JOURNAL OF MEDICAL MICROBIOLOGY

RESEARCH ARTICLE

Leigh *et al.*, *Journal of Medical Microbiology* 2022;71:001590

DOI 10.1099/jmm.0.001590

MICROBIOLOGY SOCIETY

OPEN ACCESS

Comparative genomics and pangenomics of vancomycin-resistant and susceptible *Enterococcus faecium* from Irish hospitals

Robert J. Leigh[1,*], Chloe McKenna[1], Robert McWade[2], Breda Lynch[2] and Fiona Walsh[1]

Abstract

Introduction. *Enterococcus faecium* has emerged as an important nosocomial pathogen, which is increasingly difficult to treat due to the genetic acquisition of vancomycin resistance. Ireland has a recalcitrant vancomycin-resistant bloodstream infection rate compared to other developed countries.

Hypothesis/Gap statement. Vancomycin resistance rates persist amongst *E. faecium* isolates from Irish hospitals. The evolutionary genomics governing these trends have not been fully elucidated.

Methodology. A set of 28 vancomycin-resistant isolates was sequenced to construct a dataset alongside 61 other publicly available Irish genomes. This dataset was extensively analysed using *in silico* methodologies (comparative genomics, pangenomics, phylogenetics, genotypics and comparative functional analyses) to uncover distinct evolutionary, coevolutionary and clinically relevant population trends.

Results. These results suggest that a stable (in terms of genome size, GC% and number of genes), yet genetically diverse population (in terms of gene content) of *E. faecium* persists in Ireland with acquired resistance arising via plasmid acquisition (*vanA*) or, to a lesser extent, chromosomal recombination (*vanB*). Population analysis revealed five clusters with one cluster partitioned into four clades which transcend isolation dates. Pangenomic and recombination analyses revealed an open (whole genome and chromosomal specific) pangenome illustrating a rampant evolutionary pattern. Comparative resistomics and virulomics uncovered distinct chromosomal and mobilomal propensity for multidrug resistance, widespread chromosomal point-mutation-mediated resistance and chromosomally harboured arsenals of virulence factors. Interestingly, a potential difference in biofilm formation strategies was highlighted by coevolutionary analysis, suggesting differential biofilm genotypes between *vanA* and *vanB* isolates.

Conclusions. These results highlight the evolutionary history of Irish *E. faecium* isolates and may provide insight into underlying infection dynamics in a clinical setting. Due to the apparent ease of vancomycin resistance acquisition over time, susceptible *E. faecium* should be concurrently reduced in Irish hospitals to mitigate potential resistant infections.

INTRODUCTION

The genus *Enterococcus* (*Baccilota* (*Firmicutes*); *Bacilli*; *Lactobacilli*; *Enterococcaceae*) are commonly observed in diverse biomes such as soil, surface water nd wastewater, and as commensal inhabitants of the higher chordate (inclusive of human) gastrointestinal (tract, vaginal tract and epidermis [1, 2]. Two species, *Enterococcus faecium* and *Enterococcus faecalis*, are

Received 30 April 2022; Accepted 26 July 2022; Published 27 October 2022

Author affiliations: [1]Department of Biology, Maynooth University, Mariavilla, Maynooth, Co. Kildare, Ireland; [2]Department of Microbiology, Mater Misericordiae University Hospital, Eccles St., Dublin 7, D07 R2WY, Ireland.

***Correspondence:** Robert J. Leigh, rob.leigh@mu.ie

Keywords: CC17; *Enterococcus*; evolution; Ireland; vancomycin resistance.

Abbreviations: AMR, antimicrobial resistance; CC, clonal cluster; cgMLST, core genome multi-locus sequence type; ECDC, European Centre for Disease Control; MLST, multi-locus sequence type; ST, sequence type; VF, virulence factor; VRE, vancomycin-resistant *E. faecium*; VSE, vancomycin susceptible *E. faecium*.

Genome sequences and assembled genomes are deposited at the National Center for Biotechnology Information Sequence Read Archive and GenBank under BioProject PRJNA799050.

Four supplementary figures and 21 supplementary tables are available with the online version of this article.

aetiological agents of a cohort of moderate to severe conditions when migrated from the gastrointestinal tract, especially in the immunocompromised or convalescing host [3, 4]. Both *E. faecium* and *E. faecalis* migration can lead to caries, cellulitis, cholecystitis, cystitis, endocarditis, endodontits, periodontitis, peri-implantitis, postoperational peritonitis, sepsis and neonatal meningitis [5–8]. The clinical significance of these species was amplified as antimicrobial resistance (AMR) began to evolve and disseminate throughout diverse environments [2, 9].

Vancomycin is the drug-of-last-resort for recalcitrant infections with diverse resistance profiles [10], inclusive of the prominent nosocomial pathogen methicillin-resistant *Staphylococcus aureus* (MRSA) and *E. faecium* [11, 12]. In recent years, vancomycin resistance has been observed in a number of human pathogens, often resulting in clinical complication and extended hospital stay [6, 13–16]. Ireland has one of the highest vancomycin-resistant *E. faecium* (VRE) rates in Europe where 38.4% of bloodstream infection isolates display resistance (ECDC: https://atlas.ecdc.europa.eu/). In particular, nosocomial VRE infections are of concern due to their rapid dissemination throughout this environment [4, 15, 17]. Vancomycin resistance genes have been detected both on the chromosome and on mobile elements (mobilome) and often accompany other resistance genotypes [6, 18–20]. Commonly observed *vanA* mobilome co-resistance phenotypes are observed towards tetracycline, erythromycin and aminoglycosides [21–24]. Additionally, *E. faecium* almost ubiquitously displays chromosomal-mediated resistance to aminoglycosides [via *AAC*(6')-*Ii*], macrolides (via *efmA*, *efrAB* and *msrC*), and fluoroquinolones (via *gyrA* and *parC* mutations or upregulation of efflux), rifamycin (via *efrAB*), and clindamycin, quinupristin-dalfopristin and dalfopristin (via *lsaA*) [6, 9, 24–26].

In many countries, including the Republic of Ireland, vancomycin resistance is prominently derived from plasmid-mediated *vanA*, and less commonly from chromosomalmediated *vanB* [2, 27, 28]. Both *vanA* and *vanB* are D-Ala–D-Ala ligases (EC: 6.3.2.4) and confer resistance by constructing D-Ala–D-Lac as an alternative substrate during peptidoglycan synthesis, thus reducing vancomycin–peptidoglycan binding affinity [29–31]. Aside from their associated genomic location (mobilomal vs. chromosomal), *vanA* also confers resistance to teicoplanin (another glycopeptide antibiotic) whereas *vanB* has not been observed to induce this effect [32–34].

Metal and biocide resistance is of growing concern in a plethora of environments, and especially in clinical settings [35]. Correlations have been reported between metal resistance and drug resistance in pathogenic bacteria, suggesting a potential co-evolutionary pressure on both mechanisms [36–38]. A range of metals are employed in health care for their intrinsic antimicrobial properties, for example silver-embedded plasters and copper-plated door handles [39–41]. Previous studies have reported chromosomal-mediated resistance to copper, silver, selenium and hydrogen peroxide in *E. faecium*, suggesting a growing resistance to passive protection strategies [42–44].

Further to their resistance mechanisms, *Enterococcus* species employ a small cohort of effective virulence factors during pathogenesis, allowing for adhesion, biofilm formation, invasion, immunomodulation, and the synthesis of secreted toxins, enzymes and peptides (such as bacteriocins) into local environments to inhibit competition [9, 45, 46]. Biofilm formation is of critical importance in VRE infections, due to difficulty of clearance and reduced antimicrobial penetration rates [42, 47].

When isolated from human samples, the average *E. faecium* genome contains 2765 ± 187 genes, yet the pangenome contained 12 457 when sampled from 161 genomes [26]. These results, and other pangenomic analyses (e.g. [48]), illustrate the genomic plasticity and evolutionary capacity of *E. faecium*. While a major proportion of this variance is attributed to mobile genetic elements [49], to our knowledge, no study has been conducted on chromosomal pangenomic variance, so the extent of

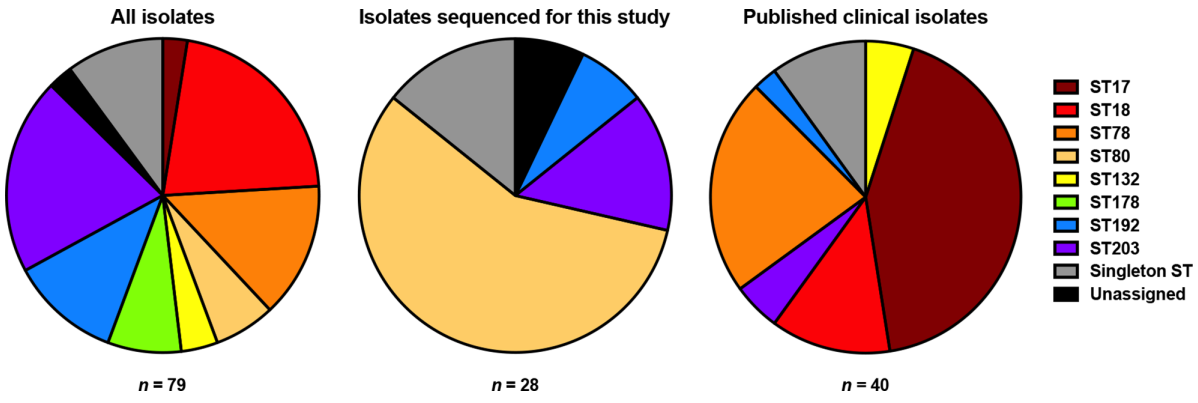

Fig. 1. Distribution of sequence types between different studies. Cork isolates (CC97; *n* = 11) all displayed identical sequence types (as all of these isolates are identical, they are not shown in a separate pie chart).

mobilome-facilitated variance has not been fully elucidated. The integration of phages (prophages) into bacterial chromosomes can introduce genetic novelty [50–52]. While a wide array of *Enterococcus* phages have been identified [53, 54], and parasite–host co-evolutionary trajectories have been explored [55], the phage impact on symbiotic co-evolutionary trajectory has yet, to our knowledge, to be explored.

The aims of this study were to compare the pan-genomes, mobilomes and chromosomes of the available genome sequences of vancomycin-resistant or susceptible *E. faecium* across the timeframe that VRE (analysed by the ECDC: 2002 to 2019) increased in prevalence from 11.1 to 38.4% in Ireland (available at https://atlas.ecdc.europa.eu/public/index.aspx?Dataset=27&HealthTopic=4). Specifically, this study provides insight into the pangenomics of *E. faecium* from three studies in the Republic of Ireland using a comparative genomic, resistomic and virulomic lens. It is now possible to derive intrapangenome correlations [56] so these approaches are explored to provide additional insight.

A recent study of the global dissemination of *E. faecium* identified that it has two main modes of genomic evolution: the acquisition and loss of genes, including AMR genes, through mobile genetic elements including plasmids, and homologous recombination of the core genome [49]. Unfortunately, there were no Irish isolates contained within this study.

METHODS

Microbiological analysis

Twenty-eight isolates identified as vancomycin-resistant *E. faecium* during 2018 and 2019 were collected by the Mater Misericordiae University Hospital (MMUH) in Dublin, Ireland. The sample metadata are described in Table S1 (available in the online version of this article). Antimicrobial susceptibility testing was performed at the MMUH according to EUCAST guidelines (https://www.eucast.org/ast_of_bacteria/) and subsequently at Maynooth University using disc testing according to the CLSI guidelines (https://clsi.org/). The isolates were investigated for resistance to the following antibiotics: ciprofloxacin, erythromycin, chloramphenicol, vancomycin, tetracycline, ampicillin and linezolid (Table S2).

DNA extractions and genome sequencing

DNA was extracted from each of the 28 isolates using the Macherey-Nagel Nucleospin microbial DNA isolation kit according to the manufacturer's instructions. The extracted DNA was quality assessed using NanoDrop and Qubit as per the sequencing centre (Novogene) guidelines. Extracted DNA was sequenced by Novogene using the 'Bacterial resequencing' service (using an Illumina NovaSeq 6000) with PE150 and Q30 ≥80%. This provided >100× coverage for each genome.

Genome assembly

The 28 samples sequenced for this study and all reads associated with Irish VRE genomes from the British Society for Antimicrobial Chemotherapy study PRJEB4344 (hereafter referred to as the 'BSAC' isolates [57]) were downloaded from the NCBI sequence read archive (SRA [58]). Each read pair (sample) was subjected to adapter removal and quality trimming using TrimGalore v.0.6.6 [59] using default settings. Adapter removal during the TrimGalore pipeline was powered by CutAdapt v.3.0 [60] and FastQC v.0.11.9 [61]. Each sample was assembled using Unicycler v.0.4.7[62] using default paired-end settings. Unicycler utilized SPAdes v.3.14.1 [63] to assemble reads and used Bowtie2 v.2.4.2 [64], Pilon v.1.23 [65], BLAST v.2.11 [66, 67] and samtools v.1.11 [68] to further complete the assembly. Each assembly was quality assessed using CheckM [69] using the *E. faecium* database (Table S3) and sequence typed using MLST v.2.19.0 [70, 71] using default settings (Table S4, Fig. 1). Where a sequence type (ST) could not be completely identified, the approximated alleles were used to approximate an ST using the 'search by locus combinations' option for *E. faecium* using PubMLST (www.pubmlst.org [72]). Each assembly was separated into 'chromosomes' and 'plasmids' (hereafter referred to as 'mobilomes' as complete plasmids were not always guaranteed and were treated collectively as an extrachromosomal entity) using Platon v.1.5.0 [73]. These partitioned assemblies were analysed alongside their concatenated 'whole genome' assemblies. Assemblies that had a reported completeness percentage ≤95% were retained for further analyses.

Additional genomes

A total of 11 *E. faecium* genome assemblies were available on NCBI assembly and attributed to Ireland in their respective metadata (from study PRJNA521309 [74]). As above, these assemblies were quality checked using CheckM, sequence typed using MLST, and separated into chromosomal and mobilomal components using Platon. These 11 isolates were all sampled in county Cork and are hereafter referred to as the 'Cork' isolates.

Plasmid containment analysis

As plasmids were contiguous, a containment analysis was used to determine the closest relatives of the mobilome. Each containment analysis was performed using the 'screen' algorithm in MASH v.2.2.2 [75] against PLSDB [76] with a *P*-value stringency cut-off of 0.1 and an identity stringency cut-off of 0.99 to replicate the parameters used during the construction of PLSDB (Table S5).

Genome annotation

Each assembly dataset (whole genome, chromosomal and mobilome) was annotated using Prokka v.1.14.6 [77] using default settings. Prokka used Prodigal v.2.6.3 [78] to predict gene sequences, Aragorn v.1.2.38 [79] to detect tRNA sequences, and Minced v.0.4.0 [80] to detect CRISPR sequences and SignalP v.4.0 [81] to detect signal peptides, and HMMER v.3.3.1 for protein similarity searching [82], BioPerl v.1.7.2 for file manipulation [83] and barrnap v.0.9 for rRNA profiling [84]. Each protein sequence in each assembly dataset was further annotated using InterProScan v.5.45–80.0 [85] using the '--appl PfamA' and '--goterms' to assign Pfam domains [86] and Gene Ontology terms [87] respectively. Finally, each assembly dataset was searched for secondary metabolite biosynthesis genes using a set of tools with their respective default settings (unless otherwise stated below): GECCO v.0.6.3 [88] and AntiSMASH v.5 [89], for transposable elements using MobileElementFinder v.1.0.3 [90], for antimicrobial resistance using ABRicate v.1.0.1 [91] with the associated CARD database [92] and PointFinder v.1 (using the *E. faecium* database [93]), for virulence factors using Abricate with the associated VFDB dataset [94], and for metal (and biocide) resistance using BacMet v.2.0 [95]. As BacMet is published as an amino acid dataset, it was first backtranslated to a representative nucleotide sequence (using translation table 11) using the 'backtranseq' function in EMBOSS v.6.6.0.0 [96] with codon usage tables for *E. faecium* (the subject of this study); *Enterococcus faecalis*, *Staphylococcus aureus*, *Klebsiella pneumoniae*, *Acinetobacter baumanii*, *Pseudomonas aeruginosa*, *Enterobacter* spp. (ESKAPE pathogens); *Escherichia coli* (model organism), *Treponema pallidium*, *Neisseria gonorrhoeae, Chlamydia trachomatis* (prevalent bacterial sexually transmitted infection pathogens), *Clostridium botulinum, Campylobacter jejuni, Listeria monocytogenes* and *Vibrio parahaemolyticus* (prominent food-bourne pathogens), *Helicobacter pylori* (common gastric pathogen) and *Clostridiodes difficle, Legionella pneumophilia* and *Mycobacterium tuberculosis* (common nosocomial infection pathogenss). Codon usage tables were obtained from the Kazusa genome research institute (https://www.kazusa.or.jp/en/). To mitigate false negatives, ABRicate was run with a 50% minimum percentage identity stringency score (as opposed to the default 80%) to allow for the detection of full-length homologues that may have been otherwise undetected (collated results for AntiSMASH and ABRicate are given in Tables S6–S10).

Genome characteristic statistical analysis

The sum of coding genes, genome size (Mbp), genome density (the mean number of genes per Mbp) and guanine–cytosine content (GC%) were calculated for the chromosomal and mobilomal datasets for each sample (Table S11, Fig. 2). Summary statistics (mean, median, standard deviation and variance) were computed for each data series (Table S12). Two-tailed Welch's *t*-tests (H_0:$\mu_a=\mu_b$; H_A:$\mu_a\neq\mu_b$ [97, 98]) were used to compare genome density and GC% between chromosomal and mobilomal datasets, a Bonferroni-Dunn correction ($P_{BD}=P\times n_{comparisons}$; $n_{comparisons}=2$ [99, 100]) was used to control Type-I errors and instances where $P_{BD}\leq 0.005$ were considered statistically significant. $P\leq 0.005$ will be used to determine significance in all pairwise test comparisons to control for potential Type I and Type II errors [101]. Summary statistics were calculated for gene length in each genomic subset (chromosomal and mobilomal), distributions were assessed for normality using a Kolmogorov–Smirnoff test (H_0:$\sim X=N(\mu,\sigma)$; H_A:$\sim X\neq N(\mu,\sigma)$ [102, 103]) and equivariance using a Levene's test (H_0:$\sigma^2_a=\sigma^2_b$; H_A:$\sigma^2_a\neq\sigma^2_b$; [104]. All samples were observed to be non-normally distributed and 26 of 79 were observed to have unequal variances (Levene's test $P\leq 0.05$). Considering these trends, a Brunner–Munzel test (H_0:$B=0.5$; H_A:$B\neq 0.5$) was used to determine whether chromosomes were statistically more likely to have longer genes than mobilomes. A Bonferroni–Dunn correction was applied ($n_{comparisons}=79$) and instances where $P_{BD}\leq 0.005$ were considered statistically significant (Table S12).

Genome relatedness

Chromosomal sequences from each isolate were all-vs.-all compared using MASH v.2.2.2 [75], and instances where the reported distance (D) ≤0.05 with a P-value ≤0.05 extracted as an edge list and visualized as a network with Gephi v.0.9.2 [105] using the Fruchterman-Reingold algorithm [106] (Fig. 3). This procedure was repeated for mobilomal sequences (Fig. 4).

Pangenome analysis

A pangenome for the whole genome and chromosomal datasets (for all isolates, and individually for the 'Cork' isolates, for the 'BSAC' isolates and for the isolates sequenced for this study) was produced using Roary v.3.1.3 [107] with the '-e' flag to align all gene clusters using PRANK v.170427 [108], and the '-z' flag to keep intermediate files (retained to produce the robust phylogeny below). To assess the effect of the mobilome on the genome diversity and complexity, each pangenomic category [n_{genes}; 'core pangenome' ($99\%\leq n_{samples}\leq 100\%$), 'soft core pangenome' ($95\%\leq n_{samples}<99\%$), 'shell pangenome' ($15\%\leq n_{samples}<95\%$) and 'cloud pangenome' ($n_{samples}<15\%$)] in the whole genome-generated pangenome was compared to the chromosomal pangenome using a two-tailed Fisher's exact test (H_0:$\rho=\pi$; H_A:$\rho\neq\pi$ [109]) with a Bonferroni–Dunn correction ($n_{comparisons}=4$), instances where $P_{BD}\leq 0.05$ were considered statistically significant, and statistically significant instances where $\rho>\pi$ were considered to be overrepresented and instances where $\rho<\pi$ were considered underrepresented (Table S13).

