Multidrug-Resistant Enterococci Lack CRISPR-cas

The Harvard community has made this article openly available. Please share how this access benefits you. Your story matters

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Published Version</td>
<td>doi:10.1128/mBio.00227-10</td>
</tr>
<tr>
<td>Citable link</td>
<td><a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:33867382">http://nrs.harvard.edu/urn-3:HUL.InstRepos:33867382</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <a href="http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA">http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA</a></td>
</tr>
</tbody>
</table>
Multidrug-Resistant Enterococci Lack CRISPR-cas

Kelli L. Palmer and Michael S. Gilmore
Departments of Ophthalmology, and Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts, USA, and the Harvard Microbial Sciences Initiative, Cambridge, Massachusetts, USA

ABSTRACT Clustered, regularly interspaced short palindromic repeats (CRISPR) provide bacteria and archaea with sequence-specific, acquired defense against plasmids and phage. Because mobile elements constitute up to 25% of the genome of multidrug-resistant (MDR) enterococci, it was of interest to examine the codistribution of CRISPR and acquired antibiotic resistance in enterococcal lineages. A database was built from 16 Enterococcus faecalis draft genome sequences to identify commonalities and polymorphisms in the location and content of CRISPR loci. With this data set, we were able to detect identities between CRISPR spacers and sequences from mobile elements, including pheromone-responsive plasmids and phage, suggesting that CRISPR regulates the flux of these elements through the E. faecalis species. Based on conserved locations of CRISPR and CRISPR-cas loci and the discovery of a new CRISPR locus with associated functional genes, CRISPR3-cas, we screened additional E. faecalis strains for CRISPR content, including isolates predating the use of antibiotics. We found a highly significant inverse correlation between the presence of a CRISPR-cas locus and acquired antibiotic resistance in E. faecalis, and examination of an additional eight E. faecium genomes yielded similar results for that species. A mechanism for CRISPR-cas loss in E. faecalis was identified. The inverse relationship between CRISPR-cas and antibiotic resistance suggests that antibiotic use inadvertently selects for enterococcal strains with compromised genome defense.

IMPORTANCE For many bacteria, including the opportunistically pathogenic enterococci, antibiotic resistance is mediated by acquisition of new DNA and is frequently encoded on mobile DNA elements such as plasmids and transposons. Certain enterococcal lineages have recently emerged that are characterized by abundant mobile DNA, including numerous viruses (phage), and plasmids and transposons encoding multiple antibiotic resistances. These lineages cause hospital infection outbreaks around the world. The striking influx of mobile DNA into these lineages is in contrast to what would be expected if a self (genome)-defense system was present. Clustered, regularly interspaced short palindromic repeat (CRISPR) defense is a recently discovered mechanism of prokaryotic self-defense that provides a type of acquired immunity. Here, we find that antibiotic resistance and possession of complete CRISPR loci are inversely related and that members of recently emerged high-risk enterococcal lineages lack complete CRISPR loci. Our results suggest that antibiotic therapy inadvertently selects for enterococci with compromised genome defense.

Enterococcus faecalis and Enterococcus faecium rank among the leading causes of antibiotic-resistant hospital-acquired bacterial infections (1–3). Resistance to last-line drugs, such as vancomycin (4), is common, and enterococci are now disseminating this resistance to methicillin-resistant Staphylococcus aureus (MRSA) (5–7). The first vancomycin-resistant E. faecalis isolate in the United States, V583 (8), is a member of clonal complex 2 (CC2) (9), a cluster of hospital-adapted lineages that emerged as early as 1981 (10). It was the first Enterococcus isolate for which a genome sequence was determined (11). Mobile elements constitute one-quarter of its genome and include three independently replicating plasmids, three chromosomally integrated plasmid remnants, seven prophage, and a pathogenicity island (PAI) (11, 12). The E. faecium hospital-adapted lineage CC17 (13), which emerged as early as 1982 (14), is similarly characterized by an abundance of exogenously acquired genes, including insertion sequences, phage, and antibiotic resistance genes (15).

Enterococci are natural inhabitants of the digestive tracts of humans and other mammals (16). Genome analysis of the natively antibiotic-sensitive human oral isolate of E. faecalis, OG1RF, revealed that this strain lacks most of the mobile elements and externally acquired DNA found in the hospital-adapted strain V583 (17). Instead, two clustered, regularly interspaced short palindromic repeat (CRISPR) loci were identified (17). CRISPR is a prokaryotic sequence-specific defense system that provides a type of acquired immunity (18, 19). Mechanistic details are emerging, but in general, a small segment of an invading mobile element is incorporated into the CRISPR array between roughly palindromic repeats (20, 21). This mobile element segment is then transcribed and processed at the palindromes to generate a small RNA termed crRNA (22). crRNA targets CRISPR-associated nucleases, encoded by the CRISPR-associated genes (cas genes), to incoming mobile elements from which the spacers derive (20, 22, 23). CRISPR defense appears to be widespread in prokaryotes, with
~90% of archaeal and ~45% of bacterial genomes possessing convincing CRISPR loci (24). Of the two CRISPR loci discovered in the *E. faecalis* OG1RF genome, one possesses the associated *cas* nuclease genes (CRISPR1- *cas*) and one is an orphan locus lacking *cas* genes (CRISPR2) (17). The hospital-adapted V583 strain possesses only the orphan locus CRISPR2, lacking the functional *cas* genes required for CRISPR defense (25).

