A Genomic Virulence Reference Map of *Enterococcus faecalis* Reveals an Important Contribution of Phage03-Like Elements in Nosocomial Genetic Lineages to Pathogenicity in a *Caenorhabditis elegans* Infection Model

Sabina Leanti La Rosa,[a,c] Lars-Gustav Snipen,[b] Barbara E. Murray,[c] Rob J. L. Willems,[d] Michael S. Gilmore,[e] Dzung B. Diep,[a] Ingolf F. Nes,[a] Dag Anders Brede[a]

Department of Chemistry, Biotechnology and Food Science, Laboratory of Microbial Gene Technology and Food Microbiology, Norwegian University of Life Sciences, Ås, Norway[a]; Department of Chemistry, Biotechnology and Food Science, Section for Biostatistics, Norwegian University of Life Sciences, Ås, Norway[b]; Department of Internal Medicine, Division of Infectious Diseases, University of Texas Health Science Center, Houston, Texas, USA[c]; Department of Medical Microbiology, University Medical Center Utrecht, Utrecht, The Netherlands[d]; Departments of Ophthalmology, Microbiology and Immunobiology, Harvard Medical School, Boston, Massachusetts, USA[e]

In the present study, the commensal and pathogenic host-microbe interaction of *Enterococcus faecalis* was explored using a *Caenorhabditis elegans* model system. The virulence of 28 *E. faecalis* isolates representing 24 multilocus sequence types (MLSTs), including human commensal and clinical isolates as well as isolates from animals and of insect origin, was investigated using *C. elegans* strain *glp-4* (*bn2ts*); *sek-1* (*km4*). This revealed that 6 *E. faecalis* isolates behaved in a commensal manner with no nematocidal effect, while the remaining strains showed a time to 50% lethality ranging from 47 to 120 h. Principal component analysis showed that the difference in nematocidal activity explained 94% of the variance in the data. Assessment of known virulence traits revealed that gelatinase and cytolysin production accounted for 40.8% and 36.5% of the observed pathogenicity, respectively. However, coproduction of gelatinase and cytolysin did not increase virulence additively, accounting for 50.6% of the pathogenicity and therefore indicating a significant (26.7%) saturation effect. We employed a comparative genomic analysis approach using the 28 isolates comprising a collection of 82,356 annotated coding sequences (CDS) to identify 2,325 patterns of presence or absence among the investigated strains. Univariate statistical analysis of variance (ANOVA) established that individual patterns positively correlated ($n = 61$) with virulence. The patterns were investigated to identify potential new virulence traits, among which we found five patterns consisting of the phage03-like gene clusters. Strains harboring phage03 showed, on average, 17% higher killing of *C. elegans* ($P = 4.4e^{-6}$). The phage03 gene cluster was also present in gelatinase-and-cytolysin-negative strain *E. faecalis* JH2-2. Deletion of this phage element from the JH2-2 clinical strain rendered the mutant apathogenic in *C. elegans*, and a similar mutant of the nosocomial V583 isolate showed significantly attenuated virulence. Bioinformatics investigation indicated that, unlike other *E. faecalis* virulence traits, phage03-like elements were found at a higher frequency among nosocomial isolates. In conclusion, our report provides a valuable virulence map that explains enhancement in *E. faecalis* virulence and contributes to a deeper comprehension of the genetic mechanism leading to the transition from commensalism to a pathogenic lifestyle.

*E*nterococcus faecalis is a Gram-positive commensal bacterium of the mammalian gastrointestinal (GI) tract (1) and a leading cause of nosocomial infections worldwide (2). In humans, the normal abundance of *E. faecalis* in the intestinal lumen ranges between 10^5 and 10^8 CFU/g of feces without causing any obvious deleterious effects on the host (3, 4). However, if either perturbations of the host/commensal balance that weaken the host immune system or environmental factors such as use of antibiotics that inadvertently facilitate outgrowth of resistant *E. faecalis* occur, life-threatening infections might arise. Moreover, the intrinsic robustness enables *E. faecalis* to withstand multiple stresses (5, 6) and provides a fitness advantage with respect to host adaptation and colonization in environments such as the hospital setting (7–9). The aptitude of this bacterium for acquisition and transfer of mobile genetic elements (plasmids, transposons, and prophages) has facilitated the spread of virulence traits and antibiotic resistance genes among isolates (10–14).

Notwithstanding several studies undertaken in the last few decades having identified putative virulence determinants that may augment the ability of *E. faecalis* to cause disease (reviewed in

reference 15), the incidence of many pathogenicity factors has been reported to be independent from the *E. faecalis* isolation source. The availability of 5 complete and more than 300 draft *E. faecalis* genome sequences (http://www.ncbi.nlm.nih.gov

/genome/808) has provided the opportunity to examine the genomic diversity among strains and identify specific traits that may contribute to virulence (16–18). However, none of these experimental investigations have performed comparative estimations of the abilities of different *E. faecalis* isolates to cause infection employing a live model system with the purpose of correlating *E. faecalis* pathogenicity to the whole-genome content.

Several studies have established the soil nematode *Caenorhabditis elegans* as a suitable surrogate animal model for *in vivo* study of host-microbe interactions with *E. faecalis*. Through comparison of the levels of longevity of nematodes infected with isogenic deletion mutants versus parental strains and screening of transposon insertion mutant libraries, several gene categories required for enterococcal pathogenicity in *C. elegans* have been identified, including virulence factors such as cytolysin (Cyl) (19) and gelatinase (Gel) and serine protease (20) and factors playing a role in cell metabolism and physiology (21, 22). Notably, these reports have demonstrated that several virulence factors required for infections of *C. elegans* and mammals coincide (21, 23).

In the present study, we investigated the pathogenicity potential of a collection of 28 *E. faecalis* strains in the *C. elegans* infection model and the correlation between the virulence phenotype and their whole-genome content. Consistent with previous reports, we found that cytolysin (Cyl) and gelatinase (Gel) are the major factors involved in *C. elegans* killing by *E. faecalis*. Further, using a comparative pangenomic analysis combined with statistical regression model testing, a virulence reference map that extends beyond the production of cytolysin and gelatinase was established. This allowed us to identify an important role of a phage03-like element in pathogenicity among nosocomial isolates.

MATERIALS AND METHODS

Bacterial strains and growth conditions. Bacterial strains used in this study are listed in Table S1 in the supplemental material. Genome-sequenced strains were selected to represent the genetic diversity of *E. faecalis* and included human, animal, and insect isolates. The human isolates comprised clinical isolates from urinary tract, liver, and blood infections, adult and baby commensal isolates from the gastrointestinal tract, and the probiotic strain Symbioflor 1. The animal-associated strains were of porcine, canine, phocine, or insect origin. The collection consisted of 24 different sequence types (ST) from 17 known clonal complexes (CC), including the most prevalent nosocomial clonal lineages. The years of isolation of the selected strains spanned 1926 to 2005, and the strains originated in most parts of the world, including the United States, New Zealand, Europe, Japan, and Solomon Islands (16).

E. faecalis strains were cultured in brain heart infusion (BHI) medium (Oxoid Ltd., United Kingdom) or Todd-Hewitt broth (THB; Oxoid Ltd., United Kingdom) at 37°C. *Escherichia coli* strains were grown in Luria Bertani (LB) medium at 37°C. When required for selective growth, chloramphenicol (Cm) was added at 10 μg/ml for *E. coli* and 15 μg/ml for *E. faecalis*. For detection of β-galactosidase activity, X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) was used at 80 μg/ml for *E. coli* and 120 μg/ml for *E. faecalis*.

Cytolysin assay. Hemolytic activity was assayed by culturing *E. faecalis* on blood agar plates supplemented with 5% (vol/vol) defibrinated horse blood, 1% (wt/vol) glucose, and 0.03% (wt/vol) L-arginine (Sigma-Aldrich) (24). Overnight cultures of the strains to be tested were diluted 1:100, spotted onto fresh plates, and incubated at 37°C or 25°C for 24 h under anaerobic conditions. Colonies able to lyse erythrocytes were surrounded by a transparent halo.

Gelatinase assay. The production of gelatinase in *E. faecalis* strains was assessed by the use of TH agar plates containing 3% gelatin (25). Overnight cultures were diluted 1:100 and spotted onto fresh plates. After

incubation at 37°C or 25°C overnight, plates were placed at 4°C for 5 h. The presence of a clear halo around the colonies indicated hydrolysis of gelatin.

Nematode strains and growth conditions. The host infection assay was carried out using the *C. elegans* mutant *glp-4* (*bn2ts*); *sek-1* (*km4*) strain, which is unable to produce offspring at the restrictive temperature of 25°C, thus allowing a prolonged experiment of 7 days without progeny interference (26). After ingestion, *E. faecalis* is able to survive the mechanical digestion and establish a long-lasting, deadly infection in the intestine of *C. elegans* (19). The temperature-sensitive sterile *C. elegans glp-4* (*bn2ts*); *sek-1* (*km4*) mutant strain was obtained from the *Caenorhabditis* Genetic Center, Minneapolis, MN, and propagated at the permissive temperature of 15°C in nematode growth medium (NGM) seeded with *E. coli* OP50 as a food source (27). After 5 days, gravid worms were harvested by rinsing the plates with M9 buffer (27), transferred to 15-ml tubes, and collected by centrifugation for 2 min at 280 × *g* at room temperature (RT). The supernatant was discarded, and sodium hypochlorite solution (Sigma-Aldrich) was added to release the eggs. Eggs were washed twice with M9 buffer and incubated to hatch at 25°C overnight. Newborn worms were transferred to NGM plates seeded with *E. coli* OP50 for 48 to 50 h to obtain synchronized sterile young adult worms to be used in the killing assay.

C. elegans-E. faecalis pathogenicity assay. The *C. elegans* mortality assay was performed as described previously (26) with some minor modifications. Briefly, a tube containing 2 ml of BHI medium was inoculated at 1:100 with the overnight culture of the appropriate *E. faecalis* strain and grown at 37°C for 4 to 5 h to reach an absorbance at 620 nm of 0.4 to 0.5. A 10-μl volume of the culture was spotted onto the center of a BHI plate. The plates were incubated overnight at 37°C and allowed to equilibrate for at least 1 h at RT before transferal of the worms. Young adult nematodes grown at 25°C were washed from plates containing *E. coli* OP50 using sterile M9 buffer. Approximately 200 to 300 worms were transferred to a BHI plate seeded with the *E. faecalis* strain to be tested. Polymyxin B (5 μg/ml) was added to the medium to selectively prevent *E. coli* growth. The plates were incubated at 25°C for 8 h. Nematodes were then carefully washed three times with 2 ml of sterile M9 buffer in order to minimize the carryover of bacterial cells. Approximately 30 nematodes were transferred to each agar plate; all experiments were performed as technical replicates. Plates were incubated at 25°C with 80% to 85% relative humidity. Worms were scored daily with a dissecting microscope for viability. The bacterial isolates were ranked from the most to the least virulent based on the time required to kill 50% of the host population (LT$_{50}$). Strains with no nematocidal effect are referred to as apathogenic.

