Elucidating vancomycin-*resistant Enterococcus faecium* outbreaks: the role of clonal spread and movement of mobile genetic elements.

7

X. Zhou[1], M.A. Chlebowicz[1], E. Bathoorn[1], S. Rosema[1], N. Couto[1], M. Lokate[1], J.P. Arends[1], A.W. Friedrich[1], J.W.A. Rossen[1*]

[1] University of Groningen, University Medical Center Groningen, Department of Medical Microbiology, The Netherlands

Short title: Transposons on the move in VREfm outbreaks

***Corresponding author**: John W.A. Rossen; Adress: Hanzeplein 1 EB80, 9713GZ Groningen, the Netherlands. Tel: +31 50 3613480; Fax: +31 50 3619105; Email: j.w.a.rossen@rug.nl

ABSTRACT

Introduction:

Vancomycin resistant *Enterococcus faecium* (VREfm) has emerged as a nosocomial pathogen worldwide. The dissemination of VREfm is due to both clonal spread and spread of mobile genetic elements (MGEs) such as transposons. We aimed to combine *vanB*-carrying transposons characteristics with core-genome MLST (cgMLST) typing and epidemiological data to understand the pathways of transmission in nosocomial outbreaks.

Methods:

Retrospectively, 36 VREfm isolates obtained from 34 patients from seven VREfm outbreak investigations in 2014 were analysed. Isolates were sequenced on a MiSeq and a MinION instrument. *De novo* assembly was performed in CLC Genomics Workbench, the hybrid assemblies were obtained through Unicycler v0.4.1. Ridom SeqSphere+ was used to extract MLST and cgMLST data. Detailed analysis of each transposon and their integration points were performed using Artemis Comparison Tool (ACT) and multiple blast analyses.

Results

Four different *vanB* transposons were found among the isolates. CgMLST divided ST80 isolates into three CTs; CT16, CT104 and CT106. ST117 isolates were divided into CT24, CT103 and CT105. Within VREfm isolates belonging to CT103, two different *vanB* transposons were found. In contrast, VREfm isolates belonging to CT104 and CT106 harboured an identical *vanB* transposon.

Conclusion:

CgMLST provides a high discriminatory power for the epidemiological analysis of VREfm. However, additional transposon analysis is needed to detect horizontal gene transfer. Combining these two methods allows to investigate both clonal spread as well as the spread of MGEs. This leads to new insights and thereby better understanding of the complex transmission routes in VREfm outbreaks.

INTRODUCTION

Enterococcus faecium has emerged as a nosocomial pathogen worldwide. Vancomycin resistant *E. faecium* (VREfm) outbreaks are mainly caused by successful hospital-associated (HA) *E. faecium* isolates that acquired the *vanA* or *vanB* gene [1]. The dissemination of VREfm is the result of both clonal spread of successful clones, mainly ST17, ST18 and ST78 [2] and the exchange of mobile genetic elements (MGEs) such as chromosomal fragments [3] and plasmids [1, 4, 5]. The *vanA* gene is part of an operon of seven genes, carried by the Tn*1546* transposon, which can be located on various plasmid types or can be integrated into the chromosome [6, 7]. Similarly to *vanA, vanB* is also a part of an operon that consists of seven genes, generally located on the conjugative transposon Tn*1549*/Tn*5382*. Like *Tn1546*, this transposon can also be located on various types of plasmids or can be integrated into the chromosome [1, 4].

In our hospital, we mainly find *vanB* VREfm. Successful HA vancomycin susceptible *E. faecium* (VSEfm) lineages may acquire the *vanB* gene by different pathways. It can occur by *de novo* acquisition of *Tn1549* from anaerobic gut microbiota [8]. Another mechanism is through the exchange of large chromosomal fragments, including Tn*1549*, between *vanB* VREfm and VSEfm [3].

In outbreak investigations, rapid and accurate typing is required to investigate the genetic relatedness between patients' isolates. This information is essential to demonstrate nosocomial transmission. Till 2014, most of VREfm isolates in our hospital were typed by multi-locus variable-number tandem repeat analysis (MLVA). MLVA is an easy, fast and highly reproducible method to type VREfm [9], but not discriminatory enough in outbreak investigations. MLST is a key tool to study the genetic relatedness and epidemiology of *E. faecium* isolates [10]. However, the discriminatory power of MLST is also insufficient in nosocomial outbreak investigations [11]. In addition to the inferior discriminatory power, MLST-based typing may be unreliable due to recombination events in the MLST loci, which can cause a high number of discrepancies between WGS based typing and MLST [8, 12, 13].

In 2014, WGS was implemented in our laboratory for outbreak investigations of multi-drug resistant microorganisms, including VREfm [14]. The challenge of using WGS is to rapidly analyse and interpret the relevant information [15, 16]. In 2015, a core genome (cg)MLST scheme (consisting 1.423 target genes) for *E. faecium* was developed [17]. This gene-by-gene typing based approach uses a defined set of genes to extract an allele-based profile which makes it scalable and comparable between laboratories. However, cgMLST may also be misleading if horizontal transfer of a single *vanB*-carrying transposon occurs between different *E. faecium* clones during a VREfm outbreak event.

In this study, we retrospectively analysed available draft genome sequences of VREfm isolates from several outbreaks in 2014 in our region and investigated relevant epidemiological data. Next, a detailed characterisation of *vanB*-carrying transposons was performed to determine possible horizontal gene transfer. Hereby spread by clonal expansion as well as by horizontal gene transfer are studied.

METHODS

Study population and infection control protocols

We retrospectively analysed VREfm outbreaks that occurred in the University Medical Center Groningen, the Netherlands in 2014. In 2014, 75 new patients with VREfm were detected. Microbiological data and infection records were used. Infection records included epidemiological information about positive VRE patients. Epidemiological data included dates of when patients were found to be positive, ward and room numbers, patient transfer data and microbiological typing data. We also made use of an epidemiological program to visualize and analyze patient transfers in more detail over several wards and rooms in time by using bed occupancy databases. Herein multiple patients and wards could be included. From 2014 on, concurrent VRE outbreaks have arisen, as experienced by many hospitals in the Netherlands.

