nature microbiology

ARTICLES

https://doi.org/10.1038/s41564-020-00806-7

Quantifying acquisition and transmission of *Enterococcus faecium* using genomic surveillance

Theodore Gouliouris[1,2,7], Francesc Coll[3,7]✉, Catherine Ludden[3], Beth Blane[1], Kathy E. Raven[1], Plamena Naydenova[1], Charles Crawley[2], Mili Estée Török[1,2], David A. Enoch[2,4,5], Nicholas M. Brown[2,4,5], Ewan M. Harrison[1,6], Julian Parkhill[1] and Sharon J. Peacock[1,2,4]✉

Nosocomial acquisition and transmission of vancomycin-resistant *Enterococcus faecium* (VREfm) is the driver for *E. faecium* carriage in hospitalized patients, which, in turn, is a risk factor for invasive infection in immunocompromised patients. In the present study, we provide a comprehensive picture of *E. faecium* transmission in an entire sampled patient population using a sequence-driven approach. We prospectively identified and followed 149 haematology patients admitted to a hospital in England for 6 months. Patient stools ($n = 376$) and environmental swabs ($n = 922$) were taken at intervals and cultured for *E. faecium*. We sequenced 1,560 isolates (1,001 stool, 559 environment) and focused our genomic analyses on 1,477 isolates (95%) in the hospital-adapted clade A1. Of 101 patients who provided two or more stool samples, 40 (40%) developed *E. faecium* carriage after admission based on culture, compared with 64 patients (63%) based on genomic analysis (73% VREfm). Half of 922 environmental swabs (447, 48%) were positive for VREfm. Network analysis showed that, of 111 patients positive for the A1 clade, 67 had strong epidemiological and genomic links with at least one other patient and/or their direct environment, supporting nosocomial transmission. Six patients (3.4%) developed an invasive *E. faecium* infection from their own gut-colonizing strain, which was preceded by nosocomial acquisition of the infecting isolate in half of these. Two informatics approaches (subtype categorization to define phylogenetic clusters and the development of an SNP cut-off for transmission) were central to our analyses, both of which will inform the future translation of *E. faecium* sequencing into routine outbreak detection and investigation. In conclusion, we showed that carriage and environmental contamination by the hospital-adapted *E. faecium* lineage were hyperendemic in our study population and that improved infection control measures will be needed to reduce hospital acquisition rates.

The gut commensal *Enterococcus faecium* is a frequent cause of nosocomial infection in immunocompromised and critically ill patients[1]. Healthcare-associated *E. faecium* is commonly resistant to numerous antibiotics including ampicillin and vancomycin, limiting treatment options[2]. The 30-day crude mortality rate for bacteraemia with VREfm is 35%, with higher mortality and hospital stay for bacteraemia caused by VREfm versus vancomycin-susceptible *E. faecium*[3]. The World Health Organization has classified VREfm as a high priority in their list of pathogens for research on new antibiotics[4]. The development of VREfm carriage in the gut is the single most important risk factor for VREfm bacteraemia[5,6]. Preventing *E. faecium* acquisition requires an understanding of reservoirs and transmission routes. It has been known since the 1990s that VREfm is shed into the hospital environment, where it may persist despite standard cleaning[7,8], and that being admitted to a room previously occupied by a VRE-positive patient is a risk factor for acquiring VRE[9]. This evidence is based on culture and bacterial typing techniques with low discrimination[10], which has limited the ability to establish relatedness between isolates from patients and those from their environment, the frequency with which patients carry more than one strain, and patterns and frequency of transmission. Whole-genome sequencing has been successfully used to study transmission of *E. faecium* between livestock and humans[11,12], in national surveillance programmes[13,14], across healthcare networks[15] and within hospitals[16–18]. Although previous hospital studies have largely focused on isolates associated with bacteraemia, a recent study[19] sequenced both carriage and clinical isolates prospectively. However, it did not measure acquisition rates, ascertain within-patient diversity, because a single isolate was sequenced per patient, or define the role of the hospital environmental as a reservoir, because it lacked environmental sampling.

In the present study, we address existing knowledge gaps by undertaking longitudinal genomic surveillance of *E. faecium* carriage, environmental contamination and transmission in a defined patient cohort.

Results

Study patients commonly carried VREfm. First, we determined the extent to which a putatively high-risk population of patients on two haematology wards carried and acquired *E. faecium* based on culture methods. The study ward characteristics, antimicrobial stewardship, infection control and cleaning policies are described in Supplementary Methods. We recruited 174 of 338 patients (51%) admitted to the two wards over 6 months (Fig. 1). Study participants were a median age of 61 years (interquartile range (IQR) 49–69, range 19–94 years), were admitted a median of once (IQR 1–2, total 281 admissions) and stayed a median of 16 d (IQR 7–27 d) (Supplementary Table 1). At least one stool for culture was provided by 149 of 174 participants, which resulted in the isolation of any *E. faecium*, ampicillin-resistant *E. faecium* (AREfm) and VREfm

[1]University of Cambridge, Cambridge, UK. [2]Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK. [3]London School of Hygiene & Tropical Medicine, London, UK. [4]Public Health England, London, UK. [5]Public Health England, Cambridge, UK. [6]Wellcome Sanger Institute, Hinxton, UK. [7]These authors contributed equally: Theodore Gouliouris, Francesc Coll. ✉e-mail: francesc.coll@lshtm.ac.uk; sjp97@medschl.cam.ac.uk

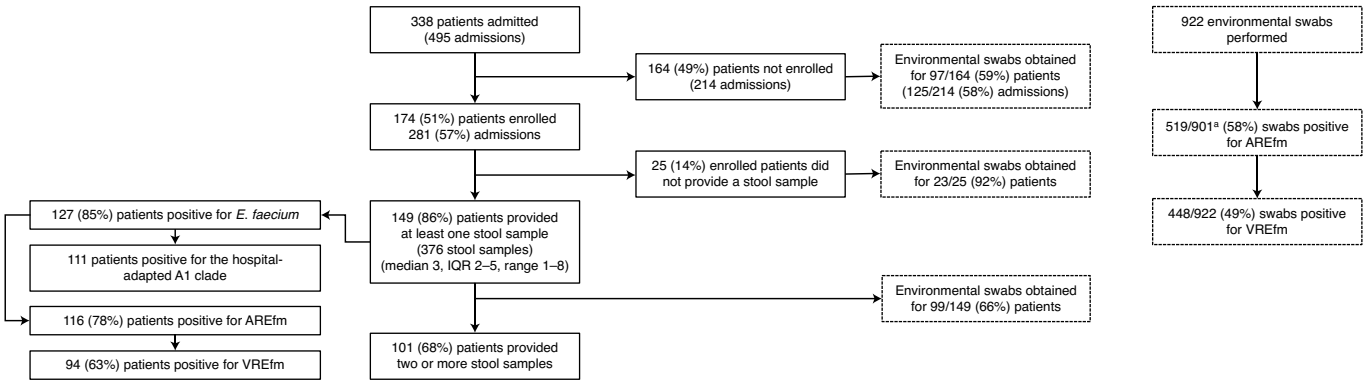


Fig. 1 | Study participants and *E. faecium* culture positivity. The number of patients sampled is shown as a subset of those enrolled, who in turn are a subset of those admitted to the two haematology wards. The overall number of environmental swabs taken from the hospital environment and for non-sampled patients is given. Culture positivity values for *E. faecium*, AREfm and VREfm are shown for both sampled participants and environmental swabs at the left- and right-hand sides of the figure, respectively. Supplementary Table 2 provides a breakdown of positivity rates for different environmental sources. [a]Twenty-one swabs were negative for VREfm and were not cultured for AREfm.

from 85%, 78% and 63% of the 149 cases, respectively. Based on 101 patients who provided two or more stools, 40 (40%) cases acquired *E. faecium* (changed from culture negative to culture positive) after admission (Extended Data Fig. 1).

VREfm was ubiquitous in the ward environment. In parallel, we evaluated the extent to which the patient environment contained *E. faecium*. Nearly half of 922 environmental swabs (447, 48%) taken over 6 months were culture positive for VREfm, the positive proportion ranging from 36% for medical devices to 76% for non-touch areas (see Supplementary Table 2 for full details). No *E. faecium* was isolated by air testing. Environmental swab VREfm positivity was similar between the two wards (237/457 (52%) versus 166/327 (51%), for wards A and B respectively; χ^2 $P=0.76$), and between individual rooms versus multiple-bed bays (119/255 (47%) versus 196/374 (52%), respectively; χ^2 $P=0.16$). Of 41 swabs taken from bedroom/bathroom areas after patient discharge and routine cleaning, 13 (32%) and 8 (20%) were positive for AREfm and VREfm, respectively. Deep cleaning was undertaken on ward B over a 3-d period during the study, when patients were moved elsewhere. This failed to eradicate *E. faecium* in 4/43 (9%) sampled locations before patients returned to the ward. These four environmental isolates collected immediately after deep cleaning were all genetically related (0–1 single nucleotide polymorphism (SNP)) to isolates collected just before cleaning, demonstrating that bacteria in these four sites persisted through decontamination. Any benefit of deep cleaning was short-lived, because around half of sampled sites were positive within 3 d of patient return (Supplementary Table 2). Isolates from positive sites within 3 d of patient return were mostly related (14/18, 0–4 SNPs) to isolates collected immediately before cleaning, demonstrating that re-establishment of environmental contamination was mostly caused by bacteria that were probably reintroduced by colonized patients.

