

**Co-carriage of diverse vancomycin-resistant *Enterococcus faecium* ST80-lineages by
70% of patients in an Irish hospital**

Nicole L. KAVANAGH¹, Peter M. KINNEVEY¹ [ORCID](#), Grainne I. BRENNAN^{2,3}, Brian
O'CONNELL^{2,3}, Richard V. GOERING⁴, David C. COLEMAN^{1*} [ORCID](#)

¹*Microbiology Research Unit, Dublin Dental University Hospital, University of Dublin,
Trinity College Dublin, Dublin, Ireland*

²*Department of Clinical Microbiology, St. James's Hospital, Dublin, Ireland*

³*National MRSA Reference Laboratory, St. James's Hospital, Dublin, Ireland.*

⁴*Department of Medical Microbiology and Immunology, Creighton University School of
Medicine, Omaha, Nebraska, USA.*

Running title: VREfm within-host diversity in Irish hospital patients

*Corresponding author. Address: Microbiology Research Unit, Division of Oral
Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College,
Lincoln Place, Dublin D02 F859, Ireland. Tel.: +353 1 6127276

E-mail address: david.coleman@dental.tcd.ie (D.C. Coleman).

SYNOPSIS

Background: Vancomycin-resistant *Enterococcus faecium* (VREfm) are significant nosocomial pathogens. Irish VREfm comprise diverse *vanA*-encoding ST80-complex type (CT) lineages. Recent studies indicate that within-patient VREfm diversity could confound surveillance.

Objectives: To investigate the intra-host VREfm genetic diversity among colonised Irish hospital patients.

Methods: Rectal VREfm (n=150) from 10 patients (15 isolates each) were investigated by WGS, core-genome MLST and split *k-mer* (SKA)-SNP analysis. Plasmids and *vanA*-transposons from 39 VREfm representative of CTs identified were resolved by hybrid assembly of short-read (Illumina) and long-read (Oxford Nanopore Technologies) sequences. Plasmid relatedness was assessed based on Mash distances. Thirty vancomycin-susceptible *E. faecium* (VSEfm) from four VREfm-positive patients were also investigated.

Results: All isolates were clade A1 and most were ST80 (VREfm, 147/150; VSEfm, 25/30). Seventy-percent of patients (7/10) harboured either two (n=4), three (n=2) or four (n=1) VREfm CTs. Individual patient isolate pairs from different CTs differed significantly (median SKA-SNPs 2933), but differences were minimal between isolate pairs of the same CT (median SKA-SNPs 0). In total, 193 plasmids were identified in 39 VREfm investigated. Near-identical plasmids ($\geq 99.5\%$ average nucleotide identity) were identified in divergent CTs from multiple patients. Most VREfm (28/39, 72%) harboured *vanA* on closely related transferable, linear plasmids. Divergent CTs within individual patients harboured either indistinguishable *vanA*-transposons or *vanA*-transposons with distinct organisational iterations. Four VSEfm from different CTs investigated harboured similar plasmids to VREfm.

Conclusion: VREfm within-host diversity is highly prevalent in Irish hospital patients, which complicates surveillance. Linear plasmids play an important role in the emergence of Irish VREfm.

Introduction

Vancomycin-resistant *Enterococcus faecium* (VREfm) are major nosocomial pathogens, associated with a wide-range of infections and persistent environmental contamination. The Republic of Ireland reported one of the highest European rates (28.4-45.8%) of invasive VREfm infection between 2006-2022.^{1,2} Irish VREfm are polyclonal with a predominance of *vanA*-encoding ST80 encompassing diverse MLST complex types (CTs) differentiated by significant genomic variation.^{3,4}

High-throughput WGS has revolutionised surveillance and phylogenetic analysis of nosocomial pathogens including VREfm. VREfm gastrointestinal colonisation is an important cofounder for subsequent infection.^{5,6} Current approaches for investigating nosocomial pathogens rely on the assumption of single-strain carriage/monostrain infection and selection of a single isolate for investigation.⁷ However, heterogeneous colonising/infecting populations could significantly complicate surveillance and outbreak investigations.⁸⁻¹⁰

Previous *E. faecium* investigations identified a well-defined hospital population, termed clade A1, distinct from community and animal isolates.¹¹⁻¹³ Clade A1 tend to be enriched with mobile genetic elements (MGEs), insertion sequences and pathogenicity islands.^{12,14} Many studies highlighted the highly recombinogenic nature of clade A1, with the possibility of new clones emerging *in situ* mediated by horizontal gene transfer (HGT) and recombination.^{15,16} The dynamic environment of the gut microbiome and hospital antibiotic selective pressures may promote multi-VREfm strain colonisation in patients with frequent hospital exposure.^{12,17-19} Additionally, patients can disseminate VREfm into the environment, which can persist and act as a reservoir of infection.⁶ This may result in novel VREfm acquisitions for colonised patients and contribute to VREfm gut diversity. Only a few studies investigated the genomic properties and frequency of multi-strain VREfm gastrointestinal carriage in hospitalised patients.²⁰⁻²⁵

This study aimed to investigate the population dynamics, intra-host genetic diversity and plasmidome characteristics of clade A1 VREfm amongst asymptomatic carriers in a large Irish hospital.

Materials and methods

Ethics

The School of Dental Science & Dublin Dental University Hospital Research Ethics Committee approved this study (Reference DSREC2020-01-02).

Patient isolates

VREfm-positive routine admission rectal screening swabs (October 2022-January 2023) from 10 patients (P1-P10) on seven wards (W) (W1 $n=1$, W2 $n=1$, W3 $n=4$, W9 $n=1$, W18 $n=1$, W20 $n=1$, W21 $n=1$) in a large Irish hospital (H1) were inoculated on CHROMID® VRE agar (bioMérieux, Marcy-l'Étoile, France). Following aerobic incubation for 48 h at 37°C, 15 presumptive VREfm colonies per swab ($n=150$) were randomly selected based on previous within-patient heterogeneity studies.²⁵ Isolates were identified as *E. faecium* by MALDI-TOF-MS (Bruker Daltonics, Bremen, Germany) and stored at -80°C in Microbank cryovials (Pro-Lab Diagnostics, Toronto, Canada). Four VREfm-positive swabs (P1, P3, P5, P8) were randomly selected to provide a snapshot of the co-carriage of vancomycin-susceptible *E. faecium* (VSEfm). Swabs were inoculated on *Enterococcus faecium* ChromoSelect agar (Merck, Ireland) and incubated aerobically for 48 h at 37°C. Multiple, separate green colonies, indicative of *E. faecium*, were subcultured on ChromoSelect and CHROMID® VRE agars. Colonies that failed to grow on CHROMID®VRE agar were identified as VSEfm by MALDI-TOF-MS and susceptibility testing.

Contextual collection

Study isolates were compared with well-characterised collections of clade A1 *E. faecium* ($n=1,286$) from seven Irish hospitals including H1 (June 2017-July 2022) and international isolates from 30 countries (1986-2016) (Supplemental (S) Table S1).^{3,4,25,26}

Susceptibility testing

Susceptibility to linezolid, vancomycin and teicoplanin was determined using the VITEK®2 system (bioMérieux) and EUCAST (v14.0) clinical breakpoints.²⁷ Plasmid-encoded antimicrobial resistance genes were detected *in silico* on WGS hybrid assemblies ($n=39$) using ABRicate v.0.9.8 (<https://github.com/tseemann/abricate>).

PFGE and S1 nuclease analysis

PFGE was performed as previously described using a Bio-Rad CHEF DRIII system (Bio-Rad, Richmond, CA, USA) with 0.8% (w/v) agarose gels and switching from 1-17 s for 17 h (120-degree angle, 6 V/cm).²⁸ Plasmid linearity was investigated by S1 nuclease (Promega Corporation, Madison, WI, USA) treatment of agarose plug slices as described previously.²⁹

Plasmid transfer

Plasmid transfer was undertaken by filter mating as described previously using the plasmid-free rifampicin- and fusidic acid-resistant recipient *E. faecium* 64/3.^{3,30}

Whole-genome sequencing

Short-read and long-read WGS and sequence analysis were performed as described previously.^{31,32} Details are provided in Supplemental Methods. VREfm and VSEfm sequence reads are available in GenBank under BioProjects PRJNA734127 and PRJNA1207810.

MLST and core genome MLST (cgMLST)

Traditional MLST STs were extracted *in silico*. Gene-by-gene allele calling and CT assignment was performed using the standardised *E. faecium* cgMLST scheme in Ridom SeqSphere+ v8.5.1 (Ridom GmbH, Münster, Germany) based on 1,423 genes.³³ cgMLST clusters were determined using single-linkage clustering and a previously recommended pairwise allelic difference threshold of ≤ 20 .³³

Split k-mer (SKA) analysis

Target free pairwise whole-genome comparisons were conducted using the SKA software package v1.0 (<https://github.com/simonrharris/SKA>).³⁴ The fastq command with default settings was used to generate *k*-mer files (*k*=15) for each isolate from short-read data. Pairwise distances and single-linkage clustering was performed using SKA distance (identity cutoff of ≥ 0.9) using a previously recommended seven SKA-SNPs threshold.³⁵ Reference free alignment of split *k*-mer files was undertaken using the SKA align command under default settings.

Phylogenetic analysis

Phylogenetic neighbour-joining trees (NJTs) were constructed using gene-by-gene allele calling for 1,423 core genes in the SeqSphere+ *E. faecium* cgMLST scheme.³³ SKA-generated maximum likelihood trees (MLTs) were generated using IQ-TREE v2.2.2.6 (<http://www.iqtree.org>) with the extended ModelFinder with tree inference (-m MFP) and 1000 ultrafast bootstrap replicates.³⁷ MLTs were mid-point rooted and visualised using Interactive Tree of Life v6.8.1(<https://itol.embl.de/>).³⁷

Plasmid analysis

Plasmid sequences from polished hybrid assemblies were visualised using SnapGene v.7.0.4. (<https://www.snapgene.com>) and characterised using the STAMAR pipeline (<https://github.com/phac-nml/staramr>). MOB Subfamilies were assigned using the MOB-Typer tool as part of the MOB-suite (<https://github.com/phac-nml/mob-suite>).³⁸ Pairwise

plasmid genomic distance estimation was calculated using Mash v2.3 (<https://github.com/marbl/Mash>) using default *k*-mer and sketch sizes.^{39,40} Clustering was performed by the complete-linkage method based on the Mash-calculated distance matrix using the `hclust` command in RStudio v2024.04.01 with a threshold of 0.06.³⁹

Results

Isolates

All 150 VREfm from rectal screening swabs from 10 patients (15 VREfm per swab) were sequenced, harboured *vanA* exclusively and were linezolid susceptible (Table S2). Thirty VSEfm sequenced from four VREfm-positive patient swabs lacked *vanA*.

Heterogeneity of within-patient VREfm

Most VREfm (147/150, 98%) were ST80. The remaining three ST80 double-locus variants were submitted to pubMLST.org for ST assignment. Recent studies of within-patient VREfm heterogeneity primarily utilised STs to differentiate between strains.^{20,24,25,41-42} As Irish VREfm are polyclonal, with a predominance of ST80 comprising genetically diverse CTs, ST assignment could not be used to investigate within-patient diversity.^{3,4} Therefore, VREfm belonging to distinct CTs were considered separate strains.

VREfm grouped into nine CT clusters (average size 16, median: 11; range: 2-48) (Figure 1) and two singletons using a previously recommended cgMLST relatedness threshold (≤ 20 allelic differences).³³ A core-genome based MST comprising study ($n=150$), national ($n=1,286$) and global ($n=1,325$) clade A1 *E. faecium* assigned all study VREfm to clade A1 (Figure S1).^{3,4,16,26} Most (7/10) patients harboured diverse populations, comprising two ($n=4$), three ($n=2$) or four ($n=1$) VREfm-CTs. Three patients each yielded one VREfm-CT (Figure 1, Table S2). The frequency of different VREfm-CTs varied in all patients with multi-strain carriage, with high-frequency, moderate-frequency and low-frequency strains identified in each patient (Figure 1, Table S2, Supplemental Results).

Isolate pairs from different CTs within individual patients exhibited significant variation. For example, P8 yielded four CTs with a median inter-CT pairwise cgMLST allelic difference of 428 (range: 169-429) (Figure 1, Table S2) and a median pairwise SKA-SNP difference of 2469 (range: 0-3232) (Figure S2, Tables S2 and S3h). Similar findings were observed for the remaining six patients harbouring multiple VREfm-CTs (Tables S3a and S3c-h). Genetic diversity was minimal among isolate pairs of the same CT from the same patient, with a median pairwise cgMLST allelic and SKA-SNP differences of 0 (range: 0-4) and 0 (range: 0-34), respectively, (Tables S2 and S3b, S3i and S3j). The median pairwise SKA-SNP difference between isolates in each of the three patients with one VREfm-CT was 0 (average: 1, range 0-8) (Tables S3b, S3i and S3j).

Related VREfm in different patients

The cgMLST-based MST revealed three CTs (CT2933, CT6489, CT6738) containing related (<20 allelic differences) VREfm from multiple patients (Figure 1). The largest cluster comprised 48 CT6489-VREfm from five patients (P1 $n=4$, P2 $n=15$, P3 $n=12$, P6 $n=2$, P9 $n=15$) with a median pairwise allelic difference of five (range: 0-11). Whole-genome comparison of CT6489-VREfm revealed greater diversity with a median pairwise SKA-SNP difference of 66 (range: 0-140) (Table S4e). Similar findings were recorded for CT2933 and CT6738 (Tables S4d and S4f). Using a previously recommended threshold of ≤ 7 SKA-SNPs to deem isolates closely related³⁵, ward W3 patients P3 and P6 harboured closely-related CT6489-VREfm (recovered 104-days apart, median pairwise SKA-SNP difference: three, range: 1-6) with no apparent epidemiological links.

Study isolates ($n=150$) were placed in a wider genomic context by comparison with 622 previously described clade A1 clinical and screening *E. faecium* (June 2017-December 2022) from patients and environmental sites from the same seven H1 wards (VREfm $n=579$, VSEfm $n=43$) (Table S1).^{3,4} The circulation and persistence of highly-related low

frequency strains across multiple wards was evidenced by a small number of examples (see Supplemental Results and Tables S5a-i).

Plasmidome of VREfm from patients harbouring multiple strains

Plasmid sequences from 39 VREfm from the 10 patients (P1 $n=6$, P2 $n=2$, P3 $n=3$, P4 $n=4$, P5 $n=5$, P6 $n=4$, P7 $n=4$, P8 $n=7$, P9 $n=2$, P10 $n=2$), representative of CTs, were resolved by hybrid assembly of short- and long-read sequences (Table S6). In total, 193 fully resolved plasmids (size range: 2,947-289,861 bp) were identified with isolates harbouring a median of 5 plasmids (range: 3-8). Ten additional plasmid sequence assemblies comprising multiple contigs, were excluded from further analysis. Plasmids were grouped by associated replication genes and nine distinct replicon types (REP) were identified, with rep11a ($n=28$) and repUS15 ($n=39$) predominating (Table S6). Twelve-percent (23/193) of plasmids were multi-replicon, encoding up to three REPs and 42% of plasmids (82/193) exhibited no known REP. The predicted transferability of plasmids was assessed using Mob and mate-pair formation (MPF) protein typing.³⁸ In total, 56.5% (109/193) of plasmids were deemed non-mobilisable, 30% (58/193) as mobilisable and 13.5% as conjugative (26/193). Most plasmids (81/84) characterised as transferable contained MOB relaxase gene family types (Table S6).

Mash distance-based plasmid clustering

Plasmid ($n=193$) relatedness was assessed using a reference-free approach based on Mash distances using a previously recommended Mash distance threshold of 0.06.³⁹ Plasmids grouped into 16 clusters (C1-C16) and two singletons (average cluster size 12, median: eight, range: 2-39) (Figure 2, Tables S7a-o). The phylogenetic clustering of plasmid sequences was mainly independent from the core genome phylogeny. Plasmid clusters were predominantly restricted to single replicon types, with 3/16 clusters exhibiting combinations of multiple plasmid family types (Tables S7a-o).

A recently described stringent Mash distance threshold of $\leq 0.0001^{43}$, equivalent to $\geq 99.5\%$ average nucleotide identity, was adopted after investigating several thresholds for differentiating plasmids into nested secondary clusters. Near-identical plasmids were observed in genetically divergent VREfm CTs indicative of widespread dissemination (Tables S7a-o). A more extensive evaluation of the plasmidome from the 39 VREfm investigated is provided in Supplementary Results.

Linear plasmids as vehicles of vanA dissemination

Among the 39 VREfm investigated for plasmids, 72% (28/39) harboured *vanA* on linear topology plasmids, most (25/28) of which were deemed non-transferable by conventional MOB typing and lacked known mobilisation or conjugation genes (Table S6). However, two linear plasmids investigated from this study (see below) were transferable by filter mating.³ Linear topology plasmids resided within genomic cluster C4 in the Mash-based dendrogram, with a median total pairwise Mash distance of 0.00797 (range: 0-0.02980) (Figure 2, Table S7d). Of the remaining 11 VREfm, three P8 ST80/CT7567-VREfm harboured *vanA* on 40,922 bp non-transferable circular plasmids, while eight harboured *vanA* on plasmids not fully closed by hybrid assembly. PFGE was performed with S1 nuclease treatment on four representative VREfm (JHM1-15 (ST80/CT2933, P1), JHM2-15 (ST80/CT6489, P2), JHM5-6 (ST80/CT1916, P5), JHM8-1 (ST80/CT2932, P8) harbouring *vanA*-encoding linear plasmids. The electrophoretic mobility of these plasmids during PFGE was comparable with or without S1 nuclease treatment, which nicks supercoiled circular plasmid DNA, thus confirming plasmid linearity (Figure S4). Transfer of *vanA*-encoding linear plasmids from two donor isolates JHM5-6 (ST80/CT1916, P5) and JHM5-13 (ST80/CT3395, P5) to the plasmid-free *E. faecium* 64/3 (ST21) recipient yielded *vanA*-positive transconjugants phenotypically resistant to vancomycin and teicoplanin, which harboured linear plasmids with near-complete homology to the donor plasmids (average nucleotide identity $>99.9\%$, Figure S5).

Most linear plasmids (82%, 23/28) encoded one resistance determinant (*vanA*).

Core gene analysis of the 28 linear plasmids (i.e., genes present in $\geq 95\%$) using Roary v13.3.046⁴⁴, with a 95% BLASTP sequence identity threshold identified 87 coding sequences, with 77/87 encoding hypothetical proteins. Seven of the genes with a known function were *vanA* operon components. The 28 linear plasmids from multiple VREfm lineages had near-identical backbones (Mash threshold ≤ 0.0001) (Figure 2, Table S7d).

VREfm diversity and vanA transposon organisation

The genetic organisation of the Tn1546-like *vanA* regions in the 39 VREfm investigated for plasmids were resolved by hybrid assembly (Table S6, Figures 3 and S6-S14). Structural analysis of the *vanA* region from multiple VREfm from individual patients (e.g., P1, P4, P7) revealed identical *vanA* regions ($\geq 99.9\%$ average nucleotide identity) in divergent CTs (Figures S6, S9 and S12). There were also many examples of distinct iterations of the *vanA* transposon in different CTs in individual patients (e.g., P5, P6, P8; Figures 3 and S10-S11). A more detailed description of clonal diversity and *vanA*-transposon structural organisation is provided in Supplementary Results.

