fragments of seven housekeeping genes (*aroE, gdh, gki, recP, spi, xpt* and *ddl*) was performed. The primers used had universal M13 tails and were previously described (Annex 1) (32, 42). The reactions had a final volume of 20 µL. Final concentrations of the reagents in the PCR were: GoTaq Flexi Buffer 1X, $MgCl_2$ 2 mM, dNTPs 0.04 mM, forward primer 0.4 pmol/µL, reverse primer 0.4 pmol/µL and GoTaq 0.025 U/µL; 2 µL of template DNA were added to each tube. The PCR program used for MLST was: initial denaturation at 94°C for 4 minutes; 30 cycles of denaturation at 94°C for 30 seconds, annealing temperature dependent of the primer used (Annex 1) for 30 seconds, elongation at 72°C for 30 seconds; final elongation at 72°C for 10 minutes and stop at 16°C. To confirm the amplification, products were visualized by agarose gel electrophoresis (Seakem LE agarose 1% in TAE 1X with, applying 3 µL of the PCR product and a voltage of 5V/cm). DNA fragments were visualized with the help of Geldoc Quantity One.

Next, to perform the PCR product purification assay, 3U of Exonuclease I and 3U of Shrimp Alkaline Phosphatase were added to 15 µL of each amplification product. Tubes were incubated for 30 min at 37°C followed by 20 min at 80°C. Purified products were sent to STAB VIDA (Caparica, Portugal) or Macrogen (Amsterdam, The Netherlands) for Sanger sequencing.

3.2.3 Sequence analysis

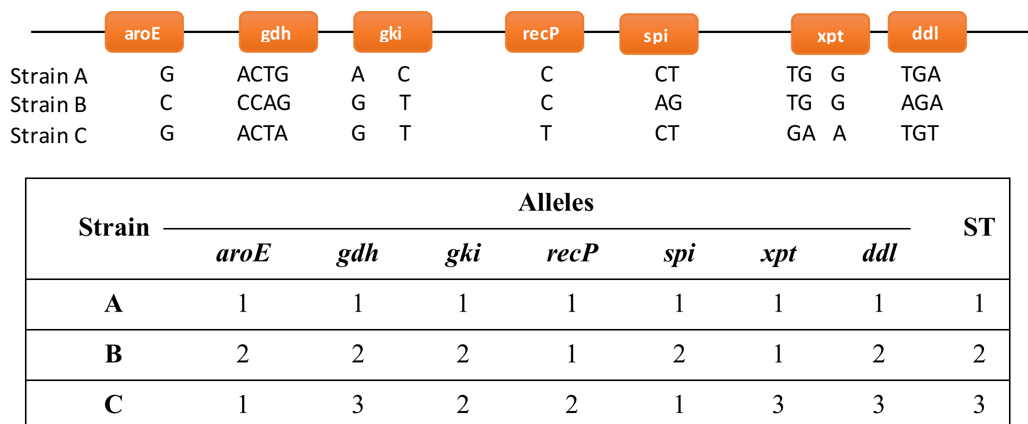
Sequences were analyzed with the Bionumerics (Applied Maths, Gent, Belgium (53)) and DNAStar (Lasergene) software, were sequences at each of the seven loci were compared with all known alleles for that locus. Sequences identical to a known allele were assigned the same allele number; sequences differing from any known allele, even at a single nucleotide site, were

assigned a new allele number (48). In the latter case, sequences and isolate information were submitted to the *S. pneumoniae* MLST database. If there was a novel combination of alleles (new sequence type - ST), the isolate information was also submitted to the database (https://pubmlst.org/spneumoniae/). According to the MLST nomenclature, a ST, which is assigned as combination of alleles at each of the seven loci, identifies a specific clone (Figure 1) (32).

3.3 Data analysis

The relationship between STs was analyzed using the PHILOViZ software with the goeBURST algorithm (46, 47).

Results were loaded and exported to the goeBURST algorithm to facilitate the understanding of probable patterns of evolution between the different allelic profiles of each sample.



Strain	Alleles							ST
	aroE	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*	
A	1	1	1	1	1	1	1	1
B	2	2	2	1	2	1	2	2
C	1	3	2	2	1	3	3	3

Figure 1: MLST scheme for allele and ST assignment.
In MLST, each sequence of each allele is screened and then compared with all the known sequences for that allele. If the sequence is different, a new allele number is assigned. The strain is characterized by the profile of the seven housekeeping genes that are been studied (ST). Adapted from (53).

STs differing in only one housekeeping gene sequence from each other were named single locus variants (SLVs). Likewise, STs differing in two or three alleles were named double locus variants (DLVs) and triple locus variants (TLVs), respectively. Graphs were generated in goeBURST: STs were represented as nodes and those that were related up to DLVs were joined forming a tree descriptive of the clonal complexes (47).

4 Results and Discussion

4.1 MLST of pneumococcal isolates

From the total collection of 657 isolates that were tested with MLST, 171 STs were identified, of which 39 were new allelic combinations (ST13406 - ST13410, ST13412 - ST13430, ST13433 - ST13436, ST13438, ST13439, ST13441 - ST13443, ST13449 - ST13452, ST13548 and ST13549) and 4 new alleles (*spi* 582, *gki* 601, *gki* 602 and *recP* 418). We also found 91 isolates with STs characteristic of PMEN clones and these were of clones Netherlands^{3}-180 (n=28), Norway^{NT}-344 (n=11), Sweden^{15A}-63 (n=11), Netherlands^{7F}-191 (n=8), USA^{NT}-448 (n=7), Greece^{21}-193 (n=6), Spain^{9V}-156 (serotype 14 variant) (n=5), Netherlands^{8}-53 (n=4), Netherlands^{15B}-199 (n=3), Portugal^{19F}-177 (n=3), Sweden^{1}-306 (n=2), Denmark^{14}-230 (serotype 24 variant) (n=2) and Colombia^{23F}-338 (n=1).

After goeBURST analysis we were able to see how the data clusters into CCs. In Table 3 we can see the descriptive analysis of each ST grouped according to serotype and CC. The results are summarized in Figure 2.

4.2 Evolution of vaccine-types

It is known that, since vaccine introduction, vaccine-types have been declining in their frequency among carriers (28). We investigated if the clones that were in circulation in the late-PCV13 period were the same found in the pre-PCV13 and early-PCV13 periods.

Table 4 summarizes the results obtained over time for both regions, regarding PCV13 serotypes. We found that, for most vaccine serotypes, it was one major clone that was maintained over time and that was responsible for the persistence of that serotype after PCV13 introduction.

Analysis of Table 5, that shows the antimicrobial resistance profiles of vaccine-type lineages, suggests that selection of antimicrobial resistant isolates may have occurred in some cases.

Serotypes 4, 6B, 9V, 14, 18C, 19F and 23F were targeted by PCV7 which was available in Portugal between 2001 and 2010.

Serotypes 4 and 18C have never been very frequent in carriage, not even before the introduction of PCV7 (10). In this study we did not detect serotype 4. Serotype 18C was only found in the late-PCV13 period and in the rural region; interestingly it was associated with a new ST (ST13409) resistant to SXT. This ST is an SLV of the widespread lineage ST113 (54).

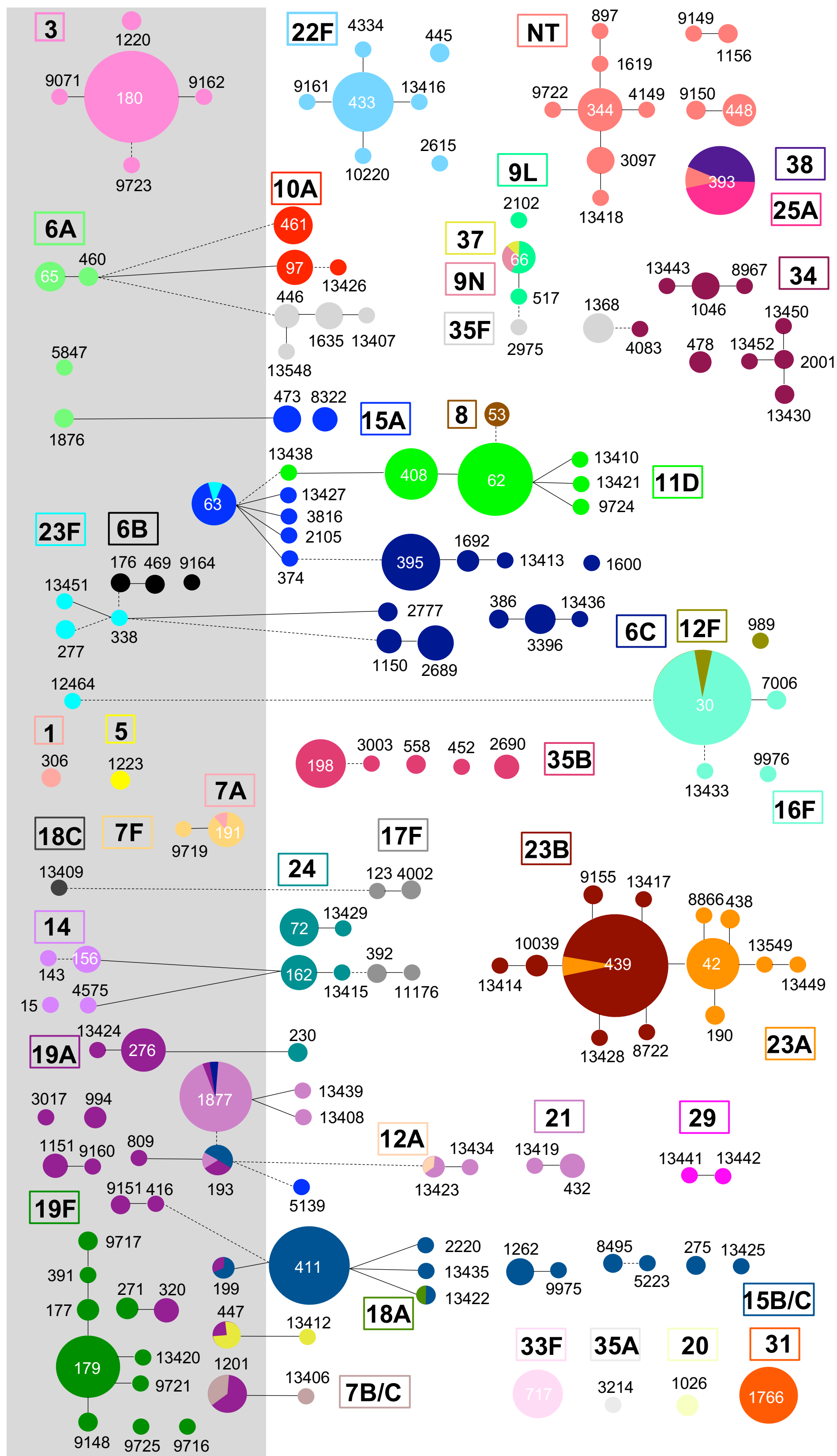


Figure 2: Clonal lineages of carriage isolates in Portugal (2009-2016).

Each circle represents one ST which is indicated inside the circle. Each color represents a serotype as indicated in the rectangles. The size of the circle is proportional to the number of isolates. Full lines indicate SLVs and dotted lines indicate DLVs. The length of the line does not have any meaning. The grey area indicates STs found in vaccine-types.

Table 3: Allelic profiles obtained in this study by serotype and geographic area.