Because of the potential for limiting entry of mobile elements, it was of interest to determine whether a correlation exists between the presence of CRISPR loci and the emergence of antibiotic-resistant enterococcal lineages. By examining enterococcal draft genomes to develop consensus information on the occurrence of CRISPR loci, and expanding the study to a historical collection of isolates that extends coverage through antibiotic and preantibiotic eras, we found a strong correlation between the absence of CRISPR- *cas* loci and the emergence of MDR enterococcal strains. We hypothesize that widespread antibiotic use has selected for enterococcal strains able to readily acquire novel traits—those with compromised genome defense—ultimately leading to the emergence of enterococcal lineages replete with antibiotic resistances and other mobile traits.

### RESULTS

#### CRISPR distribution in 16 *E. faecalis* draft genomes.

We recently collaborated in an effort to sequence the genomes of 28 enterococci, including 16 *E. faecalis* strains representing deep phylogenetic nodes in the *E. faecalis* multilocus sequence typing (MLST) dendrogram (26). CRISPRfinder (27) was used to identify putative CRISPR loci in the 16 draft *E. faecalis* genomes. The orphan CRISPR2 locus, consisting only of palindromes and spacers, was identified in all 16 *E. faecalis* genomes and was invariably located between homologues of *E. faecalis* V583 genes (Fig. 1). The *E. faecalis* V583 CRISPR2 repeat palindromes were highly conserved, in every case being nearly or perfectly identical to the *E. faecalis* OG1RF CRISPR1/2 repeat sequence (see Table S1 in the supplemental material).

![FIG 1 CRISPR loci in *E. faecalis* draft genomes. Strains are organized by date of isolation, from oldest to most recent. Homologues of *E. faecalis* V583 genes are shown as grey arrows and with V583 ORF assignments. CRISPR locus-specific genes are shown as white arrows, and CRISPR spacers are represented by black diamonds. CRISPR spacers with identity to mobile elements are starred. Note that CRISPR repeats are not shown. Red rectangles denote deletions in the X98, E1Sol, and DS5 CRISPR- *cas* loci. An “X” denotes a frameshift mutation in the D6 CRISPR1- *cas* region corresponding to OG1RF_0022. The *E. faecalis* V583 CRISPR2 locus and the OG1RF CRISPR1- *cas* and CRISPR2 loci were previously reported (17, 25). Red bars and text denote novel DNA sequences present in V583. Primer sets used in CRISPR profiling are labeled with uppercase letters: A, CRISPR1- *cas* flanking primers; B, CRISPR1- *cas* csn1 screening primers; C, CRISPR2 screening primers; D, CRISPR3- *cas* flanking primers; E, CRISPR3- *cas* csn1 screening primers. The figure is not drawn to scale.](mbio.asm.org/mbio.asf.org)
CRISPR in Enterococci