Delineation of *E. faecium* subtypes. Having established that carriage of drug-resistant *E. faecium* was highly endemic and that environmental contamination was ubiquitous, we sought to update evidence for *E. faecium* relatedness in the two reservoirs using a contemporary, genome-based approach. We sequenced 1,560 isolates (1,001 from stool, 559 from the environment) and confined further analyses to 1,477 isolates (95%) assigned to the hospital-adapted clade A1 (see Supplementary Methods for details of how clade A1 isolates were defined). Of these, 943 were from 263 stools/111 patients (Fig. 1), and 534 were from the environment. Clade A1 is an AREfm lineage which frequently acquires vancomycin resistance[20]. We restricted our genomic analysis to this clade because it accounts for the vast majority of invasive infections[14]. We further divided the clade A1 population into discrete and non-overlapping genetic clusters based on monophyletic groups in the whole-genome phylogeny, such that isolates within each group were no more than 20 SNPs different from each other, which were referred to as 'subtypes'. We first chose an arbitrary cut-off of 50 SNPs to define monophyletic groups in the phylogeny. We counted pairwise SNP differences between isolates of the same and different 50-SNP clusters within each stool sample (Fig. 2a), before and after removing recombination. With one exception, all isolates differing by fewer than 20 SNPs had limited recombination and always belonged to the same sequence type (ST). Based on these results, we selected 20 SNPs as the threshold to define *E. faecium* subtypes (see Supplementary Methods and Extended Data Fig. 2 for details). We identified 115 genetically distinct *E. faecium* subtypes distributed across patient and environmental sources (31 from stool alone, 24 from the environment alone and 60 from both) (Fig. 3a,b). We found a very good correlation between the clustering provided by subtypes ($n=115$), STs ($n=55$) and BAPS (Bayesian Analysis of Population Structure)[21] clusters ($n=25$), where subtypes provided the highest discrimination and BAPS clusters the lowest (Supplementary Data 1).

Carriage of multiple *E. faecium* subtypes was common. Carriage of multiple *E. faecium* strains is an important confounder for outbreak investigations that include an evaluation of stool carriage isolates, because a non-outbreak strain could be erroneously selected. Furthermore, standard typing methods may not distinguish between even distantly related strains carried by the same individual. We used genomic data to re-evaluate the question of mixed-strain carriage, sequencing numerous isolates cultured from 185 stools (109 patients) that had two or more primary plate colonies sequenced (median 5, IQR 3–5, total 865 colonies) (see Supplementary Methods for further details). Within the limits of detection of our methods, we found that just over half of all stools (94 stools from 63 patients) contained at least two ($n=83$), three ($n=10$) or four subtypes ($n=1$), providing clear evidence that mixed-strain *E. faecium* carriage is common. When patients were colonized with multiple subtypes, isolate pairs from the same stool belonging to different subtypes had a median of 235 SNPs (IQR 198–289), whereas isolate pairs of the same subtype had a median of 0 SNP (IQR 0–1).

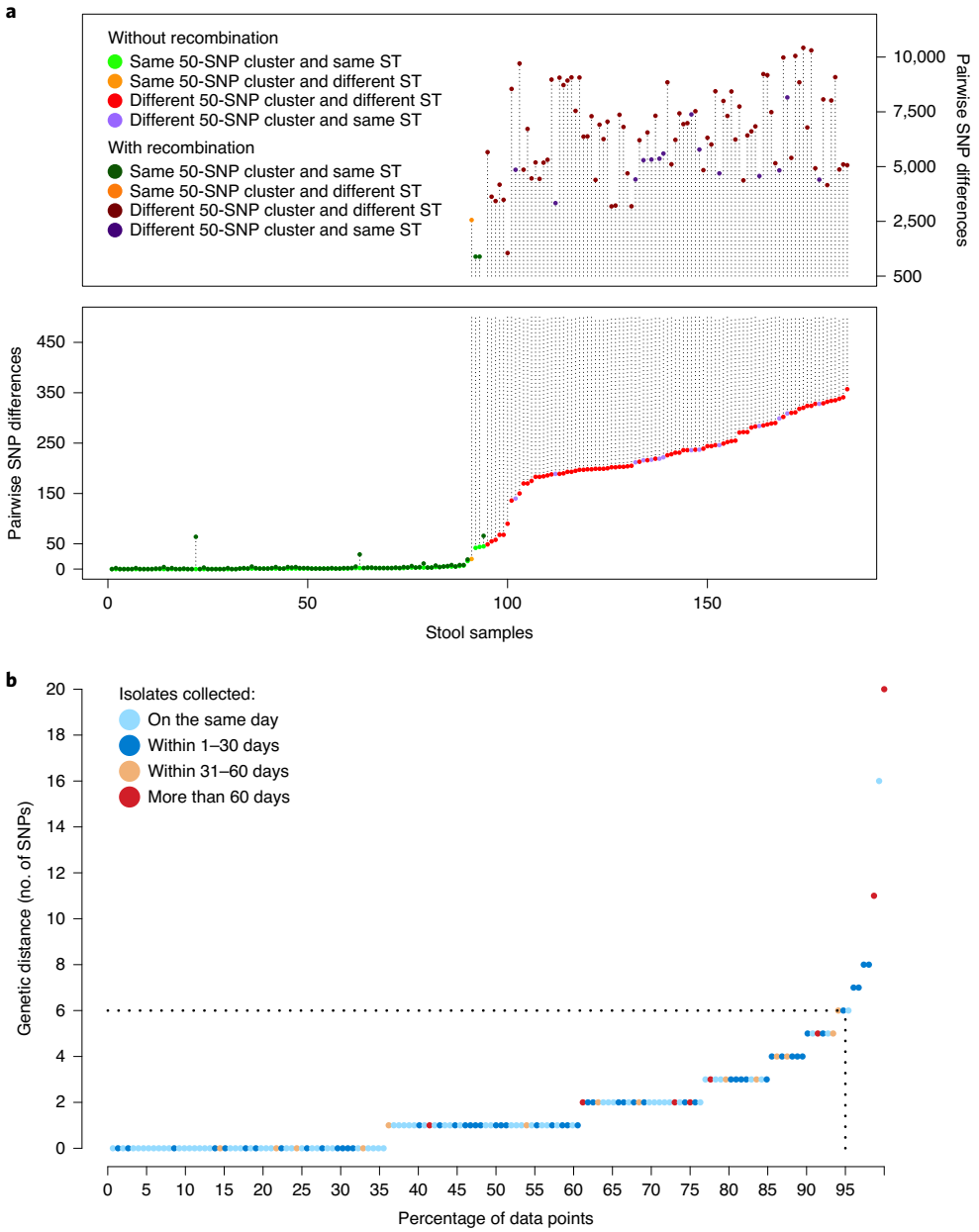


Fig. 2 | *E. faecium* within host diversity. a, Pairwise SNP differences between isolates of the same stool before and after removing recombination for a total of 185 stool samples. In samples with all isolates belonging to the same 50-SNP cluster, the maximum distance is plotted. In samples with multiple 50-SNP clusters, the minimum SNP distance between isolates from different clusters is plotted. Colours are used to indicate whether isolates belong to the same cluster and ST. The length of the vertical dotted lines denotes the number of SNPs attributable to recombination. **b**, Maximum pairwise SNP distance among isolates of the same subtype and in the same patient ($n = 152$). The colour of the dot denotes the time span between isolates.

***E. faecium* carried by patients was commonly shared with their environment.** Having defined subtypes in the stool of individual patients, we then defined the frequency with which subtypes were shared between patients; 36/91 subtypes (40%) identified in stools were isolated from two or more patients. This included two highly dominant subtypes (denoted here as 47A (sequence type (ST)78) and 15A (ST80)), which were isolated from 25 and 30 participants, respectively, spanned the entire study and accounted for 243/943 (26%) of all stool isolates (Fig. 3a). *E. faecium* subtypes in stool were often present in the ward environment (60/91, 66%). This was particularly the case for subtypes isolated from multiple patients, which were over-represented in the environment compared with subtypes isolated from single patients (32/36, 89% versus 28/55, 51%; Fisher's exact test, $P < 0.001$). Subtypes carried by multiple patients were particularly associated with contamination of communal bathrooms and medical devices (23/36 versus 9/55 and 11/36 versus 2/55, respectively; $P < 0.001$ for both).

Nosocomial acquisition of *E. faecium* and VREfm was common. Almost two-thirds of patients (64/101) acquired one or more *E. faecium* subtypes through a total of 111 acquisition events (changed from subtype negative to subtype positive), which is nearly three times the number of acquisitions detected by culture alone ($n = 40$). Culture underestimated *E. faecium* acquisition because it could not detect the acquisition of new subtypes by individuals already colonized with *E. faecium*. This equates to an acquisition rate of 59.4/100