PCV13 serotypes	Allelic profile							ST	CC	No. of isolates	Urban/rural
	aroE	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*				
1	12	8	13	5	16	4	20	306	-	2	Urban
3	7	15	2	10	6	1	22	180	180	28	Urban (20) / Rural (8)
	7	15	2	10	6	1	14	9071	180	1	Rural
	2	15	2	10	6	1	22	9162	180	1	Rural
	7	15	412	10	6	1	144	9723	180	1	Rural
	26	1	15	14	9	16	19	1220	-	2	Urban (1) / Rural (1)
5	16	12	9	1	6	33	33	1223	-	2	Rural
6A	2	7	4	10	10	1	27	65	446	6	Urban
	5	7	4	10	10	1	27	460	446	2	Urban (1) / Rural (1)
	7	25	4	12	15	20	28	1876	-	2	Rural
	1	5	74	8	13	342	17	5847	-	1	Rural
6B	7	13	8	6	10	6	14	176	338	2	Urban
	7	13	27	6	10	6	14	469	338	2	Urban
	7	8	4	18	15	4	31	9164	-	1	Rural
7F	8	9	2	1	6	1	17	191	-	7	Urban (2) / Rural (5)
	8	9	2	1	94	1	17	9719	-	1	Rural
14	7	5	10	8	6	8	1	143	162	1	Urban
	7	11	10	1	6	8	1	156	162	5	Urban (2) / Rural (3)
	7	11	10	1	6	8	248	4575	162	3	Urban
	1	5	4	5	5	3	8	15	-	1	Urban
18C	7	2	1	1	10	1	22	13409	123	1	Rural
19A	8	10	2	16	1	26	1	193	1877	2	Urban
	8	10	4	16	1	26	1	809	1877	1	Urban
	10	10	41	16	1	26	1	1877	1877	1	Urban
	8	13	14	4	17	4	14	199	411	1	Urban
	1	13	14	4	17	51	14	416	411	1	Urban
	1	392	14	4	17	51	14	9151	411	2	Urban
	2	19	2	17	6	22	14	276	276	11	Urban (7) / Rural (4)
	2	19	2	17	17	22	14	13424	276	1	Urban
	7	60	9	8	6	3	29	1151	-	4	Urban (3) / Rural (1)
	1	60	9	8	6	3	29	9160	-	1	Rural
	4	16	19	15	6	20	1	320	-	4	Urban
	29	33	19	1	36	22	31	447	-	1	Rural
	5	5	62	5	6	11	14	994	-	3	Urban (2) / Rural (1)
	1	5	1	12	17	3	8	1201	-	6	Rural
	8	11	14	1	17	230	14	3017	-	1	Rural
19F	7	14	4	12	1	1	14	177	179	3	Urban (1) / Rural (2)
	7	14	40	12	1	1	14	179	179	17	Urban (9) / Rural (8)
	7	14	4	12	1	20	14	391	179	1	Urban
	7	14	394	12	1	1	14	9148	179	2	Urban
	1	14	4	12	1	20	14	9717	179	2	Urban
	293	14	40	12	1	1	14	9721	179	1	Rural
	7	14	40	10	1	1	14	13420	179	1	Rural
	4	16	19	15	6	20	26	271	-	3	Urban
	1	5	19	15	6	20	17	9716	-	1	Urban
	15	17	4	259	6	19	17	9725	-	1	Urban
23F	7	13	8	6	6	12	8	277	338	2	Urban
	7	13	8	6	1	6	8	338	338	1	Urban
	7	13	601	6	1	6	8	13451	338	1	Urban
	2	5	36	12	17	21	14	63	-	1	Urban
	1	5	27	20	6	596	1	12464	-	1	Urban

Table 3: (cont.)

Non-PCV13 serotypes	Allelic profile							ST	CC	No. of isolates	Urban/rural
	aroE	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*				
6C	32	28	1	1	15	52	14	386	3396	2	Urban
	32	28	1	1	15	16	14	3396	3396	6	Urban (5) / Rural (1)
	32	28	1	1	15	40	14	13436	3396	1	Urban
	1	5	7	12	17	1	14	395	62	15	Urban (11) / Rural (4)
	1	5	7	12	17	158	14	1692	62	3	Urban
	1	5	27	12	17	158	14	13413	62	1	Urban
	1	5	9	12	94	1	20	1600	-	1	Urban
	7	25	8	6	25	6	8	1150	338	4	Rural
	7	25	8	6	25	28	8	2689	338	8	Urban (5) / Rural (3)
	7	13	8	6	1	1	8	2777	338	2	Rural
	10	10	41	16	1	26	1	1877	1877	1	Urban
7A	8	9	2	1	6	1	17	191	-	1	Rural
7B/C	1	5	1	12	17	3	8	1201	-	3	Urban (1) / Rural (2)
	1	5	1	12	17	26	8	13406	-	1	Rural
8	2	5	1	11	16	3	14	53	62	4	Urban (1) / Rural (3)
9L	2	8	2	4	6	1	1	66	517	4	Urban (1) / Rural (3)
	2	5	2	4	6	1	1	517	517	1	Rural
	7	60	4	8	14	12	29	2102	-	1	Rural
9N	2	8	2	4	6	1	1	66	517	2	Urban (1) / Rural (1)
10A	5	7	4	2	10	1	27	97	446	8	Urban (7) / Rural (1)
	5	40	4	1	10	1	27	461	446	9	Urban (5) / Rural (4)
	5	7	59	2	10	1	14	13426	446	1	Urban
11D	2	5	29	12	16	3	14	62	62	21	Urban (7) / Rural (14)
	2	5	6	12	16	3	14	408	62	13	Urban (10) / Rural (3)
	2	5	29	258	16	3	14	9724	62	1	Rural
	2	5	29	12	16	3	31	13410	62	1	Rural
	2	8	29	12	16	3	14	13421	62	1	Rural
	2	5	6	12	16	21	14	13438	62	1	Urban
12A	8	6	2	16	1	79	1	13423	-	1	Rural
12F	1	5	27	20	1	1	1	30	30	1	Urban
	12	5	89	8	6	112	14	989	-	1	Rural
15A	2	5	36	12	17	21	14	63	62	10	Urban (4) / Rural (6)
	2	5	36	12	17	1	14	374	62	1	Urban
	2	5	36	12	17	21	4	2105	62	2	Urban
	2	5	36	12	17	26	14	3816	62	1	Urban
	2	5	14	12	17	21	14	13427	62	1	Urban
	7	25	4	4	15	20	28	473	-	5	Rural
	8	5	2	6	1	26	1	5139	411	1	Rural
	252	13	2	16	6	12	1	8322	-	4	Urban
15B/C	8	13	14	4	17	4	14	199	411	2	Urban
	2	13	14	4	17	4	14	411	411	23	Urban (12) / Rural (11)
	10	13	14	4	17	4	14	2220	411	1	Urban
	2	13	512	4	17	4	14	13422	411	1	Rural
	2	13	14	4	17	16	14	13435	411	1	Urban
	7	41	2	6	10	26	1	1262	-	5	Urban (3) / Rural (2)
	7	41	2	6	15	26	1	9975	-	1	Urban
	43	37	4	117	36	259	74	5223	-	1	Urban
	7	5	4	117	36	259	74	8495	-	2	Urban
	10	41	47	16	6	14	17	275	-	2	Rural
	43	11	4	10	36	3	74	13425	-	1	Urban
	8	10	2	16	1	26	1	193	1877	3	Urban (1) / Rural (2)
16F	1	5	27	20	1	1	1	30	30	28	Urban (17) / Rural (11)
	1	5	27	20	1	1	476	7006	30	2	Urban
	1	5	27	20	1	596	8	13433	30	1	Rural
	7	61	5	1	6	79	6	9976	-	1	Urban

Table 3: (cont.)

Non-PCV13 serotypes	Allelic profile							ST	CC	No. of isolates	Urban/rural
	aroE	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*				
17F	7	2	40	1	10	1	45	123	123	1	Rural
	67	2	40	1	10	1	45	4002	123	2	Rural
	7	5	1	1	6	31	14	392	162	2	Urban (1) / Rural (1)
	7	5	9	1	6	31	14	11176	162	1	Urban
18A	2	13	512	4	17	4	14	13422	411	1	Rural
20	10	16	32	1	15	28	31	1026	-	3	Urban
21	8	10	2	16	1	26	1	193	1877	1	Urban
	10	10	41	16	1	26	1	1877	1877	18	Urban (16) / Rural (2)
	10	10	41	16	1	26	21	13408	1877	1	Rural
	8	6	2	16	1	79	1	13423	1877	2	Rural
	8	6	29	16	1	79	1	13434	1877	1	Rural
	10	10	41	16	1	28	1	13439	1877	1	Urban
	10	8	4	35	6	1	14	432	-	4	Urban (2) / Rural (2)
	10	5	4	35	6	1	14	13419	-	1	Rural
22F	1	1	4	1	18	58	17	433	433	16	Urban (7) / Rural (9)
	1	1	4	1	6	58	17	4334	433	1	Rural
	7	1	4	1	18	58	17	9161	433	1	Rural
	1	1	4	1	18	605	17	10220	433	1	Urban
	1	1	9	1	18	58	17	13416	433	1	Urban
	2	5	29	5	42	3	18	445	-	2	Urban
	2	5	1	16	6	192	31	2615	-	1	Urban
23A	1	8	9	9	6	4	6	42	439	13	Urban (4) / Rural (9)
	8	8	9	9	6	4	6	190	439	2	Rural
	1	5	9	9	6	4	6	438	439	2	Urban
	1	8	9	2	6	4	6	439	439	2	Urban
	241	8	9	9	6	4	6	8866	439	1	Rural
	1	8	602	9	6	3	6	13449	439	1	Rural
	1	8	602	9	6	4	6	13549	439	1	Rural
23B	1	8	9	2	6	4	6	439	439	30	Urban (16) / Rural (14)
	7	8	9	2	6	4	6	8722	439	1	Rural
	1	8	32	2	6	4	6	9155	439	2	Rural
	1	428	9	2	6	4	6	10039	439	3	Urban
	1	428	10	2	6	4	6	13414	439	1	Urban
	1	8	9	2	262	4	6	13417	439	1	Urban
	1	8	9	2	17	4	6	13428	439	1	Rural
24	7	11	10	1	6	8	14	162	162	8	Urban (5) / Rural (3)
	7	11	1	1	6	8	14	13415	162	1	Urban
	12	19	2	17	6	22	14	230	276	2	Urban
	2	13	2	4	9	4	1	72	-	9	Urban (2) / Rural (7)
	2	13	2	20	9	4	1	13429	-	1	Rural
25A	10	43	41	18	13	49	6	393	-	8	Urban (7) / Rural (1)
29	2	12	94	1	6	4	14	13441	-	1	Urban
	2	12	94	1	6	8	14	13442	-	1	Urban
31	1	5	29	1	46	14	18	1766	-	15	Urban (5) / Rural (10)
33F	5	35	29	1	45	39	18	717	-	12	Urban (5) / Rural (7)
34	15	5	2	1	9	1	18	1046	1046	5	Urban (3) / Rural (2)
	15	5	2	1	9	320	18	8967	1046	1	Rural
	15	5	2	1	9	8	18	13443	1046	1	Urban
	5	125	6	1	9	14	14	2001	2001	2	Urban (1) / Rural (1)
	5	125	6	1	9	141	14	13430	2001	1	Urban
	5	125	6	1	582	14	14	13450	2001	1	Urban
	5	125	6	418	9	14	14	13452	2001	2	Rural
	7	5	53	5	42	1	79	4083	-	1	Rural
	45	13	6	12	9	14	1	478	-	3	Urban
35A	10	17	4	38	6	14	6	3214	-	1	Rural
35B	8	13	4	8	6	22	34	198	-	12	Urban (6) / Rural (6)
	8	13	4	8	162	12	34	3003	-	1	Rural
	7	9	19	1	14	48	14	452	-	1	Urban
	18	12	4	44	14	77	97	558	-	2	Urban
	7	13	4	19	6	1	34	2690	-	4	Urban (3) / Rural (1)

Table 3: (cont.)

