

# Microbial Genomics

## Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance. --Manuscript Draft--

Manuscript Number:	MGEN-D-15-00040R1
Full Title:	Evolutionary dynamics of <i>Enterococcus faecium</i> reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance.
Article Type:	Research Paper
Section/Category:	Microbial evolution and epidemiology: Population Genomics
Corresponding Author:	Sebastiaan J. van Hal, PhD Royal Prince Alfred Hospital AUSTRALIA
First Author:	Sebastiaan J. van Hal, PhD
Order of Authors:	Sebastiaan J. van Hal, PhD Camilla L.C. Ip, PhD M. Azim Anzari, PhD Daniel J. Wilson, PhD Bjorn Espedido, PhD Slade O. Jensen, PhD Rory Bowden, PhD
Abstract:	<p>Background: <i>Enterococcus faecium</i>, a major cause of hospital-acquired infections, remains problematic because of its propensity to acquire resistance to vancomycin, which currently is considered first-line therapy. Here, we assess the evolution and resistance acquisition dynamics of <i>E. faecium</i> in a clinical context using a series of 132 blood-stream infection isolates from a single hospital. All isolates, of which 49 (37%) were vancomycin resistant, underwent whole-genome sequencing.</p> <p>Results: <i>E. faecium</i> was found to be subject to high rates of recombination with little evidence of sequence importation from outside the local <i>E. faecium</i> population. Apart from disrupting phylogenetic reconstruction, recombination was frequent enough to invalidate MLST typing in the identification of clonal expansion and transmission events, suggesting that where available, whole-genome sequencing should be used in tracing the epidemiology of <i>E. faecium</i> nosocomial infections and establishing routes of transmission. Several forms of the Tn1549-like element-vanB gene cluster, that was exclusively responsible for vancomycin resistance, appeared and spread within the hospital during the study period. Several transposon gains and losses and instances of in situ evolution were inferred and although usually chromosomal, the resistance element was also observed on a plasmid background. There was qualitative evidence for clonal expansions of both vancomycin-resistant and vancomycin-susceptible <i>E. faecium</i> with evidence of hospital specific sub-clonal expansion.</p> <p>Conclusions: Our data are consistent with continuing evolution of this established hospital pathogen and confirm hospital vancomycin-susceptible and -resistant <i>E. faecium</i> patient transmission events, underscoring the need for careful consideration before modifying current <i>E. faecium</i> infection control strategies.</p>

# MICROBIAL GENOMICS

Research paper template

## Evolutionary dynamics of *Enterococcus faecium* reveals complex genomic relationships between isolates with independent emergence of vancomycin resistance.

---

### ABSTRACT

Background: *Enterococcus faecium*, a major cause of hospital-acquired infections, remains problematic because of its propensity to acquire resistance to vancomycin, which currently is considered first-line therapy. Here, we assess the evolution and resistance acquisition dynamics of *E. faecium* in a clinical context using a series of 132 blood-stream infection isolates from a single hospital. All isolates, of which 49 (37%) were vancomycin resistant, underwent whole-genome sequencing.

Results: *E. faecium* was found to be subject to high rates of recombination with little evidence of sequence importation from outside the local *E. faecium* population. Apart from disrupting phylogenetic reconstruction, recombination was frequent enough to invalidate MLST typing in the identification of clonal expansion and transmission events, suggesting that where available, whole-genome sequencing should be used in tracing the epidemiology of *E. faecium* nosocomial infections and establishing routes of transmission. Several forms of the Tn1549-like element-*vanB* gene cluster, that was exclusively responsible for vancomycin resistance, appeared and spread within the hospital during the study period. Several transposon gains and losses and instances of *in situ* evolution were inferred and although usually chromosomal, the resistance element was also observed on a plasmid background. There was qualitative evidence for clonal expansions of both vancomycin-resistant and vancomycin-susceptible *E. faecium* with evidence of hospital specific sub-clonal expansion.

Conclusions: Our data are consistent with continuing evolution of this established hospital pathogen and confirm hospital vancomycin-susceptible and -resistant *E. faecium* patient transmission events, underscoring the need for careful consideration before modifying current *E. faecium* infection control strategies.

---

## DATA SUMMARY

1. Genome read data for all samples has been deposited into European Nucleotide Archive NA: accession PRJEB8624 (<http://www.ebi.ac.uk/ena/data/view/PRJEB8624>)
2. Annotated Plasmid pJEG050 has been deposited into GenBank accession number KR066794 (<http://www.ncbi.nlm.nih.gov/nuccore/827342679/>)

**I/We confirm all supporting data, code and protocols have been provided within the article or through supplementary data files. ☒**

---

## IMPACT STATEMENT

*Enterococcus faecium* has been identified as one of 6 key pathogens that are a threat to future health care provision. This is due to the organisms' ability to cause devastating infections in susceptible patients, which are difficult to treat as a result of *E. faecium* acquiring antimicrobial resistance especially to first line therapy like vancomycin. Not surprisingly, hospitals implement several interventions to prevent spread of this organism especially if vancomycin-resistant. However, the utility of this strategy is unknown because of a suboptimal understanding about the evolutionary dynamics of this pathogen. We present a comprehensive analysis of the largest dataset to date of *E. faecium* from a single institution and also examine the context of our isolates in relation to other Australian isolates. Not only do we document the genomic plasticity we are able to show examples of recombination leading to novel MLST emergence and recreation of old MLST types. Sources of recombination are identified with further hospital adaptation documented. Antimicrobial resistance dynamics are explored with patient transmission events observed. Overall our results substantially add to the current debates around *E. faecium* infection control and will assist in forming policies of how to curb the ongoing spread of *E. faecium*.

---

## INTRODUCTION

Enterococci form part of the normal gastrointestinal flora of healthy humans. Although two species *E. faecalis* and *E. faecium* predominate, the clinical impact of *E. faecium* as a nosocomial pathogen has increased substantially and risen to prominence over the last two decades (Arias and Murray 2012, Hidron et al. 2008). Infections caused by *E. faecium* are difficult to treat, because of both intrinsic and acquired resistance to a variety of antibiotics including vancomycin, regarded

currently as first-line therapy (McGowan and Tenover 2004). Consequently, vancomycin resistant *E. faecium* infections are associated with substantially higher patient morbidity and mortality compared to vancomycin-susceptible infections (DiazGranados et al. 2005).

Resistance to vancomycin in *E. faecium* is generally conferred by either the *vanA* operon, predominant in the United States (Arias and Murray 2012, Courvalin 2006), or the *vanB* operon common in Australia (Coombs et al. 2014), almost always carried on the transposable element Tn1549 (or a closely related derivative) (Garnier et al. 2000). Vancomycin resistant enterococcus (VRE) first emerged in Europe in the 1980s (Werner et al. 2008), and since then has been detected worldwide (Ramsey and Zilberberg 2009, Willems et al. 2005) with steadily increasing incidence in numerous countries including Australia. In recognition of the continuing threat it poses to healthcare delivery, *E. faecium* was named as one of six key problem or ESKAPE bacteria by the Infectious Diseases Society of America (Boucher et al. 2009).

The ability to easily acquire additional genetic elements including antimicrobial resistance determinants may in part explain *E. faecium*'s rise as a nosocomial pathogen (Arias and Murray 2012). Moreover, high levels of homologous recombination have been linked to the emergence of specific hospital-adapted clones (e.g. the group designated MLST clonal complex 17) with the vast majority of infection isolates clustering within clade A1 (Lebreton et al. 2013). Consequently, genome plasticity has been cited as evidence for the recent and ongoing evolution of *E. faecium* (de Been et al. 2013, Howden et al. 2013, Willems et al. 2012) with clones thought to have received at least some genetic material from (colonising) *E. faecium* isolates within the patient or the environment through cross-transmission (de Been et al. 2013). However, apart from the *vanB* operon, the identities, specific functions and overall contribution of imported sequences to a healthcare-associated phenotype remain poorly understood, reflecting perhaps the difficulty of assembling informative isolate collections and difficulties in determining specific recombination events. Although previous studies have reported signals for ongoing evolution, especially in favour of disease-causing clones (de Been et al. 2013, Willems et al. 2011, Willems et al. 2005), none has examined the alternative hypothesis of new clone emergence as a result of continuous recombination.

In locations where VRE has become endemic much effort is put into preventing cross-transmission by screening at-risk patients and isolating VRE-colonised patients. Although these are logical responses to a serious problem, the effectiveness of these costly strategies is difficult to test objectively for several reasons: they are only applied where VRE infection is already a problem, they may be difficult to disentangle from changes in antimicrobial usage, and absolute infection rates

may not be informative about the sources and transmission patterns of VRE in hospitals. One remedy for the lack of resolution provided by headline rates of infection has been the use of genetic information in the form of multi-locus sequence typing (MLST) to assess relationships between isolates and infer outbreaks. MLST remains the *de facto* standard for identifying outbreaks despite the potential availability of whole genome sequencing (WGS).

To better understand the dynamics of vancomycin resistance and its relationship to genome micro-evolution in a well delineated population, we assembled and analysed whole-genome sequences of a substantial and comprehensive single-hospital collection of bloodstream *E. faecium* isolates with a high proportion of VRE. We used this information to investigate the role of recombination in the within-hospital evolution of the organism, including the emergence of vancomycin resistance, and to assess the overall relationships to external (other Australian) *E. faecium* isolates, which will further inform the current debates about the likely effectiveness of standard approaches for VRE infection control.

## METHODS

**Isolate Collection.** One hundred and forty-six consecutive non-duplicate blood culture isolates from patients with confirmed *E. faecium* bacteraemic infection at a single institution, Royal Prince Alfred Hospital (RPA) in Sydney, Australia, from two distinct time intervals (2004-2007 and 2011-2013), were identified for study (Supplementary Table 1). Of these, 7 isolates had not been stored and 7 were later excluded because of poor sequence coverage linked to suboptimal DNA extraction, leaving 132 isolates representing 90% of identified *E. faecium* bloodstream infections at RPA over the two time periods. Initial *E. faecium* isolation from flagged positive blood-culture bottles included sub-culturing onto horse-blood agar (HBA) followed by species confirmation by routine methods, susceptibility testing using Vitek2® (bioMérieux, Australia) and *vanB* PCR to confirm phenotypic vancomycin resistance (Adams 2006).

**Sequencing and Assembly.** Genomic DNA samples were sequenced on the Illumina HiSeq 2500 platform using 151b paired-end TruSeq chemistry. Reads were mapped to the Aus0085 genome (2,994,661 bp, GenBank: CP006620, an ST203 *vanB* VRE isolated from a patient in Melbourne, Australia in 2009)(Lam et al. 2013), using Stampy v1.0.23 (Lunter and Goodson 2011) with an expected substitution rate of 0.01. Variants were called using SAMtools v0.1.19 (Li et al. 2009) and filtered for read depth (minimum 20), read base quality (minimum Phred score 20) and mapping quality (minimum 30). Variation at indels and in the presence of mobile elements was excluded from the mapping-based analysis.

