Author/s:
Lee, RS;Gonçalves Da Silva, A;Baines, SL;Strachan, J;Ballard, S;Carter, GP;Kwong, JC;Schultz, MB;Bulach, DM;Seemann, T;Stinear, TP;Howden, BP

Title:
The changing landscape of vancomycin-resistant Enterococcus faecium in Australia: A population-level genomic study

Date:
2018-12-01

Citation:
Lee, R. S., Gonçalves Da Silva, A., Baines, S. L., Strachan, J., Ballard, S., Carter, G. P., Kwong, J. C., Schultz, M. B., Bulach, D. M., Seemann, T., Stinear, T. P. & Howden, B. P. (2018). The changing landscape of vancomycin-resistant Enterococcus faecium in Australia: A population-level genomic study. Journal of Antimicrobial Chemotherapy, 73 (12), pp.3268-3278. https://doi.org/10.1093/jac/dky331.

Persistent Link:
https://hdl.handle.net/11343/290304

J Antimicrob Chemother 2018; **73**: 3268–3278
doi:10.1093/jac/dky331 Advance Access publication 4 September 2018

Journal of Antimicrobial Chemotherapy

The changing landscape of vancomycin-resistant *Enterococcus faecium* in Australia: a population-level genomic study

Robyn S. Lee[1,2], Anders Gonçalves da Silva[1], Sarah L. Baines[3], Janet Strachan[1], Susan Ballard[1], Glen P. Carter[1], Jason C. Kwong [3], Mark B. Schultz[1], Dieter M. Bulach[1], Torsten Seemann[4], Timothy P. Stinear[3] and Benjamin P. Howden[1,5]*

[1]The Microbiological Diagnostic Unit Public Health Laboratory, Department of Microbiology and Immunology, The University of Melbourne, at The Peter Doherty Institute for Infection and Immunity, 792 Elizabeth Street, Level 1, Melbourne, Victoria 3000, Australia; [2]Center for Communicable Disease Dynamics, Department of Epidemiology, Harvard T. H. Chan School of Public Health, 677 Huntington Avenue, Level 5, Boston, MA 02115, USA; [3]Department of Microbiology and Immunology, The University of Melbourne, at The Peter Doherty Institute for Infection and Immunity, 792 Elizabeth Street, Level 1, Melbourne, Victoria 3000, Australia; [4]Melbourne Bioinformatics Group, Lab-14, 700 Swanston Street, Carlton, Victoria 3053, Australia; [5]Infectious Diseases Department, Austin Health, Studley Rd, Heidelberg, Victoria, 3084, Australia

*Corresponding author. Microbiological Diagnostic Unit Public Health Laboratory, The Peter Doherty Institute for Infection and Immunity, The University of Melbourne, 792 Elizabeth Street, Level 1, Melbourne, Victoria 3000, Australia. E-mail: bhowden@unimelb.edu.au

Received 30 April 2018; returned 20 June 2018; revised 18 July 2018; accepted 23 July 2018

Background: Vancomycin-resistant *Enterococcus faecium* (VREfm) represent a major source of nosocomial infection worldwide. In Australia, there has been a recent concerning increase in bacteraemia associated with the *vanA* genotype, prompting investigation into the genomic epidemiology of VREfm.

Methods: A population-level study of VREfm (10 November–9 December 2015) was conducted. A total of 321 VREfm isolates (from 286 patients) across Victoria State were collected and sequenced with Illumina NextSeq. SNPs were used to assess relatedness. STs and genes associated with resistance and virulence were identified. The *vanA*-harbouring plasmid from an isolate from each ST was assembled using long-read data. Illumina reads from remaining isolates were then mapped to these assemblies to identify their probable *vanA*-harbouring plasmid.

Results: *vanA*-VREfm comprised 17.8% of isolates. ST203, ST80 and a *pstS*(−) clade, ST1421, predominated (30.5%, 30.5% and 37.2%, respectively). Most *vanB*-VREfm were ST796 (77.7%). *vanA*-VREfm were more closely related within hospitals versus between them [core SNPs 10 (IQR 1–357) versus 356 (179–416), respectively], suggesting discrete introductions of *vanA*-VREfm, with subsequent intra-hospital transmission. In contrast, *vanB*-VREfm had similar core SNP distributions within versus between hospitals, due to widespread dissemination of ST796. Different *vanA*-harbouring plasmids were found across STs. With the exception of ST78 and ST796, Tn*1546* transposons also varied. Phylogenetic analysis revealed Australian strains were often interspersed with those from other countries, suggesting ongoing cross-continental transmission.

Conclusions: Emerging *vanA*-VREfm in Australia is polyclonal, indicating repeat introductions of *vanA*-VREfm into hospitals and subsequent dissemination. The close relationship to global strains reinforces the need for ongoing screening and control of VREfm in Australia and abroad.

Introduction

Enterococcus faecium is a leading cause of nosocomial infections worldwide.[1,2] Resistance to glycopeptides, particularly vancomycin, is the most clinically relevant resistance for *E. faecium*;[1] mortality associated with any *E. faecium* bloodstream infection (BSI) is >30%,[3] while patients with vancomycin-resistant *E. faecium* (VREfm) bacteraemia have ∼2.5-fold higher odds of death compared with those with vancomycin-susceptible infection.[4] Resistance to vancomycin in *E. faecium* can be conferred by different *van* gene clusters (*A*, *B*, *D*, *E*, *G*, *L*, *M* and *N*[5–8]), with most infection attributed to *vanA* and *vanB* genotypes.[9] VREfm is thought to be transmitted between hospitalized patients,

healthcare staff and/or via the hospital environment,[10] with risk of colonization highest among those receiving long courses of antibiotics, the critically ill and/or immunosuppressed, and those with prolonged hospitalization or history of nursing home residence.[11] VREfm may also arise *de novo* in a patient, following transfer of the *vanB* gene cluster from commensal gut anaerobes to *E. faecium*.[12,13]

The prevalence of the *vanA* and *vanB* genotypes varies geographically; in North America, *vanA* has been shown to drive the VREfm epidemic,[14] while in Europe, both *vanA* and *vanB* genotypes play a role.[15] Australia has one of the highest rates of VREfm in the world, and until recently, was dominated by *vanB*-VREfm.

Multiple clones of *vanB*-VREfm have been detected in Australia, mostly from STs within the hospital-associated clonal complex 17.[9,16–19] A survey of sentinel laboratories from across Australia in 2011 found that 97.2% (104 of 107 isolates) of vancomycin-non-susceptible isolates had the *vanB* genotype,[16] over half of which were ST203.[18,20] Subsequent country-wide surveys showed this strain dominated until 2014, when it was surpassed by the ST796 clone[21] (largely due to its widespread dissemination throughout Victoria and southeastern Australia).[22]

Along with the spread of *vanB*-VREfm, there has been a recent, dramatic increase in the *vanA* genotype; since 2011, the prevalence of *vanA*-VREfm bacteraemia, as measured by these sentinel surveys, increased from <1% (2 of 341)[16] to 22% (90 of 408 isolates genotyped) in 2016.[23] The reasons for this increase remain unclear, but multiple STs appear to be involved. A retrospective study examining *vanA*-VREfm (screening and clinical isolates) across four Australian hospitals from 2011 to 2013[24] suggested this increase was not due to a single ST or *vanA*-harbouring plasmid; however, as it was based on a convenience sample of only 18 isolates from across Australia, this was inconclusive.

To explore further the potential reasons for the emerging dominance of *vanA*-VREfm in our region, we conducted a population-level study of VREfm. All VREfm isolates from colonization and clinical infection detected across Victoria (the second most populous state in Australia) over a month were subjected to WGS, as well as isolates from VSEfm bacteraemia. In doing so, we provide an in-depth assessment of the genomic diversity of *E. faecium* and, situating these in context with global VREfm, novel insights into the genomic epidemiology of this clinically relevant pathogen.