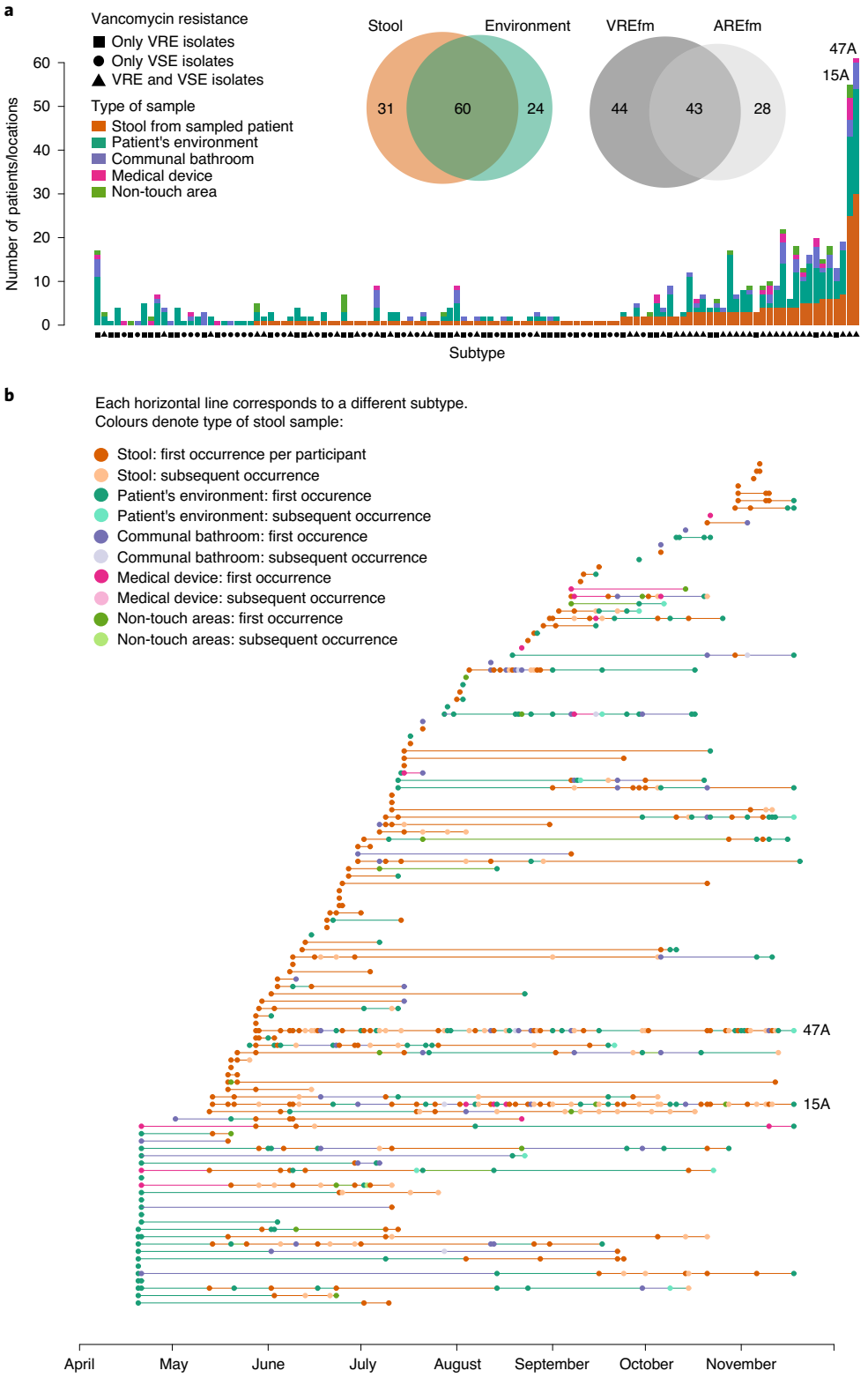


Fig. 3 | Frequency and time span of *E. faecium* subtypes. a, Frequency of *E. faecium* subtypes in stool and the environment. Each of the 115 bars represents a different subtype ordered by increasing frequency in patient stools. Venn diagrams show distribution of vancomycin resistance (right) and place of isolation (left). VSE, vancomycin-susceptible enterococcus. **b**, Isolation of 115 *E. faecium* subtypes over time. Each horizontal line represents a distinct subtype. Circles represent instances when subtypes were isolated, and the colour of the circles the source of isolation (stool or environmental). The number of dark brown circles for each subtype denotes the number of subjects colonized with it. Horizontal lines joining circles of the same subtype are coloured based on the source of the preceding sample.

admissions compared with 21.4/100 admissions based on the number of admissions/readmissions by study patients over 6 months ($n=187$). The two most common subtypes (47A (ST78) and 15A (ST80)) accounted for 28% of acquisition events; 81 acquisitions (in 52 patients) were VREfm and 30 acquisitions (in 26 patients) were AREfm. In addition to the 81 VREfm subtype acquisitions, 4 patients had a vancomycin-susceptible subtype detected in the stool that switched to the vancomycin-resistant type due to *vanA*, which can be explained by gene acquisition or the presence of a mixed population of susceptible/resistant isolates. This indicates that the dominant mode of VREfm acquisition within this study was through transfer of already vancomycin-resistant subtypes rather than horizontal gain of vancomycin resistance genes.

Common *E. faecium* subtypes were not more tolerant to hospital disinfectants. As *E. faecium* may persist in the hospital environment through increased tolerance to hospital disinfectants such as chlorhexidine and alcohol[22,23], we investigated whether the two most common subtypes had higher minimum inhibitory concentrations (MICs) to these two agents, or higher tolerance to isopropanol as previously described[23]. MICs to chlorhexidine and alcohol were not higher, on average (Extended Data Fig. 3a,b), than that of other clade A1 subtypes, and nor were they more tolerant to isopropanol (Extended Data Fig. 3c). These results suggest that factors other than increased tolerance to disinfectants may be responsible for the higher frequency of these subtypes.

SNP cut-off to detect recent *E. faecium* transmission. Although acquisition implies transmission, we applied a more stringent cut-off to each subtype to quantify recent *E. faecium* transmission between patients and/or their environment. This was based on the rationale that the maximum genetic diversity found in the same patient defines the amount of diversity that could potentially be transferred from one person to another. We thus quantified the amount of diversity in subtypes with at least two available isolates from the same patient (total of 152 patient–subtype combinations from 104 patients). Within-host subtype diversity in our study population was six or fewer SNPs in 95% of comparisons (Fig. 2b). Applying this cut-off to the acquisition events showed that 78 (70%) of the acquired subtypes were highly related (median 0, IQR 0–2 SNPs, maximum 6 SNPs) to an isolate from a previously sampled patient, supporting recent transmission. Epidemiological analysis of these 78 putative donor–recipient pairs demonstrated that 61 (78%) pairs had resided in the same location (bay, room or ward) at the same time or within 7 d, providing strong epidemiological evidence for transmission (Table 1, Supplementary Table 3 and Supplementary Data 2).

We then applied the six-SNP cut-off to all patients positive for the hospital-adapted clade A1 ($n=111$). Figure 4 shows a visual representation of transmission between patients and their environment using a network that combined bacterial relatedness and strength of epidemiological links. A high proportion of patients (67/111) had strong genetic (≤6 SNPs) and epidemiological links to one or more patients and/or their direct environment, supporting nosocomial transmission. There were four cases where two subtypes transmitted between the same donor–recipient pair, which could arise from a single event involving more than one strain or repeated transmission events. We also found evidence of multiple variants of the same subtype being transmitted in the same transmission event, as revealed by the range of SNP distances between isolates of the same transmitted subtype in the donor and the recipient patients (Supplementary Data 2).

Genetic and epidemiological links were used to reconstruct the temporal and spatial spread of each *E. faecium* transmission cluster (Supplementary Table 4). The size of transmission clusters ranged from two patients to eight patients. Eleven clusters consisted of a single transmission event, that is involving only two patients, and, with one exception, did not involve the hospital environment. The remaining 15 transmission clusters involved 3–8 patients and, in most cases (13/15), the hospital environment was found to be a plausible source of transmission (see Extended Data Fig. 4 for two examples).

Table 1 | Genomic and epidemiological evidence of nosocomial *E. faecium* transmission

	No. of subtypes (patients)	Median SNP distance (IQR)	Strong environmental links[a]
Acquired subtypes	111 (64)[b]	–	41/111
Genetically unlinked[c]	22 (19)	–	1/22
Genetically linked to[c]	89 (56)	–	40/89
Future sampled patients	11 (9)	–	1/11
Previously sampled patients	78 (52)	0 (0–2)	39/78
With weak epidemiological links	17 (15)	2 (0–2)	7/17
With strong epidemiological links	61 (43)	0 (0–1)	32/61
Same bay or room, same time	16 (11)	0 (0–0)	13/16
Same bay within 7 d	5 (5)	0 (0–1)	4/5
Same ward, same time	32 (24)	0 (0–1)	11/32
Same ward within 7 d	8 (8)	2.5 (0–3)	4/8
Index subtypes	116 (80)[d]	–	24/116
Genetically unlinked[c]	37 (32)	–	4/37
Genetically linked to[c]	79 (50)	–	20/79
Future sampled patients	28 (24)	–	5/28
Previously sampled patients	51 (43)	1 (0–3)	15/51
With weak epidemiological links	20 (20)	2 (0–3)	5/20
With strong epidemiological links	31 (25)	1 (0–2)	10/31
Same bay or room, same time	13 (11)	0 (0–2)	4/13
Same bay within 7 d	2 (2)	1 (0–2)	1/2
Same ward, same time	16 (14)	1 (0–2.5)	5/16
Same ward within 7 d	–	–	–

A total of 227 unique subtype–patient combinations were identified in 111 patients positive for clade A1, 38 of whom carried a single subtype and 73 multiple subtypes. Of the 227 subtype–patient combinations, 111 were acquired based on consecutive sampling and 116 were detected in the first available stool sample. For each subtype in each patient, evidence of nosocomial transmission was supported by genetic links to *E. faecium* isolates sampled in previous patients or environmental locations. Epidemiological data provided a second level of evidence of hospital transmission. [a]Genetically linked and with strong epidemiological links to previously sampled environmental sites. [b]Among patients with at least two available stool samples ($n=101$). [c]Using a cut-off of six SNPs to detect recent transmission of *E. faecium* subtypes during the study period. [d]Among all patients positive for the hospital-adapted clade ($n=111$).

Invasive *E. faecium* infections were associated with new VREfm acquisition. A serious consequence of *E. faecium* carriage is the development of invasive infection. This occurred in 6 study patients (3.4%), equating to 21 invasive infections per 1,000 admissions (see Table 2 and Supplementary Table 5 for details). Five of these patients had at least one stool cultured, all five of whom were positive for *E. faecium*. Comparison of stool- and disease-associated *E. faecium* genomes in each patient showed that the invasive and stool subtype was highly related in all five cases (0–5 SNPs). The invasive subtype was acquired after admission in three cases based on an earlier negative stool for the invasive subtype. Two patients had strong epidemiological links with another study case, with whom they shared an identical isolate based on a core genome comparison.