Non-PCV13 serotypes	Allelic profile							ST	CC	No. of isolates	Urban/rural
	aroE	*gdh*	*gki*	*recP*	*spi*	*xpt*	*ddl*				
35F	5	7	4	19	10	40	27	446	446	4	Rural
	10	7	4	19	10	40	27	1635	446	5	Urban (4) / Rural (1)
	10	7	4	19	10	40	1	13407	446	1	Rural
	5	7	30	19	10	40	27	13548	446	1	Urban
	2	5	4	1	6	1	1	2975	517	1	Urban
	7	5	4	5	42	92	79	1368	-	6	Urban (3) / Rural (3)
37	2	8	2	4	6	1	1	66	-	1	Rural
	29	33	19	1	36	22	31	447	-	4	Urban (1) / Rural (3)
	29	33	19	1	36	40	31	13412	-	1	Urban
38	10	43	41	18	13	49	6	393	-	9	Urban (5) / Rural (4)
NT	8	37	9	29	2	12	53	344	344	11	Urban (6) / Rural (5)
	8	37	36	29	2	12	14	897	344	1	Urban
	8	37	9	29	2	12	14	1619	344	1	Urban
	8	178	9	29	2	12	53	3097	344	5	Urban
	8	37	9	29	2	12	339	4149	344	1	Urban
	294	37	9	29	2	12	53	9722	344	1	Urban
	8	178	9	29	450	12	53	13418	344	1	Urban
	8	5	2	27	2	11	71	448	-	7	Urban (2) / Rural (5)
	8	5	2	27	2	11	583	9150	-	2	Urban
	2	13	2	29	91	19	59	1156	-	2	Rural
	2	13	2	29	91	19	582	9149	-	1	Urban
	10	43	41	18	13	49	6	393	-	2	Urban (1) / Rural (1)

Highlighted in bold are the new STs and alleles.

Serotype 9V disappeared upon introduction of PCV7 (10). This serotype was not detected in our study.

Serotypes 6B and 23F have declined in prevalence during the era of PCV7 but remained in circulation.

In our study, serotype 6B was maintained in the urban region due to STs 469 and 176 that belong to CC338. ST469 was associated with resistance to erythromycin. This clone has been described as one of the main serotype 6B clones associated with IPD, after PCV13 use, in Portugal (55). In the rural region, serotype 6B was detected only in the pre-PCV13 period and was associated with a singleton (ST9164).

Serotype 23F was not found in any period in the rural region (28). In the urban region serotype 23F was maintained in carriage after PCV7 introduction being associated with ST338 and low-level resistance to penicillin (19). After PCV13 implementation, CC338 continued to be the responsible for serotype 23F maintenance in the population.

Table 4: Clonal evolution of vaccine-types overtime in the urban and rural regions.

	Urban						Rural					
	Pre-PCV13		Early-PCV13		Late-PCV13		Pre-PCV13		Early-PCV13		Late-PCV13	
PCV13 serotypes	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**
1	1/2	306	1/2	306	0/0	-	0/0	-	0/0	-	0/0	-
3	15/68	180 (15)	5/20	180 (4) 1220	1/3	180	4/18	180 (3) 9162	6/26	180 (4) 1220 9723	2/6	180 9071
5	0/0	-	0/0	-	0/0	-	2/6	1223 (2)	0/0	-	0/0	-
6A	3/10	65 (3)	3/10	65 (2) 460	1/2	65	2/9	1876 5847	2/2	1876 460	0/0	-
6B	1/2	469	2/6	469 176	1/5	176	1/3	9164	0/0	-	0/0	-
7F	1/3	191	1/1	191	0/0	-	4/15	191 (4)	2/6	191 9719	0/0	-
14	2/6	15 143	4/16	4575 (3) 156	1/2	156	1/4	156	2/9	156 (2)	0/0	-
18C	0/0	-	0/0	-	0/0	-	0/0	-	0/0	-	1/1	**13409**
19A	18/95	276 (6) 320 (2) 994 193 1877 **13424** 1151 (3) 416 9151 (2)	5/20	276 320 (2) 994 199	2/5	193 809	11/50	276 (3) 994 9160 1201 (4) 447 3017	3/14	1201 (2) 1151	1/2	276
19F	11/48	179 (5) 177 9717 9148 (2) 271 9725	6/27	179 (2) 9717 271 (2) 9716	3/11	179 (2) 391	3/12	179 177 (2)	4/14	179 (3) 9721	5/19	179 (4) **13420**
23F	2/3	63 **13451**	2/7	277 (2)	2/9	338 12464	0/0	-	0/0	-	0/0	-

Highlighted in bold are the new STs. When a STs in underlined, it means that at least one isolate is resistant to at least one antimicrobial agent among the following: penicillin, chloramphenicol, erythromycin, clindamycin, tetracycline and sulfamethoxazole-trimethoprim.

Table 5: Resistant samples and correspondent antimicrobial resistance profiles of vaccine-types, overtime.

PCV13 serotypes	ST	CC	No. of resistant isolates	Antimicrobial resistance profile					
				Urban			Rural		
				Pre-PCV13	Early-PCV13	Late-PCV13	Pre-PCV13	Early-PCV13	Late-PCV13
6A	5847	-	1	-	-	-	S	-	-
6B	469	338	2	E (1)	E (1)	-	-	-	-
14	15	-	1	P,E,C,S	-	-	-	-	-
	143	162	1	P	-	-	-	-	-
	156	162	5	-	P,S (1)	P,S (1)	P,S (1)	P,S (2)	-
	4575	162	3	-	P,E,S (3)	-	-	-	-
18C	**13409**	123	1	-	-	-	-	-	S
19A	193	1877	2	E,C,T,Ch (1)	-	E,C,T (1)	-	-	-
	276	276	11	P,E,C,T (6)	P,E,C,T (1)	-	P,E,C,T (3)	-	P,E,C,T,S (1)
	320	-	4	P,E,C,T,S (2)	P,E,C,T,S (2)	-	-	-	-
	416	411	1	E,C,T	-	-	-	-	-
	809	1877	1	-	-	E,C,T	-	-	-
	9151	411	2	E,C,T (2)	-	-	-	-	-
	13424	276	1	P,E,T	-	-	-	-	-
19F	179	179	16	E,C,T (5)	E,C,T (2)	E,C,T (1)	E,C,T (1)	E,C,T (3)	E,C,T (4)
	271	-	3	P,E,C,S (1)	P,E,C,S (2)	-	-	-	-
	9716	-	1	-	S	-	-	-	-
	9721	179	1	-	-	-	-	E,C,T	-
	9725	-	1	S	-	-	-	-	-
23F	63	-	1	P,E,C,T	-	-	-	-	-
	277	338	2	-	P,S (2)	-	-	-	-
	338	338	1	-	-	P,S	-	-	-
	13451	338	1	P	-	-	-	-	-

S, sulfamethoxazole-trimethoprim; E, erythromycin; P, penicillin (when in bold indicates high level resistance ≥ 2 µg/mL); C, clindamycin; T, tetracycline; Ch, chloramphenicol.

Highlighted in bold are the new STs.

In parenthesis are indicated the number of isolates for that profile. When the number is not present it means that all the isolates from a specific ST have the same profile.

Serotype 14 increased significantly its prevalence in the early-PCV13 period in both regions. This was followed by a significant decrease in the late-PCV13 period (28). In our study, all but one serotype 14 isolate were mainly associated with CC162. Isolates belonging to this CC, which is associated with PMEN clone Spain^{9V}-156, have been in circulation in Portugal at least since 1996 (19). This clone has spread in the pre-PCV era globally, being an important cause of IPD worldwide (56). It has a high invasive disease potential (25) and is capable of adapting to several selective pressures (57). This clone is known for the presence of pili, an important virulence factor for pneumococci, that has been implicated in the persistence and spread of specific clones (58). CC156 has been described as a major cause of IPD in the adult population in Portugal, in a study conducted between 2008 and 2011 (57).

Serotype 19F has been found to be the most frequent vaccine-type in children attending day-care centers in Portugal, since introduction of PCVs (28). ST177 (also known as PMEN clone Portugal^{19F}-177) was one of the few vaccine clones that persisted among carriers after PCV7 introduction, being associated with serotype 19F and with multiresistance to antimicrobial

agents (macrolides, lincosamides, streptogramins, tetracyclines and low-level penicillin) (19). In our study, it is clear that, CC179 was responsible for the persistence of the vaccine-type 19F after PCV13 implementation. All isolates from ST179 were multiresistant and were present in all periods in both the urban and rural regions.

With the introduction of PVC13 in 2010, 6 additional serotypes were targeted: 1, 3, 5, 6A, 7F and 19A.

Serotypes 1, 5 and 7F have never been very prevalent in carriage studies in Portugal conducted in children (10). These serotypes have been identified as major causes of invasive disease in several European countries, especially among children aged between 2 and 5 years of age (59). These serotypes are known for having high invasive disease potential and being rarely carried in the nasopharynx by healthy individuals (56). Serotype 1 has been frequently associated with outbreaks of disease mainly due to ST306, representative of the Sweden^{1}-306 clone, particularly found in European countries (60). Interestingly, this clone expresses a non-haemolytic pneumolysin suggesting that this alteration may be neutral or even advantageous, given its high invasive potential (3, 56). After PCV7 introduction, serotypes 1 and 7F were among the most prevalent serotypes responsible for pediatric IPD in Portugal in all ages studied (61). In 2006 both these serotypes experienced an increase in asymptomatic carriage among healthy children attending day-care centers, associated with STs 306 and 191 (also known as PMEM clone Netherlands^{7F}-191), respectively (62). After being introduced in PCV10 and PCV13 their frequency in carriage and invasive disease in person aged under 18 years old significantly decreased (63). However, a study regarding invasive disease in the same population in Spain, between 2010 and 2013 still found serotypes 1 and 7F among the most common (58). The same was true for the adult population of Portugal once, when studying the clonal population causing invasive disease between 2008 and 2011, the clones more frequently found were ST191 (9.9%) and ST306 (7.0%) associated with serotypes 7F and 1, respectively (57). In our study, serotypes 1, 5 and 7F were not present in the late period in both regions, and each one was associated with one clone only: ST306, ST191 and ST1223, respectively. None was associated with antimicrobial resistance.

Serotype 3 was among the most frequently found serotypes in the pre-PCV13 period and, remained in circulation in the late-PCV13 period, although its prevalence declined (28). In our study, serotype 3 isolates from both regions belonged to CC180. ST180 (also known as PMEN clone Netherlands^{3}-180) was among the most frequent clones (n=28, 4.3%) found in the tested collection and it was never associated with antimicrobial resistance. This clone has been detected in carriage among children attending day-care centers in Portugal since the pre-PCV7

era (50). In a colonization study performed in Massachusetts, ST180 associated with serotype 3 increased in prevalence after PCV13 implementation, and the authors correlated this with the low immunoglobulin G response of this serotype (64). Regarding adult IPD in Portugal, ST180 expanded between 2008 and 2011, and, CC180 accounted for 64% of serotype 3 related IPD in the studied population (57). When analyzing the sequences of serotype 3 isolates from 20 countries, the authors found that 81% of serotype 3 isolates were from CC180 (54). ST1220 (a singleton) appeared in both regions in the early-PCV13 period. Of interest, in a study comparing the invasive disease potential of clones in circulation before PCVs, ST1220 showed a significantly high invasive disease potential (31).

