UiT The Arctic University of Norway

Faculty of Health Sciences

Department of Medical Biology

Exploring the genomes of the Norwegian vancomycin resistant enterococci

Mushtaq T. S. AL-Rubaye

A dissertation for the degree of *philosophiae Doctor*, December 2022

A dissertation for the degree of Philosophiae Doctor

Exploring the genomes of the Norwegian vancomycin resistant enterococci

Mushtaq AL-Rubaye

December 2022

Research Group of Host-Microbe Interactions

Department of Medical Biology

Faculty of Health Sciences

UiT – The Arctic University of Norway

Acknowledgement

This research project has been carried out in the research group of Host-Microbe Interaction (MHI), Faculty of Health Sciences, the Institute of Medical Biology, UiT, the Arctic University of Norway. I am grateful to all those with whom I have had the pleasure to work during this project.

First and foremost, I would like to extend my sincere gratitude to my supervisors, Kristin Hegstad, Jessin Janice James Peter, Arnfinn Sundsfjord, and Jørgen Vildershøj Bjørnholt. Thank you for the opportunity to be a part of this project and for your invaluable support from the first day that I joined this research group.

I am profoundly grateful to Kristin, my main supervisor, for her patience, motivation, and kindness, that were beyond my imagination. You were one of the most welcoming and supportive persons I have ever met.

I would like the extend my extreme thankfulness to my co-supervisors. Jessin for her endless support and kindness and for being one of the nicest people to me. I learned lots of bioinformatics from you, and Arnfinn and Jørgen, for their valuable feedback, contribution, and precise revisions.

I would like to express my deep appreciation to the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance (K-res) for excellent technical assistance and all our collaborator for their support and providing the bacterial isolates.My thanks go to all members of HMI, Labgang, my officemates and labmates. My friends, Theresa, Clement, Kenneth, Dorota and Wasifa, we shared the office for some time. Thank you for your kindness and for being so nice to me. Kjersti, Ahmed, and Bhupender, were the friendliest and most helpful in the lab. Bishnu, Adriana, and Ekaterina, you were and are great friends. A special thanks to my friend Jonathan who helped me recover my data when I lost some of my data.

My great thanks go to all who shared one of the happiest moments of my life (my son's birth) and the saddest ever moment (losing my father) that happened during my PhD.

Finally, I would like to express my special thanks to my wife (Maryam) for her support and care and to my little son (Ali), who constantly tolerated my absence. My prayers are to the soul of my father, the most remarkable, supportive, and faithful man in my life.

Tromsø, December 2022

Mushtaq

Table of Contents

1 Introduction 1
1.1 Description of genus *Enterococcus* 1
1.2 Enterococci; from a commensal to antibiotic resistant pathogens 2
1.3 The most important clinical species of enterococci 2
1.4 Population structure and phylogeny of *E. faecium* and *E. faecalis* 3
1.5 The genome of enterococci 5
1.5.1 Open pan-genome of *E. faecium* 5
1.5.2 Barriers of HGT in enterococci 5
1.6 Enterococcal infections; treatment and the relevant antimicrobial resistance 6
1.7 Acquired antibiotic resistance in enterococci 7
1.7.1 Ampicillin resistant enterococci (ARE) 7
1.7.2 High level gentamicin resistant enterococci (HLGR) 7
1.7.3 Linezolid resistant enterococci (LRE) 7
1.7.4 Vancomycin resistant enterococci (VRE) 8
1.7.5 The different *van*-types 10
vanA gene cluster 10
vanB gene cluster 11
vanC gene cluster 11
vanD gene cluster 12
vanE gene cluster 12
vanG gene cluster 12
vanL gene cluster 13
vanM gene cluster 13
vanN gene cluster 13
vanP gene cluster 13
1.8 Trends of outbreak associated *van*-types around the world 13
1.9 Hypotheses on the origin of *van*-types 14
1.10 HGT mechanisms in enterococci 15
1.11 The role of different MGEs in HGT in enterococci 17
1.11.1 Phages 17
1.11.2 Plasmids 18
The incompatibility 18 group (Inc18) plasmids 18
RepA_N plasmids 19

Rep_3 plasmids 19
Rolling-circle-replicating (RCR) plasmids 19
1.11.3 Genomic islands (GI) in enterococci 20
ICE elements in enterococci 20
1.11.4 Transposons in enterococci 21
Tn*3* family 21
Insertion sequence (IS) elements 22
Composite transposons 22
1.12 VRE; the successful resident of hospitals 22
1.13 Enterococcal virulence factors 22
1.14 Common methods and techniques for studying enterococci 23
1.15 Next generation sequencing 24
1.15.1 Short read sequencing 24
1.15.2 Long read sequencing 25
1.16 Bioinformatic analyses of NGS raw data 26
1.16.1 MLST 27
1.16.2 cgMLST 27
1.16.3 SNP-calling 28
2 Objectives of the study 30
3 Materials and methods 31
3.1 Bacterial culture collections 31
3.2 Bacterial species identification 32
3.3 Antimicrobial susceptibility testing (AST) 32
3.4 Filter mating 32
3.5 DNA extraction 33
3.6 Sequencing (Illumina and PacBio) 33
3.7 From raw reads to assemblies 33
3.8 Typing of isolates 36
3.9 Phylogeny 36
3.10 AMR and VF genes identification 36
3.11 Identification of MGEs harbouring *van* gene clusters and their genomic integration site 37
3.12 Reconstructing plasmids of *vanA*-type VRE isolates 37
4 Summary of results 39
5 General discussion 41
5.1 Setting standards for quality control of assemblies 42

5.2 Determining clonality of isolates 43
5.3 The prevalence of globally dominant STs and CTs among the Norwegian VRE*fm* 45
5.4 Relatedness between the Norwegian and global VRE 46
5.5 *van* gene clusters associated with the Norwegian VRE 47
5.6 The MGEs harbouring *van* gene cluster in the Norwegian VRE 48
5.7 Each cluster has a specific virulome; VRE are more virulent. 50
5.8 The VRE*fs* isolates. 52
6 Concluding remarks and future aspects 53
References 54

List of Tables

Table 1. The number of isolates that were whole genome sequenced in this research 31

List of Figures

Figure 1. Timeline of relevant events in the history of enterococci 1
Figure 2. Population structure of *E. faecium*. 4
Figure 3. The structure and composition of *van* gene clusters (vanA-P) 9
Figure 4. Vancomycin mechanism of action on the Gram-positive bacteria 10
Figure 5. Overview of the main mechanisms of mobile genetic element uptake in HGT 17
Figure 6. Schematic presentation of Tn1546 and Tn1549 21
Figure 7. From Illumina raw reads to assembly. 35

Abbreviations

ARE	Ampicillin Resistant enterococci
ARG	Antimicrobial Resistance Gene
AST	Antimicrobial Susceptibility Testing
BAPS	Bayesian Analysis of Population Structure
BMD	Broth Microdilution
CC	Clonal Complex
CCS	Circular Consensus Sequencing (reads)
cgMLST	Core Genome Multilocus Sequence Typing
CT	Cluster Type
CTn	Conjugative Transposon
GI	Genomic Island
HGT	Horizontal Gene Transfer
HLGR	High Level Gentamicin Resistant enterococci
ICE	Integrative Conjugative Element
IS	Insertion Sequence
LRE	Linezolid Resistant enterococci
MLST	Multilocus Sequence Typing
NGS	Next Generation Sequencing
R-M	Restriction-Modification (System)
RCR	Rolling-Circle-Replicating plasmid
ST	Sequence Type
VF	Virulence Factor
VRE	Vancomycin Resistant enterococci
VRE*fm*	Vancomycin Resistant *E. faecium*
VSE	Vancomycin Susceptible enterococci
WGS	Whole Genome Sequencing

List of papers

Paper I

Al Rubaye MTS, Janice J, Bjørnholt JV, Jakovljev A, Hultström ME, Sundsfjord A, Hegstad K. **Novel genomic islands and a new *vanD*-subtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway.** PLoS One. 2021 Jul 23;16(7):e0255187. doi: 10.1371/journal.pone.0255187. PMID: 34297779; PMCID: PMC8301612.

Paper II

Mushtaq AL Rubaye, Jessin Janice, Jørgen Vildershøj Bjørnholt, Iren H. Löhr, Arnfinn Sundsfjord, Kristin Hegstad

The first *vanE*-type vancomycin resistant *Enterococcus faecalis* isolates in Norway – phenotypic and molecular characteristics.

(Manuscript)

Paper III

Mushtaq AL Rubaye, Jessin Janice, Jørgen Vildershøj Bjørnholt, Oliver Kacelnik, Bjørg C. Haldorsen, Randi M. Nygaard, Joachim Hegstad, Arnfinn Sundsfjord, Kristin Hegstad and the Norwegian VRE study group.

The population structure of vancomycin resistant and susceptible *Enterococcus faecium* in a low prevalence antimicrobial resistance setting is highly influenced by global clones.

(Manuscript)

Summary

Enterococci are gram-positive commensals in the human gut microbiota that may cause severe infections, especially in immunocompromised and hospitalized patients. Their ability to accumulate antimicrobial resistance and virulence genes have been vital in their evolution into leading pathogens. Vancomycin resistant enterococci (VRE) are among the most important resistant pathogens causing outbreaks worldwide. In this study, we aimed to examine recent developments in the epidemiology of VRE in Norway related to the occurrence of novel *van*-types and the increasing incidence of VRE from 2010.

In **paper I**, we identified and described the first Norwegian *vanD*-type VRE isolates from two patients and reported a novel *vanD*-subtype, three novel genomic islands harbouring the *vanD* gene clusters, and the first *vanD*-type vancomycin resistant *Enterococcus casseliflavus* strain. Different subtypes of *vanD* on different genomic islands and phylogenetic distance confirmed that the isolates of the two cases are not clonally related.

In **paper II**, we identified the first *vanE*-type VRE isolates in Norway recovered from the same patient 2,5 years apart and described the mobile genetic element harbouring the *vanE* gene cluster, its insertion site, and variations in the $vanS_E$ gene which explained why one isolate expressed inducible low level and the other isolate constitutive low level vancomycin resistance.

In **paper III,** we conducted the first comprehensive study on Norwegian VRE and vancomycin susceptible *Enterococcus faecium* at the genomic level and identified the most prevalent cluster types, compared their virulomes, and mobile genetic elements harbouring the *van* gene clusters. Our result showed that the globally prevalent clones and particularly concurrent European cluster types (CTs) influence the population structure of *E. faecium* in a low prevalence antimicrobial resistance setting like Norway*,* with similar dynamic sequence type sweeps*.* The prevalent VRE *faecium* CTs have acquired more virulence determinants than the more diverse local VSE *faecium* population.

1 Introduction

1.1 Description of genus *Enterococcus*

Enterococci are Gram positive spherical or ovoid bacteria that can occur as a single cell, chains or most often in pairs (diplococci) (1). Enterococci are facultative anaerobes and chemoorganotrophs with homofermentative metabolism that are non-spore-forming (1,2). Most enterococcal species tolerate harsh environments including up to 6.5% NaCl and are resistant to 40% bile. They can grow in temperatures ranging between 5 to 50 °C and even survive at 60 °C for up to 30 minutes. Additionally, they can survive in the pH ranges of 4.8 to 9.6 (3). As ubiquitous bacteria, they have been isolated from various environmental sources, plants, fermented food products, and as part of the gut microbiota of humans and animals (4,5).

Thiercelin first reported Entérocoque in 1899, but it took almost a century to recognize it as a separate genus. At first, based on the morphological and biochemical similarities, they were classified as group D *Streptococcus* (Figure 1). This classification was valid until the 1980s (2). Although enterococci are serologically related to group D *Streptococcus*, they are phylogenetically more distant. This was the reason behind the reclassification and recognition of *Enterococcus* as a separate genus in 1986 (6).

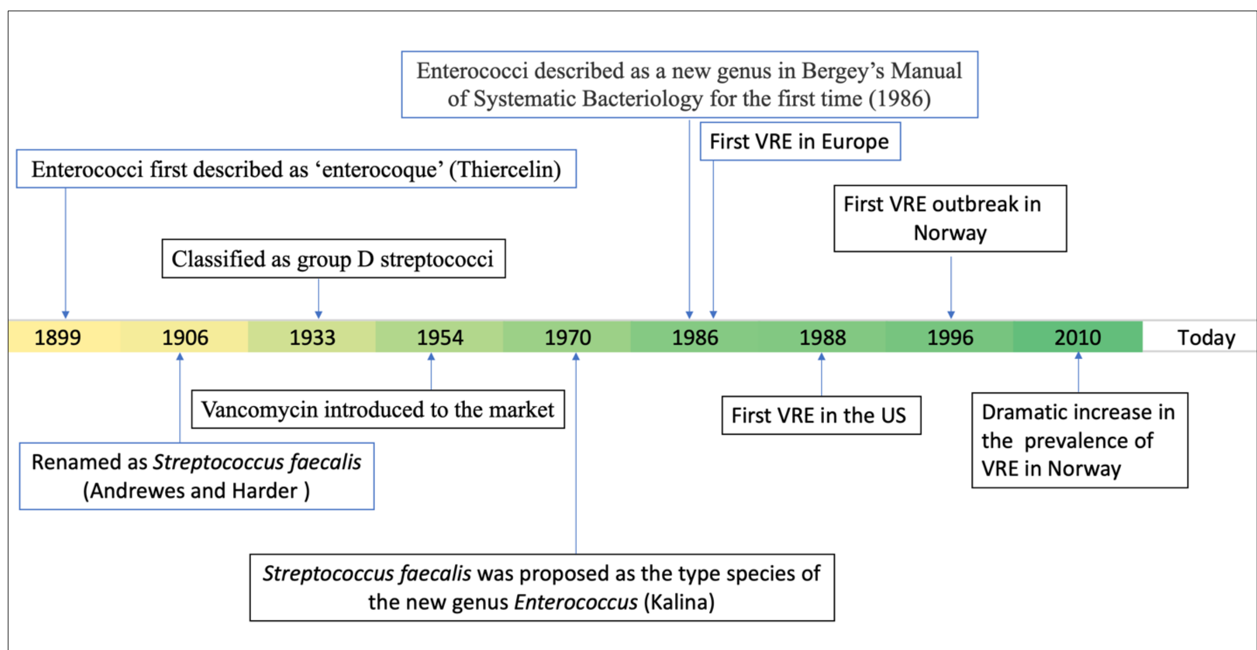


Figure 1. Timeline of relevant events in the history of enterococci and appearance of vancomycin resistant enterococci; Based on García-Solache and Rice (2).

Various enterococcal species can be isolated from different sources such as seawater (*E. aquimarinus*), surface water (*E. moraviensis*), plants (*E. plantarum*), birds (*E. alcedinis*) and humans (*E. faecium*, *E. faecalis*, *E. gallinarum*, and *E. casseliflavus*), as well as dairy products

(*E. italicus*) (2,7–11). Among 62 species taxonomically verified as enterococci in the List of Prokaryotic names with Standing in Nomenclature (LPSN) (as of 29.11.2022) (12), *E. faecium* and *E. faecalis* are the most important clinical species (13).

1.2 Enterococci; from a commensal to antibiotic resistant pathogens

As a part of the normal flora, enterococci are commensal bacteria in the human and animal gastrointestinal tract (14,15). Furthermore, they are capable of causing a range of serious infections, mainly in hospitalized patients with co-morbidities who are receiving antibiotic therapy or suffer from disturbed intestinal microbiota (14). Infections caused by enterococci include urinary tract infections, endocarditis, bacteraemia, and intra-abdominal infections (14,16,17).

The ability of enterococci to survive adverse conditions converted them into a well-adapted microorganisms within the healthcare environment (2,18). Such capability in enterococci, specifically in *E. faecium*, facilitates their spread and persistence in healthcare institutions and makes their control very difficult (19). Additionally, their intrinsic resistance to several antibiotics and remarkable capability to acquire transferable resistance make treatment more challenging. They also show an increase in virulence factors which improve their ability to colonize and infect. Moreover, their ability to acquire novel determinants for virulence and antibiotic resistance makes infection control difficult (2,20).

Enterococci are intrinsically resistant to cephalosporins, lincosamides, and streptogramins and have low-level resistance to aminoglycosides. In addition, enterococci can acquire resistance against vancomycin, teicoplanin, penicillins, linezolid, streptogramins, daptomycin, and high level of aminoglycosides (21). Vancomycin is a valid option in the treatment of invasive infections with multidrug resistant enterococci. The spread of VRE is concerning, as only a few second-line antibiotics are available to treat infections caused by them (22). Worryingly, there has been a 5% increase in vancomycin resistance in *E. faecium* (VRE*fm*) in the EU/EEA during 2016–2020 (23). Moreover, the mortality rate of bacteraemia caused by VRE can be increased by 2.5-fold (20), and Hospital costs associated with infections caused by VRE are significantly higher than those related to vancomycin-susceptible enterococci (VSE) (24).

1.3 The most important clinical species of enterococci

Among enterococci, *E. faecalis* was previously the leading cause of nosocomial infections. For example, over the last decade in the United Kingdom *E. faecalis* was the major cause of

enterococcal infection. But recently, this has changed in favour of *E. faecium,* which accounted for 51% in enterococcal bacteraemia in some regions in the UK between 2017 and 2019 (25). *E. faecalis* is the most virulent *Enterococcus* but less prone to acquire resistance to antibiotics compared to *E. faecium*. *E. faecalis* and *E. faecium* together account for about 75% of clinical enterococcal infections in the USA between 2011 to 2014 (2,26). *E. casseliflavus*, *E. gallinarum*, *E. durans*, *E. hirae*, *E. mundtii*, *E. avium*, and *E. raffinosus* are other enterococci associated with human infections, more specifically, in patients with concurrent haematological malignancies, neutropenia, and previous corticosteroid therapy (2,27).

1.4 Population structure and phylogeny of *E. faecium* and *E. faecalis*

Phylogenetically, different subpopulations exist within *E. faecium* species and are referred to as "clades". A deep phylogenetic gap divides the two main subpopulations of *E. faecium* (clade A and B) (Figure 2) (28,29). Although sub-clades in clade A remain disputed (30), clade A has so far been further divided into sub-clades A1 and A2. Subclade A1 comprises clinical strains, while subclade A2 strains are mostly recovered from livestock and domestic animals, as well as non-hospitalized persons, and clade B mainly contains human commensal isolates (29–31). Clade B strains were recently reclassified and suggested to belong to a different enterococcal species, *Enterococcus lactis* (32). Another way of analysing *E. faecium* population structure is Bayesian-based population genetic modelling. Bayesian Analysis of Population Structure (BAPS) software categorizes *E. faecium* isolates in 13 BAPS or sub-groups. The vast majority of isolates of nosocomial origin are clustered in two main sub-groups (groups 2-1 and 3-3) (33). eBURST is an older method to divide multilocus sequence typing (MLST) data into subpopulation groups and clonal complexes (34). eBURST and phylogenetic analysis of *E. faecium* BAPS (groups 2 and 3) revealed three distinct hospital lineages or clonal complexes (CC) (17,18 and 78), indicating different evolutionary paths for BAPS 2-1 (lineage 78) and 3-3 (lineage 17 and lineage 18) isolates (33). Clade A1 predominantly comprises clinical isolates and overlaps *E. faecium* sub-population clonal complex 17 (CC17) (29). All the methods mentioned above have confirmed the existence of *E. faecium* subpopulations.

Despite a small genome size and a stable large core genome in *E. faecalis*, genomic analyses reveal a population cohesively connected through homologous recombination. There is evidence that hospital-associated *E faecalis* lineages predate the “modern hospital” era, showing selection in an older niche and indicating the generalist nature of this nosocomial pathogen (35).

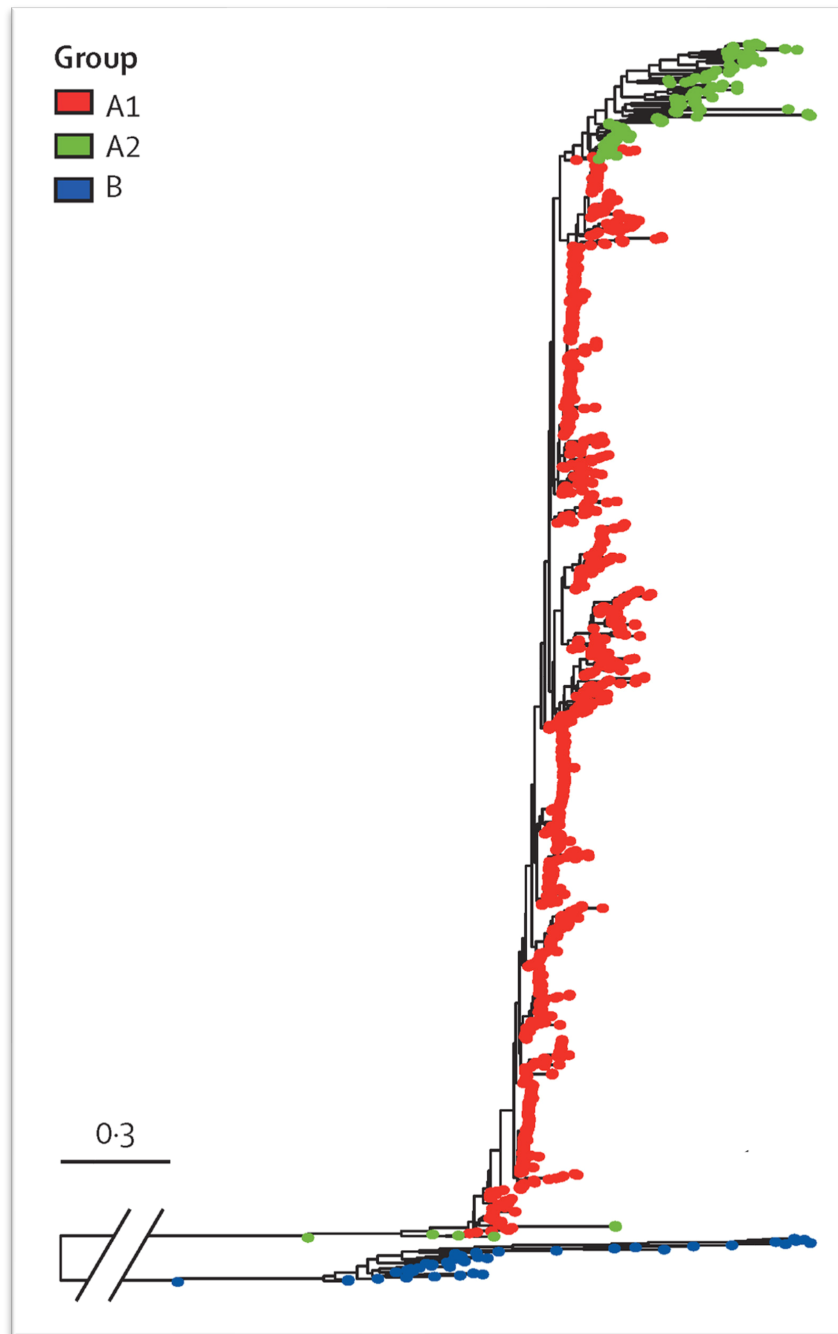


Figure 2. Population structure of *E. faecium*. Maximum-likelihood phylogenies of 1128 *E. faecium* genomes, before masking of recombination events identified using ClonalFrameML, with tips coloured by group assigned using Bayesian Analysis of Population Structure. Permission by Elsevier: slightly modified from van Hal *et al*. (30).

1.5 The genome of enterococci

Enterococcal species are known to contain plastic genomes of low GC content (34-45%) (2,36). Their genome size varies from 2.3 to 5.3 Mbp, with the predicted gene number ranging from 2154 to 5107. This variation in the genome size could have resulted from various levels of horizontal gene transfer (HGT), including gene insertions and deletions during *Enterococcus* evolution (37). The core genome of the *Enterococcus* genus contains between 605 to 1,037 genes (2).

Phylogenetic studies revealed that the environment significantly impacted the evolution of *Enterococcus*, and strains isolated from similar environments are genetically more related. Humans and other mammals are suggested to be the original host of *Enterococcus*, additionally host-shifting happened from mammals and humans to birds, plants and different environments (37).

1.5.1 Open pan-genome of *E. faecium*

E. faecium possesses a dynamic open pan-genome (28). The pan-genome includes a species' entire set of genes, including both core and accessory genome genes. Core genomes include the set of genes shared by all strains within a species, while the accessory part of the genomes is a set of additional genes present in a subset of the strains and is the result of HGT (38). In species like *E. faecium* that have an open pan-genome, a large number of genomes are needed to determine the accessory genome (28,38). The genomic events contributing to evolution in *E. faecium* are HGT and recombination rather than mutation (39). Genome plasticity can explain the variable genome size in *E. faecium*, which varies from 2.43 Mb to 3.44 Mbp (40). Mobile genetic elements (MGEs) such as plasmids, integrative conjugative elements (ICEs), transposable elements (Tns), and temperate bacteriophages are common agents of HGT in enterococci (41).

1.5.2 Barriers of HGT in enterococci

CRISPR-*cas* system in *E. faecium*

Bacteria have developed barriers to protect themselves from foreign DNA and prevent HGT. The sequence-based mechanism of clustered regularly interspaced short palindromic repeats (CRISPR) called CRISPR-*cas* system is one such defence barrier that use RNA-guided nuclease to prevent the acquisition of MGEs (42). Previously it was believed that multidrug

resistant enterococci do not possess CRISPR-*cas* system (43), but recent research revealed that it exists in the enterococcal species, most frequently in commensal isolates. In general, commensal *E. faecium* isolates contain a functional CRISPR-*cas* system, while multidrug-resistant isolates do not (41). Previously, it was believed that the lack of CRISPR-*cas* system in clade A1 isolates could explain the accumulation of plasmids. But it is unlikely to contribute to a different and higher number of plasmid content in the isolates of clade A1 recovered from hospitalized patients [50].

Restriction-modification system (R-M)

In *E. faecium*, defence system called restriction modification (R-M) systems which act as barriers for HGT through specific methylation of DNA and cleavage of DNA that does not have this specific methylation pattern have been found (44). Certain R-M systems have been hypothesised to aid formation of subspecies of *E. faecium* (clades A and B) by reducing the transformability in clade A1 isolates (45). Among the three subunits of the R-M system, S mediates specificity, M modification and R restriction. Specific variants of the S subunit are enriched in clade A1 while M and R subunits look similar in both clade A and non-clade A isolates. R-M systems thus were believed to lead to differences in plasmid content in enterococci and contribute to source specificity (46,47).

1.6 Enterococcal infections; treatment and the relevant antimicrobial resistance

As an agent of HGT, MGEs such as plasmids, transferable elements, and temperate bacteriophages facilitates the conversion of enterococcal species like *E. faecalis* and *E. faecium* into antibiotic-resistant pathogens (48,49). Among enterococci, commensal strains have a limited ability for acquiring MGEs (41). In the clinically important enterococci (*E. faecium* and *E. faecalis*), high-level ampicillin and aminoglycoside resistance, as well as vancomycin resistance are of particular importance (2). Ampicillin, gentamicin, vancomycin, and linezolid are key antibiotics in the treatment of enterococcal infections. Ampicillin alone or combined with an aminoglycoside (gentamicin or streptomycin) is the common choice of treatment for *E. faecium* infections. In cases of co-resistance to beta-lactams (penicillin, ampicillin) or aminoglycosides (gentamicin), vancomycin or linezolid will be used in place of these antibiotics (50–52). Moreover, linezolid and daptomycin-resistant isolates of *E. faecium* and *E. faecalis* have been recovered so far (41). Currently, linezolid and daptomycin are known as last-line antibiotics for treating VRE infections. Consequently, it is of great clinical concern if antimicrobial resistance develops to these last-line agents (53).

1.7 Acquired antibiotic resistance in enterococci

1.7.1 Ampicillin resistant enterococci (ARE)

Over a decade before the emergence of VRE, ampicillin resistant enterococci were first isolated in the US (19). Today, in some clinics, ampicillin resistant *E. faecium* levels exceed 70%. However, molecular analyses on the early American ARE revealed most ARE isolates belong to a few lineages, mainly to (CC17/clade A1), indicating that the acquisition of ampicillin resistance in the isolates has happened independently (54). In enterococci, ampicillin resistance is mediated either by an acquired β-lactamase gene or mutations in the intrinsic penicillin-binding protein (PBP) genes (55,56). β-lactamases are extremely rare in enterococci. In *E. faecium*, mutation in PBP5 result in a lower affinity to ampicillin while mutation in PBP4 in *E. faecalis* is the main cause of acquired ampicillin resistance (55,56). Mutations can increase PBP5 expression or lower the affinity of the protein to β-lactam antibiotics that consequently cause ampicillin resistance in enterococci (57). The transferability of PBP5 is reported in a limited number of *E. faecium* isolates as part of large chromosomal regions (58).

1.7.2 High level gentamicin resistant enterococci (HLGR)

Gentamicin is an aminoglycoside antibiotic that binds to 23S rRNA and blocks protein synthesis. Genes encoding aminoglycosides modifying enzymes (AME) such as *aac(6')-Ie-aph(2")-Ia*, *aph(2")-Ib*, *aph(2")-Ic*, *aph(2")-Id*, and *aph(3')-IIIa* are responsible for resistance to aminoglycosides in enterococci (59). AMEs modify aminoglycosides at the -OH or NH2 group of the sugar moieties or 2-deoxystreptamine nucleus. They can be classified in three types: acetyltransferases (AACs), nucleotidyltranferases (ANTs), or phosphotransferases (APHs) (60). The dominant gentamicin resistance gene in enterococci is *aac-6'-Ie-aph-2"* that typically is carried on the composite transposon Tn*4001* (2).

1.7.3 Linezolid resistant enterococci (LRE)

Linezolid is a member of the synthetic oxazolidinone drug family that was developed to combat Gram-positive bacteria resistant to multiple antibiotics. It blocks protein synthesis via binding to the translational initiation complex. Two types of acquired linezolid resistance are known (21,61). The first is caused by mutations in the ribosome causing changes in the linezolid binding site which includes mutations in 23S rRNA (G2576T or G2505A) or in the L3 and/or L4 ribosomal proteins (62). The second type is transferable linezolid resistance which involves acquisition of variants of either *optrA*, *poxtA* or *cfr* genes. Some studies challenged the role of

cfr genes in conferring linezolid resistance in enterococci showing that although the gene is expressed, it does not give phenotypic resistance in enterococci (63,64). OptrA and PoxtA belong to F-lineage of ATP-binding cassette (ABC) proteins (ABC-F proteins) that are associated with resistance to phenicols and oxazolidinones (53,65). The first LRE was discovered in 2004 in Greece. The worldwide prevalence of linezolid resistance is less than 1% in enterococci (66), but the reports of acquired LRE cases are increasing. Acquired LRE is very concerning when it co-exists with vancomycin resistance. A recent study in Ireland revealed that 26% (5 out of 19) of *poxtA*-harbouring plasmids were carrying vancomycin resistant gene cluster (*vanA*) (62).

1.7.4 Vancomycin resistant enterococci (VRE)

Vancomycin is a glycopeptide class antibiotic with a tricyclic structure. Vancomycin blocks the process of cell wall formation by inhibiting peptidoglycan synthesis. It was first isolated in 1953 from *Streptococcus orientalis*. Five years later, in 1958, it was approved for clinical use by FDA (67,68). Most Gram-positive cocci and bacilli are susceptible to vancomycin. The main medical use of vancomycin is the treatment of infections caused by amoxicillin-resistant enterococci, methicillin-resistant *Staphylococcus aureus* (MRSA), and methicillin-resistant *Staphylococcus epidermidis* (MRSE) (69). Historically, the first VRE isolate was recovered in France in 1986, followed by the UK and the US. The prevalence of VRE in Europe remained low until 2000 (70). In contrast, the hospitalization of VRE-infected patients in the US increased dramatically in the 1990s (71). Due to differences in vancomycin usage, there are significant differences between Europe and the US regarding VRE epidemiology. Compared to five different European countries (the UK, the Netherlands, France, Italy, and Germany), the use of vancomycin in the USA was five- to ten-fold higher. In European countries, a large community reservoir of VRE among livestock and healthy humans exists, which was argued to be linked to the massive use of avoparcin (a vancomycin analog used as a growth promoter) in agriculture (72). Human-associated VRE is more likely to be caused by the use of vancomycin in hospitals (73).

Vancomycin resistance in enterococci is caused by a gene cluster called "*van*". To date, ten *van*-types have been described (*vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM, vanN,* and *vanP*) (Figure 3) (74), four of them have two or more sub-types (*vanB1-3*, *vanC1-4*, *vanG1-2*, *vanD1-5*) (75–77).

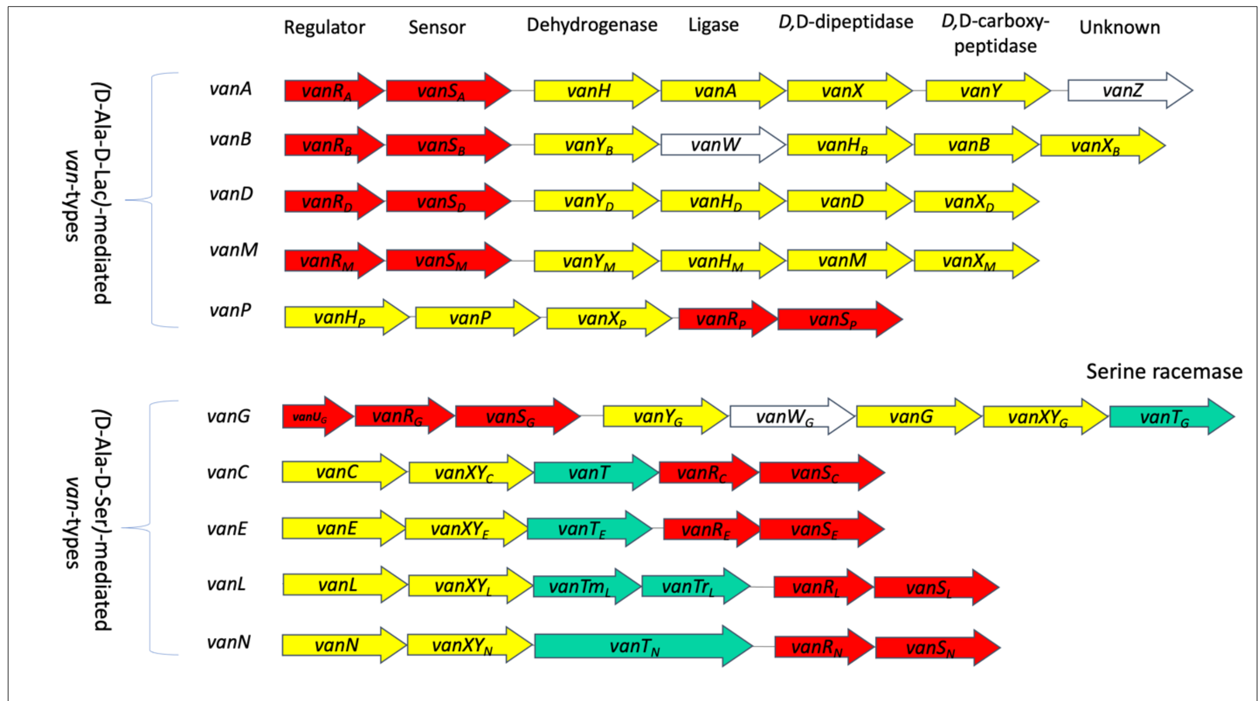


Figure 3. The structure and composition of *van* gene clusters (*vanA-P*). The cluster here are grouped in D-Ala-D-Lac-mediated (*vanA, B, D, M, and P*), and D-Ala-D-Ser-mediated (*vanG, C, E, L, and N*) resistance; based on Werner *et al* (78).

The mechanism of resistance is based on the replacement of the D-Ala-D-Ala to D-Ala-D-Lac (*vanA*, *B*, *D*, *M*, and *P*) or D-Ala-D-Ser (*vanC*, *G*, *E*, *L*, and *M*) (74). Subsequently, this reduces the affinity of vancomycin 1000-fold when changed to D-Ala-D-Lac and 7-fold in the case of D-Ala-D-Ser (Figure 4). During this process, a key hydrogen bond interactor (from cell wall peptidoglycan) with vancomycin is removed (79). The *van* gene clusters encode three products; an enzyme that removes the original D-Ala-D-Ala, enzymes that build the new peptidoglycan precursors, and two-component signal transduction systems (TCSs) for induction of resistance (80). Additionally, *van* gene clusters may include additional genes mentioned below (81).

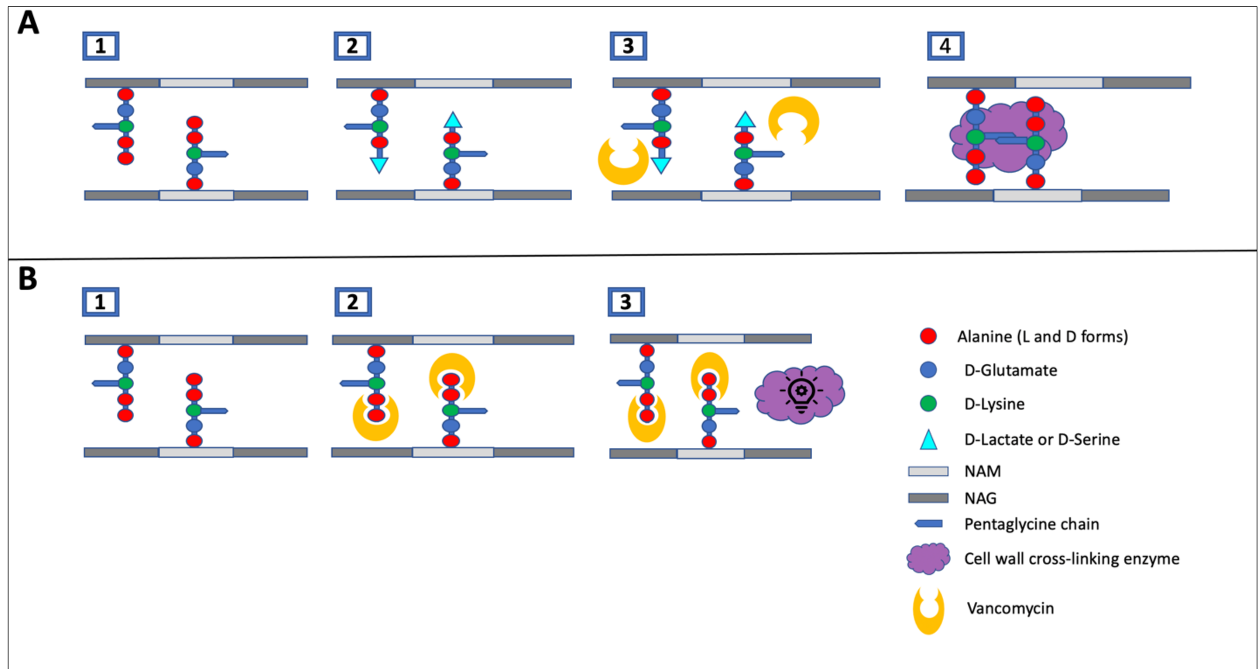


Figure 4. Vancomycin mechanism of action on the Gram-positive bacteria cell wall synthesis, based on McStrother on Wikipedia.org (2011) (82). In this figure N-acetylglucosamine (NAG) and N-acetylmuramic acid (NAM) are shown with their abbreviations.

1.7.5 The different *van*-types

vanA gene cluster

The *vanA* gene cluster was the first *van*-type identified. It confers high-level resistance to vancomycin and teicoplanin. The *vanA* and *vanB* are by far the most prevalent *van*-types worldwide (2,21). The *vanA* gene cluster is often encoded and mobilized on a Tn*3*-family transposon (Tn*1546*) which is often located on plasmids (83). *vanA* is commonly reported in *E. faecium* and E. *faecalis* (20), and sporadically identified in *E. casseliflavus* (84), *E. gallinarum* (85,86), *E. durans* (81), *E. mundii* (87), *E. hirae* (88), *E. raffinosus,* and *E. avium* (83). The cluster is composed of seven genes on two separate operons, *vanRS* and *vanHAXYZ* (Figure 3). The process of changing the dipeptide D-Ala-D-Ala starts with the *vanH* gene. It encodes a dehydrogenase that converts the cellular pyruvate to D-lactate. Next, the VanA ligase ligates D-Ala to D-Lac. Finally, the host enzymes use D-Ala-D-Lac to build the vancomycin low-affinity pentapeptide precursor (21). In addition, VanX hydrolyses the wildtype dipeptide D-Ala-D-Ala, making dipeptide D-Ala-D-Lac the only substrate for precursor synthesis. Similarly, VanY hydrolyses the D-Ala terminal residue from any normal pentapeptide precursor, making it useless for regular cell wall construction (21,89). *vanZ* plays a role in

teicoplanin resistance but its exact function is not known (90). VanR (response regulator) and VanS (sensor) form a two-component regulator system (TCS) which enable inducible expression of the *vanHAXYZ* operon (91).

***vanB* gene cluster**

The *vanB* gene cluster is responsible for different levels of inducible resistance to vancomycin. Unlike *vanA*, it normally does not produce resistance against teicoplanin (83). In enterococci, the *vanB* gene cluster can be found mainly on chromosomal elements or less frequently on plasmids (81,92). The *vanB* gene cluster is commonly found in *E. faecium* and *E. faecalis* and sporadically reported in *E. gallinarum* (85,93) and *E. hirae* (94). Moreover, the *vanB* gene cluster is common in anaerobic gut flora such as *Clostridium* spp., *Eggerthella lenta*, and *Ruminococcu*s (95). The gene organization and functionality of *vanB* are similar to *vanA* except it has an additional gene, *vanW*, with an unknown function, and lacks *vanZ*(83) (Figure 3). Based on the sequence differences, *vanB* has three sub-types (*vanB1*, *vanB2*, and *vanB3*). The sub-types in the *vanB* gene cluster have no significant differences in the resistance level (77,96,97). The most prevalent subtype cluster, *vanB2*, is carried on integrative conjugative elements (ICEs) like Tn*1549*/Tn*5382* or their variants (77,98).

***vanC* gene cluster**

The *vanC* gene cluster is characterized by low levels of vancomycin resistance (4 to 32 mg/liter) and susceptibility to teicoplanin (83,99). *vanC* has four known subtypes; *vanC1* in *E. gallinarum* and *vanC2-C4* in *E. casseliflavus*(83,100). *vanC* gene cluster is the only *van* type that can be found either acquired (in *E. faecalis* and *E. faecium*) (101,102) or intrinsic in some other enterococcal species (*E. gallinarum* and *E. casseliflavus*) (103). The gene organization in *vanC* is different from *vanA* and *B* gene clusters (Figure 3). First, the TCS genes are located downstream from *vanT*, while in *vanA* and *vanB* gene clusters, they are located upstream from the resistance genes (83,104). The *vanT* gene encodes a membrane-bound serin-racemase that converts L-serine to its enantiomer form, D-serine (105). Also, it encodes a ligase that catalyses D-ala-D-ser synthesis and VanXY that have both D,D-carboxypeptidase and D,D-dipeptidase activity and thus hydrolyse precursors ending in D-Ala (83).

vanD gene cluster

The *vanD* gene cluster is involved in moderate to high-level vancomycin resistance and varying degrees of susceptibility to teicoplanin. This cluster has a similar gene organization to *vanA* and *vanB* (Figure 3). *vanD* gene cluster contains six genes lacking *vanZ* (compared to *vanA*) and *vanW* (compared to *vanB*) (78). It has been identified in *E. faecium*, *E. faecalis*, *E. gallinarum*, *E. avium*, and *E. raffinosus* (106). It has also been reported from Gram positive anaerobic gut flora *Ruminococcus sp.* (107). All *vanD* gene clusters reported so far are non-transferable sporadic cases which are located on a chromosomal genomic island (108). With five known sub-types, *vanD* is the most diverse *van* gene cluster (109).

vanE gene cluster

vanE is one of the rarest *van*-types and has been reported only in North America and Australia in *E. faecalis* to date (81,110). The *vanE* gene cluster has an identical gene organization to *vanC* (Figure 3). It consists of five genes and causes inducible low-level vancomycin resistance and susceptibility to teicoplanin (110). Genes in the *vanE* gene cluster overlap each other. The start codon of $vanT_E$ and $vanXY_E$ overlap the stop codons of $vanXY_E$ and *vanE*, respectively. In the TCS genes of *vanE* gene cluster also, $vanS_E$ start codon overlaps the stop codon of $vanR_E$ (110).

vanG gene cluster

The *vanG* gene cluster confers low level inducible resistance to vancomycin (MIC 8-16 mg/L) but susceptibility to teicoplanin (83). Within enterococci, it has only been detected in *E. faecalis* and *E. faecium*(75), but also streptococci and some anaerobic gut flora such as *Clostridium sp* and *Ruminococcus* (107,111,112). The *vanG* gene cluster differs by its three-component regulatory system from other known *van* gene clusters, which have a TCS. The *vanS* and *vanR* genes of the *vanG* gene cluster are similar to those of the *vanD* gene cluster and the additional gene ($vanU_G$) codes for a transcriptional activator (75) (Figure 3). The resistance encoding region in the *vanG* gene cluster consists of a putative D,D-peptidases (*vanY*), the ligase (*vanG1*), a racemase (*vanT*), and a protein with unknown function (*vanW*) (113).The *vanG* gene cluster, with a few reports from Australia and Canada has two sub-types (*vanG1* and *vanG2*) (114).

***vanL* gene cluster**

The *vanL* gene cluster is characterized by a low-level resistance to vancomycin (MIC 8 µg/ml). Gene organization in *vanL* is similar to the *vanC* gene cluster except for the serin racemase gene (Figure 3). In the *vanC* gene cluster *vanT* gene code for serin racemase while in *vanL* it is encoded by two genes *vanTm*$_L$ (membrane binding) and *vanTr*$_L$ (racemase) (115). *vanL* gene cluster is located on the chromosome and causes inducible resistance. It has been identified only in *E. faecalis* and is among the rarest *van* gene clusters in enterococci (115).

***vanM* gene cluster**

The acquired *vanM* gene cluster is associated with high-level resistance to vancomycin and teicoplanin. It has been reported only in *E. faecium*. Its gene organization is similar to *vanD* (Figure 3), but in terms of sequence identity, it is more similar to *vanA* (116,117). It has been shown that *vanM* is a plasmid located gene cluster (118).

***vanN* gene cluster**

vanN is an acquired *van* gene cluster with a similar gene organization to *vanC* (Figure 3). It is responsible for low level resistance to vancomycin (MIC 16 µg/ml) but susceptibility to teicoplanin. So far, only a handful of *vanN* type VRE strains have been identified, all in *E. faecium* (119,120). It has been confirmed that *vanN* is located on plasmids (120). *vanN* is the only *van*-type that has been associated with clade B isolates VRE*fm* (121).

***vanP* gene cluster**

vanP is the newest *van* gene cluster recovered from a single *E. faecium* isolate in Belgium in 2021. It is responsible for low level vancomycin resistance (MIC 4 µg/mL) that can be increased up to 256 µg/mL when exposed to vancomycin or teicoplanin. It is located on a novel putative ICE and has been suggested to be acquired from anaerobe gut flora such as *Clostridium scidens* and *Roseburia* sp. (74).

1.8 Trends of outbreak associated *van*-types around the world

According to the World Health Organization (WHO), an outbreak is an abnormally higher disease incidence in a particular place or season (122). However, with a simple increase in the number of patient cases, the same risk factor can cause co-occurrence of the same illness and

the number of patient cases is not necessarily higher than expected (123). In the case of AMR infections, more data, including molecular and genetic typing, is needed to confirm an outbreak (124). Only two years after discovering the first VRE isolates, the first VRE outbreak was reported in 1988 (2). Because of their ability to survive in healthcare settings and low virulence, VRE can spread widely before being identified by routine microbiological methods (125). Most human VRE outbreaks are caused by *vanA* and *vanB* gene clusters (81). In the past few decades, *vanA* was the predominant *van*-type in Europe, the Americas and Asia accounting for 89% off all VRE isolates (126,127). Then around 2005, some European countries (Germany, France, Greece, and Spain) reported an increase in the numbers of *vanB*-type VRE. The *vanB*-type VRE outbreaks repeatedly happened in several European hospitals. In the Netherlands, almost 50% of all VRE cases between 2012-16 were *vanB*-type. *vanB*-type VRE exceeded *vanA* in Germany for the first time in 2016. In Australia, *vanB* has been an endemic *van*-type VRE for over 20 years with rare *vanA* isolates (128), while several *vanA*-type VRE outbreaks have been reported from Australian hospitals in recent years (2015-17), including one hospital with a *vanB*-type endemic VRE (129–131). On the other side of the world, in Asian countries like China, Japan, and (South) Korea, *vanA* has been reported to be the predominant type of VRE. Interestingly, the *vanM* gene cluster is becoming more prevalent in recent years in China (132). In several VRE studies in India, *vanA* is reported as the predominant *van* gene cluster (133–135). The incidence of *vanC* on the other hand is relatively low worldwide, although the number of outbreaks caused by *vanC*-type VRE is increasing. These are mainly associated with bacteraemia caused by intrinsic *vanC*-encoding *E. gallinarum* and *E. casseliflavus* (81). According to the United States Centers for Disease Control and Prevention (CDC), the estimated number of VRE infection cases in 2017 was 54,500, resulting in 5400 deaths. Among the AMRs, VRE death cases ranked second after MRSA in the US in 2017 (136).

1.9 Hypotheses on the origin of *van*-types

The similarity between the sequences of *van* gene clusters in vancomycin-resistant pathogens and glycopeptide antibiotic-producing (GPA) microorganisms such as *Amycolatopsis orientalis* (chloroeremomycin producer), *A. orientalis* subsp. lurida (producer of ristocetin), *Amycolatopsis coloradensis* subsp. labeda (producer of teicoplanin and avoparcin), *A. balhimycina* (producer of balhimycin), and *A. teichomyceticus* (producer of teicoplanin), point to Actinomycetes as a possible origin of the *van* gene cluster. *vanHAX* in enterococci, staphylococci, and actinomycetes often follow the same order. They are translated together and

have a high amino acid sequence identity. Regarding the *vanY*, it has a supportive role in VRE, while they are relevant in GPA-producing actinomycetes. Sequence similarity between *vanS* of enterococci and actinomycetes is very low which reflects the different modes of recognition between them (137).

A pioneering study on the phylogeny of the biosynthetic gene clusters of glycopeptide antibiotics, suggests the appearance of glycopeptide biosynthesis and resistance in Actinobacteria 150–400 million years ago (138–141). For the *vanA* gene cluster, three distinct origins are proposed. The *vanA* gene probably originated from *Amycolatopsis* genus and the *vanY* is most likely derived from the genus *Nonomuraea*, while *Actinoplanes* is the suggested origin for *vanH*, *vanX*, *vanR*, and *vanS* (138,139).

1.10 HGT mechanisms in enterococci

HGT or lateral gene transfer refers to all types of genetic material transfer from one cell to another, enabling taxonomically different organisms to share a common genetic pool (142,143). HGT can occur between any two organisms that contain DNA as their genetic material. It can happen in any possible direction between bacteria, archaea, and eukarya. But more likely, it does not occur equally in all the branches of the tree of life. For example, in bacteria and archaea, it is known that HGT is the main driver of genome evolution (144,145). Moreover, HGT happens more frequently in closely related taxa and between bacteria that share the same environment (146).

The discovery of HGT in bacteria dates to the famous story of transforming non-virulent or R forms (rough colonies) of *Streptococcus pneumonia* to virulent forms or S (smooth colonies) by Frederick Griffith in 1928. His experiment showed that an extract from dead R form pneumococci could transform the living S form into the R form (147). Later in 1951, Hotchkiss showed that bacteria can take up antimicrobial resistance genes (ARG), which transform them into antibiotic resistant bacteria (148). Moreover, he successfully induced resistance to ampicillin and streptomycin in susceptible *S. pneumonia* strains by exposure to a DNA extract of resistant strains. Further experiments proved the DNA exchange between different bacterial species (149). They revealed that MGE plays a key role in bacterial uptake, accumulation, and spread of resistance genes(150). MGE refers to segments of DNA that encode proteins necessary for the movement of DNA, either intracellular or intercellular (151). The first transferable antibiotic resistance in enterococci was reported in 1972 which was related to

plasmid-mediated tetracycline- and erythromycin-resistance (152). Finally, in 1986, the first VRE isolates were reported (49,153,154).

Among prokaryotes, three main mechanisms for HGT have been described so far: transduction, conjugation, and transformation (Figure 5). Recently some other mechanisms for HGT in prokaryotes have been discovered (155) but they will not be described further here.

a) Transduction

Transduction is bacteriophage-mediated gene transfer that certain types of phages can carry out. Bacteriophages or phages are viruses that infect bacteria. During the assembly of bacteriophages, in which capsids encapsulate phage DNA, bacterial phages can mistakenly wrap segments of host DNA in their capsids and transfer them to another bacterial cell (151).

b) Conjugation

Conjugation is the most common, and the best characterized mechanism of HGT in gut bacteria. It requires direct cell-to-cell contact between the donor and recipient mainly via the formation of a pilus bridge (156,157). Since the gut has a high microbiota and mucus layer density, it provides a suitable environment for conjugation between bacteria(157). Conjugation needs independently transferable genetic elements such as conjugative plasmids or ICEs, previously termed conjugative transposons (CTns) which encode all proteins necessary for transfer (151).

c) Transformation

Transformation is a type of genetic material transfer in which the recipient can take up foreign DNA and integrate it into its genome. To be able to take up exogenous DNA, the recipient must be in a specific state called competence. The mechanism of induction of competence and transformation varies between different species. The process of transformation is entirely controlled by recipient bacterial encoded genes (158). In enterococci conjugation and transduction are the main mechanisms of HGT, while natural transformation has not been confirmed in this genus so far (2,159).

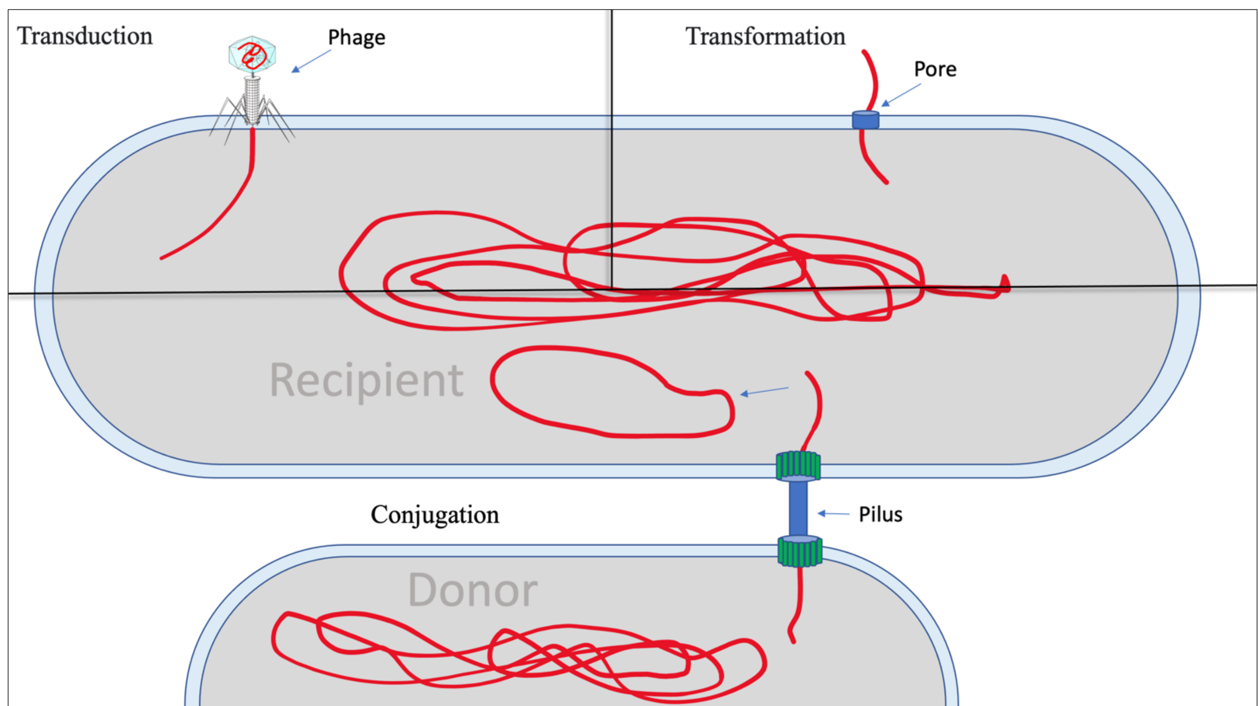


Figure 5. Overview of the main mechanisms of mobile genetic element uptake in HGT; based on Arnold et al. (143).

1.11 The role of different MGEs in HGT in enterococci

In prokaryotes, four main types of MGEs facilitate HGT: plasmids, genomic islands, bacteriophages, and transposons (49,155). Plasmids, genomic islands and transposons are the main MGEs involved in the spread of antimicrobial resistance in enterococci, specifically, *E. faecium* and *E. faecalis* (49,160).

1.11.1 Phages

The existence of phages in enterococci has been known for a century. They can transfer virulence factors and antibiotic resistance genes between enterococcal species (41). It has been confirmed that bacteriophages provided part of the accessory genome of *E. faecalis*. For instance, almost 10% of the *E. faecalis* V583 genome originated from seven prophage-like elements (pp1 - pp6 and EfCIV583) (41,161,162). Prophages are phage genomes that are integrated into the bacterial genome. It has been proved that phages can successfully transmit gentamicin resistance (*ant2*-I) and tetracycline resistance (*tetM*) genes between enterococcal species(163).

1.11.2 Plasmids

Plasmids are extrachromosomal genetic materials that can replicate independently and play an important role in the evolution of bacteria through HGT (41). Plasmids do not encode essential genes and impose a metabolic burden on the cell, but they can provide beneficial genes that boost bacterial fitness in complex environments. Many ARGs and virulence factors of enterococci, particularly *E. faecium*, are carried by plasmids (46). ARGs such as *van* gene cluster (*vanA* and *vanB*), aminoglycoside resistance (*aac(6')-Ie-aph(2")*), tetracycline resistance (*tetM*), quinupristin-dalfopristin resistance (*vat(D)* and *vat*(*E*)), and linezolid resistance (*cfr*, *optrA*, and *poxtA*) in enterococci are transmitted by plasmids (46).

Enterococcal plasmids are classified into four family groups (Inc18, RepA_N, Rep_3, and RCR) (164). This classification is based on the gene sequence of their replication initiator proteins, or other criteria, such as the mode of replication or the ability of plasmids to coexist within a bacterial cell (49,164). Recently linear plasmids were also reported in enterococci (165).

The incompatibility 18 group (Inc18) plasmids

The typing system of Inc18 plasmids is based on specific conservation functions, including the replication and post-segregation killing systems. They replicate according to the theta mechanism, which technically requires two factors encoded on the plasmid: a rate-limiting replication protein (Rep) and a short replication origin located downstream of the *rep* gene. The predominant plasmid maintenance system in Inc18 are termed toxin-antitoxin or post-segregation killing systems, which ensure persistence of these plasmid in the enterococcal population (166).

This group of plasmids is commonly isolated from enterococci and streptococci. In enterococci, Inc18 plasmids are responsible for resistance to vancomycin, chloramphenicol, and the "macrolide, lincosamide, streptogramine" (MLS) group of antibiotics. Inc18 plasmids harbouring Tn*1546* have been shown to be responsible for *vanA*-type vancomycin resistance in vancomycin-resistant *S. aureus* (VRSA). Also, Inc18-like plasmids that harbor the *vanA* gene cluster are common in enterococci (167–169). Inc18 plasmids are frequently encountered in clinical strains and are consistently recovered from the environment, particularly from livestock and sewage (170). pIP501 and pAMβ1 are two members of the Inc18 group that are very well

characterized (171). Both of these Inc18 plasmids can transfer ARGs conferring resistance to MLS-antibiotics between enterococci, staphylococci, lactococci, and streptococci (167).

RepA_N plasmids

The RepA_N family plasmids are characterized by their RepA-N domain in the replication initiator protein. The RepA_N protein has three domains. The first domain from N-terminal is the most conserved and performs some essential functions such as DNA binding and separation. In contrast, the C-terminal domain is responsible for some host-specific functions. The central domain of the RepA_N protein is highly variable and contains complex nucleotide repeats (164). RepA_N plasmids are widely distributed in low GC content Gram-positive bacteria and vary in size (3.3 to 281 kb) (164). The RepA_N family plasmids of enterococci include the pRUM-like plasmids of *E. faecium* and the pheromone responsive conjugative plasmids of *E. faecalis*(164). The pRUM plasmids were associated with *vanB2*-type transposon VRE*fm* outbreaks in Swedish hospitals (172,173), while some RepA_N-type plasmids, such as repUS15, are responsible for *vanA*-type VRE (174,175). pLG1 megaplasmids, the 281 kb large plasmids, are found in *E. faecium*, and responsible for some antibiotic resistances including *vanA* type VRE (164,176).

Rep_3 plasmids

Plasmids containing replication initiator proteins of the Rep_3 type are ubiquitous among bacteria. In enterococci, several Rep_3 plasmids have been identified to date: pMBB1, pDT1, and pCIZ2 in *E. faecium* and pS86, pAMα1, and pEF1071 in *E. faecalis*(164). The prevalence of rep_3 family plasmids in *E. faecium* is significantly higher than it in *E. faecalis* (177). Rep_3 plasmids are associated with tetracycline resistance (tet39) in different species but not in enterococci (178).

Rolling-circle-replicating (RCR) plasmids

In Gram-positive bacteria, RCRs are ubiquitous plasmids. They can also be found in Gram-negative bacteria and archaea. Numerous enterococcal cloning vector plasmids have been developed from RCR plasmids derived from other organisms. The pT181, the pMV158, and the pUB110 families are enterococcal RCR plasmids (164). RCR plasmids are not associated with AMR, but they are usually co-transferred with plasmids carrying *van* gene clusters. RCR

plasmids are present in 65% of VRE*fm* isolates whereas only 29% of vancomycin susceptible *E. faecium* (VSE*fm*) isolates contain this type of plasmids (179).

Linear plasmids

Plasmids are often circular DNA molecule structures but can be found in other forms. Linear plasmids were discovered first in the 1980s, and have since been found in only a few bacterial species, including *E. faecium* and *E. faecalis* (180). pELF1 and pELF2 are two transferable enterococcal linear plasmids reported in *E. faecium* that are associated with *vanA* and *vanM* gene clusters (165,181). Other known *van* gene clusters harbouring linear plasmids show high homology to these two plasmids (182–184).

1.11.3 Genomic islands (GI) in enterococci

GIs are horizontally transferred sequences integrated into the bacterial genomes. Usually, they have different codon usage and GC content compared to the rest of the genome. They are typically integrated close to tRNA genes and flanked by direct repeats. GIs harbour genes that may have been used for their mobilization. GIs can be mobile, non-mobile, or no longer mobile. Also, they are different in their ability in integration, excision, and transfer. They can be transferred via transformation, transduction, and conjugation (185,186). GIs can be divided into two types, integrative mobilizable elements (IMEs), and integrative conjugative elements (ICEs)(186). In enterococci, *vanD* gene cluster is mainly harboured on putative IMEs (108,187).

ICE elements in enterococci

ICE elements are characterized by their ability to encode all genes required for excision, conjugation, and integration into a recipient. They are frequently integrated into the bacterial chromosome (48,161). ICEs are responsible for resistance to *vanB2*-type vancomycin (Tn*1549*/Tn*5382*) (Figure 6), kanamycin and erythromycin (Tn*1545*), tetracycline (Tn*916*-like, Tn*6000*, Tn*5801*, Tn*5397*), and MLS antibiotics (Tn*1545,* Tn*2010,* Tn*2017*) (49,164,188). Tn*1549* is a common ICE in enterococci that confers vancomycin resistance through its *vanB2* subtype gene cluster (81). It can transfer between enterococci using either conjugative plasmids or As part of larger chromosomal elements (189). Tn*1549* has been found in different bacterial genera in the normal gut flora. The conjugative transfer of Tn*1549*/Tn*5382* from *Clostridium* to enterococci has been confirmed. Also, it can co-transfer by means of heterologous transfer

systems from *Clostridiums or Streptococcus* to enterococci pointing to the importance of non-enterococcal reservoir of Tn*1549*/Tn*5382* (48).

1.11.4 Transposons in enterococci

Transposons are the simplest MGEs that encode the essential enzymes for their transposition which is movement from one place to another within the genome. They are of three types: Tn*3* family transposons, insertion sequence (IS) elements, and composite transposons (48). Transferable plasmids and genomic islands, characterized by their ability to mobilize and conjugate, have a significant role in the transfer of bacterial genes and transposons in the HGT process (41,185).

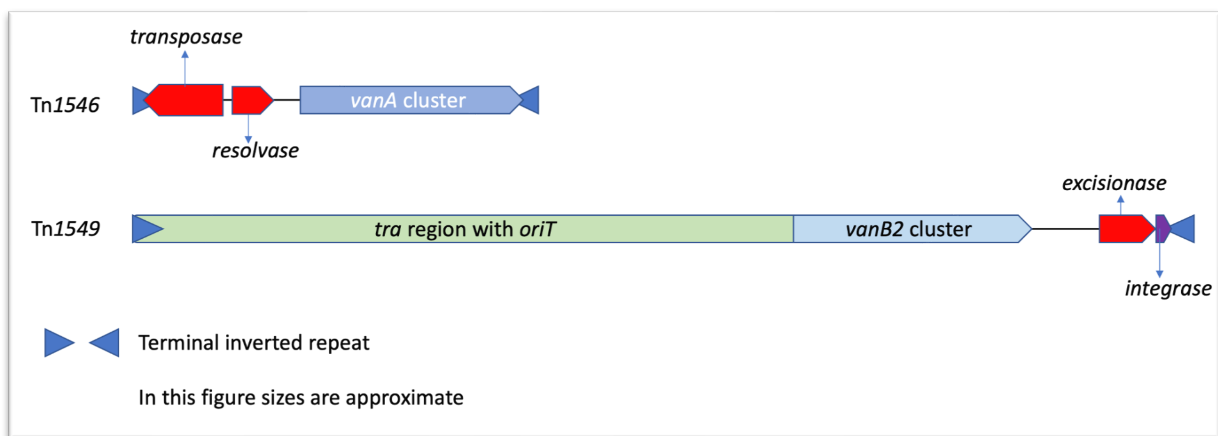


Figure 6. Schematic presentation of Tn*1546* and Tn*1549*; based on Hegstad *et al*. (49).

Tn*3* family

The transposons of Tn*3* family are intracellularly transposable by means of two enzymes, a transposase (TnpA) and a resolvase (TnpR). The Tn*3* family of transposons use a replicative transposition mechanism by which the transposon duplicates along with its insertion into the recipient DNA sequence (190). In enterococci, they are linked to resistance to vancomycin (*vanA*-type) and macrolide-lincosamide-streptogramin B (MLSB). Tn*1546* is a Tn*3* derivative that carries *vanA* gene clusters and is usually part of either conjugative or nonconjugative plasmids (49). Tn*1546* is mainly carried by RepA_N (pRUM/pLG1), Inc18 or a mosaic plasmid (175).

Insertion sequence (IS) elements

IS elements are short transferable DNA sequences that are important in shaping the bacterial genomes and are scattered in the genome of clinical *E. faecium* isolates (28). By disrupting gene promoters, CDSs, and operon's structure, they can rearrange the genome. IS elements are the simplest transposable elements that only encode enzyme necessary for their own transposition (49).

Composite transposons

The structure of composite transposons consists of adaptive features such as resistance genes bounded by a pair of IS elements (41). The mobility of composite transposons is dependent on the flanking IS elements of the same family. In the composite transposons either the entire unit can transpose, or the active IS element can do so alone (191). In enterococci, composite transposons are linked to mobilization of high-level gentamicin resistance (Tn*5281*) and one case of *vanB1*-type vancomycin resistance (Tn*1547*) (49,164). In Tn*1547*, the *vanB1* gene cluster is flanked by two IS elements (IS*16* and IS*256*-like) (192).

1.12 VRE; the successful resident of hospitals

The genome plasticity and remarkable ability to survive in different environments, as well as tolerance to detergents, antibiotic resistance, and virulence, have led to the frequent occurrence of VRE in healthcare facilities, enabling them to survive in the hospital for several years (2,193). The main source of VRE infection is an index patient (194), but also healthcare workers and fomites can serve as a VRE*fm* reservoir (128,195). The main modes of VRE transmission are the hands of healthcare workers, contacting a VRE-contaminated environment, and direct contact between patients (194).

1.13 Enterococcal virulence factors

Virulence factors (VFs) are molecular determinants that help bacteria colonize and invade their hosts resulting in infection and tissue damage (196). They also have a key role in bypassing the host immune system (197). VFs can be divided into two main groups based on how they affect the host: 1) those that affect host cell colonization and 2) secreted determinants that damage host tissue or help the bacteria evade the immune defence (198). During colonization by

Enterococcus, binding to extracellular matrix proteins is vital, which is why several enterococcal VFs are mainly associated with adhesion and biofilm formation (199).

Pilus protein (PilA, B, and C), extracellular membrane binding proteins (collagen adhesins (Acm), second collagen adhesin (Scm), Catabolite control protein A (CcpA)), fibronectin/ fibrinogen-binding proteins (Fnm and PrpA), and enterococcal surface protein (Esp) are examples of VFs involved in enterococcal colonization in their hosts. Capsular polysaccharide protein D (CapD) is a VF that helps to bypass the host immune system (199–201). Whereas VFs such as gelatinase and cytolysin damage host tissues (198). There are 30 experimentally confirmed VFs in *E. faecium* (199–204).

1.14 Common methods and techniques for studying enterococci

Enterococcal identification is of clinical importance and can be done by phenotypic or genotypic methods. Various selective media, differential experiments, and commercial kits are available for phenotypic identification of enterococci, but they are time-consuming and do not give enough resolution (2,205). In contrast, molecular methods can save time and money, increase accuracy, and provide more information. Several molecular technics such as matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS), PCR-based nucleic acid amplification tests (NAATs), and peptide nucleic acid fluorescence in situ hybridization (PNA-FISH) can identify the enterococci at the species level with high accuracy. Higher resolution inter-species identification can be obtained by methods such as multilocus sequence typing (MLST), pulsed-field gel electrophoresis (PFGE), and multiple-locus variable number tandem repeat analysis (MLVA) (2,206). Regardless of their usefulness for strain-level analysis, the discriminatory power of methods such as MLST and PFGE does not exceed 95% (206).

Next generation sequencing (NGS) is a term used to describe any sequencing technology with high throughput, massive parallelism, or deep sequencing (207). NGS can be used to sequence the entire genome of an organism by whole genome sequencing (WGS). NGS technologies facilitate WGS of multiple pathogens in one run (208). It can quickly and accurately identify a pathogen, its virulence factors, trace the transmission of a pathogen in a population, and suggest probable sources (209).

1.15 Next generation sequencing

NGS technologies can be classified based on the length of reads they produce. NGS was developed after Sanger sequencing (first generation sequencing). The second generation sequencing introduced short reads sequencing and the third generation sequencing includes long reads sequencing technologies (210).

1.15.1 Short read sequencing

Several platforms have been developed as second-generation sequencing. 454 Roche GS FLX System (2004), Illumina (2006), SOLiD sequencer (2007), and Ion Torrent (2010) are different short read sequencing. They differ in their chemistry, read length, error rate, cost, run time, and genome coverage. Technically, the sequencing process starts with DNA extraction, followed by library preparation (fragmentation of DNA and adding adaptors to both ends), amplification (cluster generation), and end with the sequencing step. In the library preparation process, genomic DNA is broken down into small, random, and overlapping fragments. A small piece of DNA called an adaptor is then ligated to the end of each fragment. Next, in the amplification of the fragments, a clonal amplification is carried out. The amplification is mainly PCR-based. Finally, the samples are sequenced, which is the step in which the different technologies vary. Sequencing can either be based on synthesis or on hybridization and ligation (211).

Illumina sequencing

By far, Illumina is the most common platform in the market and considered a highly accurate and robust sequencing platform (212,213). The maximum read length produced by Illumina is up to 300 bp (214). Briefly, to prepare the library, adaptors are added to both ends of the fragmented DNA or cDNA. Then, by means of oligos, fragments are attached to the solid surface of the flow cell. Oligos are short sequences grafted on the flow cell surface that complement the adaptors (212,215). The amplification in Illumina sequencing is a PCR-based technique called bridge amplification in which clusters (hundreds of identical DNA strings) are produced (212,216). Bridge amplification is a process wherein DNA fragments with adaptors ligated to their two ends are the substrate for repeating the amplification. Finally, the last step of sequencing in the Illumina platform is reversible termination by adding a single (fluorescent) labelled complementary deoxynucleotide triphosphate (dNTP). The fluorescent dye is identified by laser excitation and imaging, and then it will be cleaved by an enzyme to allow the next round of incorporation (217). Two types of Illumina sequencing are available: single direction, and paired end. In single direction, DNA fragments are sequenced from the 5’ end,

while in paired-end method fragments are sequenced from both 5' and 3' ends, resulting in a double number of base-pairs per reads (218). Paired-end sequencing is exclusively the capability of Illumina sequencing in which the amplification is based on bridge amplification (219). Due to the higher likelihood of alignment to a reference, paired-end sequencing can produce high quality sequences (213) with in-depth coverage and high numbers of reads (210). The most important drawback of the Illumina platform is the relatively long run-time (210). Illumina has several sequencers with different sequencing outputs and total reads/run (iSeq, MiniSeq, MiSeq, NextSeq 550, NextSeq 1000, NextSeq 2000, and NovaSeq 6000) (220).

Illumina is the most used platform in enterococcal genomic studies, and almost 89% of all genomic studies on the enterococci are performed by Illumina, followed by PacBio, Ion Torrent, and hybrid Illumina – PacBio sequencing. Cost-effectiveness, low error rates, and accessibility made Illumina platform the first choice for enterococcal genomics studies (221). Also, since Illumina can produce over 30x coverage, it is the standard for accurate SNP-calling (222).

1.15.2 Long read sequencing

On average, third-generation or long read sequencing platforms produce 30 to 50 kb reads but compared to short-read sequencing, have a higher error rate (per bp) and cost (223). In long read sequencing, read length can exceed 1 million bp. Several platforms perform long-read sequencing; Pacific Biosciences (PacBio), Oxford Nanopores technology (ONT), Quantapore, and Stratos (223). PacBio and ONT are the most used long read sequencing platforms.

PacBio sequencing

PacBio sequencing uses an approach called Single Molecule Real-Time (SMRT) for sequencing. It facilitates sequencing fragments up to 50 kb or longer (215). DNA polymerase molecules that are attached to the bottom of a well called zero-mode waveguides (ZMWs) bind to the template DNA. Each polymerase enzyme can synthesize second strand DNA in the presence of nucleotides labelled by fluorescent γ-phosphate. Because of the tiny width of the ZMW, light cannot propagate the waveguide. Fluorophores near the polymerase enzyme (at the bottom of the well) are excited and penetrated by the energy. Then, real-time fluorescence pulses are detected as each base is incorporated (224). Due to high sensitivity of PacBio SMRT systems, amplification steps are not necessary (225), and the prepared library DNA template is sequenced directly (210,225).

PacBio long reads can be combined with Illumina short reads and be hybrid assembled. This approach is useful for closing the bacterial plasmids and chromosomes (221). Additionally, PacBio can produce high fidelity (HiFi) reads with its Sequel II system. HiFi reads are long (15-20kbp) and highly accurate. Furthermore, the accuracy of PacBio can be enhanced and reach the Illumina error rate by the circular consensus sequencing (CCS) method (226). In the CCS method, a circular template DNA is created by adding ssDNA hairpin adaptors to the dsDNA, allowing multiple template sequencing (226).

ONT sequencing

In ONT sequencing, adaptors ligated to the single-strand DNA or RNA facilitate their capture by a staphylococcal protein pore (α-hemolysin), wherein the template should pass through during the sequencing process. The ONT flow cell contains a membrane of hundreds of thousands of pores. When the libraries are loaded onto the flow cells, with the help of a preloaded motor enzyme and ion current, they pass through the pores and disrupt the ion current. This can be detected by sensors and recorded (210). Compared to PacBio, ONT can produce up to 4 Mbp reads at a lower cost but with less accuracy (227). The accuracy of ONT is between 87 and 98% which is not enough for variant calling, but it is improving over time (228). Lately, ONT is advertising it's over 99% accurate sequencing (229). Furthermore, the portable ONT device (MinION) is pocket-sized and can produce real-time data (230).

1.16 Bioinformatic analyses of NGS raw data

Several bioinformatics tools and methods have been developed for the analysis of WGS data of outbreaks. Single nucleotide polymorphism (SNP)-based approaches and high-resolution sequence typing, such as core genome MLST (cgMLST), are useful bioinformatics methods for analysing WGS data of outbreaks. SNPs are highly informative markers for outbreak studies because they vary between isolates. Variant calling methods are based on a comparison of the data of interest to a reference genome. cgMLST is an upgraded MLST that provides higher resolution and accuracy. In the regular MLST, typing is based on seven housekeeping genes, whereas the cgMLST scheme includes most of the species' core genome. For *E. faecium*, the cgMLST scheme includes 1423 genes, while for *E. faecalis* it includes 1972 (231,232). Even after typing isolates using MLST or cgMLST, it is possible to achieve a finer resolution to identify probable transmission events. Some tools can calculate the pairwise SNP distance between all sample pairs based on a core genome alignment. Methods such as multi-locus sequence type core alignments (MLSTCA), cgMLST core alignments (cgMLSTCA), and

cgMLST with cluster reference core alignments (cgMLSTCRCA) can calculate the pairwise SNP distances (233). Moreover, a phylogenetic tree can also be constructed based on the core genome SNP multi alignments using the tool Parsnp (234). Other tools like Roary can generate concatenated core genome alignments for phylogenetic tree construction (235), but Parsnp is more useful because it is very fast and can align hundreds to thousands of isolates in a short time (234).