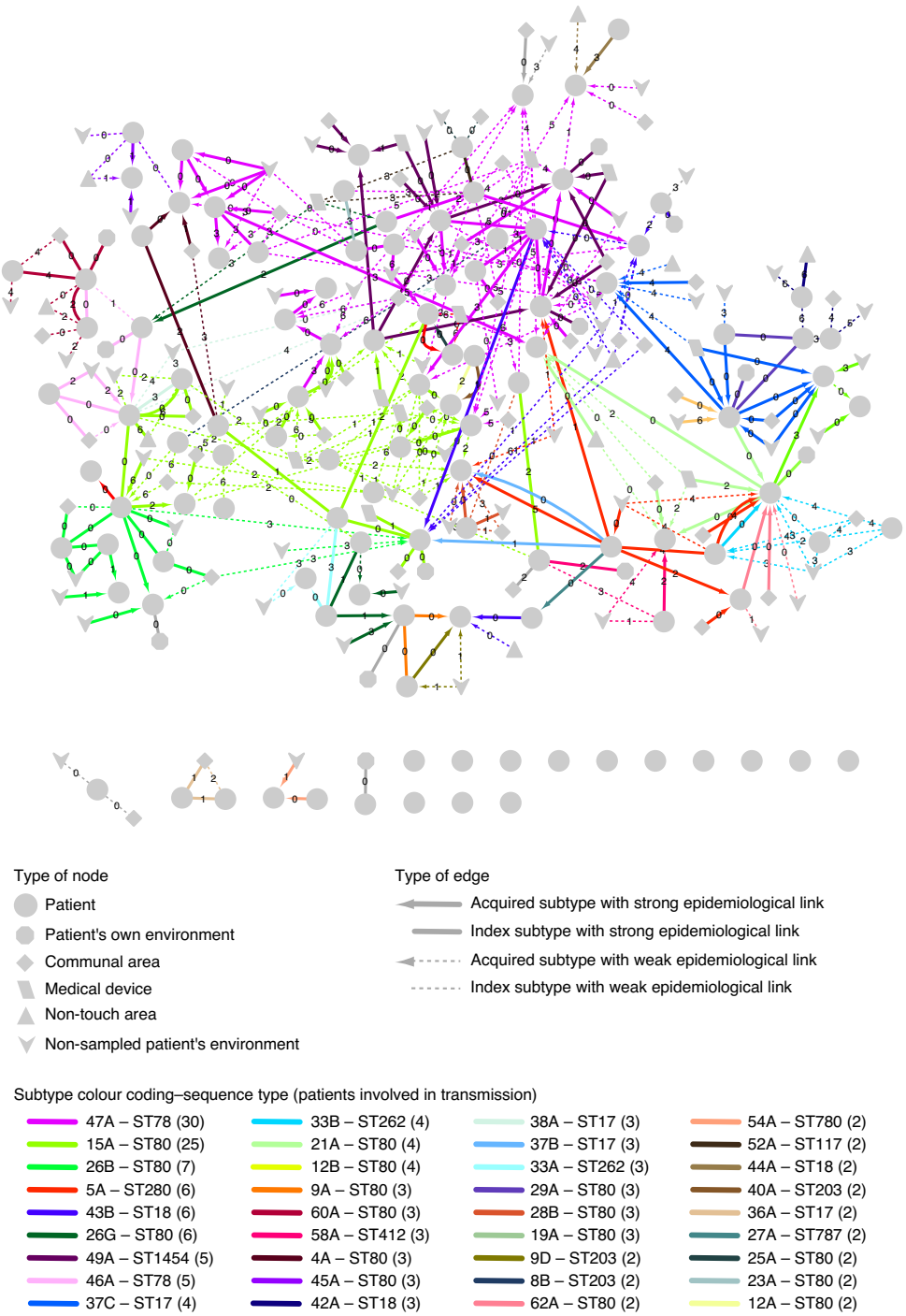


Fig. 4 | *E. faecium* transmission network. Transmission of clade A1 *E. faecium* subtypes between patients ($n = 111$) and their environment represented as a network. Only those environmental locations from which *E. faecium* was cultured that were genetically related (isolate within 0–6 SNPs) to at least one patient isolate are shown. For each patient node, lines show the shortest genetic link (within 0–6 SNPs) to a previously sampled patient or environmental location (putative transmission). Numbers on network edges show SNP distances. Edge colours show the subtype being transmitted. 'Acquired subtype' refers to subtypes not present in previously collected stool samples, whereas subtypes present in the first available stool are termed 'index subtypes'.

Discussion

Our study is a prospective observational study quantifying hospital acquisition rates for *E. faecium*, by combining in-depth longitudinal sampling and use of whole-genome sequencing. This approach allowed us to demonstrate that *E. faecium* acquisition rates were significantly higher than indicated by culture alone. We were able to establish the hospital location of such acquisitions, thanks to the density of clinical and environmental sampling and the integration of genomic and epidemiological data. Underpinning the above was the finding that mixed-strain carriage was common and the description of within-patient strain diversity at the core genome. We also provided clear evidence of invasive *E. faecium* infections originating from patients' own gut-colonizing strains, and demonstrated that nosocomial transmission is a key risk factor for subsequent

Table 2 | Relatedness of *E. faecium* associated with carriage and invasive disease

Patient ID	Clinical samples			Stool samples		Relatedness analysis		
	Sample type	No. of samples[a] (colonies sequenced)	Isolate subtype (ST) (SNP range)	No. of stool samples[a] (colonies sequenced)	Stool subtype(s) (ST)[b]	SNP distance: matching infection/stool subtype	Classification of stool subtype	Epidemiological link[c] (SNP distance with isolate from linked case)
C016	Blood	8 (56)	12B (80) (0–3)	6 (38)	**12B (80)**, 32A (262), 15A (80), 25A (80), 47A (78)	0–5	Index	None
C027	Blood	1 (1)	-	0	-	-	-	-
D041	Blood	1 (10)	47A (78) (0–0)	3 (15)	**47A (78)**, 49A (1454), 5A (280)	0–4	Acquired	Strong link with patient D034 (0)
C076	Surgical biopsy	1 (1)	15A (80)	2 (6)	**15A (80)**, 37B (17)	0–2	Index	Weak link with D012 (6)
D049	Blood	1 (10)	15A (80) (0–1)	2 (12)	**15A (80)**, 28B (80), 5A (280), 37B (17)	1–2	Acquired	Weak link with C084 (1)
C095	Blood	2 (20)	45A (80) (0–3)	1 (5)	**45A (80)**	1–4	Acquired	Strong link with C092 (0)

[a]Samples positive for *E. faecium*. [b]Subtypes identified across all stool samples from the same patient. [c]Strong link, defined as admission to the same bay, room or ward at the same time or within 7 d of discharge of the previous occupant; admission to the same ward, separated by more than 7 d, or admission to the study hospital, but to different wards, was considered to be a weak epidemiological link. Emboldened subtypes refer to the stool subtype also isolated from the invasive *E. faecium* infection.

colonization and infection. Our study confirms previous observations that VREfm can persist in the hospital environment despite standard cleaning[7,8], and that sharing hospital wards previously occupied by a VREfm-positive patient is a risk factor for acquiring VREfm, although we did not demonstrate direct acquisition from a previously contaminated room.

Carriage and environmental contamination by the hospital-adapted *E. faecium* lineage was hyperendemic in our study population, rates of which exceeded those reported previously from Cambridge University Hospitals NHS Foundation Trust (CUH)[20] and elsewhere[22,24–27]. This occurred despite the use of cleaning products and procedures with proven efficacy against VREfm, based on effectiveness in reducing rates of infection[26,27]. These high transmission rates were in sharp contrast to those detected for *Klebsiella pneumoniae* ($n=0$ patients)[25] and *Escherichia coli* ($n=20$ patients, data not published) in the same setting and patient population. Our study design supported the development of two informatics approaches (subtype categorization to define phylogenetic clusters and the development of an SNP cut-off for transmission) that will inform future genomic epidemiology studies of *E. faecium*, as well as translation of bacterial sequencing into routine outbreak detection and investigation. The use of a subtype provided important new insights into the rate at which patients carried more than one strain. Carriage of multiple *E. faecium* subtypes was common, a finding that can be explained by repeated acquisition events and/or a single transmission of a mixed population. Mixed-strain carriage indicates that the sensitivity of transmission detection based on stool testing will depend on the number of primary plate colonies tested from each sample. Subtype data also provided evidence for extensive overlap in the *E. faecium* populations residing in the patient gut and in their environment, which is consistent with a highly dynamic pattern of two-way spread between these reservoirs. The most prevalent *E. faecium* subtypes in patients were also the ones most often detected in the hospital environment and, related to this, subtypes associated with large transmission clusters were over-represented in the environment, suggesting that environmental surveillance could be an alternative to patient screening in this setting.

Our study found that the two most common subtypes accounted for 28% of acquisition events. In vitro susceptibility testing of representative isolates ruled out increased tolerance to disinfectants as one of the factors responsible for the higher frequency of these subtypes. A recent genomic epidemiological study, conducted in a German hospital[19], similarly found that the increase observed in VREfm was mostly attributable to the expansion of two major clones that were present throughout the hospital, where both intra-hospital patient-to-patient transmission and reintroduction from local hospitals had occurred. Future work is needed to elucidate the bacterial and epidemiological factors driving the expansion of particular *E. faecium* clones.

Our study has several limitations. First, we sampled fewer than 50% of patients admitted to the two haematology wards, and we did not sample healthcare workers. Unsampled carriers can explain, in part, why some acquired subtypes were not detected in any other patient (genetically unlinked), or why some patient pairs lacked strong epidemiological links despite carrying highly related isolates. Altogether, unsampled carriers would result in an underestimation of transmission as reported in the present study. Another limitation is that we did not sequence the full diversity of *E. faecium* in stool samples, but a maximum of five colonies. This can lead to some subtypes being wrongly classified as acquired instead of being present at low abundance in previous samples. Future studies will need to sequence directly from plate sweeps to capture the full heterogeneity within individuals. On the other hand, subtypes classified as 'index' could have been potentially acquired between admission and the time of sampling.

In conclusion, we have presented a most detailed genomic study of *E. faecium* hospital transmission. Whereas rates of acquisition and degree of endemicity of the global clade A1 *E. faecium* may vary between hospitals and regions, the mechanisms of transmission, infection sources and methodological developments presented here are likely to be generalizable to other settings. This is particularly true for haemato-oncology units, which are often associated with high rates of VREfm infection and colonization in the USA, Europe and Australia. The high endemicity and acquisition rates of drug-resistant *E. faecium* in haematology wards poses an important challenge to infection control. Patient screening, adequate provision of isolation and en-suite toilet facilities, improved and more frequent cleaning procedures, and stricter healthcare worker hygiene practices will all be needed, in addition to antimicrobial stewardship interventions to curtail this global epidemic.

