WHOLE-GENOME SEQUENCE ANALYSIS AND PROBIOTIC CHARACTERISTICS OF A *Lactococcus lactis* subsp. *lactis* STRAIN Lac3 ISOLATED FROM TRADITIONAL FERMENTED BUFFALO MILK (Dadih)

SYLVERE NSHIMIYIMANA

BIOTECHNOLOGY STUDY PROGRAM
GRADUATE SCHOOL
IPB UNIVERSITY
BOGOR
2021

STATEMENT CONCERNING THESIS AND SOURCES OF INFOMATION AND THE EXTRAORDINARY OF COPYRIGHT

I hereby declare that the thesis entitled " Whole-Genome Sequence Analysis and Probiotic Characteristics of a *Lactococcus lactis* subsp. *lactis* strain Lac3 Isolated from Traditional Fermented Buffalo Milk (Dadih)" is true of my work with direction from the supervisory commission and has not been submitted in any form to any tertiary institution.

Sources of information derived or quoted from works published or not published by other authors have been mentioned in the text and included in the Bibliography at the end of this thesis.

I hereby transfer the copyright of this writing to the IPB University and Indonesian Institute of Science (LIPI), Bogor.

Bogor, August 2021

Sylvère Nshimiyimana
P051198181

RINGKASAN

SYLVERE NSHIMIYIMANA. Whole-Genome Sequence Analysis dan Karakteristik Probiotik dari *Lactococcus lactis* subsp. *lactis* strain Lac3. Dibimbing oleh SRI BUDIARTI dan APON ZAENAL MUSTOPA.

Penelitian ini dilakukan karena data maupun penelitian mengenai whole-genome sequencing analysis dan karakterisasi secara *in vitro* pada probiotik potensial *L. lactis* subsp. *lactis* strain Lac3 masih sangat sedikit, padahal bakteri tersebut memiliki potensi sebagai bakteri probiotik.

L. lactis telah dimasukkan dalam status GRAS (*Generally Recognized As Safe*) oleh (FDA) (*The United States Food and Drug Administration*), *L. lactis* merupakan kandidat probiotik yang menjanjikan karena terbukti bermanfaat bagi hewan, pakan, farmasi, klinis, dan industri pangan. Bakteri tersebut telah digunakan sebagai *starter* pada fermentasi untuk produk-produk susu, seperti keju, yoghurt, dan sauerkraut. Selain itu, bakteri *L. lactis* dimanfaatkan di bidang klinis untuk menangani penyakit dengan mencegah infeksi patogen. *L. lactis* terbukti dapat meningkatkan kualitas pangan karena dapat memperkaya vitamin dan asam amino, serta mencegah alergi, dan sebagainya. Akan tetapi, pemilihan kandidat probiotik yang tepat harus melalui pengujian terhadap kemampuan resistensi bakteri tersebut dalam kondisi cekaman (asam, garam, basa, dan lain-lain) dan di dalam jalur pencernaan (daya tahan terhadap lingkungan tertentu) serta uji keamanan terkait dengan resistensi antibiotik, patogenisitas, dan virulensi. Berdasarkan paparan tersebut, dalam penelitian ini dilakukan karakterisasi akivitas probiotik terhadap *L. lactis* subsp. *lactis* strain Lac3, identifikasi gen fungsional untuk probiotik potensial, dan validasi data: uji *in vitro*. Hasil penelitian ini sangat diperlukan untuk memberikan informasi ilmiah dalam menyeleksi *L. lactis* subsp. *lactis* strain Lac3 sebagai probiotik potensial, terutama dari segi pangan dan Kesehatan.

Illumina MiSeq Next-generation sequencer digunakan untuk pembacaan genom berukuran ~3.8Mb, sekuens berukuran 150bp, dan dengan perkiraan ukuran genom sebesar ~7.6Mb. Penggabungan secara *de novo* dilakukan menggunakan Unicyler.v *assembler* yang terintegrasi pada web server PATRIC untuk selanjutnya menghasilkan 55 contigs. Anotasi genom ditambahkan dengan *prokaryotik genome annotation pipeline* (PGAP) yang tersedia di NCBI dan *web server Rapid Annotation under Subsystem Technology* (RAST). Karakteristik genom yang spesifik dipelajari atau dianotasikan dalam genom antara lain plasmid, fenotipe dari resistomes, prophages, faktor-faktor virulensi, CRISPR/Cas system, dan insertion sequences (ISs). Selain itu juga dipelajari mobilomes dengan berbagai sarana bioinformatika, antara lain *Comprehensive Antibiotic Resistance Database* (CARD), *Resistance Gene identifier* (RGI), ResFinder, PlasmidFinder, PHASTER, VirulenceFinder 2.0, *Mobile genetic element* (MGEFinder), dan sebagainya.