By protocol, we screen the following patients for VRE upon admission: patients who have been admitted in a hospital abroad within the past year; patients directly transferred from another hospital in the Netherlands; patients who are admitted to the intensive care and haematology wards; and adopted children. In the Netherlands, it is recommended to screen adopted children for MRSA, as they are frequently from countries that are highly endemic for MRSA. We have chosen to extend the screening in adopted children, by screening for all highly-resistant microorganisms (HRMOs), including VRE. Patients previously known to carry VRE of which the last positive VRE culture was less than one year ago, are treated in contact isolation and additional rectal swabs are taken for VRE screening. At least five rectal swabs are needed to discard the isolation measures in VRE positive patients and those that were known to carry VRE less than one year ago. In patients previously known to carry VRE more than more year ago are treated in contact isolation, unless one of more negative previous VRE cultures were registered. Additional one rectal swab is taken for VRE screening. If this is negative, isolation measures can be discarded. Patients carrying VRE are treated in contact isolation in a single room, using a disposable gown and gloves by the

personnel. Screening of contact patients is performed if there has been exposure of other patients in the same room, or if nosocomial acquisition of VRE is suspected. Since not all patients in our hospital are routinely screened, nosocomial acquisition (e.g. >48 hours) is difficult to define. However, in cases of VRE positive patients who were previously screened VRE negative and in situations of ongoing VRE spread, this is considered as nosocomial acquisition. Screening of contact patients is performed as follows: first, (ex-) roommates of the VRE-positive patient will be screened. If there are one or more VRE-positive contact patients, all patients at the ward and if relevant, ex-patients that have stayed in the affected ward will be screened. The screening is repeated until no new positive VRE patients are detected in at least three rounds of screening, whereas at least 48 hours between each screening round is required. On average, the last screening round will be seven days after (possible) exposure since transmission and subsequent rectal colonization takes time [18].

VRE culturing

VRE culturing was preceded by PCR-screening as described previously [19]. In short, rectal swabs were inoculated in enrichment broth. After 24 hours incubation, a *vanA/vanB* PCR (Xpert®vanA/vanB, Cepheid) was performed on a GeneXpert® XVI (Cepheid) and when positive, the broth was subcultured on VRE Brilliance agars (Oxoid®). Agars were incubated for 24-48 hours and identification and antibiotic susceptibility testing were performed on suspected colonies by MALDI-TOF Mass Spectrometry (Bruker) and VITEK®2 (Biomerieux), respectively. Additionally, we used vancomycin disk diffusion since this method is more sensitive in detecting enterococci isolates with low- and medium-level vanB-type vancomycin resistance [20]. Moreover, identified *E. faecium* isolates were again genotypically tested for the presence of *vanA* and *vanB* genes by PCR using the Xpert®vanA/vanB assay.

Standard, all first VREfm isolates of each patient were typed by MLVA, according to the method described by Top *et al.* [9]. In some cases, e.g. patients that were infected as well as colonized by VRE or harbouring *vanA* as well as *vanB* VRE, multiple VRE isolates were typed. In 2014, we started to implement WGS for VREfm outbreak investigations. In this implementation phase, only a representative subset of isolates that were typed by MLVA were selected for WGS and typed by cgMLST.

WGS and typing methods

Genomic DNA was extracted using the Ultraclean Microbial DNA Isolation Kit (MO BIO Laboratories, Carlsbad, CA, US) following the manufacturer instructions. The DNA concentration and

purity were measured by the Qubit dsDNA HS and BR assay kit (Life technologies, Carlsbad, CA, US). A DNA library was prepared using the Nextera XT v2 kit (Illumina, San Diego, CA, USA) and then run on a MiSeq sequencer (Illumina) for generating paired-end 250-bp reads. De novo assembly was performed by CLC Genomics Workbench v7.0.4 (QIAGEN, Hilden, Germany) after quality trimming (Qs ≥ 20) with optimal word sizes. All procedures were performed as previously described [21]. For the long-read sequencing, libraries of samples A13, A16, A20 and A22 were prepared without shearing to maximize sequencing read length. Samples were barcoded with the Native Barcoding Kit 1D (EXP-NBD103) and libraries were prepared using the Ligation Sequencing Kit 1D (SQK-LSK108). The library was loaded onto an FLO-MIN106 R9.4 flow cell and ran on a MinION device (48 hours). Base calling was performed using Albacore v1.2.2. Data quality was analyzed through Poretools v0.6.0. [22]. Hybrid assemblies were performed using Unicycler v0.4.1. [23]. Bandage v0.8.1 [24] was used to visualize the assembly graphics. Genes of interest were detected using ResFinder.

MLST STs and cgMLST CTs were extracted from the draft genomic sequences using SeqSphere+ version 3.0.1 (Ridom GmbH, Münster, Germany). For the cgMLST analysis, Seqsphere+ used the *E. faecium* scheme published previously [17], considering a cluster alert distance of 20 different alleles. The *vanB*-carrying transposons were identified by BLAST comparisons of *de novo* and hybrid assemblies with the reference sequence of *Tn1549* (GenBank AF192329.1) using the Webact online tool (http://www.webact.org/WebACT/home), [25] under default settings. Detailed analysis of each transposon as well as their integration points were performed using ACT [26] and multiple blast analyses.

Ethics:

The bacterial isolates used for the present analyses were collected in the course of routine diagnostics and infection prevention control. Oral consent for the use of such clinical samples for research purposes is routinely obtained upon patient admission to the UMCG, in accordance with the guidelines of the Medical Ethics Committee of the University Medical Center Groningen. All experiments were performed in accordance with the guidelines of the Declaration of Helsinki and the institutional regulations, and all samples were anonymized.

Nucleotide sequence accession numbers

Sequence data obtained in this study has been deposited at the European Nucleotide Archive (ENA) under BioProject no. PRJEB25590. The hybrid assemblies have been deposited in NCBI under BioProject no PRJNA477347.

RESULTS

Description of outbreak clusters based on epidemiological data

During the implementation period of WGS, 36 representative isolates of 34 patients were sequenced and their draft genome sequences were available for analysis. Based on epidemiological data from infection prevention records, these 34 patients were involved in six outbreak episodes in 7 different wards in 2014. All first VRE isolates of individual cases were assessed, except for two patients, from whom multiple isolates were selected for sequencing (A4 & A4.1 and A22 & A22.1).

Details of the isolates and to which outbreak investigation they belonged are presented in Table 1. Initial outbreak investigations were performed using epidemiological information as described in the methods. Outbreak investigation A took place in April 2014 on ward 1 and 12 patients were involved. Of 10 of these patients, the genome sequences of the obtained isolates (A1 and A4-A13, including A4.1) were available. One of the patients admitted to ward 1 was previously hospitalized in another hospital located in the region. Two isolates (A2 and A3) were therefore obtained from possible contact patients from the other regional hospital and were included in this analysis. Outbreak investigation B took place in July 2014 on ward 1 and 4 patients were involved. Of two patients, the genome sequences of the obtained isolates (A14 and A15) were available. Outbreak investigation C took place in July 2014 on ward 5 and 6 and 10 patients were involved. Of 5 of these patients, genome sequences of the obtained isolates (A16-A20) were available. According to epidemiological data outbreak investigation D took place in November 2014 on ward 7 and involved in total 11 patients. Of 8 of these patients, the genome sequences of the obtained isolates (A21-A28) were available. Also in November, outbreak investigation E took place on ward 2, involving 11 patients. Of 3 of these patients, the genome sequences of the obtained isolates (A29-A31) were available. Finally, outbreak investigation F took place in December 2014 on several wards, involving 7 patients. Of 3 of these patients, the genome sequences of the obtained isolates (A32-A34) were available from a selected ward (ward 4).