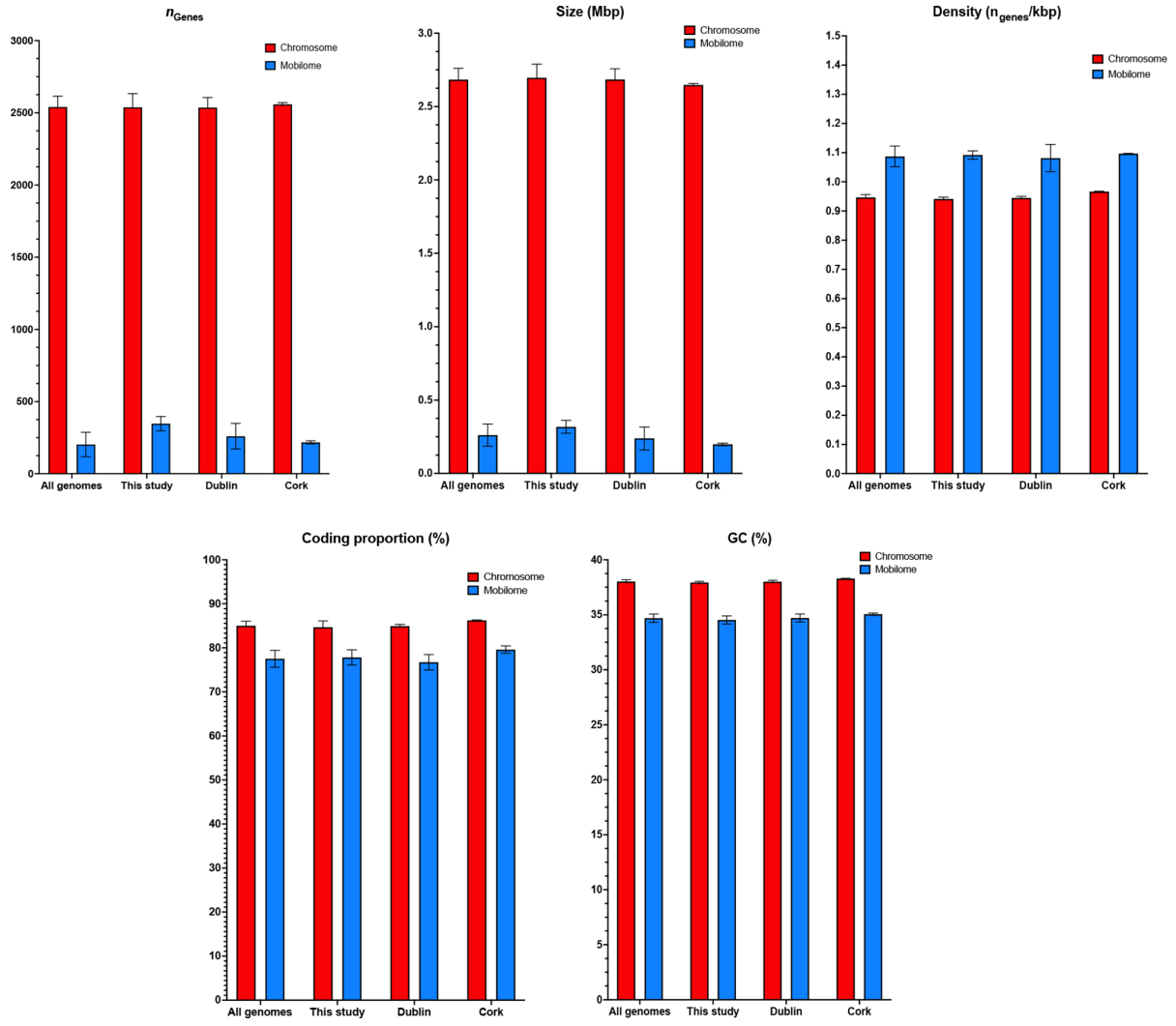

Fig. 2. Comparison of different genome characteristics between chromosomes and plasmids across different studies. Data illustrate mean ± SD for each comparison.

Pangenome function

The representative pangenome refers to the collection of representative sequences from each pangenome cluster. Gene ontology terms (as assigned by InterProScan) were extracted and slimmed from the whole genome dataset (inclusive of all isolates) using the 'map_to_slim.py' from GOATools [110] using generic .obo files. Sequences were grouped into their pangenomic categories and compared to the background population using the 'find_enrichment.py' script [using a Fisher's exact test (H_0:ρ=π; H_A:ρ≠π) and Bonferroni–Dunn correction (Table S14)].

Phylogeny reconstruction

A phylogeny was reconstructed using single-copy, ubiquitous gene alignments (as produced by PRANK during pangenome construction) from the chromosomal dataset. The chromosomal dataset was used to further minimize the likelihood of interference non-vertically inherited sequences on the phylogeny. Each alignment was quality trimmed using TrimAL v.1.4 [111] using the '-automated1' flag. A superalignment was constructed by concatenating all trimmed alignments using FASconCAT v.1.04 [112] and a consensus phylogeny was constructed using IQtree v.1.6.12 [113] with 10000 bootstrap replicates. IQTree used ModelTest-NG v.0.1.7 [114] to determine the GTR+F+R6 [115] nucleotide evolutionary model to be most appropriate for phylogenetic reconstruction. The root of the tree was determined by creating a second phylogeny with a specified outgroup of two *Staphylococcus aureus* genomes. The second phylogeny was constructed by clustering all chromosomal and staphylococcal proteins using ProteinOrtho v.6.0.24 [116] with a stringency cut off value of $E \leq 1.00e^{-50}$. All protein clusters that were ubiquitous in all *E. faecium* species and single copy for each species within each cluster were extracted and aligned using Muscle v.3.8.1551

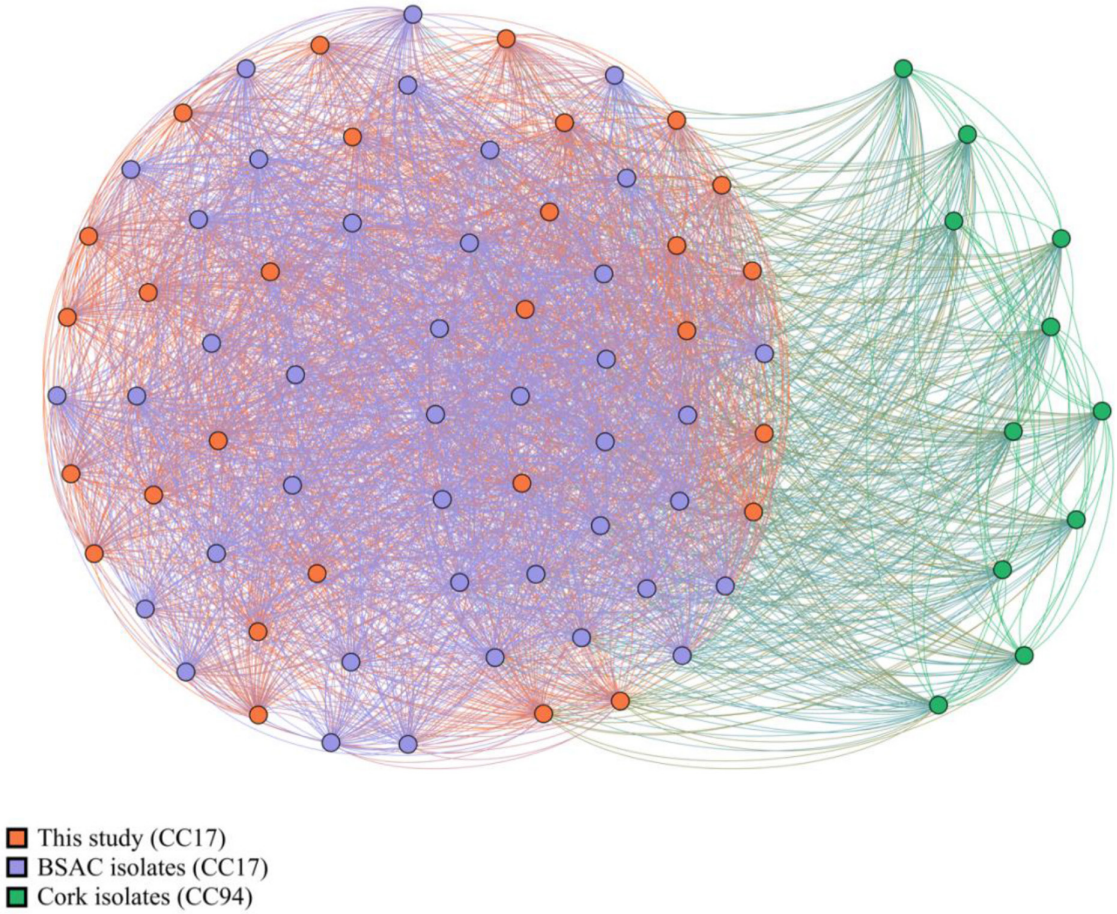

Fig. 3. Genomic distance cluster analysis of chromosomal sequences from all isolates. All CC17 isolates cluster together and all CC94 isolates cluster together (whereby all isolates share an edge). Approximately half of CC17 share an edge with CC94.

[117]. Each alignment was trimmed as above using TrimAL (using the '-automated1' flag) and a consensus tree was constructed as above using IQtree with 10000 bootstrap replicates. ModelFinder-NG determined LG+I+G [118] to be the best model of protein evolution. The second phylogeny determined ST178 (all samples isolated in Cork, Ireland) to be the outgroup of the core gene tree. The finalized phylogeny was displayed and annotated using iToL v.5 [119] (Fig. 5). The phylogeny was annotated with genome annotation data (generated using the analyses above and below) to visualise underlying trends (Figs 6–10 and S1).

Isolate clustering

Isolates were clustered using previously constructed core gene superalignment using RheirBAPS v.1.1.3 [120, 121] with a maximum depth of two, an initial population cluster of 20 and with an additional 100 rounds of processing to approximate optimal clustering (Fig. 5, Fig. 11, Table S15). All isolate-cluster assignments were determined to approach 100% likelihood (P=1 in all cases), indicating that their placement is correct.

Core genome MLST

A core genome multilocus sequence type (cgMLST) was constructed for all isolates using chewBBACA v.2.8.5 [122] using default settings and using the *E. faecium* Prodigal training file provided with the software. The cgMLST profile was displayed and annotated with metadata using GrapeTree [123] with the MSTree V2 algorithm (Fig. 12). The generated cgMLST was verified using our previously generated data pangenomic data with PANINI v.1 [124]. Briefly, as PANINI requires a minimum of 100 input taxa, we constructed a pseudo-dataset where each taxon was represented twice and processed. Pseudo-taxa were removed from the PANINI output and were visualized alongside the previously constructed phylogeny and annotated with their associated RheirBAPS-generated clade using MicroReact [125] (Fig. S2).

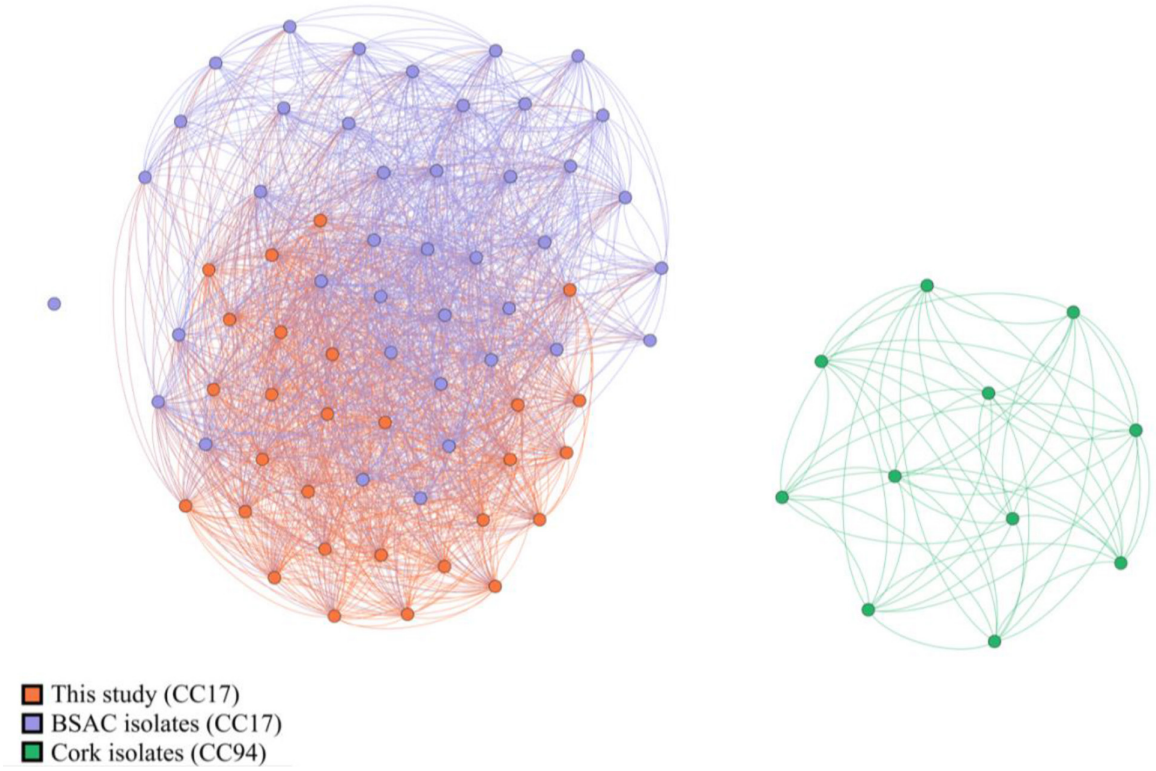

Fig. 4. Genomic distance cluster analysis of mobilomal sequences from all isolates. While some mixing occurs with BSAC isolates and isolates sequenced for this study, divergence between communities is observed. One BSAC isolate (ERR374724) is completely disconnected from the other connected components, which is unusual due to the shared presence of a vanA+ genotype between ERR374724 and many other isolates. All CC94 isolates cluster together in a separate connected component from CC17 mobilomes.

Pangenomic co-evolutionary analyses

The chromosomal phylogeny and Roary-derived pangenome (inclusive of all isolates) were used to test statistically significant associations or dissociations between any vancomycin resistance gene (*vanA, vanB, vanH, vanW, vanX, vanXB, vanY, vanYB*) and any other gene (Table S16) using Coinfinder [56]. Coinfinder uses a binomial test (H_0:$\pi=\pi_x$; H_A:$\pi\neq\pi_x$) and while the authors advise the use of a Bonferroni–Dunn correction, this was not implemented so all coincidences could be explored and instances where $P\leq0.005$ were considered statistically significant.

Recombination analysis

Using RheirBAPS resulted into two distinct clades, the 'Cork' isolates (CC94) and all other isolates (CC17; discussed in a later section). To examine the extent of core genome recombination on CC17 evolution, genomic comparisons were performed between each CC17 isolate and a 'complete genome' *E. faecium* (reference) assembly (downloaded from NCBI Assembly). The reference was selected from a pool of all available *E. faecium* complete genome assemblies (which passed taxonomy assignment) from NCBI Assembly, where each genome was assessed for *vanA* or *vanB* presence using ABRicate with the CARD database (--minid 50), assigned a sequence type using MLST, and assigned a CC using PubMLST. Any assembly with a *vanA*+ or *vanB*+ phenotype, with an indeterminant ST, or assigned to either CC17 or CC94 were discarded. Remaining isolates were processed to remove plasmids and the remaining chromosomal sequence was annotated using Prokka with default settings. Proteins from all CC17, CC94 and candidate reference assemblies were clustered using ProteinOrtho ($E\leq1e^{-50}$). Protein clusters that were both single copy and ubiquitous in all species were individually aligned using Muscle with uninformative regions using TrimAL (with the '-automated1' flag) and concatenated into a superalignment using FASConCAT. The superalignment was processed to construct a consensus tree with 10000 bootstrap replicates using IQTree. The resultant consensus tree was rooted at CC94 and the closest assembly to CC17 was selected as the reference strain. The most suitable reference was determined to be GCF_005166365.1 (strain NM213; PRJNA513159). Strain NM213 was isolated from healthy Egyptian infants in 2018 to determine its potential as a probiotic (https://www.ncbi.nlm.nih.gov/bioproject/PRJNA513159/).

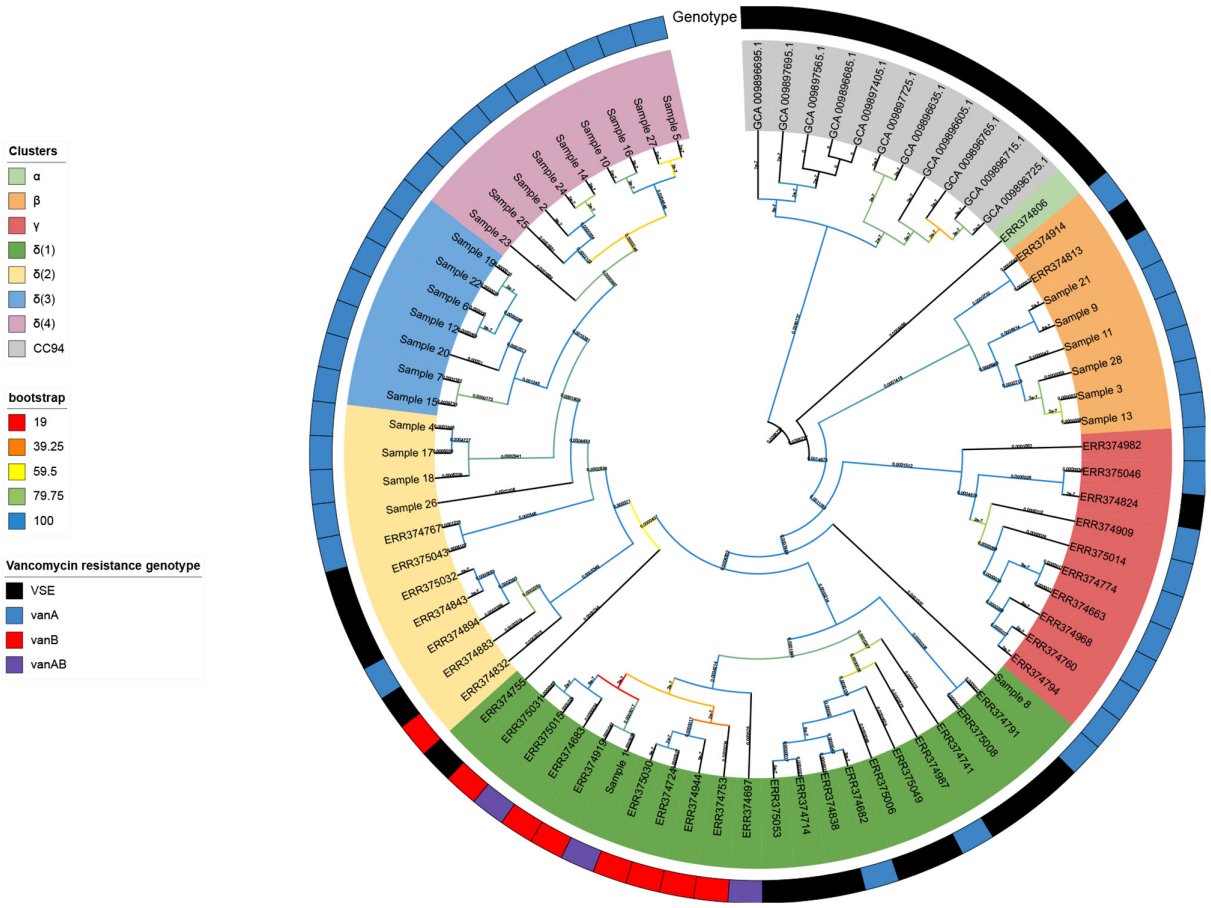

Fig. 5. A 10000 bootstrap phylogeny of isolates. Taxa are shaded based on their RheirBAPS clustering (discussed in Methods) with an outer ring denoting the associated vancomycin resistance genotypes for each taxon. Bootstrap support values are provided for internal nodes. The majority of support values=100%. Bootstrap shading is transitional, and the colours shown in the legend are equidistant scales denoting where a new colour is shown. For example, a datapoint between 79.25% and 100% would be a shade of turquoise. Deep nodes all received high bootstrap support, further corroborating RheirBAPS clustering.