Methods

Ethics

This research was conducted in accordance with the Declaration of Helsinki and Australian National and institutional standards. Data were collected as part of public health surveillance and de-identified for analysis. Individual patient consent was not required.

Study design and population

A cross-sectional survey of VREfm was conducted between 10 November and 9 December 2015 in the State of Victoria (population: 5 931 100 for 2015; http://www.abs.gov.au/AUSSTATS/). During this period, all VREfm-positive isolates (including screening and clinical samples) collected by laboratories across the state were sent to the Microbiological Diagnostic Unit Public Health Laboratory (MDU PHL), as well as all vancomycin-susceptible *E. faecium* (VSEfm) isolated from blood cultures. Final lists of isolates received were cross-checked with primary diagnostic laboratories to ensure all eligible isolates were included.

Laboratory methods

See Supplementary data (available at *JAC* Online) for detail. Isolates were sequenced using the Illumina NextSeq 500, with 150 bp paired-end reads.

Bioinformatics

In brief, sequences were analysed using the Nullarbor pipeline (Seemann T, available at: https://github.com/tseemann/nullarbor). Reads were assembled into contigs using SPAdes (v.3.10.1).[25] Resistance genes present in the ResFinder database[26] and putative virulence genes from the Virulence Factor DataBase[27] were identified *in silico* from these assemblies using ABRicate (Seemann T, https://github.com/tseemann/abricate). Assembled contigs were also searched for *E. faecium* alleles listed in the pubMLST database (https://pubmlst.org). For SNP-based analyses, reads were aligned using the Burrows–Wheeler Aligner MEM algorithm (v.0.7.16[28]) to a novel local reference (ST796, AUSMDU00004028; see Table S1). SNPs were called using Freebayes (v1.0.2).[29] Further details are given in the Supplementary data.

Phylogenetics and population structure

ClonalFrameML was used to identify potential recombination (v.1.11).[30] This was masked using a custom script (https://github.com/kwongj/cfml-maskrc) and the adjusted core SNP alignment was used to produce a maximum likelihood tree using RAxML (v.8.2.11).[31] Bayesian analysis of population structure with hierarchical model-based clustering (hierBAPS; v.6.0),[32] was used to identify major clades.

Comparative genomics of vanA-VREfm

A bacteraemia strain from each of the major *vanA*-harbouring STs was sequenced using PacBio single molecule, real-time sequencing (Pacific Biosciences, CA, USA). *vanA*-harbouring plasmids were identified using ABRicate, and annotated using Prokka (v. 1.11),[33] with the *Enterococcus* database (https://github.com/tseemann/prokka/blob/master/db/genus/Enterococcus). To compare the gene content, these plasmids were aligned to one another using progressiveMauve (build date 19 December 2014) and visualized in Geneious (v.9.1.7).[34] To investigate whether the same *vanA*-harbouring plasmids were present in the other isolates in our dataset, we mapped reads to each complete *vanA*-harbouring plasmid (see Methods in Supplementary data).

Australian isolates in global context

To assess how our *E. faecium* strains fit in the global context, we searched for publicly available short-read data; this resulted in inclusion of strains from Europe (872), North America (6), Africa (2) and Asia (Israel, 1),[35–40] and strains from the Welcome Trust Sanger BSAC Resistance Surveillance Project (from NCBI BioProject PRJEB344).

Statistical analysis

Analyses were done in Stata v.14.2 (StataCorp, College Station, TX, USA). Pairwise SNP distributions were compared using the Mann–Whitney rank-sum test. Differences in proportion were compared using the Fisher's exact test or a two-proportion *Z*-test.

Data availability

Sequencing data and PacBio assemblies are under BioProject PRJNA433676 on the NCBI Sequence Read Archive.

Results

Between 10 November and 9 December 2015, 321 VREfm-positive samples were detected at 19 primary diagnostic laboratories across Victoria (two additional laboratories reported no *E. faecium* during this time). Twelve VSEfm BSI samples were also detected, two of which were not submitted. Thus, 331 isolates were available for sequencing. All isolates passed WGS quality control (see Supplementary data).

In total, 321 isolates (from 286 patients) were VREfm, 20 of which were from BSI (from 17 patients). Ten VSEfm isolates (from 10 patients) were also received; one of these was likely a false negative, as the patient also had a *vanA*-VREfm isolated from blood collected the same day (see Supplementary data). Thus, 27 patients had *E. faecium* bacteraemia over this period. Extrapolated to 1 year, this gives an incidence of 5.5/100 000 for all *E. faecium*, and 0.6/100 000 and 3.2/100 000 for *vanA*-VREfm and *vanB*-VREfm, respectively (excluding one case of *vanA* + *vanB*-VREfm bacteraemia.

Most patients (257 of 295, 87.1%) were from hospitals in major metropolitan centres, compared with rural areas (29 of 295, 9.8%), consistent with the geographical distribution of the population in this state. Overall, 35 hospital networks (HCNs; groups of ≥1 affiliated public or private hospitals that provide services to distinct geographical areas in Victoria) were represented, with isolates received from up to five hospitals from a single network. Five isolates were collected in community medical/dental clinics (3.1%). Sex and age were not available for patients from one HCN (HCN1, 61 of 295, 20.7%); among the remaining patients, 113 of 295 were male (38.3%), and the median age was 71 (range 18–98, IQR 58–80). At sample collection, 256 persons were inpatients and 36 were outpatients. The location of sample collection was missing for four patients, who also had unknown admission status.

Twenty-seven patients had more than one sample positive for *E. faecium* (range 2–5 per patient; Table S2), with a maximum of 40 SNPs between repeat isolates of the same ST (Figure S1). Repeat samples were collected within the same hospital as the original sample for 26 of 27 patients (96.3%, Table S2).

Identification of van genotype

In silico genotyping revealed that 260 (78.5%) isolates were exclusively *vanB*, and 59 (17.8%) were exclusively *vanA*. There was no association between *van* genotype and BSI (Figure 1, Fisher's exact $P = 0.52$). Two isolates were positive for *vanA* and *vanB* (<1%). These were most closely related to *vanB*-VREfm isolates, with a minimum of zero pairwise core SNPs compared with *vanB*-VREfm isolates (versus 195 for *vanA*-VREfm), suggesting they already had *vanB* when *vanA* was acquired. As expected, all isolates with *vanA* and/or *vanB* were phenotypically resistant to vancomycin (MIC >4 mg/L[41]). Neither *vanA* nor *vanB* was found in isolates classified as VSEfm (3.0%).

Overall, *vanB*-VREfm was more widely dispersed across hospitals and HCNs compared with *vanA*-VREfm; excluding the two *vanA* + *vanB*-positive samples, *vanB*-VREfm isolates were identified in 34 of 35 HCNs, while *vanA*-VREfm was only identified in 10 of 35 during this period ($P < 0.00005$). VSEfm isolates were from six different HCNs.

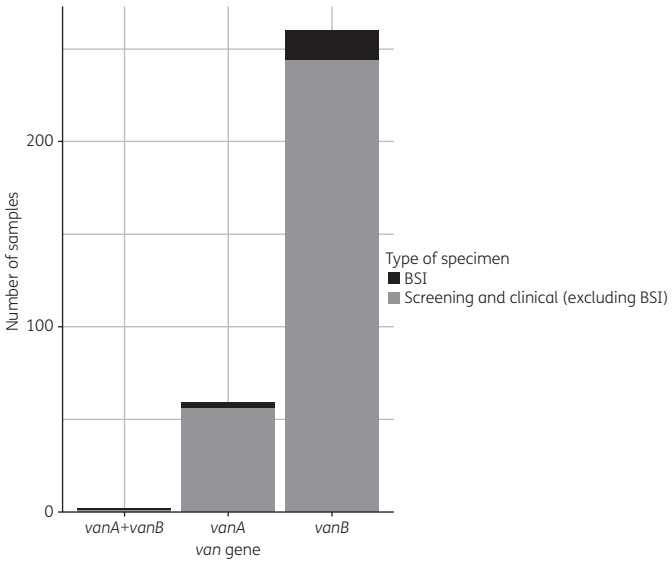

Figure 1. Specimen type by *van* gene for Australian *E. faecium* isolates.