Patients A22 and A27 were colonised with *E. faecium* isolates carrying both the *vanA* as well as the *vanB* gene. The *vanA* gene resided on the chromosome, while the *vanB* gene was located on a plasmid. This study will further focus on the *vanB* VREfm and Tn*1549*/Tn*5382* transposon analysis since the rest of the patients were colonised with only *vanB* VREfm.

Table 1: Epidemiological and molecular data of the 36 isolates from 34 patients used in this study.

Sample ID	Outbreak cluster	Month	Ward(s)	Age	Gender	VRE type	Vancomycin MIC mg/L	Isolation source	MT[a] type	ST[b] type	CT[c] type	Sample date	Location *vanB* gene	Trans poson type
A1	A	April	Ward 1	43	M	*vanB*	8	Rectal swab	12	117	24	8/3/2014	Chromosome	1
A2	A	April	*	73	M	*vanB*	8	Rectal swab	12	117	24	16/5/2014	Chromosome	1
À3	A	April	*	76	M	*vanB*	8	Sputum	12	117	24	19/5/2014	Chromosome	1
A4	A	April	Ward 1	65	V	*vanB*	>=32	Bile	12	117	24	1/4/2014	Chromosome	1
A4.1	A	April	Ward 1	65	V	*vanB*	>=32	Rectal swab	12	117	103	13/5/2014	Chromosome	1
A5	A	April	Ward 1	67	V	*vanB*	>=32	Rectal swab	12	117	24	16/4/2014	Chromosome	1
A6	A	April	Ward 1	69	V	*vanB*	>=32	Rectal swab	12	117	24	16/4/2014	Chromosome	1
A7	A	April	Ward 1	59	M	*vanB*	>=32	Rectal swab	12	117	24	21/4/2014	Chromosome	1
A8	A	April	Ward 1	82	V	*vanB*	8	Rectal swab	12	117	24	24/4/2014	Chromosome	1
A9	A	April	Ward 1	59	M	*vanB*	8	Rectal swab	12	80	16	29/4/2014	ND[d]	ND
A10	A	April	Ward 1	69	V	*vanB*	8	Rectal swab	12	117	24	1/5/2014	Chromosome	1
A11	A	April	Ward 1	67	M	*vanB*	>=32	Rectal swab	12	117	24	4/5/2014	Chromosome	1
A12	A	April	Ward 1	77	M	*vanB*	8	Rectal swab	12	117	24	2/4/2014	Chromosome	1
A13	A	April	Ward 1	61	V	*vanB*	>=32	Rectal swab	12	117	24	29/4/2014	Plasmid	2
A14	B	July	Ward 1	61	M	*vanB*	<=0.5	Rectal swab	144	262	60	29/6/2014	ND	ND
A15	B	July	Ward 1	78	M	*vanB*	8	Rectal swab	12	117	103	1/7/2014	Plasmid	2
A16	C	July	Ward 5&6	58	M	*vanB*	>=32	Rectal swab	1	80	104	22/6/2014	Plasmid	4
A17	C	July	Ward 5&6	54	M	*vanB*	8	Rectal swab	1	80	104	28/6/2014	Plasmid	4
A18	C	July	Ward 5&6	49	M	*vanB*	8	Faeces	12	117	103	30/6/2014	Plasmid	2
A19	C	July	Ward 5&6	65	V	*vanB*	>=32	Rectal swab	12	117	103	25/7/2014	Plasmid	2
A20	C	July	Ward 5&6	61	M	*vanB*	<=0.5	Rectal swab	12	117	105	3/11/2014	Chromosome	3
A21	D	November	Ward 7	68	V	*vanB*	>=32	Rectal swab	1	80	104	29/10/2014	Plasmid	4

Sample ID	Outbreak cluster	Month	Ward(s)	Age	Gender	VRE type	Vancomycin MIC mg/L	Isolation source	MT[a] type	ST[b] type	CT[c] type	Sample date	Location *vanB* gene	Trans poson type
A22	D	November	Ward 7	62	M	*vanA* and *vanB*	>=32	Faeces	1	80	106	31/10/2014	Plasmid	4
A22.1	D	November	Ward 7	62	M	*vanA* and *vanB*	>=32	Rectal swab	1	80	106	4/11/2014	Plasmid	4
A23	D	November	Ward 7	66	V	*vanB*	8	Rectal swab	1	80	104	4/11/2014	Plasmid	4
A24	D	November	Ward 7	66	V	*vanB*	1	Rectal swab	1	80	104	4/11/2014	Plasmid	4
A25	D	November	Ward 7	70	M	*vanB*	1	Rectal swab	1	80	104	4/11/2014	Plasmid	4
A26	D	November	Ward 7	59	M	*vanB*	8	Rectal swab	1	80	104	4/11/2014	Plasmid	4
A27	D	November	Ward 7	50	M	*vanA* and *vanB*	>=32	Rectal swab	1	80	106	18/11/2014	Plasmid	4
A28	D	November	Ward 7	56	V	*vanB*	8	Rectal swab	1	80	104	19/11/2014	Plasmid	4
A29	E	November	Ward 2	57	M	*vanB*	>=32	Rectal swab	12	117	103	1/12/2014	Plasmid	2
A30	E	November	Ward 2	66	M	*vanB*	>=32	Rectal swab	12	117	103	2/12/2014	Plasmid	2
A31	E	November	Ward 2	60	V	*vanB*	8	Rectal swab	12	117	103	16/12/2014	Plasmid	2
A32	F	December	Ward 4	64	M	*vanB*	>=32	Rectal swab	12	117	103	22/12/2014	Plasmid	2
A33	F	December	Ward 4	69	M	*vanB*	>=32	Rectal swab	12	117	103	28/12/2014	Plasmid	2
A34	F	December	Ward 4	87	M	*vanB*	>=32	Rectal swab	12	117	103	31/12/2014	Plasmid	2

a MT type=MLVA type, b ST type = sequence type, c CT=cluster type. d ND=not determined *=these isolates were genetically related to outbreak A, but were obtained from patients from a regional hospital.