VSEfm carriage

Thirty VSEfm from four VREfm-positive patients (P1, P3, P5, P8) were clade A1 with 25/30 identified as ST80 and 5/30 as ST203 (Table S8). Phylogenetic analysis using 1,423 cgMLST loci identified five CTs including CT2932 ($n=7$), CT2950 ($n=5$), CT2940 ($n=12$), CT3395 ($n=1$) and CT2933 ($n=5$) (Table S8). A cgMLST based MST of the 150 VREfm and 30 VSEfm using the ≤ 20 allelic differences relatedness threshold revealed 11 clusters and two singletons (average cluster size 16, median: 12, range: 3-48) (Figure S17). CT2933 and CT3395 comprised mixtures of VSEfm and VREfm from individual patients (Figure S17, Table S8). The largest cluster, CT2933, comprised 48 isolates (five VSEfm and 43 VREfm) from five patients (P1 $n=10$, P4 $n=5$, P6 $n=13$, P7 $n=14$, P8 $n=5$) with a

median cgMLST allelic difference of eight (range: 0-11). Patient P8 CT2933-VSEfm and CT2933-VREfm and Patient P5 CT3395-VSEfm and CT3395-VREfm both exhibited a median allelic difference of one, indicating close relatedness (Figure S17). SKA-SNP analysis supported these findings (median pairwise SKA-SNP difference of one, range: 1-2). Hybrid assembly of a representative Patient P8 CT3395-VSEfm (JHMS8-4) identified five plasmids (size range 2,947-167,561 bp). Comparison of the plasmidome of JHMS8-4 with a genetically indistinguishable P8 CT3395-VREfm (JHM8-4) also identified five plasmids, including four near-identical plasmids (average nucleotide identity $\geq 99.99\%$) (Figures S17-S18). VREfm isolate JHM8-4 harboured an additional *vanA*-encoding linear plasmid (110,419 bp) that was absent in JHMS8-4 (VSEfm), which harboured a 40,337 bp circular plasmid instead. No *vanA* genes were identified in the VSEfm isolate. Screening of the short-read derived FASTA contigs from the remaining four P8 CT2933-VSEfm using ABRicate revealed similar findings. See Supplementary Results for a detailed description of VSEfm carriage.

Discussion

This study provides insights into within-host genetic diversity of clade A1 VREfm in Irish hospital patients. Seventy-percent of patients carried diverse VREfm populations (≥ 2 CTs), significantly higher than previously reported (27-51% of patients with 2-4 variants).^{20,23,25,41,43} Discrepancies between reported rates likely reflect differences in the number of VREfm colonies investigated per patient (5-14).^{20,23,25,41,42}

Significant genomic variation was observed between individual patient isolate pairs from different CTs. For example, median pairwise cgMLST allelic and SKA-SNP differences of 427 and 2933, respectively, were recorded between different VREfm CTs within the same patient (Tables S2 and S3), possibly reflecting independent strain acquisitions. Longitudinal studies on genetic changes between VREfm collected sequentially from patients support this conclusion, with within-patient mutation rates of 6-

122 SNPs/genome/year reported.^{11,20,22-23} Intra-CT variation within individual patients was also observed. For example, 13 patient P6 CT2933-VREfm exhibited a pairwise SKA-SNP difference range of 0-10, highlighting micro-evolution (Table S4d).). The occurrence of CT2933 and CT6489 in 5/10 patients in each case likely reflects the institutional dominance of specific CTs widely reported in clade A1 *E. faecium*, where hospital-adapted subpopulations emerge, evolve, and subsequently predominate.^{11,19,35} We recently reported the intra- and inter-hospital dominance of three major ST80 CTs including CT2933 in Irish hospitals.⁴

Most plasmid types identified (179/193, 93%) in the 39 VREfm investigated were distributed across diverse CTs (Figure 2). This indicates a reservoir of similar plasmid backbones circulating within Irish ST80-VREfm irrespective of lineage, consistent with previous studies.^{4,25,45}

The high prevalence of within-patient VREfm heterogeneity likely drives exchange of near-identical plasmids between diverse lineages within the gut with onward dissemination to other patients. VREfm harboured a median of five plasmids (range: 3-8), in agreement with previous studies, and frequently carried antimicrobial-resistance genes (39%, 73/193, (Table S6)).⁴⁵ Interestingly, most plasmid-encoded resistance genes were co-localised on a single plasmid type. A recent investigation of within-host diversity of *Pseudomonas aeruginosa* suggested that mixed strain populations may significantly accelerate the evolution of within-host antibiotic resistance.⁴⁶

The significantly higher prevalence of transferable linear plasmids in VREfm than previously reported (72%, 28/39 isolates)⁴⁷⁻⁴⁹ suggests *vanA* acquisition is an important driver of linear plasmid dissemination/evolution in Irish VREfm. These results support the high-prevalence of transferable *vanA*-encoding linear plasmids identified in VREfm from several Irish hospitals previously reported.³ Linear plasmids were enriched with IS1216E elements, potentially contributing to translocation of the *vanA* transposon to these

plasmids.³ IS1216E-mediated instability was previously shown to be a contributory factor in the restructuring and mobility of Tn1546-like *vanA* in Irish VREfm and in vancomycin-resistant enterococci from other countries.^{3,50,51} The presence of distinct *vanA*-transposon iterations on both circular and linear plasmids in different VREfm CTs suggests multiple, separate *vanA* acquisition events by VSEfm (Figure 3, Figures S10 and S11). These findings indicate extensive recombination within VREfm *vanA* transposon regions, likely mediated by IS1216E, as reported previously.⁵²

Most (25/28) linear plasmids were deemed non-transferable by MOB-suite tools due to the absence of known relaxase and mate-pair formation type genes. However, we demonstrated the conjugative transfer of VREfm linear plasmids under selective pressure here and previously.³ Current databases do not list all mobilisation genes and the high proportion of ‘hypothetical proteins’ encoded by linear plasmids may include undescribed genes involved in plasmid transfer.

The implications of VREfm multi-strain carriage for surveillance are significant as conventional surveillance approaches rely on single strain sequencing strategies. While this study lacks epidemiological context due to the unavailability of patient movement data, examples of potential acquisition/transmission of low frequency strains were evident. For example, P3 and P6 harboured closely related CT6489-VREfm (≤ 6 SKA-SNPs) recovered within 101 days on ward W3 (Table S4e). In P6, CT6489-VREfm represented a minor frequency strain that could be overlooked using conventional surveillance (Table S2). Evaluation of study isolates with a comparator dataset of clade A *E. faecium* from H1 recovered over five years^{3,4} revealed that divergent lineages were circulating within H1 study wards and some low frequency strains identified here persisted for protracted periods (>40 months) within and between H1 wards (Tables S1, S2 and S5a-i).

Routine sequencing of multiple VREfm per sample for surveillance would entail significant costs, more complex bioinformatics and increased turnaround-time. Identifying

an optimal number of colonies for sequencing is crucial to accurately reflect within-patient diversity, while remaining justifiable financially. For example, McHugh and colleagues demonstrated that analysing <14 VREfm colonies per sample identified fewer transmission links and lower phylogenetic network clustering scores.²⁵ However, the authors concluded that single-colony picks are suitable in many settings, such as in Scotland, where 73% of cases showing limited within-patient diversity.²⁵ However, because of the extensive within-patient VREfm diversity identified in Ireland (70%), investigating multiple colonies would significantly improve the resolving power of WGS in surveillance.

Multiple colony sequencing may be challenging for hospital surveillance. A potential alternative could be a ‘one plus pool’ approach, which involves sequencing a single colony per specimen to represent the ‘dominant’ variant in conjunction with pooling and sequencing all remaining colonies to identify within-patient diversity.⁵³ Alternatively, a strain-specific ‘plate sweep’ metagenomics approach could be used, where a mixture of strains are harvested from enrichment cultures, sequenced and bioinformatic pipelines are employed to detect individual strains.^{54–56}

All 30 VSEfm investigated from four VREfm-positive patients were clade A1. Highly similar plasmidomes were evident within VREfm and VSEfm, which is unsurprising since VREfm and VSEfm were predominantly ST80. However, there were instances of genetically indistinguishable VSEfm and VREfm from the same patient with plasmidome differences. For example, P8 yielded six CT2933 isolates (five VSEfm and one VREfm), which exhibited a median SKA-SNP difference of one. Comparison of the accessory genomes of a representative CT2933-VSEfm and CT2933-VREfm revealed near-identical plasmidomes (four plasmids in both isolates sharing $\geq 99.99\%$ nucleotide identity) (Figure S18). A notable difference was the absence of a linear plasmid in the CT2933-VSEfm. It is likely that new VREfm arise in hospital patients colonised with VREfm and VSEfm by *in vivo* plasmid transfer. Patients would only have to acquire a

VREfm strain once and then subsequently exhibit within-patient diversity. However, it is likely that separate VREfm acquisition events and plasmid acquisition by VSEfm contribute to within-patient diversity.

Conclusions

VREfm within Irish hospital patients exhibit significant intra-host diversity with 70% of patients harbouring multiple strains. The presence of near-identical plasmids in Irish VREfm/VSEfm facilitates HGT across diverse lineages, with linear plasmids constituting a significant vehicle for *vanA* dissemination.

Acknowledgements

The authors acknowledge the support of the staff of the Irish National MRSA Reference Laboratory (NMRSARL) and the staff of St. James's Hospital.

Funding source

This work was supported by the Irish Health Research Board (grant number ILP-POR-2019-010 to D.C.C.).

Transparency declaration

None of the authors have any conflicts of interest to declare.

References

1. ECDC. *Antimicrobial resistance in the EU/EAA (EARS-NET)*. Annual Epidemiological Report for 2019. 2020. <https://www.ecdc.europa.eu/en/publications-data/surveillance-antimicrobial-resistance-europe-2019>
2. ECDC. *Antimicrobial resistance in the EU/EAA (EARS-NET)*. Annual Epidemiological Report for 2022. 2023. <https://www.ecdc.europa.eu/en/publications-data/surveillance-antimicrobial-resistance-europe-2022>
3. Egan SA, Kavanagh NL, Shore AC *et al*. Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar *vanA* regions via *IS1216E* and plasmid transfer in diverse genetic lineages in Ireland. *J Antimicrob Chemother* 2021; **77**: 320-30.
4. Kavanagh NL, Kinnevey PM, Egan SA *et al*. Protracted transmission and persistence of ST80 vancomycin-resistant *Enterococcus faecium* clonal complex types CT2933, CT2932 and CT1916 in a large Irish hospital: a 39-month whole-genome sequencing study. *J Hosp Infect* 2025; **151**: 11-20.
5. Arias CA, Murray BE. The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat Rev Microbiol* 2012; **10**: 266–78.
6. Cattoir V, Leclercq R. Twenty-five years of shared life with vancomycin-resistant enterococci: is it time to divorce? *J Antimicrob Chemother* 2013; **68**: 731-42.
7. Croucher NJ, Didelot X. The application of genomics to tracing bacterial pathogen transmission. *Curr Opin Microbiol* 2015; **23**: 62-7.
8. 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.
9. 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.
10. Hall MD, Holden MT, Srisomang P *et al*. Improved characterisation of MRSA transmission using within-host bacterial sequence diversity. *Elife* 2019; **8**: e46402.
11. 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.
12. Lebreton F, van Schaik W, Manson McGuire A *et al*. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* 2013; **4**: e00534-13.
13. Mundt JO. Occurrence of enterococci in animals in a wild environment. *Appl Microbiol* 1963; **11**: 136-40.
14. Gouliouris T, Raven KE, Ludden C *et al*. Genomic surveillance of *Enterococcus faecium* reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. *mBio* 2018; **9**: e01780-18.

15. 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.
16. 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.
17. Baquero F. From pieces to patterns: evolutionary engineering in bacterial pathogens. *Nat Rev Microbiol* 2004; **2**: 510–18.
18. van Schaik W, Top J, Riley DR *et al.* Pyrosequencing-based comparative genome analysis of the nosocomial pathogen *Enterococcus faecium* and identification of a large transferable pathogenicity island. *BMC Genomics* 2010; **11**: 239.
19. Howden BP, Holt KE, Lam MMC *et al.* Genomic insights to control the emergence of vancomycin-resistant enterococci. *mBio* 2013; **4**: e00412–13.
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**: 199.
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.
22. Bayjanov JR, Baan J, Rogers MRC *et al.* *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. *Microb Genom* 2019; **5**: e000277.
23. 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.
24. Both A, Kruse F, Mirwald N *et al.* Population dynamics in colonizing vancomycin-resistant *Enterococcus faecium* isolated from immunosuppressed patients. *J Glob Antimicrob Resist* 2022; **28**: 267–73.
25. McHugh MP, Pettigrew KA, Taori S *et al.* Consideration of within-patient diversity highlights transmission pathways and antimicrobial resistance gene variability in vancomycin-resistant *Enterococcus faecium*. *J Antimicrob Chemother* 2024; **79**: 656–68.
26. van Hal SJ, Willems RJL, Gouliouris T *et al.* The global dissemination of hospital clones of *Enterococcus faecium*. *Genome Med* 2021; **13**: 52.
27. EUCAST. *Breakpoint tables for interpretation of MICs and zone diameters*. Version 13.0. 2023.
http://www.eucast.org/fileadmin/src/media/PDFs/EUCAST_files/Breakpoint_tables/v_5.0_Breakpoint_Table_01.pdf.
28. Tickler IA, Tenover FC, Dewell S *et al.* *Streptococcus agalactiae* strains with chromosomal deletions evade detection with molecular methods. *J Clin Microbiol* 2019; **57**: e02040–18.

29. Freitas AR, Novais C, Tedim AP *et al.* Microevolutionary events involving narrow host plasmids influences local fixation of vancomycin-resistance in *Enterococcus* populations. *PLoS ONE* 2013; **8**: e50689.
30. Bender JK, Fleige C, Lange D *et al.* Rapid emergence of highly variable and transferable oxazolidinone and phenicol resistance gene *optrA* in German *Enterococcus spp.* clinical isolates. *Int J Antimicrob Agents* 2018; **52**: 819–27.
31. Egan SA, Shore AC, O'Connell B *et al.* Linezolid resistance in *Enterococcus faecium* and *Enterococcus faecalis* from hospitalized patients in Ireland: high prevalence of the MDR genes *optrA* and *poxA* in isolates with diverse genetic backgrounds. *J Antimicrob Chemother* 2020; **75**: 1704–11.
32. Egan SA, Corcoran S, McDermott H *et al.* Hospital outbreak of linezolid-resistant and vancomycin-resistant ST80 *Enterococcus faecium* harbouring an *optrA*-encoding conjugative plasmid investigated by whole-genome sequencing. *J Hosp Infect* 2020; **105**: 726–35
33. de Been M, Pinholt M, Top J *et al.* Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. *J Clin Microbiol* 2015; **53**: 3788–97.
34. Maechler F, Weber A, Schwengers O *et al.* Split k-mer analysis compared to cgMLST and SNP-based core genome analysis for detecting transmission of vancomycin resistant enterococci: results from routine outbreak analyses across different hospitals and hospitals networks in Berlin, Germany. *Microb Genom* 2023; **9**: mgen000937.
35. Higgs C, Sherry NL, Seemann T *et al.* Optimising genomic approaches for identifying vancomycin-resistant *Enterococcus faecium* transmission in healthcare settings. *Nat Commun* 2022; **26**: 509.
36. Nguyen LT, Schmidt HA, Von Haeseler A *et al.* IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015 **32**: 268-74.
37. 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.
38. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom* 2018; **4**: e000206.
39. Robertson J, Bessonov K, Schonfeld J *et al.* Universal whole-sequence-based plasmid typing and its utility to prediction of host range and epidemiological surveillance. *Microb Genom* 2020; **6**: mgen000435.
40. Ondov BD, Starrett GJ, Sappington A *et al.* Mash Screen: high-throughput sequence containment estimation for genome discovery. *Genome Biol* 2019; **20**: 232
41. 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.
42. Gouliouris T, Coll F, Ludden C *et al.* Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat Microbiol* 2021; **5**: 103-11.

43. Matlock W, Lipworth S, Chau KK *et al.* Enterobacterales plasmid sharing amongst human bloodstream infections, livestock, wastewater, and waterway niches in Oxfordshire, UK. *eLife* 2023; **12**: e85302
44. Page AJ, Cummins CA, Hunt M *et al.* Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015; **31**: 3691–3.
45. 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.
46. Diaz Caballero J, Wheatley RM, Kapel N *et al.* Mixed strain pathogen populations accelerate the evolution of antibiotic resistance in patients. *Nat Commun* 2023; **14**: 4083.
47. Hashimoto Y, Kita I, Suzuki M *et al.* First report of the local spread of vancomycin-resistant enterococci ascribed to the interspecies transmission of a *vanA* gene cluster-carrying linear plasmid. *mSphere* 2020; **5**: e00102-20
48. Hashimoto Y, Taniguchi M, Uesaka K *et al.* Novel multidrug-resistant enterococcal mobile linear plasmid pELF1 encoding *vanA* and *vanM* gene clusters from a Japanese vancomycin-resistant enterococci isolate. *Front Microbiol* 2019; **10**: 2568.
49. Boumasmoud M, Haunreiter VD, Schweizer TA *et al.* Genomic surveillance of vancomycin-resistant *Enterococcus faecium* reveals spread of a linear plasmid conferring a nutrient utilization advantage. *mBio* 2022; **13**: e0377121.
50. Willems RJ, Top J, van den Braak N *et al.* Molecular diversity and evolutionary relationships of Tn1546-like elements in enterococci from humans and animals. *Antimicrob Agents Chemother* 1999; **43**: 489-91.
51. Stobberingh E, van den Bogaard A, London N *et al.* Enterococci with glycopeptide resistance in Turkeys, Turkey farmers, Turkey Slaughterers, and (sub)urban residents in the south of the Netherlands: evidence for transmission of vancomycin resistance from animals to humans? *Antimicrob Agents Chemother* 1999; **43**: 2215-21.
52. Shan X, Li X-S, Wang N *et al.* Studies on the role of IS1216E in the formation and dissemination of *poxtA*-carrying plasmids in an *Enterococcus faecium* clade A1 isolate. *J Antimicrob Chemother* 2020; **75**: 3126–30
53. 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 Genom* 2023; **9**: 001111
54. 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.
55. 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.
56. Mäklin T, Kallonen T, David S *et al.* High-resolution sweep metagenomics using fast probabilistic inference. *Wellcome Open Res* 2020; **5**: 14.

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

Figure legends

Figure 1. Minimum spanning tree (MST) based on core genome multi-locus sequence typing (cgMLST) analysis of 1,423 loci for 150 clade A1 VREfm isolates recovered from a single rectal screening swab from each of 10 separate patients. Node colours represent the patient number (P1-P10) from which the isolate was recovered as indicated in the legend. Partitions within the nodes represent the presence of ≥ 2 isolates. Branch numbers indicate cgMLST allelic differences between neighbouring isolates. cgMLST segregated isolates into 11 clusters and two singletons using the previously recommended threshold of ≤ 20 allelic differences.³³

Figure 2. Pairwise Mash distance-based clustering of plasmids ($N=193$) from 39 representative VREfm investigated from 10 patient rectal screening samples. Plasmids segregated into 16 distinct clusters (C1-C16) and 2 singletons with an average cluster size of 12 (median: 8, range: 2-39) using a Mash distance of 0.06.³⁹ Metadata pertaining to the plasmid cluster, complex type (CT) of the isolate harbouring the plasmid and replicon type are indicated in the legends. The median plasmid size of each cluster is indicated on the branches. The two singleton plasmid labels are coloured red.

Figure 3. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of seven VREfm isolates from patient P8 belonging to three complex types (CTs). The figure shows an alignment of the *vanA* region encoded by seven plasmids from separate VREfm isolates as follows: (a) isolate JHM8-1 (ST80/CT2932), *vanA* region 18,974 bp encoded on pELF_JHM8-1 (linear plasmid 146,603 bp), (b) isolate JHM8-3 (ST80/CT2932), *vanA* region 18,974 bp encoded on pELF_JHM8-3 (linear plasmid 146,604 bp), (c) isolate JHM8-2 (ST80/CT7567), *vanA* region 18,974 bp encoded on pELF_JHM8-12 (linear plasmid 194,497 bp), (d) isolate JHM8-4 (ST80/CT2933), *vanA* region 13,059 bp encoded on pELF_JHM8-4 (linear

667 plasmid 167,557 bp), (e) isolate JHM8-2_b (ST80/CT7567), *vanA* region 11,317 bp encoded on
668 pJHM8-2_b (circular plasmid 40,922 bp), (f) isolate JHM8-10 (ST80/CT7567), *vanA* region 11,317
669 bp encoded on pJHM8-10_b (circular plasmid, 40,922 bp) and (g) isolate JHM8-15
670 (ST80/CT7567), *van* region, 11,317 bp encoded on pJHM8-15_b (circular plasmid, 40, 922 bp).
671 Images were generated using Easyfig.⁵⁷ Coding sequences of interest and their orientation are
672 represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences;
673 blue, proteins with a known coding function; cyan, transposase and invertase genes; grey,
674 hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in
675 Supplemental Tables A3h, A6 and A7. Regions between alignments with shared homology are
676 represented by grey connecting blocks and/or lines with the colour gradient indicating percentage
677 nucleotide identity (82-100%). A scale bar is included beneath the figure.