MLST

Eighteen previously known STs and four novel STs were identified. Most isolates were ST796 (61.3%, Table 1). The *pstS* allele was missing for 22 (6.6%) isolates; these were classified as ST1421 based on remaining alleles (Table S3).

Excluding *vanA* + *B*-positive isolates, *vanA* was found in fewer STs than *vanB* (4 versus 15, respectively, $P < 0.00005$). *vanA* was almost exclusively present in ST203, ST80 and ST1421 isolates (Table 1). Most *vanB*-positive isolates were ST796. Among the nine probable VSEfm isolates, eight different STs were represented; only ST80 ($n = 2$) was also present among VREfm.

Phylogenetics and population structure

Adjusting for recombination reduced the median core SNPs from 20 (IQR 14–2653) to 12 (IQR 9–238). The adjusted maximum likelihood tree is shown in Figure 2 (recombination blocks are shown in Figure S2). Many branches had very low bootstrap support; therefore, BAPS was used to validate clusters.

Five major BAPS groups were identified, which were largely consistent with the phylogeny (Figure 2). While all BAPS groups had at least one VSEfm, none was closely related to any VREfm within the same cluster (Table S4). Overall, BAPS-1 was predominantly ST796 (203 of 214, 94.9%), and was highly clonal, with a median of 10 core SNPs separating isolates (IQR 7–15). In contrast, BAPS-2 comprised multiple STs. The highest proportions were from ST203 (36 of 82, 43.9%), ST1421 (22 of 82, 26.8%) and ST17 (12 of 82, 14.6%). Isolates in BAPS-2 were separated by a median of 257 core SNPs (IQR 167–312). Excluding repeat samples from the same patient, 81 pairs of isolates were within one SNP. Sixty-six of these pairs (81%) were isolated at the same hospitals as one another (H1, H3, H5), while the remainder were predominantly from different HCNs (see Figure S3). BAPS-3 comprised a single VSEfm isolate from ST54, which was over 3800 SNPs from other isolates (not shown in Figure 1). BAPS-4 isolates were predominantly from ST80 (25 of 31, 80.6%), but also included isolates from ST17, ST78

Table 1. STs identified

ST	Isolates, *n* (%)	*vanA* only, *n* (%)	*vanB* only, *n* (%)	*vanA* + *B*, *n* (%)	Neither *vanA* nor *vanB*, *n* (%)
796	203 (61.3)		202 (77.7)	1 (50.0)	
203	37 (11.2)	18 (30.5)	19 (7.3)		
80	26 (7.9)	18 (30.5)	5 (1.9)		3 (30.0)[a]
17	17 (5.1)		15 (5.8)	1 (50.0)	1 (10.0)
1421 [*pstS*(−)]	22 (6.6)	22 (37.2)			
1429[b]	5 (1.5)		5 (1.9)		
Other[c]	21 (6.3)	1 (1.7)	14 (5.4)		6 (60.0)
Total, *n* (%)	331 (100)	59 (17.8)	260 (78.5)	2 (1.0)	10 (3.0)

[a]Includes the one isolate that likely lost the *vanA*-harbouring plasmid during passage.
[b]Novel single-locus variant of ST796.
[c]One isolate included here might have an interrupted *ddl* gene, but we were unable to resolve this using short-read data. Thus, we were not able to type this isolate. It clusters with ST17 isolates in the recombination-adjusted phylogenetic tree. Sixteen additional STs were identified, each represented by three or fewer isolates. Single *vanA* isolate here corresponds to ST78.

and ST262. Isolates were separated by a median of 252 core SNPs (IQR 6–346). All pairs within one core SNP were VREfm from ST80; excluding repeat isolates, 8 of 15 of these pairs were collected at the same hospitals as one another (H1, H2 and H4). BAPS-5 comprised three VSEfm isolates (ST21, ST22 and ST32), with a minimum of 369 core pairwise SNPs.

Genetic diversity across hospitals and van genotypes

Excluding comparisons within the same patient (shown in Figure S2), the median pairwise core SNP distances among VREfm were similar within and between hospitals [196 (IQR 13–297) versus 193 (IQR 13–207), respectively; Figure 3a].

As we hypothesized there were differences in diversity between isolates with different VREfm genotypes, *vanA*-VREfm and *vanB*-VREfm were analysed separately. *vanA*-VREfm were significantly less diverse within hospitals compared with between them (Figure 3b, $P = 0.0001$), suggesting multiple, independent introduction events with subsequent intra-hospital transmission. In contrast, *vanB*-VREfm was more diverse within hospitals versus between them (Figure 3c), $P = 0.0001$), consistent with widespread dissemination and long-term establishment of these strains across institutions within the Victorian healthcare system.

Virulence

To determine whether increased virulence played a role in the rise of *vanA*-VREfm, isolates were interrogated for the presence of putative virulence factors. As shown in Figure 2, only two isolates (both ST17) had the *cylA* gene, required for the expression of cytolysin.[42] No isolates had the *ptsD* gene,[43] which has been associated with clinical infection, or *bepA*, which has been implicated in biofilm formation.[44] Genes encoding agglutination substance (*agg/asa1*[45]) or gelatinase (*gelE*[46]) were also not detected; these are thought to be more prevalent in *Enterococcus faecalis* than *E. faecium*.[47]

Microbial surface components recognizing adhesive matrix molecules including *acm*,[48] *ecbA*[49] and *sgrA*[49] were also investigated. *acm* was present in 98.5% of study samples, across nearly all STs. *ecbA* was present in 62 isolates (30.5% of *vanA*-VREfm; 15.8% of *vanB*-VREfm), including all ST192, ST400 and ST893 isolates, and most ST17 and ST203 isolates (71% and 97%, respectively). *sgrA* was present in 294 isolates (91.9%, of *vanA*-VREfm; 89.2% of *vanB*-VREfm) from nearly all STs. No isolates had *pilA* and *pilB*, which are also potentially involved in adhesion.[50]

Finally, the gene *esp*, which has previously been associated with hospital outbreaks[51] and encodes enterococcal surface protein,[52] was investigated. While repeat regions made it impossible to completely assemble the *esp* gene using short-read data (see Figure S4), we found that 266 isolates had at least 80% coverage (81.6% of *vanA*-VREfm; 86.2% of *vanB*-VREfm), suggesting that these isolates were likely *esp*-positive. *esp* was more prevalent in bacteraemia due to VREfm than VSEfm (17 of 20 versus 1 of 9, $P = 0.0001$, excluding the VSEfm that had potentially lost the *vanA* operon).

The *vanA*-harbouring plasmid from one isolate per ST was completely assembled using long-read data; no putative virulence genes co-localized with *vanA* on the same plasmid (Figure S5).

