

**Pengembangan Metode Identifikasi Serotipe *Streptococcus pneumoniae* Berbasis
CRISPR/Cas9 Targeted Sequencing: Studi Serotipe dan Analisis Transposon Pembawa
Gen Resistensi Antibiotik**

Yustinus Maladan

NIM: 22/500672/SBI/00210

INTISARI

Streptococcus pneumoniae (pneumokokus) merupakan penyebab utama penyakit pneumonia, meningitis, dan sepsis, terutama pada anak-anak dan orang dewasa. Pneumonia akibat *S. pneumoniae* menyumbang sekitar 15% dari seluruh kematian anak usia di bawah lima tahun secara global. Resistensi antibiotik menjadi tantangan utama dalam pengendalian penyakit ini. Meskipun data resistensi telah banyak dilaporkan, kajian mendalam mengenai dasar genetik seperti transposon pembawa gen resistensi masih terbatas, khususnya di Indonesia. Selain itu, keragaman serotipe serta keberadaan isolat *non-typeable pneumococci* (NTP) turut menyulitkan identifikasi serotipe yang penting untuk strategi vaksinasi. Penelitian ini bertujuan untuk mengidentifikasi transposon penyebab resistensi terhadap makrolida, tetrasiklin, dan kloramfenikol menggunakan data *whole genome sequencing* (WGS), serta mengembangkan metode identifikasi serotipe *S. pneumoniae* berbasis *CRISPR/Cas9 targeted sequencing* pada platform *Oxford Nanopore Technologies* (ONT).

Sebanyak 90 isolat WGS dari berbagai wilayah Indonesia dianalisis secara *de novo* dan *resequencing* menggunakan pipeline ASA3P. Prediksi resistensi dilakukan dengan Pathogenwatch, sedangkan identifikasi dan visualisasi transposon menggunakan MicroScope, PyGenomeViz, dan DNA Features Viewer. Uji konfirmasi fenotipik dilakukan dengan metode MIC menggunakan VITEK 2, sedangkan *molecular docking* dan *molecular dynamics simulation* (MDS) dilakukan menggunakan AutoDock Vina dan NAMD 3.0.1. Untuk pengembangan metode identifikasi serotipe, *crRNA* sepanjang 20 bp didesain dengan menargetkan gen *dexB* dan *aliA* menggunakan CHOPCHOP. Metode *CRISPR/Cas9 targeted sequencing* kemudian diaplikasikan pada platform Oxford Nanopore Technologies (ONT), dikombinasikan dengan *barcoding* menggunakan *Native Barcoding Expansion*. Pipeline untuk identifikasi serotipe *S. pneumoniae* dikembangkan menggunakan skrip Python yang didesain khusus untuk menganalisis hasil sekuensing lokus *cps*.

Hasil identifikasi menunjukkan bahwa dari 90 isolat, 28 isolat resisten terhadap makrolida, 61 terhadap tetrasiklin, 15 terhadap kloramfenikol, dan 22 terhadap klindamisin. Transposon Tn916, Tn5253, Tn6002, Tn2009, dan Tn2010 teridentifikasi sebagai penyebab utama resistensi. Tn916, Tn5253, dan Tn6002 tersebar pada beberapa serotipe, sedangkan Tn2009 ditemukan pada serotipe 19 dan 23F. Sementara itu, Tn2010 hanya ditemukan pada serotipe 19F. Serotipe 19F merupakan yang paling dominan dan menunjukkan tingkat resistensi makrolida tertinggi. Hasil uji MIC menunjukkan sebagian besar konsisten dengan hasil prediksi genetik, meskipun ditemukan dua kasus ketidaksesuaian: isolat MA015B yang membawa Tn5253 tetap sensitif terhadap tetrasiklin, dan isolat BA0033B yang membawa gen *mefA* dan *ermB* namun sensitif terhadap makrolida. Analisis integrasi transposon menunjukkan bahwa Tn6002 dan Tn916 tidak memiliki pola integrasi tetap antar-isolat, sedangkan Tn5253,



UNIVERSITAS
GADJAH MADA

Pengembangan Metode Identifikasi Serotipe *Streptococcus pneumoniae* Berbasis CRISPR/Cas9 Targeted

Sequencing: Studi Serotipe dan Analisis Transposon Pembawa Gen Resistensi Antibiotik

Yustinus Maladan, Prof. Dr. Endah Retnaningrum, S.Si., M. Eng; Prof. Dr. Budi Setiadi Daryono, M.Agr.Sc.; Dodi Saf

Universitas Gadjah Mada, 2025 | Diunduh dari <http://etd.repository.ugm.ac.id/>

Tn2009, dan Tn2010 menunjukkan lokasi integrasi konservatif. Tn5253 terintegrasi pada wilayah intergenik dekat gen *rbGA*, sementara Tn2009 dan Tn2010 ditemukan pada daerah intergenik sekitar ORF pengkode *Bacteriocin-associated integral membrane protein*. Analisis MDS menunjukkan bahwa eritromisin mengalami proses efluks melalui protein efluks makrolida pada 27,6 ns. Residu Asn195, Gln196, Val197, Gln198, dan Ser199 yang terletak pada daerah loop internal menunjukkan nilai RMSF yang tinggi, mengindikasikan fleksibilitas signifikan yang kemungkinan berperan penting dalam proses efluks tersebut. Metode *CRISPR/Cas9 targeted sequencing* yang dikembangkan berhasil mengidentifikasi serotipe secara akurat, dengan hasil identik terhadap metode WGS berbasis Illumina. Hal ini menunjukkan bahwa metode ini memiliki akurasi tinggi dan berpotensi menjadi metode alternatif yang efisien untuk identifikasi serotipe *S. pneumoniae*.

