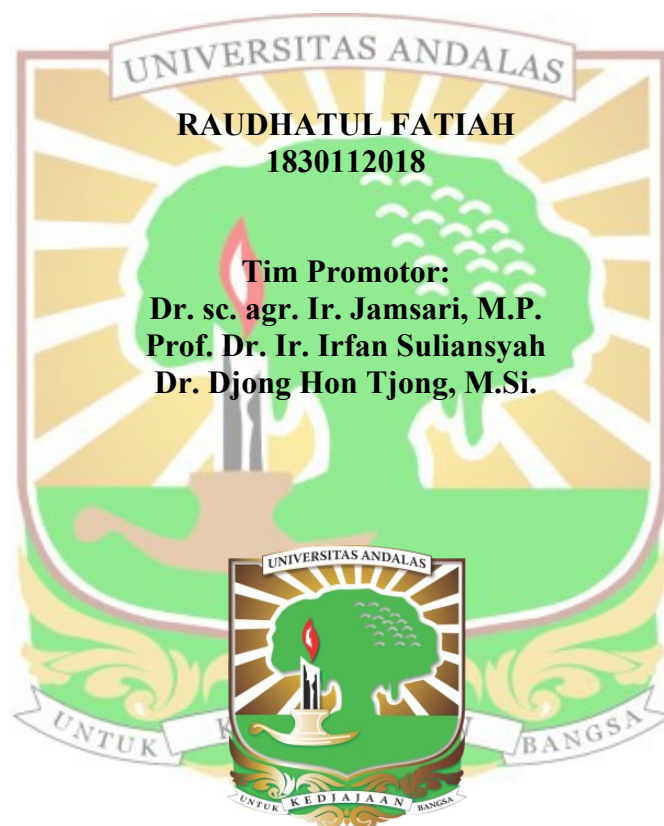


**KARAKTERISTIK GENOM DAN PEMETAAN KLASTER GEN YANG
TERLIBAT DALAM BIOSINTESIS SENYAWA METABOLIT
SEKUNDER BERBASIS *GENOME MINING* PADA
BAKTERI *Serratia plymuthica* Strain UBCF_13**

Disertasi



**RAUDHATUL FATIAH
1830112018**

**Tim Promotor:
Dr. sc. agr. Ir. Jamsari, M.P.
Prof. Dr. Ir. Irfan Suliansyah
Dr. Djong Hon Tjong, M.Si.**

**PROGRAM STUDI ILMU PERTANIAN
PROGRAM PASCASARJANA FAKULTAS PERTANIAN
UNIVERSITAS ANDALAS
2021**

**KARAKTERISTIK GENOM DAN PEMETAAN KLASTER GEN YANG
TERLIBAT DALAM BIOSINTESIS SENYAWA METABOLIT
SEKUNDER BERBASIS *GENOME MINING* PADA BAKTERI
Serratia plymuthica Strain UBCF_13**

Oleh Raudhatul Fatiah (1830112018)

(Di bawah bimbingan: Prof. Dr. sc. agr. Ir. Jamsari, M.P., Prof. Dr. Ir. Irfan
Suliansyah, M.S., dan Dr. Djong Hon Tjong, M.Si.)

RINGKASAN

Serratia plymuthica UBCF_13 adalah bakteri filoplan yang menunjukkan aktivitas antijamur. Sejumlah studi menunjukkan bahwa spesies ini memiliki potensi untuk dimanfaatkan sebagai agen biokontrol berbagai patogen tanaman seperti *Botrytis cinerea*, *Sclerotinia sclerotiorum*, dan *Colletotrichum gloeosporioides*. *S. plymuthica* memiliki mekanisme langsung dan tidak langsung untuk mengendalikan perkembangan patogen. Hal ini dapat terjadi karena peran berbagai jenis metabolit sekunder yang dapat dihasilkannya.

Selama beberapa dekade, eksplorasi senyawa metabolit sekunder dilakukan melalui pendekatan konvensional seperti *bioactivity-guided*. Keterbatasan dalam pendekatan ini adalah kebanyakan mikroba hanya memproduksi metabolit sekunder dalam jumlah yang sedikit pada kondisi yang alami. Bahkan, beberapa senyawa potensial tidak terekspresi dalam kultur di laboratorium. Oleh sebab itu, pemahaman mengenai mekanisme biosintesisnya sangat diperlukan. Pendekatan alternatif yang dapat dilakukan adalah *genome mining*. Pendekatan *Genome mining* digunakan untuk melacak klaster gene yang terlibat dalam sintesis senyawa alam (*biosynthetic gene cluster/ BGC*) dengan memanfaatkan informasi genom.

