INTEGRATING GENOMICS AND TRANSCRIPTOMICS TO UNDERSTAND THE VIRULENCE AND BIOFILM FORMING MECHANISM OF SELECTED VANCOMYCIN-RESISTANT *Enterococcus faecium*

LIM SHU YONG

FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

INTEGRATING GENOMICS AND TRANSCRIPTOMICS TO UNDERSTAND THE VIRULENCE AND BIOFILM FORMING MECHANISM OF SELECTED VANCOMYCIN-RESISTANT *Enterococcus faecium*

LIM SHU YONG

DESSERTATION SUBMITTED IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE

INSTITUTE OF BIOLOGICAL SCIENCES
FACULTY OF SCIENCE
UNIVERSITY OF MALAYA
KUALA LUMPUR

2018

UNIVERSITY OF MALAYA

ORIGINAL LITERARY WORK DECLARATION

Name of Candidate: **LIM SHU YONG**

Registration/Matric No: **SGR140051**

Name of Degree: **MASTER OF SCIENCE**

Title of Project Paper/Research Report/Dissertation/Thesis ("this Work"):

INTEGRATING GENOMICS AND TRANSCRIPTOMICS TO UNDERSTAND THE VIRULENCE AND BIOFILM FORMING MECHANISM OF SELECTED VANCOMYCIN-RESISTANT *Enterococcus faecium*

Field of Study:

MICROBIOLOGY

I do solemnly and sincerely declare that:

(1) I am the sole author/writer of this Work;
(2) This Work is original;
(3) Any use of any work in which copyright exists was done by way of fair dealing and for permitted purposes and any excerpt or extract from, or reference to or reproduction of any copyright work has been disclosed expressly and sufficiently and the title of the Work and its authorship have been acknowledged in this Work;
(4) I do not have any actual knowledge nor do I ought reasonably to know that the making of this work constitutes an infringement of any copyright work;
(5) I hereby assign all and every rights in the copyright to this Work to the University of Malaya ("UM"), who henceforth shall be owner of the copyright in this Work and that any reproduction or use in any form or by any means whatsoever is prohibited without the written consent of UM having been first had and obtained;
(6) I am fully aware that if in the course of making this Work I have infringed any copyright whether intentionally or otherwise, I may be subject to legal action or any other action as may be determined by UM.

Candidate's Signature: Date:

Subscribed and solemnly declared before,

Witness's Signature: Date:

Name: **PROFESSOR DR THONG KWAI LIN**

Designation:

INTEGRATING GENOMICS AND TRANSCRIPTOMICS TO UNDERSTAND THE VIRULENCE AND BIOFILM FORMING MECHANISM OF SELECTED VANCOMYCIN-RESISTANT *Enterococcus faecium*

ABSTRACT

Vancomycin-resistant enterococcus (VRE) is an emerging nosocomial pathogen which causes outbreaks in hospitals worldwide. It is, therefore, important to understand the virulence and how this organism persists as a nosocomial pathogen. Whole genome sequencing (WGS) provides a wealth of information to elucidate the genetic relationship, virulence potential, and resistance factors of VRE. However, such genomic studies are lacking in Malaysia. Moreover, since enterococci are often recovered from difficult-to-treat biofilm-mediated infections, a detailed study on the biofilm formed by VRE is useful to better understand its pathogenicity. The objectives of this study are to perform comparative genomics analysis on the four local vancomycin-resistant *Enterococcus faecium* (VRE_{fm}) and to elucidate the transcriptomic profile of biofilm cells in respect to the planktonic cells. Four VRE_{fm} strains were isolated from two fatal cases of nosocomial infections in a tertiary hospital in Kuala Lumpur. One of these strains (VRE2) was isolated from an index case (patient X), whereas the other three (VREr5, VREr6, VREr7) were isolated from different body sites of another patient (patient Y) at around one-week interval. WGS and comparative genomics analyses revealed that the four strains have different sequence types (STs), ST80 and ST203. Subsequent phylogenomic study showed that VREr5 was more closely related to VRE2, but was distantly related to VREr6 and VREr7 derived from the same host. Moreover, the genomic contents of VREr5 was also more similar to VRE2. The genomic data and clinical records suggested that patient Y was most probably infected by multiple strains of different clones. Alternatively, the strain that infected patient Y (VREr5) could have derived from the same clone from

patient X (VRE2), given their high genomic similarity. The four local strains were multidrug resistant. All of them carried the *vanA* genotype and showed indistinguishable Tn*1546* structure. Virulence profiling revealed that these strains harbored a total of 13 virulence genes mainly associated with adherence and biofilm formation. The transcriptomic analysis focused on the initial stage of biofilm formation to examine genes that are involved during the transition from planktonic to biofilm cells. Differential gene expression analysis revealed that the up-regulated genes in biofilm cells involved mainly in adherence, plasmid replication, and carbohydrate metabolism. Genes that have been reported to negatively regulate biofilm formation, such as the quorum sensing systems, *fsr* and *luxS*, and a transcriptional regulator gene *spx*, were highly down-regulated. The unique *bee* homolog of VREr5 was found to be down-regulated, implying a negative association of this locus to biofilm formation in VREr5. The results obtained from the gene expression study clearly reflected the attachment stage of biofilm development, including the preparation to enter the maturation stage. In conclusion, this study has contributed to the understanding of the genetic basis and diversity of local clinical strains which can be helpful to control the spread of VRE. This study also provides insight into the molecular mechanism of biofilm formation in VRE which might be useful in the development of new drugs.

Keywords: Vancomycin-resistant enterococcus, comparative genomics, transcriptomics

INTEGRASI GENOMIK DAN TRANSKRIPTOMIK DALAM MEMAHAMI KEVIRULENAN DAN PEMBENTUKAN BIOFILEM DALAM *Enterococcus faecium* TERPILIH

ABSTRAK

Enterococcus yang resistan terhadap vancomycin (vancomycin-resistant enterococcus atau VRE) adalah patogen nosokomial yang menyebabkan wabak di hospital-hospital di seluruh dunia. Oleh itu, pengetahuan mengenai kevirulenan dan bagaimana organisma ini tegar sebagai patogen nosokomial adalah penting. Penjujukan Genome Kesuluruhan (whole genome sequencing, WGS) menyediakan pelbagai maklumat bagi mengkaji hubungan genetik, potensi virulens, dan faktor-faktor resistan VRE. Namun begitu, kajian genomik tersebut masih kurang di Malaysia. Selain itu, enterococci sering diperolehi daripada jangkitan berkaitan dengan biofilem yang sukar untuk dirawat. Oleh itu, satu kajian terperinci mengenai biofilem yang dibentuk oleh VRE adalah berguna bagi lebih memahami pathogeniciti organisma ini. Objektif-objektif kajian ini adalah untuk melaksanakan analisis genomik perbandingan kepada empat *Enterococcus faecium* tempatan yang resistan terhadap vancomycin (VRE_{fm}) dan untuk mengkaji profil transkriptomik sel-sel biofilem yang berkait dengan sel-sel plankton. Keempat-empat strain yang dikaji diperolehi daripada dua kes maut jangkitan nosokomial di hospital pengajian tinggi di Kuala Lumpur. Salah satu strain yang dikaji (VRE2) diperolehi daripada kes indeks (pesakit X) manakala tiga strain lain (VREr5, VREr6, VREr7) diperolehi daripada pesakit tunggal (pesakit Y), daripada bahagian badan yang berbeza sekitar selang satu minggu. WGS dan analisis perbandingan genomik menunjukkan bahawa keempat-empat strain mempunyai pelbagai jenis jujukan (ST) iaitu ST80 dan ST203. Kajian phylogenomic seterusnya menunjukkan bahawa VREr5 lebih berkait rapat dengan VRE2 tetapi kurang berkait dengan VREr6 dan VREr7 yang

berasal daripada pesakit yang sama. Selain itu, kandungan genom VREr5 juga lebih serupa dengan VRE2. Data genomik dan rekod klinikal mencadangkan bahawa pesakit Y mungkin dijangkiti oleh pelbagai jenis klon yang berbeza. Sebagai alternatif, strain yang menjangkiti pesakit Y (VREr5) boleh diperolehi daripada klon yang sama seperti strain pesakit X (VRE2) memandangkan persamaan genomik yang tinggi. Semua strain yang dikaji mempunyai resistan terhadap pelbagai ubat. Mereka membawa genotip *vanA* dan mempunyai struktur Tn*1546* yang sama. Profil virulens mendedahkan bahawa strain ini menangandungi tiga belas gen virulens yang berkaitan dengan perekatan dan pembentukan biofilem. Analisis transkriptomik berfokus kepada peringkat awal pembentukan biofilem bagi mengkaji gen-gen yang terlibat semasa peralihan dari plankton kepada biofilem. Analisis perbezaan expresi gen mendedahkan penglibatan gen "up-regulated" dalam perekatan, replikasi plasmid, dan metabolisme karbohidrat dalam sel-sel biofilem. Gen-gen yang mempunyai pegawalan negatif terhadap pembentukan biofilem, seperti sistem kuorum sensing, *fsr* dan *luxS* dan transkripsi gen pengatur *spx*, adalah "down-regulated". Di samping itu, homolog *bee* telah didapati "down-regulated", mencadangkan bahawa locus ini mungkin tidak berkaitan dengan pembentukan biofilem bagi VREr5. Keputusan yang diperolehi daripada kajian transkriptomik menunjukkan peringkat perekatan pembentukan biofilem termasuk persediaannya untuk memasuki peringkat matang.

Secara ringkasnya, kajian ini telah menyumbang kepada pemahaman asas genetik dan kepelbagaian genetik dalam strain klinikal tempatan. Pengetahuan ini boleh digunakan untuk mengawal penyebaran VRE. Kajian ini juga menyumbang kepada pengetahuan baru berkenaan mekanisme molekul pembentukan biofilem di dalam VRE yang mungkin berguna untuk pembuatan ubat baru.

Kata Kunci: Enterococcus yang resistan terhadap vancomycin, genomik perbandingan, transkriptomik

ACKNOWLEDGEMENTS

I would like to thank University of Malaya (UM), High Impact Research (HIR) grant, and Postgraduate Research Fund (PPP) for funding this research. Thanks to Institute of Biological Sciences, Faculty of Sciences, Dean of faculty and fellow members for approving this postgraduate study. I am extremely grateful to my supervisor and co-supervisor, Professor Dr. Thong Kwai Lin and Dr. Cindy Teh Shuan Ju for their supervision, guidance, and invaluable advice throughout this period. Also, I am grateful to Dr. Kartini Abdul Jabar who helped by providing clinical data and guidance in interpreting the data. To my parents who always support and encourage me, I offer my deepest thanks and appreciation. My sincere thanks to my seniors, Tony and Wing Sze, for their support and advice for completing this study. To my research buddies in lab A407 --- Hui Key, Shiang Chiet, Hannah, Soo Tein, Soo Ling, Ah Loong, Xiu Pei, Pei Yee, Sarah, Anis --- thanks for giving me a memorable postgraduate experience.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS AND ABBREVIATIONS

AST	:	Antibiotic susceptibility test
BLAST	:	Basic Local Alignment System Tool
BRIG	:	BLAST Ring Image Generator
CA	:	Community-associated
CC17	:	Clonal complex 17
COG	:	Cluster of Orthologous Group
CSF	:	Cerebrospinal fluid
CLSM	:	Confocal laser screen microscopy
CLSI	:	Clinical and Laboratory Standard Institute
CDS	:	Coding sequences
CRISPR	:	Clustered Regularly Interspaced Short Palindromic Repeat
℃	:	Degree celcius
EVD	:	Ventricular drain insertions
GC	:	Guanine-cytosine
GO	:	Gene Ontology
HA	:	Hospital-associated
hr	:	Hour
I	:	Intermediate
IS	:	Insertion sequence
KEGG	:	Kyoto encyclopaedia of genes and genomes
M	:	Mega
MLST	:	Multilocus sequence typing
MIC	:	Minimum inhibitory concentration
ml	:	Milliliter

mg/ml	:	Milligram per milliliter
mM	:	Milli Molar
min	:	Minute
µl	:	Micro liter
µg/ml	:	Micrograms per milliliter
MSCRAMM	:	Microbial surface components recognizing adhesive matrix molecules
NCBI	:	National Center of Biotechnology Information
NGS	:	Next generation sequencing
ORF	:	Open reading frame
PAI	:	Pathogenicity island
PBP5	:	Penicillin-binding protein 5
PCR	:	Polymerase chain reaction
PCA	:	Principal component analysis
PFGE	:	Pulsed-field gel electrophoresis
RAST	:	Rapid Annotation using Subsystem Technology
RNA-seq	:	RNA sequencing
RealPhy	:	Reference Sequence Alignment-based Phylogenic Builder
R	:	Resistant
S	:	Susceptible
ST	:	Sequence type
UMMC	:	University Malaya Medical Centre
VRE	:	Vancomycin-resistant enterococci
VRE_{fm}	:	Vancomycin-resistant *Enterococcus faecium*
VFDB	:	Virulence Factors of Pathogenic Bacteria Database
WGS	:	Whole genome sequencing

LIST OF APPENDICES

CHAPTER 1: INTRODUCTION

1.1 General Introduction

Enterococcus faecium has been associated with several serious or life-threatening nosocomial diseases such as urinary tract infections, surgical-related wound infections, bacteremia, and endocarditis (Higuita & Huycke, 2014). As an opportunistic pathogen, *E. faecium* mainly targets elderly patients with underlying diseases, immunocompromised patients, and patients who have been hospitalized for prolonged periods or treated with invasive devices. The clinical significance of *E faecium* becomes more prominent with the increasing antimicrobial resistance among the clinical isolates, including high-level resistance to ampicillin, aminoglycosides, and glycopeptides (Cattoir & Giard, 2014; Leclercq et al., 1988; Padmasini et al., 2014; Uttley et al., 1988). The rapid spread of vancomycin resistant *E. faecium* (VRE_{fm}) is of particular concern as VRE is often multidrug resistant. According to a recent report from the National Healthcare Safety Network (NHSN), 82.2% and 85.1% of enterococci recovered from bloodstream and urinary tract infections are resistant to vancomycin in the year 2014 (Sievert et al., 2016). The emergence of VRE and outbreaks that occur around the world indicate the success of *E. faecium* in adapting to and surviving in the hospital environment. Worryingly, resistance to antibiotics that are used to treat VRE infections, such as linezolid, daptomycin, and tigecycline, has been reported (Edelsberg et al., 2014; Tsai et al., 2012).

With the rapid advances in next generation sequencing (NGS) technology, different NGS platforms such as Roche/*454*, SoLiD, and Illumina enable the entire genomes of bacterial pathogen including *E. faecium* to be sequenced (Lam et al., 2013, 2012; Qin et al., 2012). This technology generates massive information that is helpful in exploring the differences between strains of *E. faecium* at a genome-wide level. The whole genome sequences obtained can, therefore, provide an in-depth understanding of the genetic

contents, resistance determinants and mechanisms, as well as pathogenicity and evolution of *E. faecium*. This information is important in revealing factors that contribute to its adaptation and persistence in the clinical settings, and, hopefully, solutions to control and treat infections. To date, a number of genome sequences of *E. faecium* have been reported (García-Solache & Rice, 2016; Khan et al., 2015; Lam et al., 2013, 2012; Qin et al., 2012). Genome analyses revealed that clinical isolates are different from non-clinical isolates in that the number of mobile genetic elements, resistance and virulence genes are significantly higher in the genomes of clinical isolates (Kim & Marco, 2014; Qin et al., 2012). Apart from that, specific elements such as the *esp* gene and IS*16* are found almost exclusively in the hospital-associated isolates (Heikens et al., 2012; Leendertse et al., 2009; Werner et al., 2011; Willems et al., 2001). The findings from these genomic analyses have shed light on the virulence and persistence of clinical *E. faecium* which is important in controlling the spread of this nosocomial pathogen.

The clinical relevance of infections caused by VRE_{fm} can also be attributed to the difficult-to-treat biofilm-associated diseases. It has been reported that majority of the device-associated infections, such as infections due to central lines, urinary catheters, and ventilators, are caused by VRE_{fm} (Sievert et al., 2016). Since biofilms are highly resistant to antibiotics and phagocytosis, a detailed insight into the process and molecular mechanism of biofilm formation is pivotal for the development of new drugs against biofilm-mediated infections. Transcriptomic analysis enables the quantification of gene expression of a full transcriptome, thereby help in interpreting the functional elements of a genome that are expressed at a specific physiological condition or developmental stage (Wang et al., 2009). The introduction of RNA-seq technology provides a more sensitive and dynamics approach to study the transcriptome of various organisms as compared to the traditional hybridization method (microarray) (Hinton et al., 2004; Wang et al., 2009). In the case of biofilm formation, this high-throughput technology has generated

informative data which is helpful in understanding the molecular mechanisms of biofilm biogenesis that are important for drugs development (Rumbo-Feal et al., 2013; Tan et al., 2015).

1.2 Objectives

To the best of our knowledge, a complete genome analysis of clinical VRE_{fm} has not been reported in Malaysia. Hence, such study is needed to better understand the overall biology, and the potential resistance and virulence of the locally isolated VRE_{fm} strains. Moreover, transcriptomic studies on the biofilm formation of *E. faecium* is also lacking in Malaysia or Southeast Asia. Understanding the transcriptome of biofilm cells is important in the development of new control or treatment methods against biofilm-mediated infections.

Therefore, the objectives of this study were:

1. To perform comparative genome analyses of four selected VRE_{fm} from Malaysia;

2. To elucidate the gene expression profile of biofilm cells in respect to the planktonic cells.

CHAPTER 2: LITERATURE REVIEW

2.1 The Genus *Enterococcus*

The genus *Enterococcus* was previously classified as part of the genus *Streptococcus*. It is not until 1984 when the *Enterococcus* was recognized as a separate genus (Schleifer & Kilpper-Balz, 1984). *Streptococcus faecalis* and *Streptococcus faecium* were the first two species to be transferred to the new genus as *Enterococcus faecalis* and *Enterococcus faecium*. To date, there are more than 50 species of enterococcus being described [http://www.bacterio.net/index.html].

Enterococci compose of Gram-positive cocci that are often arranged in pairs or chains. They are catalase-negative, non-spore-forming facultative anaerobes which can survive in harsh conditions, including high salinity (6.5% NaCl) and a wide range of temperature (10°C to 45°C) (Facklam & Collins, 1989). As such, enterococci are ubiquitous in nature. They can be found in water, soil, plants, as well as fermented food and dairy products. Enterococci are also commensals in the gastrointestinal tracts of human and animals. Although accounted for only 1% of the human gut microflora (Sghir et al., 2000), enterococci can occasionally cause diseases such as urinary tract infection and endocarditis, especially in immunocompromised patients.

2.2 Emergence of vancomycin-resistant enterococci (VRE)

Enterococci have long been considered as harmless inhabitants of the human gut flora. However, in the past few decades, this organism has emerged as one of the leading cause of hospital-associated infections. The extensive use of antibiotics in the clinical settings has contributed remarkably to the transition of this organism from commensal to the

nosocomial pathogen. Following the development of resistance to beta-lactam drugs and to high concentration aminoglycosides in the 1980s, vancomycin was among the last available antibiotic for enterococcal infections.

Vancomycin acts by disrupting cell wall synthesis. Once enterococci get in contact with the compound, vancomycin binds to the D-ala-D-ala terminus of the peptidoglycan precursor, thereby inhibiting cell wall development (Courvalin, 2006). Due to its disruptive effect on peptidoglycan, vancomycin is commonly used together with other antibiotics such as aminoglycosides to effectively get rid of the pathogen. Although the use of vancomycin has successfully controlled enterococcal infections, it was later discovered that the acquisition of vancomycin resistance has become increasingly prevalent (Bonten et al., 2001; Kuo et al., 2014).

Vancomycin-resistant enterococci (VRE) was first reported in Europe in the 1980s (Leclercq et al., 1988; Uttley et al., 1988). Since then, VRE has spread worldwide, including in the United States, Europe, and East Asia regions (Bonten et al., 2001; Kuo et al., 2014). In Malaysia, the first confirmed case of hospital acquired VRE was reported in 2006 in Hospital Kuala Lumpur (HKL) in a patient with chronic renal failure (Zubaidah et al., 2006). Other local studies reported a low prevalence of VRE (1-2%) in Malaysia (Ibrahim et al., 2010, 2011). A study collecting enterococci from different countries shows that vancomycin resistance was more prevalent in *E. faecium* (Putnam et al., 2010). A similar result was observed by Edelsberg et al. (2014) who examined resistance percentage of clinically significant bacterial pathogens in 19 US hospitals. The higher rate of vancomycin resistance in *E. faecium* indicates the increasing clinical importance of this organism over the other enterococcal species.

The use of avoparcin had been suggested to contribute to the dissemination of vancomycin resistance in *E. faecium*. This correlation can be shown through the different epidemiology of VRE in Europe and the United States. In Europe, where avoparcin was massively used as a growth promoter in animal farms (Wegener, 1998), VRE are more prevalent in livestock and healthy people compared to patients (Devriese et al., 1996; Stobberingh et al., 1999; Van Braak et al., 1998). However, this type of community-reservoir in animals and healthy people is not observed in the United States, where avoparcin is banned in the animal farms. In contrast, VRE are the second most common nosocomial pathogen in the United States (Sievert et al., 2016). Since avoparcin can confer cross-resistant to vancomycin (Bager et al., 1997), the extensive use of this antibiotic in the animal farms may select for vancomycin resistant strains. Ultimately, these resistant strains could be transmitted to human through the food chain or to the farm workers due to poor hygiene practice in the animal farms. The possible association of avoparcin and the dissemination of vancomycin resistance was further strengthened by the reduction of VRE rate in Europe after the ban of avoparcin in 1997 (Aarestrup et al., 2001; Boggard & Stobberingh, 2000).

There is evidence indicating that glycopeptide-producing microorganisms, such as *Amycolatopsis orientalis* and *Streptomyces toyocaensis*, are the sources of vancomycin resistance genes (Marshall et al., 1997). The production of these resistance genes is presumed as a self-defence mechanism of these organisms. These resistance genes then can be transferred horizontally via one or more bacterial intermediates, ultimately to enterococci. In the case of Europe, resistance gene may ultimately be transferred to genogroups that colonize animals and healthy people, resulted in a large community-based reservoir. On the contrary, in the United States, vancomycin resistance gene may be transferred to genogroups which had been resistant to multiple antibiotics, thereby

increasing the chance of acquired resistance due to the selective effects of antibiotics (Bonten et al., 2001).

2.3 Molecular subtyping of *E. faecium*

For many years, majority of the human enterococcal infections were caused by *E. faecalis* (80-90%) (Jones et al., 2004; Moellering, 1992). However, in the past three decades, *E. faecium* has increasingly become predominant as a leading cause of hospital-acquired infections, particularly vancomycin-resistant *E. faecium* (VRE_{fm}). As such, molecular subtyping is important for epidemiological studies to assist in infection control as well as to trace the dissemination of this pathogen. Various subtyping methods have been developed for *E. faecium*. These include repetitive element sequence (REP)-PCR typing, plasmid typing, pulsed-field gel electrophoresis (PFGE), amplified fragment length polymorphism (AFLP), multiple locus variable number of tandem repeat analysis (MLVA), and vancomycin resistance gene cluster typing. However, each of these typing methods has its own drawbacks, mainly in terms of the ease of use, discriminatory power, cost, data reproducibility, as well as data exchange (Werner, 2013).

PFGE is a highly discriminative typing method that can distinguish strains of the same species from different sources, time, and space. This macrorestriction-based typing method utilizes rare-cutting enzymes to fragment chromosomal DNA, and an alternating electric field (i.e. pulsed field) to separate the resulting DNA fragments, generating distinct patterns which can be used to differentiate closely related bacteria. Due to its high discriminatory power, PFGE remains a "gold standard" for epidemiological studies of a large number of bacteria, including *E. faecium*. However, since the rate of recombination is relatively high in enterococci (Willems et al., 2005), extreme genomic variations may be found in strains from the same outbreak over a period of time. As such, PFGE is less suitable for long-term epidemiological study of enterococci. Apart from that, inter-

laboratory transfer of PFGE data is challenging as a standardized protocol of enterococcal PFGE has not been established.

Multi-locus sequence typing (MLST) is a recently widely accepted typing method for population and evolutionary study (Urwin & Maiden, 2003). This method determines the alleles of the internal fragment sequences of multiple housekeeping genes and assigns each allele a numerical value. In *E. faecium*, seven housekeeping genes, *atpA*, *ddl*, *gdh*, *purK*, *gyd*, *pstS*, and *adk*, are used (Homan et al., 2002). The combination of the seven allelic numbers yields a sequence type (ST) that classifies different strains. MLST allows for the exchange of data throughout the globe through the use of uniform typing method and the establishment of a large, publicly accessible database. However, this method can be relatively costly if sequencing of the seven housekeeping genes of a large number of samples is to be performed.

A distinct clonal lineage comprised of mostly clinical *E. faecium* isolates was identified when the MLST data was analyzed using the enhanced based upon related sequence types (eBURST) algorithm (Willems et al., 2005). This lineage is termed clonal complex 17 (CC17). Members of the CC17 are characterized by ampicillin resistance, a pathogenicity island and an association with hospital outbreak (Willems et al., 2005). Shortly, this lineage turns into the major genetic complex associated with hospital infections and is found to be disseminated around the world (Matsushima et al., 2012; Ryan et al., 2015; Valdezate et al., 2009; Yu et al., 2015).

2.4 Whole genome sequencing of *E. faecium*

The increased medical importance of *E. faecium* over *E. faecalis* as a nosocomial pathogen drives interest into the epidemiology and pathogenicity of this organism.

However, there are still gaps in our understanding of its virulence, pathogenicity, resistance mechanisms, and even adaptability. The availability of whole genome sequencing (WGS) allows scientists to explore in more detail the biological features of *E. faecium* to provide answers to the emergence of this pathogen.

The first complete genome of *E. faecium* was available in March 2012 from a vancomycin-resistant strain Aus0004. Aus0004 was isolated from the bloodstream of a patient in Melbourne, Australia, in 1998 (Lam et al., 2012). Following this, the second complete genome was published by (Qin et al., 2012) and was determined by strain TX16 (also named TX0016 or DO) isolated from the blood of an endocarditis patient in 1992. More genomes were later being published, which provide chances for comparative genomics to unveil the genomic variations among strains from different niches.

There are currently more than 400 publicly available *E. faecium* genomes, with 17 genomes being completely sequenced (https://www.ncbi.nlm.nih.gov/genome/genomes/871?). However, an overrepresentation of strains from Europe and North America is observed. Hence, more genomes from strains of various regions, such as from the Southeast Asia, are needed to fully understand the global diversity of *E. faecium*.

2.5 Comparative genomics

2.5.1 Core and accessory genome

Genome comparison between several *E. faecium* strains has identified a set of genes known as "core" genome that are conserved in all strains. This core genome consists of genes that are mainly responsible for housekeeping functions such as DNA and carbon

metabolism, cell structural biosynthesis, and substrate transport. The non-core genes make up the "accessory" genome which are partially shared or strain-specific. These accessory genes may contribute to specific adaptive traits or even pathogenicity of different strains from different environmental niches. For instance, comparative analyses have revealed differences in genome content between the clinical and non-clinical strains. The genome of clinical strains composes of significantly more mobile genetic elements, virulence factors and antibiotic resistance genes compared to that of non-clinical strains (Kim & Marco, 2014). The *esp* gene, encodes for the enterococcal surface protein (Esp) which is associated with endocarditis and urinary tract infections, is specifically enriched in the clinical strains (Heikens et al., 2012; Leendertse et al., 2009; Willems et al., 2001). Moreover, acquired antibiotic resistance genes, such as *van* gene that confers vancomycin resistance, are also part of the accessory genome of clinical *E. faecium* strains.

Several studies suggested that the pan-genome of *E. faecium* is open, indicating that this organism is able to acquire and incorporate novel genes efficiently into the collective gene pool of the species (Qin et al., 2012). This helps to explain the accumulation of a large number of accessory genes in the clinical strains that contribute to their pathogenicity. An open pan-genome also implies that *E. faecium* can readily acquire new genes that contribute to fitness, allowing it to adapt to various environmental niches, which might facilitate its transition from commensal to nosocomial pathogen.

2.5.2 Insertion sequence, phage, and CRISPR

Insertion sequence (IS) elements are short DNA fragments (0.7-2.5kb) that usually encode for one or two genes needed for transposition (Siguier et al., 2015). These elements are widely spread in bacterial genomes and are abundantly found in the genome of *E.*

faecium (Lam et al., 2013, 2012; Mikalsen et al., 2015; Qin et al., 2012). IS elements can play a vital role in evolution and diversification of genomes. IS*16*, for example, is almost exclusively present in the hospital-associated (HA) strains and was suggested as a marker for identification of HA *E. faecium* strains (Werner et al., 2011). Furthermore, transposition of IS elements can also lead to gene rearrangement such as duplication, deletion, and inversion, which contributes to genome plasticity among the species (Lee et al., 2016; Ooka et al., 2009). Apart from that, there have been evidence showing that IS elements can affect gene expression (Aubert et al., 2003; Sóki et al., 2013). In *E. faecium*, for example, insertion of IS elements in the vancomycin resistance gene cluster has led to reduced resistance in certain VRE strains (Gagnon et al., 2011).

Bacteriophage (phage) is another mobile genetic element which can also drive genome diversification. Phages integrated into the bacterial genomes (prophages) can introduce new genes that modify the genome content of strains within the species. A number of prophages have been identified in the genome of *E. faecium* (Lam et al., 2013, 2012; Qin et al., 2012; van Schaik et al., 2010). These prophages are distinct from strain to strain, demonstrating phage diversity of *E. faecium* (Lam et al., 2013). In a comparison between seven *E. faecium* strains, van Schaik et al. (2010) reported that most of the phages identified can be activated by the DNA cross-linking agent mitomycin. These induced phages had the typical morphology of *Siphoviridae,* which is a common bacteriophage family found in lactic acid bacteria. It remains to be determined to what extent prophages contribute to the fitness of *E. faecium* but the variety of prophages identified suggest that they play a major role in shaping the genome of *E. faecium*. Nevertheless, phages may aid in the transmission of antibiotic resistance genes among enterococci species, as reported by Mazaheri Nezhad Fard and his co-workers (Mazaheri Nezhad Fard et al., 2011).

Clustered regularly interspaced short palindromic repeats (CRISPR) are short repetitive DNA sequences that are found in both archaea and bacteria. These repeats are characterized by direct repeats of varying size interspaced by non-repetitive sequences of a similarly size known as "spacer" and flanked on one side by a conserved leader sequence (Jansen et al., 2002). The CRISPR loci are usually associated with genes known as CRISPR-associated (*cas*) genes which encode for nucleases or proteins involved in DNA and RNA processing. This CRISPR-*cas* system has been shown to provide bacteria with adaptive immunity against the integration of foreign genetic materials such as phages and plasmids (Marraffini & Sontheimer, 2011). CRISPR-*cas* system appears to be found less frequently in the HA strains, which are usually multidrug resistant (Palmer & Gilmore, 2010). This observation is consistent with the defensive role of CRISPR-*cas* system to limit the acquisition of mobile genetic elements which might carry antibiotic resistance genes. The absence of CRISPR-*cas* system in HA strains also explains the abundance of phages and other mobile elements in their genomes, highlighting the lack of barriers to horizontal gene exchange in the species.

2.5.3 Virulence factors

Comparison between the genomes of hospital-associated (HA) and community-associated (CA) strains reveals several virulence factors that are enriched in the HA strains. The most significant one being the presence of *esp* gene which is found more abundantly in the clinical isolates than in food or environmental isolates (Abriouel et al., 2008; Willems et al., 2001). The *esp* gene was initially suggested to involve in initial adherence and biofilm formation in *E. faecium* (van Wamel et al., 2007). These roles were later proven by a biofilm deficient *esp* insertion-deletion mutant (Heikens et al., 2007). Subsequent studies using animal models indicate the association of *esp* to urinary tract

infections and endocarditis (Heikens et al., 2012; Leendertse et al., 2009). The *esp* gene of *E. faecium* is located on a large transferable pathogenicity island (PAI) which ranges from 64 to 104 kb in size (van Schaik et al., 2010). Comparative genomics of three clinical *E. faecium* strains reveals that although the general architecture of *esp* PAI is conserved in all the three studied strains, variations are observed in the gene content of the PAI which are likely caused by the independent acquisition of genes through horizontal gene transfer (van Schaik et al., 2010).

A putative virulence gene hyl_{Efm} was also found to be highly prevalent in clinical *E. faecium* strains (Freitas et al., 2010; Rice et al., 2003; Soheili et al., 2014). The hyl_{Efm} gene was initially suggested to encode for hyaluronidases but was later annotated as a putative glycoside hydrolase based on sequence comparison with *spy1600* gene in *Streptococcus pyogenes* (Rice et al., 2003; Sheldon et al., 2006). It has been previously shown that the hyl_{Efm} gene is carried on transferable megaplasmids which can also carry genes conferring resistance to glycopeptides (Arias et al., 2009; Freitas et al., 2010). Some of these hyl_{Efm}-containing plasmids have been shown to enhance gastrointestinal colonization and increase virulence in an experimental peritonitis mouse model, implicating potential virulence of the hyl_{Efm} gene in *E. faecium* (Arias et al., 2009; Rice et al., 2009). However, a more recent study demonstrates that hyl_{Efm} does not mediate the increased virulence conferred by the hyl_{Efm}-containing plasmid in murine peritonitis (Panesso et al., 2011). It remains to be determined if hyl_{Efm} plays any role in other infections such as endocarditis or urinary tract infections.

In addition to *esp* and *hyl*, several virulence factors associated with surface adhesion and biofilm formation are also significantly enriched in the HA strains. These include the *acm* gene encoding a collagen adhesin, and several pili genes, such as the *ebpABC* locus

and *fms21-fms20* locus. These virulence factors will be discussed in detailed in the biofilm section.

2.6 Antibiotic resistance in *E. faecium*

The clinical importance of *E. faecium* can be linked to its resistance to a broad range of antibiotics, which can be either intrinsic or acquired. Among all, acquisition of glycopeptides resistance, particularly to vancomycin, is the most significant event contributing to the transition of *E. faceium* from commensal to nosocomial pathogen.

Glycopeptides inhibit bacterial growth by disrupting cell wall synthesis. Resistance to glycopeptides is caused by either elimination of the high-affinity peptidoglycan precursors originally produced by the host, which removes the target sites; or modification of the termini of peptidoglycan precursors from D-Ala-D-Ala to either D-Ala-D-Lac or D-Ala-D-Ser, which lowers the binding affinity of the antibiotics to the target sites (Courvalin, 2006; Guzman Prieto et al., 2016; Hollenbeck & Rice, 2012). To date, nine distinct gene clusters (*vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM*, and *vanN*) conferring glycopeptides resistance have been described in enterococci (Boyd et al., 2008; Courvalin, 2006; Lebreton et al., 2011; Xu et al., 2010), with *vanA* and *vanB* being the most predominant gene clusters found in clinical VRE_{fm}.

The *vanA* gene cluster confers a high-level resistance to vancomycin and teicoplanin. This cluster is typically encoded on Tn*1546* or related transposons and consists of seven genes. The *vanR* and *vanS* genes encode for a two-component regulatory system which regulates the expression of *vanHAXYZ*. The *vanH* gene encodes for a dehydrogenase which converts cellular pyruvate to D-Lac. The *vanA*-encoding ligase then links the D-Lac to D-Ala, creating the D-Ala-D-Lac depsipeptide which will later ligate to the

tripeptide precursor by host enzymes, yielding the low-affinity pentapeptide precursor. For a full resistance to vancomycin, elimination of normal precursors is required. This is done by *vanX* and *vanY* genes. *vanX* encodes a D-D-dipeptidase which hydrolyses the original D-Ala-D-Ala dipeptide, making D-Ala-D-Lac the sole substrate for peptidoglycan synthesis. On the other hand, *vanY* encodes a D-D-carboxypeptidase which eliminates the terminal D-Ala from the normal pentapeptides, rendering them useless for normal cell wall synthesis (Courvalin, 2006; Hollenbeck & Rice, 2012). The *vanZ* gene encodes for a protein with yet unknown function. However, an association of VanZ with low-level teicoplanin resistance had been documented before (Arthur et al., 1995). Additionally, the *vanA* cluster also includes two genes (*orf1* and *orf2*) which are responsible for transposition.

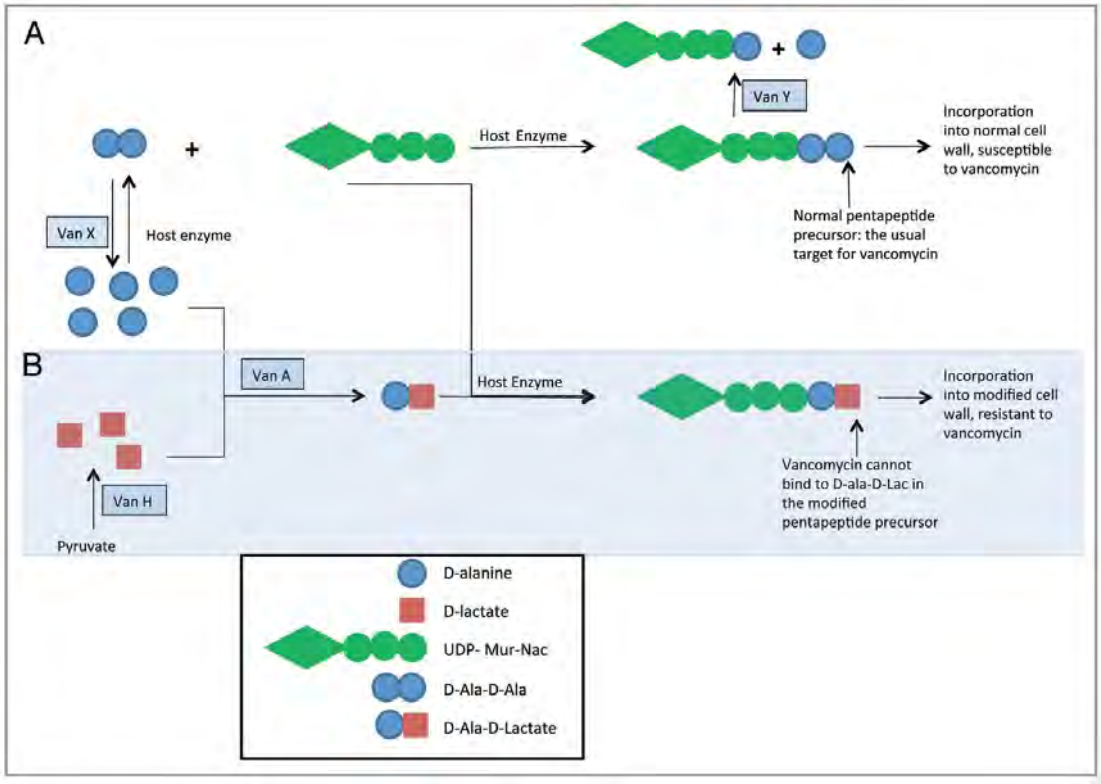


Figure 2.1: VanA-type vancomycin resistance mechanism. A) Normal cell wall synthesis pathway and disruption of the pathway by VanX and VanY. B) Construction of a modified cell wall that is resistant to vancomycin. Figure adapted from Hollenbeck and Rice (2012).

