

**UNIVERSIDADE FEDERAL FLUMINENSE  
INSTITUTO BIOMÉDICO  
GRADUAÇÃO EM BIOMEDICINA**

**PABLO ALFRADIQUE ETCHARTE**

**RESISTOMA DE *Streptococcus pneumoniae* DA LINHAGEM GENÉTICA 6C/ST386  
EMERGENTE APÓS A INTRODUÇÃO DA VACINA PNEUMOCÓCICA  
CONJUGADA 10-VALENTE NO BRASIL**

**NITERÓI  
2023**

PABLO ALFRADIQUE ETCHARTE

**RESISTOMA DE *Streptococcus pneumoniae* DA LINHAGEM GENÉTICA 6C/ST386  
EMERGENTE APÓS A INTRODUÇÃO DA VACINA PNEUMOCÓCICA  
CONJUGADA 10-VALENTE NO BRASIL**

Trabalho de conclusão de curso apresentado ao curso de graduação em Biomedicina do Instituto Biomédico da Universidade Federal Fluminense, como requisito parcial à obtenção do título de Bacharel em Biomedicina com habilitação em Análises Clínicas.

Orientador: Prof. Dr. Felipe Piedade Gonçalves Neves

Coorientadora: Prof<sup>a</sup>. Dr<sup>a</sup>. Aline Rosa Vianna de Souza

NITERÓI  
2023

Ficha catalográfica automática - SDC/BIB  
Gerada com informações fornecidas pelo autor

E83r Etcharte, Pablo Alfradique  
RESISTOMA DE Streptococcus pneumoniae DA LINHAGEM GENÉTICA  
6C/ST386 EMERGENTE APÓS A INTRODUÇÃO DA VACINA PNEUMOCÓCICA  
CONJUGADA 10-VALENTE NO BRASIL / Pablo Alfradique Etcharte. -  
2023.  
41 f.: il.

Orientador: Felipe Piedade Gonçalves Neves.  
Coorientador: Aline Rosa Vianna de Souza.  
Trabalho de Conclusão de Curso (graduação)-Universidade  
Federal Fluminense, Instituto Biomédico, Niterói, 2023.

1. Streptococcus pneumoniae. 2. Resistoma. 3. Sorotipo  
emergente. 4. Vacina pneumocócica conjugada. 5. Produção  
intelectual. I. Neves, Felipe Piedade Gonçalves, orientador.  
II. Souza, Aline Rosa Vianna de, coorientador. III.  
Universidade Federal Fluminense. Instituto Biomédico. IV.  
Título.

CDD - XXX

PABLO ALFRADIQUE ETCHARTE

**RESISTOMA DE *Streptococcus pneumoniae* DA LINHAGEM GENÉTICA 6C/ST386  
EMERGENTE APÓS A INTRODUÇÃO DA VACINA PNEUMOCÓCICA  
CONJUGADA 10-VALENTE NO BRASIL**

Trabalho de conclusão de curso apresentado ao curso de graduação em Biomedicina do Instituto Biomédico da Universidade Federal Fluminense, como requisito parcial à obtenção do título de Bacharel em Biomedicina com habilitação em Análises Clínicas.

BANCA EXAMINADORA

---

MSc Amanda Beiral da Silva (UFF)

---

Prof<sup>a</sup>. MSc. Anna Luiza Bauer Canellas (UFRJ)

---

Prof. Dr. Felipe Piedade Gonçalves Neves (UFF)

---

Prof<sup>a</sup>. Dr<sup>a</sup>. Aline Rosa Vianna de Souza (UFF)

---

Prof. Dr. Bruno De Araújo Penna – Suplente (UFF)

## AGRADECIMENTOS

Ao meu orientador, Felipe Piedade, expresso profunda gratidão por aceitar me orientar neste projeto, mesmo diante do tempo limitado disponível para realização do mesmo. Sou extremamente grato pela oportunidade e pelos valiosos ensinamentos ao longo deste período.

A Dra. Aline, por aceitar me coorientar e por todo suporte durante o projeto e todos os ensinamentos. Sendo de grande importância para o desenvolvimento do projeto e no alcance dos resultados obtidos.

Aos meus pais Valeiro e Cristina, que sempre incentivaram e me apoiaram em tudo que quis fazer, meu pai que sempre foi meu maior exemplo e me ensinou a importância dos estudos; aos meus tios Adriana e Pierre e minha avó Ana Maria, que ajudaram na minha criação com tanto carinho.

A minha namorada Isabel Vignoli, pelo apoio incondicional, companhia e toda compreensão ao longo desses dois anos.

Aos meus amigos que me acompanham desde a época da escola; Esther, Jean, Giulia, e Jesus, que apesar da distância, estamos sempre nos apoiando e celebrando as conquistas uns dos outros; Aos amigos que a UFF me deu, Ester, Maria Theresa, Yasmin e Pedro que tornaram meus dias na UFF melhores.

Um agradecimento especial a Ester, minha dupla e irmã; estamos junto nessa jornada desde o início do curso, sempre apoiando e incentivando um ao outro. Minha companhia durante os surtos da pandemia, do IC e ao longo dessa loucura que foi fazer duas monografias. Esses 6 anos teriam sido muito mais difíceis sem sua companhia, foi um enorme prazer aturar essa criatura esses anos, uma amizade que pretendo levar para toda vida.

Aos meus professores da época do Colégio Santa Mônica, Maritza, Roberto e Charles, que me inspiraram ao sonho de seguir no mundo das ciências e da pesquisa durante o ensino fundamental e médio; e aos professores do curso de graduação em Biomedicina da Universidade Federal Fluminense, por serem exemplos de professores e pesquisadores.

*“Não é o conhecimento, mas o ato de aprender, não a posse, mas o ato de chegar lá, que concede a maior satisfação.”*

***Carl Friedrich Gauss***

## RESUMO

*Streptococcus pneumoniae* é um dos principais patógenos responsáveis por doenças respiratórias graves em todo o mundo, afetando indivíduos de todas as idades. Os principais grupos de risco para doença pneumocócica invasiva (DPI) são crianças menores de 2 anos, idosos, e pacientes imunocomprometidos, sendo uma das principais causas de morbidade e mortalidade nesses grupos vulneráveis. Para minimizar o impacto das doenças pneumocócicas no sistema de saúde, diferentes formulações de vacinas pneumocócicas conjugadas foram introduzidas em muitos países desde 2000. Desde a implementação da vacina pneumocócica conjugada 10-valente (VPC10) no Programa Nacional de Imunização infantil do SUS a cobertura da vacinação atingiu mais de 90% na população elegível. Apesar da ampla cobertura vacinal no Brasil de 2010 a 2019, observou-se um total de mais de 600.000 internações anuais devido à pneumonia, enquanto a média anual de casos de meningite pneumocócica no mesmo intervalo foi de aproximadamente 1.050. A implementação da vacina VPC10 tem gerado mudanças dinâmicas na população pneumocócica, com uma notável redução nos sorotipos vacinais e um aumento correspondente nas DPIs causadas por sorotipos não vacinais, fenômeno conhecido como substituição de sorotipos. No presente estudo, foram empregados dados de sequenciamento do genoma completo de amostras multirresistentes provenientes de uma linhagem genética emergente em Niterói após a introdução da VPC10, visando determinar o seu resistoma a partir de análises utilizando diferentes ferramentas de bioinformática. A partir das análises realizadas classificamos as amostras analisadas como pertencentes ao sorotipo 6C, ao GPSC 47 e complexo clonal 386. Além disso, caracterizamos as *pbps* presentes nessas amostras, apresentando o perfil 2:53:77, exibindo resistência à penicilina (PNSP), cefotaxima e ceftriaxona. Por meio das análises *in silico*, também conseguimos prever o fenótipo de resistência em relação a vários antimicrobianos não beta-lactâmicos, identificando os genes *erm(B)* e *tet(M)*, associados à resistência à clindamicina e eritromicina, e tetraciclina, respectivamente. Adicionalmente realizamos uma análise evolutiva na qual as amostras apresentaram uma estreita relação filogenética, embora três amostras tenham ficado mais distantes das demais. Este estudo destaca a complexidade na interação entre vacinação, evolução de sorotipos e resistência. Ademais, demonstra a viabilidade do uso de ferramentas *in silico* simples e de fácil aplicação para a previsão da resistência a antimicrobianos a partir de dados de sequenciamento do genoma completo, destacando sua eficiência na análise de espécies bacterianas clinicamente relevantes.

**PALAVRAS CHAVE:** *Streptococcus pneumoniae*, substituição de sorotipos, Sorotipo 6C, resistência antimicrobiana, Análises *in silico*, VPC10, Resistoma.

## ABSTRACT

*Streptococcus pneumoniae* is one of the leading pathogens responsible for severe respiratory diseases worldwide, affecting individuals of all ages. The main risk groups for invasive pneumococcal disease (IPD) are children under 2 years old, the elderly, and immunocompromised patients, making it a major cause of morbidity and mortality in these vulnerable groups. To minimize the impact of pneumococcal diseases on the healthcare system, various formulations of pneumococcal conjugate vaccines (PCVs) have been introduced worldwide, since 2000. Vaccination coverage has reached over 90% in the eligible population since the implementation of 10-valent pneumococcal conjugate vaccine (10) in the national childhood immunization program in Brazil (PNI-SUS). Despite extensive vaccination coverage in Brazil from 2010 to 2019, there were more than 600,000 annual hospitalizations due to pneumonia, while the annual average of pneumococcal meningitis cases in the same period was approximately 1,050. The implementation of PCV10 has led to dynamic changes in the pneumococcal population, with a significant reduction in vaccine serotypes and a corresponding increase in IPDs caused by non-vaccine serotypes, a phenomenon known as serotype replacement. In this study, we employed whole-genome sequencing data from multidrug-resistant samples of an emerging genetic lineage in Niterói after the introduction of VPC10. Our aim was to determine the resistome of this lineage through analyses using various bioinformatics tools. Based on the analyses, we classified the samples as belonging to serotype 6C, GPSC 47, and clonal complex 386. Additionally, we characterized the penicillin-binding proteins present in these samples, showing *pbp* profile 2:53:77 and predicted non susceptibility to penicillin (PNSP), cefotaxime and ceftriaxone. Through *in silico* analyses, we predicted the resistance phenotype to various non-beta-lactam antimicrobials, also identifying genes *erm*(B) e *tet*(M), associated with resistance to clindamycin and erythromycin, and tetracycline, respectively. Furthermore, we conducted an evolutionary analysis, revealing a close phylogenetic relationship among the samples, although three samples stood out as more distant from the others. This study underscores the complexity of the interaction between vaccination, serotype evolution, and resistance. Moreover, it demonstrates the feasibility of using simple and easily applicable *in silico* tools for predicting antimicrobial resistance based on whole-genome sequencing data, highlighting their efficiency in the analysis of clinically relevant bacterial species

**KEY WORDS:** *Streptococcus pneumoniae*, serotype replacements, serotype 6C, antimicrobial resistance, *in silico* analyses, PCV10, resistome.



