


Bacteriocin production facilitates nosocomial emergence of vancomycin-resistant *Enterococcus faecium*

Received: 7 August 2024

Accepted: 14 February 2025

Published online: 21 March 2025



Emma G. Mills¹, Katharine Hewlett², Alexander B. Smith², Marissa P. Griffith^{1,3}, Lora Pless^{1,3}, Alexander J. Sundermann^{3,4}, Lee H. Harrison^{1,3,4}, Joseph P. Zackular^{2,5,6} & Daria Van Tyne^{1,7} 

Gastrointestinal colonization by the nosocomial pathogen vancomycin-resistant *Enterococcus faecium* (VREfm) can lead to bloodstream infections with high mortality rates. Shifts in VREfm lineages found within healthcare settings occur, but reasons underlying these changes are not understood. Here we sequenced 710 VREfm clinical isolates collected between 2017 and 2022 from a large tertiary care centre. Genomic analyses revealed a polyclonal VREfm population, although 46% of isolates formed genetically related clusters, suggesting a high transmission rate. Comparing these data to a global collection of 15,631 publicly available VREfm genomes collected between 2002 and 2022 identified replacement of the sequence type (ST) 17 VREfm lineage by emergent ST80 and ST117 lineages at the local and global level. Comparative genomic and functional analyses revealed that emergent lineages encoded bacteriocin T8, which conferred a competitive advantage over bacteriocin T8-negative strains in vitro and upon colonization of the mouse gut. Bacteriocin T8 carriage was also strongly associated with strain emergence in the global genome collection. These data suggest that bacteriocin T8-mediated competition may have contributed to VREfm lineage replacement.

Enterococcus faecium is a gastrointestinal tract commensal that can also cause serious infections, most commonly bloodstream and urinary tract infections, especially in immunocompromised and hospitalized patients¹. Hospitalized patients are often exposed to high levels of antibiotics, which decrease the diversity of commensals in the gastrointestinal (GI) tract and facilitate the overgrowth of multidrug-resistant organisms such as vancomycin-resistant *E. faecium* (VREfm)^{2–4}. VREfm overgrowth within the intestinal tract predisposes patients to invasive bloodstream infections^{2,4–7}. Furthermore, increased VREfm GI tract

burdens cause patients to shed VREfm into the environment, facilitating transmission to other patients mainly through the faecal–oral route^{8–10}.

Whole-genome sequencing facilitates the surveillance and characterization of VREfm population structure and transmission dynamics within healthcare settings. Multilocus sequence typing allows tracking of VREfm lineages both within and between healthcare facilities and on both local and global scales^{11–13}. Sequence types (STs) with similar genotypes, defined as having four or more identical loci, can be grouped into clonal complexes. VREfm lineages most often belong

¹Division of Infectious Diseases, University of Pittsburgh School of Medicine, Pittsburgh, PA, USA. ²Division of Protective Immunity, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ³Microbial Genomic Epidemiology Laboratory, Center for Genomic Epidemiology, University of Pittsburgh, Pittsburgh, PA, USA. ⁴Department of Epidemiology, School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA. ⁵Department of Pathology and Laboratory Medicine, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA. ⁶Center for Microbial Medicine, Children's Hospital of Philadelphia, Philadelphia, PA, USA. ⁷Center for Evolutionary Biology and Medicine, University of Pittsburgh, Pittsburgh, PA, USA. ✉e-mail: vantyne@pitt.edu

to clonal complex 17 (CC17), which phylogenetically resides within hospital-associated *E. faecium* clade A1. CC17 strains frequently encode antimicrobial resistance genes, mobile genetic elements and genes that enable the metabolism of amino sugars found on GI epithelia and mucin, probably contributing to the success of CC17 strains in healthcare settings^{12–18}. This success is exemplified by CC17 lineages being identified as responsible for widespread outbreaks and increased rates of invasive infection^{14,15,17,19}. Although several previous studies have investigated VREfm population structure and dynamics within healthcare settings, we know little about the factors that drive the emergence and persistence of particular VREfm lineages in the hospital.

In this study, we characterized the population structure and dynamics of VREfm within a single hospital using whole-genome sequencing-based surveillance and functional characterization of genes associated with nosocomial emergence. We systematically collected 710 VREfm clinical isolates over a 6-year period and used both genomic analysis and phenotypic testing to investigate factors contributing to population shifts observed within the facility. In addition, we compared local findings with a global collection of 15,631 publicly available VREfm genomes isolated from human sources from 2002 to 2022. We found that a bacteriocin produced by emergent VREfm lineages provided a strong competitive advantage, highlighting an adaptive mechanism that probably contributes to lineage replacement of VREfm on both local and global scales.

Results

Population structure of VREfm at a single hospital

Between 2017 and 2022, the Enhanced Detection System for Hospital-Associated Transmission (EDS-HAT) whole-genome sequencing surveillance programme collected 710 healthcare-associated VREfm isolates, that is, isolates collected from patients with hospital stays >2 days or before 30-day healthcare exposures at the University of Pittsburgh Medical Center (UPMC). The most common isolate sources were urine (42%), blood (24%) and wound sites (19%) (Supplementary Dataset 1). We investigated the genomic diversity of this collection through multilocus sequence typing, which identified 42 different STs. All isolates belonged to hospital-adapted lineages within CC17, including ST17 (23%), ST117 (13%), ST1471 (11%) and ST80 (10%) (Fig. 1a and Supplementary Dataset 1). To characterize the population structure of our collection, a core-genome phylogenetic tree was constructed based on 1,604 core genes (Extended Data Fig. 1a). Despite being entirely composed of CC17 strains, the VREfm population showed variable genetic diversity within and between STs and showed evidence that some isolates were closely related to one another. To assess genetic relatedness among the collected isolates, we performed Split Kmer Analysis to cluster isolates that had fewer than 10 single nucleotide polymorphisms (SNPs) in pairwise genome-wide comparisons. This analysis revealed 112 putative transmission clusters that contained 2–9 isolates each and encompassed 46% of the collection (Fig. 1a). Despite a high degree of clustering among all isolates, the proportion of isolates residing in putative transmission clusters was variable between STs. Although ST17 was the most prevalent lineage, it had the lowest percentage of clustered isolates (33%, 54 of 165 isolates, $P = 0.0008$). On the other hand, ST1478 showed a significantly higher percentage of clustered isolates (69%, 38 of 55 isolates, $P = 0.0002$) (Fig. 1b).

VREfm lineage replacement

We next investigated how the VREfm population changed over time within our centre by characterizing the ST distribution over the collection period in 6-month intervals (Fig. 1c). Before 2020, ST17 was the most frequently sampled ST, making up 34% of the collection between 2017 and 2019. However, during 2020, the emergence of ST1478 (23%) coincided with the decline of ST17 (17%). For the remainder of the collection period, the presence of ST17 continued to decline, and this lineage was not detected during the second half of 2022 (Fig. 1c). By contrast,

lineages ST80 and ST117 were not detected in 2017, but together rose to 81% by the end of 2022, effectively replacing ST17 and other lineages that were previously detected. We therefore designated ST80, ST117 and ST1478 as emergent lineages at our centre.

To identify factors contributing to lineage replacement, we investigated the frequency of non-susceptibility to the clinically relevant, last-line antibiotics linezolid and daptomycin (Extended Data Fig. 1a and Supplementary Dataset 1). The emergent lineages (ST80, ST117, ST1478) did not show a higher frequency of non-susceptible isolates, with non-susceptibility rates of 0–5% (linezolid) and 6–20% (daptomycin). We then investigated the distribution of genomic features such as genome length, antimicrobial resistance genes (ARGs), virulence factors and plasmid replicons among the different lineages (Extended Data Fig. 1b–e). We observed variation in the number of ARGs and virulence genes within the emergent lineages, with ST117 (mean = 15.6 ARGs) and ST1478 (mean = 16.0 ARGs) having more ARGs compared with ST17 (mean = 14.1, $P < 0.0001$) (Extended Data Fig. 1c). The macrolide efflux transporter *mefH* was found in nearly all ST117 and ST1478 isolates and was identified only in 4 other isolates (Supplementary Dataset 2). Similarly, the aminoglycoside nucleotidyltransferase *ant(6)-Ia* was highly enriched in these two lineages, being present in 99% and 93% of ST117 and ST1478 isolates, compared with 52% of other isolates. ST117 also had more virulence genes (mean = 4.0) compared with ST17 (mean = 3.7, $P < 0.0001$) (Extended Data Fig. 1d). Virulence genes enriched (>98%) in ST117 genomes included the colonization factors *acm*, *fss3*, *ecbA* and *sgrA* (Supplementary Dataset 2). ST1478 and ST117 genomes also encoded more plasmid replicons compared with the historical lineage ST17 ($P < 0.0001$) (Extended Data Fig. 1e). We further investigated the distribution of replicons among lineages and found that the rep11a replicon was present at higher frequencies in the emergent lineages ST80 (64%), ST117 (75%) and ST1478 (95%), versus only 15% of other isolates (Supplementary Dataset 2). Similarly, the repUS15_2 family replicon was enriched in ST117 (95%) and ST1478 (98%) but was seen at a lower prevalence in the remaining isolates (28%), including ST80 (10%) (Supplementary Dataset 2). Together, these data suggest that emergent lineages possess genomic features that might facilitate their emergence within the hospital.

Emergent isolates with bacteriocin T8 inhibit growth

To determine other factors contributing to lineage replacement, we investigated whether emergent lineage VREfm isolates inhibited the in vitro growth of historical lineage isolates. We performed a pairwise spot killing assay using the earliest available isolates from the historical lineage ST17 and the emergent lineage ST117. We found that the ST117 isolate was able to inhibit growth of the historical ST17 isolate, causing a large zone of inhibition in the ST17 isolate lawn surrounding the ST117 isolate spot (Fig. 2a). We then conducted pairwise spot killing assays using isolates from each of the 11 lineages having ≥10 isolates in the dataset (Fig. 2b and Supplementary Dataset 3). We found that isolates from all three emergent lineages caused growth inhibition of isolates from other lineages. Interestingly, we also observed low-level inhibition when emergent strains were spotted onto lawns of other emergent strains, possibly due to a high abundance of an antibacterial factor produced by the spot strain. Bacteriocins are antimicrobial peptides that have been widely studied in *Enterococcus* due to their ability to inhibit the growth of other bacteria and their potential role as probiotics^{20–23}. We screened the genomes in the dataset for predicted bacteriocins and found one bacteriocin, called T8, that was differentially present and found in 36% of isolates (Extended Data Fig. 2a and Supplementary Dataset 2). Bacteriocin T8 is identical to two other enterococcal bacteriocins, named hiracin JM79 (ref. 21) and bacteriocin 43 (ref. 22), and all three names have been used in previous studies to describe what is now known to be the same bacteriocin. We chose to refer to this bacteriocin as T8 because this is how it was initially and most frequently described in the literature. To assess the association between bacteriocin T8 with