Studies that characterized the *vanA* gene clusters of *E. faecium* from different geographical regions and sources revealed great structural variations of the Tn*1546*-like elements that are mainly caused by point mutations, deletions and IS insertions (Hashimoto et al., 2000; Huh et al., 2004; Kuo et al., 2014; Willems et al., 1999). Most of these modifications are observed in the genes that are not involved directly in vancomycin resistance (*orf1*, *orf2*, *vanY*, *vanZ*) and in the intergenic regions. Notably, insertion of IS*1216V,* IS*1542,* and IS*1251*, as well as truncation of the *orf1* are the most frequently observed events (Gu et al., 2009; Huh et al., 2004; Kuo et al., 2014; Schouten et al., 2001). Some of these structural changes, especially those caused by IS integrations, have been shown to contribute to changes in the vancomycin and teicoplanin resistance level (Gagnon et al., 2011; Gu et al., 2009; Hashimoto et al., 2000; Sivertsen et al., 2016). Despite its high diversity, similar *vanA* cluster variants has been found in strains from different geographical regions and sources (human, animals, environments), indicating possible horizontal spread of the vancomycin resistance elements through conjugative plasmids or as part of a larger mobilized genetic unit (Huh et al., 2004; Sletvold et al., 2010; Willems et al., 1999). In fact, polymorphism of the *vanA* gene cluster has been used as a typing tool which, coupled with epidemiological data, provides a better understanding on the dissemination of vancomycin resistance in *E. faecium* (Xu et al., 2011).

Unlike the *vanA* gene cluster, the *vanB* gene cluster confers moderate to high-level resistance to vancomycin, but not to teicoplanin (Hollenbeck & Rice, 2012). The VanB locus is usually found on Tn*1549*-like transposons, which are encoded on plasmid or chromosome. The genetic organization and resistance mechanism of VanB locus is similar to that of VanA locus. Homologs of VanH ($VanH_B$), VanX ($VanX_B$), VanY ($VanY_B$), and VanA (VanB) are found in the VanB locus. Genes encoding the two-component regulatory system in the VanB locus ($vanR_B$ and $vanS_B$) are only distantly

related (34% and 24% amino acid identity, respectively) to the *vanRS* found in the VanA locus (Evers & Courvalin, 1996). This $VanRS_B$ system regulates differently from its counterparts of the Tn*1546* in that only vancomycin, but not teicoplanin, induces resistance of the VanB locus (Evers & Courvalin, 1996). Apart from that, a gene related to *vanZ* is lacking in the VanB locus but an additional *vanW* gene with unknown function is found. Based on sequence analysis, three *vanB* subtypes (*vanB1*, *vanB2*, *vanB3*) are identified. However, there is no correlation between the different genotypes and the level of resistance to vancomycin (Dahl et al., 1999).

High-level resistance to ampicillin is another important feature of clinical *E. faecium*. Ampicillin, like other beta-lactams, binds covalently to the penicillin-binding proteins (PBPs) and disrupts cross-linking of peptidoglycan precursors, thereby impairing cell wall synthesis. *E. faecium* exhibits intrinsic resistance to ampicillin due to the expression of low-affinity penicillin-binding protein 5 (PBP5). Increased resistance to ampicillin is mediated by either acquisition of beta-lactamase or mutations in the *pbp5* gene (Hollenbeck & Rice, 2012). In *E. faecium*, high-level ampicillin resistance is mainly due to the accumulation of various point mutations in the penicillin binding regions of PBP5 (Galloway-Peña et al., 2011; Rice et al., 2004). A combination of two point mutations, a methionine-to-alanine change in position 485 and an insertion of serine at position 466 of *pbp5*, has been shown to markedly enhance ampicillin resistance (Rice et al., 2004). A later study showed that the chromosomally encoded *pbp5* is able to transfer among *E. faecium* strains through conjugation (Rice et al., 2005). This suggests that mutated *pbp5* with enhanced resistance could be transmitted in the same way among the clinical isolates. On the other hand, high-level ampicillin resistance due to overproduction of beta-lactamase is relatively rare in *E. faecium* compared to in *E. faecalis* (Hollenbeck & Rice, 2012).

Another group of antibiotics to which *E. faecium* exhibits intrinsic and high-level acquired resistance to is aminoglycosides. Aminoglycosides act by inhibiting ribosomal protein synthesis. Generally, all enterococci are intrinsically resistant to low-level of aminoglycosides due to poor uptake of the antibiotics (Bryan & Van Den Elzen, 1977). The combination of cell-wall active agents such as ampicillin or vancomycin with aminoglycosides increases the uptake of the antibiotics, thereby enhances the killing effect on enterococci (Moellering & Weinberg, 1971). Other than the natural uptake barrier, *E. faecium* also poses chromosomally encoded enzymes such as 6′-N-aminoglycoside acetyltransferase encoded by *aac(6')-Ii* and rRNA methyltransferase encoded by *efmM* that confer low to moderate intrinsic resistance to aminoglycosides (Hollenbeck & Rice, 2012).

The acquisition of various genes encoding aminoglycoside-modifying enzymes (AMEs) results in high-level aminoglycosides resistance in *E. faecium*, which abolish the synergistic killing effect that is important for the treatment of severe enterococcal infections. The bi-functional gene *aac(6′)-Ie-aph(2′′)-Ia* is the most clinically significant as strains carrying this gene is virtually resistant to all clinically available aminoglycosides except streptomycin (Chow, 2000). *aac(6′)-Ie-aph(2′′)-Ia* is the most prevalent gene that confers high-level gentamicin resistance in enterococci, although other genes such as *aph(2′′)-Ib, aph(2′′)-Ic,* and *aph(2′′)-Id* also confer resistance to gentamicin (Chow, 2000). The enzymes encoded by *aac(6′)-Ie-aph(2′′)-Ia* inactivate gentamicin by phosphorylating the 2′ hydroxyl position of gentamicin (Ferretti et al., 1986). This modification renders the antibiotic unable to bind to its target on the 30S ribosome, thereby loses its antibacterial activity. Streptomycin can be used in the synergistic therapy against enterococci when *aac(6′)-Ie-aph(2′′)-Ia* is present, provided that there is no resistance to high levels of streptomycin (MIC ≥ 1000 μg/ml). Enzymatic modification of the antibiotic or single-step point mutations in the ribosome can

contribute to high-level streptomycin resistance in enterococci (Hollenbeck & Rice, 2012). The *ant(6´)-Ia* that encodes for an adenylyl transferase is one of the well-known resistance determinants that can inactivate streptomycin. The *ant(6´)-Ia* gene is often found as part of a multi-resistance gene cassette *ant(6´)-sat4-aph(3´)* which confers resistance to streptomycin, streptothricin, and kanamycin. This gene cluster is encoded on Tn*5405* and other related transposons that are widely found in *Staphylococci* and *Enterococci* (Derbise et al., 1997). Other AMEs that contribute to acquired aminoglycosides to enterococci include those that are encoded by *aph(3´)-IIIa* and *ant(4´´)-Ia*, which confer resistance to kanamycin, tobramycin, amikacin, and neomycin.

Due to the growing problem and clonal spread of VRE_{fm}, new drugs such as linezolid, daptomycin, and tigecycline have been increasingly used to treat infections caused by VRE_{fm}. However, resistance has also been observed in these antibiotics (Edelsberg et al., 2014; Tsai et al., 2012). With very few effective therapeutic options left to treat VRE_{fm}, development of newer drugs targeting different structures in the *E. faecium* cell, as well as strict infection control measures are utmost important to combat this pathogen.

2.7 Biofilm formation by *E. faecium*

Biofilm formation is an important virulence feature responsible for the pathogenesis of enterococci. Several difficult-to-treat enterococcal diseases were often biofilm-mediated, including those associated with indwelling medical devices and urinary catheters, as well as endocarditis (Donlan et al., 2002). Several studies have also reported the association of biofilm formation to enterococci isolated from clinical settings (Baldassarri et al., 2001; Mohamed et al., 2004; Toledo-arana et al., 2001). Although clinical *E. faecalis* isolates are more frequently found to form biofilm, *E. faecium* biofilm

formation by clinical isolates is still relatively high compared to those isolated from other sources (Almohamad et al., 2014).

Multiple biofilm-associated genes have been identified, mostly originated from *E. faecalis*. In *E. faecium*, these genes are mostly associated with adhesins or pili which are involved in the initial attachment step in biofilm formation. The enterococcal surface protein (Esp) is among the first cell-wall associated proteins that are shown to involve in biofilm formation in *E. faecium* (Heikens et al., 2007). The levels of *esp* expression on cell surface have been reported to affect initial attachment and biofilm formation (van Wamel et al., 2007), thereby explaining the conflicting observations that some *esp* positive isolates are unable to form biofilm (Dupre` et al., 2003). Further genetic analysis identifies a gene encoding an AraC-type transcriptional regulator, known as *ebrB* (enterococcal biofilm regulator B), which is found upstream of the *esp* gene. Deletion of *ebrB* resulted in reduced expression of *esp*, as well as reduced biofilm formation (Top et al., 2013). These results indicate that *ebrB* is involved in regulation of *esp* expression and is implicated in biofilm formation in *E. faecium*.

Microbial surface components recognizing adhesive matrix molecules (MSCRAMMs) are a group of surface proteins that plays an important role in host-pathogen adherence, as well as binding of bacteria to abiotic surfaces coated with host-derived extracellular matrix components. Two *E. faecium* MSCRAMMs that are involved in biofilm formation are Acm and SgrA. Acm, a cell wall-anchored collagen adhesin which binds to collagen type I and type IV is highly prevalent in *E. faecium* clinical isolates (Nallapareddy et al., 2008a). Although this adhesin is also found in non-clinical isolates, the gene encoded for this protein (*acm*) shows an insertion element disruption, rendering the encoded protein non-functional (Nallapareddy et al., 2008a). An *acm* deletion mutant has been shown to reduce biofilm formation in *E. faecium*. Furthermore, this *acm* mutant is also less virulent

in a rat endocarditis model, indicating the contribution of *acm* to *E. faecium* pathogenesis (Nallapareddy et al., 2008). The surface adhesin SgrA, on the other hand, binds to extracellular matrix molecules nidogen 1 and nidogen 2. Although SgrA does not mediate binding of *E. faecium* to biotic surfaces such as human bladder and intestinal epithelial cells, its contribution to biofilm formation on polystyrene surface was demonstrated (Hendrickx et al., 2009).

Another group of surface expressed proteins that contributes to *E. faecium* biofilm formation is pili. Two pilus-like structures, PilA and PilB, are frequently found in clinical *E. faecium* isolates, including those isolated from endocarditis patients (Hendrickx et al., 2010, 2008). The *pilA* gene is part of the *pilA* (*fms21*)-*fms20* gene cluster which is encoded on a large transferable plasmid (Kim et al., 2010). On the other hand, *pilB*, also known as *ebpC* or *fms9*, is part of the $ebpABC_{fm}$ cluster and encodes for a major pilus PilB. Deletion of $ebpABC_{fm}$ eliminates cell surface expression of PilB-containing pili and reduces biofilm formation. Additionally, the deletion mutant displays reduced colonization in a murine UTI model, signifying the role of $ebpABC_{fm}$ in the pathogenesis of *E. faecium* (Sillanpää et al., 2011).

Other than factors involving in initial attachment, several proteins that are associated with maturation of biofilm are also identified in *E. faecium*. These include a major autolysin, Alt_{Efm}, and a secreted protein, SagA. Alt_{Efm} involves in the release of extracellular DNA (eDNA), which is one of the important components of extracellular polymeric substances. The importance of Alt_{Efm} in biofilm formation has been demonstrated through insertion disruption of the autolysin gene (Paganelli et al., 2013). Besides eDNA, secreted proteins are also an important component in the biofilm matrix. The secreted protein SagA presents abundantly in the biofilm of *E. faecium* strains

(Paganelli et al., 2015). Proteolytic degradation of SagA has been shown to markedly reduce biofilm formation of clinical *E. faecium* strains (Paganelli et al., 2015).

2.8 Transcriptome study

Transcriptome is the total mRNA in a cell that represents genes that are actively expressed at a specific growing stage or physiological condition (Wang et al., 2009). Transcriptomic aims to elucidate the functional elements of a genome; to determine the transcriptional structure of genes; and to quantify gene expression levels under different conditions (Wang et al., 2009). Understanding the transcriptome is, therefore, important to unravel the underlying molecular mechanisms of bacterial development and pathogenesis.

Microarray technology has been widely used in the transcriptomic analyses of different microorganisms. However, this hybridization-based method has several limitations, including a limited dynamic range of transcript detection due to problems with background noise and saturation of signals, dependency on the existing knowledge of genomes or transcripts, and complex normalization methods in comparing different experiments (Hinton et al., 2004; Wang et al., 2009). With the advent of next generation sequencing technologies, a more recent technology for transcriptome profiling termed RNA-seq (RNA sequencing) has been developed. In contrast to the microarray, transcriptomic studies using RNA-seq are not limited to previously known genes as this method directly determines the cDNA sequences (Wang et al., 2009). Moreover, RNA-seq provides better resolution and higher reproducibility compared to microarray (Wang et al., 2009). With all these advantages, RNA-seq technology has provided a large amount

of data which refined our understanding in the gene expression profile of an organism under different treatments or environmental stresses.

CHAPTER 3: METHODOLOGY

3.1 Clinical data collection and patients' background

Clinical data including the patients' ward, samples source, samples isolation date, antibiotic treatments and surgical procedures of the studied patients was collected. Ethical approval had been obtained from the University of Malaya Research Ethics Committee (UMREC) with ethical approval number 20159-1661.

Patient X was admitted into cardiology ward in 24th October 2010. Subsequently after the isolation of first VRE (VRE2) from this patient, VRE infection cases were frequently reported in the hospital. Patient Y was admitted into neuro-ICU ward due to basal ganglia bleed approximately three months (21st January 2011) after patient X was admitted. Microbiological investigation revealed that patient Y was infected by multiple pathogens, including multidrug-resistant *Acinetobacter baumannii*, methicillin-resistant *Staphylococcus aureus* (MRSA), and *Pseudomonas aeruginosa*. VREr5 was isolated from patient Y after one month of treatment with vancomycin. Following this, another two strains, VREr6 and VREr7, were isolated from the same patient who was undergoing an antibiotic treatment with meropenem, colistin, ceftazidime, and linezolid. After the linezolid treatment, VRE had not been isolated from patient Y. To note, patient Y underwent several surgical procedures such as intracranial pressure monitoring and external ventricular drain insertions (EVD) during hospitalization. These procedures may serve as the route of bacterial transmission although this association is unclear. Eventually, patient Y died of sepsis due to infective endocarditis and pneumonia.

3.2 Whole genome sequencing

3.2.1 Bacterial strains

Four clinical strains of vancomycin-resistant *E. faecium* (VRE2, VREr5, VREr6, VREr7) were collected in the year 2011 from University Malaya Medical Centre (UMMC), Kuala Lumpur. These strains were selected based on previous PFGE results, which were reconfirmed in this study. The strains were isolated and validated through PCR-amplification and sequencing of the 16S rRNA at the clinical microbiology laboratory. VRE2 was isolated from patient X and was the first VRE_{fm} isolated in the studied period. VREr5, VREr6, and VREr7 were isolated from the cerebrospinal fluid (CSF), blood, and urine of patient Y, respectively. These three strains were isolated at one-week interval, with VREr5 being the first isolate, followed by VREr6 and VREr7. A single colony of each strain was then cultured into Brain Heart Infusion broth (Oxoid Ltd., Basingstoke, UK) and incubated at 37°C for 24 hr. The culture was stored at -80°C in 50% glycerol until further experiments.

3.2.2 Pulsed-field gel electrophoresis (PFGE)

PFGE was performed as previously described (Turabelidze et al., 2000), with slight modifications. Briefly, the bacteria were first lysed in a combination of lysozyme (100mg/ml) and mutanolysin (10kU/ml) at 37°C for four hours. Chromosomal DNA was then prepared in agarose gel block and digested with restriction enzyme *Sma*I (Promega, Madison, WI, USA) at room temperature. The restriction fragments were separated by electrophoresis in 0.5 x TBE buffer for 20 hr at 14°C in a CHEF Mapper system (Bio-Rad, CA, USA) using pulsed times of 3.5-25 s and 1-5 s. *Xba*I-digested *Salmonella enterica* ser. Braenderup H9812 was used as a DNA size marker. The PFGE data were

analyzed using BioNumerics version 6.0 software (Applied Maths, Kortrijk, Belgium). The quantitative differences in the banding patterns were defined by the Dice coefficient. Cluster analysis was determined based on the unweighted pair group method with averages (UPGMA), using a position tolerance 1.5%.

3.2.3 DNA extraction

Genomic DNA of the four VRE_{fm} strains were extracted using the Wizard Genomic DNA Purification Kit (Promega, Madison, Wisconsin, United States). Briefly, 1 ml of the bacterial culture was harvested by centrifugation for 2 min at 14,000 rpm. To weaken the cell wall of gram-positive bacteria, the cell pellet formed was re-suspended in 480 µl of EDTA (50mM), after which 120 µl of lysozyme (10mg/ml) was added. The sample was incubated at 37°C for one hour, followed by centrifugation for 2 min at 14,000 rpm. The supernatant was removed and the cell pellet was re-suspended in 600 µl of lysis solution. The sample was further incubated at 80°C for 5 min. Following this, 3 µl of RNase solution was added to the re-suspended pellet, mixed, and incubated at 37°C for 15 to 60 min. After incubation, 200 µl of protein precipitation solution was added and vortexed to mix. The sample was incubated on ice for 5 min prior to centrifugation for 3 min at 14,000 rpm. The resultant supernatant was transferred to a clean microcentrifuge tube containing 600 µl of room temperature isopropanol and then followed by centrifugation at 14,000 rpm for 2 min to recover the precipitated DNA. The supernatant was discarded and 600 µl of room temperature 70% ethanol was added to clean the DNA pellet. The solution was centrifuged for another 2 min at 14,000 rpm. The ethanol was carefully aspirated and the pellet was air-dried for 10 to 15 min. Finally, the dried DNA pellet was re-suspended in 100 µl of sterile distilled water. Extracted DNA was quantified using the

spectrophotometer at OD_{260} and the purity was determined by OD_{260}/OD_{280} ratio. The extracted DNA was stored at -20°C for long term storage.

3.2.4 Genome sequencing, assembly, and annotation

Whole genome sequencing of the VRE2, VREr5, VREr6, and VREr7 was carried out by a commercial vendor using the Illumina Miseq platform, version 2.0 with reads coverage ranged from 78x to 108x. The genome sequences were then assembled using CLC Genomic Workbench version 5.1 (CLC Bio, Aarhus, Denmark). Open reading frame (ORF) prediction was performed using Prodigal (Hyatt et al., 2010). Functional annotated of the genomes was performed using RAST (Rapid Annotation using Subsystem Technology) (Aziz et al., 2008) and Blast2GO (Conesa et al., 2005).

3.2.5 Genome analyses and comparative studies

The sequence types of the four sequenced strains were determined *in silico* using the PubMLST database (http://pubmlst.org/). Genomes alignment and comparison were performed using Mauve 2.3.1 using *E. faecium* Aus0085 as the reference (Darling et al., 2004). The circular genomic map was constructed using BLAST ring image generator (BRIG) (Alikhan et al., 2011). Insertion sequence (IS) elements, prophages, and clustered regularly interspaced short palindromic repeats (CRISPR) were identified using IS Finder (Siguier et al., 2006), PHAST (Zhou et al., 2011), and CRISPRfinder (Grissa et al., 2008), respectively.

3.2.6 Phylogenomic analysis

For phylogenomic analysis, sixteen draft genomes of clinical *E. faecium* from ten different countries (Austria, United States of America, Netherlands, France, China, Africa, Portugal, United Kingdom, Israel, and Denmark) with 11 different STs (ST 203, 17, 78, 117, 18, 210, 27, 26, 414, 160, and 64) were retrieved from NCBI database and compared with our four sequenced draft genomes. The names and respective Genbank accession numbers of the sixteen strains are as follows: Aus0085 (CP006620.1), Aus0004 (CP003351.1), TX16 (CP003683.1), E1133 (AHWR00000000.1), E155 (AUWX00000000.1), E1904 (AHXQ00000000.1), E0120 (AHWI00000000.1), E2560 (AHYI00000000.1), E1185 (AHWS00000000.1), E161 (JXZA00000000.1), LCT-EF128 (AJUP00000000.1), E1731 (AHXO00000000.1), E6045 (AHYL00000000.1), E1392 (AHWV01000046), E0333 (AHWL00000000.1), VRE84 (AIVF00000000.1). Genome sequences were submitted to the Reference Sequence Alignment-based Phylogenic Builder (RealPhy) (Bertels et al., 2014) for the identification of sites that are relevant for the phylogenomic study. *E. faecium* Aus0085 was chosen as the reference genome as this strain shares similar features with our Malaysian strains and has the same sequence type (ST 203) as two of these strains. Aus0085 is a clinical VRE isolate with known complete genome sequence (Lam et al., 2013). The generated multiple genome sequence alignments were used to construct an approximately-maximum-likelihood tree using FastTreeMP as previously described (Price et al., 2010; Yap et al., 2014).

3.2.7 GenBank accession numbers

The genome sequences of *E. faecium* strains VRE2, VREr5, VREr6, and VREr7 were deposited in GenBank under accession number LTAA00000000, LTBJ00000000, LTDQ00000000, and LSZZ00000000, respectively.

3.3 Virulence factors

3.3.1 Virulence genes identification

The Virulence Factors of Pathogenic Bacteria database (Chen et al., 2005) was used to identified virulence profile of the studied strains. Only results with more than 60% coverage and 60% sequence identity were considered positive for the analysis. Selected virulence genes were further verified by PCR amplification and DNA sequencing of the amplicons.

3.3.2 Biofilm assay

To study the biofilm forming ability of the selected strains, crystal violet assay was performed as described by Baldassarri et al. (2001) with slight modifications. Briefly, overnight cultures were prepared using tryptic soy broth (TSB) (Oxoid Ltd., Basingstoke, UK) containing 6 μg/ml vancomycin. The overnight cultures were diluted 1: 100 and 200 μl was inoculated into 96-well polystyrene flat-bottom microtiter plate. The inoculated plate was incubated at 37°C for 48 hr. After incubation, the plate was tapped vigorously followed by two times washing with 1x phosphate buffered saline (PBS) to remove unbound cells. The bounded cells were fixed at 60°C for 45 min and subsequently stained

with 200 µl crystal violet (0.1%) for 10 min. The plates were then washed with PBS thrice and dried before adding 200 µl 80:20 ethanol-acetone. The absorbance of the eluted stain was measured at 590nm wavelength. The true OD readings of each strain were acquired after deducting the negative control, which contained only the growth medium.

The biofilm forming ability of the studied strains was scored as previously described (Chelvam et al., 2014; Stepanović et al., 2000). Briefly, the cut-off OD (ODc) was defined as three standard deviations above of the mean OD of the negative control. The biofilm forming ability of each strain was scored as follow: OD ≤ ODc = non-biofilm producer, ODc < OD ≤ (2x ODc) = weak-biofilm producer, (2x ODc) < OD ≤ (4x ODc) = moderate-biofilm producer, OD > (4x ODc) = strong-biofilm producer. The experiment was performed in triplicate and repeated three times for reproducibility.

3.3.3 Confocal Laser Screen Microscopy (CLSM)

To confirm the results obtained from crystal violet assay, CLSM was performed. Biofilm was grown in an eight-well chamber slide at 37°C for 48 hr. Visualization of biofilm was performed as previously described (Jurcisek et al., 2011). Briefly, planktonic cells were removed by gentle pipetting, and the biofilm was washed with sterile PBS twice. The biofilm was stained with 200 µl of BacLight Live/Dead Kit (Invitrogen Ltd., Paisley, UK) for 15 min at dark and subsequently fixed with 200 µl of neutral buffer formalin (NBF) for 30 min. The stained biofilm was later washed with sterile saline twice. An appropriate amount of sterile saline was added to the slide before applying the coverslip to keep the biofilm hydrated during imaging. Confocal images were collected using a Leica TCS SP5 microscope. The images were viewed by using the LAS AF Lite

software (Leica). The image stacks were acquired using ImageJ software (Fiji, ImageJ, Wayne Rasband National Institutes of Health).

3.3.4 Determination of the plasmid origin of *E. faecalis bee* homolog

To determine the location of the *E. faecalis bee* locus identified in VREr5, both total and plasmid DNA were extracted. Total DNA was extracted by cell boiling method. Briefly, one single colony was suspended in 100 µl of distilled water and boiled for 5 min at 99°C. The bacterial suspension was then immediately cooled on ice for 10 min followed by centrifugation for 5 min at 13,400 rpm. The supernatant was transferred to a new 500 µl Eppendorf tube and stored at -20°C.

Plasmid DNA was extracted using QIAprep spin miniprep kit (Qiagen, Hilden, Germany). 5 ml of overnight culture was centrifuged at 6800xg at 25°C for 3 min. The pelleted bacterial cell was re-suspended in 250µl Buffer P1 and transferred to a clean 2 ml microcentrifuge tube. A total of 250 µl Buffer P2 was added and the tube was inverted four to six times to mix the solution. Following this, 350 µl Buffer N3 was added and the mixture was mixed immediately by inverting the tube four to six times. After this, the mixture was centrifuged at 13,000 rpm for 10 min and the resultant supernatant was transferred to a clean QIAprep Spin Column. The supernatant was centrifuged for 60 s and the flow-through was discarded. To the QIAprep Spin Column, 500 µl Buffer PB was added, followed by centrifugation at 13,000 rpm for 60 s. After removing the flow-through, the QIAprep Spin Column was washed by adding 750 µl Buffer PE and centrifuged again for 60 s. Residual washing buffer was removed by centrifugation for another 60 s, after removing the flow-through. The QIAprep Spin Column was then transferred to a new 1.5 ml microcentrifuge tube. To elute the plasmid DNA, 50 µl of

deionized water was added directly to the membrane of the QIAprep Spin Column, let stand for 1 min, and centrifuge for 60 s. The extracted plasmid DNA was stored at -20°C.

Polymerase Chain Reaction (PCR) was performed using novel primers designed to detect the *bee*-1 and *bee*-2 genes and primers adopted from Tendolkar et al. (2006) for detection of *bee*-3 gene (P101D12-1, Bee-12) (Appendix 4). A 5 µl aliquot of each PCR product was electrophoresed on a 1.5% agarose gel for 30 min at 100V. The gel was stained with GelRed^{TM} nucleic acid gel stain (Biotium) for 15 min. The gel image was visualized with Gel documentary system (Bio-rad, Hercules, CA).

3.4 Antibiotic resistance

3.4.1 Antibiotic susceptibility tests (AST)

The Kirby-Bauer disc diffusion method was used to test the susceptibility of the *E. faecium* strains to the following ten antibiotics (Oxoid Ltd., Basingstoke, UK): vancomycin (30 µg), gentamicin (120 µg), streptomycin (10 µg), kanamycin (30 µg), ampicillin (10 µg), erythromycin (15 µg), clindamycin (2 µg), tetracycline (30 µg), teicoplanin (30 µg), and linezolid (30 µg). The antibiotic susceptibility test was carried out on Mueller-Hinton (Brcton Dickson, NJ, USA) Agar swabbed with 0.5 McFarland standard suspension of *E. faecium* and incubated at 37°C for 18 hours. For vancomycin, the plate was incubated for 24 hours. Guidelines from the Clinical and Laboratory Standards Institute (CLSI) (Clinical and Laboratory Standards Institute, 2016) were used to interpret the diameters of the inhibition zones produced.

3.4.2 MIC determination for vancomycin, teicoplanin, and gentamicin

The minimal inhibitory concentration (MIC) values of vancomycin, teicoplanin, and gentamicin were determined using E-test strips (BioMérieux, Marcy-I'Étoile, France). *E. faecium* cultures were adjusted to 0.5 McFarland standard and swabbed uniformly on Mueller-Hinton agar (Becton Dickson, NJ, USA). E-test strips were applied to the surface of the agar and the inoculated plates were incubated at 37°C for 16 to 20 hr. The MIC values were recorded and interpreted according to the guidelines from the Clinical and Laboratory Standards Institute (CLSI) (Clinical and Laboratory Standards Institute, 2016).

3.4.3 Identification and confirmation of vancomycin-resistant subtype

The vancomycin-resistant subtype (*vanA*) of the selected strains were confirmed by PCR using specific primers (Kariyama et al., 2000) and condition as shown in Appendix 4. Following PCR amplification, 5 µl of each amplicon was loaded into a 1.5% (w/v) agarose gel and electrophoresis was performed at 100V for 30 min. The gel was stained with GelRed^{TM} nucleic acid gel stain (Biotium) for 15 min. The gel image was captured and analyzed with Gel documentary system (Bio-rad, Hercules, CA)

PCR products with the desired size were purified using Wizard SV Gel and PCR Clean-up system (Promega, Madison, Wisconsin, United States). An equal volume of membrane binding solution was added to the PCR amplification product and mixed thoroughly. The solution was transferred to a Mini-column and incubated at room temperature for 1 min. The sample was centrifuged at 16,000x g for 1 min and the flow-through was discarded. Following this, 700 µl of membrane wash solution was added and the sample was centrifuged for 1 min at 16,000x g. After discarding the flow-through,

another 500 µl of membrane wash solution was added followed by centrifugation for 5 min. The empty Mini-column was further centrifuged for 1 min to remove residual ethanol. Purified DNA was eluted by adding 50 µl of sterile distilled water.

The purified DNA was sent to a laboratory (First Base Sdn. Bhd., Seri Kembangan, Malaysia) for sequencing together with the forward and reverse PCR primers. DNA sequences obtained were trimmed using Mega 6 (downloadable from http://www.megasoftware.net/) and analyzed with standard nucleotide-nucleotide BLAST at the National Center for Biotechnology Information website (http://blast.ncbi.nlm.nih.gov/Blast.cgi).

3.4.4 Antibiotic resistance genes determination

To identify antibiotic resistance genes, the nucleotide sequences of the four sequenced strains were submitted to the Antibiotic Resistance Gene Database (ARDB) (Liu and Pop, 2009). A resistance profile with corresponding resistant genes was generated for each strain. The nucleotide sequences were also submitted to ResFinder 2.1 server (Zankari et al., 2012) to identify acquired antimicrobial resistance genes in the genome of each strain.

3.4.5 Tn*1546* structural analysis

The molecular structures of the Tn*1546*-like elements were analyzed using published PCR primers (Huh et al., 2004; Simjee et al., 2002). 5235.F and 7035.R (Simjee et al., 2002) were used to fill the gap between the two contigs carrying the *van* genes of the Tn*1546*-like elements. Primer ISV650F and ISV132R (Huh et al., 2004) were used to detect the presence of IS*1216V*. To determine the DNA sequences of the right end of the

truncated Tn*1546*-like elements, primer 4511R was used with combination of the IS*1216V*-specific primers (ISV650F, ISV132R). All amplicons were sequenced commercially (First Base Sdn. Bhd., Seri Kembangan, Malaysia). The DNA sequences were analyzed using CodonCode Aligner (CodonCode Corporation, Dedham, MA) to obtain the exact structure of the Tn*1546*-like elements.

3.5 Transcriptomic analyses of *E. faecium* biofilm

3.5.1 Confirmation of biofilm developmental stage using CLSM

VREr5, the only biofilm former among the four studied strains, were selected for the transcriptomic study. In order to study the change in gene expression levels during the transition from planktonic to biofilm cells, VREr5 was allowed to form biofilm for a period of 24 hours. To verify that VREr5 was in the initial stage of biofilm development, CLSM was performed using the protocol previously described in section 3.4.3.

3.5.2 Total RNA extraction

Total RNA of both planktonic- and biofilm-state cells was extracted using Wizard SV Total RNA Extraction System (Promega, Madison, Wisconsin, United States) with slight modifications.

For planktonic-state cells, overnight cultures in TSB (Oxoid Ltd., Basingstoke, UK) were diluted 1:50 and allowed to grow until the OD_{600} of 0.8. Then, 1 ml of the culture was harvested at 4°C by centrifugation at 14,000x g for 2 min. For biofilm-state cells, biofilm was first allowed to grow in a 24-well microtiter plate for 24 hours and unbound

cells were removed by washing with PBS. The attached cells were scraped out using the pipette tip and re-suspended in PBS. The re-suspended cells were harvested at 4°C by centrifuging at 14000x g for 2 min. To the resulting pellet harvested from both biofilm and planktonic cells, 60 µl each of freshly prepared lysozyme (10mg/ml) and mutanolysin (10kU/ml) was added and vortexed. The re-suspended pellet was incubated at room temperature for 15 min. After incubation, 75 µl of RNA lysis buffer was added, followed by 350 µl of RNA dilution buffer, and the mixture was mixed by inversion. The clear lysate solution was then mixed with 200 µl of 95% ethanol, transferred to the clean Spin Column assembly, and centrifuged for 60 s at 14,000x g. After discarding the flow-through, 600 µl of the RNA wash solution was added to the spin column, followed by centrifugation at 14,000x g for 60 s. The flow through was discarded. A total of 50 µl DNase incubation mix was added directly to the membrane of the spin column and incubated at room temperature for 15 min. The DNase incubation mix was prepared according to the Table 3.1:

Table 3.1: DNase incubation mix setup

Solution	Volume per prep
Yellow Core Buffer	40 µl
$MnCl_2$	5 µl
DNase I	5 µl

After incubation, 200 µl of DNase stop solution was added to the column and centrifuged for one minute at 14,000 x g. Following this, 600 µl of RNA wash solution was added and the centrifugation step was repeated. The flow through was discarded. The spin column was centrifuged for two minutes at high speed after the addition of 250 µl of RNA wash solution. The spin basket was transferred to a clean 1.5ml microcentrifuge tube and 100 µl of nuclease-free water was added directly to the membrane of the column. The RNA was eluted by centrifugation at 14,000 x g for 60 s.

The extracted RNA was quantified using NanoDrop (IMPLEN, Germany). To access the RNA integrity, 1µl of RNA sample was mixed with 4 µl of 1x TE and electrophoresis was performed on 1% agarose gel. The extracted RNA was stored at -80°C until further use.

3.5.3 RNA-seq and data processing

The quality of RNA samples was re-checked using Qubit 2.0 (Invitrogen, California, USA) with the Qubit RNA HS assay (Life Technologies, California, USA) and Agilent 2100 bioanalyzer (Agilent Technologies, Santa Clara, California, United States) before sending for sequencing. RNA-seq was performed by a commercial vendor on the Illumina HiSeq platform. The quality of the sequenced reads was checked using Novogene in-house quality control software. The low-quality reads, which 1) contained adaptor contamination, 2) contained N> 10%, and 3) with low-quality base (>50% of the total base has quality score ≤ 5) were removed. *de novo* assembly was performed using CLC Genomic Workbench software (Version 7.5.1, CLC Bio, Denmark). Following this, coding sequences (CDS) were identified using TransDecoder and were annotated by performing BLASTp and BLASTx against the NCBI nr protein database and Swiss-Prot database with an E-Value cut-off of 1e-5. Functional annotation was performed by Blast2GO program (Conesa et al., 2005) to obtain the GO annotation and KEGG pathways. The COG assignment was carried out using the eggNOG mapper (Huerta-Cepas et al., 2016).

Reads quantification was carried out by mapping back the raw reads to the generated transcripts using the CLC Genomic Workbench software. Differentially expressed genes were determined using the Empirical analysis of DGE algorithm of the same software.

Genes with a False Discovery Rate (FDR)-corrected *p*-value < 0.001 and an absolute value of $\log_2$ fold change (log FC) > 2 were identified as being differentially expressed. A *p*-value < 0.05 was used as a cut-off for determining significantly up- and down-regulated genes.

3.5.4 Quantitative Real-time PCR for validation of RNA-seq results

Extracted RNA using protocol described in section 3.5.2 was subjected to reverse transcription using High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Foster City, California, USA).

The master mix for reverse transcription was prepared on ice as in Table 3.2. The PCR running condition for the reverse transcription is listed in Table 3.3.

Table 3.2: Master mix setup for reverse transcription

Component	Volume per reaction
Nuclease-free water	4.2 µl
10x RT buffer	2.0 µl
10x RT primer	2.0 µl
25x dNTP mix (100mM)	0.8 µl
Reverse transcriptase	1.0 µl
Total RNA template	10.0 µl

Table 3.3: PCR running condition for reverse transcription

	Temperature (°C)	Time (min)
Step 1	25	10
Step 2	37	120
Step 3	85	5
Step 4	4	∞

Real-time PCR was performed using qPCRBIO SyGreen Blue Mix (PCR Biosystems, London, UK). To determine the efficiency and specificity of the assay, standard curve with five dilution points and melt curve were performed for each of the selected up-regulated (*ebpA*, *tetS*, *repR*) and down-regulated (*arcA*, *fsrB*, *bee-2*) genes as well as housekeeping gene (*ddl*). The reaction master mix was setup according to Table 3.4. The reaction was run on a ViiA 7 Real-Time PCR system (Applied Biosystem, California, USA). Real-time cycling condition for the reactions is listed in Table 3.5. Melt curve running condition is listed in Table 3.6. During the experiment, melting curve profile for each amplification reaction and the relevant C_t value was automatically generated, using the software provided with the system. The expression values of the target genes in biofilm cells were quantified by the relative standard curve method, using *ddl* as the endogenous control and planktonic sample as reference.

Table 3.4: Master mix setup for real-time PCR

Component	Volume (µl) per reaction	Final concentration
2 x qPCRBIO SyGreen Blue Mix	10	1 x
Forward primer (10µM)	0.8	400nM
Reverse primer (10µM)	0.8	400nM
Template cDNA (100 - 0.16ng)	1	Variable
PCR grade dH_2O	7.4	-
Total reaction volume	**20**	**-**

Table 3.5: Real-time PCR cycling condition

	Polymerase activation	**PCR (40 cycles)**	
	Hold	**Denaturation**	**Annealing/Extension**
Temperature (°C)	95	95	60
Time	2 min	5s	20s

Table 3.6: Melt curve running condition

	Temperature (°C)	**Time**
Step 1	95	15 s
Step 2	60	1 min
Step 3 (Dissociation)	95	15 s

3.5.5 Raw transcriptome data deposition

The cleaned raw sequences were deposited in the NCBI's Sequence Read Archive (SRA) under the experiment accession numbers SRX2783031 to SRX2783036.