Whole genome alignments were constructed between strain NM213 and each CC17 assembly and concatenated using Snippy v.4.6.0 (snippy-multi) with default settings (https://github.com/tseemann/snippy). The resultant core whole multigenome alignment was processed to convert non-standard nucleotide characters (ATGC) to 'N' using a Snippy auxillary script ('snippy-clean_full_aln'). Genomic recombination regions were detected using Gubbins v.3.0.0 [126] using default settings, and the extent of recombination was illustrated using Phandango v.1.3.0 [127] where the previously constructed core-genome phylogeny (of CC17 and CC94) was used to scaffold recombination (Fig. 13). This procedure was replicated for the remaining 24 suitable references (Fig. S3).

Phages

Phage and prophage elements within each chromosomal and plasmid sequence were assessed using Phigaro v.2.3.0 [128] with default settings and saving the detected viral sequences to their own output files ('--save-fasta' flag) (Table S17). The sum of phages per genome were counted (Table S18). Extracted viral sequences were assessed as a source of drug resistance, metal resistance and virulence factors using abricate with the CARD, BACMET and VFDB databases as above and as a source of secondary metabolism machinery using GECCO as above; only one gene-of-interest was observed (discussed below). Extracted prophages were annotated using Prokka using the '--kingdom Viruses' flag.

Phage classification

A database of all *Caudoviridae* ICTV exemplar reference phage genomes were downloaded from NCBI Assembly. Caudoviridae were selected as this clade encompassed all viral families identified by Phigaro (Siphoviridae, Inoviridae and Myoviridae, respectively). Each extracted prophage was searched against the database using Mash v.2.2.2 [75] using the 'dist' (distance) algorithm. As

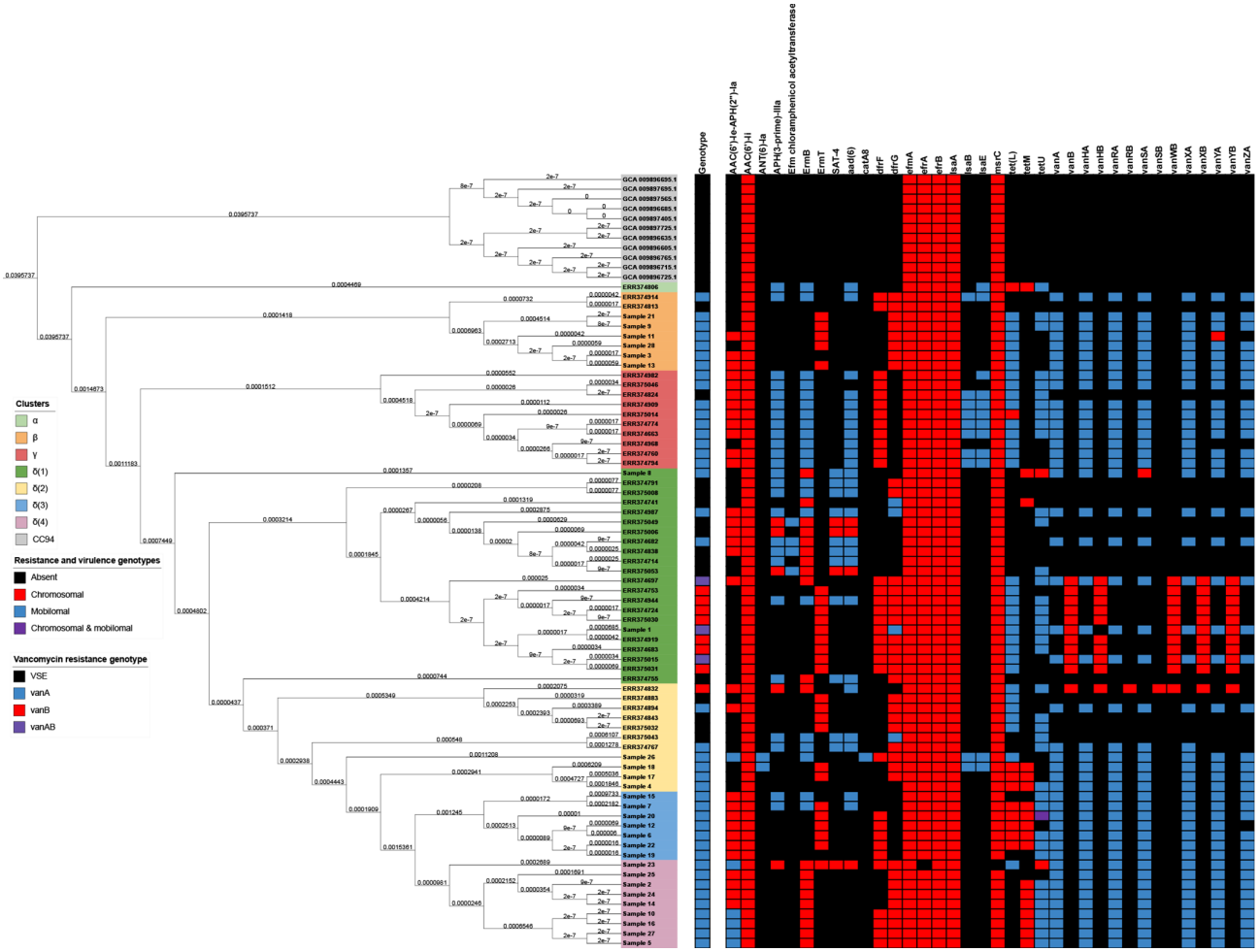

Fig. 6. A phylogeny of isolates illustrating drug resistance genotype distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

we searched prophages (partial genomes) against phages (whole genomes) we selected top hits (one per prophage) from instances with the shortest distances (D<1) and smallest P-values (P≤0.005) for each prophage (Table S18 and Fig. 14).

Integron assessment

The presence of integrons in chromosomal and plasmid sequences was determined using IntegronFinder v.2.0 [129] using the '--local-max' flag and setting '--calin-threshold' flag to 1 (instead of the default 2). We chose to set the '--calin-threshold' flag to 1 to include integrated regions with CALIN artefacts. While lowering this threshold may yield a higher number of false positives, we feel it is appropriate for exploratory purposes. The distribution of integrons per genome is given in Table S19. Again, ABRicate was used to determine whether contigs containing integron elements contained genes-of-interest (Table S20 and Fig. 15).

RESULTS

Quality control

All assemblies were reported to have a high level of completeness ($\mu_{completeness}$=99.29±0.55%; 95.16≤%≤99.61; Table S3).

General characteristics of the genomes

Genomes used in this papers were reported to have a mean chromosomal gene count of 2537.91±77.44 (η=2550.5; 2334≤n≤2730) and an aggregated (chromosomal and mobilomal) gene count of 2821.64±116.65 (η=2812; 2508≤n≤3158). For chromosomal data, we found a mean gene sum of 2537.91±77.44, a mean genome size of 2681.5±78.77 Mb, a mean density of 0.946±0.0098 genes Mb^{-1}, and a mean GC% of 38.03±0.15% (Tables S10–S11 and Fig. 2). For mobilomal data, we found a mean gene sum of

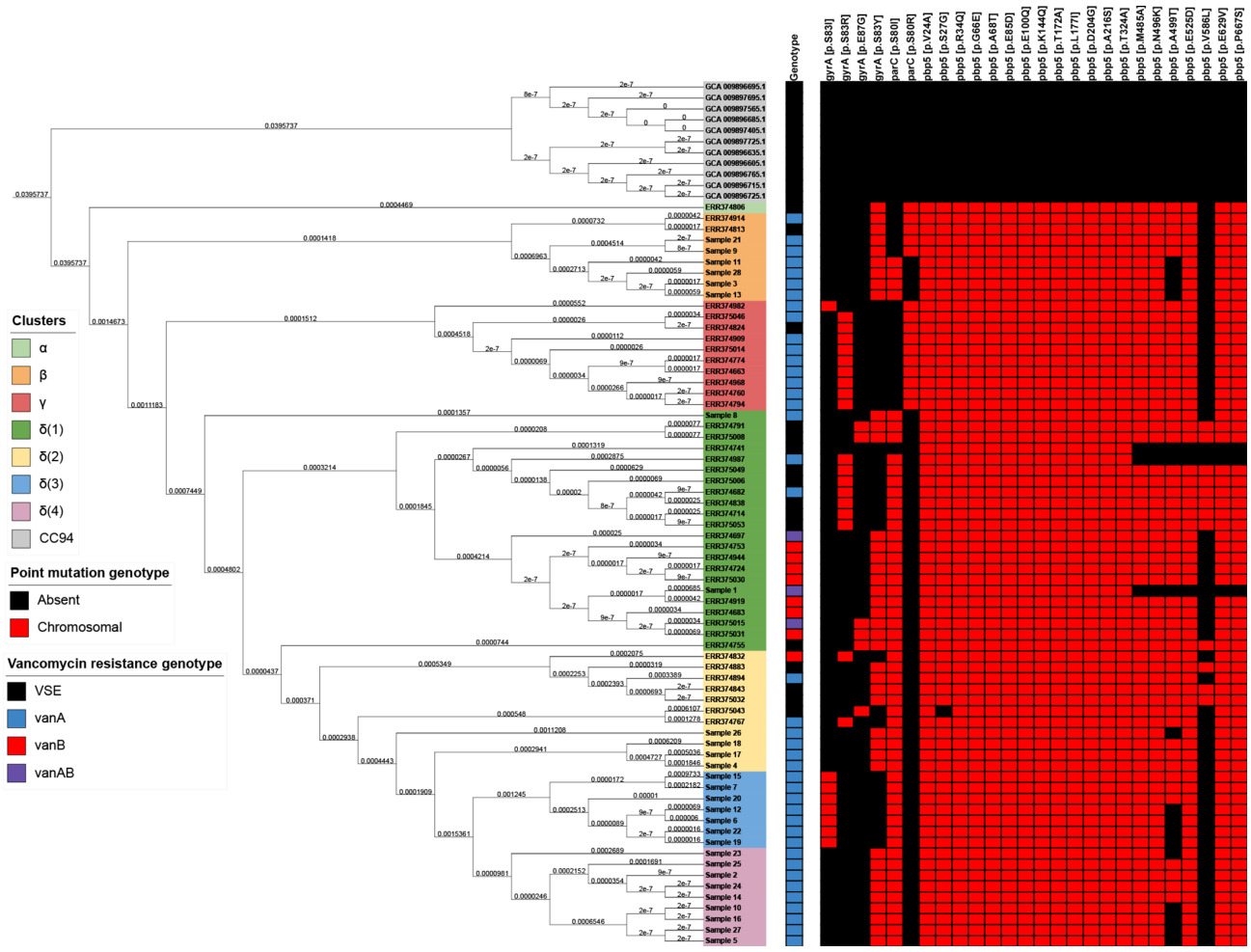

Fig. 7. A phylogeny of isolates illustrating point mutation-mediated drug resistance genotype distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

283.73±84.86, a mean genome size of 259.99±75.76 Mb, a mean density of 1.088 genes Mb^{-1}, and a mean GC% of 34.70±0.38%. When statistically examined, mobilomes were observed to be significantly denser with regard to the number of observed genes (to have a greater number of genes per Mb genome; P_{BD} <1.00e^{-06}) and had a significantly lower GC% (P_{BD} <1.00e^{-06}) When the coding proportions were compared, 19 mobilome samples were reported to be significantly less dense (in terms of the sum of coding nucleotides) than their chromosomal counterparts, but 61 comparisons displayed non-significant differences.

Genetic relatedness (MLST, cgMLST and RheirBAPS)

While a total of 16 STs were observed throughout the dataset, considerable differences were observed between the three sources (isolates sequenced for this study, previously published isolates and 'Cork' isolates). The 'Cork' isolates were exclusively ST178, and ST178 was not observed in either of the clinical sources (Fig. 1, Table S4). Isolates sequenced for this study were predominantly ST80 (16 of 28; 57.14%) and previously published isolates were predominantly ST17 (17 of 40; 42.5%) and ST192 (nine of 40; 22.5%). Interestingly, ST80 and ST17 are exclusively observed in their associated sources, while ST192 and ST203 are observed in both sources.

Isolate clustering using RheirBAPS matched clonal complexes at level one, with all CC94 isolates assigned to Cluster A and all CC17 isolates assigned to Cluster B (v.2.4.2 [64], Table S14). At level two, eight clusters were observed; however using phylogenetic inference, this can be collapsed to five clusters with four subclusters. As all CC94 isolates were again assigned to a single cluster, CC17 isolates formed four clusters (α, β, γ, δ) with four subclusters (δ_1, δ_2, δ_3, δ_4). These clusters are named based on their phylogenetic divergences (Fig. 5) . Cluster α contained one genome, ERR374806 (ST1032), and was the earliest diverging CC17 isolate in our dataset. Cluster β encompassed all ST203 isolates and all ST192 isolates except for ERR374982 (which was assigned to Cluster γ, matching its phylogenetic placement). Cluster γ encompasses all ST78 isolates and the singleton ST192. As

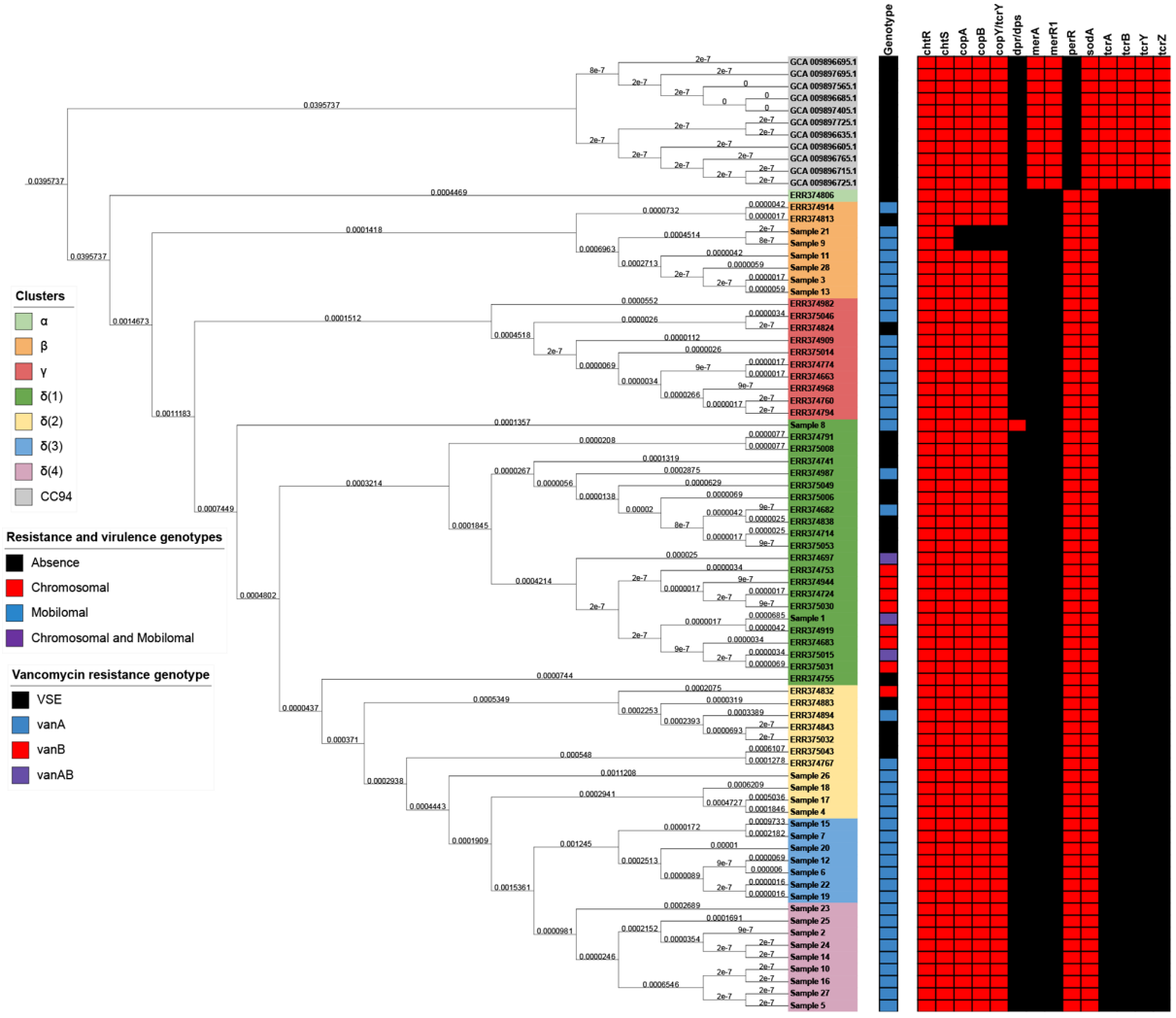

Fig. 8. A phylogeny of isolates illustrating point mutation-mediated drug resistance genotype distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

previously mentioned, Cluster δ comprised four distinct subclusters (δ_1, δ_2, δ_3 and δ_4), where δ_1 is basal to all other δ clades, and δ_2 is basal to sister clades δ_3 and δ_4. Clade δ_1 encompassed all ST17 isolates and isolates assigned to singleton ST groups (ST18, ST132, ST982, ST1038, ST1421). Clade δ_2 encompassed all ST80 isolates, two ST80 isolates and singleton ST isolates (ST16, ST64, ST132, ST202, ST612). Clade δ_3 encompassed seven ST80 isolates and clade δ_4 encompassed the remaining eight ST80 isolates and the singleton ST787. These results suggest that ST80 is a descendant of ST17 and is undergoing considerable chromosomal evolution. Clusters α, β and γ, and Subclusters δ_3 and δ_4 all displayed monophyly. Subcluster δ_2 displayed considerable paraphyly with four root nodes. Subcluster δ_1 also displayed paraphyly [if the Cluster δ earliest diverging taxon (Sample 8) is ignored] due to ERR374755. The Subcluster δ_1 paraphyly may be considered an artefact of cluster evolution, with Subclusters δ_2, δ_3 and δ_4 undergoing rapid evolution after the divergence of ERR374755 from their common ancestor. The paraphyly observed in Subcluster δ_2 may be due to sample size and may be resolved or further partitioned with additional taxa. The paraphyly observed is probably due to evolutionary expansion between sampling times, with earlier diverging isolates (BSAC isolates) being isolated prior to 2011 and later diverging isolates (isolates sequenced for this study) being sequenced in 2018.

The cgMLST graph (and associated graph analyses) displayed considerable modularity (Fig. 12; Fig.S2) with distinct clades forming, largely corroborating RheirBAPS results, regardless of the year of isolation. Two exceptions were observed for this trend: first, polyphyly in Cluster β was observed with isolates from 2009 adjacent to Cluster γ, and those from 2018 adjacent to Subcluster δ_2. Again, Subcluster δ_2 appeared less modular than other clusters (mirroring paraphyly observed in the phylogeny), but this was probably due to core genome evolution between sampling dates. A Subcluster δ_2 isolate was observed adjacent to Subcluster δ_4.

