Consideration of within-patient diversity highlights transmission pathways and antimicrobial resistance gene variability in vancomycin-resistant *Enterococcus faecium*

Martin P. McHugh [1,2]*, Kerry A. Pettigrew[1]†, Surabhi Taori[2], Thomas J. Evans [3], Alistair Leanord[3,4], Stephen H. Gillespie[1], Kate E. Templeton[2] and Matthew T. G. Holden [1]

[1]School of Medicine, University of St Andrews, St Andrews, UK; [2]Medical Microbiology, Department of Laboratory Medicine, Royal Infirmary of Edinburgh, Edinburgh, UK; [3]School of Infection and Immunity, University of Glasgow, Glasgow, UK; [4]Scottish Microbiology Reference Laboratories, Glasgow Royal Infirmary, Glasgow, UK

*Corresponding author. E-mail: mpm20@st-andrews.ac.uk
✕ @mchugh_martin, @templeton_kate, @Infection_StAnd
†Current affiliation: School of Health and Life Sciences, Teesside University, Middlesbrough, UK

Background: WGS is increasingly being applied to healthcare-associated vancomycin-resistant *Enterococcus faecium* (VREfm) outbreaks. Within-patient diversity could complicate transmission resolution if single colonies are sequenced from identified cases.

Objectives: Determine the impact of within-patient diversity on transmission resolution of VREfm.

Materials and methods: Fourteen colonies were collected from VREfm positive rectal screens, single colonies were collected from clinical samples and Illumina WGS was performed. Two isolates were selected for Oxford Nanopore sequencing and hybrid genome assembly to generate lineage-specific reference genomes. Mapping to closely related references was used to identify genetic variations and closely related genomes. A transmission network was inferred for the entire genome set using Phyloscanner.

Results and discussion: In total, 229 isolates from 11 patients were sequenced. Carriage of two or three sequence types was detected in 27% of patients. Presence of antimicrobial resistance genes and plasmids was variable within genomes from the same patient and sequence type. We identified two dominant sequence types (ST80 and ST1424), with two putative transmission clusters of two patients within ST80, and a single cluster of six patients within ST1424. We found transmission resolution was impaired using fewer than 14 colonies.

Conclusions: Patients can carry multiple sequence types of VREfm, and even within related lineages the presence of mobile genetic elements and antimicrobial resistance genes can vary. VREfm within-patient diversity could be considered in future to aid accurate resolution of transmission networks.

Introduction

Enterococcus faecium is a leading nosocomial pathogen causing opportunistic infections mostly in immunocompromised hosts. Antimicrobial resistance is a key concern, particularly against front-line anti-Gram-positive agents amoxicillin and vancomycin.[1–41] Vancomycin-resistant *E. faecium* (VREfm) infections lead to increased length of stay, cost an estimated USD200 per case per day and confer mortality of 23%–47%.[2–7] In 2020, vancomycin resistance of 45.6% was reported among all *E. faecium* bloodstream isolates in Scotland, among the highest rates in Europe.[8]

In healthcare institutions, asymptomatic intestinal carriage of VREfm can lead to shedding into the environment and transfer to other patients or staff, challenging efforts to limit the incidence of nosocomial infections.[9] WGS is increasingly applied to investigate transmission networks and identify control measures.[10,11] Many WGS-based analyses of bacterial outbreaks, however, rely on analysing single-colony picks from clinical samples assuming that this represents the entire infecting or colonizing population within individual patients.[12] It is increasingly recognized that within-patient diversity of bacterial populations can be significant and can influence transmission network resolution.[13–19] Several

studies have identified that individual patients can carry multiple strains of *E. faecium* concurrently, but few have applied this to transmission resolution.[20–24]

In this study, we aimed to identify within-patient diversity of VREfm from rectal screening swabs and determine how this affects transmission inference in a 1-month snapshot on a haematology unit. We designed a sampling strategy to reliably detect within-patient diversity and supplemented short-read and long-read sequencing to generate high-quality reference genomes to identify genomic variants in the isolate collection.

Materials and methods

Isolates

Rectal swabs were collected at admission and on all inpatients on the haematology unit developing febrile neutropenia (neutrophils <0.9× 10^9/L or <1.0×10^9/L and falling after chemotherapy, plus body temperature ≥38°C). Swabs were plated to CHROMID® VRE agar (bioMérieux, Marcy-l'Étoile, France), species identification and vancomycin resistance were confirmed with MALDI–TOF (Microflex instrument, Bruker, Billerica, USA) and VITEK-2 (bioMérieux) with EUCAST breakpoints. All purple colonies from VREfm positive plates were stored at −80°C in a Microbank cryovial (Pro-Lab Diagnostics, Birkenhead, UK). Any VREfm isolated from clinical samples within 60 days of a rectal positive were also stored. Patient metadata were retrieved from electronic records and movements visualized with HAIviz v.0.3 (https://haiviz.beatsonlab.com/). This work was approved by the NHS Scotland BioRepository Network (ref. TR000126) and the University of St Andrews Research Ethics Committee (ref. MD12651).

Genome sequencing

Cryovials were re-plated on CHROMID® VRE agar, and 14 random purple colonies were incubated overnight in 5 mL of brain heart infusion broth (Oxoid). Cells were pelleted and DNA extracted using the DNA Mini kit on a QiaSymphony instrument (Qiagen).

Short-read libraries were prepared using the Nextera XT kit (Illumina, San Diego, USA) and sequenced with a MiSeq instrument (Illumina) using 300 bp paired-end reads on a 600-cycle v3 reagent kit.

For long-read sequencing, isolates VRED06-02 (ST1424) and VRED06-10 (ST80) were selected at random from the first sample with multiple sequence types (STs) detected (sample VRED06 from patient P49). Long-read libraries were generated with the LSK109 Ligation Sequencing Kit (Oxford Nanopore Technologies, Oxford, UK) and sequenced for 8 h using an R9.4 flow cell on a GridION sequencer (Oxford Nanopore Technologies) with high accuracy basecalling in MinKNOW v.19.12.6 (Oxford Nanopore Technologies).

Sequence data from this study have been deposited in the NCBI under BioProject accession number PRJNA877253 (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA877253).

Sequence assembly and mapping

Short reads were quality trimmed with Trimmomatic v.0.32.[25] Multilocus STs (MLSTs) were determined with SRST2 v.0.2.0[26] and the *E. faecium* pubMLST database.[27] A core alignment was generated by mapping short reads to the reference genome(s) with Snippy v.4.6.0 default settings (https://github.com/tseemann/snippy) and masking all putative transposases, prophage regions and recombination blocks. Recombination blocks were identified with Gubbins v.2.4.1 and included in the Snippy mask.[28] Non-ACGT bases were converted to N with snippy-clean and a core SNP alignment generated using snp-sites v.2.5.1.[29] The 130 ST80 genomes mapped to VRED06-10 generated an initial alignment of 2 814 943 bases, 202 738 bases were masked and the final alignment contained 96 variant

sites; the 97 ST1424 genomes mapped to VRED06-02 generated an initial alignment of 2 945 113 bases, 227 540 bases were masked and 13 variant sites remained. Maximum-likelihood phylogenies were constructed with IQTree v.2.0.3 with automatic model selection and 1000 ultrafast bootstraps.[30–32] Phylogenies were visualized with iTOL.[33] Short-read assemblies were generated with Unicycler v.0.4.8[34] and searched for antimicrobial resistance genes using Abricate v.1.0.1 (https://github.com/tseemann/abricate) with default settings and the ResFinder (20 April 2021) database.[35]

Long reads <1000 bp were removed with Nanofilt v.2.7.1,[36] and adapters were trimmed and chimaeras split with Porechop v.0.2.3 (https://github.com/rrwick/Porechop). Hybrid assemblies were generated with Trycycler v.0.3.3,[37] reads were split into 12 subsamples and three assemblies made with four different long-read assemblers (Flye v.2.8.1, Redbean v.2.5, Raven v.1.1.10 and Miniasm v.0.1.3),[38–41] giving 12 assemblies in total. A consensus assembly was generated and first corrected with the long-read polisher Medaka v.0.11.5 (https://github.com/nanoporetech/medaka), and then 2–3 cycles of the short-read polisher Pilon v.1.23.[42] Assembly quality was assessed with assembly-stats v.1.0.1 (https://github.com/sanger-pathogens/assembly-stats), Ideel (https://github.com/phiweger/ideel, committed 26 June 2018) and Busco v.4.1.4 in genome mode with the lactobacillales_odb10 database.[43]

Polished assemblies were annotated with Prokka v.1.14.6[44] using the Aus0004 reference genome (Accession CP003351) with the –proteins option. Abricate identified matches to ResFinder (20 April 2021) and PlasmidFinder (12 July 2021) databases[45–47] and putative prophages were identified with PHASTER.[48] Plasmid copy numbers were estimated using short reads and Snippy: average depth for each plasmid was divided by the average depth of the chromosome.