678

Figure 1

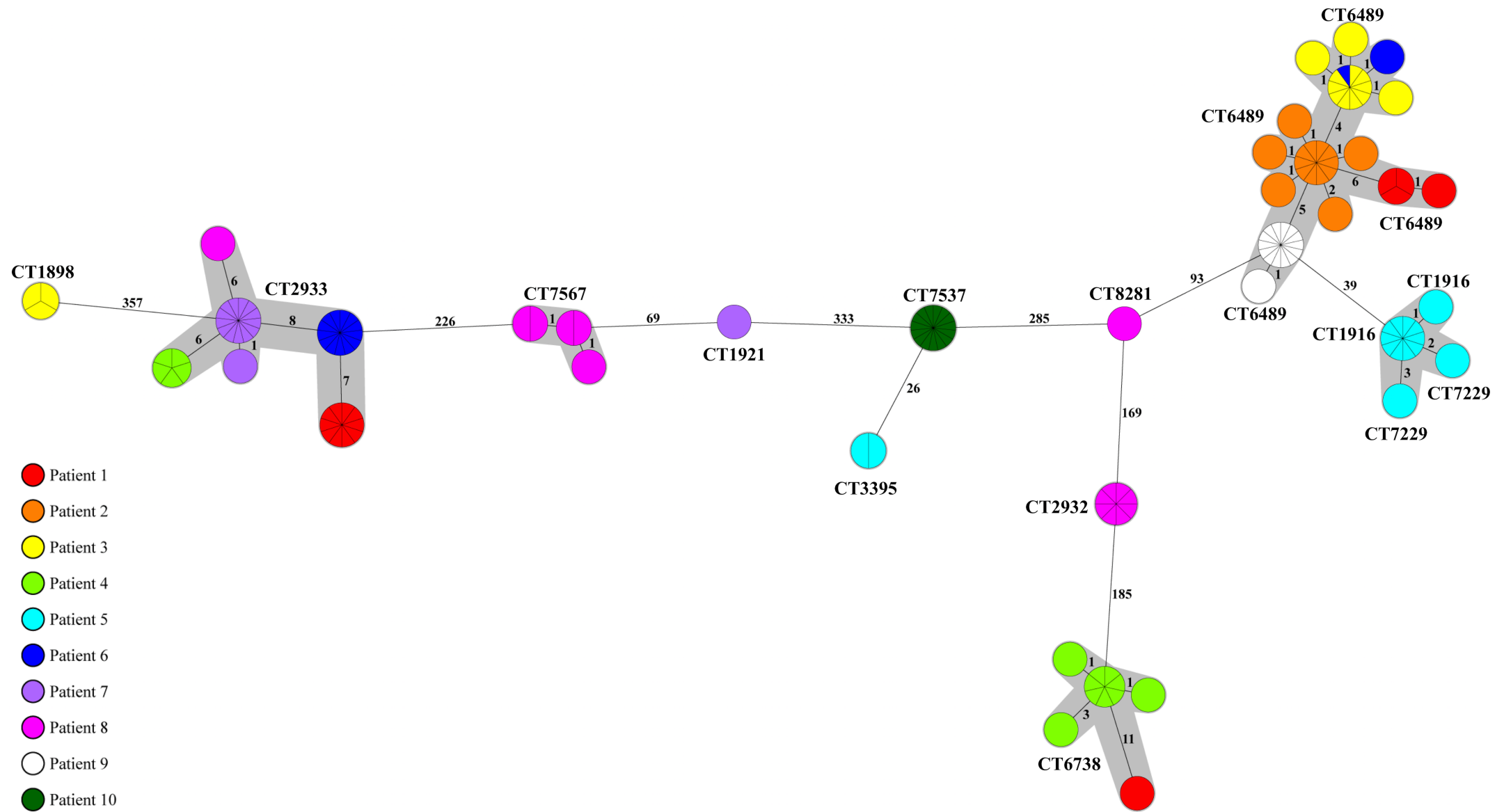


Figure 2

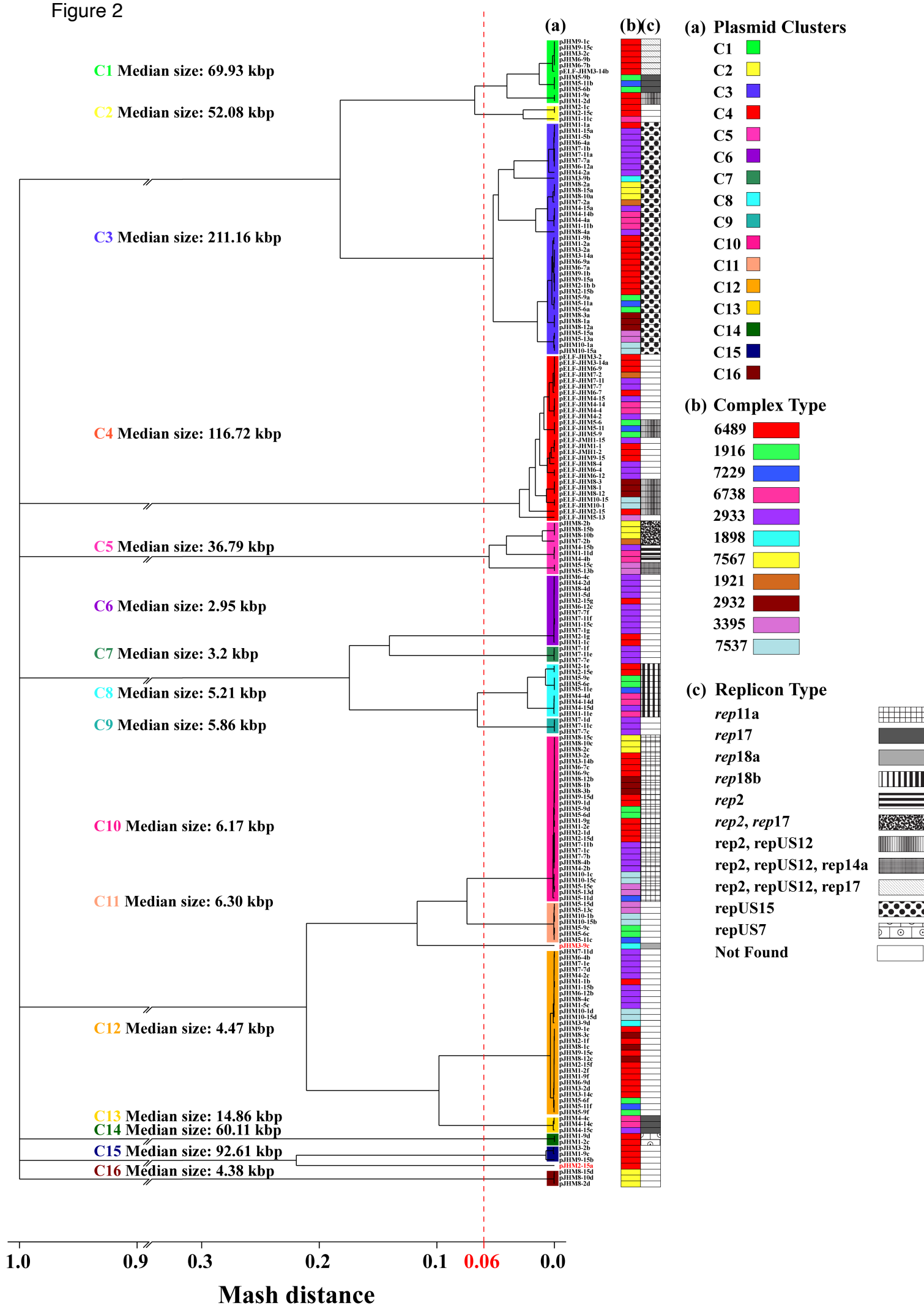
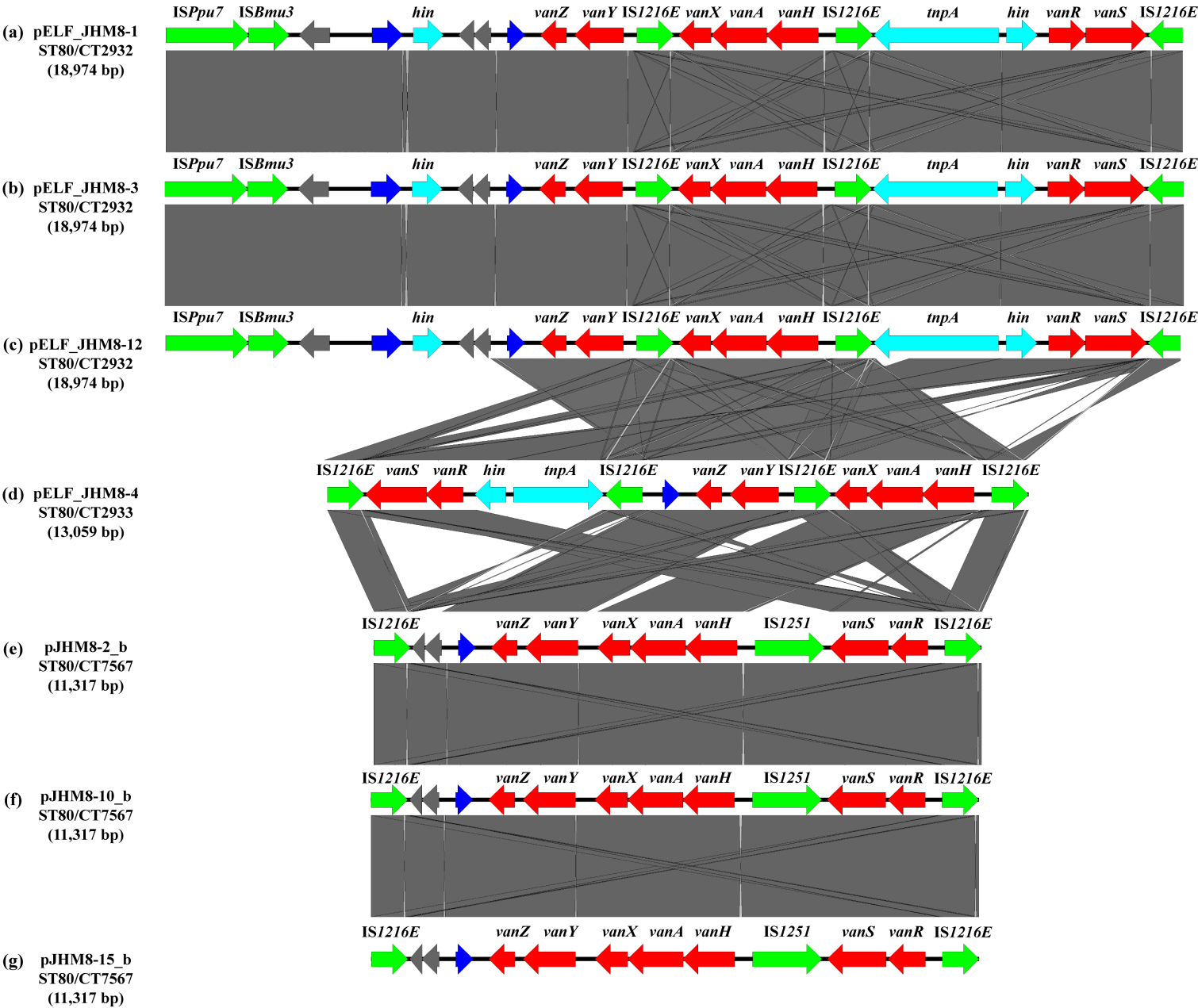


Figure 3



Average Nucleotide Identity (*vanA* transposon)

	pELF_JHM8-1	pELF_JHM8-3	pELF_JHM8-12	pELF_JHM8-4	pJHM8-2_b	pJHM8-10_b	pJHM8-15_b
pELF_JHM8-1							
pELF_JHM8-3	100						
pELF_JHM8-12	100	100					
pELF_JHM8-4	99.31	99.31	99.31				
pJHM8-2_b	98.11	98.11	98.11	97.26			
pJHM8-10_b	98.11	98.11	98.11	97.26	100		
pJHM8-15_b	98.11	98.11	98.11	97.26	100	100	

1 Kbp

100%

Supplementary Information

Co-carriage of diverse vancomycin-resistant *Enterococcus faecium* ST80-lineages by 70% of colonised patients in an Irish hospital

Nicole L. Kavanagh¹, Peter M. Kinnevey¹[ORCID](#), Grainne I. Brennan^{2,3}, Brian O'Connell^{2,3}, Richard V. Goering^{d4}, David C. Coleman^{1*} [ORCID](#)

¹*Microbiology Research Unit, Dublin Dental University Hospital, University of Dublin, Trinity College Dublin, Dublin, Ireland*

²*Department of Clinical Microbiology, St. James's Hospital, Dublin, Ireland*

³*National MRSA Reference Laboratory, St. James's Hospital, James's Street, Dublin 8, Ireland.*

⁴*Department of Medical Microbiology and Immunology, Creighton University School of Medicine Omaha, Nebraska, USA.*

*Corresponding author

Address: Microbiology Research Unit, Division of Oral Biosciences, Dublin Dental University Hospital, University of Dublin, Trinity College, Lincoln Place, Dublin D02 F859, Ireland. Tel.: +353 1 6127276; fax: + 353 1 6127295.

E-mail address: david.coleman@dental.tcd.ie (D.C. Coleman).

This file includes:

Supplemental Methods

Supplemental Results

Supplemental Figures S1-S21

Supplementary Methods

Whole genome sequencing

Genomic DNA libraries were prepared for short-read WGS using the Nextera Flex kit (Illumina, Eindhoven, The Netherlands). Sequencing was performed on the Illumina MiSeq platform (2×300 bp, paired end) using the 600-cycle Reagent Kit v3 (Illumina). Libraries were scaled to exhibit $\geq 35\times$ coverage. Long-read sequencing was undertaken by MicrobesNG (<https://microbesng.com/>). Library preparation was performed using the Oxford Nanopore Native Barcoding Kit 24 v14 (SQK-NBD114.24) and sequencing was performed on the GridION platform using a R10.4.1 Flow Cell with MinKNOW software v23.04.6 (Oxford Nanopore Technologies, United Kingdom).

Sequence analysis

Quality trimming and *de novo* assembly of short-read FASTQ files was carried out using Ridom SeqSphere+ v8.5.1 (Ridom GmbH, Münster, Germany) with default settings. Long-read sequencing files were base-called, trimmed and demultiplexed using Guppy v6.5.7 (Oxford Nanopore). Hybrid assemblies were generated using the Unicycler pipeline v0.4.0 (<https://github.com/rrwick/Unicycler>)¹. Assembly quality was assessed using the graph connectivity and node size metrics in Bandage v0.8.1 (<https://github.com/rrwick/Bandage>)². Assembled genomes were annotated using Prokka v1.14.5 (<https://github.com/tseemann/prokka>)³ and visualised using SnapGene v.7.0.4 (<https://www.snapgene.com>).

Supplementary Results

Frequency of VREfm CTs in patients with multi-strain carriage

A ‘dominant’ high-frequency VREfm CT was identified in all seven patients with multi-strain carriage that represented 53-93% of the 15 isolates investigated ($\geq 8/15$ colonies, range: 8-14). The predominant CT harboured by 3/7 patients was CT2933, whereas the remaining four patient predominant CTs belonged to CT2932, CT1916, CT6489, and CT6738. Moderate frequency CTs that represented $<35\%$ (3-5 colonies) of the 15 isolates investigated from each patient were identified in four patients (4/7), whereas minor frequency CTs that represented 15% (1-2 colonies) of the 15 isolates investigated were identified in five patients. In the three patients (P2, P9, P10) that yielded only one distinct CT, CT6849 was identified in two.

Comparison of study isolates with a five-year collection of isolates from the same hospital wards

VREfm from the present study ($n=150$) were placed in a wider genomic context with the genomes of an additional 622 previously described clade A1 clinical and screening *E. faecium* isolates (recovered between June 2017-December 2022) from individual patients and environmental sites in hospital H1 (VREfm $n=579$, VSEfm $n=43$) from the same seven wards (W1 $n=107$, W2 $n=156$, W3 $n=327$, W9 $n=7$, W18 $n=22$ and W20 $n=3$) (Supplemental Table S1)^{4,5}. The purpose of this five-year ward comparison was to highlight the

frequency at which *E. faecium* subtypes, particularly moderate and minor frequency strains, were shared between patients and their environment and to provide insights into the implications of VREfm within-patient heterogeneity on *E. faecium* population biology and surveillance practices. Previous studies employing whole-genome comparisons in VREfm outbreak analysis utilised a variety of thresholds (7-10 SNPs) to infer transmission pathways and determine genetic relatedness between isolate cohorts within specific study periods^{6,7}. However, as the comparison was not intended to determine recent transmission or infer epidemiological links, but instead to identify discrete genomic clusters and elucidate strain distribution, a less stringent threshold of ≤ 15 SKA-SNPs was adopted to deem isolates as closely related. This was determined on the basis of a previous study which delineated the population structure of clade A1 and highlighted that isolates within groups that exhibited ≤ 20 SNPs displayed limited recombination and were therefore deemed as the same sub-type descending from a common ancestor⁸. The ≤ 15 SKA-SNP threshold was established from a combination of the previously reported annual *E. faecium* diversification estimates, the five-year isolate recovery duration and previously defined thresholds describing subtypes⁹⁻¹².

cgMLST segregated isolates into 44 clusters and 38 singletons with 48% (368/772) belonging to CT2933, CT2932, CT1916 and CT6489 (Supplemental Figure S3). In the national comparison, evidence for the circulation and persistence of highly-related, minority subpopulations across multiple H1 wards was supported by a small number of examples (Supplemental Tables S5a-i). For instance, CT7567 comprised six genetically indistinguishable VREfm from two wards (W2 $n=1$, W3 $n=5$) recovered between December 2022-January 2023 (27 days) which exhibited a median pairwise SKA-SNP difference of 0 (range: 0-1) (Supplemental Table S5i). Notably, 5/6 isolates were recovered from patient P8 (W3) and represented a moderate frequency strain (33.33%, 5/15 sample colonies). Low frequency *E. faecium* sub-types were also associated with long-term ward-specific persistence; CT1898 comprised nine isolates, including a moderate frequency strain from patient P3 (20%, 3/15 sample colonies, JHM3-1, JHM3-4 and JHM3-9), recovered from two wards (W2 $n=2$, W3 $n=7$) over 39 months (June 2019-October 2022) (Supplemental Table S5f). Interestingly, the closest two neighbouring isolates (SJ178 and SJ267) to the P3 moderate frequency strains were collected 1,217 days prior and exhibited median pairwise SKA-SNP differences of 15 and 13, respectively, (Supplemental Table S5f). These findings, however, should be interpreted with caution as comparator H1 ward isolates were recovered by selecting a single bacterial colony from each sample. Consequently, a more detailed analysis is required to accurately determine the true inter and intra-ward distribution of strains and highlight the impact of low frequency strain transmission.

Pairwise Mash distance-based clustering of plasmids

The largest cluster, termed cluster C3, comprised 39 ‘mega-plasmids’ of replicon type repUS15 ranging in size from 116,547 to 274,157 bp distributed across ten diverse CTs (Supplemental Table S7c). These plasmids carried four distinct antimicrobial resistance genes conferring resistance to antibiotic classes including aminoglycosides and lincosamides. The median number of antimicrobial resistance genes per repUS15 plasmid was one (range: 1-4) (Supplemental Table S7c).