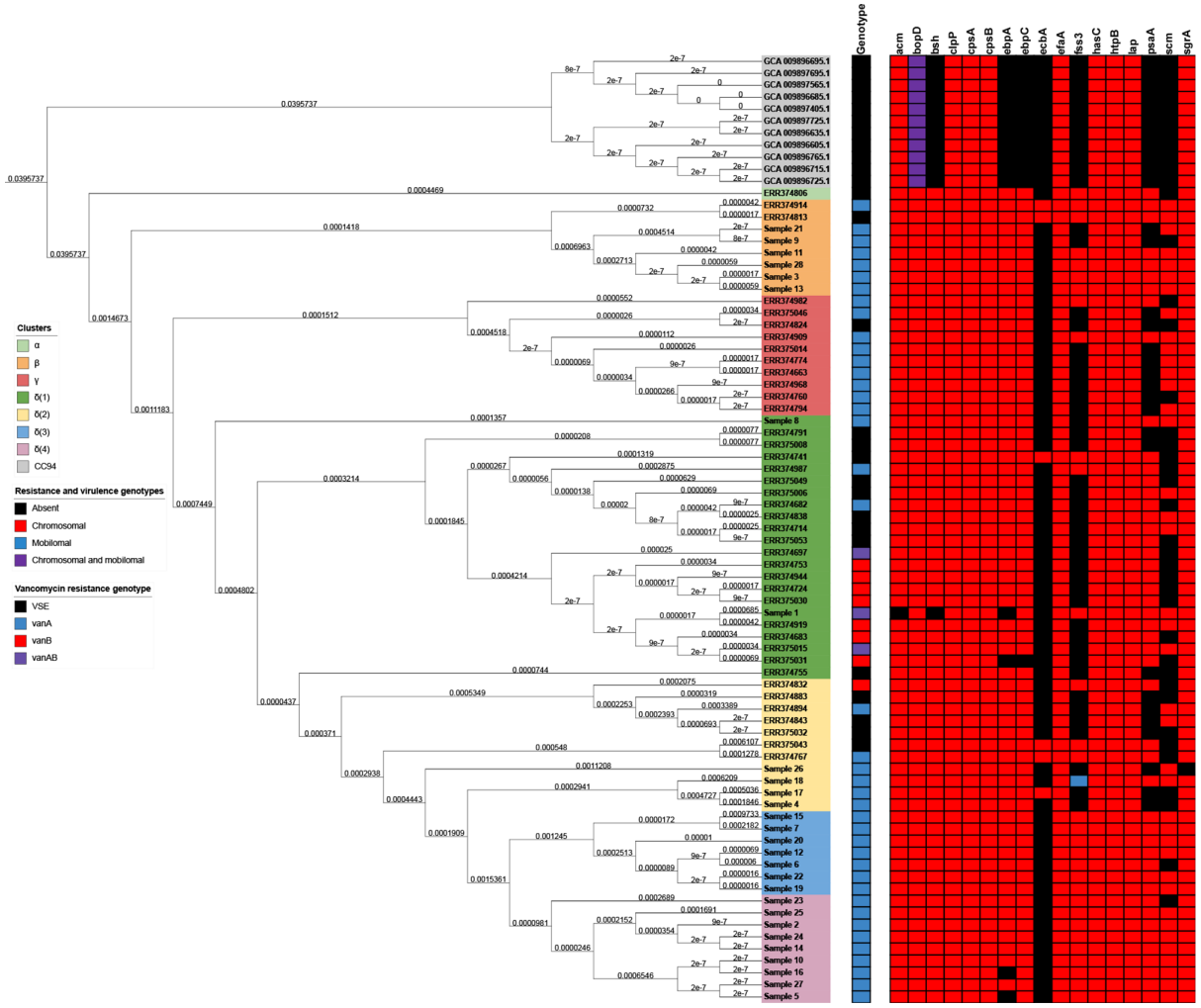

Fig. 9. A phylogeny of isolates illustrating virulence factor genotype distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

Plasmid containment

As expected, the majority of PLSDB hits were attributed to *E. faecium* (Table S5), but some other interesting observations were also noted. All 11 CC94 (Cork) isolates each displayed a hit to two *Listeria monocytogenes* strain CFSAN023459 plasmids (pCFSAN023459_01 and pCFSAN023459_02), and eight CC94 isolates also returned a hit for *Enterococcus hirae* strain CQP3-9 plasmid pCQP3-9_1. Hits pertaining to *Bacillus cereus* plasmid pBC16 were observed across CC17; interestingly, all Cluster γ isolates except ERR375014 (ST78) and all Cluster β isolates except ERR374813 and ERR374914 (ST203; only Cluster β isolates not sequenced during this study) returned a hit. Relatives of plasmid pEGM182-2 from *Enterococcus casseliflavus* strain EGM182 was observed in CC94 and in Subclusters δ_3 and δ_4, where four of six δ_3 isolates (except Samples 6 and 12; ST80) returned two hits for pEGM182-2. Hits were returned for *Escherichia coli* strain UB-ESBL31 plasmid pESBL31 from isolates across the phylogeny. Of note, all Cluster γ isolates and five δ_4 isolates returned a hit for pESBL31. Most of these isolates (14 of 15) were sampled between 2002 and 2011, and one isolate (Sample 1) was sampled in 2018, suggesting persistent survivability for this plasmid and its descendants.

Antimicrobial susceptibility profiles

The 28 vancomycin-resistant *E. faecium* were isolated from 22 hospital patients in 2018/2019 (MMUH, Dublin isolates). Samples 2 and 24 were isolated on the same day from patient 2. Samples 10, 5 and 27 were isolated from patient 5 on days 92, 100 and 107, respectively. Samples 21, 20, 9 and 26 were isolated from patient 9 on days 0, 8, 19 and 365, respectively. The antimicrobial susceptibility profiles (Table S2) of the isolates confirmed they were all vancomycin-resistant. Of the 28 vancomycin-resistant *E. faecium* (VRE*fm*), 25 were erythromycin-resistant, two were chloramphenicol resistant, 28 were ampicillin resistant, 19 were

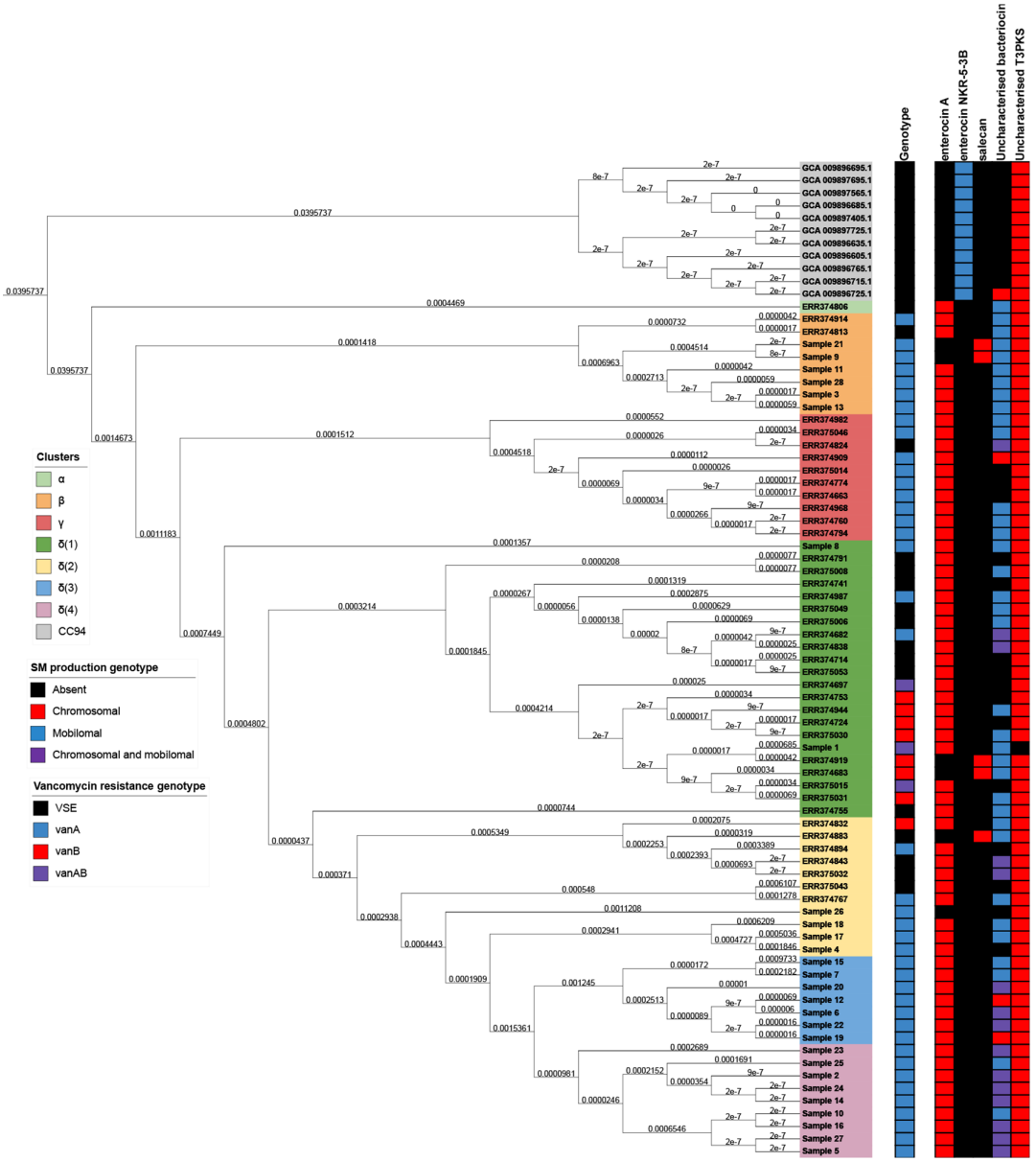

Fig. 10. A phylogeny of isolates illustrating secondary metabolite production genotype distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

tetracycline-resistant and two were linezolid-resistant. Samples 14 and 24 were the only isolates resistant to chloramphenicol or linezolid. While sample 2 was isolated from the same patient as sample 24, it did not display the same resistance profile. The three samples susceptible to erythromycin did not display a specific AMR gene absence associated only with these samples. The metadata of the BSAC study contained data for susceptibility to vancomycin only and the 'Cork' isolates contained no AMR data. Of the BSAC study isolates, 23 were vancomycin-susceptible and 17 were vancomycin-resistant (Table S1).

Antimicrobial resistance genotyping

A total of 32 resistance genes were observed throughout the dataset, of which 13 were exclusively observed in the chromosome [AAC(6')-Ii, *dfrF, ermT, efmA, msrC, tetM, vanB, vanHB, vanRB, vanSB, vanWB, vanXB,* and *vanYB*], where one pair (*vanRB* and *vanSB*) were observed exclusively in isolate ERR374834. Ten genes were observed exclusively on the mobilome (ANT6(6)-Ia, *E. faecium* chloramphenicol acetyltransferase, *catA8, lsaE, tetU, vanA, vanHA, vanRA, vanXA* and *vanZA*), and nine were observed in either the chromosome or mobilome [AAC(6′)-Ie-APH(2″)-Ia, APH(3′)-IIIa, *ermB, SAT-4, aad*(6), *dfrG, tet(L), vanSA* and

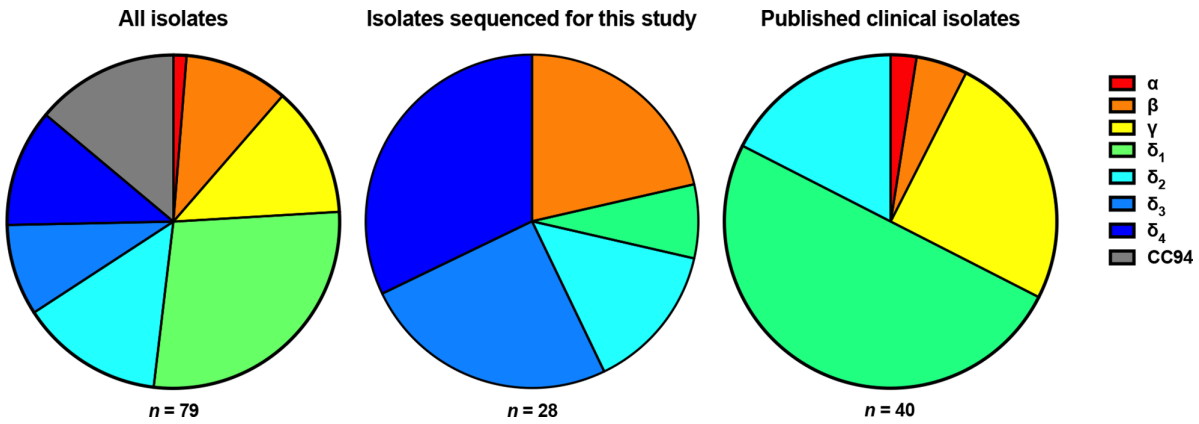

Fig. 11. Distribution of RheirBAPS clusters between studies. All Cork isolates were CC94 and not shown in a separate graph.

vanYA]. Interestingly, the nine genes present on either the chromosome or mobilome never appeared both chromosomally and mobilomally in the same isolate (Table S6 and Fig. 6).

Mobile AMR genes

The *vanA* gene was present only in the mobilome of 44 isolates, almost always with the cluster of *vanHA*, *vanRA*, *vanSA*, *vanXA*, *vanYA* and *vanZA* (Table S6 and Fig. 6). These mobilomes were sampled from across CC17, spanning Clusters β, γ and δ (encompassing 16 ST80 isolates, seven ST78 isolates, five ST203 isolates, four ST17 isolates and three ST192 genomes, and one isolate from each of ST16, ST18, ST64, ST132, ST202, ST612, ST1032, ST1421 and an unassigned ST). The *vanA* isolates comprised samples from the BSAC study (*n*=16) isolated in 2002, 2003, 2005, 2006, 2007 and 2009–2011 and the isolates from Dublin between 2018 and 2019 (*n*=28). The *vanB* gene was identified on the chromosomes in addition to the mobile *vanA* on two isolates from the BSAC study and one isolate from MMUH (sample 1). Samples 14 and 24 (ST80, Subcluster δ_4) were linezolid-resistant but did not contain the mobile linezolid resistance genes. Both samples were also the only isolates displaying chloramphenicol resistance. However, neither contained any known mobile chloramphenicol resistance genes. *E. faecium* chloramphenicol acetyltransferase (conferring resistance to phenicols) was observed in the mobilome of three of 17 ST17 isolates and in the ST1038 isolate (all Subcluster δ_4). Another phenicol resistance gene, *catA8*, was also observed in the mobilome of sample 26. However, all were phenotypically susceptible to chloramphenicol.

The resistance gene *lsaE* (conferring resistance to lincosamide) was observed on the mobilomes of samples 18 and 26 and ten isolates from the BSAC study. The genes *aad(6), ermB* and *aph(3′)III-a* were present in each of the BSAC mobilomes containing *lsaE*. The tetracycline resistance gene *tetU* was observed on the mobilome of six ST80 isolates, two ST18 isolates, and one isolate from each of ST17, ST132, ST787, ST1038 and ST1421 respectively. In addition, *tetL* (conferring resistance to tetracyclines) was frequently present on the mobilomes.

Chromosomally mediated AMR

The *vanB* gene was identified in 11 isolates and was always chromosomal. A cluster of *vanB, vanWB, vanXB* and *vanYB* (conferring resistance to glycopeptides, specifically vancomycin) was observed in nine Subcluster δ_1 isolates (accounting for one ST16, and eight of 17 ST17 isolates) and one Subcluster δ_2 isolate (ERR374832; ST132) (Table S6 and Fig. 6). With the exception of ST16, all isolates with these four genes were also observed to possess *vanHB*, and the ST132 isolate was further observed to possess *vanRB* and *vanSB*. The genes *vanSA* and *vanYA* were individually (and uniquely) observed in sample 8 (unassigned ST) and sample 23 (ST203) respectively.

One chromosomal exclusive gene, *aac(6′)-Ii* (aminoglycoside resistance), was observed to be ubiquitous in all samples and one other, *msrC* (macrolide and streptogramin B resistance), was observed in all isolates except sample 23 (ST64). The combination of these two genes is reported to confer resistance to compounds from the aminoglycoside, lincosamide, macrolide, oxazolidinone, phenicol, pleuromutilin, streptogramin and tetracycline classes [130–133]. Interestingly, CC94 isolates were not observed to possess any other AMR genes beyond these two examples. The genes *ermB* and *ermT* (conferring resistance to lincosamide, macrolide and streptogramin) were present in 46 of 68 CC17 isolates. Only one isolate, sample 23 (ST64), was observed to possess both *ermB* and *ermT* genes and each individual gene was distributed relatively unevenly throughout the dataset (e.g. *ermB* was observed in eight ST80 isolates, *ermT* in five ST80 isolates and five ST80 isolates lacked both genes); however, it was observed that, if present, an isolate only possessed one of *ermB* or *ermT* on the chromosome (with the exception of Sample 23). The distribution

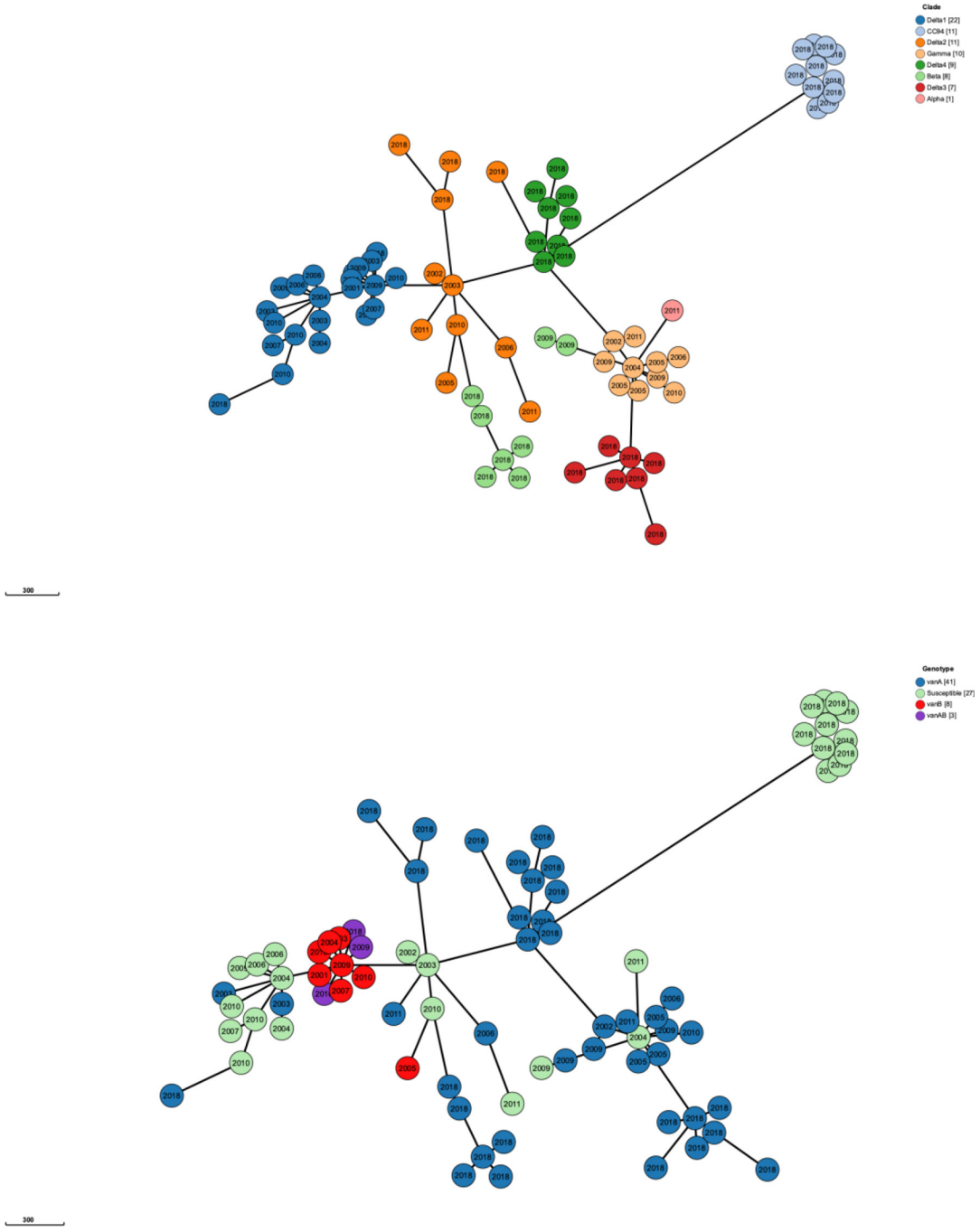

Fig. 12. Core genome MLST (cgMLST) for all isolates included in this study. The upper image is shaded based on RheirBAPS clustering and the lower is shaded based on vancomycin resistance genotypes. The year of isolation is presented in the centre of each circle, where each circle represents an isolate. Scale bar represents the sum of nucleotide substitutions.

of *erm* in this regard does seem to follow an underlying phylogenetic bias. For example, approximately half of Subcluster δ_2 possess *ermB* and the other half possess *ermT*. This transition appears to have occurred after the divergence of ERR374697 as it is retained in all later diverging taxa (Figs 5 and 6). Another example of *erm* bias can be observed between Subclusters δ_2, δ_3 and δ_4, where δ_2 and δ_3 display bias towards *ermB* and δ_4 towards *ermT*. As both genes are observed in the Sample 23 chromosome (the earliest diverging δ_4 taxon), it can be reasonably assumed that this transition occurred after the divergence of Sample 23 from the last common ancestor of the remaining δ_4 isolates.