CHAPTER 4: RESULTS

4.1 Whole genome sequencing

4.1.1 Pulsotypes of VRE$_{fm}$

PFGE subtyping of the four clinical strains revealed three pulsotypes comprising 16-27 restriction fragments (Figure 4.1). Cluster analysis at 90% similarity yielded two clusters. VREr6 and VREr7, which were isolated from the same patient (patient Y), were identical to each other (Cluster II). Interestingly, VREr5, which was also isolated from patient Y, was grouped into a different PFGE cluster (cluster I). Instead, VREr5 showed similar PFGE profile as VRE2 which was isolated from patient X, differing by only two bands.

Figure 4.1: Dendrogram showing the cluster analysis of four VRE$_{fm}$ strains based on PFGE patterns of the *Sma*I-digested chromosomal DNA. The dendrogram was constructed using the Dice coefficient and UPGMA clustering parameters at 1.5% position tolerance. The strain code, source, host, and cluster are indicated. The dotted vertical line indicates 90% similarity level.

4.1.2 Genome features

Whole genome sequence analysis indicated that the estimated genome sizes of the four VRE_{fm} strains ranged from 2.8 Mbp to 3.0 Mbp. All these strains had guanine-cytosine (GC) contents of approximately 38%. The numbers of predicted protein coding sequences (CDS) of the four genomes ranged 2853 to 3057. Both VRE2 and VREr5 harbored 55 tRNA genes whereas VREr6 and VREr7 harbored 58 tRNA genes. The general genome features of these four strains are summarized in Table 4.1.

Table 4.1: General genome features of VRE_{fm} strain VRE2, VREr5, VREr6, and VREr7.

	VRE2	**VREr5**	**VREr6**	**VREr7**
Genome size (bp)	2,862,609	2,898,367	3,014,993	3,021,201
GC%	37.85	37.8	37.61	37.6
Number of CDS	2,853	2,906	3,049	3,057
Number of tRNA	55	55	58	58
Number of rRNA	3	3	3	3

4.1.3 Multilocus sequence typing (MLST)

The sequence types (STs) of the four VRE_{fm} strains were determined by the existing *E. faecium* MLST scheme which uses the following seven housekeeping genes: *atpA*, *ddl*, *ghd*, *purK*, *gyd*, *pstS*, and *adk*. VRE2 and VREr5 were assigned to ST80 (9-1-1-1-12-1-1) whereas VREr6 and VREr7 were assigned to ST203 (15-1-1-1-1-20-1). Both these STs are grouped under clonal complex 17 (CC17), a specific lineage associated with nosocomial *E. faecium* strains (Top et al., 2008).

4.1.4 Comparative genomics

The genomes of the four Malaysian strains were compared to that of the reference strain *E. faecium* Aus0085 (GenBank accession number CP006620.1), one of the 13 *E. faecium* complete genomes known up to date. Aus0085 was chosen as the reference strain as it shares similar features with the four local VRE_{fm} strains. This clinical VRE strain was from similar geographical region as the four local strains, and was isolated from blood sample of a bacteremia patient (Lam et al., 2013). Moreover, this strain also has same ST as two of our strains (VREr6, VREr7).

Genomic comparison revealed that a total of 2148 open reading frames (ORFs) were shared among all the genomes. This conserved ORFs accounted for approximately 71% of the total ORFs present in each of the studied strains, which was comparable to other study (Lam et al., 2012). At least four chromosomal regions spanning approximately 169 kbp in Aus0085 showed low or no homology (<70%) with the four Malaysian strains, including a region carrying the *vanB*-containing Tn*1549* transposon (Figure 4.2). VREr6 and VREr7 shared higher similarity between each other and with the reference genome compared to VRE2 and VREr5, probably due to the same ST. Consistent with this result, higher number of shared genes were also observed between VREr6 and VREr7 when the four strains were compared among themselves (Figure 4.3). Similarly, VRE2 and VREr5 also shared more genes as compared to the other two strains (Figure 4.3). Some of these shared genes were associated with additional fitness and survival advantage. These included a cluster of genes encoding inositol metabolism proteins found only in VRE2 and VREr5, and a gene encoding lactococcin 972 (Lcn972) found only in VREr6 and VREr7. Moreover, each of the four strains also carried different number of strain-specific genes, ranging from 16 to 57. Most of these strain-specific genes encoded for hypothetical

proteins and mobile element proteins, which might play a crucial role in determining distinct virulence features of each strain.

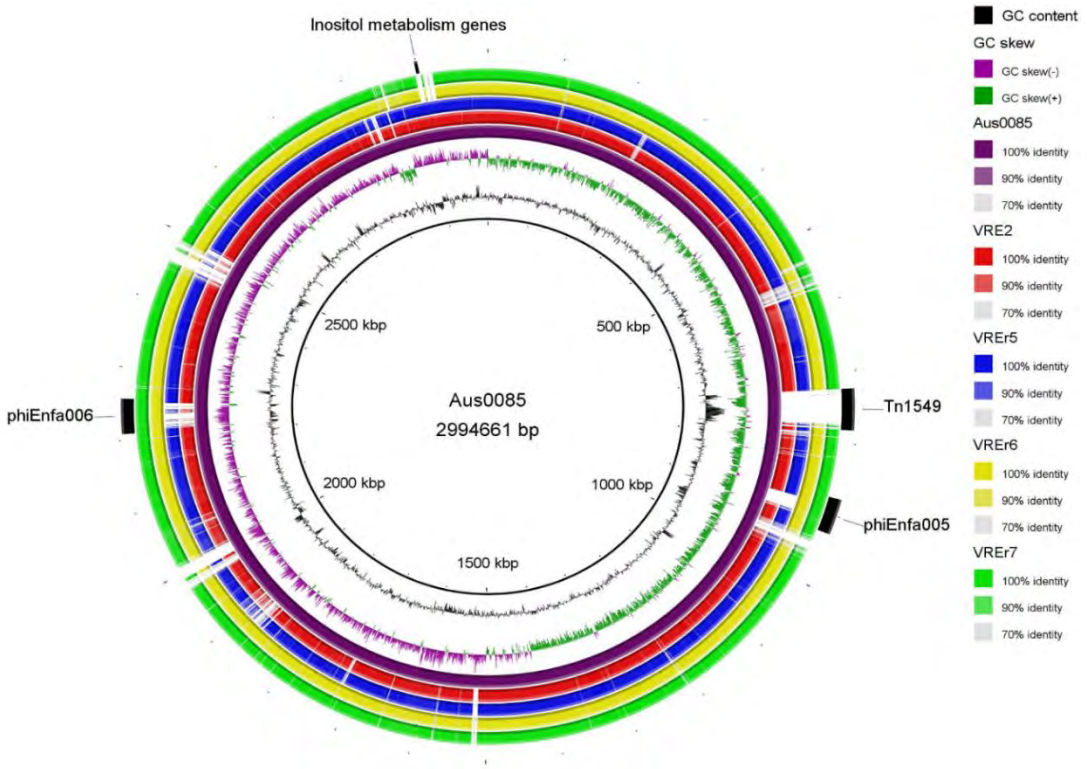


Figure 4.2: Circular genomic map and genome comparison of Aus0085, VRE2, VREr5, VREr6, and VREr7. The inner ring shows coordinate in scale and the total genome size of the reference sequence, Aus0085. The black histogram bar represents GC content whereas the purple-green histogram bar represents GC skew. Colored arches representing orthologous regions of each genome in respect to Aus0085 (purple arch) and are shown in the following order (inside to outside): Aus0085, VRE2, VREr5, VREr6, VREr7. The outermost arch (black) represents the location of Tn*1546*, phiEnfa005, phiEnfa006, and inositol catabolism genes relative to Aus0085.

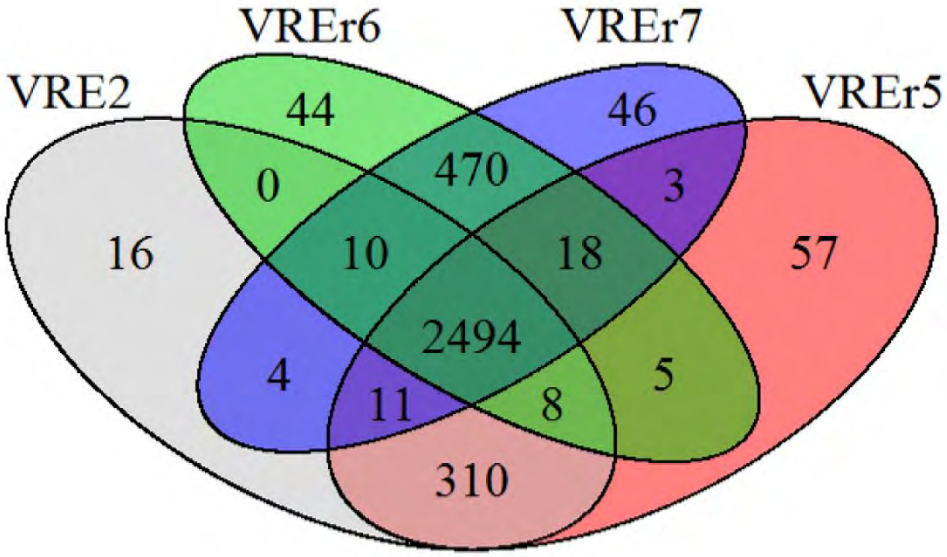


Figure 4.3: Venn diagram showing the distribution and number of core, dispensable and strain-specific genes of the Malaysian VRE_{fm} strains. Each circle is labelled with the name of its representative strain. A total of 2494 genes are shared among the analyzed strains. The number of strain-specific genes are 16, 44, 46, and 57 for VRE2, VREr6, VREr7, and VREr5, respectively.

4.1.5 Phylogenomic relationship of selected *E. faecium* strains

To better understand the relationship among the four Malaysian strains and other global strains, phylogenomic analysis was carried out using the core genome alignments of the four Malaysian strains in comparison with other 16 *E. faecium* strains from ten different countries. *E. faecium* Aus0085 was used as a reference for the analysis. Based on the phylogenomic analysis (Figure 4.4), VREr6 and VREr7 were closely related, which is expected as both were isolated from the same host (patient Y). However, VREr5, which was also derived from patient Y, was more closely related to VRE2 isolated from a different patient (patient X). The phylogenomic tree also showed that VREr6 and VREr7 were closely related to strain E161 isolated from China. Notably, these two strains were also more closely related to the reference Aus0085 compared to VRE2 and VREr5. Additionally, a distinct separation of strains of ST17, ST18, and ST78 into two different clusters was also observed. Strains from ST17 (Aus0004, E155, EnGen0180) and ST18 (DO, E1731) were clustered together and were closely related to the Malaysian strains, VRE2 and VREr5. Two strains from ST78 (E6045, E2560) formed another cluster which included VREr6 and VREr7.

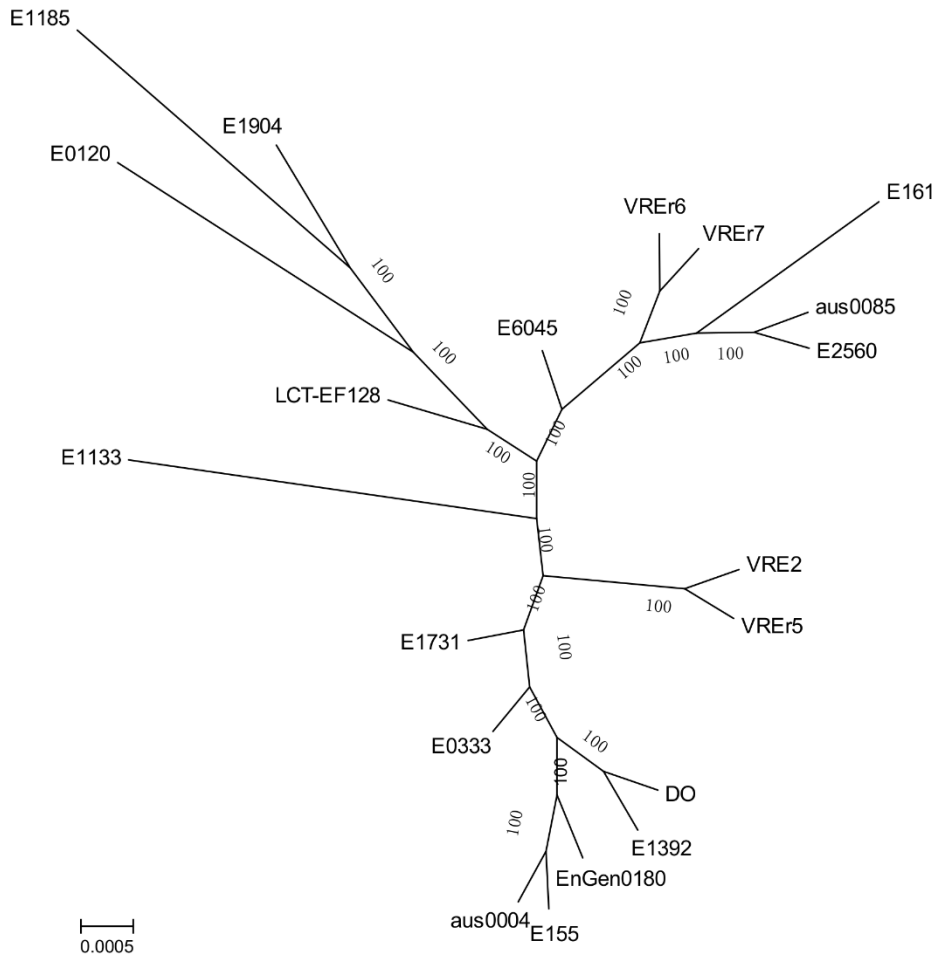


Figure 4.4: Phylogenomic tree inferred from approximately-maximum-likelihood method from aligned core genomes. Multiple genomes alignments are generated from 20 global *E. faecium* strains using Aus0085 as a reference. The unrooted phylogenomic tree is inferred via approximately-maximum likelihood method using FastTreeMP (Price et al., 2010). Bootstrap support values are shown in each node.

4.1.6 Genomic plasticity

The genomes of clinical *E. faecium* strains are known to plague with a high number of mobile genetic elements which contribute to their genome plasticity (Qin et al., 2012; van Schaik et al., 2010). Analyses of the four VRE_{fm} genomes revealed high genome dynamic among these strains contributed mainly by insertion sequences (IS), prophages, and Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR).

The number of IS elements found in the four genomes ranged from 43 to 55. These IS elements represent ten types of IS families, with members from the IS*3* (IS*1485*, IS*Efa8*, IS*Efa10*, IS*Enfa3*) and ISL*3* (IS*1251*, IS*Efa11*, IS*1476*, IS*Efa5*) families being the majority IS elements identified. Apart from that, IS*16*, which is ubiquitously found in clinical enterococcus strains (Werner et al., 2011), was also detected in the genomes of the four local strains.

One prophage region was predicted in both VRE2 and VREr5 whereas two and three prophage regions were predicted in VREr6 and VREr7, respectively. The size of the predicted prophages ranged from 39.0 kb to 47.8 kb. Interestingly, the identified phages showed higher sequence similarity (99% to 100% nucleotide identity) with those from strains of the same sequence type. This included two regions in VREr6 and VREr7 which shared 93% to 95% sequence identity with phages phiEnfa005 and phiEnfa006 of Aus0085 (Figure 4.2). Majority of the ORFs (83.3% to 98.0%) in all the predicted prophage regions encoded for phage-specific proteins and hypothetical proteins. One prophage region of VREr7 was structurally similar with that of VRE2 and VREr5, except for an additional 10.8 kb region containing genes which encode for glycopeptide resistance proteins, cadmium transporter, and RelB/RelE toxin-antitoxin system. This region was flanked by two transposases, suggesting that these additional genes were being acquired later by the prophage.

CRISPRs are repetitive sequences that, together with CRISPR-associated (cas) genes, protect bacteria against integration of exogenous DNA such as phages and plasmids into their genomes (Rakus et al., 2007). Clinical *E. faecium* are known to lack CRISPR loci (Palmer and Gilmore, 2010). This probably contributed to the higher rate of exogenous DNA in their genomes compared to those of non-clinical *E. faecium* isolates. No true CRISPR was identified in any of the studied strains. Although several CRISPR-like regions (designated as questionable CRISPR by CRISPRs Finder) were found, they probably carry no functional significance, as suggested by van Schaik et al. (2010).

4.2 Virulence factors

4.2.1 Virulence genes profiling

With the increasing incidents of VRE_{fm} infections, knowledge on the virulence and pathogenicity of this organism is important for disease treatments. Various virulence determinants could be identified from the genome sequences of the four studied VRE_{fm} strains by using the local database of enterococcal virulence genes. The established virulence genes such as *esp* and *acm* were identified in the genomes of the four strains (Table 4.2). Besides, genes encoding endocarditis specific antigen EfaA and a sugar transcriptional regulator BopD were also detected in the Malaysian VRE_{fm} strains. Additionally, these strains also harbored two capsule-related genes, *cpsA* and *cpsB*, which may provide them with resistance against phagocytosis in the host cells. An *epa* operon consists of 15 genes previously reported in *E. faecium* (Qin et al., 2012) was also determined. Apart from that, several genes encoding *E. faecium* surface proteins (*fms*) were also identified (Table 4.2). However, the *gelE* gene encoding gelatinase associated with biofilm formation was not identified in the four genomes. Similarly, hyl_{Efm} which

encodes for a glycoside hydrolase putatively involved in host colonization (Rice et al., 2009) was also not detected in any of the studied strains.

Although the four VRE_{fm} strains shared a similar set of virulence genes (Table 4.1), differences were still observed between the strains. For example, VREr6 and VREr7 differed from other two strains through the absence of the *fms21-fms20* gene cluster. On the other hand, VREr5 harboured a *bee*-like locus which was not detected in other analyzed strains. These putative virulence genes were responsible for the attachment of pathogen to both biotic and abiotic surfaces, which could be essential in the persistence and pathogenesis of *E. faecium*.

Table 4.2: Selected virulence-associated genes identified in the four Malaysian VRE_{fm} strains.

Gene	Function	Strain			
		VRE2	**VREr5**	**VREr6**	**VREr7**
esp	Biofilm formation	+	+	+	+
acm	Collagen binding	+	+	+	+
gelE	Biofilm formation	-	-	-	-
Hyl_{Efm}	Colonization/invasion	-	-	-	-
efaA	Adherence	+	+	+	+
$ebpABC_{fm}$ locus	Biofilm-formation	+	+	+	+
bee locus	Biofilm formation	-	+	-	-
pilA (*fms21*)-*fms20* locus	Adherence	+	+	-	-
fms11-19-16 locus	Adherence	+	+	+	+
fms14-17-13	Adherence	+	+	+	+
fms18 (*ecbA*)	Adherence	+	+	+	+
fms15	Adherence	+	+	+	+
bopD	Biofilm formation	+	+	+	+
epa operon (*epaA, epaB, epaC, epaD, epaE, epaF, epaG, epaH, epaL, epaM, epaN, epaO, epaP, epaQ, epaR*)	Biosynthesis of cell-surface polysaccharides	+	+	+	+
cpsA	Antiphagocytosis	+	+	+	+
cpsB	Antiphagocytosis	+	+	+	+

"+" indicates the presence of virulence genes whereas "-" indicates the absence of virulence genes.

4.2.2 Biofilm formation

The biofilm forming ability of the four VRE_{fm} strains was determined using crystal violet assay. Based on the interpretation scheme adopted from Stepanović et al. (2000), VREr5 was the only strain that was able to form biofilm (Table 4.3). To further confirm this result, CLSM was performed. The acquired images showed that VREr5 formed dense biofilm after 48hr of incubation (Figure 4.5), supporting the results observed from the crystal violet assay. Interestingly, unlike VRE2 and VREr6 which could not form biofilm, VREr7 was observed to form thin monolayer in the CLSM image acquired. This observation is possibly due to the different abiotic surfaces used to grow biofilm in the crystal violet assay and for CLSM imaging. The effect of different abiotic substrates on bacterial biofilm formation had been demonstrated in other study (Tendolkar et al., 2004).

Table 4.3: Average results of three replicates of the crystal violet assay to determine biofilm forming potential

Strain	OD_{590}	Interpretation*
VRE2	0.016	Non-biofilm producer
VREr5	1.780	Strong-biofilm producer
VREr6	0.049	Non-biofilm producer
VREr7	0.083	Non-biofilm producer

* Interpretation based on ODc of 0.218.

The cut-off OD (ODc) was defined as three standard deviations above of the mean OD of the negative control. The biofilm forming ability of each strain was scored as follow: OD $\leq$ ODc = non-biofilm producer, ODc < OD $\leq$ (2x ODc) = weak-biofilm producer, (2x ODc) < OD $\leq$ (4x ODc) = moderate-biofilm producer, OD > (4x ODc) = strong-biofilm producer (Chelvam et al., 2014; Stepanović et al., 2000).

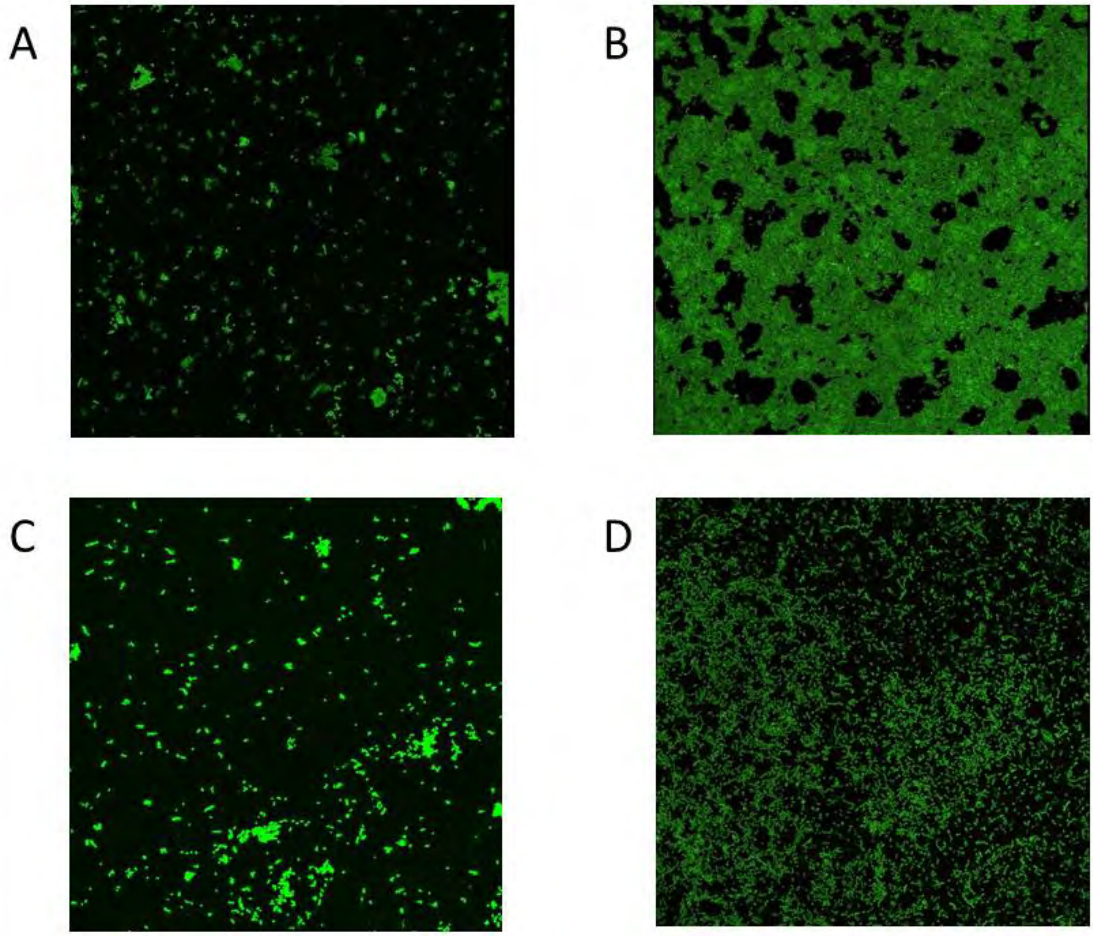


Figure 4.5: CLSM images of the four Malaysian VRE_{fm} strains grown in TSB. Each image represents the "flatten" three-dimensional (3D) Z-projection of stack images of A) VRE2, B) VREr5, C) VREr6, and D) VREr7. The non-biofilm formers (VRE2, VREr6) were either scattered around or formed small aggregates on the glass slide. VREr7 formed thin monolayer whereas VREr5 formed dense biofilm.

4.2.3 *E. faecalis bee* homolog

Since VREr5 was the only biofilm former, further analysis was performed to identify unique gene(s) that might contribute to this varying phenotype. Examination on the list of unique genes of VREr5 revealed a five-gene cluster which was homologous to the *bee* (*b*iofilm *e*nhancer in *e*nterococcus) locus of *E. faecalis*.

In *E. faecalis*, the *bee* locus composes of five genes: *bee*-1, *bee*-2, *bee*-3, *str*1, and *str*2. The three *bee* genes encode for putative cell wall-anchor proteins whereas the two *str* genes encode for sortases. Insertion mutation of the *bee-2* gene shows 70% reduction of biofilm formation in *E. faecalis* strain E99, indicating positive association of this gene cluster to biofilm formation (Tendolkar et al., 2006).

The unique gene-cluster of VREr5 shared approximately 99% amino acid identity with Bee-2, Bee-3 and the two sortase proteins of the *bee* locus but only 39% amino acid identity with Bee-1. Despite its low amino acid similarity, the structure of the putative Bee-1 homolog of VREr5 was similar to that of the Bee-1 protein. These included a region from residues 705 to 810 that showed low degree of similarity (E=1.08e-03) to collagen binding B domain of *Staphylococcus aureus*, and a region from 338 to 469 that encodes for the Von Willebrand factor type A (VWA) domain. Given the observed sequence and structural similarity, the *bee* homolog of VREr5 might serve the same role as that found in *E. faecalis*.

Coburn et al. (2010) reported that the *bee* locus of *E. faecalis* is located on a large conjugative plasmid. To study the location of the *bee* homolog of VREr5, PCR for the determination of *bee*-1, *bee*-2 and *bee*-3 gene was performed, using both total and plasmid DNA as the templates. The results showed that all the three genes were being amplified

with the expected amplicon sizes on both total and plasmid DNA (Figure 4.6), suggesting that this locus most probably located on a plasmid.

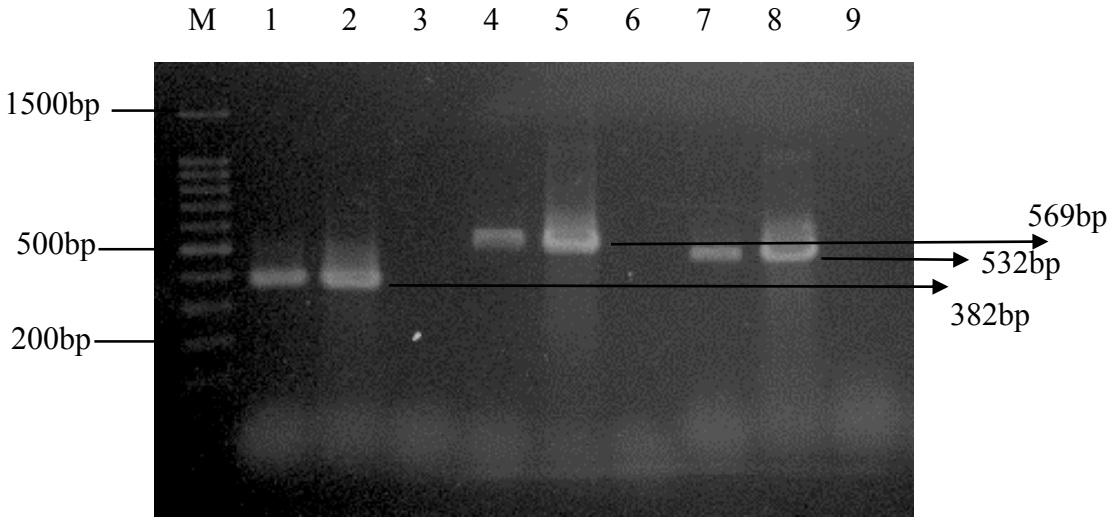


Figure 4.6: PCR results showing amplification of *bee*-1, *bee*-2 and *bee*-3 gene using total and plasmid DNA as template. Lane M: 100bp Molecular marker; lane 1: bee-1_total DNA (382bp); lane 2: *bee*-1_plasmid DNA (382bp); lane 3: *bee*-1_negative control (no DNA added); lane 4: *bee*-2_total DNA (569bp); lane 5: *bee*-2_plasmid DNA (569bp); lane 6: *bee*-2_negative control (no DNA added); lane 7: *bee*-3_total DNA (532bp); lane 8: *bee*-3_plasmid DNA (532bp); lane 9: *bee*-3_negative control (no DNA added)

4.3 Antibiotic resistance

4.3.1 Antibiotic susceptibility profile

The antibiotic susceptibility profiles of the four analyzed strains is shown in Table 4.4. The results showed that these strains were resistant to most of the antibiotics tested. All of them were resistant to ampicillin, kanamycin, streptomycin, erythromycin, clindamycin and tetracycline. VRE2 was susceptible to gentamicin, whereas the other three strains showed high-level resistance to gentamicin (M.I.C > 256 µg/ml). VRE2 and VREr5 showed intermediate resistance to teicoplanin whereas VREr6 and VREr7 were resistant to the antibiotic. All four strains were susceptible to linezolid. The MIC values indicated that all of them were highly resistant to vancomycin (M.I.C > 256 µg/ml).

Table 4.4: Antimicrobial resistance profile and the presence of corresponding resistance genes.

	VAN		**TEC**		**KAN**			**GEN**		**STM**	
	R/I/S	***vanA***	**R/I/S**	***vanZ***	**R/I/S**	***aac(6')-Ii***	***aph(3')-III***	**R/I/S**	***aac(6')-aph(2'')***	**R/I/S**	***ant(6)-Ia/aadE***
VRE2	R	+	I	+	R	+	+	S	-	R	+
VREr5	R	+	I	+	R	+	+	R	+	R	+
VREr6	R	+	R	+	R	+	+	R	+	R	+
VREr7	R	+	R	+	R	+	+	R	+	R	+

Table 4.4 continued

	AMP		**ERY**		**CLI**		**TET**			**LZD**	
	R/I/S	**Mutated *pbp5***	**R/I/S**	***ermB***	**R/I/S**	***msrC***	**R/I/S**	***tetS***	***tetM***	**R/I/S**	**G2576T mutation in 23S rRNA**
VRE2	R	+	R	+	R	+	R	+	-	R	-
VREr5	R	+	R	+	R	+	R	+	-	R	-
VREr6	R	+	R	+	R	+	R	-	+	R	-
VREr7	R	+	R	+	R	+	R	-	+	R	-

Resistance gene(s) or mutation associated with each antibiotic is (are) listed below the abbreviations of respective antibiotics. VAN: vancomycin, TEC: teicoplanin, KAN: kanamycin, GEN: gentamicin, STM: streptomycin, AMP: ampicillin, ERY: erythromycin, CLI: clindamycin, TET: tetracycline, LZD: linezolid. Plus (+) and minus (-) signs indicate presence and absence of the resistance genes or mutation, respectively. The R/I/S column gives information regarding the susceptibility of each strain to respective antibiotics: R = resistant, I = intermediate, S = susceptible.

4.3.2 Antibiotic resistance determinants

Enterococci can be resistant to virtually all antibiotics commonly found in the clinical settings. The underlying antibiotic resistance mechanisms of this organism can be intrinsic or acquired through mutation of the intrinsic resistance genes or through horizontal acquisition of the resistance determinants.

Both *vanA* and *vanB* are the predominant genes that confer acquired vancomycin resistance in enterococci. VanA phenotype is characterized by high level of inducible resistance to both vancomycin and teicoplanin whereas VanB phenotype shows inducible resistance to vancomycin but variable resistance to teicoplanin (Arthur & Courvalinn, 1993; Hollenbeck & Rice, 2012). All the local VRE_{fm} strains carried the *vanA* gene, which was validated through PCR (Figure 4.7). Despite carrying the *vanA* genotype, only VRE2 and VREr5 displaced VanA phenotype (vancomycin MIC > 256 µg/ml, teicoplanin MIC =16 µg/ml). In contrast, VREr6 and VREr7 were VanB phenotype-*vanA* genotype VRE strains (vancomycin MIC > 256 µg/ml, teicoplanin MIC =8 µg/ml). The impairment of teicoplanin resistance in *vanA* genotype VRE strains can be due to point mutations of *vanS* or disruption of *vanY* or *vanZ* (Gu et al., 2009; Hashimoto et al., 2000). However, no mutations or disruptions of these regions were observed in VREr6 and VREr7. Further analysis is needed to identify the cause of this impairment.

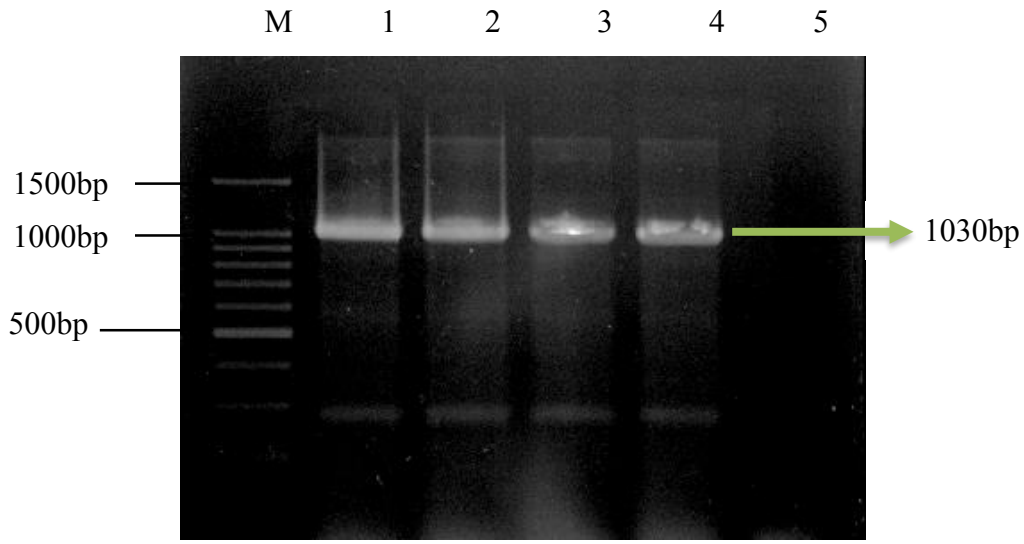


Figure 4.7: PCR results showing amplification of vanA gene. Lane M: 100bp molecular marker; lane 1: VRE2; lane 2: VREr5; lane 3: VREr6; lane 4: VREr7; lane 5: negative control (no DNA added).

The *vanA*-carrying Tn*1546* transposon varies structurally among strains (Gagnon et al., 2011; Huh et al., 2004; Willems et al., 1999). This polymorphic nature of the transposon has been employed in molecular typing of VRE to understand the dissemination of vancomycin resistance (Huh et al., 2004; Willems et al., 1999). Structural analysis revealed that the Tn*1546*-like transposons of the four studied strains were similar. All of them carried a truncated *orf1* which encodes a transposase. One IS*1216V* was found to be inserted directly before *orf2*. Moreover, insertion of an IS*1251* was observed in the intergenic region of *vanS* and *vanH*. The structure observed resembled type F reported by Willems et al. (1999).

Resistance to aminoglycoside is commonly observed in VRE strains. High-level resistance to aminoglycoside is mediated by three types of aminoglycoside modifying enzymes: N-Acetyltransferases (AAC), O-Adenyltransferases (ANT), and O-

phosphotransferases (APH) (Hollenbeck & Rice, 2012). Three genes of these enzymes, *ant(6)-Ia, aac(6')-Ii* and *aph(3')-III* were found in all four VRE_{fm} strains. *aac(6')-Ii* confers low level intrinsic resistance to tobramycin and kanamycin whereas *ant(6)-Ia* and *aph(3')-III* confer acquired resistance to streptomycin and kanamycin, respectively. The *ant(6)-Ia* (also known as *aadE*) is part of the multi-resistant cluster of *ant(6)-sat4-aphA*, where *sat4* confers resistance to streptothricin. In addition, a bifunctional gene, *aac(6')-aph(2'')*, which confers high-level acquired resistance to gentamicin was also detected. VRE2 was the only strains that did not harbor the bifunctional gene. This might explain its susceptibility towards gentamicin (M.I.C = 4μg/ml). In contrast, the other three strains were highly resistant to gentamicin (M.I.C > 256μg/ml).

E. faecium is intrinsically resistant to β-lactam drugs through the expression of low-affinity penicillin-binding protein 5 (PBP5). However, acquisition of β-lactamase or mutation of *pbp5* can contribute to increased resistance of the antibiotic. Four β-lactamases were identified in each of the studied strains. These included two from metallo-β-lactamase superfamily, one putative class C β-lactamase and one from unknown-class β-lactamase. Besides, mutations in *pbp5* which associated with high-level β-lactam resistance in *E. faecium* were also observed. These mutations included an insertion of a serine residue at position 466, after a serine residue; replacement of methionine at position 485 to alanine and replacement of glutamic acid at position 629 by valine. All these mutations have been demonstrated to increased MICs of β-lactam drugs (Rice et al., 2004).

The four local strains carried *ermB* gene which is associated with erythromycin resistance. Furthermore, the *msrC* gene encoding an ABC efflux pump for macrolides and streptogramin B (Singh et al., 2001) was also identified in all the studied strains. VRE2 and VREr5 harbored *tetS* gene whereas VREr6 and VREr7 carried *tetM* gene. Both

these genes confer tetracycline resistance through ribosomal protection (Charpentier et al., 1993; Martin et al., 1986).

4.4 Transcriptomic analysis

4.4.1 Biofilm developmental stage after 24 hours of growth

A number of enterococcal infections are associated with difficult-to-treat biofilm-mediated infections, such as those associated with indwelling medical devices, urinary catheters, orthopaedic implants, and endocarditis (Donlan et al., 2002). As such, understanding the biofilm forming mechanism, especially during the transition from planktonic to biofilm cells, is useful in controlling biofilm-associated diseases.

Since previous CLSM analysis (Figure 4.5) showed that VREr5 formed mature biofilm after 48 hr of growth, it was speculated that a shorter incubation time might correspond to the initial stage of biofilm formation. VREr5 biofilm was therefore grown for only 24 hr and the actual developmental stage was visualized using CLSM. The result showed that after 24hr of incubation, VREr5 formed cell aggregates that were loosely packed (Figure 4.8). This result indicated that at 24 hr of growth, VREr5 was in the initial stage of biofilm formation where cell attachment and aggregation occur.