Plasmid construction and generation of *E. faecalis* mutants. To achieve the construction of phage03 deletion mutants, the phage03-like integration site was PCR amplified using *E. faecalis* OG1RF chromosomal DNA as the template. The generation of an amber point mutation at the *fsrB* codon corresponding to Leu-65 was achieved by using the genomic DNA of *E. faecalis* OU510 (28) as the template for PCR amplification of a 2.0-kb fragment. The two fragments of 2.2 kb and 2.0 kb were independently digested with restriction enzymes EcoRI and PstI and ligated into similarly digested pLT06, giving pSL01 and pSL02, respectively.

pSL01 and pSL02 were electroporated into *E. coli* EC1000, which supplies RepA in *trans*, for replication of the plasmids at the permissive temperature of 30°C (29). Transformants were screened for the presence of the insertion using primers OriF/KS05seqR (see Table S2 in the supplemental material). To exclude mutations occurring during cloning, the fragments were sequenced using the same primers. The plasmids were purified using E.Z.N.A. plasmid minikit I (Omega Bio-tek). Vector pSL01 was electroporated in both *E. faecalis* JH2-2 and *E. faecalis* V583, while pSL02 was introduced only to V583, using standard procedures (30). Gene replacement was achieved by double-crossover homologous recombination following the procedure described by Thurlow et al. (31). To select for the second recombination event, the integration mutants were passaged for 2 days in THB with no selection at 30°C. The chromosomal

DNA of the white colonies arose on M9YEG plates supplemented with 10 mM p-chlorophenylalanine, and X-Gal was screened by PCR with primers int1-F/int2-R (see Table S2).

The phage03 isogenic deletion mutants were named *E. faecalis* SL100 (JH2-2 mutant) and SL110 (V583 mutant), while the frame-shifted *fsrB* V583 mutant was named SL111 (see Table S1 in the supplemental material). Gene replacement in these strains was confirmed by PCR and sequencing using primers flanking the targeted region (see Table S2).

Monitoring of growth in liquid laboratory medium. Overnight cultures of *E. faecalis* strains were diluted 100-fold in fresh GM17 medium and grown to an optical density at 620 nm (OD_{620}) of 0.2. Cells were back-diluted 100-fold in GM17, and a total volume of 300 µl of bacterial inoculum in fresh medium was added to a 96-well plate (Nunc; Thermo Fisher Scientific, Denmark). Cultures were incubated at 37°C under static conditions, and absorbance at 620 nm was measured at 15-min intervals for 11 h with a SPECTROstar Nano microplate reader (BMG Labtech, USA). The experiments were performed as independent triplicates.

Sequence analysis. At the time that this investigation began, the genome sequences of 27 of the 28 wild-type *E. faecalis* isolates included in this study were publicly available (see Table S1 in the supplemental material) whereas a 3.26-Mbp high-quality draft assembly of the *E. faecalis* MMH594 genome was generated by our group (see Text S1 in the supplemental material). For each genome, the RefSeq annotated proteins were downloaded. A sequence comparison was conducted where each annotated protein sequence was subjected to BLAST searches against all annotated proteins. The best hit for every query sequence in each genome was extracted, and a relative score was computed. This score was the obtained BLAST score divided by the maximum attainable BLAST score for the same query sequence; i.e., a relative score of 1.0 means that the query was found with 100% identity in the genome. A threshold of 0.75 was used to identify a sequence as Present (>0.75) or Absent (<0.75) in each genome. Thus, each annotated sequence has a Present/Absent pattern (0's and 1's) over the 28 genomes. Many sequences share the same Present/Absent pattern, and in the downstream analysis we considered only distinct patterns, with a lookup table assigning each sequence to its distinct pattern. Reciprocal or opposite patterns (0 replaced with 1 and vice versa) were considered reflections of the same distinct pattern. Patterns without variation across the 28 genomes (Present in all strains) were discarded since they lack discriminating power.

On the basis of the BLAST all-versus-all alignments, we identified a set of 1,006 core gene families. In cases where a gene family contained paralogs, only the ortholog (best-matching sequence) from each genome was considered. Each gene family was aligned with MUSCLE (32). The protein sequence alignments were back-translated to DNA alignments, and then sequences were concatenated. The raw distance (p-distance) between genomes was computed from this alignment, and a neighbor-joining tree (Bionj in the R-package ape) was constructed (33).

For construction of a tree based on the Present/Absent proteins, each genome was represented as a vector of 82,356 1's or 0's corresponding to whether each sequence in the pan-genome was present ($>75\%$ similar) or absent in the genome. Based on these vectors, the Manhattan distance values of comparisons of all pairs of genomes were computed, and the dendrogram tree was constructed by hierarchical clustering using average linkages.

For multiple-strain alignments, progressiveMauve software was employed (34). Alignments of phage03-like elements were performed with Easyfig (35).

PCA. The experimental data were also illustrated with a principal component analysis (PCA). A matrix with one column for each sample time (48, 72, 96, 120, 144, and 168 h) and one row for each experimental unit (strain, batch, and plate) was filled with percent survival data; i.e., row i contains the time series of survival percentages for experimental unit i over the 6 sample times. This matrix was used as the input for a standard PCA, and each experimental unit was plotted in the two first principal directions (see Fig. 2).

Analysis of variance. The percentage of nematode survival at 72 h after exposure to *E. faecalis* was used in an analysis of variance (ANOVA) to search for associations with various categorical variables. A maximum value of 100 means that all nematodes survived, while the minimum value of 0 means that all died. The experiments were conducted in different batches and time series, and a potential bias in the results from any batch was described as a random effect in an ANOVA model.

The first fixed effects included in the model were the effects of gelatinase (Gel) and cytolysin (Cyl) as well as their interaction:

$$y_{ijhk} = \alpha_i + \beta_j + (\alpha\beta)_{ij} + b_h + e_{ijhk} \qquad (1)$$

where α_i ($i = 1$ or 2) and β_j ($j = 1$ or 2) are the fixed effects of Gel and Cyl, respectively, i.e., α takes value 0 or 1 depending on whether survival y is observed for a Gel-negative (Gel$^-$) or Gel-positive (Gel$^+$) strain, and the definitions of the terms are similar for β, Cyl$^-$, and Cyl$^+$. The term b_h ($h = 1, . . ., 8$) is the random effect of the batch, and e_{ijhk} ($k = 1, 2,$ or 3) (representing triplicate plates where the nematodes were kept for each treatment) is the residual term.

A model similar to that reported in equation 1 was also used for other fixed effects in addition to those seen with Gel and Cyl, e.g., for investigating the effect of the presence or absence of a set of genes, as described in the sequence analysis subsection above:

$$y_{ghk} + \gamma_g + b_h + e_{ghk} \qquad (2)$$

where γ_g ($g = 1$ or 2) is 0 or 1, depending on whether the genes considered are absent or present in the strain from which the corresponding response y is observed. Different patterns of presence or absence were tested against survival in this model, one by one; i.e., the g fixed effect was reestimated for each distinct pattern.

Finally, we also combined model 1 and model 2, yielding the following equation:

$$y_{ijghk} = \alpha_i + \beta_j + (\alpha\beta)_{ij} + \gamma_g + b_h + e_{ijghk} \qquad (3)$$

where the effect of a gene being present or absent competes with the Gel and Cyl effects in explaining the observed variations in survival.

In all cases, the fixed effects were estimated and the null hypothesis of no effect was tested against the two-sided alternative in each case. When we searched for associated Present/Absent patterns, we conducted many hypothesis tests and, for this reason, we computed adjusted P values to correct for multiple testing, using the Benjamini-Hochberg approach of controlling the false-discovery rate (36).

Nucleotide sequence accession number. The draft assembly of the *E. faecalis* MMH594 genome generated by our group has been submitted to GenBank under accession number AOPW00000000.

RESULTS

***C. elegans* infection by *E. faecalis*.** To test the hypothesis that the degree of *E. faecalis* virulence in the *C. elegans* model correlates with the bacterial gene content, age-synchronized cohorts of *C. elegans* were challenged with the individual 28 *E. faecalis* isolates and the LT$_{50}$ (the time at which 50% nematode lethality is detected) was used to rank the strains from the most to the least virulent.