Penelitian ini berhasil mengungkap keberadaan berbagai jenis transposon pembawa gen resistensi antibiotik, menjelaskan pola integrasi dan mekanisme efluks yang dimediasi oleh gen *mefA*, serta membuktikan akurasi metode baru identifikasi serotipe *S. pneumoniae* berbasis *CRISPR/Cas9 targeted sequencing* pada platform ONT.

Kata kunci: *Streptococcus pneumoniae*, makrolida, *Cas9 targeted sequencing*, transposon

Development of a CRISPR/Cas9-Based Targeted Sequencing Method for Serotype Identification of *Streptococcus pneumoniae*: Serotype Study and Analysis of Transposons Carrying Antibiotic Resistance Genes

Yustinus Maladan

NIM: 22/500672/SBI/00210

ABSTRACT

Streptococcus pneumoniae (pneumococcus) is a major cause of pneumonia, meningitis, and sepsis, especially in children and adults. Pneumonia caused by *S. pneumoniae* contributes to around 15% of all deaths in children under five years of age globally. Antibiotic resistance is a major challenge in controlling this disease. Although resistance data have been widely reported, in-depth studies of the genetic basis such as transposons carrying resistance genes are still limited, especially in Indonesia. In addition, the diversity of serotypes and the presence of non-typeable pneumococci (NTP) isolates also make it difficult to identify serotypes that are important for vaccination strategies. This study aims to identify transposons that cause resistance to macrolides, tetracyclines, and chloramphenicol using whole genome sequencing (WGS) data, and to develop a method for identifying *S. pneumoniae* serotypes based on CRISPR/Cas9 targeted sequencing on the Oxford Nanopore Technologies (ONT) platform.

A total of 90 WGS isolates from various regions of Indonesia were analyzed de novo and resequencing using the ASA3P pipeline. Resistance prediction was performed using Pathogenwatch, while transposon identification and visualization were performed using MicroScope, PyGenomeViz, and DNA Features Viewer. Phenotypic confirmation testing was performed using the MIC method using VITEK 2, while molecular docking and molecular dynamics simulation were performed using AutoDock Vina and NAMD 3.0.1. For the development of the serotype identification method, a 20 bp crRNA was designed by targeting the *dexB* and *aliA* genes using CHOPCHOP. The CRISPR/Cas9 targeted sequencing method was then applied to the Oxford Nanopore Technologies (ONT) platform, combined with barcoding using Native Barcoding Expansion. The pipeline for *S. pneumoniae* serotype identification was developed using a Python script specifically designed to analyze the results of cps locus sequencing.

The identification results showed that out of 90 isolates, 28 isolates were resistant to macrolides, 61 to tetracycline, 15 to chloramphenicol, and 22 to clindamycin. Transposons Tn916, Tn5253, Tn6002, Tn2009, and Tn2010 were identified as the main causes of resistance. Tn916, Tn5253, and Tn6002 were distributed in several serotypes, while Tn2009 was found in serotypes 19 and 23F. Meanwhile, Tn2010 was only found in serotype 19F. Serotype 19F was the most dominant and showed the highest level of macrolide resistance. The MIC test results were mostly consistent with the genetic prediction results, although two cases of discrepancy were found: isolate MA015B carrying Tn5253 remained sensitive to tetracycline, and isolate BA0033B carrying *mefA* and *ermB* genes but was sensitive to macrolides. Transposon integration analysis showed that Tn6002 and Tn916 did not have a fixed integration pattern between isolates, while Tn5253, Tn2009, and Tn2010 showed conservative integration sites.

Tn5253 integrated in the intergenic region near the *rbGA* gene, while Tn2009 and Tn2010 were found in the intergenic region around the ORF encoding Bacteriocin-associated integral membrane protein. MDS analysis revealed that erythromycin underwent efflux through the macrolide efflux protein at 27.6 ns. Residues Asn195, Gln196, Val197, Gln198, and Ser199, located within an internal loop region, exhibited high RMSF values, suggesting notable flexibility that may be functionally relevant to the efflux process. The developed CRISPR/Cas9 targeted sequencing method successfully identified serotypes accurately, with identical results to the Illumina-based WGS method. This shows that this method has high accuracy and has the potential to be an efficient alternative method for *S. pneumoniae* serotype identification.

This study successfully revealed the presence of various types of transposons carrying antibiotic resistance genes, explained the integration pattern and efflux mechanism mediated by the *mefA* gene, and proved the accuracy of the new method for *S. pneumoniae* serotype identification based on CRISPR/Cas9 targeted sequencing on the ONT platform.

Keywords: *Streptococcus pneumoniae*, macrolides, Cas9 targeted sequencing, transposons