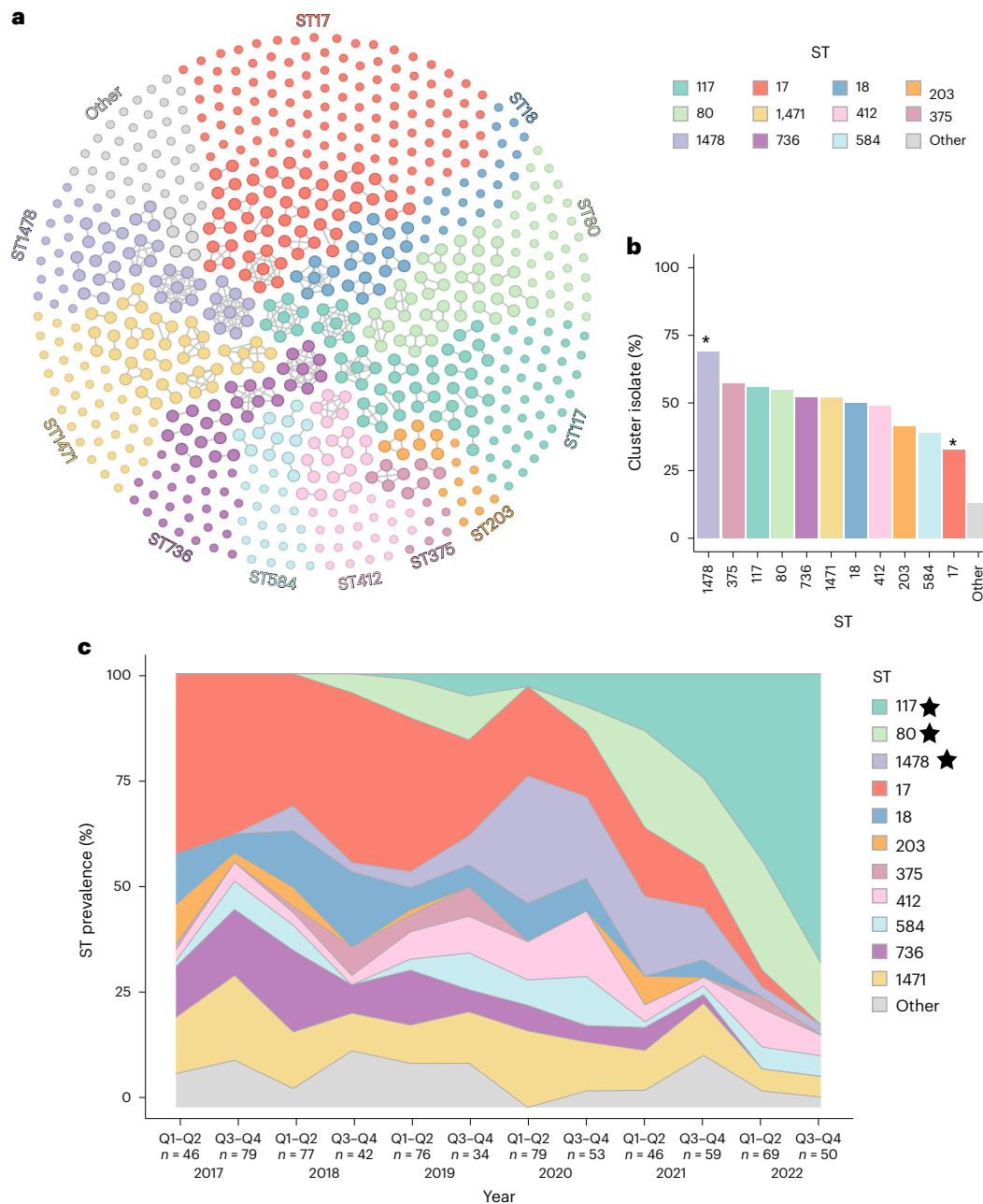


Fig. 1 | Population structure and temporal dynamics of VREfm at UPMC over 6 years. a, Cluster network diagram of 710 sequenced VREfm genomes constructed using Gephi v0.10. Isolates are grouped and coloured by multilocus ST. Isolates that fall within putative transmission clusters (≤ 10 SNPs) are connected with grey lines. **b**, Prevalence of cluster isolates within different STs. Asterisks mark STs that show a higher or lower percentage of clustered isolates

compared with the total collection (one-sided, single-proportion hypothesis test). *P* values were adjusted for multiple comparisons (* $P < 0.004$): ST1478, $P = 0.0002$; ST17, $P = 0.0008$. **c**, Biannual distribution of STs over the study period. Q1–Q2 = January–June; Q3–Q4 = July–December. The number of samples within each sample period is also noted. Emergent lineages are noted with a black star.

growth inhibition, we screened 28 VREfm isolates representing 11 STs against the same ST17 reference isolate to assess growth inhibition and found that bacteriocin T8 presence was strongly associated with growth inhibition ($P < 0.0001$) (Fig. 2c and Supplementary Dataset 4). A single isolate, called VRE36503, lacked bacteriocin T8 but still showed growth inhibition of the ST17 lawn. While no predicted bacteriocins were identified in the VRE36503 genome using the BAGEL4 prediction tool, an additional search for secondary metabolites identified a gene cluster with homology to the carnobacteriocin XY biosynthetic gene cluster²⁴, suggesting that growth inhibition by VRE36503 might be independent of bacteriocin T8.

To investigate whether bacteriocin T8 was encoded by a plasmid, we performed long-read sequencing and hybrid genome assembly on a bacteriocin T8-positive ST117 isolate. We found that bacteriocin T8 and the corresponding immunity factor were carried on a 6,173-bp rep11a-family plasmid that also encoded mobilization genes *mobABC*, allowing plasmid transfer, and was similar to plasmids previously described^{20,22} (Extended Data Fig. 2b). Of all bacteriocin T8-positive isolates ($n = 253$), rep11a was found in 236 isolates (93%) based on short-read assembly data, and 211 of these isolates encoded both bacteriocin T8 and rep11a on the same contig. We also characterized the distribution of bacteriocin T8 among the main lineages in the dataset,

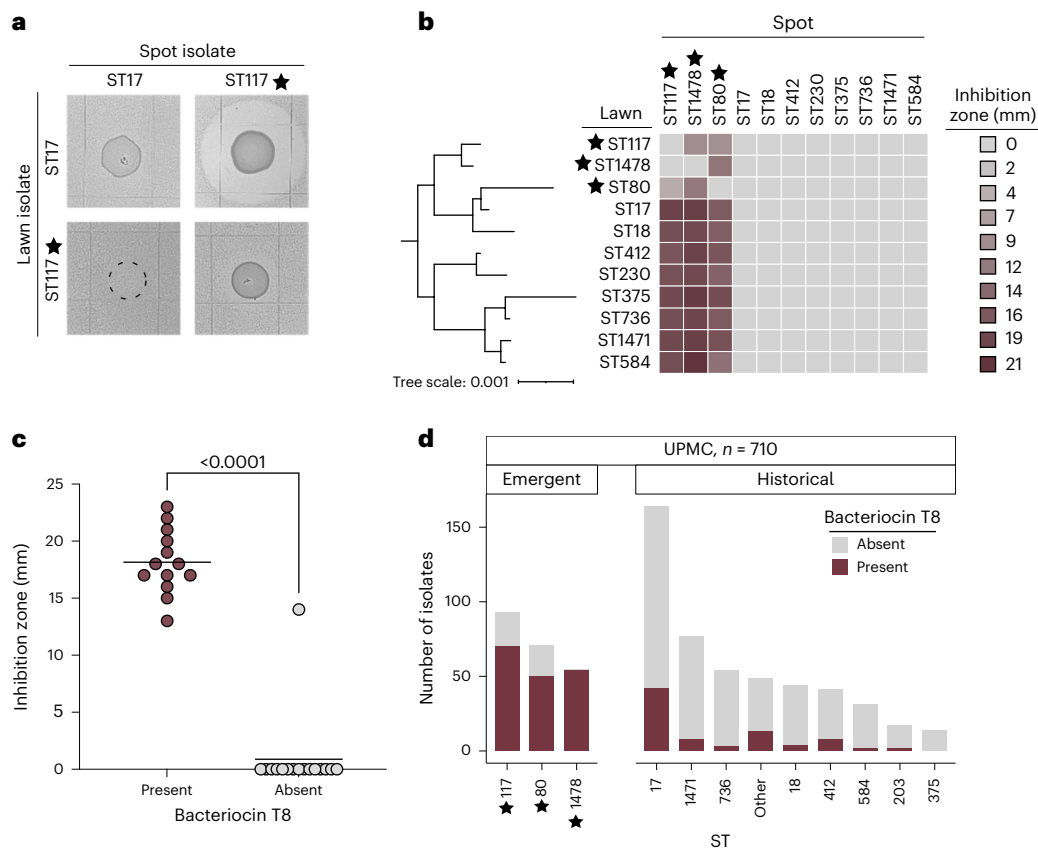


Fig. 2 | Bacteriocin T8 is associated with growth inhibition in emergent lineage isolates. **a**, Pairwise spot killing assay with representative ST17 (historical lineage) and ST117 (emergent lineage) isolates. The dashed circle shows where the ST17 isolate was spotted onto the ST117 lawn but did not grow. **b**, Pairwise spot killing assay with reference isolates from each of the 11 main lineages within the UPMC collection. The inhibition zones (mm) are shaded from highest inhibition (burgundy) to no inhibition (grey). Inhibition zone values were averaged from three biological replicates. The midpoint-rooted phylogenetic tree was constructed using RAxML with 100 bootstraps based on a single-copy

core genome alignment produced by Roary. The scale bar represents nucleotide substitutions per site. **c**, Spot killing assay results of 28 VREfm isolates spotted onto a lawn, once per isolate, of the bacteriocin T8-negative, ST17 VRE32530 representative isolate. Isolates are grouped by presence (burgundy) or absence (grey) of bacteriocin T8 in their genomes. The P value indicates significance from a two-tailed Mann–Whitney test (unadjusted, $P \leq 0.0001$). **d**, Abundance of bacteriocin T8 within main STs at UPMC. Emergent lineages are noted with a black star.

and identified a high prevalence among the emergent lineages ST80 (70%), ST117 (74%) and ST1478 (96%) (Fig. 2d). Bacteriocin T8 was found only in 16% of the remaining isolates in the collection, most of which belonged to ST17 (25%). Due to the enrichment of bacteriocin T8 in emergent lineage genomes, we hypothesized that it might provide a competitive advantage to VREfm during colonization and infection of hospitalized patients.