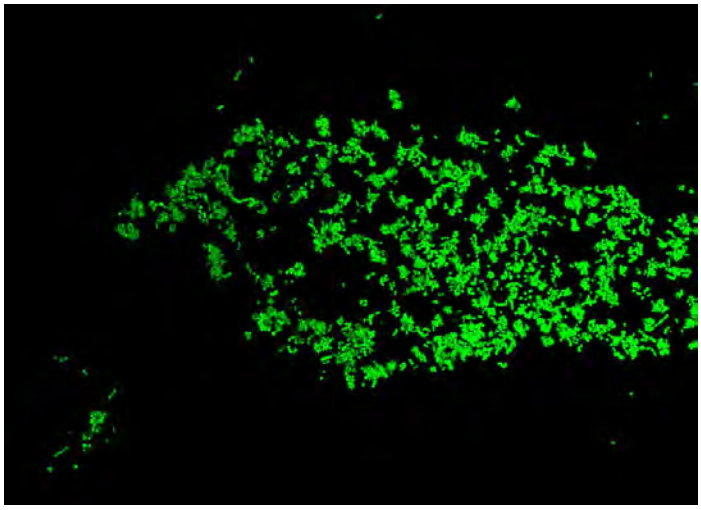


Figure 4.8: CLSM image of the biofilm formed by VREr5 after 24 hr of incubation. The image represents the "flatten" three-dimensional (3D) Z-projection of stack images acquired. VREr5 formed cell aggregates that were loosely packed at 24 hr of growth.

4.4.2 Transcriptome assembly and annotation

After removing the adaptors and low-quality sequences, a total of 7,786,583 to 11,913,906 reads were generated. Due to the lack of an appropriate reference genome, *de novo* assembly was performed, which generated a total of 1,234 transcripts. The assembled transcripts ranged from 301bp to 134,410bp, with an average of 2,887bp. From these transcripts, 2,851 coding sequences (CDSs) were generated. Out of the 2,851 CDSs, 2,833 (99.3%) were matched to the known protein sequences of the Nr database (Table 4.5).

Gene Ontology (GO) terms were subsequently assigned to the obtained CDSs. A total of 2,404 (84.3%) CDSs were assigned (Table 4.5). 1257 (44.09%) CDSs were in the cellular component category, 2001 (70.19%) CDSs were in the molecular function category, and 1832 (64.26%) CDSs were in the biological process category. Apart from that, a total of 629 (22%) CDSs were mapped to the KEGG pathways, with majority of

the CDSs mapped to the biosynthesis of antibiotics, purine metabolism, pyrimidine metabolism, and the glycolysis/gluconeogenesis pathways. To further study the functionality of the transcriptome, the annotated CDSs were also classified into different functional groups based on the COG database. Among the 2,851 CDSs, 2,411 (84.5%) were categorized into 19 COGs (Table 4.5). The largest functional group was "function unknown", which comprised of 559 (22.9%) CDSs, followed by "carbohydrate transport and metabolism" (301, 12.33%), "transcription" (207, 8.48%), "replication, recombination and repair" (195, 8%), and "cell wall/membrane/envelope biogenesis" (164. 6.72%) (Figure 4.9).

Table 4.5: Summary of annotations of the assembled CDSs in VREr5

Database	**Number of CDSs**	**Percentage**
Nr	2,833	99.3%
GO	2,404	84.3%
KEGG	629	22.0%
COG	2,411	84.5%
Total CDSs	**2, 851**	**100%**

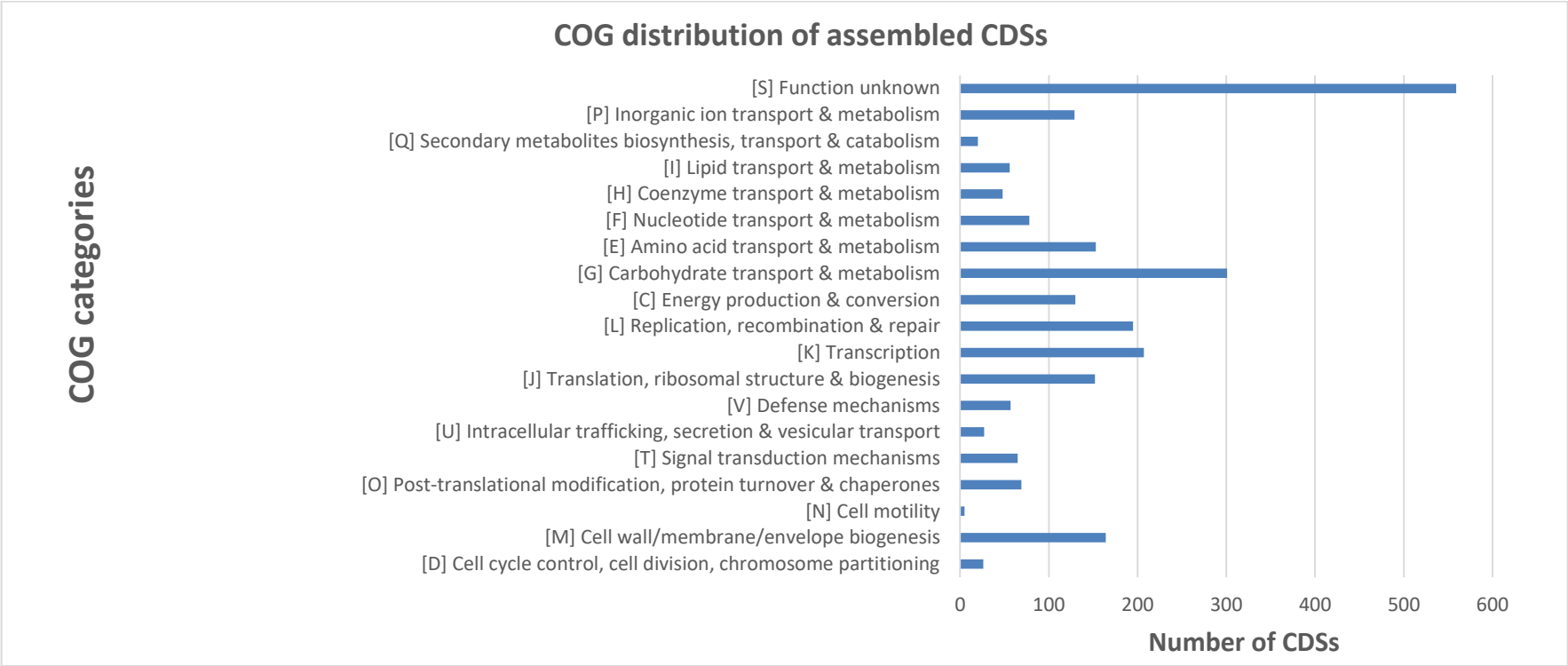


Figure 4.9: COG distribution of the assembled CDSs. The x-axis represents the number of CDSs whereas the y-axis represents the COG categories. The largest functional group is the COG group S (Function unknown) whereas the smallest group is COG group N (Cell motility).

4.4.3 Transcriptional profiles of biofilm cells

Principal component analysis (PCA) is a statistical approach to simplify a multidimensional data set into a new set of variables (the principal components) which explain the key differences of the original data set (Abdi & Williams, 2010). This technique has been used in the analysis of gene expression data (Guilhen et al., 2016; Hilsenbeck et al., 1999). In this study, PCA revealed that the first principal component (PC 1) and the second principal component (PC 2) accounted for approximately 65.5% and 31.0% of the total variance of the data set, respectively (Figure 4.10). This result clearly indicated that cells in the planktonic and biofilm conditions exhibited different gene expression profiles. These differences were also depicted in the heat map generated from the clustering of all the six samples based on their relative distance calculated using Manhattan distance (Figure 4.11). Due to the large variance observed as compared to other planktonic samples, P5 was excluded for the later analysis. To ensure equal numbers of samples were used in the following comparison, B1, one of the biofilm samples which showed slightly higher variance, was excluded for the later analysis as well.

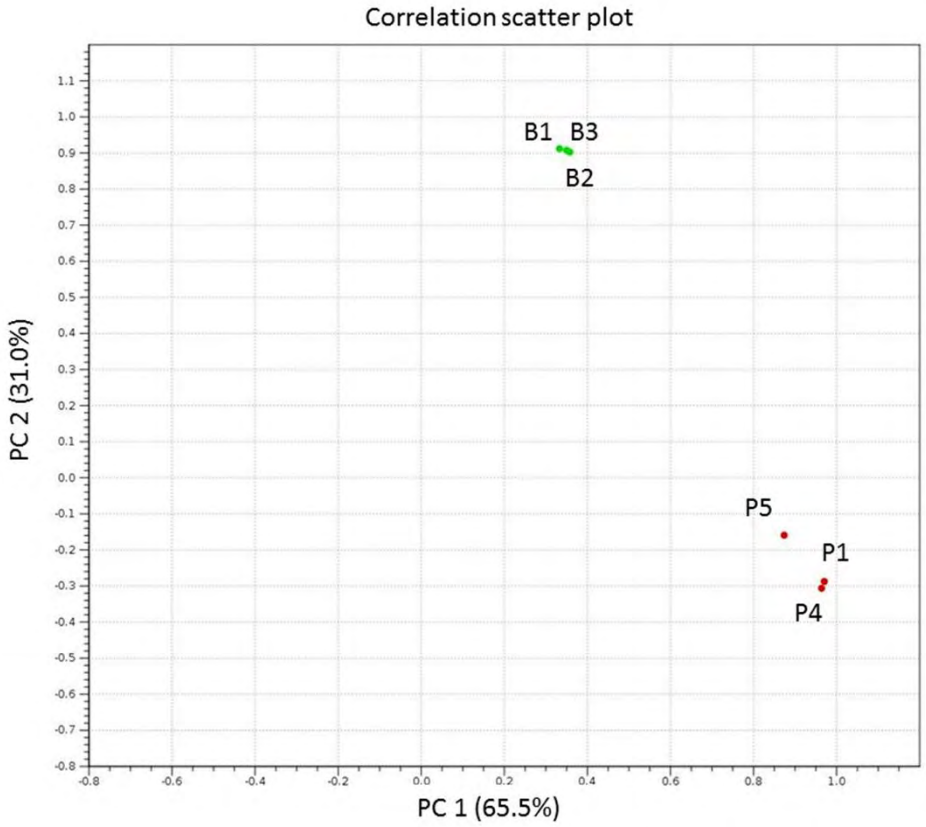


Figure 4.10: Principal component analysis (PCA) of gene expression in biofilm and planktonic cells. PCA was performed with the normalized expression values of the 2,851 CDSs. Each dot indicates a biological replicate. Green dots (B1, B2, B3) represent the biofilm samples, whereas red dots (P1, P4, P5) represent the planktonic samples.

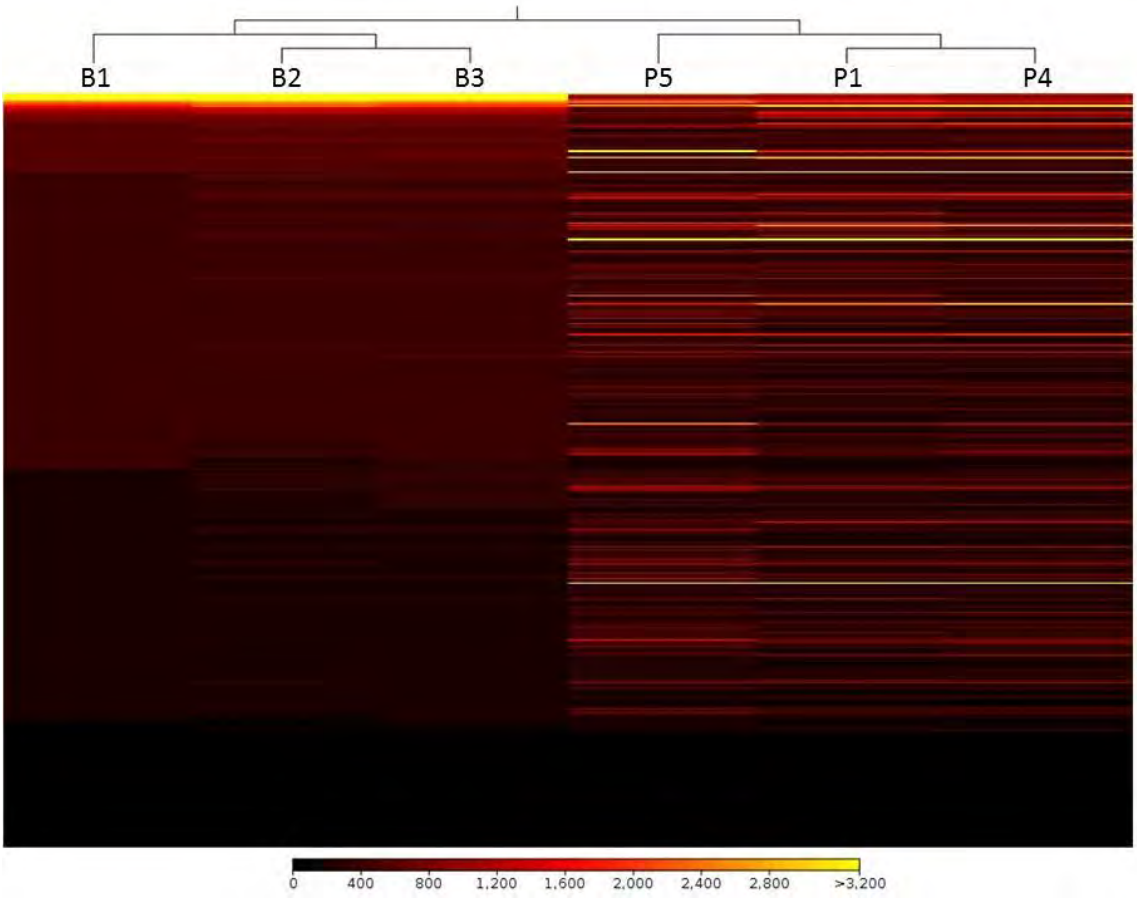


Figure 4.11: Heat map depicting the gene expression level of the six samples (three biological replicates for each growth condition). B1, B2, and B3 represent the biofilm samples whereas P1, P4, and P5 represent the planktonic samples. The bar color reflects the gene expression level from low (black), medium (red), to high (yellow). The dendrogram shows the clustering of the six samples based on similarity of their transcription patterns.

The transcriptome of the biofilm vs planktonic phase cells showed that a total of 776 genes were differentially expressed (|fold change |> 2, FDR-adjusted *P*-value < 0.001). Among these, 177 genes were up-regulated whereas 599 genes were down-regulated. COG analysis revealed that majority of the up-regulated genes were related to replication, recombination and repair (33 genes), carbohydrate transport and metabolism (23 genes), and transcription (13 genes). Moreover, genes classified as COG class S (function unknown) also represented a considerable part (29 genes) of the up-regulated genes.

A large proportion of the genes that were being induced in the biofilm mode were those that encode for transposase, recombinase and integrase, with fold change of up to 5.23 relative to the planktonic counterpart (Table 4.6). Apart from that, genes that were associated with plasmid replication were also up-regulated for 4.82 fold. Additionally, the expression of *ebpABC* encoding pilus subunit proteins increased by 2.52 fold. Several other genes putatively encoding cell wall anchor proteins, such as those carrying the LPXTG domain, were also up-regulated. Furthermore, the *tetS* gene encoding tetracycline resistance protein was also up-regulated for approximately three fold. Interestingly, two genes associated with quorum sensing, *fsr* and *luxS*, were highly down-regulated, with fold change of -133.2 and -35.9, respectively (Table 4.7). Moreover, the expression level of three genes, *arcABC*, which involves in arginine catabolism, decreased by 287.52, 259.23, and 96.1, respectively. The *spx* gene, which encodes for protein generally involves in oxidative stress response was also down-regulated for approximately 12 fold. Several other genes that are previously known to involve in enterococcal biofilm formation were either not differentially expressed or down-regulated. For example, the *esp*, *acm*, and *sgr*A, were not differentially expressed whereas the *sagA* and alt_{Efm} were down-regulated. The RNA-seq results were validated by qPCR using selected up- and down-regulated genes (Table 4.8).

Table 4.6: Selected up-regulated genes in biofilm relative to planktonic cells

Feature ID	Annotation	FDR *p*-value	Fold change
Gene.2189	IS6 family transposase	1.25E-30	5.23
Gene.1209	Plasmid replication initiation	5.91E-28	4.82
Gene.1210	Hypothetical protein	1.38E-26	4.75
Gene.417	Transposase	7.79E-26	4.36
Gene.458	Integrase	5.99E-25	4.25
Gene.889	Recombinase	8.65E-23	3.94
Gene.457	Plasmid replication protein repR	1.61E-21	3.89
Gene.888	DNA recombinase	6.05E-19	3.53
Gene.1866	Transposase for insertion sequence element IS256 in transposon	2.06E-17	3.29
Gene.1712	Transposase	4.13E-17	3.21
Gene.462	Tetracycline resistance tetS	1.40E-14	2.99
Gene.991	LPXTG-domain-containing cell wall anchor domain	1.47E-17	2.83
Gene.398	Phage head-tail adaptor	9.27E-08	2.81
Gene.1333	Initiator RepB plasmid replication protein	2.96E-13	2.74
Gene.2506	von Willebrand factor (EbpA)	8.54E-19	2.52
Gene.2511	Cell surface protein EbpB	1.78E-12	2.47
Gene.2523	AraC family transcriptional regulator	2.24E-12	2.38
Gene.2287	Hypothetical protein	2.24E-06	2.34
Gene.914	TcpC-containing conjugal transfer protein	5.39E-11	2.32
Gene.435	Mannonate dehydratase	4.13E-16	2.29
Gene.2509	Cell surface protein EbpC	1.73E-13	2.26
Gene.2436	PTS mannose transporter subunit IID	9.20E-09	2.24
Gene.895	Amino acid transporter	8.28E-17	2.2
Gene.2682	MarR family transcriptional regulator	2.48E-06	2.17
Gene.2102	Cell surface protein	8.70E-09	2.15
Gene.1013	GntR family transcriptional regulator	4.32E-11	2.12
Gene.2215	ABC transporter permease	6.46E-13	2.11
Gene.956	Hypothetical protein	6.81E-11	2.09
Gene.1296	LPXTG-domain-containing cell wall anchor protein	1.25E-10	2.08
Gene.902	Glucuronate isomerase	2.69E-13	2.06

Table 4.7: Selected down-regulated genes in biofilm relative to planktonic cells

Feature ID	Annotation	FDR *p*-value	Fold change
Gene.802	Arginine deiminase ArcA	3.09E-219	-287.52
Gene.805	Ornithine carbamoyltransferase ArcB	2.04E-217	-259.23
Gene.813	Carbamate kinase ArcC	2.71E-186	-96.1
Gene.502	Accessory regulator FsrB	8.83E-198	-133.22
Gene.1015	Alkyl hydroperoxide reductase subunit C	1.39E-111	-43.47
Gene.1973	C4-dicarboxylate anaerobic carrier	3.1E-144	-39.89
Gene.1598	S-ribosylhomocysteinase LuxS	5.53E-96	-35.99
Gene.138	Universal stress protein	4.82E-116	-27.56
Gene.1755	Bee-3-like protein	8.98E-124	-26.91
Gene.2579	Sugar ABC transporter ATP-binding	2.21E-30	-21.6
Gene.1616	Alpha-glycerophosphate oxidase	2.26E-48	-19.43
Gene.1549	NADH peroxidase	4.98E-82	-17.53
Gene.1761	Peptidase (Bee-2)	2.35E-95	-15.33
Gene.2143	Lactose-specific phosphotransferase enzyme IIA component	6.5E-84	-15.3
Gene.139	Regulatory Spx	0	-12.04
Gene.1978	Ribosomal S1	4.46E-57	-11.81
Gene.622	Formate acetyltransferase	7.7E-31	-11.55
Gene.1500	ATP-dependent Clp protease ATP-binding subunit	6.38E-46	-9.66
Gene.1752	VWA domain-containing (bee-1)	1.76E-71	-9.15
Gene.1291	Flavin reductase	2.32E-69	-8.4
Gene.1111	dTDP-glucose 4,6-dehydratase	2.4E-34	-8.03
Gene.1363	Secreated antigen SagA	7.34E-24	-6.07
Gene.2727	Acetyltransferase	8.82E-72	-5.87
Gene.553	Autolysin Alt	1.22E-08	-2.05
Gene.1113	Glucose-1-phosphate thymidylyltransferase	2.44E-08	-2.02

Table 4.8: Expression levels of selected up- and down-regulated genes in planktonic and biofilm cells as measured by qPCR.

Gene	Expression level in planktonic cells*	Expression level in biofilm cells
ebpA	1 ± 0.013	1.357 ± 0.012
tetS	1 ± 0.038	1.657 ± 0.073
repR	1 ± 0.013	1.449 ± 0.067
arcA	1 ± 0.309	0.647 ± 0.102
fsrB	1 ± 0.071	0.859 ± 0.050
bee-2	1 ± 0.096	0.613 ± 0.300

*The expression level of each of the six genes were defined with respect to that of the planktonic cells, defined as 1.

CHAPTER 5: DISCUSSION

5.1 Whole genome sequencing

The advent of next generation sequencing technology enables the whole genome sequences of bacteria to be determined in a faster and cost effective manner. Decoding the genomes of pathogens enabled the evaluation of their basic genomic features, virulence potential, and resistance mechanism, which are useful for infection control. Although many *E. faecium* genomes have been sequenced, there is still a lack of genome data from Southeast Asia, especially from the clinically important VRE_{fm}. In this study, four VRE_{fm} strains isolated from a tertiary hospital in Malaysia were selected based on the previous PFGE results. VREr6 and VREr7 isolated from patient Y displaced identical PFGE patterns. However, the PFGE profile of VREr5, which was also isolated from patient Y, was more similar to that of VRE2 isolated from another patient, patient X. In order to better elucidate the genetic differences, these strains were subjected to whole genome sequencing.

In silico MLST analysis classified the four strains into two sequence types: VRE2 and VREr5 belonged to ST80, whereas VREr6 and VREr7 belonged to ST203. These two sequence types are grouped under CC17, confirming the clinical importance of the four analyzed strains. Strains belonging to ST203 are commonly reported, including in Malaysia (Cha et al., 2012; Getachew et al., 2013; Lam et al., 2013; Yu et al., 2015). According to the records from the PubMLST database (last accessed on 23rd Jan 2017) (Homan et al., 2002), 49 entries on ST203 were found and are mainly represented by strains distributed in European (49%) and Asia Pacific countries (51%). In Australia, ST203 has been reported to replace ST17, the major founder of CC17, signifying the increasing clinical importance of this specific sequence type (Lam et al., 2013). In contrast, ST80 is rarely reported as compared to ST203. There were only 13 entries of

ST80 in the MLST database (last accessed on 23rd Jan 2017) which are mainly European strains, except for two entries with Asian strains. To the best of our knowledge, this study is the first to report ST80 VRE strain in Malaysia. The identification of ST80 has provided new insight into the dissemination of the high-risk CC17 in Malaysia. A study including more local strains will be useful to determine the prevalence and potential clinical significance of ST80 in Malaysia.

Consistent with the PFGE and MLST results, the comparative genomics analysis suggested that VREr6 and VREr7 have substantial variations from VRE2 and VREr5. The genomic contents of the four strains were further scrutinized to discover genes that may distinguish these two groups in terms of fitness and adaptive advantage. Two such genes were identified: a gene cluster encodes for *myo*-inositol metabolism proteins found only in both VRE2 and VREr5 and a gene encoding bacteriocin, lactococcin 972 (Lcn 972), found only in both VREr6 and VREr7.

Myo-inositol is commonly found in soil and can be used as a sole carbon source in various microorganisms such as *Rhizobium leguminosarum* (Fry et al., 2001) and *Bacillus subtilis* (Yoshida et al., 2008). Although several studies had been carried out on the biochemistry of this pathway, information on the molecular genetics of this pathway in bacteria is limited to *B. subtilis*. In *B. subtilis,* the *myo*-inositol catabolism pathway is mediated by the *iol* divergon that comprises of the operon *iolABCDEFGHIJ* and *iolRS* (Yoshida et al., 2008). The net result of this pathway creates acetyl-CoA which can participate in the citric acid cycle to generate energy. Hence, the ability to catabolize *myo*-inositol can be advantageous in conditions where other sugar sources are limited, such as in infected host where the rate of nutrient competition is high. The *myo*-inositol level is higher in blood and urine of patients with diabetes and diabetic-associated renal diseases (Hong et al., 2012). Therefore, pathogens that possess the ability to catabolize this sugar

have additional growth advantage to outcompete other bacteria in a diabetic patient. Although the functionality of the genes encoding inositol metabolism proteins identified in VRE2 and VREr5 was unknown, the fact that patient Y is a diabetic patient suggested the possible association of these genes in providing extra fitness to VREr5.

Bacteriocins are peptides produced by an organism to inhibit the growth of its closely related species (Héchard & Sahl, 2002). Lcn972 is known as an atypical nonlantibiotic bacteriocin which is heat sensitive and highly hydrophilic (Martínez et al., 1996). Instead of targeting the cytoplasmic membrane and forming pores, Lcn972 inhibits septum formation which leads to deformation and eventually lysis of cells (Martinez et al., 2000). Further study on the mechanistic action of this bacteriocin revealed that Lcn972 acts by interacting with cell wall precursor lipid II, which is the primary docking site for lantibiotics prior to pore formation (Martinez et al., 2008). Lcn972 is the first nonlantibiotic that specifically interacts with lipid II. Besides, Lcn972 also plays a role in the induction of prophages, although it is limited to specific prophage/host system (Madera et al., 2009). With its role in killing closely-related strains, the presence of Lcn972 in both VREr6 and VREr7 may increase the competitiveness of these two strains in colonizing and subsequently infecting the host.

The genomic differences among the four VRE_{fm} strains were further supported by the phylogenomic tree constructed. In concordance with the PFGE results, VREr5 was distantly related to VREr6 and VREr7 but was more closely related to VRE2. Apart from demonstrating the relationship among the local strains, the phylogenomic analyses using strains from CC17 also revealed a divergent of strains from ST17, ST18, and ST87 into two clusters. Members of the CC17 was initially thought to evolve from a single founder (ST17) based on the eBURST algorithm (Willems et al., 2005). However, this concept was later suggested to be erroneous by Turner et al. (2007). This group showed that the

reliability of eBURST to correctly infer ancestor-descendent link decreases in species with a high recombination to mutation ratio, such as in the case of *E. faecium*. This implies that hospital-associated strains which were previously grouped under CC17 might not be derived from a single founder, but possibly evolved independently from different ancestral strains. Subsequent analyses of the MLST data using Bayesian modelling method from the Bayesian analysis of genetic population structure (BAPS) software revealed that the three major ST of hospital-associated strains (ST17, ST18, ST78) were placed into different BAPS groups, with ST78 in BAPS 2-1 and ST17 and ST18 in BAPS group 3-3 (Willems et al., 2012). The result obtained from this study was in line with the BAPS analysis, supporting the idea that CC17 is not monophyletic.

Although PFGE, MLST, and phylogenomic analyses showed that VREr6 and VREr7 were clonally related, these two strains had diversified considerably. In addition to the differences in genome contents observed from the presence of strain-specific genes, the gain and/or loss of mobile genetic elements also plays a significant role in determining the genome dynamics of the two strains. This highlights the discriminatory power of WGS in detecting fine differences of the same clone and hence provides a better resolution in distinguishing strains with high genetic similarity.

Based on the available clinical data, patient Y underwent EVD and nasogastric tube exchange after the isolation of VREr5. Both procedures involved the insertion of catheters into the body. Together with the results obtained from the PFGE, MLST, phylogenomic, and comparative genomics analyses, the clinical data suggested that patient Y was probably infected by strains from different lineages. The two medical procedures might have introduced new VRE_{fm} strains to the patient through contaminated catheters or from hands of healthcare workers, thereby explaining the differences observed between VREr5, the initial strain isolated from patient Y, and the subsequent two strains, VREr6

and VREr7. Moreover, since patient Y was admitted when patient X was still in the hospital, the genetic similarity between VRE2 and VREr5 suggested that VRE2 could have persisted and spread in the hospital.

5.2 Virulence factors

A number of known and potential virulence factors were identified in the genomes of the four Malaysian strains. The presence of these virulence factors might explain the persistence of these organisms in the hospital settings. Furthermore, the identification of the *esp* gene in all the four genomes further confirm the previous notion that this virulence gene is significantly enriched in the clinical isolates (Abriouel et al., 2008; Willems et al., 2001).

Surface expressed proteins such as pili and microbial surface components recognizing adhesive matrix molecules (MSCRAMM) have been implicated in adherence, biofilm formation, and host tissue invasion in Gram-positive bacteria (Maisey et al., 2008; Nallapareddy et al., 2011). A total of 15 genes having the MSCRAMM features, such as the immunoglobulin-like fold, were previously identified in *E. faecium* TX16 (Nallapareddy et al., 2003; Sillanpää et al., 2009, 2008). Of these, 11 of them are grouped into four different clusters: *ebpA*$_{fm}$-*ebpB*$_{fm}$-*ebpC*$_{fm}$, *fms11-fms19-fms16*, *fms14-fms17-fms13,* and *fms21-fms20*. Among these, the *ebpABC*$_{fm}$ gene cluster encodes for biofilm-associated pili and is important in urinary tract infection in an animal model (Nallapareddy et al., 2011). On the other hand, the *fms21* (*pilA*) has been shown to form a pilus-like structure in *E. faecium* but its association to pathogenesis remains unknown (Hendrickx et al., 2008). In a study performed by Kim et al. (2010), the *fms21-fms20* cluster was found to be located on a large transferable plasmid (Kim et al., 2010). Hence,

VRE2 and VREr5 probably acquired this gene cluster horizontally, although the role of these genes in the virulence of these strains is unknown.

Cell surface polysaccharides can play an important role in the virulence and phagocytic resistance of Gram-positive bacteria. In *E. faecalis*, the *epa* gene cluster is involved in the production of an antigenic cell wall polysaccharide composed of rhamnose, glucose, galactose, N-acetylgalactosamine and N-acetylglucosamine (Teng et al., 2009). This gene cluster is important in biofilm formation, phagocytic resistance and tissue invasion in *E. faecalis* (Teng et al., 2009; Zeng et al., 2004). Similar gene cluster has been reported in *E. faecium* (Qin et al., 2012). However, the *epa* locus of *E. faecium* differs from that of *E. faecalis* in terms of the genes number (18 in *E. faecalis*, 15 in *E. faecium*) and order (Qin et al., 2012). Due to these differences, the polysaccharide synthesized by the products of *E. faecium epa* locus may have different sugar contents from that of *E. faecalis*. Nevertheless, the production of cell surface polysaccharides can provide pathogens with resistance against phagocytosis, enabling them to survive and persist in the host environment.

Despite the identification of several biofilm-associated genes in the four genomes, VREr5 was the only biofilm former among the studied strains. This observation suggested that the previously known biofilm-associated genes, such as *esp* and *ebpABC*$_{Efm}$, might not play a role in biofilm formation in the four local VRE$_{fm}$ strains. Alternatively, a combination of multiple factors and genes might be needed for the expression of biofilm phenotype in these strains. The presence of a strain-specific *bee*-like locus, for instance, may be an important factor in the complex regulatory system that contributes to the biofilm forming ability of VREr5.

5.3 Antibiotic resistance

The occurrence of VRE_{fm} has received much attention in public health, mainly because this organism is resistant to multiple available drugs commonly used in the hospitals. Despite its own intrinsic resistance to antibiotics such as penicillin and kanamycin, VRE_{fm} also possesses a remarkable ability to acquire resistance genes through either chromosomal mutations or genetic exchange of mobile elements.

Due to the clinical importance of VRE_{fm}, the vancomycin resistance mechanism has been widely studied. At present, nine types of vancomycin resistance (VanA, VanB, VanC, VanD, VanE, VanG, VanL, VanM, and VanN) have been described in enterococci (Boyd et al., 2008; Courvalin, 2006; Lebreton et al., 2011; Xu et al., 2010). Among these, the VanA and VanB types are the most commonly reported. Although widely distributed, the VanA type resistance is mainly found in strains originated from the United States, Europe, and Korea (Bonten et al., 2001; Yoo et al., 2006). On the other hand, VanB type resistance is more commonly found in Australia and Singapore (Lam et al., 2013; Molton et al., 2013). In Malaysia, the occurrence of clinical strains with VanA type resistance has been reported (Ibrahim et al., 2010; Mohamed et al., 2015). Although rare, VanB VRE has been isolated from a healthy animal-affiliated worker in Malaysia (Getachew et al., 2012). The identification of four *vanA*-carrying VRE_{fm} in this study suggested that VanA type VRE might be more prevalent in the clinical strains in Malaysia.

There have been reports documenting the discrepancies observed between the genotype and phenotype of *vanA*-genotype VRE. The differences observed are mainly due to the impairment of teicoplanin resistance which leads to the emergence of VanB phenotype-*vanA* genotype VRE (Il et al., 2008; Kuo et al., 2014; Song et al., 2006). Hashimoto et al. reported three point mutations in the sensor domain of *vanS* which contribute to impaired teicoplanin resistance among the *vanA*-genotype strains

(Hashimoto et al., 2000). Another study demonstrated that deletion of *vanY* and *vanZ* genes or insertion of IS*Efa4* at the *orf2-vanR* intergenic region may also responsible for the differences between genotype and phenotype (Gu et al., 2009). However, these mutations or gene disruptions were not observed in the Tn*1546*-like transposons of VREr6 and VREr7. Instead, the Tn*1546*-like transposons of these two strains were similar to those found in VRE2 and VREr5, which displaced VanA phenotype. Hence, the actual cause of the impaired teicoplanin resistance warrants further investigation. Apart from that, the similar Tn*1546*-like transposons carried by the four studied strains indicated a possible resistance pool in the hospital. Since no VRE strain was isolated from patient X and patient Y at the point of admission, these patients most probably acquired VRE during their prolonged hospitalization, which carried the same Tn*1546*-like transposons.

The anaerobic respiration of *E. faecium* has been thought to limit the uptake of aminoglycosides which contributes to its intrinsically low-level resistance to these drugs (Bryan & Van Den Elzen, 1977). Hence, a combination therapy involving aminoglycosides and drugs that disrupt cell wall synthesis, such as ampicillin or vancomycin, is effective in enhancing the therapeutic effect (Moellering & Weinberg, 1971). However, this synergistic effect becomes ineffective when the strain is highly resistant to aminoglycosides, especially gentamicin and streptomycin. The acquisition of the bi-functional gene *aac(6′)-Ie-aph(2′′)-Ia* confers high-level resistance to gentamicin by modifying the 2′ hydroxyl position of the antibiotic (Ferretti et al., 1986). More importantly, the presence of this gene also renders *E. faecium* resistant to virtually all clinically available aminoglycosides, except streptomycin (Chow, 2000). Therefore, streptomycin is a drug of choice in the combination therapy against stains possessing the *aac(6′)-Ie-aph(2′′)-Ia* gene, providing that these strains are not highly resistant to streptomycin (MIC $\geq$ 1000 μg/ml). High-level resistance to streptomycin is mainly determined by the presence of *ant(6′)-Ia.* Since three (VREr5, VREr6, VREr7) of the four

Malaysian strains harboured both the *aac(6′)-Ie-aph(2′′)-Ia* and *ant(6′)-Ia* genes, it was speculated that the synergistic therapy mentioned earlier cannot effectively eradicate these strains. Instead, the use of linezolid might help to clear VRE infecting patient Y, given that all the three strains were phenotypically resistant to linezolid and did not possess the mutations that confer resistance to this drug.

Except for the observed VanB phenotype-*vanA* genotype in VREr6 and VREr7, the antibiotic resistance determinants identified corresponded well with the antibiotic resistance phenotypes of the four local VRE_{fm}. The determination of resistance genes or mutations enables further understanding of the resistance mechanisms employed by the specifically studied strains. This could be useful in deciding the appropriate therapeutic dosage of the antibiotics, as well as the development of new treatment methods.

5.4 Transcriptomic analysis

Biofilm is a surface-attached community encased in a matrix of exopolymeric substances. Biofilm formation is a complicated developmental process which involves the attachment and immobilization on a surface, formation of microcolony, and development of three-dimensional structure (O`Toole et al., 2000). It has been demonstrated that the gene expression profiles of bacteria grown in biofilm and planktonic mode are different (Beenken et al., 2004; Resch et al., 2005). Determination of genes that are required for biofilm formation at different stages can, therefore, be helpful in controlling the biofilm development, which is important for therapeutic purposes.

The high-throughput sequencing technology generated a total of 1,234 transcripts, representing a comprehensive transcriptome for VREr5, the only biofilm former among the studied strains. Functional annotation showed that the transcripts obtained were involved in almost all of the biological processes. However, a large portion of them was with unknown functions, indicating the limitation in fully understand the transcriptome of this organism. The assembly of raw reads was performed *de novo*, due to the lack of an appropriate reference genome. Reads were later aligned to the generated assembly to assess the integrity and reliability of the assembly. Approximately 99.2% of the raw reads from each sample mapped to the assembled transcripts, indicating that the *de novo* assembly was performed accurately.

The transition from a free-swimming lifestyle to a surface-attached community-based growth requires the sensing of environmental cues that prompt cells to change their form of life. This shift in lifestyle involves extensive changes in structure and function which governed by a complex regulatory network. Several genes related to transcriptional regulators were up-regulated in the early stage of biofilm formed by VREr5. Most of the transcriptional regulators, such as those from the AraC, MerR, Gnt, and MarR family, are associated with carbon metabolism, stress response, and virulence (Gallegos et al., 1997; Perera & Grove, 2010). The higher gene expression levels of these proteins might due to the need to respond to environmental signals that trigger the biofilm initialization process, followed by the synthesis of sugars and energy to prepare for the next step of biofilm development. Although the exact roles of most of the up-regulated transcriptional regulators cannot be inferred in the current study, the results indicated the involvement of specific regulatory systems in the control of specific stages of biofilm formation.

Attachment to a surface is the first step in biofilm development. Therefore, it is not surprising to find out that the expression of genes related to adherence was up-regulated in the biofilm of VREr5. Most of these genes encoded for LPXTG-domain-containing cell wall anchor proteins, which are well-known in their role for attachment, catalyzed by the cleavage of sortase (Konto-Ghiorghi et al., 2009; Paterson & Mitchell, 2004). One up-regulated gene cluster which encoded such proteins was the pili-encoding *ebpABC*$_{fm}$. Prior to the discovery of *ebpABC*$_{fm}$, similar gene cluster (*ebpABC)* had been found in *E. faecalis* and its association with biofilm formation was demonstrated (Nallapareddy et al., 2006). Sillanpää et al. (2008) later identified the three-gene locus *ebpABC*$_{fm}$ in *E. faecium* which shows 48-74% similarity with that found in *E. faecalis* (Hendrickx et al., 2008). Deletion of *ebpABC*$_{fm}$ eliminates cell surface expression of the encoded pili and significantly reduces both biofilm formation and initial adherence, indicating a positive association of this gene cluster to biofilm formation (Sillanpää et al., 2010). The induction of *ebpABC*$_{fm}$ observed in VREr5 further supported the involvement of this locus in the biofilm development of *E. faecium*.