Within patient P1, all six fully resolved VREfm isolates from a total of three different CTs (CT6489, CT2933 and CT6738) harboured a repUS15 plasmid (Supplemental Tables S6 and S7c). Interestingly, isolates JHM1-1 (ST80/CT6489) and JHM1-15 (ST80/CT2933) harboured genetically identical (median Mash plasmid distance=0.0000) mega-plasmids, despite the isolates exhibiting a pairwise median SKA-SNP difference of 2822 (Supplemental Tables S3a and S7c). Further investigation into the entire plasmidome landscape of these isolates revealed a total of four distinct plasmids in each isolate that were near-identical (Figure 3, Supplemental Tables S7c, d, f, and l). Notably, none of the four-plasmid types contained known relaxase or mate-pair formation type genes involved in mobilisation and were deemed ‘non-conjugative’ by MOB-suite tools¹³.

Near-identical plasmids in diverse genetic backgrounds recovered from multiple patients were also evident. For instance, the second largest Mash distance cluster, cluster C10, contained 28 small, mobilisable plasmids (~6 kb) of replicon type *rep11a* (Figure 2). Isolates JHM3-14 (ST80/CT6489) from patient P3 and JHM8-1 (ST80/CT2932) from patient P8 harboured genetically identical *rep11a* plasmids (Mash difference of 0.0000) despite the isolates belonging to different CTs (Supplemental Table S7j). These findings indicate that there is potentially a common pool of plasmid backbones in circulation within Irish ST80 VREfm that are disseminated across diverse genetic backgrounds.

Within-patient VREfm diversity and Tn1546-like vanA transposon organisation

Three CT6489 isolates and one CT2933 isolate from patient P1 all encoded a *vanA* region of 11,916 bp with an identical structural organisation of the *van* genes, including four copies of *IS1216E* and a *tnpA* transposase gene harbouring a 1,287 bp deletion in the 3' terminal (Supplemental Figure S6). These four VREfm belonged to diverse genetic backgrounds and exhibited a median pairwise SKA-SNP distance of 1410.5 (range: 0-2818) (Supplemental Table S3a). Notably, three of the patient isolates (JHM1-1 (CT6489), JHM1-2 (CT6489) and JHM1-15 (CT2933)) harboured *vanA* on a near-identical linear plasmid (median Mash distance of 0.00017) of approximately 109,000 bp suggesting within-patient horizontal gene transfer (HGT) events (Supplemental Table S7d). An identical *vanA* region structural organisation (average nucleotide identity of 99.99%) was also identified in P1 isolate JHM1-9 (CT6489), however the plasmid topology could not be determined as it was not fully resolved during hybrid assembly and instead distributed across 9 contigs of total length 101,880 bp.

Seven VREfm isolates from three CTs were recovered from patient P8 (CT2932 n=3, CT2933 n=1 and CT7567 n=3), which exhibited a median pairwise SKA-SNP difference of 2881 (range: 0-3227) (Supplemental Table S3h, Figure 3). Limited structural variability was observed when comparing *vanA* regions from VREfm of the same CT within P8; the three CT2932 isolates (JHM8-1, JHM8-3, and JHM8-12) from patient P8 all harboured a Tn1546-like transposon of 18,973 bp encoding the *van* genes in an identical orientation, three copies of *IS1216E*, one copy of *ISPpu7* and *ISBmu3* and a *tnpA* gene with a 645 bp 3' terminal deletion (Figure 3). This Tn1536-like subtype was encoded on three near-identical linear plasmids in the three separate P8 CT2932 isolates (median pairwise Mash distance of 0.00005) of approximately 140,000 bp which also harboured *ant(6)-Ia*, *aph(3')-III* and *erm(B)* antimicrobial resistance

genes (Supplemental Table S7d). Three CT7567 patient P8 isolates (JHM8-2, JHM8-10 and JHM8-15) harboured *vanA* (11,317 bp) on a 40,922 bp circular, multi-replicon plasmid that was characterised as non-mobilisable (Figure 3, Supplemental Table S7e). The *vanA* region of the three P8 CT7567 isolates shared an average nucleotide identity of 98.11% with the corresponding *vanA* region of CT2932 isolates. Key differences in CT7567 included a reverse orientation of the *van* regulatory genes (*vanRS*), absence of the *tnpA* and *hin* genes and an insertion of *IS1251* upstream of the *vanHAX* gene operon (Figure 3). The single P8 CT2933-VREfm isolate (JHM8-4) harboured a different *vanA* structural organisation encoded on a linear plasmid of 110,419 bp with a distinct genetic backbone (Median pairwise Mash distance of 0.01061) to the CT2932 linear plasmids (Supplemental Table S7d, Figure 3). Differences included dissimilar orientations of the *vanA* operon genes and a *tnpA* gene with a 1,287 bp deletion in the 3' terminal region (Figure 3). The *vanA* region of JHM8-4 (CT2933) exhibited an average nucleotide identity of 99.31% and 97.26% with CT2932 and CT7567 isolates, respectively.

After investigating the influence of multi-strain VREfm carriage on *vanA* diversity, the next step was to analyse the structural organisation of the resolved extrachromosomal *vanA* harbouring regions from patients P1-10 within specific CT groups to determine if there was an association between transposon structural iterations and specific CTs. The two largest CTs, 2933 and 6489, comprising a total of 46% (18/39) of the within-patient diversity isolates investigated that were fully closed by hybrid assembly were investigated (Supplemental Figures S15 and S16). This analysis revealed the presence of five unique iterations of the *vanA* transposon within CT2933 isolates from five patients. The key differences included multiple insertions of different IS elements such as *IS1216E*, *ISEf1*, *ISPpu7* and *ISBmu3* throughout the operon, additional gene insertions such as cadmium efflux accessory proteins, rearrangement of the *van* operon genes and various deletions in the 3'-terminal end of the *tnpA* gene (range: 645-1,287 bp) (Supplemental Figure S15). Similar findings of multiple diverse iterations of *vanA* transposon structures were observed in CT6489 isolates from five distinct patients (Supplemental Figure S16). Interestingly, patients P3 and P6 harboured very closely related VREfm CT6489 strains (median pairwise SKA-SNP difference: 3, range: 1-6) suggestive of potential transmission. Comparison of the structural organisation of the *vanA* regions within these isolates revealed Tn1546 *vanA* diversity (Supplemental Figures S16f-i) highlighting that during transmission of clonally indistinguishable variants between patients structural variation can occur at both transposon and plasmid level with the gain/loss of segments of the *vanA* operon.

A snapshot of VSEfm carriage

CT2932 comprised 15 isolates (eight VREfm from patient P8 and seven VSEfm from patient P1) with a median cgMLST allelic difference of 2 (range: 0-11). However, SKA-SNP analysis revealed greater diversity within this cluster, with P8 and P1 isolates exhibiting a median SNP pairwise difference of 138 (range: 134-149). A representative CT2932-VSEfm isolate (JHMS1-1) was fully resolved by hybrid assembly to provide an insight into the accessory genome. This revealed the presence of six plasmids, ranging in size from 6,357-213,737 bp encoding antimicrobial resistance genes such as *ant(6)-Ia*, *aph(3')-III* and *erm(B)*. However, no *vanA* genes or remnants were found. Comparison of the structural organisation of

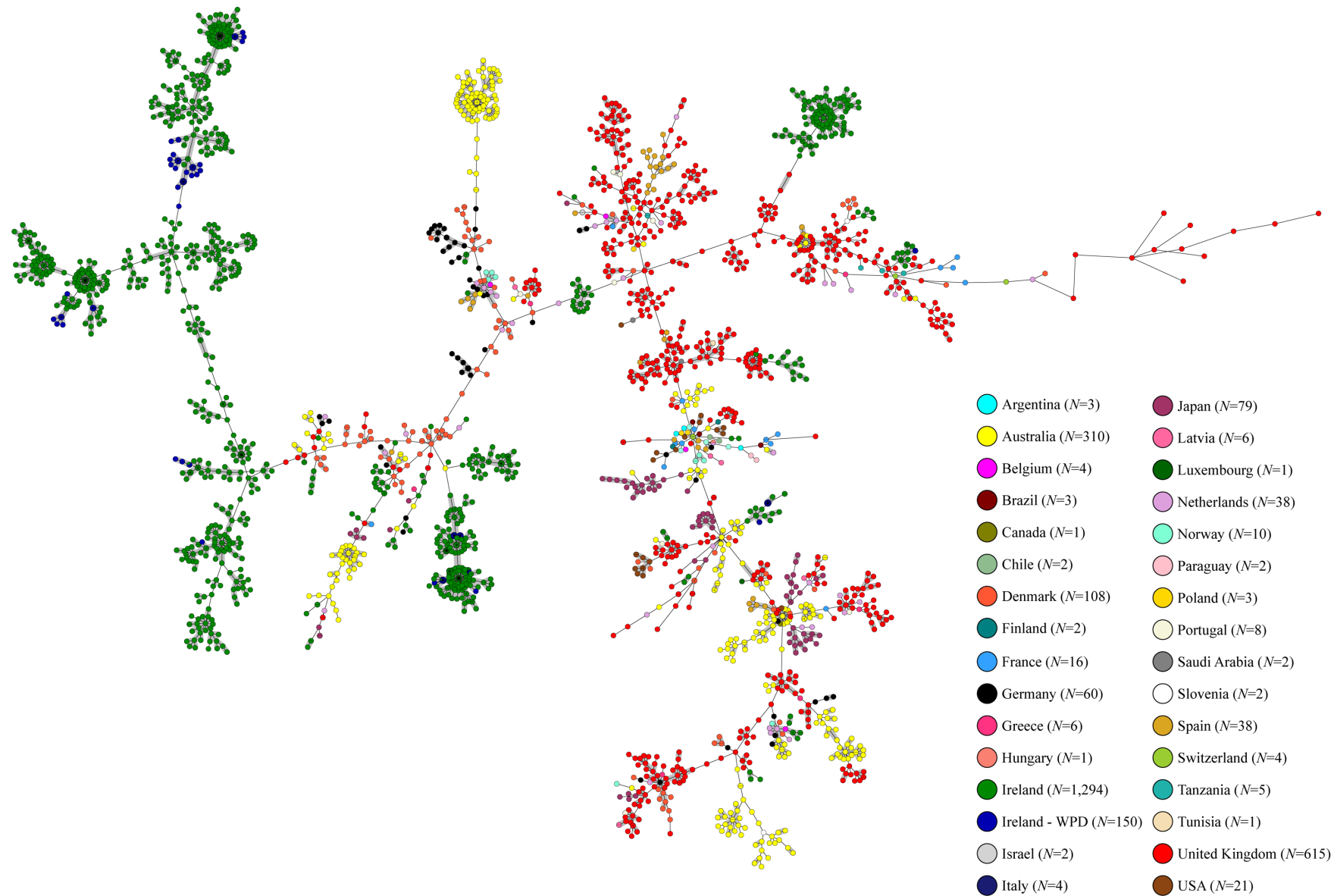
the accessory genome of the CT2943-VSEfm isolate (JHMS1-1) with a CT2932-VREfm isolate (JMH8-1) revealed the presence of some similarities (Supplemental Figure S19). Notably, JHMS1-1 harboured a linear plasmid (122,869 bp) that exhibited 99.88% average nucleotide identity to the *vanA*-encoding linear plasmid (146,603 bp) from VREfm isolate JHM8-1. Screening of the short-read derived FASTA contigs from the remaining six patient P1 CT2932 VSEfm isolates using ABRicate v.0.9.8 (<https://github.com/tseemann/abricate>) revealed similar findings, with the absence of any *vanA* operon genes.

Two representative isolates from the remaining two CT groups comprising only VSEfm isolates (CT2940 and CT2950) were investigated using hybrid assembly (Supplemental Figures S17 and S20). P3 CT2950-VSEfm isolate JHMS3-5 harboured seven plasmids ranging in size from 4,466-192,687-bp, including a plasmid of 115,993 bp that displayed linear topology. Mapping of the short-read derived FASTA contigs from JHMS3-5 using ABRicate revealed the presence of the *vanA* associated sensor-regulatory genes (*vanR* and *vanS*). Interestingly, hybrid-assembly of JHMS3-5 revealed a linear plasmid of 115,993 bp encoding a *vanA* remnant. The plasmid was enriched with four copies of *IS1216E*, with the *vanA*-remnant (*vanRS*, *tnpA* and *hin_2*) separated into a distinct segment flanked on either side by *IS1216E* (Supplemental Figure S21). Screening of the short-read derived FASTA contigs from the remaining four P5 CT2950-VSEfm isolates (Figure S17) using ABRicate failed to identify any *vanA*-associated operon genes. CT2940-VSEfm isolate JHMS5-15 harboured four distinct plasmids ranging in size from 6,173-197,142 bp (Figure S20). Plasmids in this isolate encoded antimicrobial resistance genes such as *ant(6)-Ia*, *aph(3')-III* and *erm(B)*, however no *vanA* operon remnants or genes were identified.

References

1. 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.
2. Wick RR, Schultz MB, Zobel J *et al.* Bandage: interactive visualization of de novo genome assemblies. *Bioinformatics* 2015; **31**: 3350e2.
3. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 2014; **30**: 2068e9.
4. Egan SA, Kavanagh NL, Shore AC *et al.* Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar *vanA* regions via *IS1216E* and plasmid transfer in diverse genetic lineages in Ireland. *J Antimicrob Chemother* 2021; **77**: 320-30.
5. Kavanagh NL, Kinnevey PM, Egan SA *et al.* Protracted transmission and persistence of ST80 vancomycin-resistant *Enterococcus faecium* clonal complex types CT2933, CT2932 and CT1916 in a large Irish hospital: a 39-month whole-genome sequencing study. *J Hosp Infect* 2025; **151**: 11-20.
6. Higgs C, Sherry NL, Seemann T *et al.* Optimising genomic approaches for identifying vancomycin-resistant *Enterococcus faecium* transmission in healthcare settings. *Nat Commun* 2022; **26**: 509.
7. Maechler F, Weber A, Schwengers O *et al.* Split *k*-mer analysis compared to cgMLST and SNP-based core genome analysis for detecting transmission of vancomycin-resistant enterococci: results from routine outbreak analyses across different hospitals and hospitals networks in Berlin, Germany. *Microb Genom* 2023; **9**: mgen000937.
8. Gouliouris T, Coll F, Ludden C *et al.* Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat Microbiol* 2021; **5**: 103-11

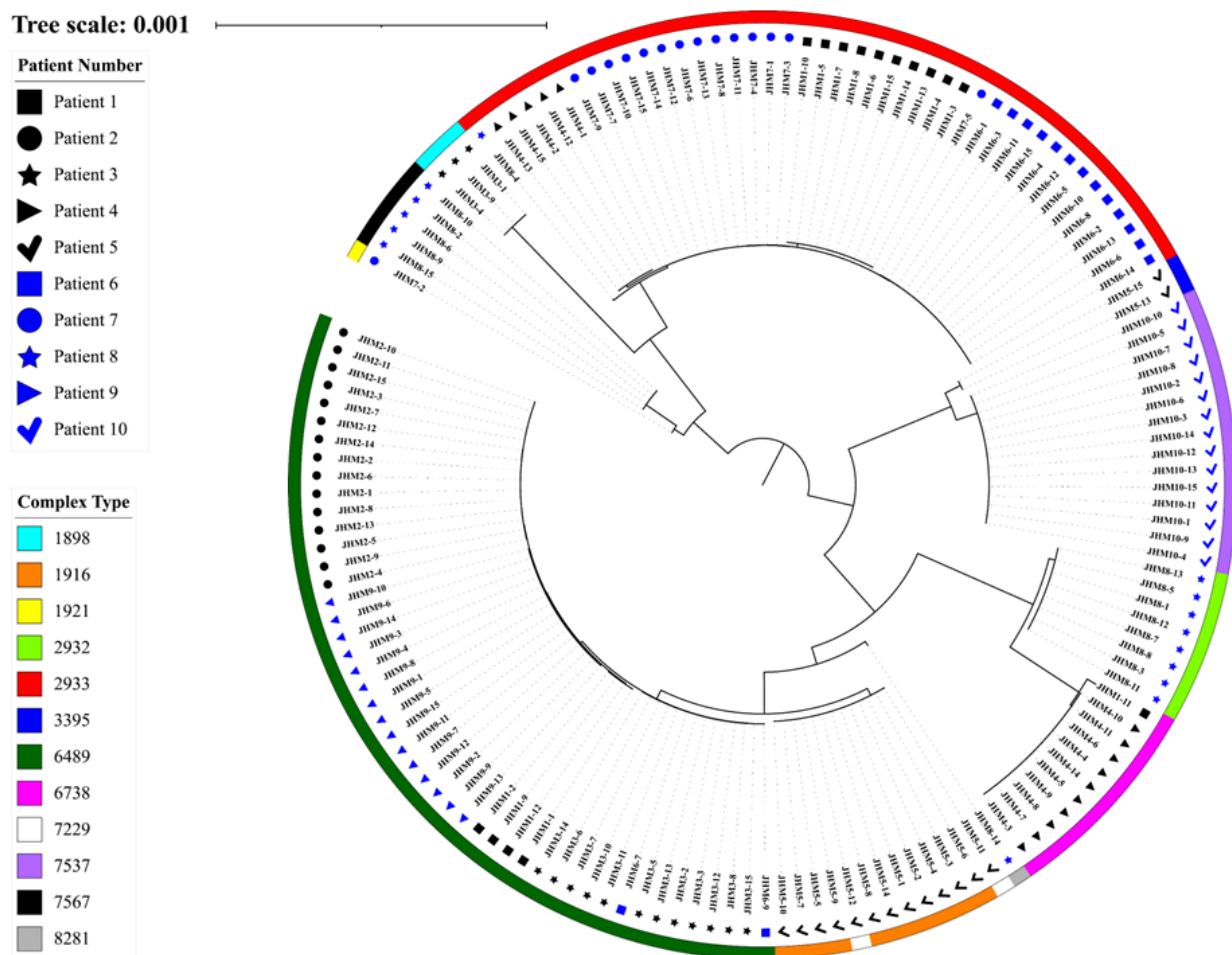
9. Moradigaravand D, Gouliouris T, Blane B *et al.* Within-host evolution of *Enterococcus faecium* during longitudinal carriage and transition to bloodstream infection in immunocompromised patients. *Genome Med* 2017; **9**: 199.
10. Raven KE, Gouliouris T, Parkhill J *et al.* Genome-based analysis of *Enterococcus faecium* bacteremia associated with recurrent and mixed-strain infection. *J Clin Microbiol* 2018; **56**: e01520-17.
11. Bayjanov JR, Baan J, Rogers MRC *et al.* *Enterococcus faecium* genome dynamics during long-term asymptomatic patient gut colonization. *Microb Genom* 2019; **5**: e000277.
12. 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.
13. Robertson J, Nash JHE. MOB-suite: software tools for clustering, reconstruction and typing of plasmids from draft assemblies. *Microb Genom* 2018; **4**: e000206.



Supplemental Figure S1. Minimum spanning tree (MST) based on core genome multi-locus sequencing typing (cgMLST) analysis of 1,4234 loci of study isolates, Irish Clade A1 *E. faecium* (n=1,444) recovered in seven Irish hospitals between June 2017 and December 2022 and global Clade A1 *E. faecium* (n=1,356) from 30 countries, spanning a 30-year time period (1986-2016)¹⁻³ Isolates are coloured by country of origin as indicated in the legend. Irish within-patient diversity isolates from the present are coloured dark blue (n=150), whilst Irish isolates described in previous studies from this laboratory (n=1,294) are coloured green.