## SUMÁRIO

<b>RESUMO</b> .....	VII
<b>ABSTRACT</b> .....	VIII
<b>1. INTRODUÇÃO</b> .....	13
1.1 - Características gerais de <i>Streptococcus pneumoniae</i> .....	13
1.2 – Epidemiologia das infecções por <i>Streptococcus pneumoniae</i> no Brasil .....	14
1.2.1 – Sorotipos emergentes de <i>Streptococcus pneumoniae</i> .....	15
1.3 – Tratamento e resistência aos antimicrobianos .....	16
1.4 - Tipificação molecular de <i>S. pneumoniae</i> .....	17
1.5 – Sequenciamento genômico e ferramentas utilizadas na análise do resistoma bacteriano .....	19
1.6 - Justificativa .....	22
<b>2. OBJETIVOS</b> .....	24
2.1 Objetivo geral .....	24
2.2 Objetivo específico .....	24
<b>3. MATERIAL E MÉTODOS</b> .....	25
3.1 – Sequências do genoma completo de <i>Streptococcus pneumoniae</i> .....	25
3.2 – Análises do genoma bacteriano .....	25
3.3 – Análise filogenética das amostras de <i>Streptococcus pneumoniae</i> do sorotipo 6C pertencente à linhagem genética ST386 .....	26
3.3.1 – Alinhamento das sequências .....	26
3.3.2 – Análise do Alinhamento .....	26
3.3.3 – Construção da árvore filogenética .....	27
<b>4. RESULTADOS</b> .....	28
4.1 Análises do genoma bacteriano .....	28
4.1.1 Tipificação de Sequências Multilocus .....	28
4.1.2 Caracterização das PBPs .....	28
4.1.3 Caracterização do perfil de resistência a penicilina e outras drogas beta-lactâmicas .....	28
4.1.4 Análise <i>in silico</i> dos genes de resistência a outros antimicrobianos .....	29

4.2 Análise filogenética das amostras de <i>Streptococcus pneumoniae</i> do sorotipo 6C pertencente à linhagem genética ST386.....	29
4.2.1 Comparação de Identidade Média de Nucleotídeos .....	29
4.3 Construção da árvore filogenética com MEGA .....	31
<b>5. DISCUSSÃO</b> .....	32
<b>6. CONCLUSÃO</b> .....	36
<b>7. REFERÊNCIAS BIBLIOGRÁFICAS</b> .....	37

**LISTA DE FIGURAS**

<b>Figura 1.</b> Interface da Plataforma Pathogenwatch .....	21
<b>Figura 2.</b> Interface do software CLC Genomics Workbench 23 .....	21
<b>Figura 3.</b> Interface do software MEGA11 .....	22
<b>Figura 4.</b> Fluxograma do projeto.....	25
<b>Figura 5.</b> Heatmap apresentando as variações nas médias de identidade entre os genomas de 17 amostras de <i>Streptococcus pneumoniae</i> , sendo 14 isoladas de crianças em Niterói e três utilizadas como amostras de referência para fins comparativos. O esquema de cores refere-se à identidade média de nucleotídeos. As amostras exibidas em azul demonstram menor identidade nucleotídica, enquanto quanto maior a intensidade do tom vermelho, maior a identidade nucleotídica entre as sequências.....	30
<b>Figura 6.</b> Árvore filogenética construída com o software MEGA 11, utilizando o algoritmo Neighbour Joining, optando pelo modelo de substituição Kimura 2p e 1000 replicatas de bootstrap.....	31

**LISTA DE TABELAS**

<b>Tabela 1.</b> Padrão de designação de alelos a partir do MLST.....	<b>27</b>
<b>Tabela 2.</b> Predições das concentrações mínimas inibitórias (CMI) e fenótipo de resistência a antimicrobianos da classe dos beta-lactâmicos para as 14 amostras de <i>Streptococcus pneumoniae</i> .....	<b>28</b>

**LISTA DE ABREVIATURAS**

**BV-BRC** - Bacterial and Viral Bioinformatics Resource Center

**CC** - Complexo Clonal

**CDC** - Centros de Prevenção e Controle de Doenças dos Estados Unidos

**CLSI** - Clinical Laboratory Standard Institute

**CMI** – Concentração mínima inibitória

**DPI** - Doença pneumocócica invasiva

**GPS** - Global Pneumococcal Sequencing

**GPSC** - Global Pneumococcal Sequence Clusters

**MEGA** - Molecular Evolutionary Genetics Analysis

**MLST** - Tipagem de sequência multilocus (do inglês, Multilocus sequence typing)

**MR** - Bactéria multirresistentes

**NGS** – Sequenciamento de Nova Geração (do inglês, Next-Generation Sequencing)

**OMS** - Organização Mundial da Saúde

**PAC** - Pneumonia adquirida na comunidade

**PBP** - Proteínas de ligação à penicilina

**VPC** – Vacinas pneumocócicas conjugadas (do inglês, Pneumococcal conjugate vaccine)

**PNI** - Programa Nacional de Imunização

**PMEN** - Pneumococcal Molecular Epidemiology Network

**RAM** - resistência antimicrobiana

**SLV** - Variante de locus único (do inglês, single locus variant)

**ST** – Sequence Type

**SUS** – Sistema Único de Saúde

**TRS** - Trato respiratório superior

## 1. INTRODUÇÃO

### 1.1 - Características gerais de *Streptococcus pneumoniae*

A espécie *Streptococcus pneumoniae*, frequentemente referida como pneumococo, é uma bactéria encapsulada Gram-positiva que se apresenta como cocos tipicamente dispostas aos pares. São anaeróbios facultativos e crescem bem em meios enriquecidos com sangue, onde suas colônias promovem hemólise parcial ( $\alpha$ -hemolíticas), imprimindo uma coloração esverdeada ao meio (TRABULSI; ALTHERTUM, 2015).

*Streptococcus pneumoniae* coloniza a mucosa do trato respiratório superior (TRS) humano, mas é também um importante patógeno humano. Apresenta cerca de 100 sorotipos identificados, que se diferenciam com base na estrutura química e antigênica da sua cápsula polissacarídica. O pneumococo é uma das principais causas bacterianas de uma ampla gama de infecções, incluindo otite média aguda, sinusite, pneumonia adquirida na comunidade (PAC), bacteremia e meningite (WEISER et al., 2018).

O genoma de *S. pneumoniae* é constituído por um único cromossomo circular, totalizando 2.189.416 pares de bases. Este genoma abriga 2216 sequências de codificação, 56 genes de RNA de transferência e 12 genes de RNA ribossômico, organizados em quatro operons (Tabela 1). Dentre as sequências de codificação, 1616 (72,9%) têm uma função biológica conhecida prevista; 145 (6,5%) apresentam semelhança com proteínas hipotéticas em outros genomas, enquanto 455 (20,5%) não demonstram uma similaridade substancial com outras proteínas previstas (CAMILLI et al., 2011; SANTORO et al., 2019). Essa complexidade genômica destaca a diversidade funcional do *S. pneumoniae* e ressalta a importância de investigações mais aprofundadas para compreender melhor sua biologia e potencial patogênico.

O pneumococo é um dos principais patógenos responsáveis por doenças respiratórias graves em todo o mundo, afetando indivíduos de todas as idades, desde crianças pequenas até idosos. Os principais grupos de risco para DPI são crianças menores de 2 anos, idosos, e pacientes imunocomprometidos, sendo uma das principais causas de morbidade e mortalidade nesses grupos vulneráveis (REIS et al., 2008; KIM et al., 2017).

A colonização do TRS por pneumococos ocorre em todos os indivíduos pelo menos uma vez durante a infância. Crianças jovens são um reservatório importante para a transmissão em toda a comunidade, que em sua maior parte são portadores assintomáticos (REIS et al., 2008; HOWARD et al., 2023).

Entre 17%–80% das crianças e <10% dos adultos são portadores de *S. pneumoniae* e a colonização envolve uma relação comensal entre a bactéria e o hospedeiro. A disseminação local, aspiração ou disseminação para a corrente sanguínea podem resultar em doenças graves, como pneumonia, bacteremia e meningite (REIS et al., 2008; KIM et al., 2017).

## 1.2 – Epidemiologia das infecções por *Streptococcus pneumoniae* no Brasil

As infecções associadas a *S. pneumoniae* são um problema global de saúde pública, representando uma das principais causas de morbidade e mortalidade em crianças e idosos em todo o mundo. Segundo a Organização Mundial da Saúde (OMS), em 2019, a pneumonia foi responsável pela morte de 740.180 crianças com menos de 5 anos, representando 14% de todas as mortes nessa faixa etária (OMS, 2022).

No Brasil, de 2000 a 2008, a meningite pneumocócica representou 11% de todas as meningites bacterianas, com incidência de 9,5 casos/100 mil habitantes em pacientes menores de um ano de idade. Dados oficiais do Brasil demonstram que, no mesmo período, 7.129.291 internações foram associadas a complicações de PAC, em que *S. pneumoniae* foi o principal agente etiológico (CAIERÃO et al., 2014).

Para minimizar o impacto das doenças pneumocócicas no sistema de saúde, diferentes formulações de vacinas pneumocócicas conjugadas foram introduzidas em muitos países desde 2000. No Brasil, VPC10 está disponível para crianças menores de 2 anos desde 2010 como parte do Programa Nacional de Imunização (PNI) infantil, oferecido pelo Sistema Único de Saúde (SUS) financiado pelo Governo Federal Brasileiro (CAIERÃO et al., 2014).

A VPC10 é composta de dez sorotipos de *Streptococcus pneumoniae*, oito deles conjugados com a proteína D do *Haemophilus influenzae* tipo b não tipável, um com o toxóide tetânico e outro com toxóide diftérico. Sendo constituída dos sorotipos: 1, 4, 5, 6B, 7F, 9V, 14, 18C, 19F e 23F (POOLMAN et al., 2013). A VPC10 é indicada para crianças com idade inferior a 2 anos, sendo aplicada em um esquema de 2 doses com intervalo de 2 meses e 1 reforço aos 12 meses de idade. Crianças com idade superior a 2 anos recebem dose única, podendo receber uma dose de reforço em casos especiais, como a presença de comorbidades (CAIERÃO et al., 2014).