### Multilocus sequence typing, resistome analysis and transposon assembly.

*De novo* assemblies constructed using Velvet v1.2.10 (Zerbino 2010) with VelvetOptimiser v2.2.5 (<http://bioinformatics.net.au/software/velvetoptimiser.shtml>) were used for *in-silico* MLST (<http://pubmlst.org/>) and to detect *vanB* sequences using BLASTN v2.2.28 against the Aus0085 genome. Susceptibility and *vanB* PCR re-testing in two isolates whose recorded phenotype and inferred genotype were discordant confirmed WGS accuracy. Tn1549-like sequences carrying the *vanB* locus were identified by comparing *de novo* contigs against complete Tn1549 sequences from Aus0085 (GenBank: CP006620 nt positions 723,825-781,312) and Aus0004 (GenBank: CP003351 nt positions 2,835,430-2,869,240) *rE. faecium* reference genomes. Obtained sequences were aligned using progressiveMauve (Darling et al. 2010) and annotated using CLC Genomics Workbench version v7.5.1 (CLC bio, Qiagen Company) with isolate-specific insertion sites identified relative to Aus0004.

**Recombination.** Homologous recombination was assessed using an iterative application of ClonalFrameML (Didelot and Wilson 2015) as follows: [1] A maximum-likelihood input phylogeny, rooted at the midpoint, was inferred using RaxML v7.0.43 (Stamatakis 2006) under the General Time-Reversible model, and analysed using ClonalFrameML; [2] Identified recombination intervals were masked in isolates containing the recombination event or descending from the branch containing the event. The modified dataset was then used to generate a new RaxML input phylogeny; [3] ClonalFrameML was rerun with the new phylogeny and the original (unmasked) genome alignment. Steps [2] and [3] were repeated until the number of recombination events and phylogeny remained stable. Estimates of recombination parameters and the set of recombination events were recorded from the final run.

**Genomic context.** All Illumina reads from another Australian study (GenBank: PRJNA205886) (Howden et al. 2013), strain DO (GenBank: NC\_017960) and Aus0004 (GenBank: CP003351) were analysed using the above pipeline in order to reconstruct a phylogeny of available Australian isolates.

**Visualisation of phylogenetic trees and recombination events.** Trees were visualised and manipulated with Figtree v1.4.2 (<http://tree.bio.ed.ac.uk/software/figtree>).

**Nucleotide sequence accession number.** The sequence reads, mapping-based variant calls and *de novo* assemblies for all isolates are available from ENA with accession PRJEB8624.

## RESULTS

### Sequencing and the Impact of Recombination

A total of 2.4 Mbp of aligned sequence, comprising 81% of the Aus0085 reference, 2,703 genes and 21,288 SNVs (reflecting the large species-wide diversity captured), was called unambiguously in all 132 successfully sequenced isolates (Supplementary Table 1) of which 49 (37%) carried the *vanB* operon and were vancomycin resistant.

Using ClonalFrameML (Didelot and Wilson 2015) as described in Methods, we detected 686 imports of homologous sequences among the 132 isolates. Recombination events were scattered across the clonal phylogeny, clustered on the genome (Figure 1) and varied widely in size, from near the practical lower limit of detection (39bp) to >100 kbp (6 events: largest 258 kbp; 10.7% of shared genome) (Supplementary Figure 1). The ratio of the estimated contributions of homologous recombination and mutation to observed substitutions (*i.e.*  $r/m$ ) across the overall phylogeny was 6.1 (the ratio of estimates for  $R$  and  $\theta$  - recombination and mutation events - was 0.22), moderately high among estimates for collections of bacterial samples representing successive cases in a host population. This number represents a credible increase compared with previous estimates (Vos and Didelot 2009) (due to expanded and denser population sampling and updated methodology), but is likely to be an under-estimate as some recombination events remain difficult to detect (Figure 1).

### Epidemiology, MLST, and Outbreak Investigations

The reconstruction of clonal relationships allows us to assess the epidemiology of *E. faecium* in a single hospital, and the effectiveness of multi-locus sequence typing (MLST, which compares the sequences of 7 gene fragments of ~500bp between isolates), in capturing those patterns.

*In silico* MLST places 126 (95%) of the 132 isolates, including 51 sequence type 203 (ST203) isolates, in the well-established clonal complex 17 (CC17). Two new ST types (2 ST990 and 1 ST991) were determined (Figure 1; MLST locus details in Supplementary Table 2) which could also be added to CC17. Quite apart from its ~1000-fold lower resolution of micro-evolutionary variation compared with WGS, the integrity of MLST-based classification compared with reconstructed clonal relationships was seriously affected by recombination. Some 28 inferred recombination events overlapped at least one (*ddl*, *atpA*, *gyd*, *pstS*) of the seven MLST loci (Figure 1). Sequence types were not monophyletic in either the naïve or the corrected clonal phylogeny, with several examples of distinct groupings by time and MLST type (*e.g.* ST203 in 2011 and ST17 in 2013; Figure 1). In fact, an



inferred recombination changed the ST within a clonal cluster at least ten times (*e.g.* ancestral ST203 and ST341, to ST17 and ST252, respectively) leading to potential false negative clusterings (Figure 1; turquoise branches). In one of these events (Figure 1; indicated by a triangle), a recombination event across a single locus (*atpA*) resulted in the emergence of a new ST, 990, in two isolates and confirms the alternative hypothesis of new clone emergence as a result of continuous recombination. Conversely, one possible false-positive clustering resulted from the re-creation of an already-present ST (Figure 1; red branches).

These observations when combined with clinical-epidemiological data would have resulted in incorrectly missing or assigning *E. faecium* outbreaks. Similarly, when examining the 7 potential VRE outbreaks (defined as >1 VRE bacteraemia of the same MLST type within the same calendar month) (Figure 1 and Supplementary Table 3), two events (outbreaks coloured red in Figure 1) would be confirmed by WGS when using a cut-off of 10 SNV differences (nominally 1 year's divergence at a molecular clock rate of ~10 SNVs per genome per year) (Howden et al. 2013, Lebreton et al. 2013).

These observations and robust evidence for recombination across housekeeping genes lead us to conclude that MLST is neither sensitive nor accurate enough for *E. faecium* molecular epidemiology. By comparison, the enhanced resolution of WGS would likely also enhance the early detection of new clonal expansion events (Miller et al. 2014).

#### **Donors and recipients of recombined sequences.**

Our dataset of 132 geographically and temporally related *E. faecium* genomes was investigated as the primary plausible source(s) of sequence imports for other genomes in the dataset. A non-descendant genome was considered a plausible within-species source of an imported fragment of ≥200nt if it contained the most similar sequence >99.9% identical to and overlapping at least 90% of the imported fragment (identically matching sequences were considered as equally likely sources). Where there was no match within the dataset, plausible sources were identified in the GenBank nucleotide database.

The potential donor of inferred imports could be ascribed to isolates sequenced within the dataset for 80.5% (552 of 686) of all events, where at least one exact match or a plausible match was inferred for 481 and 71 events, respectively. BLASTN searches of the remaining 108 imported fragments ≥200nt identified *E. faecium* (but not an isolate from the current study) as the likely source for all remaining events and places the lower bound for recent extra-species imports at zero (provided either sequence imports into our dataset has not occurred more than once or closely related lineages have not co-inherited the same recombination sequence). The lack of direct

evidence favouring a role for closely related non-faecium enterococcal species in acquisition of genetic material by *E. faecium*, however, does not exclude a more distant origin for some of these identified recombinant sequences and also does not exclude a shuttle role for other commensal species in the supply of new material. Nevertheless, the numerous within species recombination events suggest the importance of other *E. faecium* lineages in the overall adaptation of hospital-adapted CC17 isolates.

### **Distribution of recombination across the genome**

Of the 2703 annotated genes in the shared genome, 1880 (70%) underwent at least one recombination event, with several regions of the shared genome seemingly overrepresented (Supplementary Figure 2). No particular COG categories (including antimicrobial resistance) were enriched among these highly replaced loci, failing to support one possible simple explanation for recombination heterogeneity along the genome (Croucher et al. 2014). Neither were there obvious patterns of replacement in the *esp* (enterococcal surface protein) gene which encodes a virulence factor associated with hospital adaptation (Coque et al. 2002). All CC17 and none of the non-CC17 isolates (branch denoted by asterisk in Figure 1) in our dataset carried *esp*, consistent with its acquisition by an ancestor of present-day CC17 and lending support for the importance of this gene in hospital adaptation.

### **Vancomycin resistance and Tn1549.**

Four main *Tn1549*-like transposon structures carrying *vanB* were distinguishable (Figure 2) among the 49 vancomycin resistant isolates, including one with alternate secondary variants involving insertions of extra sequences into the *vanB* operon. A temporal relationship among circulating transposon structures was observed with Tn1549 structure 4d largely replacing structure 4a as the dominant form from 2011 onwards.

In three separate isolates, Tn1549 was detected on a distinct plasmid (Figure 3, denoted by the asterisks): In one isolate, Tn1549 was inserted in the *yfbU* gene of an Aus0085 p3-like plasmid designated pJEG050 (GenBank acc. no. KR066794; Supplementary Figure 3). In the other two isolates, transposons were found inserted in previously-described enterococcal plasmids (Aus0085 p1, GenBank acc. no. CP006621; DO p3, GenBank acc. no. CP003586). The detection of several Tn1549-carrying plasmids indicates the importance of extra-chromosomal elements in the transfer of resistance (Woodford et al. 1995), however the relative impact of this and other mechanisms of transmission remains uncertain.

For the remaining 46 isolates, Tn1549 was chromosomally integrated. Insertion sites occurred at least 41kb from areas of extensive recombination within the shared genome, excluding a mechanistic link between transposition and recombination. Transposons of the same structure tended to insert at the same chromosomal location, however, this was not always the case (structure 4a, Figure 2). Once the clonal phylogeny was accounted for, rather than a literal interpretation of insertion sites, there was no evidence of repeated insertion into any 'hotspot' position.

Vancomycin-resistant isolates were scattered across and in distinct branches of the clonal phylogeny, implying at least 15 distinct *vanB*-carrying transposon acquisition and 6 loss events in the history of the samples (Figure 3). In addition, taking into account the phylogeny unmasked 4 *E. faecium in situ* transposon adaptation events. This pattern of gains and losses was confirmed when temporal isolate data was included and is best illustrated by isolates clustering within MLST ST341 which were collected in 2011 and 2012, all harboured the same Tn1549 substructures (4d; Figure 3).

Notably, all detected transposon gain events occurred in isolates closely related or identical to sampled vancomycin-susceptible *E. faecium* isolates, consistent with the suggestion that VRE emerges from the circulating enterococcal population followed by VRE transmission with definitive evidence for one such transmission with two temporally linked isolates, identically matched at the core genome level and Tn1549 (Triangle in figure 3).