Bacteriocin T8 expression provides a competitive advantage

We confirmed that bacteriocin T8 caused growth inhibition by transforming a bacteriocin T8-negative clinical *E. faecium* isolate belonging to ST203 and called DVT705 with pBAC (plasmid containing bacteriocin T8 and immunity factor) or pEV (empty vector). To test whether pBAC conferred growth inhibition, we performed a pairwise spot killing assay and found that the pBAC strain caused a large zone of inhibition on a lawn of the pEV strain (Fig. 3a). Next, we quantified the competitive advantage of the pBAC strain by performing a liquid competition assay. We independently competed the pBAC and pEV strains against the parent strain at 50:50 and 10:90 starting ratios and quantified the abundance of each strain in the mixture after 24 h and 48 h. At both ratios and timepoints, the pBAC strain was able to outcompete the parent strain to a much greater extent compared with the pEV strain ($P < 0.01$) (Fig. 3b and Supplementary Dataset 5). We then evaluated whether the competitive advantage conferred by

bacteriocin T8 in vitro translated to the mammalian gut. To assess this, we pretreated C57BL/6 mice with vancomycin to deplete their endogenous *Enterococcus* population before orally gavaging mice with the pBAC or pEV strains for 2 days. We monitored the abundance of each strain in stool for 8 days following the initiation of infection (Fig. 3c and Supplementary Dataset 6A). On day 1, there was no difference in GI burden between the two groups, indicating that mice received similar inocula of each strain. However, at all subsequent timepoints, the pBAC strain was detected at a significantly higher abundance compared with the pEV strain ($P < 0.05$) (Fig. 3c). Next, to directly evaluate whether a bacteriocin T8-encoding emergent lineage strain could outcompete a historical lineage strain in the gut, we performed a competition experiment with a bacteriocin T8-positive ST117 strain and a bacteriocin T8-negative ST17 strain. To model strain emergence, mice were co-inoculated with ST117 and ST17 strains at a 10:90 ratio, with the emergent ST117 strain at an initial disadvantage (Fig. 3d and Supplementary Dataset 6B). Using culture-enriched metagenomic sequencing, we observed that mice were initially colonized on day 0 with 10.26% ST117 strain and 89.74% ST17 strain. However, at 4 days postinfection, the ST117 strain was detected at an average of 90% frequency, with half of the animals having only the bacteriocin T8-positive ST117 strain detected ($P < 0.0001$). These findings further suggest that bacteriocin T8 provides a competitive advantage to VREfm in the mammalian GI tract.

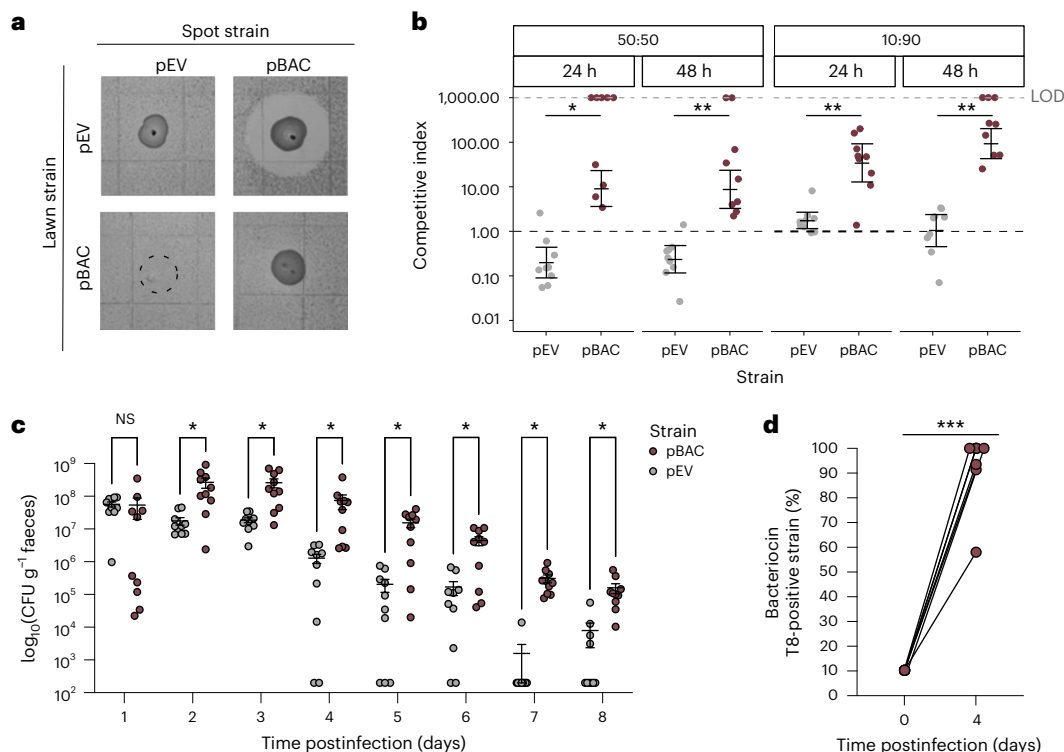


Fig. 3 | Bacteriocin T8 production provides a competitive advantage in vitro and increases *E. faecium* colonization and competition in vivo. **a**, Pairwise spot killing assay of strains pBAC, carrying bacteriocin T8 and the corresponding immunity gene, and pEV, carrying an empty vector. The dashed circle shows where the pEV strain was spotted onto the pBAC strain lawn but did not grow. **b**, Liquid competition assay. The pBAC and pEV strains were independently competed against the parent strain at 50:50 and 10:90 ratios. Samples were taken after 24 h and 48 h to calculate CFU ml⁻¹. Assays were performed in three biological replicates each consisting of three technical replicates. Instances in which the parent strain fell below the limit of detection (LOD, shown with the grey dashed line) were not included in statistical analyses. Competitive indexes were summarized using the geometric mean and a 95% confidence interval (error bars) for each timepoint and ratio⁷⁰. The distribution of competitive index values at each timepoint and ratio was compared between pBAC and pEV strains using a two-tailed Mann–Whitney test (adjusted, **P* < 0.005, ***P* < 0.001); 50:50 24 h (*P* = 0.0028), 50:50 48 h (*P* = 0.0002), 10:90 24 h (*P* = 0.0008) and 10:90 48 h

(*P* = 0.0004). Black dashed line corresponds to a competitive index of 1.0. **c**, Colonization of pBAC and pEV strains in the mouse gut. Mice were orally gavaged with either the pBAC (*n* = 10) or pEV (*n* = 10) strain for 2 days. Stool samples were collected starting on day 1 after the initiation of infection to quantify CFU g⁻¹ of each strain over time. Data were summarized using the mean with error bars representing standard error measurement. A two-tailed multiple Mann–Whitney test was used and assessed CFU g⁻¹ distribution between pBAC and pEV strains (adjusted, **P* < 0.05; NS, non-significant); day 1 (*P* = 0.6839), day 2 (*P* = 0.0230), day 3 (*P* = 0.0084), day 4 (*P* = 0.0026), day 5 (*P* = 0.0120), day 6 (*P* = 0.0410), day 7 (*P* = 0.0001) and day 8 (*P* = 0.0061). **d**, Competition between a bacteriocin T8-positive ST117 strain and a bacteriocin T8-negative ST17 strain within the mouse gut. Data show the percentages of ST117 in the inoculum on day 0 and at 4 days postinfection. Proportions of ST117 on day 4 and in the inoculum for each animal were compared using a one-sided difference in proportion hypothesis test (adjusted, ****P* ≤ 0.0001) for each mouse.

Global VREfm lineage replacement linked with bacteriocin T8

We next sought to determine whether the lineage replacement we observed at our local centre was reflective of global VREfm population dynamics. To investigate this question, we gathered 15,631 publicly available VREfm genomes collected from human sources between 2002 and 2022 (Supplementary Dataset 7). This collection consisted of genomes from 53 countries; however, the majority of isolates were from the United States (23%), Denmark (23%) and Australia (20%) (Extended Data Fig. 3 and Supplementary Dataset 7). To investigate VREfm global genomic diversity, we performed sequence typing on this collection and examined the distributions of STs by continent (Fig. 4a). Europe and Asia had relatively clonal populations, while the populations in North America and Australia were more diverse. ST80 was the single most prevalent ST (20%) and was mainly found in Europe (30%) and Asia (36%). ST117 was the second most prevalent ST (12%) and had the highest prevalence in Europe (18%) and North America (12%). ST17 accounted for 7% of the global population and was sampled predominantly in North America (18%). To characterize global population dynamics of VREfm, we investigated the prevalence of STs over the 20-year global collection period (Fig. 4b and Supplementary Dataset 7). Before 2010, ST17 and ST18 were the most prevalent lineages, while ST80 and ST117

were detected very infrequently. After 2010, however, ST117 and ST80 rose to 60% by the end of 2022, effectively replacing ST17 (3%) and ST18 (0.2%). These data suggest that the emergence of ST80 and ST117 that we observed locally was reflective of global trends.

To investigate whether bacteriocin T8 was similarly enriched in emergent lineages, we examined the distribution of bacteriocin T8 among the STs sampled in the global collection of VREfm isolated from human sources (Fig. 4c). Bacteriocin T8 was enriched in emergent lineages ST80, ST117 and ST1478 globally, with more than 79% of isolates in each ST predicted to encode the bacteriocin. Similar to our local prevalence (25%), bacteriocin T8 was found in only 30% of isolates in the previously dominant lineage ST17. We also investigated whether bacteriocin T8 was increasing over time within both collections (Fig. 4d). Within our local collection, bacteriocin T8 presence rose from 8% in 2017 to 62% in 2022. Similarly, we observed a 67% increase in the prevalence of bacteriocin T8 between 2002 and 2022 in the global collection. Within both collections, the increase in bacteriocin T8 prevalence was associated with the replacement of the historical ST17 lineage with emergent lineages ST80 and ST117. Taken together, these findings suggest that bacteriocin T8 may be a driving feature of global VREfm strain emergence and persistence in healthcare settings.

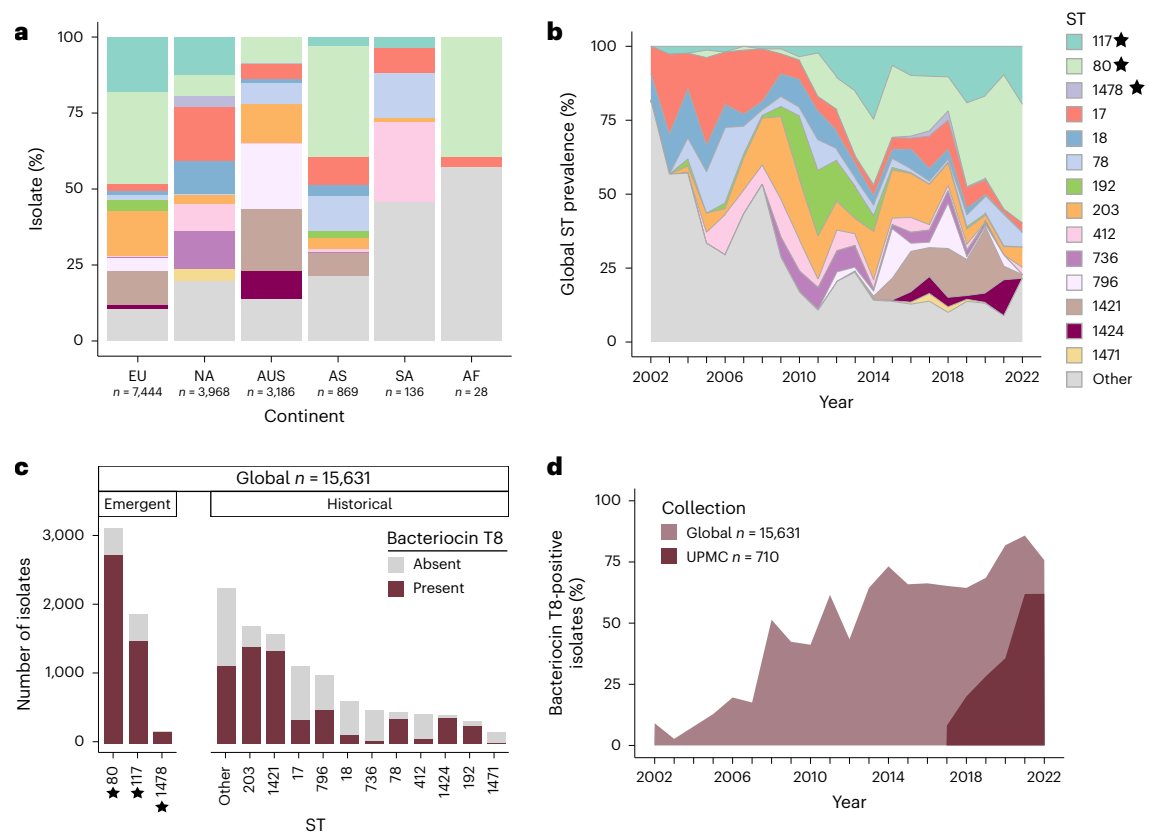


Fig. 4 | Global bacteriocin T8 prevalence increases over time and is associated with emergent lineages. **a**, Geographic distribution of 15,631 global VREfm isolates across continents. EU, Europe; NA, North America; AUS, Australia; AS, Asia; SA, South America; AF, Africa. Isolates are coloured by ST. **b**, Global ST distribution between 2002 and 2022. Isolates are coloured by ST. Emergent

lineages are noted with a black star. **c**, Abundance of bacteriocin T8 within main global STs. The bars are separated into emergent and historical lineages based on trends seen within the UPMC collection. Bacteriocin T8 presence is shaded in burgundy and absence in grey. **d**, Prevalence of bacteriocin T8 over time at UPMC (burgundy) and in the global collection (light burgundy).