Methods

Setting and study design. The study protocol was approved by the National Research Ethics Service (ref. 14/EE/1123), and the CUH Research and

Development Department (ref. A093285). We conducted a prospective observational study of consecutive patients admitted to two haematology wards at the CUH in the UK between 13 May and 13 November 2015. Patients were enrolled following informed written consent, after which stool samples were requested on admission, every week and at discharge, and cultured for *E. faecium*. Dates of hospital admission, ward transfers and bed positions were extracted electronically using the hospital bed-tracking system. Blood cultures taken for clinical reasons were recorded and cultures positive for *E. faecium* retrieved from the routine laboratory. Three weeks before the study started, environmental sampling for *E. faecium* was performed on both wards to establish baseline levels of contamination. Environmental sampling was also conducted throughout the study, in which communal bathrooms, toilets, non-touch surfaces (air vents and high-efficiency particulate air (HEPA) filters) and a range of medical devices were swabbed every fortnight. The day ward (used for outpatient chemotherapy administration) was swabbed at the start and midpoint of the study. In addition, two pooled swabs were taken in each of the patient bedside and bathroom areas on the day of discharge, for participants who provided no discharge stool sample and for non-participants. Air sampling was performed on three occasions on both wards. Supplementary Methods provide details of bed layout, swabbing and air-sampling methodology and locations, and infection control, cleaning and antibiotic-prescribing policies.

Microbiology, DNA sequencing and phylogenetic analyses. Isolation, identification and susceptibility testing of *E. faecium* from stool, environmental samples and blood cultures are described in Supplementary Methods. Antibiotic-selective media were used to isolate AREfm and VREfm from stools and the environment, with additional non-antibiotic media used in the first stool samples to isolate any *E. faecium*. Multiple *E. faecium* colonies were picked from primary cultures of positive stool samples to detect genetic diversity. Supplementary Methods describe the rationale for selecting isolates for sequencing. DNA was extracted, libraries prepared and 125-bp paired-end sequences determined on an Illumina HiSeq2000. Isolate genomes belonging to the hospital-adapted clade A1 (previously known as clonal complex (CC) 17)[28] were mapped to *E. faecium* Aus0004 strain (GenBank accession no. CP003351) using SMALT v.0.7.4 (http://www.sanger.ac.uk/science/tools/smalt-0). SNPs were identified from BAM files using SAMtools v.0.1.19 (ref. [29]) to create a whole-genome alignment. Mobile genetic elements and recombination events detected by Gubbins v.1.4.10 (ref. [30]) were removed to define the core genome. RAxML v.8.2.8 (ref. [31]) with 100 bootstraps was used to create a maximum likelihood tree from the core genome alignment. Pairwise genetic distances between isolates were calculated based on core genome SNPs. Isolates sequenced from blood cultures and stool from the same patient were compared to determine genetic relatedness and origin of the invasive isolate.

Genomic and epidemiological analyses to quantify nosocomial transmission. Genomic and epidemiological analyses were limited to the hospital-adapted A1 clade (see Extended Data Fig. 5 and Supplementary Methods for how these were identified). *E. faecium* acquisition was defined based on culture (transition from culture-negative to culture-positive stool), and genetic criteria (acquisition of a new subtype) for patients with at least two available stool samples ($n=101$). Acquisition rates were calculated as the number of *E. faecium* acquisitions divided by the number of admissions to haematology wards by these 101 patients ($n=187$). Subtypes present in the first available stool based on sequencing of a median of 5 independent colonies (IQR 5–6) were termed 'index subtypes' as opposed to 'acquired subtypes'. For each subtype in every patient, we considered previously sampled patients and environmental locations as possible sources, and identified the putative donor as the one with the genetically closest *E. faecium* isolate (within the 6-SNP cut-off). Admissions to the same bay, room or ward at the same time or within 7 d were classified as strong epidemiological links, whereas admissions in the same ward separated by more than 7 d, or to the study hospital but different wards, were classified as weak epidemiological links (see Supplementary Methods for a further explanation of this classification). A transmission network was constructed using R v.3.4.1 (ref. [32]) and visualized in Cytoscape v.3.2.0 (ref. [33]). Transmission plots were drawn using R to visualize the spatial and temporal spread of *E. faecium* subtypes.

Isopropanol and chlorhexidine MIC testing. Isolates were grown from −70 °C storage in glycerol on to Columbia blood agar plates overnight in 37 °C air. MIC testing was done using the broth microdilution method[34]. Isolates had a final dilution of 5×10^5 colony-forming units (c.f.u.) ml^{-1} in iso-sensitest broth in a 96-well, flat-bottomed microtitre plate. Isopropanol concentrations were tested at 0, 0.125, 0.25, 0.5, 1, 2, 4, 8, 12, 16, 24 and 32% (v:v). Chlorhexidine concentrations were tested at 0, 0.125, 0.25, 0.5, 1, 2, 4, 8, 12, 16, 24 and 32 $mg\,l^{-1}$. Isolates were tested in triplicate, and the MIC was found by visually inspecting the microtitre plate for the well containing the lowest concentration of biocide where no visible growth had occurred. The resulting MIC values are presented in Extended Data Fig. 3a,b.

Isopropanol tolerance assay. Isolates were grown from −70 °C storage in glycerol on to Columbia blood agar plates for overnight incubation in 37 °C air. A colony was then incubated overnight in 10 ml of brain–heart infusion broth at 37 °C in air. The overnight cultures were diluted to an optical density at 600 nm of 0.5 in phosphate-buffered saline (PBS). Then, 23% (v:v) isopropanol was added to 1 ml of the dilution and 23% PBS was added to another 1 ml dilution. Both were vortexed thoroughly and then incubated at room temperature for 5 min. The samples were then serially diluted between 10- and 1,000,000-fold in 7.5% Tween-80 in PBS, to inactivate the isopropanol; 50 µl of each dilution was evenly spread on to Muller–Hinton agar plates using an L-shaped spreader. Once dried, plates were incubated overnight at 37 °C in the air. Each isolate was tested in triplicate (biological replicate) and each dilution was plated in triplicate (technical replicate). Colonies were counted and averaged across technical triplicates, and the $\log_{10}$[c.f.u. reduction] calculated for each isolate between exposure to PBS and exposure to 23% isopropanol[23]. The $\log_{10}$[c.f.u. reduction] values are presented in Extended Data Fig. 3c.

Statistical methods. The null hypothesis (no difference between means) of median (across replicates) MICs and $\log_{10}$[c.f.u. reduction] values between lineage and subtype groups (that is, clade A1 versus basal, clade A1 versus 15A (ST80) and clade A1 versus 47A (ST78)) in Extended Data Fig. 2 was rejected for $P<0.05$ and was assessed using an unpaired Mann–Whitney *U*-test with a two-tailed *P* value. This test was performed using the *wilcox.test* function from R package stats (v.3.6.3).

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

Data availability

The whole-genome sequences from this study have been deposited at the European Nucleotide Archive (www.ebi.ac.uk/ena) under the study nos. PRJEB12937, PRJEB13191 and PRJEB13192. Individual accession nos. and isolate metadata are listed in Supplementary Data 1. Supplementary Data 2 includes the genetic and epidemiological links characterized in this study. Source data are provided with this paper.

Received: 19 February 2019; Accepted: 30 September 2020;
Published online: 26 October 2020

References

1. Arias, C. A. & Murray, B. E. The rise of the Enterococcus: beyond vancomycin resistance. *Nat. Rev. Microbiol.* **10**, 266–278 (2012).
2. Werner, G. et al. Emergence and spread of vancomycin resistance among enterococci in Europe. *Eurosurveillance* **13**, 19046 (2008).
3. Prematunge, C. et al. VRE and VSE bacteremia outcomes in the era of effective VRE therapy: a systematic review and meta-analysis. *Infect. Control Hosp. Epidemiol.* **37**, 26–35 (2016).
4. Tacconelli, E. et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327 (2017).
5. Taur, Y. et al. Intestinal domination and the risk of bacteremia in patients undergoing allogeneic hematopoietic stem cell transplantation. *Clin. Infect. Dis.* **55**, 905–914 (2012).
6. Moradigaravand, D. et al. Within-host evolution of *Enterococcus faecium* during longitudinal carriage and transition to bloodstream infection in immunocompromised patients. *Genome Med.* **9**, 119 (2017).
7. Falk, P. S., Winnike, J., Woodmansee, C., Desai, M. & Mayhall, C. G. Outbreak of vancomycin-resistant enterococci in a burn unit. *Infect. Control Hosp. Epidemiol.* **21**, 575–582 (2000).
8. Neely, A. N. & Maley, M. P. Survival of enterococci and staphylococci on hospital fabrics and plastic. *J. Clin. Mcrobiol.* **38**, 724–726 (2000).
9. Mitchell, B. G., Dancer, S. J., Anderson, M. & Dehn, E. Risk of organism acquisition from prior room occupants: a systematic review and meta-analysis. *J. Hosp. Infect.* **91**, 211–217 (2015).
10. Turabelidze, D., Kotetishvili, M., Kreger, A., Morris, J. G. & Sulakvelidze, A. Improved pulsed-field gel electrophoresis for typing vancomycin-resistant enterococci. *J. Clin. Microbiol.* **38**, 4242–4245 (2000).
11. Gouliouris, T. et al. Genomic surveillance of *Enterococcus faecium* reveals limited sharing of strains and resistance genes between livestock and humans in the United Kingdom. *mBio* **9**, e01780-18 (2018).
12. Arredondo-Alonso, S. et al. Plasmids shaped the recent emergence of the major nosocomial pathogen Enterococcus faecium. *mBio* **11**, e03284-19 (2020).
13. Lee, T. et al. A three-year whole genome sequencing perspective of *Enterococcus faecium* sepsis in Australia. *PLoS ONE* **15**, e0228781 (2020).
14. Raven, K. E. et al. A decade of genomic history for healthcare-associated *Enterococcus faecium* in the United Kingdom and Ireland. *Genome Res.* **26**, 1388–1396 (2016).