References:

1. Egan SA, Kavanagh NL, Shore AC, *et al.* Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar *vanA* regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland. *J Antimicrob Chemother* 2021; **77**: 320–30. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8809189/>.
2. 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. Available at: <https://genome.cshlp.org/content/26/10/1388>.
3. van Hal SJ, Willems RJL, Gouliouris T, *et al.* The global dissemination of hospital clones of *Enterococcus faecium*. *Genome Med* 2021; **13**: 52. Available at: <https://genomemedicine.biomedcentral.com/articles/10.1186/s13073-021-00868-0>



Supplemental Figure S2. Maximum likelihood tree (MLT) based on phylogenetic analysis of 150 Clade A1 vancomycin-resistant *Enterococcus faecium* (VREfm) isolates recovered from a single rectal screening swab from each of 10 separate patients. Detailed metadata pertaining to patient number and core-genome complex type (CT) are indicated in the legend. A tree scale bar is shown at the top left. The tree is mid-point rooted. The corresponding pairwise indices and distances for all isolates are provided in Tables S3a-j.

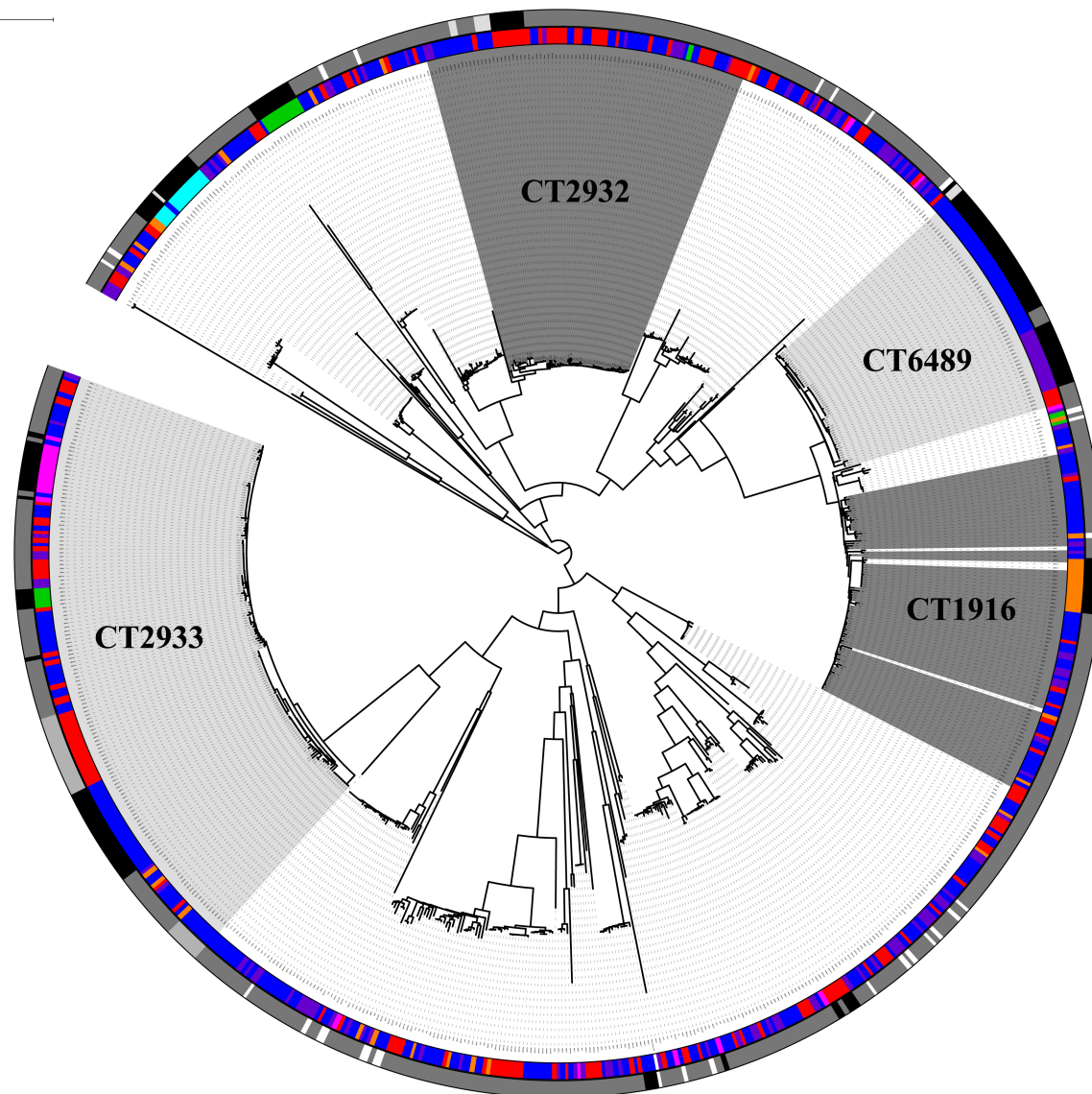
Tree scale: 0.1

Ward Location

- W1
- W2
- W3
- W9
- W18
- W20
- EMV

Isolate Source

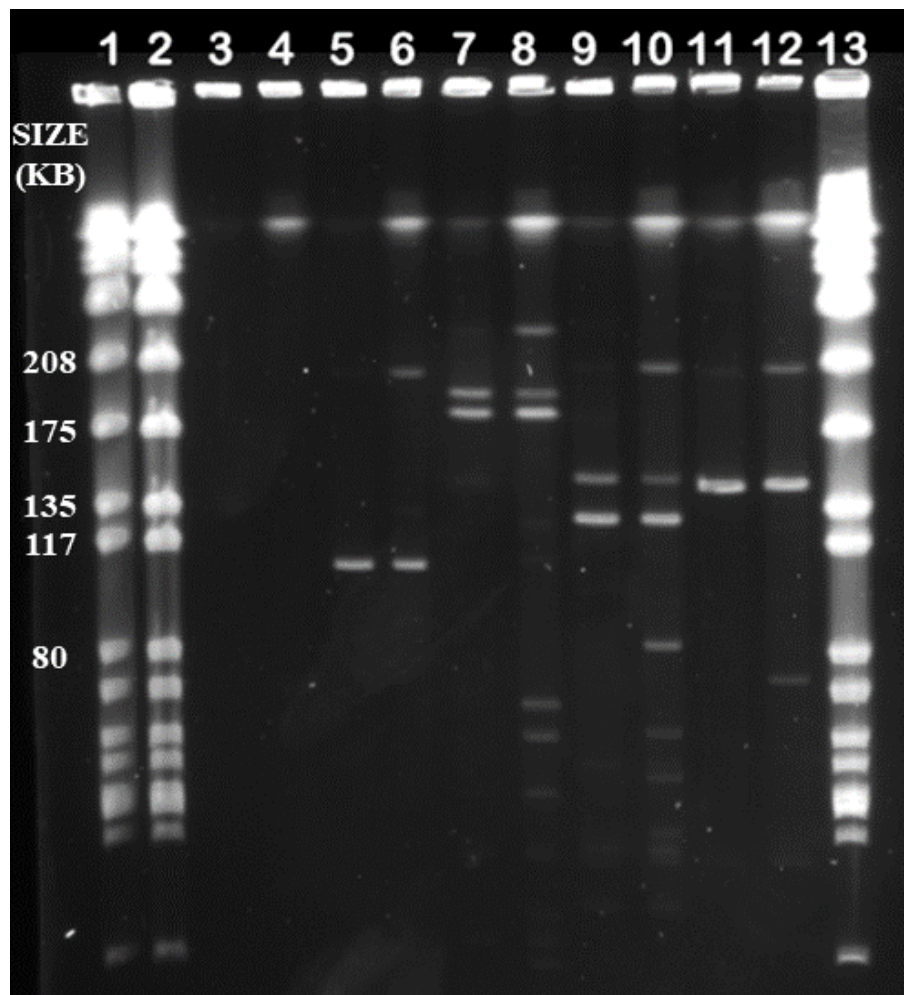
- Bloodstream Infection
- Environmental
- Handwash Basin Drain Outlet
- Rectal Screening
- Within-patient Diversity



Supplemental Figure S3. Neighbour joining tree (NJT) based on core-genome multi locus sequence type (cgMLST) analysis of 1,423 target genes for 772 Clade A1 *E. faecium* isolates recovered from wards (W)s W1, W2, W3, W9, W18, W20 and W21 in Hospital 1 between June 2017 and December 2022. The NJT comprises 622 isolates from H1 reported in previous studies from this laboratory and within-patient diversity isolates from this study ($N=150$). Isolates segregated into 44 clusters and 38 singletons^{1,2}. Four distinct clusters of closely related isolates (≤ 20 allelic differences) are highlighted, which are shaded in grey with the corresponding complex type (CT) indicated as follows: CT2933 (151 isolates): average allelic difference: 8, range: 0-22; CT2932 (80 isolates): average allelic difference: 6, range: 0-18; CT1916 (81 isolates): average allelic difference 11, range 0-28; CT6489 (56 isolates): average allelic difference: 6 range: 0-20. Metadata pertaining to ward location, isolate source and vancomycin susceptibility phenotype are indicated in the legends. A tree scale bar is shown at the top left.

References:

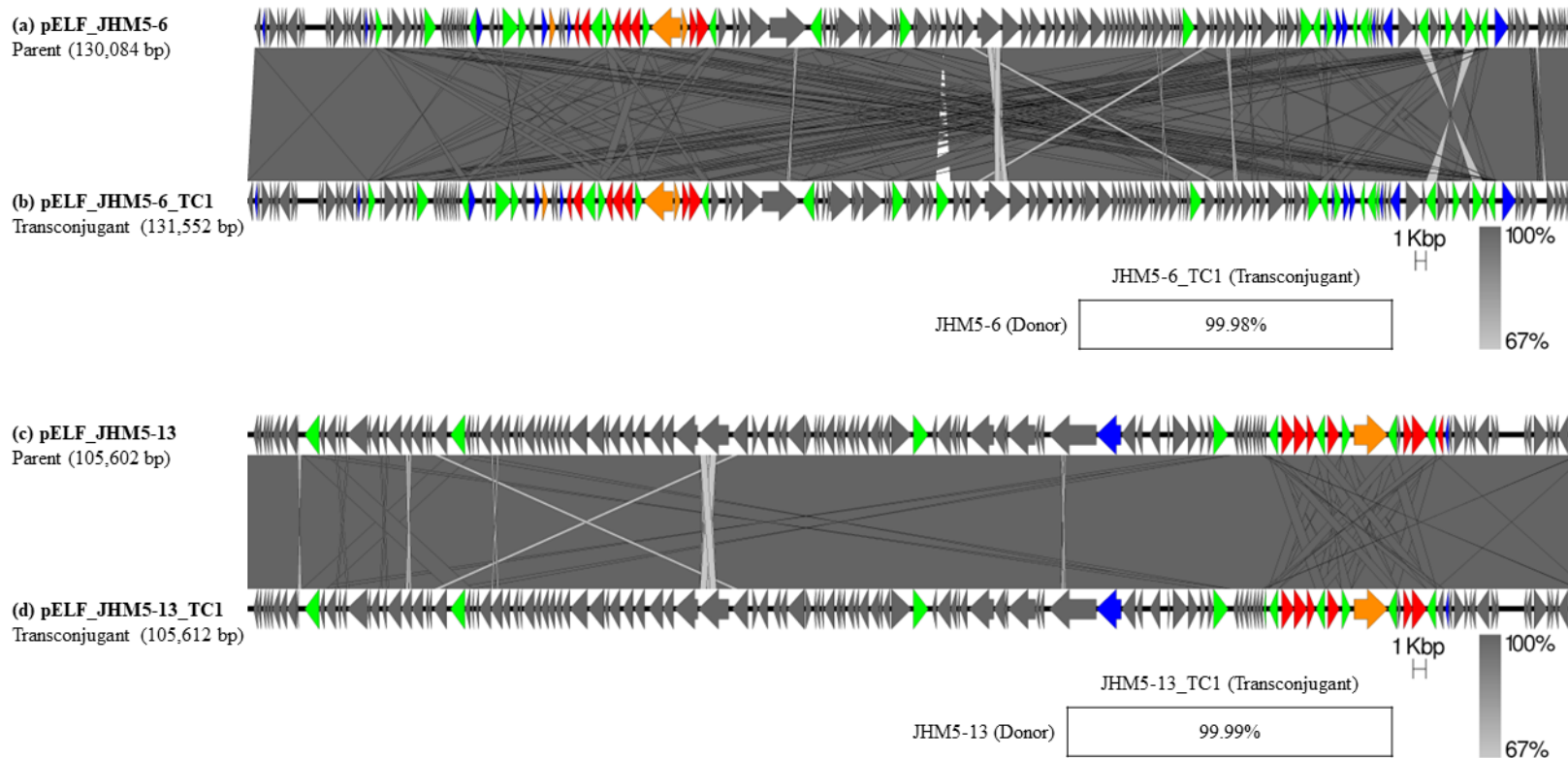
1. Egan SA, Kavanagh NL, Shore AC, *et al.* Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar *vanA* regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland. *J Antimicrob Chemother* 2021; **77**: 320–30. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8809189/>
2. Kavanagh NL, Kinnevey PM, Egan SA, *et al.* Protracted transmission and persistence of ST80 vancomycin-resistant *Enterococcus faecium* clonal complex types CT2933, CT2932 and CT1916 in a large Irish hospital: a 39-month whole-genome sequencing study. *J Hosp Infect* 2024; **151**: 11–20. Available at: <https://www.sciencedirect.com/science/article/pii/S0195670124002184>



Supplemental Figure S4. Pulsed field gel electrophoresis (PFGE) analysis of undigested and S1 nuclease-treated DNA of four representative vancomycin-resistant *Enterococcus faecium* isolates harbouring *vanA*-encoding linear plasmids. Lanes 1-13 are as follows: 1 and 2, *Sma*I-digested *S. aureus* 8325¹ genomic DNA molecular size standards; 3, S1-nuclease untreated plasmid free *E. faecium* strain Efm 64/3²; 4, S1-nuclease treated plasmid free *E. faecium* strain Efm 64/3; 5, S1 nuclease untreated JHM1-15 (ST80/CT2933, 109,014 bp linear plasmid, P1); 6, S1 nuclease-treated JHM1-15 (109,014 bp linear plasmid); 7, S1 nuclease untreated JHM2-15 (ST80/CT6489, 186,811 bp linear plasmid, P2); 8, S1 nuclease-treated JHM2-15 (186,811 bp linear plasmid); 9, S1 nuclease untreated JHM5-6 (ST80/CT1916, 130,084 bp linear plasmid, P5); 10, S1 nuclease-treated JHM5-6 (130,084 bp linear plasmid); 11, S1 nuclease untreated JHM8-1 (ST80/CT2932, 146,603 bp linear plasmid, P8); 12, S1 nuclease-treated JHM8-1 (146,603 bp linear plasmid); 13, *Sma*I-digested *S. aureus* 8325 size molecular size standards. The electrophoretic mobility of all four linear plasmids during PFGE was comparable with or without S1 nuclease treatment, which nicks supercoiled circular plasmid DNA, confirming their linearity.

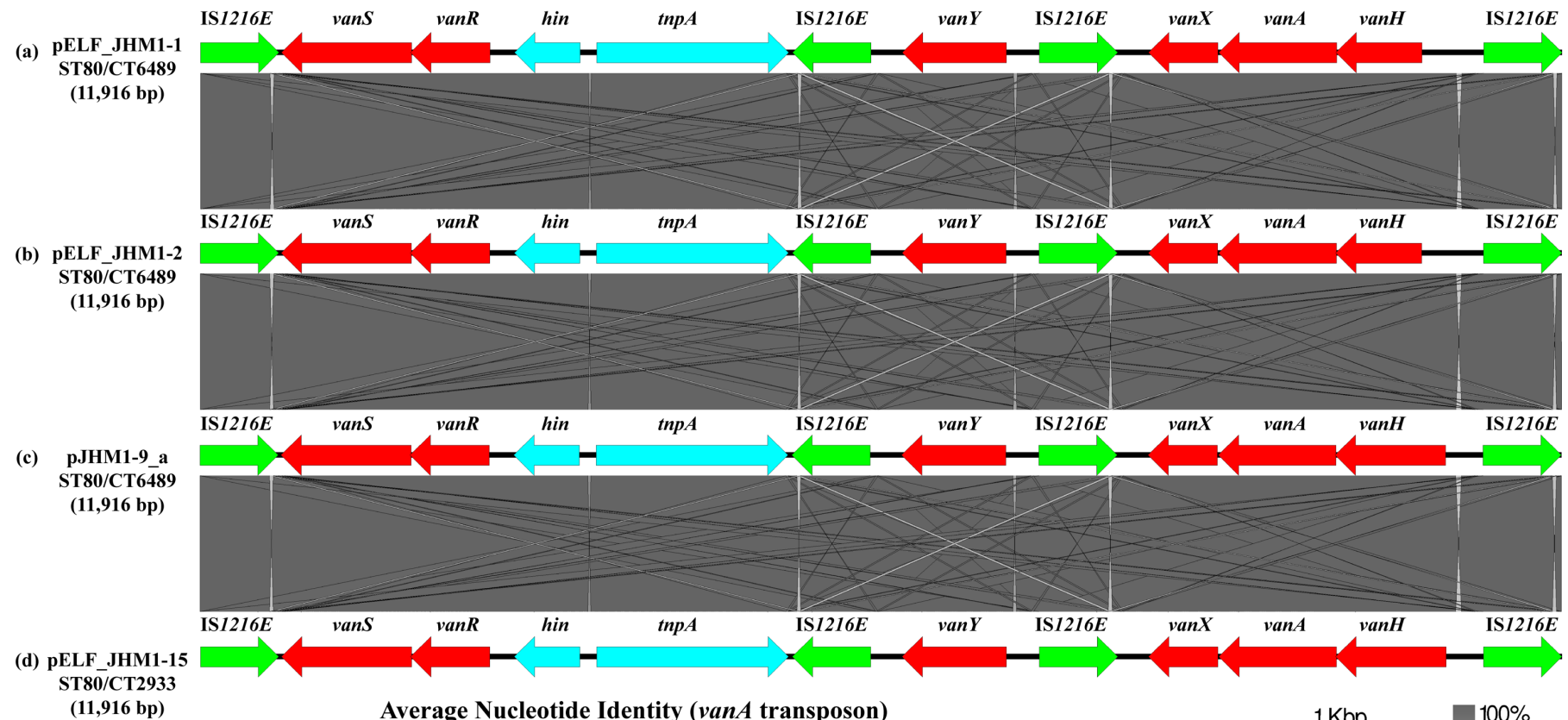
References:

1. Novick R. Properties of a cryptic high-frequency transducing phage in *Staphylococcus aureus*. *Virology* 1967; **33**: 155–66. Available at: <https://www.sciencedirect.com/science/article/pii/0042682267901055>.
2. Bender JK, Fleige C, Lange D *et al*. Rapid emergence of highly variable and transferable oxazolidinone and phenicol resistance gene *optrA* in German *Enterococcus spp.* clinical isolates. *Int J Antimicrob Agents* 2018; **52**: 819–27. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0924857918302681?via%3Dihub>



Supplemental Figure S5. Comparative genetic organisation of linear plasmids pELF_JHM5-6 and pELF_JHM5-13 from parental Irish vancomycin-resistant *Enterococcus faecium* (VREfm) isolates JHM5-6 (ST80/CT1916, P5) and JHM5-13 (ST80/CT3395, P5) and transconjugant derivatives of *E. faecium* 64/3¹ (ST21) obtained by filter-mating and closed by hybrid assembly. The figure shows a multiple alignment of (a) pELF_JHM5-6 (plasmid donor), (b) pELF_JHM5-6_TC1 (*E. faecium* 64/3 transconjugant), (c) pELF_JHM5-13 (plasmid donor) and (d) pELF_JHM5-13_TC1 (*E. faecium* 64/3 transconjugant). Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences and resolvases; orange: transposases and resolvases; blue proteins with a known coding function. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (67-100%). A scale bar is included beneath the figure. The matrix shown beneath the figure contains an $n \times n$ average nucleotide identity between the plasmids.