Discrepancies between epidemiological links and typing results

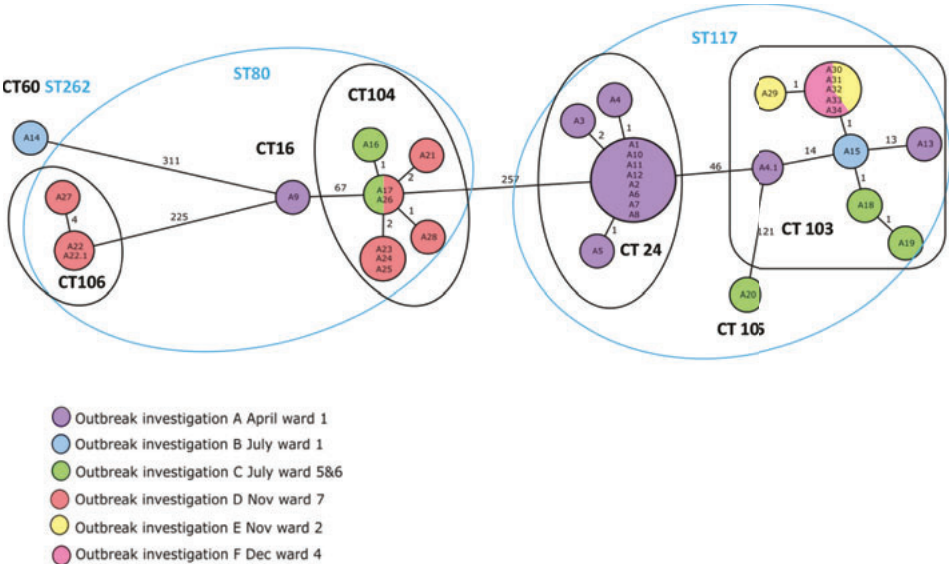
Initial MLVA typing showed three MLVA types (MT); MT1 (n=12), MT12 (n=23) and MT144 (n=1) (Table 1). Based on MLST typing, the isolates belonged to ST80 (n=12), ST117 (n=23) and ST262 (n=1). The clusters based on MLVA, and MLST matched except for isolate A9. CgMLST typing identified seven different clusters: CT103 (n=11), CT24 (n=11), CT104 (n=8), CT105 (n=1), CT106 (n=3), CT60 (n=1) and CT16 (n=1) (Table 1). The minimum spanning tree of the cgMLST typing results of the 36 sequenced isolates is shown in Figure 1.

In outbreak investigation A, the typing results of MLVA, MLST and cgMLST confirmed that 11 out of the 14 isolates were genetically related. These isolates belonged to CT24 whereas the isolates A13 and A9 were representing CT103 and CT16, respectively. Isolate A4.1 belonging to patient A4 of whom two isolates were sequenced, is discussed below. Patient A13 was initially considered as the index patient of the outbreak investigation A, because the patient was known to be colonized with VREfm already in March 2013. However, patient A13 was associated with another outbreak investigation which is discussed below. Based on the cgMLST results, patient A1 was eventually found to be most likely the index patient of the outbreak. As mentioned earlier, this patient was transferred from another regional hospital. Interestingly, the isolates of the three patients from the regional hospital (A1-A3), clustered together with the isolates (A4-A8 and A10-12) obtained from eight patients in our hospital. Isolate A9 belonged to CT16 and eventually could not be linked with any of the outbreaks. The two isolates from outbreak investigation B were totally different based on MLVA, MLST and cgMLST. In case of outbreak investigation C, MLST showed two isolates belonging to ST80 and three isolates belonging to ST117. The cgMLST results identified the presence of three CTs among the isolates in this outbreak investigation; CT103, CT104 and CT105. By MLVA and MLST typing isolates of outbreak investigation D could not be discriminated, but cgMLST divided them into two distinct clusters: five isolates belonged to CT104, and three to CT106. The isolates of CT106 were *vanA*/*vanB* co-producers. Based on cgMLST, the three isolates from outbreak investigation E belonged to CT103 as well as the three isolates from outbreak investigation F.

***vanB*-carrying transposons characterisation**

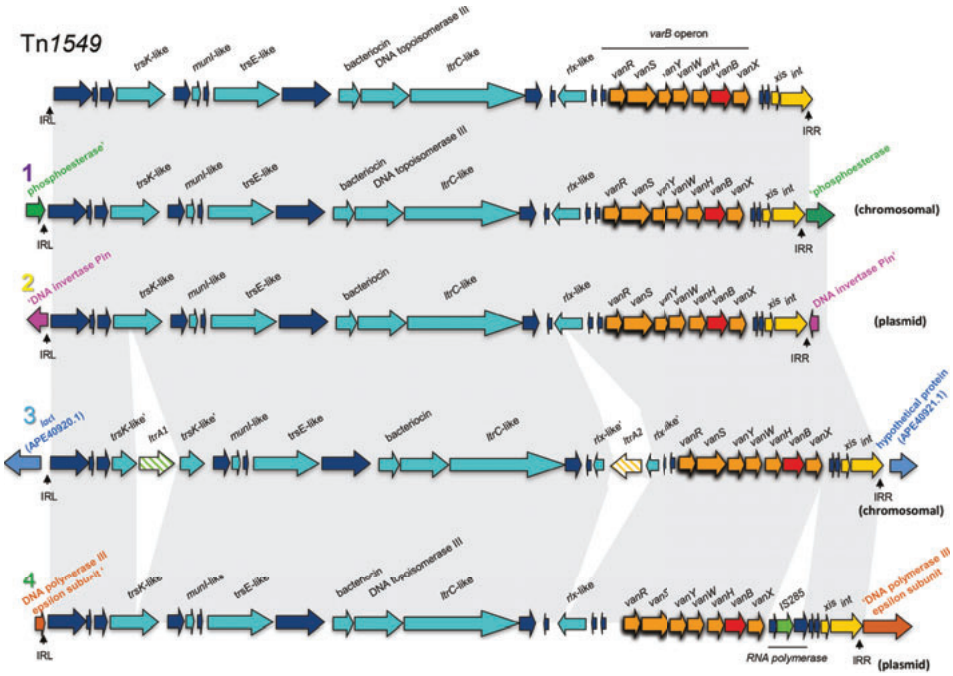
Based on the *de novo* assemblies and the hybrid assemblies generated from sequencing data of the 36 VREfm isolates, the *vanB*-carrying transposons and the genomic locations of these MGEs were investigated in more detail. Unfortunately, isolates A9 and A14 lost the *vanB* gene and were therefore excluded from this analysis. Four different transposons carrying the *vanB* operons were detected, further referred to as transposon type 1, 2, 3 and 4 (Figure 2).

Figure 1: Minimum spanning tree based on cgMLST (1,423 target genes). The different colors indicate the six different outbreaks investigations based on epidemiological data. Numbers indicate patients. Of patients 4 and 22, two samples were included, indicated as samples A4 and A4.1 and samples A22 and A22.1, respectively. The numbers next to the lines corresponds to allele differences between the isolates. ST = sequence type (blue); CT = cluster type (black).



