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


University of Groningen, University Medical Center Groningen, Department of Medical Microbiology, Groningen, The Netherlands

*Corresponding author. Tel: +31 50 3613480; Fax: +31 50 3619105; E-mail: j.w.a.rossen@rug.nl

Received 20 March 2018; returned 4 June 2018; revised 10 July 2018; accepted 6 August 2018

Background: 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.

Objectives: We aimed to combine vanB-carrying transposon data 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 and 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 was performed using the Artemis Comparison Tool (ACT) and multiple blast analyses.

Results: Four different vanB transposons were found among the isolates. cgMLST divided ST80 isolates into three cluster types (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.

Conclusions: 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 investigation of 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 have 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 fragments3 and plasmids.1,4,5 The vanA gene is part of an operon of seven genes, carried by the Tn1546 transposon, which can be located on various plasmid types or can be integrated into the chromosome.6,7 Similarly to vanA, vanB is also part of an operon that consists of seven genes, generally located on the conjugative transposon Tn1549/Tn5382. 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 Tn1549, between vanB VREfm and VSEfm.3 In outbreak investigations, rapid and accurate typing is required to investigate the genetic relatedness between patient isolates. This information is essential to demonstrate nosocomial transmission. Till 2014, most 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. However, the discriminatory power of MLST is also insufficient in nosocomial outbreak investigations. In addition to its 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.

In 2014, WGS was implemented in our laboratory for outbreak investigations of MDR microorganisms, including VREfm. The challenge of using WGS is to rapidly analyse and interpret the relevant information. In 2015, a core-genome (cg) MLST scheme (consisting of 1423 target genes) for *E. faecium* was developed. 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 characterization of vanB-carrying transposons was performed to determine possible horizontal gene transfer. We used these techniques to investigate spread by clonal expansion as well as by horizontal gene transfer.

**Methods**

**Study population and infection control protocols**

We retrospectively analysed VREfm outbreaks that occurred in the University Medical Center Groningen (UMCG), 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 VRE-positive 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 analyse 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.

As routine, we screen the following patients for VRE upon admission: patients who have been admitted to 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 <1 year ago, are treated in contact isolation and additional rectal swabs are taken for VRE screening. At least five negative rectal swabs are needed to discard the isolation measures in VRE-positive patients and those that were known to carry VRE <1 year ago. Patients previously known to carry VRE >1 year ago are treated in contact isolation, unless one or more negative previous VRE cultures were recorded. An additional 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 h after admission) 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 on 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, where at least 48 h between each screening round is required. On average, the last screening round will be 7 days after (possible) exposure since transmission and subsequent rectal colonization takes time.

**VRE culture**

VRE culture was preceded by PCR screening as described previously. In brief, rectal swabs were inoculated in enrichment broth. After 24 h 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 agar (Oxoid®). Agar was incubated for 24–48 h and identification and antibiotic susceptibility testing were performed on suspected colonies by MALDI-TOF Mass Spectrometry (Bruker) and VITEK®2 (bioMérieux), respectively. Additionally, we used vancomycin disc diffusion since this method is more sensitive in detecting enterococcal isolates with low- and medium-level vanB-type vancomycin resistance. 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.

As standard, all first VREfm isolates of each patient were typed by MLVA, according to the method described by Top et al. 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, USA) following the manufacturer’s instructions. The DNA concentration and purity were measured by the Qubit dsDNA HS and BR assay kit (Life Technologies, Carlsbad, CA, USA). A DNA library was prepared using the Nextera XT v2 kit (Illumina, San Diego, CA, USA) and then run on a MiSeq sequencer (Illumina) to generate paired-end 250 bp reads. De novo assembly was performed by CLC Genomics Workbench v7.0.4 (Qiagen, Hilden, Germany) after quality trimming (Q > 20) with optimal word sizes. All procedures were performed as previously described.

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 a FLO-MIN106 R9.4 flow cell and run on a MinION device (48 h). Base calling was performed using Albacore v1.2.2. Data quality was analysed through Poretools v0.6.0. Hybrid assemblies were performed using Unicycled v0.4.1. Bandage v0.8.1 was used to visualize the assembly graphics. Genes of interest were detected using ResFinder.

MLST STs and cgMLST cluster types (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, 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.
Nucleotide sequence accession numbers

Sequence data obtained in this study have 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 from 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 on seven 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 and A4.1; plus A22 and 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 section. Outbreak investigation A took place in April 2014 on ward 1 and 12 patients were involved. For 10 of these patients, the genome sequences of the obtained isolates (A1 and A4–A13, including A4.1) were available. One of the patients (A1) 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 four patients were involved. The genome sequences of the obtained isolates (A14 and A15) were available for two patients. Outbreak investigation C took place in July 2014 on wards 5 and 6 and 10 patients were involved. Genome sequences of the obtained isolates (A16–A20) were available for five of these patients. According to epidemiological data, outbreak investigation D took place in November 2014 on ward 7 and involved a total of 11 patients. The genome sequences of the obtained isolates (A21–A28) were available for eight of these patients. Also in November 2014, outbreak investigation E took place on ward 2, involving 11 patients. The genome sequences of the obtained isolates (A29–A31) were available for three of these patients. Finally, outbreak investigation F took place in December 2014 on several wards, involving seven patients. The genome sequences of the obtained isolates (A32–A34) were available from three patients on a selected ward (ward 4).