Plasmids in the polished assemblies were compared to each other with Mash v.2.2.2.[49] To detect plasmids, those present in the two polished assemblies were used as references against all short-read sets in Snippy and considered present if ≥85% bases were called with ≤5 SNPs/1000 bp.[50]

Transmission network inference

All short reads were mapped to the VRED06-10 ST80 reference chromosome with Snippy, the V24 *E. faecium* ST80 genome (Accession CP036151) was included as an outgroup. An alignment of 2 814 943 bases was generated and 1 418 409 bases masked as before. A posterior set of phylogenies were generated with MrBayes v.3.2.7.[51] Two MCMC runs of four coupled chains were run for 5 000 000 generations, sampling every 5000. The final standard deviation of split frequencies was 0.013, the log-likelihood was stable and the effective sample size of all parameters was >800. A random sample of 100 posterior trees was input to Phyloscanner v.1.6.6.[52] Sankoff parsimony reconstruction was performed with a *k* parameter of 281 494.5, equivalent to a within-patient diversity threshold of 10 SNPs as used in other studies.[53] A transmission network was constructed in Cytoscape v.3.9.0[54] showing edges with complex or transmission state and >0.5 probability. The role of smaller numbers of colony picks on transmission resolution was investigated by repeating the above with the first three, five and 10 genomes per sample.

Statistical analysis

To determine the optimal number of colonies to analyse for within-sample diversity a power calculation was performed as described by Huebner *et al.*:[55]

$$q\hat{\ }n = (1 - P)$$

where $q = 1$ – concentration of organisms, $\hat{\ }$ = exponentiation operator, n = number of colonies sequenced and P = probability of finding one or more variants.

The minimum number of colonies required to detect different proportions of variants in the population were calculated (Table S1, available as Supplementary data at JAC Online). Next, the expected variant proportion was derived by analysis of 135 VREfm genomes from a study by Moradigaravand et al. (details in Supplementary File and Table S2).[20] This analysis found rectal VREfm populations harbour minority variants at 20%–50% of the total population based on sequencing roughly five colonies. However, no population variants were identified in blood cultures based on sequencing ~10 colonies.[20] For rectal samples, we determined 14 colonies per sample would detect a variant at 20% of the population with 95% confidence, and for blood cultures we sequenced single colonies to identify the sequence type causing invasive disease.

Presence/absence matrices of antimicrobial resistance (AMR) genes were generated in R v.4.0.5 using ggplot2 and patchwork packages.[56–58]

Results and discussion

Epidemiological context

This study was performed over 1 month in 2017 on an inpatient unit for haematological malignancies, split into two wards (A and B). VREfm rectal screening was performed on all new admissions and any inpatients with febrile episodes to inform patient placement and antimicrobial administration. There was significant overlap between patient stays with some patients moving between the two study wards or to other wards in the hospital (Figure 1). Patients were cohorted or placed in single rooms when colonized with VREfm or other alert pathogens. However, not all rooms had ensuite bathroom facilities so risk of VREfm transmission remained. At the time of the study, surveillance systems in the hospital had not detected any suspected VREfm outbreak within the study population.

Results of VREfm screening

In total, 45 rectal swabs from 27 patients were screened for VREfm. Of these, 18 samples from 13 patients were VREfm positive (Table 1). Three (23.1%) colonized patients developed VREfm bacteraemia 9, 24 or 46 days, respectively, after being identified as VREfm carriers. We applied our sampling strategy to 16 rectal screens generating 224 isolates from 11 patients, and five blood cultures (five isolates) from two patients yielding a total of 229 isolates (Table 2). Two rectal swabs and one blood culture were not available for further study. Most patients were female, the median age was 66 years, and a range of primary diagnoses were present (Table 1). Most colonized patients had received antibiotics in the preceding 6 months and 30% had received vancomycin (Table 1).

Simultaneous carriage of multiple VREfm strains

In silico MLST typing using short reads from all 229 genomes showed ST80 (n=130), ST1424 (n=97), ST789 (n=1) and ST1659 (n=1) from the hospital-associated clade A1[59] were present (Table 2). ST80, ST789 and ST1424 are single or double locus variants of each other, while ST1659 differs at five of seven MLST loci suggesting it is more distantly related. Multiple STs were detected in three (27%) samples. Sample VRED06 from patient P49 contained 10 (71.4%) ST80, three (21.4%) ST1424 and one (7.1%) ST789 isolate; sample VRED07 from P14 contained 10 (71.4%) ST1424 and four (28.6%) ST80 isolates; sample

VRED11 from P50 contained 13 (92.9%) ST1424 and one (7.1%) ST1659 isolate. A further rectal swab sample from P49 collected 2 days after VRED06 contained only ST1424, and a blood culture collected 9 days later also contained ST1424. P9 had three rectal swab samples collected over 11 days and had positive blood cultures 1 month later: all samples contained ST80 only. Our finding of multiple strains in 27% of patients is in line with recent studies showing up to half of patients carry 2–4 different E. faecium strains, and within-patient diversity varies over time.[20,24,60,61]

Genomic population structure of VREfm suggests recent transmission events

The chromosomes of the two strain-specific genome assemblies (Table S3) were used as references for short-read mapping within each sequence type. Within-patient diversity was low when genomes of the same sequence type were compared, generally differing by zero SNPs and a maximum pairwise difference of three SNPs (Table 2). Similarly, insertions, deletions and plasmids were usually shared in genomes from the same patient. However, the presence of DEL3 (12 bp non-coding deletion) and DEL4 (11 bp deletion in a solute binding protein) were variable within 24 ST80 genomes from P20 with 0–2 differentiating SNPs (Figure 2). In genomes from P9 p1_VRED06-10 and p3_VRED06-10 were variably detected despite most genomes having no differentiating SNPs (Figure 2). Where multiple samples from the same patient were collected over time, we found low (0–3 SNPs) accumulation of SNPs and no pattern in the prevalence of other genomic variants. Estimates of diversification rates in E. faecium from single-colony sampling of national isolate collections suggest seven mutations per year,[62] other studies of longitudinal within-patient diversification have estimated higher rates of 12.6–128 mutations per year.[20,22,23] The low SNP diversity identified in our 1-month collection of carriage isolates is in keeping with these estimated mutation rates.

The ST80 genomes formed a well-structured population with five clear clusters each separated by >10 SNPs (Figure 2). Clustered genomes differed by 0–2 SNPs and were mostly from individual patients although two clusters included genomes from two different patients (patients P7 and P33, and P2 and P9). In the P7 and P33 genomes, all the reference plasmids were detected as well as four variants (INS1, INS2, DEL1 and MNP1), insertion INS3 was detected in all the clustered genomes except two from P7, and insertion INS4 was detected only in a single P33 genome (Figure 2). In the P2/P9 cluster, all P2 genomes contained all reference plasmids although in P9 there was variable detection of p1_VRED06-10 (28/48 positive) and p3_VRED06-10 (10/48 positive). The only other variant with variable presence was insertion INS3, which was not present in four genomes from P9.

Mapping of the ST1424 genomes showed a much more homogeneous population than in ST80 (Figure 3). Of the 97 ST1424 genomes, 69 had no SNPs and the remaining 28 had 1–2 SNPs differentiating them from the rest of the collection. The SNPs that were detected did not lead to any clear clustering of genomes, except for the 14 genomes from P6 that all carried a SNP in a penicillin-binding protein that differentiated them from the other ST1424 genomes. Two of the P6 genomes had further
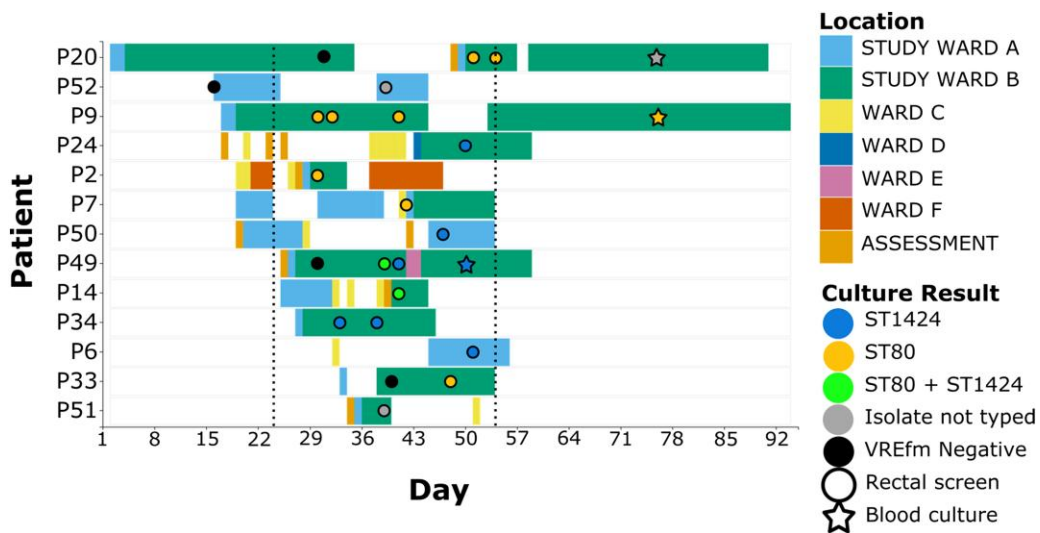
Figure 1. Patient timeline. Each row denotes the location of a patient during admission, blocks denote hospital stay, circles denote VREfm cultures, stars denote bloodstream isolates, dotted lines indicate the start and end of prospective collection of screening isolates for this study. This study was undertaken mainly within Wards A and B, although patients were moved to different wards within the hospital during their stay and were often admitted through the assessment unit. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