Serotype 6A was maintained in the urban region due to STs 65 and 460 that belong to CC446. The major ST present in the rural region was ST1876, being eliminated in the late-PCV13 period. All these STs have been described as being associated with IPD due to serotype 6A in the PCV13 era (55).

Serotype 19A was the most diverse serotype, accounting for 15 STs in the three periods and in both regions and was frequently associated with antimicrobial resistance. Serotype 19A prevalence decreased significantly over time in the PCV13 era (10), and the same was observed for clonal diversity. In the rural region, the multiresistant ST276, SLV of the PMEN clone Denmark^{14}-230, was the only clone detected in the late-PCV13 period. This is interesting once in the previous periods the major clone was ST1201, associated with susceptibility to antimicrobial agents. Besides that, it is also interesting to see that ST1201 was exclusively present in the rural region. In the urban region, the scenario was different. ST276 was a major clone in the pre-PCV13 period, was maintained through the early-PCV13 period, but was not found in the late-PCV13 period. Instead, in the late-PCV13 period we found STs 193 and 809, belonging to CC1877. ST193 belongs to the PMEN clone Greece^{21}-193 and in our study all isolates were multiresistant. This clonal complex was not present in the rural region. ST276 was the most frequently found ST in invasive samples from both children and adults in Portugal (57, 65). It was also reported that the increase in ST276 was directly related with the increase in antimicrobial resistance, as all isolates were penicillin non-susceptible (57). Serotype 19A has already been described as very diverse in terms of clonal distribution and, besides that, different clones are known to have different invasive disease potentials. For instance, ST994 has been associated with carriage isolates and STs 276, 193, 416 and 1151 have been associated with invasive disease (31). In Spain, the multiresistant ST320 was the main clone responsible for the emergence of this serotype after PCV7 introduction and, even after PCV13, this ST continued to prevail (58). ST1151 was found in a carriage study of children in Portugal between

2001 and 2003, associated with serotype 19A (51). Clonal type ST416 was first found in samples recovered from the nasopharynx of children in Portugal between 2001 and 2003 (51). In our study, this ST was only found in the urban region, in the pre-PCV13 period and in the early-PCV13 period we detected a SLV ST199. This ST199 (also known as PMEN clone Netherlands^{15B}-199) was frequent between 1996 and 2007, in Portugal, being associated with resistance to antimicrobial agents, but decline afterwards (19). Colonization studies performed after PCV13 implementation in Massachusetts found that serotype 19A was still highly prevalent but with different clonal distribution due to the decrease of ST199 and the increase of the multiresistant ST320 (64).

4.3 Non-vaccine types analysis

Introduction of PCV13 led to serotype replacement in carriage, that is, the frequency of non-vaccine types increased (28). We investigated the diversity of clones leading to this phenomenon, aiming to understand how the pneumococcal population evolved.

Table 6 summarizes the results obtained over time for both regions, regarding non-PCV13 serotypes. The role of one specific clonal lineage on the persistence of specific non-vaccine types was not as obvious as with the vaccine-types. In some cases, we observed that a major clone in the pre-PCV13 and early-PCV13 periods was not detected in the late-PCV13 periods, being replaced by new previously undetected clones, of the same serotype.

Table 7, shows the antimicrobial resistance profiles of non-vaccine type lineages. In contrast to what we observed with the vaccine-types, is not always the resistant clone that is maintained through the different periods.

PCV13 impacted differently among the NVTs, once we have serotypes that decreased, others that emerged, and some that remain in low or high frequency overtime (28).

In low frequency we found serotypes 7A, 7B/C, 9L, 9N, 12A, 12F, 17F, 18A, 20, 25A, 29, 31, 35A, 37 and 38.

Serotype 7A was only detected in the pre-PCV13 period, in the rural region, being associated with ST191.

Serotype 7B/C was only found in the late-PCV13 period, associated with ST1201 in both regions and with the new ST13406 in the rural region.

Table 6: Clonal evolution of non-vaccine types overtime in the urban and rural regions.

	Urban						Rural					
	Pre-PCV13		Early-PCV13		Late-PCV13		Pre-PCV13		Early-PCV13		Late-PCV13	
Non-PCV13 serotypes	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**	**No. of samples tested among total**	**STs (n)**
6C	18/85	395 (9) 1692 2689 (3) 1877 3396 (4)	8/36	395 (2) 1692 2689 (2) 3396 **13436** 1600	4/15	386 (2) 1692 **13413**	7/30	395 (4) 1150 (3)	6/25	3396 1150 2689 (2) 2777 (2)	1/5	2689
7A	0/0	-	0/0	-	0/0	-	1/5	191	0/0	-	0/0	-
7B/C	0/0	-	0/0	-	1/1	1201	0/0	-	0/0	-	3/11	1201 (2) **13406**
8	0/0	-	0/0	-	1/3	53	0/0	-	0/0	-	3/10	53 (3)
9L	0/0	-	0/0	-	1/1	66	4/12	66 (2) 517 2102	0/0	-	1/1	66
9N	0/0	-	1/2	66	0/0	-	1/1	66	0/0	-	0/0	-
10A	3/15	461 (2) **13426**	7/30	97 (5) 461 (2)	3/12	97 (2) 461	1/1	461	1/1	461	3/7	461 (2) 97
11D	4/18	408 (3) 62	6/24	408 (4) 62 **13438**	8/35	408 (3) 62 (5)	9/41	62 (6) 408 (2) 9724	8/35	62 (7) 408	3/13	62 **13410** **13421**
12A	1/1	**13423**	0/0	-	0/0	-	0/0	-	0/0	-	0/0	-
12F	1/1	30	0/0	-	0/0	-	0/0	-	0/0	-	1/1	989
15A	9/40	63 (2) 2105 (2) **13427** 8322 (4)	2/9	63 3816	2/6	63 374	2/7	63 (2)	4/19	63 (4)	6/27	473 (5) 5139
15B/C	10/45	411 (5) 199 5223 8495 **13425** 1262	6/27	411 (4) 8495 **13435**	9/38	411 (3) 193 199 1262 (2) 2220 9975	7/32	411 (6) **13422**	8/35	411 (4) 275 (2) 193 (2)	3/11	411 1262 (2)
16F	7/31	30 (7)	6/25	30 (5) 7006	7/31	30 (5) 7006 9976	6/26	30(6)	2/6	30 **13433**	4/17	30 (4)

Table 6: (cont.)

Non-PCV13 serotypes	Urban						Rural					
	Pre-PCV13		Early-PCV13		Late-PCV13		Pre-PCV13		Early-PCV13		Late-PCV13	
	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)
17F	0/0	-	1/1	392	1/1	11176	0/0	-	3/9	4002 (2) 392	1/1	123
18A	0/0	-	0/0	-	0/0	-	1/2	**13422**	0/0	-	0/0	-
20	1/1	1026	1/2	1026	1/1	1026	0/0	-	0/0	-	0/0	-
21	8/35	1877 (8)	6/30	1877 (3) **13439** 193 432	6/27	1877 (5) 432	3/9	1877 **13423** (2)	2/7	1877 **13434**	4/17	432 (2) **13408** **13419**
22F	4/15	433 (3) 2615	1/2	433	7/30	433 (3) 445 (2) 10220 **13416**	3/9	433 4334 9161	19/5	433 (5)	4/17	433 (3)
23A	4/18	42 (2) 439 (2)	1/5	438	3/10	42 (2) 438	4/13	42 (2) 190 8866	5/20	42 (4) 190	5/22	42 (3) **13449** **13549**
23B	7/33	439 (7)	6/29	439 (6)	8/36	439 (3) 10039 (3) **13414** **13417**	8/36	439 (5) 8722 9155 **13428**	7/31	439 (6) 9155	3/11	439 (3)
24	3/9	230 (2) 72	2/6	72 162	5/26	162 (4) **13415**	3/9	72 (2) **13429**	4/18	72 (4)	4/13	72 162 (3)
25A	1/1	393	4/16	393 (4)	2/6	393 (2)	0/0	-	0/0	-	1/2	393
29	0/0	-	2/7	**13441** **13442**	0/0	-	0/0	-	0/0	-	0/0	-
31	0/0	-	4/17	1766 (4)	1/1	1766	2/9	1766 (2)	6/22	1766 (6)	2/3	1766 (2)
33F	0/0	-	3/11	717 (3)	2/3	717 (2)	4/14	717 (4)	3/12	717 (3)	0/0	-
34	4/14	478 (3) **13450**	4/16	1046 (2) **13430** **13443**	2/10	1046 2001	2/9	1046 (2)	3/12	2001 **13452** (2)	2/5	8967 4083
35A	0/0	-	0/0	-	0/0	-	1/2	3214	0/0	-	0/0	-
35B	4/10	198 (3) 558	4/17	198 (3) 558	4/15	2690 (3) 452	4/17	198 (3) 2690	2/7	198 3003	2/7	198 (2)
35F	2/4	1368 **13548**	3/8	1368 (2) 2975	4/16	1635 (4)	1/5	1368	3/10	446 (2) 1368	5/21	446 (2) 1368 1635 **13407**

Table 6: (cont.)

Non-PCV13 serotypes	Urban						Rural					
	Pre-PCV13		Early-PCV13		Late-PCV13		Pre-PCV13		Early-PCV13		Late-PCV13	
	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)	No. of samples tested among total	STs (n)
37	0/0	-	0/0	-	2/5	447 **13412**	1/2	66	2/9	447 (2)	1/2	**13411**
38	5/21	393 (5)	0/0	-	0/0	-	2/7	393 (2)	2/6	393 (2)	0/0	-
NT	9/43	344 (2) 3097 (3) 9722 9149 9150 (2)	7/28	344 (3) 3097 897 4149 448	6/25	344 393 448 1619 3097 **13418**	7/37	344 (3) 448 (4)	4/16	344 (2) 1156 (2)	2/7	393 448

Highlighted in bold are the new STs. When a STs in underlined, it means that at least one isolate is resistant to at least one antimicrobial agent among the following: penicillin, chloramphenicol, erythromycin, clindamycin, tetracycline and sulfamethoxazole-trimethoprim.

Table 7: Resistant samples and correspondent antimicrobial resistance profile of non-vaccine types, overtime.