E. faecalis MMH594 was the most aggressive isolate in killing the nematodes, showing an LT$_{50}$ of 48 ± 3 h (Fig. 1). In contrast, ATCC 4200, EF62, T2, E1825, CH188, and Symbioflor 1 were apathogenic, as more than 85% of the *C. elegans* nematodes were alive at the end of the experiment, a percentage similar to that observed for *E. coli* OP50. Infection with *E. faecalis* V583, Merz96, D6, OG1RF, T3, E1022, HH22, Fly1, or TX0104 had a strong impact on the longevity of the nematodes, halving the *C. elegans* population in 68 ± 6 h, whereas *E. faecalis* JH-1, ARO1/DG, T8, DS5, E1Sol, HIP11704, T11, T1, and X98 exhibited intermediate
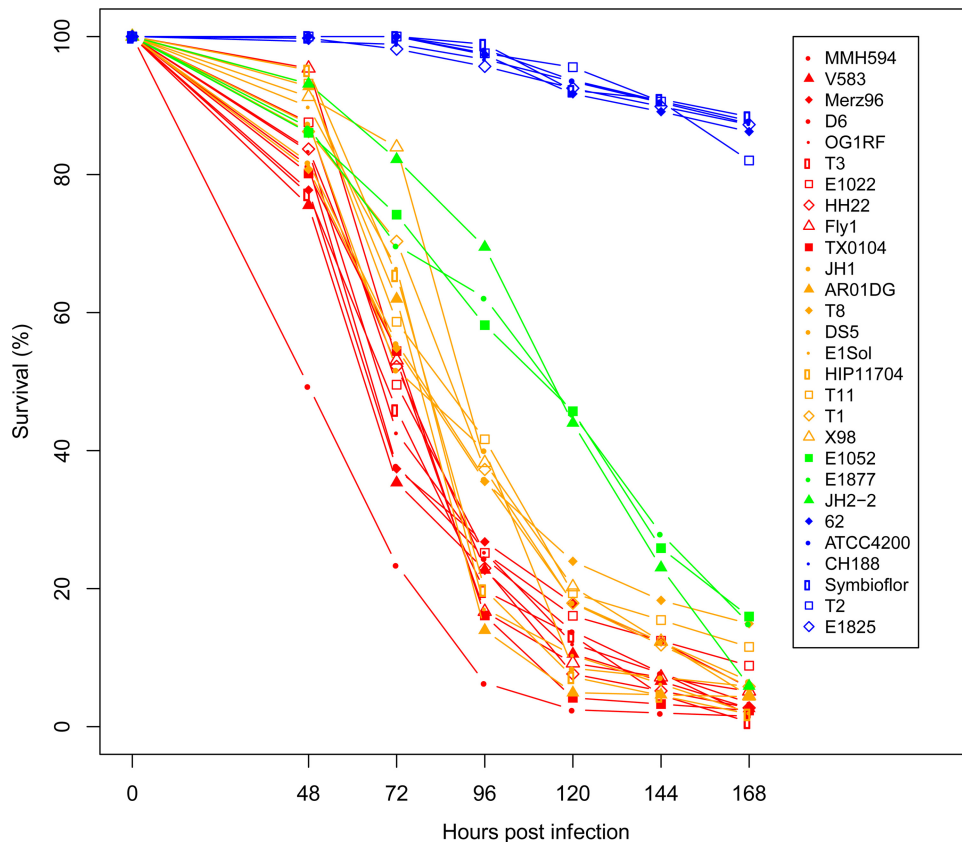
FIG 1 Survival of *C. elegans glp-4* (*bn2ts*); *sek-1* (*km4*) under conditions of feeding on *E. faecalis*. The 28 *E. faecalis* strains are ranked from the most to the least virulent according to the time required to kill 50% of the worms (LT$_{50}$). The color of the curves indicates the degree of virulence of each *E. faecalis* strain to *C. elegans*. The group designation by color has been made on the basis of the LT$_{50}$. Strains indicated in red show an LT$_{50}$ ranging from 47.4 to 72 h; strains in orange have an LT$_{50}$ ranging from 73 to 89.8 h; and strains indicated in green have an LT$_{50}$ ranging from 96 to 120 h. Strains in blue are avirulent in this infection model. The experiment was performed using three technical replicates and repeated at least two times for each strain.

killing. The slowest nematocidal activity (LT$_{50}$ >99 h) was observed for *E. faecalis* E1052, E1877, and JH2-2.

A principal component analysis (PCA) revealed that the data for all sample times were explained by two variables, where the first principal component (PC1) accounted for 94% and the second principal component (PC2) for 4% of the variability (Fig. 2). Strains with high negative values corresponding to the first component were apathogenic in the *C. elegans* model, whereas strains associated with high positive values were the most virulent (Fig. 2).

To examine whether the virulence phenotype in *C. elegans* was related to clonal lineages, the phylogenetic relationship between the *E. faecalis* strains was analyzed by constructing a core genome-based tree (see Fig. S1 in the supplemental material). However, and not surprisingly for an highly recombinogenic organism such as *E. faecalis*, only limited correlation between pathogenicity and origin could be observed, thus indicating that the clonal lineage could not be predictive of the ability of a strain to be virulent in this model.

Role of gelatinase and cytolysin pathotypes in *C. elegans* killing by *E. faecalis*. Cytolysin and gelatinase are two established enterococcal virulence factors known to play a significant role in *C. elegans* killing by *E. faecalis* (19, 20). Among the *E. faecalis* strains included in this study, we confirmed that 10/28 were cyto-

lytic on blood agar plates and that 15/28 showed a positive gelatinase phenotype, at both 37°C and 25°C. Notably, *E. faecalis* MMH594, Merz96, JH-1, and T1 were producers of both cytolysin and gelatinase, while *E. faecalis* Symbioflor 1, ATCC 4200, EF62, T2, E1825, and CH188 showed neither cytolysin nor gelatinase activity. As expected, despite the presence of the *gelE* gene, some strains (ATCC 4200, CH188, HIP11704, T2, T8, X98, JH2-2, E1052, and E1825) failed to produce gelatinase due to partial deletions of the *fsr* operon, which activates gelatinase expression (37). Strains ATCC 4200 and T2 were unable to produce cytolysin, although their genomes harbor an intact but nonfunctional *cyl* operon and its associated regulatory genes.

Figure 3 depicts the 28 *E. faecalis* strains divided in 4 subsets based on the Gel and Cyl phenotypes. The bars indicate nematode survival at 72 h. Comparing the fractions of live nematodes in each subset, it was evident that gelatinase and cytolysin are major factors of *C. elegans* killing. However, the substantial variation in survival within each subset indicated that Gel and Cyl are not the only determinants of *E. faecalis* virulence. Interestingly, all the gelatinase and cytolysin nonproducers were avirulent in *C. elegans*, with the exception of *E. faecalis* JH2-2. Observation of the pathotypes showed that the strains with a gelatinase- or cytolysin-positive phenotype were of both clinical and commensal origin.
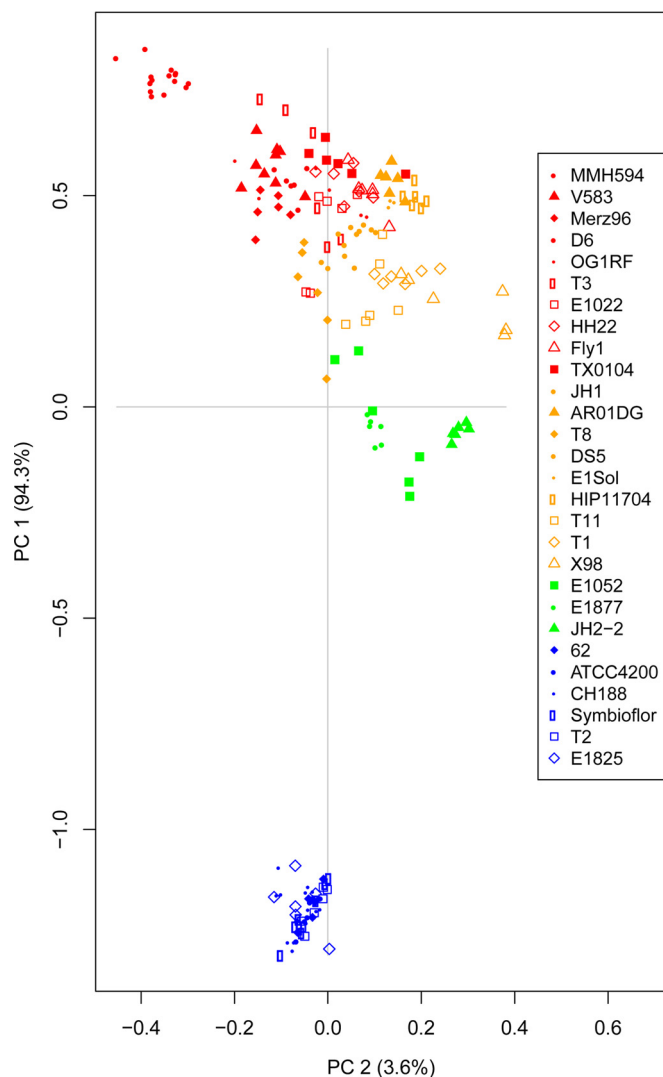
FIG 2 Principal component analysis of the observed variance in *C. elegans* survival after infection with 28 *E. faecalis* strains. The first principal component (PC1) accounts for 94.3% of the total variance, while PC2 accounts for 3.6%. The PCA result marks out four groups of strains with PC1 values ranging from highly negative (blue symbols) to highly positive (red symbols). Orange and green symbols indicate *E. faecalis* strains of moderate and low virulence, respectively.

corresponding variants of core genes) were present in all strains. The remaining sequences were organized into 2,325 protein patterns, into which proteins showing identical Present/Absent patterns across all genomes were grouped.

A Present/Absent protein pattern-based tree was constructed, and it resolved the examined strains into three lineages (Fig. 4). Highly pathogenic isolates in the *C. elegans* model (MMH594, V583, HH22, and TX0104), all belonging to the nosocomial clonal complex 2 (CC2) group, occupy a separate lineage (lineage A) in relation to the other strains employed in the study. Lineage C, although it includes commensal isolates EF62 and E1052, consists of clinical isolates (HIP11704, CH188, JH2-2, T8, ATCC 4200, and T2) exhibiting phenotypes ranging from moderately virulent to apathogenic. Lineage B consisted of two subgroups. In subgroup I, the majority of the strains are highly or moderately lethal nematode killers, with the exception of strain E1877, which is a strain of low pathogenicity. The two isolates Symbioflor1 and commensal E1825 form subgroup II in lineage B and have no nematocidal effect. Taken together, the data show a significant correlation between Present/Absent protein profiles and the *E. faecalis* phenotype of virulence toward *C. elegans*.

Next, to evaluate the contribution of cytolysin and gelatinase to *E. faecalis* pathogenicity, each of the 28 strains was coded by its Gel and Cyl phenotype (positive or negative), and an analysis of variance using a mixed-effect model was conducted as described for equation 1. This analysis revealed a strong ($P < 0.0001$) negative impact of Gel and Cyl that accounted for 40.8% and 36.5%, respectively, of the observed variations in survival of the nematodes. The interaction was not synergistic but showed that a Gel$^+$ Cyl$^+$ phenotype is (significantly) 26.7% less virulent than the sum of their separate contributions, i.e., that there is saturation or an antagonistic interaction between these two virulence traits.

Next, we fitted each of the 2,325 distinct Present/Absent patterns to worm survival, as described in equation 2, resulting in the identification of 414 patterns statistically associated with *C. elegans* killing (see Table S3 in the supplemental material). In addition, we performed a similar ANOVA considering other factors. Strains were grouped by capsular polysaccharide locus serotype, human or nonhuman origin, clinical or commensal origin, and gastrointestinal or urinary tract origin, but none of these showed a significant association with nematode survival. Similarly, the year of isolation did not reveal any significant impact on virulence. Clinical isolates from blood produced 27% higher killing ($P < 0.001$) of the worms than those from other sources.

Having observed that gelatinase and cytolysin have a strong and significant negative impact on nematode survival, the ANOVA was then extended as described in equation 3 to include both the two pathogenicity traits and the effect of Present/Absent patterns. This resulted in 376 patterns positively correlated with virulence (see Table S4 in the supplemental material); i.e., these patterns are capable of explaining significant additional variations in survival even when Gel and Cyl are also included in the model.

After a process of binning and manual curation, 61 candidate patterns remained that included 168 protein-coding sequences with a potential role during nematode infection (see Table S5 in the supplemental material). On the basis of these results, we found that 86% (52/61) of the patterns included annotated CDS identified in *E. faecalis* V583 and, importantly, that a total of 25 were associated with previously described mobile genetic elements (MGEs). These included genes that cover only parts of the MGEs

This indicates that nonclinical cytolysin- or gelatinase-producing bacteria can also be potentially infectious in this model.