Comparative genomics of vanA-VREfm

To assess whether a single *vanA*-harbouring plasmid or Tn*1546* transposon was associated with the rise in *vanA*, we compared the *vanA*-harbouring plasmids from the long-read assemblies. *vanA*-harbouring plasmids from ST78 and ST796 had the greatest homology (Figure 4), with similar plasmid backbones and 100% amino acid identity of the Tn*1546* transposon, which carries the *vanA* gene cluster (Figure S5). The main difference between these plasmids was a large insertion in ST796 compared with ST78, carrying *repB*, IS*1216* and two hypothetical proteins (Figure S5). The *vanA*-harbouring plasmid from ST80 was similar to that from ST796 (Figure S6), but had lost IS*Efa7-birA* from the plasmid backbone (Figure S5) and had a 48 bp deletion in *vanS* compared with the ST796 plasmid. Both *vanA*-harbouring plasmids from the ST203

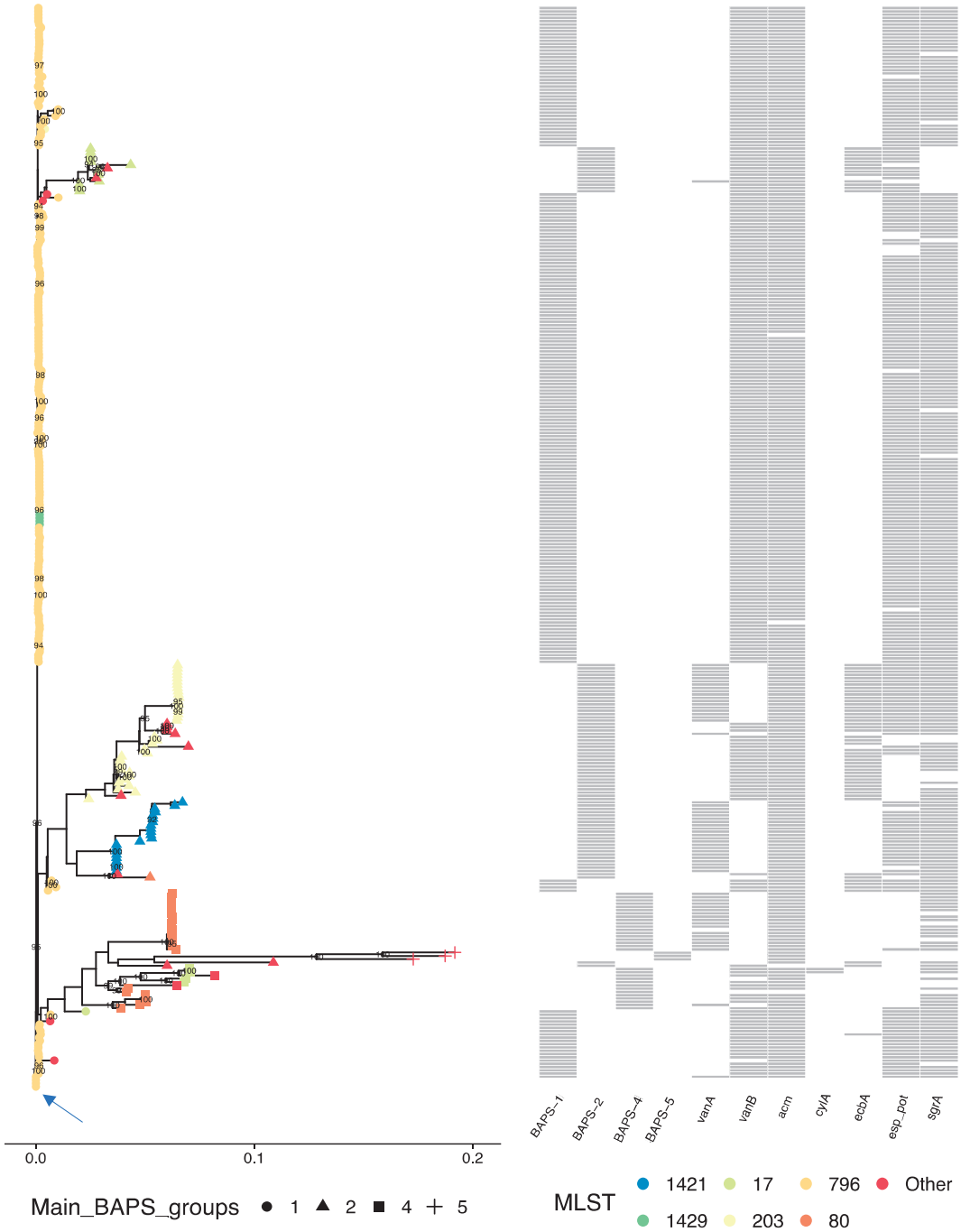

Figure 2. Maximum likelihood tree of *E. faecium* in Australia. Recombination-adjusted core SNPs (6497) from all 331 isolates (321 VREfm and 10 VSEfm) were concatenated and used to produce a maximum likelihood tree in RAxML[31] under a general time reversible model with gamma rate heterogeneity. One thousand bootstrap replicates were performed to assess confidence in the phylogeny. BAPS-3 is not shown; AUSMDU00004157 (an ST54 VSEfm) was the only isolate in this clade, and was >3800 SNPs from all other isolates, so it was excluded from this figure for easier visualization. Branches with bootstrap support >90 are shown. Many branches had very low bootstrap support, thus hierarchical Bayesian analysis of population structure (hierBAPS) was used to validate further the clusters identified.[32] Main BAPS groups are shown. Reference is identified with an arrow. No isolates had *ebpA*, *hyl*, *gelE*, *agg*, *pilA* or *pilB*; therefore, these genes have not been shown.

and ST1421 isolates had unique plasmid backbones and Tn*1546* transposons (Figure S5).

Aligning reads from each isolate to the above plasmids revealed that the majority of isolates likely shared the same plasmid as the representative isolate from their ST (see Excel file on alignments in Supplementary data). The exception to this was ST80, where nearly all isolates appeared to have the same *vanA*-harbouring plasmid as the one *vanA*+ (and *vanB*+) isolate from ST796 (Table S5).

Australian isolates in global context

Compared with published sequences, the *vanA*-harbouring plasmids from ST78, ST796 and ST80 shared closest homology to a *vanA*-harbouring plasmid from VREfm in Demark[38] (Table S6). The *vanA*-harbouring plasmid from ST203 was most closely related to plasmids from several STs of VREfm isolated in Western Australia and Queensland,[24] while the plasmid from ST1421 was closely-related to that from another ST1421 *E. faecium*, isolated in New South Wales in 2014[53] (Table S6). Phylogenetic analysis based on core SNPs from our isolates and global *E. faecium* strains showed that, with the exception of those from ST796, Australian isolates were widely distributed across the tree (Figure S7), which is suggestive of ongoing, intercontinental transmission of *E. faecium*, with subsequent clonal expansion.

Discussion

For the past two decades, the majority of VREfm in Australia have been *vanB* genotype.[13,17,18,20–22,54–56] Recently, however, *vanA*-VREfm increased dramatically. In this population-based study, we looked at factors potentially contributing to this change.

First, we examined the genetic diversity of *vanA*-VREfm across our state, to assess for potential transmission. Consistent with the hypothesis of van Hal *et al.*,[24] we found that a single clone is not driving this increase; instead, our findings suggest that multiple introductions of *vanA*-VREfm have occurred, with subsequent spread within hospitals. This pattern is similar to that of early *vanB*-VREfm in Australia,[57] though the latter has become dominated by a single ST in Victoria (ST796).[22]

Another possibility was that differences in putative virulence factors might be mediating the rise in *vanA*-VREfm. However, few putative virulence factors were identified, and their prevalence was largely consistent with previous data from *vanB*-VREfm in Australia,[58] and similar to or lower than those found in *vanA*-VREfm from settings (e.g. in China[59] or Brazil[60]). While much is still unknown about virulence factors in VREfm (compared with *E. faecalis*), this suggests that the rise in *vanA*-VREfm in Australia, and the continued success of *vanB*-VREfm, is not due to changes in virulence. Consistent with this, the rate of *E. faecium* BSI in our