Non-PCV13 serotypes	ST	CC	No. of resistant isolates	Antimicrobial resistance profile					
				Urban			Rural		
				Pre-PCV13	Early-PCV13	Late-PCV13	Pre-PCV13	Early-PCV13	Late-PCV13
6C	386	3396	1	-	-	P,E,C,T	-	-	-
	1150	338	1	-	-	-	P	-	-
	2777	338	1	-	-	-	-	P	-
	3396	3396	6	P,E,C,T (4)	E,C,T (1)	-	-	E,C,T (1)	-
	13436	3396	1	-	E,C,T (1)	-	-	-	-
10A	97	446	4	-	E,C (4)	-	-	-	-
	13426	446	1	E,C	-	-	-	-	-
11D	62	62	1	-	S	-	-	-	-
	408	62	2	S (1)	-	-	S (1)	-	-
12F	989	-	1	-	-	-	-	-	T,Ch
15A	63	62	10	P,E,C (2)	P,E,C (1)	P,E,C,T (1)	P,E,C (2)	P,E,C (4)	-
	374	62	1	-	-	P,E,C,T	-	-	-
	2105	62	2	P,E,C,T (2)	-	-	-	-	-
	3816	62	1	-	P,E,C,T	-	-	-	-
	13427	62	1	E,C	-	-	-	-	-
15B/C	193	1877	1	-	-	E,C,T	-	-	-
	275	-	2	-	-	-	-	P (2)	-
	411	411	1	-	P	-	-	-	-
	1262	-	3	S (1)	-	S (2)	-	-	-
	5223	-	1	S (1)	-	-	-	-	-
	8495	-	2	S (1)	S (1)	-	-	-	-
	13425	-	1	S	-	-	-	-	-
16F	7006	30	2	-	E,C,T,Ch (1)	E,C,T (1)	-	-	-
24	162	162	8	-	S (1)	S (4)	-	-	S (3)
	230	276	2	P,E,C,T (2)	-	-	-	-	-
	13415	162	1	-	-	S	-	-	-
	13429	-	1	-	-	-	E	-	-
33F	717	-	12	-	E,C,T (3)	E,C,T (2)	E,C (4)	E,C (3)	-
34	1046	1046	5	-	S (2)	S (1)	S (2)	-	-
	8967	1046	1	-	-	-	-	-	S
	13443	1046	1	-	S	-	-	-	-
35A	3214	-	1	-	-	-	E,T,S	-	-
35B	198	-	1	-	-	-	P	-	-
	558	-	2	P (1)	P (1)	-	-	-	-
NT	344	344	11	P,E,C,T,S (2)	P,E,C,T,S (3)	E,C,T,S (1)	P,E,C,T,S (3)	P,E,C,T,S (2)	-
	897	344	1	-	E,C,T,S	-	-	-	-
	1156	-	2	-	P,E,C,T,S (2)	-	-	-	-
	1619	344	1	-	-	P,E,C,T,S	-	-	-
	3097	344	5	P,E,C,T,S (3)	P,E,T,S (1)	E,C,T,S (1)	-	-	-
	4149	344	1	-	P,E,C,T,S	-	-	-	-
	9149	-	1	P, S	-	-	-	-	-
	9722	344	1	P,E,C,T,S	-	-	-	-	-
	13418	344	1	-	-	P,E,C,S	-	-	-

S, sulfamethoxazole-trimethoprim; E, erythromycin; P, penicillin (when in bold indicates high level resistance ≥ 2 μg/mL); C, clindamycin; T, tetracycline; Ch, chloramphenicol.

Highlighted in bold are the new STs.

In parenthesis are indicated the number of isolates for that profile. When the number is not present it means that all the isolates from a specific ST have the same profile.

Serotype 9L was detected in the pre-PCV13 period only in the rural region. In the late-PCV13 period, this serotype was present in both regions and it was always associated with CC517, mainly with ST66.

Serotype 9N was only present in the early-PCV13 period in the urban region and in the pre-PCV13 period in the rural region and it was always associated with the aforementioned ST66. This ST has been associated with low invasive disease potential (25), however, between 2008 and 2011 it was reported as the main clonal lineage related with serotype 9N IPD in adults in Portugal (57). In the PCV13 period, serotype 9N isolates recovered from IPD of adults and children in Spain, were clonally diverse, however, ST66 was the most prevalent (58). A study performed to analyze genomes from different serotypes in different time periods and locations found that, in five countries and from 1938 to 2014, CC66 was one of the two major lineages responsible for serotype 9N distribution (54).

Serotype 12A was only detected in the urban region and in the pre-PCV13 period, interestingly associated with a new ST, ST13423. The same was true for serotype 18A as it was only found in the pre-PCV13 period, but this time in the rural region, and it was associated with a new ST, ST13422. We only analyzed two isolates belonging to serotype 29 and these were present in the early-PCV13 period, in the urban region. These isolates were also associated with new STs, ST13441 and ST13442.

Serotype 12F was found in the early-PCV13 period in the urban region associated with ST30. In the rural region, this serotype was only detected in the late-PCV13 period, associated with ST989, showing resistance to tetracycline and chloramphenicol. This ST has been correlated with the increase in serotype 12F IPD found in Spain in the PCV13 era (58).

Serotype 17F was only detected in the PCV13 era: in the urban region it was associated with CC162; in the rural region, it was associated with CC123.

Serotype 20 was only present in the urban region and, in all study periods, it was associated with ST1026.

Serotype 25A was found in all periods in the urban region and, in the rural region, it was only detected in the late-PCV13 periods. All analyzed samples belonged to ST393. This ST was also detected when analyzing samples from serotype 38, however in this case it was only present in the pre-PCV13 period in the urban region and in the pre-PCV13 and early-PCV13 periods in the rural region.

Serotype 31 was only absent in the pre-PCV13 period, in the urban region. It was always associated with ST1766.

Serotype 35A was only detected in the pre-PCV13 period, in the rural region, and the isolate analyzed belonged to the multiresistant ST3214. In the MLST database, this ST was first submitted in 2007 but only in 20011 it was associated with serotype 35A (66).

Serotype 37 was found in all periods in the rural region; in the urban region it was only present in the late-PCV13 period. It was mainly associated with ST447, however, in the late-PCV13 period we found two isolates, one from the urban and other form the rural regions, with new STs, ST13412 and ST13411, respectively.

For serotypes 6C, 15A and 33F, a decrease in frequency was found after PCV13 introduction.

Serotype 6C was one of the few non-vaccine serotypes that decreased its frequency after PCV13 introduction (28) probably due to a cross-protection effect from serotype 6A, present in PCV13 (55, 64). We found that this serotype was mainly associated with three clonal complexes: CC3396, CC62 and CC338, represented by a total of three STs each. CC62, mostly represented by ST395, was the prevalent ST in the pre-PCV13 period, in both regions. ST395 has been detected in our population since 1996, always associated with serotype 6C and susceptibility to antimicrobial agents (67). In our study, in the PCV13 periods, the clonal evolution differed between the urban and the rural regions but in both regions the decline in serotype 6C prevalence was accompanied by a decline in the number of STs in circulation, a scenario similar to that observed for vaccine-types. For the urban region, in the early-PCV13 period, all CCs were represented, but in the late-PCV13, the most frequent ST was the multiresistant ST386 (a SLV of the multiresistant ST3396 that was not detected in the late-PCV13 period). ST3396 was previously found in colonization of children in Portugal, after PCV7 introduction (19, 67). In the rural region, both in the early and late-PCV13 periods, the major CC associated with serotype 6C was CC338. In the late-PCV13 period, the only clone that was found belonged to ST2689, a susceptible SLV of ST1150. ST1150 has been found as one of the major clones responsible for serotype 6C distribution along time, and some isolates were found to be non-susceptible to penicillin (67).

Serotype 15A showed a significant decrease in frequency in the urban region in the late-PCV13 period. In the rural region the trend was opposite (28). In the urban region, CC62 was the most frequent CC. In the pre-PCV13 period, the susceptible ST8322 was the most prevalent, however, the multiresistant ST63 was the only clone that was maintained overtime. After PCV7 introduction, ST63 (also known as PMEN clone Sweden^{15A}-63) associated with serotypes 15A and 19A, showed an increase that was correlated with the maintenance of the antimicrobial resistant levels among colonized children of Portugal (19). In the rural region, ST63 played an important role in the pre-PCV13 and early-PCV13 periods, however, in the late-PCV13 period this clone disappeared and the susceptible ST473 (SLV of ST1876 related with serotype 6A distribution) and ST5139 emerged.

Serotype 33F was only detected in the late-PCV13 period in the urban region, but in both regions, it was always associated with ST717, resistant to erythromycin and clindamycin in both regions, together with tetracycline in the rural region. This ST was first detected in Portugal, early in 1999, in a colonization study, already associated with serotype 33F and with antimicrobial resistance (49). Between 2001 and 2003, ST717 was found to be one of the major novel international drug-resistant clones identified in day-care centers in Portugal (51).

Serotype 8 emerged in late-PCV13 period, in both regions, being associated with ST53. ST53 (also known as PMEN clone Netherlands^{8}-53) has a high invasive disease potential (31) and was among the most frequently found clones (4.5%) in adult IPD, in Portugal, before PCV13 introduction (57). Regarding IPD in Spain, between the years 2010 and 2013, in the PCV13 era, serotype 8 was among the 10 most frequently found serotypes and this was associated with ST53 (58).

Serotypes 10A, 11D, 15B/C, 16F, 21, 22F, 23A, 23B, 24, 34, 35B, 35F and NT have been increasing in frequency in the population, overtime.

Serotype 10A was found to be related with only one clonal complex, CC446, and two variants, ST461 and ST97 (double locus variants of each other) that were detected in both regions. In the urban region, besides the high frequency of ST461 in the pre-PCV13 period, in the following periods ST97 prevailed. In the rural region ST461 was the most frequent variant. Before PCV13, ST97 was the major responsible for serotype 10A related IPD in adults, in Portugal (57). The same was true after PCV13 introduction in Spain, regarding adults and children (58).

Serotype 11D was found in high frequency in the late-PCV13 period, especially in the urban region (28). We found this serotype to be very clonal, being distributed to only one clonal complex, CC62. ST62 and ST408 were the major clones found along time. CC62 was one of the most frequent clonal complexes found before PCV13 introduction, in Portugal IPD adult cases, associated with serotypes 8 and 11A (57). In the same study, for serotype 11A, the most prevalent clone was ST408 (57). ST62 was also reported in a colonization study in Massachusetts after PCV13 use (64).

Serotype 15B/C was among the most prevalent non-vaccine serotypes found in colonization, in both regions, in the late-PCV13 period (28). In our study we found this serotype to be very diverse, in particular in the urban region. However, CC411 was the major responsible for serotype 15B/C maintenance in the population along time, especially due to ST411 and ST199. In the PCV7 period, ST199 showed an increase in prevalence in a colonization study with children, performed in Portugal, particularly among drug-resistant isolates, however it was

associated with serotype 19A (19). A study conducted in Massachusetts in the PCV13 period, showed that serotype 15B/C was a frequent serotype in colonization of young children and this was mainly due to ST199 (64). ST1262 increased its prevalence in the late-PCV13 period, and, in the rural region it became the most prevalent clonal lineage. In the study performed in Massachusetts, they also saw the emergence of other clones associated with serotype 15B/C, including ST1262 (64).

Serotype 16F was one of the most frequent non-vaccine serotypes found in the late-PCV13 period (28). The maintenance of this serotype was correlated with the susceptible ST30 and its SLV ST7006 in both regions. ST30 was one of the second most frequently found STs in our study, accounting for 4.4% of all tested isolates.

Serotype 21 was among the most frequent non-vaccine serotypes found in colonization after PCV13 implementation, in both regions (28). CC1877 was the major clonal complex responsible for this serotype maintenance along time, in both regions. However, in the early-PCV13 period in the urban, and in the late-PCV13 period in the rural region, ST432, unrelated with CC1877, was also detected. ST432 was described in a colonization study in the PCV13 period in Massachusetts, associated with serotype 21, as one of the clones that expanded after PCV13 implementation (64). Interestingly, in the rural region we found 4 new STs associated with this serotype in the different periods: in the pre-PCV13 period ST13423, in the early-PCV13 period ST13434 and in the late-PCV13 period ST13408 and ST13419; in the urban region we also found one new ST, ST13439. From all this new STs, only ST13419 was not associated in CC1877, but instead it was an SLV of ST432.

Serotype 22F increased overtime, in both regions (28). CC433 was the most frequent CC associated with this serotype, in both regions. ST433 accounted for 3.2% of IPD cases of adults found in Portugal before PCV13 (57). This ST is known for having a high invasive disease potential (31). ST433 was described in a colonization study in the PCV13 period in Massachusetts, associated with serotype 22F, as one of the clones that expanded after PCV13 implementation (64). After analyzing a large database of genome sequencing results, the authors described that 89% of serotype 22F isolates recovered from 1940 to 2014, between four countries, belonged to CC433 (54). In the urban region, in the late-PCV13 period we found ST445, a previously undetected clone, that might be emerging.