TABLE 1 E. faecalis CRISPR spacer identities to mobile genetic elements

<table>
<thead>
<tr>
<th>Strain</th>
<th>CRISPR and spacer no.</th>
<th>Sequence identity</th>
<th>Representative Blast hit</th>
<th>Area of identity in Blast hit</th>
<th>Identical Blast hits</th>
</tr>
</thead>
<tbody>
<tr>
<td>ATCC 4200</td>
<td>1-4</td>
<td>29/30</td>
<td>Enterococcus phage qEF24C</td>
<td>EFP_glp114 hypothetical protein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-4</td>
<td>29/30</td>
<td>V583 prophage 6</td>
<td>EF2813 tall tape measure protein</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-5</td>
<td>30/30</td>
<td>V583 prophage 6</td>
<td>EF2836_EF2387 intergenic region</td>
<td></td>
</tr>
<tr>
<td></td>
<td>2-6</td>
<td>29/30</td>
<td>V583 pTEF2</td>
<td>EF_B0043 sb-6</td>
<td>E. faecalis pCF10</td>
</tr>
<tr>
<td></td>
<td>29/30</td>
<td>E. faecalis pMG2200</td>
<td>ORF54 hypothetical protein BAH02364.1</td>
<td>E. faecalis pBEE99; pYT14</td>
<td></td>
</tr>
<tr>
<td>DS5</td>
<td>1-3</td>
<td>30/30</td>
<td>Enterococcus phage qFL2B</td>
<td>gp34-gp35 intergenic region</td>
<td>Enterococcus phage qFL2A</td>
</tr>
<tr>
<td></td>
<td>29/30</td>
<td>Enterococcus phage qFL3B</td>
<td>gp57-gp58 intergenic region</td>
<td>Enterococcus phage qFL3A; qfit11</td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>1-8</td>
<td>29/30</td>
<td>V583 VR1-integrated plasmid</td>
<td>EF0133-EF0134 intergenic region</td>
<td></td>
</tr>
<tr>
<td></td>
<td>29/30</td>
<td>V583 VR1-integrated plasmid</td>
<td>EF2539-EF2540 intergenic region</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>2-1</td>
<td>29/30</td>
<td>V583 prophage 3</td>
<td>EF1486 endolysin</td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>2-2</td>
<td>29/30</td>
<td>V583 prophage 6</td>
<td>EF2834 hypothetical protein</td>
<td></td>
</tr>
<tr>
<td>D6</td>
<td>2-5</td>
<td>30/30</td>
<td>V583 pTEF2</td>
<td>EF_B0047 membrane protein, putative</td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>2-5</td>
<td>30/30</td>
<td>Enterococcus phage qEF24C</td>
<td>EFP_glp116 hypothetical protein</td>
<td>E. faecalis pCF10</td>
</tr>
<tr>
<td>E1Sol</td>
<td>2-2</td>
<td>30/30</td>
<td>V583 VR11-integrated plasmid</td>
<td>EF2535 nucleotidylytransferase domain protein</td>
<td></td>
</tr>
<tr>
<td>CH188</td>
<td>2-1</td>
<td>28/30</td>
<td>V583 prophage 1</td>
<td>EF0334 portal protein</td>
<td>Enterococcus phage qFL4A</td>
</tr>
<tr>
<td>T8</td>
<td>2-3</td>
<td>30/30</td>
<td>V583 prophage 6</td>
<td>EF2836-EF2387 intergenic region</td>
<td></td>
</tr>
<tr>
<td>T8</td>
<td>2-4</td>
<td>29/30</td>
<td>V583 pTEF2</td>
<td>EF_B0043 sb-6</td>
<td>E. faecalis pCF10</td>
</tr>
<tr>
<td>AR01/DG</td>
<td>2-2</td>
<td>30/30</td>
<td>E. faecalis pMG2200</td>
<td>ORF54 hypothetical protein BAH02364.1</td>
<td>E. faecalis pBEE99; pYT14</td>
</tr>
<tr>
<td>Fly1</td>
<td>3-5</td>
<td>30/30</td>
<td>V583 prophage 6</td>
<td>EF2823 terminase, large subunit, putative</td>
<td></td>
</tr>
<tr>
<td>Fly1</td>
<td>3-6</td>
<td>30/30</td>
<td>V583 prophage 6</td>
<td>EF2838 DNA replication protein</td>
<td></td>
</tr>
<tr>
<td>T11</td>
<td>3-6</td>
<td>30/30</td>
<td>V583 pTEF1</td>
<td>EF_A0022-EF_A0024 intergenic region</td>
<td>E. faecalis pAM373; pAD1</td>
</tr>
</tbody>
</table>

* a The CRISPR locus (1, 2, or 3) followed by the spacer number is shown. CRISPR spacers are numbered in consecutive order from left to right as shown in Fig. 1.

* b Sequence identity is shown as the number of base pairs with sequence identity in GenBank/total number of base pairs in spacer.

* c V583 variable regions (VR) are from a report by McBride et al. (10).

* d This spacer is identical to ATCC 4200 CRISPR2 spacer 5.

* e This spacer is identical to ATCC 4200 CRISPR2 spacer 6.

4200, X98, E1Sol, DS5, and D6) possessed CRISPR1-cas loci, always located between homologues of the V583 ORFs EF0672 and EF0673 (Fig. 1).

The putative Cas proteins encoded by these loci are highly conserved, as are the palindromic repeat sequences (see Tables S1 and S2 in the supplemental material). As for E. faecalis OG1RF CRISPR1-cas, these loci are of the Nmeni subtype, a CRISPR-cas organizational structure found only in vertebrate-associated commensals and pathogens (28). OG1RF CRISPR1-cas includes an ORF 3’ to the CRISPR1 repeat-spacer array, OG1RF_0022, that has no homologue (17) and is not common to Nmeni subtype loci (28). This ORF varies among strains with CRISPR1-cas, with DS5, E1Sol, and X98 possessing a 435-bp deletion extending 255 bp into it (Fig. 1). A frameshift occurs in the homologue of OG1RF_0022 in strain D6 (Fig. 1). Given the variability in this region, this ORF may be dispensable for CRISPR function.