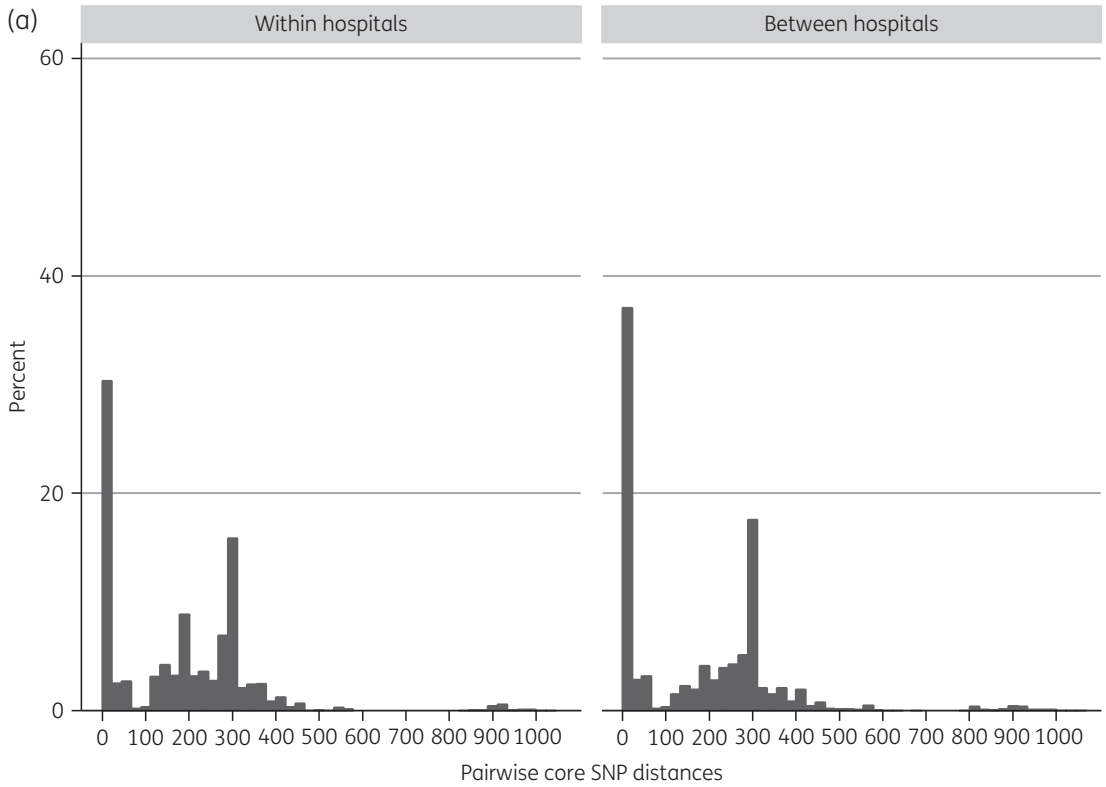

Figure 3. Recombination-adjusted pairwise core SNPs separating Australian isolates. (a) Overall pairwise comparisons, including isolates with both *vanA* and *vanB and* VSEfm. Pairwise SNP comparisons are shown regardless of *van* gene. There were 54 567 total pairwise comparisons, after excluding those involving isolates with missing data on hospital/HCN, those from unknown clinics, and pairs within the same patients. Between hospitals, the median number of pairwise SNPs was 193 (IQR 11–298), while within hospitals, the median was 196 (IQR 13–297). For easier visualization, as one VSEfm isolate (AUSMDU00004157) was a minimum of 3893 SNPs from all others, only comparisons between the 330 other isolates are shown in the figure. (b) Pairwise comparisons between isolates with *vanA* only. There were 1706 such comparisons (233 within hospitals, 1473 between hospitals). Median number of pairwise SNPS within hospitals was 10 (IQR 1–357) compared with a median of 356 (IQR 179–416) between hospitals ($P = 0.0001$). (c) Pairwise comparisons between isolates with *vanB* only. There were 33 632 such comparisons (3104 within hospitals, 30 528 between hospitals). Median number of pairwise SNPs within hospitals was 40 (IQR 9–206) compared with a median of 15 (IQR 8–174) between hospitals ($P = 0.0001$). Isolates that were both *vanA*- and *vanB*-positive ($n = 2$) and isolates that were *vanA*- and *vanB*-negative ($n = 10$) were excluded from (b) and (c).