Table 1. Characteristics of patients with rectal VREfm colonization (*n* = 13)

Demographics	*N* (%)
Female	8 (61.5)
Age, median (range) years	66 (37–77)
Primary diagnosis	
Acute myeloid leukaemia	3 (23.1)
Diffuse large B cell lymphoma	3 (23.1)
Multiple myeloma	3 (23.1)
Myelodysplasia	2 (15.4)
Chronic lymphocytic leukaemia	1 (7.7)
Mantle cell lymphoma	1 (7.7)
Antimicrobial administration	
Any antibiotics in the 7 d before positive screen	12 (92.3)
Any antibiotics in the 6 m before positive screen	12 (92.3)
Vancomycin in the 7 d before positive screen	1 (7.7)
Vancomycin in the 6 m before positive screen[a]	3 (30.0)
Outcomes within 60 d of VREfm positive screen	
VREfm bloodstream infection	3 (23.1)
ICU admission	1 (7.7)
Death	0 (0)

d, day; ICU, intensive care unit; m, month; *N*, number.
[a]Information available for 10 patients.

independent SNPs (one each) and another genome had lost p1_VRED06-02. No insertions were detected in the ST1424 collection, and of the six deletions found five were only in genomes from P49. p6_VRED06-02 was not detected in 14 P24, 14 P50 and two P49 genomes, while p1_VRED06-02 was not detected in five genomes from three patients.

Analysis of multiple VREfm colonies supports transmission resolution

A transmission network was constructed considering the phylogenetic placement of all 14 colony picks in each sample (Figure 4). The network supports transmission of ST80 between P2 and P9, and between P7 and P33, with P20 not linked to transmission. Epidemiological data support transmission from P7 to P33 on Ward B, as P33 screened negative early in their admission and then screened positive 6 days after P7 (Figures 1 and 4). P9 and P2 screened positive on the same day—no shared rooms or bed spaces were identified as this was P2's first day on Ward B so it is unclear where or when transmission may have occurred (Figures 1 and S1). P20 had two admissions during the study period, was negative at the end of first admission then screened positive on re-admission suggesting they were colonized outside of the hospital.

All ST1424 patients clustered together in the transmission analysis, P34 was strongly linked to all patients and probably had direct transmission to P6 (Figure 4). P34 was the first ST1424 identified on Ward B, P49 was positive 6 days later (having been negative earlier in admission), P14 was positive 2 days after that, and P24 was positive 9 days subsequently (Figure 1). P14 and P49 had ST1424 and ST80 in carriage samples: we did not identify sharing of the ST80 lineages in these patients suggesting there was no direct transmission between these two patients. On Ward A, P50 screened positive with ST1424 and ST1659 on day two of admission and P6 screened positive for ST1424 on day six. The ST1424 populations in P6 and P50 may derive from different hosts with P6 genomes all having a single SNP and P50 genomes having multiple different SNPs and lack the p6_VRED06-02 plasmid (Figures 3 and 4). P34 and P50 shared time on Ward A early in the study before either were known to be VREfm positive, but there is very limited overlap in time between P34 and P6 while both were in different wards (Figure 1).

Table 2. MLSTs detected

Patient ID	Sample ID	Sample type	Sample date (days from start of study)	STs detected (n, %)	Maximum pairwise SNP distance within sample	Median (IQR) pairwise SNP distance within sample
P2	VRED01	Rectal	6	80 (14, 100)	2	0 (0–1)
P6	VRED16	Rectal	27	1424 (14, 100)	2	0 (0–1)
P7	VRED10	Rectal	18	80 (14, 100)	2	0 (0–1)
P9	VRED02	Rectal	6	80 (14, 100)	2	0 (0–0)
P9	VRED03	Rectal	8	80 (14, 100)	0	0 (0–0)
P9	VRED09	Rectal	17	80 (14, 100)	0	0 (0–0)
P9	VRED18	Blood	52	80 (1, 100)	NA	NA
P9	VRED19	Blood	52	80 (1, 100)	NA	NA
P9	VRED20	Blood	52	80 (1, 100)	NA	NA
P9	VRED21	Blood	52	80 (1, 100)	NA	NA
P14	VRED07	Rectal	17	1424 (10, 71.4)	2	1 (0–1)
				80 (4, 26.6)	2	2 (1–2)
P20	VRED15	Rectal	27	80 (14, 100)	3	0 (0–2)
P20	VRED17	Rectal	30	80 (14, 100)	3	1 (0–1)
P24	VRED13	Rectal	26	1424 (14, 100)	0	0 (0–0)
P33	VRED12	Rectal	24	80 (14, 100)	1	0 (0–0)
P34	VRED04	Rectal	9	1424 (14, 100)	0	0 (0–0)
P34	VRED05	Rectal	14	1424 (14, 100)	1	0 (0–0)
P49	VRED06	Rectal	15	80 (10, 71.4)	2	0 (0–1)
				1424 (3, 21.4)	1	0 (0–0)
				789 (1, 7.1)	NA	NA
P49	VRED08	Rectal	17	1424 (14, 100)	2	0 (0–1)
P49	VRED14	Blood	30	1424 (1, 100)	NA	NA
P50	VRED11	Rectal	23	1424 (13, 92.9)	3	1 (0–2)
				1659 (1, 7.1)	NA	NA

ID, identification; IQR, inter quartile range; N, number; NA, not applicable.

None of the patients with ST1424 shared a room or used a bed space previously used by an identified ST1424-positive carrier during their stay (Figure S1).

Analysing fewer than 14 colonies per sample produced fewer transmission links and lower confidence (Table S6, Figure S2). Sequencing more than 14 colonies would improve the detection of minor variants but would increase costs, complexity and turnaround time. Given the low within-patient diversity identified in most patients this approach may not be required in every case. When Gouliouris et al.[60] analysed within-patient diversity of E. faecium in 185 stools collected from 109 patients, analysing a median of five (interquartile range 3–5, total 865) colonies. This identified 51% of stools contained multiple E. faecium subtypes. Based on our analysis, five colonies would identify 50% of the population with 95% confidence (Table S1), so may be a pragmatic choice for analysing within-patient diversity as it should identify most mixed-strain carriage. However, five colonies gave the lowest confidence in transmission analysis so the accuracy of the linkage method should be considered and verified before applying within-patient diversity estimates to routine transmission investigations in future (Table S4, Figure S2). Alternatively, strain-resolved metagenomics directly on clinical samples or sweeps of selective culture growth may be more feasible.[63–65] Further work is required to determine the optimum sampling strategy to support infection prevention and control investigations in healthcare settings.

Plasmids were mostly ST-specific

VRED06-02 (ST1424 reference) contained seven plasmids, and VRED06-10 (ST80 reference) contained five plasmids. Plasmids in the two genomes were generally distinct, suggesting limited sharing between STs within P49 (Table S5).