Desde a implementação da VPC10 no PNI infantil em todo o país, a cobertura da vacinação atingiu mais de 90% na população elegível (crianças com idade  $\geq$  2 meses), embora nos últimos anos a cobertura vacinal tenha caído consideravelmente (SILVA et al. 2023). O Brasil também conta com a VPC13-valente nas clínicas privadas e, desde 2019, na rede pública para indivíduos > 5 anos de idade considerados de alto risco para desenvolvimento de DPI, que

incluem pacientes oncológicos, transplantados ou vivendo com HIV/AIDS. A VPC13 abrange todos os sorotipos presentes na VPC10, além de incluir os sorotipos 3, 6A e 19A. (NEVES et al., 2018)

Apesar da ampla cobertura vacinal, a pneumonia, especialmente ocasionada por *S. pneumoniae*, permanece como uma das principais razões para hospitalizações de crianças com idade inferior a 6 anos. No Brasil no período de 2010 a 2019, observou-se um total de mais de 600.000 internações anuais devido à pneumonia, tendo o *S. pneumoniae* como principal agente etiológico, enquanto a média anual de casos de meningite pneumocócica no mesmo intervalo foi de aproximadamente 1.050 (FORTUNA et al., 2023)

### **1.2.1 – Sorotipos emergentes de *Streptococcus pneumoniae***

O uso de vacinas pneumocócicas conjugadas (VPCs) vem mudando continuamente a população pneumocócica na nasofaringe humana, que constitui o principal reservatório de *S. pneumoniae*. Estudos epidemiológicos realizados no Brasil demonstram os benefícios da vacinação com a VPC10, havendo uma redução significativa nos sorotipos vacinais. Contudo, embora a DPI associada aos sorotipos das vacinas tenha diminuído significativamente, estudos realizados após a introdução da VPC10 observaram um aumento de DPI causada por sorotipos não-vacinais (BRANDILEONE et al., 2018; MOTT et al., 2019).

Em um estudo conduzido pelo nosso grupo antes da introdução da VPC10 no Brasil, constatou-se uma prevalência de colonização de 49,2% em uma população de 242 crianças que frequentavam uma creche e um hospital pediátrico em Niterói/RJ. Os sorotipos de pneumococos prevalentes, apresentando-se em ordem decrescente de frequência, foram 6B, 19F, 6A, 14, 15C e 23F (NEVES et al., 2013).

Em 2014, quatro anos após a introdução das vacinas pneumocócicas conjugadas no Brasil, verificou-se que 94% das crianças avaliadas, com idade igual ou superior a 2 meses, haviam sido vacinadas com alguma VPC. Dentro desse grupo, 86% haviam sido imunizadas com a PCV10, enquanto 8% tinham recebido a VPC13. Durante este estudo, a prevalência de colonização por *S. pneumoniae* em 522 crianças foi de 22,6%. Além disso, observou-se a emergência de sorotipos que não estavam contemplados na VPC10, com destaque para os sorotipos 6C (15,3%), 15B/C (11%) e 11A/D (9,3%) (NEVES et al., 2017).

Em 2019, nosso grupo conduziu nova pesquisa com desenho de estudo semelhante ao de 2014. Neste estudo, foram recrutadas 430 crianças que frequentavam duas clínicas públicas e duas clínicas privadas na cidade de Niterói. Do total, 325 crianças eram elegíveis para a vacinação com a VPC10 (com idade superior a 2 meses). Treze crianças (4%) de 2 meses não



receberam nenhuma dose da VPC10, enquanto que 312 (96%) das 325 crianças elegíveis foram vacinadas com a VPC10. A prevalência da colonização pneumocócica foi de 17,4% (75/430). Os tipos capsulares mais frequentes associados à colonização na população investigada foram 6C/D (14,7%), 11A/D (13,3%), 23B (9,3%) e 15B/C (6,7%) (FORTUNA et al., 2023).

O sorotipo 6C foi inicialmente identificado em 2007, contudo estudos prévios sobre de *S. pneumoniae* revelaram a presença desse sorotipo ao longo de vários anos e em diversas localidades. As investigações indicaram a existência do sorotipo 6C em cepas desde o início da década de 1980. Antes da introdução da vacina conjugada, a prevalência do sorotipo 6C era baixa entre crianças saudáveis (PARK et al., 2007). Estudos moleculares iniciais sobre o sorotipo 6C relataram sua origem associada a eventos independentes de recombinação envolvendo amostras do sorotipo 6A. No entanto, estudos subsequentes sugeriram possíveis eventos recombinantes a partir de outros sorotipos, como o 6B, indicando múltiplas origens genéticas para o sorotipo 6C (JACOBS et al., 2009).

### **1.3 – Tratamento e resistência aos antimicrobianos**

Os tratamentos convencionais para doenças pneumocócicas envolvem o uso de uma ampla variedade de medicamentos, que compreendem antimicrobianos da classe dos betalactâmicos (como penicilinas e cefalosporinas de terceira geração), macrolídeos (tais como azitromicina, claritromicina e eritromicina) e fluoroquinolonas (como a levofloxacina). Estes medicamentos podem ser administrados individualmente ou em combinação terapêutica, especialmente quando se enfrenta uma resistência significativa a antimicrobianos em uma região específica ou quando se busca otimizar os resultados em comparação com o uso de apenas um medicamento (LEONARD et al., 2020).

Embora haja várias opções de medicamentos disponíveis para tratar doenças pneumocócicas, é importante destacar a crescente preocupação devido ao aumento da resistência aos antimicrobianos por parte das cepas de pneumococos. O aumento contínuo da resistência dos pneumococos aos beta-lactâmicos é um fenômeno impulsionado pelo amplo uso de antimicrobianos na comunidade, o que exerce uma constante pressão seletiva. Esse processo de resistência envolve principalmente dois mecanismos cruciais: a modificação dos genes que regulam os alvos dos medicamentos, frequentemente por meio de recombinação homóloga e a aquisição de genes de resistência especializados provenientes de elementos genéticos móveis (D'AETH et al., 2021).

As proteínas de ligação à penicilina (PBPs) são componentes naturais encontrados em muitas bactérias e estão associadas à membrana citoplasmática. Elas desempenham funções

essenciais e representam os principais alvos dos beta-lactâmicos. A resistência às penicilinas envolve a formação de genes em "mosaico", que resultam na modificação das PBPs e reduzem sua afinidade pela molécula antimicrobiana. Os genes mais frequentemente associados à resistência aos beta-lactâmicos em pneumococos são o *pbp1a*, *pbp2b* e *pbp2x* (REINERT et al., 2009).

O crescente nível de resistência à penicilina tem conduzido a um aumento na utilização de antimicrobianos de amplo espectro, como cefalosporinas, macrolídeos, fluoroquinolonas, entre outros. Esses medicamentos, que anteriormente eram principalmente reservados para pacientes alérgicos aos beta-lactâmicos ou para situações específicas de resistência ocasional, agora são empregados com maior frequência. Como resultado, tem sido observado um aumento nos mecanismos de resistência a outras categorias de antimicrobianos e o surgimento de casos de multirresistência (MDR). Estes últimos são caracterizados pela resistência a três ou mais classes de antimicrobianos e estão se tornando cada vez mais frequentes entre as cepas de pneumococos (MURRAY et al., 2019).

Em 2017, a Organização Mundial de Saúde (OMS) incluiu *S. pneumoniae* não susceptível à penicilina como um de 12 patógenos prioritários em termos de resistência a antimicrobianos, para os quais o desenvolvimento de novas opções terapêuticas é urgente (OMS, 2017). Adicionalmente, o Centro de Prevenção e Controle de Doenças dos Estados Unidos (CDC) incluiu *S. pneumoniae* resistentes aos antimicrobianos, de forma mais ampla, em sua lista de patógenos resistentes à antimicrobianos, na categoria de séria ameaça, devido ao alto custo do tratamento das doenças pneumocócicas invasivas, associado ao período prolongado de hospitalização nos casos de falhas terapêuticas (CDC, 2019).

#### **1.4 - Tipificação molecular de *S. pneumoniae***

Métodos de tipificação molecular são essenciais para compreender as mudanças na epidemiologia das infecções pneumocócicas e a dinâmica da evolução da população de pneumococos. Dentre os métodos mais amplamente utilizados, destaca-se a tipificação de sequências de múltiplos loci (MLST), empregada para caracterizar amostras de espécies bacterianas. Nesse método, são utilizadas sequências de fragmentos internos de sete genes conservados, com aproximadamente 450-500 pares de bases de cada gene, escolhidos por sua capacidade de serem sequenciados com precisão em ambas as fitas (NEVES e PINTO, 2022).

A abordagem MLST identifica diferentes sequências de cada gene conservado dentro de uma espécie bacteriana, designando-as como alelos distintos. No contexto de cada isolado, os alelos em cada um dos sete loci compõem o perfil alélico, conhecido como tipo de sequência

(ST). Esses STs podem ser organizados em complexos clonais (CC), os quais geralmente incluem clones que compartilham cinco ou seis alelos. Os clones internacionais, identificados por meio da MLST, referem-se a grupos de cepas bacterianas que apresentam STs idênticos ou altamente semelhantes em diversos loci genéticos. Sua presença não se restringe a uma única região geográfica, sendo identificada em diferentes países, o que sugere uma disseminação global de linhagens específicas. A identificação e rastreamento desses clones são cruciais para proporcionar uma compreensão mais profunda da distribuição global, transmissão e evolução das cepas bacterianas (URWIN e MAIDEN, 2023).

A combinação da MLST com a determinação do tipo capsular tem proporcionado uma compreensão mais profunda da dinâmica da estrutura genética da população pneumocócica, especialmente após a introdução da VPC. Além disso, essa abordagem revela-se um método eficaz para identificar e rastrear a troca de sorotipos. Foi através dessa técnica que, no trabalho realizado pelo laboratório em 2018, foi possível observar o fenômeno de substituição de sorotipo, atribuído ao sucesso da implementação das VPCs no programa vacinal na cidade de Niterói (NEVES et al., 2018). No estudo em questão, foram recuperadas 256 amostras de pneumococo da nasofaringe de crianças com idade inferior a 6 anos, provenientes de cinco bairros distintos de Niterói. Dentro desse conjunto, 125 amostras foram obtidas a partir de 123 amostras coletadas no período de dezembro de 2009 a agosto de 2010, antes da introdução da vacina VPC10/13. Adicionalmente, foram recuperadas 131 amostras de um total de 573 amostras coletadas de setembro a dezembro de 2014, já no período pós-implementação da vacina (NEVES et al., 2018). A partir das análises de MLST se observou 97 STs diferentes entre as 256 amostras, incluindo 36 novos STs. 33 STs foram encontrados apenas no período pré-VPC10/13, 47 apenas no período pós-VPC10/13 e 17 em ambos os períodos. Na era pós-VPC10/13, foram detectados 64 STs diferentes entre as 131 amostras, o CC386 foi o CC mais comum. O uso universal da PCV não resultou em frequências mais baixas de resistência a antimicrobianos. Clones internacionalmente disseminados, identificados pelo PMEN (Pneumococcal Molecular Epidemiology Network), foram os principais responsáveis pela resistência. As taxas mais elevadas de resistência em Niterói no período pós-VPC10/13 foram impulsionadas principalmente pelo surgimento do sorotipo MDR 6C-CC386, que diverge em apenas dois loci do clone Poland 6B-20 (NEVES et al., 2018).