Reference: 1. Bender JK, Fleige C, Lange D *et al.* Rapid emergence of highly variable and transferable oxazolidinone and phenicol resistance gene *optrA* in German *Enterococcus* spp. clinical isolates. *Int J Antimicrob Agents* 2018; 52: 819–27. Available at: <https://www.sciencedirect.com/science/article/abs/pii/S0924857918302681?via%3Dihub>

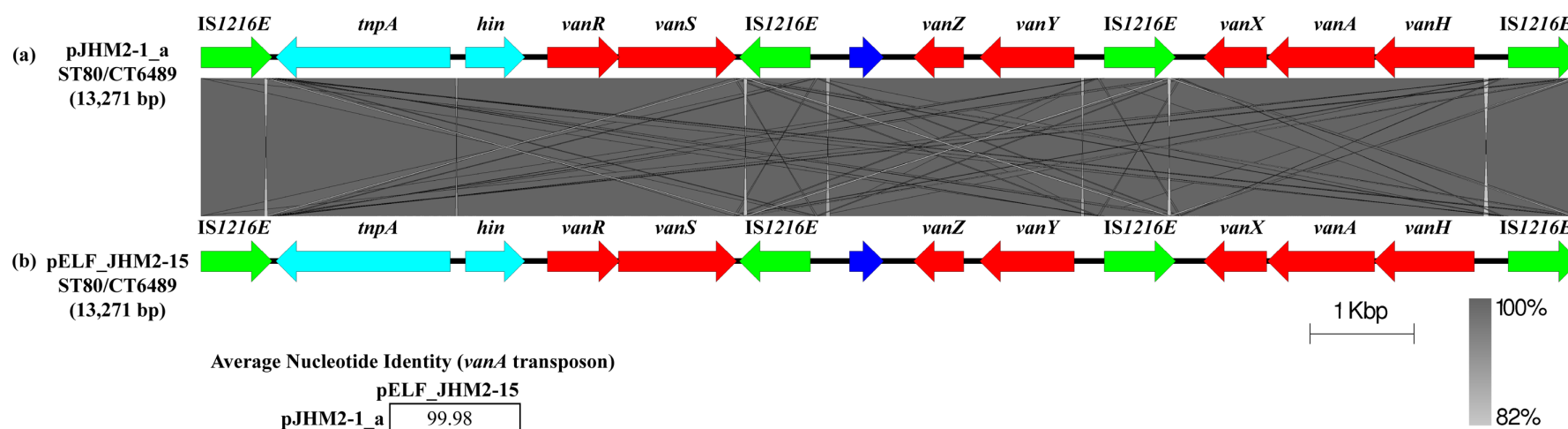


	pELF_JHM1-1	pELF_JHM1-2	pJHM1-9_a	pELF_JHM1-15
pELF_JHM1-1				
pELF_JHM1-2	100			
pJHM1-9_a	99.99	99.99		
pELF_JHM1-15	99.99	99.99	100	

Supplemental Figure S6. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of VREfm isolates belonging to two complex types (CTs) from patient P1. The figure shows an alignment of the *vanA* region encoded by four plasmids from two VREfm isolate CTs as follows: (a) linear plasmid pELF_JHM1-1 (108,997 bp) from isolate JHM1-1 (ST80/CT6489, *vanA* region 11,916 bp), (b) linear plasmid pELF_JHM1-2 (108,982 bp) from isolate JHM1-2 (ST80/CT6489, *vanA* region 11,916 bp), (c) plasmid* pJHM1-9_a from isolate JHM1-9 (ST80/CT6489, *vanA* region 11,916 bp) and (d) linear plasmid pELF_JHM1-15 (109,014 bp) from isolate JHM1-15 (ST80/CT2933, *vanA* region 11,916 bp) generated using Easyfig.¹ Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). A scale bar is included beneath the figure. *This plasmid was not fully resolved by hybrid assembly. Its sequences were spread across 25 contigs with a total length of 140,825 bp (Supplemental Table A6).

Reference:

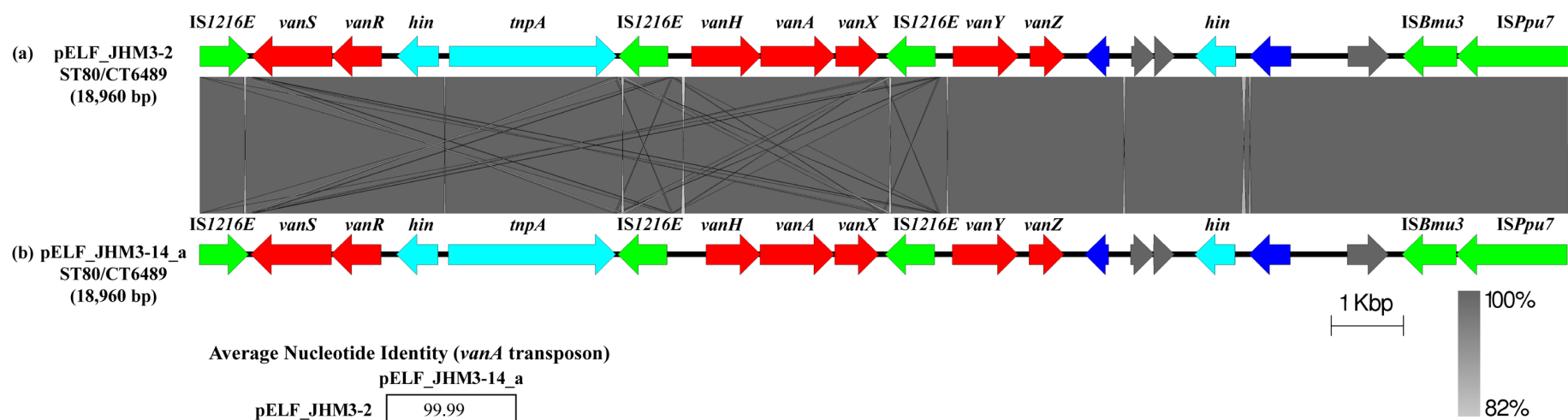
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>. A



Supplemental Figure S7. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of VREfm isolates belonging to complex type (CT) CT6489 from patient P2. The figure shows an alignment of *vanA* region encoded by plasmids from two VREfm ST80/CT6489 isolates as follows: (a) pJHM2-1_a* from isolate JHM2-1 (ST80/CT6489, *vanA* region 13,271 bp) and (b) linear plasmid pELF_JHM2-15 (186,911 bp) from isolate JHM2-15 (ST80/CT6489, *vanA* region 13,271 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). A scale bar is included beneath the figure. *This plasmid was not fully resolved by hybrid assembly. Its sequences were spread across 10 contigs with a total length of 186,915 bp (Table S6).

Reference:

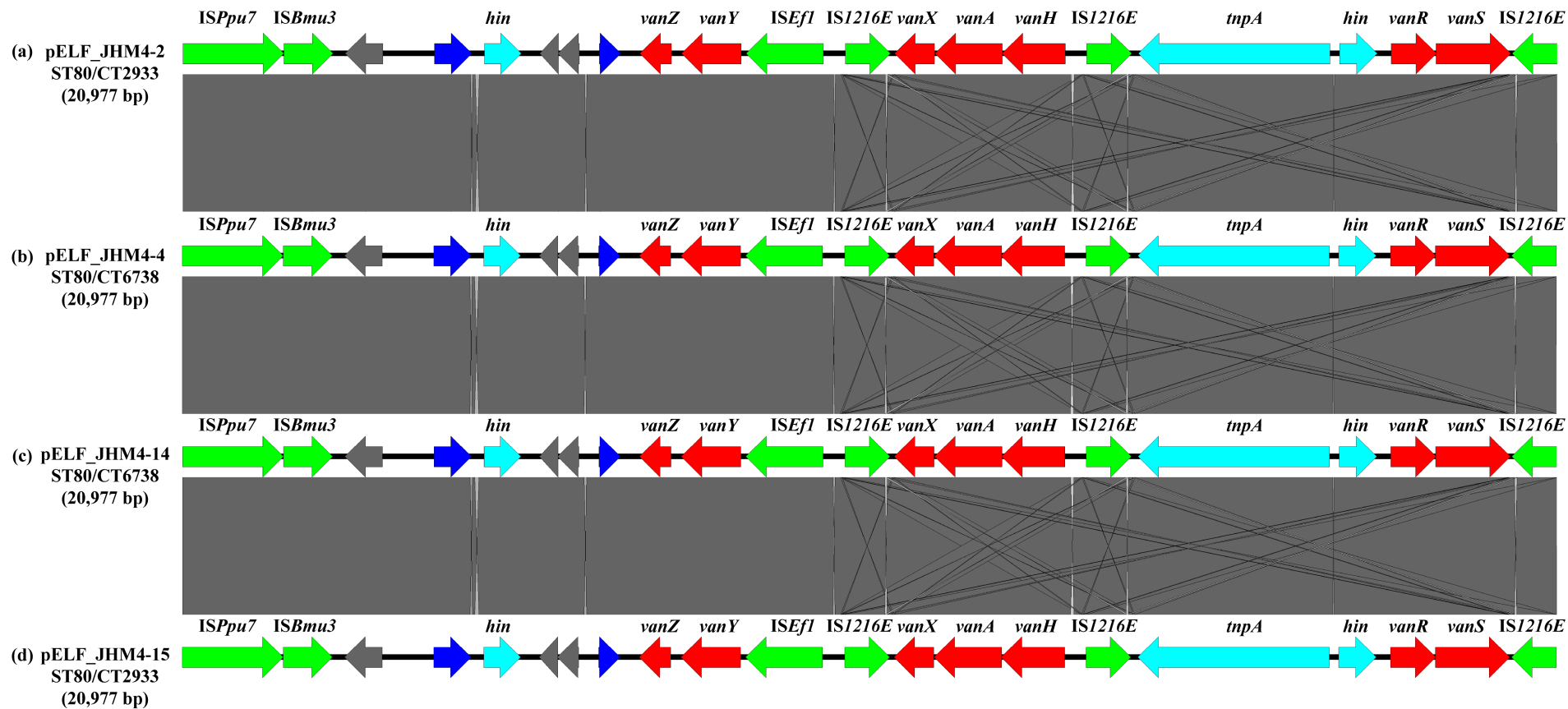
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S8. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of VREfm isolates belonging to complex type (CT) CT6489 from patient P3. The figure shows an alignment of the *vanA* region encoded by plasmids from two VREfm ST80/CT6489 isolates as follows: (a) linear plasmid pELF_JHM3-2 (116,719 bp) from isolate JHM3-2 (ST80/CT6489, *vanA* region 18,960 bp) and (b) linear plasmid pELF_JHM3-14_a (110,743 bp) from isolate JHM3-14 (ST80/CT6489, *vanA* region 18,960 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Tables 5.4-5.6. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Supplemental Table A6. A scale bar is included beneath the figure.

Reference:

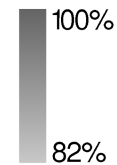
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Average Nucleotide Identity (*vanA* transposon)

	pELF_JHM4-2	pELF_JHM4-4	pELF_JHM4-14	pELF_JHM4-15
pELF_JHM4-2				
pELF_JHM4-4	99.99			
pELF_JHM4-14	100	99.99		
pELF_JHM4-15	99.99	100	99.99	

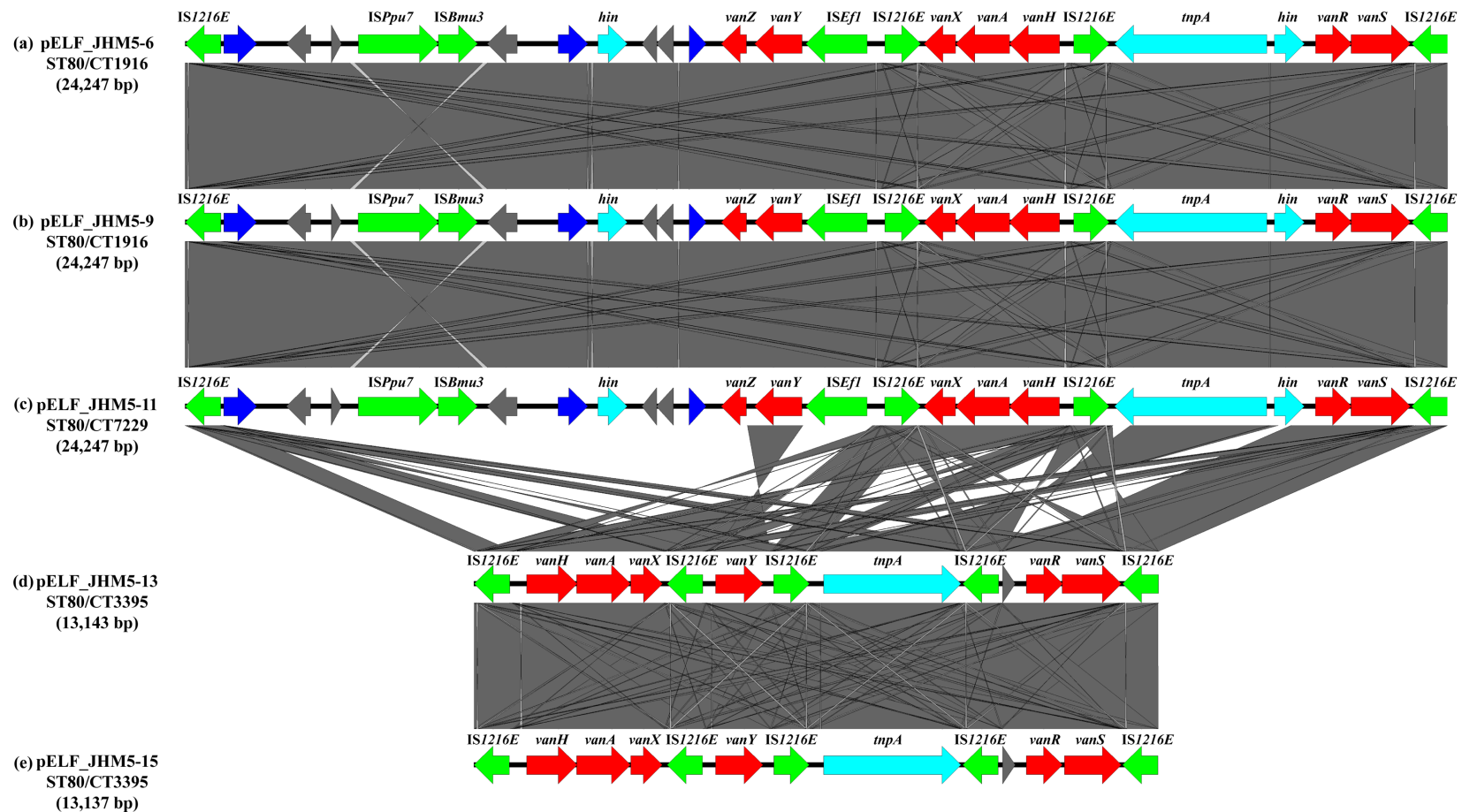
1 Kbp



Supplemental Figure S9. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of four VREfm isolates belonging to two complex types (CTs) from patient P4. The figure shows an alignment of the *vanA* region encoded by four plasmids from separate VREfm isolates belonging to two CTs as follows: (a) linear plasmid pELF_JHM4-2 (120,782 bp) from isolate JHM4-2 (ST80/CT2933, *vanA* region 20,977 bp), (b) linear plasmid pELF_JHM4-4 (119,306 bp) from isolate JHM4-4 (ST80/CT6738, *vanA* region 20,977 bp), (c) linear plasmid pELF_JHM4-14 (113,355 bp) from isolate JHM4-14 (ST80/CT6738, *vanA* region 20,977 bp) and (d) linear plasmid pELF_JHM4-15 (119,306 bp) from isolate JHM4-15 (ST80/CT2933, *vanA* region 20,977 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Table S7. A scale bar is included beneath the figure.

Reference:

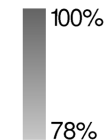
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Average Nucleotide Identity (*vanA* transposon)

	pELF_JHM5-6	pELF_JHM5-9	pELF_JHM5-11	pELF_JHM5-13	pELF_JHM5-15
pELF_JHM5-6					
pELF_JHM5-9	99.99				
pELF_JHM5-11	99.99	99.99			
pELF_JHM5-13	98.27	98.18	98.27		
pELF_JHM5-15	98.27	98.18	98.27	100	

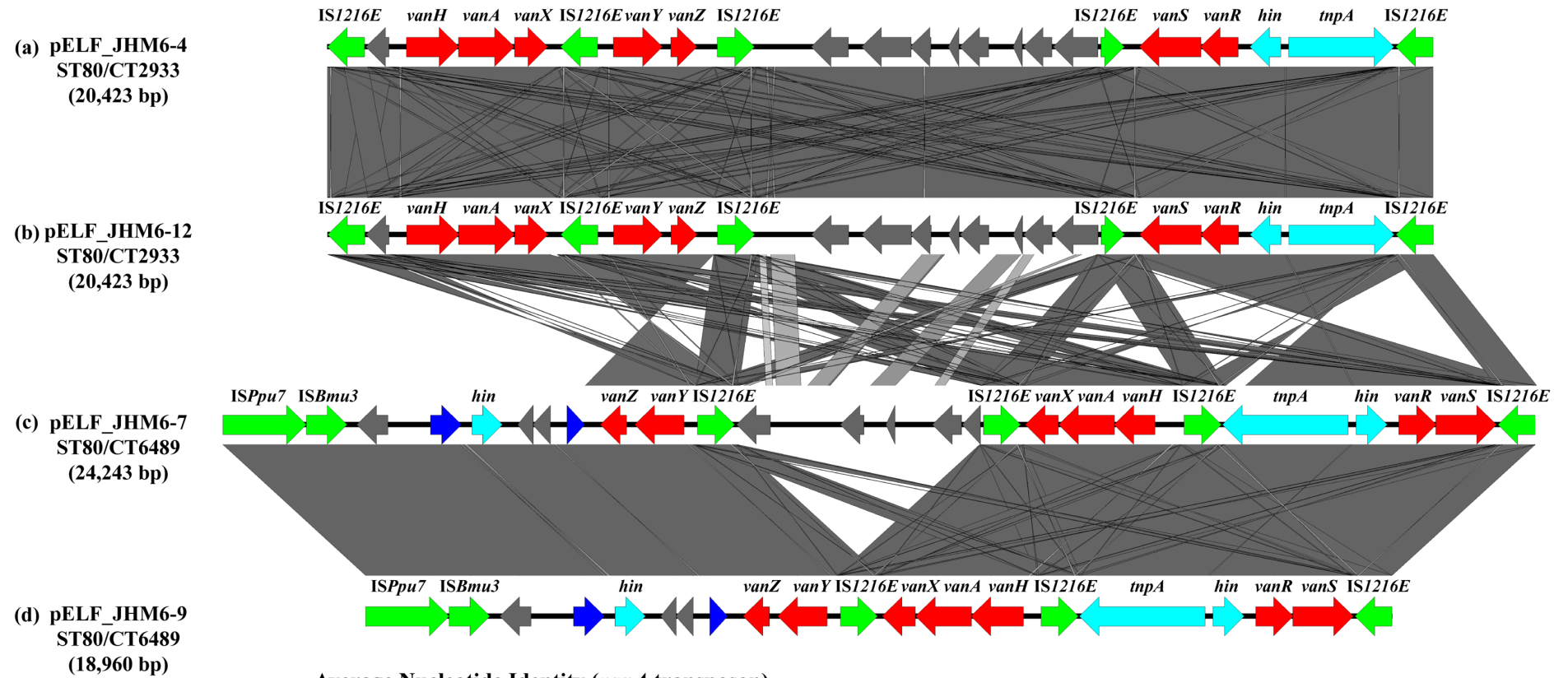
1 Kbp



Supplemental Figure S10. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of five VREfm isolates belonging to three complex types (CTs) from patient P5. The figure shows an alignment of the *vanA* region encoded by five plasmids from separate VREfm isolates belonging to three CTs as follows: (a) linear plasmid pELF_JHM5-6 (130,084 bp) from isolate JHM5-6 (ST80/CT1916, *vanA* region 24,247 bp), (b) linear plasmid pELF_JHM5-9 (130,077 bp) from isolate JHM5-9 (ST80/CT1916, *vanA* region 24,247 bp), (c) linear plasmid pELF_JHM5-11 (130,079 bp) from isolate JHM5-11 (ST80/CT7229, *vanA* region 24,247 bp), (d) linear plasmid pELF_JHM5-13 (105,602 bp) from isolate JHM5-13 (ST80/CT3395, *vanA* region 13,143 bp) and (e) plasmid pELF_JHM5-15 (105,609 bp) (ST80/CT3395, *vanA* region 13,137 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Tables S6 and S7 Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Table S6. A scale bar is included beneath the figure.