A novel CRISPR locus, CRISPR3-cas, was identified in the genomes of strains Fly1 and T11, occurring between homologues of the E. faecalis V583 ORFs EF1759 and EF1760 (Fig. 1). CRISPR3-cas also exhibits an organization consistent with Nmeni subtype CRISPR loci (28) (see Table S3 in the supplemental material). However, CRISPR3-cas is distinct from CRISPR1-cas and CRISPR2 in palindrome repeat sequence (see Fig. S1 in the supplemental material) and is distinct from CRISPR1-cas in cas gene content. Nmeni subtype loci possess the universal CRISPR marker gene cas1, the core CRISPR gene cas2, and the Nmeni subtype-specific gene csn1, while possession of the Nmeni subtype-specific gene csn2 varies (28). CRISPR1-cas possesses csn2, the fourth and final ORF in the inferred cas operon (17) (Fig. 1). Curiously, the fourth ORF in the inferred E. faecalis T11 CRISPR3-cas operon (Fig. 1) does not encode a protein with significant identity to Csn2 (TIGR01866) (28) or to any other conserved protein domains (Table S3). In Fly1, the region corresponding to this unknown ORF is interrupted by a frameshift mutation that results in two truncated ORFs (Fig. 1).

**CRISPR spacer identities with mobile elements.** No identity was reported between E. faecalis OG1RF CRISPR spacer sequences and known mobile element sequences, or any other sequences deposited in GenBank (17), and thus no evidence supporting a role for CRISPR in E. faecalis self-defense exists. Having a larger data set of 140 CRISPR spacers from 16 additional genomes, we queried the NCBI nonredundant nucleotide database to identify elements with possible identity to E. faecalis CRISPR spacer sequences. Because spacer sequences are small (30 bp), only hits with >90% identity (27 of 30 nucleotides) were considered significant. By this criterion, identities to 38 of 140 E. faecalis CRISPR spacers were found. Of the 38 hits, 19 were identical to CRISPR2 spacers of the previously sequenced E. faecalis OG1RF and V583 genomes and 19 were derived from mobile elements (Table 1). These elements include the pheromone-responsive type plasmids pAD1/pTEF1 and pCF10/pTEF2, enterococcal phage and prophage, and plasmids integrated within the E. faecalis V583 genome (Table 1). Identities to certain mobile elements were more common than others: five strains (ATCC4200, D6, T8, Fly1, and T11) possess spacers with identity to pheromone-responsive plasmids, and five strains (ATCC 4200, D6, T8, Fly1, and AR01/DG) possess spacer sequences with identities to different regions of E. faecalis V583 prophage 6.

**E. faecalis CRISPR-cas and acquired antibiotic resistance.** Examining the entire genomes of 16 E. faecalis strains allowed us...
to determine that, when present, CRISPR loci always occurred at conserved chromosomal positions: the CRISPR1-cas locus occurred between EF0672 and EF0673 homologues, CRISPR2 between EF2063 and EF2061 homologues, and CRISPR3-cas between EF1760 and EF1759 homologues. Based on this information, we extended the analysis of *E. faecalis* CRISPR by examining 29 additional isolates, including strains with isolation dates prior to the mid-1970s (Fig. 2; for additional isolation dates).

![CRISPR-cas and acquired antibiotic resistance in a historical collection of *E. faecalis* strains. *E. faecalis* strains are listed by date of isolation, from oldest to most recent. Acquired antibiotic resistance is shown in red, and CRISPR-cas presence is shown in green. Antibiotic resistance (tetracycline [tetL and tetM], erythromycin [ermB], gentamicin [aac6’-aph2’’], chloramphenicol [cat], ampicillin [blaZ], and vancomycin [vanA and vanB]) was previously profiled (red squares) (10). A single asterisk indicates that gentamicin resistance is conferred by a 3'-5' aminoglycoside phosphotransferase in this strain; double asterisks denote *E. faecalis* strains for which draft or complete genome sequences are available.](mbio.asm.org/FIG2)

**FIG 2** CRISPR-cas and acquired antibiotic resistance in a historical collection of *E. faecalis* strains. *E. faecalis* strains are listed by date of isolation, from oldest to most recent. Acquired antibiotic resistance is shown in red, and CRISPR-cas presence is shown in green. Antibiotic resistance (tetracycline [tetL and tetM], erythromycin [ermB], gentamicin [aac6’-aph2’’], chloramphenicol [cat], ampicillin [blaZ], and vancomycin [vanA and vanB]) was previously profiled (red squares) (10). A single asterisk indicates that gentamicin resistance is conferred by a 3'-5' aminoglycoside phosphotransferase in this strain; double asterisks denote *E. faecalis* strains for which draft or complete genome sequences are available.
and strain source information, see Data Set S1 in the supplemental material). *E. faecalis* strains were first screened for internal regions of CRISPR1 *csn1* and CRISPR3 *csn1* by PCR. Strains negative by PCR were rescreened with primers annealing outside the conserved CRISPR1-*cas* or CRISPR3-*cas* locus positions, and products from these reactions were sequenced to confirm CRISPR absence. Using this approach, positive PCR results indicating the presence or absence of CRISPR were obtained for all strains tested.