## Genomic context

To place our results in context, we examined the genomic relationships between isolates from single hospitals in Australia's two largest cities, Sydney (this study) and Melbourne, 700km apart, plus 3 isolates from Perth, a state capital (Howden et al. 2013). We obtained read data for all isolates sequenced on the Illumina platform, and analysed them using our framework to generate an enlarged phylogeny (Figure 4A). We verified the distant relationship between Australian isolates and the reference strain DO (an ST18 isolate from the United States) and revealed a complex relationship between Australian isolates, with interspersed study-specific clustering of samples. Inspection of the clonal phylogeny in conjunction with temporal signals of genetic clustering within our hospital (Figure 1) suggests a distinct common source or a history of progressively spreading institution-specific sub-lineages. We defined an arbitrary criterion for very close genomic and epidemiological relationships of  $\leq 20$  mutational substitutions to identify five highly probable inter-city transfer events (Figure 4A & 4B). Of these, four were suggestive of Melbourne-Sydney transfer events with the Melbourne isolates differing from Sydney isolate(s) by less than 20 SNVs. The remaining event represented a possible Perth-Sydney transfer event. More generally, we noted that patterns of

293 genomic diversity among *E. faecium* isolates in different Australian centres were not noticeably  
294 different but seem to have arisen from a common ancestor. Although this observation does not  
295 necessarily imply that the epidemiology (and role of infection control) of *E. faecium* disease in each  
296 centre is the same, it seems likely that apparent qualitative differences between centres could result  
297 from differences in sample size and sampling patterns

298

299

## DISCUSSION

Our appraisal of the extent of homologous recombination and dissection of the rapid structural evolution of Tn1549, carrier of the vancomycin resistance-conferring *vanB* operon, underscores the organism's well-known genomic plasticity (de Been et al. 2013, Lebreton et al. 2013, Willems et al. 2005) placing it towards the high end of the spectrum of the relative effects on molecular variation of homologous recombination and mutational substitution (Chewapreecha et al. 2014). By including a relatively large number of isolates, our study has clearly identified the effects of recombination at the whole-organism level, from an explicit analysis of recombination to an analysis of the recombination gene pool, with implications for our understanding of *E. faecium* evolution.

Dealing in a principled way with recombination via reconstruction of recombination events and revision of the clonal phylogeny has allowed a critical re-assessment of the utility of MLST in identifying outbreaks. We conclude that although sequence types previously may have been a useful descriptor, for newly emerging clones and local populations, the ongoing and rapid nature of homologous recombination in *E. faecium* means that MLST is suboptimal in settings where WGS is available (Howden et al. 2013, Willems et al. 2012).

In a widely accepted model, several factors contribute to the success of *E. faecium* and VRE in the hospital environment. Empiric broad-spectrum antibiotic therapy selects for enteric carriage of *E. faecium* due to the organism's intrinsic resistance to several antibiotics. The increased dominance of *E. faecium* (rather than *E. faecalis*) within the gastro-intestinal flora subsequently results in enhanced cross-transmission via surfaces and inter-personal contacts (Arias and Murray 2012) and invasive disease in a sub-set of colonised patients. This process is augmented by the hospital environment, which in turn has selected for specific *E. faecium* lineages such as CC17 that now predominate among hospital-acquired disease isolates but not among community isolates. Within this model, several questions have been considered, such as: what genetic alterations may have led to the emergence of dominant lineages, and what is the role of *de-novo* acquisition rather than transmission of vancomycin resistance in VRE epidemiology?

More recently, awareness has grown that VRE epidemiology is not as simple as the formation of VRE clones by *vanB* acquisition followed by their transmission (Howden et al. 2013). This has given rise to debates about the utility of current infection control strategies in controlling VRE, with some experts suggesting discontinuing infection control measures all together (Karki et al.

2015). It is important, however, not to consider the observed repeated *de novo* acquisition of vancomycin resistance as failure but rather the consequence of effective classical infection control measures especially given the observation of a direct case-to-case transmission and genetically clustered VRE (consistent with chains of successive asymptomatic colonisation events linking several patients).

Which strategies are most effective and should be employed to control VRE remains uncertain; for instance our data, with its evidence for multiple gains and losses of resistance, suggest that for all its importance in mortality vancomycin resistance is less crucial to the biology of *E. faecium* at the population level and is independent of core genome adaptation to the hospital environment. In contrast, the dynamics of *E. faecium* seem well established, with selection of specific adaptive traits through genome plasticity leading to the formation and expansion of successful clones within the hospital. Understanding these dynamics in conjunction with those factors, which enable *E. faecium* to gain mobile genetic elements including vancomycin resistance, may unlock the secrets that allow successful containment of VRE. To achieve this, in our view, we will need to better characterise connections between *E. faecium* colonisation and infection and substantially improve our understanding of the biology of Tn1549-*vanB* within commensals and the greater *E. faecium* population.

To address these important questions, further studies are required that include multiple different sources of isolates allowing for better characterisation of the overall hospital VRE burden and the possible extent that transmission adds to this burden. Since multiple rounds of gain and loss in non-selective circumstances (e.g. carriage within the hospital) must separate many of the infections, it is clear that future studies should include asymptomatic carriage and mild disease isolates in order to 'join the dots', both epidemiologically and in establishing how resistance and resistance-carrying organisms move around the hospital.

*Enterococcus faecium* is a highly recombining, multi-resistant organism, which appears both to have initially adapted to the hospital environment and to have repeatedly evolved new clones that have competed with each other, while frequently gaining and losing vancomycin resistance. Whole-genome sequencing in an institutional population study of disease cases has allowed us to demonstrate that classical genetic typing is unreliable and unsuitable for routine outbreak tracing where WGS is available especially given the highly complex genomic relationships observed among isolates within and between institutions is highly complex. For example, we emphasize the clear genomic clusters into which many isolates fall, finding evidence for cross transmission. In addition, VRE epidemiology may be institution-specific and is likely to be highly dynamic given the genome

plasticity observed suggesting an ongoing need for larger scale surveillance and carefully crafted genomic studies. Such studies will need to include carriage and other non-disease isolates to better determine the epidemiology of within-hospital *E. faecium* and VRE infections and the dynamics of vancomycin resistance acquisition. Together, these insights would support the development of optimised infection control strategies and interruption of transmission limiting the impact of these infections.

---

## ACKNOWLEDGEMENTS

None

---

## ABBREVIATIONS

**VSE:** vancomycin susceptible *Enterococcus faecium*

**VRE:** vancomycin resistant *Enterococcus faecium*

---

## REFERENCES

- Adams, D. N. (2006) Shortcut detection of the vanB gene cluster in enterococci by a duplex real-time PCR assay. *Pathology*, 38(4), pp. 349-52.
- Arias, C. A. and Murray, B. E. (2012) The rise of the Enterococcus: beyond vancomycin resistance. *Nature reviews. Microbiology*, 10(4), pp. 266-78.
- Boucher, H. W., Talbot, G. H., Bradley, J. S., Edwards, J. E., Gilbert, D., Rice, L. B., Scheld, M., Spellberg, B. and Bartlett, J. (2009) Bad bugs, no drugs: no ESCAPE! An update from the Infectious Diseases Society of America. *Clinical infectious diseases*, 48(1), pp. 1-12.
- Chewapreecha, C., Harris, S. R., Croucher, N. J., Turner, C., Marttinen, P., Cheng, L., Pessia, A., Aanensen, D. M., Mather, A. E., Page, A. J., Salter, S. J., Harris, D., Nosten, F., Goldblatt, D., Corander, J., Parkhill, J., Turner, P. and Bentley, S. D. (2014) Dense genomic sampling identifies highways of pneumococcal recombination. *Nature genetics*, 46(3), pp. 305-9.
- Coombs, G. W., Pearson, J. C., Daley, D. A., Le, T., Robinson, O. J., Gottlieb, T., Howden, B. P., Johnson, P. D., Bennett, C. M., Stinear, T. P., Turnidge, J. D. and Australian Group on Antimicrobial, R. (2014) Molecular epidemiology of enterococcal bacteremia in Australia. *Journal of clinical microbiology*, 52(3), pp. 897-905.
- Coque, T. M., Willems, R., Canton, R., Del Campo, R. and Baquero, F. (2002) High occurrence of esp among ampicillin-resistant and vancomycin-susceptible *Enterococcus faecium* clones from hospitalized patients. *J Antimicrob Chemother*, 50(6), pp. 1035-8.

- Courvalin, P. (2006) Vancomycin resistance in gram-positive cocci. *Clinical infectious diseases*, 42 Suppl 1, pp. S25-34.
- Croucher, N. J., Page, A. J., Connor, T. R., Delaney, A. J., Keane, J. A., Bentley, S. D., Parkhill, J. and Harris, S. R. (2014) Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.*
- Darling, A. E., Mau, B. and Perna, N. T. (2010) progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One*, 5(6), pp. e11147.
- de Been, M., van Schaik, W., Cheng, L., Corander, J. and Willems, R. J. (2013) Recent recombination events in the core genome are associated with adaptive evolution in *Enterococcus faecium*. *Genome biology and evolution*, 5(8), pp. 1524-35.
- DiazGranados, C. A., Zimmer, S. M., Klein, M. and Jernigan, J. A. (2005) Comparison of mortality associated with vancomycin-resistant and vancomycin-susceptible enterococcal bloodstream infections: a meta-analysis. *Clinical infectious diseases*, 41(3), pp. 327-33.
- Didelot, X. and Wilson, D. J. (2015) ClonalFrameML: Efficient Inference of Recombination in Whole Bacterial Genomes. *PLoS Comput Biol*, 11(2), pp. e1004041.
- Garnier, F., Taourit, S., Glaser, P., Courvalin, P. and Galimand, M. (2000) Characterization of transposon Tn1549, conferring VanB-type resistance in *Enterococcus* spp. *Microbiology*, 146, pp. 1481-9.
- Hidron, A. I., Edwards, J. R., Patel, J., Horan, T. C., Sievert, D. M., Pollock, D. A., Fridkin, S. K., National Healthcare Safety Network, T. and Participating National Healthcare Safety Network, F. (2008) NHSN annual update: antimicrobial-resistant pathogens associated with healthcare-associated infections: annual summary of data reported to the National Healthcare Safety Network at the Centers for Disease Control and Prevention, 2006-2007. *Infection control and hospital epidemiology*, 29(11), pp. 996-1011.
- Howden, B. P., Holt, K. E., Lam, M. M., Seemann, T., Ballard, S., Coombs, G. W., Tong, S. Y., Grayson, M. L., Johnson, P. D. and Stinear, T. P. (2013) Genomic insights to control the emergence of vancomycin-resistant enterococci. *MBio*, 4(4).
- Karki, S., Leder, K. and Cheng, A. C. (2015) Should we continue to isolate patients with vancomycin-resistant enterococci in hospitals? *Med J Aust*, 202(5), pp. 234-6.
- Lam, M. M., Seemann, T., Tobias, N. J., Chen, H., Haring, V., Moore, R. J., Ballard, S., Grayson, L. M., Johnson, P. D., Howden, B. P. and Stinear, T. P. (2013) Comparative analysis of the complete genome of an epidemic hospital sequence type 203 clone of vancomycin-resistant *Enterococcus faecium*. *BMC Genomics*, 14, pp. 595.
- Lebreton, F., van Schaik, W., McGuire, A. M., Godfrey, P., Griggs, A., Mazumdar, V., Corander, J., Cheng, L., Saif, S., Young, S., Zeng, Q., Wortman, J., Birren, B., Willems, R. J., Earl, A. M. and Gilmore, M. S. (2013) Emergence of epidemic multidrug-resistant *Enterococcus faecium* from animal and commensal strains. *MBio*, 4(4).
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R. and Genome Project Data Processing, S. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, 25(16), pp. 2078-9.
- Lunter, G. and Goodson, M. (2011) Stampy: a statistical algorithm for sensitive and fast mapping of Illumina sequence reads. *Genome research*, 21(6), pp. 936-9.
- McGowan, J. E., Jr. and Tenover, F. C. (2004) Confronting bacterial resistance in healthcare settings: a crucial role for microbiologists. *Nature reviews. Microbiology*, 2(3), pp. 251-8.
- Miller, R. R., Walker, A. S., Godwin, H., Fung, R., Votintseva, A., Bowden, R., Mant, D., Peto, T. E., Crook, D. W. and Knox, K. (2014) Dynamics of acquisition and loss of carriage of *Staphylococcus aureus* strains in the community: the effect of clonal complex. *J Infect*, 68(5), pp. 426-39.



- Ramsey, A. M. and Zilberberg, M. D. (2009) Secular trends of hospitalization with vancomycin-resistant enterococcus infection in the United States, 2000-2006. *Infection control and hospital epidemiology*, 30(2), pp. 184-6.
- Stamatakis, A. (2006) RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics*, 22(21), pp. 2688-90.
- Vos, M. and Didelot, X. (2009) A comparison of homologous recombination rates in bacteria and archaea. *ISME J*, 3(2), pp. 199-208.
- Werner, G., Coque, T. M., Hammerum, A. M., Hope, R., Hryniewicz, W., Johnson, A., Klare, I., Kristinsson, K. G., Leclercq, R., Lester, C. H., Lillie, M., Novais, C., Olsson-Liljequist, B., Peixe, L. V., Sadowy, E., Simonsen, G. S., Top, J., Vuopio-Varkila, J., Willems, R. J., Witte, W. and Woodford, N. (2008) Emergence and spread of vancomycin resistance among enterococci in Europe. *Euro surveillance*, 13(47).
- Willems, R. J., Hanage, W. P., Bessen, D. E. and Feil, E. J. (2011) Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. *FEMS microbiology reviews*, 35(5), pp. 872-900.
- Willems, R. J., Top, J., van Santen, M., Robinson, D. A., Coque, T. M., Baquero, F., Grundmann, H. and Bonten, M. J. (2005) Global spread of vancomycin-resistant *Enterococcus faecium* from distinct nosocomial genetic complex. *Emerging infectious diseases*, 11(6), pp. 821-8.
- Willems, R. J., Top, J., van Schaik, W., Leavis, H., Bonten, M., Siren, J., Hanage, W. P. and Corander, J. (2012) Restricted gene flow among hospital subpopulations of *Enterococcus faecium*. *MBio*, 3(4), pp. e00151-12.
- Woodford, N., Morrison, D., Johnson, A. P., Bateman, A. C., Hastings, J. G. M., Elliott, T. S. J. and Cookson, B. (1995) Plasmid-Mediated vanB Glycopeptide Resistance in Enterococci. *Microbial Drug Resistance*, 1(3), pp. 235-240.
- Zerbino, D. R. (2010) Using the Velvet de novo assembler for short-read sequencing technologies. *Current protocols in bioinformatics*, Chapter 11, pp. Unit 11.5.

---

## DATA BIBLIOGRAPHY

1. van Hal SJ, Ip CC, Ansari MA, Wilson DJ, Espedido BA, Jensen SO, Bowden R. NCBI nucleotide database KR066794, (<http://www.ncbi.nlm.nih.gov/nucore/827342679/>), 2015.
2. Wellcome Trust Centre for Human Genetics, University of Oxford. European Nucleotide Archive. PRJEB8624 (<http://www.ebi.ac.uk/ena/data/view/PRJEB8624>), 2015.

---

## FIGURES AND TABLES

**Figure 1: Clonal phylogeny and inferred recombination events.**

Midpoint-rooted clonal phylogeny of 132 *E. faecium* bloodstream isolates; metadata (centre) and recombination events (right), estimated jointly as described in methods. The sizes and genomic locations of recombination fragments (dark blue line segments) and the positions of SNVs (white ticks) occurring along branches in the phylogeny are aligned with branches in the phylogeny. Vertical grey lines mask genomic fragments with incomplete data and vertical yellow lines show the positions of 7 MLST loci. Despite the use of a plausible statistical model for recombination, marked clusters of substitutions remain in the estimated clonal phylogeny, suggesting additional undetected imports and emphasizing the analytical difficulty in distinguishing real divergence (long branches) from imports. Branch asterisk indicates three divergent, non-CC17 isolates. Turquoise and red branches: carry recombination events directly affecting sequence type i.e. MLST loci leading to false negative and false positive MLST clustering respectively.

Isolate meta-data is depicted from left to right: sequence type, *vanB* vancomycin resistance status, year of isolation and epidemiological assignment to a putative outbreak [Out] by MLST). 'Other' STs use alternating green shades for convenience for successive distinct STs. Novel ST types ST990 (Number 1) and ST991 (Number 2) are indicated. Isolates assigned to putative MLST defined outbreaks are depicted by the grey bars with evidence of two definitive transmission events (red bars) when using WGS (see text and Supplementary Table 3 for further details).

## Figure 2: Tn1549-like *van*-carrying transposons

Gene arrangements of representative Tn1549-like transposons and their included *vanB* operons (inset). Labels show the numbers of isolates sharing insertion positions (with respect to the Aus0004 genome for comparison with Howden et al (2013) – sharing is consistent with a single ancestral insertion at each position) and the insertion locus (gene, hypothetical orf, or intergenic region (IGR)) Transposon structure 2 was found in two different sites within a 32 bp region of an IGR). Where the transposon inserted into a mobile element, the name of the plasmid or IS is given; the nucleotide position of the insertion is given with respect to the reference plasmid sequence or the transposase gene of the IS. Coloured arrows indicate: red, mobilization genes; orange, transposon-related genes; green, vancomycin resistance genes (see inset); white, hypothetical open reading frames; yellow, annotated genes not in one of these categories. IS elements (white pentagons) are shown with their insertion positions.

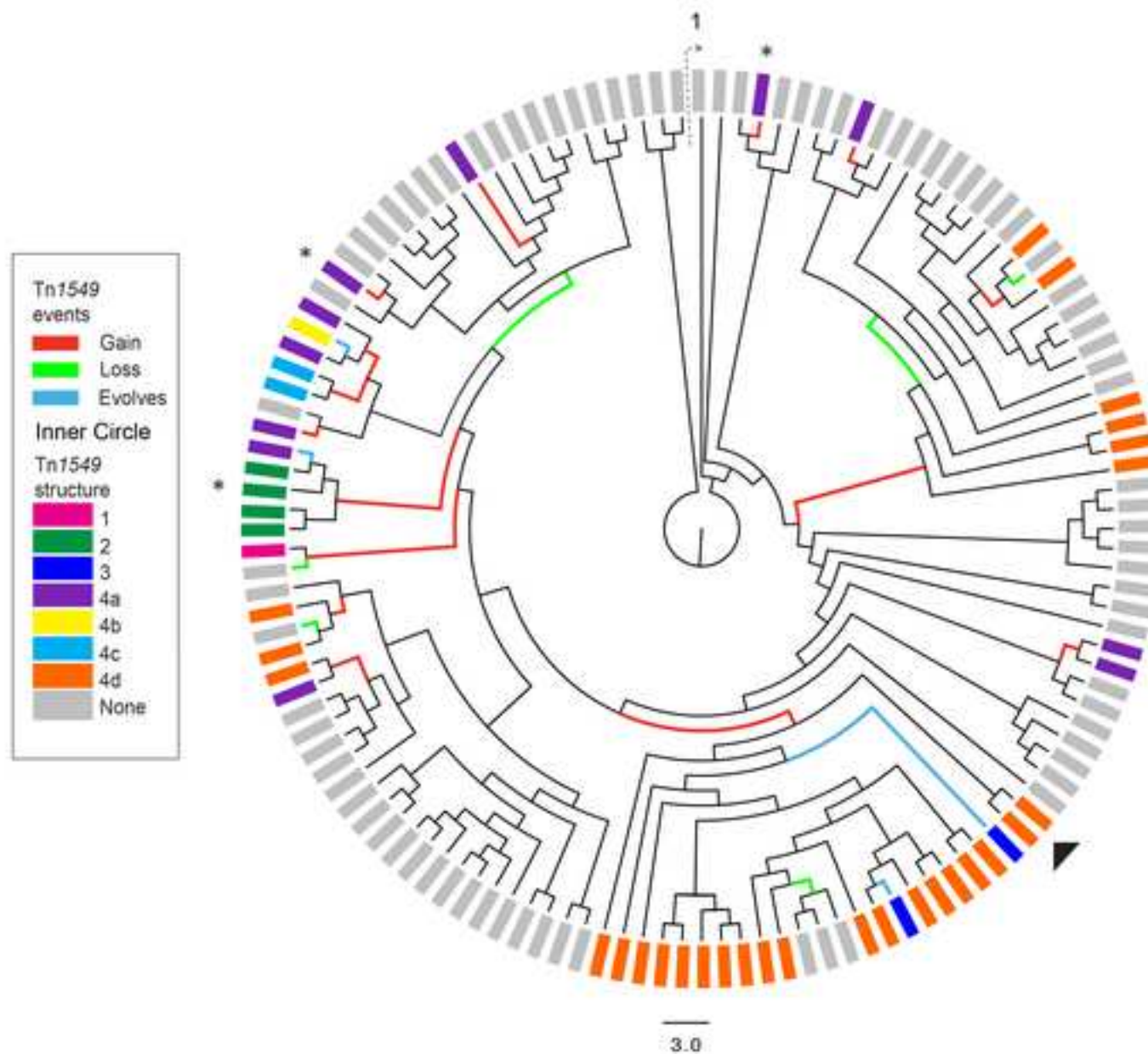
## Figure 3: *VanB* Tn1549-acquisition/loss and transposon evolution events

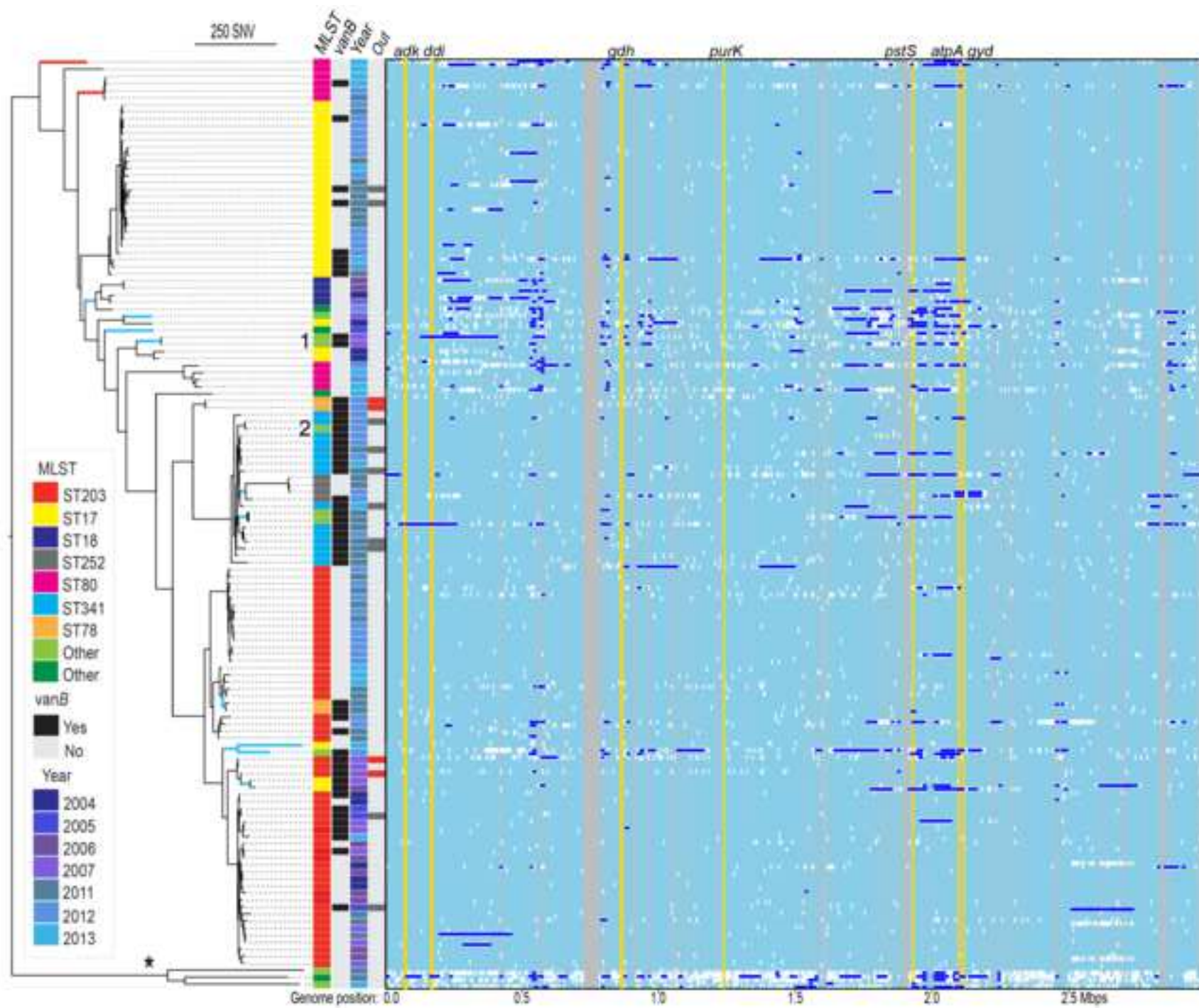
Annotated Circular phylogeny of 132 *E. faecium* isolates. Branch colours represent a parsimonious reconstruction of *vanB* transposon acquisition and loss events based on transposon structure and insertion site (green: gain; red: loss; blue: evolution of transposon *in-situ*.; from previous acquisition). Inner circle colours reflect Tn1549 structure, see figure 2 for more details. The number depicts the starting point of the tree relative to figure 1. Asterisks indicate plasmid-borne transposon. Triangle indicates two isolates that on core genome level are identical and carry the same Tn1549 consistent cross transmission.

#### Figure 4: Australian Phylogeny

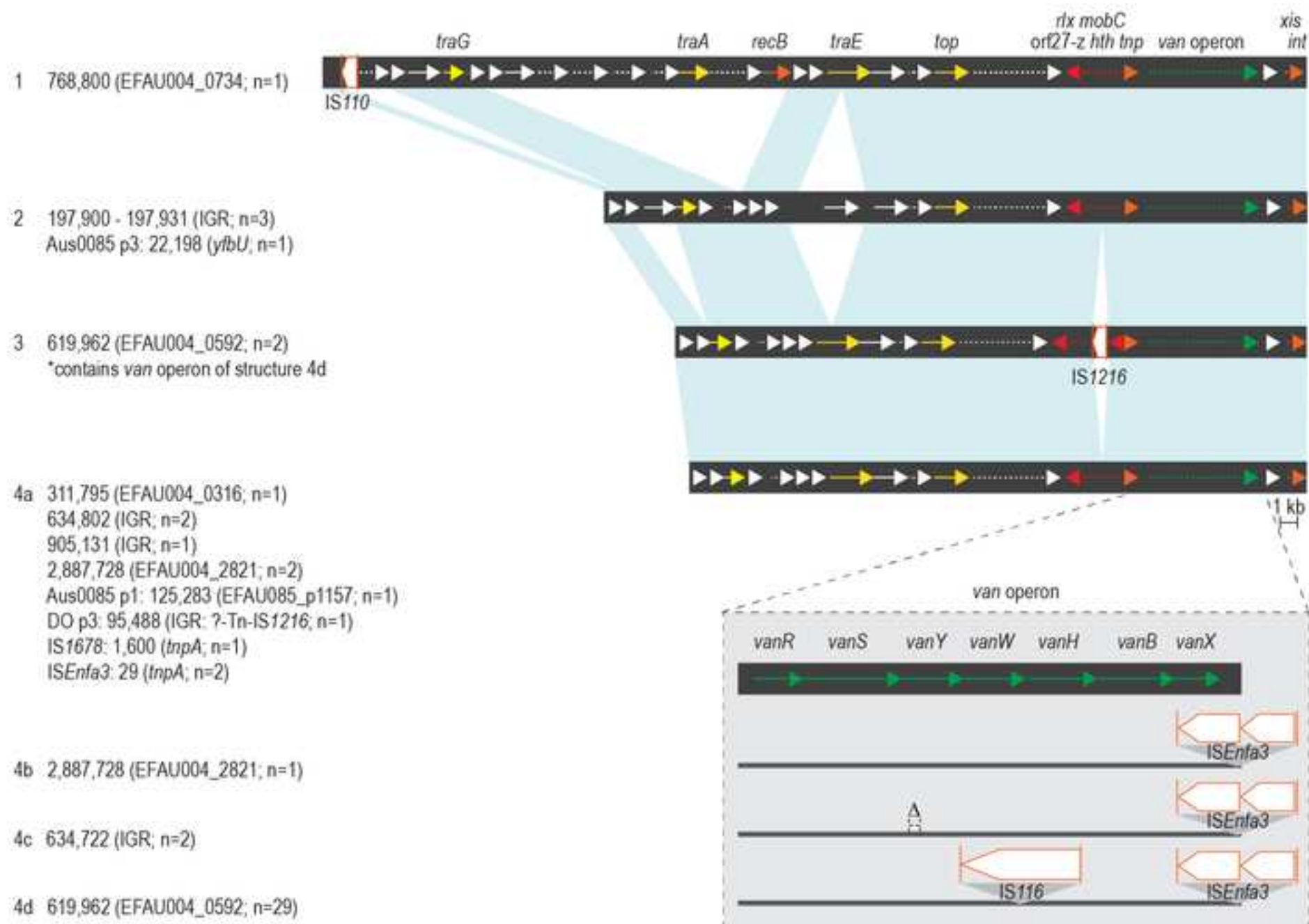
Figure 4A: Circular clonal phylogeny of 177 *E. faecium* isolates comprising of 132 bloodstream isolates from a single institution in Sydney (current study) and 45 isolates (42 from Melbourne, coloured red, and 3 vancomycin resistant *E. faecium* isolates from Perth, coloured yellow) from Howden *et al.* (Howden et al. 2013). The same *E. faecium* reference strain DO (A vancomycin susceptible ST18 *E. faecium* from USA, 1998) was included. Note that branch for strain DO has been shortened more than 4 fold as indicated by //.

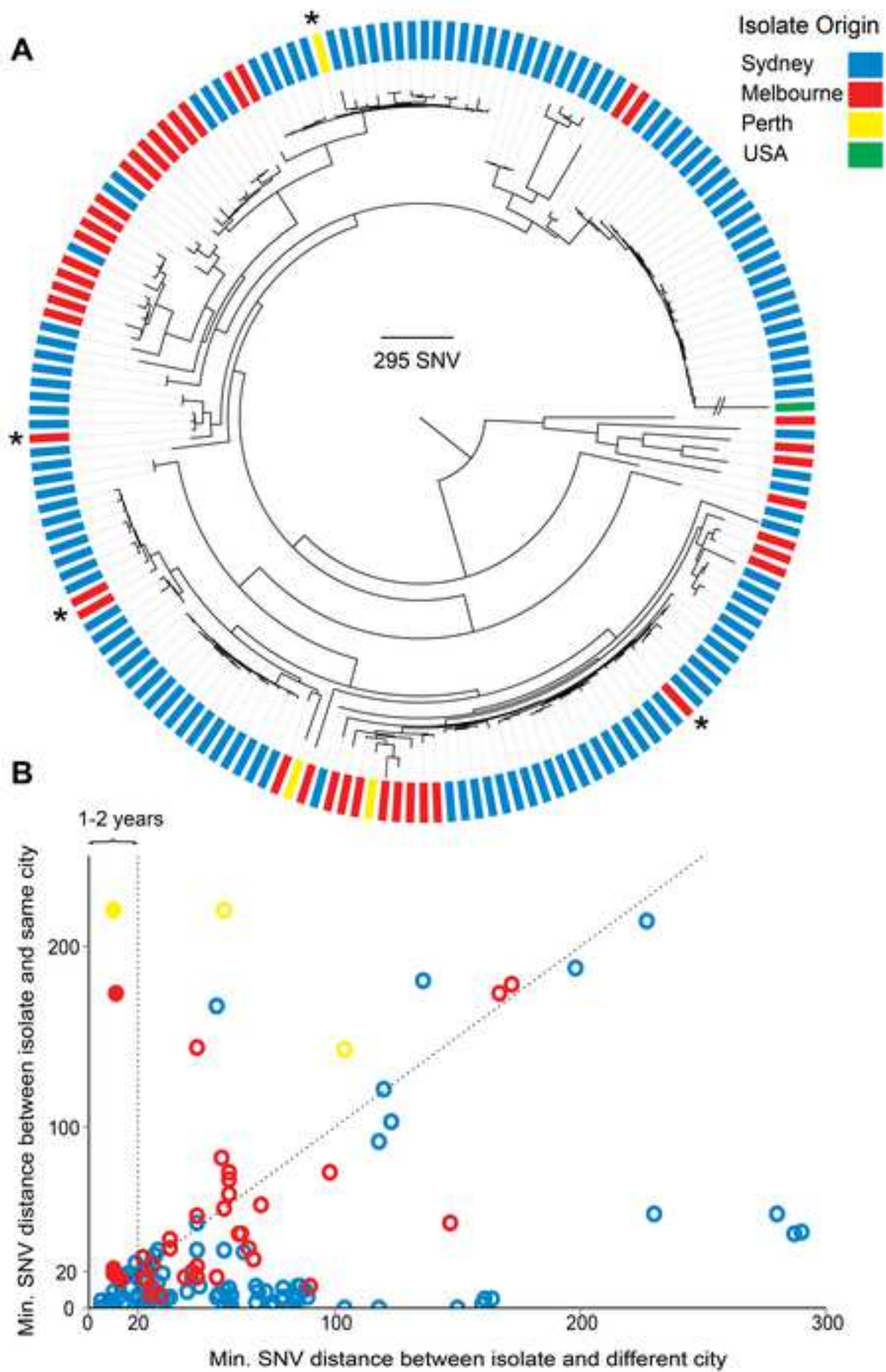
4B: Each isolate (coloured by city as in (A)) is plotted as the mutational distance in SNV to the most similar isolate from the same city (y-axis) or from a different city (x-axis). Points near the x-axis (below the diagonal) represent the majority of isolates, whose nearest neighbour is in the same population; points above the diagonal represent isolates with a nearest neighbour in another population; and points near to the diagonal ( $x=y$ ) may be interpreted as isolates in lineages (at different scales of divergence) that are shared between cities. Points plotted as filled circles correspond to the same isolates marked with asterisks as in (A), and provide some evidence for short-term inter-city spread. Seven isolates that were only distantly related to other isolates are not plotted. Vertical line represents approximately time scale of approximately 1 to 2 years between isolates.











Data ID	Sample Name	Isolate Source	Year of Isolation	Phenotype vancomycin resistant	vanB PCR	Read count	Reference positions recovered	Reference % recovered	In silico MLST	Tree position
wtchgD00002068	Efm0222	Blood culture	2013	No	Negative	4161172	2339764	78.13	ST080	1
wtchgD00002084	Efm0233	Blood culture	2013	No	Negative	3415324	2254405	75.28	ST018	2
wtchgD00002114	Efm0230	Blood culture	2013	No	Negative	4232738	2309789	77.13	ST017	3
wtchgD00002100	Efm0232	Blood culture	2013	Yes	Positive	5456810	2354970	78.64	ST080	4
wtchgD00002064	Efm0215	Blood culture	2012	No	Negative	5564376	2322215	77.55	ST080	5
wtchgD00002101	Efm0157	Blood culture	2011	No	Negative	4191410	2312592	77.22	ST080	6
wtchgD00002070	Efm0197	Blood culture	2012	No	Negative	5321582	2281931	76.20	ST017	7
wtchgD00002134	Efm0172	Blood culture	2011	No	Negative	4276814	2280932	76.17	ST017	8
wtchgD00002072	Efm0209	Blood culture	2012	Yes	Positive	5150118	2311596	77.19	ST017	9
wtchgD00002033	Efm0213	Blood culture	2012	No	Negative	3951560	2240736	74.82	ST017	10
wtchgD00002108	Efm0194	Blood culture	2013	No	Negative	4369090	2280619	76.16	ST017	11
wtchgD00002119	Efm0182	Blood culture	2012	No	Negative	3752030	2278879	76.10	ST017	12
wtchgD00002092	Efm0193	Blood culture	2011	No	Negative	3748684	2270341	75.81	ST017	13
wtchgD00002159	Efm0187	Blood culture	2012	No	Negative	3719582	2271898	75.86	ST017	14
wtchgD00002071	Efm0167	Blood culture	2011	No	Negative	4737374	2239023	74.77	ST017	15
wtchgD00002098	Efm0229	Blood culture	2013	No	Negative	5900840	2271662	75.86	ST017	16
wtchgD00002042	Efm0177	Blood culture	2012	No	Negative	3277686	2279635	76.12	ST017	17
wtchgD00002155	Efm0139	Blood culture	2011	No	Negative	4727788	2273256	75.91	ST017	18
wtchgD00002038	Efm0153	Blood culture	2011	Yes	Positive	2596522	2384209	79.62	ST017	19
wtchgD00002099	Efm0145	Blood culture	2011	No	Negative	3681096	2339746	78.13	ST017	20
wtchgD00002156	Efm0151	Blood culture	2011	Yes	Positive	4351744	2311091	77.17	ST017	21
wtchgD00002054	Efm0154	Blood culture	2011	No	Negative	5268308	2275095	75.97	ST017	22
wtchgD00002117	Efm0158	Blood culture	2011	No	Negative	4529570	2272473	75.88	ST017	23
wtchgD00002115	Efm0146	Blood culture	2011	No	Negative	4481284	2273565	75.92	ST017	24
wtchgD00002029	Efm0189	Blood culture	2012	No	Negative	2712136	2268794	75.76	ST017	25
wtchgD00002057	Efm0198	Blood culture	2012	No	Negative	5699978	2278741	76.09	ST017	26
wtchgD00002062	Efm0203	Blood culture	2012	No	Negative	4446470	2280986	76.17	ST017	27
wtchgD00002089	Efm0180	Blood culture	2012	Yes	Positive	4237000	2333391	77.92	ST017	28
wtchgD00002112	Efm0218	Blood culture	2012	Yes	Positive	5239412	2331625	77.86	ST017	29
wtchgD00002096	Efm0217	Blood culture	2012	Yes	Positive	6233762	2323726	77.60	ST017	30
wtchgD00002133	Efm0160	Blood culture	2011	Yes	Positive	3955624	2318068	77.41	ST017	31
wtchgD00002153	Efm0114	Blood culture	2006	No	Negative	4829014	2311723	77.19	ST018	32
wtchgD00002145	Efm0113	Blood culture	2006	No	Negative	4780704	2312171	77.21	ST018	33
wtchgD00002044	Efm0093	Blood culture	2004	No	Negative	4262444	2265917	75.67	ST018	34
wtchgD00002061	Efm0106	Blood culture	2005	No	Negative	4319806	2264786	75.63	ST018	35
wtchgD00002080	Efm0216	Blood culture	2012	No	Negative	3801624	2309816	77.13	ST262	36
wtchgD00002050	Efm0130	Blood culture	2007	No	Negative	5173584	2429892	81.14	ST233	37
wtchgD00002152	Efm0102	Blood culture	2004	No	Negative	4329342	2379789	79.47	ST017	38
wtchgD00002081	Efm0132	Blood culture	2007	No	Negative	3540862	2327803	77.73	ST117	39
wtchgD00002034	Efm0128	Blood culture	2007	Yes	Positive	3652926	2374484	79.29	ST990	40
wtchgD00002146	Efm0125	Blood culture	2007	Yes	Positive	4085870	2375709	79.33	ST990	41
wtchgD00002030	Efm0104	Blood culture	2004	No	Negative	3194502	2340830	78.17	ST017	42
wtchgD00002075	Efm0095	Blood culture	2004	No	Negative	4968166	2381213	79.52	ST017	43
wtchgD00002055	Efm0223	Blood culture	2013	No	Negative	5449274	2428926	81.11	ST080	44
wtchgD00002053	Efm0211	Blood culture	2012	No	Negative	5760852	2387376	79.72	ST080	45
wtchgD00002076	Efm0192	Blood culture	2012	No	Negative	4142256	2393646	79.93	ST080	46



wtchgD00002035	Efm0225	Blood culture	2013	No	Negative	5897656	2371957	79.21	ST080	47
wtchgD00002060	Efm0191	Blood culture	2012	No	Negative	4543198	2378767	79.43	ST555	48
wtchgD00002031	Efm0201	Blood culture	2012	Yes	Positive	4368982	2452881	81.91	ST078	49
wtchgD00002043	Efm0199	Blood culture	2012	Yes	Positive	4148408	2452436	81.89	ST078	50
wtchgD00002073	Efm0179	Blood culture	2012	Yes	Positive	3977232	2471052	82.52	ST341	51
wtchgD00002143	Efm0185	Blood culture	2012	Yes	Positive	3709568	2421075	80.85	ST341	52
wtchgD00002165	Efm0176	Blood culture	2012	Yes	Positive	4042238	2420950	80.84	ST991	53
wtchgD00002058	Efm0210	Blood culture	2012	Yes	Positive	5207014	2430556	81.16	ST341	54
wtchgD00002049	Efm0214	Blood culture	2012	Yes	Positive	5095280	2430223	81.15	ST341	55
wtchgD00002103	Efm0169	Blood culture	2011	Yes	Positive	3505432	2420710	80.83	ST341	56
wtchgD00002086	Efm0196	Blood culture	2012	Yes	Positive	5302084	2431163	81.18	ST341	57
wtchgD00002090	Efm0220	Blood culture	2012	Yes	Positive	4280182	2430153	81.15	ST341	58
wtchgD00002140	Efm0149	Blood culture	2011	Yes	Positive	4270178	2422775	80.90	ST341	59
wtchgD00002052	Efm0142	Blood culture	2011	No	Negative	4960698	2399000	80.11	ST252	60
wtchgD00002045	Efm0190	Blood culture	2012	No	Negative	5094332	2399869	80.14	ST252	61
wtchgD00002126	Efm0171	Blood culture	2011	No	Negative	4844904	2399546	80.13	ST252	62
wtchgD00002150	Efm0174	Blood culture	2011	Yes	Positive	4159982	2343371	78.25	ST252	63
wtchgD00002158	Efm0175	Blood culture	2012	Yes	Positive	4763402	2432079	81.21	ST341	64
wtchgD00002125	Efm0159	Blood culture	2011	Yes	Positive	4711244	2458588	82.10	ST414	65
wtchgD00002163	Efm0152	Blood culture	2011	Yes	Positive	3990756	2458298	82.09	ST414	66
wtchgD00002056	Efm0166	Blood culture	2011	Yes	Positive	4350288	2423289	80.92	ST341	67
wtchgD00002047	Efm0202	Blood culture	2012	Yes	Positive	5650236	2432997	81.24	ST341	68
wtchgD00002118	Efm0170	Blood culture	2011	Yes	Positive	3554164	2406199	80.35	ST341	69
wtchgD00002132	Efm0148	Blood culture	2011	Yes	Positive	4956032	2423321	80.92	ST341	70
wtchgD00002148	Efm0150	Blood culture	2011	Yes	Positive	3254244	2429053	81.11	ST341	71
wtchgD00002162	Efm0140	Blood culture	2011	Yes	Positive	4512894	2390217	79.82	ST341	72
wtchgD00002104	Efm0207	Blood culture	2012	No	Negative	5180208	2512310	83.89	ST203	73
wtchgD00002164	Efm0164	Blood culture	2011	No	Negative	3474648	2512054	83.88	ST203	74
wtchgD00002102	Efm0195	Blood culture	2012	No	Negative	3716328	2512076	83.89	ST203	75
wtchgD00002040	Efm0165	Blood culture	2011	No	Negative	4365366	2511843	83.88	ST203	76
wtchgD00002078	Efm0204	Blood culture	2012	No	Negative	5009590	2512527	83.90	ST203	77
wtchgD00002087	Efm0168	Blood culture	2011	No	Negative	4114298	2512838	83.91	ST203	78
wtchgD00002135	Efm0184	Blood culture	2012	No	Negative	4469040	2512784	83.91	ST203	79
wtchgD00002141	Efm0161	Blood culture	2011	No	Negative	3925882	2512461	83.90	ST203	80
wtchgD00002151	Efm0186	Blood culture	2012	No	Negative	4734862	2512923	83.91	ST203	81
wtchgD00002166	Efm0188	Blood culture	2012	No	Negative	2200148	2512132	83.89	ST203	82
wtchgD00002110	Efm0206	Blood culture	2012	No	Negative	3279842	2511772	83.88	ST203	83
wtchgD00002142	Efm0173	Blood culture	2012	No	Negative	3243984	2518013	84.08	ST203	84
wtchgD00002127	Efm0183	Blood culture	2012	No	Negative	3431158	2518827	84.11	ST203	85
wtchgD00002066	Efm0227	Blood culture	2013	No	Negative	4788350	2512383	83.90	ST203	86
wtchgD00002037	Efm0200	Blood culture	2012	No	Negative	3527236	2527460	84.40	ST203	87
wtchgD00002106	Efm0219	Blood culture	2012	No	Negative	4299346	2487997	83.08	ST203	88
wtchgD00002082	Efm0228	Blood culture	2013	No	Negative	4741482	2486285	83.02	ST203	89
wtchgD00002157	Efm0163	Blood culture	2011	No	Negative	5108048	2539494	84.80	ST203	90
wtchgD00002124	Efm0147	Blood culture	2011	No	Negative	3954910	2545457	85.00	ST203	91
wtchgD00002085	Efm0156	Blood culture	2011	Yes	Positive	3982278	2492999	83.25	ST078	92
wtchgD00002036	Efm0141	Blood culture	2011	Yes	Positive	3327096	2541346	84.86	ST078	93

wtchgD00002105	Efm0181	Blood culture	2012	Yes	Positive	3421552	2548604	85.10	ST203	94
wtchgD00002039	Efm0212	Blood culture	2012	No	Negative	3842126	2557608	85.41	ST203	95
wtchgD00002069	Efm0155	Blood culture	2011	Yes	Positive	4056592	2582911	86.25	ST203	96
wtchgD00002131	Efm0136	Blood culture	2007	No	Negative	4586176	2516350	84.03	ST203	97
wtchgD00002051	Efm0226	Blood culture	2013	No	Negative	6684914	2384388	79.62	ST017	98
wtchgD00002088	Efm0208	Blood culture	2012	Yes	Positive	5114794	2593201	86.59	ST192	99
wtchgD00002111	Efm0121	Blood culture	2007	Yes	Positive	3892196	2583125	86.26	ST203	100
wtchgD00002138	Efm0124	Blood culture	2007	Yes	Positive	3524992	2583029	86.25	ST203	101
wtchgD00002130	Efm0123	Blood culture	2007	Yes	Positive	3566890	2631317	87.87	ST203	102
wtchgD00002079	Efm0119	Blood culture	2006	Yes	Positive	3165842	2528204	84.42	ST017	103
wtchgD00002123	Efm0135	Blood culture	2007	Yes	Positive	5144942	2485718	83.00	ST017	104
wtchgD00002091	Efm0096	Blood culture	2004	Yes	Positive	4149970	2562147	85.56	ST203	105
wtchgD00002028	Efm0092	Blood culture	2003	No	Negative	3387272	2536237	84.69	ST203	106
wtchgD00002093	Efm0108	Blood culture	2005	Yes	Positive	3225388	2568682	85.78	ST203	107
wtchgD00002109	Efm0109	Blood culture	2005	Yes	Positive	4287942	2569244	85.79	ST203	108
wtchgD00002113	Efm0134	Blood culture	2005	Yes	Positive	5245794	2572345	85.90	ST203	109
wtchgD00002161	Efm0127	Blood culture	2007	Yes	Positive	3921108	2592997	86.59	ST203	110
wtchgD00002083	Efm0144	Blood culture	2011	Yes	Positive	3217408	2604167	86.96	ST203	111
wtchgD00002063	Efm0118	Blood culture	2006	No	Negative	3506442	2535867	84.68	ST203	112
wtchgD00002122	Efm0122	Blood culture	2007	Yes	Positive	4240316	2565687	85.68	ST203	113
wtchgD00002137	Efm0112	Blood culture	2006	No	Negative	4200808	2535878	84.68	ST203	114
wtchgD00002128	Efm0099	Blood culture	2004	No	Negative	4472036	2536101	84.69	ST203	115
wtchgD00002077	Efm0107	Blood culture	2007	No	Negative	3670892	2535334	84.66	ST203	116
wtchgD00002136	Efm0100	Blood culture	2004	No	Negative	3738440	2535633	84.67	ST203	117
wtchgD00002120	Efm0098	Blood culture	2004	No	Negative	4832900	2536013	84.68	ST203	118
wtchgD00002129	Efm0111	Blood culture	2005	No	Negative	2892180	2535683	84.67	ST203	119
wtchgD00002048	Efm0117	Blood culture	2006	No	Negative	3745058	2535228	84.66	ST203	120
wtchgD00002121	Efm0110	Blood culture	2005	Yes	Positive	4456736	2554821	85.31	ST203	121
wtchgD00002149	Efm0162	Blood culture	2011	No	Negative	4369836	2566822	85.71	ST203	122
wtchgD00002147	Efm0138	Blood culture	2011	No	Negative	2025366	2605310	87.00	ST203	123
wtchgD00002154	Efm0126	Blood culture	2007	No	Negative	4360104	2574448	85.97	ST203	124
wtchgD00002067	Efm0143	Blood culture	2011	No	Negative	3638724	2575029	85.99	ST203	125
wtchgD00002097	Efm0133	Blood culture	2007	No	Negative	5036258	2574353	85.96	ST203	126
wtchgD00002032	Efm0116	Blood culture	2006	No	Negative	3405482	2535107	84.65	ST203	127
wtchgD00002160	Efm0115	Blood culture	2006	No	Negative	3779648	2526245	84.36	ST203	128
wtchgD00002095	Efm0120	Blood culture	2006	No	Negative	3130464	2535509	84.67	ST203	129
wtchgD00002139	Efm0137	Blood culture	2007	No	Negative	3997940	2246177	75.01	ST682	130
wtchgD00002107	Efm0097	Blood culture	2004	No	Negative	3083806	2233646	74.59	ST032	131
wtchgD00002041	Efm0224	Blood culture	2013	No	Negative	4718050	2207245	73.71	ST092	132

**Supplementary Table 2: Details of Novel MLSTs**

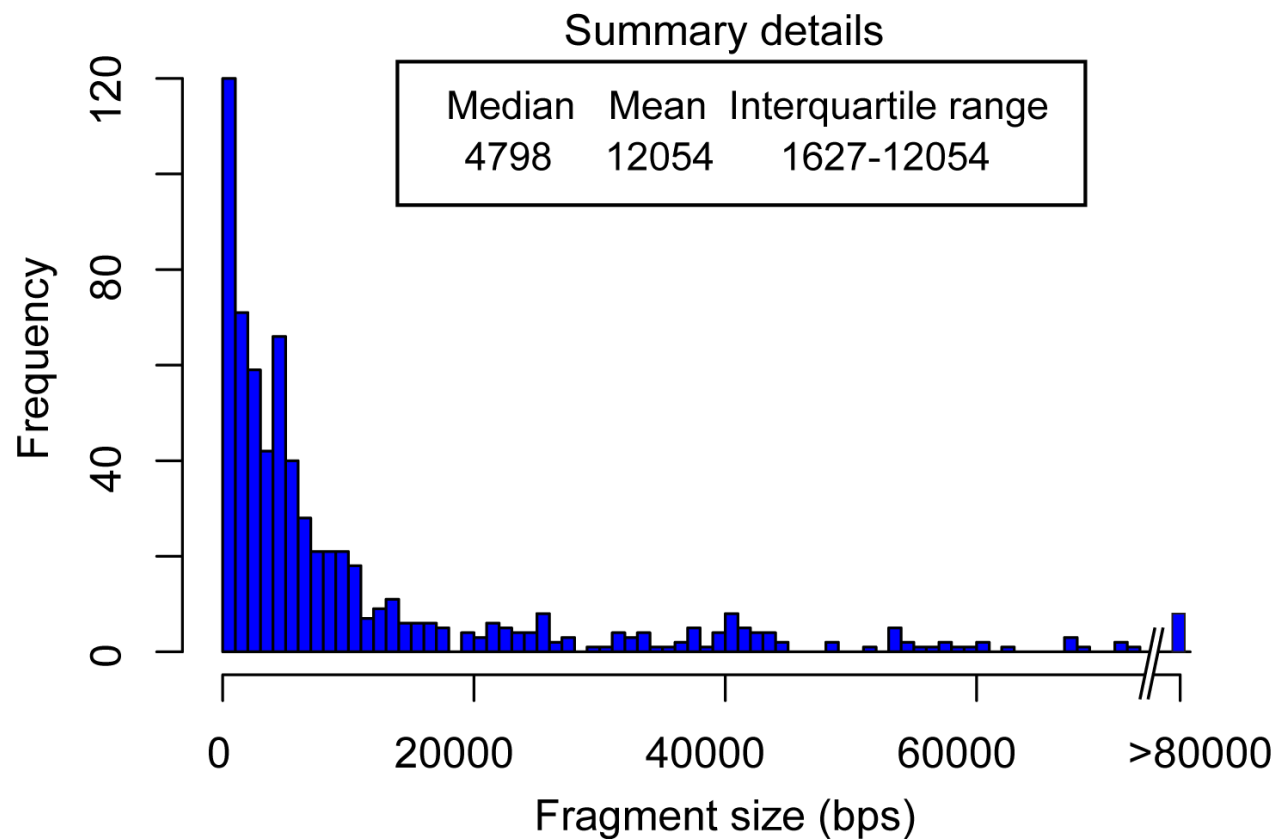
Id	country	year	isolation_ site	age_ yr	van_gene	MLST	Housekeeping loci						
							atpA	ddl	gdh	purK	gyd	pstS	adk
wtchgD00002034	Australia	2007	blood	61	vanB	990	1	7	1	1	1	1	1
wtchgD00002037	Australia	2007	blood	61	vanB	990	1	7	1	1	1	1	1
wtchgD00002165	Australia	2012	blood	77	vanB	991	15	5	1	47	1	1	1

**Supplementary Table 3:** Details of Vancomycin resistant *E. faecium* bloodstream infection outbreaks

Isolate ID	MLST	Month	Year	Outbreak Number	Outbreak by MLST	Number of SNPs between isolates on WGS	Outbreak by WGS
<b>Outbreak 1</b>					yes		no
wtchgD00002109	ST203	August	2005	1.1		24 snps	
wtchgD00002121	ST203	August	2005	1.2			
<b>Outbreak 2</b>					yes		yes
wtchgD00002130	ST203	March	2007	2.1		6 snps	
wtchgD00002111	ST203	March	2007	2.2			
<b>Outbreak 3</b>					yes		no
wtchgD00002132	ST341	July	2011	3.1		43 snps	
wtchgD00002140	ST341	July	2011	3.2			
<b>Outbreak 4</b>					yes		no
wtchgD00002156	ST17	August	2011	4.1		13 snps	
wtchgD00002038	ST17	August	2011	4.2			
<b>Outbreak 5</b>					yes		no
wtchgD00002103	ST341	December	2011	5.1		41 snps	
wtchgD00002118	ST341	December	2011	5.2			
<b>Outbreak 6</b>					yes		yes
wtchgD00002043	ST78	July	2012	6.1		1 snp	
wtchgD00002031	ST78	July	2012	6.2			
<b>Outbreak 7</b>					yes		no
wtchgD00002158	ST341	August	2012	7.1		15 snps	
wtchgD00002058	ST341	August	2012	7.2			

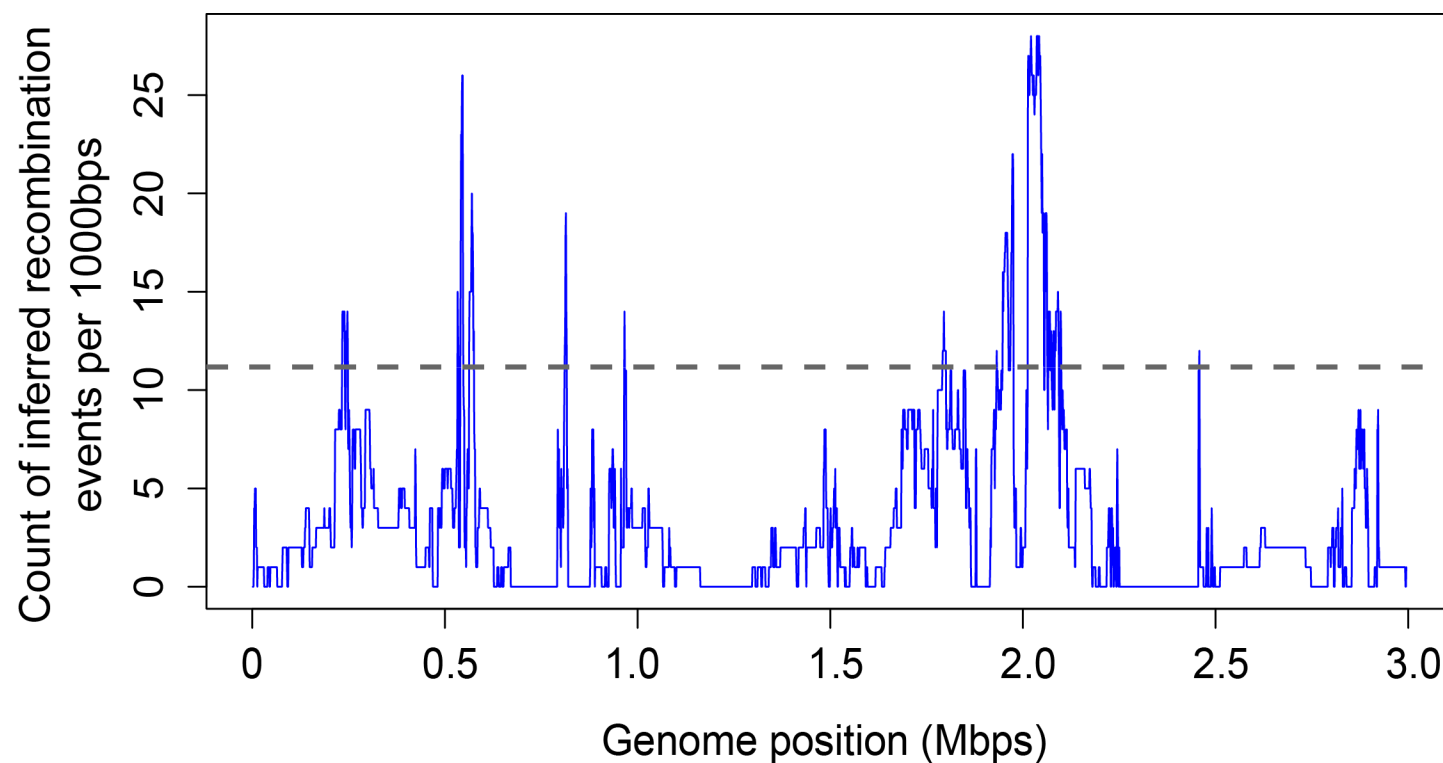
MLST: Multi-Locus Sequence Typing; SNPs: Single nucleotide polymorphisms; WGS: Whole genome sequencing

A possible outbreak was deemed to have occurred when more than one VRE bacteraemia of the same MLST type occurred in a calendar month. Of the seven events identified only two were true outbreaks if WGS criteria were applied (a SNP difference of 10 or less) and corresponds with red highlighted outbreaks in Figure 1.



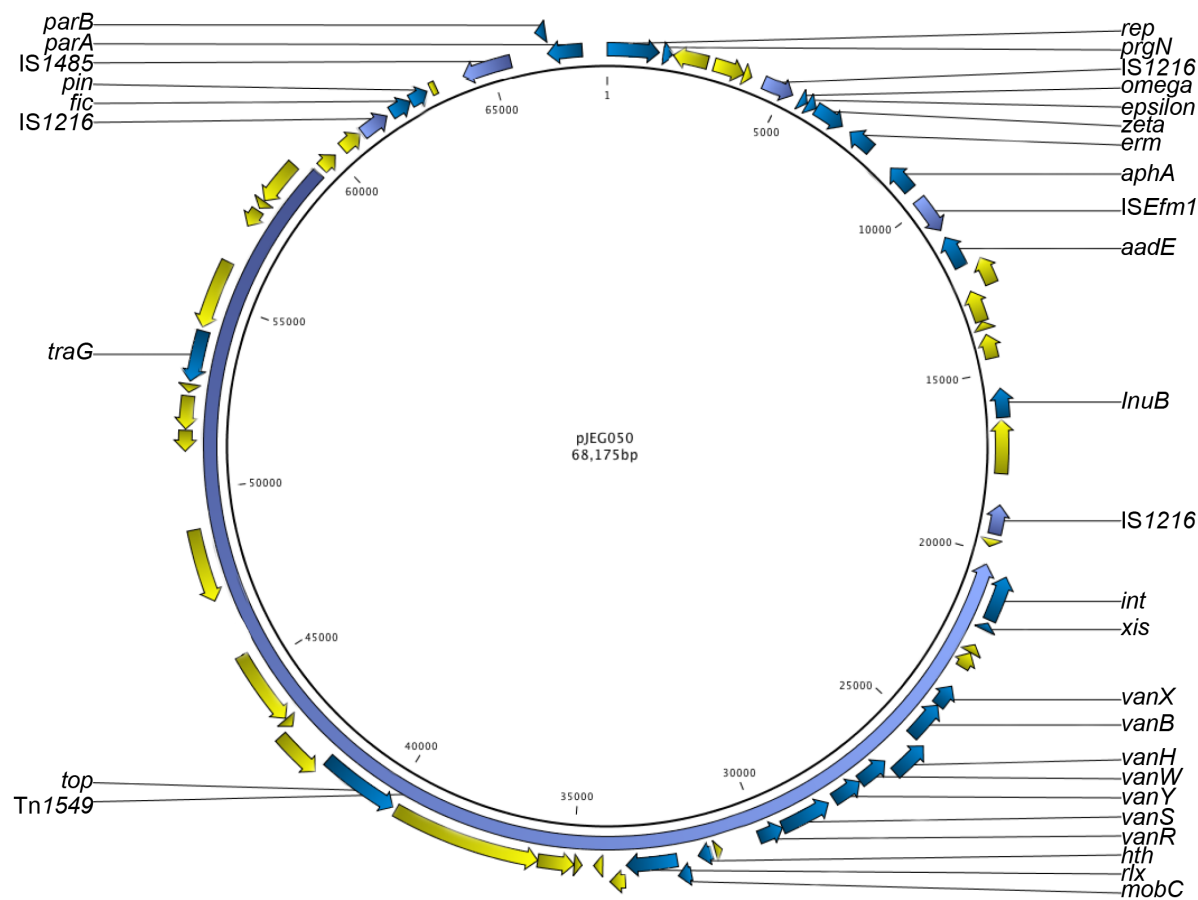
### Supplementary Figure 1: Recombination fragments

Histogram of the estimated recombination fragments lengths predicted by ClonalFrameML. Summary details of recombination fragments in bps are shown in figure textbox. Of the 20 events  $\geq 80$ kbps, three were greater than 100kbps with the largest fragment 279kbps.



**Supplementary Figure 2: Rates of recombination across genome**

Manhattan plot of the number of recombination events across the genome per 1000bps. Dashed line represents the 95<sup>th</sup> percentile of events. Examination of potential 'hotspots' (i.e. areas of the genome with events above the 95<sup>th</sup> percentile) failed to show any enrichment for any specific pathway or genes including antimicrobial resistance genes.



### Supplementary Figure 3: Annotated plasmid pJEG050

Tn1549 was inserted in the *yfbU* gene of an Aus0085 p3-like plasmid designated pJEG050 (GenBank acc. no.KR066794) and was detected in a single vancomycin-resistant *Enterococcus faecium* blood stream isolate.