Transposon type 1 was detected in all 13 VREfm isolates belonging to CT24 (A1-A8 and A10-12) and in one isolate belonging to CT103 (A4.1). The overall DNA sequence of this transposon was similar to the previously described transposon Tn*1549*/Tn*5382* (GenBank: AF192329.1) with 99 SNPs difference. In all 14 isolates, the identical *vanB* transposon was located on the bacterial chromosome integrated into the phosphoesterase gene (Genbank locus_taq: BO233_04565). Interestingly, isolates A4 with CT24 and A4.1 with CT103 were obtained from the same patient and both carried transposon type 1. In total, six isolates from rectum and bile were collected from patient A4 in the period from April till October 2014. We decided to sequence these additional six strains to verify this observation. Indeed, two isolates from rectum (A4.1 and 4.2) belonged to CT103. Two isolates from rectum (A4.3 and A4.4) and two from bile (A4 and A4.5) belonged to CT24. Details are shown in supplementary Table S1. Again, all six VREfm isolates harboured the identical *vanB* transposon (Transposon type 1) with identical insertion sites.

Figure 2: The four different *vanB* transposons in comparison to the reference *Tn1549*. Transposons were numbered as in Figure 4. All transposons have their unique insertion sites into different genes as indicated on both sides. Transposons 1 and 3 are located on the chromosome, whereas transposons 2 and 4 on the plasmids, as indicated in the Figure. IRL=left inverted repeat; IRR=right inverted repeat.



Transposon type 2 was detected in 10 isolates belonging to CT103 (A13, A15, A18, A19 and A29-A34). This transposon was found to be integrated into the plasmid DNA invertase Pin gene (Genbank locus_taq: BO233_15550). The overall DNA sequence of this transposon shared the lowest similarity in comparison with the reference Tn*1549*/Tn*5382* transposon and differed by 261 SNPs.

Transposon type 3 was detected in the single isolate of CT105 (A20). The transposon was located on the bacterial chromosome integrated between two genes; *lacI* (Genbank locus_taq :BO233_10750) and a gene encoding a hypothetical protein (GenBank locus_taq: BO233_10755). This transposon was similar to the reference Tn*1549*/Tn*5382* transposon, differing by 100 SNPs. In this transposon, two not previously reported regions were detected. A region of 2677 bp in size, was integrated into the gene encoding a *trsK*-like protein and contained a gene encoding an RNA-directed DNA polymerase sharing an 99% amino acid similarity with *Clostridioides difficile (*NCBI Reference Sequence: WP_044491975.1) The

second region of 2434 bp in size was integrated into a *Rlx* like protein and contained a gene probably responsible for encoding a group II intron reverse transcriptase/maturase. Interestingly, protein blast analysis revealed a substantial, 97% amino acid similarity, with a new identified protein homologous to a protein present in *Faecalibacterium spp.* (NCBI Reference Sequence: WP_087366583.1).

Transposon type 4 was detected in all CT104 (n=8) and CT106 (n=3) isolates. This transposon was located on a plasmid and integrated into the DNA polymerase III epsilon subunit gene. The transposon differed by 81 SNPs from the reference transposon and contained a novel insertion sequence IS*285* present downstream of *vanX*. This insertion sequence is related to *Ruminococcus* spp. as there was 98% amino acid identity with the IS*256* family transposase of *Ruminococcaceae bacterium* cv2 (NCBI Reference Sequence: WP_055079492.1).

Combining epidemiological data, cgMLST, and transposon characterisation

The analysis by cgMLST of all isolates showed clustering based on genetic relatedness of isolates which were initially grouped into different outbreak events. Isolates within CT103 belonged to outbreak clusters A, B, C, E, and F, but clustered together based on cgMLST. In addition, the identical Type 2 transposon was detected in VREfm from 10 patients, that were previously grouped into different outbreak clusters B, C, E, F. To elucidate this observation, we attempted a more detailed analysis by combining epidemiological data and to visualise patients transfer data and bed occupancies in our epidemiological program, as well as cgMLST and transposon analysis. Figure 3 shows the transfers/movements of 8 patients within and between four different hospital wards over time that were found to carry VREfm with the identical Type 2 transposon. By this approach, we identified overlaps in time and wards linking the patients A13, A15 and A29 till 34. No direct epidemiological links were found between patients A18 and A19 comparing to the other patients carrying VREfm with the Type 2 transposon.

Taking all the results together it was concluded that most likely three VREfm outbreaks took place (Figure 4). The first outbreak was caused by isolates of CT24 carrying transposon Type 1, including a case of within-patient transfer (patient A4) to CT103. A second outbreak was caused by isolates belonging to CT103 with transposon Type 2. The third outbreak was associated with isolates of CT104 and CT106 connected by horizontal transfer of transposon Type 4. All other isolates represented individual cases.

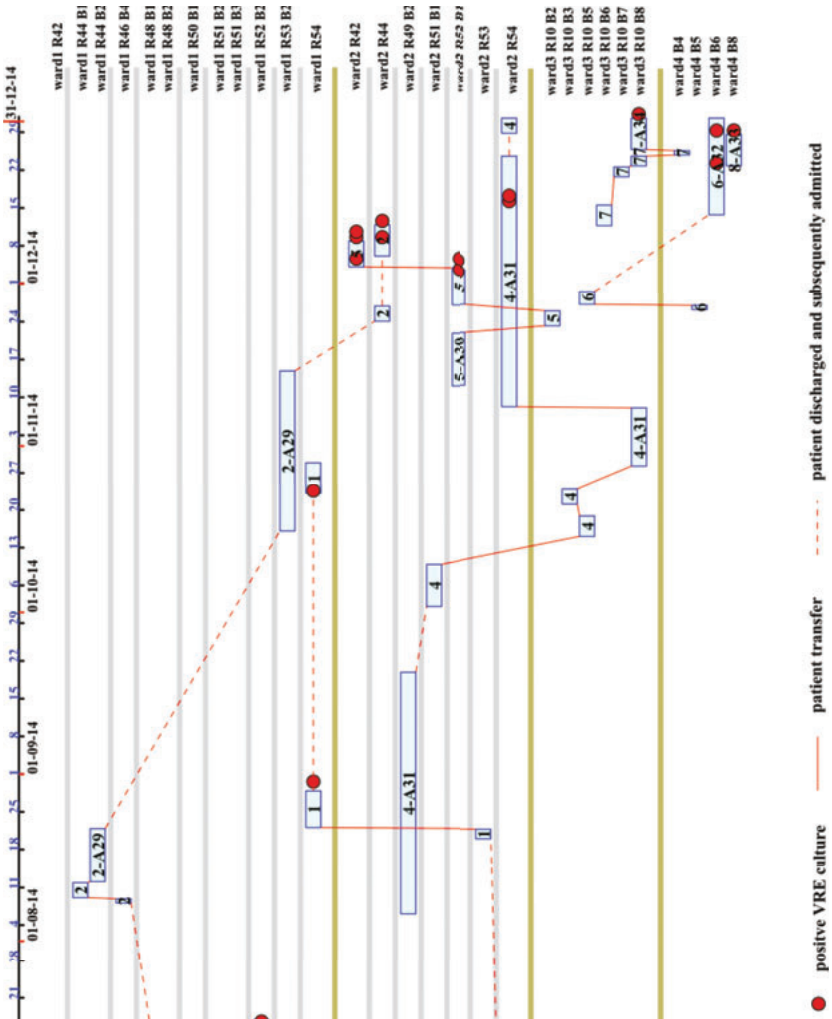
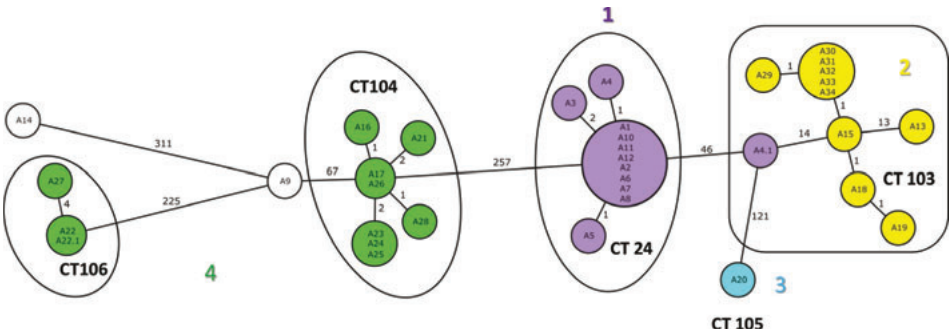