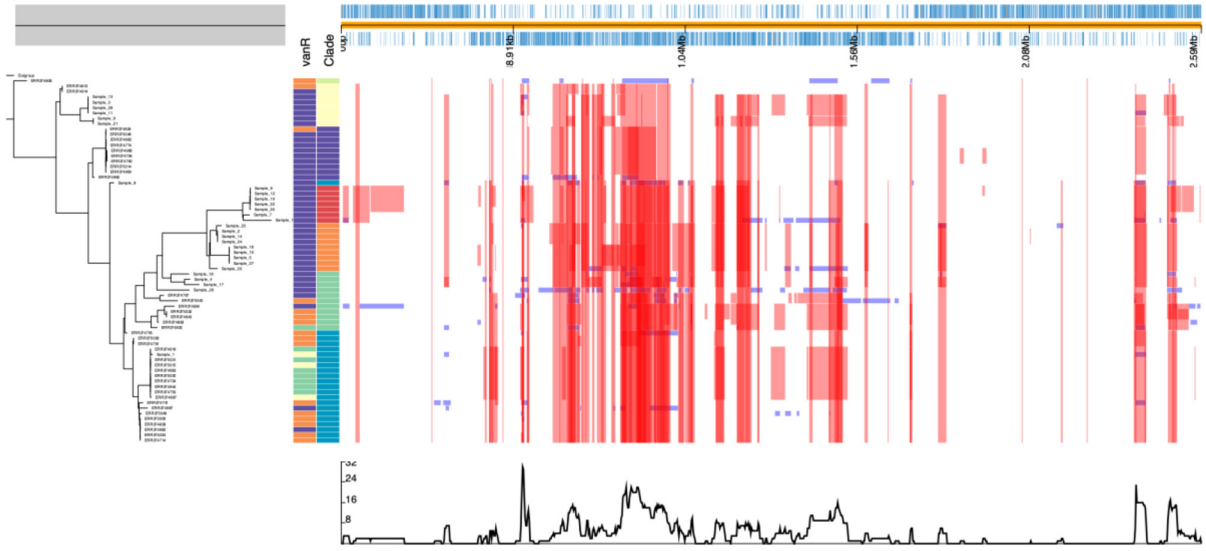

Fig. 13. Red bars illustrate hot spots compared to a reference, where a hotspot is defined as a site where more than one species reports a ≥50% likelihood for the same recombination event. Blue boxes illustrate a recombination event for a single species. Multiple different recombination events can occur at the same site (appearing as overlap in this image). The plot at the bottom shows the maximal number of species affected by a detected recombination event in a given hotspot.

The gene *dfrF* (conferring resistance to trimethoprim) was observed distributed throughout the dataset (like *ermBT*) and displayed phylogenetic biases towards Cluster γ, and *ermT*-positive members of Subclusters δ_1, δ_2 and δ_3. The gene *tetM* (conferring resistance to tetracycline) was scattered throughout the dataset but was specifically observed in 13 or 18 ST80 isolates.

AMR genes present on either mobilome or chromosome

Seven of the AMR genes detected were distributed across either the mobilome or the chromosome within the samples investigated [*AAC(6′)-Ie-APH(2″)-Ia*, *aph(3′)-IIIa*, *ermB*, *SAT-4*, *aad(6)*, *dfrG* and *tetL*]. This demonstrates the inter-connectedness of the mobilome and the chromosome as a mode of AMR gene transport within these isolates (Table S6 and Fig. 6).

Core resistome

Using default pangenomic parameters [107], a core chromosomal resistome (inclusive of the soft-core), whereby a gene is represented in ≥95% (≥75) of genomes, comprises two genes: *AAC(6′)-Ii* and *msrC*. When CC94 is excluded, the core mobile resistome (≥64 genomes) is expanded to include *efmA*. A core mobile resistome was not observed.

Single nucleotide point mutations

Linezolid resistance was identified phenotypically in sister taxa, Samples 14 and 24, but no associated plasmid-mediated resistance mechanism was identified (Table S6 and Fig. 6). In addition, the 23S rRNA, L3 and L4 gene sequences and amino acid sequences, respectively, were compared with the linezolid-susceptible isolates in this study using PointFinder (Table S7 and Fig. 7). No linezolid resistance-associated mutations were identified. Thus, the mechanism of linezolid resistance has not been identified. The linezolid resistance phenotype of the BSAC or CC94 isolates was not reported.

The ciprofloxacin resistance phenotype was not reported in the BSAC or CC94 isolate-associated metadata. Isolates sequenced for this study were all ciprofloxacin-resistant. The gyrase and topoisomerase IV genes (*gyrA* and *parC*) from each isolate were compared with the ciprofloxacin-susceptible strains using PointFinder to identify point mutations associated with ciprofloxacin resistance. Mutations in all isolates were identified at amino acid position 83 in GyrA, resulting in either an S83I change in all δ_2 isolates (Samples 6, 7, 12, 15, 19, 20 and 22) or an S83Y change in the remaining samples sequenced for this study, in *ermT*-positive δ_1 isolates, both ST203 isolates, in the only Cluster α isolate, and in a four-isolate clade of δ_2. In addition, two mutations at amino acid 709: Y709N and Y709D, outside the quinolone resistance determining region (QRDR) in the same groups of isolates. Point mutations at p.S80 were observed across all CC17 isolates except ERR37474. The majority (53 of 78) of isolates displayed an S80I mutation and the remaining 15 isolates displayed an S80R mutation. Point mutations in ParC follow a phylogenetic pattern, appearing in blocks with sister taxa. Interestingly, all Cluster γ isolates displayed the S80R mutation. As these point mutations are at the same site, it can be reasonably assumed that this point mutation is more likely to be inherited from the CC17 common ancestor than via several individual convergent events.

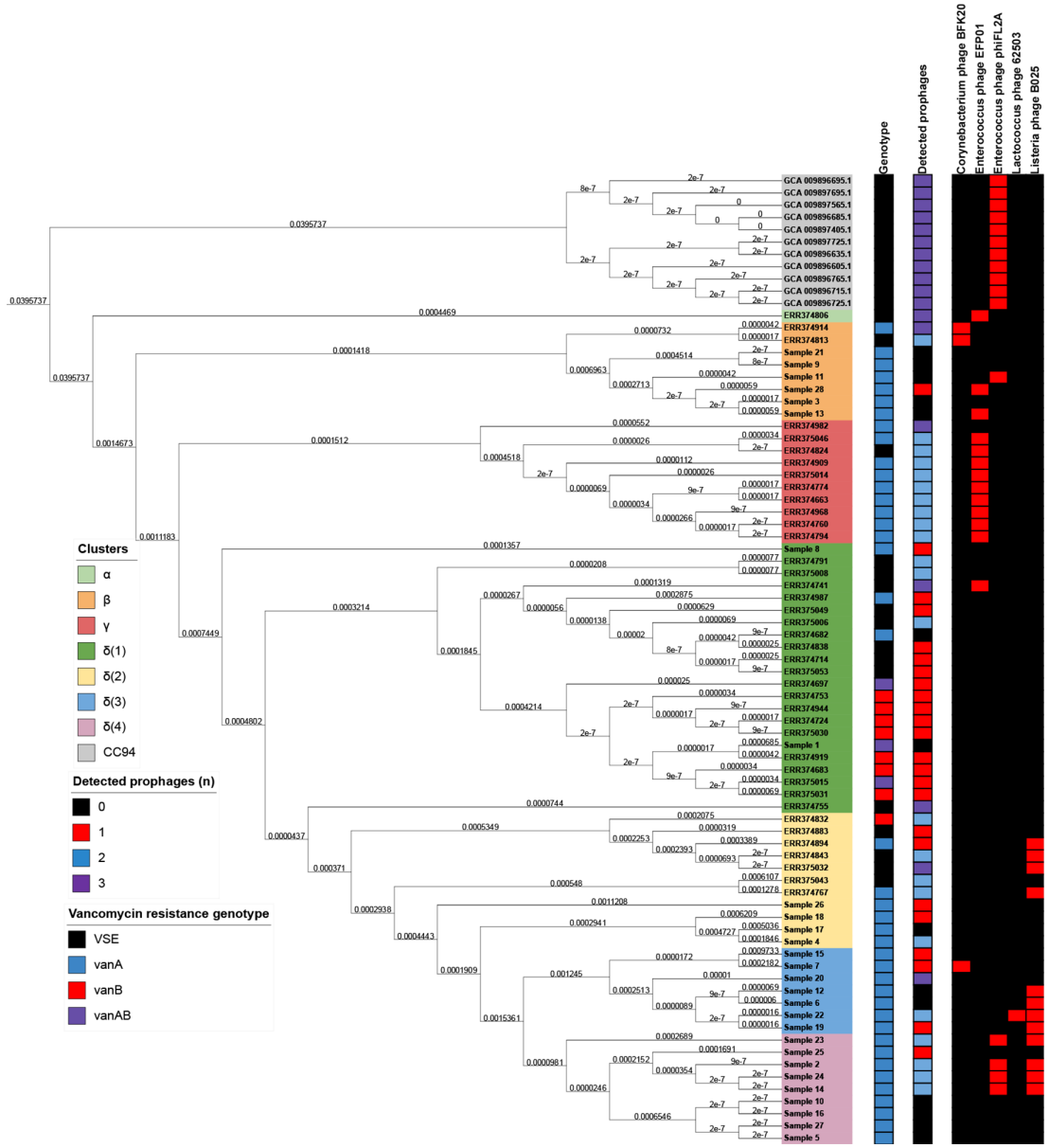

Fig. 14. A phylogeny of isolates illustrating detected prophage distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon.

Thus, ciprofloxacin resistance was due to the mutations at serine 83 in the GyrA protein and serine 80 in the ParC protein in all isolates sequenced for this study.

Metal and biocide resistance

No metal or biocide resistance genes were observed on any mobilome sequence. A 'core' metal-resistome of three single-copy genes (*chtR, chtS* and *sodA*) was observed in all chromosomes (Table S8 and Fig. 8). A further three single-copy genes (*copA*, *copB* and *copY/tcrY*) were observed in all chromosomes except for two isolates sequenced during this study, samples 9 and 21 (both from ST192). One gene, *perR*, was observed as a single-copy orthologue in all clinical isolates (inclusive of ST192) but absent in all 'Cork' samples. One gene, *dpr/dps*, was observed exclusively in sample 8 (unassigned ST). Finally, a set of six genes (*merA*, *merR1*, *tcrA*, *tcrB*, *tcrY* and *tcrZ*) were observed exclusively in the 'Cork' samples (CC94). In summation, all samples (except ST192) shared seven genes (*chtR, chtS, copA, copB, copY/tcrY, perR* and *sodA*). These seven genes confer resistance towards selenium (*chtR*), hydrogen peroxide (*chtR, perA, sodA*), chlorhexidine (*chtS*), copper (*copA*, *copB*, *copY/tcrY*) and silver (*copB*) [36, 134]. The absence of *copABY* from ST192 may indicate copper and silver susceptibility in these samples. The presence of *dpr/*

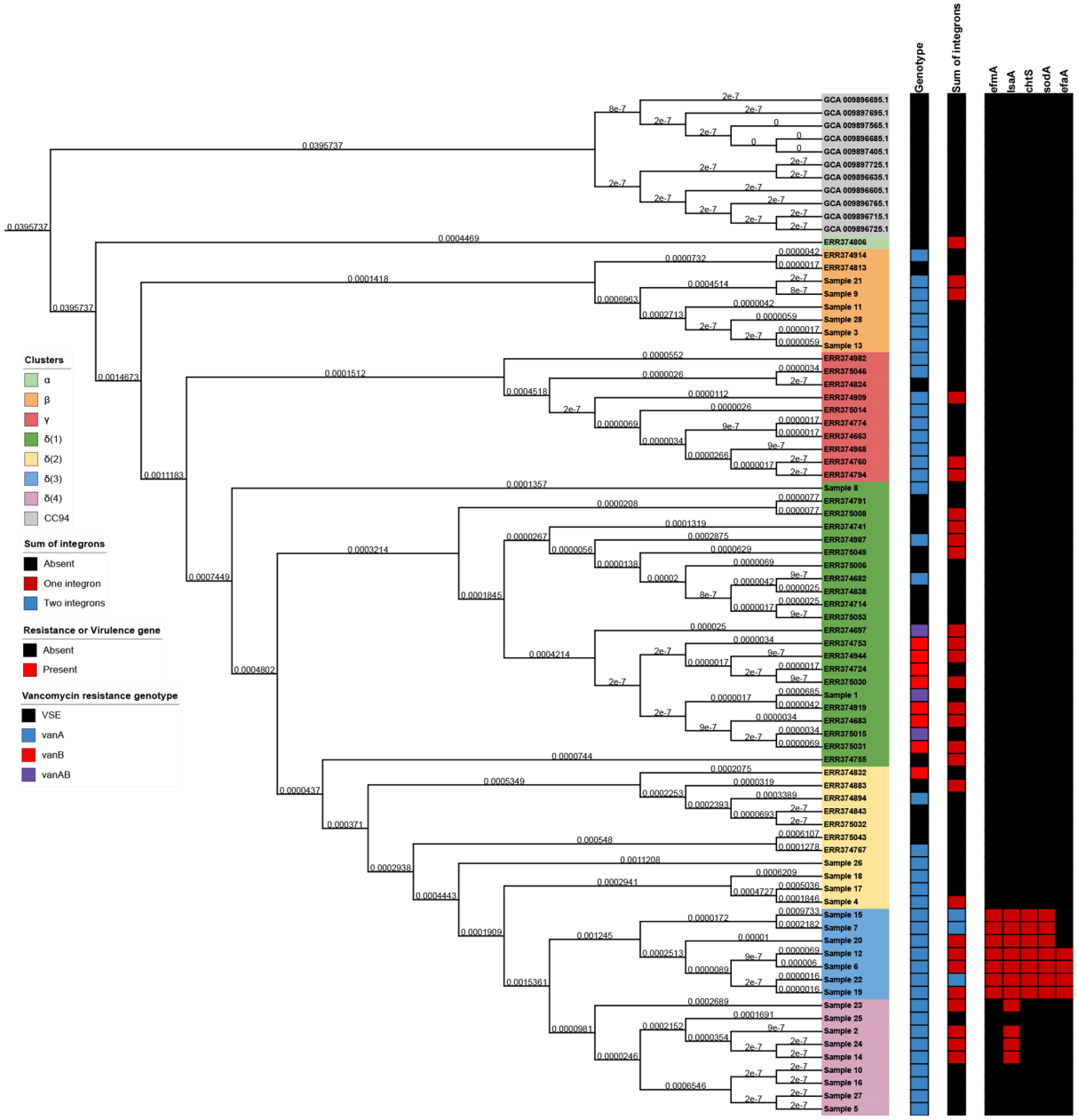

Fig. 15. A phylogeny of isolates illustrating detected integron distributions as a categorical heatmap. Again, taxa are shaded based on their associated RheirBAPS cluster and a vancomycin resistance genotype is given as a separate bar for each taxon. Gene distributions (presence or absence) are given for *efmA* and *lsaA* (drug resistance), *chtS* and *sodA* (metal and biocide resistance), and *efaA* (virulence factor).

dps in sample 8 is expected to confer iron resistance and increased resistance to hydrogen peroxide [135, 136]. The presence of *merA* and *merR1* in the CC94 samples indicates intrinsic resistance to mercury and phenylmercury acetate (an organomercuric compound), and the presence of *tcrABYZ* suggests increased copper resistance [43, 137–140].

Using default pangenomic parameters [107], a core chromosomal metal resistome (inclusive of the soft-core), whereby a gene is represented in ≥95% (≥75) of genomes, comprise six genes: *chtR, chtS, copA, copB, copY/tcrY* and *sodA*. When CC94 is excluded, the core metal resistome (≥64 genomes) is expanded to include *perR*.

Virulence factors

Each CC94 isolate contained a plasmid-mediated virulence factor (VF), the biofilm-associated transcription factor *bopD* (Table S9 and Fig. 9). Only one other VF was observed to be plasmid-mediated: the fibrinogen binding surface protein *fss3* [only observed

within sample 18 (ST80)]; all other observations were chromosomally mediated. All isolates were observed to possess *clpP* (a caseinolytic protease), *cpsA* (a biofilm-associated undecaprenyl diphosphate synthase) and *cpsB* (a capsule-associated phosphatidate cytidylyltransferase). A collagen adhesin precursor (*acm*) was observed in all isolates except sample 1 (ST1421) and the cell wall anchor protein *sgrA* was observed in every isolate except sample 26 (ST202). All isolates in CC17 [except Sample 1 (ST1421)] contained the bile salt hydrolase *bsh* and all CC17 isolates except ERR375031 (ST17) contained the endocarditis/biofilm-associated pilus *ebpC*. The collagen binding microbial surface components recognizing adhesive matrix molecule (MSCRAMM) gene *ecbA* was observed in sample 17 (ST80), ERR375043 (ST16), ERR374741 (ST17), ERR374767 (ST64), and in two ST203 samples (ERR374813 and ERR374914). The collagen adhesin protein *scm* was observed in 14 of 16 ST80 isolates, five of 17 ST17 isolates, two of five ST18, seven of nine ST78, one of three ST192 (sample 21) and in sample 26 (ST202). Finally, *fss3* was observed in 14 of 16 ST80 isolates, all six ST203 isolates, in both ST132 isolates, in one isolate each from ST16, ST17, ST64, ST78, ST192, ST787, ST1032 and ST1421, and one unclassified ST isolate (sample 22). Both *fss3* and *scm* were absent from sample 17 (ST80).

Using pangenomic parameters, a core virulome (inclusive of the soft-core), whereby a gene is represented in ≥95% (≥75) of genomes, would yield *acm, bopD, clpP, cpsA, cpsB, efaA, hasC, htpB, lap* and *sgrA*, illustrating a key pathenogenic potential for biofilm-associated proteins, capsular polysaccharide biosynthesis, caseinolytic protease, endocarditis-specific antigen and collagen adhesins [141, 142]. When only clinical isolates are considered, the core virulome (≥64 genomes) is expanded to include *bsh, epbA* and *epbC,* further increasing cardiac virulence and limiting the antibacterial function of bile acids [143, 144].