Selanjutnya dilakukan identifikasi berdasarkan 16S rRNA dengan analisis BLAST nukleotida dan *Scanning Electron Microscopy* (SEM). Pohon filogenetik dihasilkan dengan menggunakan MEGA-X, dan sekuens diurutkan menggunakan *software* MUSCLE. Bootstrap yang digunakan adalah 1000 kali, dan kekerabatan taksonomi disimpulkan dengan menggunakan metode Neighbor-joining. Selain itu, karakterisasi *in vitro L. lactis* subsp. *lactis* Lac3 dilakukan untuk menilai

karakteristik probiotik *L. lactis* subsp. *lactis* Lac3, termasuk toleransi terhadap asam, garam empedu, NaCl, auto-agregasi (sifat adhesi) dan resistensi antibiotik.

Hasil anotasi genom menunjukkan bahwa genom *L. lactis* subsp. *lactis* strain Lac3 memiliki ukuran 2411808 ~ 2.44Mb, persentase GC 34.85%, CDSs 2324, gen RNA 56, rRNA (5S(2),16S,23S), tRNA 48, ncRNA 4, dan pseudogenes 61. Anotasi spesifik menghasilkan 1 plasmid, 2 daerah prophage, 3 CRISPR arrays, dan 3 sekuens insersi terbaik (IS 3, IS150, IS6). Analisis genomik menunjukkan beberapa gen fungsional pada *L. lactis* subsp. *lactis* strain Lac3. Adanya gen fungsional tersebut membuat bakteri dapat bertahan pada kondisi cekaman, termasuk gen L-lactate dehydrogenase (EC 1.1.1.27) dan D-lactate dehydrogenase (EC 1.1.1.28). Selain itu, analisis genomik juga menunjukkan beberapa gen fungsional yang memberikan mekanisme adhesi seperti Sortase (protein permukaan yang menghambat transpeptidase), LPXTG motif, dan peregulasi pleiotropic dari eksopolisakarida (EPS) (*Ftr*). Gen-gen lainnya antara lain gen penyandi protein untuk metabolisme karbohidrat dan gliserol serta untuk biosintesis vitamin dan asam amino.

Studi ini juga memprediksi gen-gen potensial untuk probiotik, seperti colicin V (bakteriosin kelas II) dan gen yang resisten terhadap nisin dan terhadap racun logam berat. Analisis genomik menunjukkan bahwa *L. lactis* subsp. *lactis* strain Lac3 tidak bersifat patogen pada manusia. Adapun kemungkinan menjadi patogenik adalah 0.21 dan kecocokan dengan famili patogenik adalah nol. Bakteri ini tidak membawa gen yang resisten terhadap antibiotik dan tidak membawa faktor-faktor virulensi sehingga dianggap aman dan ke depannya dapat digunakan sebagai kandidat probiotik. Karakterisasi secara *in vitro* membuktikan bahwa *L. lactis* subsp. *lactis* strain Lac3 mampu bertahan lebih baik pada konsentrasi NaCl dengan kisaran antara (1, 2, 4, 5%), pH asam 2.5 dan pH 7.0. Pembentukan auto-agregasi sel meningkat dari (6.0±0.76%) ke (13.1±3.46%), artinya bahwa bakteri ini mampu melekat pada sel epitel saluran gastrointestinal. Toleransi bakteri ini terhadap garam empedu adalah sebesar (0.3, 0.5, 1%). *L. lactis* subsp. *lactis* strain Lac3 menunjukkan daya hambat pada semua antibiotik yang diujikan, kecuali antibiotik nisin dengan konsentrasi 10ng/mL. Dengan demikian, dapat disimpulkan bahwa bakteri tersebut aman karena tidak membawa gen yang resisten terhadap antibiotik (resistomes). Penelitian ini fokus pada pengujian *L. lactis* subsp. *lactis* strain Lac3 sebagai kandidat probiotik di industri pangan dan kesehatan. Penelitian secara *in vitro* ini diperlukan untuk menguji keamanan bakteri. Analisis perbandingan genom juga dibutuhkan untuk mengkarakterisasi sifat-sifat probiotik dari bakteri *L. lactis* subsp. *lactis* strain Lac3.

Kata kunci: Genome sequencing, karakterisasi *in vitro*, keamanan, patogenisitas, potensial probiotik

SUMMARY

SYLVERE NSHIMIYIMANA. Whole-Genome Sequence Analysis and Probiotic Characteristics of a *Lactococcus lactis* subsp. *lactis* strain Lac3 Isolated from Traditional Buffalo Milk (Dadih) was supervised by SRI BUDIARTI and APON ZAENAL MUSTOPA.