Reference:

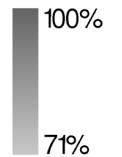
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Average Nucleotide Identity (*vanA* transposon)

	pELF_JHM6-4	pELF_JHM6-12	pELF_JHM6-7	pELF_JHM6-9
pELF_JHM6-4				
pELF_JHM6-12	99.96			
pELF_JHM6-7	97.27	97.31		
pELF_JHM6-9	99.04	99.04	99.99	

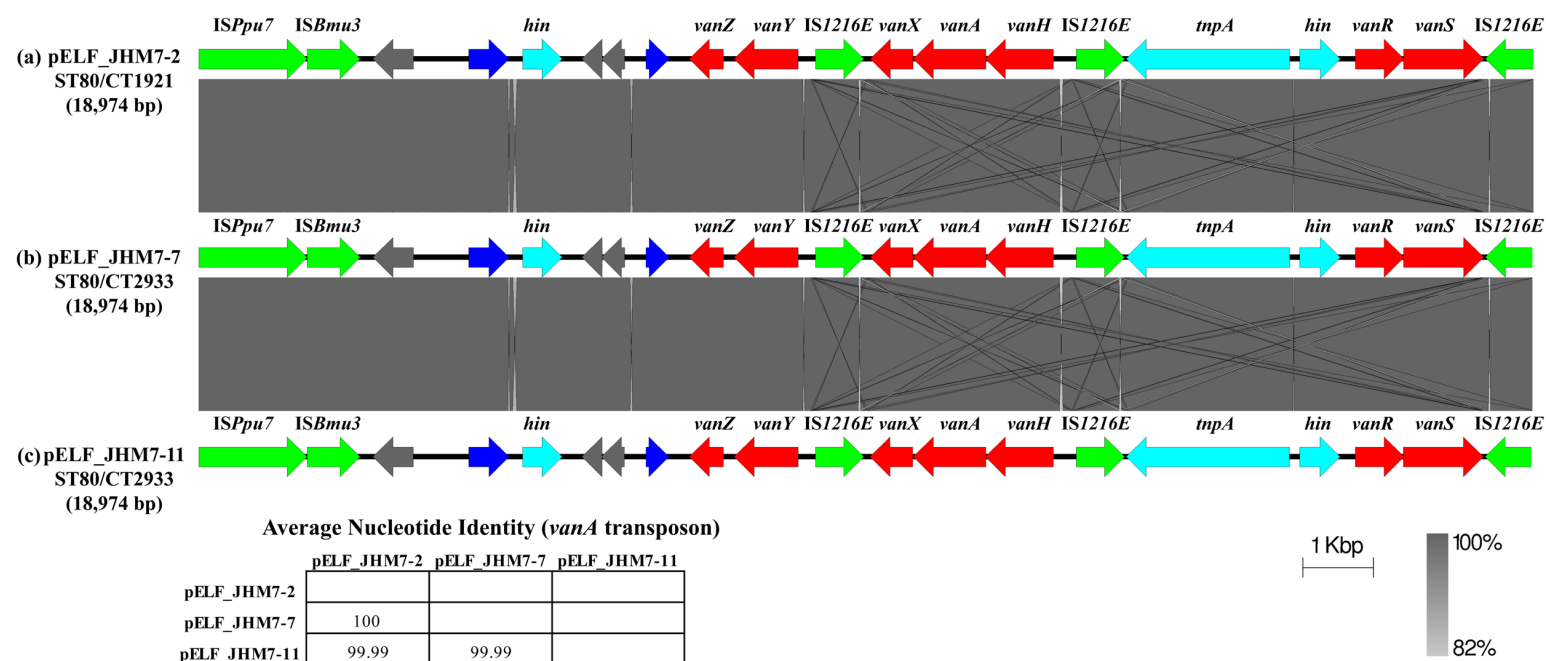
1 Kbp



Supplemental Figure S11. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of four VREfm isolates belonging to two complex types (CTs) from patient P6. The figure shows an alignment of the *vanA* region encoded by four plasmids from separate VREfm isolates belonging to two CTs as follows: (a) linear plasmid pELF_JHM6-4 (107,839 bp) from isolate JHM6-4 (ST80/CT2933, *vanA* region 20,423 bp), (b) linear plasmid pELF_JHM6-12 (107,837 bp) from isolate JHM6-12 (ST80/CT2933, *vanA* region 20,423 bp), (c) linear plasmid pELF_JHM6-7 (122,006 bp) from isolate JHM6-7 (ST80/CT6489, *vanA* region 24,247 bp) and (d) linear plasmid pELF_JHM6-9 (116,713 bp) from isolate JHM6-9 (ST80/CT6489, *vanA* region 18,960 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables (S6 & S7). Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Table S7. A scale bar is included beneath the figure.

Reference:

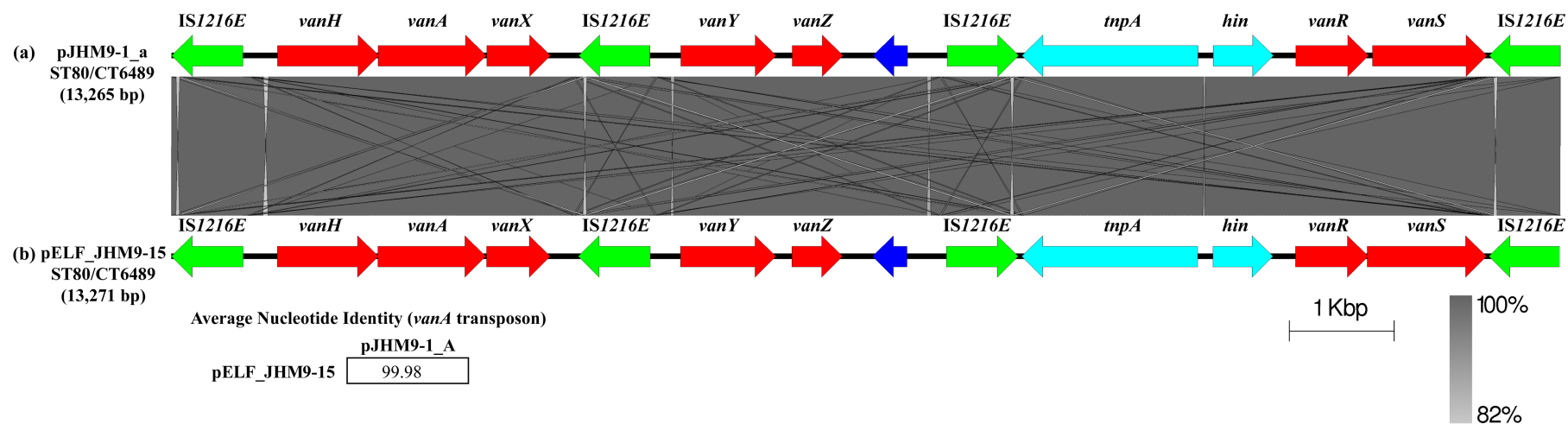
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S12. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of three VREfm isolates belonging to two complex types (CTs) from patient P7. The figure shows an alignment of the *vanA* region encoded by three plasmids from separate VREfm isolates belonging to two CTs as follows: (a) linear plasmid ELF_JHM7-2 (116,676 bp) from isolate JHM7-2 (ST80/CT1921, *vanA* region 18,974 bp), (b) linear plasmid pELF_JHM7-7 (116,675 bp) from isolate JHM7-7 (ST80/CT2933, *vanA* region 18,974 bp) and (c) linear plasmid pELF_JHM7-11 (116,674 bp) from isolate JHM7-11 (ST80/CT2933, *vanA* region 18,974 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Table S6. A scale bar is included beneath the figure.

Reference:

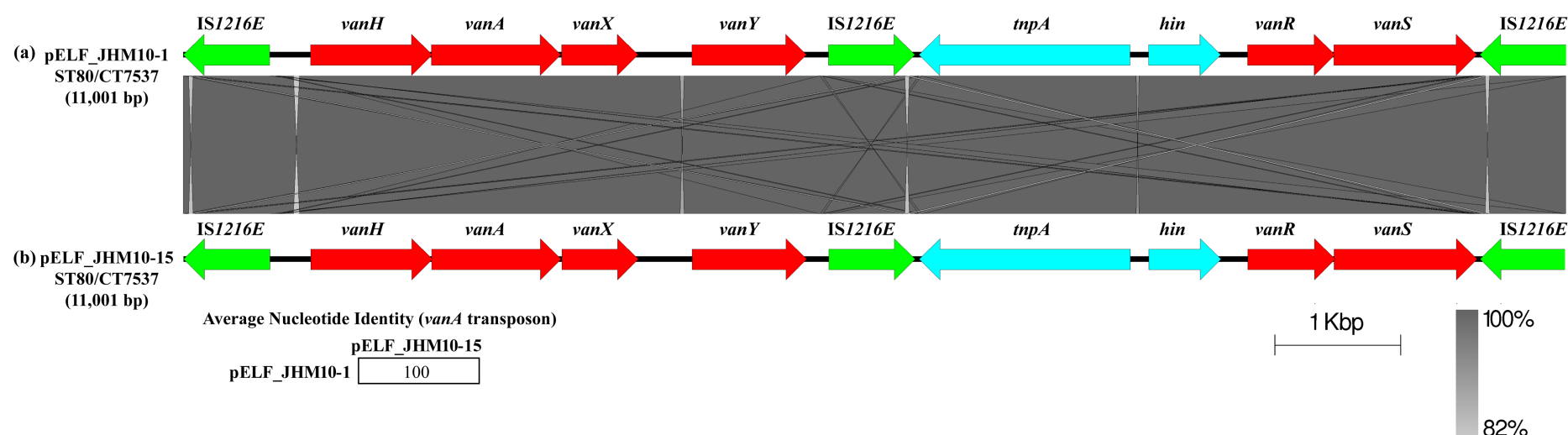
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S13. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of two VREfm isolates belonging to one complex type (CT) CT6489 from patient P9. The figure shows an alignment of the *vanA* region encoded by two plasmids from two separate VREfm isolates belonging to one CT as follows: (a) pJHM9-1_a* from isolate JHM9-1 (ST80/CT6489, *vanA* region 13,265 bp) and (b) linear plasmid pELF_JHM9-15 (108,808 bp) from isolate JHM9-15 (ST80/CT6489, *vanA* region 13,271 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). A scale bar is included beneath the figure. *This plasmid was not fully resolved by hybrid assembly. Its sequences were spread across 25 contigs with a total length of 94,436 bp (Table S6).

References.

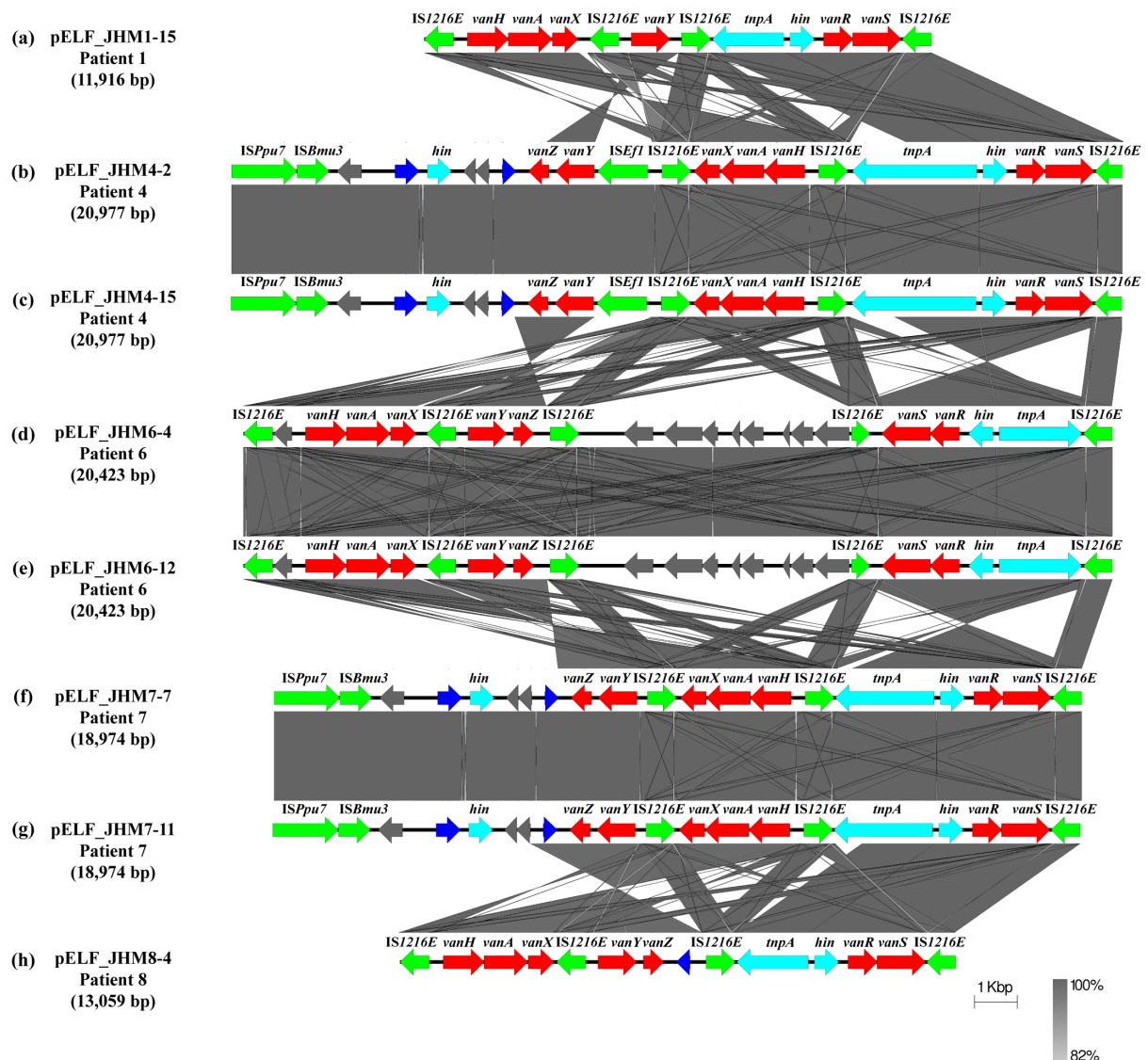
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S14. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of two VREfm isolates belonging to one complex type (CT) CT7537 from patient P10. The figure shows an alignment of the *vanA* region encoded by two plasmids from two separate VREfm isolates belonging to one CT as follows: (a) linear plasmid pELF_JHM10-1 (143,742 bp) from isolate JHM10-1 (ST80/CT7537, *vanA* region 11,001 bp) and (b) linear plasmid pELF_JHM10-15 (143,742 bp) from isolate JHM10-15 (ST80/CT7537, *vanA* region 11,001 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). Details of plasmids identified in isolates investigated are provided in Supplemental Table A6. A scale bar is included beneath the figure.

References.

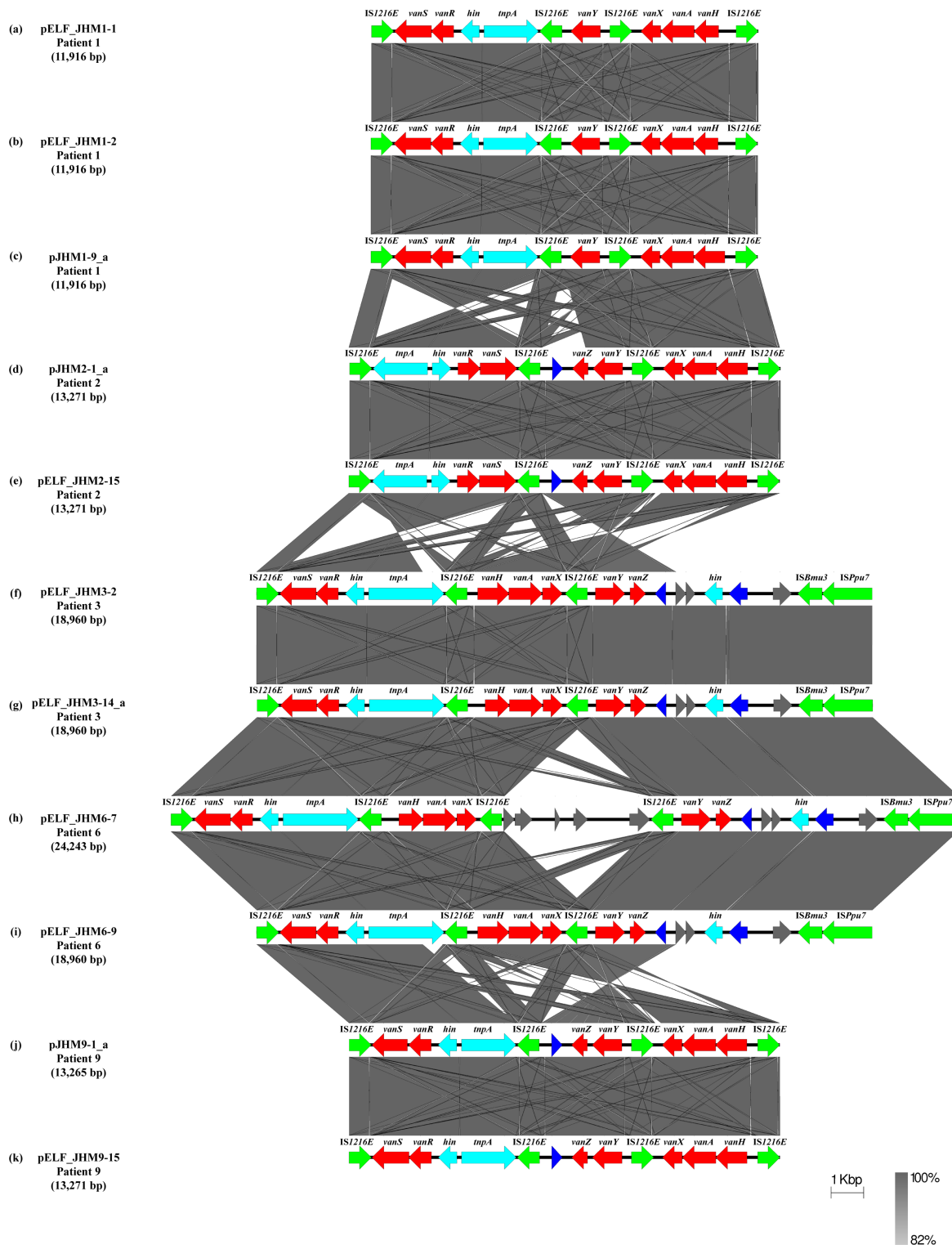
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S15. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of VREfm isolates belonging to complex type (CT) CT2933 from five patients. The figure shows an alignment of the *vanA* region encoded by eight plasmids from CT2933 isolates as follows: (a) pELF_JHM1-15 (ST80/CT2933, *vanA* region 11,916 bp), (b) pELF_JHM4-2 (ST80/CT2933, *vanA* region 20,977 bp), (c) pELF_JHM4-15 (ST80/CT2933, *vanA* region 20,977 bp), (d) pELF_JHM6-4 (ST80/CT2933, *vanA* region 20,423 bp), (e) pELF_JHM6-12 (ST80/CT2933, *vanA* region 20,423 bp), (f) pELF_JHM7-7 (ST80/CT2933, *vanA* region 18,974 bp), (g) pELF_JHM7-11 (ST80/CT2933, *vanA* region 18,974 bp) and (h) pELF_JHM8-4 (ST80/CT2933, *vanA* region 13,059 bp), generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). A scale bar is included beneath the figure.