CRISPR content for *E. faecalis* V583, OG1RF, and HH22, as previously reported (17, 25), was included in the analysis.

All 48 *E. faecalis* strains possess an orphan CRISPR2 locus (see Table S1 in the supplemental material). In contrast, CRISPR-*cas* distribution varies among antibiotic-sensitive and antibiotic-resistant strains (Fig. 2). One-third of the *E. faecalis* strains in our collection (16/48) possess a complete CRISPR-*cas* locus (CRISPR1-*cas* or CRISPR3-*cas*), and all but two of these either lack acquired antibiotic resistance genes or possess tetM, a gene commonly disseminated by the self-mobilizable conjugative transposon Tn916 (29, 30). Strikingly, of the MDR *E. faecalis* strains in this collection (defined as resistant to two or more antibiotics), all except two lack CRISPR-*cas* loci (n = 22), including all vancomycin-resistant strains (n = 8) (Fig. 2).

We used a combination of statistical analyses to evaluate the hypothesis that *E. faecalis* strains with CRISPR-*cas* possess significantly fewer acquired antibiotic resistance genes than those lacking CRISPR-*cas*. We first tallied acquired antibiotic resistance genes for each strain and performed a one-tailed Wilcoxon rank sum test to address the null hypothesis that there is no difference in the distributions of acquired antibiotic resistance genes between strains that possess CRISPR-*cas* and strains that lack CRISPR-*cas*. Because antibiotic resistance genes are often coacquired on elements conferring resistance to multiple antibiotics, we performed the Wilcoxon rank sum test using progressively stricter models for coacquisition, in which potentially coacquired genes were counted as resistance to one antibiotic, instead of two (see Data Set S1 in the supplemental material). We found that, irrespective of the model used, the null hypothesis could be rejected (P < 0.001), indicating that the distributions of acquired antibiotic resistance genes between strains that possess CRISPR-*cas* and strains that lack CRISPR-*cas*. Since antibiotic resistance genes are often coacquired on elements conferring resistance to multiple antibiotics, we performed the Wilcoxon rank sum test using progressively stricter models for coacquisition, in which potentially coacquired genes were counted as resistance to one antibiotic, instead of two (see Data Set S1 in the supplemental material). We found that, irrespective of the model used, the null hypothesis could be rejected (P < 0.001), indicating that the distributions of acquired antibiotic resistance genes significantly differ between CRISPR-*cas*-positive and -negative *E. faecalis* strains. Probability values did not change for a model in which tetM and *ermB*, which can be codisseminated by Tn916 and related elements (30), are coacquired (P values of 0.0003 if acquired jointly and 0.0004 if acquired separately) (Data Set S1). If *aac6*-aph2″ and *blaZ* are also assumed to be coacquired, as was suggested by the patterns of occurrence of antibiotic resistance in 106 *E. faecalis* strains (10), the P value increases nominally to 0.0005 (Data Set S1). Finally, if *ermB* and vancomycin resistance are assumed to be coacquired (10), in addition to tetM/*ermB* and *aac6*-aph2″/
FIG 4  EfmCRISPR1-cas loci in E. faecium draft genomes. Homologues of E. faecium DO genes are shown as grey arrows and with DO ORF assignments. EfmCRISPR1-cas locus-specific genes are shown as white arrows, and CRISPR spacers are represented by black diamonds. CRISPR spacers with identity to mobile elements are starred. Note that CRISPR repeats are not shown. The figure is not drawn to scale.

blaZ coacquisition, the P value increases to 0.0008 (Data Set S1), remaining highly significant. We additionally used a one-tailed Fisher exact test with a 2 × 2 contingency table comparing cooccurrences of CRISPR-cas and antibiotic resistance to address the null hypothesis that CRISPR-cas presence is unrelated to antibiotic resistance. The P value of 0.007 is highly significant (P < 0.01), which led us to reject the null hypothesis (Data Set S1). The results of these analyses demonstrate that a statistically significant relationship between the presence or absence of CRISPR-cas and acquired antibiotic resistance exists.