We sought to identify carriage of similar plasmids in the entire collection by short-read mapping (Table S6). Most plasmids were ST-specific with few examples of ST1424 genomes carrying plasmids from the ST80 reference, and vice versa. However, all ST80 genomes from P7 and P33 carried p7_VRED06-02 from ST1424, and almost all genomes appeared to carry p4_VRED06-10. We believe the hits against the ST1424 genomes are due to cross-mapping of reads from the related p4_VRED06-02 (Table S5). P7_VRED06-02 is unrelated to others in the collection (Table S5), but no close links to any ST1424-positive patients were identified for P7 and P33 (Figure 4). While we did not identify significant sharing of plasmids between ST80 and ST1424, we detected all five ST80 plasmids in ST789 and five of the seven ST1424 plasmids in ST1659 (Table S6). ST789 and ST80 were identified in the same sample from P49, and ST1659 and ST1424 from P50, both cases represent possible plasmid transfer between STs within individual patients during colonization.
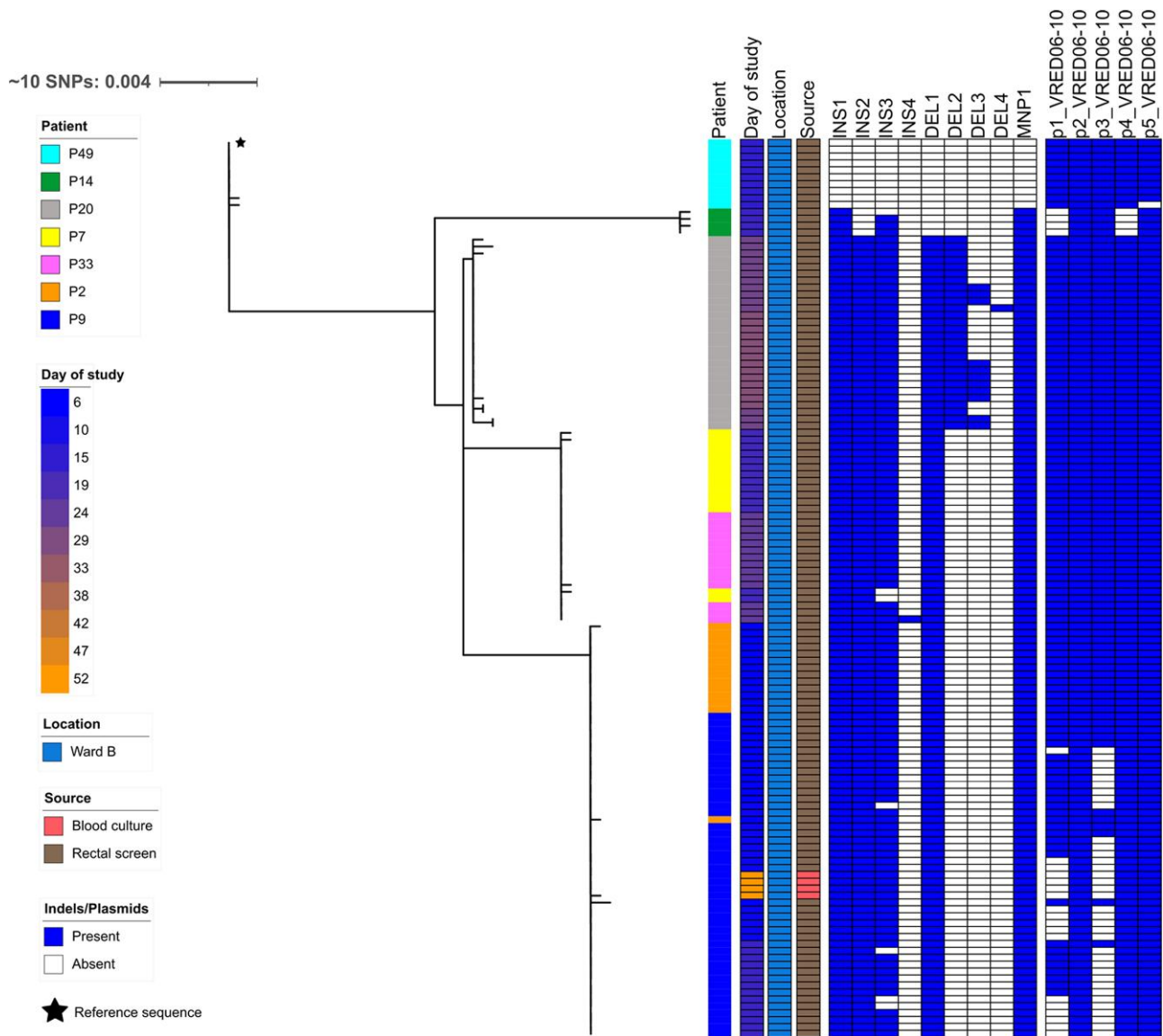
Figure 2. Phylogeny of ST80 isolates showing structured population with three patient specific clusters and two clusters indicating putative patient–patient transmission of VREfm. All ST80 isolates (*n* = 130) mapped to VRED06-10 chromosome and phylogeny built on SNP sites (*n* = 96) after removal of putative transposable and recombinant regions. INS1, 12 bp insertion at position 117 527; INS2, 2 bp insertion at position 1 156 032; INS3, 28 bp insertion at position 1 024 257; INS4, 2 bp insertion at position 447 875; DEL1, 3 bp deletion at position 2 059 113; DEL2, 17 bp deletion at position 1 124 472; DEL3, 12 bp deletion at position 1 534 712; DEL4, 11 bp deletion at position 2 723 479 and MNP1, double nucleotide polymorphism at position 2 647 613. Tree unrooted. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

A large analysis of plasmid sequences by Arredondo-Alonso *et al.*[66] identified that plasmids are shared widely and rapidly within healthcare networks, often unrelated to the genetic lineage of the host cell. This is reflected in our work by the identification of common plasmids in unrelated ST80 isolates, which suggests a wider plasmid transfer network within the lineage.

AMR gene load differs between closely related genomes

We next sought to determine the variability of AMR genes within the collection (Table 3 and Figure 5). In total, 13 AMR genes were detected with three (*aac(6')-Ii*, *msr(C)* and *vanA*) present in all genomes, two (*aph(3')-III* and *erm(B)*) in all but one genome, four (*ant(9)-Ia*, *dfrG*, *erm(A)* and *tet(M)*) only in ST1424 or ST1659 genomes, two genes (*ant(6)-Ia* and *tet(S)*) found only in ST80 and ST789 genomes, and *tet(L)* found in a single ST1659 genome. The aminoglycoside resistance gene *aac(6')-aph(2")* was variably present, found in 69.9% of all genomes.

Tetracycline resistance gene *tet(M)* was identified on the chromosome of VRED06-02 as part of Tn*6944* (Figure S3A). *tet(M)* was identified in 62.2% of ST1424 and ST1659 genomes, excision of Tn*6944* may be responsible for this variable presence. We identified variable within-patient presence of *tet(M)* and no other tetracycline resistance genes (Figure 5),
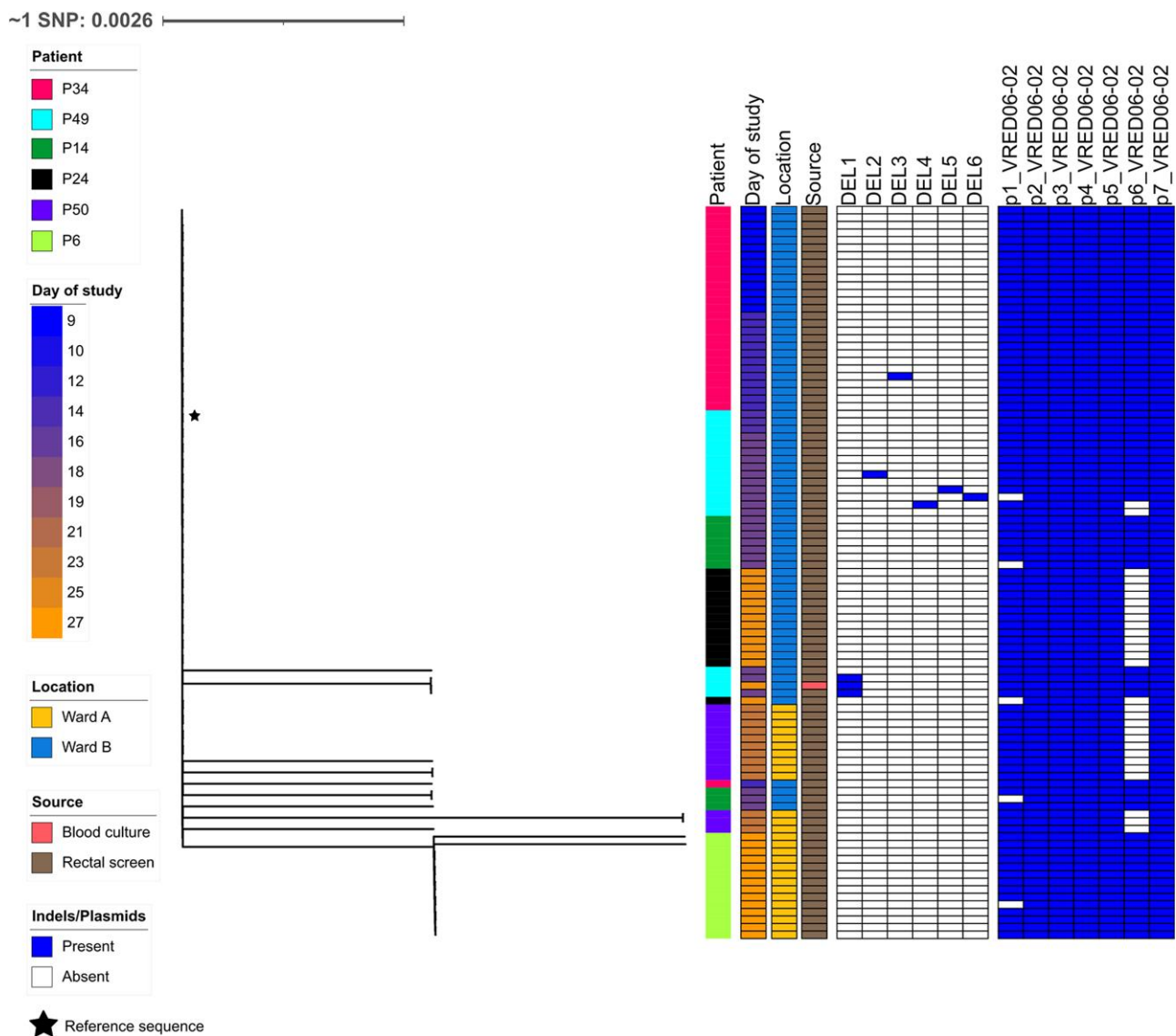
Figure 3. Phylogeny of ST1424 genomes showing homogeneous population suggestive of recent transmission outbreak. ST1424 genomes ($n=97$) mapped to VRED06-02 chromosome and phylogeny built on SNP sites ($n=13$) after removal of putative transposable and recombinant regions. DEL1, 13 bp deletion at position 1 884 038; DEL2, 1 bp deletion at position 850 433; DEL3, 2 bp deletion at position 989 380; DEL4, 11 bp deletion at position 2 184 601; DEL5, 12 bp deletion at position 2 184 601 and DEL6, 12 bp deletion at position 2 184 824. Tree unrooted. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