Reference.

1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.

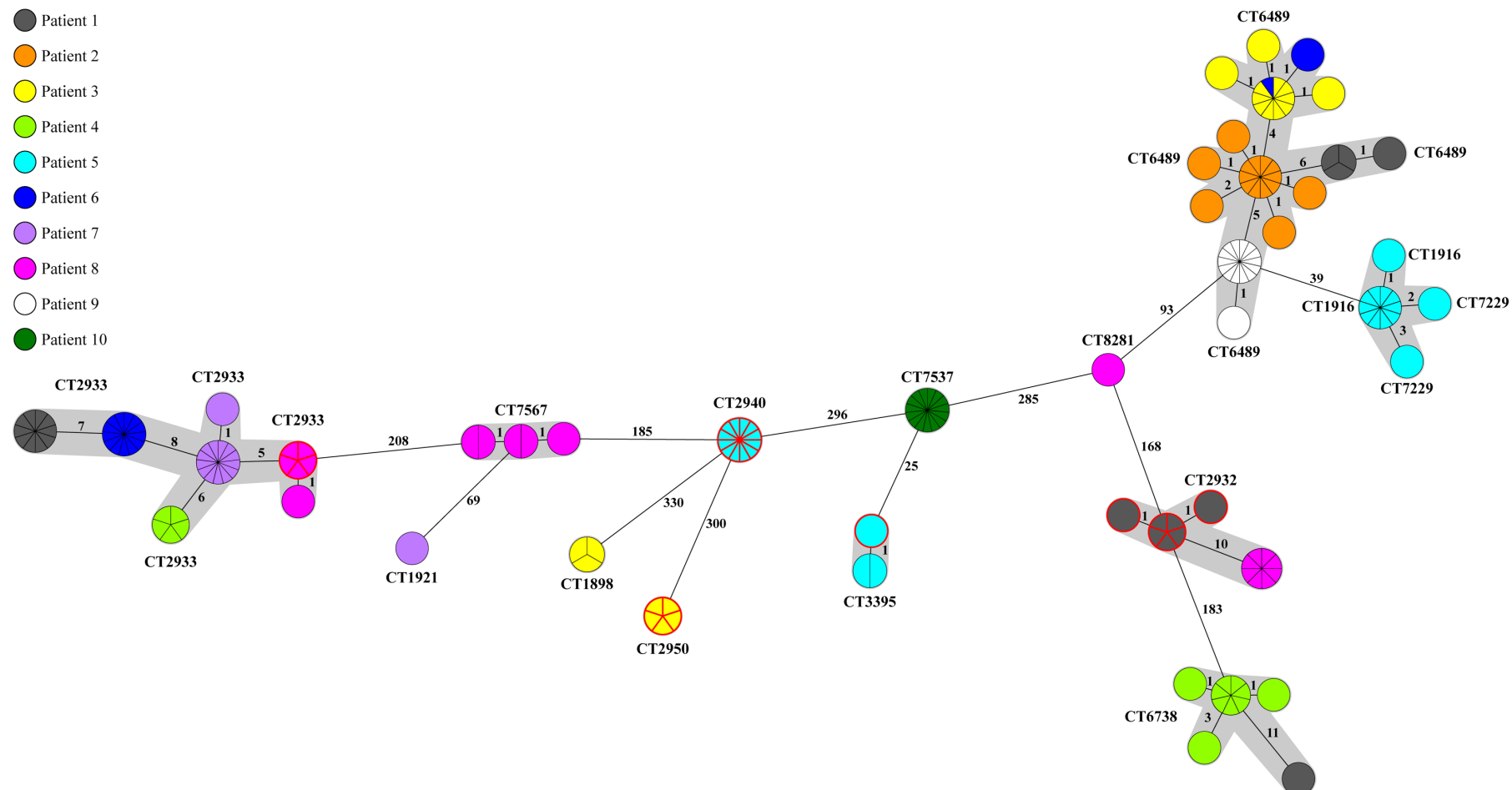


Supplemental Figure S16. Comparative genetic organisation of *vanA* transposon regions closed by hybrid assembly of VREfm isolates belonging to complex type (CT) CT6489 from five patients. The figure shows an alignment of the *vanA* region encoded by eleven plasmids from CT6489 as follows: (a) pELF_JHM1-1 (ST80/CT6489, *vanA* region 11,916 bp), (b) pELF_JHM1-2 (ST80/CT6489, *vanA* region 11,916 bp), (c) pJHM1-9_a (ST80/CT6489, *vanA* region 11,916 bp), (d) pJHM2-1_a (ST80/CT6489, *vanA* region 13,271 bp), (e) pELF_JHM2-15 (ST80/CT6489, *vanA*

region 13,271 bp), (f) pELF_JHM3-2 (ST80/CT6489, *vanA* region 18,960 bp), (g) pELF_JHM3-14_a (ST80/CT6489, *vanA* region 18,960 bp), (h) pELF_JHM6-7 (ST80/CT6489, *vanA* region 24,247 bp), (i) pELF_JHM6-9 (ST80/CT6489, *vanA* region 18,960 bp), (j) pJHM9-1_a (ST80/CT6489, *vanA* region 13,265 bp) and (k) pELF_JHM9-15 (ST80/CT6489, *vanA* region 13,271 bp) generated using Easyfig¹. Coding sequences of interest and their orientation are represented by arrows as follows: red, vancomycin resistance genes; green, insertion sequences; blue, proteins with a known coding function; cyan, transposase and invertase genes; grey, hypothetical proteins. Full details of plasmid profiles and SKA-SNP matrices are provided in Supplemental Tables S6 & S7. Regions between alignments with shared homology are represented by grey connecting blocks and/or lines with the colour gradient indicating percentage nucleotide identity (82-100%). A scale bar is included beneath the figure.

References.

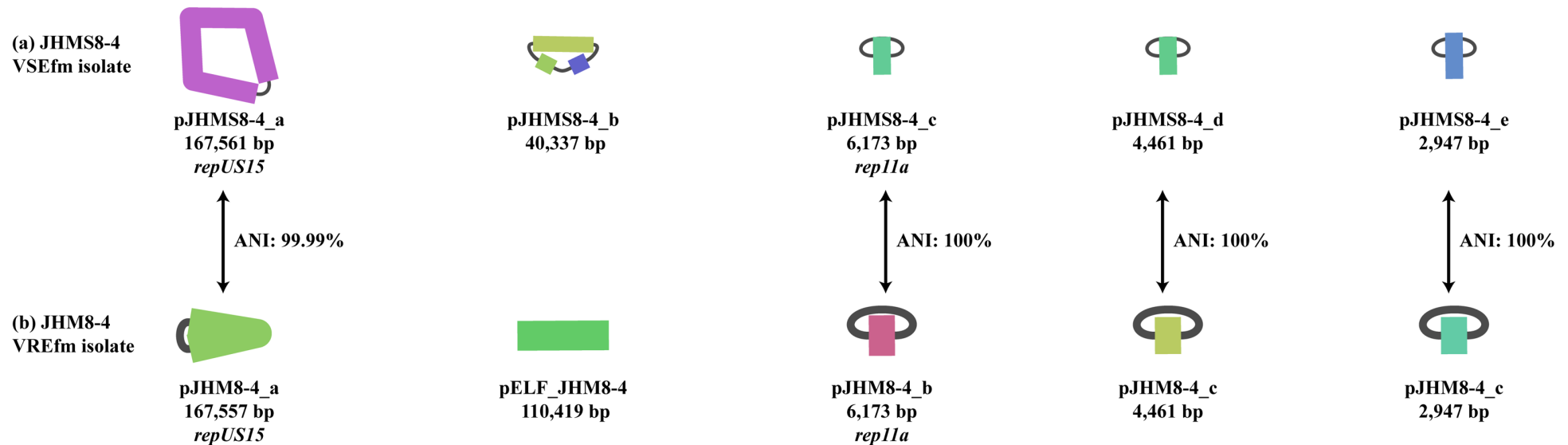
1. Sullivan MJ, Petty NK, Beatson SA. Easyfig: a genome comparison visualizer. *Bioinformatics* 2011; **27**: 1009–10. Available at: <https://doi.org/10.1093/bioinformatics/btr039>.



Supplemental Figure S17. Minimum spanning tree (MST) based on core genome multi-locus sequence typing (cgMLST) analysis of 1,423 loci for 150 clade A1 VREfm isolates recovered from a single rectal screening swab from each of 10 separate patients and 30 VSEfm isolates recovered from a single rectal swab from each of four separate patients. Node colours represent the patient number (1-10) from which the isolate was recovered as indicated in the legend. Partitions within the nodes represent the presence of ≥ 2 isolates. Branch numbers indicate cgMLST allelic differences between neighbouring isolates. cgMLST segregated isolates into 11 clusters and two singletons using the previously recommended threshold of ≤ 20 allelic differences¹. VSEfm isolates from rectal screens are indicated with a red outer node ring and red partitions within the node ($n=30$).

Reference.

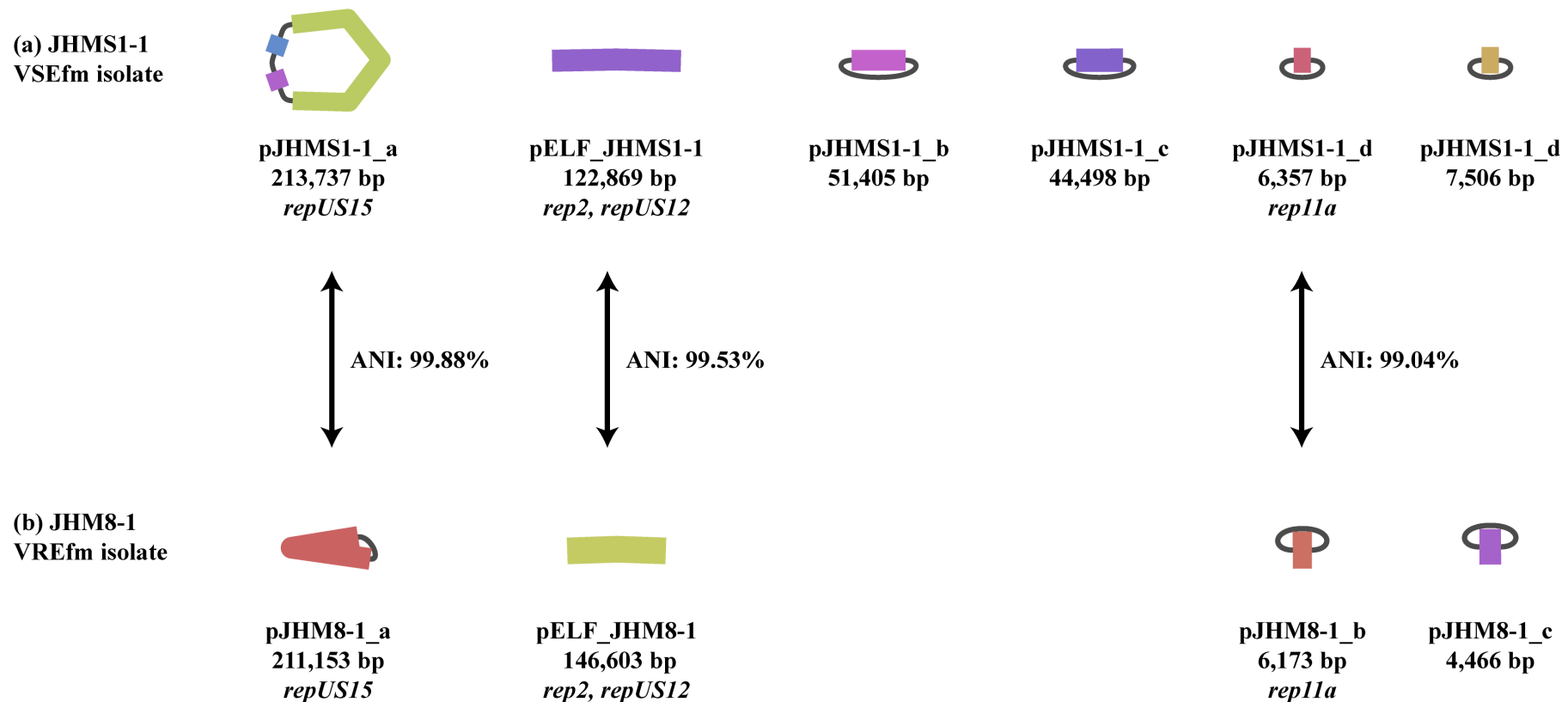
1. de Been M, Pinholt M, Top J, *et al.* Core genome multilocus sequence typing scheme for high-resolution typing of *Enterococcus faecium*. *J Clin Microbiol* 2015; **53**: 3788–97. Available at: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4652124/>.



Supplemental Figure S18. Visual representation of the complete plasmidome assembly of VSEfm isolate JHMS8-4 and VREfm isolate JHM8-4 from complex type (CT) 2933 generated using Bandage.¹ (a) VSEfm isolate JHMS8-4 recovered from a rectal screening swab from patient P8 harbouring five distinct plasmids ranging in size from 2,947-167,561 bp and (b) VREfm isolate JHM8-4 recovered from a rectal screening swab from patient P8 harbouring five distinct plasmids ranging in size from 2,947-167,557 bp. Plasmids are represented as individual circular (pJH) or linear nodes (pELF) depending on topology. Plasmid names, length and associated replicon type are indicated. The average nucleotide identity (ANI) between equivalent plasmids is indicated by a percentage beside the arrows.

Reference:

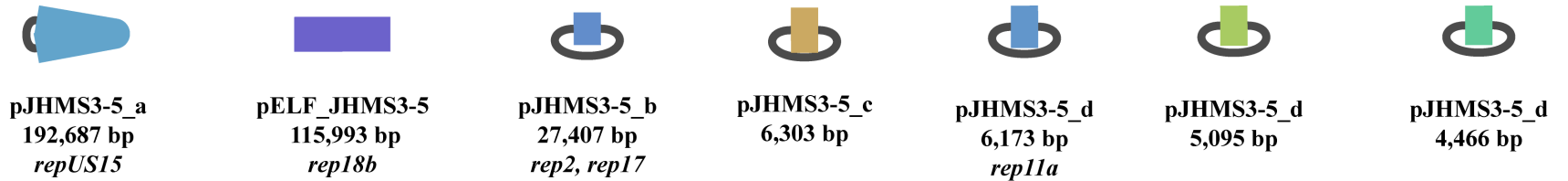
1. Wick, R.R., Schultz, M.B., *et al.* Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 2015; **31**:3350–52. Available at: <https://doi.org/10.1093/bioinformatics/btv383>



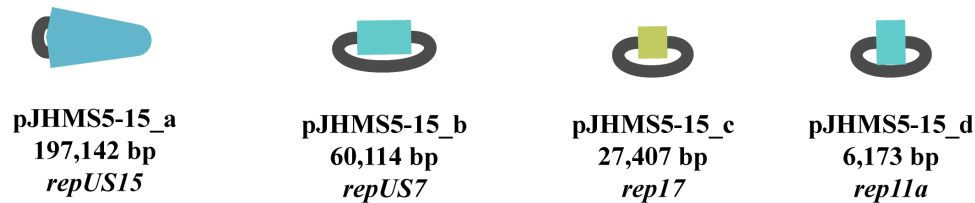
Supplemental Figure S19. Visual representation of the complete plasmidome assembly of VSEfm isolate JHMS1-1 and VREfm isolate JHM8-1 from complex type (CT) 2932 generated using Bandage.¹ (a) VSEfm isolate JHMS1-4 recovered from a rectal screening swab from patient P1 harbouring six distinct plasmids ranging in size from 6,357-213,737 bp and (b) VREfm isolate JHM8-1 recovered from a rectal screening swab from patient P8 harbouring four distinct plasmids ranging in size from 4,466-211,153 bp. Plasmids are represented as individual circular (pJH) or linear (pELF) nodes depending on topology. Plasmid names, length and associated replicon type are indicated. The average nucleotide identity (ANI) between equivalent plasmids is indicated by a percentage beside the arrows.

Reference: 1. Wick, R.R., Schultz, M.B., *et al.* Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 2015; **31**:3350–52. Available at: <https://doi.org/10.1093/bioinformatics/btv383>

(a) JHMS3-5
VSEfm isolate
ST80/CT2950

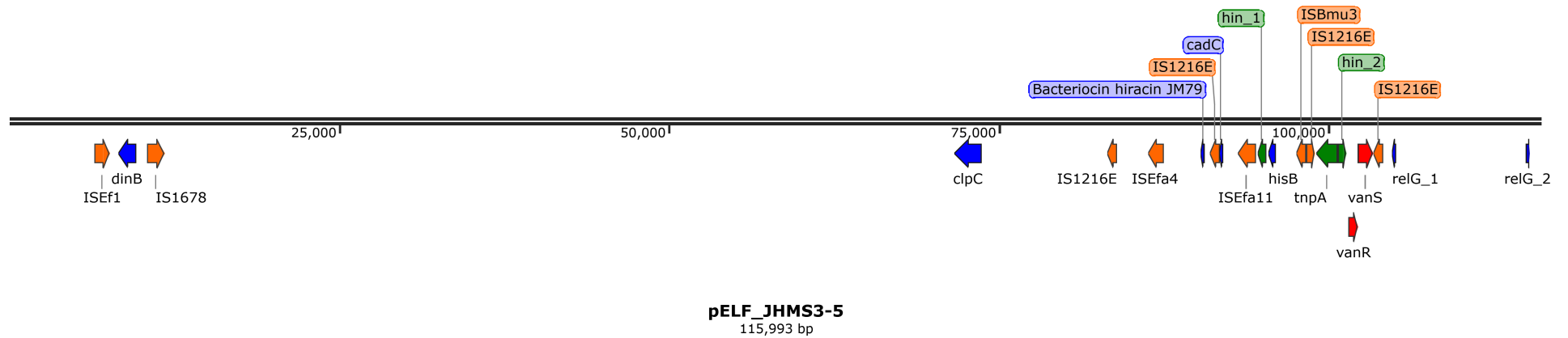


(b) JHMS5-15
VSEfm isolate
ST80/CT2940



Supplemental Figure S20. Visual representation of the complete plasmidome assembly of VSEfm isolates JHMS3-4 (CT2950) and JHMS5-15 (CT2940) generated using Bandage.¹ (a) VSEfm isolate JHMS3-5 (ST80/CT2950) recovered from a rectal screening swab from patient P3 harbouring seven distinct plasmids ranging in size from 4,466-192,687 bp and (b) VSEfm isolate JHM8-1 (ST80/CT2940) recovered from a rectal screening swab from patient P5 harbouring four distinct plasmids ranging in size from 6,173-197,142 bp. Plasmids are represented as individual circular (pJH) or linear (pELF) nodes depending on topology. Plasmid names, length and associated replicon type are indicated.

Reference: 1. Wick, R.R., Schultz, M.B., *et al.* Bandage: interactive visualization of *de novo* genome assemblies. *Bioinformatics* 2015; **31**:3350–52. Available at: <https://doi.org/10.1093/bioinformatics/btv383>



Supplemental Figure S21. Schematic diagram of the structural organisation of plasmid pELF_JHMS3-5 from the VSEfm rectal isolate JHMS3-5 (ST80/CT2950) from patient P3 encoding a *vanA*-remnant resolved by hybrid assembly of paired-end Illumina MiSeq short reads with Oxford Nanopore Technologies long reads. Genes of interest and their orientation are represented by arrows as follows: red indicates antibiotic resistance genes, orange indicates insertion sequences, green indicates transposase and invertase genes, blue indicates genes encoding known proteins. Due to spacing limitations, products annotated by RAST, Prokka and/or EGG-NOG Mapper as ‘hypothetical protein’, ‘mobile element protein’, ‘phage protein’ and ‘phage-related protein’ are not represented in this diagram.