Serotype 23A was among the most prevalent non-vaccine types found in the rural region, in the late-PCV13 period (28). In our study, it was always associated with CC439, especially ST42, in both regions.

Serotype 23B was one of the most frequently found non-vaccine types, but this increase in frequency was mainly due to its prevalence in the urban region (28). CC439 was the major clone associated with this serotype, and it was the most frequent of all clonal lineages in our study, accounting for 4.9% of all isolates. CC439, associated with serotypes 23B and 23A, showed clonal expansion among isolates causing adult IPD in Portugal, before widespread use of PCV13 (57). Interestingly, we found opposite trends in terms of clonal diversity when comparing both regions: in the urban region clonal diversity increased over time, in the rural region it decreased.

Serotype 24 has been increasing its frequency in carriage, after PCV13 introduction, mainly in the urban region (28). In our study we found differences between the urban and the rural regions, regarding the clonal lineages responsible for serotype 24 clonal distribution. In the urban region, we detected the multiresistant ST230 (also known as PMEN clone Denmark[14]-230) only in the pre-PCV13 period. Serotype 24 was between the most frequent serotypes related with IPD of children younger than 2 years of age, in Spain in the PCV13 period, and this was mainly due to the multiresistant ST230 (58). For the following periods, ST162, a SLV of the ST156 associated with the vaccine-type serotype 14, was the most frequent clone. In the rural region, the most frequent clonal lineage was the susceptible ST72, however, in the late-PCV13 period, ST162 emerged.

Serotype 34 was not associated with a single dominant clone, instead different clones were detected overtime and over regions. In the urban region, in the pre-PCV13 period, the singleton ST478 was dominant. In the following period CC1046 emerged and in the late-PCV13 period we found ST1046 and ST2001, that belongs to a different clonal complex, CC2001. In contrast, in the rural region, we found ST1046 as the major clone in the pre-PCV13 period, CC2001 as the responsible for this serotype persistence in the early-PCV13 period and, in the late-PCV13 period we detected ST8967 that belongs to CC1046, together with the previously undetected ST4083.

Serotype 35B was associated with one major clone, ST198. In the rural region, this ST was present in all studied periods and it seemed to be the responsible for this serotype maintenance in the population. In the urban region, this ST was showed importance in the pre-PCV13 and early-PCV13 periods, but in the late-PCV13 period two new, previously undetected clones appeared, ST2690 and ST452. We also detected a clone that was non-susceptible to penicillin (ST558) in this region, however it did not persist through the three periods. In a study conducted in Canada, between 2011 and 2013, using IPD samples, serotype 35B was one of the most

frequently found non-vaccine types, and this expansion was mainly associated with ST558 and, consequently, piliated isolates (68).

Serotype 35F was very frequent in the late-PCV13 period when analyzing the rural region (28). Besides the clonal diversity found in this serotype, in the both regions ST1368 was maintained between the pre-PCV13 and the early-PCV13 periods. However, in the late-PCV13 period, CC446 emerged: in the urban region with ST1635, and in the rural region with ST446, ST1635 and the new ST13407.

Nontypeables were one of the major serotypes when talking about antimicrobial resistance. Besides the appearance of other clones along time, the multiresistant CC344 was the major clone detected in the urban region. NTs have been associated with ST344 (also known as PMEN clone Norway^{NT}-344) since early 1997, in day-care centers in Portugal (50). After PCV7 introduction, this clone, associated with NTs, showed an increase in prevalence, in the same study population, especially among drug resistant isolates (19). For the rural region, interestingly, we did not detect any clone belonging to this complex in the late-PCV13 period. Instead, we found the susceptible ST448, that had already been detected in the pre-PCV13 period, and the previously undetected ST393.

4.4 Capsular switch events

For the majority of our results we found a correlation between ST and serotype, however we detected some cases in which one ST was associated with more than one serotype, suggesting a capsular switch event. This was the case for ST30, ST63, ST66, ST191, ST193, ST199, ST393, ST439, ST447, ST1877, ST1201, ST13423 and ST13422. We investigated whether we had evidence for "novel" capsular switch events and if the vaccine could have triggered/selected such events.

In some cases (7 out of 13) the associations that we detected were already described.

In our study, we found one sample from serotype 23F belonging to ST63, normally associated with the non-vaccine serotype 15A. This sample was present in the pre-PCV13 period, in the urban region. After searching the MLST database, we only found 3 samples from serotype 23F and they were submitted as being found between 2015 and 2016 (66).

We found one isolate in the rural region, in the pre-PCV13 period, from serotype 7A belonging to ST191 that is the major clone related with serotype 7F distribution. However this is already described in the MLST database, with two ST191 samples associated with serotype 7A (66). Besides that, this association was also described after sequencing analysis of the only sample from 7A recovered from the database, in 1937 (54).

ST193 is described as being a very diverse clonal lineage, and after searching in the MLST database we detected the association between this ST and 9 serotypes. In our study we found three serotypes associated with ST193: 21, 15B/C and 19A. All these serotypes were among the already described for this clone. The isolates were equally distributed along time, independently of the serotype.

ST199 has been known for being frequently associated with different serotypes: 19A, 15B/C, 3, 23F, NT and 19F (66). This genotype has been related with the increase in prevalence of serotype 19A after PCV7 introduction. However, before PCV13 introduction, this genotype was associated with the vaccine serotype 4 (56). In our collection, this ST was only found in the urban region and it was associated with serotype 19A in the early-PCV13 period (n=1) and with serotype 15B/C in both the pre-PCV13 (n=1) and late-PCV13 periods (n=1). This associations had already been in invasive disease isolates (69). Interestingly, in the same study, the authors found that this event occurred before PCV7 introduction. This supported the hypothesis that capsular switch occurs frequently and randomly and the more fitted lineages are selected; the vaccine may have an effect due to the increase of non-vaccine types, increasing the possibility of recombination events among these serotypes (69). It is not surprising that after PCV13 introduction it has been described that, in colonization, ST199 associated with 19A (targeted by PCV13) has become less abundant and, in contrast, ST199 associated with serotype 15B/C has been prevailing through time (64).

We found ST393 to be related with three serotypes: 38, 25A and NT. These serotypes were previously described in the MLST database. However it is important to highlight that there were only two NT samples described (66). In our collection we found two NT isolates in the late-PCV13 period in both regions. This clonal lineage was only detected in the late-PCV13 period associated with serotypes 25A and NT.

ST439 was mainly associated with serotype 23B, although we detected 2 isolates from the urban region and the pre-PCV13 period that belonged to serotype 23A. This was already described in the MLST database, as there were 3 samples from serotype 23A associated with this ST (66). Besides that, in one study, when analyzing the genomes of 84 pneumococcal isolates belonging to serotype 23A, from four countries between 1945 and 2014, the authors found that all samples where from CC439 (54).

ST1201 is mainly found among samples of serotypes 19A and 19F, however there are samples from serotype 7C described (66). In our study we found three samples from serotype 7B/C in the late-PCV13 period, in both regions, belonging to ST1201. This can be a case of capsular switching as a vaccine escape mechanism once we have a clone typically associated

with a vaccine serotype, associated with a non-vaccine capsule. This association had already been described in IPD samples from adults, in Portugal, before PCV13 introduction, suggesting a selection of the non-vaccine variant in colonization upon introduction of PCV13 (57).

We also found some cases that are possible capsular switching events, never described.

Regarding ST30 we found one isolate of serotype 12F when this ST is usually associated with serotype 16F. After looking into the MLST database we did not found any record correlating this clonal lineage to this serotype. This was unlikely due to PCVs use as both 16F and 12F are non-vaccine serotypes.

ST66 is described as a very diverse clonal lineage, and in the MLST database it is associated with ten serotypes. In our study we found this ST to be associated with three non-vaccine serotypes: 9N, 9L and 37. The sample of serotype 37 was found in the pre-PCV13 period, in the rural region and this appears to be a novel association.

ST447 is usually associated with serotype 37, but in the MLST database we can find this ST associated with other serotypes (66). This is the case for serotypes 3, 9N, 9L and NT. Interestingly, we found one isolate in the pre-PCV13 period, in the rural region, that belonged to the vaccine serotype 19A. In this case we would have a non-vaccine clone with a PCV13 targeted capsule that disappeared after PCV13 introduction.

Regarding ST1877 we only found it associated with serotypes 21 and 23F (66). However, in our collection, we detected one isolate from serotype 19A and another from serotype 6C, belonging to this clonal lineage. Both serotypes were detected in the pre-PCV13 period, in the urban region. This could be a case of capsular switching in the direction of non-vaccine type to vaccine-type. The fact that they were only present in the pre-PCV13 period can be due to the effect of PCV13 in vaccine serotypes.

Two of our new allelic profiles showed association with more than one serotype. ST13422 was found in isolates from serotype 18A and 15B/C. ST13423 was found among isolates belonging to serotype 12A and 21. For both cases we have one serotype that is uncommon in carriage among our studied population (18A and 12A), and another that is very frequent (15B/C and 21). As these capsular variants were only detected in the pre-PCV13 and early-PCV13 periods, these novel associations may not confer a fitness advantage.

5 Conclusions

The major aim of this study was to evaluate the impact of PCV13 introduction in the population structure of colonizing *Streptococcus pneumoniae*. We used MLST to analyze around 700 nasopharyngeal samples of children attending day-care centers in three vaccine coverage periods and in two regions of Portugal.

Several studies have been performed worldwide regarding vaccine impact on serotype distribution. It is known that after vaccine introduction serotype replacement occurs as the decrease in the serotypes covered by the vaccine leaves an empty niche that is occupied by non-vaccine types. Given that, we wanted to go beyond serotype changes and understand if the vaccine also impacts on the pneumococcal genotypes in circulation. Regarding the vaccine types that are still in circulation we wanted to know if their persistence was due to a major clonal lineage that was already in circulation before PCV13 introduction. For the non-vaccine types our focus was to understand if their increase in the population was due to the expansion of pre-existing clones or if it was due to the emergence of new clones. Besides that, we also wanted to analyze the contribution of antimicrobial resistance and capsular switching events to the patterns observed.

After analyzing the results, we can conclude that the few vaccine-types that are still in circulation are associated with clonal lineages that were already in circulation before PCV13 introduction and for some reason have been able to escape vaccine effect. One possible reason is their antimicrobial resistant profile, that may confer an advantage given the selective pressure exerted by antimicrobial consumption.

Regarding non-vaccine types we found that clonal expansion of already existing clones occurs, and, besides that, there are also novel clones emerging. In terms of antimicrobial resistance, what was found for the vaccine-types may not be true in this case once the new and previously undetected clones found associated with non-vaccine types frequently showed susceptibility to most antimicrobial agents. In this case, we can hypothesize that other factors, like increased fitness, transmission and better evading mechanisms, were responsible for the prevalence of these clones (19). Moreover, it is important to highlight that the non-vaccine serotypes are not being directly targeted by the vaccine. Of major importance is the emergence of serotype 8 after PCV13 introduction, associated with ST53, as this is known for being a highly invasive serotype already present in adult IPD in Portugal.

We also found some differences between clonal diversity, as some serotypes seemed to be very clonal (serotypes 1, 7F and 31 for example) and others were very diverse (serotypes 19A

and 15B/C, for example). In general, we found a higher diversity of lineages among non-vaccine serotypes, with the exception of serotype 19A. This may be directly related with their frequency in carriage: the clones that are infrequent in carriage tend to be more clonal once clonal diversification occurs mainly through recombination with other genotypes carried at the same time.