Construction of a genomic reference map of *E. faecalis* pathogenicity in *C. elegans*. Aiming to elucidate the genetic basis of the nematocidal activity of the individual isolates and the observed virulence variation, a bimodal mathematical and bioinformatics analysis approach was employed. A total of 79,108 annotated protein sequences for 27 sequenced strains were downloaded from the RefSeq database in addition to 3,248 proteins translated from the identified coding sequences (CDS) in the MMH594 genome. After all-against-all BLAST analysis performed as described in Materials and Methods, every sequence was given a Present/Absent pattern over the 28 genomes using a cutoff value of a 0.75 relative BLAST score. This revealed that 48,821 of the unique annotated protein sequences (encoded by the
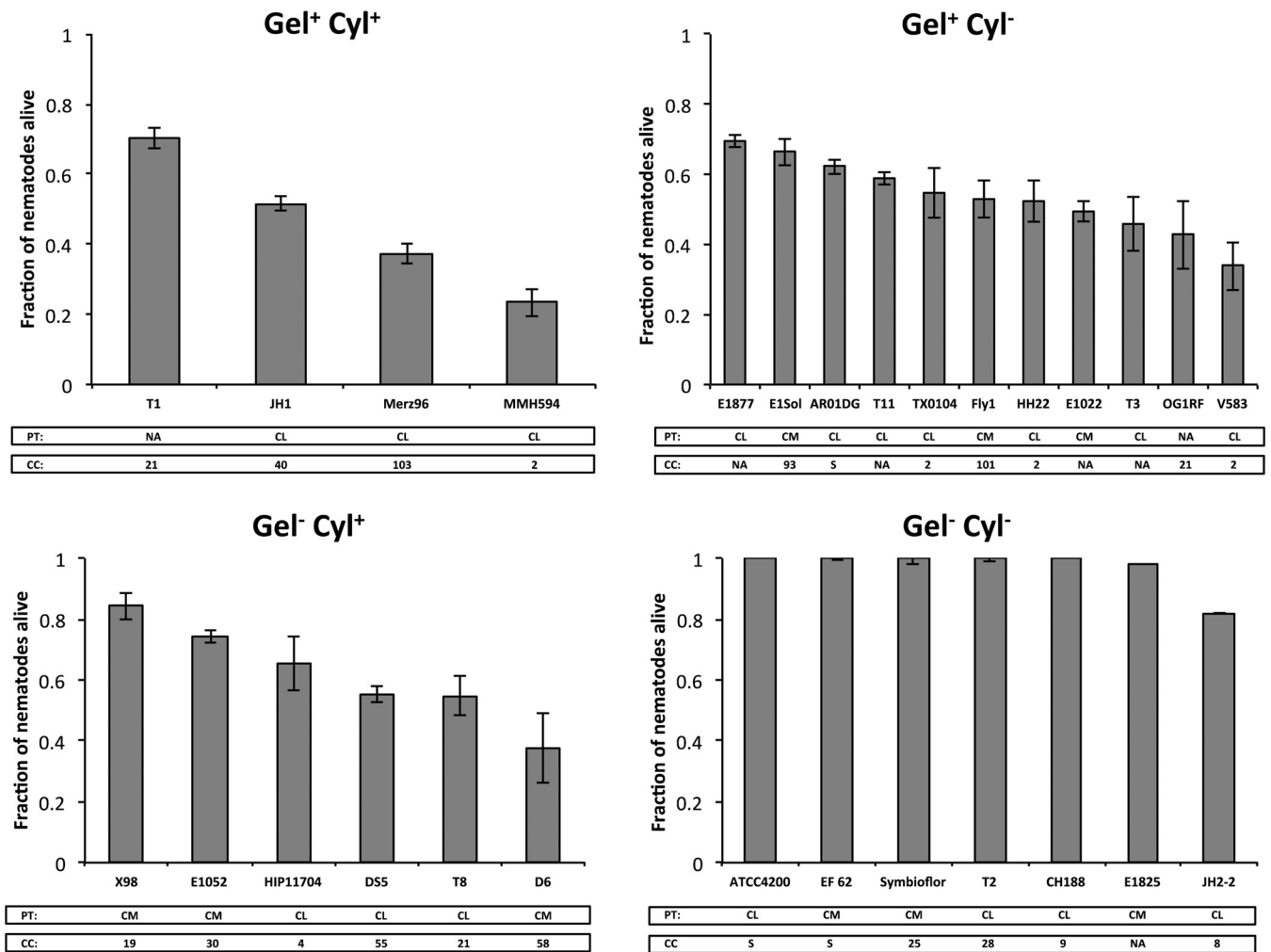
FIG 3 *C. elegans* survival at 72 h postinfection after ingestion of *E. faecalis*. The 4 subsets are based on the gelatinase (Gel) and cytolysin (Cyl) phenotypes of *E. faecalis*. +, positive phenotype; −, negative phenotype. The values represent averages of the results of two independent experiments, each comprised of three technical replicates. Standard deviations are indicated with error bars. PT, pathotype; CL, clinical; CM, commensal; NA, not available; CC, clonal complex; S, singleton.

(phage01 [$n = 5$], phage04 [$n = 11$], phage05 [$n = 10$], phage06 [$n = 6$], phage07 [$n = 1$], *E. faecalis* pathogenicity island [PAI] [$n = 3$], *efaC1* [$n = 10$], and the *vanB* vancomycin resistance region [$n = 1$]), with the exception of phage03, which was almost completely represented ($n = 54$). In addition to the MGEs, 9 patterns included a putative phage element corresponding to the gene cluster *ef3217* to *ef3227* and genes for surface structural proteins previously described to be enriched in isolates of high-risk clonal complex 2 (17).

Functional genomic analysis of pathogenesis in noncytolytic and gelatinase-negative strain JH2-2 identifies involvement of phage03 in virulence. As shown in Fig. 3, JH2-2 was the only cytolysin- and gelatinase-negative isolate that exhibited significant nematocidal activity. It was therefore of interest to identify genes that promoted *E. faecalis* JH2-2 host fitness and pathogenicity in *C. elegans*. This strain is a plasmid-free derivative of the JH2 clinical isolate (38), and its virulence, relative to that of the JH2 isolate, in the intravenous mouse infection model has been shown in a previous report (39). However, the role of genetic traits in-

volved in the development of disease has not been further elucidated.

To identify potential virulence factor-coding genes, we performed an *in silico* subtractive analysis of the genome of *E. faecalis* JH2-2 (GCA_000148445.1) against the genomes of strains showing an avirulent phenotype in the nematode model, namely, *E. faecalis* 62, *E. faecalis* CH188, *E. faecalis* ATCC 4200, *E. faecalis* T2, and *E. faecalis* E1825. The genome of *E. faecalis* V583 was used as the reference. Moreover, to rule out possible intrinsic errors in gene annotations and predictions present in the publicly available *E. faecalis* draft genomes, the search output was manually examined using BLAST. Besides, we employed Mauve 3.2.1 software to perform global alignment of the same genomes (34). These combined approaches identified 135 genes present in JH2-2 but absent from the apathogenic strains (see Table S6 in the supplemental material). A large fraction of these genes encode MGEs with high sequence homology to that identified in the *E. faecalis* V583 genome (40).

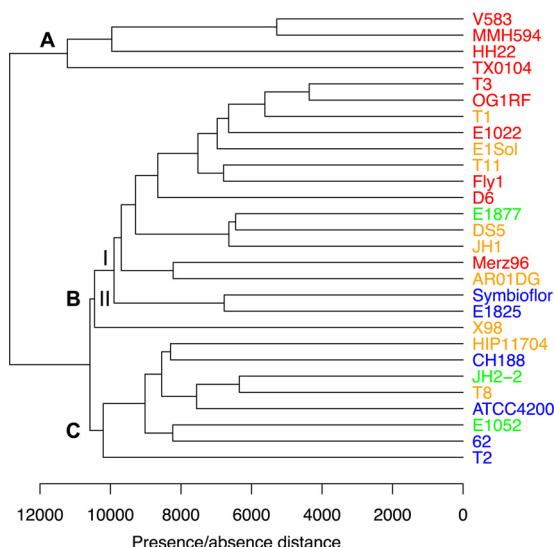Among the gene clusters identified by comparative genomics

FIG 4 Presence/absence gene pattern-based tree of the 28 *E. faecalis* isolates employed in this study. The pathogenicity of *E. faecalis* to *C. elegans* is indicated in colors, with the highly pathogenic strains in red, moderately pathogenic strains in orange, minimally pathogenic strains in green, and apathogenic strains in blue. Lineages are indicated with letters A, B, and C. Two subgroups in lineage B are indicated with I and II.

analysis, the phage03-like element in JH2-2 was among the patterns with the lowest *t* values in the ANOVA, suggesting a potential involvement in virulence. This candidate pathogenicity trait was therefore selected for further functional studies.

We designed experiments to assess whether the presence of the intact phage03-like element affects *E. faecalis* JH2-2 pathogenicity in *C. elegans*. Markerless in-frame deletion of the phage03-like region from *E. faecalis* JH2-2 was achieved by allelic replacement using the pSL100 vector (see Table S2 in the supplemental material). We compared the levels of growth in GM17 broth of the phage03-like deletion mutant, named SL100, and of the *E. faecalis* JH2-2 wild-type strain. As shown in Fig. 5a, the growth yield of SL100 was significantly higher than that of *E. faecalis* JH2-2, with increased cell density in the stationary phase of growth. This suggests that the reduced *E. faecalis* JH2-2 growth may have been due

to the metabolic burden imposed by the presence of the phage03-like DNA. Subsequently, we evaluated the contribution of the phage03-like element to nematocidal activity. SL100 was found to be severely attenuated in nematode killing compared to *E. faecalis* JH2-2 at each sampling point ($P < 0.05$, unpaired *t* test), and the results showed 76.5% increased *C. elegans* survival at the end of the experiment (Fig. 5b). Similarly, infection of *Galleria mellonella* larvae with approximately 5×10^6 CFU/ml of SL100 led to significantly ($P = 0.02$, unpaired *t* test) increased survival compared to the JH2-2 wild-type strain results (see Fig. S2 in the supplemental material). Removal of phage03 from JH2-2 resulted in a 30% reduction in killing of the larvae between 12 and 20 h postinfection. The reduced life span of the nematodes and wax moth larvae infected with JH2-2 compared to SL100 indicates that the phage03-like element contributes to increased *E. faecalis* virulence, revealing a role of the phage in the development of infection.