Discussion

In this study, we examined the population structure and dynamics of 710 VREfm clinical isolates collected over 6 years from a single hospital. A strength of our study lies in the use of a systematic collection of hospital-associated VREfm isolates collected over a multi-year period. In addition, by comparing our findings to a large global collection of over 15,000 VREfm isolates from human sources, we confirmed that many of our findings were generalizable to other settings worldwide. Our data show the emergence of ST80 and ST117 both locally and globally, highlighting the strong competitive advantage of these lineages and identifying bacteriocin T8 as a likely contributor to VREfm lineage replacement.

Similar to other studies, we found that the VREfm population at our hospital was polyclonal, with the majority (57%) of isolates belonging to 4 prevalent lineages: ST17, ST117, ST80 and ST1471 (refs. 12,13,25,26). These lineages belong to the hospital-associated clade A1 of *E. faecium* and also belong to CC17, which is known to be highly epidemic within healthcare systems and the cause of widespread outbreaks^{13–15,17,27}. While previous studies have reported nosocomial VREfm transmission rates ranging from 60% to 80% (refs. 17,28–30), we found that only 46% of isolates in our dataset belonged to putative transmission clusters. This difference is probably due to our use of a 10-SNP cut-off for clustering and not including sampling of VREfm from GI tract colonization, which might limit our ability to detect transmission³¹. We did, however, note differences in the percentage of isolates residing within putative transmission clusters among different lineages, with lineage ST1478 showing a significantly higher proportion of clustered isolates. Previous literature has found that the ST117 lineage was responsible for numerous VREfm hospital outbreaks^{12,14,32}, and that the ST1478

lineage has previously been detected in the United States, Canada and Europe^{13,19,33}. Taken together, these findings suggest that some VREfm lineages might be more efficient than others at transmitting between patients in the hospital.

Through phenotypic screening and comparative genomic analysis, we found that the antimicrobial peptide bacteriocin T8 was enriched in emergent lineages both locally and globally, and that it conferred a growth advantage to *E. faecium* both in vitro and in vivo. The enrichment of bacteriocin T8 in emergent lineages and increasing prevalence over time suggest that acquisition of this bacteriocin is highly advantageous. Bacteriocins have been shown to facilitate expansion of bacterial populations by killing susceptible bacteria, thereby carving out a stable environment for the expansion of bacteriocin-expressing bacteria^{34–39}. In polymicrobial settings that experience strong selection and associated population bottlenecks, such as within the GI tract of hospitalized patients, strains must compete with one another to dominate and persist in the population. A previous study investigated the competitive advantage conferred by bacteriocin-21 to *Enterococcus faecalis* in the mammalian GI tract³⁴. Similar to our findings, bacteriocin production in that study was associated with increased GI tract colonization and conferred a competitive advantage in the mouse gut. Furthermore, the study found that the production of bacteriocin-21 was able to clear a vancomycin-resistant *E. faecalis* strain from the GI tract³⁴. Due to the strong inhibitory activity of bacteriocins, they are an attractive avenue for development as new antimicrobial interventions, such as inclusion in probiotics and food preservation^{36,40}. A recent study showed that a genetically engineered probiotic *Escherichia coli* strain containing three bacteriocins, including bacteriocin T8 (referred to as hircin JM79), was able to clear vancomycin-resistant *E. faecium* and *E. faecalis*

in a mouse model of enterococcal colonization³⁶. Although this result is exciting, it is somewhat troubling that based on our global findings the vast majority of VREfm isolates sequenced in 2022 already encoded bacteriocin T8 and the associated immunity gene, suggesting that they would be resistant to bacteriocin T8 activity.

Bacteriocin T8 is a class IIa, secretion-dependent, heat-stable bacteriocin²⁰. It was isolated in 2006 from an *E. faecium* strain called T8 that was sampled from patients with human immunodeficiency virus²⁰. *E. faecium* appears to be the natural host for bacteriocin T8, although a single report documented a bacteriocin T8-encoding *Enterococcus hirae* strain isolated from mallard ducks²¹. Bacteriocin T8 causes bacterial cell death by irreversibly binding to the mannose phosphotransferase system resulting in pore formation^{41,42}. Inhibitory activity has been reported against *Enterococcus* spp., *Listeria* spp. and *Lactobacillus* spp., and even some *Staphylococcus aureus* and *Clostridium* spp. strains^{20–22}. Resistance to class IIa bacteriocins appears to occur largely through decreased expression of the target mannose phosphotransferase system^{43,44}. Bacteriocins are typically found on mobile genetic elements and are transmitted to compatible strains through horizontal gene transfer⁴⁵. The bacteriocin T8-encoding plasmid that we characterized in this study is highly similar to plasmids described in the early 2000s, suggesting widespread conservation and propagation of this mobile element^{20,22}.

Our study had several limitations. Within our local collection, we investigated only VREfm isolates collected from clinical infections that were suspected to be healthcare-associated. Isolates not meeting inclusion criteria, including potentially community-associated VREfm, were not included. It is very likely that we undersampled the full VREfm population diversity within our centre because many hospitalized patients have asymptomatic GI tract colonization³¹. The global collection that we analysed was also biased towards countries with high rates of VREfm infection and with the infrastructure and capacity to perform next-generation sequencing, which resulted in some continents, such as South America and Africa, to be greatly undersampled. Additionally, we focused on bacteriocin T8 as a contributor to lineage success; however, this may not be the only factor driving the lineage replacement that we observed. We did not investigate other potential adaptations among emergent lineages, such as virulence factors that were enriched in the ST117 lineage. Furthermore, it is important to note that *E. faecium* virulence factors in general are poorly characterized, limiting their identification across our collection. Moreover, additional uncharacterized mutations within the emergent lineages could contribute to antimicrobial resistance or tolerance, potentially aiding in lineage replacement. In addition, in the mouse model, we focused on the impact of bacteriocin production on enterococci, potentially overlooking other microbiome disruptions that could be explored through additional metagenomic sequencing.

In summary, we characterized the local and global population structure and temporal dynamics of VREfm using comparative genomics and functional analyses. By investigating VREfm populations sampled over 6 years at our healthcare centre and over 20 years globally, we identified lineage replacement associated with the spread of strains encoding bacteriocin T8. Phenotypic characterization showed that bacteriocin T8 probably contributes to VREfm lineage replacement by conferring a strong competitive advantage that is observed both in vitro and in vivo. Although we identified bacteriocin T8 production as a potential adaptive mechanism directing VREfm lineage replacement, this study prompts further investigation into other features driving the evolutionary dynamics in this important and difficult-to-treat pathogen.

Methods

Study setting

The study was approved by the Institutional Review Board at the University of Pittsburgh (STUDY21040126). This was a retrospective observational study of VREfm collected from patients at UPMC by

the EDS-HAT⁴⁶. UPMC is an adult tertiary care hospital with 699 beds (including 134 critical care beds) and performs >400 solid organ transplants each year. A total of 710 VREfm clinical isolates were collected from patients with a hospital admission date ≥ 2 days previously or with a recent healthcare exposure within 30 days before the culture date, from January 2017 to December 2022. Available daptomycin and linezolid minimal inhibitory concentration data were collected from patient records and interpreted using Clinical and Laboratory Standards Institute M100 guidelines.

Whole-genome sequencing and bioinformatic analyses

Genomic DNA was extracted using a DNeasy Blood and Tissue Kit (Qiagen) from VREfm isolates that were grown overnight at 37 °C on blood agar plates. Following DNA extraction, next-generation sequencing libraries were generated using the Illumina DNA Prep protocol and then sequenced (2×150 bp, paired end) on NextSeq500, NextSeq2000 or MiSeq. The resulting reads were assembled using SPAdes v3.15.5 (ref. 47). Assembly quality was determined using QUAST v5.2.0 (ref. 48). Assemblies passed quality control if the coverage was $>35\times$ and the assembly had <350 contigs. Species were identified and possible contamination was detected using Kraken2 (v2.0.8- β)⁴⁹. Multilocus STs were identified using the PubMLST database with mlst v2.11 (refs. 50,51). Isolates with undefined STs were uploaded to the PubMLST server, and if their ST was a single locus variant of a known ST, they were grouped with the latter. Clusters of genetically related isolates were identified using Split Kmer Analysis v1.0 (ref. 52) with average linkage clustering and a 10-SNP cut-off, similar to previously used thresholds^{27,53,54}. Genomes were annotated using PROKKA v1.14.5 (ref. 55). The bacteriocin T8-encoding plasmid was annotated with Bakta v1.9.2 (ref. 56). A cluster network diagram was visualized using Gephi v0.10 with the Fruchterman–Reingold layout⁵⁷. Phylogenetic trees were built using core genome alignments produced by Roary v3.13.0 (ref. 58). Gaps in the core genome alignment were masked using Geneious (Geneious Biologics 2024, <https://www.geneious.com/biopharma>). Trees were constructed using RAXML HPC v8.2.12 with 100 bootstraps⁵⁹. In the UPMC collection, bacteriocins were identified using BAGEL4 with $\geq 95\%$ coverage and identity⁶⁰. antiSMASH v7.1.0 was used to further identify secondary metabolites in the VRE36503 genome⁶¹. To identify bacteriocin T8 in the global genome collection, a custom database consisting of the nucleotide sequence for bacteriocin T8 and the corresponding immunity factor was built using ABRicate v1.01, and gene presence was defined as hits with $\geq 95\%$ coverage and identity⁶². Antimicrobial resistance genes were identified with AMRFinderPlus v3.12.8 (ref. 63). The presence of plasmid replicons and virulence factors was determined using ABRicate v1.0.01 (ref. 62) with the PlasmidFinder⁶⁴ and VFDB⁶⁵ databases, respectively. Gene presence was defined as hits with $\geq 90\%$ coverage and identity. The bacteriocin T8-encoding plasmid was resolved by using Unicycler v0.5.0 (ref. 66) to hybrid assemble Illumina and MinION sequencing data collected from isolate VRE38098. A MinION device with R9.4.1 flow cells (Oxford Nanopore Technologies) was used, with base calling performed by Dorado v0.7.2 (ref. 67).