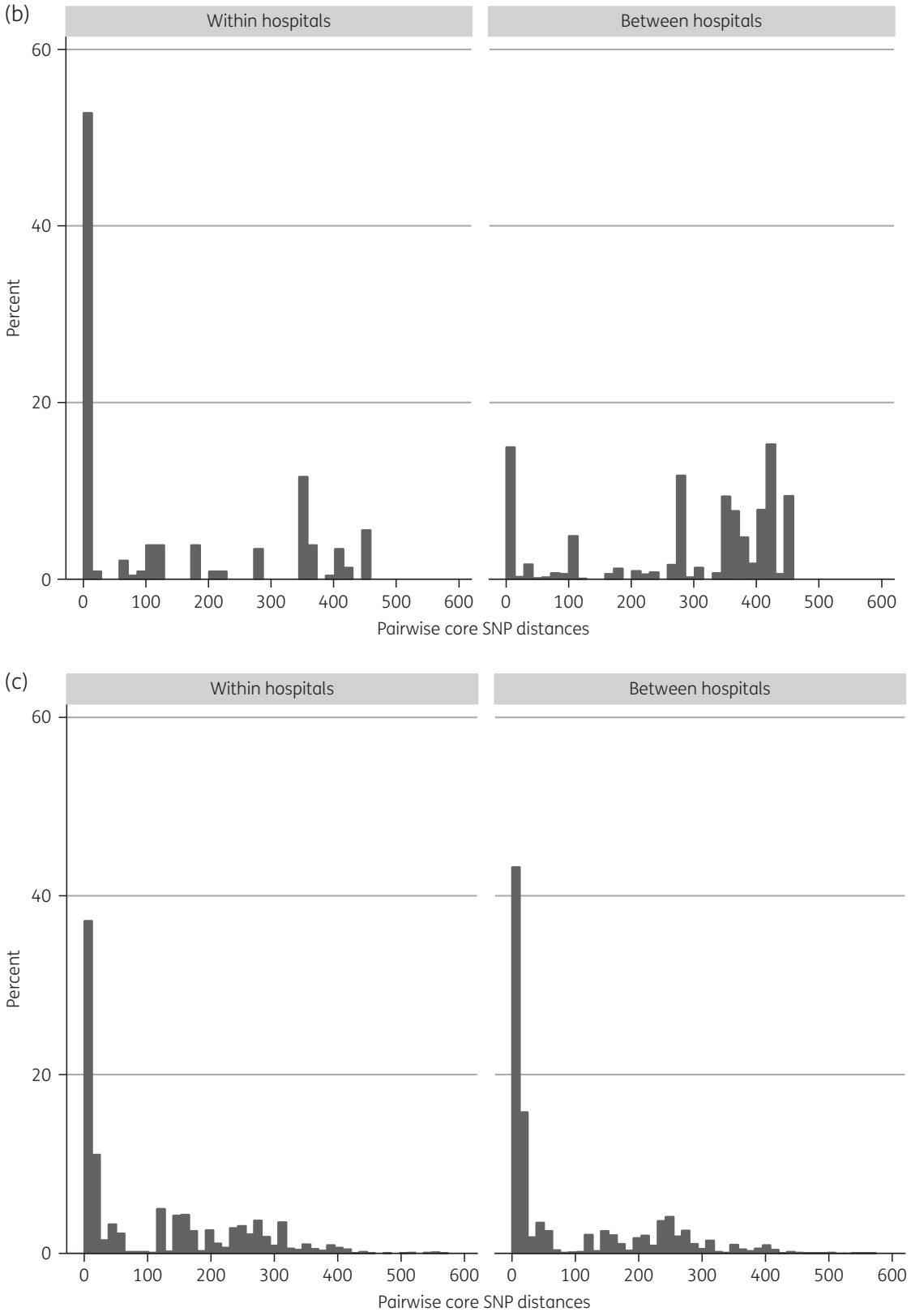

Figure 3. Continued

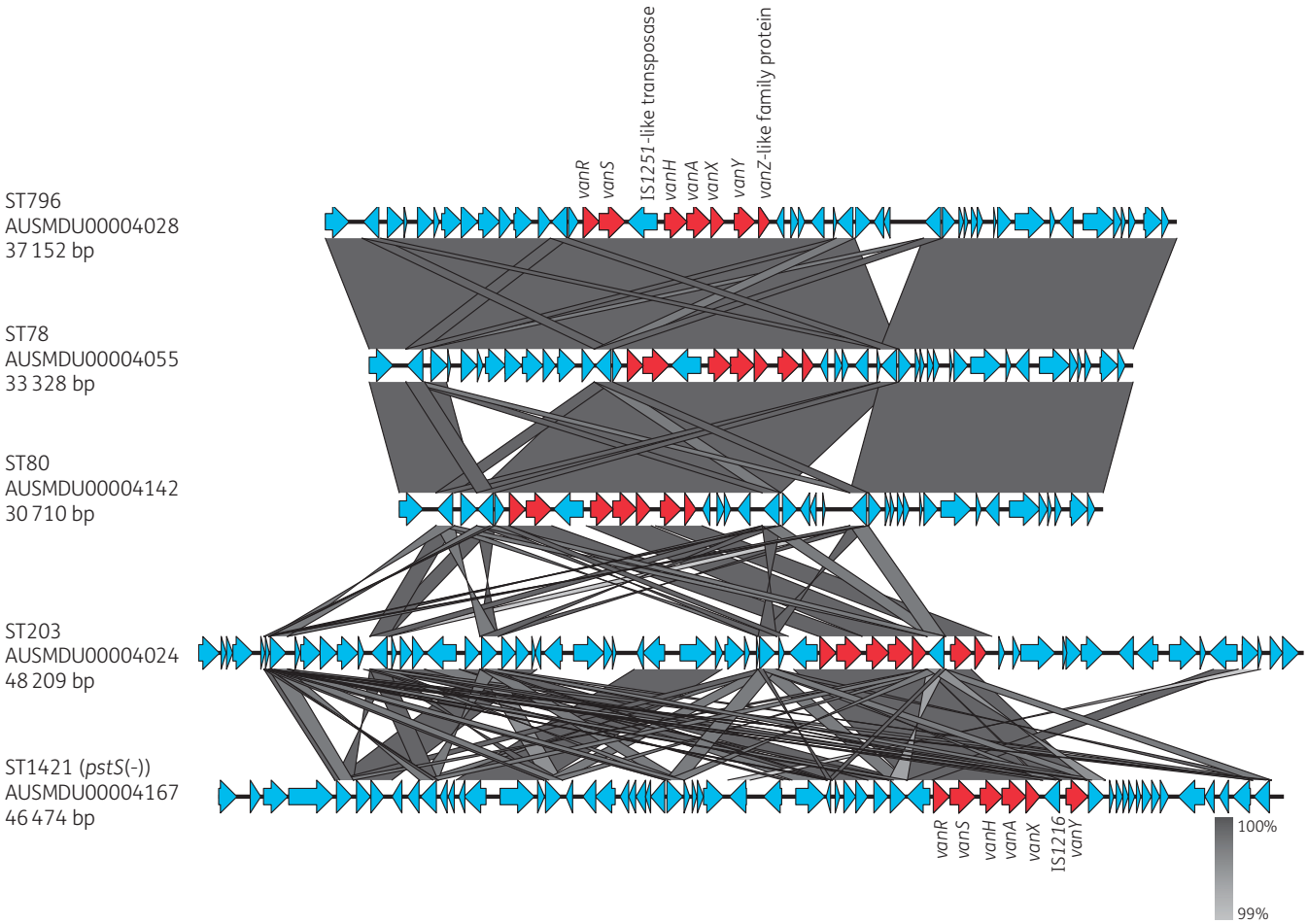

Figure 4. Homology between *vanA*-harbouring plasmids. PacBio single molecule real-time sequencing was performed on a *vanA*-positive isolate from each of the main *vanA*-associated STs (STs 78, 80, 203, 796 and 1421). Plasmids harbouring *vanA* were initially identified using ABRicate (https://github.com/tseemann/abricate), and subsequently annotated using the *Enterococcus* database (https://github.com/tseemann/prokka/blob/master/db/genus/Enterococcus). Figure was produced using EasyFig (v. 2.2.2); blastn was used to compare sequence homology. Only hits with a minimum length of 200 bp and at least 99% identity are shown for clarity. *van* genes are indicated in red, all other genes are indicated in blue.

study was similar to that of another population-based study in Denmark.[3] This has important implications for hospital infection control, as the patient population affected by VREfm has also remained consistent over time, environmental factors may be mediating the observed changes in VREfm epidemiology. Further studies are needed to investigate this hypothesis.

Finally, we investigated whether a single *vanA*-harbouring plasmid and/or Tn*1546* transposon was being transmitted within and across STs, as has been shown in Denmark.[38] A unique *vanA*-harbouring plasmid was identified using representatives from each ST, nearly all of which had different Tn*1546* transposons. Mapping reads to these assembled plasmids also suggested that each major ST was likely dominated by a different plasmid. This further suggests that horizontal gene transfer is not the sole driver of *vanA*-VREfm in this context.

In addition to gaining new insights in *vanA* epidemiology in this context, we have also provided valuable information on *vanB*-VREfm; given that we found genetic diversity was slightly lower across hospitals compared with within them, this suggests a potential community reservoir for these strains. Transmission of VREfm has previously been shown in nursing homes,[61] with carriage strains closely related to those causing infection; investigation is warranted to determine if this is also the case in Australia. The overall low diversity of *vanB*-VREfm could also be due to a 'diffusion' effect, wherein *vanB*-VREfm has slowly spread across hospitals through years of patient exchanges. Such a pattern would be reminiscent of that observed in Denmark;[62] however, while in Denmark the 'diffusion' process was marked by a hub-and-spoke pattern, with patients mainly being transferred between regional hospitals and central hospitals in Copenhagen, in Victoria, the process was likely more random. If such a 'diffusion' scenario is correct, we hypothesize that the pattern of genetic diversity of *vanA*-VREfm will eventually mirror that of *vanB*.

This work has a number of key strengths. First, by including all VREfm collected across an entire state, we have provided a comprehensive assessment of local strain diversity and the prevalence of clinically relevant *van* genotypes, as well as a baseline to monitor for changes in VREfm population. Second, we used local

reference genomes for all SNP-based analyses. As using more genetically distant reference genomes can result in false-positive SNPs, and/or loss of information due to differences in genes present in the study sample, but not in the reference, this increased the accuracy of our short-read analyses. By sharing our novel, complete VREfm genomes (from five different STs), we have also provided an invaluable resource for public health as well as future studies on *E. faecium*. Finally, the use of long-read data has also allowed us to assemble and completely characterize the *vanA*-harbouring plasmids in each of the major STs—a task not always feasible with short-read data alone. In doing so, we illustrate that horizontal gene transfer is unlikely to be the main driver of *vanA*-VREfm in this context, suggesting the conventional short-read analysis may be adequate for routine public health surveillance.

This work has several limitations. First, screening protocols for VREfm colonization are not standardized across Victoria; each hospital adhered to its own infection control guidelines, thus we may have an incomplete capture of colonizing strains. This could introduce bias and cause us to miss potential transmission. Another limitation is that we did not have any data on patient contact or intra-hospital transfer (except where isolates were collected at different hospitals), preventing us from confirming direct person-to-person transmission events. However, we do not think this affects our overall population-level inferences, given the dramatically lower genetic diversity of *vanA* within hospitals compared with *vanB*. Finally, as this was predominantly a laboratory-based study, we also did not have data on clinical presentation at the time of sample collection and were therefore unable to discriminate clinical infection aside from bacteraemia. Thus, we may underestimate the overall prevalence of infection (though the most clinically important strains, those causing bacteraemia, have been included).