Phage03-like elements are widely distributed among nosocomial isolates. Enterococcal phages contribute to bacterial evolution by facilitating the spread of virulence and antimicrobial resistance determinants via transduction (41). Paulsen and coworkers have shown that approximately 25% of the total genome of *E. faecalis* V583 consists of MGEs and seven integrated prophage regions (42). However, only a few studies have investigated how prophage-carried genes can engender virulence in *E. faecalis* (41, 43, 44).

The sequence of the phage03-like element in *E. faecalis* JH2-2 (HMPREF9496_01004 to HMPREF9496_01061 and HMPREF9496_01491 to HMPREF9496_01507) has a length of approximately 48 kb and closely resembles the region spanning *ef1417* to *ef1489* in the genome of *E. faecalis* V583. This region contains 71 coding sequences with 81% similarity to the region in *E. faecalis* V583 and an average G+C content of 35% (see Table S7 in the supplemental material). A phage03-like element is also found in the *E. faecalis* MMH594 and HH22 CC2 strains and in the *E. faecalis* Fly1 animal isolate, among the strains used in this study. ANOVA showed that the phage03-containing strains displayed on average 17% higher killing toward *C. elegans* ($P = 4.4e^{-6}$).

Among the 71 genes, only 16 have an assigned function based on homology to genes encoding annotated proteins whereas the rest share no similarity to those encoding known proteins. The
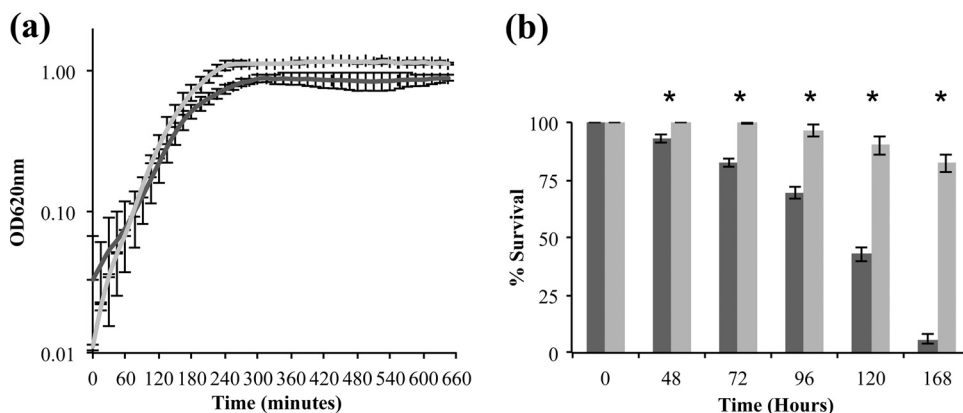

FIG 5 Effect of phage03 deletion on the growth and virulence of *E. faecalis* JH2-2. (a) Growth curves of *E. faecalis* JH2-2 (dark gray line) and SL100 (light gray line) in GM17. (b) *C. elegans* survival after ingestion of *E. faecalis* JH2-2 (dark gray bars) and SL100 (light gray bars). The values represent averages of the results of two independent experiments ± standard deviations (SD). The mean values were compared by unpaired Student *t* test (*, $P < 0.05$).
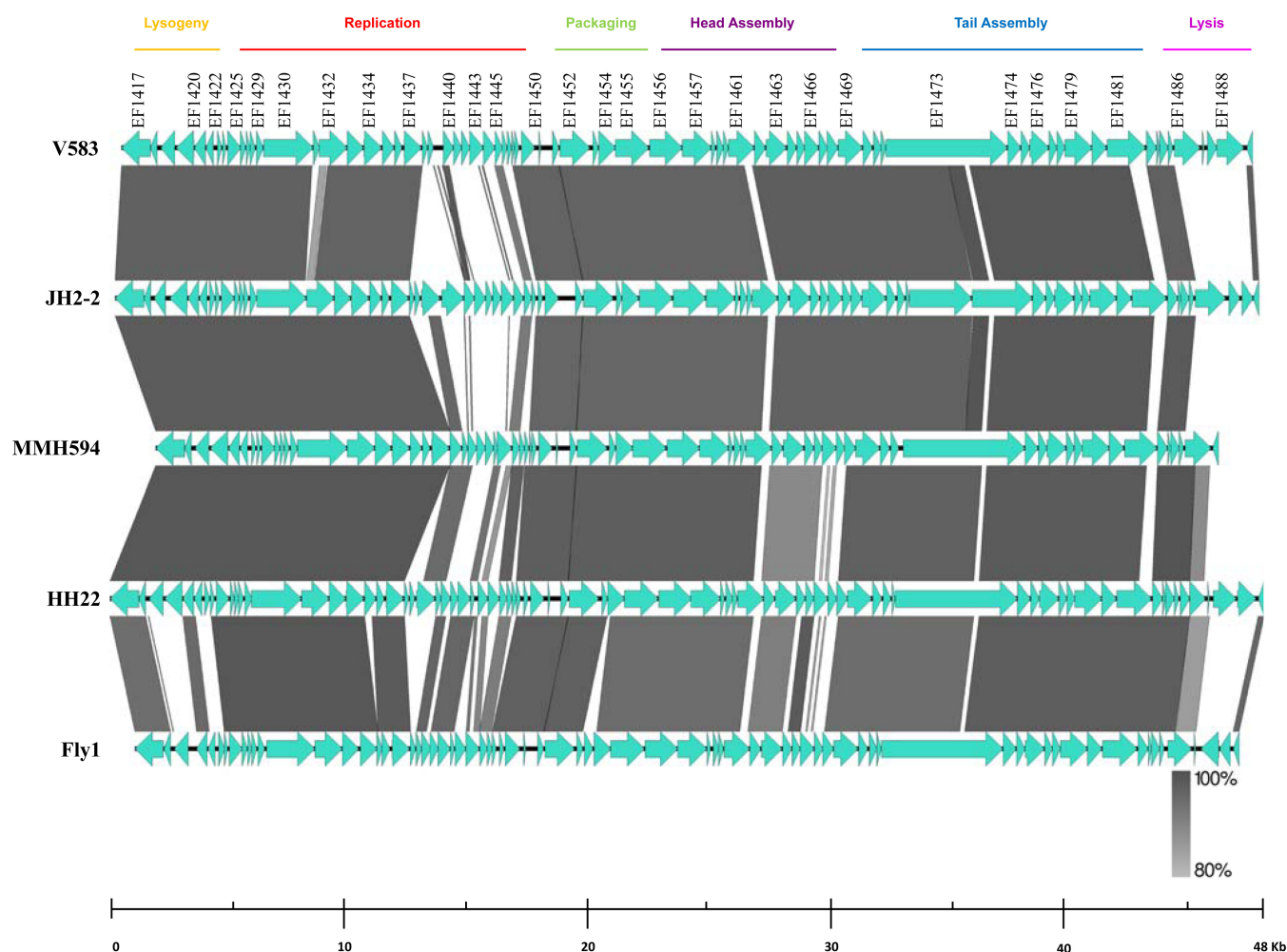
FIG 6 Comparative alignment of the phage03 elements in *E. faecalis* V583, JH2-2, MMH594, HH22, and Fly1. A BLASTX comparison was performed using EasyFig software version 2.1. Similarity of between 100% and 80% is shown in the gray gradient scale. Genes are indicated by arrows and grouped according to putative function.

phage03 element is flanked by a 76-bp inverted repeat, indicating the presence of *attL* and *attR* attachment sites (45). No phage particles could be induced by mitomycin C treatment (data not shown). The inability to detect phage particles is in accordance with previous observations that found only low production of the phage particles following mitomycin C treatment (41).

Comparative analysis of the phage integration regions in the five *E. faecalis* strains revealed different integration sites, suggesting independent acquisitions of phage DNA by different *E. faecalis* strains. The phage03 element sequences in *E. faecalis* MMH594, V583, HH22, JH2-2, and Fly1 reveal that, consistent with other temperate bacteriophages found in lactic acid bacteria, the genes are organized to form functional modules for lysogeny, DNA replication, packaging, head and tail assembly, and host lysis and that they are regulated in a coordinated manner (Fig. 6) (46). The sequences share high similarity, with the exception of the regions containing genes encoding proteins with a predicted function in virion DNA replication and host lysis.

In addition to *E. faecalis* JH2-2, MMH594, V583, HH22, JH2-2, and Fly1, BLAST searches against NCBI available genomes showed the presence of a phage03-like element in a total of 93

enterococcal isolates, of which 73 were of confirmed nosocomial origin (see Table S8 in the supplemental material).

Deconstruction of *E. faecalis* V583 virulence. Aiming to corroborate the significance of phage03 elements for the pathogenicity of *E. faecalis* in *C. elegans*, we analyzed a series of isogenic mutants of *E. faecalis* V583 (Fig. 7). The selected strain belongs to the high-risk clonal complex 2 group of strains that are prevalent in nosocomial infections (47).

As a first step, we tested the hypothesis that plasmids contribute to virulence in V583. Infection of *C. elegans* was performed by employing the *E. faecalis* V19 plasmid-cured derivative (48), and it revealed that extrachromosomal DNA found in V583 does not contribute to pathogenicity in this model ($P = 0.09$, unpaired t test). Next, we evaluated the impact of the gelatinase production in the killing of nematodes by employing the V583Δ*gelE* mutant. As depicted in Fig. 7, deletion of a gelatinase gene reduced virulence at 72 h postinfection by 40% ($P < 0.05$, unpaired t test), which is consistent with previous results (20). Then, to assess the contribution from the Fsr master regulatory system, a frameshifted *fsrB* mutant, named SL111, was constructed by allelic replacement. This strain produced 56% less killing than the parent
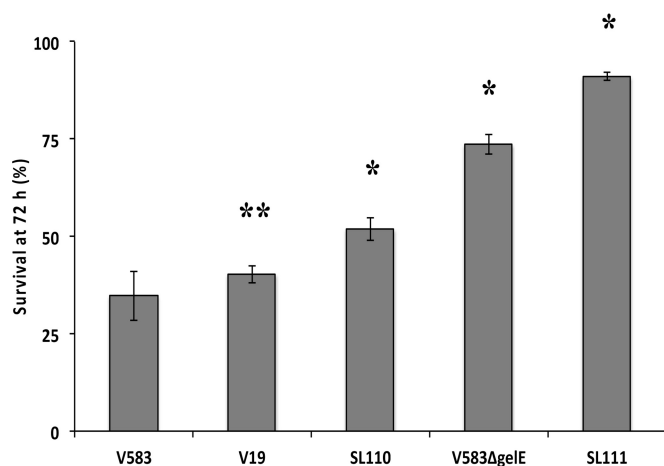
FIG 7 Impact of isogenic deletions on *E. faecalis* V583 virulence to *C. elegans*. The bars indicate the percentages of nematode survival at 72 h postinfection. Data are expressed as averages ± SD. The experiments were performed at least twice using three technical replicates. The mean values were compared by unpaired Student *t* test (*, $P < 0.05$; **, $P > 0.05$).

strain ($P < 0.05$, unpaired *t* test). We therefore hypothesized that the remaining virulence potential could be partially attributed to the phage03 element. An isogenic in-frame phage03 deletion mutant of *E. faecalis* V583 named SL110 was constructed by employing the pSL100 vector system (see Table S2 in the supplemental material) as described above for the analogous element in *E. faecalis* JH2-2. Infecting *C. elegans* with *E. faecalis* SL110, we observed a 17% reduction in the killing compared to the wild-type strain results (Fig. 7). SL110 was still capable of efficiently killing the nematode, since it harbored a functional *fsrABDC* operon and the *gelE* gene; however, the deletion of the phage03 element led to decreased killing ability by SL110 after 48 and 72 postinfection ($P = 0.0003$, unpaired *t* test). Taken together with the observation in JH2-2, these results provide evidence that the phage03 element acts as a virulence factor toward *C. elegans*.