Spot killing assay

Test strains were cultivated for 16–18 h at 37 °C in brain heart infusion (BHI) broth. Subsequently, 5 ml of a BHI top agar lawn (containing 0.35% agar) was prepared by mixing the molten agar with 100 μ l of a 1:100 dilution of the overnight culture. This mixture was poured on top of 25 ml of solid BHI agar. Competing strains were spotted (5 μ l undiluted overnight culture) onto solidified top agar lawns. Inhibition zones were measured (in mm) after 16–18 h of incubation at 37 °C.

Cloning and expression of bacteriocin T8

Bacteriocin T8 and the corresponding immunity factor were cloned into the expression vector pMSP3535, which was modified to encode the

chloramphenicol resistance gene *cat* as a positive selectable marker⁶⁸. The insert sequence was amplified from VRE38098 genomic DNA by PCR using primers 5'-AGA CCG GCC TCG AGT CTA GAA TGG GAC TGA TGA ATC AGA ATTG-3' and 5'-GCG AGC TCG TCG ACA GCG CTC AGG CGT TAC TTG GTA GTC CAT-3'. The vector was amplified by PCR using primers 5'-AGC GCT GTC GAC GAG CTC GCAT-3' and 5'-TCT AGA CTC GAG GCC GGT CTCC-3'. The amplified insert and vector were purified using a PCR Purification Kit (Qiagen), and Gibson assembly was conducted using a HiFi DNA Assembly Cloning Kit (New England Biolabs)⁶⁹. The Gibson product was then transformed into NEB 5-alpha-competent *E. coli*, and transformants were selected on BHI agar containing 10 µg ml⁻¹ chloramphenicol. Plasmids were amplified in 200-ml cultures, collected by Maxiprep and sequenced to confirm their identity. The bacteriocin T8-encoding vector (pBAC) or pMSP3535 empty vector (pEV) were then transformed into the bacteriocin T8-negative *E. faecium* ST203 strain DVT705, which is a plasmid-cured, vancomycin-susceptible derivative of the 10-10-S strain⁷. Successful transformation was confirmed with PCR using pMSP3535 backbone-specific primers 5'-CAA TAC GCA AAC CGC CTCTC-3' and 5'-TGG CAC TCG GCA CTT AATGG-3'. Inhibitory activity of DVT705 transformed with pBAC was confirmed by a pairwise spot killing assay against DVT705 transformed with pEV.

Liquid competition assay

pBAC and pEV strains were competed individually against the parent DVT705 strain at two ratios, 50:50 and 10:90. For each ratio, three biological replicates each consisting of three technical replicates were performed. Before the competition, each strain was grown separately overnight and cultures were normalized to optical density (OD)₆₀₀ = 0.5. Strains were mixed together at the above starting ratios, and the mixture was then diluted 1:100 into 5 ml of BHI and grown shaking at 37 °C. Samples were taken at 24-h and 48-h timepoints. Samples were track diluted onto BHI, and the BHI was supplemented with 10 µg ml⁻¹ of chloramphenicol to calculate colony forming units per ml (CFU ml⁻¹). The abundance of the parental strain was calculated by subtracting the CFU ml⁻¹ on the BHI plate by the CFU ml⁻¹ on the chloramphenicol plate. Parent measurements that fell below the limit of detection, 1,000 CFU ml⁻¹, were not included in competitive index calculations. The competitive index was calculated as below, and results were summarized using the geometric mean and a 95% confidence interval for each timepoint and ratio⁷⁰.

$$\text{Competitive index} = \frac{\frac{\text{24-h or 48-h ratio CFU of pBAC or pEV}}{\text{CFU of parent}}}{\frac{\text{Initial ratio CFU of pBAC or pEV}}{\text{CFU of parent}}}$$

Mouse colonization and competition experiments

Animal experiments were approved by the Animal Care and Use Committees of the Children's Hospital of Philadelphia (IAC 18–001316). Five-week-old C57BL/6 male mice were purchased from Jackson Laboratories and given 1 week to equilibrate their microbiota before experimentation. All experimental procedures were performed in a biosafety level two laminar flow hood. Mice were housed in a 12-h light–dark cycle with ambient temperature and humidity maintained at 68–74 °F and 30–70%, respectively. No statistical methods were used to predetermine sample sizes; however, our sample sizes are similar to those reported in previous publications^{2,34,36}. Mice were given vancomycin (1 mg ml⁻¹) in drinking water ad libitum for 5 days followed by a 2-day recovery period^{71,72}. Mice were then infected with 5 × 10⁸ *E. faecium* cells by oral gavage twice a day for 2 days. Mice were randomly assigned to experimental groups. Data collection and analysis were not performed blind to the conditions of the experiments.

For colonization experiments, the pBAC and pEV strains were prepared by growing them to stationary phase in liquid culture and washing them with cold PBS immediately before infection. Four cages, each

containing five mice, were independently infected with either strain, with two cages assigned to each strain. Stool samples were collected daily for quantification of bacterial CFUs. Samples were diluted and homogenized in PBS, and serially plated onto either bile esculin azide (BEA) agar to quantify the total enterococcal population or BEA agar with chloramphenicol (10 µg ml⁻¹) to quantify the pBAC and pEV strains.

For competitive colonization experiments, the bacteriocin T8-positive ST117 VRE38098 and bacteriocin T8-negative ST17 VRE32530 strains were plated onto BEA agar and grown overnight at 37 °C. Colonies of each strain were inoculated into 4 ml of BHI media and grown shaking at 37 °C until OD₆₀₀ values reached 0.6. Strains were washed with cold PBS and mixed at a 10:90 ratio of ST117:ST17 immediately before infection. Stool samples were collected daily for quantification of bacterial CFUs. Samples were diluted and homogenized in PBS and were then serially plated onto either BEA agar to quantify the total enterococcal population or BEA agar with vancomycin (50 µg ml⁻¹) to quantify the ST17 and ST117 strains. From the vancomycin-supplemented plates, 100–1,000 colonies were pooled and sequenced at deep coverage (>100×) on the Illumina platform. The initial day 0 inoculum and samples from six mice (three per cage) collected 4 days postinfection were collected and sequenced. The resulting reads were mapped to the sequences of the seven multilocus ST alleles using breseq⁷³ in population mode to assess the relative abundance of ST17 and ST117 strains.

Global isolate collection

All *E. faecium* genomes deposited in National Center for Biotechnology Information (NCBI) were downloaded on 23 May 2024. *E. faecium* genomes with collection dates between 2002 and 2022 and for which the 'host' in the BioSample metadata was listed as '*Homo sapiens*', '*Homo sapiens sapiens*', 'hospitalized patient' and 'Human being' were included. Genomes encoding the *vanA* or *vanB* operon, as identified using AMRFinderPlus, were retained for analysis.

Statistical analyses

A single-proportion hypothesis test was performed to assess the enrichment of cluster isolates within ST groups, setting the proportion of cluster isolates in the total collection as the null value, with data distribution assumed to be normal although not formally tested. The association of bacteriocin T8 presence with growth inhibition, liquid competitive advantage of pBAC versus pEV strains relative to the parent strain, and mouse GI tract colonization differences between the pBAC and pEV strains were assessed using a two-tailed Mann–Whitney test⁷¹. To assess the difference in the proportion of the ST117 strain on day 4 of the competitive colonization experiment compared with the initial inoculum, a one-sided difference in proportion hypothesis test was conducted using the average fold coverage for each sample as the sample size. Bacteriocin T8 enrichment in different STs was assessed with a single-proportion hypothesis test, using the overall proportion of bacteriocin T8-positive isolates as the null value. Differences in the number of antimicrobial resistance genes, plasmid replicons and virulence genes between lineages were assessed using two-sided *t*-tests, with data distribution assumed to be normal but not formally tested. Statistical significance was determined with an $\alpha = 0.05$, and a Bonferroni correction for multiple comparisons was applied when appropriate. Assumptions were met for analyses unless noted otherwise.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Genomic sequences for all 710 VREfm isolates can be found in BioProject [PRJNA475751](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA475751) with accession numbers listed in Supplementary Dataset 1. Sequencing reads for the mouse competition experiment can

be found in BioProject [PRJNA1217837](https://www.ncbi.nlm.nih.gov/bioproject/PRJNA1217837) with accession numbers listed in Supplementary Dataset 6B.