Novel capsular switching events were rare in our tested population and we could not correlate them with vaccine use.

Besides the fact that in Portugal we have high vaccination coverage even before the inclusion of PCV13 in the National Immunization Program, it is of high importance that its impact on the pneumococcal population continues to be monitored. Moreover, although the capsular type is a key factor in pneumococcal virulence, we could show that the genotype also plays a role on pneumococcal evolution. Thus, this feature should continue to be monitored, maybe using more advanced strategies and techniques like whole genome sequencing.

6 References

1. **Watson D, Musher D, Jacobson J, Verhoef J.** 1993. A brief history of the pneumococcus in biomedical research: a panoply of scientific discovery. *Clin Infect Dis* **17:**913-924.
2. **Hamborsky J, Kroger A, Wolfe C.** 2015. *Epidemiology and prevention of vaccine-preventable diseases.*, 13 ed, CDC.
3. **Brown J, Hammerschmidt S, Orihuela C.** 2015. *Streptococcus pneumoniae*: molecular mechanisms of host-pathogen interactions. 1st ed. *Elsevier*.
4. **Mehr S, Wood N.** 2012. *Streptococcus pneumoniae* – a review of carriage, infection, serotype replacement and vaccination. *Paediatr Respir Rev* **13:**258-264.
5. **Kadioglu A, Weiser JN, Paton JC, Andrew PW.** 2008. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. *Nat Rev Microbiol* **6:**288-301.
6. **Feldman C, Anderson R.** 2016. Epidemiology, virulence factors and management of the pneumococcus. *F1000 Research* **5:**2320.
7. **Bogaert D, de Groot R, Hermans P.** 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. *Lancet Infect Dis* **4:**144-154.
8. **Simões AS, Valente C, de Lencastre H, Sá-Leão R.** 2011. Rapid identification of noncapsulated *Streptococcus pneumoniae* in nasopharyngeal samples allowing detection of co-colonization and reevaluation of prevalence. *Diagn Microbiol Infect Dis* **71:**208-216.
9. **WHO.** 2012. Weekly epidemiological record. **14:**129-144.
10. **Nunes S, Félix S, Valente C, Simões AS, Tavares DA, Almeida ST, Paulo AC, Brito-Avô A, de Lencastre H, Sá-Leão R.** 2016. The impact of private use of PCV7 in 2009 and 2010 on serotypes and antimicrobial resistance of *Streptococcus pneumoniae* carried by young children in Portugal: Comparison with data obtained since 1996 generating a 15-year study prior to PCV13 introduction. *Vaccine* **34:**1648-1656.
11. **Hausdorff W, Bryant J, Paradiso P, Siber G.** 2000. Which pneumococcal serogroups cause the most invasive disease: implications for conjugate vaccine formulation and use, part I. *Clin Infect Dis* **30:**100-121.
12. **Valente C, Hinds J, Gould KA, Pinto FR, de Lencastre H, Sá-Leão R.** 2016. Impact of the 13-valent pneumococcal conjugate vaccine on *Streptococcus pneumoniae* multiple serotype carriage. *Vaccine* **34:**4072-4078.
13. **Almeida S, Nunes S, Santos Paulo AC, Valadares I, Martins S, Brei F, Brito-Avô A, Morais A, de Lencastre H, Sá-Leão R.** 2014. Low prevalence of pneumococcal carriage and high serotype and genotype diversity among adults over 60 years of age living in Portugal. *PLoS ONE* **9:**e90974.
14. **Sá-Leão R, Nunes S, Brito-Avô A, Alves CR, Carriço JA, Saldanha J, Almeida JS, Sanches IS, de Lencastre H.** 2008. High rates of transmission of and colonization by *Streptococcus pneumoniae* and *Haemophilus influenzae* within a day care center revealed in a longitudinal study. *J Clin Microbiol* **46:**225-234.
15. **Richter SS, Musher DM.** 2016. The ongoing genetic adaptation of *Streptococcus pneumoniae*. *J Clin Microbiol* **55:**681-685.
16. **CDC.** Pneumococcal disease. https://www.cdc.gov/pneumococcal/index.html. Accessed 05.01.2018.
17. **de Lencastre H, Kristinsson KG, Brito-Avô A, Sanches IS, Sá-Leão R, Saldanha J, Sigvaldadottir E, Karlsson S, Oliveira D, Mato R, Aires de Sousa M, Tomasz A.** 1999. Carriage of respiratory tract pathogens and molecular epidemiology of

Streptococcus pneumoniae colonization in healthy children attending day care centers in Lisbon, Portugal. *Microb Drug Resist* **5:**19-29.
18. **Cohen R, Levy C, Bonnet E, Grondin S, Desvignes V, Lecuyer A, Fritzell B, Varon E.** 2010. Dynamic of pneumococcal nasopharyngeal carriage in children with acute otitis media following PCV7 introduction in France. *Vaccine* **28:**6114-6121.
19. **Simões AS, Pereira L, Nunes S, Brito-Avô A, de Lencastre H, Sá-Leão R.** 2011. Clonal evolution leading to maintenance of antibiotic resistance rates among colonizing pneumococci in the PCV7 era in Portugal. *J Clin Microbiol* **49:**2810-2817.
20. **McGee L, McDougal L, Zhou J, Spratt B, Tenover F, George R, Hakenbeck R, Hryniewicz W, Lefévre J, Tomasz A, Klugman K.** 2001. Nomenclature of major antimicrobial-resistant clones of *Streptococcus pneumoniae* defined by the pneumococcal molecular epidemiology network. *J Clin Microbiol* **39:**2565-2571.
21. **PMEN.** 2014. Pneumococcal Molecular Epidemiology Network. http://www.pneumogen.net/pmen/. Accessed 01.03.2018.
22. **Kim YK, LaFon D, Nahm MH.** 2016. Indirect effects of pneumococcal conjugate vaccines in national immunization programs for children on adult pneumococcal disease. *Infect Chemother* **48:**257-266.
23. **Givon-Lavi N, Fraser D, Dagan R.** 2003. Vaccination of day-care center attendees reduces carriage of *Streptococcus pneumoniae* among their younger siblings. *Pediatr Infect Dis J* **22:**524-531.
24. **Weinberger DM, Malley R, Lipsitch M.** 2011. Serotype replacement in disease after pneumococcal vaccination. *Lancet* **378:**1962-1973.
25. **Brueggemann AB, Griffiths DT, Meats E, Timothy P, Crook DW, Spratt BG.** 2003. Clonal relationships between invasive and carriage *Streptococcus pneumoniae* and serotype- and clone specific differences in invasive disease potential. *J Infect Dis* **187:**1424-1432.
26. **Huang S, Hinrichsen V, Stevenson A, Rifas- Shiman S, Kleinman K, Pelton S, Lipsitch M, Hanage W, Lee G, Finkelstein J.** 2009. Continued impact of pneumococcal conjugate vaccine on carriage in young children. *Pediatrics* **124:**e1-e11.
27. **Johnson HL, Deloria-Knoll M, Levine OS, Stoszek SK, Hance LF, Reithinger R, Muenz LR, O'Brien KL.** 2010. Systematic evaluation of serotypes causing invasive pneumococcal disease among children under five: the pneumococcal global serotype project. *PLoS Med* **7:**e1000348.
28. **Félix S, Handem S, Nunes S, C. Paulo A, Valente C, Candeias C, S. Simões A, T. Almeida S, A. Tavares D, Brito-Avô A, de Lencastre H, Sá-Leão R.** 2018. Private use of the 13-valent pneumococcal conjugate vaccine (PCV13) has impacted on pneumococcal carriage among Portuguese children. Abstr: ISPPD-0434. *ISPPD-11 - 11th International Symposium on Pneumococci and Pneumococcal Disease.* Melbourne, Australia. April 15th - 19th, 2018.
29. **WHO.** 2016. Progresses and challenges with achieving universal immunization coverage: 2015 vaccine coverage.
30. **VIEW-hub.** 2017. Global vaccine introduction and implementation. https://www.jhsph.edu/research/centers-and-institutes/ivac/view-hub/ivac-view-hub-report-2017-march.pdf. Accessed 05.01.2018.
31. **Sá-Leão R, Pinto F, Aguiar S, Nunes S, Carriço JA, Frazão N, Sousa NG, Cristino JM, de Lencastre H, Ramirez M.** 2011. 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* **49:**1369-1375.

32. **Enright M, Spratt B.** 1998. A multilocus sequence typing scheme for *Streptococcus pneumoniae* : identification of clones associated with serious invasive disease. *Microbiology* **144:**3049-3060.
33. **de Lencastre H, Sanches IS, Brito-Avô A, Sá-Leão R, Saldanha J, Kristinsson KG, Tomasz A.** 1999. Carriage and antibiotic resistance of respiratory pathogens and molecular epidemiology of antibiotic-resistant *Streptococcus pneumoniae* colonizing children in day-care centers in Lisbon: the Portuguese day-care center initiative. *Clin Microbiol Infect* **5:**4S55-54S63.
34. **Sá-Leão R, Nunes S, Brito-Avô A, Frazão N, Simões AS, Crisóstomo MI, Paulo ACS, Saldanha J, Sanches IS, de Lencastre H.** 2009. Changes in pneumococcal serotypes and antibiotypes carried by vaccinated and unvaccinated day-care centre attendees in Portugal, a country with widespread use of the seven-valent pneumococcal conjugate vaccine. *Clin Microbiol Infect* **15:**1002-1007.
35. **Rodrigues F, Nunes S, Sá-Leão R, Gonçalves G, Lemos L, de Lencastre H.** 2009. *Streptococcus pneumoniae* nasopharyngeal carriage in children attending day-care centers in the central region of Portugal, in the era of 7-Valent pneumococcal conjugate vaccine. *Microb Drug Resist* **15:**269-277.
36. **Mato R, Sanches IS, Simas C, Nunes S, Carriço JA, Sousa NG, Frazão N, Saldanha J, Brito-Avô A, Almeida JS, de Lencastre H.** 2005. Natural history of drug-resistant clones of *Streptococcus pneumoniae* colonizing healthy children in Portugal. *Microb Drug Resist* **11:**309-322.
37. **Sá-Leão R, Tomasz A, de Lencastre H.** 2001. Multilocus sequence typing of *Streptococcus pneumoniae* clones with unusual drug resistance patterns: genetic backgrounds and relatedness to other epidemic clones. *J Infect Dis* **184:**1206-1210.
38. **Satzke C, Turner P, Julkunen AV, Adrian PV, Antonio M, Hare KM, Restrepo AMH, Leach AJ, Klugman KP, Porter BD, Sá-Leão R, Scott JA, Nohynek H, O'Brien KL.** 2014. Standard method for detecting upper respiratory carriage of *Streptococcus pneumoniae* : Updated recommendations from the World Health Organization Pneumococcal Carriage Working Group. *Vaccine* **32:**165-179.
39. **O'Brien KL, Bronsdon MA, Dagan R, Yagupsky P, Janco J, Elliott J, Whitney CG, Yang Y-H, Robinson L-GE, Schwartz B, Carlone GM.** 2001. Evaluation of a medium (STGG) for transport and optimal recovery of *Streptococcus pneumoniae* from nasopharyngeal secretions collected during field studies. *J Clin Microbiol* **39:**1021-1024.
40. **Murray PR.** 1979. Modification of the bile solubility test for rapid identification of *Streptococcus pneumoniae*. *J Clin Microbiol* **9:**290-291.
41. **Jauneikaite E, Tocheva A, Jefferies J, Gladstone R, Faust S, Christodoulides M, Hibberd M, Clarke S.** 2015. Current methods for capsular typing of *Streptococcus pneumoniae*. *J Microbiol Methods* **113:**41-49.
42. **Rayner RE, Savill J, Hafner LM, Huygens F.** 2015. Genotyping *Streptococcus pneumoniae*. *Future Microbiol* **10:**653-664.
43. **Pai R, Gertz RE, Beall B.** 2006. Sequential multiplex PCR approach for determining capsular serotypes of *Streptococcus pneumoniae* isolates. *J Clin Microbiol* **44:**124-131.
44. **Keller LE, Robinson DA, McDaniel LS.** 2016. Nonencapsulated *Streptococcus pneumoniae* : emergence and pathogenesis. *mBio* **7:**e01792-01715.
45. **Maiden MCJ, Bygraves JA, Feil E, Morelli G, Russell JE, Urwin R, Zhang Q, Zhou J, Zurth K, Caugant DA, Feavers IM, Achtman M, Spratt BG.** 1998. Multilocus sequence typing: A portable approach to the identification of clones within populations of pathogenic microorganisms. *Proc Natl Acad Sci USA* **95:**3140-3145.