DISCUSSION

The emergence of *E. faecalis* strains as prominent nosocomial pathogens and the occurrence of enterococcal infections not treatable with most available antibiotics emphasize the need for understanding the complex mechanism by which this organism causes life-threatening diseases (49). The availability of large amounts of whole-genome sequencing data has facilitated *in silico* comparative analyses of the gene content of enterococcal strains representing a wide spectrum of biotypes, clonal lineages, and isolation sources. As result, these investigations have contributed to the determination of the size and composition of the *E. faecalis* core genome (40, 50), the genetic background for antibiotic resistance (51), and the identification of potential virulence factors that can influence the bacterial lifestyle (16). However, there is still a lack of understanding of how the genetic variation between commensal and nosocomial strains could explain the alleged behavior associated with these.

In the present study, we combined statistical and comparative genomic approaches, followed by a functional genomic study, with a model host infection system to examine the correlation between *E. faecalis* whole-genome content and its impact in the host-microbe interaction. We showed that infection of *C. elegans*

by 28 *E. faecalis* isolates resulted in a wide range of virulence phenotypes, with six strains behaving in a commensal manner, with no nematocidal effect, and 22 strains displaying a virulence phenotype of from high to low levels of pathogenicity. Principal component analysis showed that the first and second principal components account for 98% of the variance in the virulence data, thus indicating a significant relationship between *E. faecalis* pathogenicity toward *C. elegans* and gene content. We therefore employed a comparative approach to analyze the genomes of the 28 *E. faecalis* strains to identify variations in the genetic makeup that could explain commensalism or opportunistic lifestyle in the nematode. A tree based on Present/Absent gene patterns identified by the comparative genomic analysis resolved the 28 strains into three lineages and showed a significant correlation between Present/Absent profiles and the phenotype of *E. faecalis* virulence toward *C. elegans* (Fig. 4). Moreover, the tree highlighted a deep division between isolates belonging to CC2 and the other strains employed in the study.

Statistical ANOVA established a strong contribution of cytolysin and gelatinase to nematode killing which accounted for 40.8% and 36.5%, respectively, of the observed variations in survival of the nematodes. However, the presence of the two factors explained only 50.6% of the variation. Speculatively, it is conceivable that the lack of synergism between gelatinase and cytolysin is caused by saturation or antagonistic interactions between the two factors when coproduced by a strain.

In agreement with previous studies, evaluation by ANOVA of the *E. faecalis* isolation source for the association with increased virulence toward *C. elegans* indicated that the isolate origin had no impact on virulence in the nematodes (52). The fact that non-clinically isolated strains also had the potential to be infectious (Fig. 3) is consistent with the ambivalent nature of *E. faecalis* in relation to its host, with the ability of apparently harmless strains to disrupt the commensal relationship with the host and cause severe diseases.

In addition, in line with a previous investigation that reported a low correlation between the presence of capsule among *E. faecalis* clinical isolates and the ability to establish infection (53), our data showed that the distributions of cell wall polysaccharides are equal at all levels of *E. faecalis* virulence, and no significant difference in *C. elegans* pathogenicity was observed.

Beside other reports that have clearly demonstrated that gelatinase and cytolysin are two prominent *E. faecalis* virulence factors exerting a crucial role in nematode death (19, 20), our study also identified a strong positive impact of these two factors on nematode killing. The mathematical model, computed to include the *a priori* knowledge of the cytolysin and gelatinase genotype and to fit to the virulence at 72 h, established a virulence reference map of statistically significant traits that can be predictive of *E. faecalis* virulence (see Table S5 in the supplemental material). Among the resulting patterns, a predominant (41%) fraction is comprised of MGEs. The presence of a multiplicity of mobile elements has been previously described for *E. faecalis* V583 and is considered a major driving force of genome evolution (42, 44, 50). In addition, a study from our group reported the enrichment of structures of MGEs among strains belonging to the high-risk nosocomial clonal complex 2 (CC2) group (17). The prevalence of MGEs among factors contributing to *E. faecalis* pathogenicity supports the notion that highly virulent isolates benefit from the introduction of fitness

and virulence factors that confer increased adaption under conditions of stress and infection.

Patterns 1408 and 1523, including CDS EF1827 to EF1830 and EF1833, emerged in our work as contributing to virulence in nematodes. The associated genes reside in a 23.9-kb region corresponding to the 5′-end fragment of *fsrC* (*ef1820*) and the 3′-end segment of *ef1841* of *E. faecalis* V583 whose deletion has been reported to cause the lack of a positive gelatinase phenotype among strains carrying the *gelE* gene (54). Interestingly, the genes encode sugar transport system components related to the *glpK* pathway and are involved in glycerol catabolism in *E. faecalis* (55). As described in other bacterial species, growth in the presence of glycerol as the carbon source leads to the activation of a transcription activator also modulating the expression of virulence genes (56).

The list of candidate genes that might contribute to enterococcal pathogenicity includes variable cell surface structure components, such as EF2164 (pattern 666), EF0967 (pattern 609), and EF3250 and EF3251 (pattern 1577), previously identified as enriched among isolates of the high-risk clonal complex 2 (17). Surface-exposed proteins play an important role in pathogenicity, as they are often involved in colonization of host tissues and evasion of host defenses (57, 58). Studies in *C. elegans* have ruled out the role of a number of surface structures, such as the AS aggregation substance, the Esp enterococcal surface protein, the Ace adhesin to collagen, and biofilm-associated pilus protein A, in mediating nematode killing (21, 59), suggesting that their host target may be absent. However, Creti and colleagues demonstrated that deletion of the *ef3314* gene, encoding a surface protein structurally similar to Esp, affects the pathogenicity potential of the resulting *E. faecalis* mutant in *C. elegans*, thus demonstrating that some surface factors contribute to pathogenesis in the worm (22). Further functional studies are therefore required to elucidate the significance of the identified traits in enterococcal virulence.

Among the non-V583 annotated loci, pattern 1275 included two proteins holding a CaaX amino terminus and a DJ-1/PfpI domain, respectively. These domains are common to proteases; therefore, a potential role of these genes in virulence is plausible (60).

Polylysogeny is a common feature in the *E. faecalis* genome and is considered to influence the relationship of *E. faecalis* with its host, as these elements may confer properties advantageous for the dissemination of the bacterial strain (17, 41). Duerkop and colleagues demonstrated that temperate phages of *E. faecalis* ST6 can modulate the assemblies of bacterial communities in the mammalian intestinal tract (61). A more recent study showed that five (phage01, phage03, phage04, phage05, and phage07) of the seven V583 prophages are capable of excision from the bacterial chromosome and that four of them produced infective virions and are therefore possible mediators of horizontal gene transfer (44). Bacteriophage-encoded proteins such as platelet binding proteins were established as factors that provide mechanisms to invade host tissues and damage host cells, enriching the plethora of *E. faecalis* virulence traits (43, 44).

Our comparative genomic approach pointed toward a significant enrichment of a phage03-like element in nosocomial and clinical isolates. We demonstrated that phage03 deletion resulted in a decline of *E. faecalis* JH2-2 infectivity both in *C. elegans* and in *G. mellonella* larvae. In agreement with previous reports that have described the absence of phage03 in the genome of *E. faecalis* food

isolates (45) and its enrichment among clonal lineages associated with nosocomial diseases (17, 40), we detected the presence of phage03 in the genome of CC2 strains *E. faecalis* V583, MMH594, and HH22. BLAST searches against the available *E. faecalis* genomes showed that this phage03-like element was also found with a higher incidence in nosocomial isolates whereas it is rarely found in other strains of different origins (see Table S8 in the supplemental material). We therefore examined the role of the phage in the pathogenicity of nosocomial isolates by assessing the virulence of an isogenic phage03 mutant of *E. faecalis* V583. The mutant displayed attenuation in nematode killing as well as in the *G. mellonella* model compared to the parent strain, providing additional evidence that phage03 plays a role in *E. faecalis* virulence. To our knowledge, this is the first study to have demonstrated that a phage element endows clinical and nosocomial isolates with factors promoting increased pathogenicity *in vivo* during infection in a model organism.

Though the mechanism by which phage03 impacts *E. faecalis* virulence remains to be elucidated, the ability of this element to form infective virions indicates that its genes are expressed and that it can act as an efficient vehicle for horizontal gene transfer, spreading traits contributing to boosting strain fitness (44). Excision of the phage has been associated with treatment with fluoroquinolones (44), supporting the idea that application of antibiotics in the nosocomial environment may facilitate the creation of conditions favorable for the emergence of highly pathogenic clonal lineages.

Analysis of the phage03 sequence revealed the presence of proteins that could play a role in bacterial virulence. *ef1420* encodes a protein of unknown function previously reported to significantly decrease *E. faecalis* infectivity in the *G. mellonella* model (62). This protein was predicted to contain a lipobox motif and is expected to be a surface-exposed element (45, 63). Lipoproteins represent approximately 25% of the proteins predicted to be associated with *E. faecalis* cell envelope, and they are reported to be involved in substrate binding and in delivery to ABC transporters, acquisition of sugars, protein folding, antibiotic resistance, and cell envelope stability (64, 65). In addition, lipoproteins play a significant role in virulence, as they mediate adhesion to the extracellular matrix (ECM), initiation of the inflammatory process followed by activation of the immune system, induction of phagosome escape, and translocation of virulence factors (63).