References

- Arias, C. A. & Murray, B. E. The rise of the *Enterococcus*: beyond vancomycin resistance. *Nat. Rev. Microbiol.* **10**, 266–278 (2012).
- Ubeda, C. et al. Vancomycin-resistant *Enterococcus* domination of intestinal microbiota is enabled by antibiotic treatment in mice and precedes bloodstream invasion in humans. *J. Clin. Invest.* **120**, 4332–4341 (2010).
- Paterson, D. L. et al. Acquisition of rectal colonization by vancomycin-resistant *Enterococcus* among intensive care unit patients treated with piperacillin–tazobactam versus those receiving cefepime-containing antibiotic regimens. *Antimicrob. Agents Chemother.* **52**, 465–469 (2008).
- Sjölund, M., Wreiber, K., Andersson, D. I., Blaser, M. J. & Engstrand, L. Long-term persistence of resistant *Enterococcus* species after antibiotics to eradicate *Helicobacter pylori*. *Ann. Intern. Med.* **139**, 483–487 (2003).
- Van Tyne, D. & Gilmore, M. S. Friend turned foe: evolution of enterococcal virulence and antibiotic resistance. *Annu. Rev. Microbiol.* **68**, 337–356 (2014).
- Stellfox, M. E. & Van Tyne, D. Last bacteria standing: VREfm persistence in the hospitalized gut. *mBio* **13**, e0067022 (2022).
- Chilambi, G. S. et al. Evolution of vancomycin-resistant *Enterococcus faecium* during colonization and infection in immunocompromised pediatric patients. *Proc. Natl Acad. Sci. USA* **117**, 11703–11714 (2020).
- Austin, D. J., Bonten, M. J. M., Weinstein, R. A., Slaughter, S. & Anderson, R. M. Vancomycin-resistant enterococci in intensive-care hospital settings: transmission dynamics, persistence, and the impact of infection control programs. *Proc. Natl Acad. Sci. USA* **96**, 6908–6913 (1999).
- Correa-Martinez, C. L. et al. Risk factors for long-term vancomycin-resistant enterococci persistence—a prospective longitudinal study. *Microorganisms* **7**, 400 (2019).
- Lebreton, F. et al. Tracing the enterococci from Paleozoic origins to the hospital. *Cell* **169**, 849–861.e13 (2017).
- Talaga-Ćwiertnia, K. & Bulanda, M. Analysis of the world epidemiological situation among vancomycin-resistant *Enterococcus faecium* infections and the current situation in Poland. *Przegl Epidemiol.* **72**, 3–15 (2018).
- Falgenhauer, L. et al. Near-ubiquitous presence of a vancomycin-resistant *Enterococcus faecium* ST117/CT71/vanB—clone in the Rhine-Main metropolitan area of Germany. *Antimicrob. Resist. Infect. Control* **8**, 128 (2019).
- Egan, S. A. et al. Genomic analysis of 600 vancomycin-resistant *Enterococcus faecium* reveals a high prevalence of ST80 and spread of similar vanA regions via IS1216E and plasmid transfer in diverse genetic lineages in Ireland. *J. Antimicrob. Chemother.* **77**, 320–330 (2022).
- Lisotto, P. et al. Molecular characterisation of vancomycin-resistant *Enterococcus faecium* isolates belonging to the lineage ST117/CT24 causing hospital outbreaks. *Front. Microbiol.* **12**, 728356 (2021).
- Jozefiková, A., Valček, A., Šoltys, K., Nováková, E. & Bujdáková, H. Persistence and multi-ward dissemination of vancomycin-resistant *Enterococcus faecium* ST17 clone in hospital settings in Slovakia 2017–2020. *Int. J. Antimicrob. Agents* **59**, 106561 (2022).
- McCracken, M. et al. Emergence of pstS-Null vancomycin-resistant *Enterococcus faecium* clone ST1478, Canada, 2013–2018. *Emerg. Infect. Dis.* **26**, 2247–2250 (2020).
- Fang, H., Fröding, I., Ullberg, M. & Giske, C. G. Genomic analysis revealed distinct transmission clusters of vancomycin-resistant *Enterococcus faecium* ST80 in Stockholm, Sweden. *J. Hosp. Infect.* **107**, 12–15 (2021).
- Lebreton, F. et al. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* **4**, e00534-13 (2013).
- Kleinman, D. R. et al. Vancomycin-resistant *Enterococcus* sequence type 1478 spread across hospitals participating in the Canadian Nosocomial Infection Surveillance Program from 2013 to 2018. *Infect. Control Hosp. Epidemiol.* **44**, 17–23 (2023).
- De Kwaadsteniet, M., Fraser, T., Van Reenen, C. A. & Dicks, L. M. T. Bacteriocin T8, a novel class IIa sec-dependent bacteriocin produced by *Enterococcus faecium* T8, isolated from vaginal secretions of children infected with human immunodeficiency virus. *Appl. Environ. Microbiol.* **72**, 4761–4766 (2006).
- Sánchez, J. et al. Amino acid and nucleotide sequence, adjacent genes, and heterologous expression of hircin JM79, a sec-dependent bacteriocin produced by *Enterococcus hirae* DCH5, isolated from mallard ducks (*Anas platyrhynchos*). *FEMS Microbiol. Lett.* **270**, 227–236 (2007).
- Todokoro, D., Tomita, H., Inoue, T. & Ike, Y. Genetic analysis of bacteriocin 43 of vancomycin-resistant *Enterococcus faecium*. *Appl. Environ. Microbiol.* **72**, 6955–6964 (2006).
- Herranz, C. et al. Enterocin P selectively dissipates the membrane potential of *Enterococcus faecium* T136. *Appl. Environ. Microbiol.* **67**, 1689–1692 (2001).
- Britton, A. P., van der Ende, S. R., van Belkum, M. J. & Martin-Visscher, L. A. The membrane topology of immunity proteins for the two-peptide bacteriocins carnobacteriocin XY, lactococcin G, and lactococcin MN shows structural diversity. *Microbiologyopen* **9**, e00957 (2019).
- Shokoohzadeh, L. et al. High frequency distribution of heterogeneous vancomycin resistant *Enterococcus faecium* (VREfm) in Iranian hospitals. *Diagn. Pathol.* **8**, 163 (2013).
- van Hal, S. J. et al. Polyclonal emergence of vanA vancomycin-resistant *Enterococcus faecium* in Australia. *J. Antimicrob. Chemother.* **72**, 998–1001 (2017).
- Sundermann, A. J. et al. Outbreak of vancomycin-resistant *Enterococcus faecium* in interventional radiology: detection through whole-genome sequencing-based surveillance. *Clin. Infect. Dis.* **70**, 2336–2343 (2020).
- Sherry, N. L. et al. Multi-site implementation of whole genome sequencing for hospital infection control: a prospective genomic epidemiological analysis. *Lancet Reg. Health West. Pac.* **23**, 100446 (2022).
- Eichel, V. et al. Challenges in interpretation of WGS and epidemiological data to investigate nosocomial transmission of vancomycin-resistant *Enterococcus faecium* in an endemic region: incorporation of patient movement network and admission screening. *J. Antimicrob. Chemother.* **75**, 1716–1721 (2020).
- Gouliouris, T. et al. Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance. *Nat. Microbiol.* **6**, 103–111 (2021).
- Sundermann, A. J. et al. Genomic sequencing surveillance of patients colonized with vancomycin-resistant *Enterococcus* (VRE) improves detection of hospital-associated transmission. Preprint at medRxiv <https://doi.org/10.1101/2024.05.01.24306710> (2024).
- Weber, A., Maechler, F., Schwab, F., Gastmeier, P. & Kola, A. Increase of vancomycin-resistant *Enterococcus faecium* strain type ST117 CT71 at Charité - Universitätsmedizin Berlin, 2008 to 2018. *Antimicrob. Resist. Infect. Control* **9**, 109 (2020).
- McInnes, R. S. et al. Integration of vanHAX downstream of a ribosomal RNA operon restores vancomycin resistance in a susceptible *Enterococcus faecium* strain. *NPJ Antimicrob. Resist.* **2**, 2 (2024).
- Kommineni, S. et al. Bacteriocin production augments niche competition by enterococci in the mammalian GI tract. *Nature* **526**, 719–722 (2015).

35. Corr, S. C. et al. Bacteriocin production as a mechanism for the anti-infective activity of *Lactobacillus salivarius* UCC118. *Proc. Natl Acad. Sci. USA* **104**, 7617–7621 (2007).
36. Geldart, K. G. et al. Engineered *E. coli* Nissle 1917 for the reduction of vancomycin-resistant *Enterococcus* in the intestinal tract. *Bioeng. Transl. Med.* **3**, 197–208 (2018).
37. Wagner, T. M. et al. Interactions between commensal *Enterococcus faecium* and *Enterococcus lactis* and clinical isolates of *Enterococcus faecium*. *FEMS Microbes* **5**, xtae009 (2024).
38. Ríos Colombo, N. S. et al. Impact of bacteriocin-producing strains on bacterial community composition in a simplified human intestinal microbiota. *Front. Microbiol.* **14**, 1290697 (2023).
39. Garretto, A., Dawid, S. & Woods, R. Increasing prevalence of bacteriocin carriage in a 6-year hospital cohort of *E. faecium*. *J. Bacteriol.* **206**, e00294–24 (2024).
40. Davies, E. A., Bevis, H. E. & Delves-Broughton, J. The use of the bacteriocin, nisin, as a preservative in ricotta-type cheeses to control the food-borne pathogen *Listeria monocytogenes*. *Lett. Appl. Microbiol.* **24**, 343–346 (1997).
41. Kjos, M., Salehian, Z., Nes, I. F. & Diep, D. B. An extracellular loop of the mannose phosphotransferase system component IIC is responsible for specific targeting by class IIa bacteriocins. *J. Bacteriol.* **192**, 5906–5913 (2010).
42. Kjos, M. et al. Target recognition, resistance, immunity and genome mining of class II bacteriocins from Gram-positive bacteria. *Microbiology* **157**, 3256–3267 (2011).
43. Geldart, K. & Kaznessis, Y. N. Characterization of class IIa bacteriocin resistance in *Enterococcus faecium*. *Antimicrob. Agents Chemother.* **61**, e02033-16 (2017).
44. Kjos, M., Nes, I. F. & Diep, D. B. Mechanisms of resistance to bacteriocins targeting the mannose phosphotransferase system. *Appl. Environ. Microbiol.* **77**, 3335–3342 (2011).
45. Arbulu, S. & Kjos, M. Revisiting the multifaceted roles of bacteriocins. *Microb. Ecol.* **87**, 41 (2024).
46. Sundermann, A. J. et al. Whole-genome sequencing surveillance and machine learning of the electronic health record for enhanced healthcare outbreak detection. *Clin. Infect. Dis.* **75**, 476–482 (2022).
47. Prijbelski, A., Antipov, D., Meleshko, D., Lapidus, A. & Korobeynikov, A. Using SPAdes De Novo Assembler. *Curr. Protoc. Bioinformatics* **70**, e102 (2020).
48. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
49. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
50. Jolley, K. A., Bray, J. E. & Maiden, M. C. J. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. *Wellcome Open Res.* **3**, 124 (2018).
51. Seemann, T. mlst: scan contig files against PubMLST typing schemes. *GitHub* <https://github.com/tseemann/mlst> (2015).
52. Harris, S. R. SKA: Split kmer analysis toolkit for bacterial genomic epidemiology. *GitHub* <https://doi.org/10.1101/453142> (2018).
53. Maechler, F. et al. Split k-mer analysis compared to cgMLST and SNP-based core genome analysis for detecting transmission of vancomycin-resistant enterococci: results from routine outbreak analyses across different hospitals and hospitals networks in Berlin, Germany. *Microb. Genom.* **9**, mgen000937 (2023).
54. Liese, J. et al. Expansion of vancomycin-resistant *Enterococcus faecium* in an academic tertiary hospital in southwest Germany: a large-scale whole-genome-based outbreak investigation. *Antimicrob. Agents Chemother.* **63**, e01978-18 (2019).
55. Seemann, T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics* **30**, 2068–2069 (2014).
56. Schwengers, O. et al. Bakta: rapid and standardized annotation of bacterial genomes via alignment-free sequence identification. *Microb. Genom.* **7**, 000685 (2021).
57. Bastian, M., Heymann, S. & Jacomy, M. Gephi: an open source software for exploring and manipulating networks. In *Proc. 3rd International AAAI Conf. on Web and Social Media* 361–362 (Association for the Advancement of Artificial Intelligence, 2009).
58. Page, A. J. et al. Roary: rapid large-scale prokaryote pan genome analysis. *Bioinformatics* **31**, 3691–3693 (2015).
59. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
60. van Heel, A. J. et al. BAGEL4: a user-friendly web server to thoroughly mine RiPPs and bacteriocins. *Nucleic Acids Res.* **46**, W278–W281 (2018).
61. Medema, M. H. et al. antiSMASH: rapid identification, annotation and analysis of secondary metabolite biosynthesis gene clusters in bacterial and fungal genome sequences. *Nucleic Acids Res.* **39**, W339–W346 (2011).
62. Seemann, T. Abricate: mass screening of contigs for antimicrobial and virulence genes. *GitHub* <https://github.com/tseemann/abricate> (2015).
63. Feldgarden, M. et al. AMRFinderPlus and the Reference Gene Catalog facilitate examination of the genomic links among antimicrobial resistance, stress response, and virulence. *Sci. Rep.* **11**, 12728 (2021).
64. Carattoli, A. et al. In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. *Antimicrob. Agents Chemother.* **58**, 3895–3903 (2014).
65. Chen, L., Zheng, D., Liu, B., Yang, J. & Jin, Q. VFDB 2016: hierarchical and refined dataset for big data analysis—10 years on. *Nucleic Acids Res.* **44**, D694–D697 (2016).
66. Wick, R. R., Judd, L. M., Gorrie, C. L. & Holt, K. E. Unicycler: resolving bacterial genome assemblies from short and long sequencing reads. *PLoS Comput. Biol.* **13**, e1005595 (2017).
67. Nanoporetech dorado: Oxford Nanopore's Basecaller. *GitHub* <https://github.com/nanoporetech/dorado> (2022).
68. Kim, J.-H. & Mills, D. A. Improvement of a nisin-inducible expression vector for use in lactic acid bacteria. *Plasmid* **58**, 275–283 (2007).
69. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).
70. Dammann, A. N. et al. Group B *Streptococcus* capsular serotype alters vaginal colonization fitness. *J. Infect. Dis.* **225**, 1896–1904 (2021).
71. Smith, A. B. et al. Enterococci enhance *Clostridioides difficile* pathogenesis. *Nature* **611**, 780–786 (2022).
72. Smith, A. B. et al. Liberation of host heme by *Clostridioides difficile*-mediated damage enhances *Enterococcus faecalis* fitness during infection. *mBio* **15**, e0165623 (2023).
73. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).