1.16.1 MLST

MLST is a method for isolate typing of bacterial pathogens that was proposed first in 1998. In MLST, alleles are used as a comparison unit, and each allelic change is counted as a single genetic event (236). MLST is a standardized approach that provides an unambiguous universal system for strain typing. It is useful for epidemiological investigations, population biology, pathogenicity, and bacterial evolution studies (237). There is a public database for molecular typing and microbial genome diversity (https://pubmlst.org/) in which all the new alleles and allelic profiles are registered (238). The sequences used in the scheme of MLST are either the full length or fragments of housekeeping genes (239). In general, an MSLT scheme has seven indexed loci. For each unique sequence allele, an ID number is assigned. In an allelic profile, the ID number of each locus is incorporated (239). For example, in ST17 of *E. faecium* the allelic profile is (1,1,1,1,1,1,1). The scheme of MLST for *E. faecium* was proposed in 2003 containing 300- to 600-bp length of internal fragments of the following genes; adk (adenylate kinase), *ddl* (D-alanine:D-alanine ligase), *gyd* (glyceraldehyde-3-phosphate dehydrogenase), *atpA* (ATP synthase, alpha subunit), *gdh* (glucose-6-phosphate dehydrogenase), *purK* (phosphoribosyl aminoimidazol carboxylase ATPase subunit), and *pstS* (phosphate ATP-binding cassette transporter) (240). For the scheme of *E. faecalis*, three of the housekeeping genes are in common with *E. faecium* scheme (*gdh*, *gyd*, *pstS*) and the rest are *gki* (putative glucokinase), *aroE* (shikimate 5-dehydrogenase), *xpt* (shikimate 5-dehydrogenase), and *yiqL* (acetyl-coenzyme A acetyltransferase). The internal region of *gdh* gene is slightly different from that of *E. faecium* (241).

1.16.2 cgMLST

Regardless of MLST benefits, due to the small number of genes in the MSLT scheme, its resolution is limited. The potential of WGS paved the way for the use of more genomic data in bacterial genomic studies, especially in outbreak surveillance where high resolution typing of isolates is vital (231).

SNP-based approaches can produce enough resolution, but the main drawback of SNP calling in outbreak surveillance is using different reference genomes, which complicates data comparisons between different studies(231,236). Expanding the MLST scheme from seven housekeeping genes to several hundred core-genes can give enough resolution. Allelic diversity of such a number of genes can more easily identify the outbreak related isolates and cluster them in a group known as a cluster-type (CT). cgMLST is a standardized method that translates the genomic variations (SNPs) into a portable numbering system for alleles, and compared to SNP-calling is less computationally intensive (231). To build a cgMLST scheme, a large number of isolates of one species are analysed and loci that are present in 95% of those isolates will build the scheme (242). Thus, because of different core genome sizes, the number of target genes in the cgMLST differs between species. Moreover, for technical reasons, some loci might be excluded during the process of scheme building (231). cgMLST results can be translated to phylogeny and used to build a tree using different methods such as minimum spanning (MS), neighbour-joining (NJ), or single-linkage hierarchical clustering (242).

1.16.3 SNP-calling

Although higher resolution typing can be attained by cgMLST, considerable diversity in the genome is disregarded since it is not included in the cgMLST scheme. Genes which are not included in the cgMLST schemes may affect the scheme's ability to identify closely related outbreak isolates. Here SNP-calling between bacterial isolates can be a useful method. SNP-calling can be reference based or use a reference agnostic method (243). In the reference based SNP-calling, the isolates of interest are aligned against a closely related (closed) reference genome (244). The most widely used, accurate, and simplest method for SNP-calling is reference-based, which is useful for relatively few isolates. This method also has its limitations. In reference based SNP-calling, the reference genome is very important, and not always a closely related close genome for interest isolates is available (243). To bypass the reference genome problems, k-mer based comparison methods for SNP calling were developed. K-mers are blocks of (k) length sequences that can be compared between sets of interest genomes to model intra-isolate diversity. It has been shown that k-mer based approaches using a tool (kSNP3.0) can produce consistent results without reference (245). The split K-mer analysis (SKA), is based on pairwise SNP and can identify variations among a large number of closely related genomes. SKA is a rapid method and can analyse read data or genome assemblies as input. SKA can be run on a standard personal computer and calculates pairwise distances, builds single linkage clusters, and aligns genomes using either reference-based or reference-free

approach. The notable drawback of SKA is its inability to identify SNP in repeated K-mer split (246).

2 Objectives of the study

This chapter outlines the main objective of the research project and the specific objectives for the three manuscripts. The overarching aim was to examine recent developments in the epidemiology of VRE in Norway related to the occurrence of novel *van*-types and the increasing number of VRE*fm*.

The specific objectives related to **paper I** were to:

1. Examine the genetic relatedness between the first *vanD*-type VRE isolates in Norway
2. Identify the MGE harbouring the *vanD* gene cluster, compare it to previously reported MGE associated with *vanD* gene and examine its mobility

The specific objectives related to **paper II** were to:

1. Examine the genetic relatedness between the first *vanE*-type VRE isolates in Norway
2. Identify the MGE harbouring the *vanE* gene cluster, compare it to previously reported MGE associated with *vanE* gene and examine its mobility

The specific objectives related to **paper III** were to:

1. Identify main sequence types (STs) and cluster types (CTs) in a representative concurrent selection of Norwegian VRE*fm* and VSE*fm*
2. Compare the main STs and CTs with global VRE data
3. Identify and compare MGEs harbouring *vanB* and *vanA* gene clusters in main STs/CTs
4. Elucidate specific virulome patterns in VSE*fm* and VRE*fm* and their main STs
5. Identify possible within-hospital *van*-MGE exchange between co-occurring *vanB*-type VRE
6. Examine the presence of other clinically relevant antimicrobial resistance phenotypes in the selected VRE*fm* and VSE*fm* isolates

3 Materials and methods

The methods used in this research project are listed and described in detail in the three manuscripts resulting from this research project. The following section presents a general description and rationale for the selection of methods. Some of the methods that require more detailed descriptions are explained in detail.

3.1 Bacterial culture collections

Enterococcal cultures collected at different laboratories around Norway were sent to the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance (K-res). The *vanD* and *vanE* type VRE were studied in **papers I** and **II**, respectively, and *vanA* and *vanB*-type VRE were studied together with the VSE isolates in **paper III**. The criteria for including isolates in the study varied between the first two papers and the last one. In papers I and II, all identified *vanD* and *vanE* in Norway (regardless of species, time, and geography) were studied. A random subset of *vanA*- and *vanB*-type VRE*fm* and VRE*fs* were included in paper III. In the case of VSE isolates, all available VSE*fm* isolates from NORM 2008 and 2014 were included in the study (Table 1).

Table 1. The number of isolates that were whole genome sequenced in this research based on their collection, species, and paper.

Collection	Species	*van*-type	Paper	Number of sequenced isolates
vanD	*E. faecium*	*vanD*	I	4
	E. casseliflavus	*vanD/vanC*	I	2
vanE	*E. faecalis*	*vanE*	II	2
VSE 2008	*E. faecium*	-	III	99
VSE 2014	*E. faecium*	-	III	162
VRE (2010-15)	*E. faecium*	*vanA/vanB*	III	229
	E. faecalis	*vanB*	III	12
Total				510

3.2 Bacterial species identification

Each isolate was first subjected to MALDI-TOF to confirm the species according to the manufacturer's instructions. MALDI-TOF is a fast, and robust technique to identify the microorganism's species reliably (247). The high sensitivity of MALDI-TOF method can differentiate between closely related enterococci species and around 94% of bacterial isolates can be identified at the species level. Additionally, MALDI-TOF can identify enterococci from blood culture bottles, which saves time and is vital in antibiotic therapy initiation (2).

3.3 Antimicrobial susceptibility testing (AST)

Broth microdilution method (BMD), besides agar disk diffusion, is the gold standard for AST (248). It has been argued that vancomycin's large molecular size affects disk diffusion, so trained personnel are required to interpret growth inhibition zone edges. CLSI agar screening method performs with acceptable sensitivity and specificity in detecting some low-level *vanB*-type VRE (249,250). In **papers I** and **II,** BMD methods were used to perform AST, as we were dealling with a few isolates of two rare *van*-types. While in **paper III,** the majority of VRE isolates were *vanB*-type which is known to have moderate to low level resistance to vancomycin (83,249), thus, the CLSI agar screening method was used to perform AST for vancomycin. For the other antimicrobials (gentamicin, linezolid, and ampicillin) AST were preformed using disks on MH agar plates, according to the EUCAST method.

3.4 Filter mating

Among different methods to test the ability of isolates to conjugate, filter mating was the best for our purpose. The formation of mating-pair and its efficiency are influenced by different parameters such as local cell density, the type of conjugative pili, and the outer-membrane proteins of the recipient's cells. In filter mating, cells are fixed in their place. Thus, mating is limited to neighbouring cells. Filter mating is more efficient in taxonomic related isolates since donor and recipient cells have enough time to conjugate (251). Moreover, in enterococci, most MGEs/plasmids are not transferred efficiently in liquid mating (164). In **papers I** and **II**, we performed a filter paper mating method to determine the possibility of conjugation and the transferability of MGE harbouring *van* gene clusters using the same species as a recipient.

3.5 DNA extraction

We used two different kits for genomic DNA extraction. For Illumina sequencing purposes the DNeasy Blood and tissue kit (Qiagen, Hilden, Germany) was used as it is a standard in our lab. Since the quantity of DNA needed for PacBio sequencing is much higher (3 µg), we used Wizard Genomic DNA Purification Kit (Promega, Madison, USA) to obtain enough DNA. This kit had already been used by other collaborating research groups with success for long read sequencing.

3.6 Sequencing (Illumina and PacBio)

We chose the Illumina platform to sequence all the isolates for several reasons. First, it has a low error rate and is cost effective. Also, the coverage of Illumina sequenced draft genomes is high enough for our purpose. High accuracy reads of Illumina can be mapped either against their assembly or a related genome for error correction (210,213,222). Moreover, Illumina reads can be used directly in some bioinformatic tools like ResFinder (252). Additionally, NextSeq550 is the available Illumina sequencer in the genomic support center of Tromsø (253). To increase the accuracy and coverage of the resulting draft genomes, we chose paired-end Illumina sequencing (213). Regardless of all advantages of Illumina sequencing, the Illumina sequenced genomes cannot be closed because of longer repetitive sequences in the genome (254).

To close the genomes, long read sequencing is recommended (255). Therefore, we used the PacBio sequencing platform and CCS for their higher accuracy to close several selected genomes. The aim of closing genomes was to use them as references for comparative genomic analysis such as sequence comparisons, reference-based variant calling and other analyses that need reference genomes.

3.7 From raw reads to assemblies

Bioinformatic analyses start with the assembly of raw reads generated from WGS. This starts with removing adaptor sequences and low-quality reads and end with final draft genomes. In this part, before choosing a tool, we did a benchmarking for several tools to choose the best performing one.

For adaptor removal and quality filtering of the raw reads obtained from Illumina sequencing Trimmomatic v0.39 (256) and Trim-Galore (257) were tested and the quality of the reads resulting from both tools was assessed. Trimmomatic performed the best, as the overall quality of its output was higher, and was used for all isolates. The quality of trimmed raw reads was assessed using the standard tool of reads assessment which is FastQC (258). In **Paper I**, among four assemblers (SPAdes, ABYSS, Skesa, and SOAPdenovo), SPAdes (259) performed the best as it produced a lower number of contigs, high genome completeness and slightly higher genome coverage, and was thus chosen to assemble all the samples. In **papers II** and **III,** in addition to those tested in the paper I, we tested a new assembler (Unicycler) (260). The quality of Unicycler assemblies' exceeded SPAdes assemblies and therefore it was chosen to assemble all the isolates. Finally, a maximum of 400 contigs and minimum of 40x genome coverage were set as the cut-off values to include the assemblies in the analyses. In addition, assemblies with a genome size fluctuation equal to or less than 10% of the smallest and largest completed/closed genome deposited at NCBI were considered in the bioinformatics analyses. To achieve a higher resolution and accuracy, more specifically in virulome and ARG analyses, a minimum of 40x coverage was set up (261). The *E. faecium* genome contains lots of MGEs and repeated sequences that potentially affect the assembly and can be an obstacle to generating larger and fewer contigs. Therefore, we used the maximum of 400 contig, a standard followed by K-res as well. Also, in the case of sorting contigs, assemblies with less than 400 contigs perform better. The genome size standard is set because of the open pan-genome of *E. faecium* and to avoid mixed sample sequences (Figure 7).

For PacBio sequence reads, first we tested the hybrid assembly of Illumina and PacBio reads in Unicycler, which in all samples resulted in assemblies of hundreds of contigs. Then, we tested other tools for long reads assembly such as Unicycler, Canu, Velvet, Raven, and Flye. Overall, Unicycler and Canu performed the best, but the quality of assemblies produced by Unicycler was slightly higher than Canu and had fewer errors. Thus, Unicycler was used for the PacBio sequence reads. Samples that Unicycler failed to circularize, were re-assembled using Canu, (262) then polished by Pilon (263) and circularized using Circlator (264). Assemblies that met all the standards were used for the bioinformatic analyses.

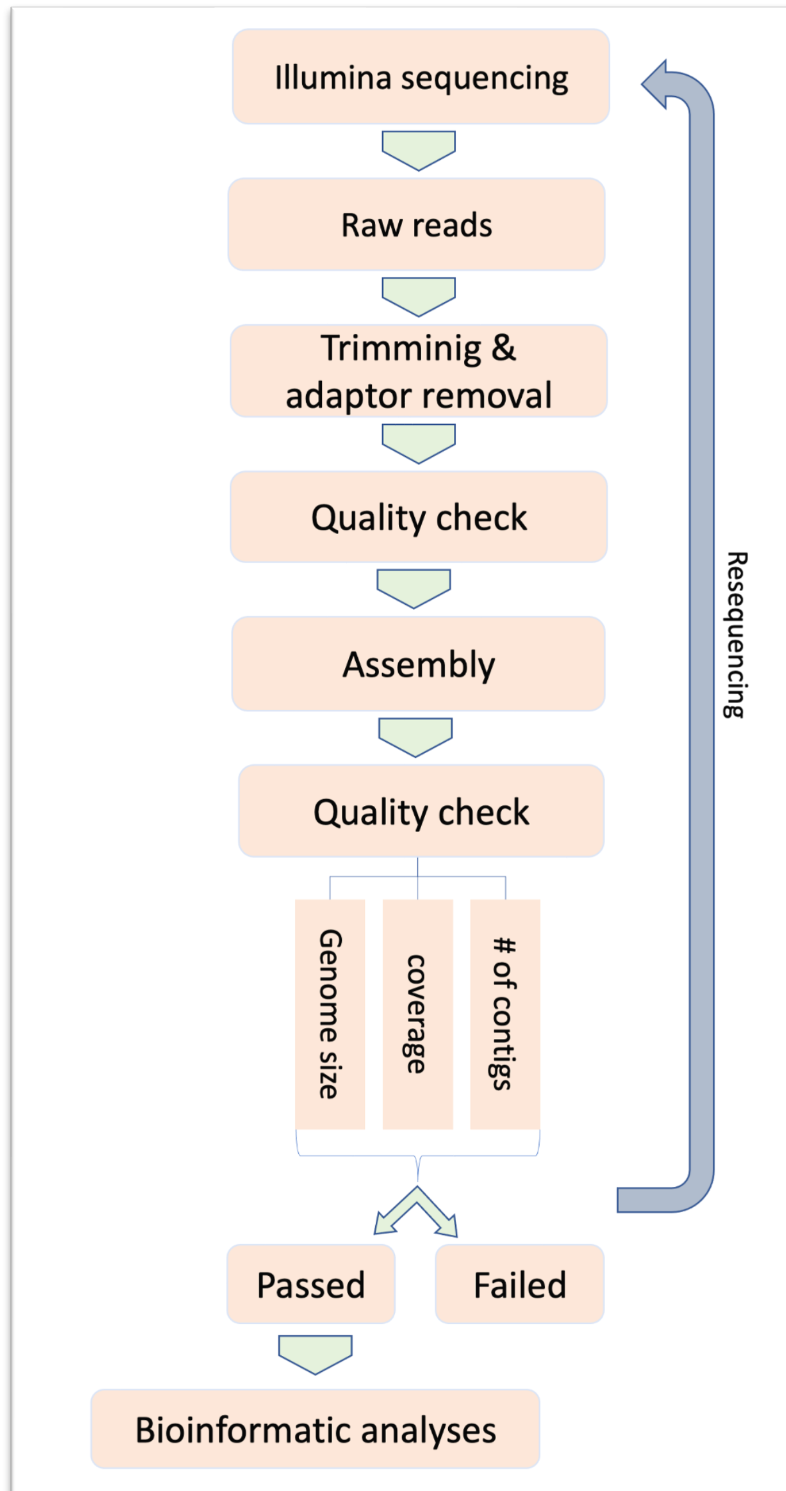


Figure 7. From Illumina raw reads to assembly.

3.8 Typing of isolates

MLST is a universal robust method for studying pathogenic bacterial population structure and epidemiology. However, cgMLST is the best typing method in terms of resolution, but not all VRE studies use cgMLST methods. Therefore, to compare the Norwegian population structure to global data, we needed a more common method such as MLST. On the other hand, we needed high resolution typing to identify the CTs to compare the isolates within the studies and for this purpose we used cgMLST. In silico MLST in (**papers I**, **II**, and **III**) was performed using MLST tool V2.19.0. As well we used SeqSphere+ software V6.0.2 (Ridom GmbH, Münster, Germany [http://www.ridom.de/seqsphere/]) with the default options to perform cgMLST. Minimum spanning trees were built based on the cgMLST scheme of *E. faecium* (papers I, III) or *E. faecalis* (papers II, III). Novel STs and CTs were submitted to pubMLST and Ridom SeqSphere+, respectively.

3.9 Phylogeny

To build the global phylogenetic trees (**papers I**, **II**, and **III**), we retrieved the available closed genomes of enterococcal species as of the date of building the trees and drew the trees using the Parsnp tool with default options, including flag (-c) to avoid excluding any samples. Local trees in **paper III** were built using the same options in Parsnp. Parsnp is the best tool for a large number of closely related species or strains and can build a tree based on the core genome alignment and SNP detection, even for low-quality draft genomes (with several hundred contigs) in a short time (234). The tree visualizing tool (Gingr) of the harvest suite that was developed for displaying Parsnp trees does not have the ability to annotate and display metadata (234). Therefore, to display the important metadata on the midpoint-rooted tree we used Interactive Tree Of Life (ITOL) v6, an online tool that can annotate the tree in multiple layers (265).

3.10 AMR and VF genes identification

To identify AMR genes in the isolates (**papers I**, **II**, and **III**), we used NCBI bacterial AMR reference gene database (PRJNA313047) in ABRicate tool V1.0.1 (266). ABRicate is an open-source tool with an acceptable running speed. Also, it is possible to use different ARG databases with the capability of updating them (266,267). PRJNA313047 has the highest number of curated ARGs among ARG databases with 6386 as of 01.11.2022 (266,268). In contrast to

AMR genes, there is no updated available database for enterococcal VFs. Thus, we used the ability of ABRicate to create our *E. faecium* VF database. We used the coding sequence of all 30 experimentally confirmed VFs of *E. faecium* identified by extensive searches in PubMed. A cut-off of 90% was used for identity and coverage before BLASTing the genomes against our VF database (paper III).

3.11 Identification of MGEs harbouring *van* gene clusters and their genomic integration site

To identify the integrated MGEs harbouring *van* gene clusters (*vanB*, *vanD*, and *vanE*) (**papers I**, **II**, and **III**) into the genomes, we used comparative genome sequence analyses (269), which is described in detail in the **paper III**. One of the most important steps in this part was choosing a proper reference genome. The reference genome should not have any MGE at the insertion site of the MGE harbouring *van* gene cluster of VRE isolate of interest. In **paper I,** we used the closet complete genomes retrieved from NCBI from the global tree built for this purpose. However, in **paper II,** the closest genome was not used as it contained another MGE at the exact location where Tn*6202* was. Thus, we used the second closest closed genome. In the **paper III,** eight VSE isolates form VSE 2008 and VSE 2014 collection were PacBio sequenced, and their genome were closed. These sequences were used as references for the *vanB*-type VRE isolates. In the case of VRE*fs,* as no VSE*fs* were included in the study, the same as **papers I** and **II,** we chose the closest complete genome from the global phylogenetic tree.

3.12 Reconstructing plasmids of *vanA*-type VRE isolates

The assembly of the accessory genome (such as plasmids) from Illumina short reads sequencing is challenging (270). Thus to identify the plasmids in Illumina sequenced isolates, we need to reconstruct them and use other approaches to compare them to the reference plasmids (271). Two approaches were used in **paper III** to reconstruct the plasmids and compare them to the references. First, the Mob-suite tool (272) was used to reconstruct the plasmids, followed by a BLAST search against the NCBI AMR database on ABRIcate. Next, plasmids with *vanA* gene clusters were typed against the PlasmidFinder online database (273). The second approach was used to compare the plasmids to the circularized reference. We mapped the Illumina reads of *vanA*-type isolates against the PacBio sequenced plasmid harbouring *vanA* type gene cluster.

Samples containing reads that covered the entire length of the reference plasmid were considered as similar to the reference plasmids.

4 Summary of results

Paper I: Novel genomic islands and a new *vanD*-subtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway.

- The first VanD-type VRE in Norway were detected in two different patients within two months.
- VanD-type VRE isolates belonged to different enterococcal species, *E. faecium* (from both A and B cases), and *E. casseliflavus* (from case B).
- We reported the first two *vanD*-type *E. casseliflavus* isolates, identified a novel subtype of *vanD* gene cluster termed *vanD6* and described three novel GIs harbouring *vanD* gene clusters of putative *Clostridiales* order origin integrated at the same chromosomal site in both *E. faecium* and *E casseliflavus*.
- Circular forms of the *vanD*-GIs were detected in all but one isolate but transfer to an *E. faecium* recipient was not detected.
- Different *vanD* subtypes on various GIs and phylogenetic distance revealed that the isolates of the two cases are not clonally related despite temporal occurrence.

Paper II: The first *vanE*-type vancomycin resistant *Enterococcus faecalis* isolates in Norway – phenotypic and molecular characteristics.

- The first VanE-type VRE isolates in Norway were recovered from the same patient 30 months apart.
- Their *vanE* gene cluster, harboured by the previously described MGE Tn*6202,* only showed difference in *vanS*$_E$.
- The *vanS*$_E$ gene was truncated in both isolates, but in the E1 isolate, the downstream histidine kinase part of the *vanS*$_E$ gene was still expressed.
- The premature stop codon in *vanS*$_E$ of E1 resulted in the histidine kinase domain still being in frame with *vanR*$_E$ while insertion of IS*6770* in *vanS*$_E$ of E2 likely resulted in inducible low level vancomycin resistance in E1 and constitutive low level vancomycin resistance in E2.
- Neither circular forms nor transfer of Tn*6202* between *E. faecalis* were detected.
- The two *vanE*-type isolates (E1 and E2) were considered clonally related as they were recovered from the same patient and had the same ST. The core genome and the allelic

differences between them were not too high (n=32), although the mutation rate in E1 was 125-fold higher than in E2. Moreover, only 60 variants existed between the two genomes that have similar organization.

Paper III: The population structure of vancomycin resistant and susceptible *Enterococcus faecium* in a low prevalence antimicrobial resistance setting is highly influenced by global clones.

- The incidence of VRE in Norway increased dramatically from 2010 to 2015.
- Whole genome sequence analyses of VSE*fm* bacteraemia isolates from 2008 and 2014 and a randomly selected subset of VRE recovered from 2010 to July 2015 showed that the Norwegian *E. faecium* population structure is influenced by globally prevalent STs and, in particular concurrent European CTs.
- The two major *vanB*-type clusters ST192-CT3/26 (n=113) and ST117-CT24 (n=31) mainly recovered from a single hospital carried ICE Tn*1549* that had been acquired independently.
- Variants of Tn*1549* were responsible for all *vanB*-type (*vanB2*) VRE*fm*.
- Although *vanB* was the most prevalent *van* type, *vanA* occurred in more diverse CTs.
- *vanA* gene clusters were carried on either Inc18 or RepA_N plasmids containing toxin-antitoxin systems, mostly as part of Tn*1546*-like elements.
- VRE*fs* incidence is much lower than VRE*fm* and were all *vanB*-type, of which eight were carried on Tn*1549* and four had a chromosomally integrated plasmid harbouring the *vanB1* gene cluster.
- Norwegian VRE*fm* and successful CTs have enriched virulomes compared to the more diverse VSE*fm* population, and each clone has its specific VF profile.
- Clade A isolates were more virulent and resistant compared to clade B isolates.

5 General discussion

As a leading cause of nosocomial infection, enterococci are receiving significant attention from researchers. The open pangenome and genome dynamics are important properties of *E. faecium* that facilitate the acquisition of MGEs (28). Although enterococci have been the subject of increasing genomic studies, further investigation is still required in some areas, such as the aspect of different virulomes (274). Concerning the diversity of CTs, novel *van* types or subtypes, and MGEs harbouring them, we implemented different methods and approaches to explain some important aspects of Norwegian VRE and VSE genomes. In general, the epidemiology of VRE in Norway has shown a similar trend as in other European countries.

In **paper I,** we studied Norway's first six *vanD*-type VRE isolates. Regardless of the rarity of *vanD*, it is one of the most diverse *van*-types in the case of subtypes, MGEs harbouring the gene cluster, and occurrence in the different *Enterococcus* species (106,108). In **paper I,** this diversity expanded further since we introduced a novel subtype, new GIs harbouring the *vanD* cluster, and reported it in a new *Enterococcus* species.

In **paper II**, we reported the first two *vanE*-type VRE*fs* isolates recovered from the same patient in Norway and, to our knowledge, in Europe. *vanE* is even rarer than *vanD* gene cluster. Furthermore, all the previously reported *vanE*-type VRE were recovered when WGS was not a common method in the study of VRE, and none have previously been whole genome sequenced (110,275,276). In **paper II** we sequenced and closed the genome of *vanE*-type isolates for the first time.

In **paper III**, we carried out the first comprehensive study on the genomes of enterococci recovered in Norway in a defined period and investigated the clonality, *van*-types, and MGEs of vancomycin resistant *E. faecium* and *E. faecalis*, as well as the virulome of VRE*fm* and VSE*fm*. Although the population structure and genomes of VRE*fm* are well studied (30), the genomic difference between VRE*fm* and VSE*fm* needs more attention. The implication of WGS in nosocomial pathogen studies, such as in VRE*fm,* can elucidate the local and global spread of VRE*fm* and genomic characteristics related to host specificity, hospital adaptation and resistance (277). One recommended approach in genomic studies is using closely related reference genomes (278). In **paper III,** we closed the genomes of eight VSE*fm* isolates from the dominant Norwegian STs to use as references for the VRE isolates. This approach helped identify and locate the MGEs harbouring the *vanA* and *vanB* gene clusters.

Moreover, we compared our results to the global data looking for clues to reveal the relatedness between the global and Norwegian VRE. As WGS is costly, researchers tend to sequence the most clinically important (VRE) rather than less important isolates (VSE) and underestimate VSE/VRE genomic evolution. VSE are important nosocomial pathogens that prop up further evolution of VRE (277). One of the remarkable differences that **paper III** pointed out is the difference between the virulomes in VRE*fm* and VSE*fm*.

5.1 Setting standards for quality control of assemblies

Nowadays, WGS has helped to elucidate the clonal spread and transmission routes of VRE*fm* and has become a widely used method in VRE studies (279). In bioinformatics analyses, the quality of the final assembly is essential, but still there is a lack of standard quality control parameters and thresholds. Therefore, we selected our own standards for quality control of assemblies.

Applying cut-offs to coverage and length will result in an assembly with a much lower level of contamination. We may have to discard a few genuine contigs in this process (280). However, most of such short contigs contain repetitive elements since Illumina reads are not long enough to cover their entire length (254).

A minimal level of coverage is required to achieve a reliable result in SNP calling. Illumina recommends an average of 30x genome coverage. Such a minimum cut-off can lead to confident SNP scores and support the genomic regions with lower coverages (281). While for the accurate prediction of ARG and VF genes as well as isolate typing, 40x coverage or more is needed (261). Thus, we set a 40x cut-off for genome coverage. The genome sequence in *E. faecium* is full of repetitive elements, which can make difficulty in the assembly process. During the read assembly, repetitive elements can affect the process of building contigs (254). In the *E. faecium* genome, IS elements, repeated regions in intergenic regions, and genes with repetitive regions like *esp* are the reason for fragmented assemblies (282). The standard we used was deleting contigs shorter than 200 bp. After this modification, assemblies with 400 contigs or less were included in the study. These standards have been used in our lab and performed well in the analyses. Moreover, in **paper I**, where we used SPAdes assembler, an additional cut-off for coverage was applied, and all contigs with less than 2x coverage were removed from the final assemblies.

400 contigs could be a high number in some species with smaller genomic sizes, but this was to include as many contigs as possible with acceptable quality in the final assembly. In addition, assemblies with a genome size fluctuation ≤ 10% of the smallest and largest genome deposited at NCBI were considered in the bioinformatics analyses. We chose this range because of the dynamic and open pangenome of *E. faecium* (28) and the difference between the smallest and biggest closed genomes deposited on NCBI (2.43 Mb to 3.44 Mbp) is more than 1 million bp (40). However, in **paper III,** the genome size ranged between 2.3 to 3.1 Mbp and did not reach the upper and lower genome size values of *E. faecium* retrieved from NCBI.

5.2 Determining clonality of isolates

Assessment of relatedness between isolates is vital in infection epidemiology. High resolution typing or genomic methods like SNP-based comparison are used for this purpose. Using epidemiological data and cgMLST can help assess the relatedness between the isolates and their transmission route (243). In all three papers **(I, II, III),** we used cgMLST data to build MS trees to assess the relatedness between isolates.

In **paper I,** among the *vanD*-type VRE*fm,* we identified two unrelated CTs (allelic difference =354) in two patients. This fact confirmed the epidemiological data indicating that the two *vanD*-type isolates in the two patients were unrelated and occurred independently. The cgMLST results are concordant with SNP-based mapping methods and are of high discriminatory power for determining the relatedness between isolates (283). However, even samples with different MLST, can be closely related. This was also the situation in **paper I**, in which two isolates that recovered from the same patient with two different STs (ST17 and ST1486) had the same CT (CT3198) differing in only 7 of 1423 alleles in the cgMLST scheme. One of the allelic differences was in the *dll* gene, which is among the MLST genes scheme of *E. faecium*. The *ddl* gene in *vanD*-type strains is often mutated, and various mutations in this gene have been identified so far (106). In such cases, if you only use the seven loci of the MLST scheme, the relatedness of isolates can be overlooked.

A cgMLST scheme is only available for 24 important clinical bacterial species. *E. casseliflavus* is not among them since it is rarely recovered from human infections (284,285). Therefore, we drew a phylogenetic tree based on the core genome SNPs using the Parsnp tool to determine the relatedness between the two isolates. The two *E. casseliflavus* isolates of patient B in the **paper I** seem clonally unrelated since they clustered in two separate branches in the global tree.

Analysis of the relatedness between *vanE*-type isolates of **paper II** was challenging. In the epidemiology of infections, three variables (time, person, and place) are studied (286). Although the two *vanE*-type isolates in **paper II** were recovered 30 months apart, they were isolated from the same patient and should thus be considered epidemiologically related. cgMLST is a robust high-resolution method for assessing genetic relatedness between epidemiologically related isolates. In cases like the *vanE*-type isolates, data produced from cgMLST need supporting data and further analyses to confirm the relatedness between the isolates. Furthermore, the assessment of relatedness between two isolates based on cgMLST and variant calling distances is highly organism-specific(287). If cgMLST and variant calling cannot provide enough resolution, investigating the accessory genome and plasmid-typing may provide more discrimination(287). Thus, in assessing the relatedness between *vanE*-type isolates, we performed accessory genome comparison in addition to cgMLST analysis and variant calling.

The cgMLST analysis showed 32 allelic differences, which is higher than the threshold of allelic differences for cluster formation in *E. faecalis* by more than four times. The cgMLST scheme of *E. faecalis* has more genes than *E. faecium*, while the threshold for determining CTs is only seven allelic differences(231,232). This can be due to the stable large core genome in *E. faecalis* (35).

The E1 isolate has a 125-fold higher mutation rate than E2, which complicated the interpretation of variant calling. Additionally, the time gap between E1 and E2 isolation was a possible reason for the differences in their accessory genomes which is reflected in different MGEs and plasmid profiles. Together, the genomic and epidemiological data suggest relatedness between E1 and E2, but the level of relatedness and the putative origin of the isolates need more supporting data and analyses.

In paper **III,** we discussed the clonality at two levels, MLST and cgMLST data. cgMLST analysis and its resulting MS tree give higher resolution, but the method is still not widely used in enterococcal studies worldwide. Therefore, we need a more widespread method to compare our results to the global data. *In silico* MLST still is a robust method and will be used in the future for broad characteristics of isolates, as it provides data for population structure using a global nomenclature and a large amount of MLST data has already been stored (288).

The comparison between MLST and cgMLST data revealed that 80% of the isolates belong to the globally dominant STs, but in the cgMLST level, only 42% (206/490) of them were from globally prevalent CTs. This is partly due to cgMLST not being a widely used typing method in enterococcal research. Moreover, the emergence of *E. faecium* clones at the local level is probably also affected by the dynamic nature of their genome (289).

5.3 The prevalence of globally dominant STs and CTs among the Norwegian VRE*fm*

The major STs (ST192, ST117, ST203, and ST80) identified in **paper III** have been dominant and responsible for many hospital associated outbreaks in different European countries (58,70,271,290–293). Moreover, the main VRE*fm* CTs identified in the cgMLST analyses have been reported from other European countries such as Denmark, the Netherlands, and Germany (184,291,293–296). The rest of the CTs were non-prevalent or novel CTs varying in size (from 21 to 1 isolate).

ST192 is one of the most important global STs associated with AMR. One of the earliest reports of this ST was among VRE isolates in Korea between 1998-2004, dominated by *vanA*. Only a single *vanB*-type VRE isolate was identified in that study, but the article did not specify its ST (297). ST192 has been one of Germany's VRE and LRE-associated STs since 2003 (298). Since then, ST192 has continued causing *vanB*-type VRE outbreaks in several German hospitals (70,299,300), while in Denmark, most of the ST192 isolates recovered between 2005-15 were *vanA*-type (294). In Sweden, ST192 was responsible for most *vanB*-type VRE between 2007 to 2011(173). In Norway, the first reported *vanB*-type ST192 isolate was from 2010, recovered from a large hospital outbreak (**paper III**).

The second most prevalent ST in Norway was the *vanB*-type outbreak associated ST117 (**paper III**), another globally dominant ST associated with both *vanA*- and *vanB*-type VRE. In Germany, it was the main VRE-mediating ST in the 1990s. In 2008, 80% of ST117 VRE isolates in Charité – Berlin University of Medicine were *vanA* but this fell to 6% after a decade (2018) (70,301). ST80 is another worldwide prevalent ST that was responsible for the largest VRE outbreak in Germany, resulting in 2900 *vanB*-type VRE isolates (70). ST203, ST17, and ST18 which followed as the next dominant STs were among the most prevalent STs in Europe between 2000 and 2009, but their prevalence began to decrease after a decade (2010-19) (70). ST203 and ST80 are mainly associated with *vanA*-type VRE(292,302).

Identification of reservoirs and transmission routes is one of the most important aspects in the prevention and understanding of the spread of VRE. Hospital outbreaks could be due to clonal expansion of an established VRE and/or due to newly acquired *van*-carrying MGE and HGT (291). Analyses and comparisons of MGEs in the VRE isolates can help to elucidate and understand the spread of VRE (291). Such analyses can provide additional information besides clonality in clarifying the spread of VRE. In **paper III,** with one exception, among each *vanB*-type CT, the Tn*1549* ICE was identical and inserted at the exact same location.

5.4 Relatedness between the Norwegian and global VRE

As outlined, globally dominant STs were prevalent among the Norwegian VRE and VSE isolates. At the CT level, only the outbreak related CTs have been reported to form important clusters elsewhere. A total of 42% of all *E. faecium* isolates in **paper III** were from known CTs, while this proportion in the VRE*fm* was 72%. The prevalence of VRE in Norway is low. The VRE problem in Norway has mainly been a local problem, with some intra-regional spread. Between 2006 to 2017, only 6% of the VRE recovered in Norway were reported to be acquired abroad (303). Genetic events like acquisition and loss of ARG via HGT and recombination in the core genomes are the main modes of genomic evolution in clade A1 *E. faecium* isolates that can produce new clusters. This pattern can lead to the emergence of novel lineages able to spread worldwide (289). This could be why the clonal composition of predominant STs like ST117 is highly diverse, which is reflected in their various CTs (304). Local evolution among CTs plays an important role in the epidemiology of VRE*fm*. For instance, a comparison between Irish and Danish VRE*fm* isolates revealed few overlapping isolates in clade A1 clusters, indicating the importance of local evolution in the epidemiology of VRE*fm* (184).

We compared our results to the available global data to find any close relatedness between the Norwegian and worldwide VRE CTs. As we do not have access to metadata of the VRE genomes deposited in NCBI, it is difficult to establish that two or more VRE isolates from two countries are epidemiologically closely related. For instance, in NCBI, the isolation date is presented as per year, and no more details are available (day or month). In addition, we did not include the draft genomes deposited into NCBI to avoid technical issues and reduce errors, which will consequently exclude most of the isolates.

Among global isolates genomes retrieved from NCBI some of the ST192-CT3, ST117-CT24, and ST203-CT20 were in common with the Norwegian CTs. Analyses showed that only two *vanB*-type isolates from the Netherlands (E7654 and E7663) showed high relatedness to the

Norwegian ST192-CT3. In the Dutch and Norwegian isolates, the Tn*1549* harbouring the *vanB2* gene cluster had an identical IS insertion and were integrated into the exact same insertion site. Although we considered them closely related, lacking the epidemiological data of the Dutch isolates was an obstacle for suggesting a possible direction of spread.

5.5 *van* gene clusters associated with the Norwegian VRE

The diversity of *van* gene clusters in enterococci has reached ten *van*-types (74). Among VRE, *vanA* and *vanB* are still the most predominant *van*-types (274), and the main reservoir of *vanA*- and *vanB*-type is *E. faecium.* In addition to *vanA* and *vanB*, Although with a lower prevalence, *vanC* and *vanM* are able to cause outbreaks, while other *van*-types are reported sporadically (70,81,132). In Norway, the main *van*-types were *vanB* and *vanA* (303). The *vanA* gene cluster has been widely spread in Europe, the USA, South America, Korea, and Africa, while *vanB* has been most prevalent in some European countries and Australian hospitals as in Norway (128,277,305,306).

In **paper I**, six *vanD*-type VRE from two enterococcal species were isolated from two separate cases. *vanD* is a relatively rare *van* gene cluster, mainly found on GIs integrated into the chromosome, but recently also reported on a highly conjugative pEF-D plasmid (109,187), homologous with the Inc18 family plasmid pMG1(109). Since 2018, *vanD*-type VRE have only been reported from the Netherlands and Japan (108,109,307,308). In Norway *vanD*-type VRE are rare and occur sporadically, similar to what was observed in the Netherlands, Germany and Japan (70,108,109,307). Localisation of the *vanD* gene cluster on a more successful MGE like pEF-D plasmid may increase further spread of it.

In **paper II**, we reported *vanE*, one of the rarest *van*-types for the first time in Norway and, to our knowledge, in Europe. Clinically it had been reported only in *E. faecalis*, but our BLAST search showed that all *E. caccae* genome sequences retrieved from the NCBI database contain the *vanE* gene cluster (Access: NZ_CABMMG010000001.1, NZ_KB946335, NZ_JXKJ01000001).

In **paper III**, a population study on VRE (*vanA*- and *vanB*-type) and VSE isolates, 63% (167/229) of all VRE*fm* were *vanB*-type. All the *vanB* gene clusters in VRE*fm* isolates were *vanB2*-subtype, while in VRE*fs* *vanB2* was found in 67% (8/12) and *vanB1* in the remaining VRE*fs*. Worldwide, *vanB2* is by far the most prevalent among the known subtypes of *vanB* (*vanB1-3*). *vanB1* has been reported from *E. faecium* and *E. faecalis*. In some studies, the

proportion of *vanB1* in *vanB*-type *E. faecium* is almost 5%(309,310). The *vanA*-type occurred in 27% of VRE*fm* and none of the VRE*fs* isolates (**paper III**). The prevalence of *vanA* type VRE is variable in different counties. For instance, in recent years the leading *van*-type in Australia and China are *vanA*-type (130,311). Although *vanB* was the most prevalent *van*-type in Norway, *vanA* was relatively more spread in different CTs with lower number of isolates per cluster. Even though *vanA*-type VRE is associated with more outbreaks in Norway, the number of *vanB*-type VRE isolates in the two large outbreaks of the W1 hospital changed the situation to the dominance of *vanB* (303).

5.6 The MGEs harbouring *van* gene cluster in the Norwegian VRE

It has been argued that the emergence of vancomycin resistance in enterococci is caused by the exorbitant use of vancomycin in the healthcare setting and similar glycopeptides (avoparcin) in animal farming. Such overuse of the antibiotic dramatically increased the pressure of natural selection in bacteria and promoted the emergence of antibiotic resistance mechanisms such as vancomycin resistance (312) via the acquisition of MGEs that carry ARGs (41). The involvement of enterococci in thriving as both commensal and pathogen is mainly due to their ARG and or VF harbouring MGEs. MGEs shuffling novel genes into enterococci is a main source for evolving them into multidrug-resistant pathogens (41). After the emergence of VRE, MGEs attracted more attention in research and have been studied in early enterococcal investigations (313). The *van* gene clusters are harboured on different MGEs. The most predominant MGEs carrying vancomycin resistance are plasmids, Tn*1546,* and Tn*1549*. *vanA* is mostly part of transposon Tn*1546* that is usually carried on a plasmid. In contrast, the *vanB* gene cluster is mostly harboured on the globally prevalent ICE Tn*1549* that can integrate into the genome of recipient enterococci (2). Detailed analyses of the MGEs harbouring the *van* gene cluster and their insertion site can increase the discriminatory power of closely related isolates when combined with cgMLST (293).

Most MGEs in enterococci carry IS elements (314). As MGEs may have various insertions of IS elements in different locations of their sequences, comparison of IS insertions in MGEs could be helpful and provide more data on the genetic relatedness of MGEs (315). Another approach that can be used to investigate differences between VRE isolates (*vanB*-type) are comparing coupling sequences, insertion sites and orientation of insertion of a MGE (316). The coupling sequence is a short nucleotide sequence (5- to 8-bp) inherited from the previous insertion site in the donor genome. During transposition, a staggered cut is made in the adjacent sequence to Tn*1549* resulting in a few nucleotides of one end of the insertion sequence being brought

together with one strand of the ICE into the recipient (316). Depending on the previous insertion site, the coupling sequence could differ and be used as a marker to investigate MGE exchange (313).

In contrast to the *vanA* and *vanB* studies, sporadically occurring *van*-types such as *vanD* and *vanE*, have (with some exceptions of *vanD*-types) not been investigated using WGS (108,187). Rather PCR or primer walking have previously been used to identify the *van*-type or sequence the MGE (106,275,276). In known *vanD* type VRE*fm* investigated using WGS, the integration site of GIs harbouring *vanD* gene cluster was identical and occurred in the *lysS* gene (108,187), which shows the site-specificity of their integrases. The origin of the *vanD* gene cluster is suggested to be a gut anaerobe (108), indicating that the reservoir of GIs harbouring *vanD* gene clusters is in the human gut. Some of the GIs identified in **paper I** show high level of identity to the published *vanD*-type VRE sequences (from The Netherlands and Japan). These GIs are variable in size, and transfer were not detected by filter mating (187).

In **paper II,** we reported and described Norway's first *vanE*-type VRE*fs* isolates, one of the least studied *van* types due to scarce occurrence and *vanE*-type isolates not being whole genome sequenced so far. Tn*6202* is the MGE that carries the *vanE* gene cluster in VRE*fs*(276). In the two Norwegian *vanE*-type isolates (E1, and E2), Tn*6202* is highly similar, and the only difference between them is the insertion of IS*6770* in the *vanE* gene cluster of E2 isolate. The insertion of Tn*6202* occurs in a specific location in the genomes of *E. faecalis* at 3′end of *guaA* gene (**paper II** and (276)). In addition, this insertion site was previously reported for Tn*5801* in *E. faecalis* (317), suggesting this site as a hot spot for insertions. The *guaA* gene is chromosomal in *E. faecalis* (276), but can be located on a plasmid in *E. faecium* (64). The site-specificity of the only known MGE harbouring *vanE* to a hot spot for insertion may partly explain the low prevalence of *vanE* among VRE. Tn*6202* may not always be able to integrate into a preferred insertion site when other MGEs are occupying this site or if the target site is only occasionally present, like on a plasmid.

In **paper III**, we found that the two most prominent VREf*m* clusters ST192-CT3/26 and ST117-CT24 were associated with Tn*1549* harbouring the *vanB* gene cluster. Tn*1549* is also by far the dominant *vanB*-type MGE in all the Norwegian *vanB*-type isolates in **paper III.** The sequences of Tn*1549* in 147 of 174 *vanB*-type isolates (both VRE*fm* and VRE*fs*) were identical to the Tn*1549* reference sequence (AF192329.1). While for the remaining variants of Tn*1549*, minor differences were observed, mainly due to insertions of IS elements or other genes (Figure 2 in

paper III). The IS insertions were used as epidemiological markers to study the possibility of Tn*1549* exchange between the two outbreaks of hospital W1. The comparison of IS insertions, as well as coupling sequence analyses of Tn*1549* in *vanB*-type VRE isolates of **paper III,** rejected the possibility of direct MGE exchange between ST192-CT3/26 and ST117-CT24.

As a Tn*916*-like element, the preferred insertion site for Tn*1549* is an AT-rich sequence with a limited sequence specificity (318). The insertion of Tn*1549* in different CTs of the Norwegian *vanB*-type VRE follows the same concept as they all occurred at AT-rich sequences of varying lengths. Recently a sequence pattern (TTTT-N6-AAAA) has been suggested as a target site for the insertion of Tn*1549* ICE in *E. coli*, but our findings do not support this insertion site pattern for enterococci (319). Although, in some CTs (ST117-CT24, ST203-CT3061, and ST17-CT6207), the insertion sites partially follow this sequence pattern (TTTT-N2-N6-AAAA) (Table 2 **paper III**).

In the Norwegian *vanA*-type VRE*fm*, the MGEs carrying the *vanA* gene clusters are more diverse. The *vanA* gene clusters in the Norwegian VRE were carried on various Inc18 or RepA_N plasmids. Most *vanA* gene clusters in RepA_N plasmids were part of Tn*1546*, a Tn*3* family transposon which is widely associated with *vanA* gene clusters (19). Recently, a novel method for typing Tn*1546* based on IS element insertions has been suggested. The IS elements insertion pattern of ST80 and ST202 is similar to the suggested BC6 subtype of Tn*1546* (320). The *vanA* gene clusters in ST17 and ST18 have the same IS elements at the exact same location but lack the transposase and resolvase genes of Tn*1546* (Figure 3 in **paper III**). In the *vanA*-type ST192 VRE, no similarity between the known Tn*1546* subtypes was observed, and a new variant of transposase gene larger than the normal gene in Tn*1546* was found. In ST203-CT20 VRE*fm,* an Inc18-type plasmid with multiple insertions of IS elements carries the *vanA* gene cluster. In this cluster, the transposase and resolvase genes of Tn*1546* are missing, but a similar pattern of IS element insertion to the BJ subtype of Tn*1546* was observed (320). The three main genes of Tn*552*, were also present in this plasmid. Tn*552* has been described in *S. aureus* encoding resistance to penicillin (150). Tn*552* is another Tn*3* family transposon which is rarely identified in enterococci carrying β-lactamase gene of penicillin resistance (321).

5.7 Each cluster has a specific virulome; VRE are more virulent.

The emergence of vancomycin resistance in enterococci is at least partly a response to selection pressure due to the use of glycopeptides, while the accumulation of VFs in *E. faecium* may reflect the presence of other drivers rather than antibiotics (199). The study on the virulome of

E. faecium still needs more attention. To date, a total of 30 virulence genes have been experimentally confirmed in *E. faecium* (199–204,322–324). However, still some putative VF genes are reported when examining the virulome of *E. faecium,* and some of the experimentally confirmed VFs are not included. For instance, a recently published study described ten putative VF genes of *E. faecium.* Only four of those were experimentally confirmed to be VFs (289). In another study published in 2022, only 6 out of 17 putative VF genes were among previously confirmed VFs of *E. faecium* (325). Moreover, the reference database for bacterial virulence factors (VFDB) provides a VF list and sequences for pathogenic bacteria. In the case of *E. faecium*, VFDB list contains only 16 VFs, including some putative VF genes (as of 20.11.2022) (326). Thus, to study the virulome of *E. faecium* in **paper III**, we built our own database and included only the experimentally confirmed VFs. While *E. faecium* is not considered a highly virulent *Enterococcus* species (2), the plastic genome and its ability to acquire VF-encoding MGEs can cause different VF profiles (327). Apart from *tirE1* and *tirE2,* which promote *E. faecium* survival in the blood, and *boNT/En* and *epx2,* which are exotoxins, all the confirmed VF genes are associated with adhesion and colonization (202,204,324). For instance, *ecbA, fms15, scm*, *ptsD* and *prpA* encode VFs supporting adherence and colonization(200,328), and *esp* encodes a surface protein involved in biofilm formation (329). The BoNT/En is a Botulinum Neurotoxin-like Toxin, encoded by a gene cluster, described to be located on conjugative repUS15 plasmids in an *E. faecium* isolate(324). Epx2 is a cytotoxic pore-forming toxin identified in two strains of *E. faecium* (204). Eukaryotic cell targeting toxins are seldom in *E. faecium* (enterococci), and the spread of neurotoxins (BoNT/En) or a toxin like epx2 within the broad reservoirs (human and livestock) of *E. faecium* could potentially have devastating consequences (204,324).

Our results showed that clusters have their specific VF profile that may differ in one or more VFs, and also between strains within a cluster. Interestingly, in the mixed VRE/VSE clusters, the VRE isolates may contain more VFs. Generally, nosocomial isolates (of clade A) successfully acquire more VF and AMR genes, which bring selective advantages for the isolates when adapting to a hospital environment (21). Like AMR and biocide resistance genes, VF genes associated with biofilm-formation, help enterococci to survive longer in the hospital and facilitate their spread (330). Among the Norwegian *E. faecium*, the virulence profiles follow this concept. Clinical clade A isolates are more virulent than clade B isolates. Among clade A isolates, the *vanB*-type ST192-CT3/26 was the most virulent, while non-prevalent STs isolates of both clade A and B were the least virulent.

5.8 The VRE*fs* isolates.

At least half of all enterococcal infections are caused by *E. faecalis*. However, the proportion of VRE is low among this species (2). The genome structure of *E. faecalis* differs from *E. faecium*. It has a larger core genome and is less apt to acquire AMR genes (35). In **paper III**, twelve *E. faecalis* isolates, all *vanB*-type VRE, were analysed. This amounts to 2.3% of all the isolates, and 5% of the VRE isolates in this study. This number was higher than the European VRE*fs* proportion (1.1%) during 2012-19 (50). The Norwegian VRE*fs* were also associated with globally dominant STs (ST6 and ST28).

6 Concluding remarks and future aspects

This chapter summarise key research findings in this study. Despite recent genomic studies on enterococci much is still unknown regarding the worldwide population structure of *E. faecium*. At the global scale, the situation of VRE in some countries especially in the Middle East, Southern Asia, Africa, and Latin America remains unknown. Most molecular epidemiology studies on enterococci are biased by outbreak related and multidrug resistant isolates. Additionally, the virulome studies often are not precise as they add several putative VF genes in their research and miss out confirmed VF genes.

The normalized and unbiased sample collection in our study is one of the main advantages clarifying the picture of VRE*fm* in Norway in comparison to contemporary VSE*fm*. Also, the first virulome study of *E. faecium* was carried out in which only experimentally confirmed VFs were included. Our results showed that virulome could vary in different CTs and the successful VRE clusters are more virulent. In addition, the virulome of VRE*fm* isolates were compared to VSE*fm* confirming that the clade A isolates are more enriched with VF genes compared to clade B. Including VSE*fm* isolates in the study gave additional benefits. In the comparative genomics part, VSE closed genomes were used as reference for VRE isolates.

Our study showed that the Norwegian *E. faecium* population is influenced by globally prevalent clusters, particularly European. A high relatedness between isolates from The Netherlands and Norway was observed. The Norwegian trend in the *van*-type (*vanA* and *vanB*), as well as STs, are following global trends. At the CT level, 72% of the VRE*fm* CTs are globally prevalent and circulating in European countries, while the VSE*fm* rather belong to local CTs.

For future research, we suggest including a larger number of commensal isolates in the genomic comparisons. Including commensal isolates in the virulome study could be a valuable step in understanding VFs acquisition in the enterococci and may help in understanding the hidden drivers supporting acquisition of virulence factors. Moreover, CRISPR-cas system and R-M system in enterococci are not fully understood. The data produced in this research can be used to study differences in these two systems in nosocomial and vancomycin resistant enterococci.

As the cost of WGS is rapidly decreasing and becoming more affordable, closing more genomes using long read sequencing is suggested. More specifically, this would be of benefit in the *vanA*-type VRE where the *van* gene clusters are mainly carried on plasmids with multiple repetitive elements.

References:

1. Hardie JM, Whiley RA. Classification and overview of the genera Streptococcus and Enterococcus. J Appl Microbiol. 1997;83(S1):1S-11S.
2. García-Solache M, Rice LB. The Enterococcus: a model of adaptability to its environment. Clin Microbiol Rev. 2019;32(2):1–28.
3. Vu J, Carvalho J. Enterococcus: review of its physiology, pathogenesis, diseases and the challenges it poses for clinical microbiology. Front Biol (Beijing). 2011;6(5):357.
4. Giraffa G. Enterococci from foods. FEMS Microbiol Rev. 2002 Jun 1;26(2):163–71.
5. Zommiti M, Chevalier S, Feuilloley MGJ, Connil N. Special Issue "Enterococci for Probiotic Use: Safety and Risk": Editorial. Microorganisms. 2022 Mar 12;10(3):604.
6. Schleifer KH, Kilpper-Bälz R. Transfer of Streptococcus faecalis and Streptococcus faecium to the Genus Enterococcus nom. rev. as Enterococcus faecalis comb. nov. and Enterococcus faecium comb. nov. Int J Syst Evol Microbiol. 1984;34:31–4.
7. Vancanneyt M, Devriese LA, Naser SM. Enterococcus aquimarinus sp . nov ., isolated from sea water. Int J Syst Evol. 2005;55(5):2183–7.
8. Vec PS, Devriese LA, Sedla I, Baele M, Vancanneyt M, Haesebrouck F, et al. Enterococcus haemoperoxidus sp. nov. and Enterococcus moraviensis sp. nov., isolated from water. Int J Syst Evol. 2001;51:1567–74.
9. Vec PS, Vandamme P, Bryndova H, Holochova P, Kosina M, Mas I. Enterococcus plantarum sp. nov., isolated from plants. Int J Syst Evol. 2012;62:1499–505.
10. Sedla I, Pavel S, Frolkova P, Litera I. Enterococcus alcedinis sp. nov., isolated from common kingfisher (Alcedo atthis). Int J Syst Evol. 2013;63:3069–74.
11. Fortina MG, Ricci G, Mora D, Manachini PL. Molecular analysis of artisanal Italian cheeses reveals Enterococcus italicus sp. nov. Int J Syst Evol Microbiol. 2004;54(5):1717–21.
12. Parte AC, Sardà Carbasse J, Meier-Kolthoff JP, Reimer LC, Göker M. List of Prokaryotic names with Standing in Nomenclature (LPSN) moves to the DSMZ. Int J Syst Evol Microbiol. 2020;70(11):5607–12.
13. Rosselli Del Turco E, Bartoletti M, Dahl A, Cervera C, Pericàs JM. How do I manage a patient with enterococcal bacteraemia? Clin Microbiol Infect. 2021;27(3):364–71.
14. Fiore E, van Tyne D, Gilmore MS. Pathogenicity of Enterococci. Microbiol Spectr. 2019 Jul;7(4):4.
15. Huycke MM, Sahm DF, Gilmore MS. Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. Emerg Infect Dis. 1998;4(2):239–49.
16. Kayser FH. Safety aspects of enterococci from the medical point of view. Int J Food Microbiol. 2003;88(2–3):255–62.
17. Barnes AMT, Frank KL, Dunny GM. Enterococcal Endocarditis: Hiding in Plain Sight. Front Cell Infect Microbiol. 2021 Aug 30;11:722482.
18. Ramos Y, Sansone S, Morales DK. Sugarcoating it: Enterococcal polysaccharides as key modulators of host–pathogen interactions. PLoS Pathog. 2021 Sep 9;17(9):e1009822.

19. Ramos S, Silva V, Dapkevicius MD, Igrejas G, Poeta P. Enterococci, from Harmless Bacteria to a Pathogen. Microorganisms. 2020;8(8):1118.
20. Faron ML, Ledeboer NA, Buchan BW. Resistance Mechanisms, Epidemiology, and Approaches to Screening for Vancomycin-Resistant Enterococcus in the Health Care Setting. J Clin Microbiol. 2016 Oct;54(10):2436–47.
21. Hollenbeck BL, Rice LB. Intrinsic and acquired resistance mechanisms in enterococcus. Virulence. 2012 Aug 15;3(5):421–569.
22. Vehreschild MJGT, Haverkamp M, Biehl LM, Lemmen S, Fätkenheuer G. Vancomycin-resistant enterococci (VRE): a reason to isolate? Infection. 2019;47(1):7–11.
23. European Centre for Disease Prevention and Control & World Health Organization. Regional Office for Europe. (2022). Antimicrobial resistance surveillance in Europe 2022 – 2020 data. World Health Organization. Regional Office for Europe.
24. Puchter L, Chaberny IF, Schwab F, Vonberg RP, Bange FC, Ebadi E. Economic burden of nosocomial infections caused by vancomycin-resistant enterococci. Antimicrob Resist Infect Control. 2018 Jan 5;7:1.
25. Horner C, Mushtaq S, Allen M, Hope R, Gerver S, Longshaw C, et al. Replacement of Enterococcus faecalis by Enterococcus faecium as the predominant enterococcus in UK bacteraemias. JAC Antimicrob Resist. 2021 Dec 1;3(4):dlab185.
26. Weiner LM, Webb AK, Limbago B, Dudeck MA, Patel J, Kallen AJ, et al. 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. Infect Control Hosp Epidemiol. 2016;37(11):1288–301.
27. Perio MA de, Yarnold PR, Warren J, Noskin GA. Risk Factors and Outcomes Associated With Non–Enterococcus faecalis, Non–Enterococcus faecium Enterococcal Bacteremia. Infect Control Hosp Epidemiol. 2006;27(1):28–33.
28. Bayjanov JR, Baan J, Rogers MRC, Troelstra A, Willems RJL, van Schaik W. Enterococcus faecium genome dynamics during long-term asymptomatic patient gut colonization. Microb Genom. 2019 Jul;5(7):e000277.
29. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, et al. Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. mBio. 2013 Aug 20;4(4):e00534-13.
30. van Hal SJ, Willems RJL, Gouliouris T, Ballard SA, Coque TM, Hammerum AM, et al. The interplay between community and hospital Enterococcus faecium clones within health-care settings: a genomic analysis. Lancet Microbe. 2022;3(2):E133–41.
31. Guzman Prieto AM, van Schaik W, Rogers MRC, Coque TM, Baquero F, Corander J, et al. Global Emergence and Dissemination of Enterococci as Nosocomial Pathogens: Attack of the Clones? Front Microbiol. 2016;7:788.
32. Belloso Daza MV, Cortimiglia C, Bassi D, Cocconcelli PS. Genome-based studies indicate that the Enterococcus faecium Clade B strains belong to Enterococcus lactis species and lack of the hospital infection associated markers. Int J Syst Evol Microbiol. 2021;71(8):004948.

33. Willems RJL, Top J, Schaik W van, Leavis H, Bonten M, Sirén J, et al. Restricted gene flow among hospital subpopulations of Enterococcus faecium. mBio. 2012;3(4):e00151-12.
34. Feil EJ, Li BC, Aanensen DM, Hanage WP, Spratt BG. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. J Bacteriol. 2004 Mar;186(5):1518–30.
35. Pöntinen AK, Top J, Arredondo-Alonso S, Tonkin-Hill G, Freitas AR, Novais C, et al. Apparent nosocomial adaptation of Enterococcus faecalis predates the modern hospital era. Nat Commun. 2021;12(1):1523.
36. Palmer KL, Gilmore MS. Selfish Elements and Self-Defense in the Enterococci. In: Genome Plasticity and Infectious Diseases. 2012. p. 125–40. (Wiley Online Books).
37. Zhong Z, Zhang W, Song Y, Liu W, Xu H, Xi X, et al. Comparative genomic analysis of the genus Enterococcus. Microbiol Res. 2017;196:95–105.
38. Vernikos G, Medini D, Riley DR, Tettelin H. Ten years of pan-genome analyses. Curr Opin Microbiol. 2015;23:148–54.
39. de Been M, van Schaik W, Cheng L, Corander J, Willems RJ. Recent recombination events in the core genome are associated with adaptive evolution in Enterococcus faecium. Genome Biol Evol. 2013;5(8):1524–35.
40. NCBI. Genome Assembly and Annotation report - Enterococcus faecium [Internet]. Available from: https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/871/
41. Johnson CN, Sheriff EK, Duerkop BA, Chatterjee A. Let Me Upgrade You: Impact of Mobile Genetic Elements on Enterococcal Adaptation and Evolution. J Bacteriol. 2021 Oct 12;203(21):e0017721.
42. Faure G, Makarova KS, Koonin E v. CRISPR–Cas: Complex Functional Networks and Multiple Roles beyond Adaptive Immunity. J Mol Biol. 2019;431(1):3–20.
43. Palmer KL, Gilmore MS, Richard. Losick. Multidrug-Resistant Enterococci Lack CRISPR-cas. mBio. 2010 Jun 23;1(4):e00227-10.
44. Vasu K, Nagaraja V. Diverse Functions of Restriction-Modification Systems in Addition to Cellular Defense. Microbiology and Molecular Biology Reviews. 2013 Mar;77(1):53–72.
45. Huo W, Adams HM, Trejo C, Badia R, Palmer KL. A Type I Restriction-Modification System Associated with Enterococcus faecium Subspecies Separation. Appl Environ Microbiol. 2019 Jan 15;85(2):e02174-18.
46. Arredondo-Alonso S, Top J, McNally A, Puranen S, Pesonen M, Pensar J, et al. Plasmids Shaped the Recent Emergence of the Major Nosocomial Pathogen Enterococcus faecium. mBio. 2020 May 28;11(1):e03284-19.
47. Willemse N, Schultsz C. Distribution of Type I Restriction–Modification Systems in Streptococcus suis: An Outlook. Pathogens. 2016;5(4):62.
48. Werner G, Coque TM, Franz CMAP, Grohmann E, Hegstad K, Jensen L, et al. Antibiotic resistant enterococci-Tales of a drug resistance gene trafficker. Int J Med Microbiol. 2013;303(6):360–79.
49. Hegstad K, Mikalsen T, Coque TM, Werner G, Sundsfjord A. Mobile genetic elements and their contribution to the emergence of antimicrobial resistant Enterococcus faecalis and Enterococcus faecium. Clin Microbiol Infect. 2010;16(6):541–54.

50. Ayobami O, Willrich N, Reuss A, Eckmanns T, Markwart R. The ongoing challenge of vancomycin-resistant Enterococcus faecium and Enterococcus faecalis in Europe: an epidemiological analysis of bloodstream infections. Emerg Microbes Infect. 2020 Jan 1;9(1):1180–93.
51. Khan Ayesha, William RM, Axell-House Dierdre, Jose MM, Cesar AA, Romney MH. Antimicrobial Susceptibility Testing for Enterococci. J Clin Microbiol. 2022 Jun 23;0(0):e00843-21.
52. Said MS, Tirthani E, Lesho E. Enterococcus Infections. Rochester General Hospital: StatPearls Publishing, Treasure Island (FL); 2022.
53. Turner AM, Lee JYH, Gorrie CL, Howden BP, Carter GP. Genomic Insights Into Last-Line Antimicrobial Resistance in Multidrug-Resistant Staphylococcus and Vancomycin-Resistant Enterococcus. Front Microbiol. 2021;12:637656.
54. Galloway-Peña JR, Nallapareddy SR, Arias CA, Eliopoulos GM, Murray BE. Analysis of Clonality and Antibiotic Resistance among Early Clinical Isolates of Enterococcus faecium in the United States. J Infect Dis. 2009;200(10):1566–73.
55. Cui P, Feng L, Zhang L, He J, An T, Fu X, et al. Antimicrobial Resistance, Virulence Genes, and Biofilm Formation Capacity Among Enterococcus species From Yaks in Aba Tibetan Autonomous Prefecture, China. Front Microbiol . 2020;11:1250.
56. Montealegre Maria Camila, Roh Jung Hyeob, Meredith R, Davlieva Milya G, Singh Kavindra V, Shamoo Yousif, et al. Differential Penicillin-Binding Protein 5 (PBP5) Levels in the Enterococcus faecium Clades with Different Levels of Ampicillin Resistance. Antimicrob Agents Chemother. 2022 May 19;61(1):e02034-16.
57. Rice Louis B, Bellais Samuel, L. CL, Hutton-Thomas Rebecca, Bonomo Robert A, Caspers Patrick, et al. Impact of Specific pbp5 Mutations on Expression of β-Lactam Resistance in Enterococcus faecium. Antimicrob Agents Chemother. 2004 Aug 1;48(8):3028–32.
58. Novais C, Tedim AP, Lanza VF, Freitas AR, Silveira E, Escada R, et al. Co-diversification of Enterococcus faecium Core Genomes and PBP5: Evidences of pbp5 Horizontal Transfer. Front Microbiol. 2016;7(6):1581.
59. Moussa AA, Md Nordin AF, Hamat RA, Jasni AS. High Level Aminoglycoside Resistance And Distribution Of The Resistance Genes In Enterococcus faecalis And Enterococcus faecium From Teaching Hospital In Malaysia. Infect Drug Resist. 2019;12:3269–74.
60. Ramirez MS, Tolmasky ME. Aminoglycoside modifying enzymes. Drug Resist Updat. 2010;13(6):151–71.
61. Michael S Gilmore, Clewell DB, Ike Y, Shankar N. Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Massachusetts Eye and Ear Infirmary; 2014.
62. Egan SA, Shore AC, O'Connell B, Brennan GI, Coleman DC. Linezolid resistance in Enterococcus faecium and Enterococcus faecalis from hospitalized patients in Ireland: high prevalence of the MDR genes optrA and poxtA in isolates with diverse genetic backgrounds. J Antimicrob Chemother. 2020 Jul 1;75(7):1704–11.
63. Ruiz-Ripa L, Feßler AT, Hanke D, Eichhorn I, Azcona-Gutiérrez JM, Pérez-Moreno MO, et al. Mechanisms of Linezolid Resistance Among Enterococci of Clinical Origin

in Spain-Detection of optrA- and cfr(D)-Carrying E. faecalis. Microorganisms. 2020;8(8):1155.

64. Guerin F, Sassi M, Dejoies L, Zouari A, Schutz S, Potrel S, et al. Molecular and functional analysis of the novel cfr(D) linezolid resistance gene identified in Enterococcus faecium. J Antimicrob Chemother. 2020 Jul 1;75(7):1699–703.
65. Crowe-McAuliffe C, Murina V, Turnbull KJ, Huch S, Kasari M, Takada H, et al. Structural basis for PoxtA-mediated resistance to phenicol and oxazolidinone antibiotics. Nat Commun. 2022;13(1):1860.
66. Mendes RE, Deshpande L, Streit JM, Sader HS, Castanheira M, Hogan PA, et al. ZAAPS programme results for 2016: an activity and spectrum analysis of linezolid using clinical isolates from medical centres in 42 countries. J Antimicrob Chemother. 2018 Jul 1;73(7):1880–7.
67. Wilhelm MP. Vancomycin. Mayo Clin Proc. 1991 Nov;66(11):1165–70.
68. Butler MS, Hansford KA, Blaskovich MAT, Halai R, Cooper MA. Glycopeptide antibiotics: Back to the future. J Antibiot (Tokyo). 2014;67(9):631–44.
69. Rubinstein E, Keynan Y. Vancomycin Revisited – 60 Years Later. Front Public Health. 2014;2(1):217.
70. Werner G, Neumann B, Weber RE, Kresken M, Wendt C, Bender JK, et al. Thirty years of VRE in Germany – "expect the unexpected": The view from the National Reference Centre for Staphylococci and Enterococci. Drug Resist Updat. 2020;53:100732.
71. Schouten MA, Hoogkamp-Korstanje JAA, Meis JFG, Voss A, Group and the EVRES. Prevalence of Vancomycin-Resistant Enterococci in Europe. Eur J Clin Microbiol Infect Dis. 2000;19(11):816–22.
72. Top J, Willems R, Bonten M. Emergence of CC17 Enterococcus faecium: from commensal to hospital-adapted pathogen. FEMS Immunol Med Microbiol. 2008 Apr 1;52(3):297–308.
73. Kühn I, Iversen A, Finn M, Greko C, Burman LG, Blanch AR, et al. Occurrence and Relatedness of Vancomycin-Resistant Enterococci in Animals, Humans, and the Environment in Different European Regions. Appl Environ Microbiol. 2005 Sep;71(9):5383–90.
74. Xavier BB, Coppens J, de Koster S, Rajakani SG, van Goethem S, Mzougui S, et al. Novel vancomycin resistance gene cluster in Enterococcus faecium ST1486, Belgium, June 2021. Euro Surveill. 2021;26(36):2100767.
75. Sassi M, Guérin F, Lesec L, Isnard C, Fines-Guyon M, Cattoir V, et al. Genetic characterization of a VanG-type vancomycin-resistant Enterococcus faecium clinical isolate. J Antimicrob Chemother. 2018 Apr 1;73(4):852–5.
76. Matsumoto T, Muratani T, Matsumoto T, Okada K, Shiraishi M, Hayashida T, et al. No regional spread of vancomycin-resistant enterococci with vanA or vanB in Kitakyushu, Japan. J Infect Chemother. 2004;10(6):331–4.
77. Dahl KH, Lundblad EW, Røkenes TP, Olsvik Ø, Sundsfjord A. Genetic linkage of the vanB2 gene cluster to Tn5382 in vancomycin-resistant enterococci and characterization of two novel insertion sequences. Microbiology (N Y). 2000;146(6):1469–79.

78. Werner G. . Current Trends of Emergence and Spread of Vancomycin-Resistant Enterococci. In: Pana, M. , editor. Antibiotic Resistant Bacteria - A Continuous Challenge in the New Millennium [Internet]. London: IntechOpen; 2012 [cited 2022 Dec 16]. Available from: https://www.intechopen.com/chapters/34696 doi: 10.5772/29791.
79. Patti GJ, Kim SJ, Schaefer J. Characterization of the peptidoglycan of vancomycin-susceptible Enterococcus faecium. Biochemistry. 2008;47(32):8378–85.
80. Munita JM, Arias CA, Unit AR, Santiago A de. Mechanisms of Antibiotic Resistance. Microbiol Spectr. 2016;4(2):1–37.
81. Ahmed MO, Baptiste KE. Vancomycin-Resistant Enterococci: A Review of Antimicrobial Resistance Mechanisms and Perspectives of Human and Animal Health. Microb Drug Resist. 2018;24(5):590–606.
82. Mcstrother. vancomycin resistance [Internet]. 2011. Available from: https://en.wikipedia.org/wiki/User:Mcstrother#/media/File:Vancomycin_resistance.svg
83. Courvalin P. Vancomycin resistance in Gram-positive cocci. Clin Infect Dis. 2006;42(Supplement 1):25–34.
84. Haenni M, Saras E, Châtre P, Meunier D, Martin S, Lepage G, et al. vanA in Enterococcus faecium, Enterococcus faecalis, and Enterococcus casseliflavus Detected in French Cattle. Foodborne Pathog Dis. 2009;6(9):1107–11.
85. Eshaghi A, Shahinas D, Li A, Kariyawasam R, Banh P, Desjardins M, et al. Characterization of an Enterococcus gallinarum Isolate Carrying a Dual vanA and vanB Cassette. J Clin Microbiol. 2015;53(7):2225–9.
86. Biavasco F, Foglia G, Paoletti C, Zandri G, Magi G, Guaglianone E, et al. VanA-Type Enterococci from Humans, Animals, and Food: Species Distribution, Population Structure, Tn1546 Typing and Location, and Virulence Determinants. Appl Environ Microbiol. 2007 May 15;73(10):3307–19.
87. Sharifi-Rad M, Shadanpour S, van Belkum A, Soltani A, Sharifi-Rad J. First case of vanA-positive Enterococcus mundtii in human urinary tract infection in Iran. New Microbes New Infect. 2016;11:68–70.
88. Liassine N, Frei R, Jan I, Auckenthaler R. Characterization of Glycopeptide-Resistant Enterococci from a Swiss Hospital. J Clin Microbiol. 1998 Jul;36(7):1853–8.
89. Reynolds PE, Arias CA, Courvalin P. Gene vanXYC encodes d,d -dipeptidase (VanX) and d,d-carboxypeptidase (VanY) activities in vancomycin-resistant Enterococcus gallinarum BM4174. Mol Microbiol. 1999;34(2):341–9.
90. Arthur M, Depardieu F, Molinas C, Reynolds P, Courvalin P. The vanZ gene of Tn1546 from enterococcus faecium BM4147 confers resistance to teicoplanin. Gene. 1995;154(1):87–92.
91. Arthur M, Molinas C, Courvalin P. The VanS-VanR two-component regulatory system controls synthesis of depsipeptide peptidoglycan precursors in Enterococcus faecium BM4147. J Bacteriol. 1992;174(8):2582–91.
92. Arthur M, Reynolds P, Courvalin P. Glycopeptide resistance in enterococci. Trends Microbiol. 1996;4(10):401–7.

93. Reid KC, Cockerill III FR, Patel R. Clinical and Epidemiological Features of Enterococcus casseliflavus/flavescens and Enterococcus gallinarum Bacteremia: A Report of 20 Cases. Clin Infect Dis. 2001 Jun 1;32(11):1540–6.
94. Bourdon N, Fines-Guyon M, Thiolet JM, Maugat S, Coignard B, Leclercq R, et al. Changing trends in vancomycin-resistant enterococci in French hospitals, 2001–08. J Antimicrob Chemother. 2011 Apr 1;66(4):713–21.
95. Ballard SA, Pertile KK, Lim M, Johnson PDR, Grayson ML. Molecular characterization of vanB elements in naturally occurring gut anaerobes. Antimicrob Agents Chemother. 2005 May;49(5):1688–94.
96. Dahl KH, Simonsen GS, Olsvik Ø, Sundsfjord A. Heterogeneity in the vanB gene cluster of genomically diverse clinical strains of vancomycin-resistant enterococci. Antimicrob Agents Chemother. 1999;43(5):1105–10.
97. Patel R, Uhl JR, Kohner P, Hopkins MK, Steckelberg JM, Kline B, et al. DNA Sequence Variation within *vanA* , *vanB* , *vanC-1* , and *vanC-2/3* Genes of Clinical *Enterococcus* Isolates. Antimicrob Agents Chemother. 1998 Jan;42(1):202–5.
98. Garnier F, Taourit S, Glaser P, Courvalin P, Galimand M. Characterization of transposon Tn1549, conferring VanB-type resistance in Enterococcus spp. Microbiology (N Y). 2000;146(6):1481–9.
99. Reynolds Peter E, Courvalin Patrice. Vancomycin Resistance in Enterococci Due to Synthesis of Precursors Terminating in d-Alanyl-d-Serine. Antimicrob Agents Chemother. 2005 Jan 1;49(1):21–5.
100. Watanabe S, Kobayashi N, Quiñones D, Hayakawa S, Nagashima S, Uehara N, et al. Genetic Diversity of the Low-Level Vancomycin Resistance Gene vanC-2/vanC-3 and Identification of a Novel vanC Subtype (vanC-4) in Enterococcus casseliflavus. Microb Drug Resist. 2009;15(1):1–9.
101. Schwaiger K, Bauer J, Hörmansdorfer S, Mölle G, Preikschat P, Kämpf P, et al. Presence of the Resistance Genes vanC1 and pbp5 in Phenotypically Vancomycin and Ampicillin Susceptible Enterococcus faecalis. Microb Drug Resist. 2012;18(4):434–9.
102. Sun M, Wang Y, Chen Z, Zhu X, Tian L, Sun Z. The first report of the $vanC_1$ gene in Enterococcus faecium isolated from a human clinical specimen. Mem Inst Oswaldo Cruz. 2014;109(6):712–715.
103. Moura TM de, Cassenego APV, Campos FS, Ribeiro AML, Franco AC, D'Azevedo PA, et al. Detection of vanC1 gene transcription in vancomycin-susceptible Enterococcus faecalis. Mem Inst Oswaldo Cruz. 2013 Jun;108(4):453–6.
104. Dutta I, Reynolds PE. The vanC-3 vancomycin resistance gene cluster of Enterococcus flavescens CCM 439. J Antimicrob Chemother. 2003 Mar 1;51(3):703–6.
105. Arias CA, Weisner J, Blackburn JM, Reynolds PE. Serine and alanine racemase activities of VanT: a protein necessary for vancomycin resistance in Enterococcus gallinarum BM4174. Microbiology (N Y). 2000;146(7):1727–34.
106. Depardieu F, Foucault M, Bell J, Dubouix A, Guibert M, Lavigne J, et al. New Combinations of Mutations in VanD-Type Vancomycin-Resistant Enterococcus faecium, Enterococcus faecalis, and Enterococcus avium Strains. Antimicrob Agents Chemother. 2009;53(5):1952–63.