The tetracycline resistance gene *tetS* had also been over-expressed in the biofilm cells. Apart from providing resistance against tetracycline, the induction of *tetS* might also associated with protein synthesis needed during biofilm formation. It has been demonstrated that protein synthesis is important for the attachment step of biofilm growth (Crouzet et al., 2014; O'Toole & Kolter, 1998). Since *tetS* confers tetracycline resistance through ribosomal protection (Charpentier et al., 1993), this action might indirectly ensure normal synthesis of proteins required for the initial step of biofilm formation. Ironically, the expression levels of genes encoding ribosomal proteins were lower in the biofilm cells as oppose to the planktonic cells. This could be attributed to a lower growth rate in cells grown in biofilm as compared to the free-floating mode (Rice et al., 2000). The reduced growth rate of biofilm cells can be supported by the decreased expression of

genes encode for DNA replication and metabolism proteins in the transcriptomic results. Similarly, the lower expression of ribosomal genes observed could be due to a slower rate of protein synthesis in the biofilm cells.

A higher level of horizontal gene transfer is a well-established phenomenon in the biofilm (Madsen et al., 2012). The up-regulation of several genes related to transposase and recombinase in VREr5 further confirm the established observation. Furthermore, the increased expression of genes encoding competence proteins in the biofilm formed by VREr5 supported the existence of genetic exchange event in the analysed stage of biofilm formation. Besides, the result obtained in this study also indicated that the increased rate of horizontal gene transfer can occur as early as the initial step of biofilm formation. However, this exchange of genetic materials is not one directional. Horizontal gene transfer, such as conjugative transfer of plasmids, can also influence biofilm formation. A study conducted by Ghigo (2001) showed that the capacity of biofilm formation was greatly enhanced in *E. coli* introduced with a conjugative F plasmid. Other studies also demonstrated enhanced biofilm formation in *E. coli* by other conjugative plasmids (Burmølle et al., 2008; Ong et al., 2009). All of the above studies indicated the role of conjugative pili in the biofilm development of Gram negative bacteria by either promoting cell aggregation or enhancing attachment to a surface. While conjugation of Gram positive bacteria is pili-independent, the close interaction between the donors and recipients can have similar effect as the conjugative pili in promoting cell aggregation, which in time may lead to biofilm formation. Alternatively, plasmids that carry biofilm-associated-pili might also be transferred to the recipient cells, leading to activation or enhancement of biofilm formation. Although gene associated with the conjugal transfer was up-regulated, the transfer of pilus-containing plasmid cannot be confirmed in VREr5. However, given the up-regulation of several genes associated with cell wall-anchored proteins, the conjugative transfer of pilus-carrying plasmids cannot be ruled out.

Apart from the interplay observed between conjugative plasmids and biofilm formation, another effect of biofilm on plasmid biology is the plasmid copy number. In the biofilm formed by VREr5, a number of genes associated with plasmid replication were found to be over-expressed, indicating active replication of plasmid. As early as in 1995, the interconnection between biofilm and the plasmid copy number had been observed (Davies & Geesey, 1995). A later study performed by May et al. (2009) showed that the copy number of pBR322, a plasmid carrying ampicillin and tetracycline resistance genes, increased in biofilm cells of *E. coli* as compared to in planktonic cells. The author also demonstrated that the increased copy number of plasmid correlates with increased antibiotic resistance (May et al., 2009). Similarly, in *E. faecalis*, the copy number of different types of plasmid was also found to be increased in cells growing in a biofilm (Cook et al., 2011; Cook & Dunny, 2013). It is postulated that the increased copy number of plasmid in biofilm cells could be related to the release of the extracellular DNA. It was reported that extracellular DNA has been substantially secreted during the early biofilm growth of *E. faecalis* by a cell-lysis-independent mechanism (Barnes et al., 2012). Cook and Dunny, therefore, propose that if chromosome, instead of plasmid, is secreted to the extracellular matrix, the fraction of cellular RNA polymerase available for plasmid genes will increase relative to that of the chromosome, thereby increasing the transcription of plasmid genes (Cook & Dunny, 2014). Although the hypothesis made by Cook and Dunny (2014) has not yet been confirmed, our observation that genes associated with plasmid replication was up-regulated during the early stages of biofilm formation can be informative for further development of the hypothesis. Furthermore, since the gene expression of the pasmid-encoded *tetS* is also increased in the biofilm cells, the up-regulation of plasmid replication genes might responsible for enhanced antibiotic resistance by increasing the number of plasmid carrying the *tetS* gene.

Several genes involved in the pentose and glucuronate interconversion pathway also displaced increased expression under biofilm condition. These included genes encode for mannonate dehydratase and glucuronate isomerase. The mannonate dehydratase catalyzes the conversion of D-mannonate to 2-keto-3-deoxy-D-gluconate (Rakus et al., 2007). The resultant product serves as a substrate for the pentose phosphate pathway, which ultimately generates energy and five-carbon sugars (Wahba et al., 1960). The enzyme glucuronate isomerase, on the other hand, catalyzes the conversion of D-glucuronate to D-fructuronate (Ashwell et al., 1960). The D-fructuronate can eventually be converted to D-mannonate and enters the pentose phosphate pathway as mentioned earlier (Wahba et al., 1960). The biofilm cells might then use the products from the pentose phosphate pathway to synthesize purines and pyrimidines, the two building blocks of nucleic acids. Given the high rate of plasmid replication in the initial stage of biofilm formation, the induction of both mannonate dehydratase and glucuronate isomerase seems to play an important role in generating substrates for the DNA replication. This observation also indicated that the metabolic pathway is regulated in the biofilm cells in respond to specific requirement during different stages of growth.

Quorum sensing is a cell-density dependent signaling system used by many bacteria species in biofilm formation (Annous et al., 2009). Quorum sensing is mediated by chemical molecules called autoinducers produced by bacteria. Binding of autoinducers to the specific transcriptional regulator(s) leads to either activation or repression of target genes. Both Gram-positive and Gram-negative bacteria possess a common quorum sensing system mediated by autoinducer 2 (AI-2). The production of AI-2 depends on a gene called *luxS* (Surette et al., 1999). This gene was found to be under-expressed in biofilm formed by VREr5. Several studies show that the *luxS* quorum sensing system has different effects on bacterial biofilm formation. For example, in *Streptococcus gordonii*, a *luxS* mutant failed to form mixed-species biofilm as compared to the wild-type strain

(McNab et al., 2003). Other studies show that deletion of *luxS* leads to altered biofilm structure (He et al., 2016; Merritt et al., 2003). Furthermore, the positive association of this quorum sensing system to biofilm formation had been demonstrated in both *E. faecalis* and *Streptococcus pneumoniae* (Shao et al., 2012; Vidal et al., 2013). Nevertheless, there are studies indicating that *luxS* suppresses biofilm formation (Sela et al., 2006; Xu et al., 2006). In VREr5, the down regulation of *luxS* suggested that this quorum sensing system might play a negative regulatory role in biofilm formation, similar to those reported by Sela et al. and Xu et al.

Another quorum sensing related gene found in enterococci is *fsr*. In *E. faecalis*, the *fsr* locus consists of three genes: *fsrA* (response regulator), *fsrB* (signalling peptide), and *fsrC* (histidine kinase). This locus regulates the expression of gelatinase and serine proteases, which contributes to biofilm formation by facilitating bacterial attachment to a surface (Hancock & Perego, 2004). Interestingly, the expression level of *fsrB* gene in the biofilm of VREr5 decreased 133.22 fold. It was found out that the *fsr* system does not always positively regulates biofilm formation in enterococci. In a study performed by Mohamed and Murray (2006), the introduction of *fsr* locus into a *gelE*-negative strong biofilm producer reduces biofilm production at 41% compared to the wild type. The same result has been observed with a medium biofilm former (Mohamed & Murray, 2006). These observations indicate that the *fsr* locus plays a different role in biofilm formation independent of the activation of gelatinase production. The author, therefore, suggested that in the absence of gelatinase, *fsr* might negatively control initial attachment, leading to reduced biofilm formation (Mohamed & Murray, 2006). Since VREr5 was gelatinase-negative, as shown by results obtained in Section 4.3.1, the decreased expression of *fsrB* in biofilms suggested that this locus might also negatively regulate biofilm growth in VREr5. In this case, *fsrB* could be the major player in the negative regulation of biofilm formation as *fsrA* and *fsrC* were not differentially expressed.

Biofilm cells are known to have a higher tolerance to environmental stresses than its free-living counterpart. A global transcriptional regulator, Spx (*spx*), is known to participate in several stress responses. Spx was initially identified in *Bacillus subtilis* as a suppressor of two proteases, ClpP and ClpX (Nakano et al., 2001). In *B. subtilis*, Spx supresses competence development while activates genes involved in oxidative stress responses (Nakano et al., 2003a, 2003b). In *E. faecalis*, this protein was shown to be involved in a variety of stress responses, where deletion of *spx* impaired growth of the bacterium under oxidative and antibiotic stresses, as well as higher killing rate by macrophage compared to the wild type (Kajfasz et al., 2012). However, this protein does not seem to play a role in general stress response in the biofilm of *E. faecium* strain VREr5 as the gene encoding it was under-expressed. Nevertheless, *spx* has been shown to negatively affect biofilm formation in both *S. aureus* and *S. epidermidis*. In both cases, *spx* acts to supress the *ica* operon encoding polysaccharide intracellular adhesin, leading to reduced primary attachment and, hence, biofilm formation (Pamp et al., 2006; Wang et al., 2010). In the same sense, *spx* in VREr5 might either supress a biofilm-related gene or more during the planktonic growth. When a decision to move into biofilm mode is made, the transcription of *spx* decreased, relieving the suppression. Wang et al. (2010) demonstrated that the protease ClpP involves in biofilm formation of *S. epidermidis* by degrading Spx. As such, one would expect the expression level of *clpP* gene to be increased in the biofilm cells. However, *clpP* was down-regulated in VREr5. This implied that *clpP* may not be involved in the *spx*-mediated biofilm formation of VREr5. Alternatively, the down-regulation of *clpP* may be due to the presence of an autoregulatory system which is activated when the level of substrate (Spx) is below the threshold.

A gene cluster encodes for putative proteins of the arginine deiminase system (*arcABC*) was among the most highly down-regulated genes in the biofilm of VREr5. This result was different from other studies which show that the *arcABC* is expressed at a higher level in biofilm cells relative to that in the planktonic cells (Beenken et al., 2004; Resch et al., 2005). Despite the differences observed, this gene cluster has been shown to be more active during the maturation of biofilm (Lindgren et al., 2014; Resch et al., 2005). When the biofilm matures, some regions become anoxic due to the enclosed matrix. In these regions, cells respire through fermentation, producing acids. The accumulation of acids leads to acid stress which can be alleviated by the *arcABC* operon. The *arc*ABC operon converts arginine to ornithine, ammonia, and carbon dioxide (Cunin et al., 1986). The resultant ammonia is then protonated into ammonium ion (NH_4^+) which increases the intracellular pH, maintaining pH homeostasis (Lindgren et al., 2014). During the initial stage of biofilm formation, however, cells are still being supplied with sufficient nutrients and oxygen for energy production through aerobic respiration. Hence, the arginine deiminase system is probably not needed during the early stage of biofilm formation, which might explain the under expression of the *arcABC* in VREr5.

The *bee* locus of *E. faecalis* has been shown to enhance biofilm formation. This five-gene locus is located on a conjugative plasmid and encodes for proteins carrying conserved domains (B domain, von Willebrand factor type A domain, LPxTG domain) that are involved in ligand binding and cell wall anchoring (Tendolkar et al., 2006). Biofilm formation was greatly reduced (70%) when *bee-2*, one of the five genes, was inactivated. Furthermore, filter mating experiment showed that the transconjugants displaced enhanced biofilm-forming ability compared to the parent strain, confirming the role of this locus in biofilm formation of *E. faecalis* (Tendolkar et al., 2006). Interestingly, although the previous genomic study suggested a possible association of a homolog of *E. faecalis bee* locus to biofilm formation in VREr5, this unique gene locus was down-

regulated. This suggested that the *bee* homolog might have a negative regulatory role in the biofilm development of VREr5. Further investigations are needed to study the association of this gene cluster to biofilm formation in VREr5 as well as in other *E. faecium* strains.

The transcriptome results also showed that some genes putatively involved in enterococcal biofilm formation, such as *esp*, *acm*, and *sgrA*, were not differentially expressed. The role of *esp* in enhancing biofilm formation in both *E. faecalis* and *E. faecium* has been reported (Heikens et al., 2007; Tendolkar et al., 2004). However, other studies suggested that *esp* is not required or sufficient for biofilm formation in *E. faecalis* and *E. faecium* (Dworniczek et al., 2005; Ramadhan & Hegedus, 2005). For *acm* and *sgrA*, both of these genes have been implicated in the biofilm formation of *E. faecium* (Hendrickx et al., 2009; Nallapareddy et al., 2008). The results obtained from the transcriptomic study of VREr5 suggested that the involvement of these genes in enterococcal biofilm development is either strain-specific, as observed for *esp*, or that these genes were only needed in the later stage of biofilm formation in VREr5. The latter postulation was supported by the under-expression of alt_{Efm} and *sagA*, which encode for an autolysin and secreted protein, respectively, that have been shown to participate in the maturation of biofilm (Paganelli et al., 2015, 2013).

CHAPTER 6: CONCLUSION

In this study, four clinical VRE_{fm} strains isolated from a tertiary hospital in Malaysia were subjected to whole genome sequencing and comparative genomics analyses. These four strains were previously shown to have different PFGE patterns, particularly three strains from a single patient (patient Y). The comparative genomics analysis indicated genetic variability among the four local strains, contributed mainly by mobile genetic elements and their respective strain-specific genes. Nevertheless, high genetic similarity was observed between VREr6 and VREr7, both isolated from patient Y. On the other hand, VREr5, the initial isolate from patient Y, was found to be genetically more similar to VRE2 isolated from an index case (patient X). Subsequent MLST and phylogenomics analyses further revealed that VREr5 was different from the other two strains isolated from the same host. Both the clinical and genomic data suggested that patient Y was most probably infected with multiple strains of VRE_{fm} from different clones. Additionally, the first isolated strain from patient Y was probably of the same clone as the strain from the index case, patient X. By having the whole genome sequences, the genetic relationship of the four local VRE_{fm} strains, previously determined through PFGE, was able to be elucidated with higher resolution. Furthermore, the genome sequences also enable the identification of a number of antibiotic resistance- and virulence genes in the VRE_{fm} strains, which could contribute to their persistence and pathogenicity. A unique *bee*-like locus was identified only in VREr5, which may associate with biofilm formation. This study, to the best of knowledge, is the first report on the comparative genomics analyses of local VRE_{fm} as well as the first to report ST80 VRE_{fm} in Malaysia.

On the other hand, the transcriptomic analysis performed in this study had provided insight into the gene expression profile of the initial stage of biofilm formation in *E. faecium*. High-quality transcripts were generated which enable reliable interpretation of

the genes that participate in different growth conditions. Genes that were induced in the biofilm cells involved mainly in surface adherence, plasmid replication, conjugation, and sugar metabolism. The up-regulation of these genes correlated with the requirements during the initial step of biofilm formation, in which attachment to a surface, cell-cell interaction, and energy production are important. Moreover, the over-expression of several transcriptional regulators indicated that cells possess specific regulatory system in controlling the different stages of biofilm development. Additionally, the transcriptomic analysis also revealed the down-regulation of genes that participate or possibly needed in the latter stage of biofilm formation, suggesting that the results obtained in this study clearly reflected the initial stage of biofilm development. The unique *bee*-like locus was found to be down-regulated, indicated a negative association of this gene to the biofilm initiation of VREr5. The transcriptomic data generated from this study can be useful in understanding the biofilm forming mechanism of *E. faecium*, which in turn may be helpful in developing new drugs against biofilm-related enterococcal diseases.

REFERENCES

Aarestrup, F.M., Seyfarth, A.M., Emborg, H.D., Pedersen, K., Hendriksen, R.S., & Bager, F. (2001). Effect of abolishment of the use of antimicrobial agents for growth promotion on occurence of antimicrobial resistance in fecal Entercocci from food animals in Denmark. *Antimicrobial Agents and Chemotherapy, 45*, 2054–2059.

Abdi, H., & Williams, L.J. (2010). Principal component analysis. *Wiley Interdisciplinary Review:Computational Statistics*, *2*, 433-459.

Abriouel, H., Omar, N. Ben, Molinos, A.C., López, R.L., Grande, M.J., Martínez-Viedma, P., … Galvez, A. (2008). Comparative analysis of genetic diversity and incidence of virulence factors and antibiotic resistance among enterococcal populations from raw fruit and vegetable foods, water and soil, and clinical samples. *International Journal of Food Microbiology, 123*, 38–49.

Alikhan, N.-F., Petty, N.K., Ben Zakour, N.L., & Beatson, S.A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. *BMC Genomics, 12*, 402.

Almohamad, S., Somarajan, S.R., Singh, K. V., Nallapareddy, S.R., & Murray, B.E. (2014). Influence of isolate origin and presence of various genes on biofilm formation by *Enterococcus faecium*. *FEMS Microbiology Letters, 353*, 151–156.

Annous, B.A., Fratamico, P.M., & Smith, J.L. (2009). Quorum sensing in biofilms: Why bacteria behave the way they do. *Journal of Food Science, 74*, R24–R37.

Arias, C.A., Panesso, D., Singh, K. V., Rice, L.B., Murray, B.E., & Bosque, U. El. (2009). Cotransfer of antibiotic resistance genes and a hyl_{Efm}-containing virulence plasmid in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 53*, 4240–4246.

Arthur, M., & Courvalinn, P.(1993). Genetics and mechanisms of glycopeptide resistance in Enterococci. *Antimicrobial Agents and Chemotherapy, 37*, 1563–1571.

Arthur, M., Depardieu, F., Molinas, C., Reynolds, P., & Courvalin, P. (1995). The *vanZ* gene of Tn*1546* from *Enterococcus faecium* BM4147 confers resistance to teicoplanin. *Gene, 154*, 87–92.

Ashwell, G., Wahba, A.J., & Hickman, J. (1960). Uronic acid metabolism in bacteria. I. Purification and properties of uronic acid isomerases in *Escherichia coli*. *Journal of Biological Chemistry*, *235*, 1559–1565.

Aubert, D., Naas, T., & Nordmann, P. (2003). IS*1999* increases expression of the extended-spectrum beta-lactamase VEB-1 in *Pseudomonas aeruginosa*. *Journal of Bacteriology, 185*, 5314–5319.

Aziz, R.K., Bartels, D., Best, A.A., DeJongh, M., Disz, T., Edwards, R.A., … Zagnitko, O. (2008). The RAST Server: Rapid annotations using subsystems technology. *BMC Genomics, 9*, 75.

Bager, F., Madsen, M., Christensen, J., & Aarestrup, F.M. (1997). Avoparcin used as a growth promoter is associated with the occurrence of vancomycin-resistant *Enterococcus faecium* on Danish poultry and pig farms. *Preventive Veterinary Medicine, 31*, 95–112.

Baldassarri, L., Cecchini, R., Bertuccini, L., Ammendolia, M.G., Iosi, F., Arciola, C.R., … Creti, R. (2001). *Enterococcus* spp. produces slime and survives in rat peritoneal macrophages. *Medical Microbiology and Immunology, 190*, 113–120.

Barnes, A.M.T., Ballering, K.S., Leibman, R.S., Wells, C.L., & Dunny, G.M. (2012). *Enterococcus faecalis* produces abundant extracellular structures containing DNA in the absence of cell lysis during early biofilm formation. *mBio, 3*, e00193-12.

Beenken, K.E., Dunman, P.M., McAleese, F., Macapagal, D., Murphy, E., Projan, S.J., … Blevins, J.S. (2004). Global gene expression in *Staphylococcus aureus* biofilms. *Journal of Bacteriology, 186*, 4665–4684.

Bertels, F., Silander, O.K., Pachkov, M., Rainey, P.B., & van Nimwegen, E. (2014). Automated reconstruction of whole genome phylogenies from short sequence reads. *Molecular Biology and Evolution, 31*, 1077–1088.

Boggard, V. den, & Stobberingh, E.E., 2000. The effect of banning avoparcin on VRE carriage in The Netherlands. *The Journal of Antimicrobial Chemotherapy*, *46*(1),146–148.

Bonten, M.J.M., Willems, R., & Weinstein, R.A. (2001). Vancomycin-resistant enterococci : Why are they here , and where do they come from ? *The Lancet Infectious Diseases*, *1*(5), 314–325.

Boyd, D.A., Willey, B.M., Fawcett, D., Gillani, N., & Mulvey, M.R. (2008). Molecular characterization of *Enterococcus faecalis* N06-0364 with low-level vancomycin resistance harboring a novel D-Ala-D-Ser gene cluster, vanL. *Antimicrobial Agents and Chemotherapy, 52*, 2667–2672.

Bryan, L.E., & Van Den Elzen, H.M. (1977). Effects of membrane-energy mutations and cations on streptomycin and gentamicin accumulation by bacteria: A model for entry of streptomycin and gentamicin in susceptible and resistant bacteria. *Antimicrobial Agents and Chemotherapy, 12*, 163–177.

Burmølle, M., Bahl, M.I., Jensen, L.B., Sørensen, S.J., & Hansen, L.H. (2008). Type 3 fimbriae, encoded by the conjugative plasmid pOLA52, enhance biofilm formation and transfer frequencies in Enterobacteriaceae strains. *Microbiology, 154*, 187–195.

Cha, J.O., Jung, Y.H., Lee, H.R., Yoo, J. Il, & Lee, Y.S. (2012). Comparison of genetic epidemiology of vancomycin-resistant *Enterococcus faecium* isolates from humans and poultry. *Journal of Medical Microbiology, 61*, 1121–1128.

Charpentier, E., Gerbaud, G., & Courvalin, P. (1993). Characterization of a new class of tetracycline-resistance gene *tet(S)* in *Listeria monocytogenes* BM4210. *Gene, 131*, 27–34.

Chelvam, K.K., Chai, L.C., & Thong, K.L. (2014). Variations in motility and biofilm formation of *Salmonella enterica* serovar Typhi. *Gut Pathogens, 6*, 2.

Chen, L., Yang, J., Yu, J., Yao, Z., Sun, L., Shen, Y., & Jin, Q. (2005). VFDB: A reference database for bacterial virulence factors. *Nucleic Acids Research, 33*, 325–328.

Chow, J.W. (2000). Aminoglycoside resistance in Enterococci. *Clinical Infectious Diseases, 31*, 586–589.

CLSI. (2016). *Performance standards for antimicrobial susceptibility testing. 26th Informational Supplement, Document M100-S26*. Wayne, PA 19087 USA: Clinical and Laboratory Standard Institute.

Coburn, P.S., Baghdayan, A.S., Craig, N., Burroughs, A., Tendolkar, P., Miller, K., … Shankar, N. (2010). A novel conjugative plasmid from *Enterococcus faecalis* E99 enhances resistance to ultraviolet radiation. *Plasmid, 64*, 18–25.

Conesa, A., Gotz, S., Garcia-Gomez, J.M., Terol, J., Talon, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics, 21*, 3674–3676.

Cook, L., Barnes, A., Dunny, G., & Yarwood, J. (2011). Biofilm growth alters regulation of conjugation by a bacterial pheromone. *Molecular Microbiology, 81*, 1499–1510.

Cook, L.C., & Dunny, G.M. (2014). The influence of biofilms in the biology of plasmids. *Microbiology Spectrum, 2*, 0012.

Cook, L.C., & Dunny, G.M. (2013). Effects of biofilm growth on plasmid copy number and expression of antibiotic resistance genes in *Enterococcus faecalis*. *Antimicrobial Agents and Chemotherapy, 57*, 1850–1856.

Courvalin, P. (2006). Vancomycin resistance in Gram-positive cocci. *Clinical Infectious Diseases, 42*, S25–34.

Crouzet, M., Le Senechal, C., Brözel, V.S., Costaglioli, P., Barthe, C., Bonneu, M., … Vilain, S. (2014). Exploring early steps in biofilm formation: Set-up of an experimental system for molecular studies. *BMC Microbiology, 14*, 1–12.

Cunin, R., Glansdorff, N., Pierard, A., & Stalon, V. (1986). Biosynthesis and metabolism of arginine in bacteria. *Microbiological Reviews, 50*, 314–352.

Dahl, K.H., Simonsen, G.S., Olsvik, Ø., & Sundsfjord, A. (1999). Heterogeneity in the *vanB* gene cluster of genomically diverse clinical strains of vancomycin-resistant enterococci. *Antimicrobial Agents and Chemotherapy, 43*, 1105–1110.

Darling, A.C.E., Mau, B., Blattner, F.R., & Perna, N.T. (2004). Mauve: Multiple alignment of conserved genomic sequence with rearrangements. *Genome Research, 14*, 1394–1403.

Davies, D.G., & Geesey, G.G. (1995). Regulation of the alginate biosynthesis gene *algC* in *Pseudomonas aeruginosa* during biofilm development in continuous culture. *Applied and Environmental Microbiology*, *61*, 860–867.

Derbise, A., Aubert, S., & El Solh, N. (1997). Mapping the regions carrying the three contiguous antibiotic resistance genes *aadE*, *sat4*, and *aphA-3* in the genomes of staphylococci. *Antimicrobial Agents and Chemotherapy, 41*, 1024–1032.

Devriese, L.A., Ieven, M., Goossens, H., Vandamme, P., Pot, B., Hommez, J., & Haesebrouck, F. (1996). Presence of vancomycin-resistant enterococci in farm and pet animals. *Antimicrobial Agents and Chemotherapy, 40*, 2285–2287.

Donlan, R.M., Costerton, J.W., Donlan, R.M., & Costerton, J.W. (2002). Biofilms : Survival mechanisms of clinically relevant microorganisms. *Journal of Clinical Microbiology, 15*, 167–193.

Dupre`, I., Zanetti, S., Schito, A.M., Fadda, G., & Sechi, L.A. (2003). Incidence of virulence determinants in clinical *Enterococcus faecium* and *Enterococcus faecalis* isolates collected in Sardinia (Italy). *Journal of Medical Microbiology, 52*, 491–498.

Dworniczek, E., Wojciech, Ł., Sobieszczańska, B., & Seniuk, A. (2005). Virulence of enterococcus isolates collected in Lower Silesia (Poland). *Scandinavian Journal of Infectious Diseases, 37*, 630–636.

Edelsberg, J., Weycker, D., Barron, R., Li, X., Wu, H., Oster, G., … Weber, D.J. (2014). Prevalence of antibiotic resistance in US hospitals. *Diagnostic Microbiology and Infectious Diseases, 78*, 255–262.

Evers, S., & Courvalin, P. (1996). Regulation of VanB-type vancomycin resistance gene expression by the VanS(B)-VanR (B) two-component regulatory system in *Enterococcus faecalis* V583. *Journal of Bacteriology, 178*, 1302–1309.

Facklam, R.R., & Collins, M.D. (1989). Identification of enterococcus species isolated from human infections by a conventional test scheme. *Journal of Clinical Microbiology, 27*, 731–734.

Ferretti, J.J., Gilmore, K.S., & Courvalin, P. (1986). Nucleotide sequence analysis of the gene specifying the bifunctional 6'-aminoglycoside acetyltransferase 2"-aminoglycoside phosphotransferase enzyme in *Streptococcus faecalis* and identification and cloning of gene regions specifying the two activities. *Journal of Bacteriology, 167*, 631–638.

Freitas, A.R., Tedim, A.P., Novais, C., Ruiz-Garbajosa, P., Werner, G., Laverde-Gomez, J.A., … Coque, T.M. (2010). Global spread of the hyl_{Efm} colonization-virulence gene in megaplasmids of the *Enterococcus faecium* CC17 polyclonal subcluster. *Antimicrobial Agents and Chemotherapy, 54*, 2660–2665.

Fry, J., Wood, M., & Poole, P.S. (2001). Investigation of myo-inositol catabolism in *Rhizobium leguminosarum bv*. viciae and its effect on nodulation competitiveness. *Molecular Plant-Microbe Interactions, 14*, 1016–1025.

Gagnon, S., Lévesque, S., Lefebvre, B., Bourgault, A.M., Labbé, A.C., & Roger, M. (2011). *vanA*-containing *Enterococcus faecium* susceptible to vancomycin and teicoplanin because of major nucleotide deletions in Tn*1546*. *The Journal of Antimicrobial Chemotherapy, 66*, 2758–2762.

Gallegos, M.T., Schleif, R., Bairoch, A., Hofmann, K., Ramos, J.L., Gallegos, A., … Bairoch, A. (1997). AraC/XylS family of transcriptional regulators. *Microbiology and Molecular Biology Reviews, 61*, 393–410.

Galloway-Peña, J.R., Rice, L.B., & Murray, B.E. (2011). Analysis of PBP5 of early U.S. isolates of *Enterococcus faecium*: Sequence variation alone does not explain increasing ampicillin resistance over time. *Antimicrobial Agents and Chemotherapy, 55*, 3272–3277.

Getachew, Y., Hassan, L., Zakaria, Z. & Saleha, A.A. (2013). Genetic variability of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* isolates from humans , chickens , and pigs in Malaysia. *Applied and Environmental Microbiology, 79*, 4528–4533.

Getachew, Y., Hassan, L., Zakaria, Z., Zaid, C.Z.M., Yardi, A., Shukor, R.A., & Marawin, L.T. (2012). Characterization and risk factors of vancomycin-resistant Enterococci (VRE) among animal-affiliated workers in Malaysia. *Journal of Applied Microbiology, 113*, 1184–1195.

Ghigo, J.M. (2001). Natural conjugative plasmids induce bacterial biofilm development. *Nature, 412*, 442–445.

Grissa, I., Vergnaud, G., & Pourcel, C. (2008). CRISPRFinder: A website to compare clustered regularly interspaced short palindromic repeats. *Nucleic Acids Research, 36*, 52–57.

Gu, L., Cao, B., Liu, Y., Guo, P., Song, S., Li, R., … Wang, C. (2009). A new Tn*1546* type of VanB phenotype-*vanA* genotype vancomycin-resistant *Enterococcus faecium* isolates in mainland China. *Diagnostic Microbiology and Infectious Diseases, 63*, 70–75.

Guilhen, C., Charbonnel, N., Parisot, N., Gueguen, N., Iltis, A., Forestier, C., & Balestrino, D. (2016). Transcriptional profiling of *Klebsiella pneumoniae* defines signatures for planktonic, sessile and biofilm-dispersed cells. *BMC Genomics, 17*, 237.

Guzman Prieto, A.M., van Schaik, W., Rogers, M.R.C., Coque, T.M., Baquero, F., Corander, J., & Willems, R.J.L. (2016). Global emergence and dissemination of Enterococci as nosocomial pathogens: Attack of the clones? *Frontiers in Microbiology, 7*, 1–15.

Hancock, L.E., & Perego, M. (2004). The *Enterococcus faecalis fsr* two-component system controls biofilm development through production of gelatinase. *Journal of Bacteriology, 186*, 5629–5639.

Hashimoto, Y., Tanimoto, K., Ozawa, Y., Murata, T., & Ike, Y. (2000). Amino acid substitutions in the VanS sensor of the VanA-type vancomycin-resistant enterococcus strains result in high-level vancomycin resistance and low-level teicoplanin resistance. *FEMS Microbiology Letters, 185*, 247–254.

He, Z., Liang, J., Zhou, W., Xie, Q., Tang, Z., Ma, R., & Huang, Z. (2016). Effect of the quorum-sensing *luxS* gene on biofilm formation by *Enterococcus faecalis*. *European Journal of Oral Sciences, 124*, 234–240.

Héchard, Y., & Sahl, H.G. (2002). Mode of action of modified and unmodified bacteriocins from Gram-positive bacteria. *Biochimie, 84*, 545–557.

Heikens, E., Bonten, M.J.M., & Willems, R.J.L. (2007). Enterococcal surface protein Esp is important for biofilm formation of *Enterococcus faecium* E1162. *Journal of Bacteriology, 189*, 8233–8240.

Heikens, E., Singh, K. V., Jacques-Pakaz, K.D., Luit-asbroek, M. Van, Oostdijk, E.A., Bonten, M.J.M., … Willems, R.J.L. (2012). Contribution of the enterococcal surface protein Esp to pathogenesis of *Enterococcus faecium* endocarditis. *Microbes and Infection, 13*, 1185–1190.

Hendrickx, A.P.A., Bonten, M.J.M., van Luit-Asbroek, M., Schapendonk, C.M.E., Kragten, A.H.M., & Willems, R.J.L. (2008). Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate. *Microbiology, 154*, 3212–3223.

Hendrickx, A.P.A., Luit-asbroek, M. Van, Schapendonk, C.M.E., Wamel, W.J.B. Van, Braat, J.C., Wijnands, L.M., … Willems, R.J.L. (2009). SgrA , a nidogen-binding LPXTG surface adhesin implicated in biofilm formation , and EcbA , a collagen binding MSCRAMM , are two novel adhesins of hospital-acquired *Enterococcus faecium*. *Infection and Immunity, 77*, 5097–5106.

Hendrickx, A.P.A., Schapendonk, C.M.E., Van Luit-Asbroek, M., Bonten, M.J.M., Van Schaik, W., & Willems, R.J.L. (2010). Differential PilA pilus assembly by a hospital-acquired and a community-derived *Enterococcus faecium* isolate. *Microbiology, 156*, 2649–2659.

Hilsenbeck, S.G., Friedrichs, W.E., Schiff, R., O'Connell, P., Hansen, R.K., Osborne, C.K., & Fuqua, S.A. (1999). Statistical analysis of array expression data as applied to the problem of tamoxifen resistance. *Journal of the National Cancer Institute, 91*, 453–459.

Hinton, J.C.D., Hautefort, I., Eriksson, S., Thompson, A., & Rhen, M. (2004). Benefits and pitfalls of using microarrays to monitor bacterial gene expression during infection. *Current Opinion in Microbiology, 7*, 277–282.

Hollenbeck, B.L., & Rice, L.B. (2012). Intrinsic and acquired resistance mechanisms in enterococcus. *Virulence, 3*, 421–433.

Homan, W.L., Tribe, D., Poznanski, S., Li, M., Hogg, G., Spalburg, E., … Willems, R.J.L. (2002). Multilocus sequence typing scheme for *Enterococcus faecium*. *Journal of Clinical Microbiology, 40*, 1963–1971.

Hong, J.H., Jang, H.W., Kang, Y.E., Lee, J.H., Kim, K.S., Kim, H.J., … Ku, B.J. (2012). Urinary chiro- and myo-inositol levels as a biological marker for type 2 diabetes mellitus. *Disease Markers, 33*, 193–199.

Huerta-Cepas, J., Szklarczyk, D., Forslund, K., Cook, H., Heller, D., Walter, M.C., … Bork, P. (2016). EGGNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences. *Nucleic Acids Research, 44*, D286–D293.

Huh, J.Y., Lee, W.G., Lee, K., & Shin, W.S. (2004). Distribution of insertion sequences associated with Tn*1546*-like elements among *Enterococcus faecium* isolates from patients in Korea. *Journal of Clinical Microbiology, 42*, 1897–1902.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., & Hauser, L.J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics, 11*, 119.

Ibrahim, R., Mohamad, M., & Rahman, M. (2010). Enterococci : Emerging drug resistant bacteria in hospital acquired infections at Hospital Kuala Lumpur , Malaysia. *International Journal of Microbiology, 9*, 1–7.

Ibrahim, R.B., Mohamad, M., & Rahman, M.M. (2011). Vancomycin resistant enterococci and detection of responsible genes. *Pakistan Journal of Medical Sciences, 27*, 784–788.

Il, J.P., Wee, G.L., Jong, H.S., Kyung, W.L., & Gun, J.W. (2008). VanB phenotype-*vanA* genotype *Enterococcus faecium* with heterogeneous expression of teicoplanin resistance. *Journal of Clinical Microbiology, 46*, 3091–3093.

Jansen, R., Van Embden, J.D.A., Gaastra, W., & Schouls, L.M. (2002). Identification of genes that are associated with DNA repeats in prokaryotes. *Molecular Microbiology, 43*, 1565–1575.

Jones, M.E., Draghi, D.C., Thornsberry, C., Karlowsky, J.A., Sahm, D.F., & Wenzel, R.P. (2004). Emerging resistance among bacterial pathogens in the intensive care unit--a European and North American Surveillance study (2000-2002). *Annals of Clinical Microbiology and Antimicrobials, 3*, 14.

Jurcisek, J. a, Dickson, A.C., Bruggeman, M.E., & Bakaletz, L.O. (2011). *In vitro* biofilm formation in an 8-well chamber slide. *Journal of Visualized Experiments*, (47), 2481.

Kajfasz, J.K., Mendoza, J.E., Gaca, A.O., Miller, J.H., Koselny, K.A., Giambiagi-deMarval, M., … Lemos, J.A. (2012). The Spx regulator modulates stress responses and virulence in *Enterococcus faecalis*. *Infection and Immunity, 80*, 2265–2275.

Kariyama, R., Mitsuhata, R., Chow, J.W., Clewell, B., & Kumon, H. (2000). Simple and reliable multiplex PCR assay for surveillance isolates of vancomycin-resistant Enterococci. *Journal of Clinical Microbiology, 38*, 3092–3095.

Kim, D.S., Singh, K. V., Nallapareddy, S.R., Qin, X., Panesso, D., Arias, C.A., & Murray, B.E. (2010). The *fms21* (*pilA*)-*fms20* locus encoding one of four distinct pili of *Enterococcus faecium* is harboured on a large transferable plasmid associated with gut colonization and virulence. *Journal of Medical Microbiology, 59*, 505–507.

Kim, E.B., & Marco, M.L. (2014). Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. *Applied and Environmental Microbiology, 80*, 154–165.

Konto-Ghiorghi, Y., Mairey, E., Mallet, A., Duménil, G., Caliot, E., Trieu-Cuot, P., & Dramsi, S. (2009). Dual role for pilus in adherence to epithelial cells and biofilm formation in *Streptococcus agalactiae*. *PLOS Pathogens, 5*.

Kuo, A.J., Su, L.H., Shu, J.C., Wang, J.T., Wang, J.H., Fung, C.P., … Wu, T.L. (2014). National Surveillance on Vancomycin-resistant *Enterococcus faecium* in Taiwan: Emergence and widespread of ST414 and a Tn*1546*-like element with simultaneous insertion of IS*1251*-like and IS*1678*. *PLOS ONE, 9*, 1–13.