Acknowledgements

We gratefully acknowledge Y. Li, K. Waggle, H. Coyle and H. Manzer for their helpful contributions to this study, and V. Cooper and A. Rokes for assistance in performing whole-genome sequencing of EDS-HAT isolates. The study was funded by National Institutes of Health grants R01AI165519 (to D.V.T.), R01AI127472 (to L.H.H.) and R35GM138369 (to J.P.Z.). The funders had no role in study design, data collection and analysis, preparation of the paper or decision to publish.

Author contributions

E.G.M., L.H.H., J.P.Z. and D.V.T. designed the study. E.G.M., A.B.S., M.P.G., K.H., L.P. and A.J.S. performed the experiments and collected data. E.G.M., J.P.Z. and D.V.T. analysed and interpreted the data. E.G.M. and D.V.T. drafted the paper and figures. All authors reviewed and edited the paper.

Competing interests

The authors declare no competing interests.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41564-025-01958-0>.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41564-025-01958-0>.

Correspondence and requests for materials should be addressed to Daria Van Tyne.

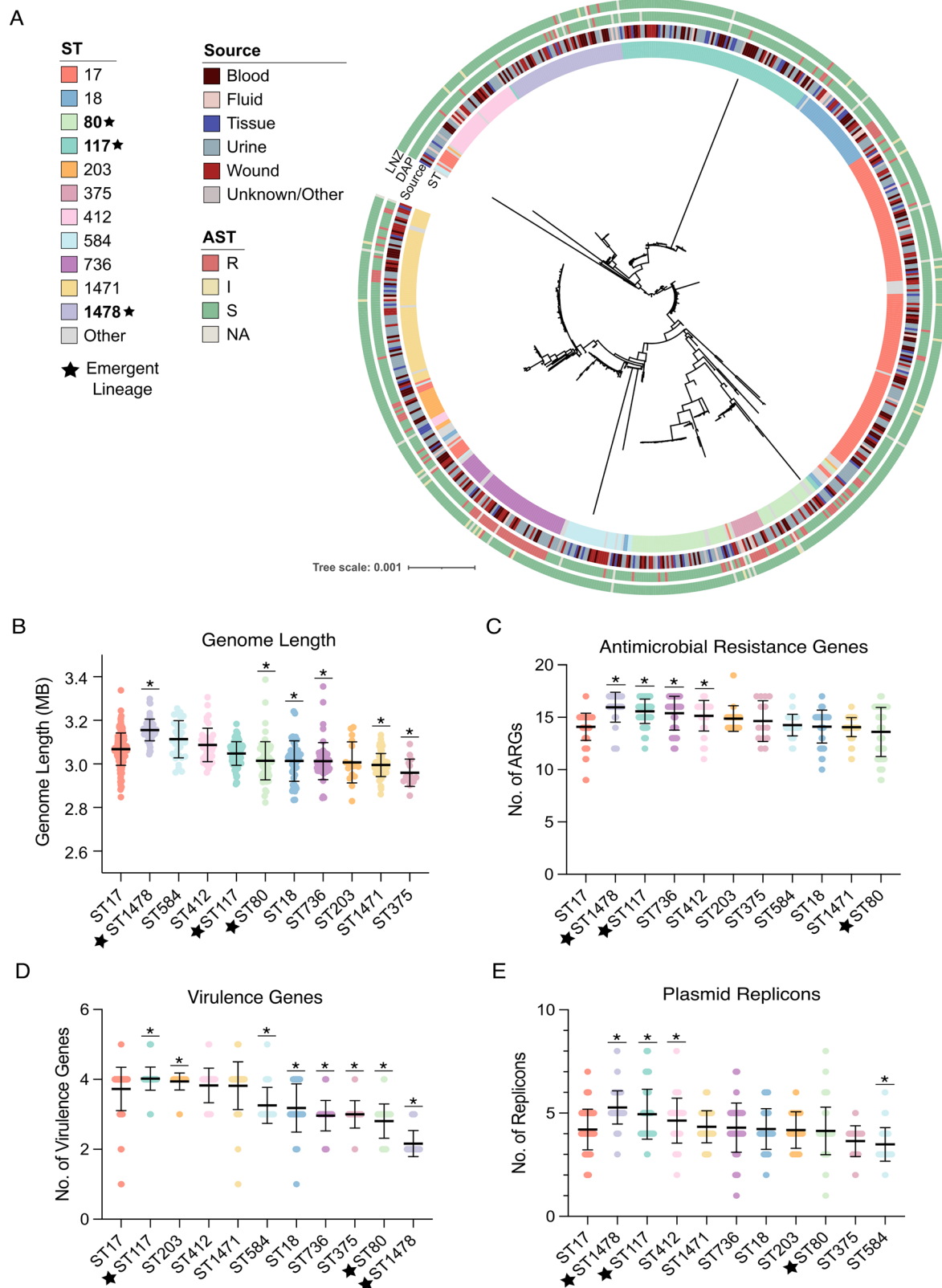
Peer review information *Nature Microbiology* thanks Morten Kjos and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

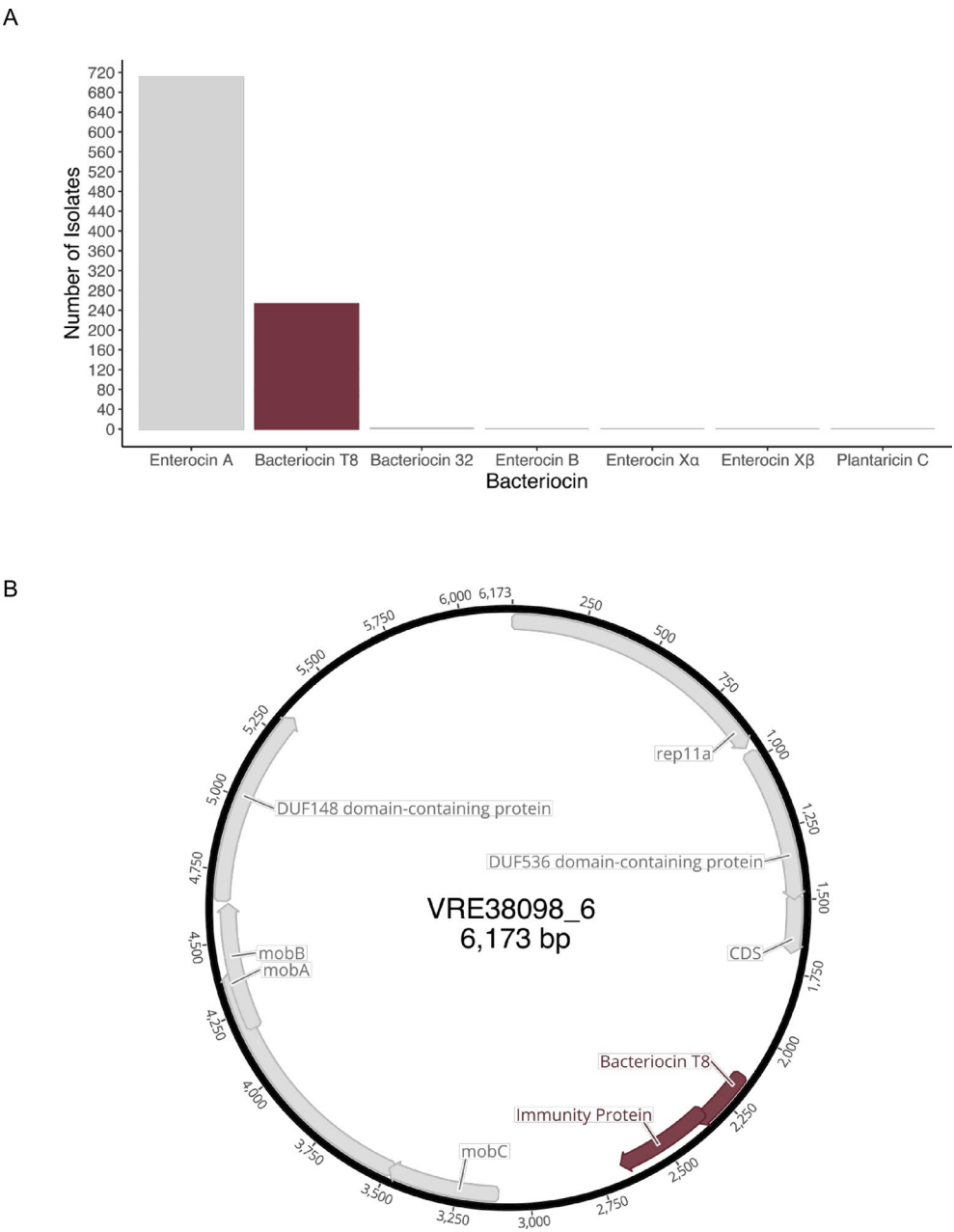
Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025



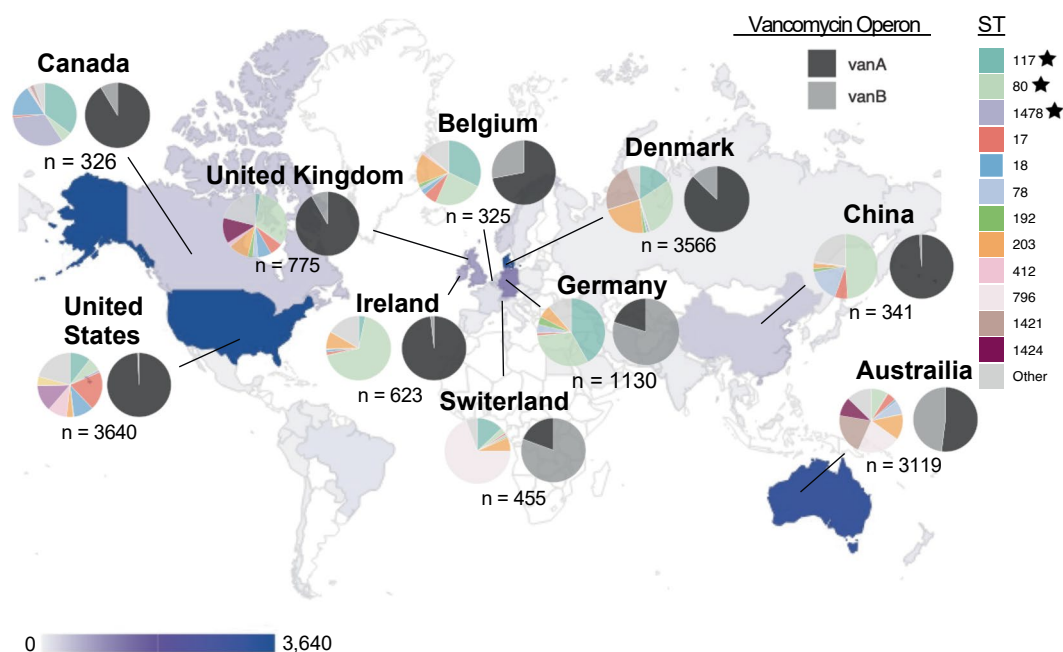
Extended Data Fig. 1 | Genetic relatedness and genomic features of 710 VREfm isolates from UPMC. (A) The midpoint-rooted phylogenetic tree was constructed from a single-copy core genome alignment. The scale bar represents nucleotide substitutions per site. Sequence types (ST) and isolation source are colored as indicated. Emergent lineages are noted by a black star. Antimicrobial susceptibility testing (AST) results for daptomycin (DAP) and linezolid (LNZ) were interpreted as resistant (R), intermediate (I), or susceptible (S). (B) Genome

length, (C) presence of antimicrobial resistance genes, (D) presence of virulence genes, and (E) presence of plasmid replicons within main VREfm lineages. Averages of genomic features within each ST were compared to the average seen in the previously dominant ST17 lineage using a two-sided t-test. Asterisks indicate p-values < 0.0045. Horizontal lines represent the average and error bars show the standard deviation for each group.