phenotypic susceptibility pattern would therefore differ based on which colony was picked. However, tetracyclines are not generally used for treatment of enterococcal human infections so the clinical impact may be limited. Similar variable presence of the vancomycin resistance element within patients has been described elsewhere and could lead to inappropriate use of vancomycin when the patient harbours a resistant sub-population.[20,21,67,68] Our study only included vancomycin-resistant isolates, so cannot resolve the potential role of variable vancomycin resistance carriage within patients or in transmission networks. Gain and loss of vancomycin resistance has been described in regional networks over periods of years.[62]

aac(6′)-aph(2″) was present on p1_VRED06-02 (ST1424) and p1_VRED06-10 (ST80). *aac(6′)-aph(2″)* was not detected in any ST80 genomes that were p1_VRED06-10 negative, although only 60.2% ($n=65$) of genomes that carried this plasmid also carried *aac(6′)-aph(2″)*. In p1_VRED06-10, *aac(6′)-aph(2″)* was surrounded by two copies of IS256 similarly to Tn6218, although the transposition machinery was missing (Figure S3B).[69] In ST1424 *aac(6′)-aph(2″)* was detected in 97.8% ($n=90$) genomes with p1_VRED06-02. Another four ST1424 genomes carried *aac(6′)-aph(2″)* but not p1_VRED06-02 (Table 3 and Figure 5). In p1_VRED06-02, two copies of *aac(6′)-aph(2″)* were surrounded by insertion sequences IS256, IS1216 and IS3, providing multiple mechanisms of excision. In the *aac(6′)-aph(2″)* positive,
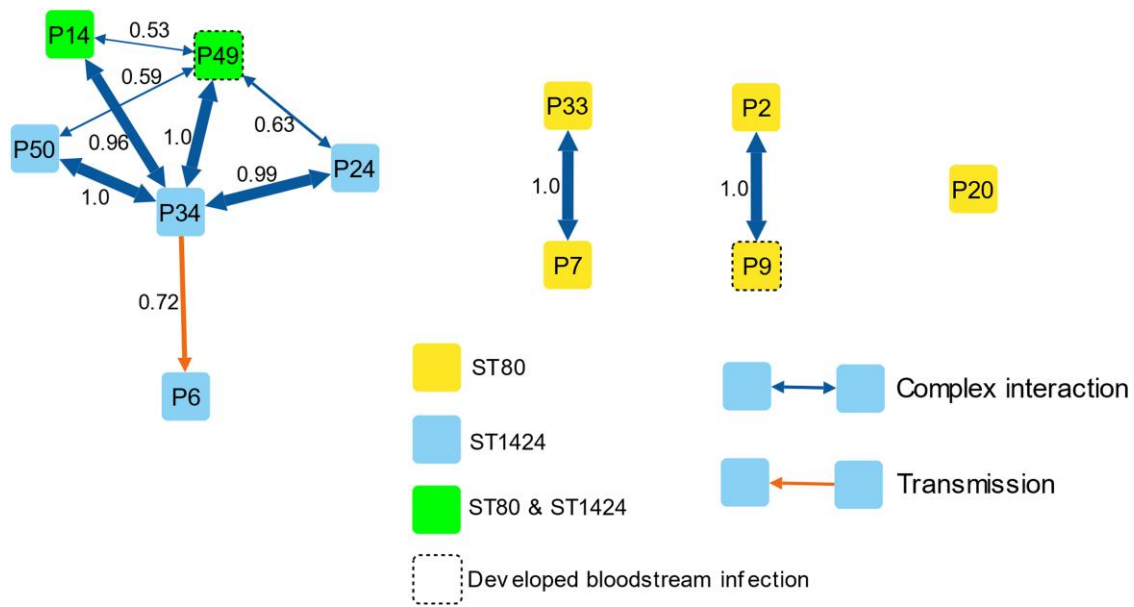
Figure 4. Phyloscanner transmission network. Each patient is represented by a node coloured by detection of the two outbreak STs. Edge thickness corresponds to fraction of Phyloscanner trees with a given relationship, relationship fraction is printed alongside each edge and edge colour is based on the type of relationship (orange, direct transmission; blue, transmission but direction unclear). This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

Table 3. Presence of AMR genes

| Gene	Phenotypic resistance	ST80, n=130		ST1424, n=97		All genomes, n=229	
		n (%)	Genetic element	n (%)	Genetic element	n (%)	Summary
aac(6′)-aph(2″)	Amikacin, Gentamicin, Kanamycin, Streptomycin, Tobramycin	65 (50.0)	p1_VRED06-10	94 (96.9)	p1_VRED06-02	160 (69.9)	Variable in ST80/ST1424/ST789
aac(6′)-Ii	Gentamicin, Tobramycin	130 (100)	Chromosome	97 (100)	Chromosome	229 (100)	All genomes
ant(6)-Ia	Streptomycin	130 (100)	p3_VRED06-10	0 (0)	NA	131 (57.2)	All ST80/ST789
ant(9)-Ia	Spectinomycin	0 (0)	NA	97 (100)	Chromosome	97 (42.4)	All ST1424
aph(3′)-III	Amikacin, Kanamycin, Neomycin	130 (100)	p3_VRED06-10	97 (100)	p2_VRED06-02	228 (99.6)	All ST80/ST789/ST1424
dfrG	Trimethoprim	0 (0)	NA	97 (100)	Chromosome	97 (42.4)	All ST1424
erm(A)	Clindamycin, Erythromycin, Quinupristin	0 (0)	NA	97 (100)	Chromosome	97 (42.4)	All ST1424
erm(B)	Clindamycin, Erythromycin, Quinupristin	129 (99.2)	p1_VRED06-10, p3_VRED06-10	97 (100)	p2_VRED06-02	228 (99.6)	All except one ST80 genome
msr(C)	Erythromycin, Quinupristin	130 (100)	Chromosome	97 (100)	Chromosome	229 (100)	All genomes
tet(L)	Doxycycline, Tetracycline	0 (0)	NA	0 (0)	NA	1 (0.4)	Only ST1659
tet(M)	Doxycycline, Minocycline, Tetracycline	0 (0)	NA	60 (61.9)	Chromosome	61 (26.6)	Variable in ST1424/ST1659
tet(S)	Doxycycline, Minocycline, Tetracycline	130 (100)	p3_VRED06-10	0 (0)	NA	131 (57.2)	All ST80/ST789
vanA	Teicoplanin, Vancomycin	97 (100)	p2_VRED06-10	97 (100)	p2_VRED06-02	229 (100)	All genomes

n, number.
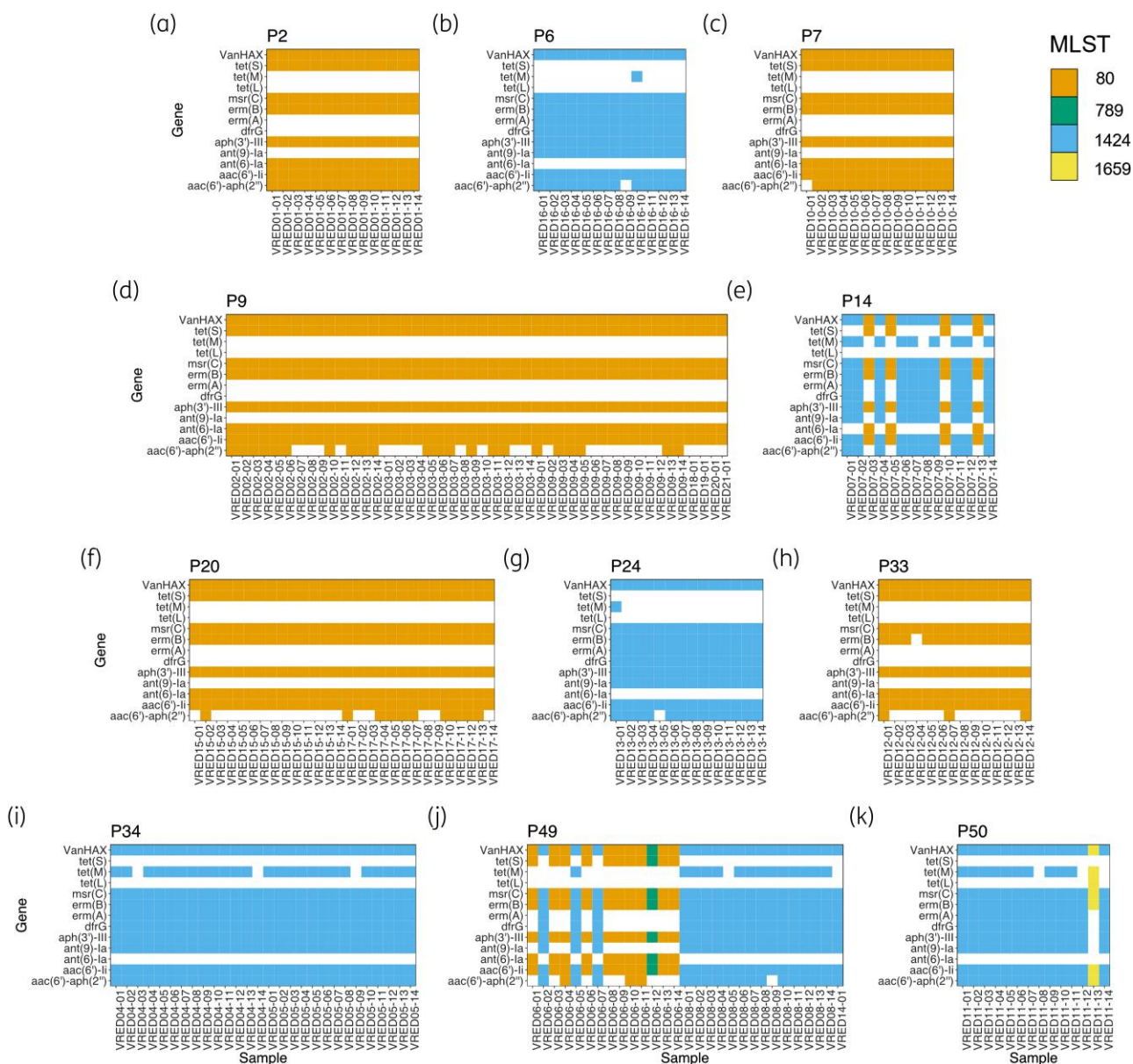
Figure 5. Detection of AMR genes. Resistance genes are plotted on the *y*-axis and isolates on the *x*-axis. Different patients are represented in each panel: (a) P2, (b) P6, (c) P7, (d) P9, (e) P14, (f) P20, (g) P24 (h) P33, (i) P34, (j) P49, and (k) P50. Presence of a gene is represented by a filled square and coloured on the basis of the MLST of the genome. This figure appears in colour in the online version of *JAC* and in black and white in the print version of *JAC*.