O projeto Global Pneumococcal Sequencing (GPS) foi concebido como uma iniciativa colaborativa entre o Instituto Wellcome Sanger e professores de diversas universidades. Seu objetivo era estabelecer um programa de vigilância genômica para sequenciar mais de 20.000 amostras de pneumococos coletados globalmente. Inicialmente voltado para a compreensão da

resistência a antibióticos, o projeto posteriormente direcionou seu foco para traçar um panorama global da evolução do pneumococo durante a introdução de vacinas, utilizando o sequenciamento completo do genoma (Wellcome Trust Sanger Institute, 2022)

No final de 2019, o projeto GPS já havia sequenciado mais de 26.000 amostras de pneumococos de mais de 50 países, resultado de uma colaboração global que envolveu quase 100 cientistas de diversas partes do mundo. Os dados coletados pelo projeto, assim como estudos anteriores, observaram que, após a introdução das vacinas, os pneumococos com sorotipos vacinais foram substituídos por aqueles com sorotipos não vacinais. No entanto, com os dados genômicos em larga escala, o projeto conseguiu ir além, realizando pela primeira vez uma análise abrangente de todas as cepas circulantes globalmente (Wellcome Trust Sanger Institute, 2022)

A análise desses dados representou um avanço significativo para o projeto GPS ao proporcionar uma definição mais precisa das cepas pneumocócicas em alta resolução, considerando variações de DNA em todo o genoma. Os Global Pneumococcal Sequence Clusters (GPSC) consistem em grupos de sequências genômicas de pneumococos que compartilham características semelhantes ou idênticas, sendo definidos com base nas variações genéticas abrangentes em todo o genoma dessas bactérias. Esse método oferece a capacidade de rastrear linhagens pneumocócicas em diferentes regiões geográficas ao longo do tempo, especialmente em relação aos períodos anterior e posterior à implementação das vacinas pneumocócicas conjugadas (GLADSTONE et al., 2019).

### **1.5 – Sequenciamento genômico e ferramentas utilizadas na análise do resistoma bacteriano**

Os avanços no campo do sequenciamento de nova geração (NGS) têm impulsionado de forma significativa a pesquisa sobre resistência antimicrobiana. O NGS, dependendo da natureza das informações requeridas, tem a capacidade de utilizar tanto o DNA quanto o RNA (na forma de cDNA) como material de sequenciamento. O sequenciamento de DNA por NGS, por exemplo, pode revelar a presença de genes relacionados à resistência antimicrobiana, enquanto o sequenciamento de RNA (RNA-Seq) detecta a expressão global de genes, abrangendo inclusive os genes envolvidos na resistência antimicrobiana (CHAN et al., 2016).

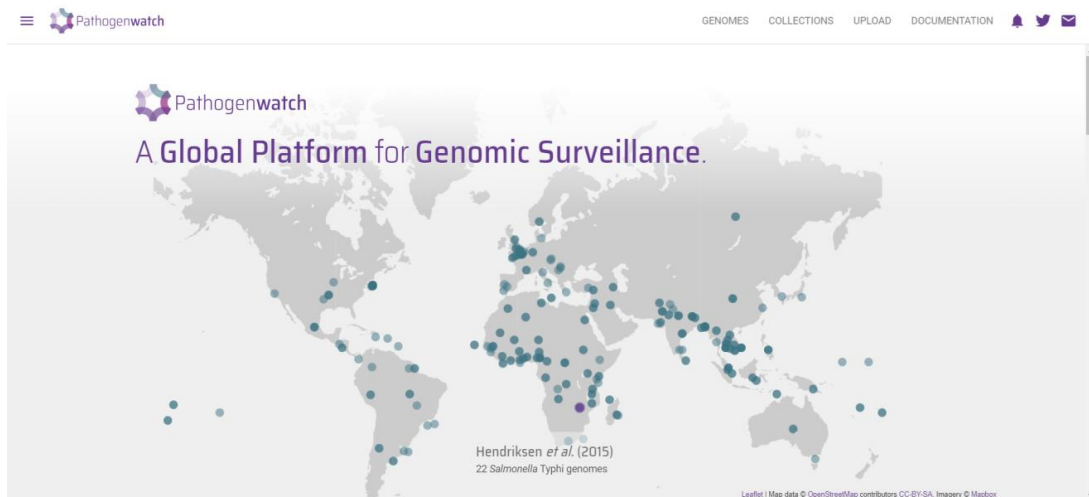
As plataformas de NGS, como a série HiSeq da Illumina, PacBio e Oxford Nanopore, possuem a capacidade de sequenciar centenas de genomas bacterianos completos de forma simultânea. Isso resultou em uma aceleração na geração de dados de sequenciamento e uma melhoria significativa na investigação dos genomas de um amplo espectro de amostras clínicas,

permitindo assim a identificação de novos genes relacionados à resistência antimicrobiana (CHAN et al., 2016).

Diferentes recursos e programas estão disponíveis na internet para análises de dados de sequenciamento de genoma completo. Muitas vezes, tais recursos são reunidos em plataformas, como a BV-BRC (Bacterial and Viral Bioinformatics Resource Center; <https://www.bv-brc.org/>) e o Center for Genomic Epidemiology (<https://www.genomicepidemiology.org/>), que reúnem diferentes programas para montagem de genomas, anotação de genomas e outros tipos de análises de sequências de genoma completo.

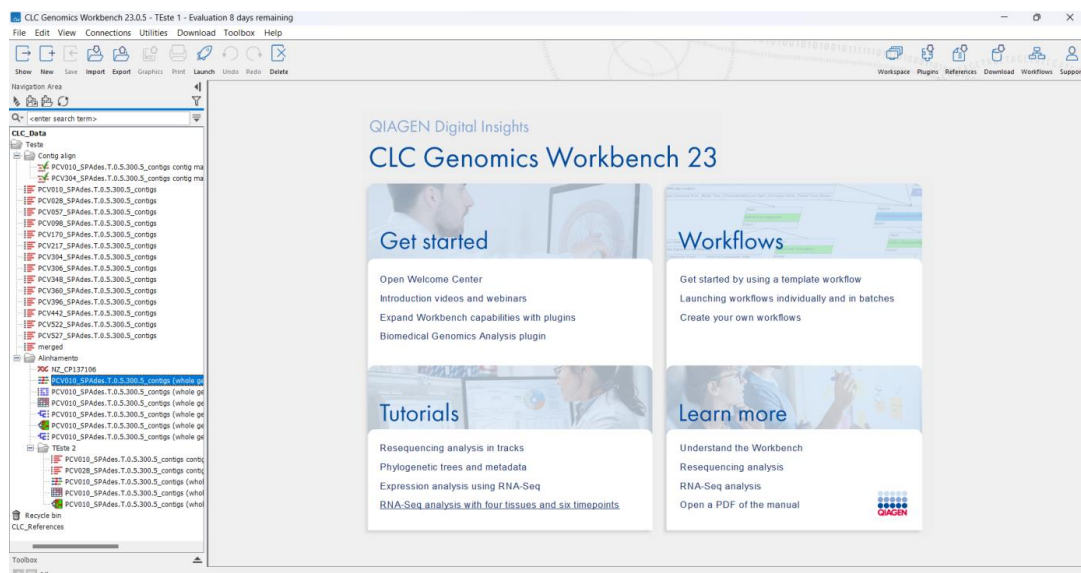
Em conjunto com pipelines de bioinformática especializados, como a ferramenta ResFinder (<https://cge.food.dtu.dk/services/ResFinder/>), disponível na plataforma do Center for Genomic Epidemiology, projetada para a identificação de genes de resistência antimicrobiana adquiridos, o NGS oferece uma vantagem sem precedentes ao proporcionar informações genéticas abrangentes de todo o genoma. Essa capacidade o torna especialmente indicado para a identificação de todos os possíveis determinantes genéticos relacionados à resistência antimicrobiana em um único microrganismo (CHAN et al., 2016; ELLINGTON et al., 2017).

Por sua vez, a PathogenWatch (<https://pathogen.watch/>) é uma plataforma pública baseada na web que auxilia na epidemiologia genômica e na vigilância de saúde pública (Figura 1). A plataforma contém dados genômicos e metadados, incluindo informações sobre resistência antimicrobiana (RAM) e o valor da concentração mínima inibitória (CMI). O PathogenWatch é capaz de processar rapidamente dados de NGS para realizar a tipificação de sequências de múltiplos *loci* (MLST), identificar genes e variações únicas de nucleotídeos (SNPs) relacionados à resistência antimicrobiana, e determinar sua suscetibilidade a antimicrobianos. Além disso, fornece informações sobre os parentes filogenéticos mais próximos e sua localização geográfica, quando disponível. Todos esses dados estão acessíveis por meio de uma interface interativa e podem ser baixados. Em abril de 2022, a plataforma contava com mais de 73.000 genomas bacterianos submetidos (AMOUTZIAS et al., 2022; ARGIMÓN et al., 2021).



**Figura 1.** Interface da Plataforma Pathogenwatch

O QIAGEN CLC Genomics Workbench é uma ferramenta de fácil utilização desenvolvida para análise e visualização eficientes de dados de NGS. Com uma interface intuitiva, oferece uma gama diversificada de recursos cruciais para a manipulação e interpretação de dados genômicos. Suas principais funcionalidades englobam a montagem de genomas e transcriptomas, identificação de variantes genéticas e estudo de expressão gênica. Esse recurso não apenas viabiliza a comparação de genomas, mas também possibilita a exploração e visualização de suas relações evolutivas. (Figura 2).