To examine the phylogenetic distribution of CRISPR-cas loci in E. faecalis, we generated an MLST dendrogram of the strains used for CRISPR analysis (10) (Fig. 3). CRISPR-cas loci are discontinuously distributed across this MLST dendrogram, and CRISPR-cas distribution varies even between strains of a single sequence type (e.g., ST21 and ST40 lineages) (Fig. 3). In other lineages, including ST6 (a CC2 lineage; 6 strains), ST9 (a CC9 lineage; 4 strains), and CC8 (8 strains) lineages, CRISPR-cas loci are uniformly absent. All three lineages are associated with MDR (9, 10); vancomycin resistance and β-lactamase production first emerged in the United States in ST6 strains of CC2 (10), and the CC2 and CC9 lineages are highly associated with nosocomial infections and hospital endemicity (9). Loss of CRISPR-cas in founders of these lineages may have precipitated their success in acquiring traits facilitating hospital adaptation.

CRISPR-cas in E. faecium. No CRISPR loci were identified in seven recently reported E. faecium draft genomes (31). We therefore examined the CRISPR locus content in eight additional E. faecium genomes (26). The E. faecium strains originated from human clinical samples (strains represented by numbers in Fig. 4 and 5) and from the feces of healthy human volunteers (Com12 and Com15). All were isolated in the early to mid-2000s. We identified an E. faecium CRISPR-cas locus (EfmCRISPR1-cas) in three genomes (Fig. 4). This locus is found between homologues of EfaeDRAFT_1858 and EfaeDRAFT_1857 from the previously sequenced clinical isolate E. faecium DO draft genome (GenBank accession no. ACIY00000000), which itself lacks CRISPR (25). EfmCRISPR1-cas possesses a palindromic repeat sequence that is divergent from those of the E. faecalis CRISPR-cas loci (see Fig. S1 in the supplemental material). The four predicted Cas proteins encoded by EfmCRISPR1-cas are consistent with the Nmeni subtype (Table S4). However, EfmCRISPR1-cas differs from the E. faecalis Nmeni subtype by the presence of an ORF of unknown function 5' to the csn1 gene which is conserved and unique to the three strains possessing EfmCRISPR-cas (Fig. 4). A Blastn comparison of the 20 EfmCRISPR1-cas spacers to sequences in the NCBI nonredundant nucleotide database reveals identity between two spacers and previously identified clostridial and lactococcal phages (Table 2).

MLST phylogeny and acquired antibiotic resistance profiles were determined for the eight E. faecium strains. Four strains belong to the hospital-adapted CC17, being either ST17 or double-locus variants of ST17 (Fig. 5; see Table S5 in the supplemental material). As expected for CC17 strains (1), all are MDR. Three of the CC17 strains possess vanA, encoding vancomycin resistance, while the fourth CC17 strain constitutes a novel sequence type variant and lacks vanA. The remaining four strains belong to other sequence types and lack acquired antibiotic resistance genes (Fig. 5). Although the E. faecalis analysis was smaller in sample size than that for E. faecalis, the distribution of EfmCRISPR1-cas in these genomes supports the conclusion that MDR enterococcal strains generally lack CRISPR-cas; in this case, three of four MDR strains lack EfmCRISPR1-cas, and these three strains are vancomycin-resistant members of the hospital-adapted CC17 lineage.

Mechanism for CRISPR-cas variability. It was previously proposed that deletion of a Streptococcus thermophilus Nmeni subtype CRISPR-cas may occur through recombination (32). Nmeni subtype CRISPR-cas loci possess a conserved but imperfect CRISPR repeat upstream of the cas operon, meaning that a Nmeni CRISPR-cas locus is flanked on both the 5' and 3' ends by repetitive sequences (28, 32). It is possible that recombination could occur at these sites, resulting in deletion (32). We analyzed the E. faecalis genome sequences corresponding to the locations where deletion of CRISPR1-cas or CRISPR3-cas would have occurred. Interestingly, E. faecalis strains lacking CRISPR-cas possess small, common sequences at these sites. E. faecalis V583 possesses 205 bp of unique sequence when aligned with the E. faecalis D6 CRISPR1-cas and flanking region and 53 bp of unique sequence compared to the Fly1 CRISPR3-cas region (Fig. 1). These 205-bp and 53-bp sequences are highly conserved in all E. faecalis

### TABLE 2

<table>
<thead>
<tr>
<th>Strain</th>
<th>Spacer</th>
<th>Sequence identity</th>
<th>Representative BLAST hit</th>
<th>Area of homology</th>
</tr>
</thead>
<tbody>
<tr>
<td>1,231,408</td>
<td>3</td>
<td>28/30</td>
<td>Clostridium novyi NT genome</td>
<td>NT01CX_2197, putative phage integrase</td>
</tr>
<tr>
<td>1,231,408</td>
<td>6</td>
<td>28/30</td>
<td>Lactococcus lactis subsp. lactis KF147 genome</td>
<td>LLKF_1066, phage protein, HNH endonuclease</td>
</tr>
</tbody>
</table>

* a CRISPR spacers are numbered in consecutive order from left to right as shown in Fig. 4.

b Sequence identity is shown as the number of base pairs with sequence identity in GenBank/total number of base pairs in spacer.
strains lacking CRISPR1-cas or CRISPR3-cas (see Fig. S2 and S3 in the supplemental material). This would not be predicted for CRISPR-cas loss by homologous recombination and excision.