p1_VRED06-02 negative cases, short-read assemblies could not resolve the full environment of *aac(6′)-aph(2″)* although in three cases *aac(6′)-aph(2″)* co-located with a single IS*3* gene suggesting mobilization to another transposable element. The impact on phenotype is unclear—all genomes carried *aac(6′)-Ii* and *aph(3′)-III* which together confer high-level resistance to the clinically relevant aminoglycosides amikacin and gentamicin, so the loss of *aac(6′)-aph(2″)* may be more efficient for the cell without an overt change in antibiotic susceptibility. Both Tn*6994* and Tn*6218* were first characterized in *C. difficile*, highlighting transmission of AMR elements between nosocomial pathogens as recently described.[70]

The tetracycline resistance gene *tet(L)* was identified in a single ST1659 genome, the gene was co-located with *tet(M)* on a 30 kb contig that was similar to Tn*6248* from *E. faecium* over ~19 kb (Figure S3C).

We recognize some limitations. Around 60% of *E. faecium* carriers can be linked to nosocomial transmission from other patients or reservoirs in the hospital environment.[60,71–74] Our study did not include environmental samples, and, although patients were mostly located in individual rooms, bathroom facilities were shared posing a significant environmental reservoir for VREfm. Our study focusses on VREfm and so may not be representative of wider populations. This may be important if

resistance is frequently gained by sensitive strains as suggested by some genomic epidemiology studies.[20,62,75,76] However, recent analysis from a single hospital in Denmark found no genomic relatedness between VSEfm and VREfm clones over a 4 year period suggesting vancomycin resistance acquisition by VSEfm is not always frequent and these may represent different transmission networks.[77] Also, we relied on direct plating to solid VREfm screening agar for inclusion in our study. Previous studies have shown a sensitivity of 58%–96% for this approach, rising to 97%–100% with a pre-enrichment step.[78–80] Our use of stored plate sweeps is a potential limitation, as different-sized colonies on the agar plate could alter the prevalence of population variants in the stored mixture and bias the final population prevalence. Plate sweeps were taken daily to avoid excessive growth on the plate, and other groups have shown high agreement in inferred population structures when comparing deep sequencing of plate sweeps with sequencing of multiple colony picks.[81,82]

Longitudinal studies in Denmark, Germany and Spain have described the rise to dominance of ST80 in the 2010s, although more recently ST1424 has been identified and in 2019 was the most common *E. faecium* lineage in Australian BSIs.[50,83–86] It is possible that our study provides a snapshot of the introduction of ST1424 into the haematology ward, on a background of previously established ST80 transmission. Genomic surveillance of enterococci in Scotland is currently limited, so we are unable to use a background dataset to identify whether the lineages identified in our work are common in the study ward and represent ongoing nosocomial transmission clusters or are new introductions into the ward. Wider adoption of a proactive sequence-based surveillance approach should avoid large infection outbreaks, reduce ward closure costs and reduce the clinical impact of invasive disease.[87–90] Recent impact modelling for the UK estimated routine WGS-based surveillance could prevent 74 408 HAIs and 1257 deaths while saving £478 million, or £7.83 per £1 invested.[91] In our setting, an outbreak of VREfm was suspected 3 weeks after the study collection period when P9 and P49 developed bloodstream infection concurrently, as this represented an excess of invasive VREfm in the patient population. On the basis of our analysis, this was many weeks after VREfm transmission had probably occurred (Figures 1 and 4). Due to our study's retrospective nature, we could not use the findings from sequencing to directly influence patient care.

To conclude, by taking account of within-patient diversity in VREfm carriage populations we identified transmission links between patients that could supplement efforts to control transmission within hospitals. However, 73% of cases in our study showed limited within-patient diversity that supports single-colony picks as a suitable method in many settings. We also show that diversity exists not just at the level of SNPs—AMR gene presence/absence, indels and plasmid presence all vary within and between patients. Accounting for within-patient diversity looks to be a useful tool for fully resolving VREfm transmission using WGS-based investigations and could be considered to help resolve outbreaks in the future.

Acknowledgements

Funding

Transparency declarations

The authors declare no competing interests.

Supplementary data

Figures S1 to S3 and Tables S1 to S6 are available as Supplementary data at *JAC* Online.

References

1 Arias CA, Murray BE. The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012; **10**: 266–78. https://doi.org/10.1038/nrmicro2761

2 WHO. WHO Publishes List of Bacteria for Which New Antibiotics are Urgently Needed. 2017. https://www.who.int/news/item/27-02-2017-who-publishes-list-of-bacteria-for-which-new-antibiotics-are-urgently-needed.

3 CDC. *Antibiotic Resistance Threats in the United States, 2019*. U.S. Department of Health and Human Services, CDC, 2019.

4 Ong DSY, Bonten MJM, Safdari K *et al.* Epidemiology, management, and risk-adjusted mortality of ICU-acquired enterococcal bacteremia. *Clin Infect Dis* 2015; **61**: 1413–20. https://doi.org/10.1093/cid/civ560

5 Billington EO, Phang SH, Gregson DB *et al.* Incidence, risk factors, and outcomes for *Enterococcus* spp. blood stream infections: a population-based study. *Int J Infect Dis* 2014; **26**: 76–82. https://doi.org/10.1016/j.ijid.2014.02.012

6 Dik JWH, Dinkelacker AG, Vemer P *et al.* Cost-analysis of seven nosocomial outbreaks in an academic hospital. *PLoS One* 2016; **11**: e0149226. https://doi.org/10.1371/journal.pone.0149226

7 Lloyd-Smith P, Younger J, Lloyd-Smith E *et al.* Economic analysis of vancomycin-resistant enterococci at a Canadian hospital: assessing attributable cost and length of stay. *J Hosp Infect* 2013; **85**: 54–9. https://doi.org/10.1016/j.jhin.2013.06.016

8 ARHAI Scotland. *Scottish One Health Antimicrobial Use and Antimicrobial Resistance in 2020*. ARHAI Scotland, 2021.

9 Zhou X, Willems RJL, Friedrich AW *et al. Enterococcus faecium*: from microbiological insights to practical recommendations for infection control and diagnostics. *Antimicrob Resist Infect Control* 2020; **9**: 130. https://doi.org/10.1186/s13756-020-00770-1

10 Quainoo S, Coolen JPM, van Hijum SAFT *et al.* Whole-genome sequencing of bacterial pathogens: the future of nosocomial outbreak analysis. *Clin Microbiol Rev* 2017; **30**: 1015–63. https://doi.org/10.1128/CMR.00016-17

11 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. https://doi.org/10.1093/jac/dkv142

12 Croucher NJ, Didelot X. The application of genomics to tracing bacterial pathogen transmission. *Curr Opin Microbiol* 2015; **23**: 62–7. https://doi.org/10.1016/j.mib.2014.11.004

13 Bryant JM, Grogono DM, Greaves D *et al*. Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. *Lancet* 2013; **381**: 1551–60. https://doi.org/10.1016/S0140-6736(13)60632-7

14 Worby CJ, Lipsitch M, Hanage WP. Within-host bacterial diversity hinders accurate reconstruction of transmission networks from genomic distance data. *PLoS Comput Biol* 2014; **10**: e1003549. https://doi.org/10.1371/journal.pcbi.1003549