15. Pinholt, M. et al. WGS of 1058 *Enterococcus faecium* from Copenhagen, Denmark, reveals rapid clonal expansion of vancomycin-resistant clone ST80 combined with widespread dissemination of a *vanA*-containing plasmid and acquisition of a heterogeneous accessory genome. *J. Antimicrob. Chemother.* **74**, 1776–1785 (2019).
16. van Hal, S. J. et al. Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. *Microbial. Genom.* **2**, e000048 (2016).
17. Wang, G. et al. Evolution and mutations predisposing to daptomycin resistance in vancomycin-resistant *Enterococcus faecium* ST736 strains. *PLoS ONE* **13**, e0209785 (2018).
18. Raven, K. E. et al. Complex routes of nosocomial vancomycin-resistant *Enterococcus faecium* transmission revealed by genome sequencing. *Clin. Infect. Dis.* **64**, 886–893 (2017).
19. 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).
20. Brown, D. F. J. & Walpole, E. Evaluation of selective and enrichment media for isolation of glycopeptide-resistant enterococci from faecal specimens. *J. Antimicrob. Chemother.* **51**, 289–296 (2003).
21. Tonkin-Hill, G., Lees, J. A., Bentley, S. D., Frost, S. D. W. & Corander, J. Fast hierarchical Bayesian analysis of population structure. *Nucleic Acids Res.* **47**, 5539–5549 (2019).
22. Prieto, A. M. G. et al. The two-component system ChtRS contributes to chlorhexidine tolerance in *Enterococcus faecium*. *Antimicrob. Agents Chemother.* **61**, e02122-16 (2017).
23. Pidot, S. J. et al. Increasing tolerance of hospital *Enterococcus faecium* to handwash alcohols. *Sci. Transl. Med.* **10**, eaar6115 (2018).
24. de Regt, M. J. A. et al. High acquisition and environmental contamination rates of CC17 ampicillin-resistant *Enterococcus faecium* in a Dutch hospital. *J. Antimicrob. Chemother.* **62**, 1401–1406 (2008).
25. Ludden, C. et al. A One Health study of the genetic relatedness of *Klebsiella pneumoniae* and their mobile elements in the east of england. *Clin. Infect. Dis.* **70**, 219–226 (2020).
26. Passaretti, C. L. et al. An evaluation of environmental decontamination with hydrogen peroxide vapor for reducing the risk of patient acquisition of multidrug-resistant organisms. *Clin. Infect. Dis.* **56**, 27–35 (2013).
27. Grabsch, E. A. et al. Significant reduction in vancomycin-resistant enterococcus colonization and bacteraemia after introduction of a bleach-based cleaning–disinfection programme. *J. Hosp. Infect.* **82**, 234–242 (2012).
28. Lebreton, F. et al. Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *mBio* **4**, e00534-13 (2013).
29. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
30. Croucher, N. J. et al. Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15 (2015).
31. Stamatakis, A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014).
32. R Core Team. R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, 2017).
33. Smoot, M. E., Ono, K., Ruscheinski, J., Wang, P.-L. & Ideker, T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* **27**, 431–432 (2011).
34. Wiegand, I., Hilpert, K. & Hancock, R. E. W. Agar and broth dilution methods to determine the minimal inhibitory concentration (MIC) of antimicrobial substances. *Nat. Protoc.* **3**, 163–175 (2008).

Acknowledgements

We thank the nurses and healthcare workers on both wards at CUH for assistance with sample collection and the ward matrons, A. Green and C. Cowling, for their support. We thank L. Chaparadza and R. Swayne for assistance with sample and clinical data collection. We thank L. Drumright, A. Chaudhry and the EPIC and Clinical Informatics teams for providing patient movement data. We thank the library construction, sequencing and Pathogen Informatics teams at the Wellcome Trust Sanger Institute for assistance with Illumina sequencing. The flocked swabs were donated by Copan Italia SpA. The present study presents independent research supported by the Health Innovation Challenge Fund (WT098600, HICF-T5-342), a parallel funding partnership between the Department of Health and the Wellcome Trust. The views expressed in this article are those of the author(s) and not necessarily those of the Department of Health or Wellcome Trust. T.G. is a Wellcome Trust Research Training Fellow (103387/Z/13/Z). F.C. is a Wellcome Trust Sir Henry Postdoctoral Fellow (201344/Z/16/Z). C.L. is a Wellcome Trust Sir Henry Postdoctoral Fellow (110243/Z/15/Z). M.E.T. is a Clinician Scientist Fellow supported by the Academy of Medical Sciences, the Health Foundation and the National Institute for Health Research (NIHR) Cambridge Biomedical Research Centre. S.J.P. is an NIHR Senior Investigator.

Author contributions

T.G. and S.J.P. designed the study, wrote the study protocol and case record forms, obtained ethical and research and development approvals for the study, and supervised the data collection. M.E.T. supported ethics approvals. T.G., C.L. and C.C. were responsible for collecting samples, and clinical and epidemiological data. T.G., C.L., B.B. and P.N. isolated and identified *E. faecium*. B.B. and P.N. undertook susceptibility testing and extracted genomic DNA. N.M.B. and D.A.E. provided access to *E. faecium* cultures in the routine diagnostic microbiology laboratory and expert opinion relating to infection control. T.G. and F.C. undertook the epidemiological and bioinformatic analyses with contributions from J.P. and K.R. B.B. undertook susceptibility testing to disinfectants with contributions from E.M.H. J.P. supervised the genomic sequencing. F.C. and S.J.P. wrote the first draft of the manuscript. S.J.P. supervised and managed the study. All authors had access to the data, and read, contributed and approved the final manuscript.

Competing interests

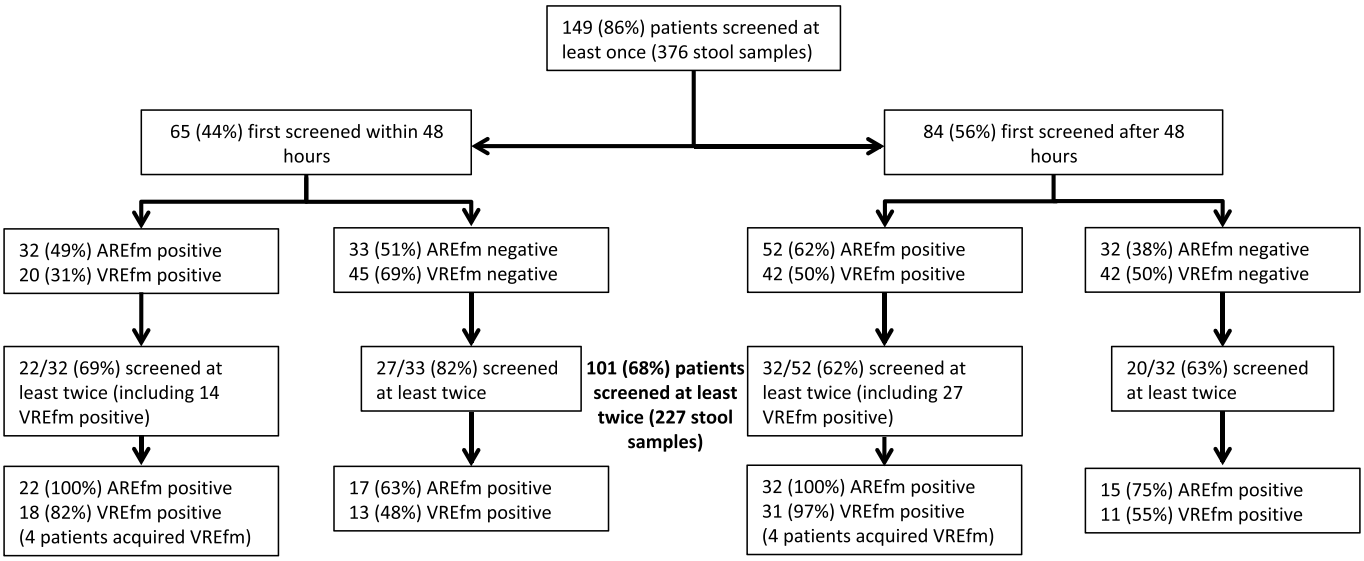
N.M.B. is on the advisory board for Discuva Ltd. S.J.P. is a consultant to Specific Technologies. S.J.P., J.P. and F.C. are consultants for Next Gen Diagnostics. The remaining authors declare no competing interests.

Additional information

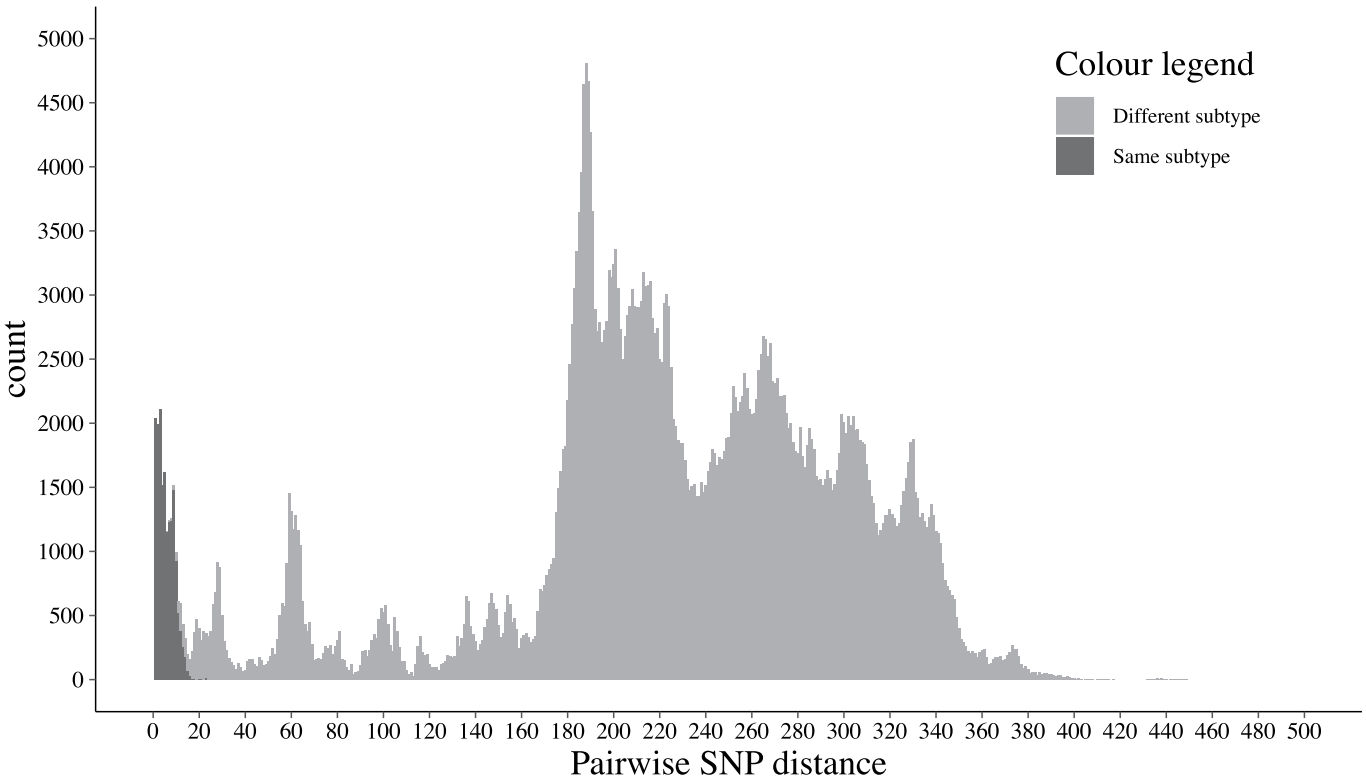
Extended data is available for this paper at https://doi.org/10.1038/s41564-020-00806-7.