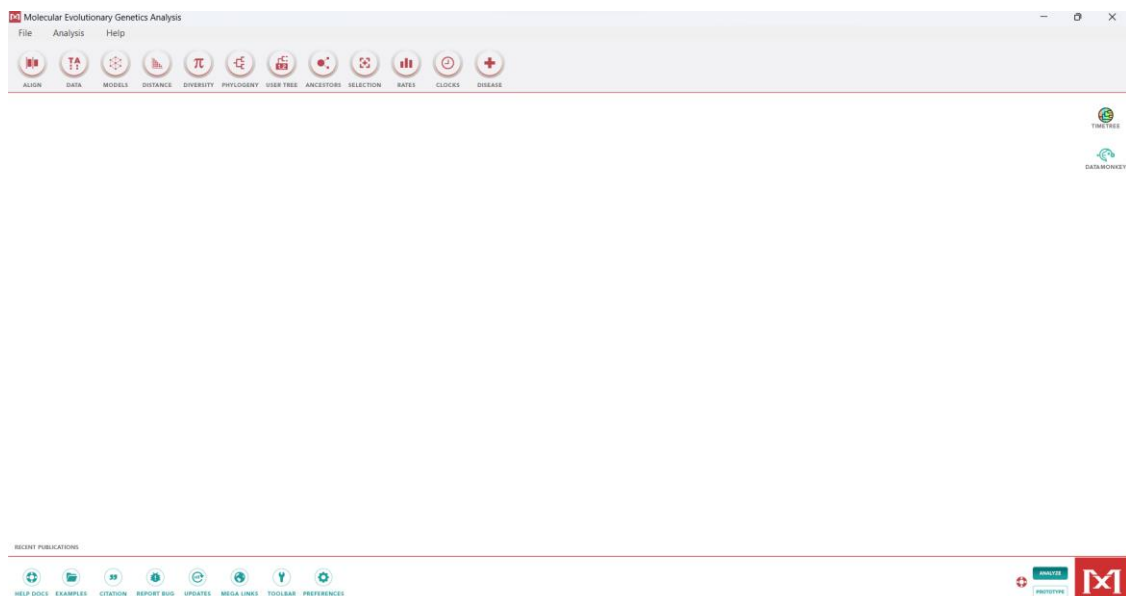


**Figura 2.** Interface do software CLC Genomics Workbench 23

Análises filogenéticas ou filogenômicas também podem ser feitas a partir de dados obtidos com o sequenciamento genômico. Uma árvore filogenética serve como uma estimativa das relações entre táxons (ou sequências) e seus ancestrais comuns hipotéticos. A construção de uma árvore filogenética envolve quatro etapas distintas: identificar e adquirir um conjunto de sequências de DNA ou proteínas homólogas, alinhar essas sequências, estimar uma árvore a partir das sequências alinhadas e apresentar essa árvore de maneira que comunique efetivamente as informações pertinentes a outros (HALL, B. G. 2013).

Tradicionalmente, realizar todas essas etapas exigiria o uso de várias ferramentas de bioinformática. O software MEGA (Molecular Evolutionary Genetics Analysis) simplifica esse processo integrando programas que executam todas as quatro etapas dentro de um ambiente unificado, apresentando uma interface única que elimina a necessidade de conversão entre formatos de arquivo. Além disso, o MEGA oferece a flexibilidade de integrar outros programas para etapas específicas, se desejado (HALL, B. G. 2013; TAMURA et al., 2021).

O MEGA é uma ferramenta integrada para alinhamento automático e manual de sequências, inferência de árvores filogenéticas, mineração de bancos de dados baseados na web, estimativa de taxas de evolução molecular e teste de hipóteses evolutivas. A primeira versão do MEGA foi lançada em 1993. Desde então, o software MEGA amadureceu e vem sendo atualizado para incluir uma ampla coleção de métodos e ferramentas de evolução molecular computacional (Figura 3) (KUMAR et al., 1994; TAMURA et al., 2021)



**Figura 3.** Interface do software MEGA11.

## 1.6 - Justificativa

*Streptococcus pneumoniae* é uma bactéria patogênica responsável por uma diversidade de infecções tanto respiratórias quanto sistêmicas em seres humanos. A habilidade dessa bactéria de evoluir e adquirir resistência a antimicrobianos enfatiza a importância da vigilância contínua, prevenção eficaz e tratamento apropriado. Essas medidas são cruciais para o controle das infecções causadas por essa bactéria e para atenuar os impactos significativos que elas podem ter na saúde pública.

Apesar da redução global nos casos de DPI, a introdução das VPCs em vários países levou a uma substituição dos sorotipos associados à colonização e doenças, um fenômeno chamado "serotype replacement". Estudos do Laboratório de Cocos Gram Positivos da UFF mostram que os sorotipos da VPC10 são agora raramente encontrados em crianças colonizadas ( $\leq 1\%$ ), indicando a eficácia da VPC10 na redução da colonização por sorotipos vacinais. No entanto, sorotipos não incluídos na vacina, especialmente aqueles ligados à resistência a antimicrobianos, estão emergindo. Destaca-se o sorotipo 6C da linhagem ST386, inicialmente caracterizado por uma baixa prevalência na população. Após a introdução da vacina conjugada, contudo, observa-se um aumento significativo na sua prevalência, associado a resistência antimicrobiana.

Assim, o presente estudo se justifica, pois utilizamos dados de sequenciamento de genoma completo de amostras multirresistentes da linhagem genética 6C ST386 emergente em Niterói após a introdução da VPC10 para determinar o resistoma. Essa estratégia tem sido cada vez mais empregada em laboratórios clínicos de grande porte, sobretudo em países de alta renda, para prever o perfil de resistência das amostras clínicas. Portanto, essa abordagem tem aplicação clínica direta.



## 2. OBJETIVOS

### 2.1 Objetivo geral

Avaliar o resistoma de amostras de *Streptococcus pneumoniae* do sorotipo 6C pertencente à linhagem genética ST386, utilizando os dados provenientes do sequenciamento do genoma completo de amostras de colonização isoladas de crianças após quatro anos da introdução da vacina pneumocócica conjugada 10-valente no Brasil.

### 2.2 Objetivo específico

- Identificar *in silico* a presença de genes de resistência a antimicrobianos nas amostras;
- Determinar a sequência dos genes que codificam as PBPs associadas à resistência à penicilina e a outras drogas beta-lactâmicas nas amostras;
- Predizer a concentração mínima inibitória das amostras frente aos beta-lactâmicos;
- Predizer o fenótipo de resistência das amostras a partir dos dados genômicos;
- Realizar uma análise filogenética das amostras.

### 3. MATERIAL E MÉTODOS



**Figura 4.** Fluxograma do projeto

#### 3.1 – Sequências do genoma completo de *Streptococcus pneumoniae*

Foram utilizados os genomas completos de 14 amostras de *S. pneumoniae* pertencentes à linhagem 6C/ST 386 isoladas da nasofaringe de crianças com idade inferior a 6 anos em Niterói em 2014, quatro anos após a introdução da VPC10 para imunização pediátrica de rotina no Brasil. Todas as sequências foram obtidas em um estudo anterior conduzido pelo Laboratório de Cocos Gram Positivos da UFF empregando a plataforma Illumina HiSeq 2000. Adicionalmente, essas sequências foram submetidas ao processo de montagem dos reads pelo software Spades (<https://cab.spbu.ru/software/spades/>) usando a plataforma BV-BRC (Bacterial and Viral Bioinformatics Resource Center; <https://www.bv-brc.org/>).

#### 3.2 – Análises do genoma bacteriano

Por meio da plataforma Pathogenwatch, foram conduzidas diversas análises, incluindo a determinação dos STS, a identificação *in silico* da presença de genes de resistência a antimicrobianos nas amostras. Adicionalmente, determinamos a sequência dos genes que codificam as proteínas de ligação à penicilina (PBPs) associadas à resistência à penicilina e outras drogas beta-lactâmicas nas amostras. Utilizando o Pathogenwatch, também fizemos a predição da concentração mínima inibitória das amostras frente aos beta-lactâmicos e

antecipamos o fenótipo de resistência com base nos dados genômicos. Todas as predições descritas são feitas de modo automático pela ferramenta comparando os dados genômicos depositados com as informações disponíveis em seu banco de dados.

### **3.3 – Análise filogenética das amostras de *Streptococcus pneumoniae* do sorotipo 6C pertencente à linhagem genética ST386**

#### **3.3.1 – Alinhamento das sequências**

Para realizar o alinhamento das sequências genômicas foi utilizado o *plugin* de Alinhamento de Genoma Completo, do software CLC Genomics Workbench.

Para efetuar o alinhamento, utilizamos as 14 amostras obtidas por meio do sequenciamento completo do genoma, conduzido previamente pelo Laboratório de Cocos Gram Positivos da UFF. Essas amostras foram alinhadas em relação a três amostras de referência extraídas do GenBank, identificadas pelos códigos NZ\_CP053210, NZ\_AP026923 e NZ\_LR216060.

A escolha das amostras baseou-se em sua qualidade como sequências de referência de *Streptococcus pneumoniae*, cobrindo o genoma completo da bactéria. As amostras NZ\_CP053210 e NZ\_AP026923, ambas pertencentes ao sorotipo 6A, não dispõem de informações acerca do seu perfil alélico identificado por meio da técnica de MLST, e sobre seu cluster GPSC. Quanto à amostra NZ\_LR216060, do sorotipo 6B, ela está associada ao Cluster GPSC47 e ao ST386. Importante destacar que, até a data de 10/11/2023, não foram identificados dados do genoma completo para a variante 6C no GenBank.

O alinhamento no CLC Workbench foi conduzido utilizando a configuração predefinida "Rearrange Contigs", de forma a otimizar o alinhamento, minimizando as conexões cruzadas entre genomas por meio de ordenação, complementação reversa e deslocamento circular dos *contigs*. Além disso, as configurações de alinhamento foram ajustadas, estabelecendo um comprimento mínimo para o *seed* inicial de 15bp e para o bloco de alinhamento de 100bp. Não foram permitidas divergências nas *seeds*. A partir dos resultados desse alinhamento inicial, procedeu-se à extração do alinhamento de múltiplas sequências.

#### **3.3.2 – Análise do Alinhamento**

Com base no alinhamento das sequências obtido, realizamos a partir do CLC Workbench uma análise da identidade média de nucleotídeos, para avaliar a similaridade

genética entre genomas. Essa análise compara a identidade média de nucleotídeos entre os genomas em questão. O procedimento envolve a divisão dos genomas em fragmentos ou janelas de tamanhos específicos e realiza o cálculo da identidade de nucleotídeos para cada par de fragmentos correspondentes. O resultado da comparação é expresso como uma porcentagem de identidade média entre os genomas comparados. Quanto maior a porcentagem, maior a similaridade genômica. Essa abordagem é especialmente útil para avaliar a diversidade genômica e a relação evolutiva entre microrganismos.

Para realizar esta análise, utilizamos as configurações predefinidas do CLC Workbench. Como não foi especificada nenhuma anotação para as amostras depositadas, a ferramenta divide todas as regiões alinhadas em fragmentos menores de 1024 nucleotídeos. Cada uma dessas regiões menores deve então atender aos critérios de filtragem de conservação. Os parâmetros de filtragem adotados foram 0.8 para a fração mínima de similaridade e 0.8 para a fração mínima de comprimento.

A partir dos dados dessa comparação, desenvolvemos um *heatmap* (mapa de calor) par a par. Na construção dos *clusters* desse mapa, utilizamos a distância Euclidiana como parâmetro e adotamos o critério de ligação completa.