Sejauh ini genom total beberapa strain *Serratia plymuthica* telah diidentifikasi melalui *platform whole-genome sequencing*. Faktanya, masing-masing strain menunjukkan karakteristik yang berbeda, baik dari segi ukuran genom, maupun gen-gen dan senyawa-senyawa bioaktif yang dihasilkannya. Penelitian ini bertujuan untuk memperoleh sekuen lengkap genom UBCF_13, sehingga dapat digunakan untuk memahami lebih jauh karakteristik genom bakteri *S. plymuthica* UBCF_13, mendapatkan lebih banyak wawasan tentang studi evolusi, sifat-sifat unik dalam genom dan kemungkinan untuk mengeksplorasi potensi mikroorganisme ini untuk studi di masa depan.

Sekuensing genom UBCF_13 dilakukan menggunakan platform Illumina menghasilkan *reads* dengan ukuran masing-masing 150 bp total *clean reads* dua arah masing-masing berjumlah 4.484.468 *reads*. Seluruh *reads* yang diperoleh dari *whole genome sequencing* disatukan untuk memperoleh sekuen genom UBCF_13 yang utuh. Metode yang digunakan adalah modifikasi dari *map based reference* dan *de novo assembly* yaitu: kombinasi *Genes References Guided Assembly* dan *De Novo*. Hasil *reads assembly* berhasil dilakukan sehingga diketahui genom sirkular UBCF_13 berukuran 5,46 Mb. Anotasi genom UBCF_13 menggunakan Prokka mengidentifikasi 4.920 gen pengkode protein dan 118 RNA (7 5S rRNA, 7 16S rRNA, 7 23S rRNA, 82 tRNA, dan 15 ncRNA).

Genom komparasi UBCF_13 dengan genom 17 strain *S. plymuthica* lainnya menggunakan pendekatan pangenom mengidentifikasi 3.315 *core gene* (CDS

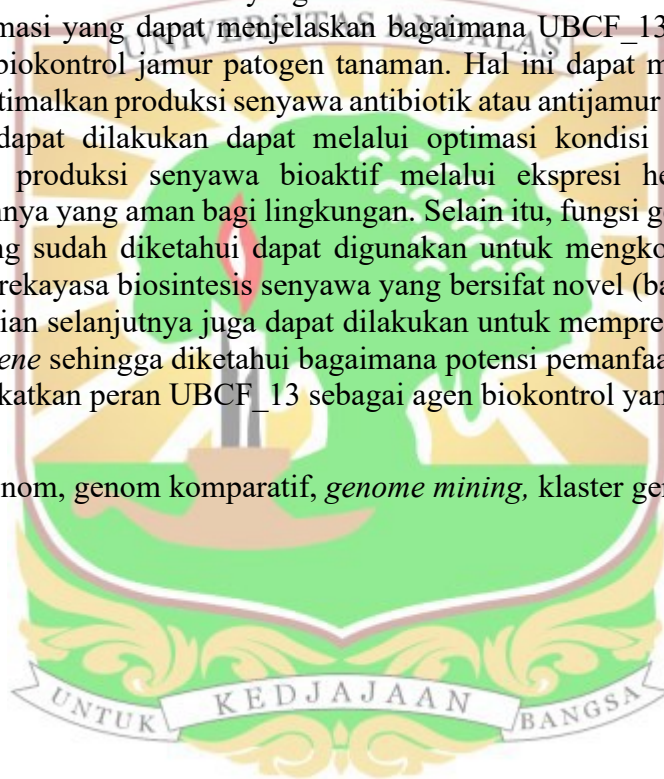
lestari). Genom UBCF_13 memiliki 1.605 *dispensable gene* (CDS selain *core gene*), dimana 300 gen diantaranya merupakan gen unik (hanya terdapat dalam UBCF_13). Analisis *Genome mining* menunjukkan bahwa UBCF_13 setidaknya memiliki kluster gen yang terlibat dalam biosintesis metabolit sekunder berupa senyawa andrimid, pirolnitrin, oocidin A, aerobactin, arylpolyene, dan BGC yang mirip dengan enterobaktin.

Dispensable gene (CDS selain *core gene*) yang ditemukan sebanyak 1.605 CDS, dimana 300 CDS diantaranya merupakan gen unik (hanya terdapat dalam UBCF_13) dapat menjadi peluang untuk ditemukannya senyawa metabolit sekunder bermanfaat selain senyawa yang diprediksi dalam penelitian ini. Upaya yang dilakukan dapat berupa melakukan identifikasi senyawa yang dihasilkan UBCF_13 dan informasi senyawa yang diketahui dapat digunakan untuk menelusuri BGC senyawa terkait.

BGC metabolit sekunder yang berhasil diidentifikasi dalam penelitian ini menjadi informasi yang dapat menjelaskan bagaimana UBCF_13 dapat berperan sebagai agen biokontrol jamur patogen tanaman. Hal ini dapat menjadi landasan untuk mengoptimalkan produksi senyawa antibiotik atau antijamur pada UBCF_13. Upaya yang dapat dilakukan dapat melalui optimasi kondisi fermentasi atau meningkatkan produksi senyawa bioaktif melalui ekspresi heterologus pada organisme lainnya yang aman bagi lingkungan. Selain itu, fungsi gen dalam genom UBCF_13 yang sudah diketahui dapat digunakan untuk mengkonstruksi biologi sintetik untuk rekayasa biosintesis senyawa yang bersifat novel (baru).