Besides EF1420, two proteins (EF1460 and EF1474) contain a LysM domain in the C terminus thought to contribute to the anchoring to peptidoglycan (66) whereas EF1426 is a VirE domain-containing protein conceivably associated with virulence by exerting a role in the type IV secretion pathway in other organisms (67). It is noteworthy that many phage03 proteins have an unknown function but nevertheless might play a role in and contribute to *E. faecalis* pathogenicity.

In conclusion, using a statistical and comparative genomic approach combined with an *in vivo* assessment of *E. faecalis* virulence during infection in *C. elegans*, our results establish a set of nonorthologous gene patterns that may be predictive of *E. faecalis* pathogenicity. Our report highlights the importance of MGEs as potential contributors to enterococcal pathogenicity and demonstrates that phage03, a prevalent element among nosocomial clonal lineages, is responsible for enhanced virulence. Future functional investigations are required to validate any involvement of the identified patterns in *E. faecalis* virulence and to gain a

deeper understanding of the genetic mechanism underlying enterococcal pathogenesis.

REFERENCES

1. **Noble CJ.** 1978. Carriage of group D streptococci in the human bowel. J Clin Pathol **31:**1182–1186. http://dx.doi.org/10.1136/jcp.31.12.1182.
2. **Richards MJ, Edwards JR, Culver DH, Gaynes RP.** 2000. Nosocomial infections in combined medical-surgical intensive care units in the United States. Infect Control Hosp Epidemiol **21:**510–515. http://dx.doi.org/10.1086/501795.
3. **Sghir A, Gramet G, Suau A, Rochet V, Pochart P, Dore J.** 2000. Quantification of bacterial groups within human fecal flora by oligonucleotide probe hybridization. Appl Environ Microbiol **66:**2263–2266. http://dx.doi.org/10.1128/AEM.66.5.2263-2266.2000.
4. **Huycke MM, Sahm DF, Gilmore MS.** 1998. Multiple-drug resistant enterococci: the nature of the problem and an agenda for the future. Emerg Infect Dis **4:**239–249. http://dx.doi.org/10.3201/eid0402.980211.
5. **Huycke MM.** 2002. Physiology of enterococci, p 133–175. *In* Gilmore MS, Clewell DB, Courvalin P, Dunny GM, Murray BE, Rice LB (ed), The enterococci: pathogenesis, molecular biology, and antibiotic resistance. ASM Press, Washington, DC.
6. **Flahaut S, Laplace JM, Frere J, Auffray Y.** 1998. The oxidative stress response in *Enterococcus faecalis*: relationship between H_2O_2 tolerance and H_2O_2 stress proteins. Lett Appl Microbiol **26:**259–264. http://dx.doi.org/10.1046/j.1472-765X.1998.00325.x.
7. **Kearns AM, Freeman R, Lightfoot NF.** 1995. Nosocomial enterococci: resistance to heat and sodium hypochlorite. J Hosp Infect **30:**193–199.
8. **Gastmeier P, Schwab F, Barwolff S, Ruden H, Grundmann H.** 2006. Correlation between the genetic diversity of nosocomial pathogens and their survival time in intensive care units. J Hosp Infect **62:**181–186. http://dx.doi.org/10.1016/j.jhin.2005.08.010.
9. **Solheim M, La Rosa SL, Mathisen T, Snipen LG, Nes IF, Brede DA.** 2014. Transcriptomic and functional analysis of NaCl-induced stress in *Enterococcus faecalis*. PLoS One **9:**e94571. http://dx.doi.org/10.1371/journal.pone.0094571.
10. **Ike Y, Clewell DB, Segarra RA, Gilmore MS.** 1990. Genetic analysis of the pAD1 hemolysin/bacteriocin determinant in *Enterococcus faecalis*: Tn917 insertional mutagenesis and cloning. J Bacteriol **172:**155–163.
11. **Dunny GM, Clewell DB.** 1975. Transmissible toxin (hemolysin) plasmid in *Streptococcus faecalis* and its mobilization of a noninfectious drug resistance plasmid. J Bacteriol **124:**784–790.
12. **Weigel LM, Clewell DB, Gill SR, Clark NC, McDougal LK, Flannagan SE, Kolonay JF, Shetty J, Killgore GE, Tenover FC.** 2003. Genetic analysis of a high-level vancomycin-resistant isolate of *Staphylococcus aureus*. Science **302:**1569–1571. http://dx.doi.org/10.1126/science.1090956.
13. **Coburn PS, Baghdayan AS, Dolan GT, Shankar N.** 2007. Horizontal transfer of virulence genes encoded on the *Enterococcus faecalis* pathogenicity island. Mol Microbiol **63:**530–544. http://dx.doi.org/10.1111/j.1365-2958.2006.05520.x.
14. **Dunny GM.** 2013. Enterococcal sex pheromones: signaling, social behavior, and evolution. Annu Rev Genet **47:**457–482. http://dx.doi.org/10.1146/annurev-genet-111212-133449.
15. **Sava IG, Heikens E, Huebner J.** 2010. Pathogenesis and immunity in enterococcal infections. Clin Microbiol Infect **16:**533–540. http://dx.doi.org/10.1111/j.1469-0691.2010.03213.x.
16. **Palmer KL, Godfrey P, Griggs A, Kos VN, Zucker J, Desjardins C, Cerqueira G, Gevers D, Walker S, Wortman J, Feldgarden M, Haas B, Birren B, Gilmore MS.** 2012. Comparative genomics of enterococci: variation in *Enterococcus faecalis*, clade structure in *E. faecium*, and defining characteristics of *E. gallinarum* and *E. casseliflavus*. mBio **3:**e00318–00311.
17. **Solheim M, Brekke MC, Snipen LG, Willems RJ, Nes IF, Brede DA.** 2011. 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 Microbiol **11:**3. http://dx.doi.org/10.1186/1471-2180-11-3.
18. **Vebo HC, Solheim M, Snipen L, Nes IF, Brede DA.** 2010. Comparative genomic analysis of pathogenic and probiotic *Enterococcus faecalis* isolates, and their transcriptional responses to growth in human urine. PLoS One **5:**e12489. http://dx.doi.org/10.1371/journal.pone.0012489.
19. **Garsin DA, Sifri CD, Mylonakis E, Qin X, Singh KV, Murray BE, Calderwood SB, Ausubel FM.** 2001. A simple model host for identifying Gram-positive virulence factors. Proc Natl Acad Sci U S A **98:**10892–10897. http://dx.doi.org/10.1073/pnas.191378698.
20. **Sifri CD, Mylonakis E, Singh KV, Qin X, Garsin DA, Murray BE, Ausubel FM, Calderwood SB.** 2002. Virulence effect of Enterococcus faecalis protease genes and the quorum-sensing locus *fsr* in *Caenorhabditis elegans* and mice. Infect Immun **70:**5647–5650. http://dx.doi.org/10.1128/IAI.70.10.5647-5650.2002.
21. **Maadani A, Fox KA, Mylonakis E, Garsin DA.** 2007. *Enterococcus faecalis* mutations affecting virulence in the *Caenorhabditis elegans* model host. Infect Immun **75:**2634–2637. http://dx.doi.org/10.1128/IAI.01372-06.
22. **Creti R, Fabretti F, Koch S, Huebner J, Garsin DA, Baldassarri L, Montanaro L, Arciola CR.** 2009. Surface protein EF3314 contributes to virulence properties of *Enterococcus faecalis*. Int J Artif Organs **32:**611–620.
23. **Sifri CD, Begun J, Ausubel FM.** 2005. The worm has turned–microbial virulence modeled in *Caenorhabditis elegans*. Trends Microbiol **13:**119–127. http://dx.doi.org/10.1016/j.tim.2005.01.003.
24. **Booth MC, Bogie CP, Sahl HG, Siezen RJ, Hatter KL, Gilmore MS.** 1996. Structural analysis and proteolytic activation of *Enterococcus faecalis* cytolysin, a novel lantibiotic. Mol Microbiol **21:**1175–1184. http://dx.doi.org/10.1046/j.1365-2958.1996.831449.x.
25. **Qin X, Singh KV, Weinstock GM, Murray BE.** 2000. Effects of Enterococcus faecalis *fsr* genes on production of gelatinase and a serine protease and virulence. Infect Immun **68:**2579–2586. http://dx.doi.org/10.1128/IAI.68.5.2579-2586.2000.
26. **Moy TI, Ball AR, Anklesaria Z, Casadei G, Lewis K, Ausubel FM.** 2006. Identification of novel antimicrobials using a live-animal infection model. Proc Natl Acad Sci U S A **103:**10414–10419. http://dx.doi.org/10.1073/pnas.0604055103.
27. **Brenner S.** 1974. The genetics of Caenorhabditis elegans. Genetics **77:**71–94.
28. **Nakayama J, Chen S, Oyama N, Nishiguchi K, Azab EA, Tanaka E, Kariyama R, Sonomoto K.** 2006. Revised model for *Enterococcus faecalis fsr* quorum-sensing system: the small open reading frame *fsrD* encodes the gelatinase biosynthesis-activating pheromone propeptide corresponding to staphylococcal *agrD*. J Bacteriol **188:**8321–8326. http://dx.doi.org/10.1128/JB.00865-06.
29. **Leenhouts K, Buist G, Bolhuis A, ten Berge A, Kiel J, Mierau I, Dabrowska M, Venema G, Kok J.** 1996. A general system for generating unlabelled gene replacements in bacterial chromosomes. Mol Gen Genet **253:**217–224. http://dx.doi.org/10.1007/s004380050315.
30. **Holo H, Nes IF.** 1989. High-Frequency Transformation, by electroporation, of *Lactococcus lactis subsp. cremoris* grown with glycine in osmotically stabilized media. Appl Environ Microbiol **55:**3119–3123.
31. **Thurlow LR, Thomas VC, Hancock LE.** 2009. Capsular polysaccharide production in *Enterococcus faecalis* and contribution of CpsF to capsule serospecificity. J Bacteriol **191:**6203–6210. http://dx.doi.org/10.1128/JB.00592-09.
32. **Edgar RC.** 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res **32:**1792–1797. http://dx.doi.org/10.1093/nar/gkh340.
33. **Paradis E, Claude J, Strimmer K.** 2004. APE: Analyses of Phylogenetics and Evolution in R language. Bioinformatics **20:**289–290. http://dx.doi.org/10.1093/bioinformatics/btg412.
34. **Darling AE, Mau B, Perna NT.** 2010. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS One **5:**e11147. http://dx.doi.org/10.1371/journal.pone.0011147.
35. **Sullivan MJ, Petty NK, Beatson SA.** 2011. Easyfig: a genome comparison visualizer. Bioinformatics **27:**1009–1010. http://dx.doi.org/10.1093/bioinformatics/btr039.
36. **Hochberg Y, Benjamini Y.** 1990. More powerful procedures for multiple significance testing. Stat Med **9:**811–818. http://dx.doi.org/10.1002/sim.4780090710.
37. **Galloway-Pena JR, Bourgogne A, Qin X, Murray BE.** 2011. Diversity of

the *fsr-gelE* region of the *Enterococcus faecalis* genome but conservation in strains with partial deletions of the fsr operon. Appl Environ Microbiol **77**:442–451. http://dx.doi.org/10.1128/AEM.00756-10.