We recently reported the low-frequency mobilization and conjugal transfer of large regions of chromosomal DNA from *E. faecalis* V583 donors to *E. faecalis* OG1RF recipients, resulting in transconjugant strains with hybrid V583-OG1RF genomes (33). This transfer was found to be dependent on either of two pheromone-responsive plasmids resident in V583, pTEF1 or pTEF2, and occurred as the result of plasmid integration and chromosome mobilization. Chromosome-to-chromosome movement of the PAI, a vancomycin resistance transposon, capsule genes, and other V583 genes to OG1RF recipients was observed. It is important to note that OG1RF does not possess CRISPR spacers with identity to pTEF1, pTEF2, or other V583 sequences (17), and in these experiments, CRISPR1-cas did not block incoming DNA from V583. Because the OG1RF CRISPR1-cas locus occurs ~41.5 kb 3' to the site of PAI incorporation, it was of interest to determine whether the CRISPR1-cas locus in PAI-containing transconjugants had been displaced by incoming DNA from V583 donors. We used five strains (TC1, TC3, TC4, TC5, and TC12) representing the five classes of PAI transconjugants (A to E) identified in our previous study (33). These transconjugants possess the V583 PAI and various amounts of flanking DNA, representing a total acquisition of 285 to 857 kb of V583 donor chromosome (33). The five PAI-containing transconjugants were screened for the occurrence of the OG1RF CRISPR1-cas locus by PCR. Strikingly, all five transconjugants lacked CRISPR1-cas and instead possessed at this locus the conserved 205-bp sequence found in the V583 donor (see Fig. S2 in the supplemental material). Thus, in a single conjugal transfer event, a novel hybrid strain that simultaneously becomes antibiotic resistant, acquires the PAI, and is rendered deficient for CRISPR-cas can be generated.

**DISCUSSION**

We examined 48 *E. faecalis* strains from a historical collection and 8 recent *E. faecium* isolates to determine the relationship between CRISPR and the emergence of multidrug resistance in enterococci. We found that CRISPR1-cas and CRISPR3-cas loci are variable among *E. faecalis* strains, but, interestingly, an orphan CRISPR2 locus occurred in all *E. faecalis* strains tested. Selective pressure for maintenance of CRISPR2 in a CRISPR-cas-deficient *E. faecalis* background is unclear, as is the origin of CRISPR2 (i.e., whether this locus is functionally and evolutionarily linked to CRISPR1-cas, or once independently possessed cas genes that have been lost). CRISPR1-cas and CRISPR2 repeat sequences are identical (17), and because Cas proteins interact with CRISPR repeat sequences following transcription (22), it is likely that CRISPR1-cas and CRISPR2 are functionally linked. Nevertheless, genome sequence analysis finds that CRISPR-cas-deficient *E. faecalis* strains lack any detectable *cst1* or *cst1*-like gene, which is essential for protection mediated by a Nmeni subtype CRISPR-cas locus (20). The lack of required functional genes in strains with the CRISPR2 orphan locus, together with the abundance of antibiotic resistance genes in E. faecalis, suggests that CRISPR2 alone does not confer immunity in *E. faecalis* hosts. Ultimately, a functional analysis of the CRISPR1-cas and CRISPR2 loci will be required to confirm this hypothesis.

By analyzing a large data set, we were able to detect identities between CRISPR spacers and sequences on known mobile elements, including pheromone-responsive plasmids and phage. Enterococci have a highly coevolved relationship with the narrow-host-range pheromone-responsive plasmids, which encode machinery to utilize extracellular signals produced by enterococcal cells to induce conjugation functions and to manipulate signal production in enterococcal cells in which they reside (34). The results of this study and our recent work (33) support a role for pheromone-responsive plasmids as important drivers of enterococcal genome plasticity, capable of promoting their own transfer, mobilizing chromosomally encoded antibiotic resistance and virulence determinants, and now also causing displacement of CRISPR-cas.