Lam, M.M.C., Seemann, T., Bulach, D.M., Gladman, S.L., Chen, H., Haring, V., … Stineara, T.P. (2012). Comparative analysis of the first complete *Enterococcus faecium* genome. *Journal of Bacteriology, 194*, 2334–2341.

Lam, M.M.C., Seemann, T., Tobias, N.J., Chen, H., Haring, V., Moore, R.J., … Stinear, T.P. (2013). Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. *BMC Genomics, 14*, 595.

Lebreton, F., Depardieu, F., Bourdon, N., Fines-Guyon, M., Berger, P., Camiade, S., … Cattoir, V. (2011). D-Ala-D-Ser VanN-type transferable vancomycin resistance in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 55*, 4606–4612.

Leclercq, R., Derlot, E., Duval, J., & Courvalin, P. (1988). Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *The New England Jounal of Medicine, 319*, 157–161.

Lee, H., Doak, T.G., Popodi, E., Foster, P.L., & Tang, H. (2016). Insertion sequence-caused large-scale rearrangements in the genome of *Escherichia coli*. *Nucleic Acids Research, 44*, 7109–7119.

Leendertse, M., Heikens, E., Wijnands, L.M., Luit-asbroek, M. Van, Teske, G.J.D., Roelofs, J.J.T.H., … Willems, R.J.L. (2009). Enterococcal surface protein transiently aggravates *Enterococcus faecium* – induced urinary tract infection in mice. *Journal of Infectious Diseases, 200*, 1–4.

Lindgren, J.K., Thomas, V.C., Olson, M.E., Chaudhari, S.S., Nuxoll, A.S., Schaeffer, C.R., … Fey, P.D.(2014). Arginine deiminase in *Staphylococcus epidermidis* functions to augment biofilm maturation through pH homeostasis. *Journal of Bacteriology*, *196*, 2277–2289.

Liu, B., & Pop, M. (2009). ARDB - Antibiotic resistance genes database. *Nucleic Acids Research, 37*, 443–447.

Madera, C., Garcia, P., Rodriguez, A., Suarez, J.E., & Martinez, B. (2009). Prophage induction in *Lactococcus lactis* by the bacteriocin Lactococcin 972. *International Journal of Food Microbiology, 129*, 99–102.

Madsen, J.S., Burmølle, M., Hansen, L.H., & Sørensen, S.J. (2012). The interconnection between biofilm formation and horizontal gene transfer. *FEMS Immunology and Medical Microbiology, 65*, 183–195.

Maisey, H.C., Quach, D., Hensler, M.E., Liu, G.Y., Gallo, R.L., Nizet, V., & Doran, K.S. (2008) A Group B Streptococcal pilus protein promotes phagocyte resistance and systemic virulence. *FASEB Journal, 22*, 1715–1724.

Marraffini, L.A., & Sontheimer, E.J. (2011). CRISPR interference: RNA-directed adaptive immunity in bacteria and archaea. *Nature Reviews Genetics, 11*, 181–190.

Marshall, C.G., Broadhead, G., Leskiw, B.K., & Wright, G.D. (1997). d-Ala-d-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. *Proceedings of the National Academy of Sciences of the United States of America, 94*, 6480–6483.

Martin, P., Trieu-cuot, P., & Courvalin, P. (1986). Nucleotide sequence of the *tetM* tetracycline resistance determinant of the streptococcal conjugative shuttle transposon Tn*1546*. *Nucleic Acids Research, 14*, 7047–7058.

Martinez, B., Bottiger, T., Schneider, T., Rodriguez, A., Sahl, H.G., & Wiedemann, I. (2008). Specific interaction of the unmodified bacteriocin lactococcin 972 with the cell wall precursor lipid II. *Applied and Environmental Microbiology, 74*, 4666–4670.

Martinez, B., Rodriguez, A., & Suarez, J.E. (2000). Lactococcin 972, a bacteriocin that inhibits septum formation in lactococci. *Microbiology, 146*, 949–955.

Martínez, B., Suárez, J.E., & Rodríguez, A. (1996). Lactococcin 972: A homodimeric lactococcal bacteriocin whose primary target is not the plasma membrane. *Microbiology, 142*, 2393–2398.

Matsushima, A., Takakura, S., Yamamoto, M., Matsumura, Y., Shirano, M., Nagao, M., … Ichiyama, S. (2012). Regional spread and control of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* in Kyoto, Japan. *European Journal of Clinical Microbiology and Infectious Diseases, 31*, 1095–1100.

May, T., Ito, A., & Okabe, S. (2009). Induction of multidrug resistance mechanism in *Escherichia coli* biofilms by interplay between tetracycline and ampicillin resistance genes. *Antimicrobial Agents and Chemotherapy, 53*, 4628–4639.

Mazaheri Nezhad Fard, R., Barton, M.D., & Heuzenroeder, M.W. (2011). Bacteriophage-mediated transduction of antibiotic resistance in enterococci. *Letters in Applied Microbiology, 52*, 559–564.

McNab, R., Ford, S.K., El-Sabaeny, A., Barbieri, B., Cook, G.S., & Lamont, R.J. (2003). LuxS-based signaling in *Streptococcus gordonii*: Autoinducer 2 controls carbohydrate metabolism and biofilm formation with *Porphyromonas gingivalis*. *Journal of Bacteriology, 185*, 274–284.

Merritt, J., Qi, F., Goodman, S.D., Anderson, M.H., & Shi, W. (2003). Mutation of *luxS* affects biofilm formation in *Streptococcus mutans*. *Infection and Immunity, 71*, 1972–1979.

Mikalsen, T., Pedersen, T., Willems, R., Coque, T.M., Werner, G., Sadowy, E., … Hegstad, K. (2015). Investigating the mobilome in clinically important lineages of *Enterococcus faecium* and *Enterococcus faecalis*. *BMC Genomics, 16*, 282.

Moellering, R.C., & Weinberg, A.N. (1971). Studies on antibiotic synergism against enterococci. *Journal of Clinical Investigation, 50*, 2580–2584.

Moellering, R.J. (1992). Emergence of Enterococcus as a significant pathogen. *Clinical Infectious Diseases, 14*, 1173–1176.

Mohamed, J.A., Huang, W., Nallapareddy, S.R., Teng, F., & Murray, B.E. (2004). Influence of origin of isolates , especially dndocarditis isolates , and various genes on biofilm formation by *Enterococcus faecalis*. *Infection and Immunity, 72*, 3658–3663.

Mohamed, J.A., & Murray, B.E. (2006). Influence of the *fsr* locus on biofilm formation by *Enterococcus faecalis* lacking *gelE*. *Journal of Medical Microbiology, 55*, 1747–1750.

Mohamed, N.A., Hussin, H., Chang, K.M., Hashim, R., & Ahmad, N. (2015). Vancomycin resistant enterococcus (VRE): Prevalence and characteristics in a tertiary hospital in malaysia. *Brunei International Medical Journal, 11*, 241–246.

Molton, J.S., Tambyah, P.A., Ang, B.S.P., Ling, M.L., & Fisher, D.A. (2013). The global spread of healthcare-associated multidrug-resistant bacteria: A perspective from Asia. *Clinical Infectious Diseases, 56*, 1310–1318.

Nakano, M.M., Hajarizadeh, F., Zhu, Y., & Zuber, P. (2001). Loss-of-function mutations in *yjbD* result in ClpX- and ClpP-independent competence development of *Bacillus subtilis*. *Molecular Microbiology, 42*, 383–394.

Nakano, S., Küster-Schöck, E., Grossman, A.D., & Zuber, P. (2003a). Spx-dependent global transcriptional control is induced by thiol-specific oxidative stress in *Bacillus subtilis*. *Proceedings of the National Academy of Sciences of the United States of America*, *100*, 13603–13608.

Nakano, S., Nakano, M.M., Zhang, Y., Leelakriangsak, M., & Zuber, P. (2003b). A regulatory protein that interferes with activator-stimulated transcription in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 100, 4233–4238.

Nallapareddy, S.R., Singh, K. V., Okhuysen, P.C., & Murray, B.E. (2008a). A functional collagen adhesin gene, *acm*, in clinical isolates of *Enterococcus faecium* correlates with the recent success of this emerging nosocomial pathogen. *Infection and Immunity, 76*, 4110–4119.

Nallapareddy, S.R., Singh, K. V, & Murray, B.E. (2008). Contribution of the collagen adhesin Acm to pathogenesis of *Enterococcus faecium* in experimental endocarditis. *Infection and Immunity, 76*, 4120–4128.

Nallapareddy, S.R., Singh, K. V., Sillanpaa, J., Zhao, M., & Murray, B.E. (2011). Relative contributions of Ebp pili and the collagen adhesin Ace to host extracellular matrix protein adherence and experimental urinary tract infection by *Enterococcus faecalis* OG1RF. *Infection and Immunity, 79*, 2901–2910.

Nallapareddy, S.R., Singh, K. V, Sillanpää, J., Garsin, D.A., Höök, M., Erlandsen, S.L., & Murray, B.E. (2006). Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *Journal of Clinical Investigation, 116*, 2799–2807.

Nallapareddy, S.R., Weinstock, G.M., & Murray, B.E. (2003). Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. *Molecular Microbiology, 47*, 1733–1747.

Ooka, T., Ogura, Y., Asadulghani, M., Ohnishi, M., Nakayama, K., Terajima, J., … Hayashi, T. (2009). Inference of the impact of insertion sequence (IS) elements on bacterial genome diversification through analysis of small-size structural polymorphisms in *Escherichia coli* O157 genomes. *Genome Research, 19*, 1809–1816.

O`Toole, G., Kaplan, H.B., & Kolter, R. (2000). Biofilm formation as microbial development. *Annual Review of Microbiology, 54*, 49–79.

O'Toole, G.A., & Kolter, R. (1998). Initiation of biofilm formation in *Pseudomans fluorescens* WC365 proceeds via multiple, convergent signalling pathways: A genetic analysis. *Molecular Microbiology, 28*, 449–461.

Ong, C.L.Y., Beatson, S.A., McEwan, A.G., & Schembri, M.A. (2009). Conjugative plasmid transfer and adhesion dynamics in an *Escherichia coli* biofilm. *Applied and Environmental Microbiology, 75*, 6783–6791.

Paganelli, F.L., de Been, M., Braat, J.C., Hoogenboezem, T., Vink, C., Bayjanov, J., … Leavis, H.L. (2015). Distinct SagA from hospital-associated clade A1 *Enterococcus faecium* strains contributes to biofilm formation. *Applied and Environmental Microbiology, 81*, 6873–6882.

Paganelli, F.L., Willems, R.J.L., Jansen, P., Hendrickx, A., Zhang, X., Bonten, M.J.M., & Leavis, H.L. (2013). *Enterococcus faecium* biofilm formation : Identification of major autolysin $AtlA_{Efm}$, associated Acm surface localization, and $AtlA_{Efm}$-independent extracellular DNA release. *mBio, 4*, 1–10.

Palmer, K.L., & Gilmore, M.S. (2010). Multidrug-resistant enterococci lack CRISPR-cas. *mBio, 1*, 1–10.

Pamp, S.J., Frees, D., Engelmann, S., Hecker, M., & Ingmer, H. (2006). Spx is a global effector impacting stress tolerance and biofilm formation in *Staphylococcus aureus*. *Journal of Bacteriology, 188*, 4861–4870.

Panesso, D., Montealegre, M.C., Rincón, S., Mojica, M.F., Rice, L.B., Singh, K. V., … Arias, C. (2011). The *hyl*$_{Efm}$ gene in pHyl$_{Efm}$ of *Enterococcus faecium* is not required in pathogenesis of murine peritonitis. *BMC Microbiology, 11*, 20.

Paterson, G.K., & Mitchell, T.J. (2004). The biology of Gram-positive sortase enzymes. *Trends Microbiology, 12*, 89–95.

Perera, I.C., & Grove, A. (2010). Molecular mechanisms of ligand-mediated attenuation of DNA binding by MarR family transcriptional regulators. *Journal of Molecular Cell Biology*, *2*, 243–254.

Price, M.N., Dehal, P.S., & Arkin, A.P. (2010). FastTree 2 - Approximately maximum-likelihood trees for large alignments. *PLOS ONE, 5*.

Putnam, S.D., Sader, H.S., Moet, G.J., Mendes, R.E., & Jones, R.N. (2010). Worldwide summary of telavancin spectrum and potency against Gram-positive pathogens: 2007 to 2008 surveillance results. *Diagnostic Microbiology and Infectious Diseases, 67*, 359–368.

Qin, X., Galloway-Peña, J.R., Sillanpaa, J., Roh, J.H., Nallapareddy, S.R., Chowdhury, S., … Murray, B.E. (2012). Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. *BMC Microbiology, 12*, 135.

Rakus, J.F., Fedorov, A.A., Fedorov, E. V, Glasner, M.E., Vick, J.E., Babbitt, P.C., … Gerlt, J.A. (2007). Evolution of enzymatic activities in the enolase superfamily: d-mannonate dehydratase from *Novosphingobium aromaticivorans*. *Biochemistry, 46*, 12896–12908.

Ramadhan, A., & Hegedus, E. (2005). Biofilm formation and *esp* gene carriage in enterococci. *Journal of Clinical Pathology, 58*, 685–686.

Resch, A., Rosenstein, R., Nerz, C., Go, F., & Götz, F. (2005). Differential gene expression profiling of *Staphylococcus aureus* cultivated under biofilm and planktonic conditions. *Applied and Environmental Microbiology, 71*, 2663–76.

Rice, A.R., Hamilton, M.A., & Camper, A.K. (2000). Apparent surface associated lag time in growth of primary biofilm cells. *Microbial Ecology, 41*, 8–15.

Rice, L.B., Bellais, S., Carias, L.L., Hutton-Thomas, R., Bonomo, R.A., Caspers, P., … Gutmann, L. (2004). Impact of specific *pbp5* mutations on expression of beta-lactam resistance in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 48*, 3028–3032.

Rice, L.B., Carias, L., Rudin, S., Vael, C., Goossens, H., Konstabel, C., … Murray, B.E. (2003). A potential virulence gene , *hyl*$_{Efm}$, predominates in *Enterococcus faecium* of clinical origin. *Journal of Infectious Diseases, 187*, 502–512.

Rice, L.B., Carias, L.L., Rudin, S., Laktičová, V., Wood, A., & Hutton-Thomas, R. (2005). *Enterococcus faecium* low-affinity *pbp5* is a transferable determinant. *Antimicrobial Agents and Chemotherapy, 49*, 5007–5012.

Rice, L.B., Lakticová, V., Carias, L.L., Rudin, S., Hutton, R., & Marshall, S.H. (2009). Transferable capacity for gastrointestinal colonization in *Enterococcus faecium* in a mouse model. *Journal of Infectious Diseases, 199*, 342–9.

Ryan, L., O'Mahony, E., Wrenn, C., FitzGerald, S., Fox, U., Boyle, B., … Klare, I. (2015). Epidemiology and molecular typing of VRE bloodstream isolates in an Irish tertiary care hospital. *Journal of Antimicrobial Chemotherapy, 70*, 2718–2724.

Schleifer, K.H., & Kilpper-Balz, R. (1984). Transfer of *Streptococcus faecalis* and *Streptococcus faecium* to the genus Enterococcus nom. rev. as *Enterococcus faecalis* comb. nov. and *Enterococcus faecium* comb. nov. *International Journal of Systematic Bacteriology, 34*, 31–34.

Schouten, M.A., Willems, R.J.L., Kraak, W.A.G., Top, J., Hoogkamp-Korstanje, J.A.A., & Voss, A. (2001). Molecular analysis of Tn*1546*-like elements in vancomycin-resistant enterococci isolated from patients in Europe shows geographic transposon type clustering. *Antimicrobial Agents and Chemotherapy, 45*, 986–989.

Sela, S., Frank, S., Belausov, E., & Pinto, R. (2006). A mutation in the *luxS* gene influences *Listeria monocytogenes* biofilm formation. *Applied and Environmental Microbiology, 72*, 5653–5658.

Sghir, A., Gramet, G., Suau, A., Rochet, V., Pochart, P., & Dore, J. (2000). Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. *Applied and Environmental Microbiology, 66*, 2263–2266.

Shao, C., Shang, W., Yang, Z., Sun, Z., Li, Y., Guo, J., … Yuan, J. (2012). LuxS-dependent AI-2 regulates versatile functions in *Enterococcus faecalis* V583. *Journal of Proteome Research, 11*, 4465–4475.

Sheldon, W.L., Macauley, M.S., Taylor, E.J., Robinson, C.E., Charnock, S.J., Davies, G.J., … Black, G.W. (2006). Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-N-acetylglucosaminidase and not a hyaluronidase. *Biochemical Journal, 399*, 241–247.

Sievert, D.M., Ricks, P., Kallen, A.J., Edwards, J.R., Schneider, A., Patel, J., … Fridkin, S. (2016). Antimicrobial-resistant pathogens associated with healthcare-associated infections: Summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2011–2014. *Infection Control and Hospital Epidemiology*, *37*, 1288–1301.

Siguier, P., Gourbeyre, E., Varani, A., Ton-hoang, B.A.O., & Chandler, M. (2015). Everyman's guide to bacterial insertion sequences. *Microbiology Spectrum, 3*, 1–35.

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., & Chandler, M. (2006). ISfinder: The reference centre for bacterial insertion sequences. *Nucleic Acids Research, 34*, D32–36.

Sillanpää, J., Nallapareddy, S.R., Prakash, V.P., Qin, X., Hook, M., Weinstock, G.M., & Murray, B.E. (2008). Identification and phenotypic characterization of a second collagen adhesin, Scm, and genome-based identification and analysis of 13 other predicted MSCRAMMs, including four distinct pilus loci, in *Enterococcus faecium*. *Microbiology, 154*, 3199–3211.

Sillanpää, J., Nallapareddy, S.R., Singh, K. V, Prakash, V.P., Fothergill, T., Ton-That, H., & Murray, B.E. (2010). Characterization of the *ebp(fm)* pilus-encoding operon of *Enterococcus faecium* and its role in biofilm formation and virulence in a murine model of urinary tract infection. *Virulence, 1*, 236–246.

Sillanpää, J., Prakash, V.P., Nallapareddy, S.R., & Murray, B.E. (2009). Distribution of genes encoding MSCRAMMs and pili in clinical and natural populations of *Enterococcus faecium*. *Journal of Clinical Microbiology, 47*, 896–901.

Simjee, S., White, D.G., McDermott, P.F., Wagner, D.D., Zervos, M.J., Donabedian, S.M., … Walker, R.D. (2002). Characterization of Tn*1546* in vancomycin-resistant *Enterococcus faecium* isolated from canine urinary tract infections : Evidence of gene exchange between human and animal Enterococci. *Journal of Clinical Microbiology, 40*, 4659–4665.

Singh, K. V., Malathum, K., & Murray, B.E. (2001). Disruption of an *Enterococcus faecium* species-specific gene, a homologue of acquired macrolide resistance genes of staphylococci, is associated with an increase in macrolide susceptibility. *Antimicrobial Agents and Chemotherapy, 45*, 263–266.

Sivertsen, A., Pedersen, T., Larssen, K.W., Bergh, K., Rønning, T.G., Radtke, A., & Hegstad, K. (2016). A silenced *vanA* gene cluster on a transferable plasmid caused an outbreak of vancomycin-variable enterococci. *Antimicrobial Agents and Chemotherapy*, *60*(7), 4119-4127.

Sletvold, H., Johnsen, P.J., Wikmark, O.G., Simonsen, G.S., Sundsfjord, A., & Nielsen, K.M. (2010). Tn*1546* is part of a larger plasmid-encoded genetic unit horizontally disseminated among clonal *Enterococcus faecium* lineages. *Journal of Antimicrobial and Chemotherapy, 65*, 1894–1906.

Soheili, S., Ghafourian, S., Sekawi, Z., Neela, V., Sadeghifard, N., Ramli, R., & Hamat, R.A. (2014). Wide distribution of virulence genes among *Enterococcus faecium* and *Enterococcus faecalis* clinical isolates. *The Scientific World Journal, 2014*, 1–6.

Sóki, J., Eitel, Z., Urbán, E., & Nagy, E. (2013). Molecular analysis of the carbapenem and metronidazole resistance mechanisms of Bacteroides strains reported in a Europe-wide antibiotic resistance survey. *International Journal of Antimicrobial Agents*, 41, 122–125.

Song, J.H., Ko, K.S., Oh, W.S., Park, S., Heo, S.T., Kwon, K.T., … Lee, N.Y. (2006). High frequency of vancomycin-resistant *Enterococcus faecium* isolates with VanB phenotype and *vanA* genotype in Korean hospitals. *Diagnostic Microbiology and Infectious Diseases, 56*, 401–406.

Sorek, R., Kunin, V., & Hugenholtz, P. (2008). CRISPR—a widespread system that provides acquired resistance against phages in bacteria and archaea. *Nature Reviews Microbiology, 6*, 181–186.

Stepanović, S., Vuković, D., Dakić, I., Savić, B., & Švabić-Vlahović, M. (2000). A modified microtiter-plate test for quantification of staphylococcal biofilm formation. *Journal of Microbiological Methods, 40*, 175–179.

Stobberingh, E., Van Den Bogaard, A., London, N., Driessen, C., Top, J., & Willems, R. (1999). Enterococci with glycopeptide resistance in turkeys, turkey farmers, turkey slaughterers, and (sub)urban residents in the south of The Netherlands: Evidence for transmission of vancomycin resistance from animals to humans? *Antimicrobial Agents and Chemotherapy, 43*, 2215–2221.

Surette, M.G., Miller, M.B., & Bassler, B.L. (1999). Quorum sensing in *Escherichia coli*, *Salmonella typhimurium*, and *Vibrio harveyi*: A new family of genes responsible for autoinducer production. *Proceedings of the National Academy of Sciences of the United States of America, 96*, 1639–44.

Tendolkar, P.M., Baghdayan, A.S., & Michael, S. (2004). Enterococcal surface protein , Esp , enhances biofilm formation by *Enterococcus faecalis*. *Infection and Immunity, 72*, 6032–6039.

Tendolkar, P.M., Baghdayan, A.S., & Shankar, N. (2006). Putative surface proteins encoded within a novel transferable locus confer a high-biofilm phenotype to *Enterococcus faecalis*. *Journal of Bacteriology, 188*, 2063–2072.

Teng, F., Singh, K. V, Bourgogne, A., Zeng, J., & Murray, B.E. (2009). Further characterization of the *epa* gene cluster and Epa polysaccharides of *Enterococcus faecalis*. *Infection and Immunity, 77*, 3759–3767.

Toledo-arana, A., Valle, J., Solano, C., Cucarella, C., Lamata, M., Amorena, B., & Penade, R. (2001). The Enterococcal surface protein , Esp , is involved in *Enterococcus faecalis* biofilm formation. *Applied and Environmental Microbiology, 67*, 4538–4545.

Top, J., Paganelli, F.L., Zhang, X., Schaik, W. Van, Leavis, H.L., Van, M., … Willems, R.J.L. (2013). The *Enterococcus faecium* enterococcal biofilm regulator , EbrB , regulates the *esp* operon and is implicated in biofilm formation and intestinal colonization. *PLOS ONE, 8*, e65224.

Top, J., Willems, R., & Bonten, M. (2008). Emergence of CC17 *Enterococcus faecium*: From commensal to hospital-adapted pathogen. *FEMS Immunology and Medical Microbiology*, 52, 297–308.

Tsai, H.Y., Liao, C.H., Chen, Y.H., Lu, P.L., Huang, C.H., Lu, C. Te., … Hsueh, P.R. (2012). Trends in susceptibility of vancomycin-resistant *Enterococcus faecium* to tigecycline, daptomycin, and linezolid and molecular epidemiology of the isolates: Results from the Tigecycline in Vitro Surveillance in Taiwan (TIST) study, 2006 to 2010. *Antimicrobial Agents and Chemotherapy, 56*, 3402–3405.

Turabelidze, D., Kotetishvili, M., Jr, J.G.M., Sulakvelidze, A., Kreger, A., & Morris, J.G. (2000). Improved pulsed-field gel electrophoresis for typing vancomycin-resistant Enterococci. *Journal of Clinical Microbiology, 38*, 4242–4245.

Turner, K.M., Hanage, W.P., Fraser, C., Connor, T.R., & Spratt, B.G. (2007). Assessing the reliability of eBURST using simulated populations with known ancestry. *BMC Microbiology, 7*, 30.

Urwin, R., & Maiden, M.C.J. (2003). Multi-locus sequence typing: A tool for global epidemiology. *Trends in Microbiology, 11*, 479–487.

Uttley, A., Collins, C., Naidoo, J., & George, R. (1988). Vancomycin-resistant enterococci. *Lancet, 331*, 57–58.

Valdezate, S., Labayru, C., Navarro, A., Mantecón, M.A., Ortega, M., Coque, T.M., … Saéz-Nieto, J.A. (2009). Large clonal outbreak of multidrug-resistant CC17 ST17 *Enterococcus faecium* containing Tn*5382* in a Spanish hospital. *Journal of Antimicrobial and Chemotherapy*, 63, 17–20.

Van Braak, N. Den, Van Belkum, A., Van Keulen, M., Vliegenthart, J., Verbrugh, H.A., & Endtz, H.P. (1998). Molecular characterization of vancomycin-resistant enterococci from hospitalized patients and poultry products in the Netherlands. *Journal of Clinical Microbiology, 36*, 1927–1932.

van Schaik, W., Top, J., Riley, D.R., Boekhorst, J., Vrijenhoek, J.E.P., Schapendonk, C.M.E., … Willems, R.J.L. (2010). Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics, 11*, 239.

van Wamel, W.J.B., Hendrickx, A.P.A., Bonten, M.J.M., Top, J., Posthuma, G., & Willems, R.J.L. (2007). Growth condition-dependent Esp expression by *Enterococcus faecium* affects initial adherence and biofilm formation. *Infection and Immunity, 75*, 924–931.

Vidal, J.E., Howery, K.E., Ludewick, H.P., Nava, P., & Klugman, K.P. (2013). Quorum-sensing systems LuxS/Autoinducer 2 and com regulate *Streptococcus pneumoniae* biofilms in a bioreactor with living cultures of human respiratory cells. *Infection and Immunity, 81*, 1341–1353.

Wahba, A.J., Hickman, J., & Ashwell, G. (1960). Uronic acid metabolism in bacteria. II. Purification and properties of D-altronic acid and D-mannonic acid dehydrogenases in *Escherichia coli*. *The Journal of Biological Chemistry, 235*, 1566–1570.

Wang, C., Fan, J., Niu, C., Wang, C., Villaruz, A.E., Otto, M., & Gao, Q. (2010). Role of *spx* in biofilm formation of *Staphylococcus epidermidis*. *FEMS Immunology and Medical Microbiology*, *59*, 152–160.

Wang, Z., Gerstein, M., & Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. *Nature Reviews Genetics, 10*, 57–63.

Wegener, H.C. (1998). Historical yearly usage of glycopeptides for animals and humans: The American-European paradox revisited. *Antimicrobial Agents and Chemotherapy, 42*, 3049.

Werner, G. (2013). Molecular typing of Enterococci/VRE. *Journal of Bacteriology and Parasitology*, *5*, 1-8.

Werner, G., Fleige, C., Geringer, U., van Schaik, W., Klare, I., & Witte, W. (2011). IS element IS*16* as a molecular screening tool to identify hospital-associated strains of *Enterococcus faecium*. *BMC Infectious Diseases, 11*, 80.

Willems, R.J.L., Homan, W., Top, J., Van Santen-Verheuvel, M., Tribe, D., Manzioros, X., … Bonten, M.J.M. (2001). Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin-resistant *Enterococcus faecium* spreading in hospitals. *Lancet, 357*, 853–855.

Willems, R.J.L., Top, J., van den Braak, N., van Belkum, A., Mevius, D.J., Hendriks, G., … Embden, J.D.A. Van. (1999). Molecular diversity and evolutionary relationships of Tn*1546*-like elements in Enterococci from humans and animals. *Society, 43*, 483–491.

Willems, R.J.L., Top, J., van Santen, M., Robinson, D.A., Coque, T.M., Baquero, F., … Bonten, M.J.M. (2005). Global spread of vancomycin resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerging Infectious Diseases, 11*, 821–828.

Willems, R.J.L., Top, J., van Schaik, W., Leavis, H., Bonten, M., Sirén, J., … Corander, J. (2012). Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. *mBio, 3*, 1–10.

Xu, H.T., Tian, R., Chen, D.K., Xiao, F., Nie, Z.Y., Hu, Y.J., … Li, J.M. (2011). Nosocomial spread of hospital-adapted CC17 vancomycinresistant *Enterococcus faecium* in a tertiary-care hospital of Beijing, China. *Chinese Medical Journal, 124*, 498–503.

Xu, L., Li, H., Vuong, C., Vadyvaloo, V., Wang, J., Yao, Y., … Gao, Q. (2006). Role of the *luxS* quorum-sensing system in biofilm formation and virulence of *Staphylococcus epidermidis*. *Infection and Immunity, 74*, 488–496.

Xu, X., Lin, D., Yan, G., Ye, X., Wu, S., Guo, Y., … Wang, M. (2010). *vanM*, a new glycopeptide resistance gene cluster found in *Enterococcus faecium*. *Antimicrobial Agents and Chemotherapy, 54*, 4643–4647.

Yap, K., Gan, H.M., Teh S.J.C., Chai, L.C., & Thong, K.L. (2014). Comparative genomics of closely related *Salmonella enterica* serovar Typhi strains reveals genome dynamics and the acquisition of novel pathogenic elements. *BMC Genomics, 15*, 1–20.

Yoo, S.J., Sung, H., Cho, Y.U., Kim, M.N., Pai, C.H., & Kim, Y.S. (2006). Role of horizontal transfer of the transposon Tn*1546* in the nosocomial spread of *vanA* vancomycin-resistant enterococci at a tertiary care hospital in Korea. *Infection Control and Hospital Epidemiology, 27*, 1081–1087.

Yoshida, K.I., Yamaguchi, M., Morinaga, T., Kinehara, M., Ikeuchi, M., Ashida, H., & Fujita, Y., 2008. myo-inositol catabolism in *Bacillus subtilis*. *The Journal of Biological Chemistry, 283*, 10415–10424.

Yu, J.J., Shi, J.F., Zhao, R.K., Han, Q.Z., Qian, X.F., Gu, G.H., … Xu, J. (2015). Molecular characterization and resistant spectrum of Enterococci isolated from a haematology unit in China. *Journal of Clinical Diagnostic Research, 9*, 7–10.

Zankari, E., Hasman, H., Cosentino, S., Vestergaard, M., Rasmussen, S., Lund, O., … Larsen, M.V. (2012). Identification of acquired antimicrobial resistance genes. *Journal of Antimicrobial Chemotherapy, 67*, 2640–2644.

Zeng, J., Teng, F., Weinstock, G.M., & Murray, B.E. (2004). Translocation of *Enterococcus faecalis* strains across a monolayer of polarized human enterocyte-like T84 cells. *Journal of Clinical Microbiology, 42*, 1149–54.

Zhou, Y., Liang, Y., Lynch, K.H., Dennis, J.J., & Wishart, D.S. (2011). PHAST: A fast phage search tool. *Nucleic Acids Research, 39*, 1–6.

Zubaidah, A.W., Ariza, A., & Azmi, S. (2006). Hospital-acquired vancomycin-resistant Enterococci : Now appearing in Kuala Lumpur Hospital. *Medical Journal of Malaysia*, *61*, 487–489.

LIST OF PUBLICATIONS AND PAPERS PRESENTED

Publications

1. Lim, S.Y., Teh, C.S.J., and Thong, K.L. (2017). Biofilm-related diseases and Omics: Global transcriptional profiling of *Enterococcus faecium* reveals different gene expression patterns in the biofilm and planktonic cells. Omics, 21(10), 592-602.

2. Lim, S.Y., Yap, K.P., Teh, C.S.J., Abdul Jabar, K., and Thong, K.L. (2017). Comparative genome analysis of multiple vancomycin-resistant *Enterococcus faecium* isolated from two fatal cases. Infection, Genetics and Evolution, 49, 55-65.

Infection, Genetics and Evolution 49 (2017) 55–65

Contents lists available at ScienceDirect

Infection, Genetics and Evolution

journal homepage: www.elsevier.com/locate/meegid

Research paper

Comparative genome analysis of multiple vancomycin-resistant *Enterococcus faecium* isolated from two fatal cases

Shu Yong Lim [a], Kien-Pong Yap [a], Cindy Shuan Ju Teh [b], Kartini Abdul Jabar [b], Kwai Lin Thong [a,*]

[a] Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia
[b] Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia

ARTICLE INFO

Article history:
Received 26 August 2016
Received in revised form 22 December 2016
Accepted 25 December 2016
Available online 28 December 2016

Keywords:
Enterococcus faecium
Vancomycin-resistant
Comparative genomics
Phylogenomic
Virulence
Multidrug resistant

ABSTRACT

Enterococcus faecium is both a commensal of the human intestinal tract and an opportunistic pathogen. The increasing incidence of enterococcal infections is mainly due to the ability of this organism to develop resistance to multiple antibiotics, including vancomycin. The aim of this study was to perform comparative genome analyses on four vancomycin-resistant *Enterococcus faecium* (VRE_{fm}) strains isolated from two fatal cases in a tertiary hospital in Malaysia. Two sequence types, ST80 and ST203, were identified which belong to the clinically important clonal complex (CC) 17. This is the first report on the emergence of ST80 strains in Malaysia. Three of the studied strains (VREr5, VREr6, VREr7) were each isolated from different body sites of a single patient (patient Y) and had different PFGE patterns. While VREr6 and VREr7 were phenotypically and genotypically similar, the initial isolate, VREr5, was found to be more similar to VRE2 isolated from another patient (patient X), in terms of the genome contents, sequence types and phylogenomic relationship. Both the clinical records and genome sequence data suggested that patient Y was infected by multiple strains from different clones and the strain that infected patient Y could have derived from the same clone from patient X. These multidrug resistant strains harbored a number of virulence genes such as the *epa* locus and pilus-associated genes which could enhance their persistence. Apart from that, a homolog of *E. faecalis bee* locus was identified in VREr5 which might be involved in biofilm formation. Overall, our comparative genomic analyses had provided insight into the genetic relatedness, as well as the virulence potential, of the four clinical strains.

1. Introduction

Enterococci are commensals in the gastrointestinal tracts of humans and animals but some members of this genus are also opportunistic nosocomial pathogens which can cause diseases associated with bloodstream and urinary tract (Willems and van Schaik, 2009). Treatment of enterococcal infections is challenging due to the intrinsic and acquired resistance of enterococci to multiple antibiotics, including the last-line drugs such as vancomycin and daptomycin (Hollenbeck and Rice, 2012).

The first reported cases of vancomycin-resistant enterococci (VRE) occurred in the 1980s in the United Kingdom and France (Leclercq et al., 1988; Uttley et al., 1988). Since then, VRE are increasingly reported worldwide, including United States, Europe, and Asia (Bonten et al., 2001; Kuo et al., 2014). Resistance to vancomycin is typically mediated by one of the nine *van* gene clusters (*vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM*, *vanN*). Among them, *vanA* and *vanB* are the predominant resistance genotypes observed (Hollenbeck and Rice, 2012). Horizontal transfer of these genes to other pathogens has been a big concern. In fact, the conjugative transfer of enterococcal *vanA* gene to *Staphylococcus aureus* strain has been reported (Zhu et al., 2013). This interspecies transfer of resistance can result in highly resistant pathogens which are difficult to treat with the currently available antibiotics.

A rapid increase in the *Enterococcus faecium* infections has been observed worldwide since late 1980s, coinciding with the acquired vancomycin resistance (Treitman et al., 2005). In the United States, 87% of *E. faecium* recovered from nosocomial infections are vancomycin-resistant (Edelsberg et al., 2014). Molecular and comparative genomic studies showed that hospital-associated (HA) *E. faecium* strains are different from community-associated (CA) strains where mobile genetic elements and antimicrobial resistance genes are enriched in the HA strains (Qin et al., 2012).

Most of the HA strains belong to the clonal complex 17 (CC17) based on multilocus sequence typing (MLST) (Top et al., 2008). Strains from the CC17 pose specific traits that enable them to persist in the clinical environment. These include resistance to ampicillin and quinolone, and a pathogenicity island which carries the *esp* gene encoding enterococcal surface protein (Esp) putatively involved in biofilm formation and endocarditis (Heikens et al., 2012; Top et al., 2008). Other genes that contribute to virulence in *E. faecium* include *acm* which encodes for collagen binding adhesin that contributes to endocarditis

* Corresponding author.
E-mail address: thongkl@um.edu.my (K.L. Thong).

http://dx.doi.org/10.1016/j.meegid.2016.12.029
1567-1348/

OMI-2017-0119-ver9-Lim_1P
Type: research-article

OMICS A Journal of Integrative Biology
Volume 21, Number 10, 2017

DOI: 10.1089/omi.2017.0119

Research Article

Biofilm-Related Diseases and Omics: Global Transcriptional Profiling of *Enterococcus faecium* Reveals Different Gene Expression Patterns in the Biofilm and Planktonic Cells

Shu Yong Lim,[1] Cindy Shuan Ju Teh,[2] and Kwai Lin Thong[1]

Abstract

Enterococcus faecium is an opportunistic pathogen with a remarkable ability to acquire resistance toward multiple antibiotics, including those of last-resort drugs such as vancomycin and daptomycin. The occurrence of vancomycin-resistant *E. faecium* is on the rise and there is a need to understand the virulence of this organism. One of the factors that contributes to the virulence is the ability to form biofilms. Since bacteria in biofilm state are more resistant to antibiotics and host immune response, understanding the molecular mechanism of biofilm development is important to control biofilm-related diseases. The aim of this study was to determine the global gene expression profiles of an *E. faecium* strain, VREr5, during the early event of sessile growth compared with its planktonic phase through RNA-sequencing approach. The results clearly illustrated distinct expression profiles of the planktonic and biofilm cells. A total of 177 genes were overexpressed in the biofilm cells. Most of them encode for proteins involved in adherence, such as the *ebpABC*$_{fm}$ locus. Genes associated with plasmid replication, gene exchange, and protein synthesis were also upregulated during the early event of biofilm development. Furthermore, the transcriptome analysis also identified genes such as *fsrB*, *luxS*, and *spx* that might suppress biofilm formation in VREr5. The putative biofilm-related *bee* locus was found to be downregulated. These new findings could provide caveats for future studies on the regulation and maintenance of biofilm and development of biomarkers for biofilm-related diseases.