38. **Jacob AE, Hobbs SJ.** 1974. Conjugal transfer of plasmid-borne multiple antibiotic-resistance in *Streptococcus faecalis* var. *zymogenes.* J Bacteriol **117**:360–372.

39. **Gentry-Weeks C, Estay M, Loui C, Baker D.** 2003. Intravenous mouse infection model for studying the pathology of *Enterococcus faecalis* infections. Infect Immun **71**:1434–1441. http://dx.doi.org/10.1128/IAI.71.3 .1434-1441.2003.

40. **Solheim M, Aakra A, Snipen LG, Brede DA, Nes IF.** 2009. Comparative genomics of *Enterococcus faecalis* from healthy Norwegian infants. BMC Genomics **10**:194. http://dx.doi.org/10.1186/1471-2164-10-194.

41. **Yasmin A, Kenny JG, Shankar J, Darby AC, Hall N, Edwards C, Horsburgh MJ.** 2010. Comparative genomics and transduction potential of *Enterococcus faecalis* temperate bacteriophages. J Bacteriol **192**:1122–1130. http://dx.doi.org/10.1128/JB.01293-09.

42. **Paulsen IT, Banerjei L, Myers GS, Nelson KE, Seshadri R, Read TD, Fouts DE, Eisen JA, Gill SR, Heidelberg JF, Tettelin H, Dodson RJ, Umayam L, Brinkac L, Beanan M, Daugherty S, DeBoy RT, Durkin S, Kolonay J, Madupu R, Nelson W, Vamathevan J, Tran B, Upton J, Hansen T, Shetty J, Khouri H, Utterback T, Radune D, Ketchum KA, Dougherty BA, Fraser CM.** 2003. Role of mobile DNA in the evolution of vancomycin-resistant *Enterococcus faecalis.* Science **299**:2071–2074. http://dx.doi.org/10.1126/science.1080613.

43. **Maddox SM, Coburn PS, Shankar N, Conway T.** 2012. Transcriptional regulator PerA influences biofilm-associated, platelet binding, and metabolic gene expression in *Enterococcus faecalis.* PLoS One **7**:e34398. http://dx.doi.org/10.1371/journal.pone.0034398.

44. **Matos RC, Lapaque N, Rigottier-Gois L, Debarbieux L, Meylheuc T, Gonzalez-Zorn B, Repoila F, Lopes Mde F, Serror P.** 2013. *Enterococcus faecalis* prophage dynamics and contributions to pathogenic traits. PLoS Genet **9**:e1003539. http://dx.doi.org/10.1371/journal.pgen.1003539.

45. **Lepage E, Brinster S, Caron C, Ducroix-Crepy C, Rigottier-Gois L, Dunny G, Hennequet-Antier C, Serror P.** 2006. Comparative genomic hybridization analysis of *Enterococcus faecalis*: identification of genes absent from food strains. J Bacteriol **188**:6858–6868. http://dx.doi.org/10 .1128/JB.00421-06.

46. **Desiere F, Lucchini S, Canchaya C, Ventura M, Brussow H.** 2002. Comparative genomics of phages and prophages in lactic acid bacteria. A Van Leeuw **82**:73–91. http://dx.doi.org/10.1023/A:1020676825358.

47. **Ruiz-Garbajosa P, Bonten MJ, Robinson DA, Top J, Nallapareddy SR, Torres C, Coque TM, Canton R, Baquero F, Murray BE, del Campo R, Willems RJ.** 2006. Multilocus sequence typing scheme for *Enterococcus faecalis* reveals hospital-adapted genetic complexes in a background of high rates of recombination. J Clin Microbiol **44**:2220–2228. http://dx.doi .org/10.1128/JCM.02596-05.

48. **Zhao C, Hartke A, La Sorda M, Posteraro B, Laplace JM, Auffray Y, Sanguinetti M.** 2010. Role of methionine sulfoxide reductases A and B of *Enterococcus faecalis* in oxidative stress and virulence. Infect Immun **78**:3889–3897. http://dx.doi.org/10.1128/IAI.00165-10.

49. **Arias CA, Murray BE.** 2012. The rise of the Enterococcus: beyond vancomycin resistance. Nat Rev Microbiol **10**:266–278. http://dx.doi.org/10 .1038/nrmicro2761.

50. **McBride SM, Fischetti VA, Leblanc DJ, Moellering RC, Jr, Gilmore MS.** 2007. Genetic diversity among *Enterococcus faecalis.* PLoS One **2**:e582. http://dx.doi.org/10.1371/journal.pone.0000582.

51. **Palmer KL, Daniel A, Hardy C, Silverman J, Gilmore MS.** 2011. Genetic basis for daptomycin resistance in enterococci. Antimicrob Agents Chemother **55**:3345–3356. http://dx.doi.org/10.1128/AAC.00207-11.

52. **Jett BD, Huycke MM, Gilmore MS.** 1994. Virulence of enterococci. Clin Microbiol Rev **7**:462–478.

53. **Chowdhury SA, Nallapareddy SR, Arias CA, Murray BE.** 2014. The

54. majority of a collection of U.S. endocarditis *Enterococcus faecalis* isolates obtained from 1974 to 2004 lack capsular genes and belong to diverse, non-hospital-associated lineages. J Clin Microbiol **52**:549–556. http://dx .doi.org/10.1128/JCM.02763-13.

54. **Nakayama J, Kariyama R, Kumon H.** 2002. Description of a 23.9-kilobase chromosomal deletion containing a region encoding *fsr* genes which mainly determines the gelatinase-negative phenotype of clinical isolates of *Enterococcus faecalis* in urine. Appl Environ Microbiol **68**:3152–3155. http://dx.doi.org/10.1128/AEM.68.6.3152-3155.2002.

55. **Bizzini A, Zhao C, Budin-Verneuil A, Sauvageot N, Giard JC, Auffray Y, Hartke A.** 2010. Glycerol is metabolized in a complex and strain-dependent manner in *Enterococcus faecalis.* J Bacteriol **192**:779–785. http: //dx.doi.org/10.1128/JB.00959-09.

56. **Joseph B, Mertins S, Stoll R, Schar J, Umesha KR, Luo Q, Muller-Altrock S, Goebel W.** 2008. Glycerol metabolism and PrfA activity in *Listeria monocytogenes.* J Bacteriol **190**:5412–5430. http://dx.doi.org/10 .1128/JB.00259-08.

57. **Xu Y, Murray BE, Weinstock GM.** 1998. A cluster of genes involved in polysaccharide biosynthesis from *Enterococcus faecalis* OG1RF. Infect Immun **66**:4313–4323.

58. **Brinster S, Furlan S, Serror P.** 2007. C-terminal WxL domain mediates cell wall binding in *Enterococcus faecalis* and other gram-positive bacteria. J Bacteriol **189**:1244–1253. http://dx.doi.org/10.1128/JB.00773-06.

59. **Garsin DA, Frank KL, Silanpaa J, Ausubel FM, Hartke A, Shankar N, Murray BE.** 2014. Pathogenesis and Models of Enterococcal Infection. *In* Gilmore MS, Clewell DB, Ike Y, Shankar N (ed), Enterococci: From Commensals to Leading Causes of Drug Resistant Infection, Boston.

60. **Pei J, Grishin NV.** 2001. Type II CAAX prenyl endopeptidases belong to a novel superfamily of putative membrane-bound metalloproteases. Trends Biochem Sci **26**:275–277. http://dx.doi.org/10.1016/S0968-0004 (01)01813-8.

61. **Duerkop BA, Clements CV, Rollins D, Rodrigues JL, Hooper LV.** 2012. A composite bacteriophage alters colonization by an intestinal commensal bacterium. Proc Natl Acad Sci U S A **109**:17621–17626. http://dx.doi.org /10.1073/pnas.1206136109.

62. **Rigottier-Gois L, Alberti A, Houel A, Taly JF, Palcy P, Manson J, Pinto D, Matos RC, Carrilero L, Montero N, Tariq M, Karsens H, Repp C, Kropec A, Budin-Verneuil A, Benachour A, Sauvageot N, Bizzini A, Gilmore MS, Bessieres P, Kok J, Huebner J, Lopes F, Gonzalez-Zorn B, Hartke A, Serror P.** 2011. Large-scale screening of a targeted *Enterococcus faecalis* mutant library identifies envelope fitness factors. PLoS One **6**:e29023. http://dx.doi.org/10.1371/journal.pone.0029023.

63. **Reffuveille F, Leneveu C, Chevalier S, Auffray Y, Rince A.** 2011. Lipoproteins of *Enterococcus faecalis*: bioinformatic identification, expression analysis and relation to virulence. Microbiol **157**:3001–3013. http://dx.doi .org/10.1099/mic.0.053314-0.

64. **Bohle LA, Riaz T, Egge-Jacobsen W, Skaugen M, Busk OL, Eijsink VG, Mathiesen G.** 2011. Identification of surface proteins in *Enterococcus faecalis* V583. BMC Genomics **12**:135. http://dx.doi.org/10.1186/1471 -2164-12-135.

65. **Reffuveille F, Serror P, Chevalier S, Budin-Verneuil A, Ladjouzi R, Bernay B, Auffray Y, Rince A.** 2012. The prolipoprotein diacylglyceryl transferase (Lgt) of *Enterococcus faecalis* contributes to virulence. Microbiol **158**:816–825. http://dx.doi.org/10.1099/mic.0.055319-0.

66. **Steen A, Buist G, Leenhouts KJ, El Khattabi M, Grijpstra F, Zomer AL, Venema G, Kuipers OP, Kok J.** 2003. Cell wall attachment of a widely distributed peptidoglycan binding domain is hindered by cell wall constituents. J Biol Chem **278**:23874–23881. http://dx.doi.org/10.1074/jbc .M211055200.

67. **Zhao Z, Sagulenko E, Ding Z, Christie PJ.** 2001. Activities of *virE1* and the VirE1 secretion chaperone in export of the multifunctional VirE2 effector via an Agrobacterium type IV secretion pathway. J Bacteriol **183**:3855–3865. http://dx.doi.org/10.1128/JB.183.13.3855-3865.2001.