15 Snitkin ES, Zelazny AM, Thomas PJ *et al*. Tracking a hospital outbreak of carbapenem-resistant *Klebsiella pneumoniae* with whole-genome sequencing. *Sci Transl Med* 2012; **4**: 148ra116. https://doi.org/10.1126/scitranslmed.3004129

16 Hall MD, Holden MT, Srisomang P *et al*. Improved characterisation of MRSA transmission using within-host bacterial sequence diversity. *eLife* 2019; **8**: e46402. https://doi.org/10.7554/eLife.46402

17 Harris SR, Cartwright EJ, Török ME *et al*. Whole-genome sequencing for analysis of an outbreak of meticillin-resistant *Staphylococcus aureus*: a descriptive study. *Lancet Infect Dis* 2013; **13**: 130–6. https://doi.org/10.1016/S1473-3099(12)70268-2

18 Okoro CK, Kingsley RA, Quail MA *et al*. High-resolution single nucleotide polymorphism analysis distinguishes recrudescence and reinfection in recurrent invasive nontyphoidal *Salmonella* Typhimurium disease. *Clin Infect Dis* 2012; **54**: 955–63. https://doi.org/10.1093/cid/cir1032

19 Paterson GK, Harrison EM, Murray GGR *et al*. Capturing the cloud of diversity reveals complexity and heterogeneity of MRSA carriage, infection and transmission. *Nat Commun* 2015; **6**: 6560. https://doi.org/10.1038/ncomms7560

20 Moradigaravand D, Gouliouris T, Blane B *et al*. Within-host evolution of *Enterococcus faecium* during longitudinal carriage and transition to bloodstream infection in immunocompromised patients. *Genome Med* 2017; **9**: 119. https://doi.org/10.1186/s13073-017-0507-0

21 Raven KE, Gouliouris T, Parkhill J *et al*. Genome-based analysis of *Enterococcus faecium* bacteremia associated with recurrent and mixed-strain infection. *J Clin Microbiol* 2018; **56**: e01520-17. https://doi.org/10.1128/JCM.01520-17

22 Dubin KA, Mathur D, McKenney PT *et al*. Diversification and evolution of vancomycin-resistant *Enterococcus faecium* during intestinal domination. *Infect Immun* 2019; **87**: e00102-19. https://doi.org/10.1128/IAI.00102-19

23 Bayjanov JR, Baan J, Rogers MRC *et al*. *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. *Microb Genomics* 2019; **5**: e000277. https://doi.org/10.1099/mgen.0.000277

24 Both A, Kruse F, Mirwald N *et al*. Population dynamics in colonizing vancomycin-resistant *E. faecium* isolated from immunosuppressed patients. *J Glob Antimicrob Resist* 2022; **28**: 267–73. https://doi.org/10.1016/j.jgar.2022.01.027

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

26 Inouye M, Dashnow H, Raven L-A *et al*. SRST2: rapid genomic surveillance for public health and hospital microbiology labs. *Genome Med* 2014; **6**: 90. https://doi.org/10.1186/s13073-014-0090-6

27 Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications [version 1; peer review: 2 approved]. *Wellcome Open Res* 2018; **3**: 124. https://doi.org/10.12688/wellcomeopenres.14826.1

28 Croucher NJ, Page AJ, Connor TR *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015; **43**: e15. https://doi.org/10.1093/nar/gku1196

29 Page AJ, Taylor B, Delaney AJ *et al*. SNP-sites: rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb Genomics* 2016; **2**: e000056. https://doi.org/10.1099/mgen.0.000056

30 Minh BQ, Schmidt HA, Chernomor O *et al*. IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol Biol Evol* 2020; **37**: 1530–4. https://doi.org/10.1093/molbev/msaa015

31 Hoang DT, Chernomor O, von Haeseler A *et al*. UFBoot2: improving the ultrafast bootstrap approximation. *Mol Biol Evol* 2018; **35**: 518–22. https://doi.org/10.1093/molbev/msx281

32 Kalyaanamoorthy S, Minh BQ, Wong TKF *et al*. ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat Methods* 2017; **14**: 587–9. https://doi.org/10.1038/nmeth.4285

33 Letunic I, Bork P. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. *Nucleic Acids Res* 2021; **49**: W293–6. https://doi.org/10.1093/nar/gkab301

34 Wick RR, Judd LM, Gorrie CL *et al*. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput Biol* 2017; **13**: e1005595. https://doi.org/10.1371/journal.pcbi.1005595

35 Zankari E, Hasman H, Kaas RS *et al*. Genotyping using whole-genome sequencing is a realistic alternative to surveillance based on phenotypic antimicrobial susceptibility testing. *J Antimicrob Chemother* 2013; **68**: 771–7. https://doi.org/10.1093/jac/dks496

36 De Coster W, D'Hert S, Schultz DT *et al*. NanoPack: visualizing and processing long-read sequencing data. *Bioinformatics* 2018; **34**: 2666–9. https://doi.org/10.1093/bioinformatics/bty149

37 Wick RR, Judd LM, Cerdeira LT *et al*. Trycycler: consensus long-read assemblies for bacterial genomes. *Genome Biol* 2021; **22**: 266. https://doi.org/10.1186/s13059-021-02483-z

38 Kolmogorov M, Yuan J, Lin Y *et al*. Assembly of long, error-prone reads using repeat graphs. *Nat Biotechnol* 2019; **37**: 540–6. https://doi.org/10.1038/s41587-019-0072-8

39 Ruan J, Li H. Fast and accurate long-read assembly with wtdbg2. *Nat Methods* 2020; **17**: 155–8. https://doi.org/10.1038/s41592-019-0669-3

40 Vaser R, Šikić M. Time- and memory-efficient genome assembly with Raven. *Nat Comput Sci* 2021; **1**: 332–6. https://doi.org/10.1038/s43588-021-00073-4

41 Li H. Minimap and miniasm: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics* 2016; **32**: 2103–10. https://doi.org/10.1093/bioinformatics/btw152

42 Walker BJ, Abeel T, Shea T *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014; **9**: e112963. https://doi.org/10.1371/journal.pone.0112963

43 Simão FA, Waterhouse RM, Ioannidis P *et al*. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 2015; **31**: 3210–2. https://doi.org/10.1093/bioinformatics/btv351

44 Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068–9. https://doi.org/10.1093/bioinformatics/btu153

45 Bortolaia V, RS K, Ruppe E *et al*. ResFinder 4.0 for predictions of phenotypes from genotypes. *J Antimicrob Chemother* 2020; **75**: 3491–500. https://doi.org/10.1093/jac/dkaa345

46 Zankari E, Allesøe R, Joensen KG *et al*. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with

chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother* 2017; **72**: 2764–8. https://doi.org/10.1093/jac/dkx217

47 Carattoli A, Zankari E, Garciá-Fernández A *et al.* In silico detection and typing of plasmids using plasmidfinder and plasmid multilocus sequence typing. *Antimicrob Agents Chemother* 2014; **58**: 3895–903. https://doi.org/10.1128/AAC.02412-14

48 Arndt D, Grant JR, Marcu A *et al.* PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* 2016; **44**: W16–21. https://doi.org/10.1093/nar/gkw387

49 Ondov BD, Treangen TJ, Melsted P *et al.* Mash: fast genome and metagenome distance estimation using MinHash. *Genome Biol* 2016; **17**: 132. https://doi.org/10.1186/s13059-016-0997-x

50 Pinholt M, Bayliss SC, Gumpert H *et al.* WGS of 1058 *Enterococcus faecium* from Copenhagen, Denmark, reveals rapid clonal expansion of vancomycin-resistant clone ST80 combined with widespread dissemination of a vanA-containing plasmid and acquisition of a heterogeneous accessory genome. *J Antimicrob Chemother* 2019; **74**: 1776–85. https://doi.org/10.1093/jac/dkz118

51 Ronquist F, Teslenko M, van der Mark P *et al.* Mrbayes 3.2: efficient Bayesian phylogenetic inference and model choice across a large model space. *Syst Biol* 2012; **61**: 539–42. https://doi.org/10.1093/sysbio/sys029

52 Wymant C, Hall M, Ratmann O *et al.* PHYLOSCANNER: inferring transmission from within- and between-host pathogen genetic diversity. *Mol Biol Evol* 2018; **35**: 719–33. https://doi.org/10.1093/molbev/msx304

53 Higgs C, Sherry NL, Seemann T *et al.* Optimising genomic approaches for identifying vancomycin-resistant *Enterococcus faecium* transmission in healthcare settings. *Nat Commun* 2022; **13**: 509. https://doi.org/10.1038/s41467-022-28156-4

54 Shannon P, Markiel A, Ozier O *et al.* Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; **13**: 2498–504. https://doi.org/10.1101/gr.1239303

55 Huebner RE, Dagan R, Porath N *et al.* Lack of utility of serotyping multiple colonies for detection of simultaneous nasopharyngeal carriage of different pneumococcal serotypes. *Pediatr Infect Dis J* 2000; **19**: 1017–20. https://doi.org/10.1097/00006454-200010000-00019

56 R Core Team. R: A Language and Environment for Statistical Computing, 2023. https://www.R-project.org/.

57 Wickham H. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016. https://ggplot2.tidyverse.org.

58 Pedersen TL. patchwork: The Composer of Plots, 2020. https://CRAN.R-project.org/package=patchwork.

59 Lebreton F, van Schaik W, McGuire AM *et al.* Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* 2013; **4**: e00534-13. https://doi.org/10.1128/mBio.00534-13