Secondary metabolism

A distinct delineation was observed in enterocin (bacteriocin) biosynthesis across the dataset (Fig. 10). Enterocin A production cassettes were observed on the chromosomes within 63 of 69 CC17 isolates and Enterocin NKR-5-3B production cassettes were observed in the mobilome of all CC94 isolates. A second (albeit uncharacterized) bacteriocin production cassette was observed across CC17. Interestingly, while mostly observed on mobilomes, this uncharacterized cassette was sometimes observed on chromosomes without a mobilome (e.g. on Samples 12 and 19); it was observed to be within both chromosomes and mobilomes in the majority of Subcluster δ_3 and δ_4 genomes. An uncharacterized type 3 polyketide synthetase (T3PKS) was observed on the chromosomes of all samples except for Sample 1. Finally, salecan (β-glucan) production cassettes were observed in the chromosomes of five isolates: Samples 9 and 21 (Cluster β), ERR374919 and ERR374683 (Subcluster δ_1), and ERR374883 (Subcluster δ_2).

Prophage detection

Prophages were observed in the chromosome of 65 isolates (2.59±1.02; $1 \leq n \leq 4$) and all were observed to be non-transposable (Tables S17–S19). The majority (196 of 202) of prophages were predicted to be Siphoviridae, three prophages were observed to be either Siphoviridae or Myoviridae in ERR374794, ERR374824 and ERR375046, one prophage (from ERR374806) was observed to be Myoviridae, one prophage (from sample 25) was observed to be Inoviridae, and one prophage (from sample 20) was of unknown taxonomy. Prophages were not observed within any mobilome. One resistance gene, *ermB*, was observed in a prophage of ERR374741, but no other resistance mechanism (antibiotic or metal/biocide) and no virulence factors were observed on any other prophage. Without this *ermB* phage, ERR374741 would be susceptible to lincosamide, macrolide and streptogramin. The prophage genome within ERR374741 was also observed to contain a probable secondary metabolism mechanism, but the product class could not be determined. One prophage within sample 13, sample 28 and ERR374909 was observed to contain a probable polyketide synthase. When these prophages were searched using Mash (Table S19), a set of four phages were mapped to 40 of 79 genomes. In the 'Cork' samples, *Enterococcus* phage ΦFL2A was ubiquitously observed. A total of 17 prophages were observed within 17 BSAC genomes, *Enterococcus* phage EFP01 was observed in 11 genomes, *Listeria* phage B025 was observed in four genomes and *Corynebacterium* phage BFK20 was observed in two genomes (Fig. 14).

When isolation dates were considered for the BSAC samples, vancomycin resistance did not appear to be affected by the presence of a prophage as both VSE and VRE were represented [57]. Both *Corynebacterium phage* BFK20 prophage-containing genomes were isolated in 2009, *Listeria* phage B025 prophage-containing genomes were isolated between 2002 and 2011, and *Enterococcus* phage EFP01 prophage-containing genomes were isolated between 2001 and 2011. In genomes sequenced for this study, *Corynebacterium* phage BFK20 was observed in one genome (sample 7), *Enterococcus* phage EFP01 was observed in two genomes (samples 13 and 28), *Enterococcus* phage ΦFL2A was observed in sample 11, *Listeria* phage B025 was observed in three genomes (samples 6, 12 and 19), and a duo of *Enterococcus* phage ΦFL2A and *Listeria* phage B025 was observed in five genomes (samples 2, 14, 22, 23 and 24).

A core phage has previously been observed in studies into *E. faecalis*, where it is implicated in virulence and genome plasticity [145], but, to our knowledge, this trend has not been observed for *E. faecium*.

Pangenomics

As mentioned above, a pangenome was constructed for this dataset and each pangenomic category was populated [n_{genes}; 'core pangenome' ($99\% \leq n_{samples} \leq 100\%$), 'soft core pangenome' ($95\% \leq n_{samples} < 99\%$), 'shell pangenome' ($15\% \leq n_{samples} < 95\%$), and 'cloud

pangenome' ($n_{samples}$ <15%)]. The number of genes in the 'whole genome' core and soft-core pangenomes and the 'chromosomal' core and soft-core pangenomes are relatively unaffected by the mobilome (1324 and 236 genes vs. 1327 and 229 genes respectively; Table S12). However, the shell pangenome (2252 vs. 1550 genes) and the cloud pangenome (3306 vs. 2762 genes) were more pronouncedly affected.

Pangenomic enrichment

The core pangenome (inclusive of all isolates) was observed to be enriched (P_{BD} ≤0.005) for major 'house-keeping' and viability functions such as metabolism [e.g. lipid metabolic process (GO:0006629), protein metabolic process (GO:0019538), carbohydrate metabolic process (GO:0005975), and RNA metabolic process (GO:0016070)], transport [e.g. protein transport (GO:0015031) and ion transport (GO:0043167)], stress response (GO:0006950), signal transduction (GO:0050794), biosynthesis (GO:0009058), cellular regulation (GO:0065007) and localization (GO:0051179) (Table S13). The soft-core pangenome was observed to be enriched for localization and for transmembrane transport (GO:0022857). The shell pangenome was enriched for similar processes as the core genome, namely metabolic processes, transport, stress response and localization. The shell pangenome was also significantly enriched for small molecule metabolism (GO:0044281) and transposition (GO:0032196). The cloud pangenome was also enriched for metabolic processes, transport and signal transduction. There was only one significant purification: transposition in the core genome. These results suggest a dynamic and non-specific evolution of the *E. faecium* pangenome and provides insight into their genomic organization and architecture.

Co-evolution

When vancomycin resistance genes were sampled from the 28-genome dataset, a significant association (*P*≤0.005) was observed between *vanAHX* and each of *cadC, dinB, gmuD, gmuR, hin, hisB, nag3, sdhB* and *uxaC*, and a significant dissociation with each of *clpC, czcD, czrA, dkgB, dps, entP, fetA, fetB, fixK, fruA, gatA, lgt, mco, metB, opuCA, panE, phoP, rihB* and *uvrA* (Table S14). These results suggest that VR*Efm* carrying *vanA* typically display an increase in carbohydrate processing (via *gmuD, gmuR, nag3, sdhB* and *uxaC*), cadmium transport (via *cadC*) and DNA repair (via *dinB*) [146–153]. Conversely, isolates with a *vanA*+ genotype were also more likely to display genotypes for a decrease in biofilm formation (via *clpC*), lipoprotein synthesis (*via lgt*), copper, iron and zinc resistance (via *czrA, fetAB,* and *mco*), pantothenate biosynthesis (via *panE*), quarternary ammonium compound resistance (via *opuCA*), oxidative resistance (via *fetAB* and *opuCA*) and stress response (*clpC, dps* and *uvrA*) [154–163]. Comparatively, *vanBWXBYB* was observed to be significantly associated with each of *agaS, bfrA, dgaR, dgoA, dgoD, epsF, fruA, gatB, gmuE, gmuR, gpr, Int-Tn, lacD2, manR, manX, metK, mhpE, mro, mshA, noxE, sorC, ssbA* and *treP* and dissociated with each of *aes, araQ, cysM, dctM, dppC, dppE, oppB* and *tuf*. These observations suggest that *vanB*+VR*Efm*, like *vanA*+VR*Efm* possess increases in carbohydrate (specifically sugar) catabolic processing genotypes (via *agaS, dgoAD, lacD2, manRX, mshA,* and *treB*), but an increase was observed in biofilm formation genotype (via *epsF*) and stress response (via *treB*) [164–170]. Isolates with a *vanB* phenotype were observed to be decreased in phenotypes for arabinose uptake (via *araQ*), cysteine synthesis (via *cysM*), dicarboxylate transport (via *dctM*) and protein transport (via *dppCE* and *oppB*) [171–175]. These protein transport genes are essential for sporulation in other *Firmicute* species, but as *Enterococcus* are non-spore forming, their exact role in relation to *vanB* has not been elucidated [175–177].

Interestingly, fructose processing genotypes (via *fruA, gatAB* and *gpr*) are significantly decreased in *vanA*+ isolates while being significantly increased in *vanB*+ isolates (Table S16), suggesting a possible fructose metabolic niche in *vanB*+ isolates.

Rampant recombination

Whole genome alignments and subsequent recombination analyses indicate massive recombination events across the CC17 genome during its divergence from non-CC17 *E. faecium* ancestors (Figs 13 and S3). Most of the recombination appears to have occurred approximately within the first 50% of the genome.

Integron evolution

A total of 31 CC17 isolates were observed to possess an integron element, where three isolates [Samples 7, 15 and 22 (all Subcluster δ_3)] were observed to possess two (Table S20 and Fig. 15). Integrons followed a phylogenetic distribution in Subclusters δ_3 and δ_4, but all other clusters were observed to be pseudorandomly distributed throughout the phylogeny. Genes-of-interest (resistance or virulence genes) were only observed on integron-containing contigs in Subclusters δ_3 and δ_4. All isolates with a gene-of-interest possessed *lsaA* (conferring resistance to clindamycin, quinupristin-dalfopristin and dalfopristin). Subcluster δ_4 isolates only possessed *lsaA*. All Subcluster δ_3 isolates also possessed *efmA* (an efflux pump conferring resistance to macrolides and fluoroquinolones), *chtS* (conferring resistance to chlorhexidine) and *sodA* (conferring resistance to peroxides). A clade within Subcluster δ_3 (comprising Samples 6, 12, 19 and 22) also all contained *efaA* (an endocarditis-specific antigen) on their integron-containing contigs.

Observation of an unusual contig

Five isolates [Samples 2, 14, 24 and 25 (all assigned to ST80; subcluster δ_4) and Sample 18 (ST80; Cluster β)] possessed a contig with a copy of *arnB,* a UDP-4-amino-4-deoxy-L-arabinose–oxoglutarate aminotransferase involved in Gram-negative lipid A biosynthesis and polymyxin resistance [178–180]. The contig in samples 2, 14 and 24 are identical (Fig $S4_{(a)}$–$4_{(c)}$), but those attributed to samples 18 and 25 are distinct (Fig $S4_{(d)}$–$S4_{(e)}$). Samples 2 and 24 were isolated from the same patient and all three of these samples are phylogenetic neighbours (Fig. 5). These contigs have a plethora of cell wall synthesis and lipopolysaccharide modification genes which may play a role in resistance. Of particular interest, sample 17 contained phage sequences suggesting a possible viral-mediated horizontal gene transfer. Instances of chromosomal segment transfer between Gram-positive and Gram-negative species (and vice versa) has previously been attributed to phage activity and plasmid integration [181–183].

DISCUSSION

Within Ireland, VRE has already spread nationally and is endemic to Irish hospitals [184]. The level of VRE in bloodstream infections in Ireland has remained at or above 30% since 2005. While some other EU countries have levels above 30%, they have not consistently had this level of VRE bloodstream infections over such a wide timeframe. In 2019, >25% of VR*Efm* bloodstream infections in 11 additional EU countries have been identified using the ECDC surveillance data. The spread and dominance of VR*Efm* within Irish hospital patients over such a long duration provides us with a unique setting to study the changing dynamics and evolution of VR*Efm*. A recent study of the global dissemination of *E. faecium* identified that it has two main modes of genomic evolution: the acquisition and loss of genes, including antimicrobial resistance genes, through mobile genetic elements including plasmids, and homologous recombination of the chromosome. Within this global study, 261 genomes contained the *vanA* gene. Within the hospital-associated A1 clade of *E. faecium* the median number of AMR genes increased in the global study between 2001 and 2019 from eight to 11 in 2015. Within this study for the *Efm* isolated between 2001 and 2011 the total AMR gene numbers matched the presence of *vanA* rather than year. The vancomycin-susceptible isolates and those with the *vanB* gene had fewer AMR genes than those harbouring a *vanA* gene. Briefly, *vanA*+ isolates had between 13 and 20 AMR genes, while those with *vanB* or no vancomycin reistance gene had between five and 12 AMR genes [except for one vancomycin-resistant isolate with *vanB* (n_{genes}=15) and one vancomycin-susceptible isolate with no vancomycin-resistant genes (n_{genes}=16)]. The timeframes for both groups were distributed across the 10 years. Those with the highest AMR gene numbers both contained *vanA* and *vanB*. The number of AMR genes from the *vanA* study in 2018/2019 was between 12 and 18, which is consistent with the number from the previous study. Thus, the number of AMR genes in Ireland did not increase over time.

The aims of this study were to compare the pan-genomes, mobilomes and chromosomes of the available genome sequences of vancomycin-resistant or susceptible *E. faecium* across the timeframe that VRE [analysed by the ECDC (2002–2019)] increased in prevalence from 11.1 to 38.4% in Ireland. To monitor and limit the spread of VR*Efm* we need to understand how it evolves and acquires vancomycin resistance, how transmission networks are operating and how VR*Efm* is developing resistance to last-line antibiotics.

The same AMR gene was not observed in both chromosome and mobilome in any instance, which indicated genomic control of genetic redundancy [185–187]. Vancomycin resistance was observed to be predominantly either plasmid-mediated (via *vanA*) or chromosomally mediated (via *vanB*). Both genotypes (*vanAB*+) were observed in hree of 79 isolates (or three of 52 *van*+ isolates). As mentioned previously, this rarity could be due to genomic redundancy control. As *vanB* is chromosomally mediated, the plasmid-mediated *vanA* is more likely to be lost due to functional redundancy unless the plasmid confers other ecological fitness benefits [188].

Tetracycline (and probable tigecycline) resistance [via *tet*(*L*)] was differentially dispersed in clinical isolates, with *tet*(*L*) being observed in the mobilomes of ST17, ST18 and ST203 and in the chromosomes of ST80 (if present within the genome). Other tetracycline resistance genes implicated in tigecycline resistance, *tet*(*M*) and *tet*(*U*), were observed in addition to *tet*(*L*) on the chromosome of most ST80 isolates. Interestingly, in contrast to previously reported trends [189, 190], increased metal resistance did not positively correlate with increased drug resistance, and in fact the opposite was observed, with the 'Cork' strains displaying diminished drug resistance and increased metal resistance.

Core-genome MLST corroborated phylogenetic and hierarchical clustering results, suggesting that core populations of *E. faecium* persist in Irish hospitals with frequent evolution towards pathogenicity via plasmid-mediated transfer of *vanA* or chromosomal incorporation of *vanB* (Figs 5 and 12). This suggestion is corroborated by the genomic similarity between isolates in different clusters and clades (regardless of isolation year) and by the fact the VSE is observed within Cluster β and clades δ_1 and δ_4 (Fig. 5).

We determined that a relatively stable 'prophagome' exists between isolates of the same phylogenetic cluster (Table S17), although, these are not, to our knowledge, implicated in the evolution of resistance or pathogenicity, as observed with closely related species [191, 192].

An interesting potential virulence mechanism difference was observed between *vanA*+ and *vanB*+ isolates, whereby *vanA*+ genotypes were unlikely to display a biofilm-forming genotype (via *clpC*) and *vanB*+ isolates were likely to display an *epsF* biofilm-forming

genotype (Table S16). Both vancomycin resistance genotypes displayed increases in carbohydrate catabolism genotypes, while *vanB*+ isolates were more likely to display specific sugar processing genotypes. Interestingly, *vanA*+ genotypes were less likely to coevolve with metal resistance, oxidative stress resistance, panthenoate or lipoprotein synthesis genotypes. These results suggest that *vanB*+ isolates are more likely to have evolved due to different stressors than *vanA*+ isolates. As *vanA* alignments used in this study did not have any point mutations (regardless of the year of isolation), it is likely that a narrow range of mobile genetic elements confer resistance to commensal VSE in the immunocompromised, prompting their pathogenicity.

The significant differences in observed gene density (higher in mobilomes), GC% (lower in mobilomes) and gene length (shorter in mobilomes) coupled with insignificant coding proportions between chromosomes and mobilomes suggest that mobilomes are more likely to evolve either through the acquisition of statistically shorter sequences (while maintaining an equivalent coding potential) or through successive gene fission and/or gene subfunctionalization events [193–196]. While not statistically significant in most samples, coding nucleotide density is, on average, approximately 7% less in mobilomes (Tables S11 – S12), where significantly variant coding proportions are also observed [$\sigma^2_{chromosomal}$=1.04; $\sigma^2_{mobilomal}$=3.62; P=1.99e^{-10} (Levene's test)]. It is possible that these variances may just be an artefact of the highly variant sequences distributed across plasmids when compared to more stable chromosomes, and these variances may also underly nucleotide deletion 'clean up' processes following subfunctionalisation or fission events, indicating evolution towards a more streamlined, and therefore less bioeconomic expensive plasmids, thus promoting their propagation [195, 197–199].

The pangenome of *E. faecium* is relatively open, with core and soft-core components only accounting for approximately 22% of the pangenome (when the mobilome is included) and for 26.5% when the mobilome is excluded (Table S12). This variation, coupled with the non-specific functional enrichments observed within the pangenome (Table S13), highlights the rampant evolutionary capabilities in 'fixed' genomic biomolecules. This variation may be partially explained by the rampant recombination observed throughout CC17 or the diversity of prophages observed throughout the respective genomes, but more research is needed to confirm the role of phages on VRE genomes (Figs 11–13 and S3). As observed for phage data, integrons were observed throughout CC17 (Table S19 and Fig. 15). When integrons were observed in phylogenetic blocks, genes-of-interest were observed (Table S20 and Fig. 15). However, as these genes are widely observed in *E. faecium*, it is possible that these integrons are faciliting mobilization of these genes. While recombination has been reported in *E. faecium* [200, 201], the extent of recombination observed during this study has not, to our knowledge, been previously reported. This recombination may have been a driving force in the evolution of *E. faecium* CC17 towards pathogenicity [202]. Despite the rampant variation observed between genomes used in this study, chromosomal size, gene content, gene density and the phylogenetic proximity of each genome remained largely stable throughout the dataset. The relatedness of the VSE and VRE indicates that vancomycin-susceptible *Efm* are transferring between patients and then acquiring plasmids to become VR*Efm* rather than or in addition to the movement of VRE between patients. The stability of *Efm* within Ireland over the 20 years indicates that the increases in VRE across patients is not due to frequent and multiple introductions of different *Efm* into the hospital but rather the maintenance of genomic-related susceptible *Efm* that acquire vancomycin resistance plasmids and the spread of a cohort of highly related VR*Efm*. Thus, the susceptible and resistant *Efm* are highly related and fixed within the hospital patients. To tackle VRE within Ireland we need to focus on reducing VSE in addition to reducing VR*Efm*. This requires analysis of the epidemiology of VSE, VR*Efm* and the plasmids containing the *vanA* gene. While this study provides some understanding of the past, we need to perform studies on the VS*Efm* and VR*Efm* within our hospitals to understand the continuing rise of VR*Efm* and the infection control required to limit the spread.

CONCLUSION

We have demonstrated that the genomic characteristics of the studied VS*Efm* and VR*Efm* (such as genome size, gene density and gene count) are relatively stable across space and time despite an open and plastic pangenome. The relatedness of the VS*Efm* and VR*Efm* across time indicates that to reduce or remove VR*Efm* from Irish hospitals we must concurrently reduce VS*Efm*. The problem of vancomycin resistance within these pathogens is plasmid-mediated *vanA* rather than chromosomally mediated *vanB*.

Funding information
This work received no specific grant from any funding agency.

Acknowledgements
We wish to thank the staff in the microbiology laboratories at MMUH and the BSAC surveillance team for making their data available.

Author contributions
Conceptualization R.L., B.L., F.W.; Methodology R.L., F.W.; Software R.L.; Validation R.L.; Formal Analysis R.L.; Investigation R.L., C.M., R.M.; Resources B.L., F.W; Data Curation R.L.; Writing – Original Draft Preparation R.L., F.W.; Writing – Review and Editing R.L., F.W.; Visualization R.L.; Supervision F.W.; Project Administration F.W.; Funding F.W.

Conflicts of interest
The author(s) declare that there are no conflicts of interest.

Ethical statement
Ethical approval was sought from, and approved by, the Mater Misericordiae University Hospital Institutional Review Board (Reference: 1/378/2134).