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-020-00806-7.

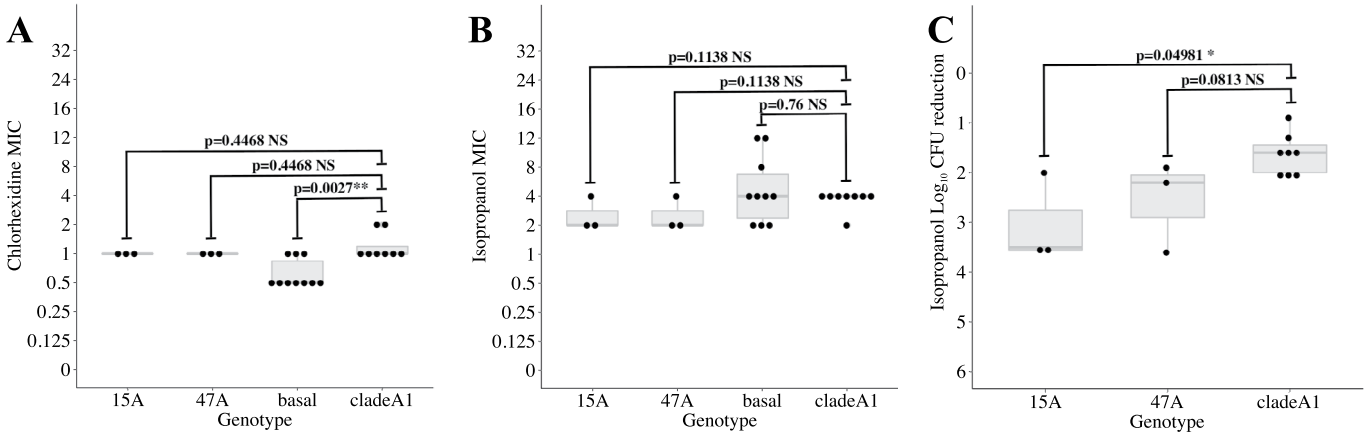
Correspondence and requests for materials should be addressed to F.C. or S.J.P.



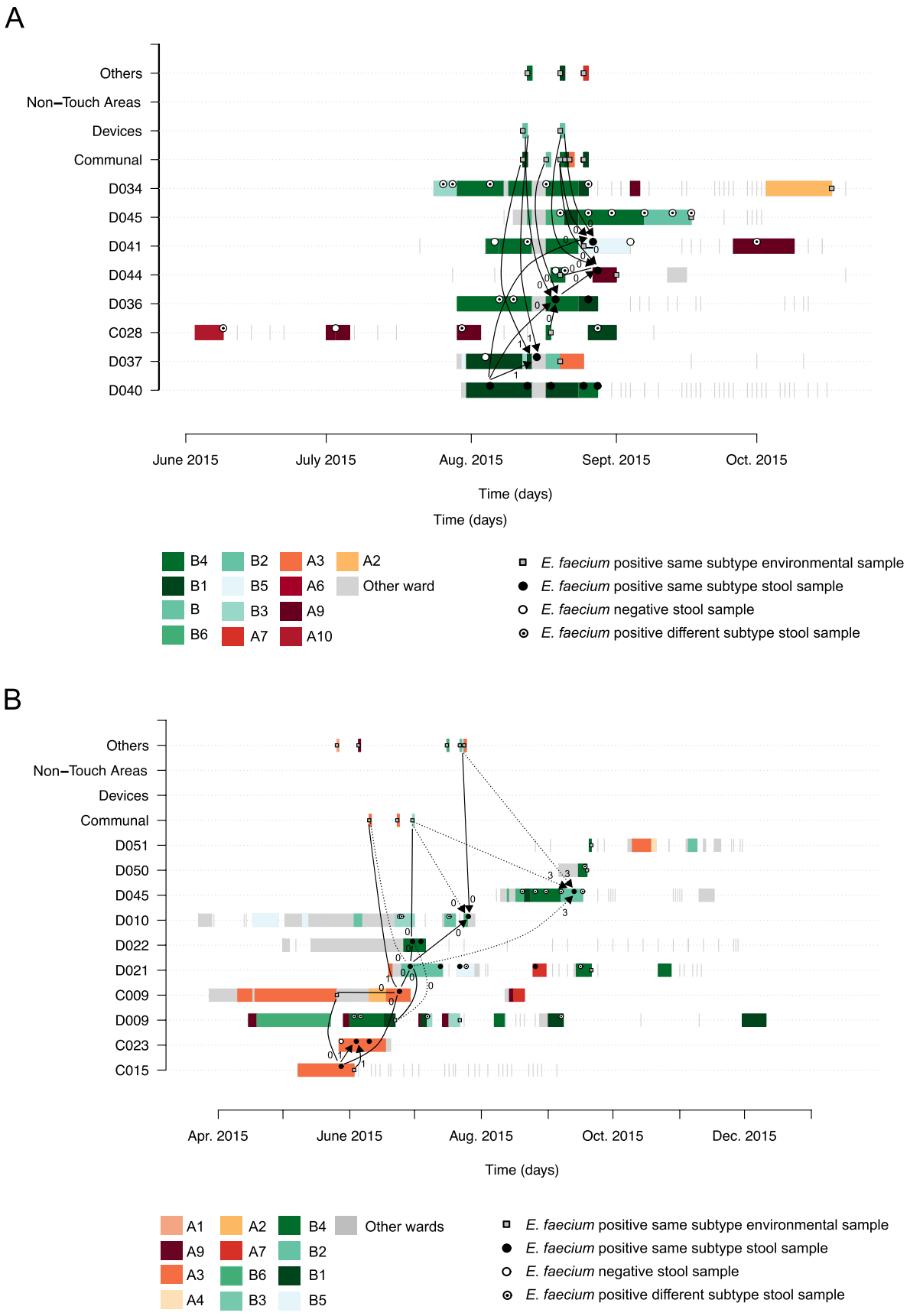
Extended Data Fig. 1 | *E. faecium* stool culture positivity during study. Diagram showing *E. faecium* positivity in patients who provided stool samples within (left branch) or after (right branch) 48 hours from index admission. Subsequent boxes show numbers of patients positive or negative for AREfm and VREfm, and, for patients screened at least twice, whether their positivity status changed, suggestive of *E. faecium* acquisition. A total of 40 cases acquired *E. faecium* based on culture, either by acquiring any type of *E. faecium* after being negative for it (17 and 15 patients in the left and right arms, respectively) or VREfm after being already positive for AREfm (4 and 4 patients in the left and right arms, respectively). Abbreviations: AREfm, ampicillin-resistant *E. faecium* (which may be vancomycin susceptible or resistant); VREfm, vancomycin-resistant *E. faecium*.



Extended Data Fig. 2 | Histogram of pairwise SNP differences between isolates of the same and different subtypes. Histogram of pairwise SNP differences between 943 clade A1 isolates from stool samples. SNP differences between isolates from the same subtype are shown in dark grey, and between isolates in different subtypes in light grey.