107. Domingo MC, Huletsky A, Giroux R, Picard FJ, Bergeron MG. vanD and vanG-Like Gene Clusters in a Ruminococcus Species Isolated from Human Bowel Flora. Antimicrob Agents Chemother. 2007;51(11):4111–7.
108. Top J, Sinnige JC, Brouwer EC, Werner G, Corander J, Severin JA, et al. Identification of a novel genomic island associated with VanD-type vancomycin resistance in six Dutch vancomycin-resistant *Enterococcus faecium* isolates. Antimicrob Agents Chemother. 2018;62(3):1–7.
109. Terada S, Harada T, Yokota M, Tsuchiya T, Adachi K, Asaka T, et al. First isolation and characterization of vancomycin-resistant *Enterococcus faecium* harboring *vanD5* gene cluster recovered from a 79-year-old female inpatient in Japan. Diagn Microbiol Infect Dis. 2019;95:1–6.
110. Abadía Patiño Lorena, Courvalin Patrice, Perichon Bruno. vanE Gene Cluster of Vancomycin-Resistant Enterococcus faecalis BM4405. J Bacteriol. 2002 Dec 1;184(23):6457–64.
111. Berthet N, Périchon B, Mazuet C, Chapetón-Montes D, Bouchier C, Bouvet P, et al. A vanG-type locus in Clostridium argentinense. J Antimicrob Chemother. 2015 Jul 1;70(7):1942–5.
112. David V, Bozdogan B, Mainardi JL, Legrand R, Gutmann L, Leclercq R. Mechanism of Intrinsic Resistance to Vancomycin in *Clostridium innocuum* NCIB 10674. J Bacteriol. 2004 Jun;186(11):3415–22.
113. McKessar SJ, Berry AM, Bell JM, Turnidge JD, Paton JC. Genetic Characterization of *vanG* , a Novel Vancomycin Resistance Locus of *Enterococcus faecalis*. Antimicrob Agents Chemother. 2000 Nov;44(11):3224–8.
114. David A Boyd , Tim Du, Romeo Hizon, Brynn Kaplen, Travis Murphy, Shaun Tyler, Shirley Brown, Frances Jamieson, Karl Weiss MRM. VanG-Type Vancomycin-Resistant Enterococcus faecalis Strains Isolated in Canada. Antimicrob Agents Chemother. 2006 Jun 1;50(6):2217–21.
115. Boyd DA, Willey BM, Fawcett D, Gillani N, Mulvey MR. Molecular characterization of Enterococcus faecalis N06-0364 with low-level vancomycin resistance harboring a novel D-Ala-D-Ser gene cluster, vanL. Antimicrob Agents Chemother. 2008 Jul;52(7):2667–72.
116. Xu X, Lin D, Yan G, Ye X, Wu S, Guo Y, et al. *vanM* , a New Glycopeptide Resistance Gene Cluster Found in *Enterococcus faecium*. Antimicrob Agents Chemother. 2010 Nov;54(11):4643–7.
117. Sun L, Chen Y, Hua X, Chen Y, Hong J, Wu X, et al. Tandem amplification of the vanM gene cluster drives vancomycin resistance in vancomycin-variable enterococci. J Antimicrob Chemother. 2020 Feb 1;75(2):283–91.
118. Teo JWP, Krishnan P, Jureen R, Lin RTP. Detection of an unusual van genotype in a vancomycin-resistant Enterococcus faecium hospital isolate. J Clin Microbiol. 2011;49(12):4297–8.
119. Lebreton F, Depardieu F, Bourdon N, Fines-Guyon M, Berger P, Camiade S, et al. d-Ala-d-Ser VanN-Type Transferable Vancomycin Resistance in Enterococcus faecium. Antimicrob Agents Chemother. 2011 Oct;55(10):4606–12.

120. Nomura T, Tanimoto K, Shibayama K, Arakawa Y, Fujimoto S, Ike Y, et al. Identification of VanN-type vancomycin resistance in an Enterococcus faecium isolate from chicken meat in Japan. Antimicrob Agents Chemother. 2012 Dec;56(12):6389–92.
121. Lebreton F, Valentino MD, Schaufler K, Earl AM, Cattoir V, Gilmore MS. Transferable vancomycin resistance in clade B commensal-type Enterococcus faecium. J Antimicrob Chemother. 2018 Jun 1;73(6):1479–86.
122. WHO. Disease outbreaks [Internet]. Available from: http://www.emro.who.int/health-topics/disease-outbreaks/index.html
123. Reintjes R, Zanuzdana A. (2009). Outbreak Investigations. In: Krämer, A., Kretzschmar, M., Krickeberg, K. (eds) Modern Infectious Disease Epidemiology. Statistics for Biology and Health. Springer, New York, NY. https://doi.org/10.1007/978-0-387-93835-6_9.
124. Friedrich AW. Control of hospital acquired infections and antimicrobial resistance in Europe: the way to go. Wien Med Wochenschr. 2019;169(1):25–30.
125. Weterings V, van Oosten A, Nieuwkoop E, Nelson J, Voss A, Wintermans B, et al. Management of a hospital-wide vancomycin-resistant Enterococcus faecium outbreak in a Dutch general hospital, 2014–2017: successful control using a restrictive screening strategy. Antimicrob Resist Infect Control. 2021;10(1):38.
126. Eisner A, Feierl G, Gorkiewicz G, Dieber F, Kessler HH, Marth E, et al. High Prevalence of VanA-Type Vancomycin-Resistant Enterococci in Austrian Poultry. Appl Environ Microbiol. 2005 Oct;71(10):6407–9.
127. Correa-Martínez CL, Jurke A, Schmitz J, Schaumburg F, Kampmeier S, Mellmann A. Molecular Epidemiology of Vancomycin-Resistant Enterococci Bloodstream Infections in Germany: A Population-Based Prospective Longitudinal Study. Microorganisms. 2022;10(1):130.
128. Zhou X, Willems RJL, Friedrich AW, Rossen JWA, Bathoorn E. Enterococcus faecium: from microbiological insights to practical recommendations for infection control and diagnostics. Antimicrob Resist Infect Control. 2020;9(1):130.
129. Andersson P, Beckingham W, Gorrie CL, Kennedy K, Daveson K, Ballard SA, et al. Vancomycin-resistant Enterococcus (VRE) outbreak in a neonatal intensive care unit and special care nursery at a tertiary-care hospital in Australia-A retrospective case-control study. Infect Control Hosp Epidemiol. 2019;40(5):551–8.
130. Pratama R, Beukers AG, McIver CJ, Keighley CL, Taylor PC, van Hal SJ. A vanA vancomycin-resistant Enterococcus faecium ST80 outbreak resulting from a single importation event. J Antimicrob Chemother. 2022 Jan 1;77(1):31–7.
131. Hughes A, Ballard S, Sullivan S, Marshall C. An outbreak of vanA vancomycin-resistant Enterococcus faecium in a hospital with endemic vanB VRE. Infect Dis Health. 2019;24(2):82–91.
132. Zhou W, Zhou H, Sun Y, Gao S, Zhang Y, Cao X, et al. Characterization of clinical enterococci isolates, focusing on the vancomycin-resistant enterococci in a tertiary hospital in China: based on the data from 2013 to 2018. BMC Infect Dis. 2020;20(1):356.

133. Sami H, Singh A, Ahmed S, Shahid M. Emergence of linezolid resistance in Enterococci: Prevalent genotypes and resistance pattern in vancomycin-resistant Enterococci in a North-Indian tertiary care hospital. N Z J Med Lab Sci. 74(1):27–30.
134. Sivaradjy M, Gunalan A, Priyadarshi K, Madigubba H, Rajshekar D, AS S. Increasing Trend of Vancomycin-resistant Enterococci Bacteremia in a Tertiary Care Hospital of South India: A Three-year Prospective Study. Indian J Crit Care Med. 2021;28(8):881–5.
135. Khudaier BY, Tewari R, Shafiani S, Sharma M, Emmanuel R, Sharma M, et al. Epidemiology and molecular characterization of vancomycin resistant Enterococci isolates in India. Scand J Infect Dis. 2007 Jan 1;39(8):662–70.
136. CDC. CDC's Antibiotic Resistance Threats in the United States, 2019 [Internet]. 2019. Available from: https://www.cdc.gov/drugresistance/biggest-threats.html#van
137. Binda E, Marinelli F, Marcone GL. Old and New Glycopeptide Antibiotics: Action and Resistance. Antibiotics (Basel). 2014 Nov 4;3(4):572–94.
138. Waglechner N, McArthur AG, Wright GD. Phylogenetic reconciliation reveals the natural history of glycopeptide antibiotic biosynthesis and resistance. Nat Microbiol. 2019;4(11):1862–71.
139. Yushchuk O, Binda E, Marinelli F. Glycopeptide Antibiotic Resistance Genes: Distribution and Function in the Producer Actinomycetes. Front Microbiol. 2020;11(17):1173.
140. Marshall CG, Lessard IA, Park I, Wright GD. Glycopeptide antibiotic resistance genes in glycopeptide-producing organisms. Antimicrob Agents Chemother. 1998 Sep;42(9):2215–20.
141. Marshall CG, Broadhead G, Leskiw BK, G D. Wright. D-Ala-D-Ala ligases from glycopeptide antibiotic-producing organisms are highly homologous to the enterococcal vancomycin-resistance ligases VanA and VanB. Proc Natl Acad Sci U S A. 1997 Jun 10;94(12):6480–3.
142. Rodríguez-Beltrán J, DelaFuente J, León-Sampedro R, MacLean RC, San Millán Á. Beyond horizontal gene transfer: the role of plasmids in bacterial evolution. Nat Rev Microbiol. 2021;19(6):347–59.
143. Arnold BJ, Huang IT, Hanage WP. Horizontal gene transfer and adaptive evolution in bacteria. Nat Rev Microbiol. 2022;20(4):206–18.
144. Husnik F, McCutcheon JP. Functional horizontal gene transfer from bacteria to eukaryotes. Nat Rev Microbiol. 2018;16(2):67–79.
145. Boto L. Horizontal gene transfer in evolution: facts and challenges. Proceedings of the Royal Society B: Biological Sciences. 2010;277(1683):819–27.
146. Dunning Hotopp JC. Horizontal gene transfer between bacteria and animals. Trends Genet. 2011;27(4):157–63.
147. Griffith F. The Significance of Pneumococcal Types. J Hyg (Lond). 1928;27(2):113–59.
148. Hotchkiss RD. Transfer of penicillin resistance in pneumococci by the desoxyribonucleate derived from resistant cultures. Cold Spring Harb Symp Quant Biol. 1951;16:7451–61.

149. von Wintersdorff CJH, Penders J, van Niekerk JM, Mills ND, Majumder S, van Alphen LB, et al. Dissemination of antimicrobial resistance in microbial ecosystems through horizontal gene transfer. Front Microbiol. 2016;7(FEB):1–10.
150. Partridge SR, Kwong SM, Firth N, Jensen SO. Mobile Genetic Elements Associated with Antimicrobial Resistance. Clin Microbiol Rev. 2018 Oct;31(4).
151. Frost LS, Leplae R, Summers AO, Toussaint A. Mobile genetic elements: the agents of open source evolution. Nat Rev Microbiol. 2005;3(9):722–32.
152. Courvalin PM, Carlier C, Chabbert YA. Plasmid-linked tetracycline and erythromycin resistance in group D "streptococcus." Ann Inst Pasteur (Paris). 1972;123(6):755–9.
153. Uttley AnneHC, Collins CH, Naidoo J, George RC. Vancomycin-resistant Enterococci. The Lancet. 1988;331(8575):57–8.
154. Leclercq R, Derlot E, Duval J, Courvalin P. Plasmid-Mediated Resistance to Vancomycin and Teicoplanin in Enterococcus Faecium. N Engl J Med. 1988 Jul 21;319(3):157–61.
155. Sitaraman R. Prokaryotic horizontal gene transfer within the human holobiont: ecological-evolutionary inferences, implications and possibilities. Microbiome. 2018;6(1):163.
156. Kohler V, Keller W, Grohmann E. Regulation of Gram-Positive Conjugation. Front Microbiol. 2019;10:1134.
157. McInnes RS, McCallum GE, Lamberte LE, van Schaik W. Horizontal transfer of antibiotic resistance genes in the human gut microbiome. Curr Opin Microbiol. 2020;53:35–43.
158. Johnston C, Martin B, Fichant G, Polard P, Claverys JP. Bacterial transformation: distribution, shared mechanisms and divergent control. Nat Rev Microbiol. 2014;12(3):181–96.
159. Li W, Wang A. Genomic islands mediate environmental adaptation and the spread of antibiotic resistance in multiresistant Enterococci - evidence from genomic sequences. BMC Microbiol. 2021;21(1):55.
160. Starikova I, Al-Haroni M, Werner G, Roberts AP, Sørum V, Nielsen KM, et al. Fitness costs of various mobile genetic elements in Enterococcus faecium and Enterococcus faecalis. J Antimicrob Chemother. 2013 Dec 1;68(12):2755–65.
161. Paulsen I T, Banerjei L, Myers G S A, Nelson K E, Seshadri R, Read T D, et al. Role of Mobile DNA in the Evolution of Vancomycin-Resistant Enterococcus faecalis. Science (1979). 2003;299(5615):2071–4.
162. McBride SM, Fischetti VA, LeBlanc DJ, Robert C. Moellering Jr, Gilmore MS. Genetic diversity among Enterococcus faecalis. PLoS One. 2007;2(7):e582.
163. Mazaheri Nezhad Fard R, Barton MD, Heuzenroeder MW. Bacteriophage-mediated transduction of antibiotic resistance in enterococci. Lett Appl Microbiol. 2011 Jun;52(6):559–64.
164. Clewell DB, Weaver KE, Dunny GM, Coque TM, Francia MV, Hayes F. Extrachromosomal and Mobile Elements in Enterococci: Transmission, Maintenance, and Epidemiology. In: Enterococci: From Commensals to Leading Causes of Drug Resistant Infection. Boston: Massachusetts Eye and Ear Infirmary; 2014.

165. Hashimoto Y, Taniguchi M, Uesaka K, Nomura T, Hirakawa H, Tanimoto K, et al. Novel Multidrug-Resistant Enterococcal Mobile Linear Plasmid pELF1 Encoding vanA and vanM Gene Clusters From a Japanese Vancomycin-Resistant Enterococci Isolate. Front Microbiol. 2019;10:2568.
166. Palmer KL, Kos VN, Gilmore MS. Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. Curr Opin Microbiol. 2010 Oct;13(5):632–9.
167. Zhu W, Murray PR, Huskins WC, Jernigan JA, McDonald LC, Clark NC, et al. Dissemination of an *Enterococcus* Inc18-Like *vanA* Plasmid Associated with Vancomycin-Resistant *Staphylococcus aureus*. Antimicrob Agents Chemother. 2010 Oct;54(10):4313–20.
168. Valdezate S, Miranda C, Navarro A, Freitas AR, Cabrera JJ, Carrasco G, et al. Clonal outbreak of ST17 multidrug-resistant Enterococcus faecium harbouring an Inc18-like::Tn1546 plasmid in a haemo-oncology ward of a Spanish hospital. J Antimicrob Chemother. 2012;67(4):832–6.
169. Flannagan SE, Chow JW, Donabedian SM, Brown WJ, Perri MB, Zervos MJ, et al. PlasmidContent of a Vancomycin-Resistant *Enterococcus faecalis* Isolate from a Patient Also Colonized by *Staphylococcusaureus* with a VanAPhenotype. Antimicrob Agents Chemother. 2003 Dec;47(12):3954–9.
170. Kohler V, Vaishampayan A, Grohmann E. Broad-host-range Inc18 plasmids: Occurrence, spread and transfer mechanisms. Plasmid. 2018;99:11–21.
171. Grohmann Elisabeth, Muth Günther, Espinosa Manuel. Conjugative Plasmid Transfer in Gram-Positive Bacteria. Microbiol Mol Biol Rev. 2003 Jun 1;67(2):277–301.
172. Bjørkeng E, Rasmussen G, Sundsfjord A, Sjöberg L, Hegstad K, Söderquist B. Clustering of polyclonal VanB-type vancomycin-resistant Enterococcus faecium in a low-endemic area was associated with CC17-genogroup strains harbouring transferable vanB2-Tn5382 and pRUM-like repA containing plasmids with axe-txe plasmid addiction system. APMIS. 2011 Apr 1;119(4–5):247–58.
173. Sivertsen A, Billström H, Melefors Ö, Liljequist BO, Wisell KT, Ullberg M, et al. A Multicentre Hospital Outbreak in Sweden Caused by Introduction of a vanB2 Transposon into a Stably Maintained pRUM-Plasmid in an Enterococcus faecium ST192 Clone. PLoS One. 2014 Aug 25;9(8):e103274.
174. Farias BO de, Bianco K, Nascimento APA, Gonçalves de Brito AS, Moreira TC, Clementino MM. Genomic Analysis of Multidrug-Resistant Enterococcus faecium Harboring vanA Gene from Wastewater Treatment Plants. Microb Drug Resist. 2022 Feb 16;28(4):444–52.
175. Biggel M, Nüesch-Inderbinen M, Raschle S, Stevens MJA, Stephan R. Spread of vancomycin-resistant Enterococcus faecium ST133 in the aquatic environment in Switzerland. J Glob Antimicrob Resist. 2021;27:31–6.
176. Wardal E, Markowska K, Żabicka D, Wróblewska M, Giemza M, Mik E, et al. Molecular Analysis of VanA Outbreak of Enterococcus faecium in Two Warsaw Hospitals: The Importance of Mobile Genetic Elements. Biomed Res Int. 2014;2014:575367.

177. Mikalsen T, Pedersen T, Willems R, Coque TM, Werner G, Sadowy E, et al. Investigating the mobilome in clinically important lineages of Enterococcus faecium and Enterococcus faecalis. BMC Genomics. 2015;16(1):282.
178. Blackwell GA, Hall RM. Mobilisation of a small Acinetobacter plasmid carrying an oriT transfer origin by conjugative RepAci6 plasmids. Plasmid. 2019;103:36–44.
179. Freitas AR, Tedim AP, Francia M v, Jensen LB, Novais C, Peixe L, et al. Multilevel population genetic analysis of vanA and vanB Enterococcus faecium causing nosocomial outbreaks in 27 countries (1986–2012). J Antimicrob Chemother. 2016 Dec 1;71(12):3351–66.
180. Hawkey J, Cottingham H, Tokolyi A, Wick RR, Judd LM, Cerdeira L, et al. Linear plasmids in Klebsiella and other Enterobacteriaceae. Microb Genom. 2022;8(4):000807.
181. Hashimoto Y, Kita I, Suzuki M, Hirakawa H, Ohtaki H, Tomita H. First Report of the Local Spread of Vancomycin-Resistant Enterococci Ascribed to the Interspecies Transmission of a *vanA* Gene Cluster-Carrying Linear Plasmid. mSphere. 2020 Apr 29;5(2):e00102-20.
182. Boumasmoud M, Dengler Haunreiter V, Schweizer TA, Meyer L, Chakrakodi B, Schreiber PW, et al. Genomic Surveillance of Vancomycin-Resistant Enterococcus faecium Reveals Spread of a Linear Plasmid Conferring a Nutrient Utilization Advantage. mBio. 2022 Apr 26;13(2):e0377121.
183. Bakthavatchalam YD, Puraswani M, Livingston A, Priya M, Venkatesan D, Sharma D, et al. Novel linear plasmids carrying vanA cluster drives the spread of vancomycin resistance in Enterococcus faecium in India. J Glob Antimicrob Resist. 2022;29:168–72.
184. Egan SA, Kavanagh NL, Shore AC, Mollerup S, Samaniego Castruita JA, O'Connell B, et al. Genomic analysis of 600 vancomycin-resistant Enterococcus faecium reveals a high prevalence of ST80 and spread of similar vanA regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland. J Antimicrob Chemother. 2022 Feb 1;77(2):320–30.
185. Juhas M, van der Meer JR, Gaillard M, Harding RM, Hood DW, Crook DW. Genomic islands: tools of bacterial horizontal gene transfer and evolution. FEMS Microbiol Rev. 2009 Mar 1;33(2):376–93.
186. Bellanger X, Payot S, Leblond-Bourget N, Guédon G. Conjugative and mobilizable genomic islands in bacteria: evolution and diversity. FEMS Microbiol Rev. 2014 Jul 1;38(4):720–60.
187. Hashimoto Y, Hisatsune J, Suzuki M, Kurushima J, Nomura T, Hirakawa H, et al. Elucidation of host diversity of the VanD-carrying genomic islands in enterococci and anaerobes. JAC Antimicrob Resist. 2022 Feb 1;4(1):dlab189.
188. Roberts AP, Mullany P. Tn916-like genetic elements: a diverse group of modular mobile elements conferring antibiotic resistance. FEMS Microbiol Rev. 2011 Sep 1;35(5):856–71.
189. Tsvetkova K, Marvaud JC, Lambert T. Analysis of the Mobilization Functions of the Vancomycin Resistance Transposon Tn *1549* , a Member of a New Family of Conjugative Elements. J Bacteriol. 2010 Feb;192(3):702–13.

190. Shkumatov A v, Aryanpour N, Oger CA, Goossens G, Hallet BF, Efremov RG. Structural insight into Tn3 family transposition mechanism. Nat Commun. 2022;13(1):6155.
191. Tansirichaiya S, Mullany P, Roberts AP. PCR-based detection of composite transposons and translocatable units from oral metagenomic DNA. FEMS Microbiol Lett. 2016 Sep 1;363(18):fnw195.
192. Quintiliani R, Courvalin P. Characterization of Tn1547, a composite transposon flanked by the IS16 and IS256-like elements, that confers vancomycin resistance in Enterococcus faecalis BM4281. Gene. 1996;172(1):1–8.
193. Lebreton F, Manson AL, Saavedra JT, Straub TJ, Earl AM, Gilmore MS. Tracing the Enterococci from Paleozoic Origins to the Hospital. Cell. 2017;169(5):849-861.e13.
194. Ulrich N, Vonberg RP, Gastmeier P. Outbreaks caused by vancomycin-resistant Enterococcus faecium in hematology and oncology departments: A systematic review. Heliyon. 2017;3(12):e00473–e00473.
195. Harris AD, Pineles L, Belton B, Johnson JK, Shardell M, Loeb M, et al. Universal Glove and Gown Use and Acquisition of Antibiotic-Resistant Bacteria in the ICU: A Randomized Trial. JAMA. 2013 Oct 16;310(15):1571–80.
196. Weiss RA. Virulence and pathogenesis. Trends Microbiol. 2002;10(7):314–7.
197. Mundy LM, Sahm DF, Gilmore M. Relationships between Enterococcal Virulence and Antimicrobial Resistance. Clin Microbiol Rev. 2000 Oct 1;13(4):513–22.
198. Chajęcka-Wierzchowska W, Zadernowska A, Łaniewska-Trokenheim Ł. Virulence factors of Enterococcus spp. presented in food. LWT. 2017;75:670–6.
199. Gao W, Howden BP, Stinear TP. Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. Curr Opin Microbiol. 2018;41:76–82.
200. Wagner T, Joshi B, Janice J, Askarian F, Škalko-Basnet N, Hagestad OC, et al. Enterococcus faecium produces membrane vesicles containing virulence factors and antimicrobial resistance related proteins. J Proteomics. 2018;187:28–38.
201. Somarajan SR, la Rosa SL, Singh K v., Roh JH, Höök M, Murray BE. The Fibronectin-Binding Protein Fnm Contributes to Adherence to Extracellular Matrix Components and Virulence of Enterococcus faecium. Infect Immun. 2015 Dec 1;83(12):4653–61.
202. Wagner TM, Janice J, Paganelli FL, Willems RJ, Askarian F, Pedersen T, et al. Enterococcus faecium TIR-Domain Genes Are Part of a Gene Cluster Which Promotes Bacterial Survival in Blood. Int J Microbiol. 2018;2018:1435820.
203. Choudhury T, Singh K v, Sillanpää J, Nallapareddy SR, Murray BE. Importance of two Enterococcus faecium loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. J Infect Dis. 2011 Apr 15;203(8):1147–54.
204. Xiong X, Tian S, Yang P, Lebreton F, Bao H, Sheng K, et al. Emerging enterococcus pore-forming toxins with MHC/HLA-I as receptors. Cell. 2022;185(7):1157-1171.e22.
205. Williams AM, Rodrigues UM, Collins MD. Intrageneric relationships ofEnterococci as determined by reverse transcriptase sequencing of small-subunit rRNA. Res Microbiol. 1991;142(1):67–74.
206. Top J, Banga NMI, Hayes R, Willems RJ, Bonten MJM, Hayden MK. Comparison of multiple-locus variable-number tandem repeat analysis and pulsed-field gel

electrophoresis in a setting of polyclonal endemicity of vancomycin-resistant Enterococcus faecium. Clin Microbiol Infect. 2008;14(4):363–9.
207. Tran A, Rowlinson MC. Application of Next-Generation Sequencing in Public Health Epidemiology and Outbreak Investigation. Adv. Mol. Pathol. 2019;2(1):89–97.
208. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, García-Cobos S, et al. Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol. 2017;243:16–24.
209. Gilchrist CA, Turner SD, Riley MF, Petri Jr WA, Hewlett EL. Whole-genome sequencing in outbreak analysis. Clin Microbiol Rev. 2015 Jul;28(3):541–63.
210. Hu T, Chitnis N, Monos D, Dinh A. Next-generation sequencing technologies: An overview. Hum Immunol. 2021;82(11):801–11.
211. Illumina. Main Steps in Next-Generation Sequencing [Internet]. Available from: https://www.illumina.com/science/technology/next-generation-sequencing/beginners/ngs-workflow.html
212. Pereira R, Oliveira J, Sousa M. Bioinformatics and Computational Tools for Next-Generation Sequencing Analysis in Clinical Genetics. J Clin Med. 2020;9(1):132.
213. Ambardar S, Gupta R, Trakroo D, Lal R, Vakhlu J. High Throughput Sequencing: An Overview of Sequencing Chemistry. Indian J Microbiol. 2016;56(4):394–404.
214. Illumina. Illumina sequencing platforms [Internet]. Available from: https://www.illumina.com/systems/sequencing-platforms.html
215. Slatko BE, Gardner AF, Ausubel FM. Overview of Next-Generation Sequencing Technologies. Curr Protoc Mol Biol. 2018;122(1):e59.
216. Heather JM, Chain B. The sequence of sequencers: The history of sequencing DNA. Genomics. 2016;107(1):1–8.
217. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008 Nov 6;456(7218):53–9.
218. Werner JJ, Zhou D, Caporaso JG, Knight R, Angenent LT. Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. ISME J. 2012;6(7):1273–6.
219. Mayer P, Farinelli L, Eric H. Kawashima. Illumina Inc, assignee. Method of nucleic acid amplification. United States patent US 7,985,565. US Patent; 7,985,565, 2011.
220. Illumina. Benchtop Sequencers. Available from: https://www.illumina.com/systems/sequencing-platforms.html
221. Rogers LA, Strong K, Cork SC, McAllister TA, Liljebjelke K, Zaheer R, et al. The Role of Whole Genome Sequencing in the Surveillance of Antimicrobial Resistant Enterococcus spp.: A Scoping Review. Front Public Health. 2021;9:599285.
222. Steinig E, Duchêne S, Aglua I, Greenhill A, Ford R, Yoannes M, et al. Phylodynamic Inference of Bacterial Outbreak Parameters Using Nanopore Sequencing. Mol Biol Evol. 2022 Mar 1;39(3):msac040.
223. Quer J, Colomer-Castell S, Campos C, Andrés C, Piñana M, Cortese MF, et al. Next-Generation Sequencing for Confronting Virus Pandemics. Viruses. 2022;14(3):600.

224. Quail MA, Smith M, Coupland P, Otto TD, Harris SR, Connor TR, et al. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. BMC Genomics. 2012;13(1):341.
225. Buermans HPJ, den Dunnen JT. Next generation sequencing technology: Advances and applications. Biochim Biophys Acta - Molecular Basis of Disease. 2014;1842(10):1932–41.
226. Kanwar N, Blanco C, Chen IA, Seelig B. PacBio sequencing output increased through uniform and directional fivefold concatenation. Sci Rep. 2021;11(1):18065.
227. de Coster W, Weissensteiner MH, Sedlazeck FJ. Towards population-scale long-read sequencing. Nat Rev Genet. 2021;22(9):572–87.
228. Logsdon GA, Vollger MR, Eichler EE. Long-read human genome sequencing and its applications. Nat Rev Genet. 2020;21(10):597–614.
229. nanoporetech. Interested in over 99% raw read accuracy? [Internet]. Available from: https://nanoporetech.com/products
230. nanoporetech. Flow cells and nanopores [Internet]. Available from: https://nanoporetech.com/how-it-works/flow-cells-and-nanopores
231. Been M de, Pinholt M, Top J, Bletz S, Mellmann A, Schaik W van, et al. Core genome multilocus sequence typing scheme for high-resolution typing of Enterococcus faecium. J Clin Microbiol. 2015;53(12):3788–97.
232. Neumann B, Prior K, Bender JK, Harmsen D, Klare I, Fuchs S, et al. A Core Genome Multilocus Sequence Typing Scheme for Enterococcus faecalis. J Clin Microbiol. 2019;57(3):e01686-18.
233. Higgs C, Sherry NL, Seemann T, Horan K, Walpola H, Kinsella P, et al. Optimising genomic approaches for identifying vancomycin-resistant Enterococcus faecium transmission in healthcare settings. Nat Commun. 2022;13(1):509.
234. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014;15(11):524.
235. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MTG, et al. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–3.
236. Maiden MCJ, van Rensburg MJJ, Bray JE, Earle SG, Ford SA, Jolley KA, et al. MLST revisited: the gene-by-gene approach to bacterial genomics. Nat Rev Microbiol. 2013;11(10):728–36.
237. Baldo Laura, Julie CDH, Keith AJ, Seth RB, Sarah AB, Rhitoban RayC, et al. Multilocus Sequence Typing System for the Endosymbiont Wolbachia pipientis. Appl Environ Microbiol. 2006 Nov 1;72(11):7098–110.
238. pubmlst. public databases for molecular typing and microbial genome diversity. Available from: https://pubmlst.org/
239. Maiden MCJ. Multilocus Sequence Typing of Bacteria. Annu Rev Microbiol. 2006 Sep 11;60(1):561–88.
240. L. Homan Wieger, Tribe David, Poznanski Simone, Li Mei, Hogg Geoff, Spalburg Emile, et al. Multilocus Sequence Typing Scheme for Enterococcus faecium. J Clin Microbiol. 2002 Jun 1;40(6):1963–71.

241. Ruiz-Garbajosa Patricia, J MBM, D AshleyR, Top Janetta, Sreedhar RN, Torres Carmen, et al. Multilocus Sequence Typing Scheme for Enterococcus faecalis Reveals Hospital-Adapted Genetic Complexes in a Background of High Rates of Recombination. J Clin Microbiol. 2006 Jun 1;44(6):2220–8.
242. Uelze L, Grützke J, Borowiak M, Hammerl JA, Juraschek K, Deneke C, et al. Typing methods based on whole genome sequencing data. One Health Outlook. 2020;2(1):3.
243. Simar SR, Hanson BM, Arias CA. Techniques in bacterial strain typing: past, present, and future. Curr Opin Infect Dis. 2021;34(4):339–45.
244. Nielsen R, Paul JS, Albrechtsen A, Song YS. Genotype and SNP calling from next-generation sequencing data. Nat Rev Genet. 2011;12(6):443–51.
245. Gardner SN, Slezak T, Hall BG. kSNP3.0: SNP detection and phylogenetic analysis of genomes without genome alignment or reference genome. Bioinformatics. 2015 Sep 1;31(17):2877–8.
246. Harris SR. SKA: Split Kmer Analysis Toolkit for Bacterial Genomic Epidemiology. bioRxiv. 2018 Jan 1;453142.
247. Han SS, Jeong YS, Choi SK. Current Scenario and Challenges in the Direct Identification of Microorganisms Using MALDI TOF MS. Microorganisms. 2021 Sep 9;9(9):1917.
248. Bellali S, Haddad G, Iwaza R, Fontanini A, Hisada A, Ominami Y, et al. Antimicrobial susceptibility testing for Gram positive cocci towards vancomycin using scanning electron microscopy. Curr Res Microb Sci. 2022;3:100154.
249. Hegstad Kristin, Christian GG, Haldorsen Bjørg, Matuschek Erika, Schønning Kristian, Truls ML, et al. Performance of the EUCAST Disk Diffusion Method, the CLSI Agar Screen Method, and the Vitek 2 Automated Antimicrobial Susceptibility Testing System for Detection of Clinical Isolates of Enterococci with Low- and Medium-Level VanB-Type Vancomycin Resistanc. J Clin Microbiol. 2014 May 1;52(5):1582–9.
250. Wilson IG, McAfee GG. Vancomycin-resistant enterococci in shellfish, unchlorinated waters, and chicken. Int J Food Microbiol. 2002;79(3):143–51.
251. Alderliesten JB, Duxbury SJN, Zwart MP, de Visser JAGM, Stegeman A, Fischer EAJ. Effect of donor-recipient relatedness on the plasmid conjugation frequency: a meta-analysis. BMC Microbiol. 2020;20(1):135.
252. Bortolaia V, Kaas RS, Ruppe E, Roberts MC, Schwarz S, Cattoir V, et al. ResFinder 4.0 for predictions of phenotypes from genotypes. J Antimicrob Chemother. 2020 Dec 1;75(12):3491–500.
253. Genomics Support Center Tromsø. Laboratory Services [Internet]. Available from: https://uit.no/prosjekter/prosjektsub?p_document_id=468291&sub_id=468396
254. Quainoo S, Coolen JPM, Hijum SAFT van, Huynen MA, Melchers WJG, Schaik W van, et al. Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev. 2017 Oct 1;30(4):1015–63.
255. Arredondo-Alonso S, Top J, Corander J, Willems RJL, Schürch AC. Mode and dynamics of vanA-type vancomycin resistance dissemination in Dutch hospitals. Genome Med. 2021;13(1):1–18.

256. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30(15):2114–20.
257. babraham Bioinformatics. Trim Galore. Available from: https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
258. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. Babraham Bioinformatics. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
259. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: A New Genome Assembly Algorithm and Its Applications to Single-Cell Sequencing. J Comput Biol. 2012 Apr 16;19(5):455–77.
260. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler : Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol. 2017;13(9):e1005595.
261. Lindsey RL, Pouseele H, Chen JC, Strockbine NA, Carleton HA. Implementation of Whole Genome Sequencing (WGS) for Identification and Characterization of Shiga Toxin-Producing Escherichia coli (STEC) in the United States. Front Microbiol. 2016;7:766.
262. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu : scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27:722–36.
263. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon : An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. PLoS One. 2014;9(11):e112963.
264. Hunt M, Silva N de, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator : automated circularization of genome assemblies using long sequencing reads. Genome Biol. 2015;16(294):1–10.
265. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021 Jul 2;49(W1):W293–6.
266. Seemann T. Abricate [Internet]. Github. Available from: https://github.com/tseemann/abricate
267. Yukgehnaish K, Rajandas H, Parimannan S, Manickam R, Marimuthu K, Petersen B, et al. PhageLeads: Rapid Assessment of Phage Therapeutic Suitability Using an Ensemble Machine Learning Approach. Viruses. 2022;14(2):342.
268. NCBI. Bacterial Antimicrobial Resistance Reference Gene Database (PRJNA313047) [Internet]. 2016. Available from: https://www.ncbi.nlm.nih.gov/bioproject/?term=PRJNA313047
269. Hughes-Games A, Roberts AP, Davis SA, Hill DJ. Identification of integrative and conjugative elements in pathogenic and commensal Neisseriaceae species via genomic distributions of DNA uptake sequence dialects. Microb Genom. 2020;6(5):e000372.
270. Margos G, Hepner S, Mang C, Marosevic D, Reynolds SE, Krebs S, et al. Lost in plasmids: next generation sequencing and the complex genome of the tick-borne pathogen Borrelia burgdorferi. BMC Genomics. 2017;18(1):422.
271. Lee RS, Gonçalves da Silva A, Baines SL, Strachan J, Ballard S, Carter GP, et al. The changing landscape of vancomycin-resistant Enterococcus faecium in Australia: a

population-level genomic study. J Antimicrob Chemother. 2018 Dec 1;73(12):3268–78.
272. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4(8):e000206.
273. Carattoli A, Zankari E, Garciá-Fernández A, Larsen MV, Lund O, Villa L, et al. In Silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. Antimicrob Agents Chemother. 2014;58(7):3895–903.
274. Riccardi N, Monticelli J, Antonello RM, di Lallo G, Frezza D, Luzzati R, et al. Therapeutic Options for Infections Due to vanB Genotype Vancomycin-Resistant Enterococci. Microb Drug Resist. 2020;27(4):536–45.
275. Abadía-Patiño L, Christiansen K, Bell J, Courvalin P, Périchon B. VanE-Type Vancomycin-Resistant *Enterococcus faecalis* Clinical Isolates from Australia. Antimicrob Agents Chemother. 2004 Dec;48(12):4882–5.
276. Boyd DA, Mulvey MR. The VanE operon in Enterococcus faecalis N00-410 is found on a putative integrative and conjugative element, Tn6202. J Antimicrob Chemother. 2013;68(2):294–9.
277. Gorrie C, Higgs C, Carter G, Stinear TP, Howden B. Genomics of vancomycin-resistant Enterococcus faecium. Microb Genom. 2019;5(7):e000283.
278. Gorrie CL, da Silva AG, Ingle DJ, Higgs C, Seemann T, Stinear TP, et al. Key parameters for genomics-based real-time detection and tracking of multidrug-resistant bacteria: a systematic analysis. Lancet Microbe. 2021;2(11):e575–83.
279. Fujiya Y, Harada T, Sugawara Y, Akeda Y, Yasuda M, Masumi A, et al. Transmission dynamics of a linear vanA-plasmid during a nosocomial multiclonal outbreak of vancomycin-resistant enterococci in a non-endemic area, Japan. Sci Rep. 2021;11(1):14780.
280. Douglass AP, O'Brien CE, Offei B, Coughlan AY, Ortiz-Merino RA, Butler G, et al. Coverage-Versus-Length Plots, a Simple Quality Control Step for de Novo Yeast Genome Sequence Assemblies. G3 (Bethesda). 2019 Mar 1;9(3):879–87.
281. Illumina. Calling Sequencing SNPs [Internet]. 2009. Available from: https://www.illumina.com/Documents/products/technotes/technote_snp_caller_sequencing.pdf
282. Palmer KL, Schaik W van, Willems RJL, Gilmore. MS. Enterococcal Genomics. In: Gilmore MS, Clewell DB, Ike Y et al., editor. Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet]. Boston: Massachusetts Eye and Ear Infirmary; 2014. PMID: 24649510.
283. Kjær Hansen S, Andersen L, Detlefsen M, Holm A, Roer L, Antoniadis P, et al. Using core genome multilocus sequence typing (cgMLST) for vancomycin-resistant Enterococcus faecium isolates to guide infection control interventions and end an outbreak. J Glob Antimicrob Resist. 2021;24:418–23.
284. Ridom SeqSphere+. cgMLST.org Nomenclature Server. Available from: https://www.cgmlst.org/ncs
285. Okumura N, Watanabe T, Teranishi S, Suzuki D, Hashimoto T, Takahashi K, et al. Successful treatment of aortic valve endocarditis caused by Enterococcus casseliflavus: a case report. BMC Infect Dis. 2021;21(1):447.

286. Ahlbom A. Modern Epidemiology, 4th edition. TL Lash, TJ VanderWeele, S Haneuse, KJ Rothman. Wolters Kluwer, 2021. Eur J Epidemiol. 2021;36(8):767–8.
287. Schürch AC, Arredondo-Alonso S, Willems RJL, Goering R v. Whole genome sequencing options for bacterial strain typing and epidemiologic analysis based on single nucleotide polymorphism versus gene-by-gene–based approaches. Clin Microbiol Infect. 2018;24(4):350–4.
288. Kimura B. Will the emergence of core genome MLST end the role of in silico MLST? Food Microbiol. 2018;75:28–36.
289. van Hal SJ, Willems RJL, Gouliouris T, Ballard SA, Coque TM, Hammerum AM, et al. The global dissemination of hospital clones of Enterococcus faecium. Genome Med. 2021;13(1):52.
290. Pinholt M, Larner-Svensson H, Littauer P, Moser CE, Pedersen M, Lemming LE, et al. Multiple hospital outbreaks of vanA Enterococcus faecium in Denmark, 2012–13, investigated by WGS, MLST and PFGE. J Antimicrob Chemother. 2015 Sep 1;70(9):2474–82.
291. Zhou X, Chlebowicz MA, Bathoorn E, Rosema S, Couto N, Lokate M, et al. Elucidating vancomycin-resistant Enterococcus faecium outbreaks: the role of clonal spread and movement of mobile genetic elements. J Antimicrob Chemother. 2018 Dec 1;73(12):3259–67.
292. Hoffmann M, Yao K, Allard M, Sanchez M, Andersen LP, Hasman H, et al. Complete Genome Sequence of a Vancomycin-Resistant Sequence Type 203 Enterococcus faecium Strain with *vanA* Belonging to Complex Type 859. Microbiol Resour Announc. 2018 Jul 19;7(2):e00815-18.
293. Lisotto P, Couto N, Rosema S, Lokate M, Zhou X, Bathoorn E, et al. Molecular Characterisation of Vancomycin-Resistant Enterococcus faecium Isolates Belonging to the Lineage ST117/CT24 Causing Hospital Outbreaks. Front Microbiol. 2021;12(27):728356.
294. Hammerum AM, Baig S, Kamel Y, Roer L, Pinholt M, Gumpert H, et al. Emergence of vanA Enterococcus faecium in Denmark, 2005–15. J Antimicrob Chemother. 2017 Aug 1;72(8):2184–90.
295. Lisotto P, Raangs EC, Couto N, Rosema S, Lokate M, Zhou X, et al. Long-read sequencing-based in silico phage typing of vancomycin-resistant Enterococcus faecium. BMC Genomics. 2021;22(1):758.
296. Neumann B, Bender JK, Maier BF, Wittig A, Fuchs S, Brockmann D, et al. Comprehensive integrated NGS-based surveillance and contact-network modeling unravels transmission dynamics of vancomycin-resistant enterococci in a high-risk population within a tertiary care hospital. PLoS One. 2020 Jun 24;15(6):e0235160.
297. Ko KS, Baek JY, Lee J young, Oh WS, Peck KR, Lee N, et al. Molecular Characterization of Vancomycin-Resistant Enterococcus faecium Isolates from Korea. J Clin Microbiol. 2005;43(5):2303–6.
298. Klare I, Konstabel C, Mueller-Bertling S, Werner G, Strommenger B, Kettlitz C, et al. Spread of ampicillin/vancomycin-resistant Enterococcus faecium of the epidemic-virulent clonal complex-17 carrying the genes esp and hyl in German hospitals. Eur J Clin Microbiol Infect Dis. 2005;24(12):815–25.

299. Werner G, Klare I, Fleige C, Geringer U, Witte W, Just HM, et al. Vancomycin-resistant vanB-type Enterococcus faecium isolates expressing varying levels of vancomycin resistance and being highly prevalent among neonatal patients in a single ICU. Antimicrob Resist Infect Control. 2012;1(1):21.
300. Bender J, Kalmbach A, Fleige C, Klare I, Fuchs S, Werner G. Population structure and acquisition of the vanB resistance determinant in German clinical isolates of Enterococcus faecium ST192. Sci Rep. 2016;6:21847.
301. Weber A, Maechler F, Schwab F, Gastmeier P, Kola A. Increase of vancomycin-resistant Enterococcus faecium strain type ST117 CT71 at Charité - Universitätsmedizin Berlin, 2008 to 2018. Antimicrob Resist Infect Control. 2020;9(1):109.
302. Hammerum AM, Justesen US, Pinholt M, Roer L, Kaya H, Worning P, et al. Surveillance of vancomycin-resistant enterococci reveals shift in dominating clones and national spread of a vancomycin-variable vanA Enterococcus faecium ST1421-CT1134 clone, Denmark, 2015 to March 2019. Euro Surveill. 2019;24(34):1900503.
303. Elstrøm P, Astrup E, Hegstad K, Samuelsen Ø, Enger H, Kacelnik O. The fight to keep resistance at bay, epidemiology of carbapenemase producing organisms (CPOs), vancomycin resistant enterococci (VRE) and methicillin resistant Staphylococcus aureus (MRSA) in Norway, 2006 - 2017. PLoS One. 2019 Feb 4;14(2):e0211741.
304. Eichel V, Klein S, Bootsveld C, Frank U, Heeg K, Boutin S, et al. Challenges in interpretation of WGS and epidemiological data to investigate nosocomial transmission of vancomycin-resistant Enterococcus faecium in an endemic region: incorporation of patient movement network and admission screening. J Antimicrob Chemother. 2020 Jul 1;75(7):1716–21.
305. Data from the ECDC Surveillance Atlas - Antimicrobial resistance: Reports 2005-2019. [Internet]. 2020. Available from: https://ecdc.europa.eu/en/antimicrobial-resistance/surveillance-and-disease-data/data-ecdc
306. Abdelbary MHH, Senn L, Greub G, Chaillou G, Moulin E, Blanc DS. Whole-genome sequencing revealed independent emergence of vancomycin-resistant Enterococcus faecium causing sequential outbreaks over 3 years in a tertiary care hospital. Eur J Clin Microbiol Infect Dis. 2019;38(6):1163–70.
307. Flipse J, Wintersdorff CJH von, Niekerk JM van, Jamin C, Tiel FH van, Hasman H, et al. Appearance of VanD-positive *Enterococcus faecium* in a tertiary hospital in the Netherlands: prevalence of VanC and VanD in hospitalized patients. Sci Rep. 2019;9(1):6949.
308. Sato T, Wada T, Shinagawa M, Fukushima Y, Nakajima C, Suzuki Y. Emergence of vancomycin- and teicoplanin-resistant *Enterococcus faecium* via *vanD5*-harbouring large genomic island. J Antimicrob Chemother. 2020;dkaa220.
309. Lee T, Pang S, Abraham S, Coombs GW. Antimicrobial-resistant CC17 Enterococcus faecium: The past, the present and the future. J Glob Antimicrob Resist. 2019;16:36–47.
310. Sadowy E. Mobile genetic elements beyond the VanB-resistance dissemination among hospital-associated enterococci and other Gram-positive bacteria. Plasmid. 2021;114:102558.

311. Yan MY, He YH, Ruan GJ, Xue F, Zheng B, Lv Y. The prevalence and molecular epidemiology of vancomycin-resistant Enterococcus (VRE) carriage in patients admitted to intensive care units in Beijing, China. J Microbiol Immunol Infect. 2022;13:S1684-1182(22)00099-8.
312. Mühlberg E, Umstätter F, Kleist C, Domhan C, Mier W, Uhl P. Renaissance of vancomycin: approaches for breaking antibiotic resistance in multidrug-resistant bacteria. Can J Microbiol. 2019 Sep 23;66(1):11–6.
313. Keith E. Weaver. Enterococcal Genetics. Microbiol Spectr. 2019 Mar 8;7(2):7.2.11.
314. Top J, Arredondo-Alonso S, Schürch AC, Puranen S, Pesonen M, Pensar J, et al. Genomic rearrangements uncovered by genome-wide co-evolution analysis of a major nosocomial pathogen, Enterococcus faecium. Microb Genom. 2020;6(12):488.
315. Islam M, Abaragu A, Sistu H, Akins RL, Palmer K. Molecular characterization of vancomycin-resistant Enterococcus faecium isolates collected from the Dallas, Texas area. bioRxiv. 2021 Jan 1;2021.02.16.431552.
316. Howden BP, Holt KE, Lam MMC, Seemann Torsten, Ballard Susan, Coombs GW, et al. Genomic Insights to Control the Emergence of Vancomycin-Resistant Enterococci. mBio. 2013 Aug 13;4(4):e00412-13.
317. Mingoia M, Morici E, Tili E, Giovanetti E, Montanari MP, Varaldo PE. Characterization of Tn *5801.Sag* , a Variant of Staphylococcus aureus Tn *916* Family Transposon Tn *5801* That Is Widespread in Clinical Isolates of Streptococcus agalactiae. Antimicrob Agents Chemother. 2013 Sep;57(9):4570–4.
318. Smyshlyaev G, Bateman A, Barabas O. Sequence analysis of tyrosine recombinases allows annotation of mobile genetic elements in prokaryotic genomes. Mol Syst Biol. 2021 May 1;17(5):e9880.
319. Lambertsen L, Rubio-Cosials A, Patil KR, Barabas O. Conjugative transposition of the vancomycin resistance carrying Tn1549: enzymatic requirements and target site preferences. Mol Microbiol. 2018 Mar 1;107(5):639–58.
320. Wardal E, Żabicka D, Hryniewicz W, Sadowy E. VanA-Enterococcus faecalis in Poland: hospital population clonal structure and vanA mobilome. Eur J Clin Microbiol Infect Dis. 2022;41(10):1245–61.
321. Kumar S, Singh S, Kumar V et al (2020b) Pathogenesis and antibiotic resistance of Staphylococcus aureus. In: Model organisms for microbial pathogenesis, biofilm formation and antimicrobial drug discovery. Springer, Singapore, pp 99–115.
322. Paganelli FL, Willems RJL, Jansen Pamela, Hendrickx Antoni, Zhang Xinglin, Bonten MJM, et al. Enterococcus faecium Biofilm Formation: Identification of Major Autolysin AtlAEfm, Associated Acm Surface Localization, and AtlAEfm-Independent Extracellular DNA Release. mBio. 2013 Apr 9;4(2):e00154-13.
323. Cacaci M, Giraud C, Leger L, Torelli R, Martini C, Posteraro B, et al. Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in Enterococcus faecium. Sci Rep. 2018;8(1):12412.
324. Zhang S, Lebreton F, Mansfield MJ, Miyashita SI, Zhang J, Schwartzman JA, et al. Identification of a Botulinum Neurotoxin-like Toxin in a Commensal Strain of Enterococcus faecium. Cell Host Microbe. 2018;23(2):169-176.e6.

325. Leigh RJ, McKenna C, McWade R, Lynch B, Walsh F. Comparative genomics and pangenomics of vancomycin-resistant and susceptible Enterococcus faecium from Irish hospitals. J Med Microbiol. 2022;71(10):1590.
326. Chen L, Yang J, Yu J, Yao Z, Sun L, Shen Y, et al. VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 2005 Jan 1;33(suppl_1):D325–8.
327. Cattoir V. The multifaceted lifestyle of enterococci: genetic diversity, ecology and risks for public health. Curr Opin Microbiol. 2022;65:73–80.
328. Revtovich A v, Tjahjono E, Singh K v, Hanson BM, Murray BE, Kirienko N v. Development and Characterization of High-Throughput Caenorhabditis elegans – Enterococcus faecium Infection Model. Front Cell Infect Microbiol. 2021;11(29):667327.
329. Heikens E, Bonten MJ WR. Enterococcal Surface Protein Esp Is Important for Biofilm Formation of Enterococcus faecium E1162. J Bacteriol. 2007 Nov 15;189(22):8233–40.
330. Geraldes C, Tavares L, Gil S, Oliveira M. Enterococcus Virulence and Resistant Traits Associated with Its Permanence in the Hospital Environment. Antibiotics. 2022;11(7):857.

Paper I

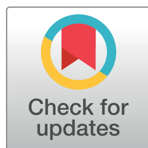
RESEARCH ARTICLE

Novel genomic islands and a new *vanD*-subtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway

Mushtaq T. S. AL Rubaye[1], Jessin Janice[1,2], Jørgen Vildershøj Bjørnholt[3,4], Aleksandra Jakovljev[5], Maria Elisabeth Hultström[6], Arnfinn Sundsfjord[1,2], Kristin Hegstad[1,2]*

1 Research Group for Host-Microbe Interactions, Department of Medical Biology, University of Tromsø (UiT) – The Arctic University of Norway, Tromsø, Norway, **2** Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North-Norway, Tromsø, Norway, **3** Department of Clinical Microbiology, Oslo University Hospital, Oslo, Norway, **4** Institute of Clinical Medicine, University of Oslo, Oslo, Norway, **5** Department of Medical Microbiology, St Olavs Hospital, Trondheim, Norway, **6** Division of Nephrology, Department of Medicine, St Olavs Hospital, Trondheim, Norway

* Kristin.Hegstad@uit.no



OPEN ACCESS

Citation: AL Rubaye MTS, Janice J, Bjørnholt JV, Jakovljev A, Hultström ME, Sundsfjord A, et al. (2021) Novel genomic islands and a new *vanD*-subtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway. PLoS ONE 16(7): e0255187. https://doi.org/10.1371/journal.pone.0255187

Editor: Rosa del Campo, Hospital Universitario Ramon y Cajal, SPAIN

Received: May 3, 2021

Accepted: July 11, 2021

Published: July 23, 2021

Copyright: © 2021 AL Rubaye et al. This is an open access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Genome raw reads and assemblies as well as the sequences of GIs generated in this study have been submitted to the NCBI. Raw reads have been deposited to the short read archive under Bioproject ID PRJNA627463. Genomic assemblies and island sequences have been deposited to NCBI under the following accession numbers: Genome sequences: JABBNS000000000, JABBNP000000000, JABBNR000000000, JABBNQ000000000,

Abstract

Background

Vancomycin-resistant enterococci (VRE) represent several types of transferable vancomycin resistance gene clusters. The *vanD* type, associated with moderate to high level vancomycin resistance, has only sporadically been described in clinical isolates. The aim of this study was to perform a genetic characterization of the first VanD-type VRE strains detected in Norway.

Methods

The VanD-type VRE-strains (n = 6) from two patient cases were examined by antimicrobial susceptibility testing and whole genome sequencing (WGS) to uncover Van-phenotype, strain phylogeny, the *vanD* gene clusters, and their genetic surroundings. The putative transferability of *vanD* was examined by circularization PCR and filter mating.

Results

The VanD-type *Enterococcus faecium* (n = 4) and *Enterococcus casseliflavus* (n = 2) strains recovered from two cases (A and B), expressed moderate to high level vancomycin resistance (MIC 64—>256 mg/L) and various levels of teicoplanin susceptibility (MIC 2—>256 mg/L). WGS analyses revealed phylogenetically different *E. faecium* strains (A1, A2, and A3 of case A and B1 from case B) as well as *vanD* gene clusters located on different novel genomic islands (GIs). The *E. casseliflavus* strains (B2 and B3 of case B) were not clonally related, but harbored nearly identical novel GIs. The *vanD* cluster of case B strains represents a novel *vanD*-subtype. All the *vanD*-GIs were integrated at the same chromosomal site and contained genes consistent with a *Clostridiales* origin. Circular forms of the

JABBNO000000000, and JABBNN000000000. GIs: MT951615, MT951616, and MT951617.

Funding: This work was supported by a grant from the Northern Norway Regional Health Authority Medical Research Programme (HNF1362-17). The publication charges for this article have been funded by a grant from the publication fund of UiT The Arctic University of Norway. The funders had no role in the study design, data collection and analysis, decision to publish or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

vanD-GIs were detected in all strains except B1. Transfer of *vanD* to an *E. faecium* recipient was unsuccessful.

Conclusions

We describe the first VanD-type *E. casseliflavus* strains, a novel *vanD*-subtype, and three novel *vanD*-GIs with a genetic content consistent with a *Clostridiales* order origin. Despite temporal occurrence, case A and B *E. faecium* strains were phylogenetically diverse and harbored different *vanD* subtypes and *vanD*-GIs.

Introduction

Vancomycin resistant enterococci (VRE) have become a global nosocomial problem three decades after the first description in the late 1980s [1]. Eight different acquired vancomycin resistance gene clusters (*vanA*, *vanB*, *vanD*, *vanG*, *vanE*, *vanL*, *vanM*, and *vanN*) have been identified [2]. The *vanC* gene cluster is intrinsic in *E. casseliflavus* and *E. gallinarum* [2]. In general, *van* gene clusters encode three groups of co-acting enzymes; 1) enzymes necessary for the synthesis of new peptidoglycan precursors, 2) enzymes that erase the inherent D-Ala-D-Ala-ending precursors, and 3) a two-component signal transduction system for inducible resistance [3]. The normal enterococcal cell wall side chain terminal residue D-Ala-D-Ala, to which vancomycin binds with high affinity, are replaced by D-Ala-D-Lac in *vanA*, *vanB*, *vanD*, and *vanM* gene clusters or D-Ala-D-Ser in the other *van* gene clusters [3]. Vancomycin binds to D-Ala-D-Ser with seven times lower affinity compared to D-Ala-D-Ala, causing low-level vancomycin resistance, while the binding affinity of vancomycin to D-Ala-D-Lac is almost 1000 times lower mediating high-level resistance [4]. The *vanA* and *vanB* clusters dominate worldwide, likely due to linkage to successful mobile genetic elements (MGEs) [5]. Although the *vanA*, *vanB*, and *vanD* clusters have a similar organization, the *vanD* gene clusters have so far only been sporadically described on chromosomal genomic islands (GIs) that have not been shown to be transferable between enterococci [6–9]. The *vanD* gene cluster has up till now been reported in five species of enterococci (*Enterococcus faecium*, *Enterococcus faecalis*, *Enterococcus gallinarum*, *Enterococcus avium*, and *Enterococcus raffinosus*) [10].

The VanD-phenotype is characterized by moderate to high level vancomycin resistance and various levels of susceptibility to teicoplanin [3, 11, 12]. The housekeeping *ddl* gene (D-Ala-D-Ala ligase) is often inactivated by mutations in *vanD* containing strains causing an impaired chromosomal peptidoglycan synthesis pathway and addiction to *vanD*-expression as the alternative peptidoglycan precursor pathway [3, 7, 13]. Based on sequence differences, there are five known subtypes of *vanD*. The sequence diversity in *vanD* gene cluster subtypes mostly is in the $vanY_D$, $vanH_D$, *vanD*, and $vanX_D$ genes and at the intergenic sequence between the two operons of the cluster [11]. VanD VRE are rare and have only been reported sporadically from the Netherlands, France, Canada, Japan, Sweden, Australia, the US, and Brazil during the last decades [7, 8, 10, 12–18].

In this study, we aim to determine the genetic relatedness between the first Norwegian VanD-type VRE strains, their Van-phenotype, and the putative MGEs harbouring the *vanD*-gene cluster.

Material and methods

Case descriptions

Case A. A middle-aged previously healthy female presented with acute hepatic failure. An urgent transplantation with an ABO-incompatible liver was performed. At week eight, a subphrenic abscess was diagnosed supported by the growth of *E. coli* and *E. faecium* and treated by local drainage. In week 16, a new subphrenic abscess was diagnosed and a *vanD E. faecium* in pure culture was isolated from the abscess drainage pigtail catheter. Screening for fecal VRE-carriage at week 20 after transplantation yielded *vanD E. faecium*. Several negative rectal VRE-screening samples were obtained during the subsequent 9 months, except for one *vanC E. casseliflavus* strain. Several screening samples were collected during linezolid treatment. Antibiotic treatment was successfully terminated almost a year after the transplantation.

Case B. An elderly female, undergoing hemodialysis for the last five years after kidney transplant failure, presented with recurrent urinary tract infections (UTIs), predominantly caused by *Klebsiella pneumoniae*, but occasionally by *E. faecium*. Due to relapsing *Clostridioides difficile* infections (CDIs), she had received oral vancomycin prophylaxis the last three years. The urine yielded *vanD E. faecium* in pure culture. Repeated fecal VRE-screening (follow-up 2 years) revealed the presence of *vanD E. casseliflavus*, but not *vanD E. faecium*. The *vanD E. faecium* UTI was successfully treated with linezolid, while the *C. difficile* prophylaxis was changed to metronidazole.

Relevant case characteristics are summarized in Table 1. Antibiotic treatment and microbiological findings for case A are presented in S1 Fig.

Ethical approval

Since this study contain only limited anonymized patient data, the study was approved by the Data Protection Officer at Oslo University Hospital and the Chief of Department of Microbiology at St Olavs Hospital. The written consents of the patients were obtained to use anonymized data from their patient journal in publication of this work.

VRE strains and data collection

The first two cases of VanD-type VRE were identified in Norway in 2017. The Norwegian National Advisory Unit on Detection of Antimicrobial Resistance received the strains for further characterization (Table 2). Three VanD-positive *E. faecium* (VanD-type VRE*fm*) (A1, A2, and A3) strains were isolated from case A. The strains of case A were recovered from a subphrenic abscess (A1 and A2) and through rectal screening (A3). A month later, a VanD-type

Table 1. Relevant case characteristics.

Case	Underlying condition	Indication antimicrobial treatment	Antimicrobial treatment	Time to isolation of *vanD E. faecium*	Infection focus	Rectal carriage[#]	Hospital
A	Acute liver Tx—otherwise healthy	Postoperative subphrenic abscesses	Broad spectrum beta-lactams, vancomycin, trimethoprim/ sulfamethoxazole (PJP prophylaxis)	19 weeks post liver tx	Subphrenic abcess	*vanD E. faecium*, *E. casseliflavus*	1 and 2
B	Tx kidney failure, hemodialysis, recurrent UTIs and CDIs	Recurrent CDI	Vancomycin p.o. (CDI prophylaxis)	3 years from start of vancomycin prophylaxis	Urinary tract infection	*vanD E. casseliflavus*	3

Tx: transplantation, UTI: Urinary tract infection, CDI: *C. difficile* infection, PJP: Pneumocystis jiroveci pneumonia, p.o.: postoperative.

[#] Fecal screening with CHROMagar™ VRE.

https://doi.org/10.1371/journal.pone.0255187.t001

Table 2. Relevant strain characteristics.

Strain ID	Strain name	Species	MLST	VAN*	TEC	AMP	LIN	GEN	Ddl ligase changes compared to *E. faecium* E1	Source	Isolation day
A1	VRE1736	*E. faecium*	1486	64	4	>8	<1	>500	S185 changed to F185	Abcess drainage	Day 1
A2	VRE1737	*E. faecium*	1486	>128	4	>8	<1	>500	S185 changed to F185	Abcess drainage	Day 1
A3	KresVRE0001	*E. faecium*	117	64	2	>8	2	<32	S319 changed to G319 #	Screening	Day 10
B1	KresVRE0002	*E. faecium*	203	>256	>256	>8	2	>500	Truncated protein of 110 aa #	Urine	Day 42
B2	KresVRE0003	*E. casseliflavus*	-	>256	>256	1	2	<2		Rectal screening	Day 65
B3	KresVRE0012	*E. casseliflavus*	-	>256	>8	<0,25	2	<32		Rectal screening	Day 665

*, MICs in mg/L for VAN (vancomycin), TEC (teicoplanin), AMP (ampicillin), LIN (linezolid), and GEN (gentamicin).

#, These changes are not within the part of the *ddl* gene used for sequence typing.

https://doi.org/10.1371/journal.pone.0255187.t002

VRE*fm* (B1) strain was isolated from the urinary tract in a hemodialysis patient (case B). Further, two *vanD*-positive *E. casseliflavus* strains were recovered from case B by rectal screening, three weeks (B2) and two years (B3) later. Both patients had received vancomycin treatment before the isolation of the VanD-type VRE.

Antimicrobial Susceptibility Testing (AST) and *van* genotype determinations

AST was performed by broth microdilution using the GPALL1F or EUENCF Sensititre plates (Thermo Fisher Scientific, Waltham, Massachusetts, USA), ComASP™ Vancomycin, and Teicoplanin MIC Test Strip (Liofilchem, Roseto Degli Abruzzi, Italy). The results (MICs) were interpreted according to EUCAST clinical breakpoints v. 10.0 2020 [19]. The *van* genotype was initially determined by a *vanDEG* multiplex PCR as described previously [20, 21] and JumpStart REDTaq ReadyMix (Merck KGaA, Darmstadt, Germany). DNA extractions for PCRs were performed using the NucliSens EasyMAG instrument and reagents (BioMeriéux, Marcy-l'Étoile, France) according to the manufacturer's instructions.

Species identification and Whole Genome Sequencing (WGS)

Strains were subcultured on blood agar to ensure pure culture. Species identification was performed by MALDI-TOF (Bruker, Billerica, USA) according to the manufacturer's instructions. Genomic DNA was extracted using DNeasy Blood and tissue kit (Qiagen, Hilden, Germany). The total DNA concentration was quantified by Qubit fluorometer (Invitrogen, Thermo Fisher Scientific). Libraries were prepared by the Nextera XT DNA library preparation kit (Illumina, San Diego, USA) and sequenced using Illumina NextSeq500 and the Mid Output 300 cycles cell.

Genomic analyses

Adapter removal and quality trimming of the raw reads were performed by trimmomatic v0.39 [22]. Later, genome assembly was done using SPAdes v3.13.0 [23] and the quality of assembled genomes was assessed using QUAST v5.0.2 [24]. The annotation of the transposons was carried out using the National Center for Biotechnology Information (NCBI) prokaryotic genome annotation pipeline (PGAP) [25]. Antimicrobial resistance (AMR) genes were

identified *in silico* from the assemblies using NCBI bacterial AMR reference gene database (PRJNA313047) [26] in ABRicate tool v0.8.7 [27]. Identification of Type IV secretion systems genes was carried out by BLASTp [28] searches against the SecReT4 database [29].

Phylogenetic analyses

To explore the phylogenetic relationship between the *vanD* strains and publically available genome sequences on NCBI, the global phylogenetic trees were generated based on the core genome. All closed genomes of *E. faecium* (n = 135) and *E. casseliflavus* (n = 3) from NCBI as of 04.04.2020 were retrieved and phylogenetic trees were constructed using Parsnp v1.2 [30]. Another core genome SNP tree was built for the publicly available VanD-type VRE*fm* genome sequences together with the Norwegian vanD-type VRE*fm*. Also, a SNP tree was generated for *vanD* gene cluster sequences using parsnp. Multilocus Sequence Typing (MLST) was performed using MLST tool version 2.11 [31]. For high-resolution typing, Minimum Spanning Tree was generated based on the 1423 core genes of *E. faecium* scheme of SeqSphere+ software V6.0.2 (Ridom GmbH, Münster, Germany [http://www.ridom.de/seqsphere/]). We used the default $\leq$ 20 allelic differences as a threshold for cluster calculation and clonal relatedness [32].

Comparative genomics

The closest non-VRE strains to each of the Norwegian VanD-type VRE were selected from the global phylogenetic tree. We used Mauve [33] to sort the contigs according to the reference genomes (E1 (NZ_CP018065.1) for A1-3 strains, E4402 (NZ_LR135174) for B1 strain, and EC20 (CP004856.1) for B2-3 strains) followed by Easyfig v2.2.2 [34] for comparison. The Artemis comparison tool [35] was used to visualize the BLASTn v2.6.0 search result and to locate the mobile genetic structures containing *vanD* gene clusters and their insertion site in the genome. Sequences of the GIs harboring the *vanD* gene clusters were BLASTed against the NCBI nr database to find the homologous sequences. Pyani v0.2.7 was used to determine the average nucleotide identity (ANI) between genomes, GIs and *vanD* gene clusters [36]. For the novel GIs, transposon numbers were registered at the Transposon Registry [37].

Excision of putative GIs

The ability of the GIs to circularize was examined by PCR using the following pair of primers which directed outwards from the GIs ends: `5´-GCGTGAGAAGCTGACAACAA-3´` and `5´-GTTTCAGCCGCCAACTATTC-3´`. Subsequent Sanger sequencing of PCR products using BigDye 3.1 technology (Applied Biosystems, CA, USA) was performed to confirm the expected sequence.

Transferability of putative GIs

Transferability of *vanD* gene clusters was examined as described previously [38] using *E. faecium* BM4105-RF [39] as a recipient. To determine transfer frequency, colony forming units were counted on Brain heart infusion agar with rifampicin (30 mg/L) and fusidic acid (20 mg/L), and/or vancomycin (8 mg/L).

Results and discussion

Most of the reported VanD-type VRE have been sporadic clinical isolates [7, 10, 12, 13, 15]. Despite an increasing prevalence of VRE in Norway since 2010, only *vanA* and *vanB* have been reported until now [40]. The detection of VanD-type VRE from two different patients within two months in 2017, therefore raised a concern of facing a VanD-type VRE outbreak in

Norway, although no obvious epidemiological link between the patients was identified. Thus, the pheno- and genotype of the six VanD VRE strains were examined (Table 2). All three VRE from case A were *E. faecium*, while in case B, one *E. faecium* and two *E. casseliflavus* were isolated. To our knowledge, B2 and B3 are the first VanD-type vancomycin resistant *E. casseliflavus* strains reported.

AST results

The AST-results are summarized in Table 2. Briefly, all strains expressed high-level vancomycin resistance (MIC $\geq$ 64 mg/L), various levels of susceptibility to teicoplanin (MIC 2 mg/L to $>$256 mg/L), and susceptibility to linezolid. All four *E. faecium* strains were ampicillin resistant and three also demonstrated high-level gentamicin resistance.

In silico analysis showed that all strains contained the *vanD* gene cluster integrated into their chromosome. The *E. casseliflavus* genomes (B2 and B3) also contained the intrinsic *vanC* gene cluster [2]. In the *E. faecium* strain B1, alignment of the housekeeping D-Ala-D-Ala ligase deduced from the *ddl* gene sequence showed a truncated protein of only 110 amino acids caused by a deletion resulting in a frameshift and a premature stop codon (Table 2 and S2 Fig). All the other VanD-type VRE*fm* strains showed point mutations in essential positions that presumably could lead to a non-functional Ddl ligase. In the literature, most VanD-type VRE strains described have had an impaired Ddl ligase and are thus dependent on the constitutively expressed *vanD* cluster to synthesise peptidoglycan [10].

The VanD *E. faecium* strains from the two cases were not closely related

The VanD VRE*fm* strains from cases A and B had different MLST profiles (Table 2). A1 and A2 genomes had an identical MLST profile which was registered as the novel ST1486, a single locus (*ddl* allele) variant of ST117 (strain A3) belonging to the hospital associated ST78 lineage. The *E. faecium* strain from case B belonged to ST203 which is part of the ST17 hospital associated lineage. Population genetic modeling based on the seven MLST genes using the Bayesian Analysis of Population Structure (BAPS) software have shown that 80% of the *E. faecium* nosocomial strains cluster in two different groups (2–1 and 3–3) [41]. *E. faecium* A and B strains belonged to lineages within these different main BAPS groups (lineage ST78 to 2–1 and lineage ST17 to 3–3) [41], confirming a large phylogenetic distance. This was further shown by cgMLST analysis which revealed that A1-3 strains belonged to the same novel cluster type (CT) 3198 (Fig 1). The B1 strain belonged to another novel CT3199 and showed at least 354 allelic differences to A1-3 strains. The two ST1486 strains had only one allelic difference, while the maximum allelic differences (eight) within CT3198 were between A1 and A3. One of these allelic differences was in the *ddl* allele which is one of the seven MLST scheme genes. Our results show that even strains with different MLST profiles could be clonally closely related and have the same CT.

For *E. casseliflavus* strains, a core genome SNP tree was constructed together with publically available closed genomes. Interestingly, the two VanD strains (B2 and B3) clustered in two separate branches, showing that they were not clonally related (S3 Fig).

The vancomycin susceptible *E. faecium* strain E1 (GCF_001886635.1) isolated from Spain in 2010, was identified as the closest genome to A1-3 strains using a core genome SNP tree of all closed *E. faecium* genomes in NCBI and the Norwegian VanD-type VRE*fm* genomes (S4 Fig). Strain E1 was therefore used as a reference genome for sorting contigs and further comparative genomic analyses. Genomic comparison using Easyfig confirmed that the A1-3 genomes were very similar. The ANI between A1 and A2 was the highest (99.99%).

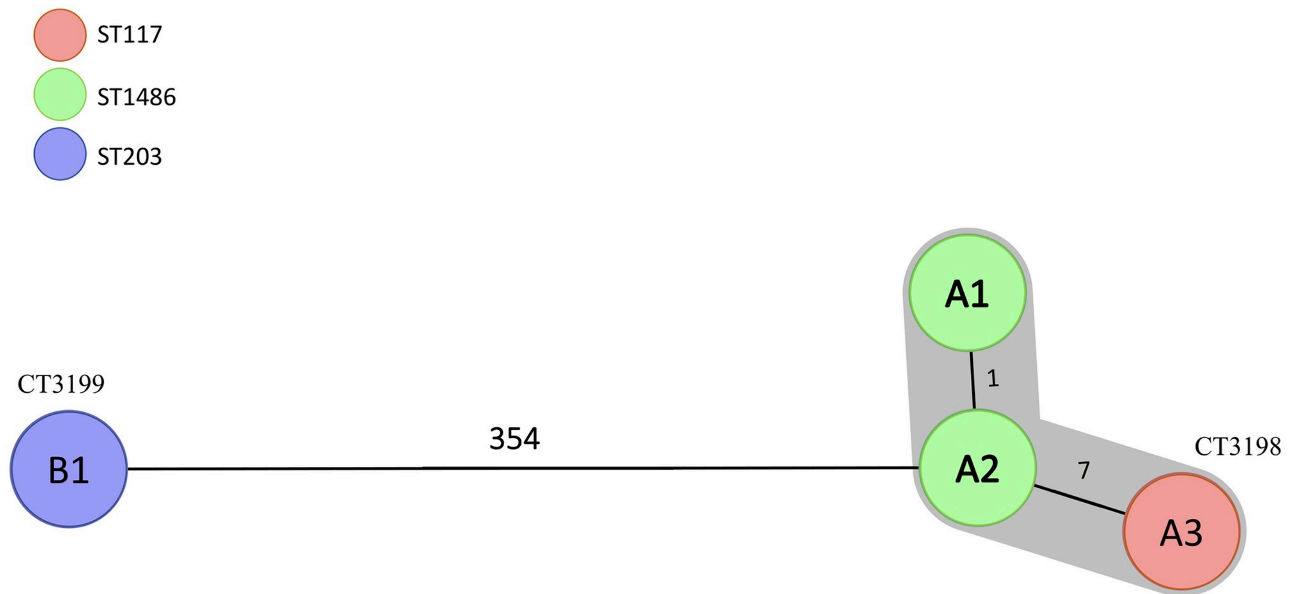


Fig 1. Minimum spanning tree based on the cgMLST typing of the Norwegian VanD VRE*fm*. Regardless of the different STs, A1-3 strains clustered together while the VRE*fm* strain B1 showed 354 allelic differences with the A2 strain.

https://doi.org/10.1371/journal.pone.0255187.g001

Comparison of case B VRE*fm* (B1) to case A VRE*fm* genomes, confirmed observed genomic differences (S5 Fig).

The significant phylogenetic difference between the *vanD E. faecium* strains from case A and B is consistent with the observed sporadic occurence of *vanD*-type VRE strains in contrast to the epidemic *vanA/B*-type VRE [7, 12, 13, 15]. Our patient characteristics with underlying diseases and long-term antibiotic exposure including vancomycin are also consistent with previous observations in *vanD* VRE cases [12, 17].

A novel *vanD*-subtype was found in strains from case B

Sequence comparison and phylogenetic analysis of complete *vanD* gene clusters from this study and reference sequences representing the five known *vanD* subtypes (*vanD1-D5*) [8, 11, 42, 43], showed that the Norwegian *vanD* gene clusters belonged to two different *vanD*-subtypes. In case A, the *vanD* gene clusters of strains A1 and A2 were 100% identical and showed 99.96% ANI to the cluster in A3. The *vanD* genes of case A clustered with the *vanD5* reference sequence (*E. faecium* strain N03-0072) (Fig 2). ANIs between the *vanD5* reference sequence and A1-3 strains were >99.9%. In case B strains, B2 and B3 *vanD* gene clusters were 99.98% identical and the B1 *vanD* gene cluster showed > 99.96% ANI with them. The ANI between case A and B *vanD* gene clusters was around 91%. B1-3 *vanD* gene clusters are significantly different from the known *vanD*-subtypes (maximum 93.7% identity to the known subtypes) (S1 Table). Thus, we propose that the B *vanD* gene cluster is a new subtype termed *vanD6*. Identification of the novel *vanD6* gene cluster in two different species of enterococci suggests inter-species genetic exchange.

Three novel *vanD*-containing GIs identified

Comparison alignments with non-VRE reference genomes using Artemis comparison tool showed that all *vanD* gene clusters in the Norwegian vanD-type VRE were part of GIs ranging

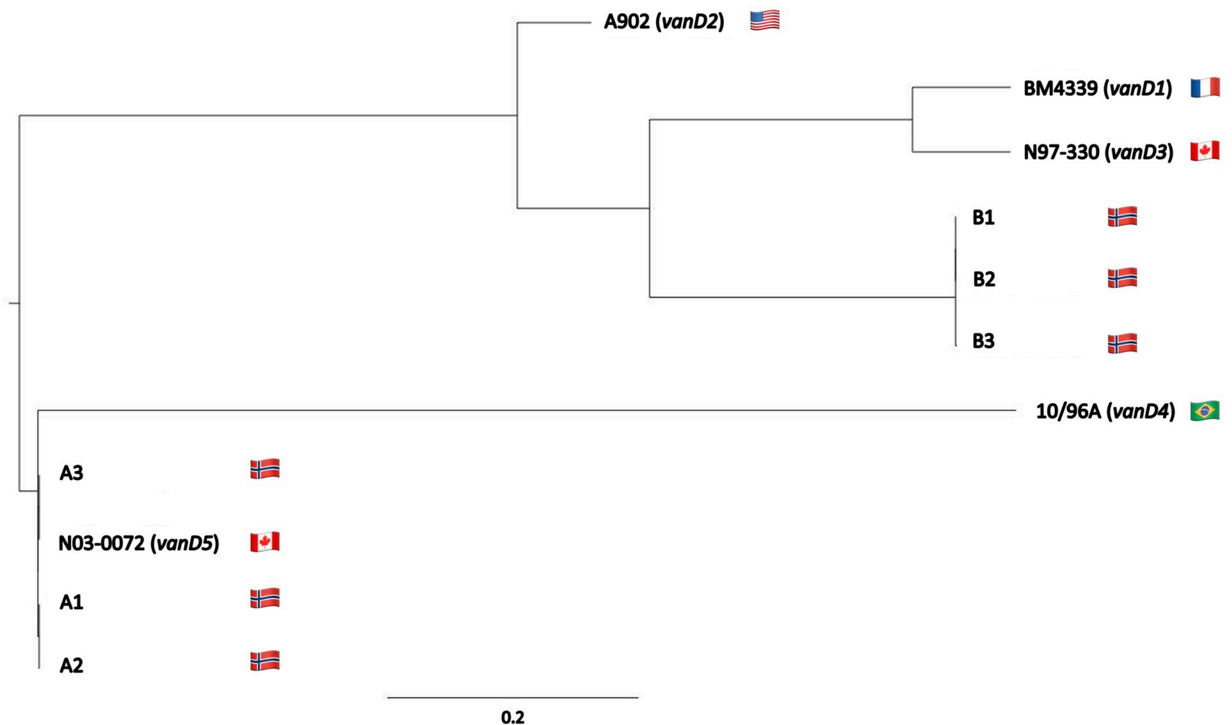


Fig 2. Phylogenetic SNP tree of the *vanD* gene clusters of the Norwegian and *vanD1*—*vanD5* subtype reference clusters retrieved from NCBI. Flags represent the countries that *vanD*-types were discovered in first. Case A strains clustered with *vanD5* reference N03-0072 while case B strains clustered separately.

https://doi.org/10.1371/journal.pone.0255187.g002

between 112–126 kb (Table 3). The GC content of the GIs was higher (44.1–44.3%) than the average GC content range of 38% of *E. faecium* strains [44–46]. For B2 and B3 *E. casseliflavus* strains, the genomic GC content was 42.4% and 42.3%, in contrast to 44.6 and 44.7% for their GIs, respectively. The GI Tn*6711* of A1-3 strains showed identical size and had an ANI above 99.99% suggesting a common origin. The GI Tn*6713* of the *E. casseliflavus* strains (B2 and B3) was identical in size and showed only 0.001% difference (S2 Table). The GI Tn*6712* in *E. faecium* strain B1 was 7230 bp larger than that of *E. casseliflavus* GI (Tn*6713*), while it was 6134 bp kb shorter and showed more rearrangements compared to Tn*6711* of strains A1-3 (Table 3 and Fig 3). ANIs were lowest (below 98%) between case A and B *E. faecium* GIs (S2 Table). Thus, the overall genetic differences between the GIs of A1-3 and B1-3, do not support a direct

Table 3. Characteristics of the GIs of the Norwegian VanD-type VRE.