46. **Francisco A, Vaz C, Monteiro P, Melo-Cristino J, Ramirez M, Carriço J.** 2012. PHYLOViZ: phylogenetic inference and data visualization for sequence based typing methods. *BMC Bioinformatics* **13**.
47. **Francisco A, Bugalho M, Ramirez M, Carriço J.** 2009. Global optimal eBURST analysis of multilocus typing data using a graphic matroid approach. *BMC Bioinformatics* **10**.
48. **ICL.** MLST: Multi Locus Sequence Typing. http://www.mlst.net. Accessed 17.01.2018.
49. **Nunes S, Sá-Leão R, Carriço J, Alves CR, Mato R, Brito-Avô A, Saldanha J, Almeida JS, Sanches IS, de Lencastre H.** 2005. Trends in drug resistance, serotypes, and molecular types of *Streptococcus pneumoniae* colonizing preschool-age children attending day care centers in Lisbon, Portugal: a summary of 4 years of annual surveillance. *J Clin Microbiol* **43:**1285-1293.
50. **Sá-Leão R, Tomasz A, Sanches IS, Nunes S, Alves CR, Brito-Avô A, Saldanha J, Kristinsson KG, de Lencastre H.** 2000. Genetic diversity and clonal patterns among antibiotic-susceptible and -resistant *Streptococcus pneumoniae* colonizing children:day care centers as autonomous epidemiological units. *J Clin Microbiol* **38:**4137-4144.
51. **Sousa NG, Sá-Leão R, Crisóstomo MI, Simas C, Nunes S, Frazão N, Carriço JA, Mato R, Santos-Sanches I, Lencastre Hd.** 2005. Properties of novel international drug-resistant pneumococcal clones identified in day-care centers of Lisbon, Portugal. *J Clin Microbiol* **43:**4696-4703.
52. **Elberse K, Nunes S, Sá-Leão R, van der Heide H, Schouls L.** 2011. Multiple-locus variable number tandem repeat analysis for *Streptococcus pneumoniae*: comparison with PFGE and MLST. *PLoS ONE* **6:**e19668.
53. **Biomerieux.** Applied Maths: a Biomérieux company. http://www.applied-maths.com/applications/mlst. Accessed 09.07.2017.
54. **van Tonder A, Bray J, Quirk S, Haraldsson G, Jolley K, Maiden M, Hoffmann S, Bentley S, Haraldsson Á, Erlendsdóttir H, Kristinsson K, Brueggemann A.** 2016. Putatively novel serotypes and the potential for reduced vaccine effectiveness: capsular locus diversity revealed among 5405 pneumococcal genomes. *Microb Genom* **2:**000090.
55. **Diamantino-Miranda J, Aguiar S, Carriço J, Melo-Cristino J, Ramirez M.** 2017. Clonal and serotype dynamics of serogroup 6 isolates causing invasive pneumococcal disease in Portugal: 1999-2012. *PLoS One* **12:**e0170354.
56. **Henriques-Normark B, Blomberg C, Dagerhamn J, Bättig P, Normark S.** 2008. The rise and fall of bacterial clones: *Streptococcus pneumoniae*. *Nat Rev Microbiol* **6:**827-837.
57. **Horácio AN, Silva-Costa C, Diamantino-Miranda J, Lopes JP, Ramirez M, Melo-Cristino J.** 2016. Population structure of *Streptococcus pneumoniae* causing invasive disease in adults in Portugal before PCV13 availability for adults: 2008-2011. *PLoS One* **11:**e0153602.
58. **del Amo E, Esteva C, Hernandez-Bou S, Galles C, Navarro M, Sauca G, Diaz A, Gassiot P, Marti C, Larrosa N, Ciruela P, Jane M, Sá-Leão R, Muñoz-Almagro C.** 2016. Serotypes and clonal diversity of *Streptococcus pneumoniae* causing invasive disease in the era of PCV13 in Catalonia, Spain. *PLoS One* **11:**e0151125.
59. **Isaacman DJ, McIntosh ED, Reinert RR.** 2010. Burden of invasive pneumococcal disease and serotype distribution among *Streptococcus pneumoniae* isolates in young children in Europe: impact of the 7-valent pneumococcal conjugate vaccine and considerations for future conjugate vaccines. *Int J Infect Dis* **14:**e197-e209.

60. **Brueggemann AB, Spratt BG.** 2003. Geographic distribution and clonal diversity of *Streptococcus pneumoniae* serotype 1 isolates. *J Clin Microbiol* **41:**4966–4970.
61. **Aguiar SI, J. Brito M, Gonçalo-Marques J, Melo-Cristino J, Ramirez Mr.** 2010. Serotypes 1, 7F and 19A became the leading causes of pediatric invasive pneumococcal infections in Portugal after 7 years of heptavalent conjugate vaccine use. *Vaccine* **28:**5167-5173.
62. **Almeida ST, de Lencastre H, Sá-Leão R.** 2013. Epidemiology and population structure of serotypes 1, 5 and 7F carried by children in Portugal from 1996-2010 before introduction of the 10-valent and 13-valent pneumococcal conjugate vaccines. *PLoS One* **8:**e75442.
63. **Aguiar SI, Brito MJ, Horácio AN, Lopes JP, Ramirez M, Melo-Cristino J.** 2014. Decreasing incidence and changes in serotype distribution of invasive pneumococcal disease in persons aged under 18 years since introduction of 10-valent and 13-valent conjugate vaccines in Portugal, July 2008 to June 2012. *Euro Surveill* **19:**20750.
64. **Chang Q, Stevenson AE, Croucher NJ, Lee GM, Pelton SI, Lipsitch M, Finkelstein JA, Hanage WP.** 2015. Stability of the pneumococcal population structure in Massachusetts as PCV13 was introduced. *BMC Infect Dis* **15**.
65. **Aguiar SI, Pinto FR, Nunes S, Serrano I, Melo-Cristino J, Sá-Leão R, Ramirez M, de Lencastre Hn.** 2010. Denmark[14]-230 clone as an increasing cause of pneumococcal infection in Portugal within a background of diverse serotype 19A lineages. *J Clin Microbiol* **48:**101-108.
66. **UO.** PubMLST. http://pubmlst.org/. Accessed 02.06.2017.
67. **Nunes S, Valente C, Sá-Leão R, de Lencastre Hn.** 2009. Temporal trends and molecular epidemiology of recently described serotype 6C of *Streptococcus pneumoniae*. *J Clin Microbiol* **47:**472-474.
68. **Golden A, Adam H, Gilmour M, Baxter M, Martin I, Nichol K, Demczuk W, Hoban D, Zhanel G.** 2015. Assessment of multidrug resistance, clonality and virulence in non-PCV-13 *Streptococcus pneumoniae* serotypes in Canada, 2011-13. *J Antimicrob Chemother* **70:**1960-1964.
69. **Makarewicz O, Lucas M, Brandt C, Herrmann L, Albersmeier A, Rückert C, Blom J, Goesmann A, van der Linden M, Kalinowski J, Pletz MW.** 2017. Whole genome sequencing of 39 invasive *Streptococcus pneumoniae s*equence type 199 isolates revealed switches from serotype 19A to 15B. *PLoS One* **12:**e0169370.

7 Anexes

Annex 1: Primers used for MLST.

List of MLST primers, and corresponding annealing temperature used in PCR (32). The second pair of primers was used whenever the first pair did not yield a positive result.

Gene	Primer name	Primer sequence	T (°C)
aroE	*aroE*/M13F	TGTAAAACGACGGCCAGTcgtttagctgcagttgttgc	63.5
	aroE/M13R	CAGGAAACAGCTATGACCcccacactggtggcattaac	63.5
	aroE/M13F2	TGTAAAACGACGGCCAGTGCCTTTGAGGCGACAGC	48
	aroE/M13R2	CAGGAAACAGCTATGACCTGCAGTTCA(G/A)AAACAT(A/T)TTCTAA	48
ddl	*ddl*/M13F	TGTAAAACGACGGCCAGTttgccatggataaaatcacgac	53
	ddl/M13R	CAGGAAACAGCTATGACCcgcgcttgtcaaaactttcc	53
	ddl-NU/M13F	TGTAAAACGACGGCCAGTaccattgattgggataagmargttg	60
	ddl-ND/S/M13R	CAGGAAACAGCTATGACCcatattntccckaagragtgg	60
gdh	*gdh*/M13F	TGTAAAACGACGGCCAGTgtgctgaaaagattaaggtct	63.5
	gdh/M13R	CAGGAAACAGCTATGACCtgcttccagctttatagtcatg	63.5
	gdh-NU/M13F	TGTAAAACGACGGCCAGTctacaacttctttcgctcct	62
	gdh-ND/M13R	CAGGAAACAGCTATGACCtgccaagtccatttggcacc	62
gki	*gki*/M13F	TGTAAAACGACGGCCAGTggcattggaatgggatcacc	63.5
	gki/M13R	CAGGAAACAGCTATGACCtctcccgcagctgacac	63.5
	gki-NU/M13F	TGTAAAACGACGGCCAGTactacacaatgaaaaaggagaatatg	60
	gki-ND/M13R	CAGGAAACAGCTATGACCcttgtgggaaactattttcatcg	60
recP	*recP*/M13F	TGTAAAACGACGGCCAGTgccaactcaggtcatccagg	63.5
	recP/M13R	CAGGAAACAGCTATGACCttcgatagcagcatggatgg	63.5
	*recP*2/M13R	CAGGAAACAGCTATGACCgcttccaagtctgttccattttc	63.5
spi	*spi*/M13F	TGTAAAACGACGGCCAGTcgcttagaaaggtaagttatg	53
	spi/M13R	CAGGAAACAGCTATGACCaggctgagattggtgattctc	53
	spi-NU/M13F	TGTAAAACGACGGCCAGTacgcttagaaaggtaagttatg	58
	spi-ND2/M13R	CAGGAAACAGCTATGACCaagaggctgagattggtgattc	58
xpt	*xpt*/M13F	TGTAAAACGACGGCCAGTggaggtcttatgaaattattag	53
	xpt/M13R	CAGGAAACAGCTATGACCagatctgcctccttaaatac	53
	*xpt*2/M13F	TGTAAAACGACGGCCAGTccactacaacgggaaatatttga	53
	*xpt*2/M13R	CAGGAAACAGCTATGACCaggatagatcctgagtacatg	53