References

1. Zhou X, Willems RJL, Friedrich AW, Rossen JWA, Bathoorn E. *Enterococcus faecium*: from microbiological insights to practical recommendations for infection control and diagnostics. *Antimicrob Resist Infect Control* 2020;9:130.
2. Howden BP, Holt KE, Lam MMC, Seemann T, Ballard S, *et al*. Genomic insights to control the emergence of vancomycin-resistant *Enterococci*. *mBio* 2013;4:e00412-13.
3. Giannakopoulos X, Sakkas H, Ragos V, Tsiambas E, Bozidis P, *et al*. Impact of *Enterococcal* urinary tract infections in immunocompromised - neoplastic patients. *J BUON* 2019;24:1768–1775.
4. Chilambi GS, Nordstrom HR, Evans DR, Ferrolino JA, Hayden RT, *et al*. Evolution of vancomycin-resistant *Enterococcus faecium* during colonization and infection in immunocompromised pediatric patients. *Proc Natl Acad Sci U S A* 2020;117:11703–11714.
5. Alghamdi F, Shakir M. The influence of *Enterococcus faecalis* as a dental root canal pathogen on endodontic treatment: a systematic review. *Cureus* 2020;12:e7257.
6. Higuita NIA, Huycke MM. *Enterococcal* Disease, epidemiology, and implications for treatment. *Enterococci* from commensals to lead causes drug resist infect. Internet; 2014 Feb 4. https://www.ncbi.nlm.nih.gov/books/NBK190429 [accessed 17 August 2021].
7. Tortosa JA, Hernández-Palazón J. *Enterococcus faecalis* meningitis after spinal anesthesia. *Anesthesiology* 2000;92:909.
8. Scapellato PG, Ormazabal C, Scapellato JL, Bottaro EG, Rodríguez Brieschke MT. Meningitis due to vancomycin-resistant *Enterococcus faecium* successfully treated with combined intravenous and intraventricular chloramphenicol. *J Clin Microbiol* 2005;43:3578–3579.
9. Van Tyne D, Gilmore MS. Friend turned foe: evolution of *Enterococcal virulence* and antibiotic resistance. *Annu Rev Microbiol* 2014;68:337–356.
10. Tsoulas C, Nathwani D. Review of meta-analyses of vancomycin compared with new treatments for gram-positive skin and soft-tissue infections: are we any clearer? *Int J Antimicrob Agents* 2015;46:1–7.
11. Liu C, Bayer A, Cosgrove SE, Daum RS, Fridkin SK, *et al*. Clinical practice guidelines by the infectious diseases society of america for the treatment of methicillin-resistant *Staphylococcus aureus* infections in adults and children. *Clin Infect Dis* 2011;52:e18–55.
12. Munita JM, Arias CA, Murray BE. *Enterococcal endocarditis*: can we win the war? *Curr Infect Dis Rep* 2012;14:339–349.
13. Top J, Sinnige JC, Brouwer EC, Werner G, Corander J, *et al*. Identification of a novel genomic island associated with *vanD*-type vancomycin resistance in six dutch vancomycin-resistant *Enterococcus faecium* isolates. *Antimicrob Agents Chemother* 2018;62:e01793-17.
14. Hayakawa K, Marchaim D, Palla M, Gudur UM, Pulluru H, *et al*. Epidemiology of vancomycin-resistant *Enterococcus faecalis*: a case-case-control study. *Antimicrob Agents Chemother* 2013;57:49–55.
15. Machado H, Seif Y, Sakoulas G, Olson CA, Hefner Y, *et al*. Environmental conditions dictate differential evolution of vancomycin resistance in *Staphylococcus aureus*. *Commun Biol* 2021;4:793.
16. Patel R. Clinical impact of vancomycin-resistant *Enterococci*. *J Antimicrob Chemother* 2003;51 Suppl 3:iii13–21.
17. Correa-Martinez CL, Tönnies H, Froböse NJ, Mellmann A, Kampmeier S. Transmission of vancomycin-resistant *Enterococci* in the hospital setting: uncovering the patient-environment interplay. *Microorganisms* 2020;8:E203.
18. Courvalin P. Vancomycin resistance in gram-positive cocci. *Clin Infect Dis* 2006;42 Suppl 1:S25–34.
19. Arias CA, Contreras GA, Murray BE. Management of multidrug-resistant *Enterococcal* infections. *Clin Microbiol Infect* 2010;16:555–562.
20. O'Driscoll T, Crank CW. Vancomycin-resistant *Enterococcal* infections: epidemiology, clinical manifestations, and optimal management. *Infect Drug Resist* 2015;8:217–230.
21. Leavis HL, Bonten MJM, Willems RJL. Identification of high-risk *Enterococcal* clonal complexes: global dispersion and antibiotic resistance. *Curr Opin Microbiol* 2006;9:454–460.
22. Getachew Y, Hassan L, Zakaria Z, Abdul Aziz S. Genetic variability of vancomycin-resistant *Enterococcus faecium* and *Enterococcus faecalis* isolates from humans, chickens, and pigs in Malaysia. *Appl Environ Microbiol* 2013;79:4528–4533.
23. Leclercq R, Derlot E, Duval J, Courvalin P. Plasmid-mediated resistance to vancomycin and teicoplanin in *Enterococcus faecium*. *N Engl J Med* 1988;319:157–161.
24. Yim J, Smith JR, Rybak MJ. Role of combination antimicrobial therapy for vancomycin-resistant *Enterococcus faecium* infections: review of the urrent Evidence. *Pharmacotherapy* 2017;37:579–592.
25. Zhong Z, Zhang W, Song Y, Liu W, Xu H, *et al*. Comparative genomic analysis of the genus *Enterococcus*. *Microbiol Res* 2017;196:95–105.
26. Zhong Z, Kwok L-Y, Hou Q, Sun Y, Li W, *et al*. Comparative genomic analysis revealed great plasticity and environmental adaptation of the genomes of *Enterococcus faecium*. *BMC Genomics* 2019;20:602.
27. Sadowy E, Gawryszewska I, Kuch A, Żabicka D, Hryniewicz W. The changing epidemiology of VanB *Enterococcus faecium* in Poland. *Eur J Clin Microbiol Infect Dis* 2018;37:927–936.
28. Paoletti C, Foglia G, Principalli MS, Magi G, Guaglianone E, *et al*. Co-transfer of vanA and aggregation substance genes from *Enterococcus faecalis* isolates in intra- and interspecies matings. *J Antimicrob Chemother* 2007;59:1005–1009.
29. Handwerger S, Pucci MJ, Volk KJ, Liu J, Lee MS. The cytoplasmic peptidoglycan precursor of vancomycin-resistant *Enterococcus faecalis* terminates in lactate. *J Bacteriol* 1992;174:5982–5984.
30. Dutka-Malen S, Molinas C, Arthur M, Courvalin P. The VANA glycopeptide resistance protein is related to D-alanyl-D-alanine ligase cell wall biosynthesis enzymes. *Mol Gen Genet* 1990;224:364–372.
31. Roper DI, Huyton T, Vagin A, Dodson G. The molecular basis of vancomycin resistance in clinically relevant *Enterococci:* crystal structure of D-alanyl-D-lactate ligase (VanA). *Proc Natl Acad Sci U S A* 2000;97:8921–8925.
32. Clark NC, Cooksey RC, Hill BC, Swenson JM, Tenover FC. Characterization of glycopeptide-resistant enterococci from U.S. hospitals. *Antimicrob Agents Chemother* 1993;37:2311–2317.
33. Arthur M, Courvalin P. Genetics and mechanisms of glycopeptide resistance in *Enterococci*. *Antimicrob Agents Chemother* 1993;37:1563–1571.
34. Park IJ, Lee WG, Shin JH, Lee KW, Woo GJ. VanB phenotype-vanA genotype *Enterococcus faecium* with heterogeneous expression of teicoplanin resistance. *J Clin Microbiol* 2008;46:3091–3093.
35. Turner RJ, Huang L-N, Viti C, Mengoni A. Metal-resistance in bacteria: why care? *Genes* 2020;11:E1470.
36. Chen J, Li J, Zhang H, Shi W, Liu Y. Bacterial heavy-metal and antibiotic resistance genes in a copper tailing dam area in northern China. *Front Microbiol* 2019;10:1916.
37. Li L-G, Xia Y, Zhang T. Co-occurrence of antibiotic and metal resistance genes revealed in complete genome collection. *ISME J* 2017;11:651–662.
38. Liu H, Li M, Dao TD, Liu Y, Zhou W, *et al*. Design of PdAu alloy plasmonic nanoparticles for improved catalytic performance in CO2 reduction with visible light irradiation. *Nano Energy* 2016;26:398–404.
39. Maillard J-Y. Antimicrobial biocides in the healthcare environment: efficacy, usage, policies, and perceived problems. *Ther Clin Risk Manag* 2021;307.

40. Wilkinson LJ, White RJ, Chipman JK. Silver and nanoparticles of silver in wound dressings: a review of efficacy and safety. *J Wound Care* 2011;20:543–549.
41. Grass G, Rensing C, Solioz M. Metallic copper as an antimicrobial surface. *Appl Environ Microbiol* 2011;77:1541–1547.
42. Lebreton F, van Schaik W, Sanguinetti M, Posteraro B, Torelli R, *et al*. AsrR is an oxidative stress sensing regulator modulating *Enterococcus faecium* opportunistic traits, antimicrobial resistance, and pathogenicity. *PLoS Pathog* 2012;8:e1002834.
43. Hasman H, Aarestrup FM. tcrB, a gene conferring transferable copper resistance in *Enterococcus faecium:* occurrence, transferability, and linkage to macrolide and glycopeptide resistance. *Antimicrob Agents Chemother* 2002;46:1410–1416.
44. Pidot S, Gao W, Buultjens A, Monk I, Guerillot R, *et al*. Increasing tolerance of hospital *Enterococcus faecium* to hand-wash alcohols. increasing toler hosp *Enterococcus faecium* to handwash alcohols. *Sci Transl Med* 2016:053728.
45. Comerlato CB, Resende MCC de, Caierão J, d'Azevedo PA. Presence of virulence factors in *Enterococcus faecalis* and *Enterococcus faecium* susceptible and resistant to vancomycin. *Mem Inst Oswaldo Cruz* 2013;108:590–595.
46. Haghi F, Lohrasbi V, Zeighami H. High incidence of virulence determinants, aminoglycoside and vancomycin resistance in *Enterococci* isolated from hospitalized patients in Northwest Iran. *BMC Infect Dis* 2019;19:744.
47. Biswas PP, Dey S, Adhikari L, Sen A. Virulence markers of vancomycin resistant *Enterococci* isolated from infected and colonized patients. *J Glob Infect Dis* 2014;6:157–163.
48. Kim EB, Marco ML. Nonclinical and clinical *Enterococcus faecium* strains, but not *Enterococcus faecalis* strains, have distinct structural and functional genomic features. *Appl Environ Microbiol* 2014;80:154–165.
49. van Hal SJ, Willems RJL, Gouliouris T, Ballard SA, Coque TM, *et al*. The global dissemination of hospital clones of *Enterococcus faecium*. *Genome Med* 2021;13:52.
50. Banfield JF, Young M. Microbiology. variety--the splice of life--in microbial communities. *Science* 2009;326:1198–1199.
51. Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. *FEMS Microbiol Rev* 2014;38:916–931.
52. Broniewski JM, Meaden S, Paterson S, Buckling A, Westra ER. The effect of phage genetic diversity on bacterial resistance evolution. *ISME J* 2020;14:828–836.
53. Bonilla N, Santiago T, Marcos P, Urdaneta M, Domingo JS, *et al*. *Enterophages*, a group of phages infecting *Enterococcus faecalis*, and their potential as alternate indicators of human faecal contamination. *Water Sci Technol* 2010;61:293–300.
54. Lee D, Im J, Na H, Ryu S, Yun C-H, *et al*. The novel *Enterococcus* Phage vB_EfaS_HEf13 has broad lytic activity against clinical isolates of *Enterococcus faecalis*. *Front Microbiol* 2019;10:2877.
55. Wandro S, Oliver A, Gallagher T, Weihe C, England W, *et al*. Predictable molecular adaptation of coevolving *Enterococcus faecium* and lytic phage EfV12-phi1. *Front Microbiol* 2018;9:3192.
56. Whelan FJ, Rusilowicz M, McInerney JO. Coinfinder: detecting significant associations and dissociations in pangenomes. *Microb Genom* 2020;6.
57. Raven KE, Reuter S, Gouliouris T, Reynolds R, Russell JE, *et al*. Genome-based characterization of hospital-adapted *Enterococcus faecalis* lineages. *Nat Microbiol* 2016;1:15033.
58. Leinonen R, Sugawara H, Shumway M. The sequence read archive. *Nucleic Acids Res* 2011;39:D19–21.
59. Krueger F. Babraham bioinformatics - Trim Galore. Internet 2012; 2021 Jan 16. https://www.bioinformatics.babraham.ac.uk/projects/trim_galore/
60. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet j* 2011;17:10.
61. Andrews S, Krueger F, Seconds-Pichon A, Biggins F, FastQC WS. A quality control tool for high throughput sequence data Babraham Bioinformatics. 1(1):undefined-undefined; 2015. https://www.mendeley.com/catalogue/8057171e-e700-36a0-b936-1c307058462d/?utm_source=desktop&utm_medium=1.19.8&utm_campaign=open_catalog&userDocumentId=%7B124edc8f-5459-3c1b-a449-8ab9bc3e95f8%7D [accessed 18 August 2021].
62. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with multiplex MinION sequencing. *Microb Genom* 2017;3:e000132.
63. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, *et al*. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* 2012;19:455–477.
64. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012;9:357–359.
65. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, *et al*. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 2014;9:e112963.
66. Camacho C, Coulouris G, Avagyan V, Ma N, Papadopoulos J, *et al*. BLAST+: architecture and applications. *BMC Bioinformatics* 2009;10:421.
67. Altschul SSF, Gish W, Miller W, Myers EEW, Lipman DJD. Basic local alignment search tool. *J Mol Biol* 1990;215:403–410.
68. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, *et al*. The sequence alignment/Map format and SAMtools. *Bioinform* 2009;25:2078–2079.
69. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 2015;25:1043–1055.
70. Seemann T. GitHub - tseemann/mlst: Scan contig files against PubMLST typing schemes. Internet 2014; 2021. https://github.com/tseemann/mlst
71. Jolley KA, Maiden MCJ. BIGSdb: calable analysis of bacterial genome variation at the population level. *BMC Bioinform* 2010;11:595.
72. Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res* 2018;3:124.
73. Schwengers O, Barth P, Falgenhauer L, Hain T, Chakraborty T, *et al*. Platon: identification and characterization of bacterial plasmid contigs in short-read draft assemblies exploiting protein sequence-based replicon distribution scores. *Microb Genom* 2020;6.
74. Angelopoulou A, Warda AK, O'Connor PM, Stockdale SR, Shkoporov AN, *et al*. Diverse bacteriocins produced by strains from the human milk microbiota. *Front Microbiol* 2020;11:788.
75. Ondov BD, Treangen TJ, Melsted P, Mallonee AB, Bergman NH, *et al*. Mash: fast genome and metagenome distance estimation using minhash. *Genome Biol* 2016;17:132.
76. Galata V, Fehlmann T, Backes C, Keller A. PLSDB: a resource of complete bacterial plasmids. *Nucleic Acids Res* 2019;47:D195–D202.
77. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinf* 2014;30:2068–2069.
78. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, *et al*. Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinform* 2010;11:111.
79. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res* 2004;32:11–16.
80. SkennertonC. minced: Mining CRISPRs in Environmental Datasets. Internet; 2013. https://github.com/ctSkennerton/minced [accessed 18 August 2021].
81. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat Methods* 2011;8:785–786.

82. Finn RD, Clements J, Eddy SR. HMMER web server: interactive sequence similarity searching. *Nucleic Acids Res* 2011;39:W29–37.
83. Stajich JE, Block D, Boulez K, Brenner SE, Chervitz SA, *et al*. The bioperl toolkit: perl modules for the life sciences. *Genome Res* 2002;12:1611–1618.
84. Seemann T. barrnap: Bacterial ribosomal RNA predictor. Internet; 2013. https://github.com/tseemann/barrnap [accessed 18 August 2021].
85. Jones P, Binns D, Chang H-Y, Fraser M, Li W, *et al*. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 2014;30:1236–1240.
86. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, *et al*. The Pfam protein families database. *Nucleic Acids Res* 2004;32:D138–41.
87. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, *et al*. Gene ontology: tool for the unification of biology. *Nat Genet* 2000;25:25–29.
88. Carroll LM, Larralde M, Fleck JS, Ponnudurai R, Milanese A. Accurate *de novo* identification of biosynthetic gene clusters with GECCO. *Bioinform* 2000. DOI: 10.1101/2021.05.03.442509.
89. Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, *et al*. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res* 2019;47:W81–W87.
90. Johansson MHK, Bortolaia V, Tansirichaiya S, Aarestrup FM, Roberts AP, *et al*. Detection of mobile genetic elements associated with antibiotic resistance in *Salmonella enterica* using a newly developed web tool: mobile element finder. *J Antimicrob Chemother* 2021;76:101–109.
91. Seemann T. ABRicate: mass screening of contigs for antimicrobial resistance or virulence genes. Internet; 2014. https://github.com/tseemann/abricate/commits/master?after=955d402a23371a61bfca48f1b9e0d30ac724e6ef+209&branch=master [accessed 18 August 2021].
92. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Bouchard M, *et al*. CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. *Nucleic Acids Res* 2020;48:D517–D525.
93. Zankari E, Allesøe R, Joensen KG, Cavaco LM, Lund O, *et al*. PointFinder: a novel web tool for WGS-based detection of antimicrobial resistance associated with chromosomal point mutations in bacterial pathogens. *J Antimicrob Chemother* 2017;72:2764–2768.
94. Chen L, Yang J, Yu J, Yao Z, Sun L, *et al*. VFDB: a reference database for bacterial virulence factors. *Nucleic Acids Res* 2005;33:D325–8.
95. Pal C, Bengtsson-Palme J, Rensing C, Kristiansson E, Larsson DGJ. BacMet: antibacterial biocide and metal resistance genes database. *Nucleic Acids Res* 2014;42:D737–43.
96. Rice P, Longden I, Bleasby A. EMBOSS: the european molecular biology open software suite. *Trends Genet* 2000;16:276–277.
97. WELCH BL. The generalisation of student's problems when several different population variances are involved. *Biometrika* 1947;34:28–35.
98. Student. The probable error of a mean. *Biometrika* 1908;6:1.
99. Bonferroni C. Teoria statistica delle classi e calcolo delle probabilita. Pubbl del R Ist uper di sci econ e commericiali di Firenze. Internet; 1936. https://ci.nii.ac.jp/naid/20001561442 [accessed 20 April 2021].
100. Dunn OJ. Multiple comparisons among means. *J Am Stat Assoc* 1961;56:52–64.
101. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers E-J, *et al*. Redefine statistical significance. *Nat Hum Behav* 2018;2:6–10.
102. Kolmogorov AN. Sulla determinazione empirica di una lgge di distribuzione – ScienceOpen Inst Ital Attuari, Giorn. Internet 1933; 2021 May 5. https://www.scienceopen.com/document?vid=c3c08573-63b2-4153-a72e-97bd1b3663a0
103. Lilliefors HW. On the Kolmogorov-smirnov test for normality with mean and variance unknown. *J Am Stat Assoc* 1967;62:399–402.
104. Levene H. Robust tests for equality of variances. In: Olkin I (eds). *Contributions to Probability and Statistics*. Stanford University Press, Palo Alto; 1960. pp. 278–292.
105. Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. in: international AAAI conference on weblogs and social media. 2009.
106. Fruchterman TMJ, Reingold EM. Graph drawing by force-directed placement. *Softw: Pract Exper* 1991;21:1129–1164.
107. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, *et al*. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* 2015;31:3691–3693.
108. Löytynoja A. Phylogeny-aware alignment with PRANK [Internet] 2014. *Methods Mol Biol* 2020.
109. Fisher RA. On the "Probable Error" of a coefficient of correlation deduced from a small sample. *Metron* 1921;1:3–32.
110. Klopfenstein DV, Zhang L, Pedersen BS, Ramírez F, Warwick Vesztrocy A, *et al*. GOATOOLS: a python library for gene ontology analyses. *Sci Rep* 2018;8:10872.
111. Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 2009;25:1972–1973.
112. Kück P, Meusemann K. FASconCAT: Convenient handling of data matrices. *Mol Phylogenet Evol* 2010;56:1115–1118.
113. Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol Biol Evol* 2015;32:268–274.
114. Darriba D, Posada D, Kozlov AM, Stamatakis A, Morel B, *et al*. ModelTest-NG: a new and scalable tool for the selection of DNA and protein evolutionary models. *Mol Biol Evol* 2020;37:291–294.
115. Tavare S. Some probabilistic and statistical problems in the analysis of DNA sequences Lect Math life. Internet 1986; 2021 Aug 18. https://books.google.com/books?hl=en&lr=&id=8al1phhOKhgC&oi=fnd&pg=PA57&ots=roMG4RHdKi&sig=csERmeVhWARlEZ2mHiyf3_zHGSA
116. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, *et al*. Proteinortho: detection of (co-)orthologs in large-scale analysis. *BMC Bioinform* 2011;12:124.
117. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–1797.
118. Le SQ, Gascuel O. An improved general amino acid replacement matrix. *Mol Biol Evol* 2008;25:1307–1320.
119. 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–W296.
120. Cheng L, Connor TR, Sirén J, Aanensen DM, Corander J. Hierarchical and spatially explicit clustering of DNA sequences with BAPS software. *Mol Biol Evol* 2013;30:1224–1228.
121. Tonkin-Hill G, Lees JA, Bentley SD, Frost SDW, Corander J. RhierBAPS: an R implementation of the population clustering algorithm hierBAPS. *Wellcome Open Res* 2018;3:93.
122. Silva M, Machado MP, Silva DN, Rossi M, Moran-Gilad J, *et al*. chewBBACA: a complete suite for gene-by-gene schema creation and strain identification. *Microb Genom* 2018;4.
123. Zhou Z, Alikhan N-F, Sergeant MJ, Luhmann N, Vaz C, *et al*. GrapeTree: visualization of core genomic relationships among 100,000 bacterial pathogens. *Genome Res* 2018;28:1395–1404.
124. Abudahab K, Prada JM, Yang Z, Bentley SD, Croucher NJ, *et al*. PANINI: pangenome neighbour identification for bacterial populations. *Microb Genom* 2019;5.
125. Argimón S, Abudahab K, Goater RJE, Fedosejev A, Bhai J, *et al*. Microreact: visualizing and sharing data for genomic epidemiology and phylogeography. *Microb Genom* 2016;2:e000093.
126. Croucher NJ, Page AJ, Connor TR, Delaney AJ, Keane JA, *et al*. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res* 2015;43:e15.