Figure 3: Patient movements among four different wards during the period from May until the end of December 2014. The figure shows the movements of patients A13, A15 and A29-A34. The numbers indicate the patients: 1=A13; 2=A29; 3=A15; 4=A31; 5=A30; 6=A32; 7=A34; 8=A33. On the right, the 4 different wards including the different room numbers (R) and beds (B) are shown.

Figure 4: Minimum spanning tree based on cgMLST (1,423 target genes). In contrast to Figure 1, colours now indicate the four different *vanB* transposon types (numbered in bold, 1-4). Isolates from A9 and A14 were excluded due to the loss of the *vanB* gene. From patient A4 and patient A22 two samples were included in the analysis (samples A4&A4.1 and A22&22.1 respectively). The numbers next to the lines correspond to allele differences between the isolates. ST=sequence type (blue), CT=cluster type (black).



DISCUSSION:

In this study, WGS and epidemiological data obtained from VREfm isolates during outbreaks in 2014 in our region were retrospectively analysed. Characterisation of *vanB*-carrying transposons in VREfm isolates was shown to be an additional value in the outbreak investigation. Transposon analysis is essential in cases where outbreaks are caused by the movement of particular MGEs. The horizontal transfer of *vanB*-carrying transposons was identified in two outbreak events. First, it was shown to occur within an individual patient, in whom isolates belonging to different clusters contained an identical transposon. Second, patients from outbreak investigation D belonging to different CTs (CT104 and CT106) carried VREfm isolates harbouring the same transposon. Thus, this study clearly shows the importance of *vanB* transposon investigation. VREfm isolates belonging to identical CTs defined by cgMLST, can acquire different *vanB* carrying transposons *de novo*, which can be incorrectly interpreted based on cgMLST only. Although this situation only occurred in one patient in our study, this phenomenon has been described [8, 27] and we hypothesize that this will happen more often if VREfm outbreaks analysis also include transposon investigation. On the contrary, VREfm isolates belonging to different CTs can also harbour the same *vanB* transposon and thereby belonging to the same outbreak cluster. Other studies have also explicitly shown the importance of transferable MGEs in VREfm outbreaks [4, 8, 27, 28]. Molecular typing methods

such as MLVA and MLST are used in the analysis of VREfm outbreaks and for epidemiological surveillance [11, 29-32]. However, these methods only allow to investigate clonal spread, as is also the case with cgMLST alone. These methods will fail in case the outbreaks are further complicated by horizontal gene transfer of MGE, like plasmids and/or transposons.

We observed the presence of the same *vanB* transposon in VREfm isolates belonging to distinct lineages, showing exchange of genomic material between VREfm and VSEfm. We also found transposons with low DNA sequence homology indicating that they originated from other species and the presence of insertion sequences originating from anaerobic bacteria which indicates transposon acquisition from anaerobic gut microbiotica to VSEfm. The occurrence of these two events are both important factors in the emergence of (*vanB*) VREfm.

In addition to the detection of horizontal gene transfer, this study shows that transposon-analysis even increases the discriminatory power of WGS compared to only using the data for cgMLST. On the other hand, cgMLST provides a higher discriminatory power than MLVA and MLST typing only. It is able to distinguish genetically closely related isolates even if they belong to the same ST lineage. This was the case for ST80 and ST117 in our study, each divided into three different CTs. Both ST117 and ST80 are frequently found in hospitals and associated with outbreaks [33-38] and typically belong the hospital associated clade A [39, 40]. CgMLST analysis also allows inter-laboratory exchange of typing data. This is important as the exchange of patients between hospitals and hospital units can contribute to the spread of VREfm within the healthcare networks. Indeed, using cgMLST allowed us not only to show clonal spread within our own hospital, but also intraregional spread via a connected hospital in our healthcare region. Recent studies from Denmark and England, where WGS for VREfm isolates was used as well, have also shown VREfm transmission within a healthcare network [11, 27, 41]. Therefore, it is wise to set up a local healthcare network surveillance program by identifying healthcare facilities that are most connected by patient traffic to allow optimal regional infection prevention measures. Such networks are currently recommended by the Ministry of Health, Wellbeing and Sports in the Netherlands, and is already well established in our Healthcare region [42].

Collecting epidemiological information is crucial to understand the transmission pathways during an outbreak [30, 43]. However, patients transfer can be quite complicated to follow as is shown in our study. Although an epidemiological link could be found for the majority of patients included in this study, some of the transmission pathways were still not fully understood. This could partially be explained by the fact that we were not able to

sequence all VREfm isolates present in all patients involved in the outbreaks investigations during the implementation of WGS in 2014. Moreover, data was not always directly available. Nowadays, WGS is fully implemented as a standardized typing methods for VRE in our institute and we have speed up the turnaround time to 48 hours (from culture to WGS data). Ideally, all WGS data should not only be used for cgMLST typing, but also in parallel for transposon analysis. Preferably, to create a complete picture of the outbreaks, all VREfm positive patients should be sequenced and included in the cgMLST analysis. Indeed, based on these preliminary results, we have now implemented WGS for every new VREfm isolate per patient. Because of horizontal gene transfer, it should also be considered to include several/all VREfm isolates per patient in outbreak investigations. This can lead to a further increase in the already enormous costs of outbreak investigations. However, advances in sequencing technologies and analysis tools, increases the output, speeds up the analysis and reduces the costs of WGS and by allowing for more focused infection control measures reducing probably overall costs [14, 15, 44]. This will lead to an increasing application of WGS, which is of great value in outbreak analysis.