### **3.3.3 – Construção da árvore filogenética**

A elaboração da árvore filogenética foi realizada por meio do método Neighbor-Joining (SAITOU et al., 1987) utilizando o software MEGA 11 (TAMURA et al., 2021). A representação da árvore foi feita em escala, com os comprimentos dos ramos nas mesmas unidades das distâncias evolutivas utilizadas para inferir a árvore filogenética. As distâncias evolutivas foram calculadas usando o método de Kimura com 2 parâmetros e estão expressas nas unidades do número de substituições de bases por sítio (KIMURA M., 1980). Todas as posições ambíguas foram removidas para cada par de sequências (opção de exclusão par a par). A confiabilidade das relações filogenéticas inferidas na árvore foi avaliada através do teste de *bootstrap* com 1.000 replicatas (FELSENSTEIN J., 1985).

## 4. RESULTADOS

### 4.1 Análises do genoma bacteriano

#### 4.1.1 Tipificação de Sequências Multilocus

Ao submeter as amostras à plataforma Pathogenwatch, foi realizada uma análise de MLST, a qual determinou o ST das amostras depositadas. Das 14 amostras analisadas, 13 apresentaram o ST 386, evidenciando o padrão de designação de alelo conforme descrito na Tabela 1. A amostra VPC360 foi associada ao ST 8587, com o padrão de designação de alelo também detalhado na tabela. Ressalta-se que esta amostra diverge em apenas um locus (gene *xpt*) das demais, sendo considerada um SLV (sigle locus variant) do ST386. Todas as 14 amostras foram classificadas como pertencentes ao sorotipo 6C e ao GPSC 47 e complexo clonal 386.

**Tabela 2.** Padrão de designação de alelos a partir do MLST

ST	Padrão de designação de alelo						
	Genes						
	aroE	gdh	gki	recP	spi	xpt	ddl
386	32	28	1	1	15	52	14
8587	32	28	1	1	15	20	14

#### 4.1.2 Caracterização das PBPs

A caracterização das PBPs, por intermédio da plataforma Pathogenwatch revelou que todas as 14 amostras apresentaram o PBP tipo 2 para PBP1a, o PBP tipo 53 para PBP2b e o PBP tipo 77 para PBP2x (perfil 2:53:77).

#### 4.1.3 Caracterização do perfil de resistência a penicilina e outras drogas beta-lactâmicas

Através da análise do tipo de PBP presente nas amostras, foi realizada a predição do fenótipo de resistência, em relação a diversos antimicrobianos beta-lactâmicos. A seleção destes beta-lactâmicos para análise através da ferramenta é guiada pelos antimicrobianos mais frequentemente empregados no tratamento de doenças pneumocócicas, proporcionando uma abordagem mais abrangente e clinicamente relevante. Todas as amostras apresentaram o mesmo perfil de susceptibilidade. Informações detalhadas encontram-se na Tabela 2.

**Tabela 3.** Predições das concentrações mínimas inibitórias (CMI) e fenótipo de resistência a antimicrobianos da classe dos beta-lactâmicos para as 14 amostras de *Streptococcus pneumoniae*.

Agente farmacológico	CMI	Fenótipo Inferido (não-meningite)	Fenótipo Inferido (meningite)
Amoxicilina	= 0.12	Sensível	Sensível
Ceftriaxona	<= 0.5	Sensível	Resistente
Cefotaxima	<= 0.06	Sensível	Resistente
Cefuroxima	<= 0.5	Sensível	Sensível
Meropeném	<= 0.06	Sensível	Sensível
Penicilina	= 0.12	Sensível	Resistente

**Legenda:** O fenótipo inferido para casos não relacionados à meningite está associado ao perfil de resistência identificado em doenças pneumocócicas não invasivas. Em contraste, o fenótipo inferido para casos de meningite está vinculado ao perfil de resistência observado em doenças invasivas, as quais frequentemente requerem terapias com doses mais elevadas.

#### 4.1.4 Análise *in silico* dos genes de resistência a outros antimicrobianos

Além das análises específicas aos beta-lactâmicos, o Pathogenwatch pode realizar previsões *in silico* do perfil de resistência a outras classes de antimicrobianos com base em seu extenso banco de dados. Todas as amostras apresentaram resistência à clindamicina, eritromicina e tetraciclina. A variação nos STs não demonstrou ter impacto no perfil de resistência das amostras.

Adicionalmente, a plataforma identificou os genes específicos responsáveis pela resistência a cada um dos fármacos. O gene *erm(B)*, por exemplo, foi associado à resistência à clindamicina e eritromicina, enquanto o gene *tet(M)\_12* foi identificado como determinante da resistência à tetraciclina.

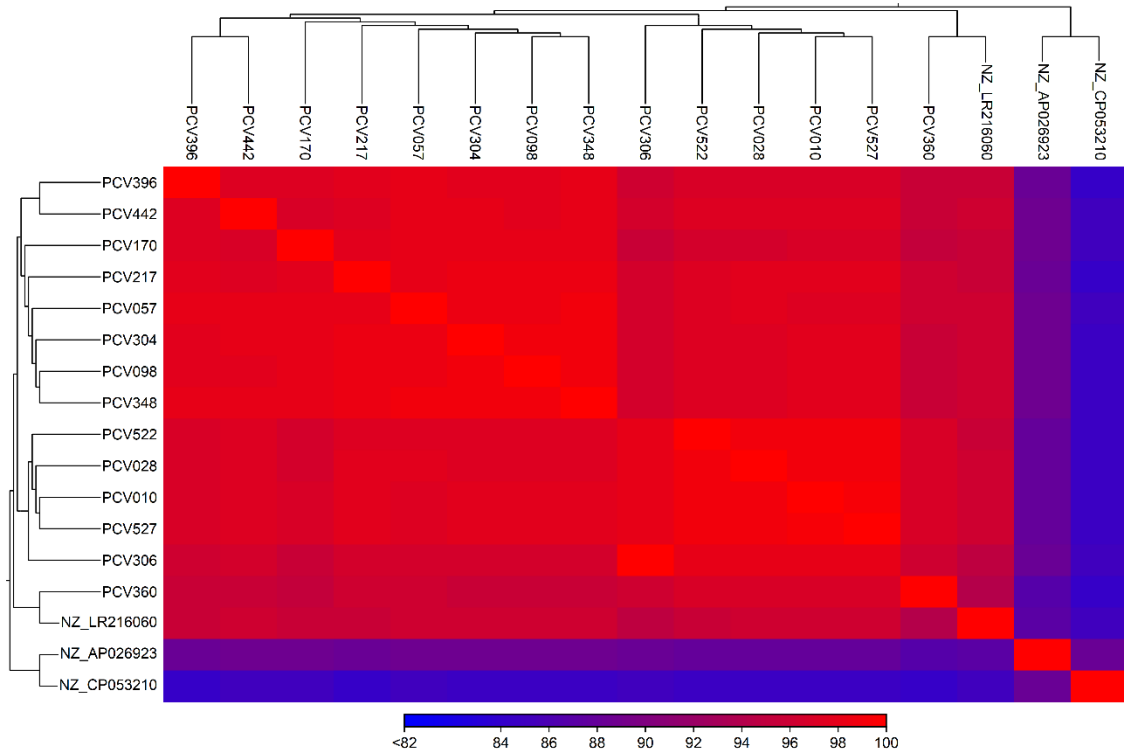
## 4.2 Análise filogenética das amostras de *Streptococcus pneumoniae* do sorotipo 6C pertencente à linhagem genética ST386

### 4.2.1 Comparação de Identidade Média de Nucleotídeos

A análise da identidade média de nucleotídeos, revelou uma identidade média acima de 94% entre nossas 14 amostras. Destaca-se que as identidades médias mais baixas (83,84% - 84,89%) foram identificadas na comparação de nossas amostras com a sequência de referência

NZ\_CP053210 (sorotipo 6A), enquanto as médias um pouco mais elevadas foram registradas na comparação com a sequência de referência NZ\_AP026923. (86,39% - 88,46%) (sorotipo 6A). A sequência NZ\_LR216060(sorotipo 6B) foi a que mais se aproximou em termos de identidade de nucleotídeos às nossas amostras (93,81%-95,87%)

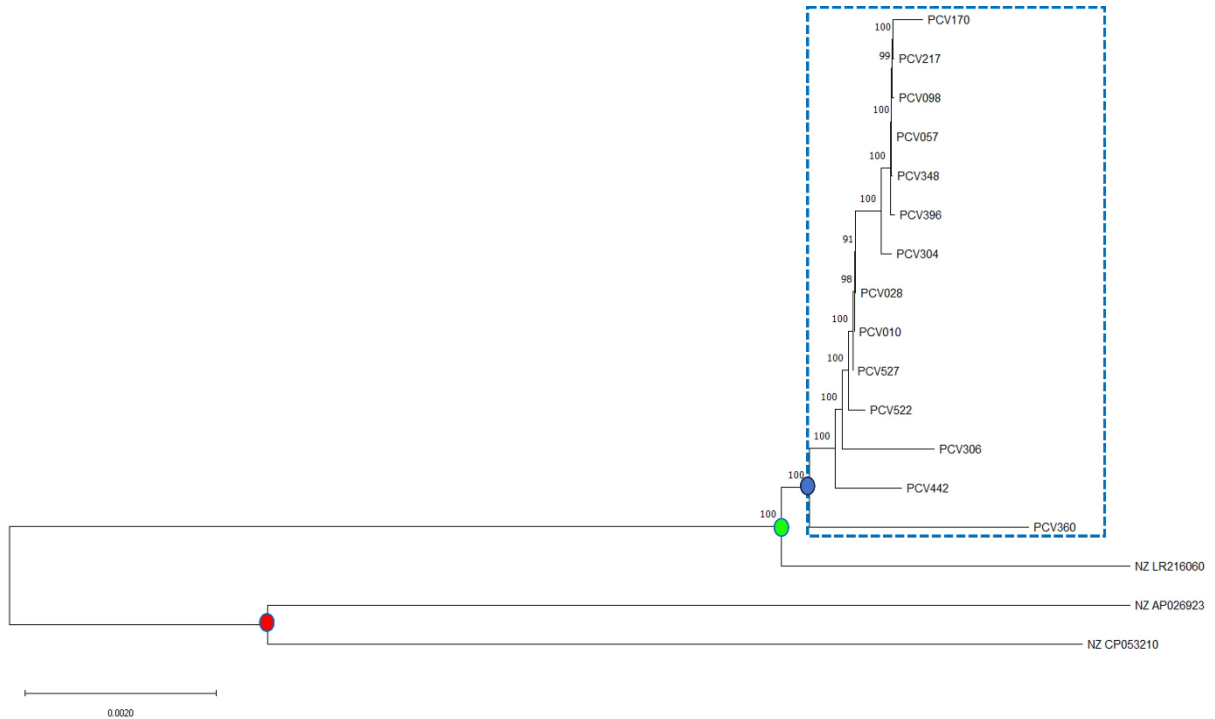
Com o intuito de proporcionar uma visualização mais clara das variações nas médias de identidade, foi criado um *heatmap* utilizando os dados provenientes da comparação de identidade média de nucleotídeos, no qual as amostras de referência do sorotipo 6A claramente podem ser visualizadas com identidades médias inferiores a 90% em relação às nossas sequências (Figura 4), sendo a sequência NZ\_LR216060 a mais relacionada geneticamente com nossas amostras, especialmente a amostras VPC360.