Extended Data Fig. 2 | Bacteriocin prevalence and genomic context of bacteriocin T8. (A) Distribution of bacteriocins within 710 VREfm isolates from UPMC. Bacteriocins were identified using BAGEL4 with sequence identity and

coverage thresholds of $\geq 95\%$. **(B)** Bacteriocin T8-encoding rep11a plasmid from the ST117 isolate VRE38098. Bacteriocin T8 and corresponding immunity factor genes are colored in burgundy.



Extended Data Fig. 3 | Global population structure of 15,631 VREfm genomes from human sources collected between 2002-2022. Geographical distribution of VREfm genomes pulled from NCBI. The number of genomes from each country

is shown from lowest (light grey) to highest (purple). Countries with >300 genomes are highlighted with the distribution of STs and vancomycin resistance alleles.

Reporting Summary

Nature Portfolio wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Portfolio policies, see our [Editorial Policies](#) and the [Editorial Policy Checklist](#).

Please do not complete any field with "not applicable" or n/a. Refer to the help text for what text to use if an item is not relevant to your study. For final submission: please carefully check your responses for accuracy; you will not be able to make changes later.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed
<input type="checkbox"/>	<input checked="" type="checkbox"/> The exact sample size (<i>n</i>) for each experimental group/condition, given as a discrete number and unit of measurement
<input type="checkbox"/>	<input checked="" type="checkbox"/> A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
<input type="checkbox"/>	<input checked="" type="checkbox"/> The statistical test(s) used AND whether they are one- or two-sided <i>Only common tests should be described solely by name; describe more complex techniques in the Methods section.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> A description of all covariates tested
<input type="checkbox"/>	<input checked="" type="checkbox"/> A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
<input type="checkbox"/>	<input checked="" type="checkbox"/> A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
<input type="checkbox"/>	<input checked="" type="checkbox"/> For null hypothesis testing, the test statistic (e.g. <i>F</i> , <i>t</i> , <i>r</i>) with confidence intervals, effect sizes, degrees of freedom and <i>P</i> value noted <i>Give P values as exact values whenever suitable.</i>
<input checked="" type="checkbox"/>	<input type="checkbox"/> For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
<input checked="" type="checkbox"/>	<input type="checkbox"/> For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Estimates of effect sizes (e.g. Cohen's <i>d</i> , Pearson's <i>r</i>), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about software and code	SPAdes v3.15.5, QUAST v5.2.0, Kraken2 (v2.0.8-β), PubMLST database with mlst v2.11, splitkrmer analysis v1.0 (SKA), PROKKA v1.14.5, Bakta v1.9.2, Roary v3.13.0, RAxML HPC v8.2.12, antiSMASH v7.1.0, ABRicate v1.01, AMRFinderPlus v3.12.8, PlasmidFinder and VFDB databases, BAGEL4 database, Unicycler v0.5.0, Dorado v 0.7.2
Data collection	
Data analysis	Data was visualized with RStudio, Prism Graph Pad, and Gephi. Statistical analyses were performed using Prism Graph Pad and RStudio.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors and reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Portfolio [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A description of any restrictions on data availability
- For clinical datasets or third party data, please ensure that the statement adheres to our [policy](#)

Genomic sequences for all 710 VREfm isolates can be found BioProject PRJNA475751 with accession numbers listed in Supplementary Table 1. Global dataset NCBI accession numbers can be found in Supplemental Table 7. Sequencing reads for mouse competition experiment can be found in BioProject PRJNA1217837 with accession numbers listed in Supplementary Datasets 6B.

Research involving human participants, their data, or biological material

Policy information about studies with [human participants or human data](#). See also policy information about [sex, gender \(identity/presentation\), and sexual orientation](#) and [race, ethnicity and racism](#).

Reporting on sex and gender

Not applicable as patient metadata was not collected

Reporting on race, ethnicity, or other socially relevant groupings

Not applicable as patient metadata was not collected

Population characteristics

Not applicable as patient metadata was not collected

Recruitment

Patient samples were collected through the Enhanced Detection System for Healthcare-Associated Transmission surveillance samples

Ethics oversight

IRB of the University of Pittsburgh gave ethical approval for this work

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences

☐ Behavioural & social sciences

☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

Data exclusions

Replication

Randomization

Blinding

Behavioural & social sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description

Research sample

Sampling strategy

Data collection

Timing

Data exclusions

Non-participation

Randomization

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	Population dynamics of vancomycin resistant <i>Enterococcus faecium</i> collected from patients over 6 years at a tertiary care center
Research sample	710 vancomycin resistant <i>Enterococcus faecium</i> isolates collected from patients between 2017-2022
Sampling strategy	Samples collected via the Enhanced Detection System for Healthcare-Associated Transmission (EDS-HAT) surveillance system at University of Pittsburgh Medical Center
Data collection	Patient samples were collected through EDSHAT
Timing and spatial scale	Collection between 2017-2020 from a tertiary care hospital, University of Pittsburgh Medical Center
Data exclusions	Not applicable
Reproducibility	Not applicable
Randomization	No randomization.
Blinding	No blinding

Did the study involve field work? ☐ Yes ☒ No

Field work, collection and transport

Field conditions	Not applicable
Location	Not applicable
Access & import/export	Not applicable
Disturbance	Not applicable

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology and archaeology
<input type="checkbox"/>	<input checked="" type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data
<input checked="" type="checkbox"/>	<input type="checkbox"/> Dual use research of concern
<input checked="" type="checkbox"/>	<input type="checkbox"/> Plants

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging

Antibodies

Antibodies used	
Validation	

Eukaryotic cell lines

Policy information about [cell lines and Sex and Gender in Research](#)

Cell line source(s)

Authentication

Mycoplasma contamination

Commonly misidentified lines
(See [ICLAC](#) register)

Palaeontology and Archaeology

Specimen provenance

Specimen deposition

Dating methods

☐ Tick this box to confirm that the raw and calibrated dates are available in the paper or in Supplementary Information.

Ethics oversight

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Animals and other research organisms

Policy information about [studies involving animals](#); [ARRIVE guidelines](#) recommended for reporting animal research, and [Sex and Gender in Research](#)

Laboratory animals

Five-week-old C57BL/six mice. Housing details can be found in the manuscript

Wild animals

no wild animals were used in this study

Reporting on sex

Male

Field-collected samples

no field collected samples were included in this study

Ethics oversight

Animal Care and Use Committees of the Children's Hospital of Philadelphia (IAC 18-001316)

Note that full information on the approval of the study protocol must also be provided in the manuscript.

Clinical data

Policy information about [clinical studies](#)

All manuscripts should comply with the ICMJE [guidelines for publication of clinical research](#) and a completed [CONSORT checklist](#) must be included with all submissions.

Clinical trial registration

Study protocol

Data collection

Outcomes

Dual use research of concern

Policy information about [dual use research of concern](#)

Hazards

Could the accidental, deliberate or reckless misuse of agents or technologies generated in the work, or the application of information presented in the manuscript, pose a threat to:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Public health
<input checked="" type="checkbox"/>	<input type="checkbox"/> National security
<input checked="" type="checkbox"/>	<input type="checkbox"/> Crops and/or livestock
<input checked="" type="checkbox"/>	<input type="checkbox"/> Ecosystems
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other significant area

Experiments of concern

Does the work involve any of these experiments of concern:

No	Yes
<input checked="" type="checkbox"/>	<input type="checkbox"/> Demonstrate how to render a vaccine ineffective
<input checked="" type="checkbox"/>	<input type="checkbox"/> Confer resistance to therapeutically useful antibiotics or antiviral agents
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enhance the virulence of a pathogen or render a nonpathogen virulent
<input checked="" type="checkbox"/>	<input type="checkbox"/> Increase transmissibility of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Alter the host range of a pathogen
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable evasion of diagnostic/detection modalities
<input checked="" type="checkbox"/>	<input type="checkbox"/> Enable the weaponization of a biological agent or toxin
<input checked="" type="checkbox"/>	<input type="checkbox"/> Any other potentially harmful combination of experiments and agents

Plants

Seed stocks	<input type="text"/>
Novel plant genotypes	<input type="text"/>
Authentication	<input type="text"/>

ChIP-seq

Data deposition

- ☐ Confirm that both raw and final processed data have been deposited in a public database such as [GEO](#).
- ☐ Confirm that you have deposited or provided access to graph files (e.g. BED files) for the called peaks.

Data access links <i>May remain private before publication.</i>	<input type="text"/>
Files in database submission	<input type="text"/>
Genome browser session (e.g. UCSC)	<input type="text"/>

Methodology

Replicates	<input type="text"/>
Sequencing depth	<input type="text"/>
Antibodies	<input type="text"/>
Peak calling parameters	<input type="text"/>
Data quality	<input type="text"/>

Software

Flow Cytometry

Plots

Confirm that:

- ☐ The axis labels state the marker and fluorochrome used (e.g. CD4-FITC).
- ☐ The axis scales are clearly visible. Include numbers along axes only for bottom left plot of group (a 'group' is an analysis of identical markers).
- ☐ All plots are contour plots with outliers or pseudocolor plots.
- ☐ A numerical value for number of cells or percentage (with statistics) is provided.

Methodology

Sample preparation

Instrument

Software

Cell population abundance

Gating strategy

- ☐ Tick this box to confirm that a figure exemplifying the gating strategy is provided in the Supplementary Information.

Magnetic resonance imaging

Experimental design

Design type

Design specifications

Behavioral performance measures

Imaging type(s)

Field strength

Sequence & imaging parameters

Area of acquisition

Diffusion MRI

☐ Used☐ Not used

Preprocessing

Preprocessing software

Normalization

Normalization template

Noise and artifact removal

Volume censoring

Statistical modeling & inference

Model type and settings

Effect(s) tested

Specify type of analysis: ☐ Whole brain ☐ ROI-based ☐ Both

Statistic type for inference

(See [Eklund et al. 2016](#))

Correction

Models & analysis

n/a | Involved in the study

- ☐ ☐ Functional and/or effective connectivity
- ☐ ☐ Graph analysis
- ☐ ☐ Multivariate modeling or predictive analysis

Functional and/or effective connectivity

Graph analysis

Multivariate modeling and predictive analysis