Herein, we have provided a comprehensive snapshot of VREfm strain diversity across Victoria, representing the first population-level study of VREfm from Australia. In doing so, we have shown that Australian *vanA*-VREfm are highly similar to those from other geographical settings, suggesting transmission of VREfm between continents is ongoing. This highlights the critical importance of continued VREfm prevention and control on a global scale. We also highlight the complexity of *E. faecium* genomic epidemiology in this setting, revealing key differences in transmission dynamics of clinically relevant *van* genotypes; while *vanB*-VRE continues to predominate in Australia, repeated introductions and dissemination of *vanA*-VREfm within Victorian hospitals suggest this is changing. To understand better the clinical significance of this shift, prospective studies with detailed corresponding clinical data are needed.

Acknowledgements

We thank the participating laboratories for their provision of isolates and the clerical and technical assistance of staff at MDU PHL. We would also like to thank Dr Sharon Peacock (Department of Medicine, University of Cambridge, Cambridge, UK; Wellcome Trust Sanger Institute, Hinxton, UK; London School of Hygiene and Tropical Medicine, London, UK) and Dr Kathy Raven (Department of Medicine, University of Cambridge, Cambridge, UK) for helpful discussions about global VREfm, and Dr Willem van Schaik (University of Birmingham, UK) for his valuable feedback on our preprint.

Funding

This study was conducted as part of routine work at the Microbiological Diagnostic Unit Public Health Laboratory. R. S. L. is supported by a Fellowship from the Canadian Institutes of Health Research (funding reference no. 152448), S. L. B. holds an Australian Government Research Training Program Scholarship, J. C. K. is supported by an early career fellowship from the National Health and Medical Research Council, Australia (GNT1142613) and B. P. H. has a Practitioner Fellowship from the National Health and Medical Research Council (GNT1105905), and is also supported by the Centre of Research Excellence on Emerging Infectious Diseases (National Health and Medical Research Council GNT1102962). The Microbiological Diagnostic Unit Public Health Laboratory is funded by the Victorian Government, Australia. These funding agencies had no role in study design, data collection and interpretation, or the decision to submit for publication.

Transparency declarations

None to declare.

Author contributions

R. S. L. designed and ran the primary analyses, interpreted results, made the tables and figures, and wrote the first draft of the manuscript. A. G. designed analyses, advised on bioinformatics approaches, helped interpret results, and contributed to writing the manuscript. S. L. B. assembled the PacBio genomes, provided input on analyses, and helped interpret results. J. S. recruited all labs, coordinated submission of isolates and data from primary diagnostic laboratories, and did the initial processing of samples at MDU PHL. S. B. designed sample collection and advised on the initial interpretation of the results. G. P. C. did the laboratory work to generate the PacBio sequences. J. C. K., M. B. S., D. M. B. and T. S. provided bioinformatics tools, and advised on bioinformatics analyses. T. P. S. and B. P. H. conceived the cross-sectional study, helped decide on analyses, and edited the manuscript. All authors critically reviewed the manuscript for content.

Supplementary data

Tables S1 to S6, Figures S1 to S7 and an Excel file with *van*A-harbouring plasmid alignment statistics appear as Supplementary data at *JAC* Online.

References

1 Gilmore MS, Lebreton F, van Schaik W. Genomic transition of enterococci from gut commensals to leading causes of multidrug-resistant hospital infection in the antibiotic era. *Curr Opin Microbiol* 2013; **16**: 10–6.

2 Guzman Prieto AM, van Schaik W, Rogers MR *et al*. Global emergence and dissemination of enterococci as nosocomial pathogens: attack of the clones? *Front Microbiol* 2016; **7**: 788.

3 Pinholt M, Ostergaard C, Arpi M *et al*. Incidence, clinical characteristics and 30-day mortality of enterococcal bacteraemia in Denmark 2006-2009: a population-based cohort study. *Clin Microbiol Infect* 2014; **20**: 145–51.

4 DiazGranados CA, Zimmer SM, Klein M *et al*. Comparison of mortality associated with vancomycin-resistant and vancomycin-susceptible enterococcal bloodstream infections: a meta-analysis. *Clin Infect Dis* 2005; **41**: 327–33.

5 Lebreton F, Depardieu F, Bourdon N *et al*. D-Ala-D-Ser *VanN*-type transferable vancomycin resistance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2011; **55**: 4606–12.

6 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.

7 Xu X, Lin D, Yan G *et al*. *vanM*, a new glycopeptide resistance gene cluster found in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2010; **54**: 4643–7.

8 Courvalin P. Vancomycin resistance in Gram-positive cocci. *Clin Infect Dis* 2006; **42**: S25–34.

9 Top J, Willems R, Bonten M. Emergence of CC17 *Enterococcus faecium*: from commensal to hospital-adapted pathogen. *FEMS Immunol Med Microbiol* 2008; **52**: 297–308.

10 Ridwan B, Mascini E, van der Reijden N *et al*. What action should be taken to prevent spread of vancomycin resistant enterococci in European hospitals? *BMJ* 2002; **324**: 666–8.

11 Zirakzadeh A, Patel R. Vancomycin-resistant enterococci: colonization, infection, detection, and treatment. *Mayo Clin Proc* 2006; **81**: 529–36.

12 Stinear TP, Olden DC, Johnson PDR *et al*. Enterococcal *vanB* resistance locus in anaerobic bacteria in human faeces. *Lancet* 2001; **357**: 855–6.

13 Howden BP, Holt KE, Lam MM *et al*. Genomic insights to control the emergence of vancomycin-resistant enterococci. *MBio* 2013; **4**: pii: e00412–13.

14 Simner PJ, Adam H, Baxter M *et al*. Epidemiology of vancomycin-resistant enterococci in Canadian hospitals (CANWARD study, 2007 to 2013). *Antimicrob Agents Chemother* 2015; **59**: 4315–7.

15 Werner G, Coque TM, Hammerum AM *et al*. Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro Surveill* 2008; **13**: pii=19046.

16 Coombs GW, Pearson JC, Le T *et al*. Australian Enterococcal Sepsis Outcome Progamme, 2011. *Commun Dis Intell Q Rep* 2014; **38**: E247–52.

17 Christiansen K, Turnidge J, Gottlieb T *et al*., on behalf of the Australian Group for Antimicrobial Resistance. Antimicrobial susceptibility and VRE characterisation report of enterococcus isolates from the Australian Group on Antimicrobial Resistance (AGAR): 2010 Surveillance Report. 2010. http://agargroup.org.au/wp-content/uploads/2017/08/ENTE-2010-Report-Final-secure.pdf.

18 Coombs GW, Pearson J, Daley DA *et al*., for the Australian Group on Antimicrobial Resistance. Australian Group on Antimicrobial Resistance (AGAR) Australian Enterococcal Sepsis Outcome Programme (AESOP) Annual Report 2013. 2013. http://agargroup.org.au/wp-content/uploads/2017/08/2013-AGAR-AESOP-Report_public.pdf.

19 Coombs GW, Pearson JC, Daley DA *et al*. Molecular epidemiology of enterococcal bacteremia in Australia. *J Clin Microbiol* 2014; **52**: 897–905.

20 Coombs GW, Daley DA, Lee YT *et al*., for the Australian Group on Antimicrobial Resistance. Australian Group on Antimicrobial Resistance (AGAR) Australian Enterococcal Sepsis Outcome Programme (AESOP) Annual Report 2014. 2014. http://agargroup.org.au/wp-content/uploads/2017/08/2014-AGAR-AESOP-CDI-FINAL.pdf.

21 Coombs GW, Daley DA, on behalf of the Australian Group for Antimicrobial Resistance AGAR. Australian Enterococcal Sepsis Outcome Program (AESOP) 2015 Final Report. 2015. http://agargroup.org.au/wp-content/uploads/2017/08/July.2016.AESOP-2015-Final-Report-2016.pdf.

22 Buultjens AH, Lam MM, Ballard S *et al*. Evolutionary origins of the emergent ST796 clone of vancomycin resistant *Enterococcus faecium*. *PeerJ* 2017; **5**: e2916.