That *E. faecalis* CRISPR loci contain spacers with identity to enterococcal mobile elements, and the distribution of these spacers, suggests that certain elements, such as pheromone-responsive plasmids, are frequently encountered by *E. faecalis* and/or have a propensity to be incorporated into CRISPR loci as spacers. It is interesting that no CRISPR spacers have yet been identified with sequence identity to conjugative transposons such as Tn916, which are also vectors of antibiotic resistance in the enterococci (29). CRISPR elements have been shown experimentally to confer immunity to plasmids and phages (20–23), although many mechanistic details remain unknown. To our knowledge, there is no experimental evidence that CRISPR defense confers protection from conjugative transposons. The observation that the tetM gene, commonly disseminated by Tn916 and related conjugative transposons (29, 30), is present in *E. faecalis* strains possessing CRISPR-cas (Fig. 2) suggests that conjugative transposons may evade this defense. Spacers targeting the Inc18 plasmid family, plasmids that are significant for their role in the dissemination of vancomycin resistance genes from enterococci to MRSA (6, 7), are also absent. Tn916 and the Inc18 plasmid family are broad host range in nature (29, 30), and the notable lack of spacers with identity to these elements may reflect the relative rarity of interspecies transfer or the relative inefficiency of transfer of elements lacking mechanisms for effective pair formation, such as the pheromone-induced aggregation mechanism.

Of the 48 *E. faecalis* and 8 *E. faecium* strains examined in this study, 7 *E. faecalis* and 2 *E. faecium* strains lacked CRISPR-cas and CRISPR2...
also lacked antibiotic resistance. Draft genome sequences are not available for six of these nine strains, so it remains a formal possibility that novel CRISPR-cas loci may be found in those strains. Alternatively, it may be that the absence of CRISPR-cas (particularly in strains isolated before, or early in, the age of antibiotics, such as *E. faecalis* D173 [isolated in 1939], ATCC 19433 [isolated before 1942], and T1 [isolated before 1950] [10]) facilitated acquisition of a mobile element carrying traits other than antibiotic resistance (for example, hemolysin/bacteriocin production and resistance or new metabolic properties), which provided a selective advantage.

Mobile elements often provide an accessory pool of genes that enhance survival in select environments. It is possible that elements such as the phenormone-responsive plasmids and putative enterococcal self-defense mechanisms such as CRISPR-cas act, respectively, to diversify and to stabilize the enterococcal genome, and that the dynamic between these opposing forces is important for the ultimate success of these microbes. Based on the data presented here, we speculate that modern antibiotic therapy has shifted this dynamic toward the facile acquisition of foreign elements conferring antibiotic resistance, among other things, decreasing genome stability/increasing plasticity, and enabling the colonization of new habitats, including the antibiotic-laden hospital environment.

**MATERIALS AND METHODS**

**Genomes, bacterial strains, and media.** Draft genome data for 16 *E. faecalis* and 8 *E. faecium* strains were generated in collaboration with the Broad Institute (26) and accessed at the Enterococcus database website (http://www.broadinstitute.org/annotation/genome/enterococcus.faecalis/MultiHome.html). *E. faecalis* strains used for CRISPR profiling are shown in Fig. 2 and in Data Set S1 in the supplemental material and have previously been described with respect to isolation dates, antibiotic resistance profiles, and MLST data (10). *E. faecalis* OG1RF transconjugants TC1, TC3, TC4, TC5, and TC12, which received the PAI and other acquisition events of the clinical and translational science center, and with helpful input from Michael Caparon, Washington University.

Mobile elements often provide an accessory pool of genes that enhance survival in select environments. It is possible that elements such as the phenormone-responsive plasmids and putative enterococcal self-defense mechanisms such as CRISPR-cas act, respectively, to diversify and to stabilize the enterococcal genome, and that the dynamic between these opposing forces is important for the ultimate success of these microbes. Based on the data presented here, we speculate that modern antibiotic therapy has shifted this dynamic toward the facile acquisition of foreign elements conferring antibiotic resistance, among other things, decreasing genome stability/increasing plasticity, and enabling the colonization of new habitats, including the antibiotic-laden hospital environment.
The content of this paper is solely the responsibility of the authors and does not necessarily represent the official views of Harvard Catalyst, Harvard University, or its affiliated academic health care centers.

We gratefully thank Ian Jorgeson, Janet Manson, and members of the Broad Institute and the Gilmore laboratory for helpful discussions.

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at http://mbio.asm.org/lookup/suppl/doi:10.1128/mBio.00227-10/-/DCSupplemental.

Table S1, DOC file, 0.108 MB.
Table S2, DOC file, 0.081 MB.
Table S3, DOC file, 0.070 MB.
Table S4, DOC file, 0.90 MB.
Table S5, DOC file, 0.056 MB.
Figure S1, PDF file, 0.045 MB.
Figure S2, PDF file, 0.117 MB.
Figure S3, PDF file, 0.050 MB.
Data Set S1, XLS file, 0.050 MB.

REFERENCES