**Figura 5.** *Heatmap* apresentando as variações nas médias de identidade entre os genomas de 17 amostras de *Streptococcus pneumoniae*, sendo 14 isoladas de crianças em Niterói e três utilizadas como amostras de referência para fins comparativos. O esquema de cores refere-se à identidade média de nucleotídeos. As amostras exibidas em azul demonstram menor identidade nucleotídica, enquanto quanto maior a intensidade do tom vermelho, maior a identidade nucleotídica entre as sequências.

### 4.3 Construção da árvore filogenética com MEGA

A árvore filogenética foi construída utilizando o método Neighbor-Joining. A análise em foco foi conduzida mediante o alinhamento das nossas 14 amostras com as três amostras de referência previamente obtidas do GenBank, conforme detalhado anteriormente (Figura 5).



**Figura 6.** Árvore filogenética construída com o software MEGA 11, utilizando o algoritmo *Neighbour Joining*, optando pelo modelo de substituição Kimura 2p e 1000 replicatas de *bootstrap*.

A análise revelou maior proximidade genética das amostras com a sequência de referência NZLR216060, do sorotipo 6B em comparação com as sequências NZAP026923 e NZCP053210 (círculo verde), que pertencem ao sorotipo 6A (círculo vermelho). Todas as genomas do sorotipo 6C incluídos na análise, independente do ST, divergiram de um último ancestral comum (círculo verde) que deu origem também à cepa NZLR216060, do sorotipo 6B. Dentro do conjunto de nossas amostras, identificamos a formação de um clado bem definido de amostras do sorotipo 6C, ST386 (círculo azul). A amostra mais divergente foi a VPC360 que exibiu o perfil alélico ST8587. Todas as amostras apresentaram uma estreita relação filogenética, embora três amostras (VPC306, VPC442 e VPC360) tenham ficado mais distantes das demais. A árvore gerada apresentou alta confiabilidade, revelada por ramos com *bootstraps* superiores a 90%.



## 5. DISCUSSÃO

Desde a introdução da VPC10, observa-se uma queda significativa na incidência de doenças pneumocócicas, acompanhada por uma redução nas taxas de mortalidade (MOTT et al., 2019). Apesar da ampla cobertura (86%) da VPC10 no Brasil, a proporção de pneumococos não suscetíveis à penicilina em crianças com menos de 6 anos, que já era elevada, apresentou um aumento de 19% após quatro anos da implementação da vacina. Além disso, as taxas de resistência à eritromicina, clindamicina e tetraciclina aumentaram em pelo menos 20% quando comparadas com o período pré-VPC10, principalmente devido ao surgimento de amostras do sorotipo 6C resistentes a múltiplos antimicrobianos (NEVES et al., 2018).

O aumento de microrganismos resistentes a antimicrobianos tem impulsionado recentes avanços na microbiologia clínica, promovendo o aprimoramento de ferramentas de diagnóstico. Além de contribuir para o aumento da mortalidade, essas bactérias resistentes acarretam custos significativos, tanto diretos quanto indiretos, nos cuidados com a saúde, representando assim uma ameaça global à saúde pública. Testes fenotípicos, como a determinação da Concentração Inibitória Mínima e a técnica de difusão em disco, desempenham papel essencial na categorização de amostras clínicas como suscetíveis ou resistentes. Contudo, em situações específicas, esses métodos podem ser insuficientes, exigindo a aplicação de abordagens adicionais para identificar o mecanismo de resistência. A confirmação da presença de genes associados, geralmente realizada por Reação em Cadeia da Polimerase (PCR), é uma prática comum. No entanto, devido à extensa diversidade de genes e suas variantes alélicas, esses métodos nem sempre conseguem identificar o padrão fenotípico correspondente (HRABAK et al., 2020)

A implementação do sequenciamento de nova geração vem se tornando parte da realidade de laboratórios de microbiologia clínica, principalmente em países mais desenvolvidos, sendo utilizado para o gerenciamento de surtos, caracterização e vigilância de patógenos, assim como para a identificação de organismos desconhecidos (DEURENBERG et al., 2017). As sequências genômicas completas apresentam uma grande vantagem nesse contexto, pois têm a capacidade de fornecer informações detalhadas sobre a classificação taxonômica, variações genômicas, cadeias de transmissão e a presença de resistência antimicrobiana ou marcadores de virulência (LEPUSCHITZ ET AL., 2020)

Neste estudo, examinamos o genoma completo de 14 amostras de *S. pneumoniae* da linhagem 6C/ST 386, isoladas da nasofaringe de crianças com menos de 6 anos. Por meio das análises realizadas no Pathogenwatch, foi observada a presença de resistências a eritromicina, clindamicina e tetraciclina em todas as 14 amostras de *S. pneumoniae* da linhagem 6C que

foram examinadas. Além disso, identificamos os genes *erm(B)* e *tet(M)* como determinantes dessas resistências.

No período que antecedeu a introdução da vacina, notamos um perfil mais suscetível a esses antimicrobianos (NEVES et al 2013). No entanto, após a implementação da vacina e a emergência de sorotipos não vacinais, observamos uma alteração no perfil de resistência do *S. pneumoniae* na população. Essas descobertas destacam a dinâmica complexa entre a vacinação, a evolução dos sorotipos e as mudanças nos padrões de resistência antimicrobiana (FORTUNA et al 2023).

Ao comparar os resultados do estudo com as amostras de 2010 realizado pelo laboratório (Neves et al. 2013) com as análises de resistência à penicilina conduzidas no Pathogenwatch em nossas amostras, destaca-se uma consistência marcante. O estudo de 2013 revelou que 27,3% das amostras apresentavam não susceptibilidade à penicilina, com CMI variando de 0,12 a 4,0 µg/ml. De acordo com os pontos de corte do *Clinical Laboratory Standard Institute (CLSI)* para a penicilina parenteral (meningite) documento dos EUA que padroniza a realização e interpretação dos testes de susceptibilidade a antimicrobianos, todas as 33 amostras seriam consideradas resistentes à penicilina. No entanto, para a penicilina parenteral (não meningite) nenhuma amostra seria resistente (NEVES et al. 2013).

Mesmo considerando que nossas amostras pertencem ao período pós-implementação da VPC10, observamos uma semelhança com os resultados encontrados em amostras prévias à vacinação. Todas as amostras do sorotipo 6C apresentaram resistência à penicilina parenteral no contexto de meningite, com uma CMI de 0,12 µg/ml, enquanto nenhuma evidenciou resistência à penicilina parenteral em cenários não relacionados à meningite.

Para realizar uma análise evolutiva das amostras, empregamos o alinhamento por meio da ferramenta CLC Workbench. A partir desse alinhamento, conduzimos uma análise preliminar da Identidade Média de Nucleotídeos, permitindo-nos avaliar a proximidade genética entre os genomas. Nessa análise, comparamos nossas 14 amostras com três genomas de referência. Devido à ausência de dados de genoma completo para o sorotipo 6C no Genbank, optamos por utilizar genomas dos sorotipos 6A e 6B, que são sorotipos geneticamente próximos ao 6C. Conforme esperado a amostra NZ\_LR216060, que pertence à mesma cepa GPSC 47 que as nossas amostras, demonstrou maior identidade nucleotídica e conseqüentemente maior proximidade filogenética com nossas amostras.

Para a análise evolutiva das amostras, inicialmente, conduzimos uma análise do *core* genoma no Pathogenwatch. Durante esse processo, identificamos a formação de um cluster que engloba as amostras VPC 057, VPC 098, VPC 170, VPC 217, VPC 348 e VPC 396. As análises

filogenéticas revelaram dentro do conjunto de nossas amostras a formação de um clado bem definido, compreendendo as amostras VPC217, VPC170, VPC098, VPC057, VPC348 e VPC396. No entanto, a maioria das amostras apresentou uma estreita relação filogenética, embora três amostras (VPC306, VPC442 e VPC360) tenham ficado mais distantes das demais.

Além disso, identificamos a formação de um clado menor, composto pelas amostras VPC 010, VPC 028 e VPC 527. Na análise do *core* genoma, notamos a constituição de um cluster envolvendo essas amostras, juntamente com a amostra VPC 522. Ao examinar a árvore filogenética, observamos que os ramos dessas amostras se originam de um mesmo nó compartilhando um ancestral comum.

A análise da árvore revela uma maior proximidade das amostras com a sequência de referência NZ\_LR216060 do sorotipo 6B em comparação com as amostras do sorotipo 6A. Este comportamento corrobora as análises de identidade média de nucleotídeos.

Nossos resultados corroboram dados recentes da literatura que sugerem amostras do ST386 (GPSC47) do período pós vacinal são pertencentes ao sorotipo 6C. Adicionalmente o estudo de Almeida e colaboradores (2021) também evidenciou a divergência de cepas do ST386 do sorotipo 6B, isoladas no período pré-vacinal das amostras do sorotipo 6C, sugerindo através de análises de datação molecular, a ocorrência de um evento de switch capsular de amostras do sorotipo 6B por volta do ano de 1994, e posteriormente e com a introdução das VPCs, ocorreu a seleção da linhagem não vacinal ST386, sorotipo 6C (GPSC47).

A principal limitação deste estudo foi a restrição computacional, que impediu a realização de análises mais robustas, como a construção de árvores por meio da técnica de Máxima Verossimilhança e a investigação de variantes de polimorfismos únicos. Devido ao tamanho das sequências das amostras, não foi viável utilizar ferramentas disponíveis em servidores web, sendo necessário conduzir a maior parte dos testes em um *desktop*, o que demanda *hardware* mais potente para análises mais complexas. Apesar das limitações de *hardware*, as análises realizadas apresentaram resultados condizentes com o que foi relatado na literatura.

*Streptococcus pneumoniae* destaca-se como um patógeno de grande relevância médica, com uma complexa interação entre a vacinação, a evolução dos sorotipos e as mudanças nos padrões de resistência antimicrobiana. Essa complexidade representa um desafio significativo na elaboração de tratamentos eficazes contra esse patógeno. As investigações genômicas realizadas *in silico* mostraram-se consistentes com os relatos encontrados na literatura, contribuindo para a compreensão das bases genéticas da resistência (ALMEIDA et al., 2021).

A análise dos dados genômicos permitiu a predição do fenótipo de resistência das amostras, indicando que, embora tenha ocorrido uma mudança nos sorotipos após a introdução da vacina, o perfil de resistência antimicrobiana manteve algumas semelhanças com amostras prévias à vacinação, especialmente em relação à penicilina.