23 Coombs GW, Daley DA, on behalf of the Australian Group on Antimicrobial Resistance. Australian Enterococcal Sepsis Outcome Program (AESOP) 2016 Final Report. 2017. http://agargroup.org.au/wp-content/uploads/2017/08/AESOP-2016-Final-Report-2017.pdf.

24 van Hal SJ, Espedido BA, Coombs GW *et al*. Polyclonal emergence of vanA vancomycin-resistant *Enterococcus faecium* in Australia. *J Antimicrob Chemother* 2017; **72**: 998–1001.

25 Bankevich A, Nurk S, Antipov D *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012; **19**: 455–77.

26 Zankari E, Hasman H, Cosentino S *et al*. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother* 2012; **67**: 2640–4.

27 Chen L, Zheng D, Liu B *et al*. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res* 2016; **44**: D694–7.

28 Li H. *Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM*. *arXiv* 2013: 1303.3997v2.

29 Garrison E, Marth G. *Haplotype-based Variant Detection from Short-read Sequencing*. *arXiv* 2012: 1207.3907v2.

30 Didelot X, Wilson DJ. ClonalFrameML: efficient inference of recombination in whole bacterial genomes. *PLoS Comput Biol* 2015; **11**: e1004041.

31 Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 2014; **30**: 1312–3.

32 Cheng L, Connor TR, Sirén, J *et al*. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol* 2013; **30**: 1224–8.

33 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–9.

34 Kearse M, Moir R, Wilson A *et al*. Geneious Basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* 2012; **28**: 1647–9.

35 Lebreton F, van Schaik W, McGuire AM *et al*. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *MBio* 2013; **4**: pii: e00534-13.

36 Raven KE, Gouliouris T, Brodrick H *et al*. Complex routes of nosocomial vancomycin-resistant *Enterococcus faecium* transmission revealed by genome sequencing. *Clin Infect Dis* 2017; **64**: 886–93.

37 Raven KE, Reuter S, Reynolds R *et al*. A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland. *Genome Res* 2016; **26**: 1388–96.

38 Pinholt M, Gumpert H, Bayliss S *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.

39 Chacko KI, Sullivan MJ, Beckford C *et al*. Genetic basis of emerging vancomycin, linezolid, and daptomycin heteroresistance in a case of persistent *Enterococcus faecium* bacteremia. *Antimicrob Agents Chemother* 2018; **62**: pii: e02007-17.

40 McGann P, Bunin JL, Snesrud E *et al*. Real time application of whole genome sequencing for outbreak investigation—what is an achievable turnaround time? *Diagn Microbiol Infect Dis* 2016; **85**: 277–82.

41 European Committee on Antimicrobial Susceptibility Testing. *Breakpoint Tables for Interpretation of MICs and Zone Diameters, Version 7.0*. 2017. http://www.eucast.org/.

42 Cox CR, Coburn PS, Gilmore MS. Enterococcal cytolysin: a novel two component peptide system that serves as a bacterial defense against eukaryotic and prokaryotic cells. *Curr Protein Pept Sci* 2005; **6**: 77–84.

43 Zhang X, Top J, de Been M *et al*. Identification of a genetic determinant in clinical *Enterococcus faecium* strains that contributes to intestinal colonization during antibiotic treatment. *J Infect Dis* 2013; **207**: 1780–6.

44 Paganelli FL, Huebner J, Singh KV *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.

45 Galli D, Lottspeich F, Wirth R. Sequence analysis of *Enterococcus faecalis* aggregation substance encoded by the sex pheromone plasmid pAD1. *Mol Microbiol* 1990; **4**: 895–904.

46 Su YA, Sulavik MC, He P *et al*. Nucleotide sequence of the gelatinase gene (*gelE*) from *Enterococcus faecalis* subsp. *liquefaciens*. *Infect Immun* 1991; **59**: 415–20.

47 Comerlato CB, de Resende MCC, Caierao J *et al*. Presence of virulence factors in *Enterococcus faecalis* and *Enterococcus faecium* susceptible and resistant to vancomycin. *Mem Inst Oswaldo Cruz* 2013; **108**: 590–5.

48 Nallapareddy SR, Weinstock GM, Murray BE. Clinical isolates of *Enterococcus faecium* exhibit strain-specific collagen binding mediated by *Acm*, a new member of the MSCRAMM family. *Mol Microbiol* 2003; **47**: 1733–47.

49 Hendrickx AP, van Luit-Asbroek M, Schapendonk CM *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* 2009; **77**: 5097–106.

50 Hendrickx AP, Bonten MJ, van Luit-Asbroek M *et al*. Expression of two distinct types of pili by a hospital-acquired *Enterococcus faecium* isolate. *Microbiology (Reading, Engl)* 2008; **154**: 3212–23.

51 Willems RJL, Homan W, Top J *et al*. Variant *esp* gene as a marker of a distinct genetic lineage of vancomycin resistant *Enterococcus faecium* spreading in hospitals. *Lancet* 2001; **357**: 853–5.

52 Leavis H, Top J, Shankar N *et al*. A novel putative enterococcal pathogenicity island linked to the *esp* virulence gene of *Enterococcus faecium* and associated with epidemicity. *J Bacteriol* 2004; **186**: 672–82.

53 Carter GP, Buultjens AH, Ballard SA *et al*. Emergence of endemic MLST non-typeable vancomycin-resistant *Enterococcus faecium*. *J Antimicrob Chemother* 2016; **71**: 3367–71.

54 Christiansen K, Turnidge J, Bell J, on behalf of the Australian Group for Antimicrobial Resistance. Enterococcus spp. Survey: 2005 Antimicrobial Susceptibility Report. 2005. http://agargroup.org.au/wp-content/uploads/2017/08/Final-Enterococcus-2005.pdf.

55 Christiansen K, Turnidge J, Gottlieb T *et al*., on behalf of the Australian Group for Antimicrobial Resistance. Enterococcus species Survey: 2007 Antimicrobial Susceptibility Report. 2007. http://agargroup.org.au/wp-content/uploads/2017/08/ENTE-2007-final-report-secure.pdf.

56 Christiansen K, Turnidge J, Gottlieb T *et al*., on behalf of the Australian Group for Antimicrobial Resistance. Enterococcus Species Survey: 2009 Antimicrobial Susceptibility Report. 2009. http://agargroup.org.au/wp-content/uploads/2017/08/ENTE-2009-report-final-protected.pdf.

57 Bell J, Turnidge J, Coombs G *et al*. Emergence and epidemiology of vancomycin-resistant enterococci in Australia. *Commun Dis Intell* 1998; **22**: 249–52.

58 Worth LJ, Slavin MA, Vankerckhoven V *et al*. Virulence determinants in vancomycin-resistant *Enterococcus faecium vanB*: clonal distribution, prevalence and significance of *esp* and *hyl* in Australian patients with haematological disorders. *J Hosp Infect* 2008; **68**: 137–44.

59 Yang J, Jiang Y, Guo L *et al*. Prevalence of diverse clones of vancomycin-resistant *Enterococcus faecium* ST78 in a Chinese hospital. *Microb Drug Resist* 2016; **22**: 294–300.

60 Ruzon FI, de Paula, SB, Kanoshiki RL *et al*. Virulence determinants in vancomycin-resistant *Enterococcus faecium vanA* isolated from different sources at University Hospital of Londrina, Paraná, Brazil. *J Microbiol* 2010; **48**: 814–21.

61 Brodrick HJ, Raven KE, Harrison EM *et al*. Whole-genome sequencing reveals transmission of vancomycin-resistant *Enterococcus faecium* in a healthcare network. *Genome Med* 2016; **8**: 4.

62 Pinholt M, Larner-Svensson H, Littauer P *et al*. Multiple hospital outbreaks of *vanA Enterococcus faecium* in Denmark, 2012–13, investigated by WGS, MLST and PFGE. *J Antimicrob Chemother* 2015; **70**: 2474–82.