Keywords: biofilm, differential gene expression, *Enterococcus faecium*, RNA-seq, transcriptomes

Introduction

Enterococcus faecium is an opportunistic pathogen, which can cause severe human diseases such as urinary tract infections, surgical-related wound infections, bacteremia, and endocarditis (Agudeo Higuita and Huycke, 2014). The clinical importance of *E. faecium* is mainly attributed to its remarkable ability to acquire resistance toward multiple antibiotics, including those of last-resort drugs such as vancomycin and daptomycin (Edelsberg et al., 2014). As the occurrence of vancomycin-resistant *E. faecium* (VRE$_{fm}$) has increasingly been reported (Bonten et al., 2001; Kuo et al., 2014), understanding the pathogenicity of this organism is of paramount importance for infection control. One of the factors that contributes to the virulence of VRE$_{fm}$ is the ability to form biofilms.

Biofilm is a surface-associated community of microorganisms encased in a matrix of exopolymeric substances. It has been shown that sessile/biofilm cells are generally more tolerant to antibiotics and host phagocytosis than their planktonic counterparts due to the multilayered structure of mature biofilm, slower metabolic rate, and the ease of gene exchange within the biofilm community (de la Fuente-Núñez et al., 2013; Donlan and Costerton, 2002). Hence, biofilm-associated infections are often difficult to eradicate. Many chronic diseases such as endocarditis, periodontitis, and otitis media are associated with bacterial biofilms (de la Fuente-Núñez et al., 2013). Moreover, a wide variety of medical devices such as catheters and prosthetic heart valves are prone to being colonized by biofilm-forming bacteria. Clinical *E. faecium* appeared to have higher occurrence of biofilm formation as compared with nonclinical isolates (Almohamad et al., 2014),

[1]Institute of Biological Sciences, Faculty of Science, University of Malaya, Kuala Lumpur, Malaysia.
[2]Department of Medical Microbiology, Faculty of Medicine, University of Malaya, Kuala Lumpur, Malaysia.

Proceedings and posters presented

1. Lim, S.Y., Yap, K.P., Teh, C.S.J., and Thong, K.L. (2016, December). Characterization of vancomycin-resistant *Enterococcus faecium* using whole genome sequencing approach. Paper presented at the 21st Biological Sciences Graduate Congress (BSGC) 2016, Kuala Lumpur, Malaysia.

2. Lim, S.Y., Yap, K.P., Teh, C.S.J., and Thong, K.L. (2015, December). Genetic variation in clinical vancomycin-resistant *Enterococcus faecium*. Paper presented at the International Congress of the Malaysian Society for Microbiology (ICMSM) 2015, Penang, Malaysia.

3. Lim, S.Y., Yap, K.P., Teh, C.S.J., and Thong, K.L. (2015, September). Comparative genomic analysis of clinical vancomycin-resistant *Enterococcus faecium.* Paper presented at the 22nd Annual Malaysian Society for Molecular Biology and Biotechnology (MSMBB) Scientific Meeting 2015, Kuala Lumpur, Malaysia.

APPENDICES

Appendix 1: Media, Buffers, and Solutions

Brain heart infusion broth

Brain heart infusion broth	14.8g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

Brain heart infusion agar

Brain heart infusion agar	18.8g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

Tryptic soy broth (TSB)

Tryptic soy broth	12.0g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

Mueller-hinton (MH) II agar

Mueller-hinton agar	8.4g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

Tryptic soy broth (TSB) + 1% glucose

Tryptic soy broth	12.0g
Glucose	4.0g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

Sodium chloride (0.85%)

Sodium chloride	3.4g
Distilled water up to	400ml

Autoclave at 121°C for 15min.

1X Phosphate buffered solution (PBS)

PBS tablet	1 tablet
Distilled water	100ml

Autoclave at 121°C for 15min.

1 M Tris, pH 8.0 [molecular weight = 121.44g]

Tris	48.45g
Deionized water	360ml

The pH of the solution was adjusted to pH8.0 by adding concentrated HCl, top up with deionized water to 400ml. Autoclave at 121°C for 15min.

0.5M Ethylenediaminetetraacetic acid (EDTA), pH 8.0 [molecular weight = 372.44g]

EDTA	74.44g
Deionized water	360ml

The pH of the solution was adjusted to pH8.0 by adding concentrated HCl, top up with deionized water to 400ml. Autoclave at 121°C for 15min.

Tris-EDTA (TE) buffer (10mM Tris: 1mM EDTA, pH 8.0)

1M Tris, pH 8.0	10ml
0.5M EDTA, pH 8.0	2ml
Deionized water to	1000ml

Autoclave at 121°C for 15min.

Cell suspension buffer (100mM Tris, 100mM EDTA, pH 8.0)

1M Tris, pH 8.0	10ml
0.5M EDTA, pH 8.0	20ml
Deionized water to	1000ml

Autoclave at 121°C for 15min.

10% Sarcosyl (N-Lauryl-Sarcosine [Molecular weight = 293.30g])

Sodium N-lauroyl-sarcosinate solution	10ml
Deionized water to	100ml

Autoclave at 121°C for 15min.

Cell lysis buffer (50mM Tris, 50mM EDTA, pH 8.0 + 1.0% Sarcosine)

1.0M Tris, pH 8.0	25ml
0.5M EDTA, pH 8.0	50ml
10% Sarcosyl	50ml
Deionized water to	500ml

Autoclave at 121°C for 15min.

10X Tris-Borate EDTA buffer (TBE), pH 8.3

Trizma base	121.1g
Othoboric acid	61.8g
EDTA (Ultra-Pure Grade)	0.745g
Deionized water to	1000ml

The pH of the stock solution was adjusted to pH 8.3 and autoclaved. It was then diluted to 0.5X for routine use. Autoclave at 121°C for 15min.

0.5X TBE

10X TBE	50ml
Deionized water	950ml

1.0% Seakem gold agarose

Seakem gold agarose	0.1g
TE buffer	10ml

1.5% Agarose for PCR

LE Agarose powder	1.5g
0.5X TBE buffer	100ml

1.0% Crystal violet solution

Crystal violet powder	0.1g
Distilled water	10ml

80:20 Ethanol-acetone

Ethanol	80ml
Acetone	20ml

Appendix 2: Strains' information

	VRE2	**VREr5**	**VREr6**	**VREr7**
Patient	X	Y	Y	Y
Isolation date	17-Jan-2011	4-Mar-2011	12-Mar-2011	18-Mar-2011
Isolation site	Blood	CSF	Blood	Urine
Genome size (bp)	2862609	2898367	3014993	3021201
GC%	37.85	37.8	37.61	37.6
CDS	2853	2906	3049	3057
tRNA	55	55	58	58
rRNA	3	3	3	3
Sequence type (ST)	80	80	203	203
Clonal complex	17	17	17	17

Appendix 3: Strains used in phylogenomic analysis

Strain	Genbank accession number	Country	Source	Sequence type	Clonal complex
Aus0085	CP006620.1	Australia	Blood	203	17
Aus0004	CP003351.1	Australia	Blood	17	17
TX16	CP003683.1	USA	Blood	18	17
E1133	AHWR00000000.1	USA	Faeces	117	17
E155	AUWX00000000.1	USA	Faeces	17	17
E1904	AHXQ00000000.1	Netherlands	Urine	210	17
E0120	AHWI00000000.1	Netherlands	Patient's ascites	27	17
E2560	AHYI00000000.1	Netherlands	Blood	78	17
E1185	AHWS00000000.1	France	Blood	26	17
E161	JXZA00000000.1	China	Urine	414	17
LCT-EF128	AJUP00000000.1	China	Bronchoalveolar lavage	160	17
E1731	AHXO00000000.1	Africa	Blood	18	17
E6045	AHYL00000000.1	Portugal	Hospitalized patient	78	17
E1392	AHWV01000046	UK	Hospitalized patient	64	17
E0333	AHWL00000000.1	Israel	Blood	80	17
VRE84	AIVF00000000.1	Denmark	Human	17	17

Appendix 4: PCR primers and cycling conditions

Primer	Sequence (5'-3')	Cycling condition	Target	Amplicon size (bp)
vanA-F	CATGAATAGAATAAAAGTTGCAATA	Initial denaturation step of 5 min at 94°C; 30 cycles of 1 min at 94°C, 1 min at 54°C, 1 min at 72°C; and a final elongation step of 10 min at 72°C	*vanA*	1030
vanA-R	CCCCTTTAACGCTAATACGATCAA			
Bee-1F	TCCCGCAAATCTTGACTGCT	Initial denaturation step of 10 min at 95°C; 35 cycles of 30 s at 94°C, 30 s at 54°C, 60 s at 72°C; and a final elongation step of 10 min at 72°C	*bee*-1	382
Bee-1R	ACAGCTTGTGTTCCATCCGT			
Bee-2F	GTGGGACCAAGCACCTGATG		*bee*-2	569
Bee-2R	TCGCCGTACCGATCAAAAGAA			
P101D12-1	CAAGATAAGCATCTGTTACATCATAGGCTG		*bee*-3	532
Bee-12	CGGACAAAGAATTAGCTACCGTTCAC			
ESP 14F	AGATTTCATCTTTGATTCTTGG	Initial denaturation step of 10 min at 95°C; 30 cycles of 30 s at 94°C, 30 s at 58°C, 30 s at 72°C; and a final elongation step of 10 min at 72°C	*esp*	510
ESP 12R	AATTGATTCTTTAGCATCTGG			
AcmF1	GATTTTTGAGAGATGATATAGTAG	Initial denaturation step of 10 min at 95°C; 30 cycles of 30 s at 94°C, 30 s at 53°C, 30 s at 72°C; and a final elongation step of 10 min at 72°C	*acm*	1600
AcmR1	ATTCTCATTTGTAACGACTAGC			

Appendix 5: Primers for gap closing and sequencing of Tn*1546*-like transposon

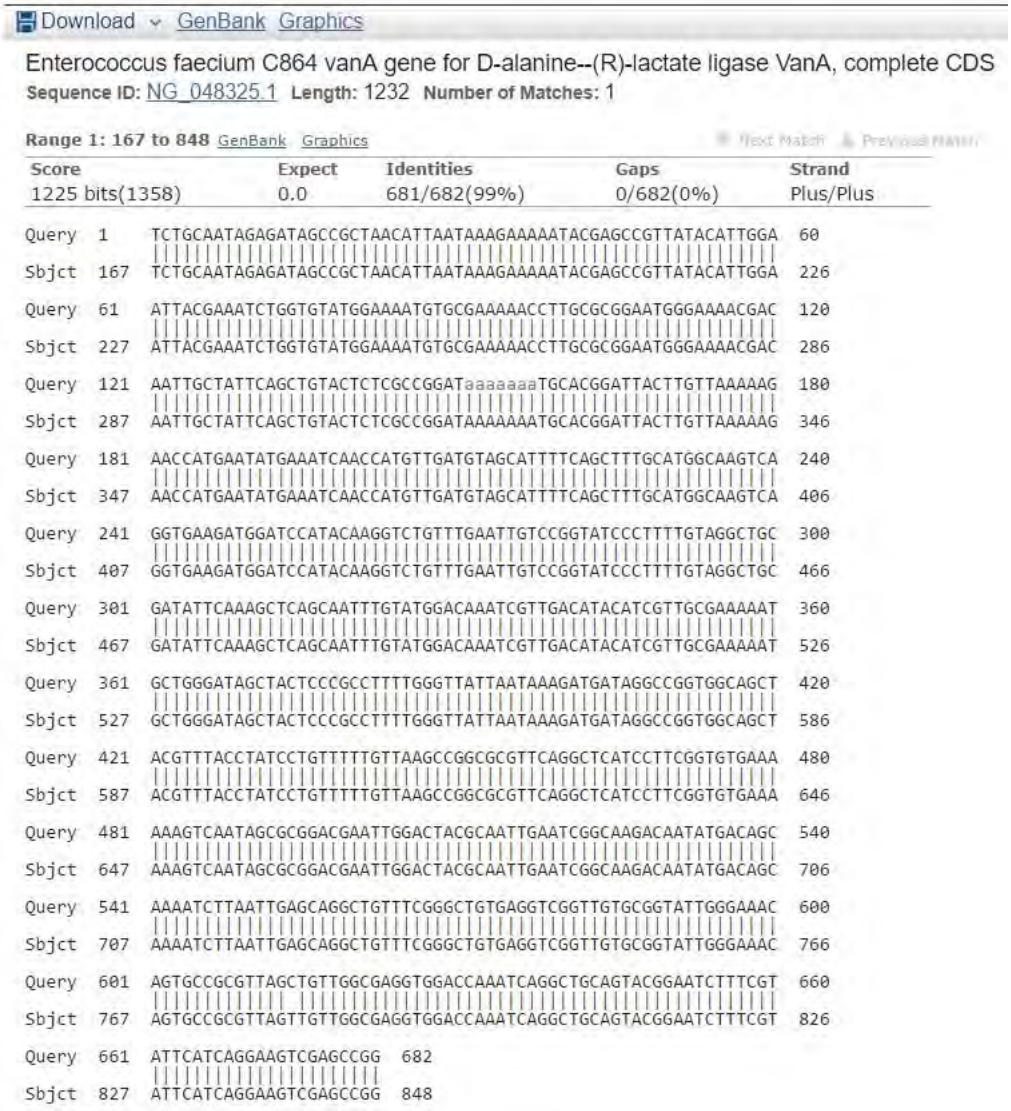
Primer	**Sequence (5'-3')**
5235.F	ATA TCA CGT TGG ACA AAG C
7035.R	TTA CGT CAT GCT CCT CTG AG
4511R	TCG GAG CTA ACC ACA TTC
ISV650F	ACC TTC ACG ATA GCT AAG GTT
ISV132R	AGG ATT ATA TAA GAA AAC CCG

Appendix 6: Primers for qPCR

Primer	Sequence (5'-3')	Product size (bp)
arcA-F	TTA GCG GCA GAA GCG ATT GA	165
arcA-R	ACG AAC CCC AGC CAT GAT TT	
fsr-F	TGG GCG GAA ACA GGA ATC AC	133
fsr-R	TCA TCA CCA AGG GAA CGC C	
bee2-F	ATC GAA GGG GTC GGC TTT AC	207
bee2-R	AAC CGC ATT TTT ACC GCC AC	
ddl-F	TTG CCT GGC GAA GTC GTA AA	162
ddl-R	TCC GCT TCC ACC TAA CAT CG	
ebpA-F	AAC GGT GCA GTA CAG ATG GG	171
ebpA-R	ACC ATC CGT CAG CAA AAC GA	
tetS-F	CGC TAT GGG TGT GAA CAA GG	109
tetS-R	TTC GGA AAT CTG CTG GCG TA	
repR-F	ACG TTG TTC GCG AGG GTT AT	137
repR-R	AGA GGG TTC AAG GGG GTT GA	

Appendix 7: Supplementary data

a) DNA sequencing BLAST result of *vanA*



Download GenBank Graphics

Enterococcus faecium C864 vanA gene for D-alanine--(R)-lactate ligase VanA, complete CDS

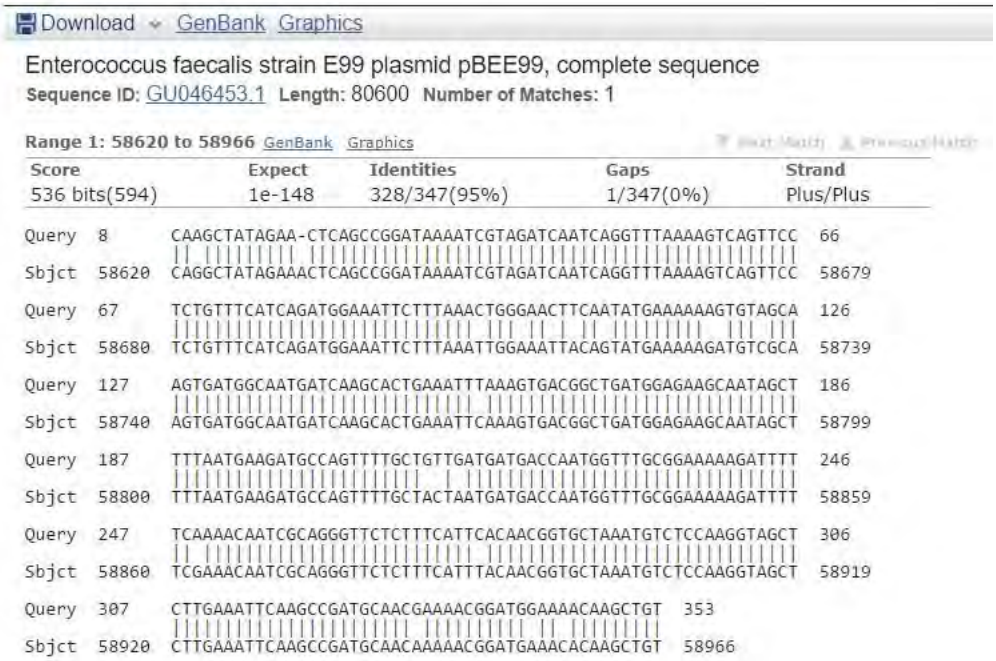
Sequence ID: NG_048325.1 **Length:** 1232 **Number of Matches:** 1

Range 1: 167 to 848 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
1225 bits(1358)	0.0	681/682(99%)	0/682(0%)	Plus/Plus

```
Query  1    TCTGCAATAGAGATAGCCGCTAACATTAATAAAGAAAAATACGAGCCGTTATACATTGGA  60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  167  TCTGCAATAGAGATAGCCGCTAACATTAATAAAGAAAAATACGAGCCGTTATACATTGGA  226

Query  61   ATTACGAAATCTGGTGTATGGAAAATGTGCGAAAAACCTTGCGCGGAATGGGAAAACGAC  120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  227  ATTACGAAATCTGGTGTATGGAAAATGTGCGAAAAACCTTGCGCGGAATGGGAAAACGAC  286

Query  121  AATTGCTATTCAGCTGTACTCTCGCCGGATaaaaaaaTGCACGGATTACTTGTTAAAAAG  180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  287  AATTGCTATTCAGCTGTACTCTCGCCGGATAAAAAAATGCACGGATTACTTGTTAAAAAG  346

Query  181  AACCATGAATATGAAATCAACCATGTTGATGTAGCATTTTCAGCTTTGCATGGCAAGTCA  240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  347  AACCATGAATATGAAATCAACCATGTTGATGTAGCATTTTCAGCTTTGCATGGCAAGTCA  406

Query  241  GGTGAAGATGGATCCATACAAGGTCTGTTTGAATTGTCCGGTATCCCTTTTGTAGGCTGC  300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  407  GGTGAAGATGGATCCATACAAGGTCTGTTTGAATTGTCCGGTATCCCTTTTGTAGGCTGC  466

Query  301  GATATTCAAAGCTCAGCAATTTGTATGGACAAATCGTTGACATACATCGTTGCGAAAAAT  360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  467  GATATTCAAAGCTCAGCAATTTGTATGGACAAATCGTTGACATACATCGTTGCGAAAAAT  526

Query  361  GCTGGGATAGCTACTCCCGCCTTTTGGGTTATTAATAAAGATGATAGGCCGGTGGCAGCT  420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  527  GCTGGGATAGCTACTCCCGCCTTTTGGGTTATTAATAAAGATGATAGGCCGGTGGCAGCT  586

Query  421  ACGTTTACCTATCCTGTTTTTGTTAAGCCGGCGCGTTCAGGCTCATCCTTCGGTGTGAAA  480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  587  ACGTTTACCTATCCTGTTTTTGTTAAGCCGGCGCGTTCAGGCTCATCCTTCGGTGTGAAA  646

Query  481  AAAGTCAATAGCGCGGACGAATTGGACTACGCAATTGAATCGGCAAGACAATATGACAGC  540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  647  AAAGTCAATAGCGCGGACGAATTGGACTACGCAATTGAATCGGCAAGACAATATGACAGC  706

Query  541  AAAATCTTAATTGAGCAGGCTGTTTCGGGCTGTGAGGTCGGTTGTGCGGTATTGGGAAAC  600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  707  AAAATCTTAATTGAGCAGGCTGTTTCGGGCTGTGAGGTCGGTTGTGCGGTATTGGGAAAC  766

Query  601  AGTGCCGCGTTAGCTGTTGGCGAGGTGGACCAAATCAGGCTGCAGTACGGAATCTTTCGT  660
            |||||||||||| ||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  767  AGTGCCGCGTTAGTTGTTGGCGAGGTGGACCAAATCAGGCTGCAGTACGGAATCTTTCGT  826

Query  661  ATTCATCAGGAAGTCGAGCCGG  682
            |||||||||||||||||||||
Sbjct  827  ATTCATCAGGAAGTCGAGCCGG  848
```

b) DNA sequencing BLAST result of *bee*-1 homolog*



Download GenBank Graphics

Enterococcus faecalis strain E99 plasmid pBEE99, complete sequence

Sequence ID: GU046453.1 Length: 80600 Number of Matches: 1

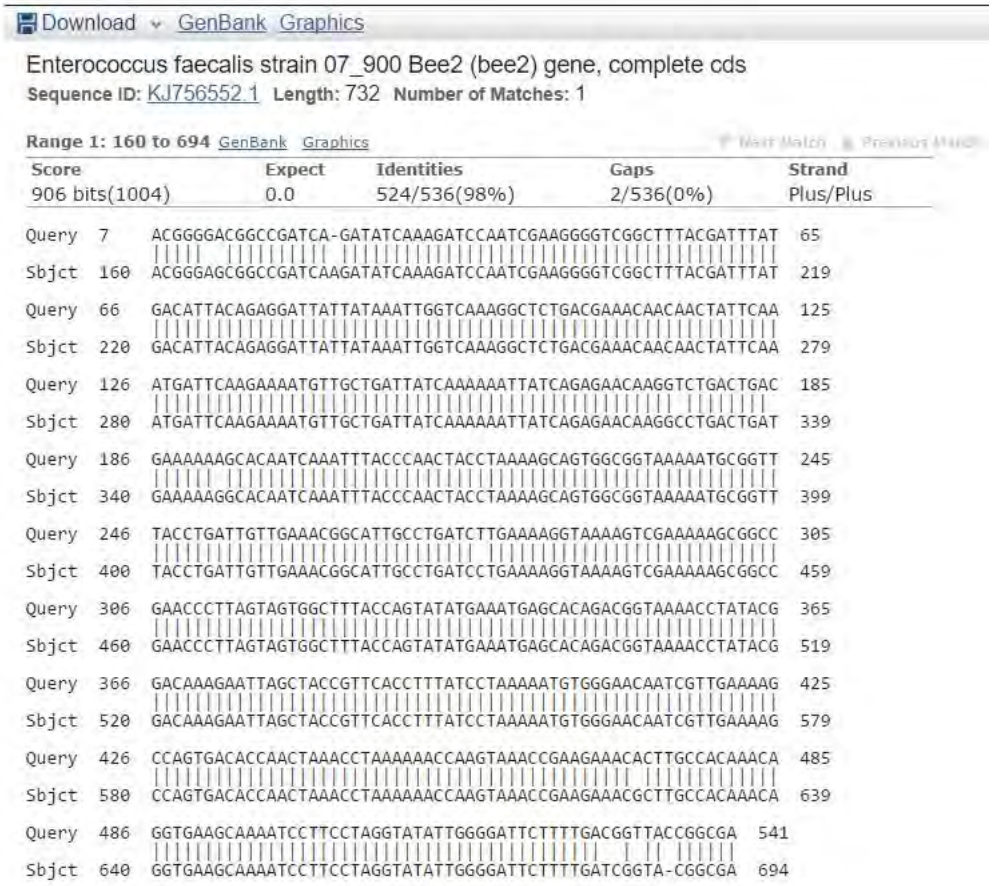
Range 1: 58620 to 58966 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
536 bits(594)	1e-148	328/347(95%)	1/347(0%)	Plus/Plus

```
Query  8      CAAGCTATAGAA-CTCAGCCGGATAAAATCGTAGATCAATCAGGTTTAAAAGTCAGTTCC  66
              || ||||||||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  58620  CAGGCTATAGAAACTCAGCCGGATAAAATCGTAGATCAATCAGGTTTAAAAGTCAGTTCC  58679

Query  67     TCTGTTTCATCAGATGGAAATTCTTTAAACTGGGAACTTCAATATGAAAAAAGTGTAGCA  126
              |||||||||||||||||||||||||||||| ||| || | || |||||||||  ||| |||
Sbjct  58680  TCTGTTTCATCAGATGGAAATTCTTTAAATTGGAAATTACAGTATGAAAAAGATGTCGCA  58739

Query  127    AGTGATGGCAATGATCAAGCACTGAAATTTAAAGTGACGGCTGATGGAGAAGCAATAGCT  186
              |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  58740  AGTGATGGCAATGATCAAGCACTGAAATTCAAAGTGACGGCTGATGGAGAAGCAATAGCT  58799

Query  187    TTTAATGAAGATGCCAGTTTTGCTGTTGATGATGACCAATGGTTTGCGGAAAAAGATTTT  246
              |||||||||||||||||||||||||  | ||||||||||||||||||||||||||||||||
Sbjct  58800  TTTAATGAAGATGCCAGTTTTGCTACTAATGATGACCAATGGTTTGCGGAAAAAGATTTT  58859

Query  247    TCAAAACAATCGCAGGGTTCTCTTTCATTCACAACGGTGCTAAATGTCTCCAAGGTAGCT  306
              || |||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  58860  TCGAAACAATCGCAGGGTTCTCTTTCATTTACAACGGTGCTAAATGTCTCCAAGGTAGCT  58919

Query  307    CTTGAAATTCAAGCCGATGCAACGAAAACGGATGGAAAACAAGCTGT  353
              ||||||||||||||||||||||| |||||||||| || |||||||||
Sbjct  58920  CTTGAAATTCAAGCCGATGCAACAAAAACGGATGAAACACAAGCTGT  58966
```

**bee*-1 homolog of VREr5 has low identity with that found in *E. faecalis.* The sequenced region represents the homologous region between the two genes. The *E. faecalis bee* locus is found in plasmid pBEE99.

c) DNA sequencing BLAST result of *bee*-2 homolog



Download GenBank Graphics

Enterococcus faecalis strain 07_900 Bee2 (bee2) gene, complete cds

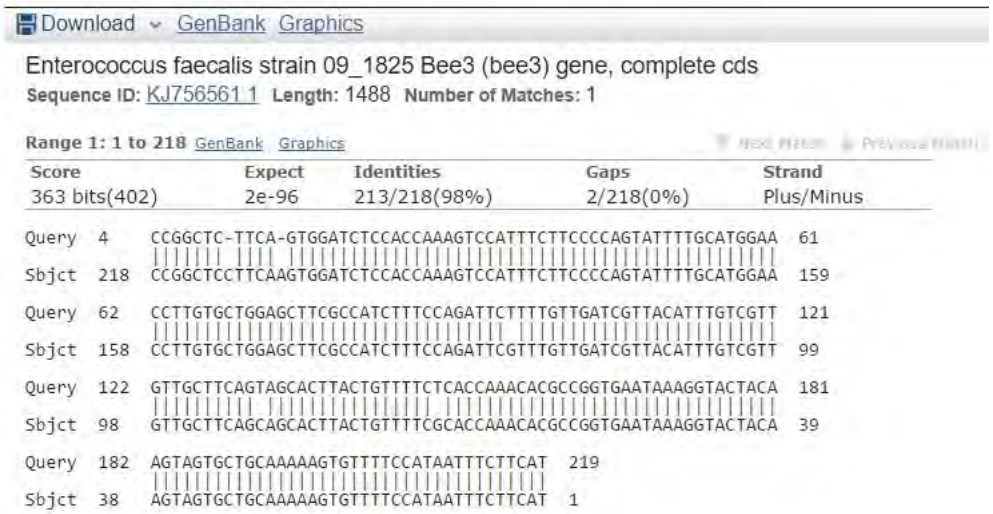
Sequence ID: KJ756552.1 **Length:** 732 **Number of Matches:** 1

Range 1: 160 to 694 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
906 bits(1004)	0.0	524/536(98%)	2/536(0%)	Plus/Plus

```
Query  7    ACGGGGACGGCCGATCA-GATATCAAAGATCCAATCGAAGGGGTCGGCTTTACGATTTAT  65
            |||||  |||||||||| ||||||||||||||||||||||||||||||||||||||||||
Sbjct  160  ACGGGAGCGGCCGATCAAGATATCAAAGATCCAATCGAAGGGGTCGGCTTTACGATTTAT  219

Query  66   GACATTACAGAGGATTATTATAAATTGGTCAAAGGCTCTGACGAAACAACAACTATTCAA  125
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  220  GACATTACAGAGGATTATTATAAATTGGTCAAAGGCTCTGACGAAACAACAACTATTCAA  279

Query  126  ATGATTCAAGAAAATGTTGCTGATTATCAAAAAATTATCAGAGAACAAGGTCTGACTGAC  185
            |||||||||||||||||||||||||||||||||||||||||||||||| ||||||||
Sbjct  280  ATGATTCAAGAAAATGTTGCTGATTATCAAAAAATTATCAGAGAACAAGGCCTGACTGAT  339

Query  186  GAAAAAAGCACAATCAAATTTACCCAACTACCTAAAAGCAGTGGCGGTAAAAATGCGGTT  245
            |||||| ||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  340  GAAAAAGGCACAATCAAATTTACCCAACTACCTAAAAGCAGTGGCGGTAAAAATGCGGTT  399

Query  246  TACCTGATTGTTGAAACGGCATTGCCTGATCTTGAAAAGGTAAAAGTCGAAAAAGCGGCC  305
            |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||
Sbjct  400  TACCTGATTGTTGAAACGGCATTGCCTGATCCTGAAAAGGTAAAAGTCGAAAAAGCGGCC  459

Query  306  GAACCCTTAGTAGTGGCTTTACCAGTATATGAAATGAGCACAGACGGTAAAACCTATACG  365
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  460  GAACCCTTAGTAGTGGCTTTACCAGTATATGAAATGAGCACAGACGGTAAAACCTATACG  519

Query  366  GACAAAGAATTAGCTACCGTTCACCTTTATCCTAAAAATGTGGGAACAATCGTTGAAAAG  425
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  520  GACAAAGAATTAGCTACCGTTCACCTTTATCCTAAAAATGTGGGAACAATCGTTGAAAAG  579

Query  426  CCAGTGACACCAACTAAACCTAAAAAACCAAGTAAACCGAAGAAACACTTGCCACAAACA  485
            |||||||||||||||||||||||||||||||||||||||||||||| ||||||||||||
Sbjct  580  CCAGTGACACCAACTAAACCTAAAAAACCAAGTAAACCGAAGAAACGCTTGCCACAAACA  639

Query  486  GGTGAAGCAAAATCCTTCCTAGGTATATTGGGGATTCTTTTGACGGTTACCGGCGA  541
            |||||||||||||||||||||||||||||||||||||||||||  | || ||||||
Sbjct  640  GGTGAAGCAAAATCCTTCCTAGGTATATTGGGGATTCTTTTGATCGGTA-CGGCGA  694
```

d) DNA sequencing BLAST result of *bee*-3 homolog



Download GenBank Graphics

Enterococcus faecalis strain 09_1825 Bee3 (bee3) gene, complete cds

Sequence ID: KJ756561.1 **Length:** 1488 **Number of Matches:** 1

Range 1: 1 to 218 GenBank Graphics

Score	Expect	Identities	Gaps	Strand
363 bits(402)	2e-96	213/218(98%)	2/218(0%)	Plus/Minus

```
Query  4    CCGGCTC-TTCA-GTGGATCTCCACCAAAGTCCATTTCTTCCCCAGTATTTTGCATGGAA  61
            ||||||| |||| ||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  218  CCGGCTCCTTCAAGTGGATCTCCACCAAAGTCCATTTCTTCCCCAGTATTTTGCATGGAA  159

Query  62   CCTTGTGCTGGAGCTTCGCCATCTTTCCAGATTCTTTTGTTGATCGTTACATTTGTCGTT  121
            |||||||||||||||||||||||||||||||| ||||||||||||||||||||||||||
Sbjct  158  CCTTGTGCTGGAGCTTCGCCATCTTTCCAGATTCGTTTGTTGATCGTTACATTTGTCGTT  99

Query  122  GTTGCTTCAGTAGCACTTACTGTTTTCTCACCAAACACGCCGGTGAATAAAGGTACTACA  181
            |||||||||| |||||||||||||| ||||||||||||||||||||||||||||||||||
Sbjct  98   GTTGCTTCAGCAGCACTTACTGTTTTCGCACCAAACACGCCGGTGAATAAAGGTACTACA  39

Query  182  AGTAGTGCTGCAAAAAGTGTTTTCCATAATTTCTTCAT  219
            |||||||||||||||||||||||||||||||||||||
Sbjct  38   AGTAGTGCTGCAAAAAGTGTTTTCCATAATTTCTTCAT  1
```

e) Complete list of up-regulated genes

Feature ID	Annotation	Fold change	FDR-corrected p-value	GO[1]	KEGG	COG[2]
Gene.1916	M trans-acting positive regulator	2	5.87E-010	NA	NA	NA
Gene.2214	AraC family transcriptional regulator	2	3.47E-008	F; P; C	NA	K
Gene.2707	Integrase	2	2.17E-009	F; P	NA	L
Gene.970	ATPase	2	6.97E-009	F	Purine metabolism	S
Gene.1781	GntR family transcriptional regulator	2.01	4.05E-011	F; P	NA	K
Gene.1888	Methyltransferase	2.01	5.24E-008	F; P	NA	Q
Gene.2030	M trans-acting positive regulator	2.01	2.54E-013	NA	NA	S
Gene.45	D-isomer specific 2-hydroxyacid dehydrogenase	2.01	5.34E-014	F; P	Pyruvate metabolism	C
Gene.75	Operon 2	2.01	5.06E-009	F; P	Pyrimidine metabolism	F
Gene.2015	Resolvase	2.02	6.15E-007	F; P	NA	L
Gene.2035	Histidine kinase	2.02	1.35E-010	C; P; F	NA	T
Gene.2257	Coproporphyrinogen III oxidase	2.02	2.89E-011	P; C; F	Porphyrin and chlorophyll metabolism	H
Gene.2710	Transcription antiterminator	2.03	2.96E-013	F; P	NA	K
Gene.28	GntR family transcriptional regulator	2.03	8.75E-012	F; P	NA	K
Gene.894	ATPase	2.03	7.41E-013	NA	NA	U
Gene.136	Competence protein	2.04	3.30E-010	C	NA	S
Gene.37	ABC transporter permease	2.04	3.81E-013	C; P; F	NA	G
Gene.439	4-phosphoerythronate dehydrogenase	2.04	8.43E-014	F; P	Glycine, serine and threonine metabolism, Methane	C

					metabolism, Biosynthesis of antibiotics	
Gene.78	N5-carboxyaminoimidazole ribonucleotide mutase (purE)	2.04	1.09E-009	F; P	Purine metabolism, Biosynthesis of antibiotics	F
Gene.2113	Hexulose-6-phosphate isomerase	2.05	1.30E-011	P; F	Pyrimidine metabolism	G
Gene.2185	Transposase	2.05	2.33E-006	F; P	NA	L
Gene.2532	MULTISPECIES: hypothetical protein	2.05	4.57E-005			NA
Gene.982	Mannitol dehydrogenase	2.05	2.28E-013	F; P	Pentose and glucuronate interconversions	G
Gene.988	ABC transporter premease	2.05	5.92E-013	F; P; C	NA	T
Gene.1149	Hypothetical protein	2.06	1.58E-012	C	NA	S
Gene.2530	Two-component system response regulator receiver	2.06	1.20E-011	F; P; C	NA	T
Gene.902	Glucuronate isomerase	2.06	2.69E-013	P; F	Pentose and glucuronate interconversions	G
Gene.143	Late competence protein	2.07	8.87E-006	C	NA	NA
Gene.1455	MULTISPECIES: hypothetical protein	2.07	2.05E-006			NA
Gene.1875	PTS galactitol transporter subunit IIC	2.07	8.43E-014	F; P; C	NA	G
Gene.1999	Integral membrane	2.07	2.31E-008	C; P; F	NA	S
Gene.2551	PTS mannose transporter subunit IID	2.07	3.90E-007	C; P; F	NA	G
Gene.1296	LPXTG-domain-containing cell wall anchor protein	2.08	1.25E-010	C	NA	M
Gene.1795	ABC transporter substrate-binding protein	2.08	7.96E-013	P	NA	G
Gene.2204	ATP-binding protein	2.08	1.11E-010	F; P; C	NA	T
Gene.2515	Major facilitator superfamily transporter	2.08	3.24E-011	P; C	NA	G
Gene.31	ABC transporter substrate-binding protein	2.08	1.06E-010	NA	NA	G

Gene.334	Phage tail tape measure TP901 core region	2.08	3.65E-014	C	NA	S
Gene.454	PTS glucitol sorbitol-specific IIA component	2.08	7.46E-010	C; P; F	NA	G
Gene.401	Hypothetical protein	2.09	5.20E-004	NA	NA	NA
Gene.875	Hypothetical protein	2.09	5.58E-014	NA	NA	S
Gene.956	Hypothetical protein	2.09	6.81E-011	NA	NA	NA
Gene.1008	Integral membrane	2.1	5.14E-010			S
Gene.1289	Collagen-binding protein	2.1	1.17E-008	C	NA	M
Gene.1786	CCS family citrate carrier protein	2.1	4.04E-013	P; C; F	NA	C
Gene.2011	DNA polymerase V	2.1	1.58E-007	P; F	Purine metabolism, Pyrimidine metabolism	L
Gene.2104	Nucleotide pyrophosphohydrolase	2.1	3.67E-005			S
Gene.2123	PTS ascorbate transporter subunit IIB	2.1	6.91E-009	P; F	NA	G
Gene.2517	Efflux ABC transporter permease	2.1	8.24E-013	C	NA	V
Gene.2689	Resolvase	2.1	4.85E-012	F; P	Purine metabolism, Pyrimidine metabolism	L
Gene.340	Phage terminase	2.1	4.75E-011	NA	NA	S
Gene.808	Ferrichrome transport system permease	2.1	6.13E-011	C; P; F	NA	P
Gene.879	Ribonuclease HI	2.1	2.13E-011	F; P	NA	L
Gene.1317	Ammonium transporter AmtB	2.11	3.25E-012	F; P; C	NA	P
Gene.144	Competence protein	2.11	4.02E-005	C	NA	NA
Gene.2215	ABC transporter permease	2.11	6.46E-013	C; P	NA	P
Gene.900	Mannitol dehydrogenase	2.11	6.76E-014	F; P	Pentose and glucuronate interconversions	G
Gene.949	Hypothetical protein	2.11	7.76E-014	NA	NA	NA
Gene.973	Transglutaminase-like superfamily	2.11	7.49E-012	C	NA	E