Strain (case)	Genomic island				Repeats in the insertion site (5'-3' strand)	
	Name	GC content (%)	Size in bp	Number of CDSs	*lysS* side	16S rRNA side*
A1 (A)	Tn*6711*	44.1	125858	157	TTCCCAACAATGA	TTCCCGACAATGA
A2 (A)	Tn*6711*	44.1	125858	157	TTCCCAACAATGA	TTCCCGACAATGA
A3 (A)	Tn*6711*	44.1	125858	157	TTCCCAACAATGA	TTCCCGACAATGA
B1 (B)	Tn*6712*	44.3	119724	149	TTCCCGACAATGA	TTCCCAACAATGA
B2 (B)	Tn*6713*	44.6	112494	143	TTCCCAACAATGA	TTCCCCACAATGA
B3 (B)	Tn*6713*	44.7	112494	143	TTCCCCACAATGA	TTCCCCACAATGA

*, difference compared to repeat on the *lysS* side is indicated by underlined nucleotide

https://doi.org/10.1371/journal.pone.0255187.t003

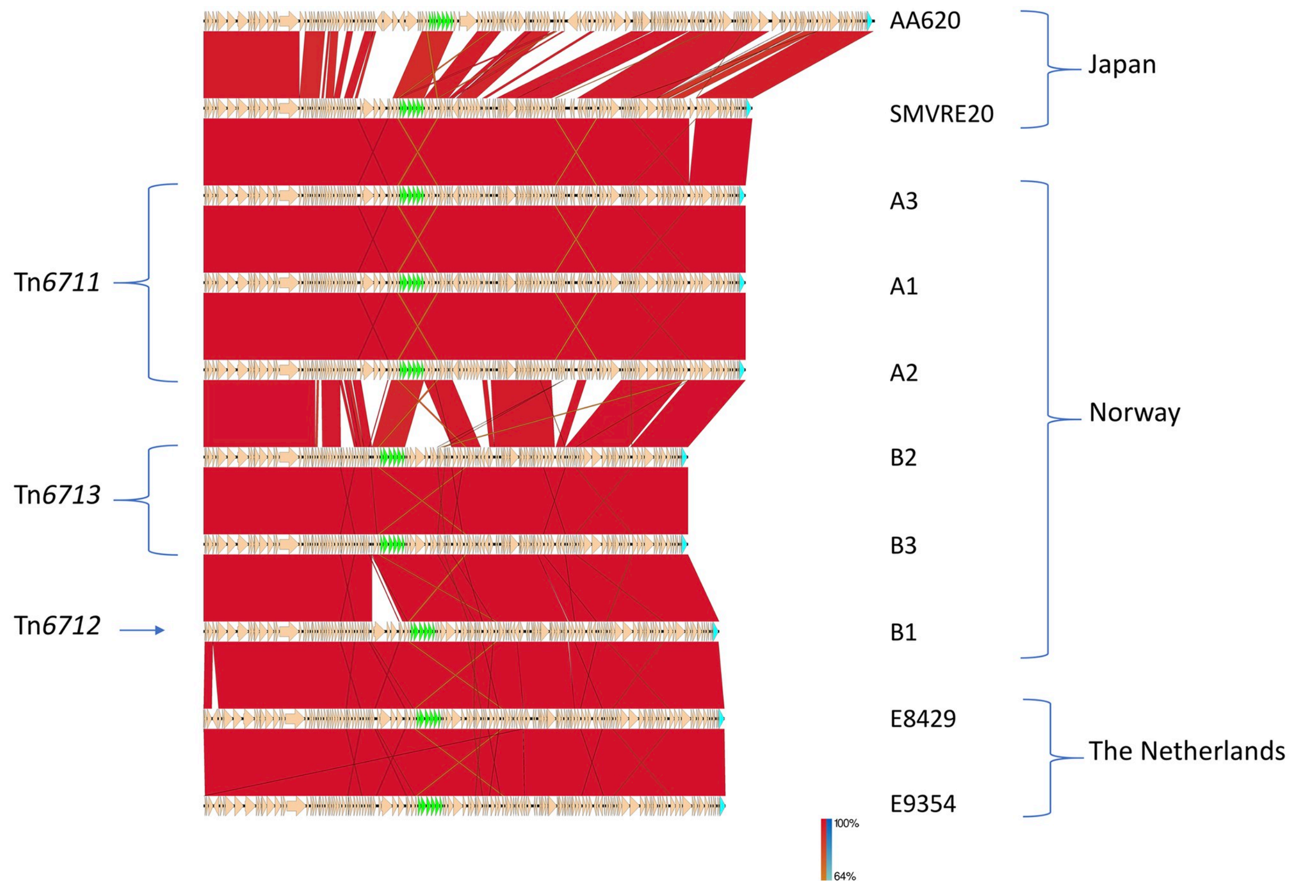


Fig 3. Comparison of the Norwegian, Dutch, and Japanese *vanD*-GIs built using Easyfig. A1-3 GIs have similar gene organization and showed high similarity with the Japanese SMVRE20 GI differing only in one hypothetical protein coding gene which contains transposase DDE domain. In case B, a high similarity exists between *E. casseliflavus* islands (B2 and B3) while the *E. faecium* island Tn*6712* of B1 is about 7.2 kb larger. The Dutch E8429 and E9354 showed the highest identity with case B GIs. *vanD* gene cluster and the integrase gene are marked in green and turquoise, respectively.

https://doi.org/10.1371/journal.pone.0255187.g003

spread between the two cases. However, in case B strains, we suggest one genetic event has evolved Tn*6713* of *E. casseliflavus* to the longer Tn*6712* in *E. faecium* or vice versa (Fig 3).

All GIs lacked conjugative apparatus genes and the *vanD* gene cluster was the only AMR gene within the islands (GenBank Acc. No. MT951615-7). The nucleotide sequence of integrase genes in Tn*6712* and Tn*6713* was identical and had only one SNP compared to Tn*6711*. Despite the existence of the same GIs in *E. casseliflavus* strains (B1 and B2) of case B, the ANI between their genomes (95.1%) was too low to be clonally related. This observation strongly suggests separate acquisitions of Tn*6713* in B2 and B3 strains.

Comparisons of the Norwegian *vanD*-GIs to those of the newly isolated VanD-type VRE*fm* from the Netherlands and Japan with publically available WGS data revealed a high rate of identity. Two VanD-type Dutch VRE strains (E8429 and E9354) [7] contained *vanD*-GIs with 99.99% sequence identity to Tn*6712* of B1. Moreover, the *vanD5*-containing GI from the Japanese *E. faecium* SMVRE20 [17] (AP019408.1) showed 99.98% sequence identity to Tn*6711* of case A. Another Japanese *vanD*-GI (157 kb) from *E. faecium* strain AA620 (LC467712.1) showed 96% identity covering 81% of Tn*6711*. Although the *vanD*-GIs are similar between the Norwegian, Dutch, and Japanese VRE*fm* strains, phylogenetic analyses based on SNPs suggest that the strains are not closely related (S6 Fig). The GI of the Japanese SMVRE20 has an additional gene compared to Tn*6711*. Likewise, Tn*6712* and the Dutch GIs show only one gene in

difference. Both these genes encode hypothetical proteins (Fig 3). The high identity between Tn*6711* and the GIs of the Japanese VanD-type VRE*fm* and between Tn*6712* and two Dutch VanD-type VRE*fm* GIs indicate a global spread of similar MGEs.

Due to the intrinsic *vanC* gene cluster of *E. casseliflavus* clinical strains, they already express low level resistance to vancomycin. Thus, *E. casseliflavus* strains often are not investigated further to see if they contain additional *van* clusters. In this study, we show that *E. casseliflavus* may be the intermediate source of the *vanD* type cluster containing GI (Tn*6713*) that spread to *E. faecium* (Tn*6712*) in case B. Based on this finding, MIC investigation of clinically important strains of *E. casseliflavus* should be considered to reveal possible acquired *van* gene clusters.

The GIs show site specific integration in *E. faecium* and *E. casseliflavus*

The insertion sites of the *vanD* GIs were identical for all six strains and located in the 3' end of the *lysS* gene which is positioned upstream of a 16S ribosomal rRNA gene. The integration resulted in a 13 bp direct repeat located 17 bp from the 3' end of the *lysS* gene. The left and right repeats in the different *vanD*-containing strains showed maximum one SNP difference. For case A GIs the imperfect direct repeats were identical. In strain B1 of case B, the repeat is identical to case A GIs but localised on opposite sides. The perfect direct repeat in B3 differed by one nucleotide compared to the other strains (Table 3). The same integration site was also found in the recently isolated Dutch and Japanese VanD-type VRE*fm* [7, 17]. Thus, this insertion site may be a hotspot in some enterococcal species including *E. faecium* and *E. casseliflavus*.

Putative origin of *vanD*-containing GIs

BLAST searches revealed 89% identity with several regions of *Blautia producta* SCSK genome covering only 59% of the Tn*6711* length. Another hit of Tn*6711* BLAST showed 89% identity to *Blautia coccoides* YL58 with 59% coverage, spanning some small fragments that were not covered by *B. producta* SCSK. An even higher identity (93%) was seen between the shorter Tn*6712* and Tn*6713* with fragments from *B. coccoides* YL58 covering 59% of these GIs. Previous reports have shown that *vanD*-type vancomycin resistance gene clusters can be found in non-enterococcal species like *Ruminococcus gauvreauii*, *Lachnospiraceae* bacterium, and *Ruthenibacterium lactatiformans* [7]. The above mentioned species and *Blautia* genus belong to the same taxonomic order of *Clostridiales* and are found in both the human and animal gut microbiome [47–49]. Thus, anaerobic *Blautia* genus or other members of the *Clostridiales* order are possible sources for *vanD* GIs.

Activity and transferability of putative GIs

Mobile chromosomal genetic elements, excise and circularize before transfer [50]. Circularization PCR and amplicon sequencing confirmed that Tn*6711* and Tn*6713* were able to circularize supporting that they are active MGEs. Agarose gel electrophoresis of PCR products repeatedly showed stronger bands for Tn*6713* in *E. casseliflavus* which could be due to higher activity compared to Tn*6711* in *E. faecium* (S7 Fig). However, we were not able to transfer *vanD* to an *E. faecium* recipient in this study (detection limit 10^{-10} to 10^{-9} transconjugants/donor cell) which is not surprising since a conjugation apparatus was not found in any of the GIs carrying the *vanD* gene clusters nor in other sites of the VanD-type VRE genomes. Type IV secretion systems play an important role in conjugation and can mediate the transfer of the conjugative plasmids and transposons. They have an impact on the spread of antimicrobial resistance among bacteria [29]. Non-conjugative MGEs can use the conjugative apparatus of other MGEs to mobilize. Thus, a mobility test can be conducted to confirm mobilization of the

GIs [38, 51]. However, the strains in this study already had several acquired resistance determinants that are used as markers in mobilization tests. Thus, we did not attempt to mobilize the islands.

Conclusions

We have performed a genetic characterization of the first VanD-type VRE strains recovered from two patients treated with broadspectrum antibiotics including vancomycin before VRE detection. All VanD-type VRE strains of case A were *E. faecium* while both *vanD E. casseliflavus* and *E. faecium* were recovered from case B. To our knowledge, this is the first two *vanD E. casseliflavus* strains reported. Based on our finding, we recommend MIC investigation of clinically important *E. casseliflavus* strains to reveal possible additional *van* gene clusters. In the VRE*fm* strains of case A, we identified a unique novel ST1486, an SLV of ST117, which were phylogenetically distant from case B VRE*fm* (ST203). Sequence analyses revealed a novel *vanD*-type cluster termed *vanD6* subtype in case B strains. The large phylogenetic distance between the VRE*fm* strain of the two cases, as well as differences in *vanD*-cluster subtypes and *vanD*-GIs, rejected the hypothesis of a clonal outbreak. We identified three novel similar *vanD*-GIs of putative *Clostridiales* order origin integrated at the same chromosomal site in both *E. faecium* and *E casseliflavus*.

Supporting information

S1 Fig. Antibiotic treatment and microbiological findings for case A. Tx: Transplantation, BAL: Bronchoalveolar lavage.
(TIF)

S2 Fig. Amino acid sequences alignment of the products deduced from the *ddl* genes of the *vanD*-containing *E. faecium* strains using Clustal omega online tool compared to the reference sequence (E1). Cov and pid represent the coverage and percent identity. The *ddl* gene of B1 showed a stop codon which resulted in a 110 amino acid protein. A1 and A2 showed a point mutation in a position involved in binding of D-Ala1 (S185 changed to F185) of the D-Ala:D-Ala ligase while A3 showed a point mutation in a position involved in binding of ATP (S319 changed to G319) [Depardieu F, Foucault M, Bell J, Dubouix A, Guibert M, Lavigne J, et al. New combinations of mutations in VanD-type vancomycin-resistant *Enterococcus faecium*, *Enterococcus faecalis*, and *Enterococcus avium* strains. Antimicrob Agents Chemother. 2009;53(5):1952–63]. The point mutations are highlighted by red boxes.
(TIF)

S3 Fig. Core genome SNP tree for the Norwegian *E. casseliflavus* strains and the available closed genomes of the species in the NCBI database on 04.04.2020.
(TIF)

S4 Fig. Extended core genome SNP tree for all *E. faecium* closed genomes retrieved from the NCBI database on 04.04.2020 and VRE*fm* of this study. The Norwegian samples are colored red and the closest genomes to them are in green.
(TIF)

S5 Fig. Genomic comparison between all Norwegian VanD-type VRE*fm* strains and *E. faecium* E1 reference genome using Easyfig tool. The red and blue gradient bars represent persent sequence matches. Red shows the direct and blue the inverted sequence matches. Arrows show the coding sequences and their direction. *vanD* gene cluster is marked in green. The similarities between case A strains (A1, A2 and A3) and their differences with case B VRE*fm* (B1)

is reflected in their machting patterns.
(TIF)

S6 Fig. Parsnp tree for the Norwegian, Dutch and Japanese VanD-type VRE*fm* genomes. Case A strains and the Japanese SMVRE20 which have the most identical GIs clustered separately. Likewise the Dutch E8429 and E9354 and B1 strain of case B also clustered separately.
(TIF)

S7 Fig. Agarose gel electrophoresis of PCR products using pairs of primers directed outwards from the GI ends to confirm the presence of the active form of the GIs (circular form). All but B1 contain the active form.
(TIF)

S1 Table. Average nucleotide identity between *vanD* gene cluster references (*vanD1*–*vanD5*) and the novel *vanD6* gene clusters from case B strains.
(DOCX)

S2 Table. Average nucleotide identity between GIs of the Norwegian VanD-type VRE samples.
(DOCX)

Acknowledgments

We would like to thank Ahmed Mekhlif, Kjersti Julin, and Theresa Wagner at UiT for their technical support, Ellen H. Josefsen, Bjørg Haldorsen, and Berit Harbak at the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance and Guro Lovise Rennan Ødegård at St Olavs Hospital for excellent technical assistance, Joachim Hegstad at University Hospital of North-Norway for his technical support in bioinformatics work, Janetta Top at University Medical Center Utrecht for providing the Dutch GI sequences, and the Genomics Support Center Tromsø™ and Statens Serum Institut in Denmark for genome sequencing.

Author Contributions

Conceptualization: Jessin Janice, Jørgen Vildershøj Bjørnholt, Arnfinn Sundsfjord, Kristin Hegstad.

Data curation: Mushtaq T. S. AL Rubaye, Jessin Janice, Aleksandra Jakovljev, Maria Elisabeth Hultström, Kristin Hegstad.

Formal analysis: Mushtaq T. S. AL Rubaye, Jessin Janice, Kristin Hegstad.

Funding acquisition: Jessin Janice, Jørgen Vildershøj Bjørnholt, Arnfinn Sundsfjord, Kristin Hegstad.

Investigation: Jørgen Vildershøj Bjørnholt, Aleksandra Jakovljev, Maria Elisabeth Hultström, Kristin Hegstad.

Methodology: Mushtaq T. S. AL Rubaye, Jessin Janice, Kristin Hegstad.

Project administration: Kristin Hegstad.

Software: Mushtaq T. S. AL Rubaye, Jessin Janice.

Supervision: Jessin Janice, Jørgen Vildershøj Bjørnholt, Arnfinn Sundsfjord, Kristin Hegstad.

Validation: Mushtaq T. S. AL Rubaye, Jessin Janice, Jørgen Vildershøj Bjørnholt, Aleksandra Jakovljev, Maria Elisabeth Hultström, Arnfinn Sundsfjord, Kristin Hegstad.

Visualization: Mushtaq T. S. AL Rubaye, Jessin Janice.

Writing – original draft: Mushtaq T. S. AL Rubaye, Jørgen Vildershøj Bjørnholt, Aleksandra Jakovljev, Maria Elisabeth Hultström.

Writing – review & editing: Jessin Janice, Jørgen Vildershøj Bjørnholt, Arnfinn Sundsfjord, Kristin Hegstad.

References

1. Patel R. Clinical impact of vancomycin-resistant enterococci. J Antimicrob Chemother. 2003; 51:13–22. https://doi.org/10.1093/jac/dkg272 PMID: 12801938
2. García-Solache M, Rice LB. The *Enterococcus*: a model of adaptability to its environment. Clin Microbiol Rev. 2019; 32:1–28.
3. Courvalin P. Vancomycin resistance in Gram-positive cocci. Clin Infect Dis. 2006; 42:25–34. https://doi.org/10.1086/491711 PMID: 16323116
4. Cattoir V, Leclercq R. Twenty-five years of shared life with vancomycin-resistant enterococci: is it time to divorce? J Antimicrob Chemother. 2013; 68:731–42. https://doi.org/10.1093/jac/dks469 PMID: 23208830
5. Pinholt M, Gumpert H, Bayliss S, Nielsen JB, Vorobieva V, Pedersen M, et al. Genomic analysis of 495 vancomycin-resistant *Enterococcus faecium* reveals broad dissemination of a *vanA* plasmid in more than 19 clones from Copenhagen, Denmark. J Antimicrob Chemother. 2017; 72:40–7. https://doi.org/10.1093/jac/dkw360 PMID: 27605596
6. Tomita H. Identification of VanD-type resistance pMG1-like conjugative plasmids in *E. faecium* clinical isolate in Japan. 5th Int Conf Enterococci. Chamonix, France; 2018.
7. Top J, Sinnige JC, Brouwer EC, Werner G, Corander J, Severin JA, et al. Identification of a novel genomic island associated with VanD-type vancomycin resistance in six Dutch vancomycin-resistant *Enterococcus faecium* isolates. Antimicrob Agents Chemother. 2018; 62:1–7. https://doi.org/10.1128/AAC.01793-17 PMID: 29311068
8. Boyd DA, Lalancette C, Lévesque S, Golding GR. Characterization of a genomic island harbouring a new *vanD* allele from *Enterococcus faecium* N15-508 isolated in Canada. J Antimicrob Chemother. 2016; 71:2052–4. https://doi.org/10.1093/jac/dkw063 PMID: 27084917
9. Depardieu F, Kolbert M, Pruul H, Bell J, Courvalin P. VanD-type vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis*. Antimicrob Agents Chemother. 2004; 48:3892–904. https://doi.org/10.1128/AAC.48.10.3892-3904.2004 PMID: 15388450
10. Depardieu F, Foucault M, Bell J, Dubouix A, Guibert M, Lavigne J, et al. New combinations of mutations in VanD-type vancomycin-resistant *Enterococcus faecium*, *Enterococcus faecalis*, and *Enterococcus avium* strains. Antimicrob Agents Chemother. 2009; 53:1952–63. https://doi.org/10.1128/AAC.01348-08 PMID: 19258279
11. Boyd DA, Kibsey P, Roscoe D, Mulvey MR. *Enterococcus faecium* N03-0072 carries a new VanD-type vancomycin resistance determinant: characterization of the *vanD5* operon. J Antimicrob Chemother. 2004; 54:680–3. https://doi.org/10.1093/jac/dkh391 PMID: 15308604
12. Flipse J, von Wintersdorff CJH, van Niekerk JM, Jamin C, van Tiel FH, Hasman H, et al. Appearance of *vanD*-positive *Enterococcus faecium* in a tertiary hospital in the Netherlands: prevalence of *vanC* and *vanD* in hospitalized patients. Sci Rep. 2019; 9:6949. https://doi.org/10.1038/s41598-019-42824-4 PMID: 31061446
13. Terada S, Harada T, Yokota M, Tsuchiya T, Adachi K, Asaka T, et al. First isolation and characterization of vancomycin-resistant *Enterococcus faecium* harboring *vanD5* gene cluster recovered from a 79-year-old female inpatient in Japan. Diagnostic Microbiol Infect Dis. 2019; 95:1–6. https://doi.org/10.1016/j.diagmicrobio.2019.114883 PMID: 31495527
14. Starlander G, Tellgren-Roth C, Melhus Å. Fatal acquisition of *vanD* gene during vancomycin treatment of septicaemia caused by *Enterococcus faecium* considerations and difficulties of stethoscope disinfection for a motivated healthcare worker. J Hosp Infect. 2016; 92:409–10. https://doi.org/10.1016/j.jhin.2016.01.002 PMID: 26876745
15. Lam PW, Kozak RA, Eshaghi A, Avaness M, Salt N, Patel SN, et al. Nosocomial outbreak of *vanD*-carrying vancomycin-resistant *Enterococcus faecium*. Infect Control Hosp Epidemiol. 2018; 39:1266–8. https://doi.org/10.1017/ice.2018.174 PMID: 30111383

16. Ostrowsky BE, Clark NC, Thauvin-Eliopoulos C, Venkataraman L, Samore MH, Tenover FC, et al. A cluster of VanD vancomycin-resistant *Enterococcus faecium*: molecular characterization and clinical epidemiology. J Infect Dis. 1999; 180:1177–85. https://doi.org/10.1086/315030 PMID: 10479146

17. Sato T, Wada T, Shinagawa M, Fukushima Y, Nakajima C, Suzuki Y. Emergence of vancomycin- and teicoplanin-resistant *Enterococcus faecium* via *vanD5*-harbouring large genomic island. J Antimicrob Chemother. 2020; 75:2411–5. https://doi.org/10.1093/jac/dkaa220 PMID: 32585683

18. Camargo ILBC, Dalla Costa LM, Woodford N, Gilmore MS, Darini ALC. Sequence analysis of *Enterococcus faecium* strain 10/96A (VanD4), the original vancomycin-resistant *E. faecium* strain in Brazil. J Clin Microbiol. 2006; 44:2635–7. https://doi.org/10.1128/JCM.00509-06 PMID: 16825401

19. Clinical breakpoints—breakpoints and guidance [Internet]. Eur. Comm. Antimicrob. Susceptibility Test. —EUCAST. 2020. https://eucast.org/clinical_breakpoints/.

20. Depardieu F, Perichon B, Courvalin P. Detection of the *van* alphabet and identification of Enterococci and Staphylococci at the species level by multiplex PCR. J Clin Microbiol. 2004; 42:5857–60. https://doi.org/10.1128/JCM.42.12.5857-5860.2004 PMID: 15583325

21. Fines M, Perichon B, Reynolds P, Sahm DF, Courvalin P. VanE, a new type of acquired glycopeptide resistance in *Enterococcus faecalis* BM4405. Antimicrob Agents Chemother. 1999; 43:2161–4. https://doi.org/10.1128/AAC.43.9.2161 PMID: 10471558

22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014; 30:2114–20. https://doi.org/10.1093/bioinformatics/btu170 PMID: 24695404

23. Nurk S, Bankevich A, Antipov D, Gurevich A, Korobeynikov A, Lapidus A, et al. Assembling genomes and mini-metagenomes from highly chimeric reads. In: Deng M, Jiang R, Sun F, Zhang X, editors. Res Comput Mol Biol. Berlin, Heidelberg: Springer Berlin Heidelberg; 2013. p. 158–70.

24. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013; 29:1072–5. https://doi.org/10.1093/bioinformatics/btt086 PMID: 23422339

25. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki P, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016; 44:6614–24. https://doi.org/10.1093/nar/gkw569 PMID: 27342282

26. Feldgarden M, Brover V, Haft DH, Prasad AB, Slotta DJ, Tolstoy I, et al. Validating the AMRFinder tool and resistance gene database by using antimicrobial resistance genotype-phenotype correlations in a collection of isolates. Antimicrob Agents Chemother. 2019; 63:e00483–19. https://doi.org/10.1128/AAC.00483-19 PMID: 31427293

27. Seemann T. Abricate [Internet]. Github. https://github.com/tseemann/abricate.

28. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990; 215:403–10. https://doi.org/10.1016/S0022-2836(05)80360-2 PMID: 2231712

29. Han N, Yu W, Qiang Y, Zhang W. T4SP database 2.0: an improved database for type IV secretion systems in bacterial genomes with new online analysis tools. Comput Math Methods Med. 2016; 2016:9415459. https://doi.org/10.1155/2016/9415459 PMID: 27738451

30. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014; 15:524. https://doi.org/10.1186/s13059-014-0524-x PMID: 25410596

31. Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics. 2010; 11:595. https://doi.org/10.1186/1471-2105-11-595 PMID: 21143983

32. de Been M, Pinholt M, Top J, Bletz S, Mellmann A, van Schaik W, et al. Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. J Clin Microbiol. 2015; 53:3788–97. https://doi.org/10.1128/JCM.01946-15 PMID: 26400782

33. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004; 14:1394–403. https://doi.org/10.1101/gr.2289704 PMID: 15231754

34. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011; 27:1009–10. https://doi.org/10.1093/bioinformatics/btr039 PMID: 21278367

35. Carver TJ, Rutherford KM, Berriman M, Rajandream M, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005; 21:3422–3. https://doi.org/10.1093/bioinformatics/bti553 PMID: 15976072

36. Pritchard L, Glover RH, Humphris S, Elphinstone JG, Toth IK. Genomics and taxonomy in diagnostics for food security: soft-rotting enterobacterial plant pathogens. Anal Methods. 2016;12–24.

37. Tansirichaiya S, Rahman A, Roberts AP. The Transposon registry. Mob DNA. Mobile DNA; 2019; 10:1–6.

38. Sivertsen A, Janice J, Pedersen T, Wagner TM, Hegstad J, Hegstad K. The Enterococcus cassette chromosome, a genomic variation enabler in enterococci. mSphere. 2018; 3:1–13. https://doi.org/10.1128/mSphere.00402-18 PMID: 30404935

39. Poyart C, Trieu-Cuot P. Heterogeneric conjugal transfer of the pheromone-responsive plasmid pIP964 (IncHlyI) of *Enterococcus faecalis* in the apparent absence of pheromone induction. FEMS Microbiol Lett. 1994; 122:173–80. https://doi.org/10.1111/j.1574-6968.1994.tb07161.x PMID: 7958769

40. Elstrøm P, Astrup E, Hegstad K, Samuelsen Ø, Enger H, Kacelnik O. The fight to keep resistance at bay, epidemiology of carbapenemase producing organisms (CPOs), vancomycin resistant enterococci (VRE) and methicillin resistant *Staphylococcus aureus* (MRSA) in Norway, 2006–2017. PLoS One. 2019; 14:e0211741. https://doi.org/10.1371/journal.pone.0211741 PMID: 30716133

41. Willems RJL, Top J, van Schaik W, Leavis H, Bonten M, Sirén J, et al. Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. MBio. 2012; 3:e00151–12. https://doi.org/10.1128/mBio.00151-12 PMID: 22807567

42. Depardieu F, Reynolds PE, Courvalin P. VanD-type vancomycin-resistant *Enterococcus faecium* 10/96A. Antimicrob Agents Chemother. 2003; 47:7–18. https://doi.org/10.1128/AAC.47.1.7-18.2003 PMID: 12499162

43. Casadewall B, Courvalin P. Characterization of the *vanD* glycopeptide resistance gene cluster from *Enterococcus faecium* BM4339. J Bacteriol. 1999; 181:3644–8. https://doi.org/10.1128/JB.181.12.3644-3648.1999 PMID: 10368136

44. Tedim AP, Lanza VF, Manrique M, Pareja E, Ruiz-Garbajosa P, Cantón R, et al. Complete genome sequences of isolates of *Enterococcus faecium* sequence type 117, a globally disseminated multidrug-resistant clone. Genome Announc. 2017; 5:e01553–16. https://doi.org/10.1128/genomeA.01553-16 PMID: 28360174

45. Qin X, Galloway-peña JR, Sillanpaa J, Roh JH, Nallapareddy SR, Chowdhury S, et al. Complete genome sequence of *Enterococcus faecium* strain TX16 and comparative genomic analysis of *Enterococcus faecium* genomes. BMC Microbiol. 2012; 12:1–20.

46. Gan YQI, Zhang TAO, Gan YQ, Zhao Z, Zhu BIN. Complete genome sequences of two *Enterococcus faecium* strains and comparative genomic analysis. Exp Ther Med. 2020; 19:2019–28. https://doi.org/10.3892/etm.2020.8447 PMID: 32104261

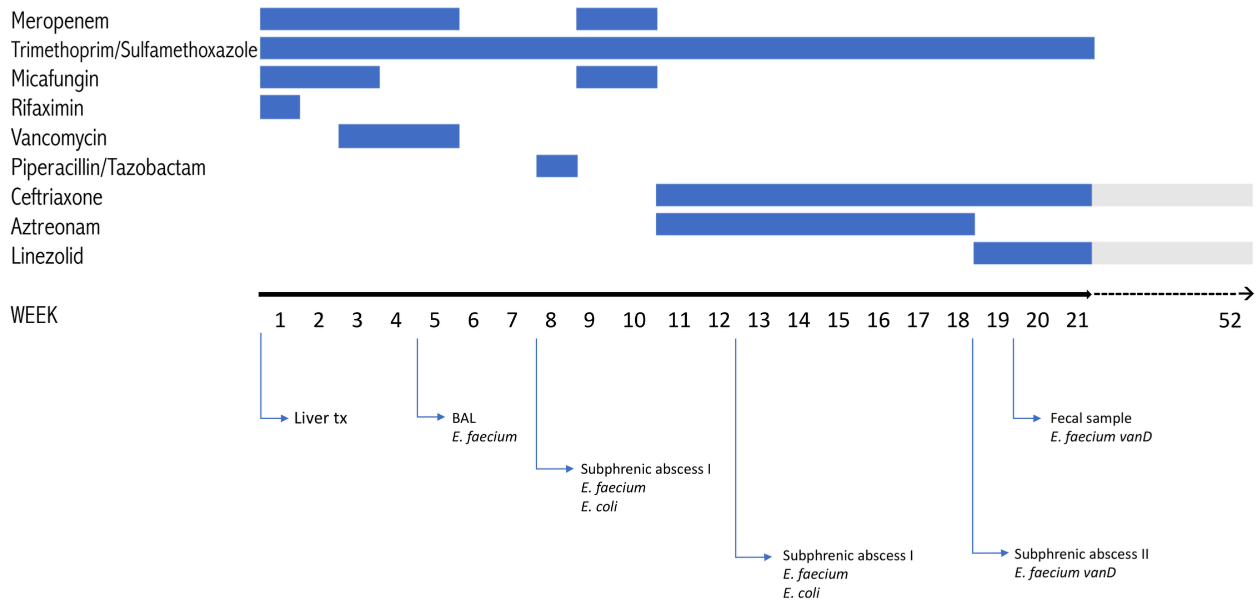
47. Shkoporov AN, Chaplin A V, Shcherbakova VA, Suzina NE, Kafarskaia LI, Bozhenko VK, et al. *Ruthenibacterium lactatiformans* gen. nov., sp. nov., an anaerobic, lactate-producing member of the family *Ruminococcaceae* isolated from human faeces. Int J Syst Evol Microbiol. 2016; 66:3041–9. https://doi.org/10.1099/ijsem.0.001143 PMID: 27154556

48. Kameyama K, Itoh K. Intestinal colonization by a *Lachnospiraceae* bacterium contributes to the development of diabetes in obese Mice. Microbes Env. 2014; 29:427–30. https://doi.org/10.1264/jsme2.ME14054 PMID: 25283478

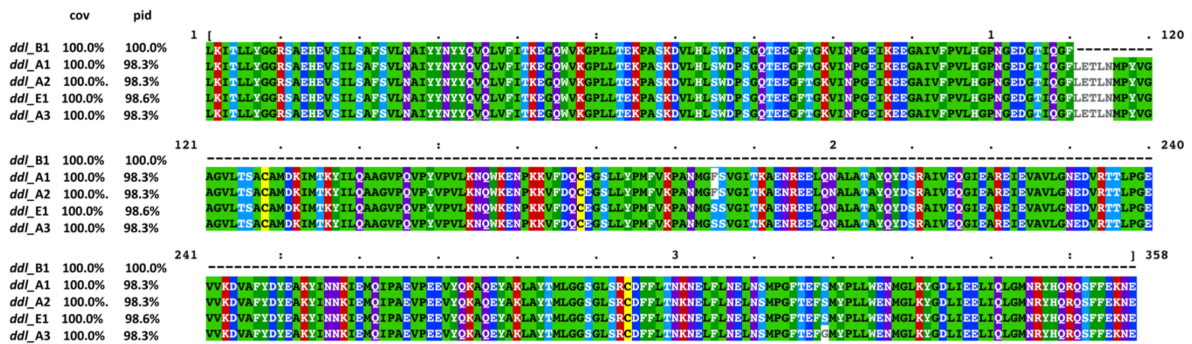
49. Domingo M, Huletsky A, Boissinot M, Bernard KA, Picard FJ, Bergeron MG. *Ruminococcus gauvreauii* sp. nov., a glycopeptide-resistant species isolated from a human faecal specimen. Int J Syst Evol Microbiol. 2008; 58:1393–7. https://doi.org/10.1099/ijs.0.65259-0 PMID: 18523184

50. Wozniak RAF, Waldor MK. Integrative and conjugative elements: mosaic mobile genetic elements enabling dynamic lateral gene flow. Nat Rev Microbiol. 2010; 8:552–563. https://doi.org/10.1038/nrmicro2382 PMID: 20601965

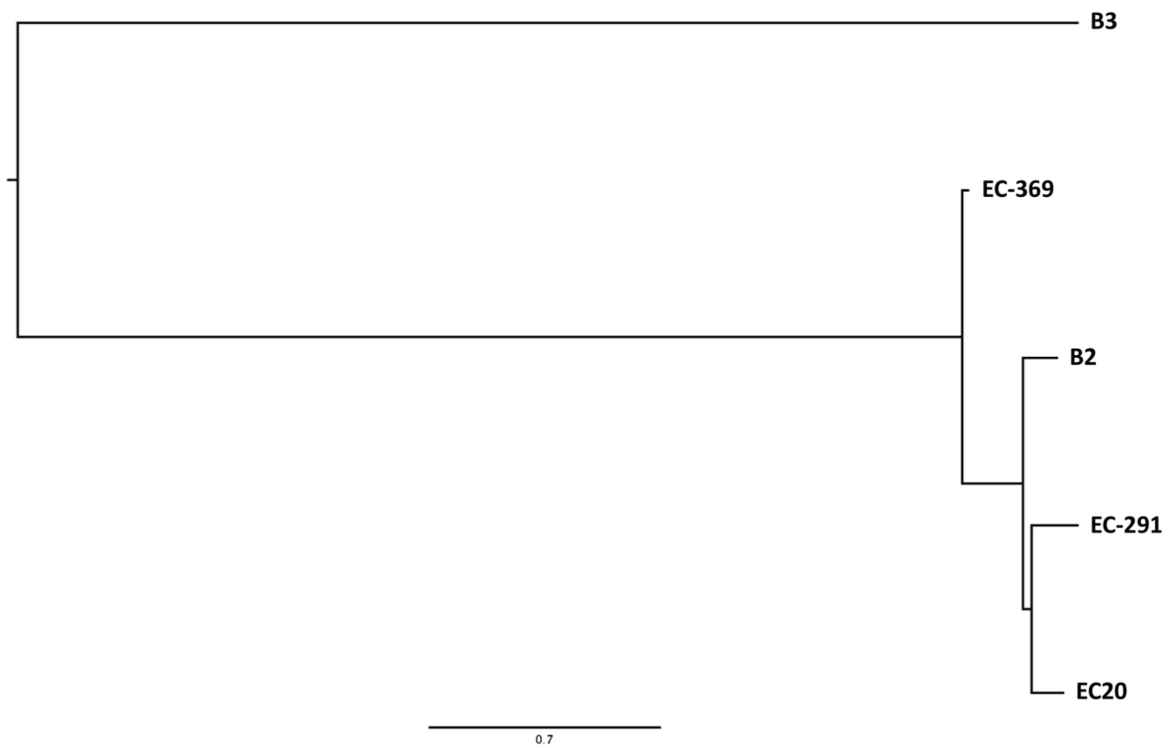
51. Ramsay JP, Firth N. Diverse mobilization strategies facilitate transfer of non-conjugative mobile genetic elements. Curr Opin Microbiol. 2017; 38:1–9. https://doi.org/10.1016/j.mib.2017.03.003 PMID: 28391142



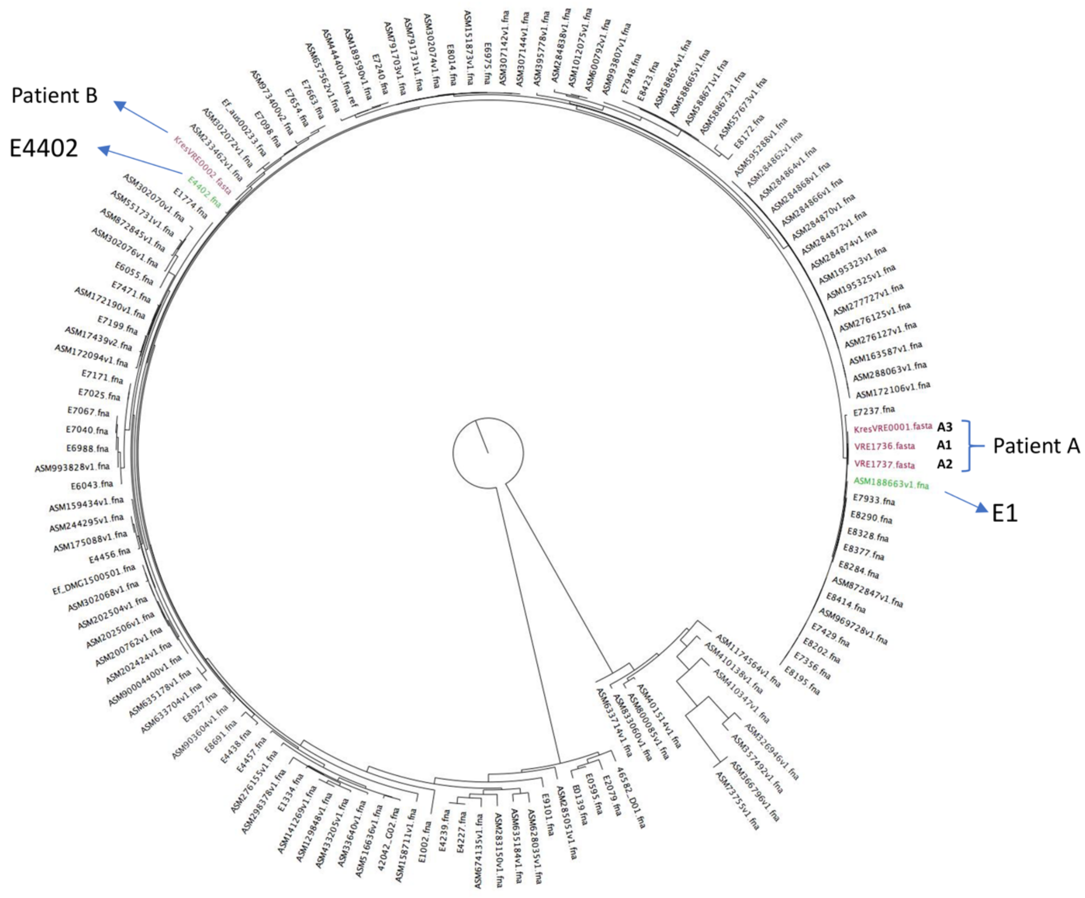
S1 Fig. Antibiotic treatment and microbiological findings for case A.



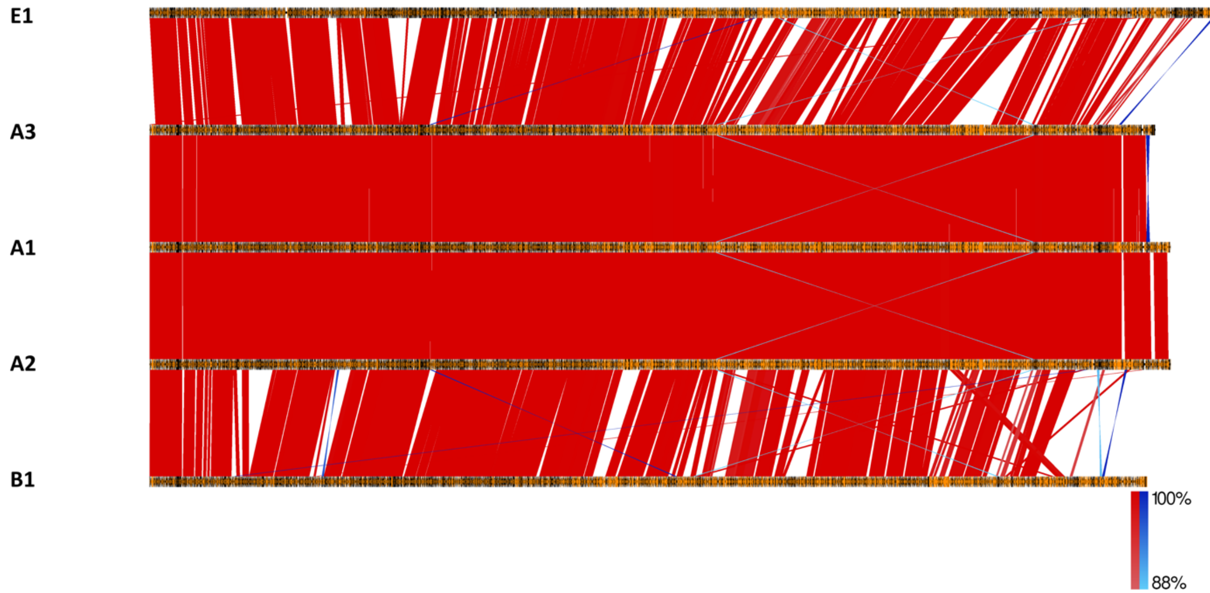
S2 Fig. Amino acid sequences alignment of the products deduced from the *ddl* genes of the *vanD*-containing *E. faecium* strains using Clustal omega online tool compared to the reference sequence (E1).



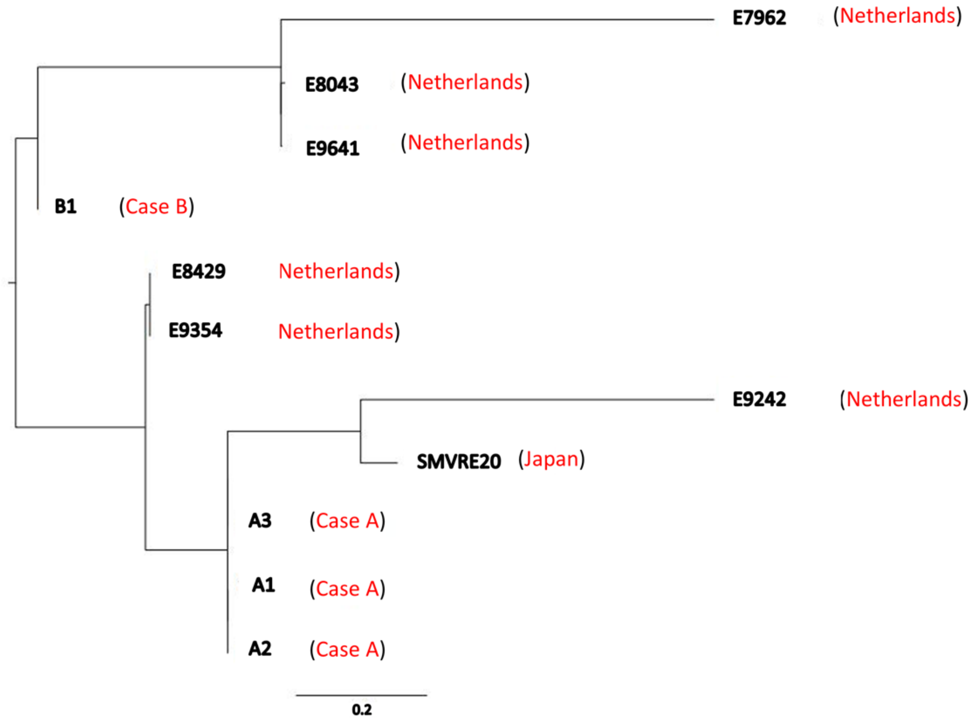
S3 Fig. Core genome SNP tree for the Norwegian *E. casseliflavus* strains and the available closed genomes of the species in the NCBI database on 04.04.2020.



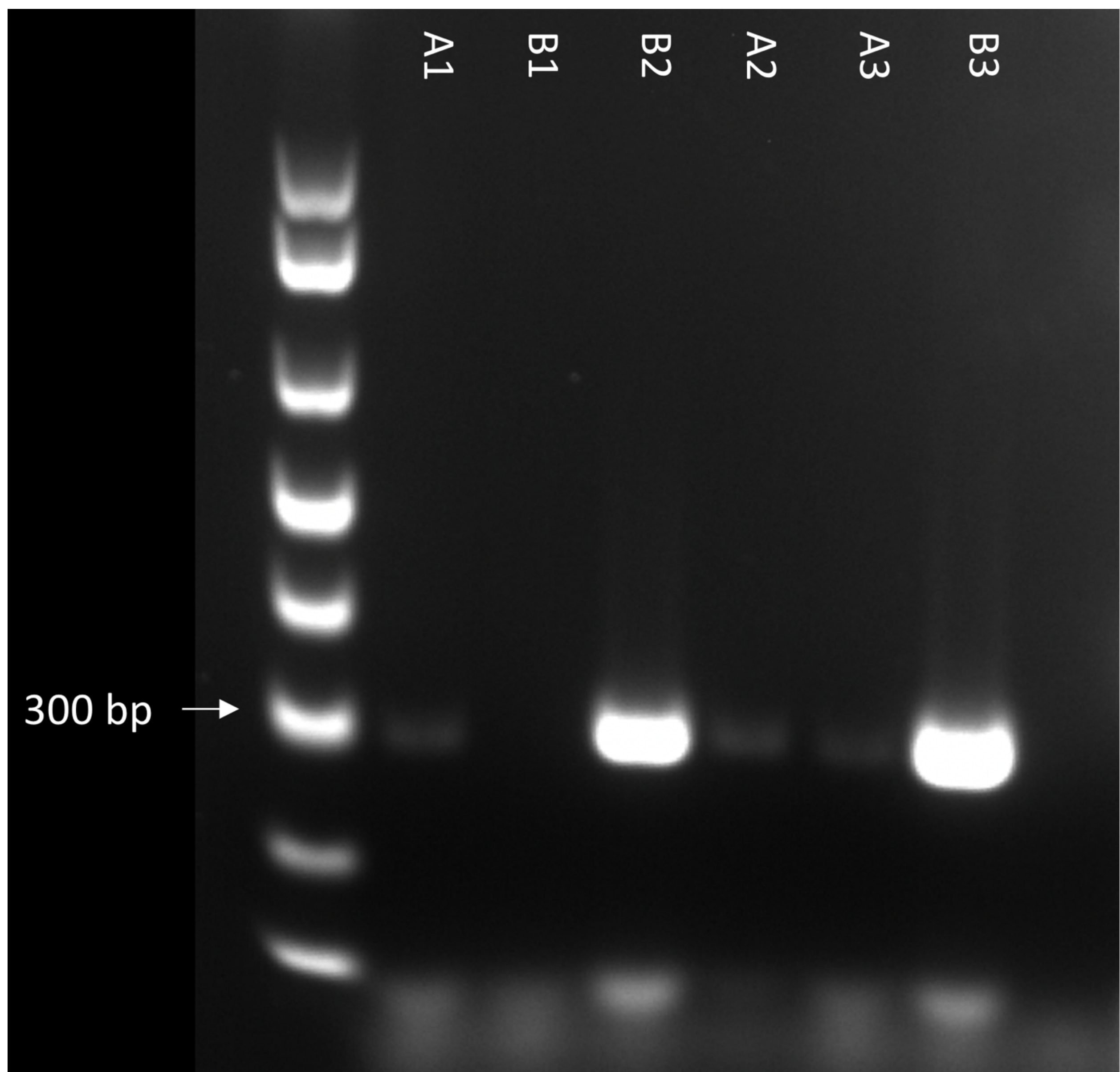
S4 Fig. Extended core genome SNP tree for all *E. faecium* closed genomes retrieved from the NCBI database on 04.04.2020 and VRE*fm* of this study.



S5 Fig. Genomic comparison between all Norwegian VanD-type VRE*fm* strains and *E. faecium* E1 reference genome using Easyfig tool.



S6 Fig. Parsnp tree for the Norwegian, Dutch and Japanese VanD-type VRE*fm* genomes.



S7 Fig. Agarose gel electrophoresis of PCR products using pairs of primers directed outwards from the GI ends to confirm the presence of the active form of the GIs (circular form).

S1 Table. Average nucleotide identity between *vanD* gene cluster references (*vanD1*–*vanD5*) and the novel *vanD6* gene clusters from patient B strains.

	***vanD1* Id (%)**	***vanD2* Id (%)**	***vanD3* Id (%)**	***vanD4* Id (%)**	***vanD5* Id (%)**
B1 *vanD6*	92.300	93.725	92.633	87.700	91.595
B2 *vanD6*	92.317	93.742	92.650	87.721	91.612
B3 *vanD6*	92.300	93.726	92.633	87.700	91.595

S2 Table. Average nucleotide identity between GIs of the Norwegian VanD-type VRE samples.

	A1 Tn*6711* Id (%)	**A2 Tn*6711* Id (%)**	**A3 Tn*6711* Id (%)**	**B1 Tn*6712* Id (%)**	**B2 Tn*6713* Id (%)**	**B3 Tn*6713* Id (%)**
A1 Tn*6711*	100	99.999	99.997	97.823	97.680	97.678
A2 Tn*6711*	99.999	100	99.998	97.824	97.681	97.679
A3 Tn*6711*	99.997	99.998	100	97.824	97.682	97.681
B1 Tn*6712*	97.823	97.824	97.824	100	99.998	99.997
B2 Tn*6713*	97.680	97.681	97.682	99.998	100	99.999
B3 Tn*6713*	97.678	97.679	97.681	99.997	99.999	100

Paper II

The first *vanE*-type vancomycin resistant *Enterococcus faecalis* isolates in Norway – phenotypic and molecular characteristics

Mushtaq AL Rubaye[1], Jessin Janice[1,2#], Jørgen Vildershøj Bjørnholt[3,4], Iren H. Löhr[5,6], Arnfinn Sundsfjord[1,2], Kristin Hegstad[1,2*]

[1]Research group for Host-Microbe Interactions, Department of Medical Biology, UiT the Arctic University of Norway, Tromsø, Norway,
[2]Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway,
[3]Department of Clinical Microbiology, Oslo University Hospital, Oslo, Norway,
[4]Institute of Clinical Medicine, University of Oslo, Oslo, Norway,
[5]Department of Medical Microbiology, Stavanger University Hospital, Stavanger, Norway
[6]Department of Clinical Science, Faculty of Medicine, University of Bergen, Bergen, Norway

* Corresponding author: Kristin Hegstad, Kristin.Hegstad@uit.no
[#] Current address: Section for Development, Department of Microbiology, Clinic for Laboratory Medicine, Oslo University Hospital, Oslo, Norway.

Keywords:
Vancomycin resistant enterococci, *Enterococcus faecalis*, *vanE*, Mobile genetic element, cgMLST, Tn*6202*

The first Norwegian *vanE*-VRE*fs* isolates

ABSTRACT

Background

Vancomycin resistance in enterococci is mainly caused by the acquisition of *van* gene clusters. The *vanE*-type is unusual among the ten known *van* gene clusters in enterococci and is associated with low-level vancomycin resistance and susceptibility to teicoplanin. The aim of this study was to assess the relatedness between the first two Norwegian *vanE*-type isolates and characterize the *vanE* gene cluster and its surroundings.

Material & Methods

The *vanE*-type vancomycin resistant *Enterococcus faecalis* isolates (E1 and E2) were recovered from one patient 30 months apart. Both isolates were whole genome sequenced using Illumina and PacBio. Detailed comparative analyses of their genomes investigating structural variation, *vanE* gene clusters and mobile genetic context were performed. The isolates were also examined for antimicrobial susceptibility, *vanE* transfer, expression of the histidine kinase part of the *vanS_E* gene, inducibility of vancomycin resistance and mutation rate.

Results

Both strains expressed low-level vancomycin resistance (MIC=16) and susceptibility to teicoplanin (MIC= 0.5). The E1 and E2 *vanE* gene clusters were part of a non-transferable Tn*6202* identical in all coding sequences except the $vanS_E$ gene. Although a premature stop codon truncated $vanS_E$ in E1 and insertion of an IS*6770* transposase truncated the gene in E2 the downstream histidine kinase part of the $vanS_E$ gene was expressed. The vancomycin resistance phenotype in E1 was inducible low level while constitutive low level in E2. E1 showed a 125-fold higher mutation rate than E2 and variant calling showed 60 variants between them, but they still belonged to the same sequence type and showed nearly identical chromosomal gene content and synteny.

Conclusions

In this study, we present *vanE*-type VRE for the first time in a whole genome context and described the chromosomal insertion site of the *vanE*-conferring Tn*6202*. Despite some differences in their genome which can be explained by the high mutation rate of E1 and acquisition of different mobile genetic elements, we believe the two isolates are related.

Introduction

The opportunistic pathogen *Enterococcus faecalis* (*Efs*) is responsible for more than half of enterococcal infections in which vancomycin is a valid treatment option in cases of penicillin allergy and resistance to other antibiotics [1]. In vancomycin resistant enterococci (VRE), the *van* gene clusters encode mechanisms that replace the D-Ala-D-Ala terminal side-chain residues of the peptidoglycan precursors and thus reduce the binding affinity of vancomycin to the cell wall. Of ten known vancomycin resistant gene clusters *vanC, E, G, L* and *N* change the terminal side-chain to D-Ala-D-Ser, while *vanA, B*, *D, M,* and *P* change it to D-Ala-D-Lac [1,2]. D-Alanine-D-Serine side-chain residue results in seven folds lower binding affinity by vancomycin. The *vanE* cluster confers low-level resistance to vancomycin and susceptibility to teicoplanin [3].

VanE-type vancomycin resistance in enterococci has so far been described in *Enterococcus faecalis* (*Efs)*, chromosomally encoded [4–7]. Among *van* gene clusters the *vanA* and *vanB* are the globally dominant, while *vanE* is one of the uncommon *van*-types, with a few reports in *Efs* from North America and Australia [4,6,8,9]. The *vanE* gene cluster consists of five genes; *vanE* (D-Ala–D-Ser ligase), $vanXY_E$ (D,D-dipeptidase/D,D-carboxypeptidase), and $vanT_E$ (serine racemase), and the regulatory operon with $vanR_E$ and $vanS_E$ [4]. VanR and VanS work as a two-component signal transduction system and regulate the expression of *van* genes in response to the extracellular glycopeptide antibiotics [10]. Glycopeptides such as vancomycin can induce autophosphorylation in the membrane bound sensor protein VanS, which consequently phosphorylates the transcription activator VanR resulting in inducible vancomycin resistance [11]. However, constitutive expression of vancomycin resistance has been reported in a VanE strain with a truncated $vanS_E$ gene [6].

Although WGS is increasingly used in VRE studies [12], none of the previously reported *vanE*-type VRE *faecalis* (*vanE*-VRE*fs*) strains have been analysed by WGS [4,5,13]. The use of WGS facilitates disclosure of microevolutionary relationships among isolates [14] which describes the evolutionary forces that genetically diversify the natural populations of bacteria [15]. The study of single nucleotide polymorphism (SNP), and core genome multilocus sequence typing (cgMLST) besides comparative genomics to confirm gain or loss of MGE are useful methods in pathogen microevolutionary analyses [16].

We describe the antimicrobial susceptibility phenotype, the genetic relatedness between the first two *vanE*-VRE*fs* isolates in Norway, their *vanE* gene cluster, and genetic support. Since the isolates showed different truncations in the $vanS_E$ leading to two coding sequences (CDSs)

instead of one, the inducibility of vancomycin resistance and expression of the histidine kinase part of the *vanS*$_E$ gene was also investigated.

Material and Methods

Case description and data collection

In 2016, an elderly male presented at a hospital in Southern Norway with lower back pain and underwent surgery for a lumbar prolapse. He had several comorbidities such as chronic obstructive pulmonary disease (COPD) and chronic renal failure. He had undergone an urostomy after cystoprostatectomy five years earlier. During the hospital stay, *vanE*-VRE*fs* (isolate E1) was detected in a urine culture sample (Table 1). The patient had no symptoms indicating a urinary tract infection, and the urostomy was considered colonized by *vanE*-VRE*fs*. The patient was therefore not treated with antibiotics and discharged from the hospital after seven days. He was readmitted twice to the hospital after ten months and fifteen months. During both hospital stays rectal and urine samples screened negative for VRE. During the last hospitalization, he received cefotaxime i. v. due to COPD exacerbation. Thirty months after the first detection a second *vanE*-VRE*fs* (isolate E2) was detected in a sample from a chronic wound obtained by the patient's general practitioner. No antibiotics were prescribed. Both isolates were sent to the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance for further analysis.

Table 1. Relevant characteristics of *vanE*-VRE*fs* isolates

Isolate	Species	ST	Source	Year	VAN*	TEC	AMP	GEN	LIN
E1	*E. faecalis*	34	Urine	2016	16	0.5	1	<32	2
E2	*E. faecalis*	34	Wound	2019	16	<0.5	2	<32	2

*MICs in mg/L for VAN (vancomycin), TEC (teicoplanin), AMP (ampicillin), GEN (gentamicin) and LIN (linezolid).

Species identification, antimicrobial susceptibility testing (AST), and *van* genotype determinations

For initial species identification, MALDI-TOF (Bruker Daltonik GmbH, Bremen, Germany) was performed according to the manufacturer's protocol. The Sensititre EUENCF plate (Thermo Scientific, Massachusetts, USA) as well as ComASP™ Vancomycin (Liofilchem,

Roseto Degli Abruzzi, Italy) were used to determine the minimum inhibitory concentrations (MICs) of vancomycin, teicoplanin, ampicillin, gentamicin and linezolid. The agar dilution method was used to determine rifampicin MIC [17]. For fusidic acid MIC determination Test Strips (Liofilchem, Roseto Degli Abruzzi, Italy) were used. MICs were interpreted according to EUCAST clinical breakpoints v.12.0. PCR was performed for initial *van*-typing using primers described previously [5,18].

DNA extraction, library preparation, and genome sequencing

To study the genomes of E1 and E2 isolates, their relatedness, and identify their MGE harbouring *vanE* gene cluster, both short- and long-read sequencing were performed. For short-read sequencing, genomic DNA extraction and quantification were performed as described previously [19]. Samples were sequenced at the Genomic Support Center Tromsø™ using NextSeq550 system of Illumina platform in Mid Output 300 cycles cell, which resulted in reads sized 35-151 bp. For long-read sequencing, the genomic DNA was extracted using Wizard Genomic DNA Purification Kit (Promega, Madison, USA). Genomic DNA quantification was done as described for Illumina sequencing. Long-read sequencing was performed in the Norwegian Sequencing Centre (NSC) using the SMRT cell 8M (Sequel II) of PacBio platform, which resulted in reads sized from 10 to 20 kb.

PacBio reads assembly and analyses (Q20 reads)

PacBio circular consensus sequence reads (Q20) were assembled using Unicycler v0.39 [20]. Then the quality of the assemblies was assessed by Quast tool v5.0.2 [21]. Prokka was used to annotate the assemblies [22]. Antimicrobial resistance (AMR) genes and plasmids were identified in the assemblies using bacterial antimicrobial resistance reference gene database (PRJNA313047) and PlasmidFinder v 2.0.1 database in ABRicate v0.8 tool [23].

Illumina reads trimming and mapping reads

Quality trimming and adapter removal of raw reads generated from Illumina sequencing were performed by trimmomatic v0.39, [24] and the quality of trimmed reads was assessed by fastQC v0.11.8 [25]. To ensure the quality of the PacBio assemblies, the trimmed paired Illumina reads were mapped on their respective PacBio assemblies. The reads were mapped using the mem algorithm in BWA tool v07.17 [26] and sorting was performed in SAMtools v1.10 [27].

Variant-calling and sequence-typing

Snippy v4.4.0 [28] was used for variant-calling between E1 and E2 isolates. Then to predict the effect of variants, variant annotation was carried out using the SNPeff tool v4.3t [29]. The minimum spanning tree (MST) was built using SeqSphere+ software v6.0.2 (Ridom GmbH, Münster, Germany [http://www.ridom.de/seqsphere/]). The MST was constructed based on 1972 gene targets of the *Efs* cgMLST scheme. The default of ≤ seven allelic differences was used as a threshold for cluster formation as defined previously [30]. To identify and locate the mobile genetic structures in the *vanE*-VRE*fs* genomes, their sequences were compared using BLASTn v2.6.0 and the comparison was visualized in Artemis Comparison Tool (ACT) v18.1.0 [31]. Additionally, Easyfig was used to visualize the comparisons between the sequences [32].

Phylogenetic analysis

To investigate the phylogenetic relationship of E1 and E2 to the global *Efs* population, all *Efs* closed genomes deposited on NCBI as of 01.08.2022 (n=458) were retrieved. A global tree based on the core genome SNP was generated using Parsnp v1.2 [33].

Excision and transferability of MGEs

To explore the ability of the MGE to circularize before the transfer, a PCR approach with pairs of primers directed outwards from their ends (Forward: 5´-TGGATTCCTGCATCAACAGA-3´ and Reverse 5´-TTGCCAATGATAAACGCTGA-3´) was carried out. Moreover, filter-mating method [34,35] was performed to determine the transferability of MGE harbouring *vanE* gene cluster to two vancomycin susceptible *Efs* strains (JH2-2 and OG1-RF) [36,37]. The transfer frequency (transconjugants frequency per donor cell) was calculated by counting the colony forming units on the BHI agar containing vancomycin (8 mg/L) and/or rifampicin (20 mg/L) and fusidic acid (10 mg/L).

Assessment of vancomycin resistance induction

To determine the inducibility of vancomycin resistance, the changes in the generation time (Tg) were calculated according to the growth rate assessment of the isolates with or without vancomycin. Growth rate comparison was performed for isolates cultured in sub-MIC concentrations (4mg/L (corresponding to ¼ of the MIC) and 10 mg/L) versus without vancomycin as described previously [8,9]. Three biological and technical replicates were tested for each strain. The optical density was measured every 30 minutes at 600 nm during the 24

hours of incubation in Synergy H1/Biospa microplate reader (Biotek Instruments, Winooski, VT, USA).

Assessment of *$vanS_E$* gene expression

Expression of the *$vanS_E$* gene after growth of the *vanE*-VRE*fs* in BHI medium overnight without and with vancomycin (10 mg/L) was assessed in three biological and technical replicates. RNA was extracted using Qiagen bacterial RNA kit (Warrington, UK) and then genomic DNA was removed using the Heat&Run gDNA removal kit (ArcticZymes, Tromsø, Norway). Next, complementary DNA (cDNA) was synthesized using Applied Biosystems™ High-Capacity cDNA Reverse Transcription Kit (Life Technologies Corporation, Carlsbad, CA, USA). The resulting cDNA was quantified by qPCR using SYBR® Green fluorescent dye PCR mix (Applied Biosystem, Kingsland Grange, UK) on a 7300 Real-time PCR System (Applied Biosystems). The *gyrB* was used as a reference for gene expression normalization. Then the mRNA levels were calculated using the $2^{-\Delta\Delta CT}$ method [38]. The qPCR primers were designed for the 3′ part of the *$vanS_E$* gene that encodes the histidine kinase domain (forward 5′-AGAAGAAGCTTGAGTGGGATTT-3′ and reverse 5′-TCGTTGTATCATTGAGCGAGTAT-3′).

Determining spontaneous mutation rate

The spontaneous mutation rate of isolates was assessed in seven biological replicates by growing them on BHI agar containing rifampicin as previously described [39]. Additionally, the protein sequence of *mutS* and *mutL* genes were compared to the reference strain sequences using Clustal Omega online tool [40] and the promoters of these genes were predicted using SAPPHIRE [41]. ATCC 29212 was chosen as a reference strain for *mutS* and *mutL* genes comparison since its rifampicin MIC was close to that of E1 and E2 (ATCC 29212 rifampicin MIC= 8 mg/L).

Results and Discussion

E1 and E2 are the first Norwegian *vanE*-type VRE isolates recovered from the same patient, and to our knowledge the first *vanE*-type VRE in Europe as well. Since these isolates came from the same patient, we investigated their relatedness as well as their *vanE* gene cluster and mobile genetic context (Tn*6202*). None of the previously reported *vanE*-type VRE isolates was whole genome sequenced [4–6].

***van* genotype and phenotype determination**

Initially in both strains, the *vanE* PCR showed presence of *vanE* type gene cluster which was confirmed by whole genome sequence data.

Phenotypic AST showed that both strains expressed low-level resistance to vancomycin (MIC 16 mg/L) and were susceptible to teicoplanin, ampicillin and linezolid and did not show high levels of gentamicin resistance (Table 1). The glycopeptide phenotype pattern is consistent with earlier reports on *vanE* [5]. Ampicillin, vancomycin and linezolid resistance in *Efs* are not common, but *Efs* often acquire high level gentamicin resistance (HLGR) [42]. To ensure correct concentration in selective media to distinguish recipients from E1 and E2 donors in filter-mating, AST was also performed for rifampicin and fusidic acid showing rifampicin MIC 16 mg/L and fusidic acid MIC 8 mg/L for both strains.

vanE* gene cluster is harboured on Tn*6202

BLAST searches of the chromosomal region containing the *vanE* gene clusters of E1 and E2 showed their similarity to the putative Integrative Conjugative Element (ICE) Tn*6202* of *Efs* strain N00-0410 (FJ872411.1) covering 100% of the sequence. The only AMR genes on Tn*6202* are those of the *vanE* gene cluster. Tn*6202* also carries seven Type IV Secretion system (T4SS) genes [43], an integrase (Int410), and a putative excisionase (Figure 1A).

In the global tree, the VSE*fs* strain 26975_1#118 (GCA_905120835.1) that was isolated from the Netherlands in 2021 was the closest closed genome to E1 and E2 isolates (supplement Figure 1). However, this strain contains an MGE precisely at the integration site of Tn*6202,* therefore the second closest genome GCA_905123845.1 (strain 28157_4#211, the Netherlands, 2021) was chosen as a reference for genomic comparison and extraction of the MGE (Tn*6202*) sequence from E1 and E2 genomes (Figure 1B). In both strains, the MGE containing the *vanE* gene cluster was inserted 7 bp before the 3′ end of the *guaA* gene (glutamine-hydrolysing) at an 11 bp direct repeat (5′-TATTCCCACTC-3′) (Table 2). The only publicly available sequence that contains the Tn*6202* harbouring *vanE* gene cluster (FJ872411.1) shows the same insertion site and perfect direct repeat. The insertion of two different MGEs in this location in the genome of *E. faecalis* suggests the location to be a hotspot for insertion of different MGEs.

The entire *vanE* gene cluster is found in all genomes at NCBI of the human gut commensal *Enterococcus caccae* (n=4) [44]. The nucleotide sequence identity between the *vanE* gene clusters of *E. caccae* and the E1 and E2 is 91% (Supplement figure 2). Since there are no *E.*

caccae genome sequences without *vanE* available, we were not able to extract the MGE harbouring the *vanE* gene cluster in this species, but we found an integrase (tyrosine based site-specific recombinase) next to the *vanE* gene cluster in *E. caccae*.

Table 2. Characteristics of Tn*6202* in E1 and E2 isolates

Strain	Tn*6202* size (bp)	Number of CDSs	Repeats in insertion site (5'-3' strand)	
			L side	R side
E1	43647	39	TATTCCCACTC	TATTCCCACTC
E2	44716	40	TATTCCCACTC	TATTCCCACTC

$vanS_E$ gene in both strains is truncated but vancomycin resistance is still low level inducible in E1

The regulatory $vanS_E$ gene in both E1 and E2 isolates was truncated due to a premature stop codon caused by a single nucleotide deletion in E1 (362delA) and insertion of IS*6770* in E2 at position 383 which in both resulted in two CDSs (Figure 1A), each containing different functional domains of the $vanS_E$ gene. The first CDS contains the two transmembrane-associated sensor domains (123 and 117 amino acids (AAs) in E1 and E2, respectively), and the second contains the histidine kinase domain (209 and 229 AA in E1 and E2, respectively). The integrated IS*6770* in the $vanS_E$ gene of E2 overlaps the last 13 nucleotides in the first CDS from the 3´end. The functional *vanS* gene in N00-140 strain (FJ872411.1) results in a 341 AA protein (Figure 2). In *vanE*-type isolates with truncated $vanS_E$ gene, inducible vancomycin resistance couldbe achieved by cross-talk of another two-component signal transduction system or by $vanR_E$ acting with a heterologous histidine kinase [13,45,46].

IS*6770* elements can be integrated at different locations of the *Efs* genome and it is reported as a common IS element in the enterococcal genomes, mostly in more than one copy [47]. Local BLAST on complete *Efs* genomes retrieved from NCBI showed that 41% (189 of 458) of the genomes contain one to 24 copies of IS*6770.* The insertion of IS*30*-like elements (IS*6770*) in the $vanS_E$ gene cluster has not been reported so far, but the insertion of IS elements has been reported before in the *vanS* gene of the *vanA* gene cluster [48].

Although the *vanE* gene cluster produces inducible vancomycin resistance, *vanE*-VRE*fs* strains may show truncated $vanS_E$ gene and constitutively expressed resistance to vancomycin [5,6]. The inducibility of vancomycin resistance in E1 and E2 was tested by assessment of growth

curves with and without vancomycin and expression of the *vanS*$_E$ gene by qPCR. qPCR plots for pre-growth in BHI without and with subMIC (10 mg/L) of vancomycin show a 2-fold increase in expression of the histidine kinase part of *vanS*$_E$ in E1, while expression in E2 isolate was reduced 3-fold. These changes were not statistically significant (p= 0.62 for E1 and p= 0.76 for E2) (Figure 3). In the *vanE*-type isolates with inducible resistance, the generation time Tg is shorter when pre-cultured in sub-MIC [9]. E1 pre-cultured in sub-MIC of vancomycin showed a shorter Tg (63,6 minutes) compared to E1 pre-cultured without vancomycin (66,6 minutes) while different pre-culturing did not affect the Tg of E2 (supplement Figure 3). However, the change in Tg of E1 was not statistically significant (P= 0.82). Since *vanR* and *vanS* genes in the *vanE* gene cluster are controlled by one promoter [45] and the CDS that contains the histidine kinase domain in E1 was in the same frame as *vanR* it is possible that this domain was still controlled by the *vanR*$_E$ promoter. The slight decrease in Tg and increase in expression of the *vanS* gene after exposure to sub-MIC of vancomycin show low-level inducibility of vancomycin resistance in E1. In E2, the insertion of IS*6770* (between the two CDSs of the *vanS* gene) with a transposase encoded in the opposite direction clearly disturbed the expression of the histidine kinase part of *vanS*$_E$ gene from the *vanR*$_E$ promoter. Still, low level constitutive expression of resistance to vancomycin was seen in E2 as has also been observed for other strains with truncated *vanS*$_E$. Insertion of IS elements carrying promoters or partial promoters can contribute to expression of downstream genes, [49] and thus may have contributed to the low level constitutive expression of the histidine kinase domain of *vanS*$_E$ in E2.

Activity and transferability of the Tn*6202*

Tn*6202* contains seven T4SS genes, as well as integrase and excisionase. So, the genes needed for excision and integration, as well as some putative genes for transfer, are available in the genome of the donor bacteria (E1 and E2 isolates). In the process of MGE transfer, after excision from the genome of the donor, many MGEs including Tn*6202* forms a circular double-stranded intermediate [50]. PCR for circularization of Tn*6202* in E1 and E2 was negative which means that the active form of Tn*6202* was not detected. The genome assembly of E1 isolate contains a contig with 100% identity to both flanking sequences of Tn*6202*. Mapping both Illumina and PacBio reads confirm the existence of this contig in the genome. This fragment indicates a heterogeneity in the E1 population. A possible explanation for this could be that the inserted MGE is not stable. Transfer of Tn*6202* from E1 and E2 to JH2-2 and OG1-RF was not detected (n>1), which is in line with previous report on Tn*6202* not being transferrable among

E. faecalis [50]. Thus, the putative ICE Tn*6202* either lack some necessary genes, a host factor is necessary for transfer, or the transfer frequency was below the detection limit.

The relatedness between E1 and E2

The isolates E1 and E2 belong to the same sequence type (ST34) but have different CTs. According to cgMLST results, there are 32 allelic differences between E1 and E2 that result in two cluster types: CT3081 and CT2880.

Variant calling (using E1 assembly as a reference for E2 reads) showed 60 variants between E1 and E2, including 49 SNP, three insertions and eight deletions. A total of 51 variants occurred in coding sequences, including a missense mutation SNP in the *vanR_E* gene. CDS variants include 30 missense and nine synonymous variants. A total of 12 variants (20%) were deemed to have a high impact on gene function. The high impact of these variants is due to gaining a stop codon (n=4), loss of start codon (n=1), and frame shift (n=7). The frame shift variants resulted in premature stop codon (n=5) or shorter product from the 5′ end (n=2) in the CDS (Supplement file 1).

Compared to the reference genome (GCA_905120835.1), E1 and E2 have 756 (454 of them on the chromosome) and 993 (699 of them on the chromosome) variants, respectively. As a result of SNPeff analysis 177 of the variants of E1 and 225 of E2 are missense SNPs while 184 and 287 synonymous variants were found in E1 and E2, respectively. SNPeff classified 11/454 (2%) and 24/699 (3%) variants in E1 and E2 as having high impact on the gene function.

Exposing the isolates to high concentrations of antibiotics such as rifampicin is a useful method to calculate the mutation rate in pathogenic bacteria [51]. For E1 and E2 the spontaneous mutation rate was 1×10^{-7} and 8×10^{-10}, respectively, while it was 3×10^{-10} for the reference strain used in the experiment (ATCC29212). Since E1 has 125-folds higher mutation rate, 60 variants between E1 and E2 in three years are not a high number. It has been shown that mutation in DNA mismatch repair genes (*mutS* and *mutL*) can cause elevated spontaneous mutation frequencies in *Efs* [39]. However, the difference in mutation rate cannot be explained by this since the DNA sequence of *mutS* and *mutL* genes and their promoters were identical in E1 and E2 isolates and showed 99.5% identity to the *mutS* and *mutL* genes of ATCC29212. Additionally, no differences in the promoter region of *mutS* and *mutL* genes between E1, E2, and the reference genomes (GCA_905120835.1) were observed.

Plasmid finder identified only one replicon in E1 (rep9c), while for E2 it detected two additional plasmid replicons (rep9a and repUS43). E1 and E2 isolates have different plasmid profiles. Local BLAST search showed that the common plasmid between the E1 and E2 (rep9c)

has 98% identity and 39% coverage. The only AMR gene on these plasmids is a *tet* gene on rep9c of E2. The two isolates also have other differences in MGEs. E1 has a 31 kb MGE (MEG-E1-A) integrated at the overlapping sequence of two hypothetical proteins genes, while E2 contains a 45 kb MGE (MGE-E2-B) integrated at the 5' end of putative tRNA sulphur transferase genes. Except for these MGEs, the genome organization of E1 and E2 is very similar (Figure 1B).

Since the isolates were recovered from the same patient, the two isolates are epidemiologically related. The similar genome organization, the same ST and only 32 allelic differences in the cgMLST as well as only 60 variations in their entire genomic sequences despite a high mutation rate in E1 suggest relatedness between them. In addition, 30 months' time between their isolations, may have contributed to the variations between these two related isolates.

Conclusion

The first *vanE*-VRE*fs* isolates (E1 and E2) reported in Norway were recovered from the same patient thirty months apart. Their *vanE* gene clusters, harboured by the previously described MGE Tn*6202,* only showed difference in the $vanS_E$ gene. In both isolates the $vanS_E$ gene was truncated resulting in two CDSs with different functional domains. In E1 the truncation resulted in the histidine kinase domain still being in frame with $vanR_E$ likely explaining why in this isolate vancomycin resistance was low-level inducible, while the IS element insertion in the $vanS_E$ gene of E2 likely explains its low-level constitutive expression of resistance to vancomycin. A total of 60 variants exists between the genomes of E1 and E2. Moreover, the accessory genomes of the isolates show different integrated MGEs and plasmid profiles. Similar chromosomal gene content and synteny of two isolates found in the same patient suggest the isolates are related. The differences in the mutation rate and long time (2,5 years) between isolation of E1 and E2 allowed independent genetic events and SNP differences between them. For the future, we suggest genomic comparison of E1 and E2 to other whole genome sequenced *vanE*-type VRE*fs* and further experiments to determine the reason for the high mutation rate in E1.

Acknowledgments

We would like to thank Kjersti Julin, Bhupender Singh, Ekaterina Mishchenko, Bishnu Joshi, and Theresa Wagner at UiT for their technical support, Siv-Heidi Barkhald and Bjørg Haldorsen at the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance for excellent technical assistance, and the Genomics Support Center Tromsø™ and Norwegian Sequencing Centre (NSC) in the University of Oslo for genome sequencing.

Funding

This work was supported by a grant from the Northern Norway Regional Health Authority Medical Research Programme (HNF1362-17). The funders had no role in the study design, data collection, and analysis, decision to publish or preparation of the manuscript.

Author contributions

MR, JJ, and KH designed the experiments, analysed, curated, and interpreted the data. JJ, JVB, AS, and KH contributed to conceptualization, supervision, and funding acquisition. JVB and IHL contributed to the acquisition of data. MR, AJ, and IHL contributed to drafting and JJ, JVB, AS, and KH to review and editing of the manuscript. All authors contributed to the validation of the work and approved the final version of the manuscript.

Figures legend

Figure 1. Comparison between Tn*6202* and genomes of *vanE*-VRE*fs* and their references using Easyfig tool. The CDSs and their directions are shown with arrows. Figure 1A illustrates the comparison between Tn*6202* and its flanking sequences of E1 and E2 to their reference (N00-410). The red bars here represent 100% identity. The Tn*6202* insertion site and its sequence are marked on the reference N00-410. Figure 1B shows similarities between the genomes of E1, E2, and the two closest closed *Efs* isolates retrieved from NCBI. The green rectangle marks the position of Tn*6202* in the genome. The red and blue gradient bars represent normal and inverted matches, respectively.

Figure 2. Amino acid sequence alignment of VanS in *vanE*-VRE*fs* isolates using Clustal omega online tool compared to the reference sequence (N00-410). Cov and pid represent the coverage and identity percentages respectively. Since in E1 and E2, the two domains of

vanS were annotated as a functional CDS in Prokka, we added each of the domains (transmembrane and histidine kinase) as separate sequences in the alignment.

Figure 3. The level of expression of *vanS*$_E$ histidine kinase part determined by RT-qPCR. The average expression levels of histidine kinase part of the *vanS*$_E$ gene ($2^{-\Delta\Delta Ct}$) of E1 and E2. Data are expressed as the mean ± standard deviation. Two-tailed t-test showed no significant changes (p= 0.62 for E1 and p= 0.76 for E2) in expression of the histidine kinase domain in E1 or E2 pre-grown in BHI broth without (control) and with 10 mg/L vancomycin (treated).

Supplement Figure 1. Global phylogenetic tree based on the core genome SNP alignment of *E. faecalis*. This tree includes the Norwegian *vanE*-VRE*fs* and the available closed genomes of *Efs* in the NCBI database as of 01.08.02022. E1 and E2 are coloured in red, and the genomes used as reference are in green.

Supplement Figure 2. Comparison between *vanE* gene cluster of *E. caccae* (strain MGYG-HGUT-02468) and *Efs* N00-410, E1 and E2. The red gradient bars indicate normal matches and genes, and their directions are shown with arrows.

Supplement Figure 3. Growth of the E1 and E2 isolates with and without vancomycin present. The cultures with vancomycin were pre-grown in BHI containing vancomycin sub-MIC and then cultured in BHI media with the same sub-MIC of vancomycin. In S figure 3A sub-MIC was 4 mg/L while in S figure 3B sub-MIC was 10 mg/L.

Supplement file 1. Annotation and predicted effect of variants between E1 and E2 using E1 assembly as reference for E2 reads (Sheet 1), and GCA_905120835.1 as reference for E1 reads (sheet 2) and E2 reads (sheet 3).

References:

1. García-Solache M, Rice LB. The Enterococcus: a model of adaptability to its environment. Clin Microbiol Rev. 2019;32:1–28.

2. Xavier BB, Coppens J, De Koster S, Rajakani SG, Van Goethem S, Mzougui S, et al. Novel vancomycin resistance gene cluster in Enterococcus faecium ST1486, Belgium, June 2021. Eurosurveillance. 2021;26:2100767.

3. Cattoir V, Leclercq R. Twenty-five years of shared life with vancomycin-resistant enterococci: is it time to divorce? J Antimicrob Chemother. 2013;68:731–42.

4. Boyd DA, Cabral T, Caeseele P Van, Wylie J, Mulvey MR. Molecular Characterization of the vanE Gene Cluster in Vancomycin-Resistant Enterococcus faecalis N00-410 Isolated in Canada. Antimicrob Agents Chemother. 2002;46:1977–9.

5. Fines M, Perichon B, Reynolds P, Sahm DF, Courvalin P. VanE , a new type of acquired Glycopeptide resistance in *Enterococcus faecalis* BM4405. Antimicrob Agents Chemother. 1999;43:2161–4.

6. Abadía-patin L, Christiansen K, Bell J, Courvalin P, Pe B. VanE-Type Vancomycin-Resistant Enterococcus faecalis Clinical Isolates from Australia. Antimicrob Agents Chemother. 2004;48:4882–5.

7. Courvalin P. Vancomycin resistance in Gram-positive cocci. Clin Infect Dis. 2006;42:25–34.

8. Patiño LA, Marc C, Courvalin. Patrice, Bruno. Périchon. Silencing of Glycopeptide Resistance in Enterococcus faecalis BM4405 by Novobiocin. Antimicrob Agents Chemother. 2005;49:1419–25.

9. Shlaes DM, Bouvet A, Devine C, Shlaes JH, Al-Obeid S, Williamson R. Inducible, transferable resistance to vancomycin in Enterococcus faecalis A256. Antimicrob Agents Chemother. 1989;33:198–203.

10. Hong H-J, Hutchings MI, Buttner MJ. Vancomycin Resistance VanS/VanR Two-Component Systems. In: Utsumi R, editor. Bact Signal Transduct Networks Drug Targets. New York, NY: Springer New York; 2008. p. 200–13.

11. Dhanda G, Sarkar P, Sandip S, Jayanta. Haldar. Battle against Vancomycin-Resistant Bacteria: Recent Developments in Chemical Strategies. J Med Chem. 2019;62:3184–205.

12. Gorrie C, Higgs C, Carter G, Stinear TP, Howden B. Genomics of vancomycin-resistant Enterococcus faecium. Microb Genomics 2019;5. 2019;5:e000283.

13. Lorena A-P, Keryn C, Jan B, Patrice C, Bruno P. VanE-Type Vancomycin-Resistant Enterococcus faecalis Clinical Isolates from Australia. Antimicrob Agents Chemother. 2004;48:4882–5.

14. Punina N V, Makridakis NM, Remnev MA, Topunov AF. Whole-genome sequencing targets drug-resistant bacterial infections. Hum Genomics. Human Genomics; 2015;9:1–20.

15. Lassalle F, Didelot X. Bacterial Microevolution and the Pangenome. In: Tettelin H, Medini D, editors. Pangenome Divers Dyn Evol Genomes. Cham: Springer International Publishing; 2020. p. 129–49.

16. Yang H, Hoffmann M, Allard MW, Brown EW, Chen Y. Microevolution and Gain or Loss of Mobile Genetic Elements of Outbreak-Related Listeria monocytogenes in Food Processing Environments Identified by Whole Genome Sequencing Analysis. Front Microbiol. 2020;11:1–29.

17. Du X, Hua X, Qu T, Jiang Y, Zhou Z, Yu Y. Molecular characterization of Rifr mutations in Enterococcus faecalis and Enterococcus faecium. J Chemother. 2014;26:217–21.

18. Depardieu F, Perichon B, Courvalin P. Detection of the Van Alphabet and identification of *Enterococci* and *Staphylococci* at the species level by multiplex PCR. J Clin Microbiol. 2004;42:5857–60.

19. AL Rubaye MTS, Janice J, Bjørnholt JV, Jakovljev A, Hultström ME, Sundsfjord A, et al. Novel genomic islands and a new vanDsubtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway. PLoS One. 2021;16:1–15.

20. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler : Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 2017;13:e1005595.

21. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.

22. Seemann T. Prokka: Rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.

23. Seemann T. Abricate [Internet]. Github. Available from: https://github.com/tseemann/abricate

24. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

25. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. Babraham Bioinforma. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/

26. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: Genomics. 2013;00:1–3.

27. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10:giab008.

28. Seemann T. Snippy, Rapid haploid variant calling and core genome alignment [Internet]. 2015. Available from: https://github.com/tseemann/snippy

29. Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, et al. A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff. Fly (Austin). 2012;6:80–92.

30. Neumann B, Prior K, Bender JK, Harmsen D, Klare I, Fuchs S, et al. A Core Genome Multilocus Sequence Typing Scheme for Enterococcus faecalis. J Clin Microbiol. 2019;57:e01686-18.

31. Carver TJ, Rutherford KM, Berriman M, Rajandream M, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21:3422–3.

32. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27:1009–10.

33. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014;15:524.

34. Sivertsen A, Janice J, Pedersen T, Wagner TM, Hegstad J, Hegstad K. The *Enterococcus* cassette chromosome, a genomic variation enabler in *Enterococci*. mSphere. 2018;3:1–13.

35. Sivertsen A, Billström H, Melefors Ö, Liljequist BO, Wisell KT, Ullberg M, et al. A Multicentre Hospital Outbreak in Sweden Caused by Introduction of a vanB2 Transposon into a Stably Maintained pRUM-Plasmid in an Enterococcus faecium ST192 Clone. PLoS One. 2014;9:e103274.

36. Dunny GM, Brown BL CD. Induced cell aggregation and mating in Streptococcus faecalis: evidence for a bacterial sex pheromone. Proc Natl Acad Sci USA. 1978;75:3479–83.

37. Jacob AE, Hobbs SJ. Conjugal transfer of plasmid-borne multiple antibiotic resistance in Streptococcus faecalis var. zymogenes. J Bacteriol. 1974;117:360–72.

38. Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. Methods. 2001;25:402–8.

39. Chilambi GS, Nordstrom HR, Evans DR, Kowalski RP, Dhaliwal DK, Jhanji V, et al. Genomic and phenotypic diversity of Enterococcus faecalis isolated from endophthalmitis. PLoS One. 2021;16:e0250084.

40. Sievers F, Wilm A, Dineen D, Gibson TJ, Karplus K, Li W, et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol Syst Biol. 2011;7:539.

41. Coppens L, Lavigne R. SAPPHIRE: a neural network based classifier for σ70 promoter prediction in Pseudomonas. BMC Bioinformatics. 2020;21:415.

42. Sparo M, Delpech G, García Allende N. Impact on Public Health of the Spread of High-Level Resistance to Gentamicin and Vancomycin in Enterococci. Front Microbiol. 2018;9:3073.

43. ICEberg. Tn6202 [Internet]. Available from: https://bioinfo-mml.sjtu.edu.cn/ICEberg2/feature_page.php?ice_id=448

44. Carvalho M da GS, Shewmaker PL, Steigerwalt AG, Morey RE, Sampson AJ, Joyce K. Enterococcus caccae sp. nov., isolated from human stools. Int J Syst Evol Microbiol. 2006;56:1505–8.

45. Abadía Patiño L, Courvalin P, Perichon B. vanE Gene Cluster of Vancomycin-Resistant Enterococcus faecalis BM4405. J Bacteriol. 2002;184:6457–64.

46. Fisher SL, Jiang W, Wanner BL, Walsh CT. Cross-talk between the Histidine Protein Kinase VanS and the Response Regulator PhoB: Characterization and identification of a VanS domain that inhibits activation of PhoB. J Biol Chem. 1995;270:23143–9.

47. Thorisdottir AS, Carias LL, Marshall SH, Green M, Zervos MJ, Giorgio C, et al. IS6770, An Enterococcal Insertion-Like Sequence Useful For Determining The Clonal Relationship Of Clinical Enterococcal Isolates. J Infect Dis. 1994;170:1539–48.

48. Monteiro da Silva BN, Faria AR, Souza S da SR, Colodette SS, Morais JM, Teixeira LM, et al. Expression of VanA-type vancomycin resistance in a clinical isolate of Enterococcus faecium showing insertion of IS19 in the vanS gene. Int J Antimicrob Agents. 2020;55:105897.

49. Vandecraen J, Chandler M, Aertsen A, Van Houdt R. The impact of insertion sequences on bacterial genome plasticity and adaptability. Crit Rev Microbiol. 2017;43:709–30.

50. Boyd DA, Mulvey MR. The VanE operon in Enterococcus faecalis N00-410 is found on a putative integrative and conjugative element, Tn6202. J Antimicrob Chemother. 2013;68:294–9.

51. O'Neill AJ, Chopra I. Insertional inactivation of mutS in Staphylococcus aureus reveals potential for elevated mutation frequencies, although the prevalence of mutators in clinical isolates is low. J Antimicrob Chemother. 2002;50:161–9.

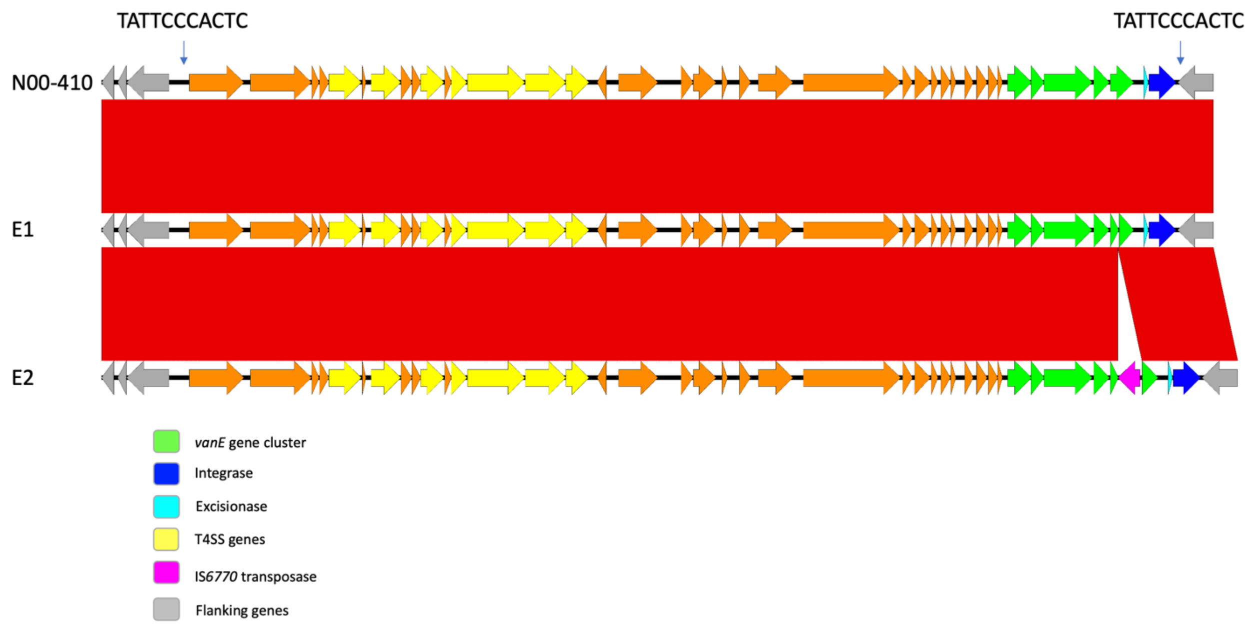


Figure 1 (A).

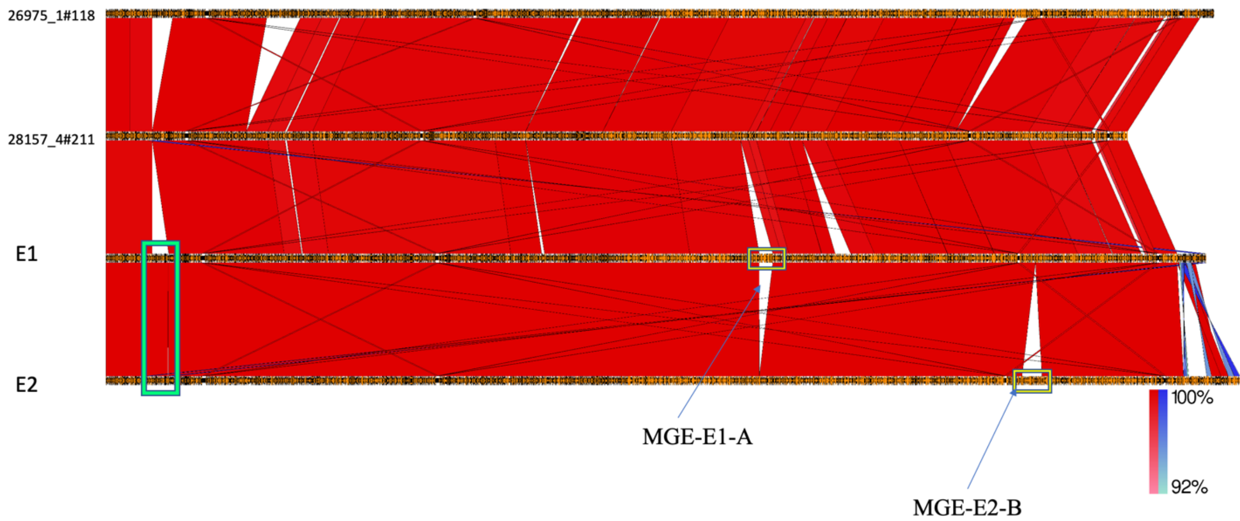


Figure 1 (B).

Figure 1 (A and B). Comparison between Tn*6202* and genomes of *vanE*-VRE*fs* and their references using Easyfig tool.

Figure 2. Amino acid sequence alignment of VanS in *vanE*-VRE*fs* isolates using Clustal omega online tool compared to the reference sequence (N00-410).

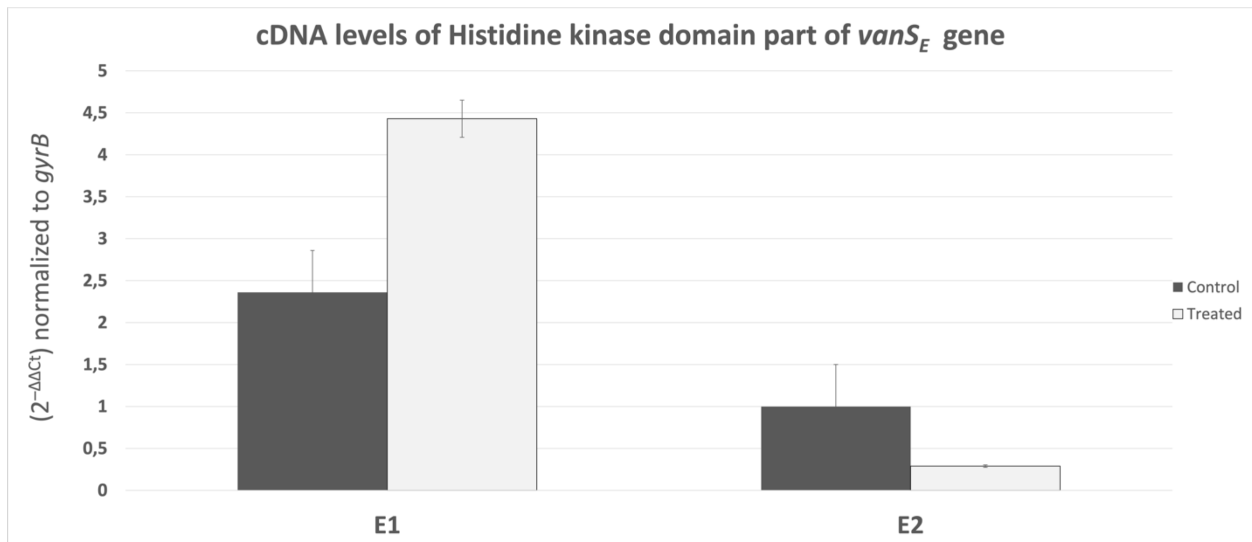
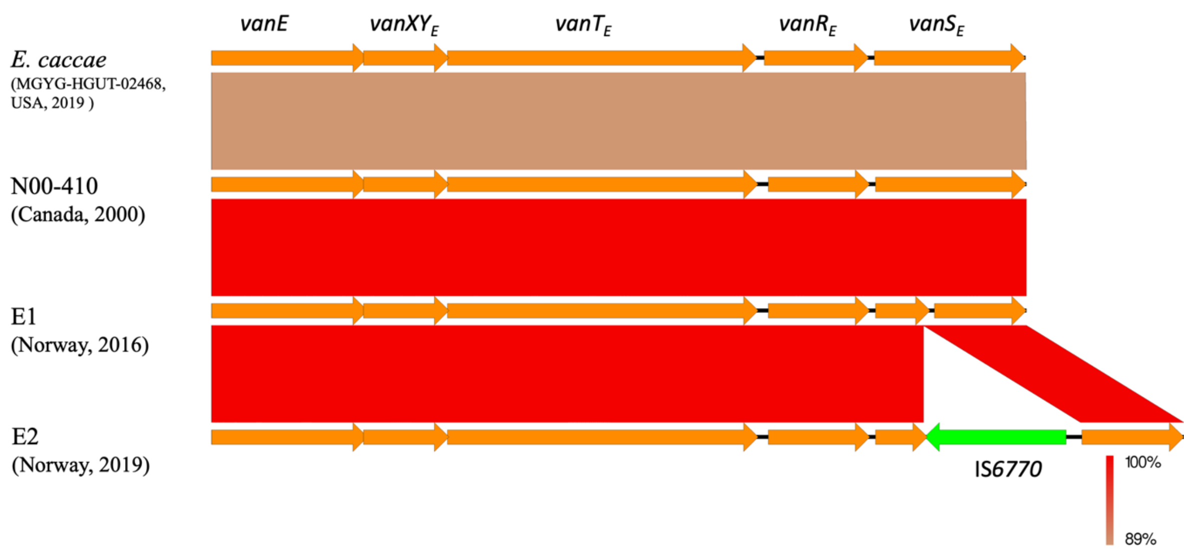
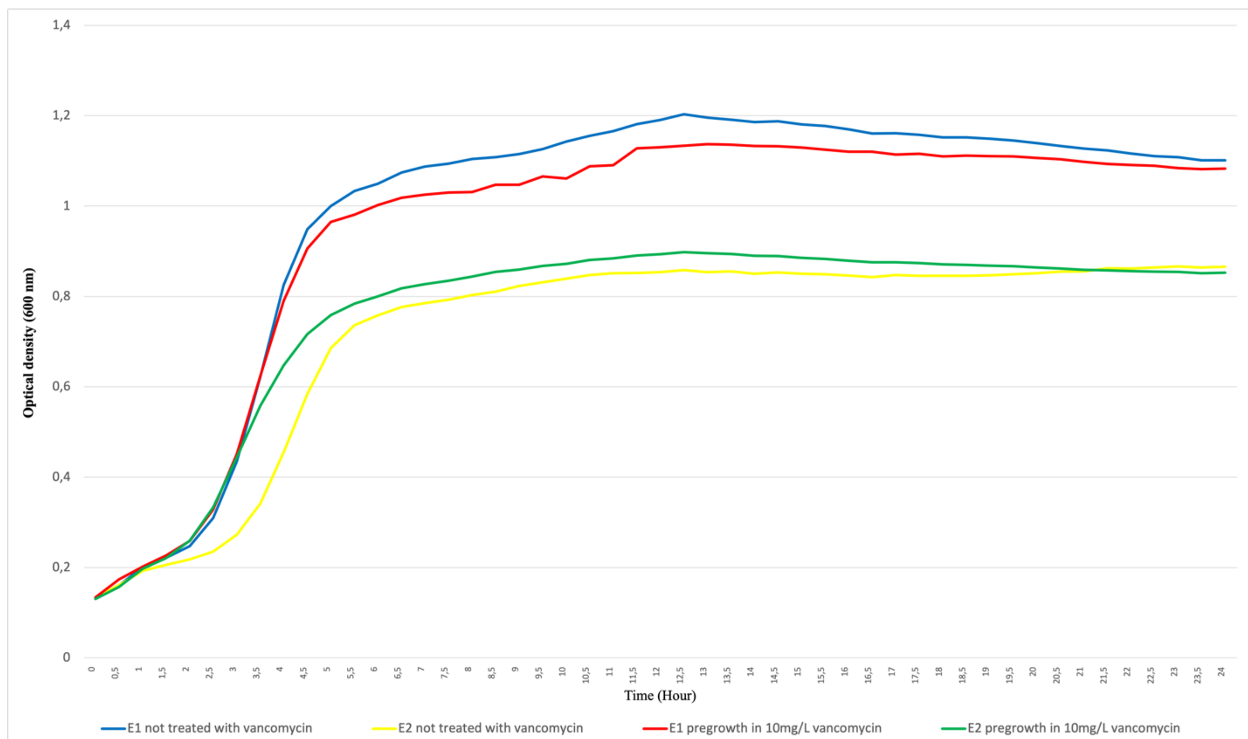


Figure 3. The level of expression of $vanS_E$ histidine kinase part determined by RT-qPCR.

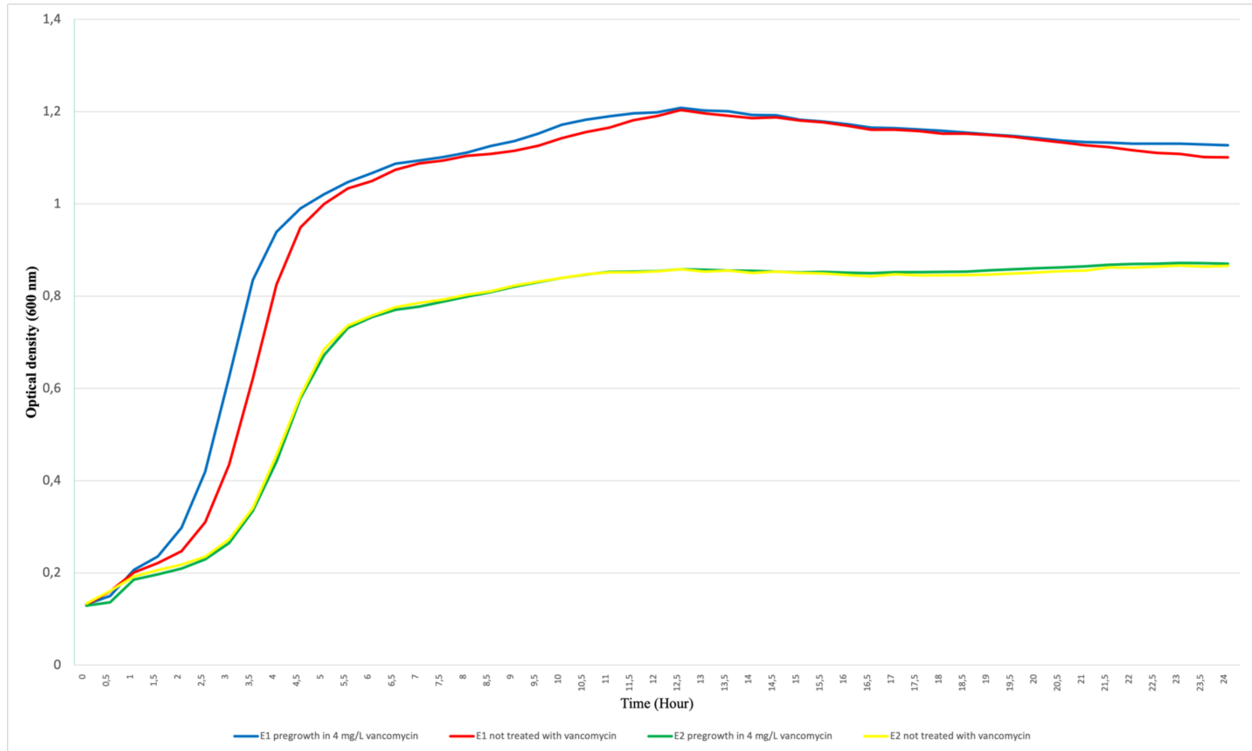
Supplement Figure 1. Global phylogenetic tree based on the core genome SNP alignment of *E. faecalis.*



Supplement Figure 2. Comparison between *vanE* gene cluster of *E. caccae* (strain MGYG-HGUT-02468) and *Efs* N00-410, E1 and E2.



Supplement Figure 3A



Supplement Figure 3B

Supplement Figure 3 (A and B). Growth of the E1 and E2 isolates with and without vancomycin present.

Supplement file 1. Annotation and predicted effect of variants between E1 and E2 using E1 assembly as reference for E2 reads

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
1	24355	A	T	snp	synonymous_variant	LOW	sorA_1
1	163190	C	G	snp	missense_variant	MODERATE	vanRE
1	233728	A	G	snp	intergenic_region	MODIFIER	niaR-ALLEEMCA_00232
1	473521	AG	A	del	frameshift_variant	HIGH	opuCA_1
1	477734	G	A	snp	missense_variant	MODERATE	ALLEEMCA_00477
1	514407	C	T	snp	missense_variant	MODERATE	ALLEEMCA_00510
1	520814	T	C	snp	synonymous_variant	LOW	tig
1	551347	T	A	snp	stop_gained	HIGH	ALLEEMCA_00542
1	554778	G	A	snp	missense_variant	MODERATE	dap
1	673581	C	T	snp	missense_variant	MODERATE	kimA
1	711138	C	A	snp	missense_variant	MODERATE	artM
1	832083	C	T	snp	synonymous_variant	LOW	gmuD
1	871502	G	A	snp	intergenic_region	MODIFIER	sasA_1-ALLEEMCA_00853
1	949500	T	C	snp	missense_variant	MODERATE	murI
1	990469	A	T	snp	stop_gained	HIGH	tagB
1	997239	G	A	snp	missense_variant	MODERATE	luxS
1	999121	C	A	snp	missense_variant	MODERATE	dapA
1	1004199	T	A	snp	synonymous_variant	LOW	yutF
1	1016816	A	T	snp	intergenic_region	MODIFIER	ALLEEMCA_00985-ALLEEMCA_00986
1	1018141	T	A	snp	intergenic_region	MODIFIER	ALLEEMCA_00988-ALLEEMCA_00989
1	1026026	T	A	snp	missense_variant	MODERATE	tagU_2
1	1056380	A	G	snp	missense_variant	MODERATE	ALLEEMCA_01022
1	1081627	T	A	snp	missense_variant	MODERATE	rlmI
1	1114469	C	T	snp	stop_gained	HIGH	cynR_1
1	1194971	C	T	snp	synonymous_variant	LOW	ALLEEMCA_01151
1	1221250	C	T	snp	missense_variant	MODERATE	moeA
1	1224990	CT	C	del	frameshift_variant	HIGH	ALLEEMCA_01179
1	1281514	AT	A	del	frameshift_variant	HIGH	mapZ
1	1302442	C	T	snp	missense_variant	MODERATE	ALLEEMCA_01249
1	1340654	A	AT	ins	intergenic_region	MODIFIER	perR-nox
1	1365561	T	C	snp	missense_variant	MODERATE	clsA_2
1	1406268	C	T	snp	missense_variant	MODERATE	ALLEEMCA_01345
1	1453032	T	A	snp	synonymous_variant	LOW	pstS1_1
1	1516936	CT	C	del	frameshift_variant	HIGH	ALLEEMCA_01457
1	1554756	A	T	snp	synonymous_variant	LOW	agaS
1	1593961	A	T	snp	missense_variant	MODERATE	larA
1	1627885	A	T	snp	missense_variant	MODERATE	gatC_4
1	1629218	C	T	snp	start_lost	HIGH	gatC_4
1	1629856	G	GT	ins	intergenic_region	MODIFIER	rpiB-srlR_2
1	1632334	G	A	snp	synonymous_variant	LOW	ALLEEMCA_01567
1	1654879	A	T	snp	missense_variant	MODERATE	ALLEEMCA_01589
1	1683251	G	A	snp	missense_variant	MODERATE	ALLEEMCA_01615
1	1708121	A	C	snp	missense_variant	MODERATE	pdxK_2
1	1796384	G	A	snp	stop_gained	HIGH	ybiT_2
1	1797879	C	T	snp	intergenic_region	MODIFIER	ybiT_2-scmP_2
1	1818212	G	A	snp	missense_variant	MODERATE	ALLEEMCA_01754

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
1	1831953	AT	A	del	frameshift_variant	HIGH	ALLEEMCA_01764
1	1855406	TA	T	del	frameshift_variant	HIGH	ALLEEMCA_01779
1	1881354	A	G	snp	synonymous_variant	LOW	rhaR_4
1	2006548	GT	G	del	intergenic_region	MODIFIER	rasP-gdh_2
1	2008188	A	C	snp	missense_variant	MODERATE	ALLEEMCA_01906
1	2026700	T	C	snp	missense_variant	MODERATE	glyS
1	2259602	GGTTTTTAAACAT	G	del	intergenic_region	MODIFIER	spxA_2-trpS
1	2309041	A	AT	ins	frameshift_variant	HIGH	ALLEEMCA_02185
1	2383261	G	T	snp	missense_variant	MODERATE	ALLEEMCA_02264
1	2383499	A	G	snp	missense_variant	MODERATE	ALLEEMCA_02265
1	2404792	C	A	snp	missense_variant	MODERATE	ALLEEMCA_02290
1	2427962	T	A	snp	missense_variant	MODERATE	liaS
1	2634607	T	C	snp	missense_variant	MODERATE	sorA_4
1	2634907	C	A	snp	missense_variant	MODERATE	sorA_4

Supplement file 1. Annotation and predicted effect of variants in E1 chromosome using GCA_905120835.1 as reference

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	2844	C	CT	snp	intergenic_region	MODIFIER	dnaN-FCKDLICC_00003
LR961994.1	21946	A	G	snp	synonymous_variant	LOW	dgaR_1
LR961994.1	44147	G	A	snp	intergenic_region	MODIFIER	radA-yacL
LR961994.1	52101	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_00045-veg
LR961994.1	57880	A	G	del	missense_variant	MODERATE	glmU
LR961994.1	65059	T	G	snp	intergenic_region	MODIFIER	dppE_1-FCKDLICC_00055
LR961994.1	75803	GTCAAA	TTTGAC	snp	intergenic_region	MODIFIER	FCKDLICC_00065-larR_1
LR961994.1	89641	C	T	snp	synonymous_variant	LOW	FCKDLICC_00081
LR961994.1	91010	A	C	snp	missense_variant	MODERATE	FCKDLICC_00081
LR961994.1	98604	A	G	complex	synonymous_variant	LOW	serS_1
LR961994.1	121935	A	T	snp	intergenic_region	MODIFIER	FCKDLICC_00110-FCKDLICC_00111
LR961994.1	186618	G	A	snp	missense_variant	MODERATE	tmpC_2
LR961994.1	202573	G	A	snp	missense_variant	MODERATE	yxdL_2
LR961994.1	231227	C	T	complex	missense_variant	MODERATE	scmP_1
LR961994.1	232892	A	G	snp	synonymous_variant	LOW	ecfA2
LR961994.1	240320	T	C	complex	missense_variant	MODERATE	mltF_1
LR961994.1	242538	T	G	snp	missense_variant	MODERATE	dapH_1
LR961994.1	248554	C	T	snp	intergenic_region	MODIFIER	ldh_1-pth
LR961994.1	251483	C	A	snp	synonymous_variant	LOW	mfd
LR961994.1	255248	T	C	snp	missense_variant	MODERATE	FCKDLICC_00260
LR961994.1	256921	A	G	snp	missense_variant	MODERATE	tilS
LR961994.1	257901	C	T	snp	synonymous_variant	LOW	hpt
LR961994.1	278518	C	T	snp	synonymous_variant	LOW	gpmA_2
LR961994.1	296169	G	GA	snp	intergenic_region	MODIFIER	gmuC-murR_1
LR961994.1	313462	G	A	complex	missense_variant	MODERATE	FCKDLICC_00339
LR961994.1	324836	T	C	snp	synonymous_variant	LOW	FCKDLICC_00348
LR961994.1	341535	T	C	snp	missense_variant	MODERATE	atzC_1
LR961994.1	388790	T	G	snp	intergenic_region	MODIFIER	FCKDLICC_00405-dtpT
LR961994.1	400489	A	C	snp	intergenic_region	MODIFIER	nagB_1-fcbA2
LR961994.1	404808	A	G	snp	missense_variant	MODERATE	manZ_2
LR961994.1	420028	A	C	snp	missense_variant	MODERATE	nrdE2
LR961994.1	424441	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_00439-xerC_3
LR961994.1	426825	GG	AA	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426835	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426849	G	C	ins	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426867	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426883	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426893	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426901	CTT	TCC	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426908	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426921	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426948	C	A	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426957	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426982	A	C	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	426991	C	T	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427009	T	A	snp	missense_variant	MODERATE	FCKDLICC_00444

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	427014	G	T	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427021	C	T	del	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427038	TA	CG	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427045	CTCCG	ACCTA	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427054	A	G	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427082	ACT	GCC	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427091	CT	GA	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427133	T	C	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427148	TGA	AAG	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427155	C	T	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427162	C	T	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427172	GGA	AGG	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427183	T	C	del	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427189	TTTGTTG	AGATTTA	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427208	A	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427213	TCCG	CCCA	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427221	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427239	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427266	A	G	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427272	G	A	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427295	TATT	CATC	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427329	GAC	AAT	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427347	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427353	A	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427362	G	T	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427367	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427443	AGAGTACT	GGAATACC	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427459	G	T	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427469	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427487	A	C	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427499	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427519	GTAAT	TTAGC	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427545	G	A	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427576	TTC	ATG	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427584	A	G	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427589	T	G	mnp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427616	T	TC	complex	intergenic_region	MODIFIER	FCKDLICC_00445-FCKDLICC_00446
LR961994.1	431151	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	431205	TCT	CCA	snp	missense_variant	MODERATE	asa1
LR961994.1	431213	AAATGGAG	TAAAGGTA	snp	missense_variant	MODERATE	asa1
LR961994.1	431228	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431243	C	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431275	TAGG	CAGA	complex	missense_variant	MODERATE	asa1
LR961994.1	431289	TCAG	T	ins	conservative_inframe_deletion	MODERATE	asa1
LR961994.1	431315	A	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431325	G	A	snp	missense_variant	MODERATE	asa1
LR961994.1	431367	A	G	complex	missense_variant	MODERATE	asa1

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	431378	TGCT	AGCG	snp	missense_variant	MODERATE	asa1
LR961994.1	431394	ACT	TCC	snp	missense_variant	MODERATE	asa1
LR961994.1	431429	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	431447	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431459	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431468	TGTC	CGTT	snp	synonymous_variant	LOW	asa1
LR961994.1	431485	T	C	snp	missense_variant	MODERATE	asa1
LR961994.1	431492	T	C	del	synonymous_variant	LOW	asa1
LR961994.1	431498	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431523	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	431564	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	431576	G	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431585	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431648	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431653	T	C	snp	missense_variant	MODERATE	asa1
LR961994.1	431666	CTAC	TTAT	snp	synonymous_variant	LOW	asa1
LR961994.1	431705	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	431777	G	T	snp	synonymous_variant	LOW	asa1
LR961994.1	431873	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432306	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	432318	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	432413	T	A	snp	synonymous_variant	LOW	asa1
LR961994.1	432692	T	C	ins	synonymous_variant	LOW	asa1
LR961994.1	432710	C	T	del	synonymous_variant	LOW	asa1
LR961994.1	432749	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	432773	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	432902	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432935	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432977	TCTTC	CCTTA	snp	missense_variant	MODERATE	asa1
LR961994.1	432989	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433001	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433031	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433124	AG	GC	snp	missense_variant	MODERATE	asa1
LR961994.1	433133	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433139	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433154	T	C	del	synonymous_variant	LOW	asa1
LR961994.1	433160	GATTAAC	TATCAAT	snp	synonymous_variant	LOW	asa1
LR961994.1	433176	G	A	complex	missense_variant	MODERATE	asa1
LR961994.1	433181	TAAT	CAAC	snp	synonymous_variant	LOW	asa1
LR961994.1	433196	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433244	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433262	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	433346	C	G	snp	synonymous_variant	LOW	asa1
LR961994.1	433403	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433490	T	G	snp	synonymous_variant	LOW	asa1
LR961994.1	433586	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433640	T	G	snp	synonymous_variant	LOW	asa1

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	433687	C	T	snp	missense_variant	MODERATE	asa1
LR961994.1	433692	GCC	TCT	snp	missense_variant	MODERATE	asa1
LR961994.1	433769	T	G	snp	synonymous_variant	LOW	asa1
LR961994.1	433802	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433853	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433874	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433901	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433921	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	433937	A	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433964	GATT	AATC	snp	synonymous_variant	LOW	asa1
LR961994.1	434018	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434027	ACTCCCT	TTTACCA	snp	synonymous_variant	LOW	asa1
LR961994.1	434042	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434048	TG	CA	snp	missense_variant	MODERATE	asa1
LR961994.1	434060	G	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434066	T	G	snp	synonymous_variant	LOW	asa1
LR961994.1	434096	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434117	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434144	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434153	GGAT	AGAC	snp	synonymous_variant	LOW	asa1
LR961994.1	434162	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434171	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434195	TGCC	CGCT	snp	synonymous_variant	LOW	asa1
LR961994.1	434219	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434267	A	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434286	A	C	snp	missense_variant	MODERATE	asa1
LR961994.1	434291	CGAT	TGAC	snp	synonymous_variant	LOW	asa1
LR961994.1	434327	A	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434369	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434393	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434426	CA	AC	snp	missense_variant	MODERATE	asa1
LR961994.1	434465	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434493	A	C	snp	missense_variant	MODERATE	asa1
LR961994.1	434501	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	434542	C	T	complex	missense_variant	MODERATE	asa1
LR961994.1	434723	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434738	AC	GA	snp	missense_variant	MODERATE	asa1
LR961994.1	434771	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	434783	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434788	TTA	CCG	complex	missense_variant	MODERATE	asa1
LR961994.1	434869	A	G	snp	intergenic_region	MODIFIER	asa1-FCKDLICC_00450
LR961994.1	435087	C	T	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435099	TATA	CATT	del	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435108	T	C	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435118	T	C	complex	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435219	GATT	AGTA	snp	missense_variant	MODERATE	FCKDLICC_00450
LR961994.1	435246	T	G	snp	synonymous_variant	LOW	FCKDLICC_00450

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	435253	T	G	mnp	missense_variant	MODERATE	FCKDLICC_00450
LR961994.1	435259	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00450-FCKDLICC_00451
LR961994.1	457336	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	457343	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457463	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457469	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457481	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457544	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457556	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457688	A	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457735	A	G	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457757	A	G	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457769	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457781	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	457808	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457814	T	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457847	C	T	complex	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458017	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	458198	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458204	T	C	complex	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458222	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458264	G	A	mnp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458537	G	C	complex	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458573	T	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458623	G	C	snp	intergenic_region	MODIFIER	FCKDLICC_00475-FCKDLICC_00476
LR961994.1	515091	T	TA	snp	frameshift_variant	HIGH	FCKDLICC_00542
LR961994.1	517497	T	TA	snp	frameshift_variant	HIGH	ytrE
LR961994.1	524022	T	A	snp	missense_variant	MODERATE	clsA_1
LR961994.1	532536	G	T	snp	intergenic_region	MODIFIER	nhaC_2-FCKDLICC_00553
LR961994.1	545233	G	A	snp	missense_variant	MODERATE	FCKDLICC_00569
LR961994.1	549960	T	G	snp	synonymous_variant	LOW	mepA
LR961994.1	569950	C	T	snp	intergenic_region	MODIFIER	csn2-FCKDLICC_00592
LR961994.1	574637	T	C	complex	synonymous_variant	LOW	argR_3
LR961994.1	616669	A	G	complex	missense_variant	MODERATE	FCKDLICC_00634
LR961994.1	617119	C	A	snp	missense_variant	MODERATE	FCKDLICC_00634
LR961994.1	625795	T	C	complex	intergenic_region	MODIFIER	srlR_1-clcA
LR961994.1	627338	A	G	snp	synonymous_variant	LOW	clcA
LR961994.1	643744	G	A	snp	missense_variant	MODERATE	aguA
LR961994.1	660918	G	T	snp	stop_gained	HIGH	FCKDLICC_00672
LR961994.1	661884	T	C	snp	synonymous_variant	LOW	FCKDLICC_00673
LR961994.1	689899	C	T	complex	missense_variant	MODERATE	FCKDLICC_00695
LR961994.1	696162	T	G	snp	intergenic_region	MODIFIER	metK-bmr3
LR961994.1	703391	C	A	snp	intergenic_region	MODIFIER	FCKDLICC_00705-FCKDLICC_00706
LR961994.1	711543	A	G	snp	missense_variant	MODERATE	FCKDLICC_00714
LR961994.1	721813	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00722-FCKDLICC_00723
LR961994.1	735381	C	T	snp	missense_variant	MODERATE	rplY
LR961994.1	736535	G	C	snp	missense_variant	MODERATE	rsuA_2

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	779726	G	A	complex	missense_variant	MODERATE	queA
LR961994.1	799604	C	T	snp	missense_variant	MODERATE	FCKDLICC_00790
LR961994.1	803977	C	G	snp	missense_variant	MODERATE	arnC_1
LR961994.1	839110	T	G	snp	stop_lost&splice_region_variant	HIGH	FCKDLICC_00828
LR961994.1	852797	G	GT	snp	intergenic_region	MODIFIER	niaX-msmX_1
LR961994.1	854595	C	T	snp	intergenic_region	MODIFIER	mgsA-folT
LR961994.1	864272	A	G	snp	missense_variant	MODERATE	ung
LR961994.1	871876	T	C	mnp	missense_variant	MODERATE	malP
LR961994.1	873833	C	T	snp	synonymous_variant	LOW	ptsG_1
LR961994.1	913901	T	C	snp	synonymous_variant	LOW	ileS
LR961994.1	924932	A	G	snp	synonymous_variant	LOW	dgaR_2
LR961994.1	928397	A	G	snp	missense_variant	MODERATE	FCKDLICC_00920
LR961994.1	986601	G	A	snp	synonymous_variant	LOW	dppE_5
LR961994.1	1001928	G	GA	snp	frameshift_variant	HIGH	dinB_1
LR961994.1	1007069	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_01005-FCKDLICC_01006
LR961994.1	1008555	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_01006-FCKDLICC_01007
LR961994.1	1033348	T	C	snp	missense_variant	MODERATE	addB
LR961994.1	1043806	T	C	snp	missense_variant	MODERATE	pheT
LR961994.1	1044990	C	A	snp	missense_variant	MODERATE	glnP
LR961994.1	1047589	T	C	snp	missense_variant	MODERATE	glnQ_4
LR961994.1	1049004	T	C	mnp	missense_variant	MODERATE	rph
LR961994.1	1049251	G	C	snp	missense_variant	MODERATE	rph
LR961994.1	1068258	C	T	snp	synonymous_variant	LOW	gpsB
LR961994.1	1072485	T	C	snp	missense_variant	MODERATE	dnaD
LR961994.1	1080468	G	A	snp	synonymous_variant	LOW	dinG_1
LR961994.1	1082789	T	C	snp	synonymous_variant	LOW	FCKDLICC_01070
LR961994.1	1089212	T	C	snp	intergenic_region	MODIFIER	rpmE2-tagB
LR961994.1	1128322	C	CA	snp	intergenic_region	MODIFIER	aldC-FCKDLICC_01118
LR961994.1	1134378	C	T	mnp	missense_variant	MODERATE	spuD
LR961994.1	1149632	AT	A	complex	frameshift_variant	HIGH	FCKDLICC_01138
LR961994.1	1155782	G	C	snp	missense_variant	MODERATE	FCKDLICC_01142
LR961994.1	1158048	C	T	snp	missense_variant	MODERATE	exuR
LR961994.1	1170838	G	A	complex	missense_variant	MODERATE	FCKDLICC_01157
LR961994.1	1172033	C	T	complex	intergenic_region	MODIFIER	FCKDLICC_01158-FCKDLICC_01159
LR961994.1	1186428	A	G	complex	intergenic_region	MODIFIER	FCKDLICC_01169-rimP
LR961994.1	1193201	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_01178-FCKDLICC_01179
LR961994.1	1198061	C	T	snp	missense_variant	MODERATE	FCKDLICC_01189
LR961994.1	1200106	G	A	snp	missense_variant	MODERATE	FCKDLICC_01189
LR961994.1	1206670	ATAAAAG	CTTTTAT	complex	intergenic_region	MODIFIER	FCKDLICC_01194-truB
LR961994.1	1218553	C	T	complex	missense_variant	MODERATE	hemW
LR961994.1	1245496	G	T	snp	missense_variant	MODERATE	FCKDLICC_01232
LR961994.1	1246673	A	T	snp	missense_variant	MODERATE	FCKDLICC_01232
LR961994.1	1257719	A	T	snp	missense_variant	MODERATE	trxB
LR961994.1	1275293	G	T	complex	intergenic_region	MODIFIER	mgtA-pdhA
LR961994.1	1284584	G	A	snp	missense_variant	MODERATE	dhaK
LR961994.1	1285355	C	T	snp	intergenic_region	MODIFIER	dhaL-FCKDLICC_01265
LR961994.1	1285414	GTAACAAAAAA	G	mnp	intergenic_region	MODIFIER	dhaL-FCKDLICC_01265

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	1298837	T	A	snp	missense_variant	MODERATE	nrnA
LR961994.1	1303673	T	G	snp	intergenic_region	MODIFIER	cshB-FCKDLICC_01281
LR961994.1	1310846	T	C	snp	synonymous_variant	LOW	FCKDLICC_01286
LR961994.1	1328762	TA	T	snp	intergenic_region	MODIFIER	cadA-zapA
LR961994.1	1350325	T	C	snp	missense_variant	MODERATE	ntpK
LR961994.1	1377002	A	G	snp	missense_variant	MODERATE	dnaG
LR961994.1	1377462	T	G	snp	missense_variant	MODERATE	dnaG
LR961994.1	1386307	T	C	snp	synonymous_variant	LOW	FCKDLICC_01353
LR961994.1	1392879	C	T	snp	missense_variant	MODERATE	xerD
LR961994.1	1405691	C	T	complex	intergenic_region	MODIFIER	hup-FCKDLICC_01374
LR961994.1	1419699	G	A	snp	synonymous_variant	LOW	aroA
LR961994.1	1424476	T	C	snp	synonymous_variant	LOW	FCKDLICC_01394
LR961994.1	1429163	G	T	snp	missense_variant	MODERATE	ettA
LR961994.1	1436490	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_01404-FCKDLICC_01405
LR961994.1	1439861	G	A	snp	missense_variant	MODERATE	nox
LR961994.1	1441920	A	G	snp	missense_variant	MODERATE	paiA_1
LR961994.1	1459207	T	A	snp	missense_variant	MODERATE	degA_2
LR961994.1	1468810	T	A	snp	missense_variant	MODERATE	clsA_2
LR961994.1	1468861	A	C	snp	missense_variant	MODERATE	clsA_2
LR961994.1	1483986	CA	C	complex	intergenic_region	MODIFIER	FCKDLICC_01442-pduA
LR961994.1	1487314	T	C	snp	missense_variant	MODERATE	eutC
LR961994.1	1517534	T	C	snp	missense_variant	MODERATE	bfmBAB
LR961994.1	1546750	C	A	snp	missense_variant	MODERATE	FCKDLICC_01508
LR961994.1	1552135	T	G	snp	missense_variant	MODERATE	FCKDLICC_01516
LR961994.1	1586374	T	C	snp	missense_variant	MODERATE	yheI
LR961994.1	1587575	G	A	snp	missense_variant	MODERATE	nfo
LR961994.1	1608514	G	T	snp	missense_variant	MODERATE	FCKDLICC_01569
LR961994.1	1620218	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_01579-celA_3
LR961994.1	1628987	C	T	snp	missense_variant	MODERATE	purH
LR961994.1	1635969	G	A	snp	missense_variant	MODERATE	purC
LR961994.1	1636114	C	CA	complex	intergenic_region	MODIFIER	purC-purK_1
LR961994.1	1638733	CATAA	C	snp	frameshift_variant	HIGH	FCKDLICC_01602
LR961994.1	1659351	G	A	snp	intergenic_region	MODIFIER	nagR_3-gspA_1
LR961994.1	1661492	A	G	snp	intergenic_region	MODIFIER	gspA_2-ltaS1_2
LR961994.1	1677168	A	T	snp	missense_variant	MODERATE	FCKDLICC_01633
LR961994.1	1683386	C	T	snp	missense_variant	MODERATE	adhA
LR961994.1	1701880	T	G	ins	missense_variant	MODERATE	sorC
LR961994.1	1712764	G	C	ins	missense_variant	MODERATE	xerC_5
LR961994.1	1715398	T	G	snp	synonymous_variant	LOW	FCKDLICC_01666
LR961994.1	1725450	T	C	snp	missense_variant	MODERATE	gatC_4
LR961994.1	1729110	C	T	snp	synonymous_variant	LOW	hxlB_2
LR961994.1	1732516	TG	T	snp	intergenic_region	MODIFIER	rpiB-srlR_2
LR961994.1	1739619	T	C	snp	intergenic_region	MODIFIER	manR_3-FCKDLICC_01695
LR961994.1	1742160	T	G	snp	synonymous_variant	LOW	FCKDLICC_01697
LR961994.1	1745527	G	A	complex	synonymous_variant	LOW	FCKDLICC_01700
LR961994.1	1752075	T	C	snp	missense_variant	MODERATE	FCKDLICC_01707
LR961994.1	1774745	C	T	snp	missense_variant	MODERATE	glpO

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	1799647	A	G	complex	intergenic_region	MODIFIER	dgaR_3-FCKDLICC_01753
LR961994.1	1832300	T	C	complex	intergenic_region	MODIFIER	rlmN-yxdM
LR961994.1	1870555	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_01814-FCKDLICC_01816
LR961994.1	1916148	T	A	snp	missense_variant	MODERATE	FCKDLICC_01849
LR961994.1	1927650	T	C	snp	missense_variant	MODERATE	FCKDLICC_01860
LR961994.1	1965488	G	T	snp	missense_variant	MODERATE	FCKDLICC_01892
LR961994.1	1967768	T	C	snp	missense_variant	MODERATE	FCKDLICC_01892
LR961994.1	1969749	C	A	complex	synonymous_variant	LOW	FCKDLICC_01894
LR961994.1	1972044	C	A	snp	intergenic_region	MODIFIER	ngcF-FCKDLICC_01897
LR961994.1	2001182	T	C	snp	missense_variant	MODERATE	FCKDLICC_01919
LR961994.1	2001229	T	G	snp	missense_variant	MODERATE	FCKDLICC_01919
LR961994.1	2016412	T	G	snp	missense_variant	MODERATE	aspB
LR961994.1	2019099	C	T	snp	synonymous_variant	LOW	dinG_2
LR961994.1	2021297	G	A	snp	stop_gained	HIGH	FCKDLICC_01934
LR961994.1	2033006	G	A	snp	missense_variant	MODERATE	gdh_2
LR961994.1	2038805	A	G	snp	synonymous_variant	LOW	FCKDLICC_01944
LR961994.1	2048135	A	G	snp	missense_variant	MODERATE	FCKDLICC_01954
LR961994.1	2055955	G	A	snp	synonymous_variant	LOW	dgkA
LR961994.1	2062908	G	T	snp	missense_variant	MODERATE	thrB
LR961994.1	2067722	G	A	snp	synonymous_variant	LOW	pgcA
LR961994.1	2075872	T	A	snp	intergenic_region	MODIFIER	guaD-FCKDLICC_01982
LR961994.1	2077658	T	G	snp	missense_variant	MODERATE	ybbH_3
LR961994.1	2077761	C	A	snp	missense_variant	MODERATE	ybbH_3
LR961994.1	2080014	C	T	snp	missense_variant	MODERATE	murQ_2
LR961994.1	2085700	C	T	snp	missense_variant	MODERATE	FCKDLICC_01992
LR961994.1	2087089	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_01994-panE_2
LR961994.1	2090742	A	G	snp	stop_lost&splice_region_variant	HIGH	comEC_2
LR961994.1	2095730	A	C	snp	missense_variant	MODERATE	FCKDLICC_02005
LR961994.1	2103162	A	G	complex	synonymous_variant	LOW	recQ_2
LR961994.1	2145116	A	G	snp	synonymous_variant	LOW	metN2
LR961994.1	2163543	A	G	complex	synonymous_variant	LOW	prmC
LR961994.1	2167433	C	T	complex	intergenic_region	MODIFIER	FCKDLICC_02066-ntpJ_2
LR961994.1	2174691	C	A	snp	missense_variant	MODERATE	FCKDLICC_02070
LR961994.1	2183361	A	T	snp	missense_variant	MODERATE	FCKDLICC_02079
LR961994.1	2191006	C	T	snp	missense_variant	MODERATE	dpaL
LR961994.1	2252647	T	C	snp	missense_variant	MODERATE	FCKDLICC_02146
LR961994.1	2270738	TC	T	snp	frameshift_variant	HIGH	pspA_1
LR961994.1	2273616	A	C	snp	synonymous_variant	LOW	cutC
LR961994.1	2309187	CAAAGT	ACTTTG	snp	intergenic_region	MODIFIER	FCKDLICC_02197-FCKDLICC_02198
LR961994.1	2317394	C	T	snp	missense_variant	MODERATE	ebgA
LR961994.1	2329674	G	A	snp	intergenic_region	MODIFIER	sdhA_2-dppA
LR961994.1	2335288	A	G	snp	missense_variant	MODERATE	nusG
LR961994.1	2337698	T	C	ins	missense_variant	MODERATE	FCKDLICC_02227
LR961994.1	2342547	T	C	snp	intergenic_region	MODIFIER	polC_2-FCKDLICC_02232
LR961994.1	2348864	G	A	snp	synonymous_variant	LOW	FCKDLICC_02237
LR961994.1	2355342	A	G	snp	synonymous_variant	LOW	dltA
LR961994.1	2370389	A	G	snp	missense_variant	MODERATE	tenA

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	2374821	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_02264-emrB_2
LR961994.1	2381645	G	T	snp	missense_variant	MODERATE	FCKDLICC_02273
LR961994.1	2384760	T	A	snp	missense_variant	MODERATE	galE_2
LR961994.1	2389844	C	A	snp	missense_variant	MODERATE	glcK
LR961994.1	2398227	G	T	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398238	CTTGA	TTTGG	snp	missense_variant	MODERATE	FCKDLICC_02291
LR961994.1	2398254	T	A	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398266	AGC	TGT	snp	missense_variant	MODERATE	FCKDLICC_02291
LR961994.1	2398284	C	A	complex	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398290	CG	TC	snp	missense_variant	MODERATE	FCKDLICC_02291
LR961994.1	2398299	ATC	GTT	snp	missense_variant	MODERATE	FCKDLICC_02291
LR961994.1	2398314	G	A	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398323	A	G	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398350	TACGTGTGTT	CACATGAGTC	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398371	C	T	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398383	T	A	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398389	T	C	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398398	T	C	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2398458	C	T	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2399031	G	A	ins	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2399040	G	A	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2399258	T	C	ins	missense_variant	MODERATE	FCKDLICC_02291
LR961994.1	2399289	C	T	snp	synonymous_variant	LOW	FCKDLICC_02291
LR961994.1	2399411	A	G	snp	synonymous_variant	LOW	FCKDLICC_02292
LR961994.1	2399460	T	G	snp	synonymous_variant	LOW	FCKDLICC_02292
LR961994.1	2399713	C	G	snp	synonymous_variant	LOW	FCKDLICC_02293
LR961994.1	2399816	T	C	snp	synonymous_variant	LOW	FCKDLICC_02294
LR961994.1	2399831	A	G	ins	synonymous_variant	LOW	FCKDLICC_02294
LR961994.1	2400329	A	G	snp	missense_variant	MODERATE	FCKDLICC_02296
LR961994.1	2400450	T	C	snp	missense_variant	MODERATE	FCKDLICC_02296
LR961994.1	2400592	CCCAT	ACCAC	snp	missense_variant	MODERATE	FCKDLICC_02296
LR961994.1	2427816	ACGA	GCGG	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2427828	A	T	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2427860	A	G	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2427867	C	T	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2427903	T	A	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2427912	TACTTCGATA	AACATCAATC	snp	missense_variant	MODERATE	FCKDLICC_02337
LR961994.1	2427927	A	G	snp	synonymous_variant	LOW	FCKDLICC_02337
LR961994.1	2428427	CATCTCTAAT	TATTTTTAAA	complex	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428462	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428471	T	C	complex	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428541	A	T	snp	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428571	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428583	A	G	ins	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428662	CA	TG	snp	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428691	G	A	complex	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340
LR961994.1	2428712	T	A	complex	intergenic_region	MODIFIER	FCKDLICC_02339-FCKDLICC_02340

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	2428726	T	C	snp	synonymous_variant	LOW	FCKDLICC_02340
LR961994.1	2428942	C	T	ins	synonymous_variant	LOW	FCKDLICC_02340
LR961994.1	2442668	C	T	snp	synonymous_variant	LOW	aadK
LR961994.1	2451141	A	G	ins	synonymous_variant	LOW	yqeH
LR961994.1	2455150	A	G	snp	missense_variant	MODERATE	accC
LR961994.1	2489750	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_02404-greA
LR961994.1	2493239	G	GCTT	snp	conservative_inframe_insertion	MODERATE	mnaA
LR961994.1	2513889	G	A	snp	synonymous_variant	LOW	pbuG
LR961994.1	2522524	A	C	snp	missense_variant	MODERATE	ulaA
LR961994.1	2529461	T	G	snp	synonymous_variant	LOW	strB1
LR961994.1	2560282	T	C	snp	intergenic_region	MODIFIER	ybbW-paiA_2
LR961994.1	2577179	T	C	snp	missense_variant	MODERATE	FCKDLICC_02481
LR961994.1	2601249	C	T	snp	synonymous_variant	LOW	pepA
LR961994.1	2606657	T	TCATGATTGG	snp	conservative_inframe_insertion	MODERATE	nagA
LR961994.1	2608247	T	C	snp	missense_variant	MODERATE	FCKDLICC_02508
LR961994.1	2613414	G	A	snp	synonymous_variant	LOW	thiT
LR961994.1	2630013	G	T	snp	missense_variant	MODERATE	FCKDLICC_02533
LR961994.1	2677885	A	G	snp	missense_variant	MODERATE	pknD
LR961994.1	2693059	T	C	snp	missense_variant	MODERATE	manX_5
LR961994.1	2694831	C	T	snp	missense_variant	MODERATE	sorA_4
LR961994.1	2696268	A	G	snp	missense_variant	MODERATE	hcxA_2
LR961994.1	2699528	T	TA	snp	frameshift_variant	HIGH	hexR
LR961994.1	2746745	T	G	snp	synonymous_variant	LOW	FCKDLICC_02642
LR961994.1	2748083	C	CA	complex	intergenic_region	MODIFIER	FCKDLICC_02642-aes
LR961994.1	2767209	A	G	snp	synonymous_variant	LOW	dgaR_5
LR961994.1	2783177	T	A	ins	missense_variant	MODERATE	xylB
LR961994.1	2791633	A	G	snp	synonymous_variant	LOW	rpoB
LR961994.1	2800928	T	C	snp	missense_variant	MODERATE	FCKDLICC_02687
LR961994.1	2832117	A	G	del	missense_variant	MODERATE	licC_6
LR961994.1	2840661	C	T	snp	synonymous_variant	LOW	phoP_2
LR961994.1	2878169	A	G	snp	missense_variant	MODERATE	citC
LR961994.1	2880686	A	C	snp	missense_variant	MODERATE	FCKDLICC_02760

Supplement file 1. Annotation and predicted effect of variants in E2 chromosome using GCA_905120835.1 as reference

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	2844	C	CT	ins	intergenic_region	MODIFIER	dnaN-FCKDLICC_00003
LR961994.1	21946	A	G	snp	synonymous_variant	LOW	dgaR_1
LR961994.1	24354	A	T	snp	synonymous_variant	LOW	sorA_1
LR961994.1	44147	G	A	snp	intergenic_region	MODIFIER	radA-yacL
LR961994.1	52101	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_00045-veg
LR961994.1	57880	A	G	snp	missense_variant	MODERATE	glmU
LR961994.1	65059	T	G	snp	intergenic_region	MODIFIER	dppE_1-FCKDLICC_00055
LR961994.1	75803	GTCAAA	TTTGAC	complex	intergenic_region	MODIFIER	FCKDLICC_00065-larR_1
LR961994.1	89641	C	T	snp	synonymous_variant	LOW	FCKDLICC_00081
LR961994.1	91010	A	C	snp	missense_variant	MODERATE	FCKDLICC_00081
LR961994.1	98604	A	G	snp	synonymous_variant	LOW	serS_1
LR961994.1	121935	A	T	snp	intergenic_region	MODIFIER	FCKDLICC_00110-FCKDLICC_00111
LR961994.1	128239	A	G	snp	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128268	T	C	snp	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128293	CTCTCCCCC	TTCCCCACT	complex	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128308	A	G	snp	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128316	CACAAATAAG	AAC	complex	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128335	AACATA	GCCAAT	complex	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128345	C	A	snp	intergenic_region	MODIFIER	rep-FCKDLICC_00115
LR961994.1	128378	ATGTC	TTGTG	complex	missense_variant	MODERATE	FCKDLICC_00115
LR961994.1	142141	G	T	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142149	ATAC	CTAT	complex	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142163	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142171	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142230	GA	AT	complex	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142296	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142310	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142387	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142411	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142444	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142480	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00129-tet(M)
LR961994.1	142991	A	G	snp	missense_variant	MODERATE	tet(M)
LR961994.1	143956	C	T	snp	synonymous_variant	LOW	tet(M)
LR961994.1	143968	A	G	snp	synonymous_variant	LOW	tet(M)
LR961994.1	143989	T	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144010	CGGT	TGGC	complex	synonymous_variant	LOW	tet(M)
LR961994.1	144037	T	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144059	A	G	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144071	C	T	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144076	G	A	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144081	CCTTTAG	TCTTAAA	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144131	G	A	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144139	A	G	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144160	G	A	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144166	T	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144178	C	T	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144190	AAAT	GAAC	complex	synonymous_variant	LOW	tet(M)
LR961994.1	144200	AAT	GAC	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144209	C	T	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144230	ATTAT	CTTAG	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144244	T	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144253	A	G	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144265	T	A	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144271	CAA	TAG	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144286	T	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144298	GCT	ACG	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144314	G	A	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144320	C	T	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144328	ATATCAG	GTACCAT	complex	missense_variant	MODERATE	tet(M)
LR961994.1	144340	CACTGGC	TACCGGT	complex	synonymous_variant	LOW	tet(M)
LR961994.1	144362	A	C	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144372	T	C	snp	missense_variant	MODERATE	tet(M)
LR961994.1	144395	A	C	snp	synonymous_variant	LOW	tet(M)
LR961994.1	144424	CG	TA	complex	intergenic_region	MODIFIER	tet(M)-FCKDLICC_00131
LR961994.1	144456	A	G	snp	intergenic_region	MODIFIER	tet(M)-FCKDLICC_00131
LR961994.1	144476	C	T	snp	intergenic_region	MODIFIER	tet(M)-FCKDLICC_00131
LR961994.1	144484	T	A	snp	intergenic_region	MODIFIER	tet(M)-FCKDLICC_00131
LR961994.1	144502	G	T	snp	intergenic_region	MODIFIER	tet(M)-FCKDLICC_00131
LR961994.1	144855	T	C	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144860	C	T	snp	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	144867	CCTTCGA	TCGTCGG	complex	synonymous_variant	LOW	FCKDLICC_00131

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	144882	AATCAAA	TACCAAG	complex	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	144893	CTGT	TAGC	complex	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	144917	TAGAT	CGGAA	complex	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	144927	G	A	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144933	T	C	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144942	A	G	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144948	CAGC	AAGA	complex	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144958	GTACTT	ATGCTG	complex	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	144972	T	C	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144978	G	A	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144984	T	A	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144990	T	A	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	144996	ATAA	GTAC	complex	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	145019	TTGT	ATGC	complex	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	145041	GAGACCC	TAAACCT	complex	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	145055	T	A	snp	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	145062	C	T	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	145071	C	G	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	145082	C	A	snp	missense_variant	MODERATE	FCKDLICC_00131
LR961994.1	145095	G	A	snp	synonymous_variant	LOW	FCKDLICC_00131
LR961994.1	145120	CTAAAT	GTAGAC	complex	intergenic_region	MODIFIER	FCKDLICC_00131-FCKDLICC_00132
LR961994.1	145143	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_00131-FCKDLICC_00132
LR961994.1	146522	AATTCGAAAGT	GATCCGAGAAGC	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146549	G	GC	ins	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146554	CATGACT	AACGACA	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146586	GTATAGCC	ATACAGTCA	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146607	TAGAAG	GATAAAAC	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146617	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146627	GGTAC	AGTAT	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146653	AAGTGAAATTCCT	GGGTGGAACTCCC	complex	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	146679	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_00133-FCKDLICC_00134
LR961994.1	151595	C	T	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151604	A	G	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151625	T	A	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151635	GCATTC	ACTTTT	complex	missense_variant	MODERATE	FCKDLICC_00140
LR961994.1	151661	A	G	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151670	TTTG	CTTA	complex	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151700	TGTAAAA	GGTTAAG	complex	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151724	T	C	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151761	C	T	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151766	G	A	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151772	C	A	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151796	G	A	snp	synonymous_variant	LOW	FCKDLICC_00140
LR961994.1	151818	TAGCAGTAAT	AGTTAATAAG	complex	stop_lost&splice_region_variant	HIGH	FCKDLICC_00140
LR961994.1	151833	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151840	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151858	GTT	ATG	complex	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151865	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151872	TTTTATAA	ATTTTTGTAT	complex	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151888	AT	G	complex	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151903	G	T	snp	intergenic_region	MODIFIER	FCKDLICC_00140-FCKDLICC_00141
LR961994.1	151940	C	T	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152000	G	A	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152012	AGTGAGAGGA	TGTTCGTGGT	complex	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152027	TTTTACT	GTTCACC	complex	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152039	T	A	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152048	A	G	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152063	A	G	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152081	GA	AC	mnp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152105	T	C	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152117	T	A	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152144	TAACGTG	CAATGTA	complex	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152168	T	A	snp	synonymous_variant	LOW	FCKDLICC_00141
LR961994.1	152239	GTTAA	ATTAT	complex	missense_variant	MODERATE	FCKDLICC_00142
LR961994.1	152257	TACA	CACG	complex	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152279	CA	GT	mnp	missense_variant	MODERATE	FCKDLICC_00142
LR961994.1	152317	G	T	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152350	C	G	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152359	A	T	snp	missense_variant	MODERATE	FCKDLICC_00142
LR961994.1	152377	G	A	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152467	TGAA	CGAG	complex	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152485	GACT	AACG	complex	synonymous_variant	LOW	FCKDLICC_00142

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	152497	T	G	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152507	ACTTTC	GCTTAT	complex	missense_variant	MODERATE	FCKDLICC_00142
LR961994.1	152518	C	T	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152542	G	A	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152548	T	A	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152557	ACGTTCAGCA	CCGTGCAGTG	complex	missense_variant	MODERATE	FCKDLICC_00142
LR961994.1	152578	G	A	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	152584	G	A	snp	synonymous_variant	LOW	FCKDLICC_00142
LR961994.1	154291	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154310	C	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154351	A	T	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154363	G	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154385	A	C	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154390	TCCTTTA	CCCATTG	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154402	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154411	G	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154426	ATACA	TTATG	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154444	GATA	AATC	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154457	A	G	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154462	T	G	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154493	A	G	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154498	A	G	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154510	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154516	GGGG	AGGC	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154534	G	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154555	ACA	GCG	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154573	CGGG	TGGC	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154582	CACAAAAAAA	TACGAAGAAG	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154604	AGTGACGGC	TCTGATGGT	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154620	A	G	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154630	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154675	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154693	GGGT	AGGA	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154708	G	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154714	A	G	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154745	A	G	snp	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154771	G	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154786	TTTTG	CTTTA	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154802	T	C	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154807	AACA	TTCT	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154816	G	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154825	G	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	154833	GG	CC	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154839	AG	GA	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154855	CG	TA	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	154861	C	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155317	T	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155335	A	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155342	CGT	AGA	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155353	AGAATTTTTA	CGAGTTTATC	complex	missense_variant	MODERATE	FCKDLICC_00147
LR961994.1	155368	A	G	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155375	CGTGAC	AGAGAT	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155395	C	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155401	T	G	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155407	ACGTC	CCGTT	complex	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155416	T	A	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	155422	A	T	snp	synonymous_variant	LOW	FCKDLICC_00147
LR961994.1	156475	A	C	snp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156486	ATTA	GTCT	complex	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156495	AAT	GAC	complex	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156502	G	A	snp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156520	A	T	snp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156525	C	T	snp	synonymous_variant	LOW	FCKDLICC_00149
LR961994.1	156541	A	G	snp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156558	T	C	snp	synonymous_variant	LOW	FCKDLICC_00149
LR961994.1	156606	T	G	snp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156637	AA	GG	mnp	missense_variant	MODERATE	FCKDLICC_00149
LR961994.1	156651	GAAA	AAGC	complex	intergenic_region	MODIFIER	FCKDLICC_00149-FCKDLICC_00150
LR961994.1	156659	A	T	snp	missense_variant	MODERATE	FCKDLICC_00150
LR961994.1	156670	C	G	snp	missense_variant	MODERATE	FCKDLICC_00150
LR961994.1	156683	AGC	CGT	complex	missense_variant	MODERATE	FCKDLICC_00150
LR961994.1	156690	C	T	snp	synonymous_variant	LOW	FCKDLICC_00150

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	156700	G	T	snp	missense_variant	MODERATE	FCKDLICC_00150
LR961994.1	156705	G	A	snp	synonymous_variant	LOW	FCKDLICC_00150
LR961994.1	157240	GC	AT	complex	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157269	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157289	GTAA	ATGT	complex	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157308	TTAAATG	AGTAGATA	complex	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157323	C	CT	ins	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157339	TTTGAAA	GTTAAAT	complex	intergenic_region	MODIFIER	FCKDLICC_00150-FCKDLICC_00151
LR961994.1	157352	A	G	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157494	ATGACAA	GTGGAAG	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	157505	AA	GC	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	157511	C	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157520	A	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157541	GATCGATTTTT	AATTGACTTTC	complex	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157556	GATTCGT	AATACGA	complex	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157568	GACT	AACA	complex	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157582	AA	GG	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	157595	A	G	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157607	T	C	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157619	T	C	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157625	C	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157637	A	G	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157649	C	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157658	T	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157664	T	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157973	A	G	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	157987	ATGGT	GTGGG	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	158007	CATC	AATA	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	158015	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158021	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158033	A	G	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158060	TCCTTTG	GCCATTA	complex	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158084	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158090	C	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158096	ATTG	GTTA	complex	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158108	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158113	ATATGAGCGTCGAA	GTGTAAGTGTTGAG	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	158132	TCGAA	AAGAG	complex	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	158142	C	T	snp	missense_variant	MODERATE	FCKDLICC_00151
LR961994.1	158177	C	T	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158183	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158195	G	A	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158201	A	C	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	158211	T	C	snp	synonymous_variant	LOW	FCKDLICC_00151
LR961994.1	159109	T	A	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159118	G	A	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159129	A	G	snp	missense_variant	MODERATE	FCKDLICC_00152
LR961994.1	159139	T	C	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159148	A	T	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159154	C	T	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159169	A	T	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159182	GT	AC	complex	missense_variant	MODERATE	FCKDLICC_00152
LR961994.1	159198	TTGCTTTG	CTGGATTA	complex	missense_variant	MODERATE	FCKDLICC_00152
LR961994.1	159211	GC	AT	complex	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	159217	TAGCATT	CAGTGTA	complex	missense_variant	MODERATE	FCKDLICC_00152
LR961994.1	159229	T	A	snp	synonymous_variant	LOW	FCKDLICC_00152
LR961994.1	167163	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167199	ATC	GTT	complex	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167292	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167331	GGCGC	AGCGT	complex	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167346	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167378	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167399	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167404	GA	CT	complex	intergenic_region	MODIFIER	FCKDLICC_00161-FCKDLICC_00162
LR961994.1	167492	CCCA	TCCG	complex	synonymous_variant	LOW	FCKDLICC_00162
LR961994.1	167507	G	A	snp	synonymous_variant	LOW	FCKDLICC_00162
LR961994.1	167531	T	A	snp	synonymous_variant	LOW	FCKDLICC_00162
LR961994.1	167564	T	G	snp	synonymous_variant	LOW	FCKDLICC_00162
LR961994.1	167579	G	A	snp	synonymous_variant	LOW	FCKDLICC_00162
LR961994.1	167625	C	A	snp	missense_variant	MODERATE	FCKDLICC_00162
LR961994.1	167643	TACG	CACT	complex	intergenic_region	MODIFIER	FCKDLICC_00162-FCKDLICC_00163
LR961994.1	167679	ATTG	TTTA	complex	intergenic_region	MODIFIER	FCKDLICC_00162-FCKDLICC_00163

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	176385	T	C	snp	splice_region_variant&stop_retained_variant	LOW	guaA
LR961994.1	177447	G	T	snp	missense_variant	MODERATE	guaA
LR961994.1	186618	G	A	snp	missense_variant	MODERATE	tmpC_2
LR961994.1	202573	G	A	snp	missense_variant	MODERATE	yxdL_2
LR961994.1	231227	C	T	snp	missense_variant	MODERATE	scmP_1
LR961994.1	232892	A	G	snp	synonymous_variant	LOW	ecfA2
LR961994.1	240320	T	C	snp	missense_variant	MODERATE	mltF_1
LR961994.1	242538	T	G	snp	missense_variant	MODERATE	dapH_1
LR961994.1	243901	A	G	snp	intergenic_region	MODIFIER	niaR-FCKDLICC_00253
LR961994.1	248554	C	T	snp	intergenic_region	MODIFIER	ldh_1-pth
LR961994.1	251483	C	A	snp	synonymous_variant	LOW	mfd
LR961994.1	255248	T	C	snp	missense_variant	MODERATE	FCKDLICC_00260
LR961994.1	256921	A	G	snp	missense_variant	MODERATE	tilS
LR961994.1	257901	C	T	snp	synonymous_variant	LOW	hpt
LR961994.1	278518	C	T	snp	synonymous_variant	LOW	gpmA_2
LR961994.1	296169	G	GA	ins	intergenic_region	MODIFIER	gmuC-murR_1
LR961994.1	313462	G	A	snp	missense_variant	MODERATE	FCKDLICC_00339
LR961994.1	324836	T	C	snp	synonymous_variant	LOW	FCKDLICC_00348
LR961994.1	341535	T	C	snp	missense_variant	MODERATE	atzC_1
LR961994.1	388790	T	G	snp	intergenic_region	MODIFIER	FCKDLICC_00405-dtpT
LR961994.1	400489	A	C	snp	intergenic_region	MODIFIER	nagB_1-fcbA2
LR961994.1	404808	A	G	snp	missense_variant	MODERATE	manZ_2
LR961994.1	420028	A	C	snp	missense_variant	MODERATE	nrdE2
LR961994.1	426883	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426893	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426901	CTT	TCC	complex	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426908	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426921	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426957	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_00443-FCKDLICC_00444
LR961994.1	426982	A	C	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	426991	C	T	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427009	T	A	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427014	G	T	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427021	C	T	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427038	TA	CG	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427045	CTCCG	ACCTA	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427054	A	G	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427082	ACT	GCC	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427091	CT	GA	complex	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427133	T	C	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427148	TGA	AAG	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427155	C	T	snp	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427162	C	T	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427172	GGA	AGG	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427183	T	C	snp	synonymous_variant	LOW	FCKDLICC_00444
LR961994.1	427189	TTTGTTG	AGATTTA	complex	missense_variant	MODERATE	FCKDLICC_00444
LR961994.1	427208	A	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427213	TCCG	CCCA	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427221	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427239	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427266	A	G	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427272	G	A	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427295	TATT	CATC	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427329	GAC	AAT	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427347	T	C	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427353	A	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427362	G	T	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427367	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427443	AGAGTACT	GGAATACC	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427459	G	T	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427469	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427487	A	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427499	T	C	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427519	GTAAT	TTAGC	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427545	G	A	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427576	TTC	ATG	complex	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427584	A	G	snp	synonymous_variant	LOW	FCKDLICC_00445
LR961994.1	427589	T	G	snp	missense_variant	MODERATE	FCKDLICC_00445
LR961994.1	427616	T	TC	ins	intergenic_region	MODIFIER	FCKDLICC_00445-FCKDLICC_00446
LR961994.1	427644	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_00445-FCKDLICC_00446
LR961994.1	431205	TCT	CCA	complex	missense_variant	MODERATE	asa1

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	431213	AAATGGAG	TAAAGGTA	complex	missense_variant	MODERATE	asa1
LR961994.1	431228	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431243	C	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431275	TAGG	CAGA	complex	missense_variant	MODERATE	asa1
LR961994.1	431289	TCAG	T	del	conservative_inframe_deletion	MODERATE	asa1
LR961994.1	431315	A	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431325	G	A	snp	missense_variant	MODERATE	asa1
LR961994.1	431367	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	431378	TGCT	AGCG	complex	missense_variant	MODERATE	asa1
LR961994.1	431394	ACT	TCC	complex	missense_variant	MODERATE	asa1
LR961994.1	431429	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	431447	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431459	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431468	TGTC	CGTT	complex	synonymous_variant	LOW	asa1
LR961994.1	431492	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431498	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431523	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	431564	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	431576	G	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431585	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	431648	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	431653	T	C	snp	missense_variant	MODERATE	asa1
LR961994.1	431873	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432306	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	432318	A	G	snp	missense_variant	MODERATE	asa1
LR961994.1	432710	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	432749	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	432773	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	432902	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432935	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	432977	TCTTC	CCTTA	complex	missense_variant	MODERATE	asa1
LR961994.1	432989	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433001	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433031	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433133	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	433139	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433490	T	G	snp	synonymous_variant	LOW	asa1
LR961994.1	433586	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	433874	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434096	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434117	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434144	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434171	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434195	TGCC	CGCT	complex	synonymous_variant	LOW	asa1
LR961994.1	434219	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434267	A	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434286	A	C	snp	missense_variant	MODERATE	asa1
LR961994.1	434291	CGAT	TGAC	complex	synonymous_variant	LOW	asa1
LR961994.1	434327	A	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434369	T	C	snp	synonymous_variant	LOW	asa1
LR961994.1	434393	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434426	CA	AC	mnp	missense_variant	MODERATE	asa1
LR961994.1	434465	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434493	A	C	snp	missense_variant	MODERATE	asa1
LR961994.1	434501	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	434542	C	T	snp	missense_variant	MODERATE	asa1
LR961994.1	434723	G	A	snp	synonymous_variant	LOW	asa1
LR961994.1	434738	AC	GA	mnp	missense_variant	MODERATE	asa1
LR961994.1	434771	A	G	snp	synonymous_variant	LOW	asa1
LR961994.1	434783	C	T	snp	synonymous_variant	LOW	asa1
LR961994.1	434788	TTA	CCG	mnp	missense_variant	MODERATE	asa1
LR961994.1	434869	A	G	snp	intergenic_region	MODIFIER	asa1-FCKDLICC_00450
LR961994.1	435087	C	T	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435099	TATA	CATT	complex	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435108	T	C	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435118	T	C	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435219	GATT	AGTA	complex	missense_variant	MODERATE	FCKDLICC_00450
LR961994.1	435246	T	G	snp	synonymous_variant	LOW	FCKDLICC_00450
LR961994.1	435253	T	G	snp	missense_variant	MODERATE	FCKDLICC_00450
LR961994.1	435259	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00450-FCKDLICC_00451
LR961994.1	457336	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	457343	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	457463	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457469	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457481	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457544	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457556	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457688	A	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457735	A	G	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457757	A	G	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457769	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457781	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	457808	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457814	T	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	457847	C	T	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458017	T	C	snp	missense_variant	MODERATE	FCKDLICC_00475
LR961994.1	458198	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458204	T	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458222	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458264	G	A	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458537	G	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	458573	T	C	snp	synonymous_variant	LOW	FCKDLICC_00475
LR961994.1	515091	T	TA	ins	frameshift_variant	HIGH	FCKDLICC_00542
LR961994.1	517497	T	TA	ins	frameshift_variant	HIGH	ytrE
LR961994.1	524022	T	A	snp	missense_variant	MODERATE	clsA_1
LR961994.1	532536	G	T	snp	intergenic_region	MODIFIER	nhaC_2-FCKDLICC_00553
LR961994.1	545233	G	A	snp	missense_variant	MODERATE	FCKDLICC_00569
LR961994.1	549960	T	G	snp	synonymous_variant	LOW	mepA
LR961994.1	569950	C	T	snp	intergenic_region	MODIFIER	csn2-FCKDLICC_00592
LR961994.1	572409	AG	A	del	frameshift_variant	HIGH	opuCA_1
LR961994.1	574637	T	C	snp	synonymous_variant	LOW	argR_3
LR961994.1	576622	G	A	snp	missense_variant	MODERATE	FCKDLICC_00597
LR961994.1	613295	C	T	snp	missense_variant	MODERATE	FCKDLICC_00630
LR961994.1	616669	A	G	snp	missense_variant	MODERATE	FCKDLICC_00634
LR961994.1	617119	C	A	snp	missense_variant	MODERATE	FCKDLICC_00634
LR961994.1	619702	T	C	snp	synonymous_variant	LOW	tig
LR961994.1	625795	T	C	snp	intergenic_region	MODIFIER	srlR_1-clcA
LR961994.1	627338	A	G	snp	synonymous_variant	LOW	clcA
LR961994.1	643744	G	A	snp	missense_variant	MODERATE	aguA
LR961994.1	650235	T	A	snp	stop_gained	HIGH	FCKDLICC_00662
LR961994.1	653666	G	A	snp	missense_variant	MODERATE	dap
LR961994.1	660918	G	T	snp	stop_gained	HIGH	FCKDLICC_00672
LR961994.1	661884	T	C	snp	synonymous_variant	LOW	FCKDLICC_00673
LR961994.1	689899	C	T	snp	missense_variant	MODERATE	FCKDLICC_00695
LR961994.1	696162	T	G	snp	intergenic_region	MODIFIER	metK-bmr3
LR961994.1	703391	C	A	snp	intergenic_region	MODIFIER	FCKDLICC_00705-FCKDLICC_00706
LR961994.1	711543	A	G	snp	missense_variant	MODERATE	FCKDLICC_00714
LR961994.1	721813	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_00722-FCKDLICC_00723
LR961994.1	735381	C	T	snp	missense_variant	MODERATE	rplY
LR961994.1	736535	G	C	snp	missense_variant	MODERATE	rsuA_2
LR961994.1	772469	C	T	snp	missense_variant	MODERATE	kimA
LR961994.1	779726	G	A	snp	missense_variant	MODERATE	queA
LR961994.1	799604	C	T	snp	missense_variant	MODERATE	FCKDLICC_00790
LR961994.1	803977	C	G	snp	missense_variant	MODERATE	arnC_1
LR961994.1	810026	C	A	snp	missense_variant	MODERATE	artM
LR961994.1	839110	T	G	snp	stop_lost	HIGH	FCKDLICC_00828
LR961994.1	852797	G	GT	ins	intergenic_region	MODIFIER	niaX-msmX_1
LR961994.1	854595	C	T	snp	intergenic_region	MODIFIER	mgsA-folT
LR961994.1	864272	A	G	snp	missense_variant	MODERATE	ung
LR961994.1	871876	T	C	snp	missense_variant	MODERATE	malP
LR961994.1	873833	C	T	snp	synonymous_variant	LOW	ptsG_1
LR961994.1	913901	T	C	snp	synonymous_variant	LOW	ileS
LR961994.1	924932	A	G	snp	synonymous_variant	LOW	dgaR_2
LR961994.1	928397	A	G	snp	missense_variant	MODERATE	FCKDLICC_00920
LR961994.1	930970	C	T	snp	synonymous_variant	LOW	gmuD
LR961994.1	970389	G	A	snp	intergenic_region	MODIFIER	sasA_1-FCKDLICC_00972
LR961994.1	986601	G	A	snp	synonymous_variant	LOW	dppE_5
LR961994.1	1001928	G	GA	ins	frameshift_variant	HIGH	dinB_1
LR961994.1	1007069	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_01005-FCKDLICC_01006
LR961994.1	1008555	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_01006-FCKDLICC_01007
LR961994.1	1033348	T	C	snp	missense_variant	MODERATE	addB
LR961994.1	1043806	T	C	snp	missense_variant	MODERATE	pheT
LR961994.1	1044990	C	A	snp	missense_variant	MODERATE	glnP
LR961994.1	1047589	T	C	snp	missense_variant	MODERATE	glnQ_4

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	1048386	T	C	snp	missense_variant	MODERATE	murI
LR961994.1	1049004	T	C	snp	missense_variant	MODERATE	rph
LR961994.1	1049251	G	C	snp	missense_variant	MODERATE	rph
LR961994.1	1068258	C	T	snp	synonymous_variant	LOW	gpsB
LR961994.1	1072485	T	C	snp	missense_variant	MODERATE	dnaD
LR961994.1	1080468	G	A	snp	synonymous_variant	LOW	dinG_1
LR961994.1	1082789	T	C	snp	synonymous_variant	LOW	FCKDLICC_01070
LR961994.1	1089212	T	C	snp	intergenic_region	MODIFIER	rpmE2-tagB
LR961994.1	1089355	A	T	snp	stop_gained	HIGH	tagB
LR961994.1	1096125	G	A	snp	missense_variant	MODERATE	luxS
LR961994.1	1098007	C	A	snp	missense_variant	MODERATE	dapA
LR961994.1	1103085	T	A	snp	synonymous_variant	LOW	yutF
LR961994.1	1115702	A	T	snp	intergenic_region	MODIFIER	FCKDLICC_01105-FCKDLICC_01106
LR961994.1	1117027	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_01108-FCKDLICC_01109
LR961994.1	1124912	T	A	snp	missense_variant	MODERATE	tagU_2
LR961994.1	1128322	C	CA	ins	intergenic_region	MODIFIER	aldC-FCKDLICC_01118
LR961994.1	1134378	C	T	snp	missense_variant	MODERATE	spuD
LR961994.1	1149632	AT	A	del	frameshift_variant	HIGH	FCKDLICC_01138
LR961994.1	1155266	A	G	snp	missense_variant	MODERATE	FCKDLICC_01142
LR961994.1	1155782	G	C	snp	missense_variant	MODERATE	FCKDLICC_01142
LR961994.1	1158048	C	T	snp	missense_variant	MODERATE	exuR
LR961994.1	1170838	G	A	snp	missense_variant	MODERATE	FCKDLICC_01157
LR961994.1	1172033	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_01158-FCKDLICC_01159
LR961994.1	1180334	T	A	snp	missense_variant	MODERATE	rlmI
LR961994.1	1186428	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_01169-rimP
LR961994.1	1193201	G	A	snp	intergenic_region	MODIFIER	FCKDLICC_01178-FCKDLICC_01179
LR961994.1	1198061	C	T	snp	missense_variant	MODERATE	FCKDLICC_01189
LR961994.1	1200106	G	A	snp	missense_variant	MODERATE	FCKDLICC_01189
LR961994.1	1206670	ATAAAAG	CTTTTAT	complex	intergenic_region	MODIFIER	FCKDLICC_01194-truB
LR961994.1	1213176	C	T	snp	stop_gained	HIGH	cynR_1
LR961994.1	1218553	C	T	snp	missense_variant	MODERATE	hemW
LR961994.1	1245496	G	T	snp	missense_variant	MODERATE	FCKDLICC_01232
LR961994.1	1246673	A	T	snp	missense_variant	MODERATE	FCKDLICC_01232
LR961994.1	1257719	A	T	snp	missense_variant	MODERATE	trxB
LR961994.1	1275293	G	T	snp	intergenic_region	MODIFIER	mgtA-pdhA
LR961994.1	1284584	G	A	snp	missense_variant	MODERATE	dhaK
LR961994.1	1285355	C	T	snp	intergenic_region	MODIFIER	dhaL-FCKDLICC_01265
LR961994.1	1285414	GTAACAAAAAA	G	del	intergenic_region	MODIFIER	dhaL-FCKDLICC_01265
LR961994.1	1293688	C	T	snp	synonymous_variant	LOW	FCKDLICC_01271
LR961994.1	1298837	T	A	snp	missense_variant	MODERATE	nrnA
LR961994.1	1303673	T	G	snp	intergenic_region	MODIFIER	cshB-FCKDLICC_01281
LR961994.1	1310846	T	C	snp	synonymous_variant	LOW	FCKDLICC_01286
LR961994.1	1319967	C	T	snp	missense_variant	MODERATE	moeA
LR961994.1	1323707	CT	C	del	frameshift_variant	HIGH	FCKDLICC_01299
LR961994.1	1328762	TA	T	del	intergenic_region	MODIFIER	cadA-zapA
LR961994.1	1350325	T	C	snp	missense_variant	MODERATE	ntpK
LR961994.1	1377002	A	G	snp	missense_variant	MODERATE	dnaG
LR961994.1	1377462	T	G	snp	missense_variant	MODERATE	dnaG
LR961994.1	1380232	AT	A	del	frameshift_variant	HIGH	mapZ
LR961994.1	1386307	T	C	snp	synonymous_variant	LOW	FCKDLICC_01353
LR961994.1	1392879	C	T	snp	missense_variant	MODERATE	xerD
LR961994.1	1401160	C	T	snp	missense_variant	MODERATE	FCKDLICC_01369
LR961994.1	1405691	C	T	snp	intergenic_region	MODIFIER	hup-FCKDLICC_01374
LR961994.1	1419699	G	A	snp	synonymous_variant	LOW	aroA
LR961994.1	1424476	T	C	snp	synonymous_variant	LOW	FCKDLICC_01394
LR961994.1	1429163	G	T	snp	missense_variant	MODERATE	ettA
LR961994.1	1436490	T	A	snp	intergenic_region	MODIFIER	FCKDLICC_01404-FCKDLICC_01405
LR961994.1	1439372	A	AT	ins	intergenic_region	MODIFIER	perR-nox
LR961994.1	1439861	G	A	snp	missense_variant	MODERATE	nox
LR961994.1	1441920	A	G	snp	missense_variant	MODERATE	paiA_1
LR961994.1	1459207	T	A	snp	missense_variant	MODERATE	degA_2
LR961994.1	1468218	T	C	snp	missense_variant	MODERATE	clsA_2
LR961994.1	1468810	T	A	snp	missense_variant	MODERATE	clsA_2
LR961994.1	1468861	A	C	snp	missense_variant	MODERATE	clsA_2
LR961994.1	1483986	CA	C	del	intergenic_region	MODIFIER	FCKDLICC_01442-pduA
LR961994.1	1487314	T	C	snp	missense_variant	MODERATE	eutC
LR961994.1	1508926	C	T	snp	missense_variant	MODERATE	FCKDLICC_01467
LR961994.1	1517534	T	C	snp	missense_variant	MODERATE	bfmBAB
LR961994.1	1546750	C	A	snp	missense_variant	MODERATE	FCKDLICC_01508
LR961994.1	1552135	T	G	snp	missense_variant	MODERATE	FCKDLICC_01516
LR961994.1	1555690	T	A	snp	synonymous_variant	LOW	pstS1_1
LR961994.1	1586374	T	C	snp	missense_variant	MODERATE	yheI

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	1587575	G	A	snp	missense_variant	MODERATE	nfo
LR961994.1	1608514	G	T	snp	missense_variant	MODERATE	FCKDLICC_01569
LR961994.1	1619594	CT	C	del	frameshift_variant	HIGH	FCKDLICC_01579
LR961994.1	1620218	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_01579-celA_3
LR961994.1	1628987	C	T	snp	missense_variant	MODERATE	purH
LR961994.1	1635969	G	A	snp	missense_variant	MODERATE	purC
LR961994.1	1636114	C	CA	ins	intergenic_region	MODIFIER	purC-purK_1
LR961994.1	1638733	CATAA	C	del	frameshift_variant	HIGH	FCKDLICC_01602
LR961994.1	1657417	A	T	snp	synonymous_variant	LOW	agaS
LR961994.1	1659351	G	A	snp	intergenic_region	MODIFIER	nagR_3-gspA_1
LR961994.1	1677168	A	T	snp	missense_variant	MODERATE	FCKDLICC_01633
LR961994.1	1683386	C	T	snp	missense_variant	MODERATE	adhA
LR961994.1	1696622	A	T	snp	missense_variant	MODERATE	larA
LR961994.1	1701880	T	G	snp	missense_variant	MODERATE	sorC
LR961994.1	1712764	G	C	snp	missense_variant	MODERATE	xerC_5
LR961994.1	1715398	T	G	snp	synonymous_variant	LOW	FCKDLICC_01666
LR961994.1	1725450	T	C	snp	missense_variant	MODERATE	gatC_4
LR961994.1	1729110	C	T	snp	synonymous_variant	LOW	hxlB_2
LR961994.1	1730546	A	T	snp	missense_variant	MODERATE	gatC_5
LR961994.1	1731879	C	T	snp	start_lost	HIGH	gatC_5
LR961994.1	1732518	G	T	snp	intergenic_region	MODIFIER	rpiB-srlR_2
LR961994.1	1734996	G	A	snp	synonymous_variant	LOW	FCKDLICC_01690
LR961994.1	1739619	T	C	snp	intergenic_region	MODIFIER	manR_3-FCKDLICC_01695
LR961994.1	1742160	T	G	snp	synonymous_variant	LOW	FCKDLICC_01697
LR961994.1	1752075	T	C	snp	missense_variant	MODERATE	FCKDLICC_01707
LR961994.1	1757541	A	T	snp	missense_variant	MODERATE	FCKDLICC_01712
LR961994.1	1774745	C	T	snp	missense_variant	MODERATE	glpO
LR961994.1	1785913	G	A	snp	missense_variant	MODERATE	FCKDLICC_01738
LR961994.1	1799647	A	G	snp	intergenic_region	MODIFIER	dgaR_3-FCKDLICC_01753
LR961994.1	1810783	A	C	snp	missense_variant	MODERATE	pdxK_2
LR961994.1	1832300	T	C	snp	intergenic_region	MODIFIER	rlmN-yxdM
LR961994.1	1862316	G	A	snp	stop_gained	HIGH	ybiT_2
LR961994.1	1863811	C	T	snp	intergenic_region	MODIFIER	ybiT_2-scmP_2
LR961994.1	1870555	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_01814-FCKDLICC_01816
LR961994.1	1885212	G	A	snp	missense_variant	MODERATE	FCKDLICC_01828
LR961994.1	1898953	AT	A	del	frameshift_variant	HIGH	FCKDLICC_01838
LR961994.1	1916148	T	A	snp	missense_variant	MODERATE	FCKDLICC_01849
LR961994.1	1921890	TA	T	del	frameshift_variant	HIGH	FCKDLICC_01853
LR961994.1	1927650	T	C	snp	missense_variant	MODERATE	FCKDLICC_01860
LR961994.1	1947838	A	G	snp	synonymous_variant	LOW	rhaR_4
LR961994.1	1965488	G	T	snp	missense_variant	MODERATE	FCKDLICC_01892
LR961994.1	1967768	T	C	snp	missense_variant	MODERATE	FCKDLICC_01892
LR961994.1	1969749	C	A	snp	synonymous_variant	LOW	FCKDLICC_01894
LR961994.1	1972044	C	A	snp	intergenic_region	MODIFIER	ngcF-FCKDLICC_01897
LR961994.1	2001182	T	C	snp	missense_variant	MODERATE	FCKDLICC_01919
LR961994.1	2001229	T	G	snp	missense_variant	MODERATE	FCKDLICC_01919
LR961994.1	2016412	T	G	snp	missense_variant	MODERATE	aspB
LR961994.1	2019099	C	T	snp	synonymous_variant	LOW	dinG_2
LR961994.1	2021297	G	A	snp	stop_gained	HIGH	FCKDLICC_01934
LR961994.1	2032146	GT	G	del	intergenic_region	MODIFIER	rasP-gdh_2
LR961994.1	2033006	G	A	snp	missense_variant	MODERATE	gdh_2
LR961994.1	2033786	A	C	snp	missense_variant	MODERATE	FCKDLICC_01940
LR961994.1	2038805	A	G	snp	synonymous_variant	LOW	FCKDLICC_01944
LR961994.1	2048135	A	G	snp	missense_variant	MODERATE	FCKDLICC_01954
LR961994.1	2052298	T	C	snp	missense_variant	MODERATE	glyS
LR961994.1	2055955	G	A	snp	synonymous_variant	LOW	dgkA
LR961994.1	2062908	G	T	snp	missense_variant	MODERATE	thrB
LR961994.1	2067722	G	A	snp	synonymous_variant	LOW	pgcA
LR961994.1	2075872	T	A	snp	intergenic_region	MODIFIER	guaD-FCKDLICC_01982
LR961994.1	2077658	T	G	snp	missense_variant	MODERATE	ybbH_3
LR961994.1	2077761	C	A	snp	missense_variant	MODERATE	ybbH_3
LR961994.1	2080014	C	T	snp	missense_variant	MODERATE	murQ_2
LR961994.1	2085700	C	T	snp	missense_variant	MODERATE	FCKDLICC_01992
LR961994.1	2087089	T	C	snp	intergenic_region	MODIFIER	FCKDLICC_01994-panE_2
LR961994.1	2090742	A	G	snp	stop_lost	HIGH	comEC_2
LR961994.1	2095730	A	C	snp	missense_variant	MODERATE	FCKDLICC_02005
LR961994.1	2103162	A	G	snp	synonymous_variant	LOW	recQ_2
LR961994.1	2145116	A	G	snp	synonymous_variant	LOW	metN2
LR961994.1	2163543	A	G	snp	synonymous_variant	LOW	prmC
LR961994.1	2167433	C	T	snp	intergenic_region	MODIFIER	FCKDLICC_02066-ntpJ_2
LR961994.1	2174691	C	A	snp	missense_variant	MODERATE	FCKDLICC_02070
LR961994.1	2183361	A	T	snp	missense_variant	MODERATE	FCKDLICC_02079

#Chromosome	Position	Reference	ALT	Catagory	Type of variant	Impact	CDS
LR961994.1	2191006	C	T	snp	missense_variant	MODERATE	dpaL
LR961994.1	2252647	T	C	snp	missense_variant	MODERATE	FCKDLICC_02146
LR961994.1	2270738	TC	T	del	frameshift_variant	HIGH	pspA_1
LR961994.1	2273616	A	C	snp	synonymous_variant	LOW	cutC
LR961994.1	2285201	GGTTTTTAAACAT	G	del	intergenic_region	MODIFIER	spxA_2-trpS
LR961994.1	2309187	CAAAGT	ACTTTG	complex	intergenic_region	MODIFIER	FCKDLICC_02197-FCKDLICC_02198
LR961994.1	2317394	C	T	snp	missense_variant	MODERATE	ebgA
LR961994.1	2329674	G	A	snp	intergenic_region	MODIFIER	sdhA_2-dppA
LR961994.1	2334640	A	AT	ins	frameshift_variant	HIGH	FCKDLICC_02221
LR961994.1	2335288	A	G	snp	missense_variant	MODERATE	nusG
LR961994.1	2337698	T	C	snp	missense_variant	MODERATE	FCKDLICC_02227
LR961994.1	2342547	T	C	snp	intergenic_region	MODIFIER	polC_2-FCKDLICC_02232
LR961994.1	2348864	G	A	snp	synonymous_variant	LOW	FCKDLICC_02237
LR961994.1	2355342	A	G	snp	synonymous_variant	LOW	dltA
LR961994.1	2370389	A	G	snp	missense_variant	MODERATE	tenA
LR961994.1	2374821	A	G	snp	intergenic_region	MODIFIER	FCKDLICC_02264-emrB_2
LR961994.1	2381645	G	T	snp	missense_variant	MODERATE	FCKDLICC_02273
LR961994.1	2384760	T	A	snp	missense_variant	MODERATE	galE_2
LR961994.1	2389844	C	A	snp	missense_variant	MODERATE	glcK
LR961994.1	2442668	C	T	snp	synonymous_variant	LOW	aadK
LR961994.1	2443497	G	T	snp	missense_variant	MODERATE	FCKDLICC_02355
LR961994.1	2443735	A	G	snp	missense_variant	MODERATE	FCKDLICC_02356
LR961994.1	2451141	A	G	snp	synonymous_variant	LOW	yqeH
LR961994.1	2455150	A	G	snp	missense_variant	MODERATE	accC
LR961994.1	2465028	C	A	snp	missense_variant	MODERATE	FCKDLICC_02380
LR961994.1	2488198	T	A	snp	missense_variant	MODERATE	liaS
LR961994.1	2489750	A	C	snp	intergenic_region	MODIFIER	FCKDLICC_02404-greA
LR961994.1	2493239	G	GCTT	ins	conservative_inframe_insertion	MODERATE	mnaA
LR961994.1	2513889	G	A	snp	synonymous_variant	LOW	pbuG
LR961994.1	2522524	A	C	snp	missense_variant	MODERATE	ulaA
LR961994.1	2529461	T	G	snp	synonymous_variant	LOW	strB1
LR961994.1	2560282	T	C	snp	intergenic_region	MODIFIER	ybbW-paiA_2
LR961994.1	2577179	T	C	snp	missense_variant	MODERATE	FCKDLICC_02481
LR961994.1	2601249	C	T	snp	synonymous_variant	LOW	pepA
LR961994.1	2606657	T	TCATGATTGG	ins	conservative_inframe_insertion	MODERATE	nagA
LR961994.1	2608247	T	C	snp	missense_variant	MODERATE	FCKDLICC_02508
LR961994.1	2613414	G	A	snp	synonymous_variant	LOW	thiT
LR961994.1	2630013	G	T	snp	missense_variant	MODERATE	FCKDLICC_02533
LR961994.1	2677885	A	G	snp	missense_variant	MODERATE	pknD
LR961994.1	2693059	T	C	snp	missense_variant	MODERATE	manX_5
LR961994.1	2695131	C	A	snp	missense_variant	MODERATE	sorA_4
LR961994.1	2696268	A	G	snp	missense_variant	MODERATE	hcxA_2
LR961994.1	2699528	T	TA	ins	frameshift_variant	HIGH	hexR
LR961994.1	2746745	T	G	snp	synonymous_variant	LOW	FCKDLICC_02642
LR961994.1	2748083	C	CA	ins	intergenic_region	MODIFIER	FCKDLICC_02642-aes
LR961994.1	2767209	A	G	snp	synonymous_variant	LOW	dgaR_5
LR961994.1	2783177	T	A	snp	missense_variant	MODERATE	xylB
LR961994.1	2791633	A	G	snp	synonymous_variant	LOW	rpoB
LR961994.1	2800928	T	C	snp	missense_variant	MODERATE	FCKDLICC_02687
LR961994.1	2832117	A	G	snp	missense_variant	MODERATE	licC_6
LR961994.1	2840661	C	T	snp	synonymous_variant	LOW	phoP_2
LR961994.1	2878169	A	G	snp	missense_variant	MODERATE	citC
LR961994.1	2880686	A	C	snp	missense_variant	MODERATE	FCKDLICC_02760

Paper III

The population structure of vancomycin resistant and susceptible *Enterococcus faecium* in a low prevalence antimicrobial resistance setting is highly influenced by global clones

Mushtaq AL Rubaye[1], Jessin Janice[1,2#], Jørgen Vildershøj Bjørnholt[3,4], Oliver Kacelnik[5], Bjørg C. Haldorsen[2], Randi M. Nygaard[6], Joachim Hegstad[7], Arnfinn Sundsfjord[1,2], Kristin Hegstad[1,2]* and the Norwegian VRE study group.

[1] Research group for Host-Microbe Interactions, Department of Medical Biology, UiT the Arctic University of Norway, Tromsø, Norway

[2] Norwegian National Advisory Unit on Detection of Antimicrobial Resistance, Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway

[3] Department of Clinical Microbiology, Oslo University Hospital, Oslo, Norway

[4] Institute of Clinical Medicine, University of Oslo, Oslo, Norway

[5] Department of Antibiotic Resistance and Infection Prevention, Norwegian Institute of Public Health, Oslo, Norway

[6] Department of Microbiology, Haukeland University Hospital, Bergen, Norway

[7] Department of Microbiology and Infection Control, University Hospital of North Norway, Tromsø, Norway

* Corresponding author: Kristin Hegstad, Kristin.Hegstad@uit.no

[#] **Current address:** Section for development, Department of Microbiology, Clinic for Laboratory Medicine, Oslo University Hospital

Keywords: Vancomycin resistant enterococci, *E. faecium*, *E. faecalis*, VRE outbreak, *vanA* gene cluster, *vanB* gene cluster.

Repositories:

Illumina and PacBio reads, and their assemblies are available under the following project numbers PRJNA858233, PRJNA407052, PRJNA393251, and PRJNA306646. Biosample ID and metadata have been provided in Supplement file 1. All other supporting data have been provided within the article.

Abstract

The incidence of vancomycin resistant enterococci (VRE) increased dramatically between 2010 and 2015 in Norway. Thus, we examined the population structure of Norwegian vancomycin susceptible *Enterococcus faecium* (VSE*fm*) and resistant enterococci (VRE) and global *E. faecium* strains to explore the emergence of VRE. VSE*fm* bacteraemia isolates collected in 2008 and 2014 (n=261) through the Norwegian surveillance programme for antimicrobial resistance as well as a randomly selected subset of VRE recovered from 2010 to July 2015 underwent phenotypic susceptibility testing and Illumina whole genome sequencing. The genomic data were used for typing and in-depth molecular analyses of *van* gene clusters, mobile genetic elements, and virulome of *E. faecium*. All Norwegian VRE *faecium* (VRE*fm*) and most of the VSE*fm* belonged to globally prominent hospital associated sequence types (STs). The *vanB2* subtype carried by variants of the Tn*1549* integrative conjugative element was the dominant *van*-type. The major cluster types (CTs) of VRE*fm* have been reported concurrently in other European countries. The dominant *vanB*-type VRE*fm* CTs, ST192-CT3/26 and ST117-CT24, were mostly linked to a single hospital in Norway and had acquired Tn*1549* independently. Although the total number of *vanA* was lower compared to *vanB* VRE, their CTs were more diverse. The *vanA* gene clusters were carried by either Inc18 or RepA_N plasmids which harboured toxin-antitoxin systems that support their persistence. Only 5% of the VRE were *Enterococcus faecalis* which all contained *vanB*. While the Norwegian VRE*fm* and VSE*fm* isolates are overloaded with virulence factors (VFs) supporting biofilm formation and colonization, each CT has specific VF profiles. Successful VRE*fm* CTs generally harbour more virulence determinants than VSE*fm* and clade A more VFs than clade B isolates which is in line with clade A occurring much more frequently in infections. In conclusion, globally prevalent clones and particularly European CTs influence the population structure of *E. faecium* in a low prevalence antimicrobial resistance setting like Norway. Prevalent VRE*fm* CTs contain more VFs than VSE*fm*.

Impact statement

This is the first comprehensive study on the population structure of *E. faecium* in Norway in which 241 VRE and 261 VSE were sequenced. The inclusion of invasive and representative outbreak related isolates gave a clear picture of the population structure of VRE*fm* and VSE*fm* in Norway. Most enterococcal studies are biased by antibiotic resistant outbreak isolates thus both vancomycin resistant and susceptible isolates were included to reveal the dominant STs, circulating MGE harbouring *van* clusters in the Norwegian hospitals, as well as differences in the virulome profiles. This study provides new insights into the dominant STs, the circulating MGEs harbouring *van* clusters in the Norwegian hospitals, as well as the virulome profiles of VRE*fm* and VSE*fm*. To the best of our knowledge, this is the first population study on enterococci in which the virulomes of concurrent VRE*fm* and VSE*fm* were compared. The findings contribute to our understanding of the genomic evolution of clinical strains of VRE*fm* and VSE*fm*.

INTRODUCTION

Enterococcus faecium and *Enterococcus faecalis* are commensals in the human gut microbiota and may cause severe infections, especially in immunocompromised and hospitalized patients [1]. The flexibility of enterococcal genomes and the ability to acquire antimicrobial resistance (AMR) genes have contributed to their development into first-class opportunistic pathogens [2–4]. While *E. faecalis* causes most infections, the expected *E. faecium* phenotype is more resistant and more often acquires resistance to vancomycin [3]. The global phylogeny of *E. faecium* is dominated by two separate phylogenetic clades (A and B). Clade A can be divided into two sub-clades, A1 comprising mainly clinical strains, and A2 strains mainly recovered from animals but also non-hospitalized persons. Clade B encompasses community isolates [3,5,6], and genomic analyses have suggested reclassification as new enterococcal species (*Enterococcus lactis*) [7].

The treatment of enterococcal infections is challenging due to intrinsic and acquired antimicrobial resistance. Vancomycin is used in the treatment of infections with multidrug resistant (MDR) enterococci [1]. The increasing prevalence of enterococcal infections is associated with the rise of vancomycin resistance among enterococci [8]. Ten different *van* gene clusters (*vanA*, *vanB*, *vanC*, *vanD*, *vanE*, *vanG*, *vanL*, *vanM, vanN,* and *vanP*) are responsible for vancomycin resistance in enterococci [9]. The *vanC* gene cluster is intrinsic in *Enterococcus casseliflavus* and *Enterococcus gallinarum,* but may rarely be acquired by *E. faecium* and *E. faecalis* [3,10]. The other *van*-types have only been associated with acquired vancomycin resistance in enterococci [9,11]. The *van* gene clusters contain three types of genes encoding; 1) enzymes that remove the inherent D-Ala-D-Ala-ending precursors, 2) enzymes acting to synthesize new peptidoglycan precursors, and 3) a two-component signal transduction system for inducible resistance [3].

VanA-type VRE has been the most prevalent VRE worldwide [3], but in recent years *vanB* has repeatedly caused hospital associated outbreaks, particularly across Europe. The successful spread of *vanA* and *vanB* is partly due to their linkage to promiscuous mobile genetic elements (MGEs) [12]. The *vanA* gene cluster is usually part of the Tn*1546* transposon often found on plasmids [13]. In contrast, the widespread *vanB2* subtype gene cluster is associated with Tn*1549* integrative conjugative elements (ICE) acquired from gut anaerobes [14]. However, the mechanisms driving the dissemination of VRE*fm* are complex and both clonal spread and MGE exchange likely play important roles [15].

Virulence factors (VFs) have been divided into two main groups, (i) those enhancing colonization and (ii) those mediating invasion and host tissue damage [16]. Although *E. faecium* and *E. faecalis* are not considered highly virulent, both species have VFs associated with colonization and host invasion and/or tissue damage [3,17] Some of these VFs have a key role in bypassing the host immune system [18]. In *E. faecium* most of the VFs are involved in interactions with extracellular matrix proteins which is vital in biofilm formation and colonization [19].

Genomic analyses are important in understanding pathogen evolution and epidemiology [20,21]. Whole genome sequencing (WGS) has become the most widely used method in outbreak investigation and surveillance of bacterial pathogens [22] allowing core genome MLST (cgMLST) analyses for high-resolution typing, variant calling, and drawing phylogenetic trees based on SNP alignment [23–25].

Clinical infections and carriage of VRE have been notifiable in Norway since 1996. The annual number of reported VRE cases was less than ten before 2010. After 2010 the prevalence of VRE increased significantly and peaked up until 2015, before ceasing, though never returning to the situation before 2010 [26]. The majority of VRE were acquired in hospitals and 85% of them were associated with outbreaks [26]. The Norwegian monitoring system for antibiotic resistance in microbes (NORM/NORM-VET) monitors the overall use of antibiotics and the prevalence of antimicrobial resistance in human and animal pathogens including *E. faecium* and *E. faecalis* [27]. The program involves all diagnostic laboratories which collect, perform standardized antimicrobial susceptibility testing, store strains, and make them available for research.

In the current study, we aimed to investigate the molecular epidemiology of Norwegian VRE and compare them to invasive VSE isolates collected through the NORM-program, as well as global strain genomes. We describe the dominant outbreak clones, their MGEs harbouring *van* gene clusters, and the VF-profile of *E. faecium*.

MATERIALS AND METHODS

Samples size, collections description and data collection

A total of *E. faecium* (n=490) and *E. faecalis* (n=12) isolates from three different collections are included in this study: VSE 2008 from NORM 2008 [28], VSE 2014 from NORM 2014 and VRE (2010-throughout June 2015) (Table 1) [29,30]. The study period was chosen

because of a sudden increase in the VRE incidence from 2010 (0,12 in 2009 to 1,10 in 2010 and 5,87 in 2011 per 100,000 person years), which then gradually decreased to 1,5 in 2015 (Supplement Fig. 1). Inclusion of two VSE collections allowed us to compare vancomycin susceptible *E. faecium* (VSE*fm*) and vancomycin resistant *E. faecium* (VRE*fm*) genomes before and after the increase of VRE. All VSE isolates were VSE*fm* recovered from blood culture samples in the nine first months of the year [29,30]. A total of 99/110 (90%) and 162 out of 174 (93%) VSE*fm* isolates were available for inclusion from 2008 and 2014, respectively. The third collection encompassed a random selection of 239 isolates out of 783 (31%) VRE reported between 2010 and 2015, VRE*fm* (n=227) and *E. faecalis* VRE (VRE*fs*; n=12). A total of 87/783 (11%) VRE were clinical isolates. The random selection included all clinical isolates (blood, urine, wounds, other) and up to three carrier isolates (faeces) per clinical isolate if available, weighted across geography and time (n = 261) [31]. Twenty-two isolates were excluded (5 wrong ID, 14 not available for sequencing, 3 with repeated low quality of assemblies). Thus, a total of 227 VRE*fm* and 12 VRE*fs* were included in the study. The relative proportion of included VRE compared to the total numbers of VRE reported in Norway is illustrated in Supplement Fig. 1. In addition we also included two VRE*fm* isolates recovered in 1996 from the first VRE outbreak reported in Norway [32] for phylogenetic analyses. All the isolates of the study are listed in supplement file 1 with anonymized IDs, and the name of hospitals was changed to the ID of a letter (N, M, SE, and W) which refer to Northern, Middle, South-Eastern, and Western health regions of Norway respectively, and a digit. Moreover, an overview of sequence types (STs) for VRE from 2019-20 was obtained from the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance (K-res) to compare the ST distribution to more recent data.

Table 1. Isolates included in the study and occurrence of resistance to vancomycin, ampicillin, and high levels of gentamicin

Collection and year	Isolates n	Ampicillin resistant n (%)	High-level gentamicin resistant n (%)
VRE			
E. faecium 2010-2015	227	226 (99.5)	85 (37)
E. faecalis 2010-2015	12	0	8 (67)
E. faecium 1996	2	2 (100)	0
VSE			
E. faecium 2008	99	83 (84)	56 (55)
E. faecalis 2014	162	162 (100)	71 (44)

Species identification and antimicrobial susceptibility testing (AST)

A single colony of each blood agar culture sample was used for subculturing and subsequent antimicrobial susceptibility testing (AST), genomic DNA extraction for WGS, and species identification by MALDI-TOF (Bruker Daltonik GmbH, Bremen, Germany).

For the VSE isolates, AST data were collected as part of the NORM program. Susceptibility testing within NORM follows a standard protocol defined in the yearly surveillance reports (appendix 5) [30]. For the VRE collection AST were performed at K-res using the same methods as in NORM. Briefly, gentamicin, linezolid, and ampicillin susceptibility testing was performed and interpreted according to the EUCAST disc diffusion method [33], and EUCAST clinical breakpoints [34], respectively. The CLSI agar screening method was used for detection of reduced susceptibility to vancomycin[35].

Whole genome sequencing

Initially, all samples were subjected to short-read sequencing. First, DNeasy Blood and tissue kit (Qiagen, Hilden, Germany) was used to extract the genomic DNA. Next, Qubit fluorometer (Invitrogen) was used to quantify the concentration of total genomic DNA. The Genomic support center Tromsø^{TM} sequenced the samples using Illumina NextSeq550 system as described previously [21]. A selection of 21 isolates was subsequently chosen for long-read sequencing to use as reference genomes. The selection was based on their position in the phylogenetic tree. Wizard Genomic DNA Purification Kit (Promega, Madison, USA) was used to

extract a large quantity of genomic DNA for long read sequencing. Then, the genomic DNA concentration was quantified with Qubit fluorometer. Long-read sequencing was performed at the Norwegian Sequencing Centre (University of Oslo). To prepare multiplexed microbial libraries, SMRTbell Express Template prep Kit 2.0 was used according to pacific biosciences protocol. Fragmentation of DNA was carried out using g-tubes (Covaries) resulting in 10-16 kb sized fragments. To select the final library, BluePippin with an 8 kb cut-off was used. Libraries were sequenced on ~90% of 8M SMRT cell on Sequel II using Sequel II Banding kit 2.0 and sequencing chemistry v2.0. Demultiplex Barcodes pipeline was carried out using SMRT Tools (SMRT Link v9.0.0.92188) to demultiplex the reads (Minimum barcode score 26). Finally, the circular consensus sequencing (CCS) sequences were produced for demultiplexed data using CCS pipeline (SMRT Link v9.0.0.92188). The resulting PacBio reads length ranged from 10 to 20 kb.

Genomic analyses

For Illumina sequenced samples Trimmomatic v0.39 was used to perform quality trimming and adaptor removal [36] before output reads files were assessed using FastQC [37]. Next, Unicycler v0.39 was used for genome assembly [38], and finally, quality assessment of the genome assemblies was performed using Quast v5.0.2 [39]. A cut-off maximum of 400 contigs and minimum of 40x genome coverage was used for Illumina sequenced samples to consider the assemblies as eligible to be included in the analyses (with exception of three samples in coverage 30-37x). Moreover, the genome size, should not show more than ± 10% fluctuation compared to the smallest and biggest complete *E. Faecium* or *E. faecalis* genome assemblies in National Center for Biotechnology Information (NCBI)'s Refseq database.

For PacBio sequenced samples, Unicycler was used to assemble the CCS reads. The assemblies that Unicycler was unable to circularize were reassembled using Canu v2.2 [40], corrected with Pilon v1.23 [41] and circularized using circulator v1.5.5 [42]. Finally, we performed quality assessment using QUAST. The prokaryotic genome annotation pipeline (PGAP) of NCBI was used to annotate the assemblies, MGEs, and plasmids [43]. Snippy v3.1 was used for variant calling between sequences [44].

Multilocus sequence typing

Multilocus sequence typing (MLST) was carried out for all samples using MLST v2.19.0 [45]. To generate minimum spanning trees, cgMLST was performed using SeqSphere+ software V6.0.2 (Ridom GmbH, Münster, Germany [http://www.ridom.de/seqsphere/]). For *E. faecium* isolates, the scheme included 1423 core genes and a threshold of ≤ 20 allelic differences for cluster calculation and determination of clonal relatedness [23]. The scheme of 1,972 gene targets with ≤ seven allelic differences was set up for cluster calculation and clonal relatedness of *E. faecalis* genomes [46]. Novel sequence types (STs) and cluster types (CTs) were obtained by submission of assemblies for allelic profiling to PubMLST [47] and Ridom SeqSphere+, respectively.

Phylogenetic trees

Phylogenetic trees based on the core genome of the Norwegian *E. faecium*, were constructed using Parsnp v1.2 [25]. The global tree included all Norwegian *E. faecium* isolates of the study (n=490) as well as all publicly available complete genomes of *E. faecium* retrieved from NCBI as of 29.11.2021 (n=237). In addition, a local tree which included the 490 Norwegian isolates was built. A tree for VRE*fs* (n=12) isolates, was generated by core genome phylogeny using Parsnp. Finally, Interactive Tree Of Life (iTOL) was applied to display metadata in the trees [48].

MGEs harbouring the *vanB* gene cluster

To identify the *van*-type in VRE assemblies, NCBI bacterial AMR reference gene database (PRJNA313047) was used in the ABRicate tool v1.0.1 [49]. To locate and extract the sequences of MGEs harbouring *vanB* gene clusters in individual isolates, the closest PacBio closed VSE genome was used as a reference. The contigs of the Illumina assemblies were sorted according to the references using Mauve [50]. Next, sorted Illumina assemblies were concatenated and BLASTed against their reference genomes using BLASTn tool v2.6.0 [51]. Artemis Comparison Tool (ACT) [52] was used to visualize the BLASTs and locate the MGEs harbouring the *vanB* gene cluster. Finally, one representative from each MGE type was chosen to perform a BLAST and visualize the results using Easyfig v2.2.2 [53].

Plasmids harbouring the *vanA* gene cluster

Mob-suite was used to reconstruct plasmids in VanA-type VRE*fm* isolates [54]. Plasmid-typing was performed using PlasmidFinder v2.0.1 online database (https://cge.food.dtu.dk/services/PlasmidFinder/). Then plasmids were BLASTed against NCBI bacterial AMR reference gene database (PRJNA313047) in ABRicate tool v1.0 to find those containing the *vanA* gene cluster. To compare the plasmids and determine the identity between them, a closed Pacbio sequenced *vanA* plasmid of each cluster type was utilized as a reference for reads mapping. The mem algorithm in BWA tool v07.17 [55] was used to map the reads against the reference sequence. Indexing and sorting were performed in SAMtools v1.10 [56] and the resulting BAM file was visualized using Artemis v18.1.0 [52]. Samples whose reads fully covered the reference *vanA* plasmid were considered to contain plasmids similar to the reference. EasyFig v2.2.2 was used to BLAST the closed plasmids and generate a comparison figure.

Virulence factor profile

All *E. faecium* genomes were investigated for the presence of the determinants of 30 experimentally confirmed VFs (Supplement file 2) [19,57–65]. The coding sequences of all 30 VFs were used to build a database in ABRicate v1.0.1 [49]. BLASTing the *E. faecium* genomes against the database was performed using the minimum cut-off for identity and coverage at 90%. Next, the local phylogenetic tree of *E. faecium* was annotated using iTOL [48]. Since the *esp* gene contains several repeats which make trouble in the assembly process in the Illumina sequencing technology [66], only the conserved part of this gene (2190 bp) was used to BLAST against the assemblies. For *scm*, a new allele was found in our samples which is 173 bp longer than the reference allele. These extra nucleotides are in the linker region and between the two conserved domains of the gene. For *scm*, both alleles were used for BLAST searches.

RESULTS AND DISCUSSION

Norwegian VRE*fm* are dominated by prominent global STs

The VRE*fm* 2010-15 isolates (n=227) were dominated by ST192 (55%), ST117 (15%), ST203 (14%), ST80 (7%), and ST17 (3%). Non-prevalent STs (npSTs) including ST18, ST78 and ST202, added 6% (Fig. 1A). A marked shift in the relative proportions of STs was observed when compared to Norwegian VRE-data of 2019-20 [27,67] (Fig. 1B). While VRE*fm* ST192 was most dominant during 2010-12 it was not observed in 2019-20. In contrast, the prevalence of VRE*fm*

ST17 and ST80 increased. All the prevalent STs have been or still are among the dominant STs in European countries. For instance, ST192 was a globally dominant ST mostly related to *vanB* type VRE in the 2010s [68–70]. ST117 was a dominant ST in Germany over 1990s and its prevalence increased again after 2010 [68,71]. ST80 was responsible for the largest VRE outbreak recorded in Germany between 2015 to 2017 with 2900 (*vanB*-type) cases [68]. ST203, ST17, and ST18 were among the most common STs in Germany from 2000-9, but they began to fade away after a decade (2010-19) [68]. STs most prevalent in Norway at the beginning of the 2010s were gradually replaced by other STs over time showing clonal sweeps of new STs and reintroduction of some STs (Fig. 1B) which has also been observed in other countries like Germany and Denmark [68,72].

Norwegian VRE*fm* are dominated by concurrent major European clusters

Among VRE*fm* 2010-15 and 1996, *vanB*-type (n=167) were recovered from patients in nine hospitals mainly in Western and South-Eastern Norway while *vanA*-type VRE*fm* isolates (n=62) were recovered from patients in eleven hospitals all around Norway. In total, 25 *vanA*-type CTs (including 19 singletons) and 19 *vanB*-type CTs (12 of these singletons) were detected among the VRE*fm* (n=229) (Supplement Fig. 2 and Supplement file 3). Although *vanB* was dominant, the diversity of *vanA*-type CTs was slightly higher, consistent with smaller outbreaks. We identified four major clusters associated with Norwegian hospital VRE outbreaks during the study period. ST192-CT3/CT26 and ST117-CT24 associated with *vanB*-type VRE (Table 2), and ST203-CT20, and ST80-CT3097 associated with *vanA*-type (Table 3)

ST192-CT3/CT26 *vanB*-type VRE (n=113) caused the largest VRE outbreak affecting hospital W1 (109/113) and W2 (n=4) in Western Norway during 2010-13. The mixed ST192-CT3/CT26 is an artifact due to a combination of some of the alleles in the cgMLST scheme not being detected and the allelic profiles of the two chosen CT static founders being similar [73]. During the study period (2010-15), several outbreaks of *vanB*-type ST192 occurred in other countries such as Germany (2008-2009), Denmark (2012-13), and Sweden (2007-2011) [74]. In the global tree, two *vanB*-type VRE*fm* isolates from the Netherlands in 2019 (GCA_900639515.1 (E7654), and GCA_900639525.1 (E7663)) cluster with the Norwegian *vanB*-type ST192 clade (Supplement Fig. 3). These two strains have only fifteen allelic differences to the closest Norwegian ST192-CT3 (data not shown). In 2013 there was a shift towards the ST117-CT24 *vanB*-type VRE. This second largest VRE cluster was also mostly recovered from hospital W1 (28/31)

but belonged to a mixed VRE-VSE cluster encompassing 31/51 (61%) *vanB*-type VRE isolates. ST117-CT24 *vanB*-type VRE has been reported to cause outbreaks in the Netherlands from 2011-2017[15,75]. This CT was also linked to outbreaks of *vanA* in Denmark and Germany [68,76] and linezolid in Austria [77].

The ST203-CT20 *vanA* VRE*fm* cluster (n=19) was recovered from hospitals in Mid, South-Eastern and Northern Norway in 2013-15 while the ST80-CT3097 *vanA* VRE*fm* cluster (n=10) was found in three hospitals in the South-Eastern Norway in 2010-11 (Supplement Fig. 2 and Supplement file 3). The ST80-CT3097 *vanA* VRE*fm* cluster has not been reported elsewhere yet. An ST203-CT20 *vanA* VRE*fm* cluster has been reported in Ireland with only seven allelic differences from the Norwegian ST203-CT20 [78]. The *vanA*-type ST203-CT20 was also prevalent in Germany in 2015-18 among blood culture isolates and was even reported as a vancomycin variable isolate from Sweden [79]. Additionally, *vanA*-type ST203-CT20 isolates are reported from The Netherlands, Denmark, Belgium, and Australia [78].

vanB* gene clusters in VRE*fm* were carried on *de novo* acquired variants of ICE Tn*1549

MGE-analyses revealed that *vanB* was carried on variants of the dominant ICE Tn*1549* (Table 2 and Fig. 2) in all VRE*fm* from 2010-15 and 1996 [74]. In ST192-CT3/CT26 all but one isolate had an IS*L3* element integrated inside the *vanB* gene cluster in the intergenic region between the $vanS_B$ and $vanY_B$ genes (variant A in Fig. 2). The Tn*1549* in ST17, ST80, and ST203 were larger, mainly due to different IS element insertions (variants B, C, and E in Fig. 2).

Acquisitions of Tn*1549* have been shown to occur *de novo* from anaerobic gut microbiota to enterococci but may also occur through exchange of Tn*1549* between VRE and VSE [80,81]. It has been confirmed that Tn*1549* can transfer between enterococci as part of large chromosomal elements (90-250 kb), in which case the flanking region of Tn*1549* should be identical in the donor and recipient isolates [80,82]. If only Tn*1549* transfers between or into enterococci, this should be associated with the transfer of a short coupling sequence from the donor into the recipient genome (5 - 6 bp) on either the left or right flank of Tn*1549* [83]. Since the identical prototypic Tn*1549* was found in both one isolate of ST192-CT3/CT26 and ST117-CT24 that took over as the dominant clone in the same hospital, we investigated whether the ICE could have been transferred directly between the two *E. faecium* clusters. Tn*1549* was integrated into different genomic locations with different flanking sequences in ST192-CT3/CT26 compared to ST117-CT24. This

difference reflects that they both have different coupling sequences, as well as are not transferred as part of a larger chromosomal element and suggests two independent ICE Tn*1549* acquisitions in ST192-CT3/CT26 and ST117-CT24. A high prevalence of Tn*1549* has been demonstrated in the non-enterococcal gut flora of patients at hospital W1, which also supports this interpretation [84].

The integration site of Tn*1549* in ST192-CT3/CT26, was identified in an AT rich sequence in the *sir* gene of the *tirE* operon (Table 2) [58]. Tn*1549* insertion in the exact same position in *sir* was also reported in ST192 as well as other STs from Germany [80] indicating this is a preferred sequence site. E7654 and E7663, the two Dutch isolates closest to the Norwegian ST192-CT3/CT26 in the global tree, have IS*L3* insertion in the *vanB* cluster and Tn*1549* inserted at the identical AT rich sequence inside the *sir* gene. The high sequence identity, a *vanB* gene cluster with identical IS*L3* and insertion site and the same Tn*1549* insertion sequence support the idea of a common ancestor of *vanB*-type ST192-CT3 from Norway and the Netherlands.

The integration site of Tn*1549* in cluster ST117-CT24 occurred in another AT rich sequence in the overlapping 3´end of *btuD* and 5´end of *ndvA* (Table 2). Among the closed genomes retrieved from NCBI, six ST117-CT24 exist (Supplement Fig. 3); One VSE from Spain (GCF_001886635.1), one *vanB*-type VRE from Norway isolated in 2017 in hospital W1, as well as one *vanB*-type VRE (GCA_900639505.1_E7356) and three VSE (GCA_900639465.1, GCA_900639535.1, and GCA_900639565.1) from the Netherlands. The insertion site of Tn*1549* in this Dutch VRE is different and inside a hypothetical protein gene. Moreover, this Dutch *vanB*-type isolate has IS*30* family transposase (IS*1062*) integrated just upstream of the $vanR_B$ gene, which the Norwegian ST117-CT24 lack. Additionally, Tn*1549* show yet other insertion sites on the chromosome in other Dutch ST117-CT24 *vanB*-type VRE*fm* outbreak samples from 2014 and 2017 [75]. Thus, the overall genetic differences between the ST117-CT24 isolates of Norway and the Netherlands do not support relatedness between them.

Table 2. Characteristics of Norwegian VRE*fm* clusters and their *vanB* gene harbouring MGEs. Singleton VRE*fs* isolates and 15 VRE*fm* isolates with low quality assembly in the insertion site of Tn*1549* are not included in this table.

Cluster	Isolates n	MGE	MGE insertion location	Insertion sequence on reference genome (5' - 3')
E. faecium				
ST192-CT3/CT26	113	Tn*1549*	*sir* gene of *tirE* operon	AATATTAAAGGAA
ST117-CT24	31	Tn*1549*	*btuD* gene encoding vitamin B12 import ATP-binding protein	AAAAGTTTTT
ST203-CT3061	3	Tn*1549*	Between two CDSs encoding hypothetical proteins (HPs)	TTTTTATAAAAAAA
ST17-CT1709	2	Tn*1549*	Between CDSs encoding ribonucleoside-diphosphate reductase 2 subunit beta and HP	TTCAAAAATTTT
ST17-CT6207	1	Tn*1549*	IS*3* family transposase gene	TTTTTTCTTAAAA
ST80-CT16	1	Tn*1549*	Between tRNA-Gly and CDS encoding HP	ATTTTACT
E. faecalis				
ST6-CT107	4	Plasmid	CDS encoding HP	GATGATGT
ST6-CT1160	3	Tn*1549*	Between peptidase propeptide and oligopeptide-binding protein (*oppA)* genes	TTTTGACA
ST28-CT1162	2	Tn*1549*	CDS encoding catechol-2,3-dioxygenase	TTTTAT

vanA gene clusters and toxin-antitoxin systems were carried by different plasmids in unrelated CTs

In ST203-CT20 VRE*fm* an Inc18 plasmid of 55 kb with multiple IS-integrations carried the *vanA* gene cluster. Mapping *vanA*-type VRE isolate reads from this CT against the PacBio sequenced ST203-CT20 isolates showed that 17 out of 19 *vanA* plasmids have 100% coverage to our reference Inc18 plasmid. Characteristics of the different *vanA* containing plasmids are summarized in Table 3 and Fig. 3. The *vanA* gene cluster in this Inc18 plasmid was not part of Tn*1546* while other *vanA*-type clusters like ST80-CT3097, ST192-CT188, and ST202-CT3079 were associated with Tn*1546*. In the second largest cluster ST80-CT3097 *vanA* was carried by a RepA_N (rep17) plasmid of 32 kb. Other clusters showed variants of Inc18 and RepA-N with different sizes (Fig. 3 and Table 3).

Both plasmid types found carrying *vanA* in the Norwegian VRE*fm* are typically associated with *vanA* and may confer increased fitness costs. The maintenance of such plasmids has been linked to loss of phenotypic resistance, partial plasmid deletions, decreased copy number and toxin-antitoxin systems [74,79,85]. The partial homology and different sizes of the RepA_N *vanA* containing plasmids in our study (Fig. 3) suggest rearrangements have occurred. Moreover, all the Norwegian VRE*fm* *vanA* plasmids encoded at least one putative toxin-antitoxin system (Table 3). The *vanA* RepA_N plasmids all contained Axe-Txe typically found in *rep17* (pRUM) plasmids while the two *vanA* Inc18 plasmids contained Epsilon-Zeta typically associated with different Inc18 replicons [86,87].

Table 3. Characteristics of *vanA* gene clusters and plasmids in the PacBio sequenced Norwegian VRE*fm*

CT (Reference isolate)	Isolates n	Plasmid size	CDSs n	Plasmid type	Toxin-antitoxin systems	Transposon in plasmid
ST203-CT20 (51271218)	19	55 kb	73	Inc18	Epsilon-Zeta	Tn*552*
ST80-CT3097 (51271936)	10	32 kb	42	RepA_N (rep17)	Axe-Txe	Tn*1546*
ST192-CT188 (51271057)	4	62 kb	72	Inc18	Epsilon-Zeta	Tn*1546*
ST18-CT3042 (51276509)	2	43 kb	51	RepA_N (rep17)	Axe-Txe	
ST17-CT3037 (51271928)	2	38 kb	47	RepA_N (rep17)	Axe-Txe & Epsilon-Zeta	
ST202-CT3079 (51271933)	1	35 kb	43	RepA_N (rep17)	Axe-Txe	Tn*1546*

Successful VRE*fm* CTs generally have a high number of virulence determinants genes

Fig. 4 illustrates the distribution of 26 out of 30 virulence determinants genes in the

Norwegian *E. faecium*. Local BLAST of the isolates against our VF database showed that all isolates were negative for *boNT*/En and *epx2* genes while positive for *fnm* and *lysM4*. *E. faecium* contain a variety of cell surface components, many of them involved in colonization and biofilm formation [19]. Apart from *boNT*/En and *epx2* which are exotoxins [61,63], most of the mentioned genes encode products involved in biofilm formation and colonization. *Esp* encodes a surface protein involved in biofilm formation [88]. *Acm, capD, ecbA, fms15, fmn, lysM, pilA2, prpA*, *ptsD* and *scm* encode VFs supporting adherence and colonization [57,89] (supplement file 2).

In *vanB*-type ST192 isolates, the insertion of the MGE harbouring the *vanB* gene cluster occurred in the second gene of the *tirE* operon, associated with increased blood survival [58], which may affect the functional expression of this operon. Some of the other VFs are also encoded by a gene cluster. Thus, lacking one gene can affect the overall function. For instance, *empA*, *empB*, and *empC* are encoded by one operon coding for the pilus subunits. Deletion in *empA* and *empB* causes a reduction in biofilm formation while *empC* seems to be dispensable [90]. In the Norwegian VRE*fm*, if an isolate was positive for the *empABC* operon, the entire operon was present with all three genes (Supplement file 4). For STs containing a mix of VRE and VSE isolates, some VSE lacked *empA* or *empB*. In clade B, 5/21 (24%) of the isolates lacked the entire *empABC* operon (Fig. 4 and Supplement file 4).

A total of 484/490 (99%) *E. faecium* isolates, including all VRE*fm*, contained *bepA,* which is associated with biofilm formation and endocarditis [91]. Globally more than 80% of all clinical *E. faecium* isolates have been shown to carry this VF determinant [91]. Other genes encoding VFs involved in biofilm formation are also highly prevalent in the Norwegian VRE*fm; sagA* 99%, $atlA_{Efm}$ 99%, *sgrA* 91%, and *esp* 80%. Thus, several VFs associated with biofilm formation are highly prevalent in the Norwegian VRE*fm,* and the successful VRE*fm* CTs generally have a high but slightly variable number of virulence determinants (Table 4 and Fig. 4).

Table 4. VF genes and their distributions (%) in dominant Norwegian *E. faecium* CTs

Virulence factor gene	**Percent containing VF within major CTs**					**Percent with VF in all E*fm* (n=490)**
	ST192-CT3/26 (n=113)	ST117-CT24 (n=51)	ST203-CT20 (n=19)	ST80-CT16 (n=23)	ST80-CT3097 (n=10)	
	VRE	VRE/VSE	VRE	VRE/VSE	VRE	
atlA$_{Efm}$	99	98	100	100	100	99
bepA	100	100	100	100	100	99
ccpA	100	98	100	100	100	99
empA	100	100	100	100	100	98
empB	100	98	100	100	100	98
empC	100	98	100	100	100	98
sgrA	100	100	100	100	100	91
fnm	100	100	100	100	100	100
ptsD	99	100	100	100	100	93
sagA	100	100	100	100	90	99
gls20	100	100	94	91	100	96
gls33	100	100	94	91	100	95
glsB	100	100	94	91	100	95
glsB1	100	100	94	91	100	96
lysM1	71	68	89	69	90	70
lysM2	100	100	100	100	100	99
lysM3	84	86	42	43	50	45
lysM4	100	100	100	100	100	100
acm	97	100	100	10	100	94
ecbA	0	100	100	0	100	41
fms15	31	45	15	34	70	34
pilA2	99	13	89	100	50	68
scm	61	72	94	65	50	51
esp	99	100	68	0	10	80
capD	0	80	94	0	0	48
prpA	100	0	100	0	0	65
tirE1	94	0	0	0	0	46
tirE2	83	0	0	0	0	42
bonT/En	0	0	0	0	0	0
epx2	0	0	0	0	0	0

VRE*fs* incidence is much lower than VRE*fm*

Only 5% of the Norwegian VRE 2010-15 isolates were VRE*fs*. Previous VRE studies also in Norway have shown that VRE*fm* is more prevalent than VRE*fs* [26]. The VRE*fs* isolates (n=12), all *vanB*-type, clustered in ST6 (n=10) and ST28 (n=2), and nine of them formed three clusters, ST6-CT107 (n=4) and ST6-CT1160 (n=3), ST28-CT1162 (n=2) (Supplement Fig. 4). ST6 and ST28 are among the most prevalent clinical STs of *E. faecalis* [92]. ARE and linezolid resistant enterococci (LRE) were not observed among VRE*fs* isolates but most of them were HLGR (n=8) (Table 1). The Norwegian VRE*fs* are mainly associated with Tn*1549* (8/12), while in ST6-CT107 VRE*fs* (n=4), the *vanB* gene cluster was on a pTEF1 plasmid which is integrated in a hypothetical protein in the chromosome (Table 2). The latter gene cluster has 100% identity and coverage to the typical *vanB1*-type VRE isolates of V583 (AE016830.1) [93]. A BLAST search showed that this integrated plasmid was also found in the genomes of four other VRE*fs* with 100% identity and 100% coverage to isolates from France (CP039296.1, CP039548.1, and CP039549.1) [94] and the Netherlands (LR961935.1). Although the incidence of VRE*fs* in Norway was low compared to the VRE*fm*, it showed more diversity in *vanB*-subtypes and MGEs harbouring the *vanB* gene clusters. On the other side, the much more frequent occurrence of VRE*fm* than VRE*fs* indicate that *E. faecium* more prone to acquire and maintain vancomycin resistance and also being a hub for further spread by horizontal gene transfer or clonal expansion. Transfer of *vanB* ICE Tn*1549* has been shown to occur from anaerobes to *E. faecium* [83]. ICE Tn*1549* has not been shown to transfer on its own between enterococci and is only occasionally integrated plasmids.

How *vanA* enters enterococci is less clear. *E. faecalis* probably mostly acquire vancomycin resistance from *E. faecium* [11,95] which require yet another conjugative transfer by plasmids or ICEs. In addition, successful hospital associated clones of *E. faecium* ensure further spread of VRE*fm* [5] and toxin-antitoxin systems linked to VRE*fm* plasmids support their maintenance [96]. Moreover, integration of the *vanB* containing pTEF1 plasmid in the genome of *E. faecalis* has been shown to be involved in increased fitness cost [94]. All this is adding to our understanding of why the frequency of occurrence of VRE*fs* is lower than VRE*fm*.

Both VSE*fm* and VRE*fm* are dominated by globally prevalent STs

The main STs in the VSE*fm* collections are the same as in VRE 2010-15 but in a different order of prevalence; ST203 (26%), ST17 (13%), ST117 (10%), ST192 (10%), ST80 (9%), and

ST18 (5%). npSTs in the VSE*fm* collection accounted for 27% of the isolates including ST32, ST78 and ST202 (Fig. 1A). The presence of each ST varies over time and between VRE*fm* and VSE*fm*. For instance, ST80 and ST117 were absent in VSE 2008 but appeared in VRE in 2010 and 2013, respectively, and were prevalent STs in VSE 2014 (Fig. 1). ST117 and ST203, the dominant STs in VRE*fm* in 2014, were also present in VSE 2014. In contrast, ST17, ST18, ST32, and ST202 were present in VSE 2014 but absent in VRE of the same year. Moreover, in VRE 2014, only two isolates out of 47 belonged to npSTs, while in VSE 2014, 29 out of 162 isolates have npSTs. Thus, the VSE*fm* is much more diverse in STs while the VRE*fm* is dominated by typical global STs.

The most prevalent VSE CTs are from the mixed *vanB* VRE-VSE clusters, ST117-CT24, ST203-CT3061, and ST80-CT16. While the VRE isolates of ST117-CT24 are mainly from one hospital in western Norway, the corresponding VSE isolates (n=21) were recovered from 9 different hospitals covering all four health regions in the VSE 2014 collection. Thus, this clone has been successful in spreading but only picked up the *vanB* ICE Tn*1549* in hospital W1 where we know there has been a high prevalence of Tn*1549* in the non-enterococcal gut flora of patients [84] . The pure VSE clusters, ST203-CT3056 (n=12), ST203-CT3067 (n=9), and ST203-CT3062 (n=9), are smaller than the mixed clusters and are only present in either VSE 2008 or 2014 (supplement file 3) indicating the VSE population structure is dynamic.

Norwegian VRE*fm* and successful CTs have enriched virulomes compared to the more diverse VSE*fm* population

The main VRE*fm* clusters (ST192-CT3/26, ST117-CT24, ST203-CT20, and ST80-CT3097) have enriched VF profiles compared to VSE isolates although all VSE are from blood cultures and VRE are a mix of carrier and clinical isolates. ST192-CT3/26 (n=113) is the VRE*fm* most overloaded with VFs, while in the ST80/CT16 cluster (containing only one VRE out of 23 isolates) all the isolates lack the eight VFs *capD*, *ecbA*, *esp*, *prpA*, *tirE1*, *tirE2*, *boNT*/*En*, and *epx2* and between 34 to 69% of them lack *fms15*, *lysM1*, *lysm3*, and *scm*, thereby being the least virulent cluster. The prevalence of VFs in the five main clusters is depicted in Table 4.

In general, similar VF profiles were observed within a CT irrespective of presence of a *van* gene cluster. Interestingly, the clinical VSE*fm* isolates may contain fewer VFs than VRE isolates belonging to the same CT and npST isolates have fewer VF genes compared to predominant STs

(Fig. 4). For instance, the third largest cluster ST203-CT3061 (n=25) included three *vanB*-type VRE isolates. The VSE*fm* isolates (n=22) of this cluster were from both VSE 2008 and 2014 collections. The three VRE isolates lack *fms15*, *lysM3*, *tirE1, tirE2*, *epx2* and *boNT*/*En* genes, while two of them lack one more gene (*scm)*. Three of the VSE ST203-CT3061 isolates lacked *fms15*, *gls20*, *gls33*, *glsB*, *glsB1, lysM1/3*, *tirE1, tirE2*, *epx2* and *boNT*/*En* genes, and two of them lacked one (*scm*) or two more genes (*capD* and *sgrA*). Since the virulome of mixed VRE/VSE clusters was highly variable, it was impossible to confirm the significance of the differences statistically.

Trends in antimicrobial susceptibility patterns in *E. faecium*

Regarding antimicrobial resistance, there is a trend toward increasing ampicillin resistance over the years. ARE increased from 84% (VSE 2008) to 99.5% among VRE (2010-15) and 100% among VSE 2014. HLGR on the other hand, showed a slightly decreasing trend over collections and years. 56 % and 44 % of VSE 2008 and 2014 were HLGR, but only 37% of VRE*fm* isolates are HLGR (Table 1). The only linezolid resistant isolate in this study was a VRE*fm* isolate from 2011 that harboured a G2576T mutation in the 23S rRNA gene. The co-resistance pattern of ampicillin and gentamicin in the Norwegian VRE*fm* and VSE*fm* coincide with the proportions observed in EU/EEA between 2012 and 2018 [97]. In the European study, 99% and 49% of the VRE*fm* showed resistance to ampicillin and gentamicin, respectively, while the corresponding proportions were 49% and 43% in VSE*fm*.

Clade B is less resistant and has fewer known VFs than clade A

The A and B clades are illustrated in the phylogenetic tree of local *E. faecium* (n=490) (Fig. 5). Clade A is mainly formed by globally dominant STs, and no clear separation within clade A (A1 and A2 sub-clades) was observed. Sub-clades in clade A are still a disputable topic in the *E. faecium* population structure [4] and may be affected by geographical context. For instance, in *E. faecium* isolates from Latin America, further subclading of A1 was proposed [98]. Globally clade A isolates have shown to be more prone to acquire resistance (ampicillin, vancomycin, and aminoglycosides), while clade B isolates usually are susceptible [4,6].

All the clade B (n=21) isolates (Fig. 4 and 5) in this study were npST VSE from 2008/2014 (black colour in ST ring of Fig. 4 and 5). Recently it has been suggested that clade B strains belong

to a different enterococcal species, *E. lactis* [7]. Our results confirm that there are clear differences between clade B and clade A isolates. For instance, clade B isolates were all susceptible to vancomycin (Fig. 5), aminoglycosides and linezolid, and only two isolates were ampicillin resistant (Supplement file 1). Clade B VRE isolates are rarely reported, however, clade B *vanN*-type VRE*fm* has been observed in Japan (ST669) and the US (ST240) [99].

None of the samples in clade B contain IS*L3*. However, the sample numbers are not large enough to draw any conclusions of IS*L3* being absent from clade B in general. Moreover, the lower number of VFs is especially pronounced in clade B isolates (n=21) that lack from 13 to 19 of the investigated VF genes. None of the clade B isolates were shown to harbour *ecbA*, *esp*, *fms15*, *prpA*, *ptsD*, *scm*, *tirE1*, *tirE2*, *boNT*/En, or *epx2* (Fig. 4 and Supplement file 4).

Strengths and limitations of the study

One of the main issues in the global molecular epidemiology of enterococci is the bias caused by the skewed geographical representation. Most of the examined VRE and VSE genomes are submitted from Europe followed by Japan, Australia, and the US. Thus, the epidemiology of VRE is less known in other parts of the world (Africa, the Middle East, and south Asia). Moreover, most of the studies are biased by antibiotic resistant outbreak isolates. In this study, the sample selection of VRE was done randomly across time and region, including different types of infection sources and carriers. Additionally, we included VSE isolates for genomic comparison. Thus, the current strain collections are representative of the VSE and VRE in a low prevalence AMR setting.

In the global trees and genomic comparisons, we used the closed genomes of the enterococci which barely includes 2% of the available *E. faecium* genomes in (NCBI) [100]. Excluding 98% of the genomes as well as missing data from the rest of the world can introduce a new concept of "missing link" in the identification of globally related enterococci.

In some recent studies, several putative VFs are included in the *E. faecium* virulome [98,101]. All VF genes included in this study have been confirmed as virulence determinants experimentally (references listed in Supplement file 2) which we believe is most appropriate.

CONCLUSIONS

To our knowledge this study is the first comprehensive genomic study on the population structure of enterococci in a low prevalence AMR setting. Our study shows that globally prevalent clones and particularly concurrent European CTs influence the population structure of Norwegian *E. faecium,* with similar dynamic ST sweeps. The prevalent VRE*fm* CTs have acquired more virulence determinants than the more diverse local VSE*fm* population.

Acknowledgments

We would like to express our deep gratitude to Kjersti Julin, Ahmed Mekhlif, and Theresa Wagner at UiT and Ellen H. Josefsen, Martin O. Christensen, and Berit Harbak at the Norwegian National Advisory Unit on Detection of Antimicrobial Resistance for their technical support. We would like to express our appreciation to The Norwegian VRE study group members Silje Bakken Jørgensen at Akershus universitetssykehus, Annette Onken at Bærum sykehus, Einar Tollaksen Weme at Drammen sykehus, Reidar Hjetland at Førde sentralsjukehus, Ghantous Milad Chedid at Haugesund sjukehus, Fabian Åhrberg at Molde sjukehus, Hege Elisabeth Larsen at Nordlandssykehuset, Karianne Wiger Gammelsrud at Oslo universitetssykehus Ullevål, Iren Löhr at Stavanger universitetssjukehus, Hege Enger at St. Olavs hospital, Kari Ødegaard at Sykehuset Innlandet, Åshild Marvik-Rødland at Sykehuset i Vestfold, Kyriakos Zaragkoulias at Sykehuset Levanger, Andreas Emmert at Sykehuset Østfold, Ståle Tofteland at Sørlandet sykehus, Gunnar Skov Simonsen at Universitetssykehuset Nord-Norge, and Einar Nilsen at Ålesund sjukehus for their help in collecting the samples.

We would also like to extend our thank to the Genomics Support Center Tromsø™ and The Norwegian Sequencing Centre in UIO for genome sequencing.

Legends

Fig. 1. The frequencies of STs based on collection and year. 1A. Frequencies of sequence types per sample collection. The chart illustrates the STs containing at least 1% of the total number of isolates in this study. STs with less than 1% are shown together as non-prevalent STs (npSTs). 1B. The prevalence of STs per year for VRE*fm*. Data for 2019 and 2020 was added to compare shifts of STs from the period of the study (2010-2015) to more recent data (2019-2020).

Fig. 2. Comparison of the MGEs harbouring a *vanB* gene cluster. The red and blue gradient bars represent direct and reverted sequence matches, respectively. One representative from each variant of MGE harbouring *vanB* gene cluster is shown in the Figure. Arrows symbolise the coding sequences (CDSs) and indicate the direction of transcription. Genes with different functions are shown in different colours according to the legend beside the figure. Genes of *vanB* cluster and the IS elements are marked in green and red, respectively.

Fig. 3. Comparison of *E. faecium* plasmids carrying the *vanA* gene cluster. Red shows the direct and blue the inverted sequence matches. In ST80, ST202, and ST192, the *vanA* gene cluster is carried by Tn*1546* integrated into the plasmids. Genes with different functions are shown in different colours according to the legend beside the figure. The transposase gene in ST192 is larger than the transposase gene in ST202 and ST80 by 268 amino acids.

Fig. 4. Core genome SNP tree of Norwegian VRE*fm* annotated with 26 virulence factor genes of *E. faecium*. Genes of one operon or some genes with similar functional categories are marked with same colours and red is for remaining genes. All the Norwegian isolates in this study were positive for *fnm* and LysM4 and negative for BonT/En and epx2, which are not shown in the tree. Annotations shown from the inner layer are sample collection, ST, CT, and VF genes. Clade B is highlighted with red-coloured branches.

Fig. 5. Norwegian *E. faecium* core genome SNP tree. Metadata added from the inner layer are year of isolation, sample collection, ST, *van*-type (*vanA* or *vanB*), and information about which isolates were sequenced by long read (PacBio) technology. The nine most prevalent STs are highlighted in different colours, while all non-prevalent STs are marked in black. Clade B is highlighted with red-coloured branches.

Supplement Fig 1. Norwegian VRE total numbers per year versus the number of VRE included in this study. The total numbers of resistant enterococci collected from the Norwegian Surveillance System for Communicable Diseases (MSIS) included linezolid resistant isolates which were subtracted to get the total VRE numbers.

Supplement Fig 2. Minimum spanning tree of the 490 Norwegian *E. faecium* isolates based on the cgMLST target gene scheme. Samples are double-labelled (*van*-type and cluster type) and the colour of the leaves is based on *van*-type. In each cluster, the isolates are connected by a grey area. The number of allelic differences is shown next to the black lines.

Supplement Fig 3. Global core genome SNP tree. The midpoint rooted tree includes all complete *E. faecium* assemblies from NCBI (n=272) as of 11.05.2022 in addition to the Norwegian isolates of our study (n=490). Annotations shown from the inner layer are ST, sample collection, and *van*-type. The eight most prevalent STs are highlighted in different colours, while all non-prevalent STs are marked in black. Clade B is highlighted with red-coloured branches.

Supplement Fig 4. Minimum spanning tree of the 12 Norwegian *E. faecalis vanB*-type VRE isolates based on the cgMLST target gene scheme. A grey area connects isolates belonging to the same cluster. The number of allelic differences is shown next to the black lines.

Supplement file 1. Excel file with assembly quality, metadata and repository numbers for each sample included in this study.

Supplement file 2. Table with 30 experimentally confirmed VFs in *E. faecium* including a short description of the VFs and the accession number of the sequence that was used to build our VF database.

Supplement file 3. Table showing the prevalence of the Norwegian *E. faecium* cluster types, which collection they belong to, their *van*-types and geographical region.

Supplement file 4. Excel file showing the VF gene profile of all *E. faecium* (n=490) in this study which was used to annotate the local phylogenetic tree.

References

1. Fiore E, Van Tyne D, Gilmore MS. Pathogenicity of Enterococci. Microbiol Spectr. 2019;7:4.
2. Ramos S, Silva V, Dapkevicius M de LE, Igrejas G, Poeta P. Enterococci, from harmless bacteria to a pathogen. Microorganisms. 2020;8:1118.
3. García-Solache M, Rice LB. The Enterococcus: a model of adaptability to its environment. Clin Microbiol Rev. 2019;32:1–28.
4. van Hal SJ, Willems RJL, Gouliouris T, Ballard SA, Coque TM, Hammerum AM, et al. The interplay between community and hospital Enterococcus faecium clones within health-care settings: a genomic analysis. The Lancet Microbe. 2022;3:E133–41.
5. Arredondo-Alonso S, Top J, McNally A, Puranen S, Pesonen M, Pensar J, et al. Plasmids Shaped the Recent Emergence of the Major Nosocomial Pathogen Enterococcus faecium. MBio. 2020;11:e03284-19.
6. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, et al. Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio. 2013;4:e00534-13.
7. Belloso Daza MV, Cortimiglia C, Bassi D, Cocconcelli PS. Genome-based studies indicate that the Enterococcus faecium Clade B strains belong to Enterococcus lactis species and lack of the hospital infection associated markers. Int J Syst Evol Microbiol. 2021;71:004948.
8. Sparo M, Delpech G, García Allende N. Impact on Public Health of the Spread of High-Level Resistance to Gentamicin and Vancomycin in Enterococci. Front Microbiol. 2018;9:3073.
9. Xavier BB, Coppens J, De Koster S, Rajakani SG, Van Goethem S, Mzougui S, et al. Novel vancomycin resistance gene cluster in Enterococcus faecium ST1486, Belgium, June 2021. Eurosurveillance. 2021;26:2100767.
10. Moura TM de, Cassenego APV, Campos FS, Ribeiro AML, Franco AC, D'Azevedo PA, et al. Detection of vanC1 gene transcription in vancomycin-susceptible Enterococcus faecalis. Mem Inst Oswaldo Cruz. 2013;108:453–6.
11. Sivertsen A, Pedersen T, Larssen KW, Bergh K, Rønning TG, Radtke A, et al. A Silenced vanA Gene Cluster on a Transferable Plasmid Caused an Outbreak of Vancomycin-Variable Enterococci. Antimicrob Agents Chemother. 2022;60:4119–27.
12. Pinholt M, Gumpert H, Bayliss S, Nielsen JB, Vorobieva V, Pedersen M, et al. Genomic analysis of 495 vancomycin-resistant Enterococcus faecium reveals broad dissemination of a vanA plasmid in more than 19 clones from Copenhagen, Denmark. J Antimicrob Chemother. 2017;72:40–7.
13. Arredondo-Alonso S, Top J, Corander J, Willems RJL, Schürch AC. Mode and dynamics of vanA-type vancomycin resistance dissemination in Dutch hospitals. Genome Med. Genome Medicine;

2021;13:1–18.
14. Hegstad K, Mikalsen T, Coque TM, Werner G, Sundsfjord A. Mobile genetic elements and their contribution to the emergence of antimicrobial resistant Enterococcus faecalis and Enterococcus faecium. Clin Microbiol Infect. 2010;16:541–54.
15. Zhou X, Chlebowicz MA, Bathoorn E, Rosema S, Couto N, Lokate M, et al. Elucidating vancomycin-resistant Enterococcus faecium outbreaks: the role of clonal spread and movement of mobile genetic elements. J Antimicrob Chemother. 2018;73:3259–67.
16. Chajęcka-Wierzchowska W, Zadernowska A, Łaniewska-Trokenheim Ł. Virulence factors of Enterococcus spp. presented in food. LWT. 2017;75:670–6.
17. Weiss RA. Virulence and pathogenesis. Trends Microbiol. 2002;10:314–7.
18. Mundy L, Sahm D, Gilmore M. Relationships between Enterococcal Virulence and Antimicrobial Resistance. Clin Microbiol Rev. 2000;13:513–22.
19. Gao W, Howden BP, Stinear TP. Evolution of virulence in Enterococcus faecium, a hospital-adapted opportunistic pathogen. Curr Opin Microbiol. 2018;41:76–82.
20. Homan W, Tribe D, Poznanski S, Li M, G. Hogg, Spalburg E, et al. Multilocus sequence typing scheme for Enterococcus faecium. J Clin Microbiol. 2002;40:1963–71.
21. AL Rubaye MTS, Janice J, Bjørnholt JV, Jakovljev A, Hultström ME, Sundsfjord A, et al. Novel genomic islands and a new vanDsubtype in the first sporadic VanD-type vancomycin resistant enterococci in Norway. PLoS One. 2021;16:1–15.
22. Deneke C, Uelze L, Brendebach H, Tausch SH, Malorny B. Decentralized Investigation of Bacterial Outbreaks Based on Hashed cgMLST. Front Microbiol. 2021;12:649517.
23. Been M De, Pinholt M, Top J, Bletz S, Mellmann A, Schaik W Van, et al. Core genome multilocus sequence typing scheme for high-resolution typing of Enterococcus faecium. J Clin Microbiol. 2015;53:3788–97.
24. Simar SR, Hanson BM, Arias CA. Techniques in bacterial strain typing: past, present, and future. Curr Opin Infect Dis. 2021;34:339–45.
25. Treangen TJ, Ondov BD, Koren S, Phillippy AM. The Harvest suite for rapid core-genome alignment and visualization of thousands of intraspecific microbial genomes. Genome Biol. 2014;15:524.
26. Elstrøm P, Astrup E, Hegstad K, Samuelsen Ø, Enger H, Kacelnik O. The fight to keep resistance at bay, epidemiology of carbapenemase producing organisms (CPOs), vancomycin resistant enterococci (VRE) and methicillin resistant Staphylococcus aureus (MRSA) in Norway, 2006 - 2017. PLoS One. 2019;14:e0211741.
27. NORM/NORM-VET 2020. Usage and Occurence of Antimicrobial Resistance in Norway. ISSN:1502-2307 (print) / 1890-9965 (electronic). Tromsø / Oslo; 2020.

28. Rosvoll TCS, Lindstad BL, Lunde TM, Hegstad K, Aasnaes B, Hammerum AM, et al. Increased high-level gentamicin resistance in invasive Enterococcus faecium is associated with aac(6')Ie-aph(2″)Ia-encoding transferable megaplasmids hosted by major hospital-adapted lineages. FEMS Immunol Med Microbiol. 2012;66 2:166–76.
29. NORM/NORM-VET 2008. Usage of Antimicrobial Agents and Occurrence of Antimicrobial Resistance in Norway. ISSN: 1502-2307 (print) / 1890-9965 (electronic). Tromsø / Oslo; 2008.
30. NORM/NORM-VET 2014. Usage of Antimicrobial Agents and Occurrence of Antimicrobial Resistance in Norway. ISSN: 1502-2307 (print) / 1890-9965 (electronic). ISSN:1502-2307 (electronic). Tromsø / Oslo; 2015.
31. Grimes DA, Schulz KF. Compared to what? Finding controls for case-control studies. Lancet. 2005;365:1429–33.
32. Dahl KH, Røkenes TP, Lundblad EW, Sundsfjord A. Nonconjugative transposition of the vanB-containing Tn5382-like element in Enterococcus faecium. Antimicrob Agents Chemother. 2003;47:786–9.
33. The European Committee on Antimicrobial Susceptibility Testing - EUCAST [Internet]. Antimicrob. susceptibility Test. EUCAST Disk Diffus. method, version 7.0. 2019. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Disk_test_documents/2019_manuals/Manual_v_7.0_EUCAST_Disk_Test_2019.pdf
34. European Committee on Antimicrobial Susceptibility Testing. [Internet]. Break. tables Interpret. MICs Zo. diameters, version 9.0, 2019. 2019. Available from: https://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_9.0_Breakpoint_Tables.pdf
35. Swenson JM, Clark NC, Ferraro MJ, Sahm DF, Doern G, Pfaller MA, et al. Development of a standardized screening method for detection of vancomycin-resistant enterococci. J Clin Microbiol. 1994;32:1700–4.
36. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.
37. Andrews S. FastQC: A Quality Control Tool for High Throughput Sequence Data [Internet]. Babraham Bioinforma. 2010. Available from: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/
38. Wick RR, Judd LM, Gorrie CL, Holt KE. Unicycler : Resolving bacterial genome assemblies from short and long sequencing reads. PLOS Comput Biol. 2017;13:e1005595.
39. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. Bioinformatics. 2013;29:1072–5.
40. Koren S, Walenz BP, Berlin K, Miller JR, Bergman NH, Phillippy AM. Canu : scalable and accurate

long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res. 2017;27:722–36.
41. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon : An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. PLoS One. 2014;9:e112963.
42. Hunt M, Silva N De, Otto TD, Parkhill J, Keane JA, Harris SR. Circlator : automated circularization of genome assemblies using long sequencing reads. Genome Biol. 2015;16:1–10.
43. Tatusova T, Dicuccio M, Badretdin A, Chetvernin V, Nawrocki P, Zaslavsky L, et al. NCBI prokaryotic genome annotation pipeline. Nucleic Acids Res. 2016;44:6614–24.
44. Seemann T. Snippy, Rapid haploid variant calling and core genome alignment [Internet]. 2015. Available from: https://github.com/tseemann/snippy
45. Jolley KA, Maiden MCJ. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics. 2010;11:595.
46. Neumann B, Prior K, Bender JK, Harmsen D, Klare I, Fuchs S, et al. A Core Genome Multilocus Sequence Typing Scheme for Enterococcus faecalis. J Clin Microbiol. 2019;57:e01686-18.
47. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome open Res. 2018;3:124.
48. Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49:W293–6.
49. Seemann T. Abricate [Internet]. Github. Available from: https://github.com/tseemann/abricate
50. Darling ACE, Mau B, Blattner FR, Perna NT. Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res. 2004;14:1394–403.
51. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.
52. Carver TJ, Rutherford KM, Berriman M, Rajandream M, Barrell BG, Parkhill J. ACT: the Artemis comparison tool. Bioinformatics. 2005;21:3422–3.
53. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. Bioinformatics. 2011;27:1009–10.
54. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. Microb Genom. 2018;4:e000206.
55. Li H. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv: Genomics. 2013;00:1–3.
56. Danecek P, Bonfield JK, Liddle J, Marshall J, Ohan V, Pollard MO, et al. Twelve years of SAMtools and BCFtools. Gigascience. 2021;10:giab008.
57. Wagner T, Joshi B, Janice J, Askarian F, Škalko-Basnet N, Hagestad OC, et al. Enterococcus faecium

produces membrane vesicles containing virulence factors and antimicrobial resistance related proteins. J Proteomics. 2018;187:28–38.
58. Wagner TM, Janice J, Paganelli FL, Willems RJ, Askarian F, Pedersen T, et al. Enterococcus faecium TIR-Domain Genes Are Part of a Gene Cluster Which Promotes Bacterial Survival in Blood. Int J Microbiol. 2018;2018:1435820.
59. Somarajan SR, La Rosa SL, Singh K V., Roh JH, Höök M, Murray BE. The Fibronectin-Binding Protein Fnm Contributes to Adherence to Extracellular Matrix Components and Virulence of Enterococcus faecium. Infect Immun. 2015;83:4653–61.
60. Choudhury T, Singh K V, Sillanpää J, Nallapareddy SR, Murray BE. Importance of two Enterococcus faecium loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. J Infect Dis. 2011;203:1147–54.
61. Xiong X, Tian S, Yang P, Lebreton F, Bao H, Sheng K, et al. Emerging enterococcus pore-forming toxins with MHC/HLA-I as receptors. Cell. 2022;185:1157-1171.e22.
62. Paganelli FL, Willems RJL, Jansen P, Hendrickx A, Zhang X, Bonten MJM, et al. Enterococcus faecium Biofilm Formation: Identification of Major Autolysin AtlAEfm, Associated Acm Surface Localization, and AtlAEfm-Independent Extracellular DNA Release. MBio. 2013;4:e00154-13.
63. Zhang S, Lebreton F, Mansfield MJ, Miyashita S-I, Zhang J, Schwartzman JA, et al. Identification of a Botulinum Neurotoxin-like Toxin in a Commensal Strain of Enterococcus faecium. Cell Host Microbe. 2018;23:169-176.e6.
64. Somarajan SR, Roh JH, Singh KV, Weinstock GM, Murray BE, Pirofski L. CcpA Is Important for Growth and Virulence of Enterococcus faecium. Infect Immun. 2014;82:3580–7.
65. Cacaci M, Giraud C, Leger L, Torelli R, Martini C, Posteraro B, et al. Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in Enterococcus faecium. Sci Rep. 2018;8:12412.
66. Palmer KL, Schaik W van, Willems RJL, Gilmore. MS. Enterococcal Genomics. In: Gilmore MS, Clewell DB, Ike Y et al., editor. Enterococci From Commensals to Lead Causes Drug Resist Infect [Internet]. Boston: Massachusetts Eye and Ear Infirmary; 2014.
67. NORM/NORM-VET 2019. Usage and Occurence of Antimicrobial Resistance in Norway. ISSN:1502-2307 (print) / 1890-9965 (electronic). Tromsø / Oslo; 2019.
68. Werner G, Neumann B, Weber RE, Kresken M, Wendt C, Bender JK, et al. Thirty years of VRE in Germany – “expect the unexpected”: The view from the National Reference Centre for Staphylococci and Enterococci. Drug Resist Updat. 2020;53:100732.
69. Zhou W, Zhou H, Sun Y, Gao S, Zhang Y, Cao X, et al. Characterization of clinical enterococci isolates, focusing on the vancomycin-resistant enterococci in a tertiary hospital in China: based on the

data from 2013 to 2018. BMC Infect Dis. 2020;20:356.
70. Sivertsen A, Billström H, Melefors Ö, Liljequist BO, Wisell KT, Ullberg M, et al. A Multicentre Hospital Outbreak in Sweden Caused by Introduction of a vanB2 Transposon into a Stably Maintained pRUM-Plasmid in an Enterococcus faecium ST192 Clone. PLoS One. 2014;9:e103274.
71. Weber A, Maechler F, Schwab F, Gastmeier P, Kola A. Increase of vancomycin-resistant Enterococcus faecium strain type ST117 CT71 at Charité - Universitätsmedizin Berlin, 2008 to 2018. Antimicrob Resist Infect Control. 2020;9:109.
72. Hammerum AM, Justesen US, Pinholt M, Roer L, Kaya H, Worning P, et al. Surveillance of vancomycin-resistant enterococci reveals shift in dominating clones and national spread of a vancomycin-variable vanA Enterococcus faecium ST1421-CT1134 clone, Denmark, 2015 to March 2019. Euro Surveill. 2019;24:1900503.
73. Https://www.ridom.de/. Core Genome MLST Complex Type.
74. Freitas AR, Tedim AP, Francia M V, Jensen LB, Novais C, Peixe L, et al. Multilevel population genetic analysis of vanA and vanB Enterococcus faecium causing nosocomial outbreaks in 27 countries (1986–2012). J Antimicrob Chemother. 2016;71:3351–66.
75. Lisotto P, Couto N, Rosema S, Lokate M, Zhou X, Bathoorn E, et al. Molecular Characterisation of Vancomycin-Resistant Enterococcus faecium Isolates Belonging to the Lineage ST117/CT24 Causing Hospital Outbreaks. Front Microbiol. 2021;12:728356.
76. Hammerum AM, Baig S, Kamel Y, Roer L, Pinholt M, Gumpert H, et al. Emergence of vanA Enterococcus faecium in Denmark, 2005–15. J Antimicrob Chemother. 2017;72:2184–90.
77. Kerschner H, Cabal A, Hartl R, Machherndl-Spandl S, Allerberger F, Ruppitsch W, et al. Hospital outbreak caused by linezolid resistant Enterococcus faecium in Upper Austria. Antimicrob Resist Infect Control. 2019;8:150.
78. Egan SA, Kavanagh NL, Shore AC, Mollerup S, Samaniego Castruita JA, O'Connell B, et al. Genomic analysis of 600 vancomycin-resistant Enterococcus faecium reveals a high prevalence of ST80 and spread of similar vanA regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland. J Antimicrob Chemother. 2022;77:320–30.
79. Wagner TM, Janice J, Sivertsen A, Sjögren I, Sundsfjord A, Hegstad K. Alternative vanHAX promoters and increased vanA-plasmid copy number resurrect silenced glycopeptide resistance in Enterococcus faecium. J Antimicrob Chemother. 2021;76:876–82.
80. Bender J, Kalmbach A, Fleige C, Klare I, Fuchs S, Werner G. Population structure and acquisition of the vanB resistance determinant in German clinical isolates of Enterococcus faecium ST192. Sci Rep. 2016;6:21847.
81. Howden BP, Holt KE, Lam MMC, Seemann T, Ballard S, Coombs GW, et al. Genomic Insights to

Control the Emergence of Vancomycin-Resistant Enterococci. MBio. 2013;4:e00412-13.
82. Quintiliani Jr. R, Courvalin P. Conjugal transfer of the vancomycin resistance determinant vanB between enterococci involves the movement of large genetic elements from chromosome to chromosome. FEMS Microbiol Lett. 1994;119:359–63.
83. Launay A, Ballard SA, Johnson PD, Grayson ML LT. Transfer of Vancomycin Resistance Transposon Tn1549 from Clostridium symbiosum to Enterococcus spp. in the Gut of Gnotobiotic Mice. Antimicrob Agents Chemother. 2006;50:1054–62.
84. Nygaard RM, Hegstad K, Kommedal Ø, Lindemann PC. High prevalence of vanB in non-enterococcal flora including a novel species contributes to persistent outbreaks of vancomycin resistant enterococci in a Norwegian hospital. 12th Int Meet Microb Epidemiol Markers. Dubrovnik, Croatia; 2019.
85. Tedim AP, Lanza VF, Rodríguez CM, Freitas AR, Novais C, Peixe L, et al. Fitness cost of vancomycin-resistant Enterococcus faecium plasmids associated with hospital infection outbreaks. J Antimicrob Chemother. 2021;76:2757–64.
86. Grady R, Hayes F. Axe–Txe, a broad-spectrum proteic toxin–antitoxin system specified by a multidrug-resistant, clinical isolate of Enterococcus faecium. Mol Microbiol. 2003;47:1419–32.
87. Rosvoll TCS, Pedersen T, Sletvold H, Johnsen PJ, Sollid JE, Simonsen GS, et al. PCR-based plasmid typing in Enterococcus faecium strains reveals widely distributed pRE25-, pRUM-, pIP501- and pHTβ-related replicons associated with glycopeptide resistance and stabilizing toxin–antitoxin systems. FEMS Immunol Med Microbiol. 2010;58:254–68.
88. Heikens E, Bonten MJ WR. Enterococcal Surface Protein Esp Is Important for Biofilm Formation of Enterococcus faecium E1162. J Bacteriol. 2007;189:8233–40.
89. Revtovich A V, Tjahjono E, Singh K V, Hanson BM, Murray BE, Kirienko N V. Development and Characterization of High-Throughput Caenorhabditis elegans – Enterococcus faecium Infection Model. Front Cell Infect Microbiol. 2021;11:667327.
90. Montealegre MC, Singh KV, Somarajan SR, Yadav P, Chang C, Spencer R, Sillanpää J, Ton-That H MB. Role of the Emp Pilus Subunits of Enterococcus faecium in Biofilm Formation, Adherence to Host Extracellular Matrix Components, and Experimental Infection. Infect Immun. 2022;84:1491–500.
91. Paganelli FL, Huebner J, Singh K V, Zhang X, van Schaik W, Wobser D, et al. Genome-wide Screening Identifies Phosphotransferase System Permease BepA to Be Involved in Enterococcus faecium Endocarditis and Biofilm Formation. J Infect Dis. 2016;214:189–95.
92. Pöntinen AK, Top J, Arredondo-Alonso S, Tonkin-Hill G, Freitas AR, Novais C, et al. Apparent nosocomial adaptation of Enterococcus faecalis predates the modern hospital era. Nat Commun. 2021;12:1523.

93. McGregor KF, Young H-K. Identification and Characterization ofvanB2 Glycopeptide Resistance Elements in Enterococci Isolated in Scotland. Antimicrob Agents Chemother. 2000;44:2341–8.
94. Furlan S, Matos RC, Kennedy SP, Doublet B, Serror P R-GL. Fitness Restoration of a Genetically Tractable Enterococcus faecalis V583 Derivative To Study Decoration-Related Phenotypes of the Enterococcal Polysaccharide Antigen. mSphere. 2022;4:e00310-19.
95. Hashimoto Y, Kurushima J, Nomura T, Tanimoto K, Tamai K, Yanagisawa H, et al. Dissemination and genetic analysis of the stealthy vanB gene clusters of Enterococcus faecium clinical isolates in Japan. BMC Microbiol. 2018;18:213.
96. Moritz EM, Hergenrother PJ. Toxin–antitoxin systems are ubiquitous and plasmid-encoded in vancomycin-resistant enterococci. Proc Natl Acad Sci. 2007;104:311–6.
97. Ayobami O, Willrich N, Reuss A, Eckmanns T, Markwart R. The ongoing challenge of vancomycin-resistant Enterococcus faecium and Enterococcus faecalis in Europe: an epidemiological analysis of bloodstream infections. Emerg Microbes Infect. 2020;9:1180–93.
98. Rios R, Reyes J, Carvajal LP, Rincon S, Panesso D, Echeverri AM, et al. Genomic Epidemiology of Vancomycin-Resistant Enterococcus faecium (VREfm) in Latin America: Revisiting The Global VRE Population Structure. Sci Rep. 2020;10:5636.
99. Lebreton F, Valentino MD, Schaufler K, Earl AM, Cattoir V, Gilmore MS. Transferable vancomycin resistance in clade B commensal-type Enterococcus faecium. J Antimicrob Chemother. 2018;73:1479–86.
100. NCBI. Enterococcus faecium assemblies [Internet]. Available from: https://www.ncbi.nlm.nih.gov/assembly/?term=enterococcus faecium
101. van Hal SJ, Willems RJL, Gouliouris T, Ballard SA, Coque TM, Hammerum AM, et al. The global dissemination of hospital clones of Enterococcus faecium. Genome Med. 2021;13:52.

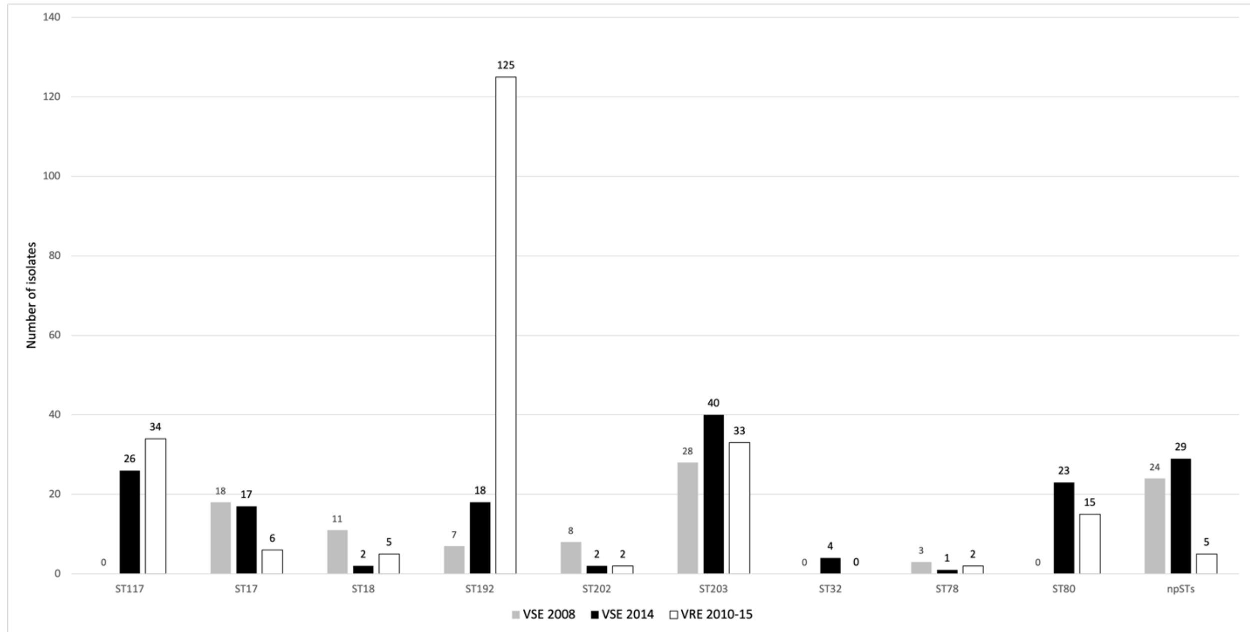


Fig. 1 (A).

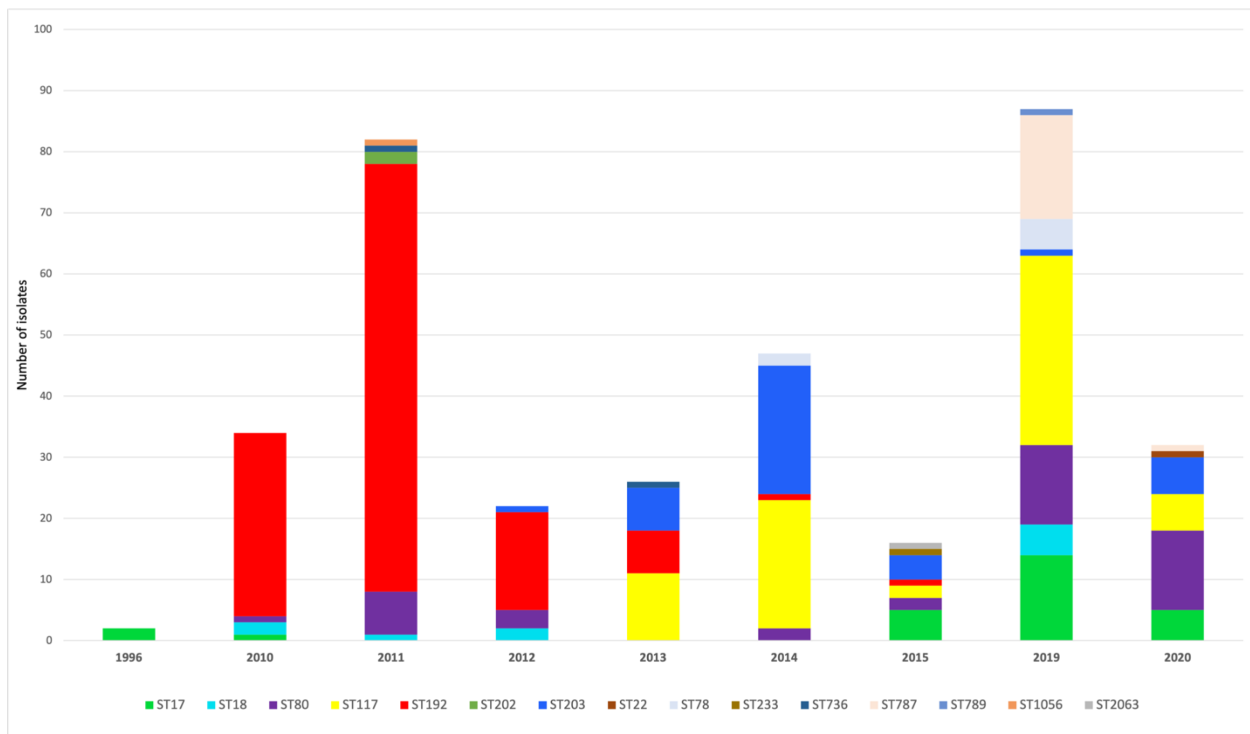


Fig. 1 (B).

Fig. 1 (A and B). The frequencies of STs based on collection and year.

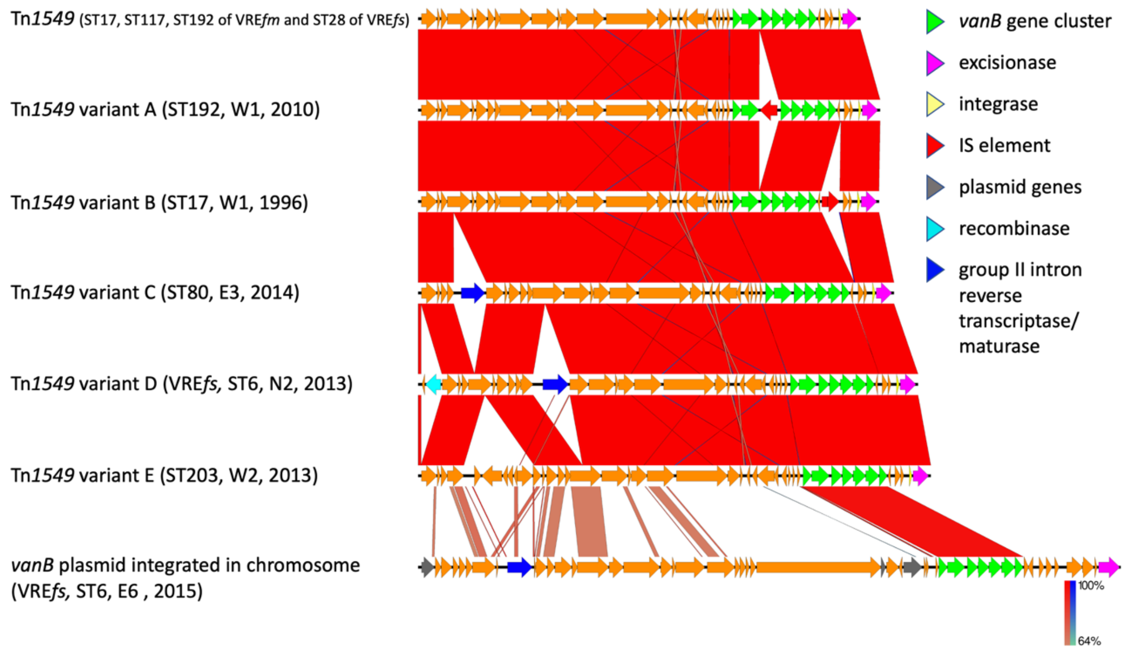


Fig. 2. Comparison of the MGEs harbouring a *vanB* gene cluster.

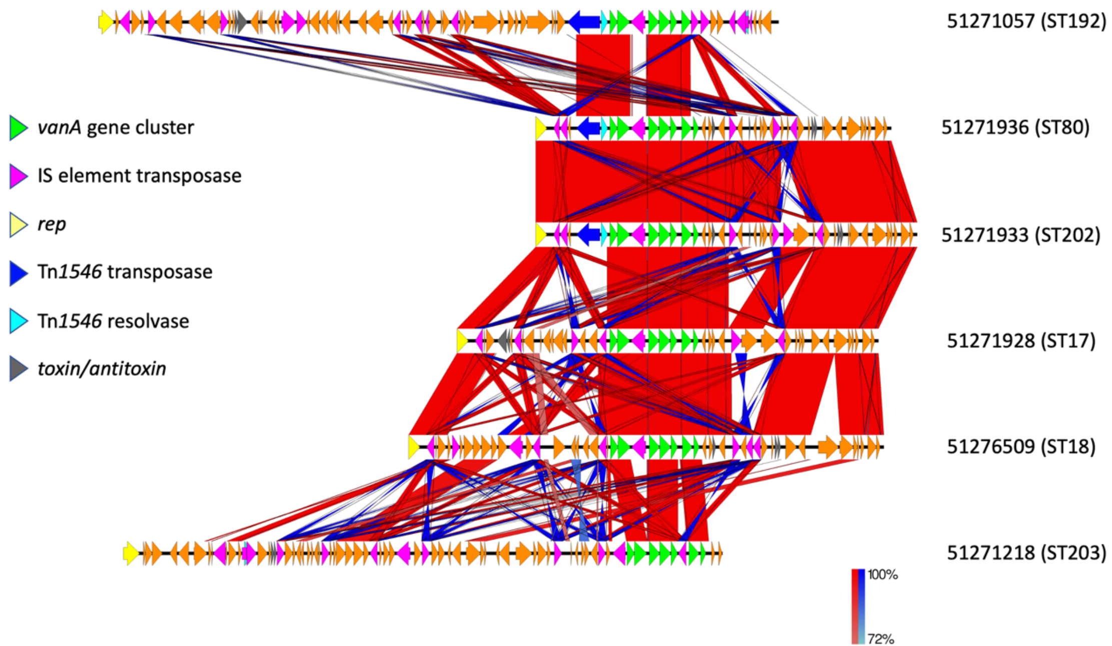


Fig. 3. Comparison of *E. faecium* plasmids carrying the *vanA* gene cluster.

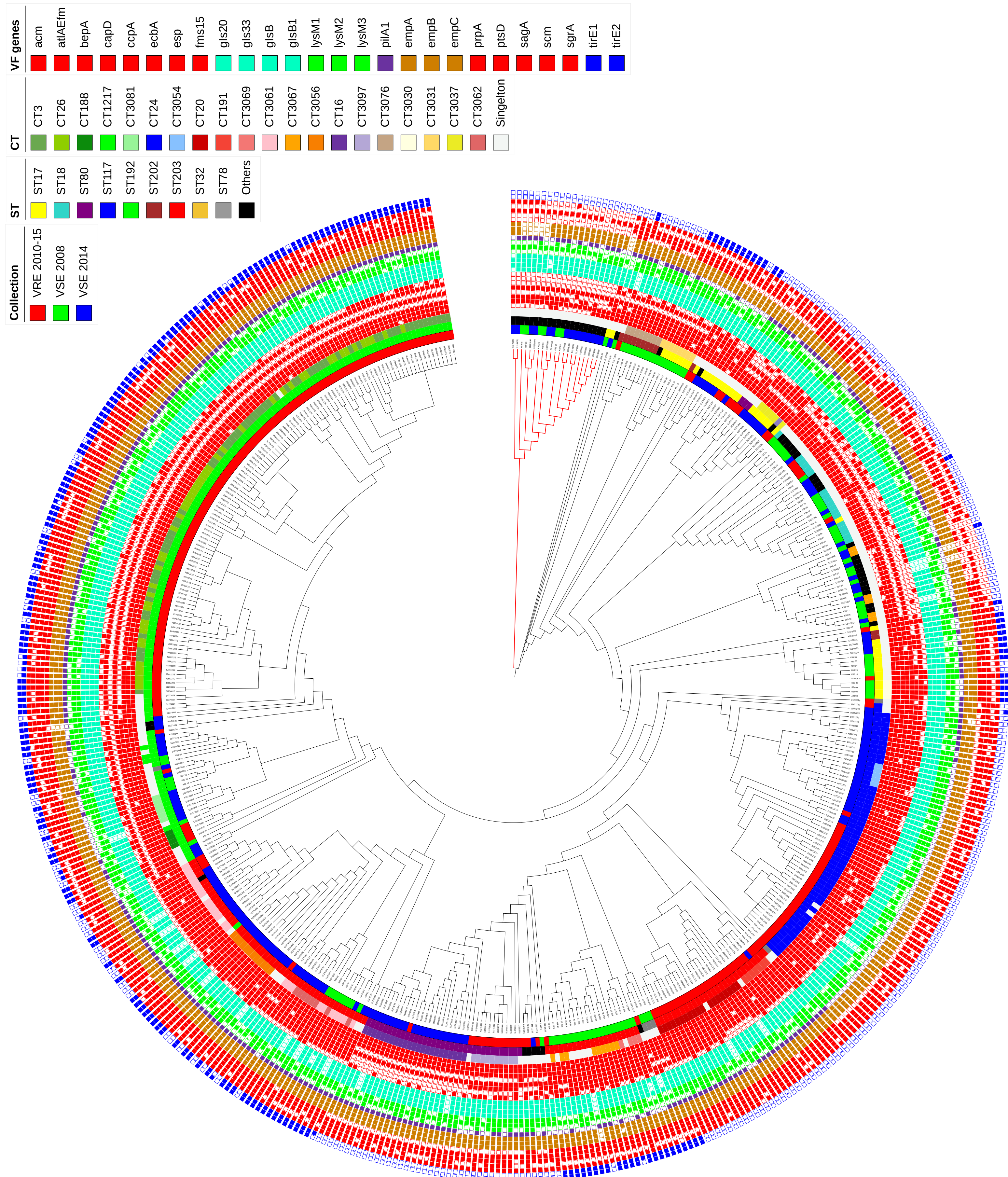


Fig. 4. Core genome SNP tree of Norwegian VRE*fm* annotated with 26 virulence factor genes of *E. faecium.*

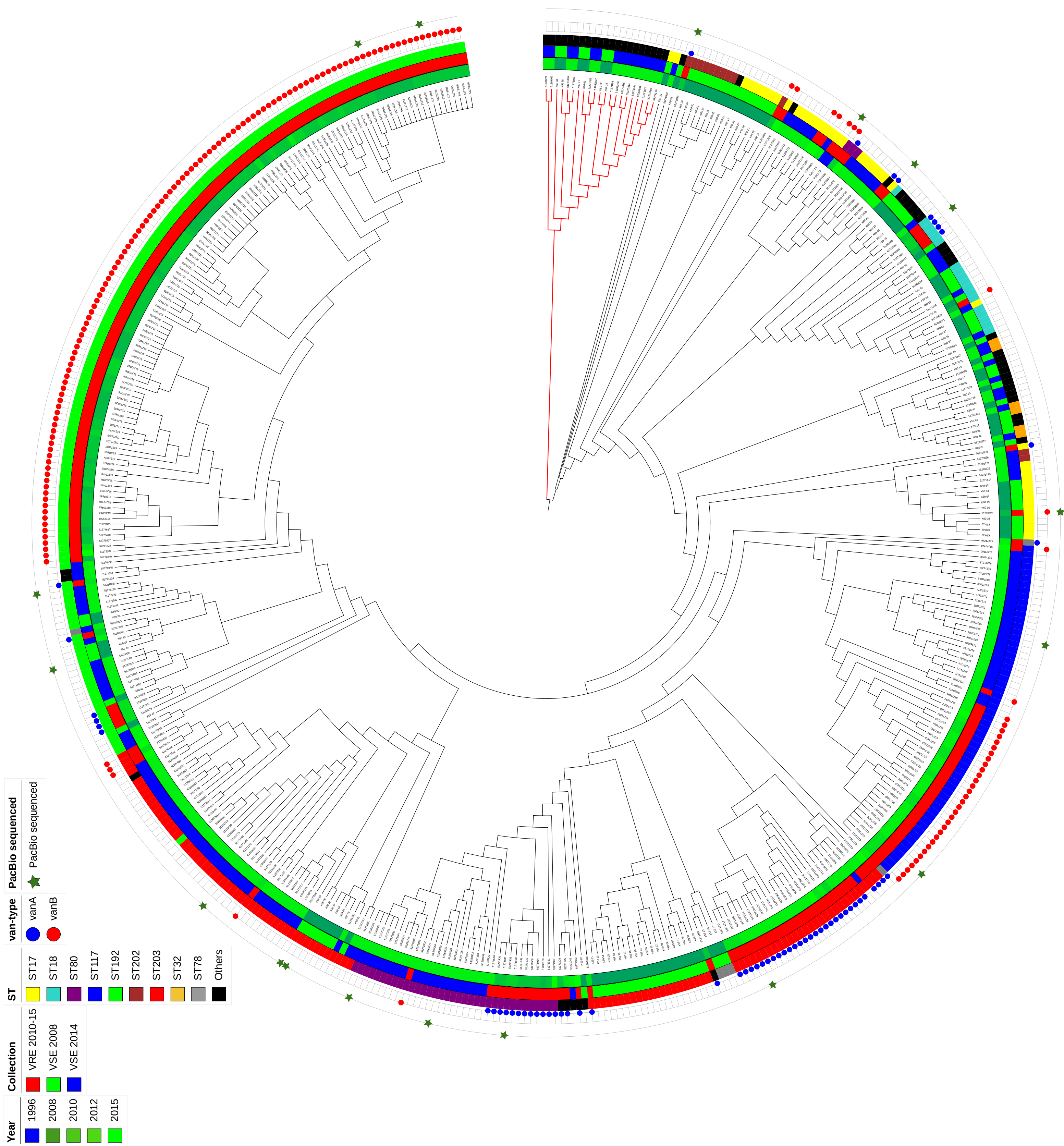
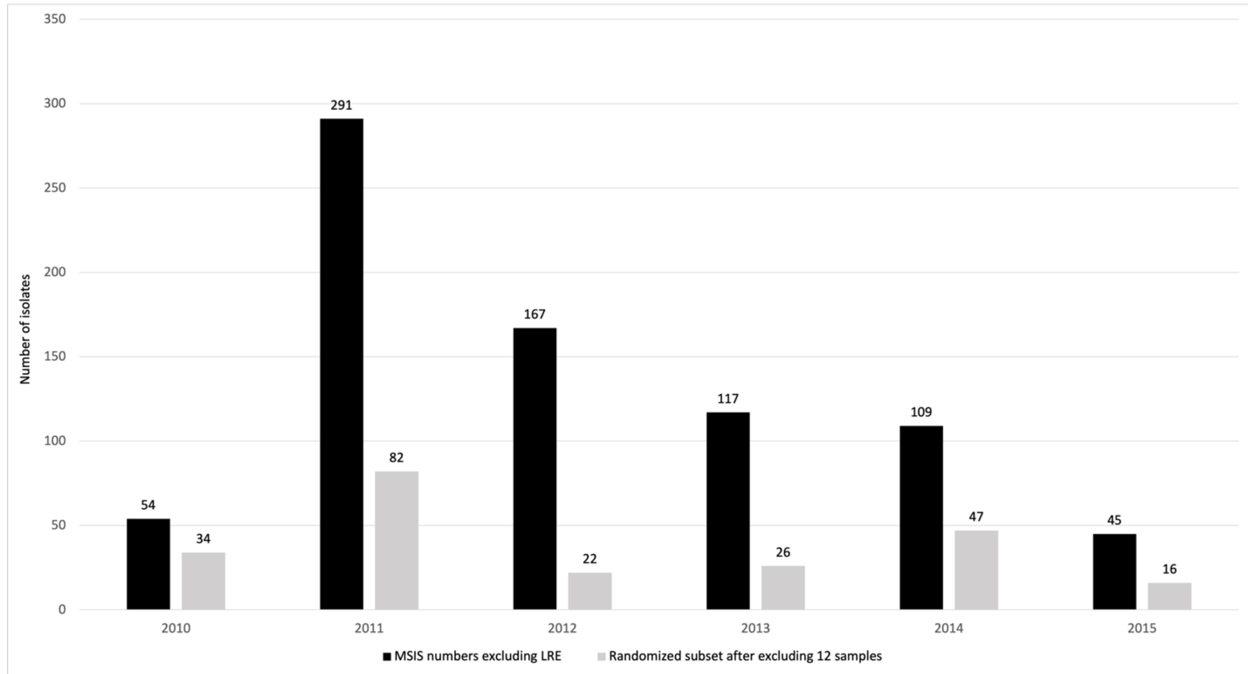
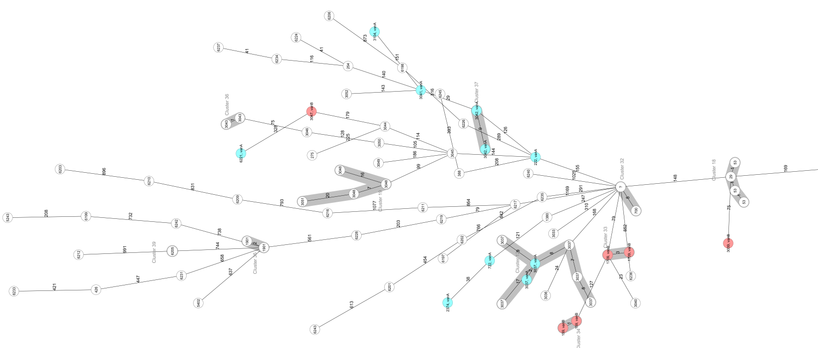


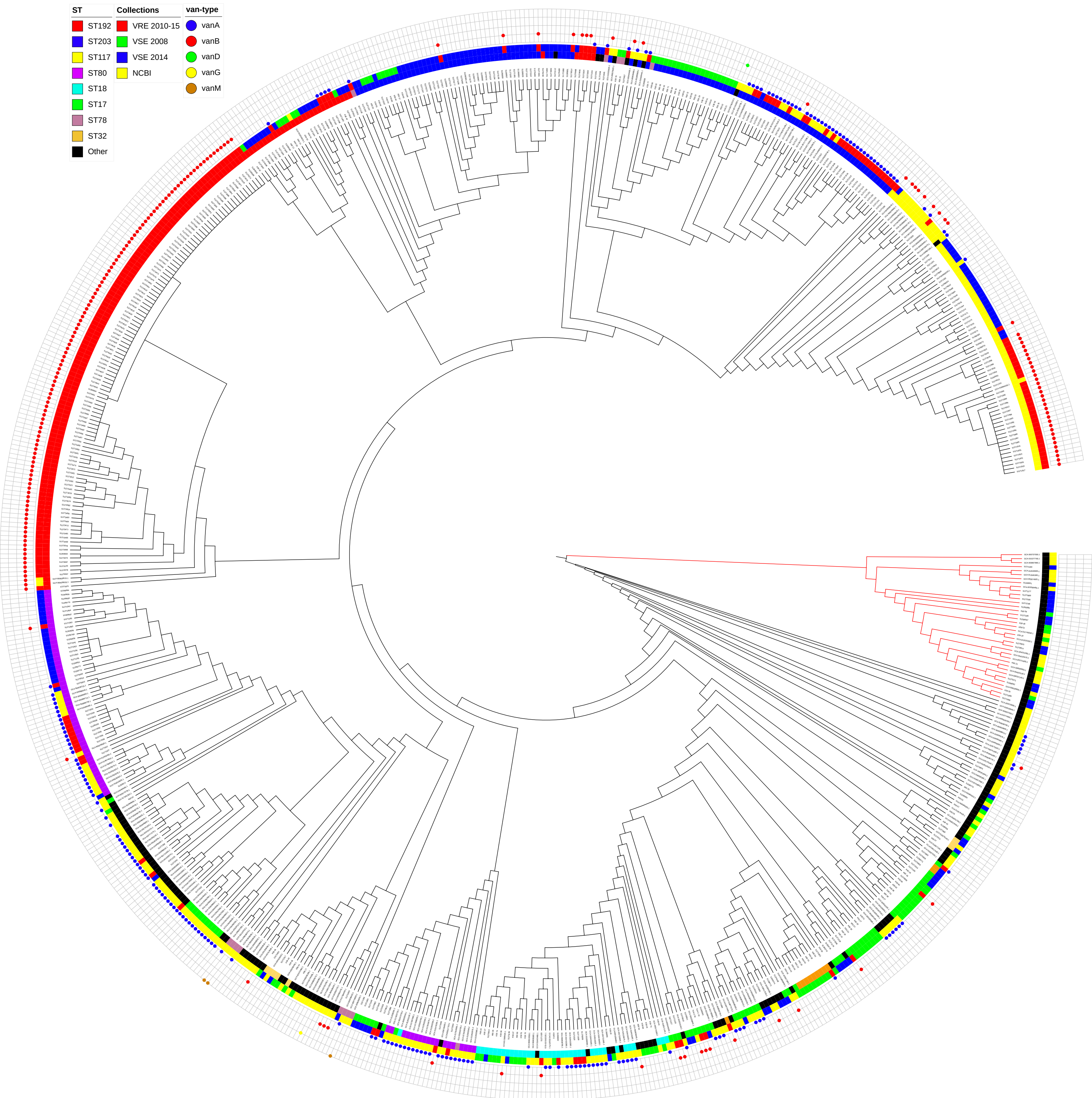
Fig. 5. Norwegian *E. faecium* core genome SNP tree.



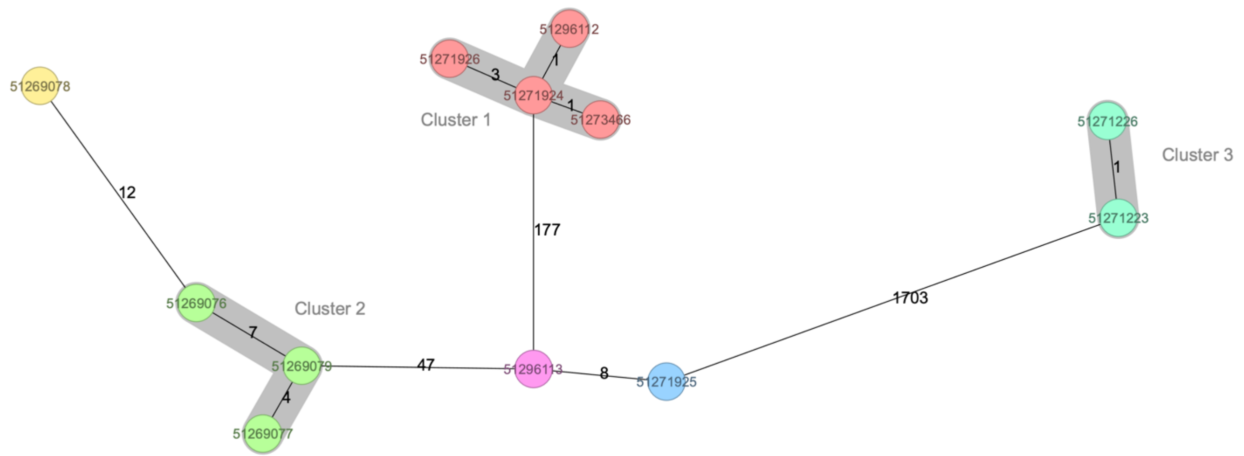
Supplement Fig 1. Norwegian VRE total numbers per year versus the number of VRE included in this study.



Supplement Fig 2. Minimum spanning tree of the 490 Norwegian *E. faecium* isolates based on the cgMLST target gene scheme.



Supplement Fig 3. Global core genome SNP tree.



Supplement Fig 4. Minimum spanning tree of the 12 Norwegian *E. faecalis vanB*-type VRE isolates based on the cgMLST target gene scheme.

Supplement file 1. Assembly quality, metadata as well as repository numbers for each *E. faecium* sample

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51268383	180	2986191	260.185	*E. faecium*	203	3056		Yes	No	No	2014	W3	Blood	VSE 2014	JANDWB000000000	SAMN29678716	PRJNA858233
51268385	204	2948703	267.375	*E. faecium*	117	24		Yes	Yes	No	2014	W3	Blood	VSE 2014	JANDWA000000000	SAMN29678717	PRJNA858233
51268386	52	2449687	323.163	*E. faecium*	1033	6197		No	No	No	2014	W3	Blood	VSE 2014	JANDVZ000000000	SAMN29678718	PRJNA858233
51269029	263	2816456	182.215	*E. faecium*	203	3061		Yes	Yes	No	2014	N2	Blood	VSE 2014	JANDVY000000000	SAMN29678719	PRJNA858233
51269051	176	2762419	252.677	*E. faecium*	80	16		Yes	No	No	2014	N2	Blood	VSE 2014	JANDVX000000000	SAMN29678720	PRJNA858233
51269053	197	2913264	294.992	*E. faecium*	203	3061		Yes	Yes	No	2014	N2	Blood	VSE 2014	JANDVW000000000	SAMN29678721	PRJNA858233
51269054	179	2921773	252.259	*E. faecium*	117	3054		Yes	Yes	No	2014	N2	Blood	VSE 2014	JANDVV000000000	SAMN29678722	PRJNA858233
51269055	82	2696792	247.855	*E. faecium*	1262	6198		No	No	No	2014	N2	Blood	VSE 2014	JANDVU000000000	SAMN29678723	PRJNA858233
51269056	186	2702017	183.336	*E. faecium*	80	16		Yes	No	No	2014	N2	Blood	VSE 2014	JANDVT000000000	SAMN29678724	PRJNA858233
51269057	85	2696920	280.072	*E. faecium*	94	6199		No	No	No	2014	N2	Blood	VSE 2014	JANDVS000000000	SAMN29678725	PRJNA858233
51269058	33	2525923	215.737	*E. faecium*	214	6200		No	No	No	2014	N2	Blood	VSE 2014	JANDVR000000000	SAMN29678726	PRJNA858233
51269059	188	3008516	62.7322	*E. faecium*	203	3056		Yes	No	No	2014	W1	Blood	VSE 2014	JANEVY000000000	SAMN29681574	PRJNA858233
51269060	188	3069601	193.986	*E. faecium*	203	3056		Yes	No	No	2014	W1	Blood	VSE 2014	JANDVQ000000000	SAMN29678727	PRJNA858233
51269061	195	2930062	242.545	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Blood	VRE 2010-15	JANDVP000000000	SAMN29678728	PRJNA858233
51269062	83	2590933	247.809	*E. faecium*	361	1901		No	No	No	2014	W1	Blood	VSE 2014	JANDVO000000000	SAMN29678729	PRJNA858233
51269063	208	2760983	130.096	*E. faecium*	17	3040		Yes	No	No	2014	W1	Blood	VSE 2014	JANDVN000000000	SAMN29678730	PRJNA858233
51269064	167	2720911	277.962	*E. faecium*	80	16		Yes	No	No	2014	W1	Blood	VSE 2014	JANDVM000000000	SAMN29678731	PRJNA858233
51269065	122	2330471	118.809	*E. faecium*	29	6201		No	No	No	2014	W1	Blood	VSE 2014	JANDVL000000000	SAMN29678732	PRJNA858233
51269066	226	3057345	251.722	*E. faecium*	203	3061		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVK000000000	SAMN29678733	PRJNA858233
51269067	181	2989587	258.753	*E. faecium*	203	3056		Yes	No	No	2014	W1	Blood	VSE 2014	JANDVJ000000000	SAMN29678734	PRJNA858233
51269068	202	2920775	35.5477	*E. faecium*	192	3083		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANEVX000000000	SAMN29681575	PRJNA858233
51269069	177	2839179	179.459	*E. faecium*	192	3086		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVI000000000	SAMN29678735	PRJNA858233
51269070	198	2926694	241.24	*E. faecium*	117	24		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVH000000000	SAMN29678736	PRJNA858233
51269071	194	2846804	220.996	*E. faecium*	17	275		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVG000000000	SAMN29678737	PRJNA858233
51269072	142	2710648	253.577	*E. faecium*	80	3096		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVF000000000	SAMN29678738	PRJNA858233
51269073	178	2880838	252.687	*E. faecium*	117	24		Yes	Yes	No	2014	W1	Blood	VSE 2014	JANDVE000000000	SAMN29678739	PRJNA858233
51269075	173	2881112	196.84	*E. faecium*	192	188	*vanA*	Yes	No	No	2013	N2	Feces	VRE 2010-15	JANDVD000000000	SAMN29678740	PRJNA858233
51269769	180	2820439	33.6819	*E. faecium*	80	16		Yes	No	No	2014	E8	Blood	VSE 2014	JANEVS000000000	SAMN29681580	PRJNA858233
51269770	170	2762377	209.347	*E. faecium*	80	16		Yes	Yes	No	2014	E8	Blood	VSE 2014	JANDVC000000000	SAMN29678741	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51269771	212	2917979	253.876	*E. faecium*	202	3077		Yes	Yes	No	2014	E8	Blood	VSE 2014	JANDVB000000000	SAMN29678742	PRJNA858233
51269772	243	3052659	186.71	*E. faecium*	262	1016		Yes	Yes	No	2014	E8	Blood	VSE 2014	JANDVA000000000	SAMN29678743	PRJNA858233
51269773	194	2858341	231.652	*E. faecium*	17	53		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUZ000000000	SAMN29678744	PRJNA858233
51269774	250	3056802	199.531	*E. faecium*	262	1016		Yes	Yes	No	2014	E8	Blood	VSE 2014	JANDUY000000000	SAMN29678745	PRJNA858233
51269775	232	2913586	113.663	*E. faecium*	203	3056		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUX000000000	SAMN29678746	PRJNA858233
51269776	91	2339300	176.272	*E. faecium*	822	6202		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUW000000000	SAMN29678747	PRJNA858233
51269777	193	2744134	142.238	*E. faecium*	80	16		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUV000000000	SAMN29678748	PRJNA858233
51269778	226	2826734	96.8037	*E. faecium*	17	53		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUU000000000	SAMN29678749	PRJNA858233
51269779	268	2728708	90.2721	*E. faecium*	80	16		Yes	No	No	2014	E8	Blood	VSE 2014	JANDUT000000000	SAMN29678750	PRJNA858233
51269928	243	2633991	78.4424	*E. faecium*	80	16		Yes	No	No	2014	W2	Blood	VSE 2014	JANDUS000000000	SAMN29678751	PRJNA858233
51269929	294	2626435	81.8218	*E. faecium*	80	16		Yes	No	No	2014	W2	Blood	VSE 2014	JANDUR000000000	SAMN29678752	PRJNA858233
51269930	172	2700796	38.4251	*E. faecium*	80	16		Yes	No	No	2014	W2	Blood	VSE 2014	JANEVR000000000	SAMN29681581	PRJNA858233
51269931	143	2690775	124.05	*E. faecium*	1102	6203		No	No	No	2014	W2	Blood	VSE 2014	JANDUQ000000000	SAMN29678753	PRJNA858233
51269932	159	2589142	102.455	*E. faecium*	18	222	*vanA*	Yes	No	No	2010	W2	Urine	VRE 2010-15	JANDUP000000000	SAMN29678754	PRJNA858233
51269933	327	2871267	40.172	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W2	Feces	VRE 2010-15	JANDUO000000000	SAMN29678755	PRJNA858233
51269934	282	2891603	122.659	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W2	Feces	VRE 2010-15	JANDUN000000000	SAMN29678756	PRJNA858233
51269935	293	2810321	117.338	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W2	Urine	VRE 2010-15	JANDUM000000000	SAMN29678757	PRJNA858233
51269936	367	2745063	113.266	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W2	Feces	VRE 2010-15	JANDUL000000000	SAMN29678758	PRJNA858233
51269937	256	2894153	117.338	*E. faecium*	203	3061	*vanB*	Yes	No	No	2013	W2	Feces	VRE 2010-15	JANDUK000000000	SAMN29678759	PRJNA858233
51269938	349	2853105	114.294	*E. faecium*	203	3057	*vanA*	Yes	Yes	No	2013	W2	Feces	VRE 2010-15	JANDUJ000000000	SAMN29678760	PRJNA858233
51269939	254	2947113	120.813	*E. faecium*	203	3058	*vanB*	Yes	Yes	No	2013	W2	Urine	VRE 2010-15	JANDUI000000000	SAMN29678761	PRJNA858233
51270240	141	2783241	226.032	*E. faecium*	192	3083		Yes	No	No	2014	M4	Blood	VSE 2014	JANDUH000000000	SAMN29678762	PRJNA858233
51270243	204	2921984	236.022	*E. faecium*	192	1217		Yes	No	No	2014	M4	Blood	VSE 2014	JANDUG000000000	SAMN29678763	PRJNA858233
51270244	172	2854162	295.687	*E. faecium*	262	6204		Yes	No	No	2014	M4	Blood	VSE 2014	JANDUF000000000	SAMN29678764	PRJNA858233
51270271	77	2461696	219.843	*E. faecium*	361	1901		No	No	No	2014	M2	Blood	VSE 2014	JANDUE000000000	SAMN29678765	PRJNA858233
51270272	246	2974127	291.408	*E. faecium*	203	3062		Yes	Yes	No	2014	M2	Blood	VSE 2014	JANDUD000000000	SAMN29678766	PRJNA858233
51270273	212	2939913	286.259	*E. faecium*	203	3062		Yes	Yes	No	2014	M2	Blood	VSE 2014	JANDUC000000000	SAMN29678767	PRJNA858233
51270274	175	2986962	221.481	*E. faecium*	117	24		Yes	No	No	2014	M2	Blood	VSE 2014	JANDUB000000000	SAMN29678768	PRJNA858233
51270275	207	2923071	249.269	*E. faecium*	203	3062		Yes	Yes	No	2014	M2	Blood	VSE 2014	JANDUA000000000	SAMN29678769	PRJNA858233
51270276	179	2921714	257.005	*E. faecium*	117	24		Yes	Yes	No	2014	M2	Blood	VSE 2014	JANDTZ000000000	SAMN29678770	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51270277	184	2869979	141.644	*E. faecium*	203	3059		Yes	No	No	2014	M2	Blood	VSE 2014	JANDTY000000000	SAMN29678771	PRJNA858233
51270809	274	2938179	86.0988	*E. faecium*	117	24		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTX000000000	SAMN29678772	PRJNA858233
51270810	272	2941247	84.3264	*E. faecium*	117	24		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTW000000000	SAMN29678773	PRJNA858233
51270811	216	2924621	116.751	*E. faecium*	192	3084		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTV000000000	SAMN29678774	PRJNA858233
51270812	198	2938269	317.044	*E. faecium*	117	24		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTU000000000	SAMN29678775	PRJNA858233
51270813	215	2848764	110.378	*E. faecium*	18	3048		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTT000000000	SAMN29678776	PRJNA858233
51270814	111	2668938	82.1332	*E. faecium*	2042	6205		No	No	No	2014	E2	Blood	VSE 2014	JANDTS000000000	SAMN29678777	PRJNA858233
51270815	110	2668386	149.93	*E. faecium*	2042	6205		No	No	No	2014	E2	Blood	VSE 2014	JANDTR000000000	SAMN29678778	PRJNA858233
51270816	314	2767339	82.2062	*E. faecium*	192	3084		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTQ000000000	SAMN29678779	PRJNA858233
51270817	187	2976029	175.667	*E. faecium*	203	3060		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTP000000000	SAMN29678780	PRJNA858233
51270818	379	2465735	62.1725	*E. faecium*	17	700		Yes	No	No	2014	?	Blood	VSE 2014	JANDTO000000000	SAMN29678781	PRJNA858233
51270819	280	2788331	96.6311	*E. faecium*	192	3084		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTN000000000	SAMN29678782	PRJNA858233
51270820	287	2941522	109.652	*E. faecium*	203	3061		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTM000000000	SAMN29678783	PRJNA858233
51270821	148	2765202	127.052	*E. faecium*	17	700		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTL000000000	SAMN29678784	PRJNA858233
51270822	172	2922355	115.017	*E. faecium*	17	3037		Yes	No	No	2014	E2	Blood	VSE 2014	JANDTK000000000	SAMN29678785	PRJNA858233
51270823	162	2780017	128.342	*E. faecium*	80	16		Yes	Yes	No	2014	E2	Blood	VSE 2014	JANDTJ000000000	SAMN29678786	PRJNA858233
51270824	84	2503339	122.011	*E. faecium*	2042	6206		No	No	No	2014	W4	Blood	VSE 2014	JANDTI000000000	SAMN29678787	PRJNA858233
51270825	178	2964451	89.1333	*E. faecium*	17	3038		Yes	No	No	2014	E5	Blood	VSE 2014	JANDTH000000000	SAMN29678788	PRJNA858233
51270826	233	2917699	85.7919	*E. faecium*	202	3077		Yes	Yes	No	2014	E5	Blood	VSE 2014	JANDTG000000000	SAMN29678789	PRJNA858233
51270828	258	2909982	79.8092	*E. faecium*	17	6207	*vanB*	Yes	Yes	No	2010	E5	Urine	VRE 2010-15	JANEVQ000000000	SAMN29681582	PRJNA858233
51271041	341	2749650	70.9176	*E. faecium*	280	1380		Yes	Yes	No	2014	N1	Blood	VSE 2014	JANDTF000000000	SAMN29678790	PRJNA858233
51271042	321	2615406	93.111	*E. faecium*	80	16		Yes	No	No	2014	N1	Blood	VSE 2014	JANDTE000000000	SAMN29678791	PRJNA858233
51271044	168	2813115	101.942	*E. faecium*	192	1217		Yes	No	No	2014	N1	Blood	VSE 2014	JANDTD000000000	SAMN29678792	PRJNA858233
51271045	198	2965105	110.37	*E. faecium*	17	3037		Yes	No	No	2014	N1	Blood	VSE 2014	JANDTC000000000	SAMN29678793	PRJNA858233
51271046	169	2836690	119.895	*E. faecium*	117	3054		Yes	Yes	No	2014	N1	Blood	VSE 2014	JANDTB000000000	SAMN29678794	PRJNA858233
51271047	157	2718918	249.366	*E. faecium*	80	16		Yes	No	No	2014	N1	Blood	VSE 2014	JANDTA000000000	SAMN29678795	PRJNA858233
51271048	215	2813500	108.727	*E. faecium*	167	6208		Yes	Yes	No	2014	N1	Blood	VSE 2014	JANDSZ000000000	SAMN29678796	PRJNA858233
51271049	179	2872211	95.2149	*E. faecium*	17	3037		Yes	No	No	2014	N1	Blood	VSE 2014	JANDSY000000000	SAMN29678797	PRJNA858233
51271050	204	2947008	120.134	*E. faecium*	117	24		Yes	Yes	No	2014	N1	Blood	VSE 2014	JANDSX000000000	SAMN29678798	PRJNA858233
51271051	169	2864626	112.33	*E. faecium*	167	6208		Yes	Yes	No	2014	N1	Blood	VSE 2014	JANDSW000000000	SAMN29678799	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271053	175	2783006	87.162	*E. faecium*	80	16		Yes	No	No	2014	N1	Blood	VSE 2014	JANDSV000000000	SAMN29678800	PRJNA858233
51271054	186	2926614	113.68	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2013	N1	Feces	VRE 2010-15	JANDSU000000000	SAMN29678801	PRJNA858233
51271055	208	2838451	85.3818	*E. faecium*	192	188	*vanA*	Yes	Yes	No	2013	N1	Urine	VRE 2010-15	JANDST000000000	SAMN29678802	PRJNA858233
51271056	334	2776267	59.1588	*E. faecium*	192	188	*vanA*	Yes	No	No	2013	N1	Feces	VRE 2010-15	JANDSS000000000	SAMN29678803	PRJNA858233
51271057	172	2881129	88.2029	*E. faecium*	192	188	*vanA*	Yes	No	No	2013	N1	Feces	VRE 2010-15	JANEVP000000000	SAMN29681583	PRJNA858233
51271164	186	2925955	270.228	*E. faecium*	117	24		Yes	No	No	2014	M1	Blood	VSE 2014	JANDSR000000000	SAMN29678804	PRJNA858233
51271165	182	2910997	204.286	*E. faecium*	117	24		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSQ000000000	SAMN29678805	PRJNA858233
51271166	205	3034026	229.198	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Blood	VRE 2010-15	JANDSP000000000	SAMN29678806	PRJNA858233
51271167	209	2933645	275.728	*E. faecium*	203	3062		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSO000000000	SAMN29678807	PRJNA858233
51271168	202	2923032	254.231	*E. faecium*	203	3062		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSN000000000	SAMN29678808	PRJNA858233
51271169	63	2730416	127.661	*E. faecium*	328	6209		No	No	No	2014	M1	Blood	VSE 2014	JANDSM000000000	SAMN29678809	PRJNA858233
51271170	173	2921171	225.115	*E. faecium*	117	24		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSL000000000	SAMN29678810	PRJNA858233
51271171	229	2912320	101.132	*E. faecium*	117	24		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSK000000000	SAMN29678811	PRJNA858233
51271172	284	2812949	130.004	*E. faecium*	203	3062		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSJ000000000	SAMN29678812	PRJNA858233
51271173	169	2889241	246.562	*E. faecium*	117	24		Yes	No	No	2014	M1	Blood	VSE 2014	JANDSI000000000	SAMN29678813	PRJNA858233
51271174	288	2759858	138.932	*E. faecium*	117	24		Yes	No	No	2014	M1	Blood	VSE 2014	JANDSH000000000	SAMN29678814	PRJNA858233
51271175	265	2682551	133.282	*E. faecium*	203	3059		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSG000000000	SAMN29678815	PRJNA858233
51271176	264	2741370	126.655	*E. faecium*	17	29		Yes	No	No	2014	M1	Blood	VSE 2014	JANDSF000000000	SAMN29678816	PRJNA858233
51271177	116	2769425	122.363	*E. faecium*	2047	6210		No	No	No	2014	M1	Blood	VSE 2014	JANDSE000000000	SAMN29678817	PRJNA858233
51271178	299	2815606	57.2663	*E. faecium*	192	3083		Yes	Yes	No	2014	M1	Blood	VSE 2014	JANDSD000000000	SAMN29678818	PRJNA858233
51271179	308	2885237	85.453	*E. faecium*	203	3056		Yes	No	No	2014	E10	Blood	VSE 2014	JANDSC000000000	SAMN29678819	PRJNA858233
51271180	227	2932669	72.470	*E. faecium*	117	1543		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDSB000000000	SAMN29678820	PRJNA858233
51271181	259	2847341	117.088	*E. faecium*	203	3061		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDSA000000000	SAMN29678821	PRJNA858233
51271182	277	2869532	130.448	*E. faecium*	203	3061		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRZ000000000	SAMN29678822	PRJNA858233
51271183	272	2641583	101.768	*E. faecium*	80	16		No	No	No	2014	E10	Blood	VSE 2014	JANDRY000000000	SAMN29678823	PRJNA858233
51271184	244	2463727	116.17	*E. faecium*	773	6211		No	No	No	2014	E10	Blood	VSE 2014	JANDRX000000000	SAMN29678824	PRJNA858233
51271185	156	2576845	106.467	*E. faecium*	289	5243		No	No	No	2014	E10	Blood	VSE 2014	JANDRW000000000	SAMN29678825	PRJNA858233
51271186	335	2739030	108.585	*E. faecium*	203	3056		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRV000000000	SAMN29678826	PRJNA858233
51271187	365	2716751	103.909	*E. faecium*	203	3061		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRU000000000	SAMN29678827	PRJNA858233
51271188	266	2881374	75.7962	*E. faecium*	203	3056		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRT000000000	SAMN29678828	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271189	344	2647897	88.717	*E. faecium*	192	3081		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRS000000000	SAMN29678829	PRJNA858233
51271190	233	2675300	134.192	*E. faecium*	80	16		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRR000000000	SAMN29678830	PRJNA858233
51271191	234	2903144	108.016	*E. faecium*	117	24		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRQ000000000	SAMN29678831	PRJNA858233
51271192	199	3012915	146.171	*E. faecium*	203	3062		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRP000000000	SAMN29678832	PRJNA858233
51271193	174	2910174	220.336	*E. faecium*	17	3038		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRO000000000	SAMN29678833	PRJNA858233
51271194	184	3074255	250.188	*E. faecium*	203	3056		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRN000000000	SAMN29678834	PRJNA858233
51271195	294	2685925	59.3006	*E. faecium*	80	16		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRM000000000	SAMN29678835	PRJNA858233
51271196	188	2809886	265.572	*E. faecium*	18	3049		Yes	Yes	No	2014	E10	Blood	VSE 2014	JANDRL000000000	SAMN29678836	PRJNA858233
51271197	233	2924920	289.242	*E. faecium*	203	3063		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRK000000000	SAMN29678837	PRJNA858233
51271198	48	2756591	267.716	*E. faecium*	583	6212		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRJ000000000	SAMN29678838	PRJNA858233
51271199	181	2861979	287.317	*E. faecium*	17	3037		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRI000000000	SAMN29678839	PRJNA858233
51271200	274	2882497	58.2468	*E. faecium*	203	3063		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRH000000000	SAMN29678840	PRJNA858233
51271201	183	2930504	215.558	*E. faecium*	117	24		Yes	No	No	2014	E10	Blood	VSE 2014	JANDRG000000000	SAMN29678841	PRJNA858233
51271208	217	2955317	219.461	*E. faecium*	736	722	*vanA*	Yes	No	No	2011	M1	Urine	VRE 2010-15	JANDRF000000000	SAMN29678842	PRJNA858233
51271210	206	2995712	185.128	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	M1	Clinical si	VRE 2010-15	JANDRE000000000	SAMN29678843	PRJNA858233
51271211	198	2926732	259.976	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2013	M1	Urine	VRE 2010-15	JANDRD000000000	SAMN29678844	PRJNA858233
51271212	190	2973876	114	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2013	M1	Feces	VRE 2010-15	GCA_025073215.1	SAMN04358604	PRJNA306646
51271213	194	2943483	196.425	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2013	M1	Urine	VRE 2010-15	JANDRC000000000	SAMN29678845	PRJNA858233
51271214	186	2973390	180	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	M1	Feces	VRE 2010-15	GCA_025073195.1	SAMN04358607	PRJNA306646
51271215	187	2960087	162.254	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Feces	VRE 2010-15	JANDRB000000000	SAMN29678846	PRJNA858233
51271216	161	2831037	190.183	*E. faecium*	78	6213	*vanA*	Yes	Yes	No	2014	M1	Clinical si	VRE 2010-15	JANDRA000000000	SAMN29678847	PRJNA858233
51271217	331	2712654	100.015	*E. faecium*	203	3064	*vanA*	Yes	Yes	No	2014	M1	Feces	VRE 2010-15	JANDQZ000000000	SAMN29678848	PRJNA858233
51271218	192	2972260	91.6233	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Blood	VRE 2010-15	JANEVO000000000	SAMN29681584	PRJNA858233
51271219	196	2895526	80.0519	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Feces	VRE 2010-15	JANDQY000000000	SAMN29678849	PRJNA858233
51271220	174	2880472	138.525	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Feces	VRE 2010-15	JANDQX000000000	SAMN29678850	PRJNA858233
51271221	223	2922464	194.672	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	M1	Feces	VRE 2010-15	JANDQW000000000	SAMN29678851	PRJNA858233
51271224	172	2902720	100.126	*E. faecium*	192	4737	*vanA*	Yes	No	No	2013	E10	Feces	VRE 2010-15	JANDQV000000000	SAMN29678852	PRJNA858233
51271225	215	2995515	182.503	*E. faecium*	736	2374	*vanA*	Yes	Yes	No	2013	E10	Blood	VRE 2010-15	JANDQU000000000	SAMN29678853	PRJNA858233
51271227	209	2960318	674.995	*E. faecium*	17	159	*vanB*	Yes	Yes	No	2015	E10	Blood	VRE 2010-15	JANDQT000000000	SAMN29678854	PRJNA858233
51271228	170	2915549	245.366	*E. faecium*	17	3039	*vanB*	Yes	No	No	2015	E10	Feces	VRE 2010-15	JANDQS000000000	SAMN29678855	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271229	210	2963153	562.746	*E. faecium*	17	159	*vanB*	Yes	Yes	No	2015	E10	Feces	VRE 2010-15	JANDQR000000000	SAMN29678856	PRJNA858233
51271509	188	2851377	213.604	*E. faecium*	192	3081		Yes	No	No	2014	E10	Blood	VSE 2014	JANDQQ000000000	SAMN29678857	PRJNA858233
51271510	211	2873542	175.939	*E. faecium*	203	3061		Yes	Yes	No	2014	E9	Blood	VSE 2014	JANDQP000000000	SAMN29678858	PRJNA858233
51271511	23	2478396	310.326	*E. faecium*	32	6214		No	No	No	2014	E9	Blood	VSE 2014	JANDQO000000000	SAMN29678859	PRJNA858233
51271512	194	2950086	257.832	*E. faecium*	2045	3061		Yes	No	No	2014	E9	Blood	VSE 2014	JANDQN000000000	SAMN29678860	PRJNA858233
51271513	202	2942109	137.4	*E. faecium*	117	24		Yes	No	No	2014	E9	Blood	VSE 2014	JANDQM000000000	SAMN29678861	PRJNA858233
51271514	254	2893150	64.6267	*E. faecium*	17	3038		Yes	No	No	2014	E9	Blood	VSE 2014	JANDQL000000000	SAMN29678862	PRJNA858233
51271515	181	2918323	228.366	*E. faecium*	117	24		Yes	Yes	No	2014	E9	Blood	VSE 2014	JANDQK000000000	SAMN29678863	PRJNA858233
51271516	154	2865309	292.224	*E. faecium*	203	3061		Yes	Yes	No	2014	E9	Blood	VSE 2014	JANDQJ000000000	SAMN29678864	PRJNA858233
51271517	53	2605345	185.298	*E. faecium*	32	1424		No	No	No	2014	E9	Blood	VSE 2014	JANDQI000000000	SAMN29678865	PRJNA858233
51271518	189	2956426	237.475	*E. faecium*	203	20	*vanA*	No	No	No	2014	E9	Blood	VRE 2010-15	JANDQH000000000	SAMN29678866	PRJNA858233
51271519	210	2973079	181.086	*E. faecium*	203	3056		Yes	No	No	2014	E9	Blood	VSE 2014	JANDQG000000000	SAMN29678867	PRJNA858233
51271520	274	2645307	107.789	*E. faecium*	18	3041	*vanA*	Yes	No	No	2011	E9	Urine	VRE 2010-15	JANDQF000000000	SAMN29678868	PRJNA858233
51271521	197	2910165	165.495	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	E9	Urine	VRE 2010-15	JANDQE000000000	SAMN29678869	PRJNA858233
51271522	246	2870629	179.92	*E. faecium*	203	20	*vanA*	Yes	Yes	No	2014	E9	Feces	VRE 2010-15	JANDQD000000000	SAMN29678870	PRJNA858233
51271523	177	2959501	277.808	*E. faecium*	203	3065	*vanA*	Yes	No	No	2015	E9	Feces	VRE 2010-15	JANDQC000000000	SAMN29678871	PRJNA858233
51271524	191	2947299	177.249	*E. faecium*	203	20	*vanA*	Yes	No	No	2015	E9	Urine	VRE 2010-15	JANDQB000000000	SAMN29678872	PRJNA858233
51271815	240	2989807	135.67	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Clinical si	VRE 2010-15	JANDQA000000000	SAMN29678873	PRJNA858233
51271816	251	2900461	126.134	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Blood	VRE 2010-15	JANDPZ000000000	SAMN29678874	PRJNA858233
51271817	240	2952790	101.907	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPY000000000	SAMN29678875	PRJNA858233
51271818	340	2824358	63.1183	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPX000000000	SAMN29678876	PRJNA858233
51271819	392	2744951	73.8543	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPW000000000	SAMN29678877	PRJNA858233
51271820	351	2821524	112.088	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPV000000000	SAMN29678878	PRJNA858233
51271821	231	3071769	98.8843	*E. faecium*	192	3087	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPU000000000	SAMN29678879	PRJNA858233
51271822	254	2929873	138.614	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPT000000000	SAMN29678880	PRJNA858233
51271823	329	2876548	79.0431	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Blood	VRE 2010-15	JANDPS000000000	SAMN29678881	PRJNA858233
51271824	215	2972803	105.583	*E. faecium*	192	3082	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPR000000000	SAMN29678882	PRJNA858233
51271825	212	2977959	83.3942	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANEVL000000000	SAMN29681587	PRJNA858233
51271826	263	2918621	102.185	*E. faecium*	117	3053	*vanB*	Yes	Yes	No	2013	W1	Feces	VRE 2010-15	JANDPQ000000000	SAMN29678883	PRJNA858233
51271880	284	2757211	105.234	*E. faecium*	203	3061		Yes	Yes	No	2014	E6	Blood	VSE 2014	JANDPP000000000	SAMN29678884	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271881	213	2840892	125.806	*E. faecium*	17	3037		Yes	No	No	2014	E6	Blood	VSE 2014	JANDPO000000000	SAMN29678885	PRJNA858233
51271882	90	2638741	123.698	*E. faecium*	32	6215		Yes	No	No	2014	E6	Blood	VSE 2014	JANDPN000000000	SAMN29678886	PRJNA858233
51271883	212	2777199	110.091	*E. faecium*	80	16		Yes	Yes	No	2014	E6	Blood	VSE 2014	JANDPM000000000	SAMN29678887	PRJNA858233
51271884	209	2869762	124.723	*E. faecium*	117	3054		Yes	No	No	2014	E6	Blood	VSE 2014	JANDPL000000000	SAMN29678888	PRJNA858233
51271885	264	2667659	108.808	*E. faecium*	262	2646		No	No	No	2014	E6	Blood	VSE 2014	JANDPK000000000	SAMN29678889	PRJNA858233
51271886	223	2792865	99.5447	*E. faecium*	192	3081		Yes	No	No	2014	E6	Blood	VSE 2014	JANDPJ000000000	SAMN29678890	PRJNA858233
51271887	284	2831632	96.7842	*E. faecium*	192	3088		Yes	Yes	No	2014	E6	Blood	VSE 2014	JANDPI000000000	SAMN29678891	PRJNA858233
51271888	215	2822502	156.882	*E. faecium*	17	3036		Yes	No	No	2014	E3	Blood	VSE 2014	JANDPH000000000	SAMN29678892	PRJNA858233
51271889	173	2744618	101.651	*E. faecium*	108	6216		No	No	No	2014	E3	Blood	VSE 2014	JANDPG000000000	SAMN29678893	PRJNA858233
51271890	236	2478795	68.710	*E. faecium*	178	6217		No	No	No	2014	E3	Blood	VSE 2014	JANDPF000000000	SAMN29678894	PRJNA858233
51271891	291	2812956	135.281	*E. faecium*	203	3061		Yes	Yes	No	2014	E3	Blood	VSE 2014	JANDPE000000000	SAMN29678895	PRJNA858233
51271892	237	2936001	283.108	*E. faecium*	192	3081		Yes	No	No	2014	E3	Blood	VSE 2014	JANDPD000000000	SAMN29678896	PRJNA858233
51271893	242	2895901	88.923	*E. faecium*	78	6218		Yes	No	No	2014	E3	Blood	VSE 2014	JANDPC000000000	SAMN29678897	PRJNA858233
51271894	331	2721993	141.396	*E. faecium*	203	3061		Yes	Yes	No	2014	E3	Blood	VSE 2014	JANDPB000000000	SAMN29678898	PRJNA858233
51271895	215	2838182	128.471	*E. faecium*	2046	53		Yes	No	No	2014	E3	Blood	VSE 2014	JANDPA000000000	SAMN29678899	PRJNA858233
51271896	124	2690478	131.971	*E. faecium*	178	6219		No	No	No	2014	E3	Blood	VSE 2014	JANDOZ000000000	SAMN29678900	PRJNA858233
51271898	199	2851279	152.882	*E. faecium*	192	3081		Yes	No	No	2014	E3	Blood	VSE 2014	JANDOY000000000	SAMN29678901	PRJNA858233
51271900	268	2967040	112.01	*E. faecium*	203	191	*vanA*	Yes	No	No	2014	E3	Blood	VRE 2010-15	JANDOX000000000	SAMN29678902	PRJNA858233
51271901	183	2823869	97.388	*E. faecium*	80	16	*vanB*	Yes	No	No	2014	E3	Blood	VRE 2010-15	JANDOW000000000	SAMN29678903	PRJNA858233
51271902	94	2576387	134.239	*E. faecium*	32	6220		No	No	No	2014	E3	Blood	VSE 2014	JANDOV000000000	SAMN29678904	PRJNA858233
51271903	178	2772880	403.074	*E. faecium*	80	16		Yes	Yes	No	2014	E3	Blood	VSE 2014	JANDOU000000000	SAMN29678905	PRJNA858233
51271904	233	2753956	120.3	*E. faecium*	80	16		Yes	No	No	2014	E3	Blood	VSE 2014	JANDOT000000000	SAMN29678906	PRJNA858233
51271919	250	3012639	143.078	*E. faecium*	203	191	*vanA*	Yes	Yes	No	2014	E6	Feces	VRE 2010-15	JANDOS000000000	SAMN29678907	PRJNA858233
51271920	275	2985209	141.452	*E. faecium*	203	191	*vanA*	Yes	Yes	No	2014	E6	Urine	VRE 2010-15	JANDOR000000000	SAMN29678908	PRJNA858233
51271921	247	3016298	126.638	*E. faecium*	203	191	*vanA*	Yes	Yes	No	2014	E6	Feces	VRE 2010-15	JANDOQ000000000	SAMN29678909	PRJNA858233
51271922	246	2922477	120.306	*E. faecium*	233	6221	*vanA*	Yes	Yes	No	2015	E6	Feces	VRE 2010-15	JANDOP000000000	SAMN29678910	PRJNA858233
51271923	286	2984150	115.23	*E. faecium*	117	71	*vanB*	Yes	No	No	2015	E6	Feces	VRE 2010-15	JANDOO000000000	SAMN29678911	PRJNA858233
51271927	204	2838612	107.98	*E. faecium*	80	2632	*vanA*	Yes	No	No	2015	E6	Feces	VRE 2010-15	JANDON000000000	SAMN29678912	PRJNA858233
51271928	226	2964696	64.6779	*E. faecium*	17	3037	*vanA*	Yes	No	No	2015	E6	Feces	VRE 2010-15	JANEVJ000000000	SAMN29681590	PRJNA858233
51271929	217	2869990	121.786	*E. faecium*	18	3047	*vanB*	Yes	No	No	2010	E3	Blood	VRE 2010-15	JANDOM000000000	SAMN29678913	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271930	311	2836404	74.625	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2011	E3	Feces	VRE 2010-15	JANDOL000000000	SAMN29678914	PRJNA858233
51271931	274	2757824	143.238	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2011	E3	Urine	VRE 2010-15	JANDOK000000000	SAMN29678915	PRJNA858233
51271932	255	2911220	93.9158	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2011	E3	Feces	VRE 2010-15	JANDOJ000000000	SAMN29678916	PRJNA858233
51271933	337	2779969	40.0921	*E. faecium*	202	3079	*vanA*	Yes	Yes	No	2011	E3	Urine	VRE 2010-15	JANEVI000000000	SAMN29681591	PRJNA858233
51271934	307	2791676	134.789	*E. faecium*	80	3097	*vanA*	Yes	No	No	2011	E3	Feces	VRE 2010-15	JANDOI000000000	SAMN29678917	PRJNA858233
51271935	204	2868768	81.0327	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2011	E3	Feces	VRE 2010-15	JANDOH000000000	SAMN29678918	PRJNA858233
51271936	250	2881592	69.8998	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2012	E3	Urine	VRE 2010-15	JANEVH000000000	SAMN29681592	PRJNA858233
51271937	233	2870412	77.4812	*E. faecium*	203	3066	*vanA*	Yes	Yes	No	2012	E3	Urine	VRE 2010-15	JANDOG000000000	SAMN29678919	PRJNA858233
51271938	217	2827743	87.4822	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2012	E3	Blood	VRE 2010-15	JANDOF000000000	SAMN29678920	PRJNA858233
51271939	334	2857128	133.739	*E. faecium*	18	3042	*vanA*	Yes	No	No	2012	E3	Feces	VRE 2010-15	JANDOE000000000	SAMN29678921	PRJNA858233
51271940	326	2789439	118.17	*E. faecium*	192	3090	*vanA*	Yes	Yes	No	2012	E3	Clinical si	VRE 2010-15	JANDOD000000000	SAMN29678922	PRJNA858233
51271942	291	2910177	121.298	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	E3	Feces	VRE 2010-15	JANDOC000000000	SAMN29678923	PRJNA858233
51271943	241	2956149	79.6872	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	E3	Feces	VRE 2010-15	JANDOB000000000	SAMN29678924	PRJNA858233
51271944	290	2809219	141.296	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	E3	Feces	VRE 2010-15	JANDOA000000000	SAMN29678925	PRJNA858233
51271945	180	2922868	208.792	*E. faecium*	78	6222	*vanA*	Yes	No	No	2014	E3	Urine	VRE 2010-15	JANDNZ000000000	SAMN29678926	PRJNA858233
51271946	200	2894333	164.544	*E. faecium*	203	20	*vanA*	Yes	No	No	2014	E3	Urine	VRE 2010-15	JANDNY000000000	SAMN29678927	PRJNA858233
51271978	212	3010001	316.396	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNX000000000	SAMN29678928	PRJNA858233
51271982	275	3028131	39.8528	*E. faecium*	117	24	*vanB*	Yes	No	No	2014	W1	Feces	VRE 2010-15	JANDNW000000000	SAMN29678929	PRJNA858233
51271984	356	2863870	80.7859	*E. faecium*	203	191	*vanA*	Yes	Yes	No	2014	W1	Clinical si	VRE 2010-15	JANDNV000000000	SAMN29678930	PRJNA858233
51271985	209	2995879	526.717	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNU000000000	SAMN29678931	PRJNA858233
51271986	289	2864375	64.9184	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNT000000000	SAMN29678932	PRJNA858233
51271987	386	2819184	95.8828	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNS000000000	SAMN29678933	PRJNA858233
51271988	201	2997225	362.269	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNR000000000	SAMN29678934	PRJNA858233
51271989	246	2862463	71.683	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNQ000000000	SAMN29678935	PRJNA858233
51271990	271	2820354	103.149	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNP000000000	SAMN29678936	PRJNA858233
51271991	316	2815592	127.118	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNO000000000	SAMN29678937	PRJNA858233
51271992	327	2818555	140.61	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Clinical si	VRE 2010-15	JANDNN000000000	SAMN29678938	PRJNA858233
51271993	308	2827555	125.985	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNM000000000	SAMN29678939	PRJNA858233
51271994	257	2885658	114.683	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNL000000000	SAMN29678940	PRJNA858233
51271995	360	2837906	118.784	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Urine	VRE 2010-15	JANDNK000000000	SAMN29678941	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51271996	237	2903419	178.625	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Urine	VRE 2010-15	JANDNJ000000000	SAMN29678942	PRJNA858233
51271997	250	2913134	165.577	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANDNI000000000	SAMN29678943	PRJNA858233
51271998	218	2935646	104.419	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Urine	VRE 2010-15	JANDNH000000000	SAMN29678944	PRJNA858233
51271999	196	2980092	93.8833	*E. faecium*	117	3055	*vanB*	Yes	Yes	No	2014	W1	Urine	VRE 2010-15	JANDNG000000000	SAMN29678945	PRJNA858233
51272000	177	2954309	85.9017	*E. faecium*	192	26	*vanB*	Yes	No	No	2014	W1	Feces	VRE 2010-15	JANDNF000000000	SAMN29678946	PRJNA858233
51272001	341	2868897	85.6789	*E. faecium*	203	3061	*vanB*	Yes	Yes	No	2015	W1	Urine	VRE 2010-15	JANDNE000000000	SAMN29678947	PRJNA858233
51272002	264	2860168	109.376	*E. faecium*	192	3082	*vanB*	Yes	No	No	2015	W1	Feces	VRE 2010-15	JANDND000000000	SAMN29678948	PRJNA858233
51272003	322	2883506	148.295	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2015	W1	Feces	VRE 2010-15	JANDNC000000000	SAMN29678949	PRJNA858233
51272004	256	2943208	162.874	*E. faecium*	80	3098	*vanA*	Yes	Yes	No	2015	W1	Feces	VRE 2010-15	JANDNB000000000	SAMN29678950	PRJNA858233
51273071	186	2830960	198.299	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANDNA000000000	SAMN29678951	PRJNA858233
51273073	207	2891929	193.586	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANDMZ000000000	SAMN29678952	PRJNA858233
51273074	195	2870901	149.978	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANDMY000000000	SAMN29678953	PRJNA858233
51273075	245	2794906	136.353	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMX000000000	SAMN29678954	PRJNA858233
51273076	269	2835661	148.773	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMW000000000	SAMN29678955	PRJNA858233
51273077	316	2840395	157.834	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMV000000000	SAMN29678956	PRJNA858233
51273078	276	2863562	164.325	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANDMU000000000	SAMN29678957	PRJNA858233
51273079	249	2823027	157.374	*E. faecium*	192	26	*vanB*	Yes	No	Yes	2011	W1	Feces	VRE 2010-15	JANDMT000000000	SAMN29678958	PRJNA858233
51273080	187	2831531	166.507	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMS000000000	SAMN29678959	PRJNA858233
51273081	175	2880540	141.268	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMR000000000	SAMN29678960	PRJNA858233
51273082	182	2887241	164.332	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANDMQ000000000	SAMN29678961	PRJNA858233
51273083	179	2852813	152.839	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMP000000000	SAMN29678962	PRJNA858233
51273084	176	2907054	163.192	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMO000000000	SAMN29678963	PRJNA858233
51273085	200	2877598	130.434	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMN000000000	SAMN29678964	PRJNA858233
51273086	218	2951618	166.531	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANDMM000000000	SAMN29678965	PRJNA858233
51273087	267	2832631	176.004	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANEVG000000000	SAMN29681593	PRJNA858233
51273088	199	2876575	232.078	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANEVF000000000	SAMN29681594	PRJNA858233
51273089	178	2849550	271.776	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEVE000000000	SAMN29681595	PRJNA858233
51273090	184	2849206	189.066	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEVD000000000	SAMN29681596	PRJNA858233
51273091	179	2904696	270.258	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEVC000000000	SAMN29681597	PRJNA858233
51273092	188	2892485	209.47	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEVB000000000	SAMN29681598	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51273093	183	2902574	237.68	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEVA000000000	SAMN29681599	PRJNA858233
51273094	177	2905138	220.416	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUZ000000000	SAMN29681600	PRJNA858233
51273095	211	2954003	162.64	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANEUY000000000	SAMN29681601	PRJNA858233
51273096	273	2816920	100.5	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUX000000000	SAMN29681602	PRJNA858233
51273097	333	2847655	142.925	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUW000000000	SAMN29681603	PRJNA858233
51273098	171	2858588	71.1981	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUV000000000	SAMN29681604	PRJNA858233
51273099	210	3012813	282.386	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUU000000000	SAMN29681605	PRJNA858233
51273100	202	2869768	280.451	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUT000000000	SAMN29681606	PRJNA858233
51273101	171	2851727	253.8	*E. faecium*	192	3	*vanB*	Yes	Yes	No	2011	W1	Blood	VRE 2010-15	JANEUS000000000	SAMN29681607	PRJNA858233
51273102	176	2905254	283.779	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUR000000000	SAMN29681608	PRJNA858233
51273103	196	3055187	249.47	*E. faecium*	192	3	*vanB*	Yes	Yes	No	2011	W1	Feces	VRE 2010-15	JANEUQ000000000	SAMN29681609	PRJNA858233
51273104	170	2904835	233.561	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUP000000000	SAMN29681610	PRJNA858233
51273105	226	3125006	272.564	*E. faecium*	192	3	*vanB*	Yes	Yes	No	2011	W1	Unknowr	VRE 2010-15	JANEUO000000000	SAMN29681611	PRJNA858233
51273106	170	2860059	36.028	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUN000000000	SAMN29681612	PRJNA858233
51273107	176	2899797	240.21	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUM000000000	SAMN29681613	PRJNA858233
51273108	233	2926366	181.716	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANEUL000000000	SAMN29681614	PRJNA858233
51273451	285	2853885	115.086	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUK000000000	SAMN29681615	PRJNA858233
51273452	319	2879464	193.957	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUJ000000000	SAMN29681616	PRJNA858233
51273453	176	2907745	105.725	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUI000000000	SAMN29681617	PRJNA858233
51273454	176	2997215	150.309	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUH000000000	SAMN29681618	PRJNA858233
51273455	280	2860908	156.846	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUG000000000	SAMN29681619	PRJNA858233
51273456	295	2784820	124.69	*E. faecium*	202	3078	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUF000000000	SAMN29681620	PRJNA858233
51273457	208	2827489	164.861	*E. faecium*	1056	6223	*vanA*	Yes	Yes	No	2011	W1	Urine	VRE 2010-15	JANEUE000000000	SAMN29681621	PRJNA858233
51273458	239	2928970	108.989	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUD000000000	SAMN29681622	PRJNA858233
51273459	230	2807207	99.034	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUC000000000	SAMN29681623	PRJNA858233
51273460	209	2950482	96.4548	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUB000000000	SAMN29681624	PRJNA858233
51273461	334	2789691	124.541	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANEUA000000000	SAMN29681625	PRJNA858233
51273462	210	2823408	212.272	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETZ000000000	SAMN29681626	PRJNA858233
51273464	332	2841079	124.892	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETY000000000	SAMN29681627	PRJNA858233
51273465	264	2878038	120.442	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETX000000000	SAMN29681628	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51273467	208	2836643	128.301	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANETV000000000	SAMN29681630	PRJNA858233
51273468	353	2748401	158.272	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANETU000000000	SAMN29681631	PRJNA858233
51273469	144	2778677	99.1431	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETT000000000	SAMN29681632	PRJNA858233
51273470	231	2922523	29.6823	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETS000000000	SAMN29681633	PRJNA858233
51273471	317	2870419	63.1386	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETR000000000	SAMN29681634	PRJNA858233
51273472	224	2970231	86.6235	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETQ000000000	SAMN29681635	PRJNA858233
51273473	208	2957204	127.303	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Blood	VRE 2010-15	JANETP000000000	SAMN29681636	PRJNA858233
51273474	288	2732905	79.3805	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETO000000000	SAMN29681637	PRJNA858233
51273475	302	2874981	136.12	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANETN000000000	SAMN29681638	PRJNA858233
51273476	286	2748477	137.974	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Urine	VRE 2010-15	JANETM000000000	SAMN29681639	PRJNA858233
51273477	329	2857531	52.7804	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETL000000000	SAMN29681640	PRJNA858233
51273478	268	2984788	121.722	*E. faecium*	192	3080	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETK000000000	SAMN29681641	PRJNA858233
51273479	231	2921761	97.2021	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETJ000000000	SAMN29681642	PRJNA858233
51273480	192	2792716	110.788	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETI000000000	SAMN29681643	PRJNA858233
51273481	185	2993565	55.205	*E. faecium*	192	26	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETH000000000	SAMN29681644	PRJNA858233
51273482	189	2856673	40.2883	*E. faecium*	192	3	*vanB*	Yes	No	No	2011	W1	Feces	VRE 2010-15	JANETG000000000	SAMN29681645	PRJNA858233
51273483	246	2884361	89.114	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANETF000000000	SAMN29681646	PRJNA858233
51273484	264	2915623	65.9714	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Urine	VRE 2010-15	JANETE000000000	SAMN29681647	PRJNA858233
51273485	176	2994753	49.964	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Urine	VRE 2010-15	JANETD000000000	SAMN29681648	PRJNA858233
51273486	207	3014616	104.62	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANETC000000000	SAMN29681649	PRJNA858233
51273487	191	3009331	200.28	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANETB000000000	SAMN29681650	PRJNA858233
51273488	241	2953100	202.83	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Urine	VRE 2010-15	JANETA000000000	SAMN29681651	PRJNA858233
51273489	254	3022470	223.448	*E. faecium*	192	3	*vanB*	Yes	Yes	No	2012	W1	Feces	VRE 2010-15	JANESZ000000000	SAMN29681652	PRJNA858233
51273490	358	2954843	109.052	*E. faecium*	192	26	*vanB*	Yes	Yes	No	2012	W1	Feces	VRE 2010-15	JANESY000000000	SAMN29681653	PRJNA858233
51273491	151	2722586	230.124	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESX000000000	SAMN29681654	PRJNA858233
51273492	191	2876641	241.771	*E. faecium*	192	26	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESW000000000	SAMN29681655	PRJNA858233
51273493	181	3008424	172.729	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESV000000000	SAMN29681656	PRJNA858233
51273494	183	2974401	213.963	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Urine	VRE 2010-15	JANESU000000000	SAMN29681657	PRJNA858233
51273495	176	2916393	198.779	*E. faecium*	80	3099	*vanB*	Yes	Yes	No	2012	W1	Feces	VRE 2010-15	JANEST000000000	SAMN29681658	PRJNA858233
51273496	198	2996496	196.056	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESS000000000	SAMN29681659	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51273497	235	2935420	203.241	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESR000000000	SAMN29681660	PRJNA858233
51273498	201	2830589	218.158	*E. faecium*	192	3	*vanB*	Yes	No	No	2012	W1	Feces	VRE 2010-15	JANESQ000000000	SAMN29681661	PRJNA858233
51273553	231	2920560	161.101	*E. faecium*	17	3104	*vanA*	not determined	not determined	not determined	2015	E12	Feces	VRE 2010-15	JANESP000000000	SAMN29681662	PRJNA858233
51273875	196	2908367	113.653	*E. faecium*	80	3097	*vanA*	Yes	No	No	2011	E3	Unknowr	VRE 2010-15	JANESO000000000	SAMN29681663	PRJNA858233
51274612	371	2686122	192.739	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Blood	VRE 2010-15	JANESN000000000	SAMN29681664	PRJNA858233
51274613	174	2902654	80.5701	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESM000000000	SAMN29681665	PRJNA858233
51274614	304	2748006	155.431	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESL000000000	SAMN29681666	PRJNA858233
51274615	224	2835580	198.794	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESK000000000	SAMN29681667	PRJNA858233
51274616	189	2865355	135.606	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANESJ000000000	SAMN29681668	PRJNA858233
51274617	181	2878617	134.049	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESI000000000	SAMN29681669	PRJNA858233
51274618	175	2893173	133.55	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESH000000000	SAMN29681670	PRJNA858233
51274619	197	2880338	140.327	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANESG000000000	SAMN29681671	PRJNA858233
51274620	162	2881287	149.979	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESF000000000	SAMN29681672	PRJNA858233
51274621	174	2942192	164.41	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANESE000000000	SAMN29681673	PRJNA858233
51274622	212	2846674	144.639	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESD000000000	SAMN29681674	PRJNA858233
51274623	319	2828680	148.116	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESC000000000	SAMN29681675	PRJNA858233
51274625	223	2831250	191.899	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Clinical si	VRE 2010-15	JANESB000000000	SAMN29681676	PRJNA858233
51274626	186	2868714	162.972	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANESA000000000	SAMN29681677	PRJNA858233
51274627	196	2953370	157.624	*E. faecium*	192	26	*vanB*	Yes	Yes	No	2010	W1	Feces	VRE 2010-15	JANERZ000000000	SAMN29681678	PRJNA858233
51274628	158	2904243	137.223	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERY000000000	SAMN29681679	PRJNA858233
51274629	167	2907545	124.582	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERX000000000	SAMN29681680	PRJNA858233
51274630	167	2907394	85.2872	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERW000000000	SAMN29681681	PRJNA858233
51274631	177	2989289	129.687	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Unknowr	VRE 2010-15	JANERV000000000	SAMN29681682	PRJNA858233
51274632	186	2878464	128.836	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERU000000000	SAMN29681683	PRJNA858233
51274633	180	2896748	147.427	*E. faecium*	192	26	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERT000000000	SAMN29681684	PRJNA858233
51274634	329	2733997	117.662	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERS000000000	SAMN29681685	PRJNA858233
51274635	166	2891835	143.826	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANERR000000000	SAMN29681686	PRJNA858233
51274636	231	2838901	146.308	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANERQ000000000	SAMN29681687	PRJNA858233
51274637	295	2734272	140.792	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERP000000000	SAMN29681688	PRJNA858233
51274638	332	2753587	85.5504	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERO000000000	SAMN29681689	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
51274639	228	2851980	163.018	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERN000000000	SAMN29681690	PRJNA858233
51274640	229	2874044	197.223	*E. faecium*	192	3091	*vanB*	Yes	No	No	2010	W1	Urine	VRE 2010-15	JANERM000000000	SAMN29681691	PRJNA858233
51274641	238	2891227	172.37	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERL000000000	SAMN29681692	PRJNA858233
51274642	172	2900289	55.7552	*E. faecium*	192	3	*vanB*	Yes	No	No	2010	W1	Feces	VRE 2010-15	JANERK000000000	SAMN29681693	PRJNA858233
51274643	197	2930902	226.256	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	W1	Feces	VRE 2010-15	JANERJ000000000	SAMN29681694	PRJNA858233
51276488	176	2895257	49.2809	*E. faecium*	117	3054		Yes	No	No	2014	E1	Blood	VSE 2014	JANERI000000000	SAMN29681695	PRJNA858233
51276490	187	2910322	146.506	*E. faecium*	192	3089		Yes	Yes	No	2014	E1	Blood	VSE 2014	JANERH000000000	SAMN29681696	PRJNA858233
51276491	247	2876178	97.3974	*E. faecium*	203	3056		Yes	No	No	2014	E1	Blood	VSE 2014	JANERG000000000	SAMN29681697	PRJNA858233
51276492	267	2750395	97.3315	*E. faecium*	17	3035		Yes	No	No	2014	E1	Blood	VSE 2014	JANERF000000000	SAMN29681698	PRJNA858233
51276494	195	2919623	261.116	*E. faecium*	203	3061		Yes	No	No	2014	E1	Blood	VSE 2014	JANERE000000000	SAMN29681699	PRJNA858233
51276495	197	2934013	81.4537	*E. faecium*	203	3056		Yes	No	No	2014	E1	Blood	VSE 2014	JANERD000000000	SAMN29681700	PRJNA858233
51276496	178	2823313	120.7	*E. faecium*	192	3081		Yes	No	No	2014	E1	Blood	VSE 2014	JANERC000000000	SAMN29681701	PRJNA858233
51276497	208	2876551	99.7973	*E. faecium*	117	24		Yes	Yes	No	2014	E1	Blood	VSE 2014	JANERB000000000	SAMN29681702	PRJNA858233
51276498	198	2736347	98.7253	*E. faecium*	192	17		Yes	Yes	No	2014	E1	Blood	VSE 2014	JANERA000000000	SAMN29681703	PRJNA858233
51276499	280	2793683	91.703	*E. faecium*	203	3061		Yes	Yes	No	2014	E1	Blood	VSE 2014	JANEQZ000000000	SAMN29681704	PRJNA858233
51276500	263	2684203	98.1823	*E. faecium*	80	16		Yes	No	No	2014	E1	Blood	VSE 2014	JANEQY000000000	SAMN29681705	PRJNA858233
51276501	219	3030667	342.028	*E. faecium*	203	191		Yes	No	No	2014	E1	Blood	VSE 2014	JANEQX000000000	SAMN29681706	PRJNA858233
51276502	185	2962399	268.852	*E. faecium*	117	3054		Yes	No	No	2014	E1	Blood	VSE 2014	JANEQW000000000	SAMN29681707	PRJNA858233
51276507	219	3063006	141.963	*E. faecium*	192	3080	*vanB*	Yes	No	No	2011	E1	Blood	VRE 2010-15	JANEQV000000000	SAMN29681708	PRJNA858233
51276508	205	2922702	130.415	*E. faecium*	80	3097	*vanA*	Yes	Yes	No	2011	E1	Feces	VRE 2010-15	JANEQU000000000	SAMN29681709	PRJNA858233
51276509	270	2897322	47.482	*E. faecium*	18	3042	*vanA*	Yes	No	No	2012	E1	Blood	VRE 2010-15	JANEQT000000000	SAMN29681710	PRJNA858233
51276510	244	2842863	142.257	*E. faecium*	80	3100	*vanA*	Yes	Yes	No	2014	E1	Urine	VRE 2010-15	JANEQS000000000	SAMN29681711	PRJNA858233
51276511	177	2936372	158.921	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	E1	Feces	VRE 2010-15	JANEQR000000000	SAMN29681712	PRJNA858233
51276512	193	2981433	201.919	*E. faecium*	117	24	*vanB*	Yes	Yes	No	2014	E1	Feces	VRE 2010-15	JANEQQ000000000	SAMN29681713	PRJNA858233
51276513	254	2965575	130.567	*E. faecium*	2063	3037	*vanA*	Yes	No	No	2015	E1	Feces	VRE 2010-15	JANEQP000000000	SAMN29681714	PRJNA858233
51276514	215	2959065	202.783	*E. faecium*	203	3061	*vanB*	Yes	Yes	No	2015	E1	Clinical si	VRE 2010-15	JANEQO000000000	SAMN29681715	PRJNA858233
51280801	211	3005375	362.475	*E. faecium*	203	3074		Yes	Yes	No	2014	E1	Blood	VSE 2014	JANEQN000000000	SAMN29681716	PRJNA858233
51280802	184	2812951	358.166	*E. faecium*	80	16		Yes	No	No	2014	E1	Blood	VSE 2014	JANEQM000000000	SAMN29681717	PRJNA858233
51296109	230	2926974	347.964	*E. faecium*	80	3097	*vanA*	Yes	No	No	2010	E6	Clinical si	VRE 2010-15	JANEQL000000000	SAMN29681718	PRJNA858233
51296110	259	3066079	124.947	*E. faecium*	203	191	*vanA*	Yes	Yes	No	2014	E6	Feces	VRE 2010-15	JANEQK000000000	SAMN29681719	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
K59-16	290	2990330	258.358	*E. faecium*	440	6224		Yes	Yes	No	2008	N2	Blood	VSE 2008	JANEQH000000000	SAMN29681722	PRJNA858233
K59-17	61	2575051	302.488	*E. faecium*	22	6225		No	No	No	2008	M3	Blood	VSE 2008	JANEQG000000000	SAMN29681723	PRJNA858233
K59-18	210	2927590	214.275	*E. faecium*	574	6226		Yes	No	No	2008	M3	Blood	VSE 2008	JANEQF000000000	SAMN29681724	PRJNA858233
K59-19	75	2634831	248.932	*E. faecium*	296	426		Yes	No	No	2008	M3	Blood	VSE 2008	JANEQE000000000	SAMN29681725	PRJNA858233
K59-20	240	3004093	278.206	*E. faecium*	203	3062		Yes	Yes	No	2008	M4	Blood	VSE 2008	JANEQD000000000	SAMN29681726	PRJNA858233
K59-21	319	3062938	239.03	*E. faecium*	203	3067		Yes	Yes	No	2008	M4	Blood	VSE 2008	JANEQC000000000	SAMN29681727	PRJNA858233
K59-22	231	2961828	279.45	*E. faecium*	78	6227		Yes	No	No	2008	M4	Blood	VSE 2008	JANEQB000000000	SAMN29681728	PRJNA858233
K59-23	335	3087049	236.827	*E. faecium*	203	3068		Yes	Yes	No	2008	W3	Blood	VSE 2008	JANEQA000000000	SAMN29681729	PRJNA858233
K59-25	271	3087924	182.724	*E. faecium*	203	3067		Yes	Yes	No	2008	W4	Blood	VSE 2008	JANEPZ000000000	SAMN29681730	PRJNA858233
K59-26	156	2770594	271.699	*E. faecium*	94	6228		No	No	No	2008	W2	Blood	VSE 2008	JANEPY000000000	SAMN29681731	PRJNA858233
K59-27	320	3081279	192.167	*E. faecium*	17	3030		Yes	Yes	No	2008	W2	Blood	VSE 2008	JANEPX000000000	SAMN29681732	PRJNA858233
K59-28	277	2928766	266.729	*E. faecium*	17	3033		No	No	No	2008	E12	Blood	VSE 2008	JANEPW000000000	SAMN29681733	PRJNA858233
K59-29	287	3053361	239.884	*E. faecium*	203	3067		Yes	No	No	2008	E12	Blood	VSE 2008	JANEPV000000000	SAMN29681734	PRJNA858233
K59-30	218	2804996	233.089	*E. faecium*	192	1217		Yes	No	No	2008	E12	Blood	VSE 2008	JANEPU000000000	SAMN29681735	PRJNA858233
K59-31	274	3021544	258.685	*E. faecium*	1421	3031		Yes	Yes	No	2008	E12	Blood	VSE 2008	JANEPT000000000	SAMN29681736	PRJNA858233
K59-32	260	3129468	255.789	*E. faecium*	203	3067		Yes	No	No	2008	E11	Blood	VSE 2008	JANEPS000000000	SAMN29681737	PRJNA858233
K59-33	318	3053706	279.972	*E. faecium*	203	3069		Yes	Yes	No	2008	E11	Blood	VSE 2008	JANEPR000000000	SAMN29681738	PRJNA858233
K59-34	304	2907717	230.889	*E. faecium*	192	397		Yes	No	No	2008	E11	Blood	VSE 2008	JANEPQ000000000	SAMN29681739	PRJNA858233
K59-35	297	3080443	186.309	*E. faecium*	203	3067		Yes	Yes	No	2008	E11	Blood	VSE 2008	JANEPP000000000	SAMN29681740	PRJNA858233
K59-36	189	2800168	216.194	*E. faecium*	575	6229		No	Yes	No	2008	E10	Blood	VSE 2008	JANEPO000000000	SAMN29681741	PRJNA858233
K59-37	207	3007305	242.948	*E. faecium*	203	3061		Yes	Yes	No	2008	E10	Blood	VSE 2008	JANEPN000000000	SAMN29681742	PRJNA858233
K59-40	242	2990557	263.633	*E. faecium*	203	3061		Yes	Yes	No	2008	E10	Blood	VSE 2008	JANEPM000000000	SAMN29681743	PRJNA858233
K59-41	394	3086988	242.901	*E. faecium*	203	3067		Yes	Yes	No	2008	E10	Blood	VSE 2008	JANEPL000000000	SAMN29681744	PRJNA858233
K59-42	362	2910043	172.104	*E. faecium*	192	1217		Yes	Yes	No	2008	E10	Blood	VSE 2008	JANEPK000000000	SAMN29681745	PRJNA858233
K59-43	379	2912322	195.744	*E. faecium*	192	3092		Yes	Yes	No	2008	E10	Blood	VSE 2008	JANEPJ000000000	SAMN29681746	PRJNA858233
K59-44	61	2527591	202.273	*E. faecium*	32	5262		No	No	No	2008	E10	Blood	VSE 2008	JANEPI000000000	SAMN29681747	PRJNA858233
K59-46	67	2439555	235.708	*E. faecium*	533	6230		No	No	No	2008	E10	Blood	VSE 2008	JANEPH000000000	SAMN29681748	PRJNA858233
K59-48	105	2757331	259.489	*E. faecium*	94	6231		No	No	No	2008	E6	Blood	VSE 2008	JANEPG000000000	SAMN29681749	PRJNA858233
K59-49	52	2570603	216.377	*E. faecium*	32	6232		No	No	No	2008	E6	Blood	VSE 2008	JANEPF000000000	SAMN29681750	PRJNA858233
K59-50	124	2663717	237.842	*E. faecium*	202	3076		Yes	No	No	2008	E4	Blood	VSE 2008	JANEPE000000000	SAMN29681751	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
K59-51	237	2893437	216.748	*E. faecium*	18	3052		Yes	Yes	No	2008	W1	Blood	VSE 2008	JANEPD000000000	SAMN29681752	PRJNA858233
K59-52	138	2729428	195.849	*E. faecium*	576	6233		No	No	No	2008	W1	Blood	VSE 2008	JANEPC000000000	SAMN29681753	PRJNA858233
K59-53	302	2907485	218.159	*E. faecium*	132	6234		Yes	No	No	2008	W1	Blood	VSE 2008	JANEPB000000000	SAMN29681754	PRJNA858233
K59-54	319	3084363	196.339	*E. faecium*	203	3067		Yes	Yes	No	2008	W1	Blood	VSE 2008	JANEPA000000000	SAMN29681755	PRJNA858233
K59-55	196	2796820	222.401	*E. faecium*	279	6235		Yes	Yes	No	2008	W1	Blood	VSE 2008	JANEOZ000000000	SAMN29681756	PRJNA858233
K59-56	235	3048099	213.051	*E. faecium*	203	3073		Yes	No	No	2008	W1	Blood	VSE 2008	JANEOY000000000	SAMN29681757	PRJNA858233
K59-57	93	2753638	279.92	*E. faecium*	38	6236		No	No	No	2008	W1	Blood	VSE 2008	JANEOX000000000	SAMN29681758	PRJNA858233
K59-58	271	3007634	181.069	*E. faecium*	132	6237		Yes	Yes	No	2008	W1	Blood	VSE 2008	JANEOW000000000	SAMN29681759	PRJNA858233
K59-59	218	3000267	62.1764	*E. faecium*	203	3061		Yes	Yes	No	2008	W1	Blood	VSE 2008	JANEOV000000000	SAMN29681760	PRJNA858233
K59-60	213	3020566	86.3268	*E. faecium*	203	3061		No	Yes	No	2008	M1	Blood	VSE 2008	JANEOU000000000	SAMN29681761	PRJNA858233
K59-62	204	2882134	255.462	*E. faecium*	282	6238		Yes	No	No	2008	M1	Blood	VSE 2008	JANEOT000000000	SAMN29681762	PRJNA858233
K59-63	329	3112934	198.325	*E. faecium*	17	3034		Yes	Yes	No	2008	M1	Blood	VSE 2008	JANEOS000000000	SAMN29681763	PRJNA858233
K59-64	324	3042247	214.309	*E. faecium*	17	3034		Yes	Yes	No	2008	M1	Blood	VSE 2008	JANEOR000000000	SAMN29681764	PRJNA858233
K59-65	335	3079066	163.414	*E. faecium*	17	3034		Yes	Yes	No	2008	M1	Blood	VSE 2008	JANEOQ000000000	SAMN29681765	PRJNA858233
K59-66	333	2867857	248.597	*E. faecium*	18	3044		Yes	No	No	2008	M1	Blood	VSE 2008	JANEOP000000000	SAMN29681766	PRJNA858233
K59-67	310	2910113	196.508	*E. faecium*	18	3043		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOO000000000	SAMN29681767	PRJNA858233
K59-68	4	2947119	300	*E. faecium*	203	3069		Yes	Yes	No	2008	N1	Blood	VSE 2008	GCA_002263115.1	SAMN07326775	PRJNA393251
K59-69	263	2913649	220.9	*E. faecium*	18	3043		Yes	No	No	2008	N1	Blood	VSE 2008	JANEON000000000	SAMN29681768	PRJNA858233
K59-70	198	2791638	261.52	*E. faecium*	18	3050		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOM000000000	SAMN29681769	PRJNA858233
K59-71	286	3097155	262.158	*E. faecium*	203	3070		Yes	Yes	No	2008	N1	Blood	VSE 2008	JANEOL000000000	SAMN29681770	PRJNA858233
K59-72	355	3116945	237.753	*E. faecium*	203	3072		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOK000000000	SAMN29681771	PRJNA858233
K59-73	257	2831998	255.047	*E. faecium*	440	254		Yes	Yes	No	2008	N1	Blood	VSE 2008	JANEOJ000000000	SAMN29681772	PRJNA858233
K59-74	209	2830680	201.761	*E. faecium*	18	3045		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOI000000000	SAMN29681773	PRJNA858233
K59-75	211	3022267	218.124	*E. faecium*	203	3071		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOH000000000	SAMN29681774	PRJNA858233
K59-76	170	2794386	251.29	*E. faecium*	18	3046		Yes	No	No	2008	N1	Blood	VSE 2008	JANEOG000000000	SAMN29681775	PRJNA858233
K59-77	241	3065442	166.009	*E. faecium*	78	6239		Yes	Yes	No	2008	N1	Blood	VSE 2008	JANEOF000000000	SAMN29681776	PRJNA858233
K59-78	183	2952048	190.741	*E. faecium*	78	1196		Yes	Yes	No	2008	E1	Blood	VSE 2008	JANEOE000000000	SAMN29681777	PRJNA858233
K59-79	230	3134658	169.021	*E. faecium*	203	3075		Yes	No	No	2008	E1	Blood	VSE 2008	JANEOD000000000	SAMN29681778	PRJNA858233
K59-80	192	3001319	160.087	*E. faecium*	203	3061		Yes	Yes	No	2008	E1	Blood	VSE 2008	JANEOC000000000	SAMN29681779	PRJNA858233
K59-81	190	3000916	203.51	*E. faecium*	203	3061		Yes	Yes	No	2008	E1	Blood	VSE 2008	JANEOB000000000	SAMN29681780	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
K60-01	253	3108434	192.387	*E. faecium*	203	3072		Yes	No	No	2008	E1	Blood	VSE 2008	JANEOA000000000	SAMN29681781	PRJNA858233
K60-02	77	2821274	224.468	*E. faecium*	52	6240		No	No	No	2008	E1	Blood	VSE 2008	JANENZ000000000	SAMN29681782	PRJNA858233
K60-03	149	2887974	286.943	*E. faecium*	202	3076		Yes	Yes	No	2008	E1	Blood	VSE 2008	JANENY000000000	SAMN29681783	PRJNA858233
K60-04	210	3019594	246.668	*E. faecium*	203	3061		Yes	Yes	No	2008	E9	Blood	VSE 2008	JANENX000000000	SAMN29681784	PRJNA858233
K60-05	151	2861958	214.035	*E. faecium*	202	3076		Yes	Yes	No	2008	E9	Blood	VSE 2008	JANENW000000000	SAMN29681785	PRJNA858233
K60-06	247	3067598	173.224	*E. faecium*	203	3067		Yes	No	No	2008	E9	Blood	VSE 2008	JANENV000000000	SAMN29681786	PRJNA858233
K60-07	221	2966068	215.075	*E. faecium*	578	6241		Yes	No	No	2008	E9	Blood	VSE 2008	JANENU000000000	SAMN29681787	PRJNA858233
K60-08	271	3098662	171.317	*E. faecium*	17	3030		Yes	Yes	No	2008	E9	Blood	VSE 2008	JANENT000000000	SAMN29681788	PRJNA858233
K60-09	80	2609025	284.655	*E. faecium*	579	6242		No	No	No	2008	E5	Blood	VSE 2008	JANENS000000000	SAMN29681789	PRJNA858233
K60-10	283	3095658	142.579	*E. faecium*	17	3030		Yes	Yes	No	2008	E5	Blood	VSE 2008	JANENR000000000	SAMN29681790	PRJNA858233
K60-12	226	2914725	131.794	*E. faecium*	192	397		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENQ000000000	SAMN29681791	PRJNA858233
K60-13	211	3025633	190.448	*E. faecium*	17	3031		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENP000000000	SAMN29681792	PRJNA858233
K60-14	198	2911834	25.1478	*E. faecium*	192	397		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENO000000000	SAMN29681793	PRJNA858233
K60-15	207	2854892	156.119	*E. faecium*	18	3048		Yes	No	No	2008	E8	Blood	VSE 2008	JANENN000000000	SAMN29681794	PRJNA858233
K60-16	273	3081760	166.863	*E. faecium*	17	3032		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENM000000000	SAMN29681795	PRJNA858233
K60-17	177	2988494	260.905	*E. faecium*	17	3031		Yes	No	No	2008	E8	Blood	VSE 2008	JANENL000000000	SAMN29681796	PRJNA858233
K60-18	271	3017368	102.239	*E. faecium*	17	3031		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENK000000000	SAMN29681797	PRJNA858233
K60-19	262	3072318	222.897	*E. faecium*	17	3032		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENJ000000000	SAMN29681798	PRJNA858233
K60-20	211	3008138	247.646	*E. faecium*	17	3031		Yes	Yes	No	2008	E8	Blood	VSE 2008	JANENI000000000	SAMN29681799	PRJNA858233
K60-21	72	2451393	189.232	*E. faecium*	580	5402		No	No	No	2008	E8	Blood	VSE 2008	JANENH000000000	SAMN29681800	PRJNA858233
K60-22	157	2823062	238.582	*E. faecium*	202	3076		Yes	Yes	No	2008	E2	Blood	VSE 2008	JANENG000000000	SAMN29681801	PRJNA858233
K60-23	170	2913003	116.848	*E. faecium*	202	3076		Yes	Yes	No	2008	E2	Blood	VSE 2008	JANENF000000000	SAMN29681802	PRJNA858233
K60-24	148	2834838	227.079	*E. faecium*	202	3076		Yes	Yes	No	2008	E2	Blood	VSE 2008	JANENE000000000	SAMN29681803	PRJNA858233
K60-25	79	2499225	284.41	*E. faecium*	581	6243		No	No	No	2008	E2	Blood	VSE 2008	JANEND000000000	SAMN29681804	PRJNA858233
K60-26	175	2918445	262.663	*E. faecium*	202	3076		Yes	Yes	No	2008	E2	Blood	VSE 2008	JANENC000000000	SAMN29681805	PRJNA858233
K60-27	216	2825223	213.358	*E. faecium*	18	3051		Yes	No	No	2008	E2	Blood	VSE 2008	JANENB000000000	SAMN29681806	PRJNA858233
K60-29	223	2818036	217.618	*E. faecium*	19	6245		Yes	No	No	2008	E2	Blood	VSE 2008	JANENA000000000	SAMN29681807	PRJNA858233
K60-30	215	3001303	235.085	*E. faecium*	17	3031		Yes	Yes	No	2008	E2	Blood	VSE 2008	JANEMZ000000000	SAMN29681808	PRJNA858233
K60-31	289	3130377	140.796	*E. faecium*	203	3067		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMY000000000	SAMN29681809	PRJNA858233
K60-32	293	3129397	232.004	*E. faecium*	203	3069		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMX000000000	SAMN29681810	PRJNA858233

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster type	*van* type	Ampicillin restance	Gentamicin restance	Linezolid restance	Year	Hospital	Source	Collection	Genome/GenBank assembly accession	Biosample accession	Bioproject
K60-33	262	3025667	208.432	*E. faecium*	17	3031		Yes	No	No	2008	E3	Blood	VSE 2008	JANEMW000000000	SAMN29681811	PRJNA858233
K60-35	269	3105187	179.244	*E. faecium*	203	3069		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMV000000000	SAMN29681812	PRJNA858233
K60-36	181	2800179	236.47	*E. faecium*	18	3048		Yes	No	No	2008	E3	Blood	VSE 2008	JANEMU000000000	SAMN29681813	PRJNA858233
K60-37	218	3027563	166.295	*E. faecium*	17	3031		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMT000000000	SAMN29681814	PRJNA858233
K60-38	237	3083830	198.345	*E. faecium*	17	3030		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMS000000000	SAMN29681815	PRJNA858233
K60-39	5	2739582	256.0	*E. faecium*	192	397		Yes	Yes	No	2008	E3	Blood	VSE 2008	GCA_002334625.1	SAMN07638053	PRJNA407052
K60-40	164	2887060	100.709	*E. faecium*	202	3076		Yes	Yes	No	2008	E3	Blood	VSE 2008	JANEMR000000000	SAMN29681816	PRJNA858233
K60-42	60	2703026	146.353	*E. faecium*	22	6246		No	No	No	2008	E3	Blood	VSE 2008	JANEMQ000000000	SAMN29681817	PRJNA858233
K60-43	119	2844027	244.561	*E. faecium*	18	388		Yes	No	No	2008	E3	Blood	VSE 2008	JANEMN000000000	SAMN29681818	PRJNA858233
TUH2_18	136	2830096	69.4951	*E. faecium*	17	1709	*vanB*	Yes	No	No	1996	W1	Urine	VRE 1996	JANEMP000000000	SAMN29681819	PRJNA858233
TUH2_19	138	2827369	219.183	*E. faecium*	17	1709	*vanB*	Yes	No	No	1996	W1	Clinical si	VRE 1996	JANEMO000000000	SAMN29681820	PRJNA858233

Supplement file 1. Assembly quality, metadata as well as repository numbers for each *E. faecalis* sample

ID	Number of contigs	Genome size	Coverage	Species	ST	Cluster Type	*van* type	Ampicillin resistance	Gentamicin resistance	Linezolid resistance	Year	Hospital	Source	Collection	Genome accession	Biosample accession	Bioproject
51269076	147	3263574	205.059	*E. faecalis*	6	1160	*vanB*	No	Yes	No	2013	N2	Urine	VRE 2010-15	JANEVW000000000	SAMN29681576	PRJNA858233
51269077	141	3131613	20.672	*E. faecalis*	6	1160	*vanB*	No	Yes	No	2013	N2	Urine	VRE 2010-15	JANEVV000000000	SAMN29681577	PRJNA858233
51269078	165	3237561	194.858	*E. faecalis*	6	1159	*vanB*	No	Yes	No	2015	N2	Clinical site other materials	VRE 2010-15	JANEVU000000000	SAMN29681578	PRJNA858233
51269079	136	3255678	220.302	*E. faecalis*	6	1160	*vanB*	No	Yes	No	2015	N2	Urine	VRE 2010-15	JANEVT000000000	SAMN29681579	PRJNA858233
51271223	149	3079089	37.3833	*E. faecalis*	28	1162	*vanB*	No	No	No	2012	E10	Clinical site other materials	VRE 2010-15	JANEVN000000000	SAMN29681585	PRJNA858233
51271226	72	3104404	191.121	*E. faecalis*	28	1162	*vanB*	No	No	No	2013	E10	Urine	VRE 2010-15	JANEVM000000000	SAMN29681586	PRJNA858233
51271924	157	3227596	89.6763	*E. faecalis*	6	107	*vanB*	No	Yes	No	2015	E6	Urine	VRE 2010-15	JANHGF000000000	SAMN29884051	PRJNA858233
51271925	97	3156253	61.2812	*E. faecalis*	6	1164	*vanB*	No	Yes	No	2015	E6	Feces	VRE 2010-15	JANEVK000000000	SAMN29681588	PRJNA858233
51271926	146	3243803	107.093	*E. faecalis*	6	107	*vanB*	No	Yes	No	2015	E6	Clinical site other materials	VRE 2010-15	JANEYI000000000	SAMN29681589	PRJNA858233
51273466	136	3219235	119.938	*E. faecalis*	6	107	*vanB*	No	Yes	No	2011	W1	Clinical site other materials	VRE 2010-15	JANETW000000000	SAMN29681629	PRJNA858233
51296112	105	3324700	246.167	*E. faecalis*	6	107	*vanB*	No	Yes	No	2010	E10	Urine	VRE 2010-15	JANEQJ000000000	SAMN29681720	PRJNA858233
51296113	153	3226181	95.7768	*E. faecalis*	6	1167	*vanB*	No	Yes	No	2012	E10	Urine	VRE 2010-15	JANEQI000000000	SAMN29681721	PRJNA858233

Supplement file 2. Table with 30 experimentally confirmed VF genes in *E. faecium* including a short description of the VFs and the accession number of the sequence that was used to build our VF database.

	Virulence factor genes	Alternative gene names	DNA tag	Function	Reference
1	$atlA_{Efm}$		CP003583.1:c2230282-2228117	Major autolysin, biofilm formation	1
2	*acm*	*fms8*	ABQJ01000138.1:27210-29228	Collagen binding adhesin, MSCRAMM, similar to Ace in *E. faecalis*	2, 17
3	*bepA*	*fruA*	ABQJ01000011.1 :42962-44380	Biofilm and endocarditis-associated permease A (PTS associated)	4
4	*boNT*/En		NGLI01000004.1:167397-171236	Botulinum neurotoxin-like toxin	20
5	*capD*		CP003583.1:892837-893844	Capsular polysaccharide biosynthesis protein, adhesion, avoid opsonic killing	9, 19
6	*ccpA*		AEBU01000039.1:1370-2389	Catabolite control protein A, growth, virulence	3
7	*ecbA*	*orf2430*	ABQJ01000112.1 :5675-8902	Collagen binding MSCRAMMM, adhesion	8
8	*empABC*	*empA* (pilin) previously $ebpA_{fm}$ and *fms1*	AAAK03000002.1 :97025-100414	Pili, biofilm formation (mainly EmpA) and adherence to ECM proteins (EmpAB), reduced virulence in murine UIT (EmpABC) and infective endocarditis (EmpA) models	22, 11, 6
9		*EempB* (pilin) previously $ebpB_{fm}$ and *fms5*	AAAK03000002.1 :95601-97022		
10		*empC* (major pilus subunit) previously $ebpC_{fm}$, *pilB* and *fms9*	AAAK03000002.1 :93727-95604		
11	*epx2*		LGAN01000048.1:c51744-50740	Cytotoxic pore-forming toxin, preferred receptor human leukocyte antigen class I (HLA-I) complex	26
12	*esp*	*espfm*, *esp*(fm)	ABQJ01000139.1:70993-76920	Enterococcal surface protein, biofilm formation	5, 23

13	*fms15*		AAAK03000002.1:24432-25460	Adhesin; E. faecium surface protein of the MSCRAMM family	6, 21
14	*fnm*		AAAK03000054.1 :14273-15979	Fibrionectin binding protein, matrix adhesion	7
15 - 16	General stress proteins (*gls*)	*gls33* and *gls20* homologous to *gls24* of *E. faecalis*	CP003583.1 :1462262-1462819 CP003583.1 :1467075-1467977	Mutants lacking both *gls33-glsB*, *gls20-glsB1* or both show increased sensitivity for bile salts, maybe important for adaptation to the intestinal environment in addition to virulence. Clade B isolates also contain these loci but with lower identity (93-97%)	25
17-18		*glsB* and *glsB1* homologous to *glsB* of *E. faecalis*	AY548799.1 :2876-3118 CP003583.1 :1461988-1462230		
19-21	*lysM*-containing proteins	*lysM1* *lysM2* *lysM3*	CP003351.1:1080180-1080818 CP003351.1:509550-510158 CP003351.1 :1177130-1177753	Tissue adhesion	18
22		*lysM4*	CP003351.1 :1237718-1238353	Host colonization, role in peptidoglycan synthesis	18
23	*pilA2*	fms21	ABSW01000038.1:25993-27969	Pilus subunit protein A: initial adherence?	10
24	*prpA*		ABQJ01000017.1 :31606-32775	Prolin rich protein A, binding to the extracellular matrix proteins fibrinogen and fibronectin	13
25	*ptsD*	*pts*_clin	ABQJ01000092.1 :14293-15114	Phosphotransferase system subunit IID, intestinal colonization determinant during antibiotic treatment	12
26	*sagA*		AF242196.1 :1549-3123	Secreted Antigen A, biofilm formation	14, 15
27	*scm*	*fms10*	CP003583.1:2656348-2658159	Collagen adhesion, MSCRAMM	6
28	*sgrA*	*orf2351*	ABQJ01000055.1:3784-4758	Nidogen-binding surface adhesin implicated in biofilm formation	8
29	*tirE1*		Z_ABQJ01000097.1:c6111-5626	TIR-domain containing protein, promotes survival in blood	16
30	*tirE2*		NZ_ABQJ01000097.1:c8433-7582	TIR-domain containing protein, promotes survival in blood	16

References

1. Paganelli, F. L. *et al.* Enterococcus faecium Biofilm Formation: Identification of Major Autolysin AtlAEfm, Associated Acm Surface Localization, and AtlAEfm-Independent Extracellular DNA Release. *mBio* **4,** (2013). DOI: 10.1128/mBio.00154-13
2. Nallapareddy, S. R., Singh, K. V. & Murray, B. E. Construction of improved temperature-sensitive and mobilizable vectors and their use for constructing mutations in the adhesin-encoding acm gene of poorly transformable clinical Enterococcus faecium strains. *Appl. Environ. Microbiol.* **72,** 334–345 (2006). DOI: 10.1128/AEM.72.1.334-345.2006
3. Somarajan, S. R., Roh, J. H., Singh, K. V., Weinstock, G. M. & Murray, B. E. CcpA Is Important for Growth and Virulence of Enterococcus faecium. *Infect. Immun.* **82,** 3580–3587 (2014). DOI: 10.1128/IAI.01911-14
4. Paganelli, F. L. *et al.* Genome-wide Screening Identifies Phosphotransferase System Permease BepA to Be Involved in Enterococcus faecium Endocarditis and Biofilm Formation. *J. Infect. Dis.* **214,** 189–195 (2016). DOI: 10.1093/infdis/jiw108
5. Willems, R. J. *et al.* Variant esp gene as a marker of a distinct genetic lineage of vancomycin resistant Enterococcus faecium spreading in hospitals. *The Lancet* **357,** 853–855 (2001). DOI: 10.1016/S0140-6736(00)04205-7
6. Sillanpää, J. *et al.* 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. *Microbiol. Read. Engl.* **154,** 3199–3211 (2008). DOI: 10.1099/mic.0.2008/017319-0
7. Somarajan, S. R. *et al.* The fibronectin-binding protein Fnm contributes to adherence to extracellular matrix components and virulence of Enterococcus faecium. *Infect. Immun.* **83,** 4653–4661 (2015). DOI: 10.1128/IAI.00885-15
8. Hendrickx, A. P. A. *et al.* 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. *Infect. Immun.* **77,** 5097–5106 (2009). DOI: 10.1128/IAI.00275-09
9. Ali, L. *et al.* Identification and functional characterization of the putative polysaccharide biosynthesis protein (CapD) of Enterococcus faecium U0317. *Infect. Genet. Evol.* **37,** 215–224 (2016). DOI: 10.1016/j.meegid.2015.11.020
10. Hendrickx, A. P. A. *et al.* Expression of two distinct types of pili by a hospital-acquired Enterococcus faecium isolate. *Microbiology* **154,** 3212–3223 (2008). DOI: 10.1099/mic.0.2008/020891-0
11. Sillanpää, J. *et al.* Characterization of the ebpfm 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 (2010). DOI: 10.4161/viru.1.4.11966
12. Zhang, X. *et al.* Identification of a genetic determinant in clinical Enterococcus faecium strains that contributes to intestinal colonization during antibiotic treatment. J. Infect. Dis. 207, 1780–1786 (2013). DOI: 10.1093/infdis/jit076
13. Prieto, A. M. G. *et al.* The N-terminal domain of the thermo-regulated surface protein PrpA of Enterococcus faecium binds to fibrinogen, fibronectin and platelets. Sci. Rep. 5, srep18255 (2015). DOI: 10.1038/srep18255
14. Rangan, K. J. *et al.* A secreted bacterial peptidoglycan hydrolase enhances tolerance to enteric pathogens. Science 353, 1434-1437 (2016). DOI: 10.1126/science.aaf3552
15. Paganelli, F. L. *et al.* Distinct SagA from Hospital-Associated Clade A1 Enterococcus faecium Strains Contributes to Biofilm Formation. Appl. Environ. Microbiol. 81, 6873–6882 (2015). DOI: 10.1128/AEM.01716-15
16. Wagner, T. M. *et al.* Enterococcus faecium TIR-domain genes are part of a gene dluster which promotes bacterial survival in blood. Int J Microbiol ID 1435820 (2018). DOI: 10.1155/2018/1435820
17. Nallapareddy, S. R. *et al.* Clinical isolates of Enterococcus faecium exhibit strain-specific collagen binding mediated by Acm, a new member of the MSCRAMM family. Mol Microbiol 47, 1733-1747 (2003). DOI: 10.1046/j.1365-2958.2003.03417.x
18. Cacaci, M. *et al*. Expression profiling in a mammalian host reveals the strong induction of genes encoding LysM domain-containing proteins in Enterococcus faecium. Sci Rep 8, 12412 (2018). DOI: 10.1038/s41598-018-30882-z

19. Ali, L. *et al.* Detection and characterization of bacterial polysaccharides in drug-resistant enterococci. Glycoconjugate Journal 36, 429-38 (2019). DOI: 10.1007/s10719-019-09881-3
20. Zhang *et al.* Identification of a botulinum neurotoxin-like toxin in a commensal strain of Enterococcus faecium. Cell Host Microbe 23, 169-76 (2018). DOI: 10.1016/j.chom.2017.12.018
21. Revtovich *et al*. Development and Characterization of High-Throughput *Caenorhabditis elegans - Enterococcus faecium* Infection Model. Front Cell Infect Microbiol. 2021 Apr 29;11:667327. DOI: 10.3389/fcimb.2021.667327
22. Montealegre MC *et al.* Role of the Emp Pilus Subunits of Enterococcus faecium in Biofilm Formation, Adherence to Host Extracellular Matrix Components, and Experimental Infection. Infect Immun. 2016 Apr 22;84(5):1491-1500. DOI: 10.1128/IAI.01396-15
23. Heikens E *et al.* Enterococcal surface protein Esp is important for biofilm formation of Enterococcus faecium E1162. J Bacteriol. 2007 Nov;189(22):8233-40. DOI: 10.1128/JB.01205-07
25. Choudhury et al. Importance of two Enterococcus faecium loci encoding Gls-like proteins for in vitro bile salts stress response and virulence. J Infect Dis. 2011 Apr 15;203(8):1147-54. DOI: 10.1093/infdis/jiq160
26. Xiong et al. Emerging enterococcus pore-forming toxins with MHC/HLA-I as receptors. Cell. 2022 Mar 31;185(7):1157-1171.e22. doi: 10.1016/j.cell.2022.02.002.

Supplement file 3. Prevalence of the Norwegian *E. faecium* cluster types (CT), strain collection, *van*-types, and geographical region.

CT	Number of isolates	Strain collection	*van*-type	Geographical region
ST192-CT3/26	113	VRE 2010-15	*vanB* (n=113)	West
ST117-CT24	51	VRE 2010-15, VSE 2014	*vanB* (n=31)	VRE; West, East, Midle VSE; All health regions
ST203-CT3061	25	All collections	*vanB* (n=3)	VRE; West, East VSE; All health regions
ST80-CT16	23	VRE 2010-15, VSE 2014	*vanB* (n=1)	West, North, East
ST203-CT20	19	VRE 2010-15	*vanA* (n=19)	East, Midle, North
ST203-CT3056	12	VSE 2014		East, West
ST80-CT3097	10	VRE 2010-15	*vanA* (n=10)	East
ST203-CT3067	9	VSE 2008		East, West, Midle
ST203-CT3062	8	VSE 2014, VSE 2008		Middle, East
ST17-CT3031	8	VSE 2008		East
ST202-CT3076	8	VSE 2008		East
ST203-CT191	7	VRE 2010-15, VSE 2014	*vanA* (n=6)	East, West
ST17-CT3037	7	VRE 2010-15, VSE 2014	*vanA* (n=2)	East, North
ST192-CT3081	6	VSE 2014		East
ST117-CT3054	5	VSE 2014		East, North
ST203-CT3069	4	VSE 2008		East, North
ST192-CT188	4	VRE 2010-15	*vanA* (n=4)	North
ST17-CT3030	4	VSE 2008		East, West
ST192-CT1217	4	VSE 2008, VSE 2014		East, Midle, North
ST192_CT397	4	VSE 2008		East
ST17-CT1709	2	VRE 1996	*vanB* (n=2)	West
ST192-CT3080	2	VRE 2010-15	*vanB* (n=2)	West, East
ST192-CT3082	2	VRE 2010-15	*vanB* (n=2)	West
ST17-CT159	2	VRE 2010-15	*vanB* (n=2)	East
ST18-CT3042	2	VRE 2010-15	*vanA* (n=2)	East
Singleton VRE*fms* and CTs with ≤ 3 isolates	149	All collections	*vanA* (n=19) *vanB* (n=11)	All health regions

Supplement file 4. VF gene profiles of all *E. faecium* (n=490) in this study

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51268383	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51268385	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51268386	0	1	0	1	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51269029	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51269051	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269053	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269054	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269055	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	0	0	0	1	1	0	0
51269056	1	1	1	0	1	0	0	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269057	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51269058	0	1	1	0	1	0	0	1	1	1	1	1	0	1	1	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0
51269059	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269060	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269061	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51269062	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	1	0	0	1	1	0	0
51269063	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	1	0	0	1	1	0	0
51269064	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269065	0	1	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51269066	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51269067	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269068	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269069	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51269070	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269071	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51269072	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269073	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269075	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51269769	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269770	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51269771	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51269772	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
51269773	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	1	1	1	1	0	0
51269774	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
51269775	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51269776	0	1	1	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0
51269777	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51269778	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0
51269779	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269928	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51269929	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51269930	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269931	1	1	1	0	1	0	0	0	1	1	1	1	1	0	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51269932	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51269933	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51269934	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51269935	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51269936	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51269937	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51269938	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51269939	1	1	1	0	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270240	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270243	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51270244	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270271	0	1	1	0	1	0	0	0	0	1	1	0	0	1	0	0	1	1	1	0	0	1	0	1	0	0	1	1	0	0
51270272	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270273	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270274	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270275	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270276	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51270277	1	1	1	1	1	1	1	1	0	0	0	0	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51270809	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51270810	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270811	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51270812	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51270813	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51270814	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51270815	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51270816	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51270817	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270818	0	1	1	0	1	1	0	0	1	1	1	1	0	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270819	1	1	1	1	1	0	1	0	1	1	1	1	0	0	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51270820	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51270821	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270822	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51270823	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51270824	0	1	1	0	1	0	0	1	1	1	1	1	0	1	1	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0
51270825	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270826	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51270828	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271041	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271042	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271044	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271045	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271046	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271047	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271048	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271049	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271050	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	0	1	1	1	1	0	0	1	1	0	0
51271051	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271053	1	1	1	0	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271054	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271055	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51271056	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271057	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271164	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271165	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271166	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271167	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271168	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271169	1	1	0	1	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271170	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271171	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271172	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271173	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271174	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271175	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271176	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271177	1	1	0	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271178	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271179	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271180	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271181	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0
51271182	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271183	1	1	1	0	1	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271184	0	1	0	0	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	0	0	0	0	0	0	1	1	0	0
51271185	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271186	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271187	1	1	1	0	1	1	1	0	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271188	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271189	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271190	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271191	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271192	1	1	1	1	1	1	1	0	0	0	0	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0
51271193	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271194	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271195	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271196	0	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0
51271197	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271198	0	1	0	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271199	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271200	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51271201	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271208	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271210	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271211	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271212	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271213	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271214	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271215	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271216	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271217	1	1	1	1	1	1	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0
51271218	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271219	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271220	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271221	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271224	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0
51271225	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271227	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271228	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271229	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271509	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271510	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0
51271511	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	0	1	1	1	0	0	1	1	0	0	0	1	1	0	0
51271512	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271513	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271514	1	1	1	0	1	1	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271515	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271516	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271517	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	0	0	1	1	0	0
51271518	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271519	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271520	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271521	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271522	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271523	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271524	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271815	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271816	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271817	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271818	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271819	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271820	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271821	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51271822	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271823	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271824	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271825	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271826	1	1	1	0	1	1	1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271880	1	1	1	0	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271881	1	1	1	0	1	1	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271882	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	0	0	1	1	0	0
51271883	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271884	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271885	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271886	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51271887	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271888	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271889	1	1	0	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271890	0	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0
51271891	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271892	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271893	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271894	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271895	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	1	1	1	1	0	0
51271896	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0
51271898	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51271900	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271901	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271902	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
51271903	1	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271904	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271919	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271920	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271921	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271922	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271923	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0
51271927	1	1	1	1	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271928	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271929	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271930	1	1	1	0	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271931	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	0	0	1	1	0	0
51271932	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271933	1	1	1	0	1	0	1	1	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
51271934	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271935	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271936	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271937	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271938	1	1	1	0	1	1	0	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271939	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271940	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271942	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271943	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271944	1	1	1	1	1	1	0	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271945	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271946	1	1	1	0	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51271978	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271982	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271984	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51271985	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271986	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271987	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271988	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271989	1	1	1	0	0	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271990	1	0	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271991	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271992	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271993	1	1	1	0	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271994	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271995	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51271996	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271997	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51271998	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51271999	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51272000	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51272001	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51272002	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51272003	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51272004	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51273071	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273073	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273074	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273075	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51273076	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273077	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51273078	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273079	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273080	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273081	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273082	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273083	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273084	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273085	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273086	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273087	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273088	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273089	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273090	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273091	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273092	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51273093	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273094	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273095	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273096	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273097	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273098	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273099	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273100	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273101	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273102	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273103	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273104	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273105	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273106.ref	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273107	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273108	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273451	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273452	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51273453	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273454	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273455	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273456	1	1	1	0	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273457	1	1	1	0	1	0	1	1	0	0	0	1	0	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	0	0
51273458	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273459	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273460	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273461	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51273462	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273464	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51273465	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51273467	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273468	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51273469	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273470	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273471	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273472	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273473	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273474	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51273475	0	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51273476	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51273477	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51273478	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273479	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0
51273480	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273481	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273482	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273483	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273484	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51273485	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273486	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273487	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273488	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273489	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273490	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273491	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0
51273492	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273493	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273494	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273495	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51273496	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51273497	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273498	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51273553	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51273875	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51274612	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51274613	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274614	0	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274615	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274616	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274617	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274618	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274619	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274620	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274621	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274622	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274623	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51274625	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274626	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274627	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274628	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274629	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274630	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274631	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274632	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274633	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274634	0	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	0	1	1	0	0
51274635	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274636	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	0	0
51274637	1	1	1	0	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	0	1	1	1	0	0
51274638	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
51274639	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274640	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274641	1	1	1	0	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
51274642	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51274643	1	1	1	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276488	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276490	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276491	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276492	1	1	1	0	1	0	1	0	1	1	1	1	0	0	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51276494	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276495	1	1	1	1	1	1	1	0	0	0	0	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276496	1	1	1	1	1	0	1	0	0	0	0	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276497	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51276498	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	0	0	0	1	1	1	1	1	1	1	1	0	0
51276499	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51276500	1	1	1	0	1	0	0	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276501	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51276502	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276507	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
51276508	1	1	1	0	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276509	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51276510	1	1	1	0	1	0	0	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
51276511	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276512	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51276513	1	1	1	0	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
51276514	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0
51280801	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51280802	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51296109	1	1	1	0	1	1	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
51296110	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
K59_16	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_17	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
K59_18	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K59_19	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	0	1	1	1	0	0	1	0	1	0	0	1	1	0	0
K59_20	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_21	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_22	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_23	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_25	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_26	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0
K59_27	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_28	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	0	0	0	0	1	1	0	1	0	0	1	1	0	0
K59_29	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	1	1	1	0	1	1	1	1	1	0	0
K59_30	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_31	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_32	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_33	1	1	1	1	1	1	1	1	1	0	0	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_34	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_35	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	0	1	1	0	0
K59_36	1	1	1	0	1	0	1	0	0	0	0	0	0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
K59_37	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_40	1	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_41	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_42	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	0	0
K59_43	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_44	0	1	1	0	1	0	0	1	1	1	1	1	1	1	0	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
K59_46	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	0	0	1	1	0	0
K59_48	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	0	0	0	0	0	1	0	1	0	0	1	1	0	0
K59_49	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
K59_50	1	1	1	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
K59_51	1	1	1	0	1	0	1	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_52	0	1	1	0	1	0	0	0	1	1	1	1	0	1	1	1	1	1	1	0	0	1	0	0	0	0	1	1	0	0
K59_53	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	0	0	0	1	1	0	0
K59_54	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_55	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0	0	1	1	0	0
K59_56	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_57	0	1	1	0	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	1	0	0	1	1	1	1	0	0
K59_58	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_59	1	1	1	1	1	1	1	1	0	0	0	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	0	0
K59_60	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_62	1	1	1	0	1	1	0	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_63	1	1	1	1	1	0	1	0	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_64	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_65	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_66	1	1	1	0	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K59_67	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_68	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
K59_69	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_70	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K59_71	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_72	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_73	1	1	1	1	1	0	0	0	1	1	1	1	0	1	1	0	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_74	1	1	1	0	1	0	0	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K59_75	1	1	1	1	1	1	1	1	0	0	0	0	0	1	0	1	0	0	0	1	1	1	0	1	1	1	1	1	0	0
K59_76	1	1	1	0	1	0	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K59_77	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_78	1	1	1	1	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_79	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K59_80	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K59_81	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
K60_01	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_02	0	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	0	0	1	1	0	0
K60_03	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_04	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_05	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_06	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_07	1	1	1	1	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_08	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_09	0	1	1	0	1	0	0	0	1	0	0	1	0	1	0	0	1	1	1	0	0	1	0	0	0	0	1	1	0	0
K60_10	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_12	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	0	1	1	0	0
K60_13	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_14	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	0
K60_15	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_16	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K60_17	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_18	1	0	1	1	1	0	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_19	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K60_20	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_21	0	1	1	0	1	0	0	0	1	1	1	1	0	1	0	0	0	0	0	0	0	1	0	0	0	0	1	1	0	0
K60_22	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_23	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K60_24	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_25	0	1	1	0	1	0	0	0	0	0	0	0	0	1	1	0	1	1	1	1	0	0	0	0	0	0	1	1	0	0
K60_26	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_27	1	1	1	0	1	0	1	0	1	1	1	1	1	1	0	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K60_29	1	1	1	0	1	0	0	1	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	0	0	0	1	1	0	0
K60_30	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_31	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_32	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0

ID	*acm*	*atlAEfm*	*bepA*	*capD*	*ccpA*	*ecbA*	*esp*	*fms15*	*gls20*	*gls33*	*glsB*	*glsB1*	*lysM1*	*lysM2*	*lysM3*	*pilA2*	*empA*	*empB*	*empC*	*prpA*	*ptsD*	*sagA*	*scm (both alleles)*	*sgrA*	*tirE1*	*tirE2*	*fnm*	*lysM4*	*bont/En*	*epx2*
K60_33	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_35	1	1	1	1	1	1	1	0	1	1	1	1	0	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_36	1	1	1	0	1	0	1	0	1	1	1	1	0	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_37	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_38	1	1	1	1	1	1	1	0	1	1	1	1	1	1	0	0	1	1	1	1	1	1	0	1	1	1	1	1	0	0
K60_39	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	0	1	1	1	1	1	1	1	1	1	1	1	1	0	0
K60_40	1	1	1	1	1	0	1	0	1	1	1	1	1	1	1	1	1	1	1	1	1	1	0	1	0	0	1	1	0	0
K60_42	0	0	1	0	1	0	0	0	1	1	1	1	1	1	0	0	1	1	1	0	1	1	0	1	0	0	1	1	0	0
K60_43	1	1	1	1	1	0	0	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	0	1	0	0	1	1	0	0
TUH2_18	1	1	1	0	1	0	0	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0
TUH2_19	1	1	1	0	1	0	0	0	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	0	1	1	0	0