Penelitian selanjutnya juga dapat dilakukan untuk memprediksi fungsi dari *hypothetical gene* sehingga diketahui bagaimana potensi pemanfaatan gen tersebut dalam meningkatkan peran UBCF_13 sebagai agen biokontrol yang lebih optimal.

Kata kunci: genom, genom komparatif, *genome mining*, kluster gen



**GENOME CHARACTERIZATION AND MAPPING OF GENE
CLUSTERS INVOLVED IN THE BIOSYNTHESIS OF
SECONDARY METABOLITE COMPOUNDS BASED ON
GENOME MINING IN BACTERIA
Serratia plymuthica Strain UBCF_13**

By Raudhatul Fatiah (1830112018)

(Under the supervision of: Prof. Dr. sc. agr. Ir. Jamsari, M.P., Prof. Dr. Ir. Irfan
Suliansyah, M.S., and Dr. Djong Hon Tjong, M.Si.)

SUMMARY

Serratia plymuthica UBCF_13 is a phyloplan bacteria that exhibits antifungal activity. A number of studies have shown that this species has the potential to be utilized as a biocontrol agent for various plant pathogens such as *Botrytis cinerea*, *Sclerotinia sclerotiorum*, and *Colletotrichum gloeosporioides*. *S. plymuthica* has direct and indirect mechanisms to control pathogen development. This can occur due to the role of various types of secondary metabolites that it can produce.

For decades, the exploration of secondary metabolite compounds has been conducted through conventional approaches such as bioactivity-guided. Some limitations of this approach are that most microbes only produce small amounts of secondary metabolites under natural conditions. Moreover, some potential compounds are not expressed in laboratory cultures. Therefore, an understanding of their biosynthetic mechanisms is necessary. An alternative approach is genome mining. Genome mining approach is used to trace gene clusters involved in biosynthetic gene cluster (BGC) by utilizing genomic information.

The complete genomes of several strains of *S. plymuthica* have been identified through whole-genome sequencing platforms. In fact, each strain shows different characteristics, either in terms of genome size, or the genes and bioactive compounds it produces. This study aims to obtain the complete genome sequence of *S. plymuthica* UBCF_13, so that it can be used to further understand the genome characteristics of the bacterium *S. plymuthica* UBCF_13, gain more insight into evolutionary studies, unique traits in the genome and the possibility to explore the potential of this microorganism for future studies.

Sequencing of the *S. plymuthica* UBCF_13 genome was carried out using the Illumina platform resulting in reads with a size of 150 bp each, totaling 4,484,468 two-way clean reads each. All reads obtained from whole genome sequencing were put together to obtain the complete *S. plymuthica* UBCF_13 genome sequence. The method used is a modification of map-based reference and de novo assembly, namely: a combination of Genes References Guided Assembly and De Novo. The results of reads assembly were successfully carried out so that the UBCF_13 circular genome was known to be 5.46 Mb in size. Annotation of the UBCF_13 genome using Prokka identified 4,920 protein-coding genes and 118 RNAs (7 5S rRNA, 7 16S rRNA, 7 23S rRNA, 82 tRNA, and 15 ncRNA).

Comparative genomic of *S. plymuthica* UBCF_13 with the genomes of 17 other *S. plymuthica* strains using the pangenom approach identified 3,315 core genes (conserved CDS). The UBCF_13 genome has 1,605 dispensable genes (CDS other than core genes), of which 300 genes are unique (only found in UBCF_13).

Genome mining analysis shows that UBCF_13 at least has a gene cluster involved in the biosynthesis of secondary metabolites in the form of andrimid compounds, pyrrolnitrin, oocidin A, aerobactin, arylpolyene, and BGC similar to enterobactin.

Dispensable genes (CDS other than core genes) found as many as 1,605 CDS, of which 300 CDS are unique genes (only found in UBCF_13) can be an opportunity to find useful secondary metabolite compounds other than the compounds predicted in this study. Efforts can be made to identify compounds produced by UBCF_13 and information on known compounds can be used to trace the BGC of related compounds.

The secondary metabolite BGCs identified in this study provide information that can explain how UBCF_13 can act as a biocontrol agent for plant pathogenic fungi. This can be the basis for optimizing the production of antibiotic or antifungal compounds in UBCF_13. This could be done through optimizing fermentation conditions or increasing the production of bioactive compounds through heterologous expression in other organisms that are safe for the environment. In addition, the known functions of genes in the UBCF_13 genome can be used to construct synthetic biology for engineering the biosynthesis of novel compounds.

Further research can also be conducted to predict the function of the hypothetical gene so that it is known how the potential utilization of the gene in enhancing the role of UBCF_13 as a more optimal biocontrol agent.

Keywords: genome, comparative genomic, genome mining, gene cluster