In conclusion, this study shows that although cgMLST provides a high discriminatory power in the epidemiological analysis of VREfm, transposon analysis increases the power of WGS and allows the detection of horizontal gene transfer. Combining these two methods allows to investigate both clonal spread as well as concomitant spread of MGEs which will lead to a better insight and understanding of the highly complex transmission routes during in-hospital and regional VREfm outbreaks.

Funding

This work was a part of the routine work.

Transparency declarations

None to declare

REFERENCES

1. Freitas AR, Tedim AP, Francia MV, Jensen LB, Novais C, Peixe L, Sanchez-Valenzuela A, Sundsfjord A, Hegstad K, Werner G, Sadowy E, Hammerum AM, Garcia-Migura L, Willems RJ, Baquero F, Coque TM: Multilevel population genetic analysis of vanA and vanB Enterococcus faecium causing nosocomial outbreaks in 27 countries (1986-2012). J Antimicrob Chemother 2016, 71(12):3351-3366.
2. Willems RJ, Hanage WP, Bessen DE, Feil EJ: Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. FEMS Microbiol Rev 2011, 35(5):872-900.
3. Bender JK, 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.
4. Sivertsen A, Billstrom H, Melefors O, Liljequist BO, Wisell KT, Ullberg M, Ozenci V, Sundsfjord A, Hegstad K: 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(8):e103274.
5. Pinholt M, Gumpert H, Bayliss S, Nielsen JB, Vorobieva V, Pedersen M, Feil E, Worning P, Westh H: 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(1):40-47.
6. Novais C, Freitas AR, Sousa JC, Baquero F, Coque TM, Peixe LV: Diversity of Tn1546 and its role in the dissemination of vancomycin-resistant enterococci in Portugal. Antimicrob Agents Chemother 2008, 52(3):1001-1008.
7. Courvalin P: Vancomycin resistance in gram-positive cocci. Clin Infect Dis 2006, 42 Suppl 1:S25-34.
8. Howden BP, Holt KE, Lam MM, Seemann T, Ballard S, Coombs GW, Tong SY, Grayson ML, Johnson PD, Stinear TP: Genomic insights to control the emergence of vancomycin-resistant enterococci. MBio 2013, 4(4):10.1128/mBio.00412-13.
9. Top J, Schouls LM, Bonten MJ, Willems RJ: Multiple-locus variable-number tandem repeat analysis, a novel typing scheme to study the genetic relatedness and epidemiology of Enterococcus faecium isolates. J Clin Microbiol 2004, 42(10):4503-4511.
10. Homan WL, Tribe D, Poznanski S, Li M, Hogg G, Spalburg E, Van Embden JD, Willems RJ: Multilocus sequence typing scheme for Enterococcus faecium. J Clin Microbiol 2002, 40(6):1963-1971.
11. Pinholt M, Larner-Svensson H, Littauer P, Moser CE, Pedersen M, Lemming LE, Ejlertsen T, Sondergaard TS, Holzknecht BJ, Justesen US, Dzajic E, Olsen SS, Nielsen JB, Worning P, Hammerum AM, Westh H, Jakobsen L: Multiple hospital outbreaks of vanA Enterococcus faecium in Denmark, 2012-13, investigated by WGS, MLST and PFGE. J Antimicrob Chemother 2015, 70(9):2474-2482.
12. Raven KE, Reuter S, Reynolds R, Brodrick HJ, Russell JE, Torok ME, Parkhill J, Peacock SJ: A decade of genomic history for healthcare-associated Enterococcus faecium in the United Kingdom and Ireland. Genome Res 2016, 26(10):1388-1396.
13. van Hal SJ, Ip CL, Ansari MA, Wilson DJ, Espedido BA, Jensen SO, Bowden R: Evolutionary dynamics of Enterococcus faecium reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. Microb Genom 2016, 2(1):10.1099/mgen.0.000048.
14. Deurenberg RH, Bathoorn E, Chlebowicz MA, Couto N, Ferdous M, Garcia-Cobos S, Kooistra-Smid AM, Raangs EC, Rosema S, Veloo AC, Zhou K, Friedrich AW, Rossen JW: Application of next generation sequencing in clinical microbiology and infection prevention. J Biotechnol 2017, 243:16-24.