The intention to conduct the current study was associated with the insufficiency of data availability on the whole-genome sequencing and *in vitro* characterization of probiotic potential of *L. lactis* subsp. *lactis* strain Lac3 in order to select the *L. lactis* subsp. *lactis* Lac3 as the future probiotic candidate.

Most *L. lactis* has been included in the GRAS (*Generally Recognized As Safe*) status by (FDA) (the United States Food and Drug Administration). *L. lactis* has been considered as a promising probiotic candidate that has the versatile applications in animal, animal feeds, pharmaceuticals, clinical and food industries. *L. lactis* has been used as the starter cultures in the fermentation of dairy products, like cheese, yoghurt, sauerkraut, as well as in clinical to treat diseases by preventing pathogenic infections from the host. Moreover, *L. lactis* proved to enrich the foods into vitamins and amino acids and prevent allergy, etc. However, the selection of suitable probiotic candidate should require the assessment of its ability of tolerance against stressful conditions, such as acid, salts, bile salt tolerance and adhesion properties either in the gastrointestinal transit or survival in the niche environments, as well as the safety analysis like antibiotic resistance, pathogenicity, virulency, hemolytic activity and biogenic amines production. Besides, probiotic candidate should be evaluated for its ability to produce antimicrobial compounds. By this reason, we aimed at carrying whole-genome sequence analysis and *in vitro* characterization of probiotic potential of a *L. lactis* subsp. *lactis* Lac3. The results of this study are indispensable to provide information on genome diversity, stability, molecular evolution, safety, and probiotic potential properties which are suitable for the selection of *L. lactis* subsp. *lactis* Lac3 as the future probiotic for industrial food exploitation and health.

Illumina MiSeq Next-generation sequencer was used to produce the genome reads with the size of ~3.8Mb, each sequence had the length of 150bp, and with an estimated genome size approximately ~7.6Mb. Sequencing was processed by (Novogene Co., Ltd). The *de novo* assembly was conducted using Unicyler.v assembler algorithm integrated into the Pathosystems Resource Integration Center (PATRIC) platform, and produced 55 contigs. Genome annotation was added by prokaryotic genome annotation pipeline (PGAP) available in NCBI and Rapid Annotations under Subsystem Technology (RAST) online webserver. Specific genome features were annotated within the genome, including antibiotic resistance phenotypes (resistomes), mobilomes (prophages, insertion sequences (ISs), plasmid), virulence factors and CRISPR/Cas system. The specific annotation was done by using various online bioinformatic tools, such as Comprehensive Antibiotic resistance Database (CARD), Resistance Gene identifier (RGI), ResFinder 4.1, PlasmidFinder, PHASTER, VirulenceFinder 2.0 and Mobile genetic element (MGEsFinder).

Furthermore, identification was done based on 16S rRNA by performing nucleotide BLAST analysis and scanning electron microscopy. Phylogenetic tree was generated by using MEGA-X, and sequences were aligned by using MUSCLE. The bootstrap was inferred from 1000 replications, and taxonomic connection was inferred by using Neighbor-joining method.

Besides, *in vitro* characterization of *L. lactis* subsp. *lactis* Lac3 was conducted to assess the probiotic characteristics of *L. lactis* subsp. *lactis* Lac3, including tolerance to acid, bile salts, NaCl, auto-aggregation (adhesion properties) and antibiotic resistance.