60 Gouliouris T, Coll F, Ludden C *et al.* Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat Microbiol* 2021; **6**: 103–11. https://doi.org/10.1038/s41564-020-00806-7

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. https://doi.org/10.1186/s13073-015-0259-7

62 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. https://doi.org/10.1093/cid/ciw872

63 Mäklin T, Kallonen T, David S *et al.* High-resolution sweep metagenomics using fast probabilistic inference [version 2; peer review: 2 approved]. *Wellcome Open Res* 2021; **5**: 14. https://doi.org/10.12688/wellcomeopenres.15639.2

64 Anyansi C, Straub TJ, Manson AL *et al.* Computational methods for strain-level microbial detection in colony and metagenome sequencing data. *Front Microbiol* 2020; **11**: 1925. https://doi.org/10.3389/fmicb.2020.01925

65 Mu A, Kwong JC, Isles NS *et al.* Reconstruction of the genomes of drug-resistant pathogens for outbreak investigation through metagenomic sequencing. *mSphere* 2019; **4**: e00529-18. https://doi.org/10.1128/mSphere.00529-18

66 Arredondo-Alonso S, Top J, McNally A *et al.* Plasmids shaped the recent emergence of the major nosocomial pathogen *Enterococcus faecium*. *mBio* 2020; **11**: e03284-19. https://doi.org/10.1128/mBio.03284-19

67 Xanthopoulou K, Wille J, Zweigner J *et al.* Characterization of a vancomycin-resistant *Enterococcus faecium* isolate and a vancomycin-susceptible *E. faecium* isolate from the same blood culture. *J Antimicrob Chemother* 2020; **76**: 883–6. https://doi.org/10.1093/jac/dkaa532

68 Cárdenas AM, Andreacchio KA, Edelstein PH. Prevalence and detection of mixed-population enterococcal bacteremia. *J Clin Microbiol* 2014; **52**: 2604–8. https://doi.org/10.1128/JCM.00802-14

69 Dingle KE, Elliott B, Robinson E *et al.* Evolutionary history of the *Clostridium difficile* pathogenicity locus. *Genome Biol Evol* 2014; **6**: 36–52. https://doi.org/10.1093/gbe/evt204

70 Smith AB, Jenior ML, Keenan O *et al.* Enterococci enhance *Clostridioides difficile* pathogenesis. *Nature* 2022; **611**: 780–6. https://doi.org/10.1038/s41586-022-05438-x

71 Lee AS, White E, Monahan LG *et al.* Defining the role of the environment in the emergence and persistence of vanA Vancomycin-Resistant Enterococcus (VRE) in an intensive care unit: a molecular epidemiological study. *Infect Control Hosp Epidemiol* 2018; **39**: 668–75. https://doi.org/10.1017/ice.2018.29

72 Ford CD, Lopansri BK, Gazdik MA *et al.* Room contamination, patient colonization pressure, and the risk of vancomycin-resistant *Enterococcus* colonization on a unit dedicated to the treatment of hematologic malignancies and hematopoietic stem cell transplantation. *Am J Infect Control* 2016; **44**: 1110–5. https://doi.org/10.1016/j.ajic.2016.03.044

73 McDermott H, Skally M, O'Rourke J *et al.* Near-patient environmental contamination of an intensive care unit with vancomycin-resistant enterococci (VRE) and extended-spectrum beta-lactamase–producing Enterobacteriaceae (ESBL-E) before and after the introduction of chlorhexidine bathing for patients. *Infect Control Hosp Epidemiol* 2018; **39**: 1131–2. https://doi.org/10.1017/ice.2018.146

74 Bonten MJM, Hayden MK, Nathan C *et al.* Epidemiology of colonisation of patients and environment with vancomycin-resistant enterococci. *Lancet* 1996; **348**: 1615–9. https://doi.org/10.1016/S0140-6736(96)02331-8

75 Leigh RJ, McKenna C, McWade R *et al.* Comparative genomics and pangenomics of vancomycin-resistant and susceptible *Enterococcus faecium* from Irish hospitals. *J Med Microbiol* 2022; **71**: 001590. https://doi.org/10.1099/jmm.0.001590

76 van Hal SJ, Ip CLC, Ansari MA *et al.* Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. *Microb Genomics* 2016; **2**: e000048. https://doi.org/10.1099/mgen.0.000048

77 Hansen SGK, Roer L, Karstensen KT *et al.* Vancomycin-sensitive *Enterococcus faecium* bacteraemia—hospital transmission and mortality in a Danish University Hospital. *J Med Microbiol* 2023; **72**: 001731. https://doi.org/10.1099/jmm.0.001731

78 Gouliouris T, Blane B, Brodrick HJ *et al.* Comparison of two chromogenic media for the detection of vancomycin-resistant enterococcal carriage by nursing home residents. *Diagn Microbiol Infect Dis* 2016; **85**: 409–12. https://doi.org/10.1016/j.diagmicrobio.2016.04.026

79 D'Agata EMC, Gautam S, Green WK *et al.* High rate of false-negative results of the rectal swab culture method in detection of gastrointestinal colonization with vancomycin-resistant enterococci. *Clin Infect Dis* 2002; **34**: 167–72. https://doi.org/10.1086/338234

80 Linfield RY, Campeau S, Injean P *et al.* Practical methods for effective vancomycin-resistant enterococci (VRE) surveillance: experience in a liver transplant surgical intensive care unit. *Infect Control Hosp Epidemiol* 2018; **39**: 1178–82. https://doi.org/10.1017/ice.2018.178

81 Raghuram V, Gunoskey JJ, Hofstetter KS *et al.* Comparison of genomic diversity between single and pooled *Staphylococcus aureus* colonies isolated from human colonization cultures. *Microb Genomics* 2023; **9**: 001111. https://doi.org/10.1099/mgen.0.001111

82 Tonkin-Hill G, Ling C, Chaguza C *et al.* Pneumococcal within-host diversity during colonization, transmission and treatment. *Nat Microbiol* 2022; **7**: 1791–804. https://doi.org/10.1038/s41564-022-01238-1

83 Hammerum AM, Baig S, Kamel Y *et al.* Emergence of vanA *Enterococcus faecium* in Denmark, 2005–15. *J Antimicrob Chemother* 2017; **72**: 2184–90. https://doi.org/10.1093/jac/dkx138

84 Tedim AP, Ruíz-Garbajosa P, Rodríguez MC *et al.* Long-term clonal dynamics of *Enterococcus faecium* strains causing bloodstream infections (1995–2015) in Spain. *J Antimicrob Chemother* 2017; **72**: 48–55. https://doi.org/10.1093/jac/dkw366

85 Werner G, Neumann B, Weber RE *et al.* Thirty years of VRE in Germany—"expect the unexpected": the view from the National Reference Centre for Staphylococci and Enterococci. *Drug Resist Updat* 2020; **53**: 100732. https://doi.org/10.1016/j.drup.2020.100732

86 Coombs GW, Daley DA, Mowlaboccus S *et al.* Australian group on Antimicrobial Resistance (AGAR) Australian Enterococcal Sepsis Outcome Programme (AESOP) Annual Report 2019. *Commun Dis Intell* 2020; **44**. https://www1.health.gov.au/internet/main/publishing.nsf/Content/AD2DF748753AFDE1CA2584E2008009BA/$File/australian_group_on_antimicrobial_resistance_agar_australian_enterococcal_sepsis_outcome_programme_aesop_annual_report_2019.pdf. https://doi.org/10.33321/cdi.2020.44.72

87 Gordon LG, Elliott TM, Forde B *et al.* Budget impact analysis of routinely using whole-genomic sequencing of six multidrug-resistant bacterial pathogens in Queensland, Australia. *BMJ Open* 2021; **11**: e041968. https://doi.org/10.1136/bmjopen-2020-041968

88 Torok E, Brodrick H, Khokhar F *et al.* Prospective surveillance and rapid whole-genome sequencing detects two unsuspected outbreaks of carbapenemase-producing *Klebsiella pneumoniae* in a UK teaching hospital. *Open Forum Infect Dis* 2017; **4**: S43–4. https://doi.org/10.1093/ofid/ofx162.104

89 Sherry NL, Gorrie CL, Kwong JC *et al.* Multi-site implementation of whole genome sequencing for hospital infection control: a prospective genomic epidemiological analysis. *Lancet Reg Health West Pac* 2022; **23**: 100446. https://doi.org/10.1016/j.lanwpc.2022.100446

90 Forde BM, Bergh H, Cuddihy T *et al.* Clinical implementation of routine whole-genome sequencing for hospital infection control of multi-drug resistant pathogens. *Clin Infect Dis* 2023; **76**: e1277-84. https://doi.org/10.1093/cid/ciac726

91 Fox JM, Saunders NJ, Jerwood SH. Economic and health impact modelling of a whole genome sequencing-led intervention strategy for bacterial healthcare-associated infections for England and for the USA. *Microb Genomics* 2023; **9**: 001087. https://doi.org/10.1099/mgen.0.001087