15. Rossen JWA, Friedrich AW, Moran-Gilad J, ESCMID Study Group for Genomic and Molecular Diagnostics (ESGMD): Practical issues in implementing whole-genome-sequencing in routine diagnostic microbiology. Clin Microbiol Infect 2017, .
16. Sabat AJ, Budimir A, Nashev D, Sa-Leao R, van Dijl J, Laurent F, Grundmann H, Friedrich AW, ESCMID Study Group of Epidemiological Markers (ESGEM): Overview of molecular typing methods for outbreak detection and epidemiological surveillance. Euro Surveill 2013, 18(4):20380.
17. de Been M, Pinholt M, Top J, Bletz S, Mellmann A, van Schaik W, Brouwer E, Rogers M, Kraat Y, Bonten M, Corander J, Westh H, Harmsen D, Willems RJ: A core genome MLST scheme for high-resolution typing of Enterococcus faecium. J Clin Microbiol 2015, .
18. Frakking FNJ, Bril WS, Sinnige JC, Klooster JEV, de Jong BAW, van Hannen EJ, Tersmette M: Recommendations for the successful control of a large outbreak of vancomycin-resistant Enterococcus faecium (VRE) in a non-endemic hospital setting. J Hosp Infect 2018, .
19. Zhou X, Arends JP, Kampinga GA, Ahmad HM, Dijkhuizen B, van Barneveld P, Rossen JW, Friedrich AW: Evaluation of the Xpert vanA/vanB assay using enriched inoculated broths for direct detection of vanB vancomycin-resistant Enterococci. J Clin Microbiol 2014, 52(12):4293-4297.
20. Hegstad K, Giske CG, Haldorsen B, Matuschek E, Schonning K, Leegaard TM, Kahlmeter G, Sundsfjord A, NordicAST VRE Detection Study Group: 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 resistance: a multicenter study. J Clin Microbiol 2014, 52(5):1582-1589.
21. Kluytmans-van den Bergh MF, Rossen JW, Bruijning-Verhagen PC, Bonten MJ, Friedrich AW, Vandenbroucke-Grauls CM, Willems RJ, Kluytmans JA: Whole-Genome Multilocus Sequence Typing of Extended-Spectrum-Beta-Lactamase-Producing Enterobacteriaceae. J Clin Microbiol 2016, 54(12):2919-2927.
22. Loman NJ, Quinlan AR: Poretools: a toolkit for analyzing nanopore sequence data. Bioinformatics 2014, 30(23):3399-3401.
23. Wick RR, Judd LM, Gorrie CL, Holt KE: Unicycler: Resolving bacterial genome assemblies from short and long sequencing reads. PLoS Comput Biol 2017, 13(6):e1005595.
24. Wick RR, Schultz MB, Zobel J, Holt KE: Bandage: interactive visualization of de novo genome assemblies. Bioinformatics 2015, 31(20):3350-3352.
25. Abbott JC, Aanensen DM, Rutherford K, Butcher S, Spratt BG: WebACT--an online companion for the Artemis Comparison Tool. Bioinformatics 2005, 21(18):3665-3666.
26. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream MA, Barrell B: Artemis: sequence visualization and annotation. Bioinformatics 2000, 16(10):944-945.
27. Brodrick HJ, Raven KE, Harrison EM, Blane B, Reuter S, Torok ME, Parkhill J, Peacock SJ: Whole-genome sequencing reveals transmission of vancomycin-resistant Enterococcus faecium in a healthcare network. Genome Med 2016, 8(1):4-015-0259-7.
28. Wardal E, Markowska K, Zabicka D, Wroblewska M, Giemza M, Mik E, Polowniak-Pracka H, Wozniak A, Hryniewicz W, Sadowy E: Molecular analysis of vanA outbreak of Enterococcus faecium in two Warsaw hospitals: the importance of mobile genetic elements. Biomed Res Int 2014, 2014:575367.
29. Borgmann S, Schulte B, Wolz C, Gruber H, Werner G, Goerke C, Klare I, Beyser K, Heeg P, Autenrieth IB: Discrimination between epidemic and non-epidemic glycopeptide-resistant E. faecium in a post-outbreak situation. J Hosp Infect 2007, 67(1):49-55.

30. Marcade G, Micol JB, Jacquier H, Raskine L, Donay JL, Nicolas-Viaud S, Rouveau M, Ribaud P, Dombret H, Leclercq R, Cambau E: Outbreak in a haematology unit involving an unusual strain of glycopeptide-resistant Enterococcus faecium carrying both vanA and vanB genes. J Antimicrob Chemother 2014, 69(2):500-505.
31. Lytsy B, Engstrand L, Gustafsson A, Kaden R: Time to review the gold standard for genotyping vancomycin-resistant enterococci in epidemiology: Comparing whole-genome sequencing with PFGE and MLST in three suspected outbreaks in Sweden during 2013-2015. Infect Genet Evol 2017, 54:74-80.
32. Werner G: Molecular Typing of Enterococci/VRE. J Bacteriol Parasitol 2013, S5-001.
33. Klare I, Konstabel C, Mueller-Bertling S, Werner G, Strommenger B, Kettlitz C, Borgmann S, Schulte B, Jonas D, Serr A, Fahr AM, Eigner U, Witte W: 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):815-825.
34. Sanchez-Diaz AM, Cuartero C, Rodriguez JD, Lozano S, Alonso JM, Rodriguez-Dominguez M, Tedim AP, Del Campo R, Lopez J, Canton R, Ruiz-Garbajosa P: The rise of ampicillin-resistant Enterococcus faecium high-risk clones as a frequent intestinal colonizer in oncohaematological neutropenic patients on levofloxacin prophylaxis: a risk for bacteraemia? Clin Microbiol Infect 2016, 22(1):59.e1-59.e8.
35. Tedim AP, Lanza VF, Manrique M, Pareja E, Ruiz-Garbajosa P, Canton R, Baquero F, Coque TM, Tobes R: Complete Genome Sequences of Isolates of Enterococcus faecium Sequence Type 117, a Globally Disseminated Multidrug-Resistant Clone. Genome Announc 2017, 5(13):10.1128/genomeA.01553-16.
36. Tedim AP, Ruiz-Garbajosa P, Rodriguez MC, Rodriguez-Banos M, Lanza VF, Derdoy L, Cardenas Zurita G, Loza E, Canton R, Baquero F, Coque TM: Long-term clonal dynamics of Enterococcus faecium strains causing bloodstream infections (1995-2015) in Spain. J Antimicrob Chemother 2017, 72(1):48-55.
37. McCracken M, Wong A, Mitchell R, Gravel D, Conly J, Embil J, Johnston L, Matlow A, Ormiston D, Simor AE, Smith S, Du T, Hizon R, Mulvey MR, Canadian Nosocomial Infection Surveillance Program: Molecular epidemiology of vancomycin-resistant enterococcal bacteraemia: results from the Canadian Nosocomial Infection Surveillance Program, 1999-2009. J Antimicrob Chemother 2013, 68(7):1505-1509.
38. Papagiannitsis CC, Malli E, Florou Z, Medvecky M, Sarrou S, Hrabak J, Petinaki E: First description in Europe of the emergence of Enterococcus faecium ST117 carrying both vanA and vanB genes, isolated in Greece. J Glob Antimicrob Resist 2017, 11:68-70.
39. Lebreton F, van Schaik W, McGuire AM, Godfrey P, Griggs A, Mazumdar V, Corander J, Cheng L, Saif S, Young S, Zeng Q, Wortman J, Birren B, Willems RJ, Earl AM, Gilmore MS: Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio 2013, 4(4):10.1128/mBio.00534-13.
40. Willems RJ, Top J, van Schaik W, Leavis H, Bonten M, Siren J, Hanage WP, Corander J: Restricted gene flow among hospital subpopulations of Enterococcus faecium. MBio 2012, 3(4):e00151-12.
41. Raven KE, Gouliouris T, Brodrick H, Coll F, Brown NM, Reynolds R, Reuter S, Torok ME, Parkhill J, Peacock SJ: Complex Routes of Nosocomial Vancomycin-Resistant Enterococcus faecium Transmission Revealed by Genome Sequencing. Clin Infect Dis 2017, 64(7):886-893.
42. Http://remis-plus.net/.
43. Pearman JW: 2004 Lowbury Lecture: the Western Australian experience with vancomycin-resistant enterococci - from disaster to ongoing control. J Hosp Infect 2006, 63(1):14-26.
44. Quainoo S, Coolen JPM, van Hijum SAFT, Huynen MA, Melchers WJG, van Schaik W, Wertheim HFL: Whole-Genome Sequencing of Bacterial Pathogens: the Future of Nosocomial Outbreak Analysis. Clin Microbiol Rev 2017, 30(4):1015-1063.