Gene.996	Mannonate dehydratase	2.11	1.09E-013	P; F	Pentose and glucuronate interconversions	G
Gene.1013	GntR family transcriptional regulator	2.12	4.32E-011	F; P	NA	K
Gene.1300	MerR family transcriptional regulator	2.12	1.68E-007	F; P	NA	K
Gene.2681	DNA replication protein DnaD	2.12	2.82E-011	NA	NA	L
Gene.312	Integrase core domain partial	2.12	1.00E-011	F; P	NA	L
Gene.2127	PTS glucitol sorbitol- IIBC component	2.13	7.47E-015	C; P; F	NA	G
Gene.850	Amino acid permease/serine:threonine exchanger SteT	2.13	1.85E-013	C; F; P	NA	E
Gene.1279	ISL3 family transposase	2.14	3.14E-012	NA	NA	L
Gene.1281	Sortase	2.14	2.15E-008	C	NA	M
Gene.2508	alpha-L-rhamnosidase	2.14	2.12E-015	P; F	NA	S
Gene.1175	ComF family protein	2.15	1.52E-013	P; F	NA	S
Gene.1981	Hypothetical protein	2.15	3.42E-014	NA	NA	S
Gene.2102	Cell surface protein	2.15	8.70E-009	C	NA	NA
Gene.2109	PTS ascorbate transporter subunit IIC	2.15	8.97E-016	C; P; F	NA	G
Gene.69	Hypothetical protein	2.15	1.10E-014	C	NA	NA
Gene.2130	Glucitol operon activator/transcriptional regulator	2.16	6.74E-014	NA	NA	S
Gene.24	ABC transporter ATP-binding	2.16	5.38E-011	F	Purine metabolism	P
Gene.2526	Shikimate dehydrogenase	2.16	1.93E-013	F; P	Phenylalanine, tyrosine and tryptophan biosynthesis, Biosynthesis of antibiotics	E
Gene.2567	Hypothetical protein	2.16	8.78E-007	NA	NA	NA
Gene.62	Two-component system response regulator receiver	2.16	3.89E-011	F; P; C	NA	T
Gene.2682	MarR family transcriptional regulator	2.17	2.48E-006	F; P; C	NA	K

Gene.1290	DNA-3-methyladenine glycosylase I	2.18	1.95E-010	F; P	NA	L
Gene.1329	X-Pro dipeptidyl-peptidase	2.18	2.29E-010	C; F	NA	S
Gene.1676	ABC transporter permease	2.18	3.20E-010	C	NA	NA
Gene.2128	Short chain dehydrogenase reductase family oxidoreductase	2.18	1.07E-008	F; P	Fructose and mannose metabolism	S
Gene.2283	Restriction endonuclease	2.19	2.40E-004	F; P	NA	V
Gene.2711	Carbohydrate deacetylase	2.19	2.20E-014	F; P	NA	G
Gene.449	PTS system IIB component	2.19	2.53E-015	C; P; F	NA	G
Gene.1258	PTS fructose IIA component	2.2	6.14E-011	C; P; F	NA	G
Gene.1283	Sugar ABC transporter ATP-binding	2.2	1.79E-015	C; P	NA	G
Gene.895	Amino acid transporter	2.2	8.28E-017	C	NA	S
Gene.921	Haloacid dehalogenase	2.2	2.14E-010	NA	NA	S
Gene.981	Neopullulanase	2.2	5.62E-014	P; F	NA	G
Gene.56	ABC transporter permease	2.22	2.94E-011	C; P	NA	P
Gene.2718	Hypothetical protein	2.23	1.59E-013	NA	NA	NA
Gene.933	Hypothetical protein	2.23	2.90E-012	NA	NA	NA
Gene.2436	PTS mannose transporter subunit IID	2.24	9.20E-009	C; P; F	NA	G
Gene.53	ABC transporter permease	2.24	6.02E-014	C; P	NA	P
Gene.1316	(gadC) glutamate:gamma-aminobutyrate antiporter	2.25	8.35E-018	C; F; P	NA	E
Gene.17	Competence protein comEC/Rec2	2.26	2.74E-018	C; P	NA	S
Gene.1810	Two-component system response regulator	2.26	1.44E-011	F; P; C	NA	T
Gene.2081	Cell wall surface anchor protein	2.26	1.95E-013	C	NA	M
Gene.2509	Cell surface protein ebpC	2.26	1.73E-013	C	NA	M
Gene.379	Hypothetical protein	2.26	2.12E-009	NA	NA	S
Gene.890	Hypothetical protein	2.26	7.40E-009	NA	NA	NA
Gene.2626	Transposase	2.27	4.73E-009	F; P	NA	L
Gene.1319	Glutaminase A	2.29	1.46E-013	F; P	D-Glutamine and D-glutamate metabolism,	E

					Arginine biosynthesis, Alanine, aspartate and glutamate metabolism	
Gene.2507	DNA topoisomerase III	2.29	9.54E-013	F; C; P	NA	L
Gene.435	Mannonate dehydratase	2.29	4.13E-016	P; F	Pentose and glucuronate interconversions	G
Gene.1257	RpiR family transcriptional regulator	2.3	1.11E-015	F; P	NA	K
Gene.2220	Response regulator receiver	2.3	1.32E-012	F; P; C	NA	T
Gene.906	cro CI family transcriptional regulator	2.3	6.09E-018	P; F	NA	K
Gene.396	Hypothetical protein	2.31	6.40E-007	NA	NA	NA
Gene.903	MFS transporter	2.31	1.56E-013	P; C; F	NA	G
Gene.1156	Transcription antiterminator	2.32	2.26E-015	F; P	NA	K
Gene.448	Alcohol dehydrogenase	2.32	3.74E-014	F; P	Metabolism of xenobiotics by cytochrome P450, Naphthalene degradation, Glycine, serine and threonine metabolism, Chloroalkane and chloroalkene degradation, Glycolysis / Gluconeogenesis, Tyrosine metabolism, Retinol	C

					metabolism, Fatty acid degradation, Drug metabolism - cytochrome P450, Biosynthesis of antibiotics, alpha-Linolenic acid metabolism	
Gene.891	LPXTG-domain-containing cell wall anchor domain	2.32	1.33E-017	C	NA	M
Gene.914	TcpC-containing conjugal transfer protein	2.32	5.39E-011	C	NA	S
Gene.985	Hypothetical protein	2.32	2.23E-009	NA	NA	NA
Gene.2024	Bacteriocin-associated protein	2.34	4.57E-015	C	NA	S
Gene.2287	Hypothetical protein	2.34	2.24E-006	NA	NA	NA
Gene.2420	DNA invertase Pin	2.34	1.79E-009	F; P	NA	L
Gene.2421	Integrase	2.35	8.55E-010	F; P	NA	L
Gene.1337	Mobilization protein (plasmid)	2.36	4.63E-010	C	NA	NA
Gene.947	Hypothetical protein	2.36	4.11E-008	NA	NA	NA
Gene.441	Foldase	2.37	6.02E-014	F; P; C	NA	O
Gene.2523	AraC family transcriptional regulator	2.38	2.24E-012	F; P	NA	K
Gene.2422	IS3 family transposase	2.39	4.14E-010	NA	NA	L
Gene.2669	Integrase core domain partial	2.42	2.87E-004	F; P	NA	L
Gene.2512	M trans-acting positive regulator	2.45	2.46E-012	NA	NA	S
Gene.1342	Hypothetical protein	2.47	7.14E-011	NA	NA	NA
Gene.2511	Cell surface protein EbpB	2.47	1.78E-012	C	NA	M
Gene.2606	Integrase	2.47	7.55E-011	F; P	NA	L
Gene.54	Histidine kinase	2.5	3.31E-016	F; P; C	NA	T
Gene.409	Phage terminase	2.51	2.24E-009	NA	NA	NA
Gene.2506	von Willebrand factor (EbpA)	2.52	8.54E-019	C	NA	S

Gene.2648	Integrase	2.52	4.87E-019	F; P	NA	L
Gene.2013	Plasmid replication initiator A	2.54	6.85E-011	NA	NA	S
Gene.1016	Guanylate kinase	2.55	9.81E-019	P; F	Purine metabolism	F
Gene.905	Cell division protein FtsK	2.55	1.78E-016	F; C	NA	D
Gene.1847	Phage repressor protein	2.61	5.01E-009	NA	NA	K
Gene.1951	Hypothetical protein	2.62	6.49E-012	NA	NA	NA
Gene.2018	IS3 family transposase	2.63	1.73E-012	NA	NA	L
Gene.1496	Replication protein	2.66	8.89E-012	P; F; C	Purine metabolism, Pyrimidine metabolism	L
Gene.2546	Cobalt transporter	2.68	6.17E-012	F; P	NA	L
Gene.1497	Hypothetical protein	2.69	3.64E-012	NA	NA	NA
Gene.1336	Hypothetical protein	2.7	4.72E-013	NA	NA	NA
Gene.1288	DUF624 domain-containing protein	2.71	2.36E-020	C	NA	NA
Gene.1339	Mobilization protein	2.71	1.17E-011	NA	NA	S
Gene.1333	Initiator RepB plasmid replication protein	2.74	2.96E-013	P; F; C	Purine metabolism, Pyrimidine metabolism	L
Gene.405	DNA-binding protein	2.79	1.23E-007	F	NA	K
Gene.398	Phage head-tail adaptor	2.81	9.27E-008	NA	NA	NA
Gene.456	MULTISPECIES: hypothetical protein	2.81	1.56E-012	C	NA	S
Gene.991	LPXTG-domain-containing cell wall anchor domain protein	2.83	1.47E-017	C	NA	NA
Gene.938	Hypothetical protein	2.92	1.08E-013	NA	NA	NA
Gene.1344	Hypothetical protein	2.97	7.39E-015	C	NA	NA
Gene.462	Tetracycline resistance tetS	2.99	1.40E-014	F; P	NA	T
Gene.1335	MULTISPECIES: hypothetical protein	3.03	5.14E-016	C	NA	NA

Gene.460	Hypothetical protein	3.03	2.10E-014	F; C; P	Purine metabolism, Pyrimidine metabolism	NA
Gene.887	Transposase	3.03	3.99E-015	F; P	NA	L
Gene.2419	Cell filamentation protein Fic	3.2	7.02E-015	P	NA	D
Gene.1712	Transposase	3.21	4.13E-017	F	NA	L
Gene.1866	Transposase for insertion sequence element IS256 in transposon	3.29	2.06E-017	F; P	NA	L
Gene.1948	Replication initiation factor	3.31	3.49E-018	F; P	NA	NA
Gene.1947	Recombinase	3.34	2.02E-018	F; C; P	NA	S
Gene.2227	Replication initiation factor	3.52	2.17E-019	F; P	NA	NA
Gene.888	DNA recombinase	3.53	6.05E-019	F; P	NA	L
Gene.2605	Integrase	3.65	1.08E-019	F; P	nA	L
Gene.457	Plasmid replication protein RepR	3.89	1.61E-021	F; P	Purine metabolism, Pyrimidine metabolism	S
Gene.889	Recombinase	3.94	8.65E-023	F; P	NA	L
Gene.458	Integrase	4.25	5.99E-025	F; P	NA	L
Gene.641	Transposase	4.31	3.44E-025	F; P	NA	L
Gene.1334	Relaxase	4.33	5.14E-026	F; C; P	NA	U
Gene.417	Transposase	4.36	7.79E-026	F	NA	L
Gene.1338	Bacteriocin precursor	4.71	1.34E-027	C; P	NA	NA
Gene.1210	Hypothetical protein	4.75	1.38E-026	NA	NA	NA
Gene.1209	Plasmid replication initiation	4.82	5.91E-028	P; F; C	Purine metabolism, Pyrimidine metabolism	L
Gene.2189	IS6 family transposase	5.23	1.25E-030	F; P	NA	L

[1]C, cellular component; P, biological process; F, molecular function; NA, not applicable

[2][D] Cell cycle control, cell division, chromosome partitioning; [M] Cell wall/membrane/envelope biogenesis; [N] Cell motility; [O] Post-translational modification, protein turnover & chaperones; [T] Signal transduction mechanisms; [U] Intracellular trafficking, secretion & vesicular transport; [V] Defense mechanisms; [J] Translation, ribosomal structure & biogenesis; [K] Transcription; [L] Replication, recombination & repair; [C] Energy production & conversion; [G] Carbohydrate transport & metabolism; [E] Amino acid transport & metabolism; [F] Nucleotide transport & metabolism; [H] Coenzyme transport & metabolism; [I] Lipid transport & metabolism; [Q] Secondary metabolites biosynthesis, transport & catabolism; [P] Inorganic ion transport & metabolism; [S] Function unknown

f) Partial list of down-regulated genes

*Only the first 200 genes are shown here

Feature ID	Annotation	Fold change	FDR-corrected p-value	GO[1]	KEGG	COG[2]
Gene.802	Arginine deiminase ArcA	-287.52	3.09E-219	C; P; F	Biosynthesis of antibiotics, Arginine biosynthesis	E
Gene.805	Ornithine carbamoyltransferase ArcB	-259.23	2.04E-217	C; F; P	Biosynthesis of antibiotics, Arginine biosynthesis	E
Gene.502	Accessory regulator FsrB	-133.22	8.83E-198	P; C; F	NA	T
Gene.813	Carbamate kinase ArcC	-96.1	2.71E-186	F; P	Nitrogen metabolism, Arginine biosynthesis, Purine metabolism	E
Gene.1407	PTS mannose fructose sorbose IID component	-81.2	4.29E-143	C; P; F	NA	G
Gene.1201	Ribosomal Ctc-form	-78.1	7.80E-173	F; C; P	NA	J
Gene.1172	Ribosomal subunit interface	-68.03	1.91E-155	C; P	NA	J
Gene.1071	MULTISPECIES: hypothetical protein	-59.77	1.75E-119	C	NA	S
Gene.790	ATP synthase epsilon chain	-45.28	3.66E-134	C; F; P	NA	C
Gene.1182	Membrane protein	-43.83	1.73E-125	C	NA	S
Gene.1015	Alkyl hydroperoxide reductase subunit C	-43.47	1.39E-111	F; P	Phenylpropanoid biosynthesis	O
Gene.1788	Enolase	-42.28	5.25E-114	F; P; C comple	Biosynthesis of antibiotics, Methane	G

				x; C:cell surface	metabolism, Glycolysis / Gluconeogenesis	
Gene.2002	Hypothetical protein	-41.31	1.02E-125	NA	NA	S
Gene.1187	PspC family transcriptional regulator	-40.26	6.66E-139	C	NA	S
Gene.1973	C4-dicarboxylate anaerobic carrier	-39.89	3.10E-144	C	NA	S
Gene.986	Alkyl hydroperoxide F subunit	-38.9	7.88E-137	F; C; P	Phenylpropanoid biosynthesis, Glutathione metabolism	O
Gene.1817	Triosephosphate isomerase	-38.59	1.39E-116	P; C; F	Inositol phosphate metabolism, Glycolysis / Gluconeogenesis, Fructose and mannose metabolism, Carbon fixation in photosynthetic organisms, Biosynthesis of antibiotics	G
Gene.1413	PTS system mannose-specific IIC component	-37.69	1.44E-116	C; P; F	NA	G
Gene.1404	PTS mannose fructose sorbose IIB component	-37.14	2.70E-111	C; P; F	NA	G
Gene.1100	Tryptophan-rich sensory	-36.3	7.22E-121	C	NA	T
Gene.1598	S-ribosylhomocysteinase LuxS	-35.99	5.53E-096	P; F	Cysteine and methionine metabolism	T
Gene.1801	Glyceraldehyde-3-phosphate dehydrogenase	-35.86	1.21E-111	F; P	NA	G

Gene.1791	Phosphoglycerate kinase	-34.99	2.22E-114	F; C; P	Glycolysis / Gluconeogenesis, Carbon fixation in photosynthetic organisms, Biosynthesis of antibiotics	G
Gene.781	ATP synthase subunit beta	-33.54	3.00E-122	C; F; P	NA	C
Gene.784	ATP synthase gamma chain	-31.78	6.89E-108	C; F; P	NA	C
Gene.2652	Glucose-6-phosphate isomerase	-30.3	3.88E-099	F; P; C	Glycolysis / Gluconeogenesis, Amino sugar and nucleotide sugar metabolism, Starch and sucrose metabolism, Biosynthesis of antibiotics, Pentose phosphate pathway	G
Gene.675	Elongation factor Tu	-29.47	5.95E-119	F; C; P	Purine metabolism, Thiamine metabolism	J
Gene.780	ATP synthase subunit alpha	-28.77	4.66E-101	C; F; P	NA	C
Gene.2093	Nucleoside diphosphate kinase	-27.94	6.68E-120	F; P; C	Purine metabolism, Pyrimidine metabolism, Biosynthesis of antibiotics	F
Gene.2411	Preprotein translocase subunit	-27.81	2.78E-108	C	NA	U
Gene.1975	Dipeptidase	-27.76	4.79E-121	F; P	NA	E

Gene.138	Universal stress protein	-27.56	4.82E-116	C; P	NA	T
Gene.1851	Lactoylglutathione lyase	-27.34	6.22E-104	F; P	Pyruvate metabolism	E
Gene.1755	Bee3-like protein	-26.91	8.98E-124	F; C	NA	M
Gene.1231	Hypothetical protein	-25.37	1.66E-119	NA	NA	S
Gene.1177	Glutamyl-tRNA(Gln) amidotransferase subunit E	-24.65	6.57E-101	F	Aminoacyl-tRNA biosynthesis	S
Gene.787	ATP synthase subunit delta	-24.36	1.89E-107	C; F; P	NA	C
Gene.1138	Chromosome partition Smc	-24.26	4.80E-123	NA	NA	S
Gene.2332	DNA protection during starvation	-23.73	0	F; P; C	NA	P
Gene.2579	Sugar ABC transporter ATP-binding protein	-21.6	2.21E-030	F; P; C	Purine metabolism	G
Gene.1894	Universal stress protein	-21.48	1.52E-107	C; P	NA	T
Gene.2173	50S ribosomal L7 L12	-21.04	3.35E-100	F; C; P	NA	J
Gene.1170	General stress protein	-19.87	3.81E-112	C	NA	S
Gene.1152	UTP--glucose-1-phosphate uridylyltransferase	-19.72	4.25E-110	P: F	Galactose metabolism, Pentose and glucuronate interconversions, Amino sugar and nucleotide sugar metabolism, Starch and sucrose metabolism, Biosynthesis of antibiotics	M
Gene.1616	alpha-glycerophosphate oxidase	-19.43	2.26E-048	P; C; F	Glycerophospholipid metabolism	C
Gene.1599	Glutathione peroxidase	-19.4	3.66E-096	P; F	Glutathione metabolism, Arachidonic acid metabolism	O

Gene.1504	Phosphoenolpyruvate- phosphotransferase	-19.11	1.00E-069	C; F; P	NA	G
Gene.1097	Superoxide dismutase	-19.08	3.03E-104	F; P	NA	P
Gene.2171	50S ribosomal L10	-18.89	1.04E-108	F; P; C	NA	J
Gene.1622	Glycerol kinase	-18.86	3.19E-038	F P	Glycerolipid metabolism	C
Gene.2479	PTS trehalose-specific IIBC component	-17.92	1.36E-100	F; P; C	NA	G
Gene.1549	NADH peroxidase	-17.53	4.98E-082	C; F; P	NA	P
Gene.367	Cell division protein	-17.43	8.83E-098	P; C	NA	D
Gene.721	2,3-bisphosphoglycerate-dependent phosphoglycerate mutase	-17.09	9.29E-097	P; F	NA	G
Gene.140	Histidine triad	-16.92	6.55E-096	F	NA	F, G
Gene.764	50S ribosomal L17	-16.9	4.46E-088	F; C; P	NA	J
Gene.1490	Large conductance mechanosensitive channel	-16.85	4.94E-099	F; P; C	NA	M
Gene.2677	NADH oxidase	-16.46	3.74E-048	C; F; P	Oxidative phosphorylation	P
Gene.1900	Pyruvate kinase	-16.37	2.73E-039	F; P	Glycolysis / Gluconeogenesis, Purine metabolism, Pyruvate metabolism, Biosynthesis of antibiotics	G
Gene.118	Foldase	-16.08	6.78E-100	F; P; C	NA	O
Gene.1858	Thioredoxin	-15.86	4.48E-079	C; F; P	NA	O
Gene.1761	Peptidase (Bee-2)	-15.33	2.35E-095	C	NA	M
Gene.2165	Hypothetical protein	-15.33	4.10E-071	NA	NA	S
Gene.2143	Lactose-specific phosphotransferase enzyme IIA component	-15.3	6.50E-084	C; F; P	NA	G
Gene.738	Ribosomal S7	-15.13	4.21E-078	F; C; P	NA	J

Gene.47	Pyruvate dehydrogenase E1 component subunit beta	-14.89	5.27E-083	F; P	Glycolysis / Gluconeogenesis, Pyruvate metabolism, Citrate cycle (TCA cycle), Biosynthesis of antibiotics	C
Gene.1018	Membrane protein	-14.45	1.74E-083	C	NA	S
Gene.2135	Tagatose 1,6-diphosphate aldolase	-14.35	1.73E-097	F; P	Galactose metabolism	G
Gene.1136	Phosphoglucomutase	-13.61	2.81E-055	F; P	Glycolysis / Gluconeogenesis, Purine metabolism, Galactose metabolism, Streptomycin biosynthesis, Amino sugar and nucleotide sugar metabolism, Starch and sucrose metabolism, Biosynthesis of antibiotics, Pentose phosphate pathway	G
Gene.22	Dihydrolipoamide acetyltransferase	-13.58	2.18E-084	P; F	NA	C
Gene.1559	Galactokinase	-13.36	1.26E-068	F; C; P	Galactose metabolism, Amino sugar and nucleotide sugar metabolism	G
Gene.658	Pyruvate oxidase	-13.35	9.57E-073	F; P	Pyruvate metabolism	E

Gene.1039	Tyrosine decarboxylase	-13.23	1.86E-084	F; P	Tyrosine metabolism, Isoquinoline alkaloid biosynthesis, Methane metabolism	E
Gene.29	Dihydrolipoyl dehydrogenase	-13.08	2.65E-085	C; F; P	Glycine, serine and threonine metabolism, Glycolysis / Gluconeogenesis, Propanoate metabolism, Pyruvate metabolism, Citrate cycle (TCA cycle), Valine, leucine and isoleucine degradation, Biosynthesis of antibiotics	C
Gene.1713	Transketolase	-12.87	6.07E-056	F; P	Biosynthesis of ansamycins, Carbon fixation in photosynthetic organisms, Biosynthesis of antibiotics, Pentose phosphate pathway	G
Gene.1414	Hypothetical protein	-12.84	2.45E-094	C	NA	S
Gene.607	ABC transporter permease	-12.31	5.79E-086	C; P	NA	E

Gene.2140	Galactose-6-phosphate isomerase subunit lacB	-12.3	1.43E-079	F; P	Galactose metabolism	G
Gene.734	50S ribosomal L5	-12.3	9.28E-071	F; C; P	NA	J
Gene.2292	DNA-directed RNA polymerase subunit beta	-12.2	1.26E-074	F; P	Purine metabolism, Pyrimidine metabolism	K
Gene.1941	Hypothetical protein	-12.12	1.26E-065	C	NA	S
Gene.139	Regulatory Spx	-12.04	0	P; C	NA	K
Gene.1530	Ribose-5-phosphate isomerase	-11.91	1.28E-026	F	NA	S
Gene.348	Cell division protein	-11.91	7.10E-070	F; P; C	NA	D
Gene.1978	Ribosomal S1	-11.81	4.46E-057	F; C; P	NA	J
Gene.2160	50S ribosomal L1	-11.75	1.24E-076	P; F C	NA	J
Gene.622	Formate acetyltransferase	-11.55	7.70E-031	P; C; F	Propanoate metabolism, Pyruvate metabolism, Butanoate metabolism	C
Gene.1322	Trypsin-like serine protease	-11.24	7.82E-081	P; C; F	NA	S
Gene.872	Orotate phosphoribosyltransferase	-11.21	2.41E-054	F; P	Drug metabolism - other enzymes, Pyrimidine metabolism	F
Gene.370	YggS family pyridoxal phosphate enzyme	-11.11	7.03E-062	F	D-Alanine metabolism	F
Gene.1427	Transcription antitermination factor	-11.05	1.65E-073	F; P	NA	K
Gene.652	Elongation factor G	-10.92	9.04E-063	F; C; P	NA	J
Gene.1147	Glycerol-3-phosphate dehydrogenase	-10.74	1.81E-079	P; C; F	Glycerophospholipid metabolism	C

Gene.1528	Acetyl- biotin carboxyl carrier	-10.68	1.62E-061	F; P; C	Fatty acid biosynthesis, Tetracycline biosynthesis, Propanoate metabolism, Aflatoxin biosynthesis, Carbon fixation pathways in prokaryotes, Pyruvate metabolism, Biosynthesis of antibiotics	I
Gene.810	LacI family transcriptional regulator	-10.63	8.14E-076	F; P	NA	K
Gene.1303	50S ribosomal L19	-10.49	8.77E-075	F; C; P	NA	J
Gene.1400	ATP-binding protein	-10.45	3.55E-083	C	NA	S
Gene.788	ATP synthase subunit B	-10.41	2.34E-076	C; F; P	NA	C
Gene.1323	General stress protein	-10.23	3.89E-075	NA	NA	S
Gene.2389	Aldehyde-alcohol dehydrogenase	-10.15	4.47E-053	F; P	Glycine, serine and threonine metabolism, Tyrosine metabolism, Retinol metabolism, Fatty acid degradation, Butanoate metabolism, Biosynthesis of antibiotics, Benzoate	C

					degradation, Dioxin degradation, alpha-Linolenic acid metabolism, Metabolism of xenobiotics by cytochrome P450, Naphthalene degradation, Chloroalkane and chloroalkene degradation, Glycolysis / Gluconeogenesis, Xylene degradation, Pyruvate metabolism, Phenylalanine metabolism, Drug metabolism - cytochrome P450	
Gene.749	50S ribosomal L6	-10.12	3.80E-072	F; C; P	NA	J
Gene.809	UDP-glucose 4-epimerase	-10.12	1.32E-067	F; C; P	Galactose metabolism, Amino sugar and nucleotide sugar metabolism	M
Gene.38	Pyruvate dehydrogenase (acetyl-transferring) E1 alpha subunit	-10.05	1.80E-079	F; P	Glycolysis / Gluconeogenesis, Pyruvate metabolism, Citrate	C

					cycle (TCA cycle), Biosynthesis of antibiotics	
Gene.472	Alpha,alpha-phosphotrehalase	-9.96	3.08E-070	C; P; F	Starch and sucrose metabolism	G
Gene.773	50S ribosomal L24	-9.87	6.75E-070	F; C; P	NA	J
Gene.1140	Metal ion (Mn2+ Fe2+) transporter	-9.78	3.34E-054	C; P; F	NA	P
Gene.467	N-acetylmuramoyl-L-alanine amidase	-9.76	3.19E-077	F; P	Tryptophan metabolism, Aminobenzoate degradation, Arginine and proline metabolism, Styrene degradation, Phenylalanine metabolism	N, U
Gene.737	30S ribosomal S5	-9.68	4.80E-070	F; C; P	NA	J
Gene.1500	ATP-dependent Clp protease ATP-binding subunit	-9.66	6.38E-046	F; P	NA	O
Gene.2077	GNAT family acetyltransferase	-9.66	1.31E-073	F	NA	S
Gene.122	Aldo keto reductase	-9.49	2.28E-075	F; P	NA	C
Gene.1696	Uracil phosphoribosyltransferase	-9.47	2.68E-041	F; P	Pyrimidine metabolism	F
Gene.1360	ATP-dependent metallopeptidase	-9.43	6.71E-031	F; P; C	Purine metabolism	O
Gene.2136	Tagatose-6-phosphate kinase	-9.42	4.54E-079	F; P	Galactose metabolism	G
Gene.1439	Hypothetical protein	-9.38	2.45E-084	NA	NA	S
Gene.1752	VWA domain-containing (Bee-1)	-9.15	1.76E-071	C	NA	S
Gene.1405	Fructose-bisphosphate aldolase	-9.14	1.14E-058	F; P	Glycolysis / Gluconeogenesis, Fructose and	G

					mannose metabolism, Carbon fixation in photosynthetic organisms, Methane metabolism, Biosynthesis of antibiotics, Pentose phosphate pathway	
Gene.1318	Cell division protein	-9.04	1.91E-077	P; C	NA	D
Gene.2133	PTS system lactose-specific EIICB component	-9.02	1.61E-073	F; C; P	NA	G
Gene.1630	Tyrosine--tRNA ligase	-8.91	4.41E-063	F; P; C	Aminoacyl-tRNA biosynthesis	J
Gene.166	Hypothetical protein	-8.87	8.98E-016	NA	NA	S
Gene.342	Glucose-6-phosphate dehydrogenase	-8.87	1.10E-061	F	Ubiquinone and other terpenoid-quinone biosynthesis, Benzoate degradation	S
Gene.105	Oligoendopeptidase F	-8.85	4.67E-071	P; F	NA	E
Gene.1743	CAAX amino terminal protease	-8.85	1.74E-071	P; C; F	NA	S
Gene.1636	MIP family channel	-8.84	4.29E-025	C; P; F	NA	G
Gene.2467	ATP-dependent Clp protease proteolytic subunit	-8.49	6.94E-072	C; P; F	NA	O
Gene.2501	Hypothetical protein	-8.44	1.92E-027	NA	NA	NA
Gene.293	Alpha3-beta1 integrin-binding family	-8.43	3.27E-057	F	NA	S
Gene.377	S4 domain-containing protein	-8.41	1.30E-074	F; C	NA	J
Gene.1291	Flavin reductase	-8.4	2.32E-069	F; P	Riboflavin metabolism	S
Gene.2232	Universal stress	-8.38	3.95E-061	C; P	NA	T

Gene.1515	Acetyl-coenzyme A carboxylase carboxyl transferase subunit beta	-8.36	1.10E-066	F; P; C	Fatty acid biosynthesis, Tetracycline biosynthesis, Propanoate metabolism, Aflatoxin biosynthesis, Carbon fixation pathways in prokaryotes, Pyruvate metabolism, Biosynthesis of antibiotics	I
Gene.1053	Zinc-containing alcohol dehydrogenase	-8.34	6.36E-075	F; P	Metabolism of xenobiotics by cytochrome P450, Naphthalene degradation, Glycine, serine and threonine metabolism, Chloroalkane and chloroalkene degradation, Glycolysis / Gluconeogenesis, Tyrosine metabolism, Retinol metabolism, Fatty	C

					acid degradation, Drug metabolism - cytochrome P450, Methane metabolism, Biosynthesis of antibiotics, alpha-Linolenic acid metabolism	
Gene.1526	MarR family transcriptional regulator	-8.3	1.07E-069	F; P	NA	K
Gene.735	Single-stranded DNA-binding protein	-8.3	1.99E-073	P; F	NA	L
Gene.798	Arginyl-tRNA synthetase	-8.28	2.03E-071	F; C; P	Aminoacyl-tRNA biosynthesis	J
Gene.1397	Virion core protein	-8.21	1.56E-070	NA	NA	S
Gene.851	Carboxypeptidase Taq metallopeptidase	-8.14	6.21E-055	F; P	NA	E
Gene.785	ATP synthase subunit A	-8.1	5.44E-068	C; F; P	NA	C
Gene.1174	General stress protein	-8.08	1.25E-069	C	NA	S
Gene.707	50S ribosomal L2	-8.08	1.17E-058	F; C; P	NA	J
Gene.2405	Integral membrane protein	-8.07	4.82E-075	C	NA	S
Gene.980	Penicillin-binding protein	-8.07	1.00E-068	C; P; F	NA	M
Gene.1111	dTDP-glucose 4,6-dehydratase	-8.03	2.40E-034	F; P	Polyketide sugar unit biosynthesis, Acarbose and validamycin biosynthesis, Biosynthesis of vancomycin group antibiotics, Streptomycin biosynthesis,	M

					Biosynthesis of antibiotics	
Gene.742	50S ribosomal L15	-7.94	1.61E-061	F; C; P	NA	J
Gene.1744	NADH-flavin reductase	-7.81	1.91E-089	F; P	Riboflavin metabolism	S
Gene.378	Cell division protein	-7.79	4.07E-035	P; C	NA	S
Gene.1168	Phosphate transport system regulatory	-7.76	1.02E-067	P; C	NA	P
Gene.1558	Tellurite resistance protein TelA	-7.76	8.19E-072	NA	NA	P
Gene.1863	Hypothetical protein	-7.66	1.60E-067	NA	NA	S
Gene.1502	Threonine--tRNA ligase	-7.6	1.38E-057	F; C; P	Aminoacyl-tRNA biosynthesis	J
Gene.799	Galactose-1-phosphate uridylyltransferase	-7.57	4.02E-066	C; P; F	Galactose metabolism, Amino sugar and nucleotide sugar metabolism	G
Gene.2329	Hypothetical protein	-7.51	2.25E-065	NA	NA	S
Gene.87	Ribosomal S17	-7.5	3.28E-082	C	NA	S
Gene.2259	Lipid kinase	-7.46	1.16E-034	F; P	Nicotinate and nicotinamide metabolism, Phosphatidylinositol signaling system, Glycerophospholipid metabolism, Glycerolipid metabolism	I
Gene.2726	Bifunctional AAC/APH aminoglycoside modifying enzyme	-7.37	4.08E-065	F; C; P	NA	S
Gene.650	DNA gyrase subunit A	-7.21	1.65E-066	F; C; P	NA	L

Gene.114	Acetate kinase	-7.16	2.27E-058	F; P; C	Taurine and hypotaurine metabolism, Propanoate metabolism, Carbon fixation pathways in prokaryotes, Pyruvate metabolism, Methane metabolism	C
Gene.2017	Hypothetical protein	-7.15	0	NA	NA	NA
Gene.86	Hypothetical Cytosolic	-7.13	3.75E-064	C	NA	S
Gene.1196	Amino acid ABC transporter ATP-binding	-7.1	6.41E-033	F; P	NA	E
Gene.1272	Hypothetical protein	-7.1	1.19E-048	NA	NA	NA
Gene.1022	Transcriptional regulator	-7.09	6.25E-072	NA	NA	K
Gene.1232	NAD dependent epimerase dehydratase family	-7.07	2.98E-066	P; F	NA	G, M
Gene.1505	Acetyl- biotin carboxylase subunit	-7.01	1.58E-033	F	Fatty acid biosynthesis, Tetracycline biosynthesis, Propanoate metabolism, Aflatoxin biosynthesis, Carbon fixation pathways in prokaryotes, Pyruvate metabolism, Biosynthesis of antibiotics	I

Gene.1877	Elongation factor Tu	-7.01	8.60E-043	F; C; P	NA	J
Gene.1691	Adenosylcobyric acid synthase	-6.94	1.90E-034	P; F	Porphyrin and chlorophyll metabolism	S
Gene.1351	Ribonucleotide-diphosphate reductase subunit beta	-6.89	2.13E-059	F; P; C	Purine metabolism, Glutathione metabolism, Pyrimidine metabolism	F
Gene.1977	Hypothetical protein	-6.87	2.15E-062	NA	NA	S
Gene.727	30S ribosomal S3	-6.87	1.01E-044	F; C; P	NA	J
Gene.1844	Holiday junction resolvase	-6.86	3.10E-041	F; P; C	NA	L
Gene.83	Membrane protein	-6.86	8.28E-064	C	NA	S
Gene.605	Glycine betaine carnitine choline ABC transporter	-6.81	4.68E-046	P; F	NA	M
Gene.885	Membrane protein	-6.79	9.71E-058	C	NA	NA
Gene.111	Cytochrome P450	-6.74	3.59E-067	F; P	NA	Q
Gene.1438	Ribosomal protein	-6.74	6.84E-066	C	NA	J
Gene.2616	UDP-glucose 4-epimerase	-6.67	9.86E-035	F	NA	M
Gene.2141	Galactose-6-phosphate isomerase subunit LacA	-6.59	6.82E-046	F; P:	Galactose metabolism	G
Gene.1235	Asparagine--tRNA ligase	-6.58	1.23E-021	F; C; P	Aminoacyl-tRNA biosynthesis	J
Gene.1037	Calcium-translocating P-type PMCA-type	-6.48	4.41E-064	F; C	NA	P
Gene.821	30S ribosomal S2	-6.46	3.65E-033	F; C; P	NA	J
Gene.838	Ribosome-binding factor A	-6.46	2.79E-043	C; P	NA	J
Gene.1208	GTP cyclohydrolase	-6.45	9.30E-062	F	NA	S
Gene.1565	Aldose 1-epimerase	-6.41	1.53E-020	P; F	Glycolysis / Gluconeogenesis, Galactose	G

					metabolism, Biosynthesis of antibiotics	
Gene.271	PTS Fru IIC component	-6.37	2.90E-021	P; C; F	NA	G
Gene.1996	Hypothetical protein	-6.33	1.33E-061	NA	NA	S
Gene.1403	Lipoyltransferase and lipoate- ligase	-6.32	5.40E-025	F; P	NA	H
Gene.672	Preprotein translocase subunit	-6.31	1.71E-031	P; C	NA	U
Gene.147	Hypothetical protein	-6.29	2.51E-045	NA	NA	S
Gene.608	Glycine betaine carnitine choline ABC transporter permease	-6.29	3.34E-028	C; P	NA	E
Gene.1652	Dihydroxyacetone phosphotransfer subunit	-6.25	7.63E-062	C; P; F	NA	G
Gene.751	30S ribosomal S8	-6.25	4.45E-030	F; C; P	NA	J
Gene.1980	3-dehydroquinate synthase	-6.19	2.79E-045	NA	NA	NA
Gene.1216	Glycyl-tRNA synthetase beta subunit	-6.13	4.65E-034	F; P; C	Aminoacyl-tRNA biosynthesis	J
Gene.1425	PTS mannose fructose sorbose IIB component	-6.08	9.09E-033	C; P; F	NA	G
Gene.1363	Secreated antigen SagA	-6.07	7.34E-024	NA	NA	M
Gene.934	Peptidase	-6.07	1.15E-070	NA	NA	NA
Gene.419	Dipeptidase	-6.05	1.47E-029	F; P:	NA	E
Gene.747	30S ribosomal S12	-6.03	1.50E-030	F; C; P	NA	J

[1]C, cellular component; P, biological process; F, molecular function; NA, not applicable

[2][D] Cell cycle control, cell division, chromosome partitioning; [M] Cell wall/membrane/envelope biogenesis; [N] Cell motility; [O] Post-translational modification, protein turnover & chaperones; [T] Signal transduction mechanisms; [U] Intracellular trafficking, secretion & vesicular transport; [V] Defense mechanisms; [J] Translation, ribosomal structure & biogenesis; [K] Transcription; [L] Replication, recombination & repair; [C] Energy production & conversion; [G] Carbohydrate transport & metabolism; [E] Amino acid transport & metabolism; [F] Nucleotide transport & metabolism; [H] Coenzyme transport & metabolism; [I] Lipid transport & metabolism; [Q] Secondary metabolites biosynthesis, transport & catabolism; [P] Inorganic ion transport & metabolism; [S] Function unknown