127. Hadfield J, Croucher NJ, Goater RJ, Abudahab K, Aanensen DM, *et al*. Phandango: an interactive viewer for bacterial population genomics. *Bioinformatics* 2018;34:292–293.
128. Starikova EV, Tikhonova PO, Prianichnikov NA, Rands CM, Zdobnov EM, *et al*. Phigaro: high-throughput prophage sequence annotation. *Bioinformatics* 2020;36:3882–3884.
129. Cury J, Jové T, Touchon M, Néron B, Rocha EP. Identification and analysis of integrons and cassette arrays in bacterial genomes. *Nucleic Acids Res* 2016;44:4539–4550.
130. Singh KV, Malathum K, Murray BE. Disruption of an *Enterococcus faecium* species-specific gene, a homologue of acquired macrolide resistance genes of *Staphylococci*, is associated with an increase in macrolide susceptibility. *Antimicrob Agents Chemother* 2001;45:263–266.
131. Costa Y, Galimand M, Leclercq R, Duval J, Courvalin P. Characterization of the chromosomal aac(6')-Ii gene specific for *Enterococcus faecium*. *Antimicrob Agents Chemother* 1993;37:1896–1903.
132. Draker K, Northrop DB, Wright GD. Kinetic mechanism of the GCN5-related chromosomal aminoglycoside acetyltransferase AAC(6')-Ii from *Enterococcus faecium*: evidence of dimer subunit cooperativity. *Biochemistry* 2003;42:6565–6574.
133. Wybenga-Groot LE, Draker K, Wright GD, Berghuis AM. Crystal structure of an aminoglycoside 6'-N-acetyltransferase: defining the GCN5-related N-acetyltransferase superfamily fold. *Structure* 1999;7:497–507.
134. Guzmán Prieto AM, Wijngaarden J, Braat JC, Rogers MRC, Majoor E, *et al*. The two-component system ChtRS contributes to chlorhexidine tolerance in *Enterococcus faecium*. *Antimicrob Agents Chemother* 2017;61:e02122-16.
135. Fujishima K, Kawada-Matsuo M, Oogai Y, Tokuda M, Torii M, *et al*. dpr and sod in *Streptococcus mutans* are involved in coexistence with S. sanguinis, and PerR is associated with resistance to H2O2. *Appl Environ Microbiol* 2013;79:1436–1443.
136. Yamamoto Y, Poole LB, Hantgan RR, Kamio Y. An iron-binding protein, Dpr, from *Streptococcus mutans* prevents iron-dependent hydroxyl radical formation in vitro. *J Bacteriol* 2002;184:2931–2939.
137. Hasman H, Kempf I, Chidaine B, Cariolet R, Ersbøll AK, *et al*. Copper resistance in *Enterococcus faecium,* mediated by the tcrB gene, is selected by supplementation of pig feed with copper sulfate. *Appl Environ Microbiol* 2006;72:5784–5789.
138. Fard RMN, Heuzenroeder MW, Barton MD. Antimicrobial and heavy metal resistance in commensal *Enterococci* isolated from pigs. *Vet Microbiol* 2011;148:276–282.
139. Horsburgh MJ, Clements MO, Crossley H, Ingham E, Foster SJ. PerR controls oxidative stress resistance and iron storage proteins and is required for virulence in *Staphylococcus aureus*. *Infect Immun* 2001;69:3744–3754.
140. Møller AK, Barkay T, Hansen MA, Norman A, Hansen LH, *et al*. Mercuric reductase genes (merA) and mercury resistance plasmids in high arctic snow, freshwater and sea-ice brine. *FEMS Microbiol Ecol* 2014;87:52–63.
141. Zaheer R, Cook SR, Barbieri R, Goji N, Cameron A, *et al*. Surveillance of *Enterococcus* spp. reveals distinct species and antimicrobial resistance diversity across a one-health continuum. *Sci Reports* 2021.
142. Creti R, Koch S, Fabretti F, Baldassarri L, Huebner J. *Enterococcal* colonization of the gastro-intestinal tract: role of biofilm and environmental oligosaccharides. *BMC Microbiol* 2006;6:60.
143. Nallapareddy SR, Singh KV, Sillanpää J, Garsin DA, Höök M, *et al*. Endocarditis and biofilm-associated pili of *Enterococcus faecalis*. *J Clin Invest* 2006;116:2799–2807.
144. Dahl A, Rasmussen RV, Bundgaard H, Hassager C, Bruun LE, *et al*. *Enterococcus faecalis* infective endocarditis. *Circulation* 2021.
145. Matos RC, Lapaque N, Rigottier-Gois L, Debarbieux L, Meylheuc T, *et al*. *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. *PLoS Genet* 2013;9:e1003539.
146. Mayer C, Vocadlo DJ, Mah M, Rupitz K, Stoll D, *et al*. Characterization of a beta-N-acetylhexosaminidase and a beta-N-acetylglucosaminidase/beta-glucosidase from *Cellulomonas fimi*. *FEBS J* 2006;273:2929–2941.
147. Rivolta C, Soldo B, Lazarevic V, Joris B, Mauël C, *et al*. A 35.7 kb DNA fragment from the *Bacillus* subtilis chromosome containing a putative 12.3 kb operon involved in hexuronate catabolism and a perfectly symmetrical hypothetical catabolite-responsive element. *Microb* 1998;144 (Pt 4):877–884.
148. Phillips MK, Hederstedt L, Hasnain S, Rutberg L, Guest JR. Nucleotide sequence encoding the flavoprotein and iron-sulfur protein subunits of the *Bacillus* subtilis PY79 succinate dehydrogenase complex. *J Bacteriol* 1987;169:864–873.
149. Wipat A, Carter N, Brignell SC, Guy BJ, Piper K, *et al*. The dnaB-pheA (256 degrees-240 degrees) region of the *Bacillus* subtilis chromosome containing genes responsible for stress responses, the utilization of plant cell walls and primary metabolism. *Microbio* 1996;142 (Pt 11):3067–3078.
150. Mekjian KR, Bryan EM, Beall BW, Moran CP. Regulation of hexuronate utilization in *Bacillus* subtilis. *J Bacteriol* 1999;181:426–433.
151. Setlow B, Cabrera-Hernandez A, Cabrera-Martinez RM, Setlow P. Identification of aryl-phospho-beta-D-glucosidases in *Bacillus* subtilis. *Arch Microbiol* 2004;181:60–67.
152. Sadaie Y, Nakadate H, Fukui R, Yee LM, Asai K. Glucomannan utilization operon of *Bacillus* subtilis. *FEMS Microbiol Lett* 2008;279:103–109.
153. Wong MD, Lin Y-F, Rosen BP. The soft metal Ion binding sites in the *Staphylococcus aureus* pI258 CadC Cd(II)/Pb(II)/Zn(II)-responsive repressor are formed between subunits of the homodimer. *J Biol Chem* 2002;277:40930–40936.
154. Msadek T, Kunst F, Rapoport G. MecB of *Bacillus* subtilis, a member of the ClpC ATPase family, is a pleiotropic regulator controlling competence gene expression and growth at high temperature. *Proc Natl Acad Sci U S A* 1994;91:5788–5792.
155. Krüger E, Völker U, Hecker M. Stress induction of clpC in *Bacillus* subtilis and its involvement in stress tolerance. *J Bacteriol* 1994;176:3360–3367.
156. Leskelä S, Wahlström E, Kontinen VP, Sarvas M. Lipid modification of prelipoproteins is dispensable for growth but essential for efficient protein secretion in *Bacillus* subtilis: characterization of the lgt gene. *Mol Microbiol* 1999;31:1075–1085.
157. Harvie DR, Andreini C, Cavallaro G, Meng W, Connolly BA, *et al*. Predicting metals sensed by ArsR-SmtB repressors: allosteric interference by a non-effector metal. *Mol Microbiol* 2006;59:1341–1356.
158. Nicolaou SA, Fast AG, Nakamaru-Ogiso E, Papoutsakis ET. Overexpression of fetA (ybbL) and fetB (ybbM), encoding an iron exporter, enhances resistance to oxidative stress in *Escherichia coli*. *Appl Environ Microbiol* 2013;79:7210–7219.
159. Zheng R, Blanchard JS. Kinetic and mechanistic analysis of the E. coli panE-encoded ketopantoate reductase. *Biochemistry* 2000;39:3708–3717.
160. Fraser KR, Harvie D, Coote PJ, O'Byrne CP. Identification and characterization of an ATP binding cassette L-carnitine transporter in listeria monocytogenes. *Appl Environ Microbiol* 2000;66:4696–4704.
161. Huynh TN, Choi PH, Sureka K, Ledvina HE, Campillo J, *et al*. Cyclic di-AMP targets the cystathionine beta-synthase domain of the osmolyte transporter OpuC. *Mol Microbiol* 2016;102:233–243.
162. Pruteanu M, Baker TA. Controlled degradation by ClpXP protease tunes the levels of the excision repair protein UvrA to the extent of DNA damage. *Mol Microbiol* 2009;71:912–924.
163. Myles GM, Sancar A. Isolation and characterization of functional domains of UvrA. *Biochemistry* 1991;30:3834–3840.
164. Cooper RA. The utilisation of D-galactonate and D-2-oxo-3-deoxygalactonate by *Escherichia coli* K-12. biochemical and genetical studies. *Arch Microbiol* 1978;118:199–206.

165. LowKam C, Liotard B, Sygusch J. Structure of a class I tagatose-1,6-bisphosphate aldolase: investigation into an apparent loss of stereospecificity. *J Biol Chem* 2010;285:21143–21152.
166. Erni B, Zanolari B, Kocher HP. The mannose permease of *Escherichia coli* consists of three different proteins. amino acid sequence and function in sugar transport, sugar phosphorylation, and penetration of phage lambda DNA. *J Biol Chem* 1987;262:5238–5247.
167. Stolz B, Huber M, Marković-Housley Z, Erni B. The mannose transporter of *Escherichia coli*. Structure and function of the IIABMan subunit. *J Biol Chem* 1993;268:27094–27099.
168. Newton GL, Koledin T, Gorovitz B, Rawat M, Fahey RC, *et al*. The glycosyltransferase gene encoding the enzyme catalyzing the first step of mycothiol biosynthesis (mshA). *J Bacteriol* 2003;185:3476–3479.
169. Schöck F, Dahl MK. Analysis of DNA flanking the treA gene of *Bacillus* subtilis reveals genes encoding a putative specific enzyme IITre and a potential regulator of the trehalose operon. *Gene* 1996;175:59–63.
170. Kearns DB, Chu F, Branda SS, Kolter R, Losick R. A master regulator for biofilm formation by *Bacillus* subtilis. *Mol Microbiol* 2005;55:739–749.
171. S-Nogueira I, Nogueira TV, Soares S, de Lencastre H. The *Bacillus subtilis* L-arabinose (ara) operon: nucleotide sequence, genetic organization and expression. *Microbiology*1997;143 (Pt 3):957–969.
172. Sirko A, Hryniewicz M, Hulanicka D, Böck A. Sulfate and thiosulfate transport in *Escherichia coli* K-12: nucleotide sequence and expression of the cysTWAM gene cluster. *J Bacteriol* 1990;172:3351–3357.
173. Chowdhury N, Norris J, McAlister E, Lau SYK, Thomas GH, *et al*. The VC1777-VC1779 proteins are members of a sialic acid-specific subfamily of TRAP transporters (SiaPQM) and constitute the sole route of sialic acid uptake in the human pathogen Vibrio cholerae. *Microbiol* 2012;158:2158–2167.
174. Mathiopoulos C, Mueller JP, Slack FJ, Murphy CG, Patankar S, *et al*. A *Bacillus subtilis* dipeptide transport system expressed early during sporulation. *Mol Microbiol* 1991;5:1903–1913.
175. Perego M, Higgins CF, Pearce SR, Gallagher MP, Hoch JA. The oligopeptide transport system of *Bacillus subtilis* plays a role in the initiation of sporulation. *Mol Microbiol* 1991;5:173–185.
176. Koide A, Hoch JA. Identification of a second oligopeptide transport system in *Bacillus subtilis* and determination of its role in sporulation. *Mol Microbiol* 1994;13:417–426.
177. Parra-Lopez C, Baer MT, Groisman EA. Molecular genetic analysis of a locus required for resistance to antimicrobial peptides in *Salmonella typhimurium*. *EMBO J* 1993;12:4053–4062.
178. Breazeale SD, Ribeiro AA, Raetz CRH. Origin of lipid a species modified with 4-AMINO-4-DEOXY-L-ARABINOSE in polymyxin-resistant mutants of *Escherichia coli*. AN AMINOTRANSFERASE (ArnB) THAT GENERATES UDP-4-deoxyl-L-ARABINOSE. *J Biol Chem* 2003;278:24731–24739.
179. Yan A, Guan Z, Raetz CRH. An undecaprenyl phosphate-aminoarabinose flippase required for polymyxin resistance in *Escherichia coli*. *J Biol Chem* 2007;282:36077–36089.
180. Liu Y-Y, Wang Y, Walsh TR, Yi L-X, Zhang R, *et al*. Emergence of plasmid-mediated colistin resistance mechanism MCR-1 in animals and human beings in China: a microbiological and molecular biological study. *Lancet Infect Dis* 2016;16:161–168.
181. Alonso A, Sanchez P, Martínez JL. *Stenotrophomonas maltophilia* D457R contains a cluster of genes from gram-positive bacteria involved in antibiotic and heavy metal resistance. *Antimicrob Agents Chemother* 2000;44:1778–1782.
182. Humphrey S, Fillol-Salom A, Quiles-Puchalt N, Ibarra-Chávez R, Haag AF, *et al*. Bacterial chromosomal mobility via lateral transduction exceeds that of classical mobile genetic elements. *Nat Commun* 2021;12:6509.
183. Chiang YN, Penadés JR, Chen J. Genetic transduction by phages and chromosomal islands: The new and noncanonical. *PLoS Pathog* 2019;15:e1007878.
184. Royal College of Physicians I, Health Service Executive I. Guidelines for the prevention and control of MDRO excluding MRSA in the healthcare setting. *J R Coll Physicians Edinb* 2012.
185. Wein T, Wang Y, Barz M, Stücker FT, Hammerschmidt K, *et al*. Essential gene acquisition destabilizes plasmid inheritance. *PLoS Genet* 2021;17:e1009656.
186. Alice AF, López CS, Crosa JH. Plasmid- and chromosome-encoded redundant and specific functions are involved in biosynthesis of the siderophore anguibactin in *Vibrio anguillarum* 775: a case of chance and necessity? *J Bacteriol* 2005;187:2209–2214.
187. Zheng J, Guan Z, Cao S, Peng D, Ruan L, *et al*. Plasmids are vectors for redundant chromosomal genes in the *Bacillus cereus* group. *BMC Genomics* 2015;16:6.
188. Carroll AC, Wong A. Plasmid persistence: costs, benefits, and the plasmid paradox. *Can J Microbiol* 2018;64:293–304.
189. Baker-Austin C, Wright MS, Stepanauskas R, McArthur JV. Co-selection of antibiotic and metal resistance. *Trends Microbiol* 2006;14:176–182.
190. Bae T, Baba T, Hiramatsu K, Schneewind O. Prophages of *Staphylococcus aureus* Newman and their contribution to virulence. *Mol Microbiol* 2006 Nov;62:1035–1047.
191. Bae T, Baba T, Hiramatsu K, Schneewind O. Prophages of *Staphylococcus aureus* newman and their contribution to virulence. *Mol Microbiol* 2006;62:1035–1047.
192. Xia G, Wolz C. Phages of *Staphylococcus aureus* and their impact on host evolution. *Infect Genet Evol* 2014;21:593–601.
193. Pasek S, Risler J-L, Brézellec P. Gene fusion/fission is a major contributor to evolution of multi-domain bacterial proteins. *Bioinformatics* 2006;22:1418–1423.
194. Des Marais DL, Rausher MD. Escape from adaptive conflict after duplication in an anthocyanin pathway gene. *Nature* 2008;454:762–765.
195. Rastogi S, Liberles DA. Subfunctionalization of duplicated genes as a transition state to neofunctionalization. *BMC Evol Biol* 2005;5:28.
196. Leonard G, Richards TA. Genome-scale comparative analysis of gene fusions, gene fissions, and the fungal tree of life. *Proc Natl Acad Sci U S A* 2012;109:21402–21407.
197. Conant GC, Birchler JA, Pires JC. Dosage, duplication, and diploidization: clarifying the interplay of multiple models for duplicate gene evolution over time. *Curr Opin Plant Biol* 2014;19:91–98.
198. Wang W, Yu H, Long M. Duplication-degeneration as a mechanism of gene fission and the origin of new genes in *Drosophila species*. *Nat Genet* 2004;36:523–527.
199. Snel B, Bork P, Huynen M. Genome evolution. gene fusion versus gene fission. *Trends Genet* 2000;16:9–11.
200. de Been M, van Schaik W, Cheng L, Corander J, Willems RJ. Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. *Genome Biol Evol* 2013;5:1524–1535.
201. Leavis HL, Willems RJL, van Wamel WJB, Schuren FH, Caspers MPM, *et al*. Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of *E. faecium*. *PLoS Pathog* 2007;3:0075–0096.
202. Arenas M, Araujo NM, Branco C, Castelhano N, Castro-Nallar E, *et al*. Mutation and recombination in pathogen evolution: relevance, methods and controversies. *Infect Genet Evol* 2018;63:295–306.