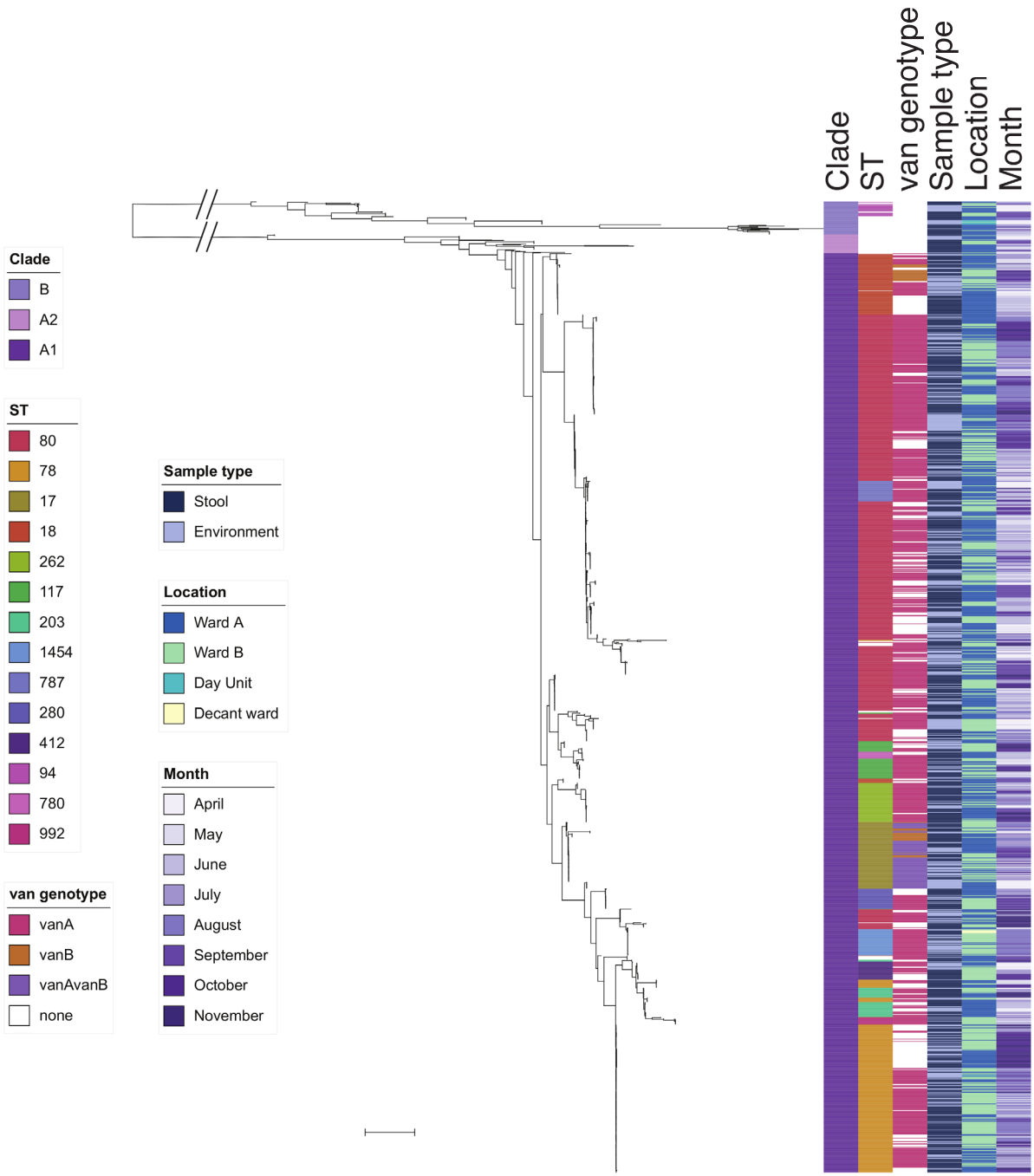


Extended Data Fig. 3 | Chlorhexidine and isopropanol susceptibility among selected E. faecium isolates. Chlorhexidine and isopropanol susceptibility testing results for a subset of phylogenetically representative *E. faecium* isolates (n=24 biologically independent samples) from the two major subtypes (15A/ST80 (n=3) and 47A/ST78 (n=3)), rest of subtypes in 'clade A1' (n=8) and 'basal' isolates (n=10) to clade A1. Each dot denotes the median MIC value (panels **a** and **b**) or median reduction in colony forming units (CFU) (panel **c**) across three independent replicates for each isolate tested. In the boxplots, the lower and upper hinges correspond to the first and third quartiles (the 25th and 75th percentiles), and the middle horizontal line to the median. P-values for two-tailed, unpaired Mann-Whitney are showed as NS (non-significant, $P > 0.05$), * ($P < 0.05$) or ** ($P < 0.01$).



Extended Data Fig. 4 | See next page for caption.

Extended Data Fig. 4 | Exemplars of *E. faecium* transmission clusters. Each row represents the hospital admission period(s) of patients with the exception of the top four rows, which show different environmental sources. Ward of admission is denoted as A or B, and the room numbered and color-coded. Visits to other hospital wards or areas are colored in grey. Positivity results for stool and environmental samples are shown as circles and squares, respectively. Blunt lines and arrowed lines are drawn to point to the putative sources of index and acquired subtypes respectively, the numbers adjacent to these lines indicating the minimum genetic distance observed between connected samples, which ranged from 0 to 6 SNPs. Solid and dotted lines denote strong and weak epidemiological links, respectively. (**a**) Exemplar of transmission cluster in the same ward (subtype 49A – ST1454). Strong genetic and epidemiological links point to transmission of this subtype in different rooms of ward B among patients D040, D037, D036, D044 and D041. Strong links to the hospital environment, including communal bathrooms and medical devices, suggest their involvement as reservoirs for onward transmission to patients. (**b**) Exemplar of transmission cluster spanning both hematology wards and involving 7 patients (subtype 26B – ST80). Strong genetic and epidemiological links point to transmission of this subtype in room A3 among patients C015, C023, C009 and D021, followed by spread in different rooms of ward B among patients D021, D022, D010 and D045.



Extended Data Fig. 5 | Midpoint rooted maximum likelihood tree based on SNPs in the core genes of 1,560 E. faecium isolates. *E. faecium* genomes (1,001 stool, 559 environmental) labeled by clade (B, A2, and A1), commonest sequence types (STs) (only those with more than 10 isolates shown), *van* genotype, source, ward of origin and month of isolation. Scale bar, ~10,000 SNPs.

nature research

Corresponding author(s): Francesc Coll; Sharon J. Peacock

Last updated by author(s): Aug 12, 2020

Reporting Summary

Nature Research 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 Research policies, see Authors & Referees and the Editorial Policy Checklist.

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
☐	☒ The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒ The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐ A description of all covariates tested
☒	☐ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☒	☐ 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)
☐	☒ For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐ Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), 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 availability of computer code

Data collection	Microsoft Access, Microsoft Excel, Stata v13 (StataCorp)
Data analysis	The following open-source bioinformatics tools were used: Prokka v1.11, SMALT v0.7.4, SAMtools v0.1.19, RAxML v8.2.8, Gubbins v1.4.10, Cytoscape v3.2.0, R v3.4.1, Roary (https://github.com/sanger-pathogens/Roary), snp-sites (https://github.com/sanger-pathogens/snp-sites), MLST Check (https://github.com/sanger-pathogens/mlst_check), ARIBA v2.5.0 (https://github.com/sanger-pathogens/ariba), pairwise-difference-count (https://github.com/simonrharris/pairwise_difference_count), fastbaps v1.0.0. The custom bash code used to run the bioinformatic pipelines and R code used to analyse the epidemiological data and create the figures in this manuscript are available through the corresponding author.

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/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research 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 list of figures that have associated raw data
- A description of any restrictions on data availability

Sequence data for all isolates have been deposited in the European Nucleotide Archive (ENA) (www.ebi.ac.uk/ena) under the study numbers PRJEB12937, PRJEB13191 and PRJEB13192. Individual accession numbers and isolate metadata are listed in Supplementary Data 1. Supplementary Data 2 includes the genetic and epidemiological links characterized in this study. The following databases were used: ResFinder 2012 database (https://cge.cbs.dtu.dk/services/ResFinder/), PHASTer (https://phaster.ca/), ISFinder (https://isfinder.biotoul.fr/), pubMLST website (https://pubmlst.org/efaecium/), RefSeq and Kraken.

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

Ecological, evolutionary & environmental sciences study design

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

Study description	The study was a prospective observational study of consecutive patients admitted to two hematology wards at the Cambridge University Hospitals NHS Foundation Trust (CUH).
Research sample	The sample is meant to represent the population of hematology patients. We did not preselect the initial research sample (number of patients) as this was determined by the number of patients admitted to the hematology wards (n=338) over a 6-month period. The final sample size was determined by the number of patients we could enroll (n=174) and, out of those, the ones we could obtain a stool sample from (n=149). A formal sample size calculation was not performed as this was an observational study with no interventions. The predicted E. faecium positivity rate was 50%.
Sampling strategy	Representative E. faecium specimens were isolated from patients' stool samples by sequencing multiple primary plate colonies (median 3, range 1-10, IQR 1-5) to ascertain carriage of more than one subtype, within-sample diversity and strain acquisition. Stools were cultured on each of a range of enterococcal selective media and broths with increasing antibiotic selective pressure. Up to 2 colonies from antibiotic-free plates, and up to 5 colonies from ampicillin-supplemented and VRE plates were selected. One colony of VREfm, or VSEfm if no VREfm grew was selected from culture plates from environmental swabs.
Data collection	Patients' data on hospital admission and ward transfers were extracted electronically from the hospital bed tracking system. Clinical and demographic data were extracted manually by trained clinicians.
Timing and spatial scale	Timing and spatial scale: Sampling took place between 13 May and 13 Nov 2015 in the two hematology wards. Stool samples were taken when new patients were admitted to the hematology wards, every week thereafter and at discharge. Environmental samples were obtained on discharge for study participants if stool samples were not available, and for non-participants. Additional environmental sampling was conducted every fortnight, except for the chemotherapy day unit which was swabbed at the start and midpoint of the study.
Data exclusions	No patient exclusion criteria were applied. Out of a total of 1682 E. faecium isolate genomes, we excluded those that did not pass genomic QC (n=24) or did not belong to the hospital-adapted clade (n=83). The rest were included.
Reproducibility	Not applicable as this is an observational study.
Randomization	Not applicable as this is an observational study.
Blinding	Not applicable as this is an observational study.

Did the study involve field work? ☐ Yes ☒ No

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
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☐	☒ Human research participants
☐	☒ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Human research participants

Policy information about studies involving human research participants

Population characteristics	Patients admitted to two hematology wards at the Cambridge University Hospitals NHS Foundation Trust (CUH) in the UK between 13 May and 13 Nov 2015. This focused on two adult hematology wards with a total of 27 beds. High-risk chemotherapy patients, allogeneic and most autologous hematopoietic stem cell transplant (HSCT) recipients were nursed on ward A.
Recruitment	Patients were enrolled following informed written consent, after which stool samples were requested on admission, every week and at discharge, and cultured for E. faecium. We recruited 174 of 338 patients (51%) admitted to the two wards over 6 months. Study participants were a median age of 61 years (IQR 49 to 69, range 19-94), were admitted a median of once (IQR 1 to 2, total 281 admissions), and stayed a median of 16 days (IQR 7 to 27 days).
Ethics oversight	The study protocol was approved by the National Research Ethics Service (ref: 14/EE/1123), and the Cambridge University Hospitals NHS Foundation Trust Research and Development Department (ref: A093285).

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	Not applicable
Study protocol	The study protocol was approved by the National Research Ethics Service (ref: 14/EE/1123), and the Cambridge University Hospitals NHS Foundation Trust Research and Development Department (ref: A093285).
Data collection	The study was conducted in two hematology wards at the Cambridge University Hospitals NHS Foundation Trust (CUH)
Outcomes	Not applicable. Outcomes were not measured