Discrepancies between epidemiological links and typing results

Initial MLVA typing showed three MLVA types (MTs); 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 from CT103 and CT16, respectively. Isolate A4.1 from patient A4, from 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 the most likely 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 the 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 transposon characterization

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 types 1, 2, 3 and 4 (Figure 2).
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 Tn1549/Tn5382 (GenBank: AF192329.1) with 99 SNP differences. In all 14 isolates, the identical vanB transposon was located on the bacterial chromosome integrated into the phosphoesterase gene (Genbank locus_tag: 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 the 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 the rectum (A4.1 and A4.2) belonged to CT103. Two isolates from the rectum (A4.3 and A4.4) and two from bile (A4 and A4.5) belonged to CT24. Details are shown in Table S1 (available as Supplementary data at JAC Online). Again, all six VREfm isolates harboured the identical vanB transposon (transposon type 1) with identical insertion sites.

Transposon type 2 was detected in 10 isolates belonging to CT103 (A1–A8 and A10–12) and in one isolate belonging to CT24 (A13–A15, A18, A19 and A29–A34). This transposon was found to be integrated into the plasmid DNA invertase Pin gene.
The overall DNA sequence of this transposon shared the lowest similarity in comparison with the reference Tn1549/Tn5382 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_ta: BO233_10750) and a gene encoding a hypothetical protein (GenBank locus_ta: BO233_10755). This transposon was similar to the reference Tn1549/Tn5382 transposon, differing by 100 SNPs. In this transposon, two previously unreported regions were detected. A region of 2677 bp was integrated into the gene encoding a TrsK-like protein and contained a gene encoding an RNA-directed DNA polymerase sharing 99% amino acid similarity with Clostridioides difficile (NCBI Reference Sequence: WP_044491975.1). The second region of 2434 bp was integrated into an Rlx-like protein and contained a gene probably responsible for encoding a group II intron reverse transcriptase/maturase. Interestingly, protein blast analysis revealed 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 IS, IS285, present downstream of vanX. This insertion sequence is related to Ruminococcus spp. as there was 98% amino acid identity with the IS256 family transposase of Ruminococcaceae bacterium cv2 (NCBI Reference Sequence: WP_055079492.1).

Combining epidemiological data, cgMLST and transposon characterization

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 and F. To elucidate this observation, we attempted a more detailed analysis by combining epidemiological data and visualization of patient transfer data and bed occupancies in our epidemiological program, as well as cgMLST and transposon analysis. Figure 3 shows the transfers/movements of eight 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–A34. No direct epidemiological links were found between patients A18 and A19 compared with 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.

**Discussion**

In this study, WGS and epidemiological data obtained from VREfm isolates during outbreaks in 2014 in our region were retrospectively analysed. Characterization of vanB-carrying transposons in VREfm isolates was shown to be of 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 individual patients, 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 already been described\(^8,27\) and we hypothesize that this will be observed more often if VREfm outbreak analysis also includes transposon investigation. In contrast, VREfm isolates belonging to different CTS can also harbour the same vanB transposon and thereby belong to the same outbreak cluster. Other studies have also explicitly shown the importance of transferable MGEs in VREfm outbreaks.\(^5,8,27,28\) 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 investigation of clonal spread, as is also the case with cgMLST alone. These methods will fail in the case of outbreaks being further complicated by horizontal gene transfer of MGE, such as mediated by 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 ISs originating from anaerobic bacteria, which indicates transposon acquisition from the anaerobic gut microbiota by 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 also increases the discriminatory power of WGS compared with 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

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; and 8, A33. On the right, the four different wards including the different room numbers (R) and beds (B) are shown. This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.

Figure 4. Minimum spanning tree based on cgMLST (1423 target genes). In contrast to Figure 1, colours now indicate the four different vanB transposon types (numbered in colour, 1–4). Isolates from A9 and A14 were excluded due to the loss of the vanB gene. Patient A4 and patient A22 had two samples included in the analysis (samples A4 and A4.1, and A22 and 22.1 respectively). The numbers next to the lines correspond to allele differences between the isolates. This figure appears in colour in the online version of JAC and in black and white in the print version of JAC.
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, are associated with outbreaks, and typically belong to the HA clade A. cgMLST analysis also allows inter-laboratory exchange of typing data. This is important as the exchange of patient data between hospitals and hospital units can contribute to the exchange of typing data. This is important as the exchange of optimal regional infection prevention measures. Such networks are currently recommended by the Ministry of Health, Wellbeing and Sports in the Netherlands and one is already well established in our healthcare region.

Collecting epidemiological information is crucial to understand the transmission pathways during an outbreak. However, patient transfers 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 outbreak investigations during the implementation of WGS in 2014. Moreover, data were not always directly available. Nowadays, WGS is fully implemented as a standardized typing method for VRE in our institute and we have speeded up the turnaround time to 48 h (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 have their isolates 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 it may reduce overall costs. 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 investigation of both clonal spread as well as concomitant spread of MGEs which will lead to better insights into and understanding of the highly complex transmission routes during in-hospital and regional VREfm outbreaks.

**Funding**

This study was carried out as part of our routine work.

**Supplementary data**

Table S1 appears as Supplementary data at JAC Online.

**Transparency declarations**

None to declare.

**References**

Transposons on the move in VREfm outbreaks