A resistência antimicrobiana, especialmente à penicilina, permanece uma preocupação, apesar dos benefícios observados com a vacinação. Este estudo destaca a complexidade na interação entre vacinação, evolução de sorotipos e resistência, enfatizando a necessidade contínua de monitoramento e estratégias adaptativas no controle das infecções por *S. pneumoniae*, bem como demonstra a possibilidade da utilização de ferramentas *in silico* simples e de fácil aplicação para predição da resistência a antimicrobianos a partir de dados de sequenciamento de genoma completo de uma espécie bacteriana de grande relevância clínica.

O uso de dados provenientes do sequenciamento do genoma completo associados a ferramenta de bioinformática se mostra uma metodologia cada vez mais valiosa na análise de características genômicas e padrões de resistência antimicrobiana. O progresso nas ferramentas de fácil utilização permite uma abordagem prática e dinâmica, facilitando cada vez mais a incorporação dessas ferramentas em laboratórios de análises clínicas, muitas vezes carentes de mão de obra capacitada para análises *in silico* mais robustas. Os resultados obtidos por meio dessas ferramentas proporcionaram uma visão abrangente das características genômicas e dos padrões de resistência antimicrobiana da linhagem utilizada no estudo, possibilitando uma melhor compreensão da dinâmica evolutiva da linhagem 6C ST386 após a implementação da vacina.

## 6. CONCLUSÃO

- Através do uso de ferramentas bioinformáticas, foi possível identificar *in silico* a presença de genes de resistência a antimicrobianos nas amostras de *Streptococcus pneumoniae* do sorotipo 6C. Essa análise revelou a presença de genes específicos, como *erm(B)* e *tet(M)\_12*, associados à resistência a clindamicina, eritromicina e tetraciclina, contribuindo para a compreensão do perfil de resistência antimicrobiana dessa linhagem.
- A caracterização das proteínas ligadoras de penicilina (PBPs) por meio de análises bioinformáticas identificou sequências específicas, revelando a presença de PBP tipo 2 para PBP1a, PBP tipo 53 para PBP2b, e PBP tipo 77 para PBP2x em todas as amostras. Esses achados contribuem para a compreensão dos mecanismos genéticos subjacentes à resistência a penicilina e outras drogas beta-lactâmicas.
- A análise dos dados genômicos permitiu a predição do fenótipo de resistência das amostras, incluindo a determinação da concentração mínima inibitória, o que ressalta a importância deste tipo de estudo como base para a proposta de utilização de ferramentas NGS na rotina clínica.
- A análise filogenética, conduzida com base no sequenciamento do genoma completo, proporcionou uma compreensão das relações evolutivas entre as amostras, corroborando o íntimo relacionamento genético entre as amostras do sorotipo 6C, complexo clonal 386

## 7. REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, Ana Lucia et al. Population-based surveillance for invasive pneumococcal disease and pneumonia in infants and young children in Goiania, Brazil. **Vaccine**, v. 30, n. 10, p. 1901-1909, 2012.

ALMEIDA, Samanta CG et al. Genomic surveillance of invasive *Streptococcus pneumoniae* isolates in the period pre-PCV10 and post-PCV10 introduction in Brazil. **Microbial Genomics**, v. 7, n. 10, 2021.

BORTOLAIA, Valeria et al. ResFinder 4.0 for predictions of phenotypes from genotypes. **Journal of Antimicrobial Chemotherapy**, v. 75, n. 12, p. 3491-3500, 2020.

BRANDILEONE, Maria-Cristina C. et al. Distribution of invasive *Streptococcus pneumoniae* serotypes before and 5 years after the introduction of 10-valent pneumococcal conjugate vaccine in Brazil. **Vaccine**, v. 36, n. 19, p. 2559-2566, 2018.

CAIERAO, Juliana et al. Serotypes and genotypes of invasive *Streptococcus pneumoniae* before and after PCV10 implementation in southern Brazil. **PLoS One**, v. 9, n. 10, p. e111129, 2014.

CDC. Antibiotic Resistance Threats in the United States, 2019. Atlanta, GA: U.S. Department of Health and Human Services, CDC; 2019. Disponível em: [www.cdc.gov/DrugResistance/Biggest-Threats.html](http://www.cdc.gov/DrugResistance/Biggest-Threats.html). Acesso em 01 de dez. 2023

CHAN, Kok-Gan. Whole-genome sequencing in the prediction of antimicrobial resistance. **Expert review of anti-infective therapy**, v. 14, n. 7, p. 617-619, 2016.

CAMILLI, Romina et al. Complete genome sequence of a serotype 11A, ST62 *Streptococcus pneumoniae* invasive isolate. **BMC microbiology**, v. 11, n. 1, p. 1-13, 2011.

DA SILVA, Amanda Beiral et al. Carriage prevalence, serotype distribution, antimicrobial resistance, *pspA* typing and pilus islets of *Streptococcus pneumoniae* isolated from adults living in a Brazilian urban slum. **Vaccine**, v. 41, n. 8, p. 1431-1437, 2023.

D'Aeth JC, van der Linden MP, McGee L, Lencastre H, Turner P, Hoon CJ, et al., The role of interspecies recombination in the evolution of antibiotic-resistant pneumococci. **Elife**. 2021;10:e67113.

DEURENBERG, Ruud H. et al. Application of next generation sequencing in clinical microbiology and infection prevention. *Journal of biotechnology*, v. 243, p. 16-24, 2017.

ELLINGTON, M. J. et al. The role of whole genome sequencing in antimicrobial susceptibility testing of bacteria: report from the EUCAST Subcommittee. **Clinical microbiology and infection**, v. 23, n. 1, p. 2-22, 2017.

GLADSTONE, Rebecca A. et al. International genomic definition of pneumococcal lineages, to contextualise disease, antibiotic resistance and vaccine impact. **EBioMedicine**, v. 43, p. 338-346, 2019.

HALLGREN, Malte B. et al. MINTyper: an outbreak-detection method for accurate and rapid SNP typing of clonal clusters with noisy long reads. **Biology Methods and Protocols**, v. 6, n. 1, p. bpab008, 2021.

HOWARD, Leigh M. et al. Association between nasopharyngeal colonization with multiple pneumococcal serotypes and total pneumococcal colonization density in young Peruvian children. **International Journal of Infectious Diseases**, v. 134, p. 248-255, 2023.

HRABAK, Jaroslav; BITAR, Ibrahim; PAPAGIANNITSIS, Costas C. Combination of mass spectrometry and DNA sequencing for detection of antibiotic resistance in diagnostic laboratories. **Folia Microbiologica**, v. 65, p. 233-243, 2020.

JACOBS, Michael R. et al. Occurrence, distribution, and origins of *Streptococcus pneumoniae* serotype 6C, a recently recognized serotype. **Journal of clinical microbiology**, v. 47, n. 1, p. 64-72, 2009.

KIM, Gyu-Lee; SEON, Seung-Han; RHEE, Dong-Kwon. Pneumonia and *Streptococcus pneumoniae* vaccine. *Archives of pharmacal research*, v. 40, p. 885-893, 2017.

Leonard A, Möhlis K, Schlüter R, Taylor E, Lalk M, Methling K. Exploring metabolic adaptation of *Streptococcus pneumoniae* to antibiotics. **J Antibiot** (Tokyo). 2020;73(7):441-454.

LEPUSCHITZ, Sarah et al. Analytical performance validation of next-generation sequencing based clinical microbiology assays using a K-mer analysis workflow. **Frontiers in Microbiology**, v. 11, p. 1883, 2020.

MOTT, M. P. et al. Emergence of serotype 19A *Streptococcus pneumoniae* after PCV10 associated with a ST320 in adult population, in Porto Alegre, Brazil. **Epidemiology & Infection**, v. 147, p. e93, 2019.

Murray CJL, Ikuta KS, Sharara F, Swetschinski L, Aguilar GR, Gray A, et al., Antimicrobial Resistance Collaborators. Global burden of bacterial antimicrobial resistance in 2019: a systematic analysis. *Lancet*. 2022;399(10325):629-655.

NEVES, Felipe Piedade Gonçalves et al. Nasopharyngeal carriage, serotype distribution and antimicrobial resistance of *Streptococcus pneumoniae* among children from Brazil before the introduction of the 10-valent conjugate vaccine. **BMC Infectious Diseases**, v. 13, n. 1, p. 1-7, 2013.

NEVES, Felipe PG et al. Pneumococcal carriage among children after four years of routine 10-valent pneumococcal conjugate vaccine use in Brazil: the emergence of multidrug resistant serotype 6C. **Vaccine**, v. 35, n. 21, p. 2794-2800, 2017.

NEVES, Felipe; PINTO, Tatiana. *Streptococcus Pneumoniae*. In: FILIPPIS, Ivano. **Molecular Typing in Bacterial Infections, Volume I**. Rio de Janeiro: Springer, 2022.p. 139-152.

Organização Mundial de Saúde (OMS). **WHO publishes list of bacteria for which new antibiotics are urgently needed**, 2017. Disponível em: <https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed>. Acesso em: 02 set. 2023.



Organização Mundial de Saúde (OMS). **WHO Fact sheets about Pneumonia in children, 2022**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/pneumonia>. Acesso em: 26 nov. 2023.

PARK, In Ho et al. Discovery of a new capsular serotype (6C) within serogroup 6 of *Streptococcus pneumoniae*. **Journal of clinical microbiology**, v. 45, n. 4, p. 1225-1233, 2007.

POOLMAN, Jan T.; PEETERS, Carla CAM; VAN DEN DOBBELSTEEN, Germie PJM. The history of pneumococcal conjugate vaccine development: dose selection. **Expert review of vaccines**, v. 12, n. 12, p. 1379-1394, 2013.

REINERT, R. R. The antimicrobial resistance profile of *Streptococcus pneumoniae*. *Clinical Microbiology and Infection*, v. 15, p. 7-11, 2009. REIS, Joice Neves et al. Transmission of *Streptococcus pneumoniae* in an urban slum community. **Journal of Infection**, v. 57, n. 3, p. 204-213, 2008.

SANTORO, Francesco; IANNELLI, Francesco; POZZI, Gianni. Genomics and Genetics of *Streptococcus pneumoniae*. **Microbiology Spectrum**, v. 7, n. 3, p. 7.3. 5, 2019.

URWIN, Rachel; MAIDEN, Martin CJ. Multi-locus sequence typing: a tool for global epidemiology. **Trends in microbiology**, v. 11, n. 10, p. 479-487, 2003.

WEISER, Jeffrey N.; FERREIRA, Daniela M.; PATON, James C. *Streptococcus pneumoniae*: transmission, colonization and invasion. **Nature Reviews Microbiology**, v. 16, n. 6, p. 355-367, 2018.

Wellcome Trust Sanger Institute. Global Pneumococcal Sequencing Project, 2022. Project Outline. Disponível em: [https://www.pneumogen.net/gps/project\\_outline.html](https://www.pneumogen.net/gps/project_outline.html). Acesso em: 27 de nov. de 2023