The genome annotation predicted *L. lactis* subsp. *lactis Lac*3 to have the genome size of 2411808bp ~ 2.44Mb, with an average %GC content of 34.85%, 2324 CDSs, 56 RNA genes, (5S (2), 16S, 23S) rRNAs, 48 tRNAs, 4 ncRNAs, and 61 pseudogenes. Specific annotation consisted of 1plasmid, 2prophage regions encompassing 6 most hits phages, 3 CRISPR arrays and 3 most hits insertions sequences (IS 3, IS150, IS6). Genomic analysis revealed the ability of *L. lactis* subsp. *lactis* Lac3 to harbor several functional genes conferring its ability to survive in stressful conditions, like L-lactate dehydrogenase (EC 1.1.1.27) and D-lactate dehydrogenase (EC 1.1.1.28); several genes and associated-proteins conferring the adhesion mechanisms, such as Sortase (the surface protein anchoring transpeptidase), the LPXTG motif, and pleiotropic regulator of exopolysaccharide (EPS) (*Ftr*), etc.; numerous proteins involved in carbohydrates and glycerol metabolisms, and functional genes for vitamins and amino acids biosynthesis. We predicted also probiotic potential genes, such as colicin V as a class II bacteriocin, nisin-resistance protein and heavy metal toxic resistance. Moreover, genomic analysis revealed that *L. lactis* subsp. *lactis* Lac3 was not a human pathogen with a probability of being a human pathogen scored 0.21 and matching pathogenic families scored zero. Furthermore, *L. lactis* subsp. *lactis* Lac3 showed neither to harbor antibiotic resistance genes nor virulence factors. This confirms that it is safe to be selected as a probiotic candidate for the future application. The *in vitro* study showed the ability of *L. lactis* subsp. *lactis* Lac3 to resist in NaCl concentrations ranged between (1, 2, 4, 5%), acid (pH 2.5) and (pH7.0) and adhere to the epithelial cells of the gastrointestinal tract (GIT) with the auto-aggregation capacity that increased from (6.0±0.76%) to (13.1±3.46%). *L. lactis* subsp. *lactis* Lac3 showed tolerance against the bile salt at concentrations of (0.3, 0.5, 1%). However, *L. lactis* subsp. *lactis* Lac3 was unable to resist against all tested antibiotics. The inability of the strain Lac3 to grow in the presence of all tested antibiotics explained its safety from not carrying antibiotic resistance genes (resistomes). Dislike *L. lactis* subsp. *lactis* Lac3 was able to resist in the presence of nisin at 10ng/mL. Overall, this study highlighted the possibility to select *L. lactis* subsp. *lactis* Lac3 as the future probiotic candidate in the food industries and health. There is still a need *in vitro* investigation and carrying comparative genome analysis to maximize the probiotic properties of *L. lactis* subsp. *lactis* Lac3.

Key words: Genome sequence, *in vitro* characterization, pathogenicity, probiotic potential, safety

WHOLE-GENOME SEQUENCE ANALYSIS AND PROBIOTIC CHARACTERISTICS OF A *Lactococcus lactis* subsp. *lactis* STRAIN Lac3 ISOLATED FROM TRADITIONAL FERMENTED BUFFALO MILK (Dadih).

SYLVERE NSHIMIYIMANA

Thesis
Submitted in partial fulfillment for the award of
Master of Science (M.Sc.) degree
in
Biotechnology (PS-BTK)

BIOTECHNOLOGY STUDY PROGRAM
GRADUATE SCHOOL
IPB UNIVERSITY
BOGOR
2021

External examiner:
Dr. Ir. Aris Tjahjoleksono, DEA

Thesis Title : Whole-Genome Sequence Analysis and Probiotic Characteristics of a *Lactococcus lactis* subsp. *lactis* strain Lac3 Isolated from Traditional Fermented Buffalo Milk (Dadih)

Name : Sylvère Nshimiyimana

Nim : P051198181

Approved by
Supervisory committee,

Head supervisor:
Prof. Dr. dr.Sri Budiarti

Digitally signed by:
Sri Budiarti
[25857D90AHC03E3]
Date: 2 Sep 2021 05.04.45 WIB
Verify at disign.ipb.ac.id

Co-supervisor:
Dr. Apon Zaenal Mustopa, MSc.

Endorsed by,

Head of study program of biotechnology:
Prof. Dr. Ir. Sony Suharsono, DEA

Digitally signed by:
Suharsono
[035AAF98305014DE]
Date: 7 Sep 2021 20.46.28 WIB
Verify at disign.ipb.ac.id

Dean of graduate school:
Prof. Dr. Ir. Anas Miftah Fauzi, M.eng

Digitally signed by:
Anas Miftah Fauzi
[3FE49AA995DD0C4F]
Date: 8 Sep 2021 10.25.17 WIB
Verify at disign.ipb.ac.id

Date of Exam: 30th July 2021

Graduation date:

FOREWORD

Glory be to Almighty God by whose grace all righteousness is accomplished. My special thanks to my Mum and Brothers and all members of my beloved family. I specially thank the Government of Indonesia for awarding me the scholarship (KNB) to study the master degree in Indonesia. Special recognition for the program study of biotechnology who let me feel as a member of that beloved family, as well as Lembaga Ilmu Pengetahuan Indonesia (LIPI) for providing me the opportunity to conduct my lab activities. Warm recognition to DIPA Prioritas Nasional Lembaga Ilmu Pengetahuan Indonesia (LIPI, 2021), for supporting this research. Then, I thank and commend my revered supervisors, Prof. Dr. dr. Sri Budiarti, Dr. Apon Zaenal Mustopa for their relentless efforts of guidance and advice throughout. I also appreciate the love and kindness from all my Laboratory assistants, especially Lita Meillina, Shasmita Irawan, Anika, Lita Triratna for providing me guidance. I thank all Indonesians whom I have met at any point during my studies in IPB University. And finally, great thankfulness for the best university in Indonesia, IPB University.

Bogor, August 2021

Sylvère Nshimiyimana

LIST OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES