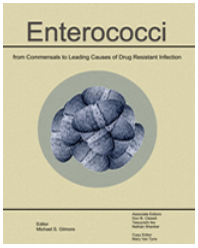
U.S. National Library of Medicine
National Center for Biotechnology Information

NLM Citation: Palmer KL, van Schaik W, Willems RJL, et al. Enterococcal Genomics. 2014 Feb 8. In: Gilmore MS, Clewell DB, Ike Y, et al., editors. Enterococci: From Commensals to Leading Causes of Drug Resistant Infection [Internet]. Boston: Massachusetts Eye and Ear Infirmary; 2014-.
Bookshelf URL: https://www.ncbi.nlm.nih.gov/books/



Enterococcal Genomics

Kelli L. Palmer,[1] Willem van Schaik,[2] Rob J. L. Willems,[3] and Michael S. Gilmore[4]

Created: February 8, 2014.

Introduction

Enterococcal genomics is a rapidly growing area of study. The first enterococcal genome sequence—that of *Enterococcus faecalis* V583—was published ten years ago (McShan & Shankar, 2002; Paulsen, et al., 2003), and complete or draft genome sequences of various enterococcal strains and species now number in the hundreds (http://www.ncbi.nlm.nih.gov/genome). Concurrent with rapid advances in genome sequencing, the sequencing-based classification scheme of multilocus sequence typing (MLST) has been used to interrogate population structures of the two enterococcal species that are most associated with human health and disease, *E. faecalis* and *Enterococcus faecium*. These two species also constitute the bulk of enterococcal genome sequence data that has been generated to date. This wealth of genomic data has allowed for an investigation of enterococcal diversity at a depth not previously achievable. Genomic studies in enterococci have been driven by overarching questions, such as: Why do multiple species of enterococci exist that inhabit seemingly identical niches, such as *E. faecalis* and *E. faecium* in the human gut, and what ecological factors have contributed to their divergence from a common ancestor? Within an enterococcal species such as *E. faecalis* or *E. faecium*, what qualities distinguish one strain from another? Are infection- or hospital-derived strains evolutionarily distinct from strains that benignly co-exist in the complex microbial consortium of the healthy human intestine? Related to this, have antibiotic use and the nosocomial environment led to changes in the enterococcal genome and/or its population structure?

This chapter highlights major advances in enterococcal genomics, including the development of MLST schemes to study the population structures of *E. faecalis* and *E. faecium;* comparative genome hybridization (CGH) studies to catalog the genomic contents of hundreds of *E. faecalis* and *E. faecium* strains; and significant findings from genome sequencing of multiple enterococcal species, beginning with the discovery and sequencing of the *E. faecalis* pathogenicity island (PAI). Additionally, we review the use of genome resequencing as a tool to study the short-term evolution of *E. faecalis* and the use of metagenomics to assemble *in situ* enterococcal genomes. In concluding the chapter, we discuss future perspectives in enterococcal genomics, including pressing questions that should drive future research in this field. While comparative genomics in enterococci has rapidly advanced over the last ten years, the number of genomes discussed here pales in comparison to what has been emerging—

Author Affiliations: 1 Department of Molecular and Cell Biology, University of Texas at Dallas, Richardson, Texas, USA; Email: Kelli.Palmer@utdallas.edu. 2 Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands. 3 Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands. 4 Departments of Ophthalmology and of Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA.

Corresponding author.

136 enterococcal genomes have been sequenced as part of the Human Microbiome Project (http://www.hmpdacc.org/), and 406 more were sequenced in a large-scale enterococcal genome sequencing endeavor performed in a multi-national collaboration with the Broad Institute (Cambridge, MA). Clearly, our foray into enterococcal genomics has only just begun.

Population structures of *E. faecalis* and *E. faecium* defined by MLST

MLST is a sequencing-based typing method that was designed to infer the global and long-term epidemiology of a particular bacterial species. In addition, MLST has been widely used in both population and evolutionary biology (Urwin & Maiden, 2003). The technique relies upon the amplification and sequencing of internal fragments of multiple housekeeping genes that are present at different locations on a chromosome. MLST is similar in concept to 16S rRNA sequencing, in that both techniques rely upon the sequencing of genomic loci for which diversifying selection is low, but is distinguished from 16S rRNA sequencing in its ability to resolve relationships at a subspecies level. Although MLST is able to infer bacterial population structure using hundreds or thousands of strains, one of its drawbacks is the limited number of alleles that are analyzed—only seven for *E. faecium* and *E. faecalis*—which are presumed to represent the evolutionary history of the entire genome. This limitation can be overcome by comparing full genome sequences, as has been shown by studies of the pathogens *Staphylococcus aureus* and *Streptococcus pneumoniae*, in which the recent evolution of pathogenic lineages was described by the analysis of several dozen genome sequences (Croucher, et al., 2011; Harris, et al., 2010). However, as a facile classification method, MLST has shown that certain lineages within *E. faecalis* and *E. faecium* are consistently associated with infections, multidrug resistance, and hospital persistence, and this information has helped guide the selection of strains for genome sequencing projects.

E. *faecium* MLST

The utility of MLST in evaluating enterococcal population structures was first explored in *E. faecium*. The need to understand epidemiology and population structure has been an important driver for research into this organism. In particular, the recognition that hospital-acquired infections are not caused by *E. faecium* clones that normally reside in the gut, but rather are caused by specific *E. faecium* clones that are acquired during hospitalization and are apparently well adapted to thrive in the perturbed microbiota of hospitalized patients, prompted research into potential virulence genes that were significantly enriched in these hospital clones. The first molecular epidemiological studies that showed host specificity of *E. faecium* and the existence of a distinct genogroup of hospital associated isolates used techniques such as amplified fragment length polymorphism (AFLP) and ribotyping (Borgen, Wasteson, Kruse, & Willems, 2002; Brisse, Fussing, Ridwan, Verhoef, & Willems, 2002; Coque, et al., 2005; Jureen, et al., 2003; Vancanneyt, et al., 2002; Willems, et al., 2000). However, these electrophoresis-band migration-based techniques are difficult to standardize and bands of different sequences may migrate similar distances in a gel, which makes these methods less than ideal for studying long term or global epidemiology. MLST overcomes these limitations (Maiden, et al., 1998).

In *E. faecium* MLST, the sequences of internal gene fragments (between 395 and 583 nucleotides [nt] in size) of seven housekeeping genes (*atpA*, *ddl*, *gdh*, *purK*, *gyd*, *pstS*, *adk*) are determined (Homan, et al., 2002). An allele number is assigned to the sequence of each MLST gene, which results in a "barcode" of seven numbers for each strain. This barcode is represented by a sequence type (ST). One of the major advantages of MLST is its portable nature: the presence of a large and publicly accessible database of *E. faecium* MLST profiles (http://efaecium.mlst.net/) enables researchers from around the world to add allelic profiles for strains that they have collected and typed by using MLST. The use of a uniform typing method and the establishment of databases that can be consulted through the Internet also leads to a unambiguous nomenclature for clones, which is pivotal for studying the global spread and long-term transmission of particular strains. This has resulted in a database that consists of thousands of strains that comprise nearly a thousand distinct STs. This dataset, which for all strains

combined consists of more than 6 Mbp of sequence, also allows for the study of evolutionary relationships between STs within the species *E. faecium*.

Initially, analyses of the genetic relatedness of *E. faecium* MLST data were performed using eBurst (Feil, Li, Aanensen, Hanage, & Spratt, 2004), an algorithm widely used to study relatedness of isolates and to generate hypotheses about patterns of recent evolutionary descent in bacteria. This resulted in the identification of a large cluster of clinical *E. faecium* isolates that was first termed lineage C1 and later renamed clonal complex 17 (CC17), after ST17, from which these clinical isolates presumably evolved (Willems, et al., 2005). Practically all strains from CC17 were found to be resistant to ampicillin and ciprofloxacin (Leavis H. L., Willems, Top, & Bonten, 2006; Willems, et al., 2005). Strains from CC17 were enriched for several genes with putative roles in virulence, such as the large surface protein *esp*, and carbohydrate metabolism (Heikens, van Schaik, Leavis, Bonten, & Willems, 2008; Leavis H. L., et al., 2007). The most prominent marker for such CC17 strains was found to be the presence of IS*16* (Leavis H. L., et al., 2007; Werner, et al., 2011). However, the assumption that CC17 isolates have emerged from a single "founder" ST appears to be erroneous. This was first suggested by Turner *et al*. (Turner, Hanage, Fraser, Connor, & Spratt, 2007), who showed that eBurst performs unreliably in species in which recombination is a more prominent driver of genetic diversity than mutation. *E. faecium* has a high ratio of recombination over mutation (Willems, et al., 2005) which indicates that algorithms other than eBURST would better reflect *E. faecium*'s population structure. Indeed, analysis of MLST data by algorithms that account for homologous recombination events indicate that clinical *E. faecium* isolates are not closely linked in purely evolutionary terms (Willems & van Schaik, 2009).

An alternate approach is to perform an analysis of population structure that has the power to identify deep-branching lineages as well as recombination between them. This can be implemented by using Bayesian analysis of genetic population structure (BAPS) software (Corander, Marttinen, Sirén, & Tang, 2008; Corander & Tang, 2007; Tang, Hanage, Fraser, & Corander, 2009). BAPS uses a statistical model to partition molecular variation in a population into different groups based on both clonal ancestry and recombination patterns, as identified from their DNA sequence. BAPS analysis of 491 distinct STs found among 1720 *E. faecium* isolates identified 13 groups of related *E. faecium* strains (Willems, et al., 2012). Phylogenetic analysis based on concatenated MLST gene sequences of isolates contained in the two largest BAPS groups (BAPS 2-1 and BAPS 3-3) showed that currently circulating clinical isolates belong almost exclusively to three different lineages (lineage-17, lineage-18, and lineage-78) (Willems, et al., 2012). BAPS analysis indicated that lineage-78, which belongs to BAPS 2-1 and which has emerged in the first decade of the 21st century as a major cause for nosocomial infections, seems to have an evolutionary history which is distinct from lineages-17 and -18 that are both contained in BAPS 3-3.

This means that these three major hospital lineages stem from at least three different ancestral strains, and have independently acquired genes that characterize clinical isolates through convergent evolution (Willems, et al., 2012). This observation raises questions as to exactly how monophyletic strains belonging to the group previously termed CC17 are, and precisely how long ago they may have emerged. BAPS analysis also indicated that clinical isolates are unrelated to *E. faecium* strains that commonly colonize healthy humans as commensals, which belong to BAPS group 1, but instead that the genetic evolution of hospital clones possibly involved animals (Willems, et al., 2012). BAPS analysis showed that hospital-associated *E. faecium* have undergone lower levels of admixture (cases in which sequence types contain sequence characteristics of more than one subpopulation) as compared to animal isolates, which is indicative of either fairly restricted recombination or recent emergence. This suggests that although the evolutionary development of hospital-associated *E. faecium* from animal isolates most probably included horizontal gene transfer with successive acquisition of adaptive elements like *esp*, once adapted to a distinct pathogenic niche, hospital-associated *E. faecium* may have become isolated (Willems, et al., 2012).

The availability of large amounts of MLST gene sequences has also allowed for a detailed analysis of the origin of genetic diversity in the species. Observations that the majority (60%) of gene tree topologies of individual MLST

genes are not congruent and that a high average number of nucleotide differences occur in single-locus variants indicate that genetic diversity in *E. faecium* has mainly been driven by recombination and not by mutation (Willems, 2010).

E. *faecalis* MLST

An MLST scheme has also been developed to study *E. faecalis* population structure (http://efaecalis.mlst.net) (Ruiz-Garbajosa, et al., 2006). Similar to *E. faecium*, this scheme relies on the amplification and sequencing of internal regions of seven housekeeping genes (*gdh*, *gyd*, *pstS*, *gki*, *aroE*, *xpt*, and *yqiL*) for which diversifying selection signatures are low (Ruiz-Garbajosa, et al., 2006). Using this system, clonal complexes CC2 and CC9 were found to be enriched among hospital-derived strains (Ruiz-Garbajosa, et al., 2006). (Note that there is no relationship between *E. faecium* and *E. faecalis* ST and CC number designations.) Earlier studies that relied on different MLST schemes indicated that several infection-derived, β-lactamase-producing, or vancomycin-resistant *E. faecalis* strains were clonally related (Nallapareddy, Duh, Singh, & Murray, 2002; Nallapareddy, Wenxiang, Weinstock, & Murray, 2005). This lineage was designated the BVE (Bla^{+}-Van^{R}-endocarditis) clonal complex (Nallapareddy, Wenxiang, Weinstock, & Murray, 2005), which overlaps *E. faecalis* CC2 in the more commonly used MLST scheme (Ruiz-Garbajosa, et al., 2006). Analysis of 106 *E. faecalis* strains isolated worldwide from the early 1900s to 2006 revealed that acquired antibiotic resistance is enriched in the CC2, CC8, and CC9 lineages (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007). Analysis of more recently isolated strains from Europe (primarily isolated from 2006-2009) found that multidrug resistance is enriched in the CC2, CC16, and CC87 lineages, while CC2 and CC87 lineages were found nearly exclusively in hospitals (Kuch, et al., 2012). Collectively, these studies and others (Freitas, Novais, Ruiz-Garbajosa, Coque, & Peixe, 2009; Kawalec, et al., 2007; Willems, Hanage, Bessen, & Feil, 2011) highlight the success of certain MLST lineages—in particular, CC2—in acquiring multidrug resistance, as well as in hospital persistence and infection. CC2 and CC9 have been identified as potential "high-risk" *E. faecalis* lineages (Leavis, Bonten, & Willems, 2006), although CC87 may displace CC9 as a potential high-risk lineage in Europe (Kuch, et al., 2012).

MLST analyses have revealed a potentially important difference between *E. faecalis* and *E. faecium*. For both species, particular CCs appear to be significantly associated with hospital-derived isolates. However, most *E. faecalis* CCs contain both clinical and commensal isolates, while the same is not true for *E. faecium*. This suggests that human commensal and clinical *E. faecalis* isolates are not as evolutionarily distinct as those in *E. faecium*.

CGH, PCR, and blot-based screening methods have shown that horizontally acquired traits other than antibiotic resistance are enriched in certain *E. faecalis* MLST lineages. PCR and blot screens found that certain *E. faecalis* PAI genes are enriched in infection-derived *E. faecalis* isolates (Nallapareddy, Duh, Singh, & Murray, 2002; Shankar, Baghdayan, & Gilmore, 2002), and in the CC2, CC8, and CC9 lineages (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007). The genome sequence of *E. faecalis* V583, a bloodstream infection-derived, vancomycin-resistant ST6/CC2 strain and the first *Enterococcus* strain to be sequenced (Paulsen, et al., 2003), allowed for the development of microarrays that could be used for CGH studies (Aakra, Nyquist, Snipen, Reiersen, & Nes, 2007; Lepage, et al., 2006; Solheim, Aakra, Snipen, Brede, & Nes, 2009; Solheim, et al., 2011). CGH studies of *E. faecalis* strains of diverse provenance (Aakra, Nyquist, Snipen, Reiersen, & Nes, 2007; McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007), as well as more focused studies on food (Lepage, et al., 2006) and infant fecal isolates (Solheim, Aakra, Snipen, Brede, & Nes, 2009), have found that most variation between these strains and a V583 reference occurs in regions of the V583 genome that were thought to be horizontally acquired (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007; Paulsen, et al., 2003). One CGH study identified a prophage (V583 prophage 3), a putative genomic island (V583 ORFs EF1847-EF1897) sometimes referred to as *efaB5* (Lepage, et al., 2006), and a putative genomic island flanking the V583 *vanB* transposon as being among genes enriched in CC2 strains of hospital and non-hospital origins, as compared to non-CC2 strains (Solheim, et al., 2011). In this study, none of the well-characterized *E. faecalis*

virulence factors were among the CC2-enriched genes (Solheim, et al., 2011), which highlights the nature and occurrence of *E. faecalis* strains related to hospital isolates in other ecologies. Collectively, CGH studies appear to show that the acquisition and exchange of mobile genetic elements are major contributors to the genomic diversity of *E. faecalis*. Further, the extent of horizontally acquired material in *faecalis* genomes varies among different MLST lineages, with certain elements being enriched in CC2.

Similar to *E. faecium*, population genetic analyses of MLST alleles indicate that the *faecalis* species is highly recombinogenic, which may contribute to its genome diversity. Alleles at different MLST loci are significantly associated with each other when considering both distantly related STs and STs within CCs that may have recently diversified, which is suggestive of a clonal population that evolves primarily by mutation (Ruiz-Garbajosa, et al., 2006). However, this association disappears if only distantly related STs are considered (Ruiz-Garbajosa, et al., 2006). Additionally, gene trees for individual MLST loci are incongruent, and individual MLST alleles are distributed throughout the concatenated MLST phylogeny (Ruiz-Garbajosa, et al., 2006). These observations are consistent with an epidemic population structure, where recombination occurs frequently among all members of the population—yet signatures of recombination are masked by the emergence of a highly successful lineage (or perhaps multiple lineages) that rapidly increase in frequency (Smith, Smith, O'Rourke, & Spratt, 1993). The CC2 lineage has not been detected in isolates collected prior to the 1980s (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007), suggesting that it has recently emerged and rapidly disseminated. Notably, because of the recombinogenic nature of *E. faecalis*, eBurst analysis may oversimplify the population structure of the species, as has occurred for *E. faecium* MLST analyses. It remains to be determined whether BAPS analysis will alter our current view of *E. faecalis* population structure.

Early Comparative Genome Analysis Reveals the Existence of a Pathogenicity Island in *E. faecalis*

In a retrospective study of an outbreak of 206 enterococcal bacteremias at the University of Wisconsin Hospitals and Clinics over a 17-month period, 190 were found to be caused by *E. faecalis*, most of which were resistant to high levels of aminoglycosides and macrolides (Huycke, Spiegel, & Gilmore, 1991). Moreover, nearly half of the 190 infections were caused by a single *E. faecalis* lineage, as defined by pulsed-field electrophoresis (PFGE) patterns (now known to be ST6/CC2 (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007)), which was multidrug resistant as well as hemolytic, whereas the remainder were caused by largely non-hemolytic idiosyncratic strains with few identities by PFGE (Huycke, Spiegel, & Gilmore, 1991). This indicated that one strain, termed MMH594, had become highly hospital-adapted, antibiotic resistant, and unusually pathogenic. Nearly simultaneously, the first vancomycin-resistant *Enterococcus* in the United States was isolated at Barnes-Jewish Hospital at Washington University (Sahm, et al., 1989). This strain, V583 (now also known to be ST6/CC2 (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007)) was identified as *E. faecalis*, and represented a novel vancomycin resistance phenotype, termed VanB. Identical VanB strains were isolated over a span of three months from the bloodstream, urine, and feces of a chronically infected patient who had received vancomycin therapy prior to the first isolation (Sahm, et al., 1989). An identical isolate was later obtained from another patient in the same ICU, but who had not received vancomycin therapy, and identical isolates were obtained from urine and blood over the course of two weeks (Sahm, et al., 1989).

In examining the genomes of the multidrug-resistant hospital outbreak strain from Wisconsin and the vancomycin-resistant strains from St. Louis by sequence analysis, a number of new traits for enterococci were discovered and found to be common to the hospital strains from both locations. These included a capsule (Hancock & Gilmore, 2002; Hancock, Shepard, & Gilmore, 2003), a novel adhesin termed enterococcal surface protein (Esp) that was enriched in clinical isolates (Shankar, et al., 2001; Shankar V. , Baghdayan, Huycke, Lindahl, & Gilmore, 1999), and a bile acid hydrolase. An operon for the enterococcal cytolysin was also found, which is a factor that confers lethality to enterococcal infection in humans (Huycke, Spiegel, & Gilmore, 1991) and animals (Chow, et al., 1993; Garsin, et al., 2001; Ike, Hashimoto, & Clewell, 1984; Ike, Hashimoto, & Clewell,

1987; Jett, Jensen, Nordquist, & Gilmore, 1992; Singh, Qin, Weinstock, & Murray, 1998); is enriched in bloodstream *E. faecalis* isolate collections (Huycke & Gilmore, 1995; Huycke, Spiegel, & Gilmore, 1991; Ike, Hashimoto, & Clewell, 1987); is novel in structure (Bogie, Hancock, & Gilmore, 1995; Booth, et al., 1996; Coburn & Gilmore, 2003; Coburn, Hancock, Booth, & Gilmore, 1999; Cox, Coburn, & Gilmore, 2005; Gilmore, Segarra, & Booth, An HlyB-type function is required for expression of the Enterococcus faecalis hemolysin/bacteriocin, 1990); and is regulated by a novel quorum-sensing mechanism (Haas, Shepard, & Gilmore, 2002) that enables enterococci to detect target cells at a distance (Coburn, Pillar, Jett, Haas, & Gilmore, 2004). Aggregation substance, an enterococcal surface factor that causes clumping, was first described in the Clewell lab (Yagi, et al., 1983), and was shown to be lethally synergistic with the cytolysin in endocarditis (Chow, et al., 1993), was also present in the genomes of these clinical isolates. Intriguingly, all of these factors, except the capsule, occurred on a single macrorestriction fragment (Shankar, Baghdayan, & Gilmore, 2002). Further sequence analysis revealed that these traits were organized into an island of over 150 kb in size that included a phage-like integrase and excisionase, differed in G+C content from the rest of the chromosome, possessed flanking 10 bp direct repeats, and was inserted at a lysyl-tRNA locus (Shankar, Baghdayan, & Gilmore, 2002), which was reminiscent of the organization of PAIs in Gram-negative bacteria (Langille, Hsiao, & Brinkman, 2010). The *E. faecalis* PAI was the first canonical PAI to be identified in a Gram-positive bacterium.

Compared to those occurring in V583 and V586, the 153,571 bp PAI of strain MMH594 was found to be closest to the prototype, in that the islands from the vancomycin-resistant St. Louis strains V583 and V586 from the chronically infected patient possessed additional IS elements (IS*256* and IS*905*) that disrupted the cytolysin operon (Shankar, Baghdayan, & Gilmore, 2002) (Figure 1). Moreover, these IS elements volatilized this region of the island, which resulted in a 17 kb deletion that distinguishes V583 from V586 and occurs at a very high frequency of 1 in 10^3 V586 cells (Shankar, Baghdayan, & Gilmore, 2002). This indicates that the chronically infected St. Louis HIV/AIDS patient was infected with a mixed population of V583/V586 cells. Adaptation of the microbe in chronic infection, as seen with V583/V586, is similar to that seen with *P. aeruginosa* in chronically infected cystic fibrosis lungs (Hoboth, et al., 2009). Thus, from comparative sequence analysis and *in vitro* experiments, it is clear that the PAI is a dynamic component of *E. faecalis* genomes, and rapid changes in its structure are likely to impact *E. faecalis* virulence (Figure 1).

Further studies have found that the PAI has a modular structure and likely evolves by accretion of modules obtained through horizontal gene transfer (McBride, et al., 2009). A collection of 53 *esp*$^+$ *E. faecalis* strains were interrogated for the presence of MMH594 PAI genes using dot blot hybridization, PCR, and a novel pathoarray with probes that corresponded to each of the 129 open reading frames of the MMH594 PAI. Presence of *esp* was chosen as a marker for PAI presence in the 53 strains, as *esp* in *E. faecalis* had only been detected within the PAI. Incongruence between phylogenies built from MLST alleles and from PAI gene content in the 53 strains suggested that the PAI and the core *E. faecalis* genome evolve through different mechanisms. Transversal clustering of PAI hybridization data revealed that the canonical PAI contains six modules (named A through F, moving from 5′′ to 3′ PAI ends) of consecutive genes (with the exception of module E) that appear to be acquired or inherited as discrete units. Module A is composed of the pAM373-like plasmid remnant that is present at the 5′ end of the MMH594 PAI (PAI ORFs EF0005-EF0033; (Shankar, Baghdayan, & Gilmore, 2002)). That this module has high identity to a pheromone-responsive plasmid provides a straightforward mechanism for its horizontal transfer. It is flanked on one end by a putative phage integrase and excionase and by transposases on the other end, which could also contribute to the mobility of this module. Module B (EF0042-EF0049) contains the cytolysin operon and occurs 3′ to a conjugated bile salt hydrolase gene that is flanked by transposases (EF0039-EF0041). Module C (EF0051-EF0074) possesses *esp* and the *gls24*-like gene, among others, as well as an internal recombinase gene and flanking transposase (EF0075). This transposase also flanks Module D (EF0076-EF0092), which is composed of sugar uptake and metabolism genes. Module E is non-contiguous (EF0093-EF0108 and EF0124-EF0128), and encodes various genes that are likely involved in metal transport and cofactor biosynthesis. Module F (EF0108-EF0122) is internal to Module E and possesses several transposases that may provide mobility to this module.

Collectively, these data suggest that the PAI evolves by the accretion of discrete modules bounded by transposases, and is transmitted by horizontal gene transfer, as opposed to a model where a single progenitor PAI entered an ancestral *E. faecalis* strain and was diversified by differential gene loss in diverging MLST lineages. If an evolution-by-accretion model is correct, we could expect to discover novel PAI modules in comparative genome analyses. Consistent with this proposal, a novel PAI module has been identified in genomes of the CC2 strains HH22 and TX0104 (Solheim, et al., 2011). This module encodes several hypothetical proteins and a predicted mucin-binding domain protein (Solheim, et al., 2011), and occurs between MMH594 PAI ORFs EF0040 and EF0065, which correspond to Module B and part of Module C as defined above. Despite its absence in MMH594 and V583 genomes, the presence of this module was found to be enriched among CC2 strains, compared to non-CC2 strains, in a PCR-based screening analysis (Solheim, et al., 2011). This result suggests that comparative genomics will reveal much about the module diversity and evolutionary history of the PAI. Indeed, one *E. faecalis* strain for which genome sequence is available, T1, possesses a ~60 kb PAI located at the same PAI *att* site, which is composed entirely of novel modules not present in the prototypical MMH594 PAI (Palmer and Gilmore, unpublished).

The First Complete Enterococcal Genome, *E. faecalis* V583, and Comparison With OG1RF

E. faecalis V583 was provided to TIGR in 1998 and its genome sequence was published in 2003 (Paulsen, et al., 2003). The release of this genome was significant for many reasons, including its provision of a complete, high-quality reference for subsequent genome projects. V583 was selected because it was the first vancomycin-resistant *Enterococcus* isolated in the United States, and as noted above, was found to be similar to MMH594 and representative of the ST6/CC2 lineage (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007; Sahm, et al., 1989). A significant and striking finding of the V583 genome is the amount of mobile DNA contained within it—calculated to be over 25% of the genomic content (Paulsen, et al., 2003). This includes seven predicted prophages, multiple integrated plasmids, IS elements and genomic islands (including the PAI), a *vanB*-type transposon that confers vancomycin resistance, and three extrachromosomal plasmids (pTEF1, pTEF2, and pTEF3), including one that encodes multiple antibiotic resistance genes (pTEF1) (Lepage, et al., 2006; McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007; Paulsen, et al., 2003) (Figure 2). Two of the pTEF plasmids, pTEF1 (66.3 kb) and pTEF2 (57.7 kb), are predicted pheromone-responsive plasmids similar to the extensively studied pheromone-responsive plasmids pAD1 (Clewell, 2007) and pCF10 (Dunny, 2007), respectively (see Extrachromosomal and Mobile Elements in Enteroocci). These plasmids detect small peptides produced by plasmid-free cells to induce high-efficiency transfer (transfer frequencies of up to one transconjugant per 10–100 donors; (Hirt, Schlievert, & Dunny, 2002; Huycke, Gilmore, Jett, & Booth, 1992; Licht, Laugesen, Jensen, & Jacobsen, 2002)), and their replication mechanisms appear to be functional only in the *faecalis* species (Palmer, Kos, & Gilmore, 2010). pTEF3 (18.0 kb) possesses the broad host range Inc18 replicon of pAMβ1 (Paulsen, et al., 2003). The remarkable amount of horizontally acquired material in the V583 genome leads to the proposition that enterococci have a special propensity to acquire and disseminate mobile elements, such as those that encode antibiotic resistance genes (Paulsen, et al., 2003).

A second complete *E. faecalis* genome was published in 2008 (Bourgogne, et al., 2008), which provides a comparator to V583. *E. faecalis* OG1RF is a laboratory-generated rifampicin and fusidic acid-resistant derivative of the non-antibiotic resistant human oral isolate OG1 (Gold, Jordan, & van Houte, 1975), which was originally isolated prior to 1975. OG1RF is an ST1 strain, and is not a member of any of the high-risk *E. faecalis* MLST lineages. Compared to V583, the OG1RF genome has little mobile content, lacks plasmids and the PAI, and possesses only one prophage that is core to the *faecalis* species (prophage 2; (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007)), as well as a putative transposable element that lacks antibiotic resistance genes (Bourgogne, et al., 2008). The abundance of horizontally acquired DNA in V583, as compared to OG1RF, results in a large size discrepancy between their genomes: 3.36 Mb for V583, compared to 2.74 Mb for OG1RF.

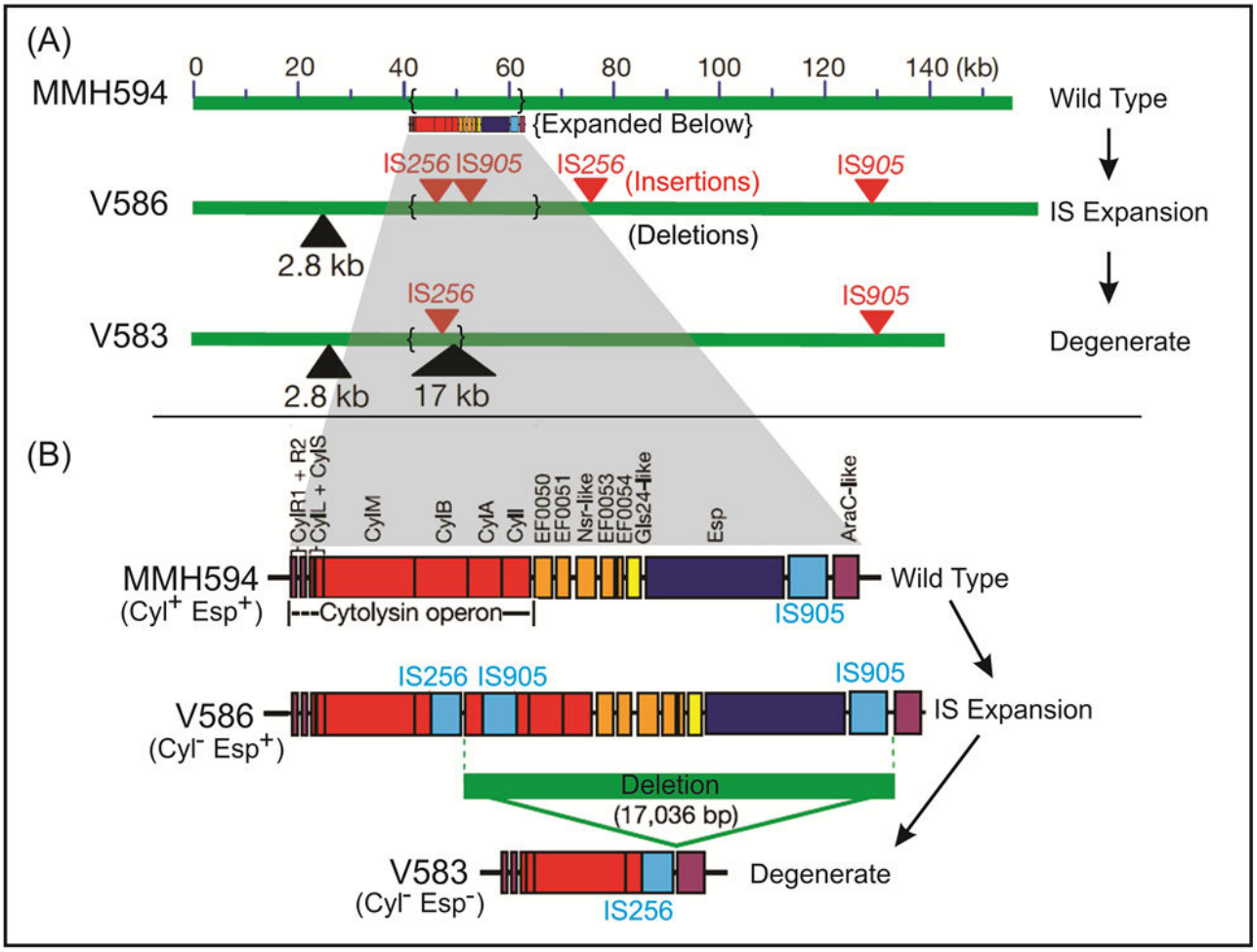


Figure 1. A dynamic PAI in hospital-adapted *E. faecalis.* Alignment of PAI sequences from hospital infection isolates E. faecalis MMH594, V586, and V583. (A) Overall comparison of PAIs reveals expansion of the prototypical MMH594 PAI by IS acquisition (indicated by inverted red triangles) in strain V586, followed by ~17 kb deletion of an IS-flanked region (indicated by a black triangle) in strain V583. (B) Expanded view of PAI showing genes affected by PAI structural variation. The prototypical PAI of MMH594 encodes cytolysin and Esp, among other factors important for virulence. The cytolysin operon of V586 is disrupted as a result of IS expansion; this strain does not produce cytolysin. V583 does not produce cytolysin or Esp as a result of IS expansion in the V586 progenitor and a subsequent ~17 kb internal deletion.

Interestingly, despite the 620 kb of additional coding potential in the V583 genome, including the PAI, OG1RF can cause disease in animal models of infection (Bourgogne, et al., 2008), which highlights the point that even commensal strains of *E. faecalis* can be opportunists, and that basic fitness functions are important to the ability of members of this species (and probably genus) to cause disease.

Dynamic *E. faecalis* Genomes and CRISPR-Cas Defense

Further genome sequencing has illuminated the extent of mobile content and range of genome sizes in the *faecalis* species. A third complete *E. faecalis* genome has been announced (Brede, Snipen, Ussery, Nederbragt, & Nes, 2011), although a detailed comparative analysis has not yet been published. *E. faecalis* 62 was isolated from the feces of a healthy human infant and is a ST66 strain, which is a lineage that has not been associated with either infection or hospitals (Solheim, Aakra, Snipen, Brede, & Nes, 2009; Solheim, et al., 2011). The extrachromosomal elements of this strain, and to some extent the PAI, have been described (Brede, Snipen, Ussery, Nederbragt, & Nes, 2011), and their abundance and presence, respectively, indicate that the *E. faecalis* 62

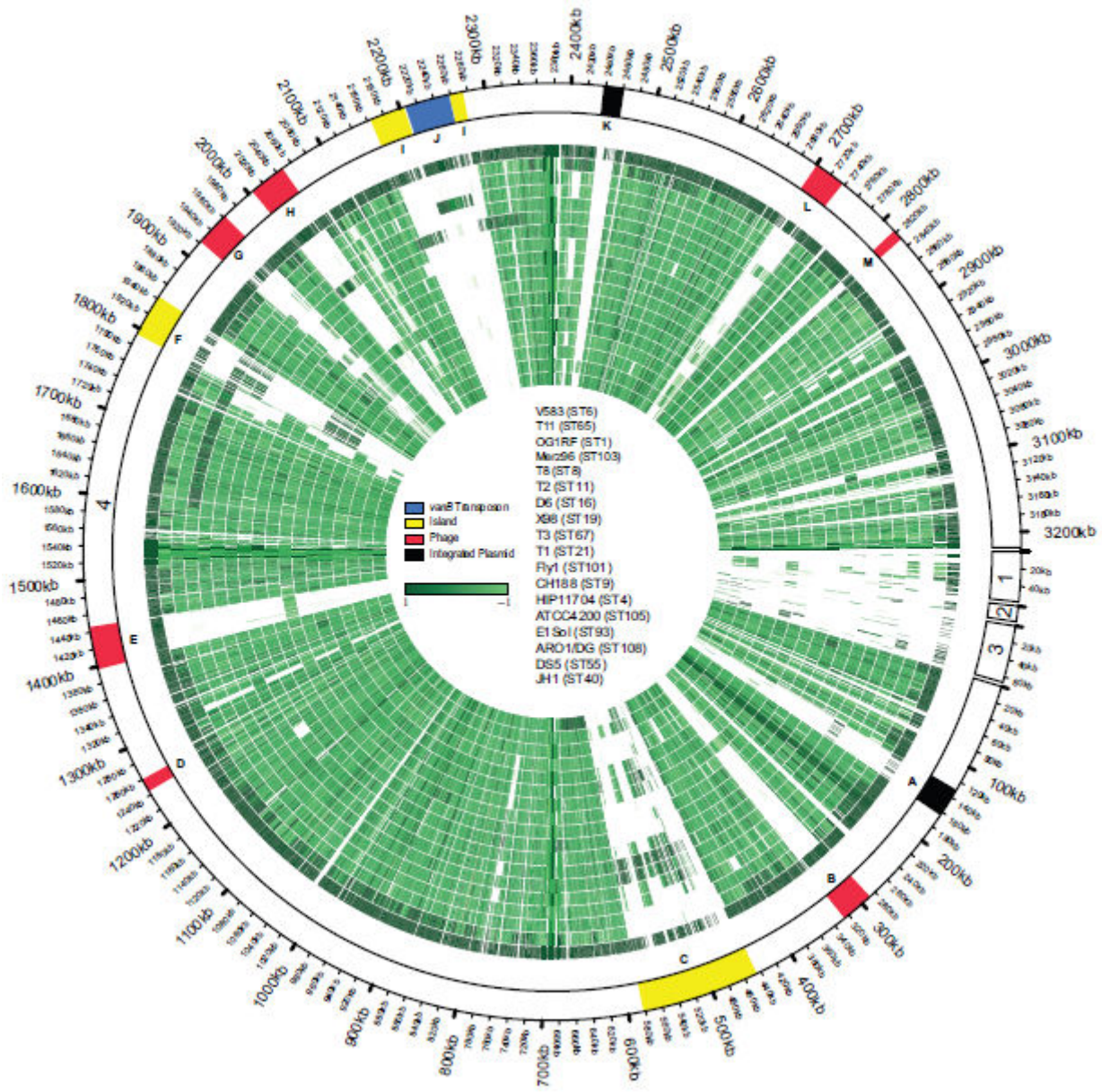


Figure 2. *E. faecalis* genome plot. The outermost ring shows E. faecalis V583 chromosomal (scaffold 4) and plasmid scaffolds (scaffold 1, pTEF2; scaffold 2, pTEF3; scaffold 3, pTEF1), with each V583 gene represented as a radial position along the ring. The locations of V583 variable regions are shown (98): A, integrated plasmid; B, prophage 1; C, PAI; D, prophage 2; E, prophage 3; F, putative island efaB5; G, prophage 4; H, prophage 5; I, putative island; J, vancomycin resistance (vanB) transposon; K, integrated plasmid; L, prophage 6; M, prophage 7. Seventeen additional E. faecalis genomes are represented by the rings below V583. Genes are colored by phylogenetic distance from V583 (from dark to light green with increasing phylogenetic distance). Strains shown, from outermost to innermost rings, are V583, T11, OG1RF, Merz96, T8, T2, D6, X98, T3, T1, Fly1, CH188, HIP11704, ATCC4200, E1Sol, AR01/DG, DS5, and JH1. Patterns suggestive of recombination (darker colored genes indicate close phylogenetic relatedness to V583) occur in Merz96, T2, and JH1, which flank genomic islands.

genome is more similar to that of V583 than OG1RF in its mobile genome content. *E. faecalis* 62 possesses three predicted plasmids—one similar to the cryptic plasmid pS86 (EF62pA; 5.1 kb) and two putative pheromone-responsive plasmids that are similar to pCF10 (EF62pB; 51.1 kb) and pAM373 (EF62pC; 55.4 kb) (Brede, Snipen, Ussery, Nederbragt, & Nes, 2011). A pseudotemperate linear phage (EF62φ; 30.5 kb) is also present, as is a chromosomally encoded Tn*916* and a variant PAI encoding *esp* (Brede, Snipen, Ussery, Nederbragt, & Nes,

2011). Collectively, this results in a genome size of 3.13 Mb, which is intermediate between those of V583 and OG1RF.

Sixteen additional *E. faecalis* genomes were sequenced and assembled to high-quality draft status as part of a enterococcal genome sequencing project that increased the amount of available sequencing data for the genus by approximately 10-fold (Palmer, et al., 2010). The 16 *E. faecalis* strains were selected to represent the deepest nodes in the *faecalis* MLST phylogeny, and were isolated over an approximately 80-year period from clinical and non-clinical sites around the world (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007). Analysis of the 16 draft *E. faecalis* genomes has further supported the proposition that mobile element acquisition is a major contributor to genome diversity in the *faecalis* species. The genome sizes of all 16 strains are intermediate between the extremes of V583 and OG1RF, with multidrug-resistant strains generally possessing larger genomes (Palmer, et al., 2012). Strains with genomes larger than 3 Mb encode significantly more protein domains associated with mobile elements, including plasmid replication initiation and mobilization domains, a plasmid addiction toxin domain, a transposase domain, and the anti-restriction protein ArdA domain, which is encoded by the self-mobilizable transposon Tn*916* (Palmer, et al., 2012). This is consistent with these strains acquiring foreign genetic elements conferring new traits, which results in larger genomes. The observed range of genome sizes suggests that the propensity to acquire mobile elements *in situ* varies within the *faecalis* species.

Genome size and the extent of acquired antibiotic resistance appear to be related to the CRISPR-*cas* status of *E. faecalis* strains. Clustered, regularly interspaced short palindromic repeats (CRISPR) with CRISPR-associated (*cas*) genes provide a type of acquired immune defense in prokaryotes (Wiedenheft, Sternberg, & Doudna, 2012). CRISPR-Cas systems have been shown to block phage infections and plasmid uptake, and thus are endogenous barriers to horizontal gene transfer (see Enterococcal bacteriophages and genome defense). In the comparative analysis of V583 and OG1RF, it was noted that OG1RF possesses CRISPR-*cas* while V583 lacks it, which suggests that CRISPR-Cas defense prevented the accumulation of mobile genetic elements in OG1 prior to its isolation from the human mouth (Bourgogne, et al., 2008). Supporting a role for CRISPR-Cas systems in *E. faecalis* genome defense, genome size distributions significantly differ between *E. faecalis* strains that possess or lack CRISPR-*cas*, with strains that lack CRISPR-*cas* having larger genomes (with an average size of 3.1 Mb, compared to 2.9 Mb for strains that possess CRISPR-*cas*) (Palmer, et al., 2012). Similarly, strains that lack CRISPR-*cas* in their genomes encode significantly more conserved protein domains associated with mobile elements (Palmer, et al., unpublished). Analysis of 48 *E. faecalis* strains using comparative genomics and a PCR-based screening approach found a significant relationship between acquired antibiotic resistance and CRISPR-*cas*, with strains that lacked CRISPR-*cas* possessing more acquired antibiotic-resistance genes (Palmer & Gilmore, 2010). Strains deficient for CRISPR-*cas* are additionally enriched for cytolysin and aggregation-substance genes (Lindenstrauss, et al., 2011), both of which are encoded on pheromone-responsive plasmids and the PAI. Finally, all CC2 strains characterized thus far lack CRISPR-*cas* (Palmer & Gilmore, 2010), which suggests that the absence of this line of defense is a core property of the lineage that perhaps contributes to its success in acquiring novel genomic content. The role of CRISPR-Cas defense in regulating the influx of mobile elements into *E. faecalis* genomes is of high interest.

The Recombinogenic Nature of *E. faecalis*

Genomics has illuminated the extent of recombination in *E. faecalis*. As might be expected from MLST analyses, the inter-relationships of 18 *E. faecalis* strains (16 draft genomes discussed above, along with V583 and OG1RF) could not be confidently resolved by a phylogeny built from 847 core gene sequences (Palmer, et al., 2012). In pairwise comparisons of the same 18 strains, the average nucleotide identity in shared genes varied little (97.8-99.5% average identity). However, total shared gene content varied widely, which is likely a function of the varying genome sizes of the strains (Palmer, et al., 2012). This is consistent with *E. faecalis* strains generally being closely related in their core genes, with most variation introduced through the acquisition of horizontally transferred material. That said, signatures of recombination are apparent in some *E. faecalis* genomes (Fig. 2).

Three *E. faecalis* strains, JH1 (ST40), Merz96 (ST103), and T2 (ST11), appear to have acquired genomic islands (the PAI, *efaB5*, and a putative genomic island that flanks the V583 *vanB* transposon, respectively), as well as a sequence adjacent to the islands, from strains closely related to V583 (Palmer, et al., 2012). The recombinant region of Merz96 overlaps the *fsr* locus, of interest because variation in the *fsr* region influences production of gelatinase (42). The Merz96 *fsr* region was previously shown to be similar to that present in V583, MMH594, and HH22—all ST6 strains (Galloway-Peña, Bourgogne, Qin, & Murray, 2011). It is likely that Merz96 acquired this region, as well as *efaB5,* through recombination with a ST6 strain.

A mechanism for recombinatorial transfer of large chromosomal regions between *E. faecalis* strains has been elucidated, and provides an explanation for decades of observations that implicate chromosomal DNA exchange in enterococci (Manson, Hancock, & Gilmore, 2010). In laboratory mating experiments with V583 as donor and OG1RF as recipient, hybrid V583-OG1RF transconjugants were observed that possessed up to 857 kb of the V583 genomic sequence. Depending on the location of the selectable marker in the V583 donor genome, transconjugants were observed that possessed the entire PAI, the V583 capsule biosynthesis operon, the V583 vancomycin resistance operon, and/or V583 *gdh* and *yqiL* MLST markers, which generated a double locus MLST variant of OG1RF in one mating experiment (Manson, Hancock, & Gilmore, 2010). Displacement of the OG1RF CRISPR-*cas* locus by V583 sequence was also observed (Palmer & Gilmore, 2010). Transfer of large chromosomal regions was dependent upon the presence and transfer functions of either of the pheromone-responsive plasmids pTEF1 or pTEF2 (Manson, Hancock, & Gilmore, 2010). Both pTEF1 and pTEF2 were found to integrate into the chromosome at shared IS*256* sequences, and from there, likely initiated transfer induced by OG1RF pheromones, which resulted in mobilization of the V583 chromosome. The PAI and *efaB5* appear to be hotspots for IS element accumulation, and IS elements occur in the genomic island that surrounds *vanB* (Paulsen, et al., 2003), which suggests that pheromone-responsive plasmids that possess similar IS elements could readily integrate into these regions. It is likely that Merz96, JH1, and T2 have acquired genomic islands, along with surrounding sequences, by a similar mechanism. This pheromone-responsive plasmid-dependent chromosome transfer mechanism likely contributes to the recombinogenic population structure of the *faecalis* species.

E. faecium Genome Sequencing

Table 1 provides an overview of several important characteristics of the first 31 publicly available *E. faecium* genomes. Most (23/31) of these isolates originated from human infection or from hospitalized patients that were colonized by antibiotic resistant strains. Six genomes are from strains that were isolated from feces of healthy humans (and likely represent human commensal isolates). *E. faecium* strains from non-human sources are poorly represented in this first set of *E. faecium* genomes, with genomes of only two representative strains sequenced. Both strains had been isolated from dog feces (de Regt, et al., 2012). There is a large sample bias toward clinical isolates in *E. faecium* strains that have been selected for genome sequencing studies. However, clinical isolates form only a minute fraction of the total *E. faecium* population, and it is important that genomes of *E. faecium* strains from non-human sources be included in sequencing projects. Another bias that is evident from Table 1 is the overrepresentation of strains from Europe and North America. In fact, only three strains (a clinical isolate from Brazil and a commensal and a clinical isolate from Australia) do not originate from these two continents. More balance is needed to fully understand the global diversity of *E. faecium*.

The first completely finished *E. faecium* genome sequence became available in March 2012 and was determined for strain Aus0004. Aus0004 is a vancomycin-resistant strain that was isolated from the bloodstream of a patient in Melbourne, Australia, in 1998 (Lam, et al., 2012). The Aus0004 genome sequence was determined by a combination of 454 and Illumina sequencing and was finished by primer walking. A bacterial artificial chromosome (BAC) library was prepared to resolve chromosomal duplications. A second complete genome sequence was determined for strain TX16 (also named TX0016 or DO), which was isolated from the blood of a patient with endocarditis in 1992. Unassembled fragmentary sequences were first sequenced by Sanger

sequencing in 1999, and were completed with additional 454 sequencing and PCR to close gaps (Qin, et al., 2012).

Other draft *E. faecium* genomes have been generated by either 454 or Illumina sequencing. Most fragmentary assemblies are a consequence of the large number of repetitive elements in the genomes of *E. faecium*, which are challenging to resolve bioinformatically. These elements include IS elements (for example, there are 95 IS elements in Aus0004 [77]), but also repeats within protein coding sequences, such as those that are present in the gene that encodes the Esp surface protein (Leavis, et al., 2004). Practically all *E. faecium* strains with available genome sequences were isolated in the last 20 years. The oldest strain is E1636, which was isolated in 1961 from a bloodstream infection in the Netherlands. Interestingly, this strain encodes the tetracycline resistance gene *tetM* (van Schaik, et al., 2010), though this antibiotic had only been on the market since 1952. It would be of particular interest to sequence the genomes of *E. faecium* strains from the pre-antibiotic era, as this would definitively answer the question if *E. faecium* strains are a natural source of antibiotic-resistance genes or have acquired all their resistances through lateral gene transfer in the last 70 years. While limitations exist for the first available set of *E. faecium* genome sequences, these data have provided us with important new insights into the evolution and basic biology of this organism.

Table 1. Initial 31 publicly available *E. faecium* genome sequences

Strain	Source[a]	Country	Year	Sequencing Center[b]	Assembly size (nt)[c]	No. contigs	Resistances, virulence genes and remarks[d]	Ref
Aus0004	Clinical isolate from blood	AUS	1998	UM	2.96 Mbp	1	VAN; *esp*+	(Johnson, et al., 2010; Lam, et al., 2012)
TX16	Clinical isolate from blood	USA	1992	WU	2.70 Mbp	1	Also named DO or TX0016. AMP; hyl+	(Qin, et al., 2012)
1,141,733	Clinical isolate from blood	USA	2005	BI	2.87 Mbp	101		(Palmer, et al., 2010; Palmer, et al., 2012)
1,230,933	Clinical isolate from wound	USA	2005	BI	2.95 Mbp	304	AMP, CIP and VAN; *esp*+ and *hyl*+	(Palmer, et al., 2010)
1,231,408	Clinical isolate from blood	USA	2005	BI	2.89 Mbp	379	AMP and CIP	(Palmer, et al., 2010; Palmer, et al., 2012)
1,231,410	Clinical isolate from skin and soft tissue	USA	2005	BI	2.94 Mbp	230	AMP, CIP and VAN; *esp*+ and *hyl*+	(Palmer, et al., 2010; Palmer, et al., 2012)
1,231,501	Clinical isolate from blood	USA	2005	BI	2.80 Mbp	140		(Palmer, et al., 2010; Palmer, et al., 2012)
1,231,502	Clinical isolate from blood	USA	2005	BI	2.92 Mbp	220	AMP, CIP and VAN;*esp*+ and *hyl*+	(Palmer, et al., 2010; Palmer, et al., 2012)
C68	Isolated from feces of hospitalized individual	USA	1996	BI	2.94 Mbp	182	VAN, AMP; *esp*+ and *hyl*+	(Carias, Rudin, Donskey, & Rice, 1998)

Table 1. continued from previous page.

Strain	Source[a]	Country	Year	Sequencing Center[b]	Assembly size (nt)[c]	No. contigs	Resistances, virulence genes and remarks[d]	Ref
Com12	Isolated from feces of healthy individual	USA	2006	BI	2.69 Mbp	67		(Palmer, et al., 2010; Palmer, et al., 2012)
Com15	Isolated from feces of healthy individual	USA	2007	BI	2.77 Mbp	70		(Palmer, et al., 2010; Palmer, et al., 2012)
D344SRF	Spontaneous mutant of strain D344R	USA	?	BI	2.75 Mbp	215	ERY, CLI, FUS and RIF.	(Rice, et al., 2009)
E1039	Isolated from feces of healthy individual	NLD	1998	UMCU	2.50 Mbp	124		(van Schaik, et al., 2010)
E1071	Isolated from feces of hospitalized individual	NLD	2000	UMCU	2.70 Mbp	96	VAN, CHL, ERY, SPC, STR and TET	(van Schaik, et al., 2010)
E1162	Isolated from a bloodstream infection	FRA	1997	UMCU	2.71 Mbp	139	AMP and TET, *esp*+	(van Schaik, et al., 2010)
E1636	Isolated from a bloodstream infection	NLD	1961	UMCU	2.84 Mbp	223	AMP and TET	(van Schaik, et al., 2010)
E1679	Isolated from the tip of a vascular catheter	BRA	1998	UMCU	2.93 Mbp	340	AMP, CIP, ERY, GEN, SPC and VAN *esp*+	(van Schaik, et al., 2010)
E4452	Isolated from dog feces	NLD	2008	UMCU	2.77 Mbp	268	AMP	(de Regt, et al., 2012)
E4453	Isolated from dog feces	NLD	2008	UMCU	2.82 Mbp	374	AMP	(de Regt, et al., 2012)
E980	Isolated from feces of healthy individual	NLD	1998	UMCU	2.79 Mbp	131	ERY	(van Schaik, et al., 2010)
LCT-EF90	Derived from *E. faecium* type strain[e]	SWI	1940s?	BGI	2.77 Mbp	118		(Chang, et al., 2012)
PC4.1	Isolated from feces of healthy individual	AUS	2008	JCVI	2.81 Mbp	78		
TC 6	Transconjugant of mating between C68 and D344SRF	USA	?	BI	2.88 Mbp	278	Resistant to TET. *hyl*+	(Rice, et al., 2009)

Table 1. continued from previous page.

Strain	Source[a]	Country	Year	Sequencing Center[b]	Assembly size (nt)[c]	No. contigs	Resistances, virulence genes and remarks[d]	Ref
TX0082	Isolated from bloodstream of patient with infective endocarditis	USA	1999	WU	2.69 Mbp	151	AMP, ERY and VAN	(Nallapareddy, Singh, & Murray, 2006)
TX0133A[f]	Isolated from bloodstream of patient with infective endocarditis	USA	2006	WU	2.93 Mbp	223	VAN; *esp*+	(Arias, et al., 2007)
TX0133a01[f]	Isolated from bloodstream of patient with infective endocarditis	USA	2006	WU	3.07 Mbp	252	VAN; *esp*+; *hyl*+	(Arias, et al., 2007)
TX0133a04[f]	Isolated from bloodstream of patient with infective endocarditis	USA	2006	WU	2.92 Mbp	211	*esp*+	(Arias, et al., 2007)
TX0133B[f]	Isolated from bloodstream of patient with infective endocarditis	USA	2006	WU	2.93 Mbp	221	*esp*+	(Arias, et al., 2007)
TX0133C[e]	Isolated from bloodstream of patient with infective endocarditis	USA	2006	WU	2.91 Mbp	215	VAN; *esp*+	(Arias, et al., 2007)
TX1330	Isolated from feces of healthy individual	USA	1994	BCM	2.72 Mbp	156		(Singh, Weinstock, & Murray, 2002)
U0317	Isolated from urine of patient with a urinary tract infection	NLD	2005	UMCU	2.89 Mbp	227	AMP, CIP, ERY, GEN and STR. *esp*+ *hyl*+	(van Schaik, et al., 2010)

[a] All isolates are from human subjects, unless mentioned otherwise.
[b] Abbreviations used: UM: University of Melbourne; BI: Broad Institute; JCVI: J. Craig Venter Institute; UMCU: University Medical Center Utrecht; BGI: Beijing Genomics Institute; BCM: Baylor College of Medicine; WU: Washington University School of Medicine.
[c] The size of the assembly does not include the estimated sizes of gaps between contigs.
[d] The abbreviations used for antibiotic resistances are AMP: ampicillin; CIP: ciprofloxacin; VAN: vancomycin; ERY: erythromycin; CLI: clindamycin; RIF: rifampicin; FUS: fusidic acid; CHL: chloramphenicol, SPC: spectinomycin, STR: streptomycin; GEN: gentamicin.
[e] This strain was derived from the *E. faecium* type strain. The type strain was cultured at 15°C and 37°C for more than four weeks, after which DNA was isolated and sequenced. The *E. faecium* type strain was first deposited to culture collections in 1946. We have been unable to determine the isolation date and the exact source of this isolate.
[f] These strains originate from a study in which the evolution of an *E. faecium* strain was followed during antibiotic therapy and during *in vitro* culture. See (Arias, et al., 2007) for further details.

Toward a Genome-Based Reconstruction of *E. faecium* Evolution

In the first genome-based study of *E. faecium*, seven isolates were sequenced (148). Based on a phylogenetic analysis of the concatenated alignments of 649 protein sequences, it was shown that the human commensal strain E980 was relatively distantly related to the other six *E. faecium* isolates, which suggested the presence of a deeply branching division within the species. Interestingly, a similar division between two major clades within *E. faecium* had previously been suggested by non-sequence based studies, in which random amplified polymorphic DNA (RAPD)-PCR, AFLP, and PFGE were used to characterize a collection of *E. faecium* strains (Vancanneyt, et al., 2002; Vankerckhoven, et al., 2008).

A divide between two major clades in *E. faecium* was also found by analyses of MLST data and additional genome sequence-based studies (Galloway-Peña, Roh, Latorre, Qin, & Murray, 2012; Galloway-Peña, Rice, & Murray, 2011; Palmer, et al., 2012). BAPS analysis of MLST data (Willems, et al., 2012), combined with genome sequence data, confirmed a major split in the *E. faecium* population (Figure 3). To further analyze the differences between the two clades, termed A and B, Palmer *et al.* (Palmer, et al., 2012) performed pairwise comparisons of 8 *E. faecium* genomes to determine the average nucleotide identity (ANI) in genes shared between the genomes. Pairwise genome comparisons between clade A and clade B strains revealed ANI values that ranged from 93.9 to 95.6%. In examining other species, ANI values of 95 ± 0.5% have been found to correlate with 70% DNA-DNA hybridization values, which is the threshold usually used to distinguish different species (50). By this criterion, clade A and clade B would represent separate species-level subdivisions of *E. faecium.* A recent study including 51 newly sequenced *E. faecium* genomes (Lebreton, et al., 2013) provided evidence for a second split in the *E. faecium* population, which distinguishes modern clinical isolates (clade A1) from most animal-derived strains (clade A2) in the dataset. After removing recombined regions in the genomes, mutation rates were calculated using the BEAST algorithm (Drummond, Suchard, Xie, & Rambaut, 2012) for the three clades (A1, A2 and B) of the *E. faecium* population. Interestingly, a significantly higher mutation rate was found for clade A1 strains ($4.9 \times 10^{-5} \pm 0.3 \times 10^{-5}$ substitutions per nucleotide per year) than for both clade A2 ($3.6 \times 10^{-6} \pm 0.6 \times 10^{-6}$ substitutions per nucleotide per year) and clade B ($1.3 \times 10^{-5} \pm 0.2 \times 10^{-5}$ substitutions per nucleotide per year). The inferred mutation rates for each clade were used to estimate the time of divergence between the different *E. faecium* clades A1, A2, and B, placing the time of the split between clade A and B at 2,776 ± 818 years ago and the split between clade A1 and clade A2 at 74 ± 30 years ago (Lebreton, et al., 2013).

Human commensal isolates overwhelmingly belong to clade B (which corresponds to BAPS-group 1 defined by Willems *et al.* (Willems, et al., 2012)), while strains from other sources, including hospital-acquired infections and animal isolates, cluster in clade A1 and A2, respectively (Lebreton, et al., 2013). However, this ecological distinction between two clades is not absolute: of the strains with sequenced genomes, strain 1,141,733 was isolated from a bloodstream infection but clusters in clade B, while strain E1039, which was isolated from the feces of a healthy volunteer, clusters in clade A. Notably, the presence of evolutionarily distinct clades within what we currently characterize as *E. faecium* is very different from what genomics has revealed for *E. faecalis*, where isolates can be so closely related in core genes as to be virtually indistinguishable from one another.

Recently, a lack of CRISPR-Cas defense has been implicated in the evolution of multidrug-resistant *E. faecalis* (Palmer & Gilmore, 2010). CRISPR-Cas systems appear to be relatively rare in *E. faecium*. Among the first 31 publicly available *E. faecium* genomes, a CRISPR-Cas system can be identified in only five isolates (strains TX1330, Com12, 1,141,733, PC4.1, and 1,231,408) (Palmer & Gilmore, 2010) and unpublished observations). Except for strain 1,231,408, each of these strains belongs to clade B; strain 1,231,408 is a hybrid clade A-clade B strain that appears to have acquired CRISPR-*cas* through recombination with a clade B strain (Palmer, et al., 2012). It has been hypothesized that the absence of CRISPR-Cas in *E. faecium* would allow for the facile acquisition of foreign DNA, such as plasmids that encode genes for antibiotic resistance (Palmer & Gilmore, 2010). However, even commensal *E. faecium* isolates with very few resistances (such as E1039 and E980) lack a CRISPR-Cas system, which indicates that the absence of CRISPR-Cas does not necessarily lead to the rapid

accumulation of antibiotic resistance genes (van Schaik, et al., 2010). Further genome sequencing of antibiotic-resistant and antibiotic-sensitive isolates is required to determine if CRISPR-Cas defense contributes significantly to the evolution of multidrug resistance in *E. faecium*. Nevertheless, based on available genome data, it appears that CRISPR-Cas defense is an ancestral aspect of the clade B *E. faecium* lineage and not of clade A.

Intraspecies Diversity of *E. faecium*

Bacterial genomes differ as the result of drift in the sequences of shared genes (with some mutations becoming fixed by selection), as well as from the acquisition of new genes by horizontal gene transfer. A first insight into variation in gene content of different *E. faecium* isolates was provided by a study in which comparative genomic hybridization was used to characterize the accessory gene pool of *E. faecium* (Leavis H. L., et al., 2007). This led to the identification of 175 genes found more frequently in clinical isolates than in non-clinical isolates. This approach identified the transposon IS*16* as most specifically enriched in clinical isolates (Leavis H. L., et al., 2007). This study also showed that, based on gene content, hospital isolates were more similar to each other than to isolates from other sources, such as healthy humans and animals.

Clinical *E. faecium* isolates are predominantly found within clade A1, and these appear to have emerged from a background of animal isolates, while human commensal isolates are more distantly related (Lebreton, et al., 2013). Genome sequence analysis of two ampicillin-resistant strains (E4452 and E4453) that were isolated from dogs (de Regt, et al., 2012) appear to corroborate the hypothesis that strains from lineage-78 in clade A1 have evolved from animal isolates, since strains E4452 (ST266) and E4453 (ST192) both belong to lineage-78, as defined in the previously described BAPS analysis of *E. faecium* MLST data (159). Phylogenomic analyses underscore the relatively close relationship of E4453 with strains that belong to either lineage-78 or lineage-18 (de Regt, et al., 2012). Both these canine *E. faecium* strains carry IS*16*, which was previously shown to be highly enriched among clinical isolates; but lack virulence genes, such as *hyl* or *esp*, which are commonly found in clinical isolates. Strain E4453, but not strain E4452, carries a genomic island that is putatively involved in the breakdown of complex sugars (Heikens, van Schaik, Leavis, Bonten, & Willems, 2008). However, the two canine *E. faecium* strains share more genes with each other than with the clinical strains E1162 and U0317, which indicates that niche adaptation is occurring on the level of the accessory genome of these isolates. For instance, both strains carry an element that is putatively involved in the breakdown and metabolism of xylopolysaccharides, which is absent from all other isolates for which the genome sequences are publicly available, and appears to be relatively rare in clinical isolates (de Regt, et al., 2012). A PCR-based analysis of the presence of this element in ampicillin-resistant strains isolated from the community showed that this element is relatively common among ampicillin-resistant *E. faecium* strains originating from dogs (18 positives from 25 tested strains) and cats (8 positives from 11 tested strains). This element was also found in two of only three ampicillin-resistant strains that were isolated from a total of 40 healthy volunteers who were given a week-long course of amoxicillin. All ampicillin-resistant isolates from the community were typed with MLST and found to be related to ampicillin-resistant canine *E. faecium* strains. The low number of ampicillin-resistant enterococci that were isolated from healthy humans underscores the relatively low colonization rate of this group by ampicillin-resistant enterococci, and indicates that ampicillin-resistant *E. faecium* are more frequently present in pets than in healthy humans. The finding that ampicillin-resistant isolates from pets are phylogenetically related to ampicillin-resistant clinical isolates of lineage-78 may indicate that ampicillin-resistant isolates have first emerged in pets or other animals and have acquired additional elements (such as virulence determinants, like the *esp* gene and transposons that confer vancomycin resistance) during their adaptation to the lifestyle of nosocomial pathogens. Further genome sequencing of ampicillin-resistant and ampicillin-sensitive strains from non-clinical and non-human reservoirs is required to clarify the evolutionary pathways that have led to the emergence of multidrug-resistant *E. faecium* strains that can successfully colonize and infect hospitalized patients.

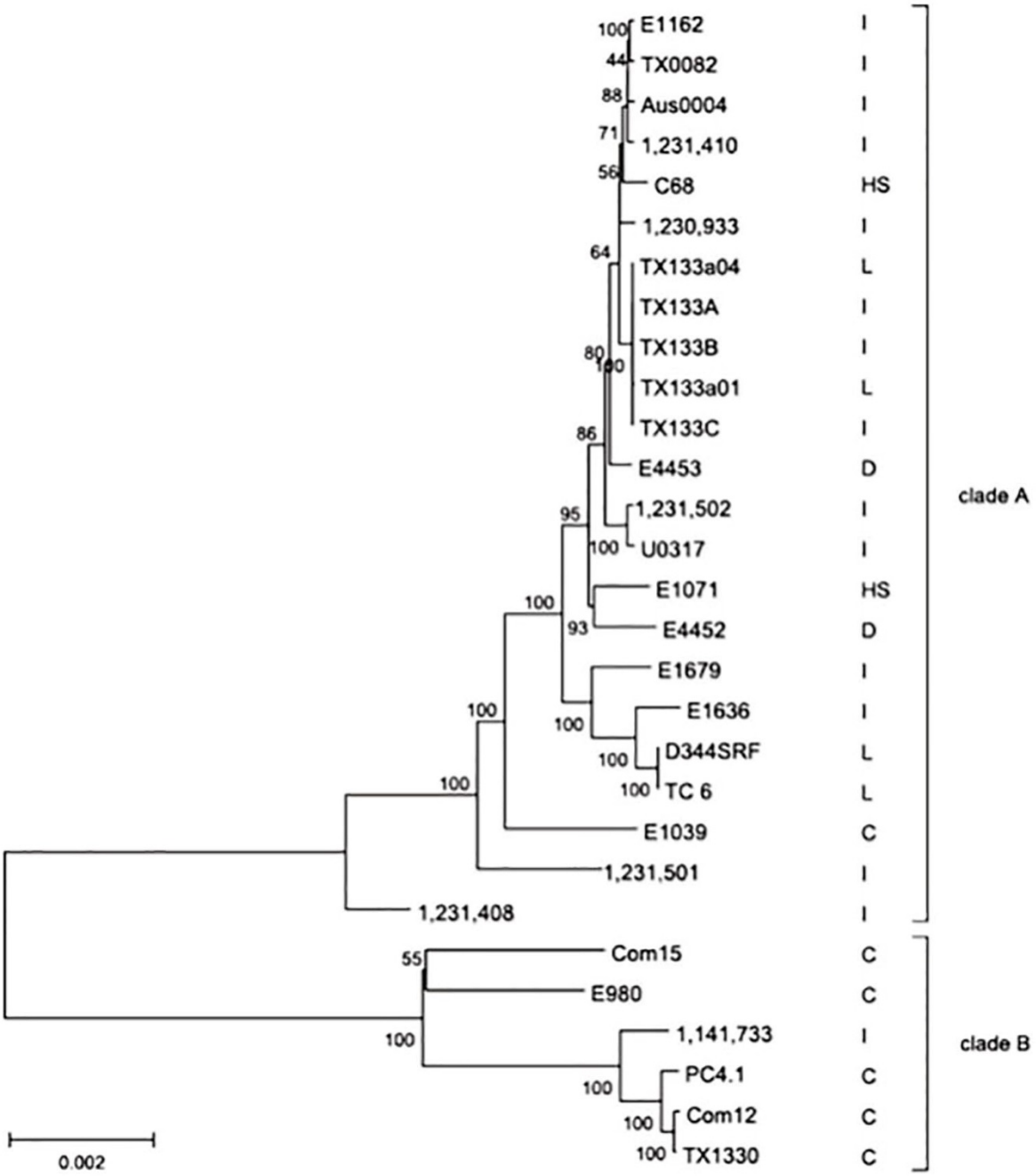


Figure 3. Phylogenomic analysis of *E. faecium*. Minimum evolution tree based on the concatenated alignments of 274 orthologous proteins conserved in draft genome sequences of 30 *E. faecium* isolates. Bootstrap values are indicated and are based on 1000 permutations. The two *E. faecium* clades termed A and B by Palmer *et al.* (110) are indicated. The origins of the strains are also provided (I: isolated from site of infection; HS: hospital surveillance isolate; C: human commensal isolate; D: isolated from dog feces; L: strain was experimentally generated in the laboratory; see Table 1 for further details).

To further quantify intraspecies genomic diversity and to estimate the size of the total gene pool that is accessible to a single species, pan-genome analyses are required (reviewed in (Tettelin, Riley, Cattuto, & Medini, 2008)). In these analyses, pairwise comparisons are made between all possible combinations of genome sequences of a single species, leading to estimates of the total gene pool that is accessible to the organism, the size of the core

genome (namely, the number of genes that are conserved in all strains of the same species), and the number of new genes that might be expected to be discovered when additional genomes are sequenced. By this approach and by using 73 *E. faecium* genomes, the size of the *E. faecium* core genome is estimated to consist of 1597 coding sequences (Lebreton, et al., unpublished data). The total number of coding sequences in an *E. faecium* genome can range from 2272 (for strain EnGen0014) to 3318 (for strain TX0133a01), which indicates that between 29% and 52% of the DNA in any *E. faecium* isolate is non-core and is thus variably present between strains. For strain Aus0004, 38% of the genome is accessory (Lam, et al., 2012). These analyses have also found that the *E. faecium* pan-genome is comparatively open, which signifies that *E. faecium* can efficiently acquire and incorporate novel DNA into the collective gene pool of the species. This indicates that *E. faecium* can rapidly acquire genes that contribute to fitness, which enables it to inhabit a wide variety of environmental niches.

Diversity in the *E. faecium* Accessory Genome: Plasmids, Phage and Genomic Islands

Genome sequencing has shown that plasmids, prophages, and genomic islands are important drivers of diversity in the accessory genome of *E. faecium*.

Plasmids

As in *E. faecalis*, most *E. faecium* isolates carry plasmids that significantly contribute to the genome. Plasmids in *E. faecium* are mostly studied because of their conveyance of antibiotic resistance genes (Palmer, Kos, & Gilmore, 2010), but plasmids can also confer traits that contribute to gastrointestinal colonization (Arias C. A., Panesso, Singh, Rice, & Murray, 2009; Rice, et al., 2009) and the metabolism of carbohydrates (Zhang, Vrijenhoek, Bonten, Willems, & van Schaik, 2011). A broad discussion of enterococcal plasmids is presented in Extrachromosomal and Mobile Elements in Enteroocci.

Megaplasmids occur in some *E. faecium* isolates, including those that encode the hyl_{Efm} gene. hyl_{Efm} is more frequently present in clinical than in non-clinical *E. faecium* isolates, and was initially erroneously classified as a hyaluronidase (Rice, et al., 2003). However, hyaluronidase activity has not been detected in strains that carry this gene (Laverde Gomez, et al., 2011). More recent sequence analyses have shown that the protein encoded by hyl_{Efm} is likely to encode a family 84 glycosyl hydrolase, which may function as a β-N-acetylglucosaminidase (Freitas, et al., 2010; Sheldon, et al., 2006). Transfer of a plasmid that encodes hyl_{Efm} increases the virulence of *E. faecium* in a mouse peritonitis model (Arias C. A., Panesso, Singh, Rice, & Murray, 2009), as well as in a mouse model of gastrointestinal colonization (Rice, et al., 2009). However, specific deletion of the hyl_{Efm} gene did not lead to lower virulence in a mouse peritonitis model (Rice, et al., 2009), which indicates that other genes encoded by the plasmid contribute to virulence. Plasmids that encode hyl_{Efm} range in size from 150 to 375 kb (Arias C. A., Panesso, Singh, Rice, & Murray, 2009; Freitas, et al., 2010; Rice, et al., 2009). Sequencing of the 280 kb plasmid pLG1 revealed that hyl_{Efm} is encoded within a conserved 18-kb gene cluster that is flanked by incomplete inverted copies of the IS*1476* putative transposase gene. Additional pLG1 genes have functions in plasmid maintenance, replication, and conjugation. Genes involved in carbon uptake and metabolism, as well as those that confer resistance to heavy metals and antibiotics, were also identified (Laverde Gomez, et al., 2011). The *pilA*/*fms21-fms20* gene cluster that encodes a pilus structure in *E. faecium* is also present on pLG1 (Kim, et al., 2010; Laverde Gomez, et al., 2011). Large multifunctional plasmids in *E. faecium* appear to play an important role in the adaptation of *E. faecium* to new environments, as in a single event (such as the acquisition of a plasmid), many genes are acquired with diverse roles in gastrointestinal colonization, virulence, and antibiotic survival. Sequencing of additional antibiotic resistance plasmids has shed new light on plasmid maintenance systems and the genetic context of the vancomycin-resistance transposon Tn*1546* (Li, Alvarez, Harper, & Wang, 2011; Sletvold, et al., 2008; Sletvold, et al., 2007; Sletvold, et al., 2010).

Bacteriophages

Enterococcal bacteriophages are comprehensively discussed in Enterococcal bacteriophages and genome defense. In the genomes of *E. faecium* strains, integrated prophages appear to be important drivers of diversity. Genes of phage origin make up 2.3% to 5.1% of seven sequenced *E. faecium* genomes, and in most of those strains, integrated prophages can be activated by the addition of the DNA cross-linking agent mitomycin, which leads to the release of phage particles that can be visualized by electron microscopy (van Schaik, et al., 2010). In this study, all bacteriophages activated by mitomycin had the typical morphology of *Siphoviridae*, with an isometric head and a long non-contractile segmented tail. This family of bacteriophages is common in lactic acid bacteria, and a study aimed at identifying temperate bacteriophages in *E. faecalis* also exclusively found *Siphoviridae* (Yasmin, et al., 2010). However, lysogenic phages from other families have also been detected in *E. faecium* and other enterococci (Mazaheri Nezhad Fard, Barton, & Heuzenroeder, 2010). In strain Aus0004, three distinct prophages can be identified, of which one is present as a 50-kb inverted repeat (Lam, et al., 2012). These two identical prophages appear to have facilitated a 683 kb chromosomal inversion, which leads to a shift in the predicted replication terminus, which is normally positioned 180° from the origin of replication, but is at 115° in Aus0004 instead.

The presence of integrated prophages poses a potential threat to survival of the *E. faecium* cell, as the activation of the phage may lead to cell death. Even the presence of phage DNA may come at a cost, in terms of the increased metabolic burden required for the replication of phage DNA during cell division. For these reasons, bacteria have evolved mechanisms by which they can withstand attack by bacteriophages, such as the CRISPR-Cas system previously discussed. Alternately, prophages can be retained in the population if the phage confers a fitness benefit to the host, in a process termed lysogenic conversion (Brüssow, Canchaya, & Hardt, 2004). There are currently no data available that would indicate that phages carry genes that contribute significantly to the fitness of *E. faecium*, but systematic studies aimed at characterizing phage genes have not been performed. Phages may also be involved in the spread of antibiotic resistance genes among enterococci, as they can transfer antibiotic resistance genes between different enterococcal species, which expands the mobility of resistance genes in this genus (Mazaheri Nezhad Fard, Barton, & Heuzenroeder, 2011).

Genomic islands

Similar to what was found in *E. faecalis* (Schwarz, Perreten, & Teuber, 2001), genome sequencing of *E. faecium* led to the identification of two genomic islands that are enriched in clinical *E. faecium* isolates. The most thoroughly studied genomic island in *E. faecium* carries the *esp* gene, which encodes a large (~200 kDa) peptidoglycan-anchored surface protein that contributes to biofilm formation (Heikens, Bonten, & Willems, 2007). In animal models, it has been shown to contribute to urinary tract infection (Leendertse, et al., 2009) and endocarditis (Heikens, et al., 2011), but not to gastrointestinal colonization in the model tested (Heikens, et al., 2009). In *E. faecium*, the *esp* gene is part of a large genomic island that ranges from 64 to 125 kbp in size (Figure 4), and which has been sequenced to completion in the strains E1162, E1679, U0317, and Aus0004 (Lam, et al., 2012; van Schaik, et al., 2010). As in *E. faecalis* (124), the *esp* genomic island in *E. faecium* carries all the hallmarks of a prototypical genomic island, as its %G+C content is 1.1% to 2.3% lower than the entire genome of this strain, and it is flanked by two 54-bp direct repeats. In all strains, the *esp* genomic island has been integrated between a small gene that encodes a hypothetical protein and the *rpsI* gene, which encodes a small ribosomal protein. In strain U0317, a genomic rearrangement event appears to have occurred after the initial integration of the *esp* genomic island, which resulted in the nearly complete island reintegrating 5′ of the *tuf* gene (van Schaik, et al., 2010). The locus in which the *esp* island integrates appears to be a hotspot for the integration of foreign elements; in several strains that lack the *esp* island, smaller elements of possible phage or plasmid origin have integrated at this position.

In *E. faecium* strains E1162 and UW3308, the *esp* island can be mobilized and horizontally transferred to *esp*$^-$ acceptor strains by conjugation (Oancea, Klare, Witte, & Werner, 2004; van Schaik, et al., 2010). The observation that the *esp* island is mobilizable and self-transmissible has led to the renaming of this island to the integrative conjugative element (ICE) ICE*Efm1* (Top, et al., 2011). Sequencing of ICE*Efm1* in different *E. faecium* strains has revealed that this element encodes many functions that are potentially beneficial for the bacterial cell during infection of the human host. Therefore, the conjugative spread of ICE*Efm1* may significantly contribute to the recent emergence of *E. faecium* as a nosocomial pathogen of major importance.

The overall structure of ICE*Efm1* is similar in strains E1162, E1679, U0317 and Aus0004, but several insertions and rearrangements have occurred. The 5′ end of ICE*Efm1* consists of an element that resembles the conjugation module of the widely spread conjugative transposon Tn*916*. At the extreme 5′ end of ICE*Efm1* in strains E1162, E1679, and U0317, a predicted peptidoglycan-anchored surface protein is encoded. This protein is 91% identical on the amino-acid level to the collagen adhesin EcbA (Hendrickx, et al., 2009). This gene is also present in strain Aus0004, but an additional 23.8 kb fragment is present between the gene that encodes the *ecbA* homolog and the 54 bp sequence that forms the outer boundary of ICE*Efm1*. The central part of ICE*Efm1* consists of the *esp* gene and several other genes that have predicted roles in regulation of transcription, transport, and other functions. There is considerable heterogeneity in the *esp* gene among different strains, particularly in the 3′ end of the gene where repeats are encoded (Leavis, et al., 2004). In the 3′ end of ICE*Efm1*, a 10-kb region can be identified that is essentially identical (98% nucleotide sequence identity) to a region of the *E. faecalis* PAI, which also encodes *esp* (Shankar, Baghdayan, & Gilmore, 2002). Several of the genes in this region have a predicted role in the uptake of manganese, which is an important element for virulence in both Gram-negative and Gram-positive bacteria (Papp-Wallace & Maguire, 2006). On the extreme 3′ end of ICE*Efm1*, an integrase gene is present, which is indispensable for excision of ICE*Efm1* (Top, et al., 2011). Multiple transposases are scattered throughout ICE*Efm1*, and the presence of other elements (such as a group II intron in strain E1162) can vary from strain to strain. The most striking difference among the currently sequenced ICE*Efm1* elements is the presence of a Tn*1549*-like transposon (Garnier, Taourit, Glaser, Courvalin, & Galimand, 2000) in the ICE*Efm1* of Aus0004. This transposon carries the *vanB* gene cluster, which confers resistance to vancomycin. Another similar integration in ICE*Efm1* has occurred in strain E1679, where a gene cluster putatively encoding a pathway for the metabolism of inositol has been inserted. This element appears to be functional, as strain E1679 can grow on inositol as carbon source, while *E. faecium* strains that lack these genes cannot (van Schaik *et al.*, unpublished data). Genome sequencing of strain 1,231,502 also revealed the presence of this inositol utilization gene cluster as part of ICE*Efm1* in this strain.

A second genomic island was first identified in the draft TX16 genome, due to its aberrant dinucleotide frequency and %G+C content. The island is an 8.4 kb element that is more commonly found in clinical isolates than in strains isolated from non-clinical settings (Heikens, van Schaik, Leavis, Bonten, & Willems, 2008). This genomic island is flanked by inverted and direct repeats, contains an integrase, has an anomalous %G+C; and in addition to DO, is present in the genomes of the clinical isolates 1,230,933, 1,231,502, 1,231,410, Aus0004, C68, E1162, E1679, TX0133, TX0082, and U0317, all of which are clade A strains. The only non-clinical isolate in which this element is also found in the genome sequence is the canine *E. faecium* strain E4453. The functional role of this element remains to be experimentally determined, but sequence analysis points towards its role in the uptake and metabolism of complex carbohydrate substrates.

Genomics of Other Enterococci

Although most sequencing efforts have has focused on genomes of *faecalis* and *faecium* species, data are increasingly available for other enterococcal species as well. These genomes, in addition to being interesting in their own right, are important for understanding the diversity of the genus *Enterococcus,* and to understand the evolution of the species within this genus. Annotated and unannotated draft genome sequences are available for

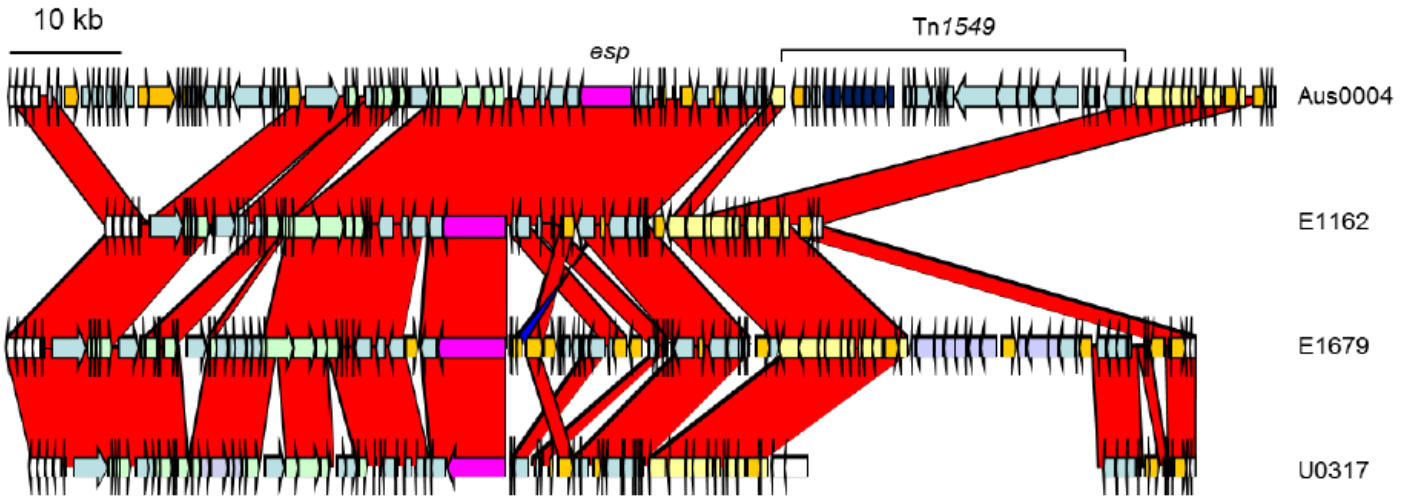


Figure 4. ICE*Efm1* in *E. faecium* strains Aus0004, E1162, E1679 and U0317. The ICE*Efm1* elements of strains Aus0004, E1162, E1679, and U0317 were aligned using the Artemis Comparison Tool (15). Arrows indicate coding sequences (CDS) and the direction of transcription. Red-colored bands indicate matching regions between ICE*Efm1* of different strains. A small blue band represents an inversion of a transposase fragment between E. faecium E1162 and E1679. The yellow arrows indicate a region that is identical to a fragment of the E. faecalis MMH594 PAI (126). The green arrows represent CDS that are homologous to the conjugation module of Tn*916*. The purple arrows in E1679 represent CDS putatively encoding a functional pathway for inositol uptake and metabolism. Orange arrows represent transposases and integrases. Dark blue arrows represent the genes of the *vanB* cluster on Tn*1549*. The *esp* gene is colored pink. White arrows indicate flanking genes that are not part of ICE*Efm1*.

strains of *E. casseliflavus, E. gallinarum*, *E. italicus*, *E. saccharolyticus,* and *E. mundtii* (Magni, et al., 2012; Palmer, et al., 2010; Palmer, et al., 2012) and unpublished data in GenBank).

E. casseliflavus and *E. gallinarum*, historically thought to be primarily associated with plants (Mundt & Graham, 1968) and birds (Collins, Jones, Farrow, Kilpper-Balz, & Schleifer, 1984), respectively, have also been associated with life-threatening infections, such as bacteremia and meningitis (Choi, et al., 2004; de Perio, Yarnold, Warren, & Noskin, 2006; Khan & Elshafi, 2011; Prakash, Rao, & Parija, 2005). A hospital outbreak of *E. gallinarum* has been reported (Contreras, et al., 2008). The endogenous low-level vancomycin resistance of *E. casseliflavus* and *E. gallinarum*, combined with the intrinsic resistance of all enterococci to antibiotics such as cephalosporins (Tannock & Cook, 2002), may contribute to a rise in opportunistic infections caused by these species. Three of the sequenced *E. casseliflavus* genomes (EC10, EC20, and EC30) and the *E. gallinarum* EG2 genome derive from clinical isolates obtained from patients in the United States (Palmer, et al., 2012). *E. casseliflavus* ATCC 12755 (Genbank accession AEWT00000000) is a reference genome for the Human Microbiome Project. Of the remaining sequenced species, to our knowledge, only *E. mundtii* has been associated with human infection, and reports of these types of infection are few (de Perio, Yarnold, Warren, & Noskin, 2006; Prakash, Rao, & Parija, 2005). *E. mundtii* ERL1656, for which a genome sequence is available, was isolated from cow′s milk and is a bacteriocin-producing strain that is thought to have probiotic properties (Magni, et al., 2012). *E. saccharolyticus* 30_1 (Genbank accession ADLY00000000) was isolated from inflamed ileum tissue taken from a patient with Crohn's disease (http://www.broadinstitute.org/annotation/genome/Enterococcus_group), and is a reference genome for the Human Microbiome Project. Finally, *E. italicus* DSM 15952 (Genbank accession AEPV00000000), which is also a reference genome for the Human Microbiome Project, was isolated from an Italian cheese and is the type strain for the *italicus* species (Fortina, Ricci, Mora, & Manachini, 2004).

Two phenotypes typically absent in *E. faecalis* and *E. faecium* have been of particular interest for clinical discrimination of enterococcal species. *E. casseliflavus* and *E. gallinarum* are distinguished from *E. faecalis* and *E. faecium* by their motility, while *E. casseliflavus* and *E. mundtii* are further distinguished by their production of a

yellow pigment (Facklam, Carvalho, & Teixeira, 2002). However, variability in these phenotypes occurs, which complicates species designations when phenotypic characteristics are used (Devriese, Pot, & Collins, 1993; Tannock & Cook, 2002; Vincent, Knight, Green, Sahm, & Shlaes, 1991). The genetic basis for these traits, as well as whether they are truly species-specific, are of interest. Genomic analysis has revealed genetic bases for motility and pigment production in *E. casseliflavus* and *E. gallinarum* (Palmer, et al., 2012), among other traits.

Flagellar motility

A set of 48 predicted flagellar biosynthesis and chemotaxis genes are co-localized on one contig of the draft *E. casseliflavus* EC10 genome (contig 22; ORFs ECAG_01613-ECAG_01660) (Palmer, et al., 2012). Orthologs of all or most of these genes are present in *E. casseliflavus* EC20, EC30, and *E. gallinarum* EG2 (Palmer, et al., 2012). The *E. casseliflavus* ATCC 12755 genome possesses a highly conserved homologue of the putative flagellin protein (ECAG_01617) encoded by *E. casseliflavus* EC10 (Palmer, unpublished), which suggests that a similar gene cluster is present in that strain. Most top BLASTP hits to the *E. casseliflavus* EC10 motility cluster proteins are from *Lactobacillus ruminis* strains, and a few are from *Carnobacterium* sp. strains (Palmer, et al., 2012); both genera are within families that are closely related to the *Enterococcaceae*. The distribution of flagellar motility genes among available enterococcal genomes suggest either that motility was once a core property of the *Enterococcus* genus, and was subsequently lost by the progenitors of the *faecalis* and *faecium* lineages; or that flagellar motility was laterally acquired by the progenitor of the *casseliflavus/gallinarum* lineage from closely related bacteria. Because motility is a rare phenotype in the *Enterococcus* genus (Facklam, Carvalho, & Teixeira, 2002), the most parsimonious explanation would appear to be lateral transfer, although we note that motility is also present in some species of *Vagococcus*, another genus within the *Enterococcaceae* family (Lawson, Foster, Falsen, Ohlén, & Collins, 1999; Shewmaker, et al., 2004; Teixeira, et al., 1997).

Surprisingly, a homolog of *E. casseliflavus* EC10 flagellin is present in the draft *E. saccharolyticus* 30_1 genome (HMPREF9478_02662; 94% amino acid identity to ECAG_01617), and is flanked by additional genes with likely roles in flagellar biosynthesis and chemotaxis (Palmer, unpublished). The presence of a flagellin homologue in this bacterium is surprising, because motility has not been noted as a characteristic of the *saccharolyticus* species (Facklam, Carvalho, & Teixeira, 2002). Analysis of the 16S rRNA sequence of the *E. saccharolyticus* 30_1 genome (HMPREF9478_03015) reveals that it is 100% identical to the 16S rRNA sequence of *E. gallinarum* EG2 and 99% identical to several partial *E. gallinarum* 16S rRNA sequences in GenBank. Further, preliminary analyses indicate that the housekeeping genes of *E. saccharolyticus* 30_1 are >99% identical to those of *E. gallinarum*, which suggests that strain 30_1 may have been misclassified (Palmer and van Schaik, unpublished). Whether the *E. saccharolyticus* 30_1 strain is a motile variant of the *saccharolyticus* species or is in fact an *E. gallinarum* strain should be investigated further.

Carotenoid production

Pigment production is a property of only a few described *Enterococcus* species, including *E. casseliflavus* and *E. mundtii* (Facklam, Carvalho, & Teixeira, 2002), although pigment production appears to be an unreliable phenotypic marker for species designations (Tannock & Cook, 2002). The yellow pigment elaborated by *E. casseliflavus* is a carotenoid pigment that likely protects it, as well as other species that produce the pigment, from photo-oxidation (Maraccini, Ferguson, & Boehm, 2012; Taylor, Ikawa, & Chesbro, 1971). *Staphylococcus aureus* also produces a carotenoid pigment, staphyloxanthin, which has been more extensively studied. Staphyloxanthin biosynthesis is directed by the *crtOPQMN* operon, and *crtOPQMN* expression *in trans* is sufficient to confer staphyloxanthin production on the natively non-pigmented *Staphylococcus carnosus* (Pelz, et al., 2005). Biochemical characterization of the staphyloxanthin biosynthesis pathway expressed in *S. carnosus* revealed that *crtMN* expression alone is sufficient for production of the yellow pigment and pathway intermediate 4,4-diaponeurosporene (Pelz, et al., 2005). The additional expression of *crtOPQ* is required to complete the biosynthetic pathway to staphyloxanthin (Pelz, et al., 2005).

Homologues of all five *crt* genes are present in the *E. casseliflavus* EC10, EC20, and EC30 genomes, while only *crtM* and *crtN* are present in *E. gallinarum* EG2 (Palmer, et al., 2012). For the enterococci, *crt* gene organization differs from that of *S. aureus*. In *E. gallinarum* and *E. casseliflavus*, *crtNM* are putatively co-transcribed, while *crtPQO* are putatively transcribed and are encoded elsewhere on the *E. casseliflavus* genome. The significance of this is unknown, but is suggestive of two "modules" of pigment biosynthesis genes, one composed of *crtNM* and one composed of *crtPQO*—although note that CrtM and CrtN activities should be required for the later steps to be functional. *Lactobacillus plantarum* also possesses the *crtNM* module, and a collection of 18 *L. plantarum* strains that possessed *crtNM* were found to be variable for production of the yellow pigment 4,4-diaponeurosporene, with some strains forming yellow cell pellets with high amounts of pigment detectable by chromatographic analysis, and others forming white cell pellets with low or trace amounts of pigment detected (Garrido-Fernández, Maldonado-Barragán, Caballero-Guerrero, Hornero-Méndez, & Ruiz-Barba, 2010). For *L. plantarum*, 4,4-diaponeurosporene production varies with the growth medium used (Garrido-Fernández, Maldonado-Barragán, Caballero-Guerrero, Hornero-Méndez, & Ruiz-Barba, 2010). The presence of *crtNM* in *E. gallinarum* EG2 is surprising, as it suggests that this strain could also produce the yellow pigment 4,4-diaponeurosporene, which has not been observed in routine laboratory culture. Whether these genes are functional and expressed, and perhaps induced under conditions so far not encountered in laboratory culture, remains to be determined. Regardless, it appears that yellow pigment production by certain enterococcal strains is due to the growth medium-dependent production of one or both of two structurally distinct carotenoids, which likely explains the inconsistencies that have been observed in this phenotype.

BLAST analysis reveals the distribution of the *crt* genes among other sequenced enterococci and closely related bacteria (Palmer, unpublished). Consistent with a conserved pigment biosynthesis pathway in *E. casseliflavus*, putative *crtNM* and *crtPQO* genes are present in *E. casseliflavus* ATCC12755. Nucleotide sequences with high identity to *crtN* and *crtPQO* are also present in the unannotated *E. mundtii* ERL1656 genome, although some sequences are partial and operon structures are unclear. This suggests that a similar pathway is present in *E. mundtii*, although further bioinformatic analysis and perhaps additional sequencing will be required for confirmation. Interestingly, *crt* genes are also present in *Carnobacterium* sp. strains 17-4 and AT7, arranged in a cluster as *crtNMP-ywdH-crtQO* (*ywdH* encodes a putative aldehyde dehydrogenase). This further highlights the similarities between *Carnobacterium* sp. and motile, pigmented enterococci.

Differences Between Enterococcal Species Revealed by Genomics

Beyond the phylogenetic classification, what does it mean to be an *E. faecalis*, an *E. faecium*, an *E. casseliflavus*, or another enterococcal species? Are there certain traits that define these species relative to one another? A great deal of work in the clinical microbiology of enterococci has focused on biochemical or phenotypic tests that are capable of discriminating one species from another (Facklam, Carvalho, & Teixeira, 2002). However, phenotypic variability within a species and atypical isolates complicate those analyses. Comparative genomics, through the interrogation of tens or hundreds of genomes, offers a new way to search for traits that are characteristic of a species.

Species-specific metabolism

Comparative analysis of metabolic pathways encoded by enterococcal genomes has been limited, although some insights in this area have emerged. A COG (clusters of orthologous genes)-based analysis of seven *E. faecium* and six *E. faecalis* genomes identified 70 COGs specific to *E. faecium* and 140 COGs specific to *E. faecalis* (van Schaik, et al., 2010). Several of these belong to pathways known to differ between the two species, including that of arabinose catabolism (present in *faecium*, but not *faecalis*) and of heme-dependent cytochrome biosynthesis (present in *faecalis*, but not *faecium*). Other species-specific differences identified by COG analysis were the presence of putative pectin and lignocellulose catabolism genes in *faecium* but not *faecalis*, and ethanolamine utilization genes in *faecalis* but not *faecium*. Ethanolamine catabolism genes were confirmed to be specific to the

faecalis species in a separate study evaluating 30 *faecalis*, *faecium*, *casseliflavus* and *gallinarum* genomes (Palmer, et al., 2012). It is interesting that genes for the catabolism of complex plant polysaccharides and their degradation products are specific to *E. faecium*, while genes for catabolism of the membrane phospholipid head group ethanolamine are specific to *E. faecalis*. This suggests that these two species can occupy fundamentally different nutrient niches in the human intestine.

A combination of Biolog carbon utilization analysis and comparative genomics identified additional metabolic reactions and pathways that discriminate different enterococcal species (Palmer, et al., 2012). Genes that putatively direct synthesis of the bacterial second messenger signal cyclic-di-GMP, and genes for an acetoin dehydrogenase were identified in *E. casseliflavus,* and not *E. faecalis*, *E. faecium*, or *E. gallinarum*. Inulin catabolism was additionally found to be specific to *E. casseliflavus* using Biolog analysis; however, no species-specific genes conferring this phenotype have been identified. As mentioned above, ethanolamine catabolism genes were found to be specific to *E. faecalis*, as are the *bkdDABC* genes encoding a branched-chain α-keto acid dehydrogenase complex, and a putative formate dehydrogenase gene (*fdhA*). Consistent with the distribution of the *bkdDABC* genes, catabolism of α-ketovalerate in Biolog assays is specific to *E. faecalis*. Finally, a gene that encodes a putative glutaminase was found to be *E. faecium*-specific. Each of these genes and metabolic reactions present opportunities for the development of novel biochemical assays or genetic screens capable of rapidly discriminating enterococcal species. Biolog analysis has also provided information on *E. faecium* intraspecies metabolic diversity by aiding in the identification of genetic determinants that confer raffinose catabolism (Zhang, Vrijenhoek, Bonten, Willems, & van Schaik, 2011).

Extracellular polysaccharide biosynthesis

Extracellular polymer biosynthesis has been most studied for *E. faecalis*, which in addition to lipoteichoic and wall teichoic acids, produces a rhamnopolysaccharide called the enterococcal polysaccharide antigen (Epa), and a capsule (Hancock & Gilmore, 2002). Biosynthesis of the Epa polymer is directed by the *epa* locus, a cluster of genes that encompasses *epaA* to *epaR,* which encode functions that are likely to be important for synthesis and export of the rhamnopolysaccharide polymer (Teng, Singh, Bourgogne, Zeng, & Murray, 2009; Xu, Murray, & Weinstock, 1998). Based on PCR-based screening and dot blot hybridization studies, it was reported that the *epa* locus is core to the *faecalis* species (Teng, Singh, Bourgogne, Zeng, & Murray, 2009). Conversely, the presence of the *cps* genes that direct capsule biosynthesis varies among *E. faecalis* strains (McBride, Fischetti, Leblanc, Moellering, Jr., & Gilmore, 2007). Little is known about extracellular polymer synthesis by enterococcal species other than *faecalis*. A comparative analysis of 30 *E. faecalis*, *E. faecium*, *E. casseliflavus*, and *E. gallinarum* genomes confirmed that the *epa* locus is core to the *faecalis* species, and also identified a similar *epa* locus as being core to *E. faecium* (Palmer, et al., 2012). The presence of *epa* loci in both *E. faecalis* and *E. faecium*, but not in *E. casseliflavus* or *E. gallinarum*, suggests that the Epa polymer might contribute in some significant way to the colonization of mammalian hosts, with which these two species are most closely associated. Surprisingly, the *E. faecalis cps* locus was not detected outside of that species (Palmer, et al., 2012). Instead, a second putative capsule biosynthesis mechanism was identified, being present in the genomes of *E. faecium*, *E. casseliflavus*, and *E. gallinarum*, as well as the *E. saccharolyticus* 30_1 and *E. italicus* DSM 15952 genomes (Palmer, et al., 2012). This putative capsule biosynthesis mechanism appears to be similar to that of *Streptococcus pneumoniae*, and consists of a core phosphoregulatory system encoded by *wzg-wzd-wze-wzh* and a variable set of genes that occur downstream of *wzg-wzd-wze-wzh,* which appear to direct biosynthesis of a polysaccharide polymer. Other than *E. faecalis*, *wzg-wzd-wze-wzh* is core to all enterococcal species for which a genome sequence is available, which highlights a significant evolutionary deviation of *faecalis* from other enterococci.

Additional Genomic Applications: Metagenomics and Genome Resequencing

To this point, we have discussed comparative genomic analyses of closed and draft enterococcal genomes, each obtained by *de novo* sequencing of DNA obtained from an isolate grown in pure culture. Here, we discuss two additional genomic techniques that have been applied to *E. faecalis*: metagenomics and genome resequencing.

Metagenomic assembly of *E. faecalis* UC1ENC from a complex microbial community

Enterococci are abundant in the feces of formula-fed infants and are generally more prevalent in infant intestinal communities compared to adult communities, which suggests that their presence contributes to the conditioning of the infant immune system (Tannock & Cook, 2002). A molecular phylogenetic analysis of the fecal community of a premature human infant revealed diet-dependent fluctuations in community structure, with enterococcal 16S rRNA sequences being most abundant after a dietary switch from enteral feedings with maternal breast milk to commercial infant formula (Morowitz, et al., 2011). A complete enterococcal genome, designated UC1ENC, was assembled from metagenomic sequence generated from DNA extracted from fecal samples obtained from this infant. The UC1ENC 16S rRNA gene sequence is 100% identical to that of multiple *E. faecalis* strains, which indicates that the UC1ENC genome originated from an *E. faecalis* strain abundant in the infant gut. *In silico* MLST analysis of UC1ENC assigned the genome to ST179, an ST previously recovered from hospitalized patients in Spain and the Netherlands (http://efaecalis.mlst.net), and non-hospitalized individuals in Spain (Kuch, et al., 2012). Two putative enterococcal plasmids (plasmid 1 and plasmid 2) and two putative enterococcal phages (phage 1 and phage 2) were also assembled from the infant metagenomic data (Morowitz, et al., 2011).

Comparative analysis of the UC1ENC genome with the V583 genome revealed substantial relatedness in core functions as well as shared mobile genetic elements. Predicted protein sequences encoded by the two genomes possess an average amino acid identity of 98.7% (Morowitz, et al., 2011). UC1ENC shares two prophage (prophages 2 and 4) with V583, as well as portions of the PAI (Morowitz, et al., 2011) that lack the module F and the 3′ region of module D, as defined by previous studies of the PAI (McBride, et al., 2009). BLAST analysis of predicted enterococcal plasmid and phage sequences additionally assembled from infant fecal metagenomic data reveal substantial identity to previously characterized *E. faecalis* mobile elements (Palmer, unpublished analysis). Plasmid 1 (68.7 kb) is similar to the pheromone-responsive plasmid pMG2200 (Zheng, Tomita, Inoue, & Ike, 2009), although it lacks the Tn*1549 vanB*-type vancomycin resistance transposon that is present in pMG2200. Plasmid 2 (8.3 kb) appears to be derived from the conjugative plasmid pRE25 (124) and possesses a similar zeta-epsilon-delta toxin-antitoxin system. Phage 1 (8.3 kb) is similar to chromosomal sequence of the infant fecal E. faecalis strain 62, while phage 2 (28.9 kb) is similar to the pseudotemperate linear phage present in E. faecalis 62, EF62ϕ (Brede, Snipen, Ussery, Nederbragt, & Nes, 2011). In situ copy numbers of the putative UC1ENC plasmids and phages were calculated using metagenomic reads from the premature infant fecal analysis, using samples extracted at 16, 18, and 21 days post-birth (Morowitz, et al., 2011). Plasmid 1 copy numbers were calculated to be approximately 2 at each day of sampling, which is consistent with the known copy numbers of pheromone-responsive plasmids (Weaver, Rice, & Churchward, 2002). Copy numbers of plasmid 2 and phage 1 were around 1, perhaps indicating that these elements are integrated into the UC1ENC genome. Copy numbers of phage 2 fluctuated from 5 to 12, which is consistent with this element being a pseudotemperate phage that replicates independently from the chromosome, or perhaps with a virulent phage being liberated from UC1ENC cells. To our knowledge, this study is the first instance in which copy numbers of enterococcal extrachromosomal elements have been calculated for an in vivo environment.

In vitro evolution and resequencing identifies a genetic basis for daptomycin resistance

In prokaryotic genome resequencing, derivatives or populations of a reference strain are sequenced to high coverage, and reads are aligned to the reference genome to identify polymorphisms that occur within. A notable use of this technique was the identification of mutations that occurred over a long-term evolution experiment with *Escherichia coli* (Barrick, et al., 2009). This technique has also proven useful for identifying a genetic basis for enterococcal resistance to the last-line antibiotic daptomycin (Palmer, Daniel, Hardy, Silverman, & Gilmore, 2011). Daptomycin is a lipopeptide antibiotic with a unique mechanism of action that causes cell death in the absence of lysis (Cotroneo, Harris, Perlmutter, Beveridge, & Silverman, 2008). Daptomycin was approved by the US Food and Drug Administration in 2003 for the treatment of certain staphylococcal infections and is used off-label to treat enterococcal infections, such as bacteremia and endocarditis (Cantón, Ruiz-Garbajosa, Chaves, & Johnson, 2010). Enterococcal resistance to daptomycin has emerged in the clinic, but occurs sporadically and does not appear to be conferred by the acquisition of mobile elements, such as plasmids and transposons (Kelesidis, Humphries, Uslan, & Pegues, 2011; Munoz-Price, Lolans, & Quinn, 2005).

In triplicate *in vitro* experiments, *E. faecalis* V583 was passaged through increasing concentrations of daptomycin, which resulted in the emergence of three high-level daptomycin resistant strains (Palmer, Daniel, Hardy, Silverman, & Gilmore, 2011). Illumina resequencing of these strains identified mutations that were putatively responsible for this phenotype, including single nucleotide polymorphisms, deletions, duplications, and the transposition of an IS*256* element. Mutations in one gene, a putative cardiolipin synthase (*cls*; V583 ORF EF0631), were among the first to occur in each of the three evolution experiments. Strikingly, EF0631 mutations were observed in daptomycin-resistant *E. faecalis* clinical isolates, as well as in a homolog of EF0631 in daptomycin-resistant *E. faecium* clinical isolates, but not their susceptible parent strains isolated prior to daptomycin therapy. Expression in OG1RF of EF0631 alleles from daptomycin-resistant strains conferred resistance to OG1RF, which conclusively demonstrated that the EF0631 mutations identified by resequencing analysis conferred resistance. Collectively, *in vitro* evolution experiments and genome resequencing identified mutations that confer resistance to a last line antibiotic, with those mutations also occurring in resistant clinical isolates. This study illustrates the utility of genomic approaches to study rapid adaptation in enterococci.

Not all daptomycin-resistant *E. faecalis* clinical isolates possess EF0631 mutations, which indicates that mutational pathways to resistance exist that are independent from this gene (Palmer, Daniel, Hardy, Silverman, & Gilmore, 2011). *De novo* genome sequencing of the daptomycin-resistant *E. faecalis* human bloodstream infection isolate R712 and its daptomycin-susceptible parent strain S613, isolated from the same patient prior to daptomycin therapy, revealed mutations in the R712 EF0631 *cls* and in two additional genes, *liaF* (EF2913) and *gdpD* (EF1904) (Arias C. A., et al., 2011). *liaF* and *gdpD* mutations were not observed in the resequencing study (Palmer, Daniel, Hardy, Silverman, & Gilmore, 2011). Experimental interrogation of the *liaF* and *gdpD* mutant alleles revealed that the *liaF* mutation alone, or *liaF* and *gdpD* mutations together, were sufficient to confer daptomycin resistance to S613 (Arias C. A., et al., 2011), which demonstrates that a *cls*-independent pathway to resistance exists. In the case of enterococcal daptomycin resistance, genomic approaches have significantly contributed to an understanding of how resistance emerges.

Conclusion and Perspectives

Genome sequencing and comparative genomics has significantly advanced our understanding of enterococcal virulence, biology, and evolution. Insights discussed here include the identification of adaptive elements that likely contribute to the diverse ecology of the enterococci, analyses illuminating the fluidity of the *E. faecalis* and *E. faecium* genomes, and the discovery that the species *E. faecium* is actually composed of distinct phylogenetic clades between which recombination occurs. Despite rapid and significant advances in this area, much remains to be learned from existing genome data.

For example, for currently available *E. faecalis* genomes, most comparative analysis has focused on recombination and horizontal gene transfer, which is precipitated by multiple lines of evidence that indicate that horizontal gene transfer is rampant in the species. However, a detailed analysis of the mobile elements that populate these genomes remains to be performed. What novel traits do these plasmids, phage, and genomic islands encode? What impact does the insertion of mobile elements have on the genome; namely, are core functions disrupted? What are the core functions of the *faecalis* genome, and what roles might these genes play in infection, niche colonization, and persistence? Beyond *faecalis*, what functions are core to the other enterococcal species for which multiple genome sequences are available—*faecium* and *casseliflavus*—and what do they tell us about what it means to be classified as one of these species? Further, what can comparative genomics tell us about phylogenetic relationships and gene exchange among genera closely related to *Enterococcus*, such as *Melissococcus* and *Tetragenococcus*, each of which already have sequenced representatives available in GenBank? These are a handful of questions that are meant to illustrate that we have only scratched the surface of this field.

References

Aakra Å., Nyquist O. L., Snipen L., Reiersen T. S., Nes I. F. Survey of genomic diversity among Enterococcus faecalis strains by microarray-based comparative genomic hybridization. Applied and Environmental Microbiology. 2007;73(7):2207–2217. PubMed PMID: 17220255.

Arias C. A., Panesso D., McGrath D. M., Qin X., Mojica M. F., Miller C., et al. Genetic basis for in vivo daptomycin resistance in enterococci. The New England Journal of Medicine. 2011;365:892–900. PubMed PMID: 21899450.

Arias C. A., Panesso D., Singh K. V., Rice L. B., Murray B. E. Cotransfer of antibiotic resistance genes and a hylEfm-containing virulence plasmid in Enterococcus faecium. Antimicrobial Agents and Chemotherapy. 2009;53(10):4240–4246. PubMed PMID: 19667280.

Arias C. A., Torres H. A., Singh K. V., Panesso D., Moore J., Wanger A., et al. Failure of daptomycin monotherapy for endocarditis caused by an Enterococcus faecium strain with vancomycin-resistant and vancomycin-susceptible subpopulations and evidence of in vivo loss of the vanA gene cluster. Clinical Infectious Diseases. 2007;45(10):1343–1346. PubMed PMID: 17968832.

Barrick J. E., Yu D. S., Yoon S. H., Jeong H., Oh T. K., Schneider D., et al. Genome evolution and adaptation in a long-term experiment with Escherichia coli. Nature. 2009;461:1243–1247. PubMed PMID: 19838166.

Bogie C. P., Hancock L. E., Gilmore M. S. The Enterococcus faecalis cytolysin determinant and its relationship to those encoding lantibiotics. Developments in Biological Standardization. 1995;85:627–634. PubMed PMID: 8586241.

Booth M. C., Bogie C. P., Sahl H. G., Siezen R. J., Hatter K. L., Gilmore M. S. Structural analysis and proteolytic activation of Enterococcus faecalis cytolysin, a novel lantibiotic. Molecular Microbiology. 1996;21(6):1175–1184. PubMed PMID: 8898386.

Borgen K., Wasteson Y., Kruse H., Willems R. J. Vancomycin-resistant Enterococcus faecium (VREF) from Norwegian poultry cluster with VREF from poultry from the United Kingdom and The Netherlands in an amplified fragment length polymorphism genogroup. Applied and Environmental Microbiology. 2002;68(6):3133–3137. PubMed PMID: 12039778.

Bourgogne A., Garsin D. A., Qin X., Singh K. V., Sillanpää J., Yerrapragada S., et al. Large scale variation in Enterococcus faecalis illustrated by the genome analysis of strain OG1RF. Genome Biology. 2008;9(7):R110. PubMed PMID: 18611278.

Brede D. A., Snipen L. G., Ussery D. W., Nederbragt A. J., Nes I. F. Complete genome sequence of the commensal Enterococcus faecalis 62, isolated from a healthy Norwegian infant. Journal of Bacteriology. 2011;193(9):2377. PubMed PMID: 21398545.

Brisse S., Fussing V., Ridwan B., Verhoef J., Willems R. J. Automated ribotyping of vancomycin-resistant Enterococcus faecium isolates. Journal of Clinical Microbiology. 2002;40(6):1977–1984. PubMed PMID: 12037051.

Brüssow H., Canchaya C., Hardt W. D. Phages and the evolution of bacterial pathogens: from genomic rearrangements to lysogenic conversion. Microbiology and Molecular Biology Reviews. 2004;68(3):560–602. PubMed PMID: 15353570.

Cantón R., Ruiz-Garbajosa P., Chaves R. L., Johnson A. P. A potential role for daptomycin in enterococcal infections: what is the evidence? Journal of Antimicrobial Chemotherapy. 2010;65(6):1126–1136. PubMed PMID: 20363805.

Carias L. L., Rudin S. D., Donskey C. J., Rice L. B. Genetic linkage and cotransfer of a novel, vanB-containing transposon (Tn5382) and a low-affinity penicillin-binding protein 5 gene in a clinical vancomycin-resistant Enterococcus faecium isolate. Journal of Bacteriology. 1998;180(17):4426–4434. PubMed PMID: 9721279.

Carver T. J., Rutherford K. M., Berriman M., Rajandream M.-A., Barrell B. G., Parkhill J. ACT: the Artemis Comparison Tool. Bioinformatics. 2005;21(16):3422–3423. PubMed PMID: 15976072.

Chang D., Zhu Y., Zou Y., Fang X., Li T., Wang J., et al. Draft Genome Sequence of Enterococcus faecium Strain LCT-EF90. Journal of Bacteriology. 2012;194(13):3556–3557. PubMed PMID: 22689242.

Choi S. H., Lee S. O., Kim T. H., Chung J. W., Choo E. J., Kwak Y. G., et al. Clinical features and outcomes of bacteremia caused by Enterococcus casseliflavus and Enterococcus gallinarum: analysis of 56 cases. Clinical Infectious Diseases. 2004;38:53–61. PubMed PMID: 14679448.

Chow J. W., Thal L. A., Perri M. B., Vasquez J. A., Donabedian S. M., Clewell D. B., et al. Plasmid-associated hemolysin and aggregation substance production contribute to virulence in experimental enterococcal endocarditis. Antimicrobial Agents and Chemotherapy. 1993;37(11):2474–2477. PubMed PMID: 8285637.

Clewell D. B. Properties of Enterococcus faecalis plasmid pAD1, a member of a widely disseminated family of pheromone-responding, conjugative, virulence elements encoding cytolysin. Plasmid. 2007;58(3):205–227. PubMed PMID: 17590438.

Coburn P. S., Gilmore M. S. The Enterococcus faecalis cytolysin: a novel toxin active against eukaryotic and prokaryotic cells. Cellular Microbiology. 2003;5(10):661–669. PubMed PMID: 12969372.

Coburn P. S., Hancock L. E., Booth M. C., Gilmore M. S. A novel means of self-protection, unrelated to toxin activation, confers immunity to the bactericidal effects of the Enterococcus faecalis cytolysin. Infection and Immunity. 1999;67(7):3339–3347. PubMed PMID: 10377111.

Coburn P. S., Pillar C. M., Jett B. D., Haas W., Gilmore M. S. Enterococcus faecalis senses target cells and in response expresses cytolysin. Science. 2004;306(5705):2270–2272. PubMed PMID: 15618522.

Collins M. D., Jones D., Farrow J. A., Kilpper-Balz R., Schleifer K. H. Enterococcus avium nom. rev., comb. nov.; E. casseliflavus nom. rev., comb. nov.; E. durans nom. rev., comb. nov.; E. gallinarum comb. nov.; and E. malodoratus sp. nov. International Journal of Systematic and Evolutionary Microbiology. 1984;34(2):220–223.

Contreras G. A. Nosocomial outbreak of Enteroccocus gallinarum: untaming of rare species of enterococci. Journal of Hospital Infection. 2008;70(4):346–352. DiazGranados, C. A., Cortes, L., Reyes, J., Vanegas, S., Panesso, D., et al. PubMed PMID: 18799242.

Coque T. M., Willems R. J., Fortún J., Top J., Diz S., Loza E., et al. Population structure of Enterococcus faecium causing bacteremia in a Spanish university hospital: setting the scene for a future increase in vancomycin resistance? Antimicrobial Agents and Chemotherapy. 2005;49(7):2693–2700. PubMed PMID: 15980338.

Corander J., Tang J. Bayesian analysis of population structure based on linked molecular information. Mathematical Biosciences. 2007;205(1):19–31. PubMed PMID: 17087977.

Corander J., Marttinen P., Sirén J., Tang J. Enhanced Bayesian modelling in BAPS software for learning genetic structures of populations. BMC Bioinformatics. 2008;9:539. PubMed PMID: 19087322.

Cotroneo N., Harris R., Perlmutter N., Beveridge T., Silverman J. A. Daptomycin exerts bactericidal activity without lysis of Staphylococcus aureus. Antimicrobial Agents and Chemotherapy. 2008;52(6):2223–2225. PubMed PMID: 18378708.

Cox C. R., Coburn P. S., Gilmore M. S. Enterococcal cytolysin: a novel two component peptide system that serves as a bacterial defense against eukaryotic and prokaryotic cells. Current Protein and Peptide Science. 2005;6(1):77–84. PubMed PMID: 15638770.

Croucher N. J., Harris S. R., Fraser C., Quail M. A., Burton J., van der Linden M., et al. Rapid pneumococcal evolution in response to clinical interventions. Science. 2011;331(6016):430–434. PubMed PMID: 21273480.

de Perio M. A., Yarnold P. R., Warren J., Noskin G. A. Risk factors and outcomes associated with non-Enterococcus faecalis, non-Enterococcus faecium enterococcal bacteremia. Infection Control and Hospital Epidemiology. 2006;27(1):28–33. PubMed PMID: 16418983.

de Regt M. J., van Schaik W., van Luit-Asbroek M., Dekker H. A., van Duijkeren E., Koning C. J., et al. Hospital and community ampicillin-resistant Enterococcus faecium are evolutionarily closely linked but have diversified through niche adaptation. PLoS One. 2012;7(2):e30319. PubMed PMID: 22363425.

Devriese L. A., Pot B., Collins M. D. Phenotypic identification of the genus Enterococcus and differentiation of phylogenetically distinct enterococcal species and species groups. Journal of Applied Bacteriology. 1993;75(5):399–408. PubMed PMID: 8300442.

Drummond A. J., Suchard M. A., Xie D., Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Molecular Biology and Evolution. 2012;29(8):1969–1973. PubMed PMID: 22367748.

Dunny G. M. The peptide pheromone-inducible conjugation system of Enterococcus faecalis plasmid pCF10: cell-cell signalling, gene transfer, complexity and evolution. Philosophical Transactions of the Royal Society B. 2007;362(1483):1185–1193. PubMed PMID: 17360276.

Facklam, R. R., Carvalho, M. d., & Teixeira, M. (2002). History, taxonomy, biochemical characteristics, and antibiotic susceptibility testing of enterococci. In M. S. Gilmore (Ed.), The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance (pp. 1-54). Washington, D. C.: ASM Press.

Feil E. J., Li B. C., Aanensen D. M., Hanage W. P., Spratt B. G. eBURST: inferring patterns of evolutionary descent among clusters of related bacterial genotypes from multilocus sequence typing data. Journal of Bacteriology. 2004;186(5):1518–1530. PubMed PMID: 14973027.

Fortina M. G., Ricci G., Mora D., Manachini P. L. Molecular analysis of artisanal Italian cheeses reveals Enterococcus italicus sp. nov. International Journal of Systematic and Evolutionary Microbiology. 2004;54(Pt 5):1717–1721. PubMed PMID: 15388734.

Freitas A. R., Novais C., Ruiz-Garbajosa P., Coque T. M., Peixe L. Clonal expansion within clonal complex 2 and spread of vancomycin-resistant plasmids among different genetic lineages of Enterococcus faecalis from Portugal. Journal of Antimicrobial Chemotherapy. 2009;63(6):1104–1111. PubMed PMID: 19329507.

Freitas A. R., Tedim A. P., Novais C., Ruiz-Garbajosa P., Werner G., Laverde-Gomez J. A., et al. Global spread of the hyl(Efm) colonization-virulence gene in megaplasmids of the Enterococcus faecium CC17 polyclonal subcluster. Antimicrobial Agents and Chemotherapy. 2010;54(6):2660–2665. PubMed PMID: 20385861.

Galloway-Peña J., Bourgogne A., Qin X., Murray B. E. Diversity of the fsr-gelE region of the Enterococcus faecalis genome but conservation in strains with partial deletions of the fsr operon. Applied and Environmental Microbiology. 2011;77(2):442–451. PubMed PMID: 21097591.

Galloway-Peña J., Rice L. B., Murray B. E. Analysis of PBP5 of early U.S. isolates of Enterococcus faecium: sequence variation alone does not explain increasing ampicillin resistance over time. Antimicrobial Agents and Chemotherapy. 2011;55(7):3272–3277. PubMed PMID: 21576454.

Galloway-Peña J., Roh J. H., Latorre M., Qin X., Murray B. E. Genomic and SNP analyses demonstrate a distant separation of the hospital and community-associated clades of Enterococcus faecium. PLoS One. 2012;7(1):e30187. PubMed PMID: 22291916.

Garnier F., Taourit S., Glaser P., Courvalin P., Galimand M. Characterization of transposon Tn1549, conferring VanB-type resistance in Enterococcus spp. Microbiology. 2000;146(Pt 6):1481–1489. PubMed PMID: 10846226.

Garrido-Fernández J., Maldonado-Barragán A., Caballero-Guerrero B., Hornero-Méndez D., Ruiz-Barba J. L. Carotenoid production in Lactobacillus plantarum. International Journal of Food Microbiology. 2010;140(1):34–39. PubMed PMID: 20303609.

Garsin D. A., Sifri C. D., Mylonakis E., Qin X., Singh K. V., Murray B. E., et al. A simple model host for identifying Gram-positive virulence factors. Proceedings of the National Academy of Sciences. 2001;98(19):10892–10897. PubMed PMID: 11535834.

Gilmore M. S., Segarra R. A., Booth M. C. An HlyB-type function is required for expression of the Enterococcus faecalis hemolysin/bacteriocin. Infection and Immunity. 1990;58(12):3914–3923. PubMed PMID: 2123826.

Gilmore M. S., Segarra R. A., Booth M. C., Bogie C. P., Hall L. R., Clewell D. B. Genetic structure of the Enterococcus faecalis plasmid pAD1-encoded cytolytic toxin system and its relationship to lantibiotic determinants. Journal of Bacteriology. 1994;176(23):7335–7344. PubMed PMID: 7961506.

Gold O. G., Jordan H. V., van Houte J. The prevalence of enterococci in the human mouth and their pathogenicity in animal models. Archives of Oral Biology. 1975;20(7):473–477. PubMed PMID: 807189.

Goris J., Konstantinidis K. T., Klappenbach J. A., Coenye T., Vandamme P., Tiedje J. M. DNA-DNA hybridization values and their relationship to whole-genome sequence similarities. International Journal of Systematic and Evolutionary Microbiology. 2007;57(Pt 1):81–91. PubMed PMID: 17220447.

Haas W., Shepard B. D., Gilmore M. S. Two-component regulator of Enterococcus faecalis cytolysin responds to quorum-sensing autoinduction. Nature. 2002;415(6867):84–87. PubMed PMID: 11780122.

Hancock L. E., Gilmore M. S. The capsular polysaccharide of Enterococcus faecalis and its relationship to other polysaccharides in the cell wall. Proceedings of the National Academy of Sciences. 2002;99(3):1574–1579. PubMed PMID: 11830672.

Hancock L. E., Shepard B. D., Gilmore M. S. Molecular analysis of the Enterococcus faecalis serotype 2 polysaccharide determinant. Journal of Bacteriology. 2003;185(15):4393–4401. PubMed PMID: 12867447.

Harris S. R., Feil E. J., Holden M. T., Quail M. A., Nickerson E. K., Chantratita N., et al. Evolution of MRSA during hospital transmission and intercontinental spread. Science. 2010;327(5964):469–474. PubMed PMID: 20093474.

Heikens E., Bonten M. J., Willems R. J. Enterococcal surface protein Esp is important for biofilm formation of Enterococcus faecium E1162. Journal of Bacteriology. 2007;189(22):8233–8240. PubMed PMID: 17827282.

Heikens E., Leendertse M., Wijnands M., van Luit-Asbroek M., Bonten M. J., van der Poll T., et al. Enterococcal surface protein Esp is not essential for cell adhesion and intestinal colonization of Enterococcus faecium in mice. BMC Microbiology. 2009;9:19. PubMed PMID: 19178704.

Heikens E., Singh K. V., Jacques-Palaz K. D., van Luit-Asbroek M., Oostdijk E. A., Bonten M. J., et al. Contribution of the enterococcal surface protein Esp to pathogenesis of Enterococcus faecium endocarditis. Microbes and Infection. 2011;13(14-15):1185–1190. PubMed PMID: 21911077.

Heikens E., van Schaik W., Leavis H. L., Bonten M. J., Willems R. J. Identification of a novel genomic island specific to hospital-acquired clonal complex 17 Enterococcus faecium isolates. Applied and Environmental Microbiology. 2008;74(22):7094–7097. PubMed PMID: 18836023.

Hendrickx A. P., van Luit-Asbroek M., Schapendonk C. M., van Wamel M. J., Braat J. C., Wijnands L. M., et al. SgrA, a nidogen-binding LPXTG surface adhesin implicated in biofilm formation, and EcbA, a collagen

binding MSCRAMM, are two novel adhesins of hospital-acquired Enterococcus faecium. Infection and Immunity. 2009;77(11):5097–5106. PubMed PMID: 19737906.

Hirt H., Schlievert P. M., Dunny G. M. In vivo induction of virulence and antibiotic resistance transfer in Enterococcus faecalis mediated by the sex pheromone-sensing system of pCF10. Infection and Immunity. 2002;70(2):716–723. PubMed PMID: 11796604.

Hoboth C., Hoffmann R., Eichner A., Henke C., Schmoldt S., Imhof A., et al. Dynamics of adaptive microevolution of hypermutable Pseudomonas aeruginosa during chronic pulmonary infection in patients with cystic fibrosis. The Journal of Infectious Diseases. 2009;200(1):118–130. PubMed PMID: 19459782.

Homan W. L., Tribe D., Poznanski S., Li M., Hogg G., Spalburg E., et al. Multilocus sequence typing scheme for Enterococcus faecium. Journal of Clinical Microbiology. 2002;40(6):1963–1971. PubMed PMID: 12037049.

Huycke M. M., Gilmore M. S. Frequency of aggregation substance and cytolysin genes among enterococcal endocarditis isolates. Plasmid. 1995;34(2):152–156. PubMed PMID: 8559804.

Huycke M. M., Gilmore M. S., Jett B. D., Booth J. L. Transfer of pheromone-inducible plasmids between Enterococcus faecalis in the Syrian hamster gastrointestinal tract. The Journal of Infectious Diseases. 1992;166(5):1188–1191. PubMed PMID: 1402034.

Huycke M. M., Spiegel C. A., Gilmore M. S. Bacteremia caused by hemolytic, high-level gentamicin-resistant Enterococcus faecalis. Antimicrobial Agents and Chemotherapy. 1991;35(8):1626–1634. PubMed PMID: 1929336.

Ike Y., Clewell D. B., Segarra R. A., Gilmore M. S. Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in Enterococcus faecalis: Tn917 insertional mutagenesis and cloning. Journal of Bacteriology. 1990;172(1):155–163. PubMed PMID: 2152897.

Ike Y., Hashimoto H., Clewell D. B. Hemolysin of Streptococcus faecalis subspecies zymogenes contributes to virulence in mice. Infection and Immunity. 1984;45(2):528–530. PubMed PMID: 6086531.

Ike Y., Hashimoto H., Clewell D. B. High incidence of hemolysin production by Enterococcus (Streptococcus) faecalis strains associated with human parenteral infections. Journal of Clinical Microbiology. 1987;25(8):1524–1528. PubMed PMID: 3114322.

Jett B. D., Jensen H. G., Nordquist R. E., Gilmore M. S. Contribution of the pAD1-encoded cytolysin to the severity of experimental Enterococcus faecalis endophthalmitis. Infection and Immunity. 1992;60(6):2445–2452. PubMed PMID: 1587612.

Johnson P. D., Ballard S. A., Grabsch E. A., Stinear T. P., Seemann T., Young H. L., et al. A sustained hospital outbreak of vancomycin-resistant Enterococcus faecium bacteremia due to emergence of vanB E. faecium sequence type 203. The Journal of Infectious Diseases. 2010;202(8):1278–1286. PubMed PMID: 20812846.

Jureen R., Top J., Mhn S. C., Harthug S., Langeland N., Willems R. J. Molecular characterization of ampicillin-resistant Enterococcus faecium isolates from hospitalized patients in Norway. Journal of Clinical Microbiology. 2003;41(6):2330–2336. PubMed PMID: 12791845.

Kawalec M., Pietras Z., Daniłowicz E., Jakubczak A., Gniadkowski M., Hyrniewicz W., et al. Clonal structure of Enterococcus faecalis isolated from Polish hospitals: characterization of epidemic clones. Journal of Clinical Microbiology. 2007;45(1):147–153. PubMed PMID: 17093020.

Kelesidis T., Humphries R., Uslan D. Z., Pegues D. A. Daptomycin nonsusceptible enterococci: An emerging challenge for clinicians. Clinical Infectious Diseases. 2011;52:228–234. PubMed PMID: 21288849.

Khan F. Y., Elshafi S. S. Enterococcus gallinarum meningitis: a case report and literature review. The Journal of Infection in Developing Countries. 2011;5(3):231–234. PubMed PMID: 21444994.

Kim D. S., Singh K. V., Nallapareddy S. R., Qin X., Panesso D., Arias C. A., et al. The fms21 (pilA)-fms20 locus encoding one of four distinct pili of Enterococcus faecium is harboured on a large transferable plasmid

associated with gut colonization and virulence. Journal of Medical Microbiology. 2010;59(Pt 4):505–507. PubMed PMID: 20075113.

Kuch A., Willems R. J., Werner G., Coque T. M., Hammerum A. M., Sundsfjord A., et al. Insight into antimicrobial susceptibility and population structure of contemporary human Enterococcus faecalis isolates from Europe. Journal of Antimicrobial Chemotherapy. 2012;67(3):551–558. PubMed PMID: 22207599.

Lam M. M., Seemann T., Bulach D. M., Gladman S. L., Chen H., Haring V., et al. Comparative Analysis of the First Complete Enterococcus faecium Genome. Journal of Bacteriology. 2012;194:2334–2341. PubMed PMID: 22366422.

Langille M. G., Hsiao W. W., Brinkman F. S. Detecting genomic islands using bioinformatics approaches. Nature Reviews Microbiology. 2010;8(5):373–382. PubMed PMID: 20395967.

Laverde Gomez J. A., van Schaik W., Freitas A. R., Coque T. M., Weaver K. E., Francia M. V., et al. A multiresistance megaplasmid pLG1 bearing a hylEfm genomic island in hospital Enterococcus faecium isolates. International Journal of Medical Microbiology. 2011;301(2):165–175. PubMed PMID: 20951641.

Lawson P. A., Foster G., Falsen E., Ohlén M., Collins M. D. Vagococcus lutrae sp. nov., isolated from the common otter (Lutra lutra). International Journal of Systematic and Evolutionary Microbiology. 1999;49(Pt 3):1251–1254. PubMed PMID: 10425788.

Leavis H. L., Bonten M. J., Willems R. J. Identification of high-risk enterococcal clonal complexes: global dispersion and antibiotic resistance. Current Opinion in Microbiology. 2006;9(5):454–460. PubMed PMID: 16880002.

Leavis H. L., Willems R. J., Top J., Bonten M. J. High-level ciprofloxacin resistance from point mutations in gyrA and parC confined to global hospital-adapted clonal lineage CC17 of Enterococcus faecium. Journal of Clinical Microbiology. 2006;44(3):1059–1064. PubMed PMID: 16517894.

Leavis H. L., Willems R. J., van Wamel W. J., Schuren F. H., Caspers M. P., Bonten M. J. Insertion sequence-driven diversification creates a globally dispersed emerging multiresistant subspecies of E. faecium. PLOS Pathogens. 2007;3(1):e7. PubMed PMID: 17257059.

Leavis H., Top J., Shankar N., Borgen K., Bonten M., van Embden J., et al. A novel putative enterococcal pathogenicity island linked to the esp virulence gene of Enterococcus faecium and associated with epidemicity. Journal of Bacteriology. 2004;186(3):672–682. PubMed PMID: 14729692.

Lebreton F., van Schaik W., McGuire A. M., Godfrey P., Griggs A., Mazumdar V., et al. Emergence of epidemic multidrug-resistant Enterococcus faecium from animal and commensal strains. MBio. 2013;4(4):e00534–00513. PubMed PMID: 23963180.

Leendertse M., Heikens E., Wijnands L. M., van Luit-Asbroek M., Teske G. J., Roelofs J. J., et al. Enterococcal surface protein transiently aggravates Enterococcus faecium-induced urinary tract infection in mice. The Journal of Infectious Diseases. 2009;200(7):1162–1165. PubMed PMID: 19702507.

Lepage E., Brinster S., Caron C., Ducroix-Crepy C., Rigottier-Gois L., Dunny G., et al. Comparative genomic hybridization analysis of Enterococcus faecalis: identification of genes absent from food strains. Journal of Bacteriology. 2006;188(19):6858–6868. PubMed PMID: 16980489.

Li X., Alvarez V., Harper W. J., Wang H. H. Persistent, toxin-antitoxin system-independent, tetracycline resistance-encoding plasmid from a dairy Enterococcus faecium isolate. Applied and Environmental Microbiology. 2011;77(20):7096–7103. PubMed PMID: 21784909.

Licht T. R., Laugesen D., Jensen L. B., Jacobsen B. L. Transfer of the pheromone-inducible plasmid pCF10 among Enterococcus faecalis microorganisms colonizing the intestine of mini-pigs. Applied and Environmental Microbiology. 2002;68(1):187–193. PubMed PMID: 11772626.

Lindenstrauss A. G., Pavlovic M., Bringmann A., Behr J., Ehrmann M. A., Vogel R. F. Comparison of genotypic and phenotypic cluster analyses of virulence determinants and possible role of CRISPR elements towards

their incidence in Enterococcus faecalis and Enterococcus faecium. Systematic and Applied Microbiology. 2011;34(8):553–560. PubMed PMID: 21943678.

Magni C., Espeche C., Repizo G. D., Saavedra L., Suárez C. A., Blancato V. S., et al. Draft genome sequence of Enterococcus mundtii CRL1656. Journal of Bacteriology. 2012;194(2):550. PubMed PMID: 22207752.

Maiden M. C., Bygraves J. A., Feil E., Morelli G., Russell J. E., Urwin R., et al. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proceedings of the National Academy of Sciences. 1998;95(6):3140–3145. PubMed PMID: 9501229.

Manson J. M., Hancock L. E., Gilmore M. S. Mechanism of chromosomal transfer of Enterococcus faecalis pathogenicity island, capsule, antimicrobial resistance, and other traits. Proceedings of the National Academy of Sciences. 2010;107(27):12269–12274. PubMed PMID: 20566881.

Maraccini P. A., Ferguson D. M., Boehm A. B. Diurnal variation in Enterococcus species composition in polluted ocean water and a potential role for the enterococcal carotenoid in protection against photoinactivation. Applied and Environmental Microbiology. 2012;78(2):305–310. PubMed PMID: 22081569.

Mazaheri Nezhad Fard R., Barton M. D., Heuzenroeder M. W. Novel Bacteriophages in Enterococcus spp. Current Microbiology. 2010;60(6):400–406. PubMed PMID: 19967374.

Mazaheri Nezhad Fard R., Barton M. D., Heuzenroeder M. W. Bacteriophage-mediated transduction of antibiotic resistance in enterococci. Letters in Applied Microbiology. 2011;52(6):559–564. PubMed PMID: 21395627.

McBride S. M., Coburn P. S., Baghdayan A. S., Willems R. J., Grande M. J., Shankar N., et al. Genetic variation and evolution of the pathogenicity island of Enterococcus faecalis. Journal of Bacteriology. 2009;191(10):3392–3402. PubMed PMID: 19270086.

McBride S. M., Fischetti V. A., Leblanc D. J., Moellering R. C. Jr, Gilmore M. S. Genetic diversity among Enterococcus faecalis. PLoS One. 2007;2(7):e582. PubMed PMID: 17611618.

McShan, W. M., & Shankar, N. (2002). The genome of Enterococcus faecalis V583: a tool for discovery. In M. S. Gilmore (Ed.), The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance (pp. 409-415). Washington, D. C.: ASM Press.

Morowitz M. J., Denef V. J., Costello E. K., Thomas B. C., Poroyko V., Relman D. A., et al. Strain-resolved community genomic analysis of gut microbial colonization in a premature infant. Proceedings of the National Academy of Sciences. 2011;108:1128–1133. PubMed PMID: 21191099.

Mundt J. O., Graham W. F. Streptococcus faecium var. casselifavus, nov. var. Journal of Bacteriology. 1968;95(6):2005–2009. PubMed PMID: 5669887.

Munoz-Price L. S., Lolans K., Quinn J. P. Emergence of resistance to daptomycin during treatment of vancomycin-resistant Enterococcus faecalis infection. Clinical Infectious Diseases. 2005;41(4):565–566. PubMed PMID: 16028170.

Nallapareddy S. R., Duh R. W., Singh K. V., Murray B. E. Molecular typing of selected Enterococcus faecalis isolates: pilot study using multilocus sequence typing and pulsed-field gel electrophoresis. Journal of Clinical Microbiology. 2002;40(3):868–876. PubMed PMID: 11880407.

Nallapareddy S. R., Singh K. V., Murray B. E. Construction of improved temperature-sensitive and mobilizable vectors and their use for constructing mutations in the adhesin-encoding acm gene of poorly transformable clinical Enterococcus faecium strains. Applied and Environmental Microbiology. 2006;72(1):334–345. PubMed PMID: 16391062.

Nallapareddy S. R., Wenxiang H., Weinstock G. M., Murray B. E. Molecular characterization of a widespread, pathogenic, and antibiotic resistance-receptive Enterococcus faecalis lineage and dissemination of its putative pathogenicity island. Journal of Bacteriology. 2005;187(16):5709–5718. PubMed PMID: 16077117.

Oancea C., Klare I., Witte W., Werner G. Conjugative transfer of the virulence gene, esp, among isolates of Enterococcus faecium and Enterococcus faecalis. Journal of Antimicrobial Chemotherapy. 2004;54:232–235. PubMed PMID: 15192049.

Palmer K. L., Gilmore M. S. Multidrug-resistant enterococci lack CRISPR-cas. MBio. 2010;1(4):e00227–e00210. PubMed PMID: 21060735.

Palmer K. L., Carniol K., Manson J. M., Heiman D., Shea T., Young S., et al. High-quality draft genome sequences of 28 Enterococcus sp. isolates. Journal of Bacteriology. 2010;192(9):2469–2470. PubMed PMID: 20207762.

Palmer K. L., Daniel A., Hardy C., Silverman J., Gilmore M. S. Genetic basis for daptomycin resistance in enterococci. Antimicrobial Agents and Chemotherapy. 2011;55(7):3345–3356. PubMed PMID: 21502617.

Palmer K. L., Godfrey P., Griggs A., Kos V. N., Zucker J., Desjardins C., et al. Comparative genomics of enterococci: variation in Enterococcus faecalis, clade structure in E. faecium, and defining characteristics of E. gallinarum and E. casseliflavus. MBio. 2012;3(1):e00318–e00311. PubMed PMID: 22354958.

Palmer K. L., Kos V. N., Gilmore M. S. Horizontal gene transfer and the genomics of enterococcal antibiotic resistance. Current Opinion in Microbiology. 2010;13(5):632–639. PubMed PMID: 20837397.

Panesso D., Montealegre M. C., Rincón S., Mojica M. F., Rice L. B., Singh K. V., et al. The hylEfm gene in pHylEfm of Enterococcus faecium is not required in pathogenesis of murine peritonitis. BMC Microbiology. 2011;11:20. PubMed PMID: 21266081.

Papp-Wallace K. M., Maguire M. E. Manganese transport and the role of manganese in virulence. Annual Review of Microbiology. 2006;60:187–209. PubMed PMID: 16704341.

Paulsen I. T., Banerjei L., Myers G. S., Nelson K. E., Seshadri R., Read T. D., et al. Role of mobile DNA in the evolution of vancomycin-resistant Enterococcus faecalis. Science. 2003;299(5615):2071–2074. PubMed PMID: 12663927.

Pelz A., Wieland K. P., Putzbach K., Hentschel P., Albert K., Götz F. Structure and biosynthesis of staphyloxanthin from Staphylococcus aureus. The Journal of Biological Chemistry. 2005;280(37):32493–32498. PubMed PMID: 16020541.

Prakash V. P., Rao S. R., Parija S. C. Emergence of unusual species of enterococci causing infections, South India. BMC Infectious Diseases. 2005;5:14. PubMed PMID: 15774018.

Qin X., Galloway-Peña J. R., Sillanpää J., Roh J. H., Nallapareddy S. R., Chowdhury S., et al. Complete genome sequence of Enterococcus faecium strain TX16 and comparative genomic analysis of Enterococcus faecium genomes. BMC Microbiology. 2012;12:135. PubMed PMID: 22769602.

Razeto A., Giller K., Haas W., Gilmore M. S., Zweckstetter M., Becker S. Expression, purification, crystallization and preliminary crystallographic studies of the Enterococcus faecalis cytolysin repressor CylR2. Acta Crystallographica Section D: Biological Crystallography. 2004;60(Pt 4):746–748. PubMed PMID: 15039573.

Rice L. B., Carias L., Rudin S., Vael C., Goossens H., Konstabel C., et al. A potential virulence gene, hylEfm, predominates in Enterococcus faecium of clinical origin. The Journal of Infectious Diseases. 2003;187(3):508–512. PubMed PMID: 12552437.

Rice L. B., Lakticová V., Carias L. L., Rudin S., Hutton R., Marshall S. M. Transferable capacity for gastrointestinal colonization in Enterococcus faecium in a mouse model. The Journal of Infectious Diseases. 2009;199(3):342–349. PubMed PMID: 19049434.

Ruiz-Garbajosa P., Bonten M. J., Robinson D. A., Top J., Nallapareddy S. R., Torres C., et al. Multilocus sequence typing scheme for Enterococcus faecalis reveals hospital-adapted genetic complexes in a background of high rates of recombination. Journal of Clinical Microbiology. 2006;44(6):2220–2228. PubMed PMID: 16757624.

Rumpel S., Razeto A., Pillar C. M., Vijayan V., Taylor A., Giller K., et al. Structure and DNA-binding properties of the cytolysin regulator CylR2 from Enterococcus faecalis. The EMBO Journal. 2004;23(18):3632–3642. PubMed PMID: 15359276.

Sahm D. F., Kissinger J., Gilmore M. S., Murray P. R., Mulder R., Solliday J., et al. In vitro susceptibility studies of vancomycin-resistant Enterococcus faecalis. Antimicrobial Agents and Chemotherapy. 1989;33(9):1588–1591. PubMed PMID: 2554802.

Schwarz F. V., Perreten V., Teuber M. Sequence of the 50-kb conjugative multiresistance plasmid pRE25 from Enterococcus faecalis RE25. Plasmid. 2001;46(3):170–187. PubMed PMID: 11735367.

Segarra R. A., Booth M. C., Morales D. A., Huycke M. M., Gilmore M. S. Molecular characterization of the Enterococcus faecalis cytolysin activator. Infection and Immunity. 1991;59(4):1239–1246. PubMed PMID: 1900808.

Shankar N., Baghdayan A. S., Gilmore M. S. Modulation of virulence within a pathogenicity island in vancomycin-resistant Enterococcus faecalis. Nature. 2002;417(6890):746–750. PubMed PMID: 12066186.

Shankar N., Lockatell C. V., Baghdayan A. S., Drachenberg C., Gilmore M. S., Johnson D. E. Role of Enterococcus faecalis surface protein Esp in the pathogenesis of ascending urinary tract infection. Infection and Immunity. 2001;69(7):4366–4372. PubMed PMID: 11401975.

Shankar V., Baghdayan A. S., Huycke M. M., Lindahl G., Gilmore M. S. Infection-derived Enterococcus faecalis strains are enriched in esp, a gene encoding a novel surface protein. Infection and Immunity. 1999;67(1):193–200. PubMed PMID: 9864215.

Sheldon W. L., Macauley M. S., Tayor E. J., Robinson C. E., Charnock S. J., Davies G. J., et al. Functional analysis of a group A streptococcal glycoside hydrolase Spy1600 from family 84 reveals it is a beta-N-acetylglucosaminidase and not a hyaluronidase. Biochemical Journal. 2006;399(2):241–247. PubMed PMID: 16822234.

Shewmaker P. L., Steigerwalt A. G., Morey R. E., Carvalho M. G., Elliott J. A., Joyce K., et al. Vagococcus carniphilus sp. nov., isolated from ground beef. International Journal of Systematic and Evolutionary Microbiology. 2004;54(Pt 5):1505–1510. PubMed PMID: 15388702.

Singh K. V., Qin X., Weinstock G. M., Murray B. E. Generation and testing of mutants of Enterococcus faecalis in a mouse peritonitis model. The Journal of Infectious Diseases. 1998;178(5):1416–1420. PubMed PMID: 9780263.

Singh K. V., Weinstock G. M., Murray B. E. An Enterococcus faecalis ABC homologue (Lsa) is required for the resistance of this species to clindamycin and quinupristin-dalfopristin. Antimicrobial Agents and Chemotherapy. 2002;46(6):1845–1850. PubMed PMID: 12019099.

Sletvold H., Johnsen P. J., Hamre I., Simonsen G. S., Sundsfjord A., Nielsen K. M. Complete sequence of Enterococcus faecium pVEF3 and the detection of an omega-epsilon-zeta toxin-antitoxin module and an ABC transporter. Plasmid. 2008;60(1):75–85. PubMed PMID: 18511120.

Sletvold H., Johnsen P. J., Simonsen G. S., Aasnæs B., Sundsfjord A., Nielsen K. M. Comparative DNA analysis of two vanA plasmids from Enterococcus faecium strains isolated from poultry and a poultry farmer in Norway. Antimicrobial Agents and Chemotherapy. 2007;51(2):736–739. PubMed PMID: 17116680.

Sletvold H., Johnsen P. J., Wikmark O. G., Simonsen G. S., Sundsfjord A., Nielsen K. M. Tn1546 is part of a larger plasmid-encoded genetic unit horizontally disseminated among clonal Enterococcus faecium lineages. Journal of Antimicrobial Chemotherapy. 2010;65(9):1894–1906. PubMed PMID: 20558469.

Smith J. M., Smith N. H., O'Rourke M., Spratt B. G. How clonal are bacteria? Proceedings of the National Academy of Sciences. 1993;90(10):4384–4388. PubMed PMID: 8506277.

Solheim M., Aakra Å., Snipen L. G., Brede D. A., Nes I. F. Comparative genomics of Enterococcus faecalis from healthy Norwegian infants. BMC Genomics. 2009;10:194. PubMed PMID: 19393078.

Solheim M., Brekke M. C., Snipen L. G., Willems R. J., Nes I. F., Brede D. A. Comparative genomic analysis reveals significant enrichment of mobile genetic elements and genes encoding surface structure-proteins in hospital-associated clonal complex 2 Enterococcus faecalis. BMC Microbiology. 2011;11:3. PubMed PMID: 21205308.

Tang J., Hanage W. P., Fraser C., Corander J. Identifying currents in the gene pool for bacterial populations using an integrative approach. PLoS Computational Biology. 2009;5(8):e1000455. PubMed PMID: 19662158.

Tannock, G. W., & Cook, G. (2002). Enterococci as members of the intestinal microflora of humans. In M. S. Gilmore (Ed.), The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance (pp. 101-132). Washington, D. C.

Taylor R. F., Ikawa M., Chesbro W. Carotenoids in yellow-pigmented enterococci. Journal of Bacteriology. 1971;105(2):676–678. PubMed PMID: 5541541.

Teixeira L. M., Carvalho M. G., Merquior V. L., Steigerwalt A. G., Brenner D. J., Facklam R. R. Phenotypic and genotypic characterization of Vagococcus fluvialis, including strains isolated from human sources. Journal of Clinical Microbiology. 1997;35(11):2778–2781. PubMed PMID: 9350732.

Teng F., Singh K. V., Bourgogne A., Zeng J., Murray B. E. Further characterization of the epa gene cluster and Epa polysaccharides of Enterococcus faecalis. Infection and Immunity. 2009;77(9):3759–3767. PubMed PMID: 19581393.

Tettelin H., Riley D., Cattuto C., Medini D. Comparative genomics: the bacterial pan-genome. Current Opinion in Microbiology. 2008;11(5):472–477. PubMed PMID: 19086349.

Top J., Sinnige J. C., Majoor E. A., Bonten M. J., Willems R. J., van Schaik W. The recombinase IntA is required for excision of esp-containing ICEEfm1 in Enterococcus faecium. Journal of Bacteriology. 2011;193(4):1003–1006. PubMed PMID: 21148730.

Turner K. M., Hanage W. P., Fraser C., Connor T. R., Spratt B. G. Assessing the reliability of eBURST using simulated populations with known ancestry. BMC Microbiology. 2007;7:30. PubMed PMID: 17430587.

Urwin R., Maiden M. C. Multi-locus sequence typing: a tool for global epidemiology. Trends in Microbiology. 2003;11(10):479–487. PubMed PMID: 14557031.

van Schaik W., Top J., Riley D. R., Boekhorst J., Vrijenhoek J. E., Schapendonk C. M., et al. Pyrosequencing-based comparative genome analysis of the nosocomial pathogen Enterococcus faecium and identification of a large transferable pathogenicity island. BMC Genomics. 2010;11:239. PubMed PMID: 20398277.

Vancanneyt M., Lombardi A., Andrighetto C., Knijff E., Torriani S., Björkroth K. J., et al. Intraspecies genomic groups in Enterococcus faecium and their correlation with origin and pathogenicity. Applied and Environmental Microbiology. 2002;68(3):1381–1391. PubMed PMID: 11872491.

Vankerckhoven V., Huys G., Vancanneyt M., Snauwaert C., Swings J., Klare I., et al. Genotypic diversity, antimicrobial resistance, and virulence factors of human isolates and probiotic cultures constituting two intraspecific groups of Enterococcus faecium isolates. Applied and Environmental Microbiology. 2008;74(14):4247–4255. PubMed PMID: 18487397.

Vincent S., Knight R. G., Green M., Sahm D. F., Shlaes D. M. Vancomycin susceptibility and identification of motile enterococci. Journal of Clinical Microbiology. 1991;29(10):2335–2337. PubMed PMID: 1939593.

Weaver, K. E., Rice, L. B., & Churchward, G. (2002). Plasmids and Transposons. In M. S. Gilmore (Ed.), The Enterococci: Pathogenesis, Molecular Biology, and Antibiotic Resistance (pp. 219-263). Washington, D. C.

Werner G., Fleige C., Geringer U., van Schaik W., Klare I., Witte W. IS element IS16 as a molecular screening tool to identify hospital-associated strains of Enterococcus faecium. BMC Infectious Diseases. 2011;11:80. PubMed PMID: 21453466.

Wiedenheft B., Sternberg S. H., Doudna J. A. RNA-guided genetic silencing systems in bacteria and archaea. Nature. 2012;482:331–338. PubMed PMID: 22337052.

Willems, R. J. (2010). Population genetics of Enterococcus. In D. A. Robinson, D. Falush, & E. J. Feil (Eds.), Bacterial Population Genetics in Infectious Diseases. Hoboken, New Jersey: John Wiley & Sons.

Willems R. J., van Schaik W. Transition of Enterococcus faecium from commensal organism to nosocomial pathogen. Future Microbiology. 2009;4(9):1125–1135. PubMed PMID: 19895216.

Willems R. J., Hanage W. P., Bessen D. E., Feil E. J. Population biology of Gram-positive pathogens: high-risk clones for dissemination of antibiotic resistance. FEMS Microbiology Review. 2011;35:872–900. PubMed PMID: 21658083.

Willems R. J., Top J., van Den Braak N., van Belkum A., Endtz H., Mevius D., et al. Host specificity of vancomycin-resistant Enterococcus faecium. The Journal of Infectious Diseases. 2000;182(3):816–823. PubMed PMID: 10950776.

Willems R. J., Top J., van Santen M., Robinson D. A., Coque T. M., Baquero F., et al. Global spread of vancomycin-resistant Enterococcus faecium from distinct nosocomial genetic complex. Emerging Infectious Diseases. 2005;11(6):821–828. PubMed PMID: 15963275.

Willems R. J., Top J., van Schaik W., Leavis H., Bonten M., Sirén J., et al. Restricted gene flow among hospital subpopulations of Enterococcus faecium. MBio. 2012;3(4):e00151–e00112. PubMed PMID: 22807567.

Xu Y., Murray B. E., Weinstock G. M. A cluster of genes involved in polysaccharide biosynthesis from Enterococcus faecalis OG1RF. Infection and Immunity. 1998;66(9):4313–4323. PubMed PMID: 9712783.

Yagi Y., Kessler R. E., Shaw J. H., Lopatin D. E., An F., Clewell D. B. Plasmid content of Streptococcus faecalis strain 39-5 and identification of a pheromone (cPD1)-induced surface antigen. Microbiology. 1983;129(4):1207–1215. PubMed PMID: 6411857.

Yasmin A., Kenny J. G., Shankar J., Darby A. C., Hall N., Edwards C., et al. Comparative genomics and transduction potential of Enterococcus faecalis temperate bacteriophages. Journal of Bacteriology. 2010;192(4):1122–1130. PubMed PMID: 20008075.

Zhang X., Vrijenhoek J. E., Bonten M. J., Willems R. J., van Schaik W. A genetic element present on megaplasmids allows Enterococcus faecium to use raffinose as carbon source. Environmental Microbiology. 2011;13(2):518–528. PubMed PMID: 20946531.

Zheng B., Tomita H., Inoue T., Ike Y. Isolation of VanB-type Enterococcus faecalis strains from nosocomial infections: first report of the isolation and identification of the pheromone-responsive plasmids pMG2200, Encoding VanB-type vancomycin resistance and a Bac41-type bacteriocin... Antimicrobial Agents and Chemotherapy. 2009;53(2):735–747. PubMed PMID: 19029325.

License